Introduction to Topology

Thomas Kwok-Keung Au

Contents

Chapter 1. Metric Spaces 1
1.1. The Space with Distance 1
1.2. Balls, Interior, and Open sets 5
1.3. Metric Topology 9

Chapter 2. Transition to Topology 13
2.1. Cluster, Accumulation, Closed sets 13
2.2. Continuous Mappings 16
2.3. Sequence 20
2.4. Complete Metric Space 23
2.5. Continuity and Sequences 26
2.6. Baire and Countability 30
2.7. Further Continuity 34

Chapter 3. Topological Spaces 39
3.1. Topology, Open and Closed 39
3.2. Base and Subbase 42
3.3. Countability 45

Chapter 4. Space Constructions 49
4.1. Subspaces 49
4.2. Finite Product 50
4.3. Quotient Spaces 54
4.4. Examples of Spaces 59
4.5. Digression: Quotient Group 63
4.6. Infinite Product 67

Chapter 5. Compactness 71
5.1. Compact Spaces and Sets 71
5.2. Compactness 74
5.3. Compactness and Separation 77
5.4. Locally Compactness 80
5.5. Equivalences 83

Chapter 6. Connectedness 87
6.1. Disconnected and Connected 87
6.2. Components 91
6.3. Other Connectivity 95

Chapter 7. Algebraic Topology 99
7.1. Idea of Invariant 99

7.2. Homotopy 104
7.3. Homotopy Classes and Homotopy Equivalences 108
7.4. Fundamental Group 113
7.5. Useful Examples 118
7.6. Homotopy Invariance 126
7.7. Brouwer Fixed Point Theorem 129

CHAPTER 1

Metric Spaces

In mathematics, there are many occasions that we need to compare objects, approximate one object by another, or take limit of objects. Many of times, this can be easily achieved if there is a distance between any pair of objects. In other words, a set with a measurement of distance is what we need. Of course, there are certain natural rules about the distance. The rules become the definition of a distance, or called metric. Moreover, properties are developed from those rules. The aim of this chapter is an introduction to those essential properties that will be frequently used in all branches of mathematics.

In this chapter, we will start with the definition of metric spaces in §1.1, continued with the most basic concept of open sets in §1.2. Using open sets, we will pave our way towards topology in §1.3 by defining open sets and interior.

1.1. The Space with Distance

We choose to start the study of topology from a natural extension of absolute value or modulus between two numbers, that is, a distance measurement on a set. This provides an easy intuition of the study.

Let X be a nonempty set. A *metric* on X is a function $d\colon X \times X \to [0, \infty)$, that is, $d(x, y) \geq 0$, satisfying the followings

- $d(x, y) = 0$ if and only if $x = y$;
- $d(x, y) = d(y, x)$ for all $x, y \in X$;
- $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$.

Definition 1.1. The pair (X, d) is called a *metric space*.

The first criterion emphasizes that a zero distance is exactly equivalent to being the same point. The second symmetry criterion is natural. The third criterion is usually referred to as the *triangle inequality*.

The concept of metric space is trivially motivated by the easiest example, the Euclidean space. Namely, the metric space $(\mathbb{R}^n, d)$ with

$$d(x, y) = \|x - y\| = \left[\sum_{k=1}^{n} (x_k - y_k)^2\right]^{1/2} ,$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. This is usually referred to as the *standard metric* on $\mathbb{R}^n$.

Example 1.2. There are other metrics on $\mathbb{R}^n$, customarily called ℓ_p-metric, for $p \geq 1$, where

$$d_p(x, y) = \|x - y\|_p = \left[\sum_{k=1}^{n} (x_k - y_k)^p\right]^{1/p} .$$

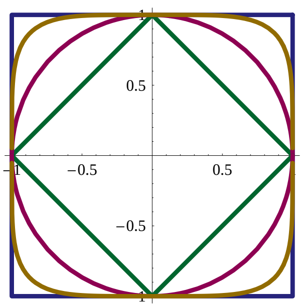
In this sense, the standard metric is actually the ℓ_2-metric. There is also the ℓ_∞-metric given by

$$d_\infty(x, y) = \max\left\{|x_k - y_k| : k = 1, \ldots, n\right\} .$$

The properties of metric can be easily verified in these cases. Perhaps, the hardest one will be left as an exercise.

Exercise 1.1.1. Prove that the triangle inequality is satisfied by the ℓ_p-metric on $\mathbb{R}^n$ for all $p \geq 1$ and $p = \infty$.

We will discuss more about the relationship between ℓ_p-metric for different values of p. For this moment, let us consider the the pictures for several values of p about the sets $\{\, x \in \mathbb{R}^n : d_p(x, 0) = 1 \,\}$. The pictures are illustrative that they are convex when $p \geq 1$.



The pictures for $p = 1$ (green), $p = 2$ (purple), $p = 5$ (brown), and $p = \infty$ (blue).

EXAMPLE 1.3. The *discrete metric* on any nonempty set X is defined by

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

The criteria of being a metric can be readily verified case by case (EXERCISE 1.1.2). This is kind of an uninteresting metric because any two distinct points will have a fixed distance afar. However, it often serves as an example to check certain property of a space. The following exercise is often a good way to understand a metric.

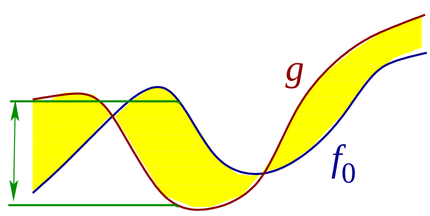
EXERCISE 1.1.3. Let (X,d) be the discrete metric space and $x_0 \in X$. Determine the sets $\{ x \in X : d(x,x_0) < r \}$ for different values of $r > 0$.

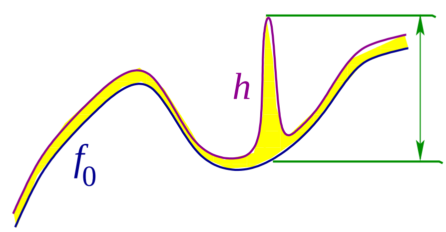
EXAMPLE 1.4. Similar to the situation of $\mathbb{R}^n$, there are several metrics on a function space. For simplicity, let $X = \mathcal{C}([a,b],\mathbb{R})$ be the set of all continuous real valued functions defined on an interval $[a,b]$. We have metrics d_p for $p \geq 1$ and $p = \infty$, namely, for $f, g \in X$,

$$d_p(f,g) = \left[\int_a^b |f(t) - g(t)|^p \, dt \right]^{1/p},$$
$$d_\infty(f,g) = \sup \{ |f(t) - g(t)| : t \in [a,b] \}.$$

The proof for that these are metrics is similar to the Euclidean cases. In fact, d_∞ is a metric on $\mathcal{B}([a,b],\mathbb{R})$, the set of bounded functions on $[a,b]$.

In the following pictures, we will show a comparison between $p = 1$ and $p = \infty$ to illustrate their differences. Let $f_0, g, h \in X$. The yellow area between the curves illustrate the distances, $d_1(g,f_0)$ is large while $d_1(h,f_0)$ is small. On the other hand, the green arrows illustrate the sup-distances, both $d_\infty(g,f_0)$ and $d_\infty(h,f_0)$ are large.





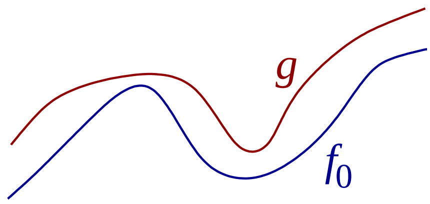
Again, it is beneficial to think about the set $\{ f \in X : \; d_p(f,0) < 1 \}$ where 0 is the constant zero function.

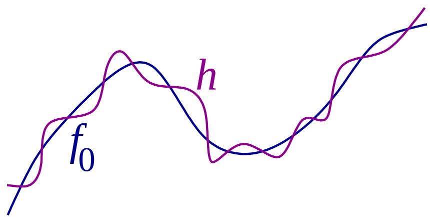
EXERCISE 1.1.4. Is d_1 defined above a metric for the set $\mathcal{L}([a,b],\mathbb{R})$ of all integrable functions on an interval $[a,b]$?

EXAMPLE 1.5. A suitable choice of metric may have the effect of good comparison. Let X be the set of all continuously differentiable (C^1) functions on an interval $[a, b]$ and

$$d(f, g) = \sup \{ |f(t) - g(t)| : t \in [a, b] \} + \sup \left\{ \left|f'(t) - g'(t)\right| : t \in [a, b] \right\} .$$

With this choice of metric, for the functions illustrated below, $d(f, g) < d(f, h)$ because the contribution of derivatives $|f'(t) - h'(t)|$ is large.





EXERCISE 1.1.5. (1) Let $\mathbb{R}^\infty$ be the set of all sequences $x = (x_k)_{k=1}^\infty$ in $\mathbb{R}$ of which only finitely many terms are nonzero. Show that the ℓ_p-metrics for all $p \geq 1, \infty$ are well-defined metrics on $\mathbb{R}^\infty$.

(2) Prove that if (X, d) is a metric space, then d^* is also a metric on X, where

$$d^*(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \qquad x, y \in X .$$

(3) Let (X, d) be a metric space such that $0 \leq d < 1$. Determine whether

$$d^\#(x, y) = \sum_{k=1}^\infty \frac{d(x, y)^k}{2^k} \qquad \text{is also a metric.}$$

(4) Let (X, d) be a metric space and $f \colon [0, \infty) \to [0, \infty)$. Try to explore the conditions on f such that $d_f = f \circ d$ is also a metric on X. From the above, $f(r) = r/(1 + r)$ is an example. Try to give another example.

(5) Let (X, d) be a metric space and $A \subset X$. Define d_A on $A \times A$ by

$$d_A(a_1, a_2) = d(a_1, a_2), \qquad a_1, a_2 \in A .$$

Show that (A, d_A) is a metric space. The metric d_A is called the *induced metric* on A.

(6) Let (X, d_X) and (Y, d_Y) be metric spaces. Prove that both

$$d_\infty((x_1, y_1), (x_2, y_2)) = \max \{ d_X(x_1, x_2), d_Y(y_1, y_2) \} ,$$
$$d_1((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$$

for $x_1, x_2 \in X$ and $y_1, y_2 \in Y$, are metrics on $X \times Y$. Both may be called the *product metric*. Do you think there are metrics d_p?

To finish this section, we will give two examples. Both examples are related to error-correcting applications. You are encouraged to understand them by the two standard tasks: verify the metric conditions and think of the typical situation of $d(x, x_0) < r$.

EXAMPLE 1.6. Let $X = \{0, 1\}^n$, i.e., it contains points of n-coordinates of 0 and 1; $d(x, y)$ is the number of different coordinates between x and y.

EXAMPLE 1.7. Let X be the set of finite sequences of alphabets. For example, "homomorphic", "homeomorphic", "holomorphic", "homotopic", "homologous" are elements of X. Suppose there are three valid operations, inserting an alphabet, deleting an alphabet, and replacing an alphabet by another. For two elements $x, y \in X$, define $d_3(x, y)$ by the minimum number of operations required to transform x to y.

We may also consider replacing an alphabet equivalent to deleting then inserting. In this case, we only accept two types of operations and let $d_2(x, y)$ be the minimum number of such operations.

1.2. Balls, Interior, and Open sets

In this section, we will discuss an important concept of open sets in a metric space. Open set is the most fundamental notion in topology and it will be used often in more general context.

Let (X, d) be a metric space. The discussion starts with open balls.

DEFINITION 1.8. At any point $x \in X$ and $\varepsilon > 0$, an *open ball* at x with radius ε is the set $B(x, \varepsilon) = \{ y \in X : d(y, x) < \varepsilon \}$.

An open ball in the standard $\mathbb{R}$ is merely an open interval while that in $\mathbb{R}^n$ with ℓ_2-metric is the usual idea of circular ball. For balls of other ℓ_p-metrics, please refer to the typical pictures shown after Example 1.2 in Section 1.1.

EXERCISE 1.2.1.
(1) Find $B(x, \varepsilon)$ in the discrete metric space.
(2) Let (X, d) be a metric space and $d^* = d/(1 + d)$ be another metric (see Exercise 1.1.5). Find the relation between the open balls determined by the two metrics.
(3) Similar to the above, compare the balls of a metric d and $d_f = f \circ d$ assuming that it is still a metric.

(4) Let (X, d) be a metric space and $A \subset X$ be given the induced metric (see Exercise 1.1.5). Show that $B_A(a, \varepsilon) = B_X(a, \varepsilon) \cap A$ for any $a \in A$, where B_A and B_X denote the open balls in A and X respectively.

(5) Let X and Y be metric spaces. Express an open ball in the product metric space $X \times Y$ (see Exercise 1.1.5) in terms of open balls in X and Y. Note that the answer may depends on which product metric you are choosing.

(6) Is there a metric d on $\mathbb{R}^2$ such that all the open balls $B(x, \varepsilon)$ are ellipses with center at x? What if we change the requirement that x is at one of the foci of the ellipses?

(7) Let (X, d) be a metric space and $A \subset X$. Define the *diameter* of A by

$$\operatorname{diam}(A) = \sup \{ d(a_1, a_2) : a_1, a_2 \in A \} .$$

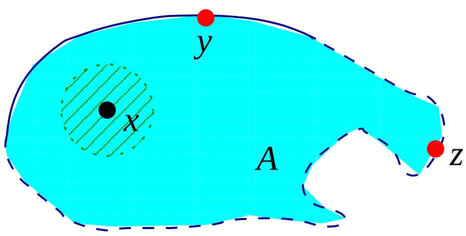
Is it true that $\operatorname{diam}(B(x, \varepsilon)) = \varepsilon$?

(8) If $A \subset X$ has $\operatorname{diam}(A) < \varepsilon$ and $A \cap B(x, \varepsilon) \neq \emptyset$, then $A \subset B(x, 2\varepsilon)$.

Definition 1.9. Let (X, d) be a metric space, $A \subset X$, and $x \in A$. The point x is called an *interior point* of A if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset A$.

Note that in the definition, the radius ε depends on the "position" of x in A.

Remark . As a good practice of beginning in topology, let us write down the negation of the definition. That is, a point $w \in X$ is *not* an interior point of $A \subset X$ if for all $\varepsilon > 0$, $B(w, \varepsilon) \cap (X \setminus A)$ is nonempty. Certainly, this negation is satisfied in the case $w \in X \setminus A$ as $B(w, \varepsilon) \cap (X \setminus A) \supset \{w\}$.



In the above illustration, the set A includes the solid boundary curve but not the dashed boundary curve. The point $x \in A$ is an interior point because the dotted green ball also belongs to A; the point $y \in A$ but it is not an interior point; the point $z \notin A$ is definitely not an interior point. The points y, z in the above picture in fact satisfy a bit more than the negation, namely, for all $\varepsilon > 0$, $B(y, \varepsilon) \cap (X \setminus A) \neq \emptyset$ and $B(y, \varepsilon) \cap A \neq \emptyset$. Later, we will see that these are called *frontier points* of A.

EXERCISE 1.2.2.
(1) Prove that in any metric space (X, d), any point $y \in B(x, \varepsilon)$ is an interior point of $B(x, \varepsilon)$. Note that triangle inequality of d must be used in the proof.
(2) Show that if $A \subset \mathbb{R}^n$ and $x \in A$, then x is an interior point of A in the standard $\mathbb{R}^n$ if and only if x is an interior point of A in $(\mathbb{R}^n, d_p)$ for all ℓ_p-metric with $p \geq 1$ or $p = \infty$ (Example 1.2).

In the second problem above, if we denote $B_p(x, \varepsilon)$ the open ball in the space $(\mathbb{R}^n, d_p)$, then the whole problem comes down to a comparison of open balls. More precisely, we need to establish such a statement: given p, q, for all $\varepsilon > 0$, there exists $\delta > 0$ such that $B_q(x, \delta) \subset B_p(x, \varepsilon)$. This statement is trivial if $p \geq q$ as one may take $\delta = \varepsilon$. It is still easy for $p < q$. The pictures in Example 1.2 may be helpful to find δ in terms of ε. In fact, such δ may not be found if the space is infinitely dimensional.

DEFINITION 1.10. Let $A \subset X$ in a metric space (X, d). The *interior* of A, denoted $\mathring{A}$ or $\mathrm{Int}(A)$, is the set of all interior points of A. The set $G \subset X$ is an *open set* if $G = \mathring{G} = \mathrm{Int}(G)$, i.e., every point of G is an interior point of G.

Such definitions match the intuition of the standard $\mathbb{R}^n$. The interior of a subset in $\mathbb{R}^n$ is exactly those points that is not on the "boundary". For example, for $A = \{ (x, y) : x^2 + y^2 \leq 1 \}$, its interior is the set $\{ (x, y) : x^2 + y^2 < 1 \}$. By Exercise 1.2.2, the interior of a set in $(\mathbb{R}^n, d_p)$ is independent of $p \geq 1$ or $p = \infty$.

EXAMPLE 1.11.
(1) In any metric space (X, d), $\emptyset$ and X are always open sets.
(2) From Exercise 1.2.2, any open ball $B(x, \varepsilon)$ of a metric space is open.
(3) It can be easily shown (also given in Exercise 1.2.1) that in a discrete space (X, d), an open ball $B(x, \varepsilon)$ is either $\{x\}$ or X. Thus, it follows that any subset $A \subset X$ is open.
(4) Let $A \subset X$ be given the induced metric d_A from the metric space (X, d). Note that the interior of A in (A, d_A) is always A, while the interior of A in (X, d) is only a subset of A.

    In addition, if $B \subset A$, then it may happen that $\mathrm{Int}_A(B) \neq \mathrm{Int}_X(B)$, where $\mathrm{Int}_A(B)$ denotes the interior of B in (A, d_A). For this reason, if B is open in (X, d) then it is open in (A, d_A). However, the converse is not true.

    EXERCISE 1.2.3. Give a condition on A such that B is open in A if and only if it is open in X.

(5) An open set in the product metric space $X \times Y$ is always of the form $U \times V$ where U is open in X and V is open in Y. Note that this may not be true for infinite product.

The standard argument of showing a subset U in (X, d) is open must involve the following. Basically, it is to show $U \subset \mathrm{Int}(U)$, i.e., every point x of U is an interior point of U. Thus, we begin by taking arbitrary point $x \in U$, then try to find an $\varepsilon > 0$ with $B(x, \varepsilon) \subset U$. To get the inclusion, we take any $y \in B(x, \varepsilon)$, i.e., $d(y, x) < \varepsilon$ and try to argue that $y \in U$. Here, we will demonstrate such a process by an example.

Example 1.12. We will show that for any subset A in (X, d), its interior $\mathrm{Int}(A)$ is an open set.

Let $x \in \mathrm{Int}(A)$. We need to find a suitable ε with $B(x, \varepsilon) \subset \mathrm{Int}(A)$.

By definition of $x \in \mathrm{Int}(A)$, there exists $\delta > 0$ such that $B(x, \delta) \subset A$. In the following, we will show that just taking $\varepsilon = \delta$ will be enough for our need. In other words, indeed, $B(x, \delta) \subset \mathrm{Int}(A)$.

Take arbitrary $y \in B(x, \delta)$. Since $B(x, \delta)$ is open (or by directly using triangle inequality argument), there exists $\xi > 0$ such that $y \in B(y, \xi) \subset B(x, \delta)$. Hence $y \in B(y, \xi) \subset A$. This shows that $y \in \mathrm{Int}(A)$ for arbitrary $y \in B(x, \delta)$. This simply means $B(x, \delta) \subset \mathrm{Int}(A)$. □

Equivalently, we have $\mathrm{Int}(\mathrm{Int}(A)) = \mathrm{Int}(A)$. Note that the proof above is very similar to showing that an open ball is open. The same argument is useful to show that if $B \subset A$, then $\mathrm{Int}(B) \subset \mathrm{Int}(A)$.

In addition, one may prove the following (which is left as an exercise). It shows that in the future, the concept of open balls can be replaced by open sets.

Proposition 1.13. *Let $A \subset X$. A point $x \in A$ is an interior point of A if and only if there exists an open set U in X such that $x \in U \subset A$.*

Exercise 1.2.4.

(1) Show that if each U_α is open then $\bigcup_\alpha U_\alpha$ is also open.
(2) Show that if each $U_1, \ldots, U_n$ is open then $U_1 \cap \cdots \cap U_n$ is also open.
(3) Give an example of infinitely many open sets U_α which has $(\bigcap_\alpha U_\alpha)$ not open. Also, give an example that $(\bigcap_\alpha U_\alpha)$ is still open.
(4) Show that the interior $\mathrm{Int}(A)$ is the largest open set contained in A.
(5) Prove that a subset A is open if and only if it is a union of open balls.

1.3. Metric Topology

In this section, we will introduce the concept of topology arisen from a metric. In the future, it will be the foundation of study even when there is no metric. In principle, it tells us what the open sets are.

DEFINITION 1.14. Let (X, d) be a metric space. The *topology* of the metric d is the set, $\mathfrak{T}_d$ or simply $\mathfrak{T}$, containing all the open subsets of X. Thus, by definition, $\mathfrak{T}$ is a subset of the power set $\mathcal{P}(X)$ of X.

From Exercise 1.2.1, in the discrete metric space, every singleton $\{ x \}$ is an open set. Thus, by taking their unions, every set is open. Hence, the topology with respect to the discrete metric is the power set.

EXAMPLE 1.15. On $\mathbb{R}^n$, no matter which ℓ_p-metric is considered, $p \geq 1$ or $p = \infty$, it does not change whether a subset is open or not. In other words, the topologies corresponding to all ℓ_p-metrics is the same; they contain the same open sets. This common topology is called the *standard topology* of Euclidean space. In fact, consider the metric on $\mathbb{R}^2$ that every open ball $B(x, \varepsilon)$ is an ellipse with center at x. This metric also gives rise to the standard topology.

As we have mentioned before, on an infinite dimensional space, all ℓ_p-metrics may not give the same topology. In fact, if X is the space of all continuous functions on an interval $[a, b]$, the topologies determined by the integral metric and the maximum metric are not the same.

EXERCISE 1.3.1. (1) If (X, d) has $\mathfrak{T}$ equals the power set, what is d?
(2) Let (X, d) be a metric space and $d^* = d/(1+d)$. Show that the topologies $\mathfrak{T}_d = \mathfrak{T}_{d^*}$.

DEFINITION 1.16. Let d_1 and d_2 are two metrics on X with their corresponding topologies $\mathfrak{T}_1$ and $\mathfrak{T}_2$. The metrics are *equivalent* if $\mathfrak{T}_1 = \mathfrak{T}_2$.

Now, one can see that all the ℓ_p-metrics on $\mathbb{R}^n$ are equivalent. Moreover, for any metric d, one can always have an equivalent metric $d/(1 + d)$ with distances in the range $[0, 1)$. In the future, one may see that the topology is the true basis of study. It does not always require a definition of metric. This can be illustrated by the following fact.

THEOREM 1.17. *Two metrics d_1, d_2 are equivalent if and only if every open ball in (X, d_1) is an open set in (X, d_2) and vice versa.*

Proof. Assume that $\mathfrak{T}_1 \subset \mathfrak{T}_2$. Let $x \in X$ and take any open ball $B_1(x, \varepsilon)$ wrt d_1. Obviously, $B_1(x, \varepsilon) \in \mathfrak{T}_1$ and, by assumption, it is also in $\mathfrak{T}_2$. For the point $x \in B_1(x, \varepsilon)$, it must be an interior point wrt d_2. Thus, there exists $\delta > 0$ such that $x \in B_2(x, \delta) \subset B_1(x, \varepsilon)$. So, we have proved one direction of the following proposition, which is essentially the theorem. The other direction is an easy exercise. $\square$

Proposition 1.18. *Two topologies $\mathfrak{T}_1 \subset \mathfrak{T}_2$ if and only if for every d_1-open ball $B_1(x, \varepsilon)$, there exists a d_2-open ball $B_2(x, \delta) \subset B_1(x, \varepsilon)$.*

There are several terminologies about two topologies that $\mathfrak{T}_1 \subset \mathfrak{T}_2$. In different books, it may be called weaker versus stronger, smaller versus larger, coarser versus finer. We will try to avoid confusion by specifying such comparison by the mathematical expressions $\subset$ or $\supset$.

For future convenience, a set $N \subset X$ with $x \in \mathrm{Int}(N)$ is called a *neighborhood* of x. Very often, it is equivalent to simply take an *open neighborhood* U of x, that is, $x \in U \in \mathfrak{T}$. The following exercise demonstrates that the role of open ball may be replaced with neighborhoods. One may verify it by simply writing equivalent statements with patience.

Exercise 1.3.2. Given a metric space (X, d), $A \subset X$ and $x \in X$. There exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subset A$ if and only if there exists a neighborhood N of x with $x \in N \subset A$. On the other hand, $B(x, \varepsilon) \cap A \neq \emptyset$ for each $\varepsilon > 0$ if and only if for each neighborhood N of x, $N \cap A \neq \emptyset$.

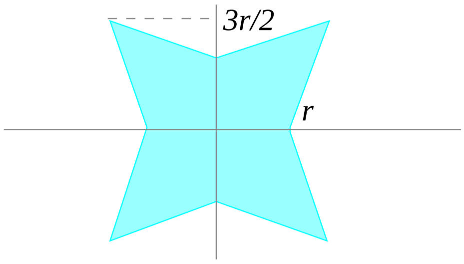
With the concept of topology and neighborhoods, one sees that the concept of open ball is not totally essential. Once we know what it means by open sets, we are able to use neighborhoods instead of open balls. The is the essential key to the understanding in the following chapters. Here is a situation that illustrates that a metric is not essential to define topology.

Example 1.19. A function $\rho : X \times X \to [0, \infty)$ is temporarily called a pseudo-metric if it satisfies all the conditions of a metric except replacing the triangle inequality to

$$\rho(x, y) + \rho(y, z) \geq \alpha\, \rho(x, z), \qquad \text{where } 0 < \alpha < 1 \text{ is a fixed number.}$$

Show that we may also define open ball $B(x, \varepsilon)$, interior points, and open sets; and hence topology similarly just as the case for metric.

Exercise 1.3.3. Try to define a pseudo-metric ρ on $\mathbb{R}^2$ such that its $B(0,r)$ is given by the star-shaped picture shown below.



Note that for this ρ, it satisfies the above inequality with $\alpha = 2/3$.

Exercise 1.3.4. From the observation in the above example and exercise, is it true that "if d is a metric on $\mathbb{R}^2$ such that it is preserved by translation, i.e., $d(x,y) = d(x+z, y+z)$, then $B(0,r)$ is convex"?

CHAPTER 2

Transition to Topology

In this chapter, we continue to introduce several important notions related to a metric space. However, as mentioned at the end of the previous chapter, we will try to discuss without referring to the metric. From this point onwards, we will mainly use open sets and neighborhoods. Metric will only be mentioned when a property or a theorem is only true in metric space.

In §2.1, closed sets will be introduced as the counterpart of open sets. The next topic is about relation between spaces, which is studied through continuity of mappings in §2.2. Sequences and approximation are discussed in §2.3 in relation to previous notions. Next, important properties about completeness is presented in §2.4. Then we return to the intimate relation between continuity and sequence in §2.5. A brief discussion of Baire category theory and dense sets is given in §2.6. This chapter is ended by further uniform properties of continuity in §2.7.

2.1. Cluster, Accumulation, Closed sets

In many branches of mathematics, there is the need of studying an object which is arbitrarily close to a type of objects. For example, a smooth curve is arbitrarily close to segments of straight lines. This provides a basis for studying approximation. In this section, such concept is explored in the context of metric spaces or more general topological spaces.

Let $A \subset X$. A point $x \in X$ (not necessarily in A) is called a cluster point or accumulation point (we do not use limit point) of A if for every $\varepsilon > 0$, the punctured ball $B(x, \varepsilon) \setminus \{x\}$ always intersects A. However, as we mentioned before, we will use the following equivalent definition.

Definition 2.1. A point $x \in X$ is a *cluster point* of $A \subset X$ if for every neighborhood $U \in \mathfrak{T}$ of x, $(U \setminus \{x\}) \cap A \neq \emptyset$. The *derived set* of A, denoted A', is the set of all cluster points of A.

By writing down the negation, it is easy to see that a point $x \notin A$ is not a cluster point of A if and only if $x \in \mathrm{Int}(X \setminus A)$. In simple mathematics language, $X \setminus (A \cup A') = \mathrm{Int}(X \setminus A)$. Also, if a point $a \in A$ is not a cluster point of A, it is an isolated point belong to A.

EXAMPLE 2.2.
(1) In the standard $\mathbb{R}^n$, if $x \in \mathrm{Int}(A)$ then x is a cluster point of A.
(2) The above statement is not true for all metric spaces. Consider a discrete metric space and see why it does not work.
(3) In the standard $\mathbb{R}$, let $A = \{ 1/n : 1 \leq n \in \mathbb{Z} \}$. Then the point 0 is a cluster point of A and it does not belong to A.

DEFINITION 2.3. The *closure* of A, denoted by $\overline{A}$ or $\mathrm{Cl}(A)$, is the set $A \cup A'$. A subset $F \subset X$ is *closed* if $F = \overline{F} = \mathrm{Cl}(F)$, i.e., every cluster point of F must be inside F.

Comparing the definitions of $\mathrm{Cl}(A)$ and A', it is easy to see that $x \in \mathrm{Cl}(A)$ if and only if for every neighborhood $U \in \mathfrak{T}$ of x, $U \cap A \neq \emptyset$. This is a statement often used in the future.

By definition, F is closed if and only if $F = \mathrm{Cl}(F) = F \cup F'$ if and only if $F \supset F'$. This is equivalent to the following,

$$X \setminus F = X \setminus (F \cup F') = \mathrm{Int}(X \setminus F) .$$

Recall that $A = \mathrm{Int}(A)$ if and only if A is open. Thus, we have shown the following,

PROPOSITION 2.4. *F is closed $\iff$ $X \setminus F$ is open.*

Note that a set may be neither open nor closed, an interval $[a, b)$ in the standard $\mathbb{R}$ is an example. Also, a set may be both open and closed. The easiest example is $\emptyset$ and X. Other nontrivial examples may be introduced in the discussion of connectedness.

EXAMPLE 2.5. The concept of whether a subset A is closed should be considered in the context of the whole space X. Let $X = \mathbb{R}$ and $Y = (0,1) \cup (2,3) \subset \mathbb{R}$, both with the standard topology. Let $A = (0,1)$. If A is considered as a subset of $X = \mathbb{R}$, it is obvious open but not closed. However, if A is considered as a subset of Y, then it is both open and closed. It is open because every point is an interior point. On the other hand, for the same reason, its complement $Y \setminus A = (2,3)$ is also open. Thus, A itself is also closed in Y.

Definition 2.6. Let A be a subset in a space X. A point $x \in X$ belongs to the *frontier* or boundary of A, denoted by $\mathrm{Frt}(A)$ if for every neighborhood $U \in \mathfrak{T}$ of x, $U \cap A \neq \emptyset$ and $U \cap (X \setminus A) \neq \emptyset$.

In other words, x belongs to the frontier of A if x both belongs to the $\mathrm{Cl}(A)$ and $\mathrm{Cl}(X \setminus A)$. We avoid using "boundary" here because it has a different meaning in manifold theory.

Example 2.7. Let us give a number of examples of closed sets.

(1) The sets $\emptyset$ and X are always closed, because their complements are open.
(2) For any subset A, the sets A', $\overline{A}$, and $\mathrm{Frt}(A)$ are closed.
(3) If (a_n) is a sequence in $\mathbb{R}^n$ and it converges to $a \in \mathbb{R}^n$, then $A = \{a\} \cup \{\, a_n \in \mathbb{R}^n \ : \ n \in \mathbb{N} \,\}$ is closed in the standard $\mathbb{R}^n$.

To show the standard way of argument in point set topology, let us demonstrate how to prove that $\overline{A} = \mathrm{Cl}(A)$ is a closed set.

We need to prove $\mathrm{Cl}(\overline{A}) = \overline{A}$. Since $S \subset \overline{S}$ is always true, it is sufficient to show $\mathrm{Cl}(\overline{A}) \subset \overline{A}$. Let $x \in \mathrm{Cl}\left(\overline{A}\right)$. Then by definition, for all neighborhood $U \in \mathfrak{T}$ of x, $U \cap \overline{A} \neq \emptyset$. And we are done if we can show $U \cap A \neq \emptyset$.

As $U \cap \overline{A} \neq \emptyset$, there is a point $y \in U \cap \overline{A}$. Now, $y \in U \in \mathfrak{T}$ means that U is a neighborhood of y. By $y \in \overline{A}$, this neighborhood U must have $U \cap A \neq \emptyset$. That is what desired. □

Exercise 2.1.1. Exercises about closure of a set are often analogous to interior.

(1) Show that if each F_α is closed then $\bigcap_\alpha F_\alpha$ is closed.
(2) Show that if each $F_1, \dots, F_n$ is closed then $\bigcup_{k=1}^n F_k$ is closed.
(3) Prove that $\overline{A}$ is the smallest closed set containing A.
(4) We already have $\mathrm{Int}(\mathrm{Int}(A)) = \mathrm{Int}(A)$, $\mathrm{Cl}(\mathrm{Cl}(A)) = \mathrm{Cl}(A)$, what about $(A')'$ and A'?
(5) Use $\mathrm{Int}(\cdot)$, $\mathrm{Cl}(\cdot)$, $\mathrm{Frt}(\cdot)$, taking complement, etc., to explore equations about them. For examples,
    (a) $\overline{A} = X \setminus \mathrm{Int}(X \setminus A)$
    (b) $\mathring{A} = X \setminus \mathrm{Cl}(X \setminus A)$
    (c) $\mathrm{Int}(A) \cup \mathrm{Frt}(A) = \overline{A}$.
(6) Show that $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ but they are not necessarily equal.

(7) On a metric space (X, d), given $A \subset X$ and $x \in X$, define

$$d(x, A) = \inf \{ d(x, a) \ : \ a \in A \} \ .$$

Show that if $x \in A$ then $d(x, A) = 0$. On the other hand, show that if $d(x, A) = 0$ then $x \in \overline{A}$.

EXERCISE 2.1.2. Let X be a metric space and $\mathcal{B}(X, \mathbb{R})$ be the set of all bounded functions from X to $\mathbb{R}$ and consider the metric space $(\mathcal{B}, d)$ where $d(f, g) = \sup_{t \in X} |f(t) - g(t)|$. Show that $\mathcal{C}(X, \mathbb{R})$, the set of all continuous functions, is a closed set in $\mathcal{B}$.

EXAMPLE 2.8. In topology, there are always surprising results even in simple examples. Here is an example that may help clarify certain concepts. Let $X = (-\infty, 1) \cup [2, 4] \subset \mathbb{R}$ and $d(x, y) = |x - y|$.

As previously discussed, every open ball $B(x, \varepsilon)$ is an open set. So, $B(3, 0.5) = (2.5, 3.5)$ is clearly open. Likewise, $B(3, 1.5)$ is open. However, by definition

$$B(3, 1.5) = \{ x \in X \ : \ d(x, 3) < 1.5 \} = [2, 4] \, .$$

It is a closed interval in $\mathbb{R}$ but an open set in X. Thus, its complement in X, $(-\infty, 1)$ is a closed set in X. Furthermore, consider the open ball $B(0, 2)$, it is the set $(-2, 1) \subset X$. Note that we have $\{ x \in X \ : \ d(x, 0) \leq 2 \} = (-2, 1) \cup \{ 2 \}$. One may also verify that the closure $\mathrm{Cl}(B(0, 2)) = [-2, 1)$ is different from the above set.

2.2. Continuous Mappings

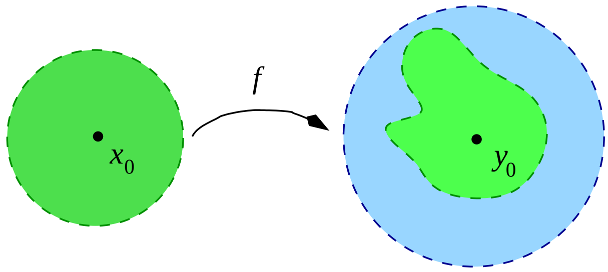
In every branch of mathematics, relations between objects are studied by mappings. Suitable requirements are imposed on the mappings for special purpose of the context. In the study of topology, we study continuous mappings. In this section, we will only give a short and brief description about mappings between metric spaces. Later, more attention will be given on mappings between topological spaces.

It is natural to start with copying the definition of continuous functions on $\mathbb{R}^n$. Let (X, d_X) and (Y, d_Y) be metric spaces. Let $f \colon X \to Y$ be a mapping with $x_0 \in X$ and $y_0 = f(x_0) \in Y$. It is continuous at x_0 if for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \varepsilon$.

Let us start to rewrite the statements. Note that the statement "if $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \varepsilon$" can be rephrased as

$$\text{for all } x \in B_X(x_0, \delta), \qquad f(x) \in B_Y(f(x_0), \varepsilon)\,.$$

Or, simply in set language, $f\,(\,B_X(x_0, \delta)\,) \subset B_Y(f(x_0), \varepsilon)$.



Now, we can rephrase continuity in terms of the topologies $\mathfrak{T}_X$ and $\mathfrak{T}_Y$, which are sets of all open sets determined by the corresponding metric. The definition does not refer to metric and is valid in the future.

Definition 2.9. A mapping $f\colon X \to Y$ is *continuous at* $x_0 \in X$ if for every $V \in \mathfrak{T}_Y$ with $f(x_0) \in V$, there exists $U \in \mathfrak{T}_X$ with $x_0 \in U$ such that $f(U) \subset V$.

For a definition for f being continuous on the whole space, a more useful version is given here. It will be justified below that this is equivalent to continuity at every point.

Definition 2.10. A mapping $f\colon X \to Y$ is continuous if for every $V \in \mathfrak{T}_Y$, the pre-image $f^{-1}(V) \in \mathfrak{T}_X$.

Example 2.11. (1) Let $f\colon \mathbb{R} \to \mathbb{R}$ be a mapping on the standard $\mathbb{R}$ where

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q}\,, \\ 1 & x \notin \mathbb{Q}\,. \end{cases}$$

This mapping is well-known to be discontinuous at every $x_0 \in \mathbb{R}$.

(2) Consider the same mapping $f\colon (\mathbb{R}, d) \to \mathbb{R}$ from $(\mathbb{R}, d)$ with discrete metric d to standard $\mathbb{R}$. Then f is continuous everywhere. It is because $\{x_0\}$ is always an open set in discrete metric, so $f\,(\,\{x_0\}\,) \subset V$ for every V containing $f(x_0)$.

(3) If (X, d) is the discrete metric, then any mapping $f\colon X \to Y$ is continuous.

(4) Recall that on $\mathbb{R}^n$, we have the ℓ_p-metric,

$$d_p(x,y) = \|x-y\|_p = \left[\sum_{k=1}^{n} |x_k - y_k|^p\right]^{1/p},$$
$$d_\infty(x,y) = \max\{\,|x_k - y_k| \;:\; k = 1,\dots,n\,\}.$$

The identity mapping $\mathrm{id}\colon (\mathbb{R}^n, d_p) \to (\mathbb{R}^n, d_q)$ is continuous everywhere for all $p, q \geq 1$ or ∞. At the end, it involves proving the inequality

$$\|z\|_p \leq \|z\|_q \leq C\,\|z\|_p\,, \qquad z \in \mathbb{R}^n, \quad p \leq q\,,$$

where C is a fixed constant depending on p, q, n. With this inequality, one may see that

$$\mathrm{id}\,(\,B_q(x_0,\varepsilon)\,) = B_q(x_0,\varepsilon) \subset B_p(x_0,\varepsilon)\,;$$
$$\mathrm{id}\left(\,B_p(x_0,\frac{\varepsilon}{C})\right) = B_p(x_0,\frac{\varepsilon}{C}) \subset B_q(x_0,\varepsilon)\,.$$

(5) There are also metrics d_p on the space X of all continuous functions from a fixed interval $[a,b]$ to $\mathbb{R}$. The mapping $\mathrm{id}\colon (X, d_1) \to (X, d_\infty)$ is not continuous. This can be seen from the illustration of the difference of the metrics given in Example 1.4 on page 3.

To the above examples, the reader is encouraged to think both in terms of continuity at a point or on the whole.

Exercise 2.2.1. (1) When will a mapping $f\colon X \to Y$ be continuous if Y is the discrete metric space?

(2) Let (X,d) be a metric space. Consider $X \times X$ with any product metric (see page 4) and standard $[0,\infty)$. Show that the mapping

$$d\colon X \times X \to [0,\infty) \quad \text{is continuous.}$$

(3) Let (X, d_X) and (Y, d_Y) be metric spaces; $X \times Y$ be given a product metric. Show that the projection mappings $\pi_X\colon X \times Y \to X$ and $\pi_Y\colon X \times Y \to Y$ are continuous.

In addition, let Z be another metric space with $f\colon Z \to X \times Y$. Prove that f is continuous if and only if both $\pi_X \circ f$ and $\pi_Y \circ f$ are so.

(4) Let $f, g\colon X \to \mathbb{R}$ be continuous functions into standard $\mathbb{R}$. Show that $\{\, x \in X : f(x) < g(x)\,\}$ is open while $\{\, x \in X : f(x) \leq g(x)\,\}$ is closed.

Proposition 2.12. *The following statements are equivalent.*

(1) *$f\colon X \to Y$ is continuous.*

(2) *For each $x_0 \in X$, f is continuous at x_0.*

(3) *For every closed set H in Y, the pre-image $f^{-1}(H)$ is closed in X.*

PROOF. "(2) $\Longrightarrow$ (1)". Let $V \in \mathfrak{T}_Y$. To prove $f^{-1}(V) \in \mathfrak{T}_X$, we need to show that every $x_0 \in f^{-1}(V)$ is an interior point. By (2), there exists $U \in \mathfrak{T}_X$ with $x_0 \in U$ and $f(U) \subset V$. Therefore, $x_0 \in U \subset f^{-1}(V)$ and hence it is an interior point.

"(1) $\Longrightarrow$ (3)". It can be done by simply considering $Y \setminus H \in \mathfrak{T}_Y$. Therefore

$$X \setminus f^{-1}(H) = f^{-1}(Y \setminus H) \in \mathfrak{T}_X.$$

"(3) $\Longrightarrow$ (2)". This is left as an exercise. □

EXERCISE 2.2.2. Let $f\colon X \to Y$ be continuous. Prove that for each $A \subset X$, $f\left(\overline{A}\right) \subset \overline{f(A)}$. Is it true that $f^{-1}\left(\mathring{B}\right) = \operatorname{Int}\left(f^{-1}(B)\right)$ for $B \subset Y$?

EXERCISE 2.2.3. Determine whether there is such an example. Let $f : X \to Y$ be a continuous mapping and $B_n \subset Y$ are closed subsets for $n \in \mathbb{N}$ such that $\bigcup_{n=1}^{\infty} B_n$ is still closed but, $\bigcup_{n=1}^{\infty} f^{-1}(B_n)$ is not closed in X.

EXERCISE 2.2.4. Let $X = A \cup B$ and $f\colon X \to Y$ such that both $f|_A$ and $f|_B$ are continuous. Detect the fallacy in the following argument:

> Let $V \subset Y$ be an open set. Then its pre-image $f^{-1}(V) = f^{-1}(V) \cap (A \cup B) = (f|_A)^{-1}(V) \cup (f|_B)^{-1}(V)$, which is a union of two open sets because both $f|_A$ and $f|_B$ are continuous. Thus, f is continuous.

Give an example that f is not continuous. Show that if in addition both $A, B \subset X$ are open then f is continuous. Finally, would similar conclusion hold if both A, B are closed?

There is another property that looks like continuity; however, it is different.

DEFINITION 2.13. A mapping $f\colon X \to Y$ is called an *open mapping* if for any open set U in X, its image $f(U)$ is open in Y.

Clearly, since $f(X \setminus A) \neq Y \setminus f(A)$, we cannot conclude that $f(A)$ is closed if $A \subset X$ is so.

EXERCISE 2.2.5. Construct some examples of open mappings $f\colon X \to Y$ for $X, Y \subset \mathbb{R}$.

Exercise 2.2.6. Give examples of the following mappings $f\colon X \to Y$.

(1) It is both open and continuous.
(2) It is continuous but not open.
(3) It is open but not continuous.

Definition 2.14. A mapping $f\colon X \to Y$ is called a *homeomorphism* if it is a bijection and it is both open and continuous. Equivalently, its inverse mapping f^{-1} is also continuous. The spaces X and Y are *homeomorphic* or *topologically the same*. A mapping $f : (X, d_X) \to (Y, d_Y)$ is called an *isometry* if for all $x_1, x_2 \in X$, $d_X(x_1, x_2) = d_Y(f(x_1), f(x_2))$.

2.3. Sequence

The concept of sequence is tremendously important in analysis of $\mathbb{R}$ or $\mathbb{R}^n$ or $\mathbb{C}$. The underlying reason is because each of them is a metric space. In fact, sequences play a particular role in certain topological spaces, satisfying so-called first countability. Such topological spaces include metric spaces.

A sequence in a space X is a mapping from $\mathbb{N}$ to X, usually denoted by $(x_n)_{n\in\mathbb{N}}$. By mimicking the definition of convergence in $\mathbb{R}^n$, we may say that a sequence $(x_n)_{n\in\mathbb{N}}$ converges to $x \in X$ if for all $\varepsilon > 0$, there is an integer $N \in \mathbb{N}$ such that for all $n \geq N$, $d(x_n, x) < \varepsilon$, i.e., $x_n \in B(x, \varepsilon)$. Again, for smooth migration to general topological spaces, we will take the following equivalent definition.

Definition 2.15. A sequence $(x_n)_{n\in\mathbb{N}}$ *converges* to $x \in X$, denoted $x_n \to x$, if for all neighborhood $U \in \mathfrak{T}$ of x, there is an integer $N \in \mathbb{N}$ such that for every $n \geq N$, $x_n \in U$. The point $x \in X$ is called the *limit* of the sequence.

Exercise 2.3.1. A sequence $x_n \to x$ in a metric space (X, d) if and only if $\lim_{n\to\infty} d(x_n, x) = 0$ in $\mathbb{R}$.

One should develop a good habit of clarify the metric when the convergence of sequence is discussed. The following exercise shows that it may be confusing if the context is not clear.

Example 2.16. Let $X = \mathcal{C}([a, b], \mathbb{R})$ be the set of continuous functions on a closed interval $[a, b]$. Recall that we have the metrics

$$d_1(f, g) = \int_a^b |f(t) - g(t)|\, dt, \qquad d_\infty(f, g) = \sup_{t\in[a,b]} |f(t) - g(t)| \,.$$

In this case, there are three concepts of sequence convergence, namely, pointwisely, or under d_1, or under d_∞.

EXERCISE 2.3.2. In the above example of (X, d_1) and (X, d_∞), let $\mathbf{0}$ be the constant zero function.

(1) Prove that any convergent sequence under d_∞ must converge under d_1.
(2) Find a sequence $f_n \in X$ such that $d(f_n, \mathbf{0}) = 1/n$ but f_n does not converge *pointwisely*.

EXERCISE 2.3.3. (1) Let (X, d) be a metric space with sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ such that $a_n \to a$ and $b_n \to b$. Show that $d(a_n, b_n) \to d(a, b)$ in $\mathbb{R}$.
(2) Let (X, d) and (X, d^*) be metric spaces where $d^* = d/(1+d)$. Show that $x_n \to x$ in (X, d) if and only if $x_n \to x$ in (X, d^*).
(3) Let (X, d_X) and (Y, d_Y) be metric spaces and $X \times Y$ is given any product metric. Prove that a sequence $(x_n, y_n) \to (x, y)$ in $X \times Y$ if and only if $x_n \to x$ in X and $y_n \to y$ in Y.

PROPOSITION 2.17. *If a sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) converges, then its limit is unique.*

PROOF. Suppose there are $x, y \in X$ such that $x_n \to x$ and $x_n \to y$. Then take arbitrary $\varepsilon > 0$, and consider $B(x, \varepsilon/2)$ and $B(y, \varepsilon/2)$. Since x_n converges to both x and y, there exists $N \in \mathbb{N}$ such that for every $n \geq N$, we have both $x_n \in B(x, \varepsilon/2)$ and $x_n \in B(y, \varepsilon/2)$. By triangle inequality, we have $d(x, y) < \varepsilon$. Thus, we have shown for arbitrary $\varepsilon > 0$, $d(x, y) < \varepsilon$. Consequently, $d(x, y) = 0$ and $x = y$. $\square$

Note that in this proof, using the metric d, we have constructed two neighborhoods U_x of x and U_y of y such that $U_x \cap U_y = \emptyset$. In a metric space, this can always be done for every $x \neq y$. Equivalently, we may say that if each pair of neighborhoods, U_x of x and U_y of y, intersect each other, then $x = y$. This is a major property true for metric spaces but not necessary for general topological spaces. This property is the foundation of uniqueness of limit.

The following tells us the relationship between limit of a sequence and closure. We deliberately write it in two statements to emphasize that one part of it requires the metric and the other does not.

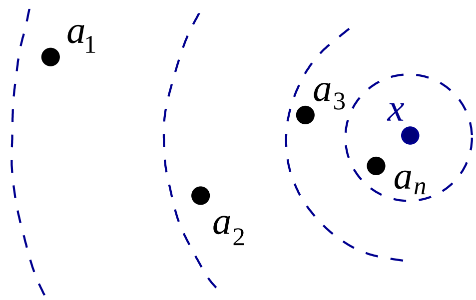
PROPOSITION 2.18. *Let X be a space with $A \subset X$ and $x \in X$.*

- *If there is a sequence $(a_n)_{n\in\mathbb{N}}$ in A and $a_n \to x$, then $x \in \overline{A}$.*
- *If $x \in \overline{A}$ and (X, d) is a metric space, then there is a sequence $(a_n)_{n\in\mathbb{N}}$ in A and $a_n \to x$.*

In the proof of the first, we will only use the concept of neighborhood instead of ball. This is because the first statement is independent of metric.

PROOF. In order to prove $x \in \overline{A}$, we need to show $U \cap A \neq \emptyset$ for every neighborhood $U \in \mathfrak{T}$ of x. Take any arbitrary $U \in \mathfrak{T}$ with $x \in U$. By definition of $a_n \to x$, there exists an integer $N \in \mathbb{N}$ such that whenever $n \geq N$, $a_n \in U$. Clearly $a_n \in U \cap A$.

For the second statement, by definition of $x \in \overline{A}$, it is known that one may pick a point $a \in A \cap U$ whenever $x \in U \in \mathfrak{T}$. So, if there is countably many neighborhoods of x "shrinking" down to x, then we have a sequence.



With a given metric, pick $a_n \in B(x, 1/n) \cap A$. Then $a_n \to x$. □

EXERCISE 2.3.4. In a metric space (X, d), is it true that

$$\overline{B(x_0, \varepsilon)} = \{\, x \in X : d(x, x_0) \leq \varepsilon \,\}?$$

COROLLARY 2.19. *Let (X, d) be a metric space. $F \subset X$ is closed, i.e., $\overline{F} = F$, if and only if every convergent sequence in F must have its limit also in F.*

PROOF. Assume that F is closed and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in F with $x_n \to x$. Using the first statement above, we have $x \in \overline{F} = F$.

To prove the "if" part, let $x \in \overline{F}$. Now, use the fact that (X, d) is a metric space and the second statement above, we can construct a sequence in F converging to x. By assumption, $x \in F$. Thus, $\overline{F} \subset F$ and F is closed. □

EXERCISE 2.3.5. Let d and ρ are two metrics on the same set. Prove that $x_n \to x$ wrt d if and only if $x_n \to x$ wrt ρ provided either there exists fixed constants $C_1 < C_2$ such that $C_1\,\rho \le d \le C_2\,\rho$; or the topologies of d and ρ are the same.

EXERCISE 2.3.6. In a metric space, it is true that if a sequence $x_n \to x$ then every subsequence of (x_n) converges to x; and conversely, if every convergent subsequence has limit x, then $x_n \to x$. Try to formulate the proof (from the known one in $\mathbb{R}^n$). Ask yourself whether the proof requires a metric or simply open balls.

2.4. Complete Metric Space

The standard Euclidean space is an important example of metric space. In addition, $\mathbb{R}^n$ has a property that makes it different from $\mathbb{Q}^n$ analytically. It worths particular attention and becomes an important type of metric space.

DEFINITION 2.20. A sequence $(x_n)_{n\in\mathbb{N}}$ in a metric space (X, d) is *Cauchy* if for every $\varepsilon > 0$, there exists an integer $N \in \mathbb{Z}$ such that if $m, n \ge N$, $d(x_m, x_n) < \varepsilon$.

Obviously, this definition requires a metric and it is a very familiar concept in $\mathbb{R}^n$. Thus, a natural question is how it is related to convergence of a sequence.

EXERCISE 2.4.1. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence such that the set $\{\, x_n : n \in \mathbb{N} \,\}$ has a cluster point. What can you conclude about the sequence?

Using the triangle inequality and the same argument as in $\mathbb{R}^n$, it can be easily proved that a convergent sequence in a metric space is always Cauchy. The following example shows that the converse is not true.

EXAMPLE 2.21. Let $X = \mathbb{R}^2$ with the standard metric and $x_n = \left(\frac{1}{n}, 0\right)$. Then $(x_n)_{n\in\mathbb{N}}$ is Cauchy and convergent.

However, on $A = \mathbb{R}^2 \setminus \{(0,0)\}$ with the standard metric. The same $(x_n)_{n\in\mathbb{N}}$ is still Cauchy but not convergent.

DEFINITION 2.22. A metric space is *complete* if every Cauchy sequence is convergent.

EXERCISE 2.4.2. If both X and Y are complete metric space, is the product metric space $X \times Y$ complete?

EXERCISE 2.4.3. Let $(\mathcal{B}, d_\infty)$ be the metric space of bounded functions on the interval $[a, b]$; $(\mathcal{C}, d_1)$ be the one of continuous functions, where

$$d_\infty(f, g) = \sup_{t\in[a,b]} \|f(t) - g(t)\| , \qquad d_1(f, g) = \int_a^b |f(t) - g(t)| \, dt .$$

(1) Show that $(\mathcal{B}, d_\infty)$ is a complete metric space. *Hint.* Use the completeness of $\mathbb{R}$ to get a pointwise limit first.

From the proof, one should see that the functions need not be real-valued and the domain need not be $[a, b]$.

(2) Show that $(\mathcal{C}, d_1)$ is not complete. *Hint.* Consider functions that are mostly 0 or 1 on $[a, b]$.

The above Example 2.21 of the puncture plane leads us to the following discussion about *metric subspace*. Given a metric space (X, d) and $A \subset X$, one may define an *induced metric* d_A on A simply by

$$d_A(a_1, a_2) = d(a_1, a_2) , \qquad \text{seeing } a_1, a_2 \in X.$$

We will not spend too much effort on discussing metric subspace here. Most of the properties will be discussed later in the context of topological space. Here, we will only focus on the properties about sequences.

Let (A, d_A) be a metric subspace of the metric space (X, d) and $(a_n)_{n\in\mathbb{N}}$ is a sequence in A. The following is trivial.

PROPOSITION 2.23.
- *The sequence $(a_n)_{n\in\mathbb{N}}$ is Cauchy in (X, d) if and only if it is Cauchy in (A, d_A).*
- *If $(a_n)_{n\in\mathbb{N}}$ converges in (A, d_A), then it is convergent in (X, d).*

Since the converse of the second statement is not true, the completeness of (X, d) does not imply that of (A, d_A). Nevertheless, we have

PROPOSITION 2.24. *Let A be a subset in a complete metric space (X, d). Then (A, d_A) is complete if and only if A is closed in X.*

From this result, $\mathcal{C}$ in Exercise 2.4.3 is complete.

PROOF. Assume that (A, d_A) is complete and let $x \in \overline{A}$. Then there exists a sequence $a_n \in A$ with $a_n \to x$ in (X, d). Thus, $(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in (X, d), and also in (A, d_A). By assumption, it converges in (A, d_A), say $a_n \to$

$a \in A$ wrt d_A, and also wrt d. By uniqueness of limit in (X,d), $a = x$ and so $x \in A$.

Assume that A is closed in X and let $(a_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in (A, d_A). Then it is also Cauchy in (X,d) and it converges in (X,d), say $a_n \to x \in X$. Since $a_n \in A$, $x \in \overline{A}$. By assumption, A is closed and so $x \in A$ and hence the sequence converges in (A, d_A). $\square$

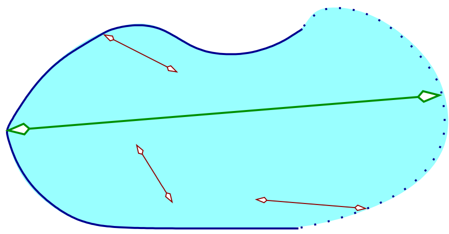
The completeness of $\mathbb{R}$ is a very important feature of the real line. It leads to many useful results. One of them is the so-called Nested Interval Theorem.

Exercise 2.4.4. Recall the statement of the Nested Interval Theorem and how it is used, e.g., to proving Intermediate Value Theorem.

In a complete metric space, an analogous theorem is expected. Since there may not be an order, we need some other notions instead of “shrinking” intervals.

On a metric space (X,d), for a subset $A \subset X$, define the *diameter* of A by

$$\operatorname{diam}(A) = \sup\{\, d(a_1,a_2) \;:\; a_1,a_2 \in A \,\}\ .$$



Exercise 2.4.5. Show that if A is closed in $\mathbb{R}^n$ and $\operatorname{diam}(A) < \infty$, then there are two points $a_1, a_2 \in A$ such that $d(a_1,a_2) = \operatorname{diam}(A)$, i.e., the diameter can be attained. *Remark.* The same statement is not true for metric space. In that case, we need the set A to be compact, which is a concept to be discussed later.

Theorem 2.25 (Cantor Intersection Theorem). *Let (X,d) be a complete metric space; $F_n \subset X$ be nonempty closed sets such that*

$$F_n \supset F_{n+1} \quad \text{for all } n; \qquad \operatorname{diam}(F_n) \to 0 \quad \text{as } n \to \infty.$$

Then $\bigcap_{n=1}^{\infty} F_n$ is a singleton set.

Proof. In order to prove the existence of a point in $\bigcap_{n=1}^{\infty} F_n$, knowing that X is complete, the natural idea is to construct a Cauchy sequence and obtain its limit. Just pick $x_n \in F_n$. The following exercises will yield the proof.

Exercise 2.4.6. Use the fact that $\operatorname{diam}(F_n) \to 0$ to show that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

Thus, by completeness of X, $x_n \to x \in X$. Observe that for any $p \in \mathbb{N}$, $(x_n)_{n=p}^{\infty}$ is a sequence in F_p convergent to x. Therefore, $x \in \overline{F}_p = F_p$ for any $p \in \mathbb{N}$.

Exercise 2.4.7. Use $\operatorname{diam}(F_n) \to 0$ again to conclude that $\bigcap_{n=1}^{\infty} F_n = \{x\}$.

This completes the proof. □

Exercise 2.4.8. In the Cantor Intersection Theorem above, if it is not given that $\operatorname{diam}(F_n) \to 0$, can we prove that $\bigcap_{n=1}^{\infty} F_n$ is non-empty? In addtion, give an example to illustrate that one must need the condition on "each F_n is closed".

Exercise 2.4.9. Let (X, d) be a metric space. For Cauchy sequences $(x_n), (y_n)$ in X, define an equivalence relation $(x_n) \sim (y_n)$ if $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$ and denote the equivalence class of (x_n) by $\mathbf{x}$. Let $\hat{X}$ be the set of such equivalence classes and $\hat{d}(\mathbf{x}, \mathbf{y}) = \lim\limits_{n\to\infty} d(x_n, y_n)$.

(1) Show that $(\hat{X}, \hat{d})$ is a complete metric space.
(2) Show that there is a natural continuous one-to-one mapping $j : X \to \hat{X}$ such that $j(X)$ is isometric to X.
(3) Show that if X itself is complete, then X and $\hat{X}$ are isometric.

2.5. Continuity and Sequences

As it is mentioned in Proposition 2.12, there are several equivalent ways of describing the continuity of a mapping between metric spaces. Similar to mappings between Euclidean spaces, continuity of the mapping is related to limit of sequences. In the following, two statements are given to emphasize that one is independent of metric.

Proposition 2.26. (1) *If $f : X \to Y$ is continuous, then for each sequence $(x_n)_{n\in\mathbb{N}}$ in X with $x_n \to x$, $f(x_n) \to f(x)$ in Y.*
(2) *If (X, d) is metric space and every sequence $(x_n)_{n\in\mathbb{N}}$ in X with $x_n \to x$ must also have $f(x_n) \to f(x)$ in Y, then f is continuous.*

Proof. For statement (1), let $V \in \mathfrak{T}_Y$ with $f(x) \in V$. By continuity of f, $f^{-1}(V)$ is a neighborhood of x in $\mathfrak{T}_X$. Since $x_n \to x$, there exist $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \in f^{-1}(V)$ and hence $f(x_n) \in V$.

For statement (2), we will prove the contra-positive. Assume that f is not continuous, i.e., there exists a $V \in \mathfrak{T}_Y$ such that $f^{-1}(V) \notin \mathfrak{T}_X$. Therefore, there is $x \in f^{-1}(V)$ which is not an interior point. As a consequence, every neighborhood $U \in \mathfrak{T}_X$ of x contains a point outside $f^{-1}(V)$. We will use this fact to construct a sequence $x_n \to x$ and here is exactly why a metric d on X is needed.

For each $0 < n \in \mathbb{Z}$, $B(x, 1/n)$ is a neighborhood of x. So, there exists $x_n \in B(x, 1/n) \setminus f^{-1}(V)$. Clearly, $x_n \to x$. However, $x_n \notin f^{-1}(V)$ and so $f(x_n) \notin V$. Hence, $f(x_n)$ does not converge to $f(x)$. $\square$

Proposition 2.27. *Let X, Y be metric spaces and $f\colon X \to Y$ be a mapping. Then f is continuous if and only if for each set $A \subset X$, $f\left(\overline{A}\right) \subset \overline{f(A)}$.*

Proof. If f is given to be continuous, the proof is straight forward without using properties of metric. It has been done in Exercise 2.2.2.

We will use Proposition 2.26 to the contra-positive of the "if" part. Suppose that f is not continuous at $x_0 \in X$ and there is a sequence $x_n \to x_0$ but $f(x_n) \not\to f(x_0)$. Let us consider the intuitive idea and though it is not accurate. Naturally, take $A = \{\, x_n \;:\; n \in \mathbb{N} \,\}$. Then $f(A) = \{\, f(x_n) \;:\; n \in \mathbb{N} \,\}$ and $\overline{A} = A \cup \{\, x_0 \,\}$. We hope that $f(x_0) \notin \overline{f(A)}$ because $f(x_n) \not\to f(x_0)$. Obviously, such a conclusion is not possible because it may happen that $f(x_k) = f(x_0)$ for some $k \in \mathbb{N}$. Even that does not occur, there may be a subsequence $f(x_{n_k}) \to f(x_0)$ as $k \to \infty$ and so $f(x_0) \in \overline{f(A)}$. Thus, one has to construct A in a clever way to avoid these two situations. The details is left as an exercise below. $\square$

Exercise 2.5.1. Justify that there exists $A \subset X$ with $f\left(\overline{A}\right) \not\subset \overline{f(A)}$. In addition, consider whether the metrics on X or Y is necessary in the proof.

Studies of approximation are important in many branches of mathematics. Examples are abundant and let us consider a few. The Weierstrass Approximation Theorem is roughly saying that continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$ can be "approximated" by polynomials. This can be rephrased in the language of topology. Let X be the set of continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$ with a certain topology. Let A be the subset of all polynomials. Then the Weierstrass Approximation Theorem is essentially $\overline{A} = X$. Such an A is called dense in X, which will be further discussed in the next section.

Analogous situation occurs in the study of knots (strings tied into a knot with the two ends glued up). Then any knot can be approximated by polygonal knots, which is formed by finitely many segments of straight lines. This also can be phrased in topology as the set of polygonal knots is dense in the space of all knots.

Clearly, certain calculations are easier to be done on polynomials or polygonal knots. Moreover, these calculations may be indeed well-defined for general continuous functions or knots, but it is harder to calculate. It is certainly hoped that those results of easier calculations determine the harder ones. In mathematical terms, we already have $\overline{A} = X$ and a calculation on A, $f\mid_A\colon A \to \mathbb{R}$ is known. Can we comfortably say that a unique $f\colon X \to \mathbb{R}$ is behind? Luckily, the answer is mostly yes.

THEOREM 2.28. *Let $f, g\colon X \to Y$ be continuous mappings and $\overline{A} = X$. If $f\mid_A \equiv g\mid_A$ on A then $f \equiv g$ on X.*

PROOF I. Take arbitrary $x \in X$ and try to show $f(x) = g(x)$. Since $x \in X = \overline{A}$, there is a sequence $(a_n)_{n\in\mathbb{N}}$ in A with $a_n \to x$. Note that this requires X to be a metric space. By continuity of both f and g, we have

$$f(a_n) \to f(x) \qquad \text{and} \qquad g(a_n) \to g(x)\,.$$

By assumption that $f\mid_A \equiv g\mid_A$, $f(a_n) = g(a_n)$ for all $n \in \mathbb{N}$. Thus, by uniqueness of limit (this requires certain condition on Y, e.g., Y is metric space), we have $f(x) = g(x)$. □

The above theorem only guarantees that the uniqueness of the extension if it exists. It is not known whether such extension exists. A related result on this direction will be given later in Theorem 2.42 in Section 2.7.

In fact, the theorem is true in a more general situation, in which the following proof works.

PROOF II. Suppose otherwise, then there exists $x \in X$ such that $y_1 = f(x) \neq g(x) = y_2$. Choose neighborhoods $V_1 \in \mathfrak{T}_Y$ of y_1 and $V_2 \in \mathfrak{T}_Y$ of y_2 with $V_1 \cap V_2 = \emptyset$. This is possible if Y is a metric space and V_1, V_2 are open balls with radius $\frac{1}{3}d_Y(y_1, y_2)$.

By continuity of f and g, $f^{-1}(V_1)$ and $g^{-1}(V_2)$ are neighborhoods of x in X, and so is $f^{-1}(V_1) \cap g^{-1}(V_2)$. Since $x \in X = \overline{A}$, there is $a \in A \cap f^{-1}(V_1) \cap g^{-1}(V_2)$. Thus, $f(a) = g(a) \in V_1 \cap V_2$ which contradicts $V_1 \cap V_2 = \emptyset$. □

Note that in second proof, there is not any condition on X (only neighborhood argument is used) and only a mild condition on Y to get the disjoint neighborhoods at $y_1 \neq y_2$. The space Y is called Hausdorff or T_2 for the existence of such disjoint neighborhoods. Such technique had been used in proving uniqueness of limit of sequence. We will discuss more about this property in the future.

In many branches of mathematics or other discipline, studying the fixed point of a mapping is a key issue. The existence of a fixed point could lead to theoretical development or a practical method. This study is often done in the context of a metric space. Let $f\colon X \to X$ be a mapping into a space itself, a point $x_* \in X$ is called a *fixed point* of f if $f(x_*) = x_*$.

In many cases, there is additional requirement on the space X. For example, the Newton's Method, which is used to find the root of the equation $g(x) = 0$ in $\mathbb{R}$. Essentially, it transforms the problem to find a fixed point for $f(x) = x - \dfrac{g(x)}{g'(x)}$. The fixed point is guaranteed to exist if $g'(x) \neq 0$. The other example is the Inverse Function Theorem for a function $h\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$. There are several proofs, one of them makes use of the existence of a fixed point. Here we will introduce a situation where a fixed point must exist.

Definition 2.29. A mapping $f\colon (X, d_X) \to (Y, d_Y)$ is called a *contraction* or *contraction mapping* if there exists a fixed $0 < \alpha < 1$ such that for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) < \alpha\, d(x_1, x_2)$.

Exercise 2.5.2. A mapping is called *Lipschitz* is there exists a fixed $0 < C$ (not necessarily < 1) such that for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) < C\, d_X(x_1, x_2)$. Prove that a Lipschitz mapping (and hence a contraction) is continuous.

Theorem 2.30 (Banach Fixed Point Theorem). *Let $f\colon (X, d) \to (X, d)$ be a contraction mapping on a complete metric space (X, d). Then it has a fixed point.*

The main idea of getting such a fixed point in a complete metric space is naturally by constructing a Cauchy sequence and considering its limit.

Proof. Start at any $x_0 \in X$ and define a sequence recursively by

$$x_{n+1} = f(x_n), \quad 0 \le n \in \mathbb{Z}\,.$$

First, we will try to show that $(x_n)_{n\in\mathbb{N}}$ is Cauchy. As f is a contraction,

$$\begin{aligned} d(x_m, x_{m+p}) &= d(f(x_{m-1}, f(x_{m+p-1})) < \alpha\, d(x_{m-1}, x_{m+p-1}) \\ &< \cdots\cdots \quad < \alpha^m\, d(x_0, x_p)\,. \end{aligned}$$

Unfortunately, the last term depends on the value of p even though α^m can be very small. Therefore, the method is changed a little with triangle inequality.

$$\begin{aligned} d(x_m, x_{m+p}) &\le d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{m+p-1}, x_{m+p}) \\ &< \alpha^m\, d(x_0, x_1) + \alpha^{m+1}\, d(x_0, x_1) + \cdots + \alpha^{m+p-1}\, d(x_0, x_1) \\ &= \left(\alpha^m + \alpha^{m+1} + \cdots + \alpha^{m+p-1}\right) d(x_0, x_1)\,. \end{aligned}$$

Now, by the convergence of geometric series $\sum_{n=1}^{\infty} \alpha^n$, for each $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for all $m \ge N$, $\left(\alpha^m + \alpha^{m+1} + \cdots + \alpha^{m+p-1}\right) d(x_0, x_1) < \varepsilon$.

Thus, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and $x_n \to x_*$ in the complete metric space X. By continuity of f, we have $f(x_*) = x_*$. □

Exercise 2.5.3. Is this fixed point theorem still valid if the mapping is only "partially contraction", i.e., $\alpha \le 1$?

2.6. Baire and Countability

In this section, we slightly touch on the theory of Baire. Most of the spaces studied in analysis, geometry, and topology satisfy certain conditions of Baire. There is a simple but powerful classification of spaces, which is somewhat related to countability.

Definition 2.31. Let X be a space. A subset $A \subset X$ is called *dense* in X if $\mathrm{Cl}(A) = \overline{A} = X$. A subset $N \subset X$ is called *nowhere dense* in X if $\mathrm{Int}\,(\mathrm{Cl}(N)) = \emptyset$.

These two concepts are in a way dichotomized, but not logically negation to each other.

Example 2.32. In standard $\mathbb{R}$, both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense subsets; while $\mathbb{Z}$ is nowhere dense. However, $\mathrm{Int}(\mathbb{Q}) = \emptyset = \mathrm{Int}(\mathbb{R} \setminus \mathbb{Q})$, as well as $\mathrm{Int}(\mathbb{Z}) = \emptyset$. This shows why we need to take closure first in the definition of nowhere dense.

By explicitly writing out the definition, $A \subset X$ is dense if and only if for each $x \in X$ and each neighborhood $U \in \mathfrak{T}$ with $x \in X$, $U \cap A \neq \emptyset$. In fact, it can be further simplified to the commonly used statement "for each $U \in \mathfrak{T}$, $U \cap A \neq \emptyset$".

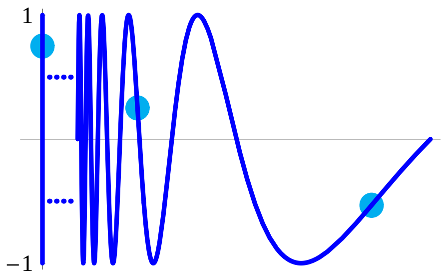
Exercise 2.6.1. Show that $A \subset X$ is dense if and only if the only open set contained in $X \setminus A$ is $\emptyset$; and if and only if the only closed set containing A is X.

Exercise 2.6.2.
(1) Let $N \subset X$ be nowhere dense. Show that every open set $U \subset X$ contains an open set $V \subset U$ such that $V \cap N = \emptyset$.

(2) If $A \subset X$ is dense, what can you say about $X \setminus A$? Similarly, if $N \subset X$ is nowhere dense, what about $X \setminus N$?
(3) Find all the dense sets and nowhere dense sets in a discrete space.

Example 2.33. Many nowhere dense subsets come from "lower dimension". For example, a straight line in $\mathbb{R}^2$ is nowhere dense. A slight variation is the image of a parametrized curve, $\{\,(x(t), y(t)) \;:\; t \in [a,b]\,\}$, is *usually* a "nice" curve; if x, y are differentiable functions satisfying some technical conditions. Most of them are nowhere dense but it is not always.

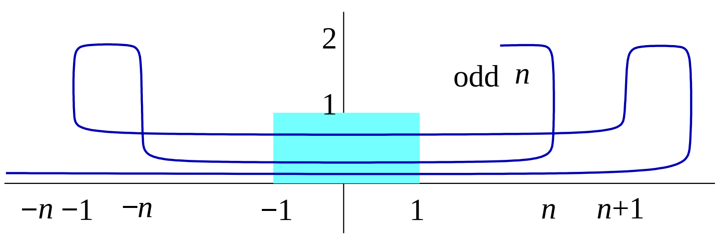
Let $N = \left\{\, \left(x, \sin\frac{1}{x}\right) \in \mathbb{R}^2 : x > 0 \,\right\}$ be a subset of the standard $\mathbb{R}^2$. It can be seen that $\overline{N} = N \cup \left\{\, (0,y) \in \mathbb{R}^2 : y \in [-1,1] \,\right\}$.



Then, N is nowhere dense because $\operatorname{Int}\left(\overline{N}\right) = \emptyset$. In fact, for any point in $\overline{N}$, any neighborhood of it intersects $\mathbb{R}^2 \setminus N$. This is illustrated by the above picture.

The space filling curve has its closure equal to the whole square, which has nonempty interior. However, it is the result of an infinite process so it is difficult to parametrize the curve by differentiable functions. The following example may be easier to visualize.

Example 2.34. Let q_n be rational numbers in the interval $[0,1]$. It is easy to construct a U-shape differentiable curve containing the segments from the point $(-n,2)$ to $(-n,q_n)$ to (n,q_n) to $(n,2)$ by smoothing the corners. Each one will have a horizontal segment at the height of q_n. Then suitably join these U-shapes curves with smooth corners as shown in the picture below. Then the closure of the curve contains the rectangle $[-1,1] \times [0,1]$.



EXAMPLE 2.35. There is a famous important theorem in differential topology. In a simple form, it is about an infinitely differentiable function $f\colon [a,b] \to \mathbb{R}$. We may ask the following two questions.

- Can it happen that the set of critical points $\{ x \in [a,b] : f'(x) = 0 \}$ is dense? The answer is yes; for example, the constant function, and many others.
- What about the image set $\{ f(x) \in \mathbb{R} : f'(x) = 0,\ x \in [a,b] \}$? It can never be dense. In fact, it must be nowhere dense.

Let $N \subset X$ be a nowhere dense subset. Then $\mathrm{Int}(\overline{N}) = \emptyset$ and so not only $X \setminus \overline{N}$ is nonempty, indeed, for any open set U, $U \setminus \overline{N}$ must be nonempty. This will be a property that we will use later.

DEFINITION 2.36. A subset $A \subset X$ is of *first category*, Cat-I, if there exists countably many nowhere dense sets N_k such that $A = \bigcup_{k=1}^{\infty} N_k$. Otherwise, it is of *second category*, Cat-II.

EXAMPLE 2.37.
(1) In the standard $\mathbb{R}$, $\mathbb{Q}$ is dense but it is trivially of Cat-I. However, $\mathbb{R}\setminus\mathbb{Q}$ is of Cat-II. This may be seen from Theorem 2.39 below.
(2) In the discrete space X, any nonempty subset is of Cat-II. The reason is that the only nowhere dense set is the empty set. In other words, if $\emptyset \neq A \subset X$, then $\overline{A} = A$ and $\mathrm{Int}(\mathrm{Cl}(A)) = A \neq \emptyset$.

EXERCISE 2.6.3.
(1) Show that $\mathbb{Z}$ with the standard metric $d(m,n) = |m-n|$ is of second category. Note: this does not contradict that $\mathbb{Z}$ is nowhere dense in $\mathbb{R}$.
(2) Show that if A is of Cat-I, then $B \subset A$ will also be of Cat-I.
(3) Show that if N_1, N_2 are nowhere dense sets, then so is $N_1 \cup N_2$.

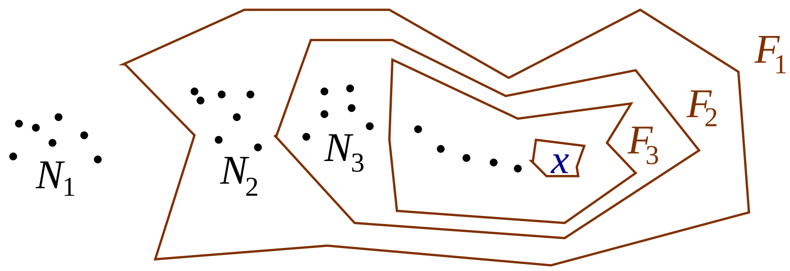
EXAMPLE 2.38. Assuming the fact that standard $\mathbb{R}^n$ is of Cat-II (see below), then one may conclude that any open ball in $\mathbb{R}^n$ is of Cat-II. In turns, every set in $\mathbb{R}^n$ with non-empty interior is of Cat-II.

There are deeper discussions of Baire's theory. However, we will only stop at the following conclusion about complete metric spaces.

THEOREM 2.39 (Baire Category Theorem). *A complete metric space is of second category.*

The idea is to consider nonempty closed sets $F_n \subset X \setminus (\bigcup_{k=1}^n N_k)$. Then try to establish that

$$\emptyset \neq \bigcap_{n=1}^{\infty} F_n \subset X \setminus \left(\bigcup_{k=1}^{\infty} N_k \right) .$$



From the Cantor Intersection Theorem, we need that $F_n \supset F_{n+1}$ and $\operatorname{diam}(F_n) \to 0$. A natural choice of closed sets is $F_n = \{ x \in X \ : \ d(x, x_n) \le r_n \}$. Thus, the key is to pick the suitable points x_n and radii r_n. Indeed, we will prove the following equivalent statement.

Proposition 2.40. *Let X be a complete metric space and N_k, $k \in \mathbb{N}$, be a sequence of nowhere dense subsets in X. Then $X \neq \bigcup_{k=1}^{\infty} N_k$.*

Proof. Observe that $\operatorname{Int}(\operatorname{Cl}(N_1)) = \emptyset$. So, $\operatorname{Cl}(N_1) \neq X$. In other words, $X \setminus \operatorname{Cl}(N_1)$ is a nonempty open set. There exists a neighborhood $B(x_1, 2r_1)$ lying in $X \setminus \operatorname{Cl}(N_1)$ for some point $x_1 \in X$ and a radius $r_1 > 0$. Then

$$F_1 = \{ x \in X : d(x, x_1) \le r_1 \} \subset X \setminus \operatorname{Cl}(N_1) \subset X \setminus N_1 .$$

Now, since $\operatorname{Int}(\operatorname{Cl}(N_2)) = \emptyset$, so $B(x_1, r_1) \not\subset \operatorname{Cl}(N_2)$; and one may have $x_2 \in \operatorname{Int}(F_1)$ and a radius $0 < r_2 < r_1$ such that $B(x_2, 2r_2) \subset B(x_1, r_1) \setminus \operatorname{Cl}(N_2)$. Then

$$F_2 = \{ x \in X : d(x, x_2) \le r_2 \} \subset F_1 \setminus \operatorname{Cl}(N_2) \subset X \setminus (\operatorname{Cl}(N_1) \cup \operatorname{Cl}(N_2)) \subset X \setminus (N_1 \cup N_2) .$$

This argument can be continued and thus there exists nonempty closed sets F_k satisfying

$$F_1 \supset F_2 \supset \cdots \supset F_n, \qquad \operatorname{diam}(F_n) < r_1/2^n, \qquad F_n \subset X \setminus \left(\bigcup_{k=1}^{n} N_k \right) .$$

By the Cantor Intersection Theorem, $\bigcap_{n \in \mathbb{N}} F_n \neq \emptyset$ and the result follows. □

Exercise 2.6.4.
(1) Are there statements about first and second category of $X \times Y$ with reference to the categories of X and Y?
(2) Show that $A \subset X$ is open dense if and only if $X \setminus A$ is closed nowhere dense. Give counter examples if the open/closed condition is dropped.
(3) Let $f : X \to Y$ be a continuous mapping.

(a) If $D \subset X$ is dense, is $f(D) \subset Y$ dense?
(b) If $N \subset X$ is nowhere dense, is $f(N) \subset Y$ nowhere dense?
(c) What about pre-images of a dense set and a nowhere dense set?
(d) What can you conclude about image or pre-image of a set of first or second category?

(4) Prove that if $A \subset \mathbb{R}^n$ with measure$(A) \neq 0$, then A is of Cat-II.

2.7. Further Continuity

In the study of continuous functions, if the spaces involved are metric spaces, there are additional properties. Two of which are related to "uniformity". They are particularly important in analysis. The following is exactly the direct generalization from Euclidean spaces.

DEFINITION 2.41. A mapping $f\colon (X, d_X) \to (Y, d_Y)$ is *uniformly continuous* if for every $\varepsilon > 0$, there exists $\delta > 0$ (depending only on ε) such that for all $x_1, x_2 \in X$ with $d_X(x_1, x_2) < \delta$, $d_Y(f(x_1), f(x_2)) < \varepsilon$.

EXERCISE 2.7.1.

(1) Show that a Lipschitz mapping is uniformly continuous.
(2) Is the composition of two uniformly continuous mappings still uniformly continuous?
(3) Let (X, d) be a metric space. Let $f\colon X \to \mathbb{R}$ be given by $f(x) = d(x, x_0)$ for a fixed $x_0 \in X$. Is this function uniformly continuous?

   Also, $d\colon X \times X \to \mathbb{R}$ where $X \times X$ is given a product metric. Is it uniformly continuous?
(4) If $f\colon A \to Y$ is uniformly continuous and $(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in A, then $(\,f(a_n)\,)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. Give a counter example for f is only continuous.

Uniformity plays an important role in analysis and approximation. In a later chapter, we will establish the concept of compactness. Then, it is easy to see, similar to Euclidean case, that every continuous mapping from a compact space to a metric space is uniformly continuous. The following is a version about extension of a uniformly continuous mapping.

THEOREM 2.42. *Let $A \subset \overline{A} = X$ be given the induced metric from (X, d_X) and (Y, d_Y) be complete. If $f\colon A \to Y$ is uniformly continuous, then there exists* *a unique continuous extension* $\tilde{f}\colon X \to Y$ *such that* $\tilde{f}|_A \equiv f$*. Indeed,* $\tilde{f}$ *is uniformly continuous.*

One should be reminded of the difference between this theorem and Theorem 2.28, which is a uniqueness theorem and it starts with functions defined on X. However, in this theorem, the function f is only defined on a dense set A and one is resulted with getting an extension on X. So, this is an existence theorem, though the extension is unique by virtue of Theorem 2.28.

Before studying the proof, it would be beneficial to understand the theorem from another angle.

EXERCISE 2.7.2. Give an example that the theorem fails if f is only continuous. Also, give an example that the theorem fails if Y is not complete.

PROOF. Our aim is to define the value $\tilde{f}(x) \in Y$ for each $x \in X = \overline{A}$.

Since X is a metric space and $x \in \overline{A}$, there exists a sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \to x$. Correspondingly, we have $f(a_n) \in Y$ but $f(x)$ may not be defined because x may not be in A. The sequence (a_n) is Cauchy in A (because it converges in X). By uniform continuity of f (on the set A), $f(a_n)$ is a Cauchy sequence also. Therefore, $f(a_n)$ converges, say, to $y \in Y$. Define $y = \tilde{f}(x)$.

It is sufficient to show that the above $\tilde{f}(x)$ is well-defined. That means it is independent of choice of the sequence $(a_n)_{n\in\mathbb{N}}$. But, this will be seen shortly.

Let us temporarily assume $\tilde{f}$ depends on the choices of sequences. That is, for each $x \in X$, one particular sequence converging to x is chosen and we use these sequences to define $\tilde{f}$. In other words, we have $a_n(x) \in A$ with $a_n(x) \to x$ and $f(a_n(x)) \to \tilde{f}(x)$. We will first prove the uniform continuity of this $\tilde{f}$ and later prove its independence of choice of the sequences.

Given any $\varepsilon > 0$, by uniform continuity of f on A, there is a $\delta > 0$ such that whenever $a, a' \in A$ with $d_X(a, a') < 3\delta$, $d_Y(f(a), f(a')) < \varepsilon/3$. We will take this $\delta > 0$ which only depends on ε.

Let $x_1, x_2 \in X$ with $d_X(x_1, x_2) < \delta$. In the following, we are going to establish that $d_Y(\tilde{f}(x_1), \tilde{f}(x_2)) < \varepsilon$. Since $x_1 \in X = \overline{A}$, from the definition of $\tilde{f}$, we have chosen a sequence $a_n(x_1) \in A$ such that

$$a_n(x_1) \to x_1 \qquad \text{and} \qquad f(a_n(x_1)) \to \tilde{f}(x_1)\,.$$

Thus, for $\delta > 0$ and $\varepsilon/3 > 0$, by taking large enough $N \in \mathbb{N}$, we have $a_N(x_1)$ with $d_X(a_N(x_1), x_1) < \delta$ and $d_Y(f(a_N(x_1)), \tilde{f}(x_1)) < \varepsilon/3$. Similarly, we have $a_M(x_2)$ in the defining sequence of $\tilde{f}(x_2)$ with $d_X(a_M(x_2), x_2) < \delta$ and $d_Y(f(a_M(x_2)), \tilde{f}(x_2)) < \varepsilon/3$.

By triangle inequality on X, we have $d_X(a_N(x_1), a_M(x_2)) < 3\delta$, which leads to $d_Y(f(a_N(x_1)), f(a_M(x_2))) < \varepsilon/3$ by the choice of δ. Now, apply the triangle inequality on Y, we have $d_Y(\tilde{f}(x_1), \tilde{f}(x_2)) < \varepsilon$.

At this point, it is proved that this particular choice of the function $\tilde{f}$ on X is uniformly continuous, no matter how we choose the sequences $a_n(x) \to x$ to define $\tilde{f}(x)$. By Theorem 2.28, the continuous extension of f from A to $\overline{A} = X$ is unique. Thus, there is only one such $\tilde{f}$. Its uniqueness indeed guarantees its independent of choice of the sequences. □

Another important concept of uniformity is about convergence of a sequence of mappings. That usually guarantees the limiting mapping has the same analytical properties as the mappings in the sequence.

Definition 2.43. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of mappings $X \to (Y, d)$. It *converges uniformly* to a mapping $f\colon X \to Y$ if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ (depending only on ε) such that whenever $n \geq n$, for all $x \in X$, $d(f_n(x), f(x)) < \varepsilon$.

Equivalently, we may say for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that whenever $n \geq N$, $\sup_{x\in X} d(f_n(x), f(x)) < \varepsilon$.

One may have seen uniform convergence in studying sequence of functions on $\mathbb{R}$. The situation on metric spaces is very similar. The value of $N \in \mathbb{N}$ (for a fixed ε) can be seen as the speed of convergence. The larger the N, the slower the speed. In general, this speed depends on $x \in X$ and the difference between speeds at various x's can be very large. Uniform convergence means that there is a bounded difference, or there is a minimum speed.

Theorem 2.44. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of mappings $X \to (Y, d)$ uniformly converges to f. If every $f_n\colon X \to Y$ is continuous, then $f\colon X \to Y$ is also continuous.*

Remark . In words, we may say that continuity is carried to the uniform limit. In other context, when one can define differentiability or integrability, these properties can also be carried to uniform limit.

Proof. Let $V \in \mathfrak{T}_Y$ and we will prove $f^{-1}(V) \in \mathfrak{T}_X$. For this, we take arbitrary $x_* \in f^{-1}(V)$ and it is desired that x_* is an interior point of $f^{-1}(V)$.

Let $y = f(x_*) \in Y$ and $B(y, \varepsilon) \subset V$ for some $\varepsilon > 0$. First, let us take a smaller $W \subset V$ and try to get $x_* \in f_n^{-1}(W) \subset f^{-1}(V)$ for large n.

Indeed, $W = B(y, \varepsilon/2)$. Then for $\varepsilon/2 > 0$, choose $N \in \mathbb{N}$ such that for all $n \geq N$, $\sup_{x \in X} d(f_n(x), f(x)) < \varepsilon/2$. In particular, $d(f_n(x_*), f(x_*)) < \varepsilon/2$, i.e., $x_* \in f_n^{-1}(W)$. Let $x \in f_n^{-1}(W)$ where $n \geq N$. Then $d(f_n(x), y) < \varepsilon/2$ and $d(f_n(x), f(x)) < \varepsilon/2$. By triangle inequality, $f(x) \in B(y, \varepsilon) = V$, i.e., $x \in f^{-1}(V)$. Thus, we have shown that $x_* \in f_n^{-1}(W) \subset f^{-1}(V)$.

By continuity of each f_n for $n \geq N$, $f_n^{-1}(W)$ is open and x_* is an interior point of $f^{-1}(V)$. □

CHAPTER 3

Topological Spaces

In the previous chapter, one has already seen that in many discussions and deductions, metric is not always necessary. On the other hand, open sets or its complements closed sets are often used. It gives us an idea that something is more fundamental in the study of approximation and relative positions.

Historically, there had been several attempts. The most famous ones are neighborhood systems or convergence. At the end, people chose the elegant concept of topology. It is already enough to determine concepts of neighborhood and convergence; as well as continuity of mappings.

We will begin by introducing the definition and basics of topology in §3.1. Then in §3.2, fundamental building blocks of a topology called base and subbase are discussed. In §3.3, spaces with good property due to countability conditions are introduced.

3.1. Topology, Open and Closed

In the previous chapter, one defines open sets by interior points and then topology as the set of all open sets. Here, we will start with a collection of sets.

Definition 3.1. Let X be a nonempty set. A *topology* $\mathfrak{T}$ on X is a subset of the power set of X satisfying

- Both $\emptyset, X \in \mathfrak{T}$;
- For each family $\{U_\alpha\}_{\alpha\in I}$ of sets $U_\alpha \in \mathfrak{T}$, the union $\bigcup_{\alpha\in I} U_\alpha \in \mathfrak{T}$;
- For any finitely many $U_1, \ldots, U_n \in \mathfrak{T}$, the intersection $\bigcap_{k=1}^n U_k \in \mathfrak{T}$.

The pair $(X, \mathfrak{T})$ is called a *topological space*.

A set $G \subset X$ is called *open* in $(X, \mathfrak{T})$ if $G \in \mathfrak{T}$. A set $F \subset X$ is *closed* in $(X, \mathfrak{T})$ if $X \setminus F \in \mathfrak{T}$.

The two conditions about a topology is usually referred to as *closed under arbitrary union and finite intersection*. In fact, by accepting empty collection of sets, these two conditions will lead to $\emptyset, X \in \mathfrak{T}$. It is clear that they are also closed.

EXAMPLE 3.2.
(1) Let (X, d) be a metric space. If open balls, interior points, and open sets are defined as before, and $\mathfrak{T}$ is the set of open sets, then $\mathfrak{T}$ satisfies the two conditions and $(X, \mathfrak{T})$ is a topological space.
(2) Let X be a nonempty set and $\mathfrak{T} = \mathcal{P}(X)$, the power set of X. In other words, every subset is an open set. This is called the *discrete topology*. It is the same as the one given by the discrete metric.
(3) Let X be a nonempty set and $\mathfrak{T} = \{\emptyset, X\}$. The two conditions are obviously satisfied. This is called the *indiscrete topology*, another extreme from the discrete topology. One may expect that the indiscrete topology is not a metric topology, i.e., it does not come from a metric.
(4) On $\mathbb{R}$, one has the topology wrt the standard metric. It is called the *standard topology*, $\mathfrak{T}_{\text{std}}$. Let $\mathfrak{T} \overset{\text{def}}{:=} \{\emptyset, \mathbb{R}\} \cup \{(a, \infty) : a \in \mathbb{R}\}$. This $\mathfrak{T}$ is also a topology on $\mathbb{R}$ and $\mathfrak{T} \subset \mathfrak{T}_{\text{std}}$. Do you think this is a metric topology?
(5) Let X be a nonempty set and $\mathfrak{T} = \{G \subset X : G = \emptyset \text{ or } X \setminus G \text{ is finite}\}$. This is called the *co-finite topology* on X. If X itself is finite, it is simply the discrete topology.

EXERCISE 3.1.1.
(1) Show that an arbitrary intersection of closed sets is still closed and a finite union of closed sets is closed. On the other hand, any collection of sets satisfying these two conditions indeed defines a topology by their complements.
(2) Let Y be a closed set in $(X, \mathfrak{T})$ and be given the induced topology $\mathfrak{T}|_Y$. If $A \subset (Y, \mathfrak{T}|_Y)$ is closed, show that A is also closed in $(X, \mathfrak{T})$.
(3) Suppose $\mathfrak{T}_1$ and $\mathfrak{T}_2$ are two topologies on the set X. Which one is still a topology, $\mathfrak{T}_1 \cup \mathfrak{T}_2$ or $\mathfrak{T}_1 \cap \mathfrak{T}_2$?
(4) Can we replace the word "finite" above by "countable" to define a "co-countable" topology?
(5) Given a topological space $(X, \mathfrak{T})$ and $A \subset X$. Show that $\mathfrak{T}|_A \overset{\text{def}}{:=} \{U \cap A : U \in \mathfrak{T}\}$ is a topology for A. It is called the *induced* or *relative topology*.

In abstract topology space, open sets are simply elements of the topology. The concept of interior has not yet been established. Therefore, we will now define interior and closure with reference to the topology.

DEFINITION 3.3. Let $A \subset X$ where $(X, \mathfrak{T})$ is a topological space. The *interior* of A is given by $\mathrm{Int}(A)$ or $\mathring{A} \stackrel{\text{def}}{:=} \bigcup \{ G \in \mathfrak{T} \ : \ G \subset A \}$. A point $x \in A$ is an *interior point* of A if $x \in \mathrm{Int}(A)$.

Similarly, the *closure* of A is given by

$$\mathrm{Cl}(A) \text{ or } \overline{A} = \bigcap \{ F \ : \ F \supset A, X \setminus F \in \mathfrak{T} \} .$$

At any point $x \in X$, a *neighborhood* of x is a set N such that $x \in \mathrm{Int}(N) \subset N$. When the context is clear, we often assume N itself is open.

EXERCISE 3.1.2.
(1) Based on the above definition, show that $x \in \mathrm{Int}(A)$ if and only if there is a set $U \in \mathfrak{T}$ such that $x \in U \subset A$.
(2) Similarly, show that $x \in \overline{A}$ if and only if for all open set $U \in \mathfrak{T}$ with $x \in U$, $U \cap A \neq \emptyset$.
(3) Give a suitable definition of *derived set* A' such that $\overline{A} = A \cup A'$.
(4) Define the convergence of a sequence in a topological space.
(5) For a general topological space $(X, \mathfrak{T})$,
   (a) Is there an example of $(X, \mathfrak{T})$ such that $\mathrm{Frt}(A) \neq \overline{A} \setminus \mathrm{Int}(A)$?
   (b) For an open set $G \in \mathfrak{T}$, is it true that $G = \mathrm{Int}(\mathrm{Cl}(G))$?
   (c) For a closed set $F \subset X$, is it true that $F = \mathrm{Cl}(\mathrm{Int}(F))$?
   (d) Is it true that $\overline{A \setminus B} = \overline{A} \setminus \mathrm{Int}\, B$?
(6) Define dense and nowhere dense subsets in a topological space.

Once the concept of open sets is established, many notions in the previous chapter can be carried to the setting of topological space.

DEFINITION 3.4. Let $(X, \mathfrak{T}_X)$ and $(Y, \mathfrak{T}_Y)$ be topological spaces. A mapping $f \colon X \to Y$ is *continuous* if for every $V \in \mathfrak{T}_Y$, $f^{-1}(V) \in \mathfrak{T}_X$. It is *continuous at* $x_0 \in X$ if for every neighborhood V of $f(x_0)$, $f{-}1(V)$ is a neighborhood of x_0.

EXERCISE 3.1.3. Note that a mapping f is continuous at x_0 does not imply that for every $f(x_0) \in V$ with $V \in \mathfrak{T}_Y$, $x_0 \in f^{-1}(V) \in \mathfrak{T}_X$. Give an example of $f \colon \mathbb{R} \to \mathbb{R}$ to illustrate this.

EXERCISE 3.1.4. Let $A \subset (X, \mathfrak{T})$ be given a topology $\mathfrak{T}_A$. Formulate an equivalent condition for $\mathfrak{T}_A = \mathfrak{T}|_A$ in terms of the inclusion map $\iota_A \colon A \subset X \to X$.

EXAMPLE 3.5. Let $(X, \mathfrak{T})$ ba a topological space with a proper subset, $\emptyset \neq A \subset X$. Then a topology slightly larger than the indiscrete topology is $\{ \emptyset, A, X \}$.

EXERCISE 3.1.5. In light of the above example, if there are two proper subsets, $\emptyset \neq A \neq B \subset X$. What should be included in the topology $\mathfrak{T}$ in which both $A, B \in \mathfrak{T}$? More generally, if there are several proper subsets $A_1, \ldots, A_n$, how can we create a topology $\mathfrak{T}$ such that every $A_k \in \mathfrak{T}$?

To end this section, let us consider the following example as an exercise.

EXERCISE 3.1.6. Let $\mathcal{C} = \mathcal{C}([a, b], \mathbb{R})$ be the set of continuous functions on a closed interval $[a, b]$. For any open subset U of $[a, b] \times \mathbb{R}$ with standard topology, define a subset $\mathcal{W}_U \subset \mathcal{C}$ by $\mathcal{W}_U = \{ f \in \mathcal{C} : \ \text{graph}(f) \subset U \}$.

Show that $\mathfrak{T} = \{ \mathcal{W}_U : \ U \text{ is an open subset of } [a, b] \times \mathbb{R} \}$ defines a topology for $\mathcal{C}$. Moreover, verify that $\mathfrak{T}$ is the same as the metric topology of d_∞.

3.2. Base and Subbase

In the previous section, a topological space is described by its topology, i.e., listing all the open sets. In many examples, certain important open sets already determine all the other open sets, for example, all open intervals in $\mathbb{R}$ or all open balls in a metric space. That is the concept of basic open sets. Moreover, people usually only write down the typical open sets when defining a topology. Thus, there is a need to consider certain subsets of a topology which are already enough to determine the topology.

DEFINITION 3.6. Let $(X, \mathfrak{T})$ be a topological space. A set $\mathcal{B} \subset \mathfrak{T}$ is called a *base* for $\mathfrak{T}$ if $\mathfrak{T} = \{ \bigcup \mathcal{A} \ : \ \mathcal{A} \subset \mathcal{B} \}$. Any $G \in \mathcal{B}$ is called a *basic open set*.

In the above notation, $\bigcup \mathcal{A}$ is a union of some of the sets in $\mathcal{B}$. In other words, taking all arbitrary unions of sets in $\mathcal{B}$ will form the topology $\mathfrak{T}$. Often, we will write in terms of families, i.e., $\mathcal{B} = \{ B_\alpha \in \mathfrak{T} : \ \alpha \in I \}$ where I is an index set such that $\mathfrak{T} = \left\{ \bigcup_{\beta \in J} B_\beta : \ J \subset I \right\}$.

EXAMPLE 3.7. In a metric space (X, d) with metric topology $\mathfrak{T}$, the set of all open balls $\mathcal{B}_1$ is a base. Also, the set consisting of all open balls with rational radius, $\mathcal{B}_2$ is a base. Let $\mathcal{B}_3$ be the set of all open balls with radius of the form $1/n$, $0 < n \in \mathbb{N}$. Is $\mathcal{B}_3$ a base? From this example, we see that a topology may have many bases.

The advantage of using basic sets is usually the following. Whenever it is needed to verify a statement involving open sets, it is sufficient to only verify it for basic open sets (see the exercise below for example).

EXERCISE 3.2.1.

(1) Let $(X, \mathfrak{T})$ be a topological space which has a base $\mathcal{B}$ and $A \subset X$. Show that $x \in \operatorname{Int}(A)$ if and only if there exists $B \in \mathcal{B}$ such that $x \in B \subset A$. Moreover, $x \in \operatorname{Cl}(A)$ if and only if for each $B \in \mathcal{B}$ with $x \in B$, $B \cap A \neq \emptyset$. Formulate and prove similar statements for $x \in A'$ or $x \in \operatorname{Frt}(A)$.
(2) Given two topologies $\mathfrak{T}_1$ and $\mathfrak{T}_2$ for X, each has base $\mathcal{B}_1$ and $\mathcal{B}_2$ respectively. Determine an equivalent condition on the bases for $\mathfrak{T}_1 \subset \mathfrak{T}_2$.
(3) Let $f\colon (X, \mathfrak{T}_X) \to (Y, \mathfrak{T}_Y)$ be mapping; $\mathcal{B}_X$ and $\mathcal{B}_Y$ are respectively bases for $\mathfrak{T}_X$ and $\mathfrak{T}_Y$. Then f is continuous at $x \in X$ if and only if for every $V \in \mathcal{B}_Y$ with $f(x) \in V$, there exists $U \in \mathcal{B}_X$ with $x \in U$ such that $f(U) \subset V$.
(4) Prove or disprove: f is continuous if and only if for each $V \in \mathcal{B}_Y$, $f^{-1}(V) \in \mathcal{B}_X$.

A topology $\mathfrak{T}$ has to satisfy two conditions. What is the implication of these two conditions on a base $\mathcal{B}$ for $\mathfrak{T}$? An arbitrary union or a finite intersection of sets in $\mathfrak{T}$ is also in $\mathfrak{T}$, so it must be a union of sets in $\mathcal{B}$. Thus, these may impose restrictions on $\mathcal{B}$.

EXERCISE 3.2.2. Let $\mathcal{B}$ denote a base for the topology $\mathfrak{T}$ of X.

(1) Give an example of $\mathcal{B}$ that $X \notin \mathcal{B}$. However, $X \in \mathfrak{T}$, for this reason, a base must satisfy that $\bigcup \mathcal{B} = X$.
(2) Show that for any $x \in G \in \mathfrak{T}$, there is a $B \in \mathcal{B}$ such that $x \in B \subset G$.
(3) Let $U, V \in \mathcal{B}$. Is it true that $U \cap V \in \mathcal{B}$?
(4) Show that if $U \cap V \neq \emptyset$, then there is $B \in \mathcal{B}$ such that $B \subset U \cap V$.

A topology may be determined by a “smaller” subset than a base. Let us consider the following collection of subsets in $\mathbb{R}$,

$$\mathcal{S} = \{\, (a, \infty) : a \in \mathbb{R} \,\} \cup \{\, (-\infty, b) : b \in \mathbb{R} \,\} \ .$$

Is this a base for the standard topology of $\mathbb{R}$? No, it is because the intersection of two such subsets may be a finite interval and cannot contain a set in $\mathcal{S}$ (violating (4) in Exercise 3.2.2). However, these finite intersections are all the open intervals, and hence they form a base. This becomes the typical example for a subbase.

DEFINITION 3.8. Let $(X, \mathfrak{T})$ be a topological space. A set $\mathcal{S} \subset \mathfrak{T}$ is called a *subbase* if $\{ S_1 \cap \cdots \cap S_n : S_j \in \mathcal{S}, n \in \mathbb{N} \}$ is a base for $\mathfrak{T}$. Equivalently, for any $G \in \mathfrak{T}$,

$$G = \bigcup_{\alpha} B_\alpha, \quad \text{where} \quad B_\alpha = \cap \mathcal{A}_\alpha \text{ for some finite subset } \mathcal{A}_\alpha \subset \mathcal{S}.$$

In this situation, we also say that $\mathfrak{T}$ is generated by $\mathcal{S}$.

Obviously, the same $\mathfrak{T}$ may be generated by different subbases. A topological space $(X, \mathfrak{T})$ may have different bases or subbases. This allows us to work with a good choice of bases or subbases according to the need.

EXERCISE 3.2.3. Is the following statement true: Let $f : (X, \mathfrak{T}_X) \to (Y, \mathfrak{T}_Y)$ be mapping; $\mathcal{S}_X$ and $\mathcal{S}_Y$ are respectively subbases for $\mathfrak{T}_X$ and $\mathfrak{T}_Y$. Then f is continuous if and only if for every $x \in X$ and $V \in \mathcal{S}_Y$ with $f(x) \in V$, there exists $U \in \mathcal{S}_X$ with $x \in U$ such that $f(U) \subset V$? Justify your answer.

In the above, given a topology $\mathfrak{T}$, we may take a subset $\mathcal{B} \subset \mathfrak{T}$ which may be sufficient to represent the whole $\mathfrak{T}$ in some sense. Alternatively, there is another consideration from a different point of view. Let X be a nonempty set, not yet with a topology, and take a subset $\mathcal{C} \subset \mathcal{P}(X)$. Could $\mathcal{C}$ determine a topology on X? The question can be addressed in three different levels. At the first level, would $\mathcal{C}$ itself be a topology. That simply depends on whether arbitrary unions and finite intersections remain inside $\mathcal{C}$.

At the second level, we may ask whether $\mathcal{C}$ is a base for an unknown topology. In other words, could the set $\{ \cup \mathcal{A} : \mathcal{A} \subset \mathcal{C} \}$ satisfy the two conditions of topology. We observe that any topology must have the whole set X, thus $\mathcal{C}$ must be sufficient to create the whole space by an arbitrary union; mathematically, $X = \cup \mathcal{A}$ for some $\mathcal{A} \subset \mathcal{C}$. Next, using the argument in Exercise 3.2.2, it is easy to show the useful fact given in the exercise below.

EXERCISE 3.2.4. A subset $\mathcal{C}$ is a base for a topology if and only if (a) there is $\mathcal{A} \subset \mathcal{C}$ such that $\cup \mathcal{A} = X$ and (b) for each $U, V \in \mathcal{C}$, there is $B \in \mathcal{C}$ such that $x \in B \subset U \cap V$.

At the third level, one asks whether $\mathcal{C}$ is a subbase for an unknown topology. Let us consider the simple case in Example 3.5. Given a set X and a proper subset, $\emptyset \neq A \subset X$, $\{ \emptyset, A, X \}$ is a topology containing A. In fact, it is the smallest one. Further, in Exercise 3.1.5, suppose $\emptyset \neq B \neq A \subset X$, the smallest possible topology that contains both A, B is clearly $\{\emptyset, A, B, A\cap B, A\cup B, X\}$. Likewise, if there are several different sets $A_1, \dots, A_n$ in the topology, many other sets must be also there. This is the idea behind generating a topology.

Suppose we would like to construct a "topology" $\mathcal{W}$ from $\mathcal{C}$. We need to check whether the constructed collection $\mathcal{W}$ satisfies the two properties of being a topology, namely, $\bigcup \mathcal{A} \in \mathcal{W}$ for each $\mathcal{A} \subset \mathcal{W}$ and $\bigcap \mathcal{F} \in \mathcal{W}$ for each finite $\mathcal{F} \subset \mathcal{W}$. This may pose additional requirements on the original $\mathcal{C}$. It turns out that no condition is required for $\mathcal{C}$ to be a subbase.

PROPOSITION 3.9. *Any nonempty collection $\mathcal{C} \subset \mathcal{P}(X)$ defines a topology $\mathfrak{T}$ such that $\mathcal{C}$ is a subbase for $\mathfrak{T}$.*

PROOF. Let $\mathcal{B} = \{\bigcap \mathcal{F} : \mathcal{F} \subset \mathcal{C},\quad \#(\mathcal{F}) < \infty\}$ and then $\mathfrak{T} = \{\bigcup \mathcal{A} : \mathcal{A} \subset \mathcal{B}\}$, which is called the topology generated by $\mathcal{C}$. In order to verify that $\mathfrak{T}$ is a topology, one only needs to use the de Morgan's law several times. $\square$

EXERCISE 3.2.5.
(1) Show that $\mathcal{B} = \{(a, b) \subset \mathbb{R} : a < b\} \cup \{\{r\} : r \in \mathbb{R} \setminus \mathbb{Q}\}$ form a base for a topology on $\mathbb{R}$.
(2) Show that $\mathcal{S} = \{[a, b) \subset \mathbb{R} : a <\leq b\}$ form a subbase for a topology on $\mathbb{R}$.

EXAMPLE 3.10. Let X be a nonempty set with a simple order "$<$". That is, it is a transitive relation satisfying if $x \neq y$, then either $x < y$ or $y < x$. Suppose that there is no smallest element nor largest element in X. The topology determined by the base $\mathcal{B} = \{(a, b) :\ a < b \in X\}$, where $(a, b) :\stackrel{\text{def}}{=\!=} \{x \in X :\ a < x < b\}$, is called the *ordered topology*. If there is smallest element $s \in X$ or largest element $\ell \in X$, then $\mathcal{B}$ needs to contain also sets of the form $[s, b)$ and $(a, \ell]$.

3.3. Countability

Metric spaces have many special properties which general topological spaces may not have. It is because the system of open balls always provides a good way of keeping records. Some important properties may be retained in topological spaces in which things may be counted.

DEFINITION 3.11. Let $(X, \mathfrak{T})$ be topological space and $x \in X$. A *neighborhood base* or *local base* at x is a set $\mathcal{B}_x$ containing neighborhoods of x such that for all neighborhood U of x, there is $B \in \mathcal{B}_x$ such that $x \in B \subset U$.

Clearly, if every $x \in X$ has a local base $\mathcal{B}_x$, then $\mathcal{B} = \bigcup_{x\in X} \mathcal{B}_x$ is a base for $\mathfrak{T}$, though it may be a very big collection. On the other hand, any base indeed contains enough open sets to have local bases everywhere.

Exercise 3.3.1. Let $\mathcal{B} \subset \mathfrak{T}$ be a base for $\mathfrak{T}$. Construct a local base for each $x \in X$.

Definition 3.12. A topological space $(X, \mathfrak{T})$ is called *second countable*, denoted $\mathrm{C_{II}}$, if it has a countable base. It is called *first countable*, denoted $\mathrm{C_I}$, if every $x \in X$ has a countable local base.

Exercise 3.3.2. Let $(A, \mathfrak{T}|_A)$ be the induced topology of $(X, \mathfrak{T})$. Show that if X is $\mathrm{C_I}$ or $\mathrm{C_{II}}$, then so is A.

Now, it is obvious that any metric space is $\mathrm{C_I}$ because one may choose balls of rational radii to form local bases. More precisely, $\mathcal{B}_x = \{\, B(x,q) \ : \ q \in \mathbb{Q} \,\}$ is a local base at $x \in X$. Then, $\mathcal{B} = \bigcup_{x\in X} \mathcal{B}_x$ is a base for the metric space. However, this $\mathcal{B}$ may be uncountable. In the case of $\mathbb{R}$ or $\mathbb{R}^n$, it is possible to take points with rational coordinates as centers, thus the standard Euclidean space is $\mathrm{C_{II}}$. Here is the analogous concept of $\mathbb{Q}$ in $\mathbb{R}$ in a topological space.

Definition 3.13. A topological space $(X, \mathfrak{T})$ is *separable* if it has a countable dense subset D. That is, $\overline{D} = X$ and D is countable.

Exercise 3.3.3. Is $A \subset X$ separable if X is so?

Naturally, one would expect that a separable $\mathrm{C_I}$ space behaves similarly as $\mathbb{R}^n$ and so it is $\mathrm{C_{II}}$. But, it is NOT true and we will discuss it below. Anyway, let us observe the relation between these countability concepts. First, that $\mathrm{C_{II}} \implies \mathrm{C_I}$ trivially follows from Exercise 3.3.1. The countable base $\mathcal{B}$ is also a local base for any $x \in X$. Second, one would naturally expect $\mathrm{C_{II}}$ leads to separable.

Theorem 3.14. *If a space $(X, \mathfrak{T})$ is C_{II} then it is separable.*

Proof. Let $\mathcal{B} = \{\, B_j \subset \mathfrak{T} : j \in \mathbb{N} \,\}$ be a countable base. The key of the proof is to construct a countable set. Naturally, pick $x_j \in B_j$ and let $D = \{\, x_j : j \in \mathbb{N} \,\}$.

In order to show that D is dense, we take any open set $G \in \mathfrak{T}$ and try to establish $G \cap D \neq \emptyset$. However, as $G \in \mathfrak{T}$ and $\mathcal{B}$ is a base, G is a union of some sets in $\mathcal{B}$. The desired result easily follows. □

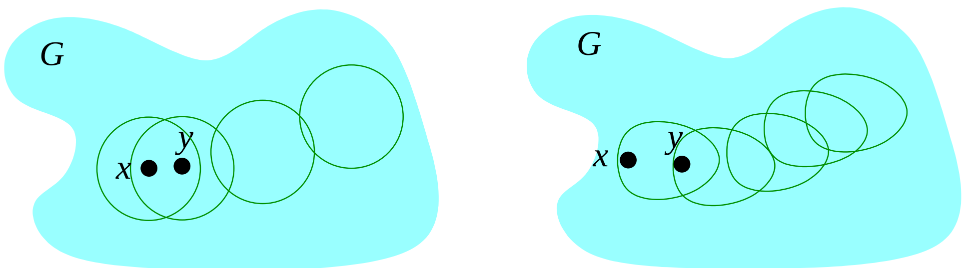
Now, let us start with a topological space $(X, \mathfrak{T})$ which is $\mathrm{C_I}$ and separable, i.e., there is $D \subset X$ with $\overline{D} = X$. We will see how it "almost" gets to $\mathrm{C_{II}}$ but fails. Let us write the countable dense set as $D = \{\, x_m \ : \ m \in \mathbb{N} \,\}$ and the countable local base at $x \in X$ as $\mathcal{B}_x = \{\, U_k^x \ : \ k \in \mathbb{N} \,\}$. It is natural to hope that

$$\mathcal{B} = \bigcup_{m \in \mathbb{N}} \mathcal{B}_{x_m} = \{\, U_k^{x_m} \ : \ m \in \mathbb{N}, \ k \in \mathbb{N} \,\}$$

is the desired countable base. In order to show that $\mathcal{B}$ is a base, an arbitrary open set $G \in \mathfrak{T}$ must be a union of some sets in $\mathcal{B}$. Since D is dense, $G \cap D \neq \emptyset$. Let $G \cap D = \{\, y_j = x_{mj} \ : \ j \in \mathbb{N} \,\}$, i.e., y_j are all the points of D inside G. Since $y_j \in G$ and $\mathcal{B}_{y_j}$ is a local base, we always have some $U \in \mathcal{B}_{y_j}$ such that $y_j \in U \subset G$. Taking all such U's in B_{y_j} and running over all y_j, we have

$$\mathcal{A} = \bigcup_{j \in \mathbb{N}} \{\, U \in \mathcal{B}_{y_j} \ : \ U \subset G \,\} = \{\, U_k^{y_j} \ : \ U_k^{y_j} \subset G, \ k \in \mathbb{N} \,\} \subset \mathcal{B}.$$

Since every open sets in $\mathcal{A}$ is contained in G, we have $\cup\mathcal{A} \subset G$. The difficulty occurs for $G \subset \cup\mathcal{A}$. Let $x \in G$, we hope that there exists $U_k^{y_j}$ such that $x \in U_k^{y_j} \subset G$. As D is dense, we may have y_j arbitrarily very near x. The natural idea is to take y_j inside a small neighborhood of x so that a neighborhood $U_k^{y_j}$ of y_j may cover x (as in the picture on the left). However, a bad situation (as in the picture on the right) may occur, in which $y_j \in U^x$ while $x \notin U^{y_j}$.



In a metric space, we always have situation such as the picture on the left. Mathematically, the metric guarantees that when $y_j \in B(x, \varepsilon)$ if and only if $x \in B(y_j, \varepsilon)$.

THEOREM 3.15. *A separable metric space is of second countable.*

EXERCISE 3.3.4. Finish the details of showing a separable metric space is second countable.

The countability of bases reveals certain facts about the topology. Another type of countability is related to how a space is covered by open sets. Let $\mathcal{C} \subset \mathfrak{T}$ be a collection of open sets. It is called an *open cover* if $X = \cup\mathcal{C}$. That is, the union of all sets in $\mathcal{C}$ equals X. A *subcover* of an open cover $\mathcal{C}$ is a subset $\mathcal{E} \subset \mathcal{C}$ such that $\mathcal{E}$ is itself an open cover, i.e., $\cup\mathcal{E} = X$.

DEFINITION 3.16. A topological space $(X, \mathfrak{T})$ is *Lindelöf* if every open cover has a countable subcover. It is called *compact* if every open cover has a finite subcover.

REMARK . We often say $A \subset X$ is a *compact subset* of X. That actually means under the induced topology, $(A, \mathfrak{T}|_A)$ is compact. In this situation, an open cover for A is usually taken as $\mathcal{C} \subset \mathfrak{T}$ with $\cup\mathcal{C} \supset A$. For this reason, A is compact if and only if every $\mathcal{C} \subset \mathfrak{T}$ with $\cup\mathcal{C} \supset A$ has a finite $\mathcal{E} \subset \mathcal{C}$ with $\cup\mathcal{E} \supset A$.

We will finish the section by relating the concept of Lindelöf with countability.

THEOREM 3.17. *If $(X, \mathfrak{T})$ is C_{II}, then it is Lindelöf.*

PROOF. Let $\mathcal{C} \subset \mathfrak{T}$ be an open cover for X and $\mathcal{B} \subset \mathfrak{T}$ be a countable base. Our aim is to select sets in $\mathcal{C}$ to form a countable subcover $\mathcal{E}$. Let

$$\mathcal{B}^* = \{ B \in \mathcal{B} : B \subset V \text{ for some } V \in \mathcal{C} \}.$$

Then there is a natural mapping $\mathcal{B}^* \to \mathcal{C}$, which has a countable image, say $\mathcal{E}$. We will show that $\bigcup \mathcal{E} = X$.

Take any $x \in X$. Since $\mathcal{C}$ is an open cover, there is $U \in \mathcal{C}$ such that $x \in U$. In turns, U is a union of some sets in $\mathcal{B}$. Thus, there is $B \in \mathcal{B}$ with $x \in B \subset U$. Let the image of B under the mapping $\mathcal{B}^* \to \mathcal{E}$ be V. Then, we have $V \in \mathcal{E}$ with $x \in V$. □

THEOREM 3.18. *A Lindelöf metric space is C_{II}.*

PROOF. For each $n \in \mathbb{N}$, let $\mathcal{C}_n = \{ B(x, 2^{-n}) : x \in X \}$. It is clearly an open cover for X. Since X is Lindelöf, there is a countable subcover $\mathcal{E}_n \subset \mathcal{C}_n$. Let $\mathcal{B} = \bigcup_{n\in\mathbb{N}} \mathcal{E}_n$. It remains to show that $\mathcal{B}$ is actually a base. □

EXERCISE 3.3.5. Show that $\mathcal{B}$ obtained above is a base.

CHAPTER 4

Space Constructions

In this chapter, we will discuss several usual ways of constructing new spaces from existing topological spaces.

The first notion of subspaces in §4.1 has frequently occurred in the previous sections. The product space is introduced in §4.2 where infinite product is highlighted. Then, some attention is given to quotient space in §4.3, in which two equivalent views of quotient topology are given. An abundance of examples of quotient spaces are given in §4.4. In §4.5, some matrix topological groups are given with highlight on the quotient space structures. To conclude this chapter, in §4.6, infinite product spaces are introduced by their featuring properties, which is analogous to quotient spaces.

4.1. Subspaces

We have seen before that in a topological space $(X, \mathfrak{T})$, a subset $Y \subset X$ may be given the *induced* or *relative topology*,

$$\mathfrak{T}|_Y = \{ G \cap Y : G \in \mathfrak{T} \} .$$

In this situation, $(Y, \mathfrak{T}|_Y)$ is called a *subspace* of $(X, \mathfrak{T})$.

It is obvious that $E \subset Y$ is a closed set in $(Y, \mathfrak{T}|_Y)$ if and only if there exists closed set F in $(X, \mathfrak{T})$ such that $E = F \cap Y$. Since $E \subset Y$ is also a subset of X. We have to careful about the two concepts of open or closed subset. More precisely, E is open/closed in $(Y, \mathfrak{T}|_Y)$ may not be open/closed in $(X, \mathfrak{T})$. However, if Y is a closed set in X, then $E \subset Y$ is closed in Y if and only if it is closed in X.

Therefore, the main concern about a subspace Y in X is so-called *heredity*. For examples, let $A \subset Y \subset X$,

- Is it true that $\mathrm{Int}_X(A) \cap Y = \mathrm{Int}_Y(A)$?
- What about similar questions about closure or sequences in Y versus in X?

- What is the relation between continuity of $f : X \to Z$ and $f|_Y : Y \to Z$?
- If X is Hausdorff, then is Y Hausdorff?
- If (X, d) is a complete metric space, then what about $(Y, d|_Y)$?

Exercise 4.1.1. Write down some properties or statements about a topological space. Then figure out a question about heredity, prove or disprove it.

Exercise 4.1.2. Given a base $\mathcal{B}$ and a subbase $\mathcal{S}$ for the topological space $(X, \mathfrak{T})$. Let $Y \subset X$. Is there a natural way to construct base and subbase for the induced topology $\mathfrak{T}|_Y$?

4.2. Finite Product

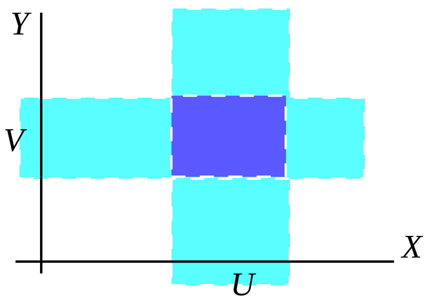
Given topological spaces $(X, \mathfrak{T}_X)$ and $(Y, \mathfrak{T}_Y)$, the *product topology* $\mathfrak{T}_{X\times Y}$ on $X \times Y$ is generated by the collection

$$\mathcal{S} = \{ X \times V : V \in \mathfrak{T}_Y \} \cup \{ U \times Y : U \in \mathfrak{T}_X \} .$$

Finite intersections of sets in $\mathcal{S}$ gives the following base

$$\mathcal{B} = \{ U \times V : U \in \mathfrak{T}_X,\ V \in \mathfrak{T}_Y \} .$$

In general, a finite product is defined similarly, only with more factors in the product. However, this format of a base is not valid for *infinite* product, which we will further discuss below.



Definition 4.1. Let $(X_j, \mathfrak{T}_j)$ be topological spaces, $j = 1, \ldots, n$. The *product topology* on $X_1 \times X_2 \times \cdots X_n$ is generated by the subbase

$$\bigcup_{k=1}^{n} \{ X_1 \times \cdots \times U_k \times \cdots \times X_n \ : \ U_k \in \mathfrak{T}_k \} .$$

Essentially, the following result is no longer true in infinite product.

Exercise 4.2.1. In a finite product topology, the subbase given in the definition determined the base $\{ U_1 \times U_2 \times \cdots U_n \ : \ U_j \in \mathfrak{T}_j \}$.

In terms of topological spaces, one also expects that $(X\times Y)\times Z = X\times(Y\times Z)$.

EXAMPLE 4.2. The standard Euclidean space $\mathbb{R}^n$ is a product of finite copies of standard $\mathbb{R}$. Note that a simple statement $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$ as topological spaces may require some thoughts. The standard $\mathbb{R}^2$ is defined as a metric space (any ℓ_p-metric). An obvious base for the topology is the set of open balls. However, the open balls of the ℓ_2-metric on $\mathbb{R}^2$ are circular disks while those of ℓ_∞-metric are squares. Neither of them contains a rectangular set, which is of the form $U\times V$ and so belongs to the base of $\mathbb{R}\times\mathbb{R}$. In addition to the above, $\mathbb{R}^n$ can also be obtained by an inductive process of $\mathbb{R}^n = \mathbb{R}\times\mathbb{R}^{n-1}$.

EXERCISE 4.2.2.
(1) Let $\mathcal{B}_X$ and $\mathcal{B}_Y$ are bases for the topologies $\mathfrak{T}_X$ and $\mathfrak{T}_Y$ respectively. Show that $\{U\times V : U\in\mathcal{B}_X,\ V\in\mathcal{B}_Y\}$ is a base for the product topology.
(2) Let $X\times Y$ be given the product topology of X and Y; $A\subset X$, $B\subset Y$. Show that $\mathrm{Int}(A\times B) = \mathrm{Int}(A)\times\mathrm{Int}(B)$; $\mathrm{Cl}(A\times B) = \mathrm{Cl}(A)\times\mathrm{Cl}(B)$. What about $\mathrm{Frt}(A\times B)$?
(3) Let X has topologies $\mathfrak{T}_X\subset\mathfrak{T}'_X$ and Y has topologies $\mathfrak{T}_Y\subset\mathfrak{T}'_Y$. Compare the product topologies $\mathfrak{T}_X\times\mathfrak{T}_Y$ and $\mathfrak{T}'_X\times\mathfrak{T}'_Y$.
(4) From the above, if A, B are separable in X, Y respectively, show that $A\times B$ is separable in the product space.
(5) Assume that both X and Y are second countable, is $X\times Y$ the same?

There are two natural projections from the product to each factor,

$$\pi_X\colon X\times Y \to X, \qquad \pi_Y\colon X\times Y \to Y.$$

For any open set $U\in\mathfrak{T}_X$, $\pi_X^{-1}(U) = U\times Y\in\mathfrak{T}_{X\times Y}$. Thus, π_X, and similarly π_Y is continuous. In fact, one may prove the minimality/maximality statements in the following exercise.

EXERCISE 4.2.3.
(1) *Minimality.* Let $\mathfrak{T}$ be a topology on $X\times Y$ such that both π_X and π_Y are continuous. Then $\mathfrak{T}_{X\times Y}\subset\mathfrak{T}$.
(2) For any $f\colon Z \to (X\times Y, \mathfrak{T}_{X\times Y})$ from a topological space Z, f is continuous if and only if both $\pi_X\circ f$ and $\pi_Y\circ f$ are so.
(3) *Maximality.* Let $\mathfrak{T}$ be a topology on $X\times Y$ such that for any space Z, $f\colon Z \to (X\times Y, \mathfrak{T})$ is continuous if both $\pi_X\circ f$ and $\pi_Y\circ f$ are so. Prove that $\mathfrak{T}\subset\mathfrak{T}_{X\times Y}$.

The projections π_X and π_Y are also *open mappings*, in the sense that, for every open set $G \in \mathfrak{T}_{X\times Y}$, $\pi_X(G) \in \mathfrak{T}_X$ and $\pi_Y(G) \in \mathfrak{T}_Y$. This can be easily observed by considering images of the set $\bigcup_\alpha (U_\alpha \times V_\alpha)$ where $U_\alpha \in \mathfrak{T}_X$ and $V_\alpha \in \mathfrak{T}_Y$.

Something in the product topology may be surprising, as illustrated in the following exercises.

EXERCISE 4.2.4. (1) Let $[0,1]$ and $(0,1]$ be given the standard topology. Prove that the product spaces $[0,1]\times(0,1]$ and $(0,1]\times(0,1]$ are homeomorphic.

(2) Given an example of spaces X such that the diagonal set $\{\,(x,x) \;:\; x \in X\,\}$ is not closed in the product space $X \times X$.

EXERCISE 4.2.5. Let $\mathbb{R}_L$ denote $\mathbb{R}$ with the topology generated by intervals of the form $[a,b)$ and $\mathbb{R}\times\mathbb{R}_L$ be the product topological space. What is the induced topology on the diagonal $\{\,(x,x) : x \in \mathbb{R}\,\}$?

4.2.1. The Annulus or Cylinder. Consider the complex plane $\mathbb{C}$ with the standard topology, i.e., just as $\mathbb{R}^2$. The annulus may be defined as the subset

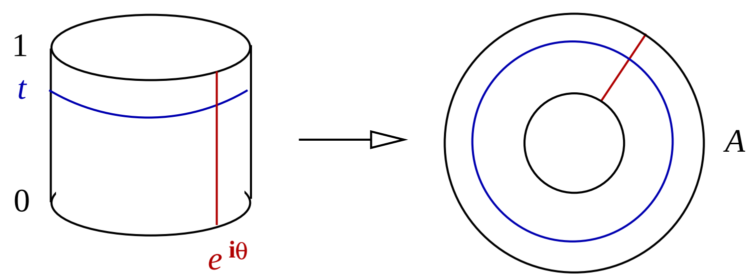
$$A = \{\, z \in \mathbb{C} \;:\; 1 \le |z| \le 2\,\}$$

with the induced topology. On the other hand, if for $\mathbb{S}^1 = \{\, z \in \mathbb{C} \;:\; |z| = 1\,\}$ and $[0,1]$ with standard topology, we may show that $\mathbb{S}^1 \times [0,1]$ with the product topology is homeomorphic to A. The homeomorphism $h\colon \mathbb{S}^1 \times [0,1] \to A$ is simply given by

$$h\left(e^{\mathbf{i}\theta}, t\right) = (1+t)e^{\mathbf{i}\theta}\,.$$

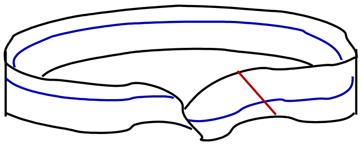
EXERCISE 4.2.6. Show that h is a homeomorphism. *Hint.* pick suitable bases for the topologies of $\mathbb{S}^1 \times [0,1]$ and $\mathbb{C}$.

The mapping is illustrated by the following picture.



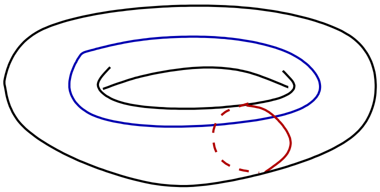
It should be noted that a similar bijection can be defined from the Möbius band to the annulus. However, the mapping is not a homeomorphism. In fact, the

Möbius band is not homeomorphic to the annulus. It usually requires algebraic topology to prove this fact.



Intuitively speaking, one may imagine the red interval in the above picture as $[-1,1]$. After going one round of the strip, $[-1,1]$ is inverted. Algebraically, this is captured by $n \sim -n$ for $n \in \mathbb{Z}$. The resulting group is $\mathbb{Z}/2 = \{0,1\}$. But, this group does not occur in the product $\mathbb{S}^1 \times [0,1]$.

4.2.2. The Torus. Start with a small circle of radius r with its center at a distance of $R > r$ from the rotation axis. The surface of revolution obtained is called a torus $\mathbb{T}$. Therefore, it can be seen as a subset of $\mathbb{R}^3$ and so there is an induced topology.



Let $\mathbb{S}^1 \subset \mathbb{C}$ be given the standard induced topology. The product space $\mathbb{S}^1 \times \mathbb{S}^1$ is homeomorphic to $\mathbb{T}$ by the mapping

$$\left(e^{\mathbf{i}\theta}, e^{\mathbf{i}\phi}\right) \mapsto \left((R + r\cos\theta)\cos\phi, (R + r\cos\theta)\sin\phi, r\sin\theta \right) .$$

EXERCISE 4.2.7. Verify that the above mapping is a homeomorphism.

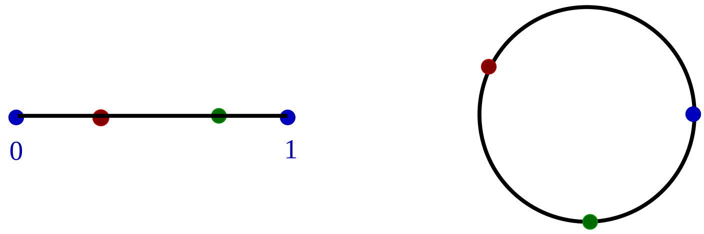
Similarly, one can define the *n-dimensional torus*, which is the product of n copies of circles, i.e., $\mathbb{T}^n \stackrel{\text{def}}{:=} \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$. Moreover, a 3-dimensional *solid torus* is the product space $\mathbb{S}^1 \times \mathbb{D}^2$, where $\mathbb{D}^2 = \{ z \in \mathbb{C} \ : \ |z| \leq 1 \}$. The solid torus has a boundary, which is exactly the torus, i.e.,

$$\partial(\mathbb{S}^1 \times \mathbb{D}^2) = \mathbb{S}^1 \times \mathbb{S}^1 .$$

EXERCISE 4.2.8. Let X be an ordered topological space (see Example 3.10). Then $X \times X$ has a *dictionary order* by $(a_1, a_2) < (b_1, b_2)$ if $a_1 < a_2$ or $a_1 = a_2$ with $b_1 < b_2$. In this way, $X \times X$ also becomes an ordered topological space. Show that $\mathbb{R} \times \mathbb{R}$ with order topology is homeomorphic to $\mathbb{R}_d \times \mathbb{R}$ where $\mathbb{R}_d$ is discrete and the second $\mathbb{R}$ is standard.

4.3. Quotient Spaces

Let us start with an example. Take an interval $[0, 1]$, imagine it as a wire and glue its two end-points. Then a circle, $\mathbb{S}^1$ is obtained.



In terms of set theory, this gluing process can be seen as the quotient set of an equivalence relation. More precisely, define $\sim$ on $[0, 1]$ by

$$s \sim t \qquad \text{if and only if} \qquad |s - t| = 0 \text{ or } 1\,.$$

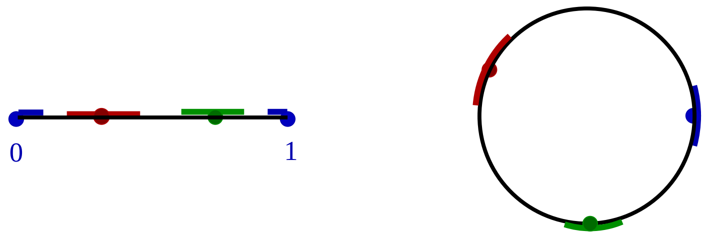
Since $s, t \in [0, 1]$, this really means either $s = t$ or $s = 0, t = 1$ or $s = 1, t = 0$. In other words, a point is only identified with itself and the two end-points will be identified. With this equivalence relation $\sim$, there are two types of equivalence classes,

$$\{\, x \,\}\,, \quad 0 < x < 1; \qquad \text{and} \qquad \{\, 0, 1 \,\}\ .$$

There is obviously a bijection between the quotient set $[0, 1]/\!\sim$ and $\mathbb{S}^1$.

Exercise 4.3.1. Let $[s]$ denote the equivalence class of $s \in [0, 1]$. Give an explicit expression of an bijection from $[0, 1]/\!\sim$ to $\mathbb{S}^1$.

Now, can we define a topology for $[0, 1]/\sim$ such that it is "the same as" the standard circle $\mathbb{S}^1$? The answer is yes and this is called the *quotient topology*. Let us compare the neighborhoods of corresponding points in $[0, 1]$ and $\mathbb{S}^1$ in the following picture.



The neighborhood at any point of $\mathbb{S}^1$ is always a short arc. The green or red neighborhoods correspond to short open intervals (green or red) in the interior of $[0, 1]$. However, the blue neighborhood in $\mathbb{S}^1$ corresponds to a union of two half open-closed short intervals at the end-points 0 and 1. This is because the blue point in $\mathbb{S}^1$ corresponds to the equivalence class $\{0,1\}$. This example serves as a good illustration of the following general theory.

4.3.1. Quotient Topology I. Let $(X, \mathfrak{T})$ be a topological space and $\sim$ be an equivalence relation on X with a quotient map $q\colon X \to X/{\sim}$ from X to the quotient set $X/{\sim}$ taking $x \mapsto [x]$.

DEFINITION 4.3. The *quotient topology* on the quotient space $X/{\sim}$ is

$$\mathfrak{T}_q = \left\{ V \subset X/{\sim} :\ q^{-1}(V) \in \mathfrak{T} \right\} .$$

It can be easily verified that $\mathfrak{T}_q$ is a topology for $X/{\sim}$, i.e., it is closed under arbitrary union and finite intersection.

Here is another way to understand an open set in $\mathfrak{T}_q$. Let $V \subset X/{\sim}$. Then it is a collection of equivalence classes (each one is a subset of X). Take all these equivalence classes and its union is a subset of X. We ask whether this union is an open set in X to determine whether $V \in \mathfrak{T}_q$. That is,

$$q^{-1}(V) = \bigcup_{[x]\in V} [x] = \bigcup_{[x]\in V} \{ y \in X \ :\ y \sim x \} .$$

There are several important properties.

PROPOSITION 4.4. (1) *The mapping $q\colon X \to X/{\sim}$ is surjective.*
(2) *The mapping $q\colon (X, \mathfrak{T}) \to (X/{\sim}, \mathfrak{T}_q)$ is continuous.*
(3) *If $\mathfrak{T}'$ is a topology on $X/{\sim}$ such that $q\colon (X, \mathfrak{T}) \to (X/{\sim}, \mathfrak{T}')$ is continuous, then $\mathfrak{T}' \subset \mathfrak{T}_q$.*

The third property is also called MAXIMALITY of the quotient topology; in fact, it means that the quotient topology is the LARGEST topology to have the mapping q continuous.

EXERCISE 4.3.2. Prove the above properties of $\mathfrak{T}_q$. Moreover, show that $\mathfrak{T}_q$ is the only topology on $X/{\sim}$ that satisfies all the properties.

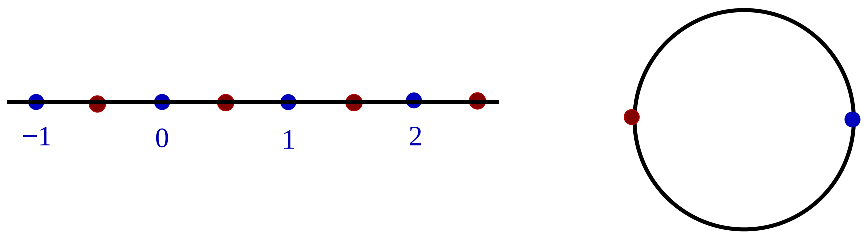
EXERCISE 4.3.3. (1) Let $X = \{(x,0) : x \in \mathbb{R}\} \cup \{(x,1) : x \in \mathbb{R}\} \subset \mathbb{R}^2$. Such a space is called the disjoint union of two copies of $\mathbb{R}$, denoted $X = \mathbb{R} \sqcup \mathbb{R}$. Define an equivalence relation on X by identifying $(0,0)$ and $(0,1)$. Rigorously, this means $(s_1,t_1) \sim (s_2,t_2)$ iff $(s_1,t_1) = (s_2,t_2)$ or $(s_1,t_1) = (0,0)$ while $(s_2,t_2) = (0,1)$ or vice versa. Show that $X/{\sim}$ is homeomorphic to the union of the two axes in $\mathbb{R}^2$.

(2) Let $X = \{ (s, n) \in \mathbb{R}^2 : 0 \neq n \in \mathbb{Z} \}$ and $Y = \{ (s, 1/n) \in \mathbb{R}^2 : n \in \mathbb{Z} \}$ be given the standard induced topology. Define an equivalence relation on both X and Y by $(s_1, t_1) \sim (s_2, t_2)$ iff $(s_1, t_1) = (s_2, t_2)$ or $s_1 = s_2 = 0$. That is, all the points on the y-axis are identified to be one point. Is it true that $X/\sim$ and $Y/\sim$ are homeomorphic?

EXAMPLE 4.5. There is another way to construct the circle. Define an equivalence relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \qquad \text{if} \qquad |x - y| \in \mathbb{Z}\,.$$

That means x and y have the same decimal part, so they both can be represented by a number $r \in [0, 1)$. The quotient set in this construction is usually denoted $\mathbb{R}/\mathbb{Z}$.

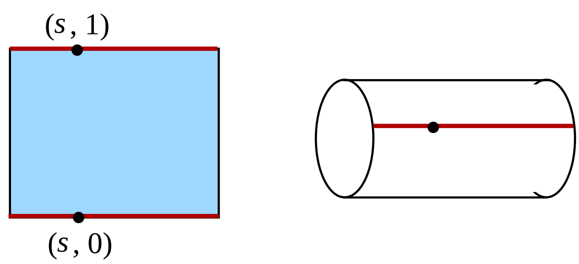


There is an obvious homeomorphism between $\mathbb{R}/\mathbb{Z}$ and $\mathbb{S}^1$ given in the following,

$$\begin{array}{ccccc} \mathbb{R} & \longrightarrow & \mathbb{R}/\mathbb{Z} & \longrightarrow & \mathbb{S}^1 \\ x & \mapsto & [x] & \mapsto & e^{2\pi \mathrm{i} x} \end{array}$$

Note that $(\mathbb{R}, +)$ is also an abelian group and $\mathbb{R}/\mathbb{Z}$ is indeed the factor group.

EXAMPLE 4.6. A cylinder can be obtained as a quotient space by a construction similar to that of the circle.



Define an equivalence relation $\sim$ on $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ with standard topology by

$$(s_1, t_1) \sim (s_2, t_2) \qquad \text{if} \qquad (s_1, t_1) = (s_2, t_2) \quad \text{or} \quad s_1 = s_2 \,\&\, |t_1 - t_2| = 1\,.$$

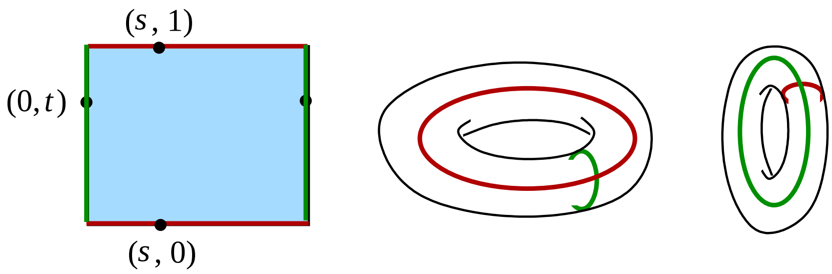
Alternatively, one may define $\sim$ on $[0, 1] \times \mathbb{R}$ similar to $\mathbb{R}/\mathbb{Z}$ to obtain the same space.

EXERCISE 4.3.4. Show that the cylinder with the above quotient topology is homeomorphic to $[0,1]\times\mathbb{S}^1$ with the product topology.

EXAMPLE 4.7. The torus can also be seen as a quotient space, in addition to being a product space or subspace of $\mathbb{R}^3$ as previously discussed. Define an equivalence relation on $\mathbb{R}\times\mathbb{R}$ by $(s_1,t_1)\sim(s_2,t_2)$ if

$$(s_1,t_1)=(s_2,t_2) \qquad \text{or} \qquad s_1-s_2\in\mathbb{Z} \quad \& \quad t_1-t_2\in\mathbb{Z}\,.$$

There is an analogous equivalence relation on $[0,1]\times[0,1]$ but the above is much simpler. One can show that $(\mathbb{R}\times\mathbb{R})/\!\sim$ is homeomorphic to $\mathbb{S}^1\times\mathbb{S}^1$.



Note that one may either glue $(s,0)\sim(s,1)$ first or $(0,t)\sim(1,t)$ first. Apparently in $\mathbb{R}^3$, it may result in the second or third picture above. However, they are topologically the same. In fact, there are many other ways of gluing.

EXERCISE 4.3.5. (1) Define an equivalence relation on $\mathbb{R}$ by identifying n with $1/n$ for all $n\in\mathbb{Z}$.
   (a) Sketch a picture to represent the space $\mathbb{R}/\!\sim$.
   (b) Find a sequence $x_n\in\mathbb{R}$ such that $[x_n]\in\mathbb{R}/\!\sim$ converges but x_n does not.
   (c) Can a sequence $x_n\in\mathbb{R}$ converge but $[x_n]\in\mathbb{R}/\!\sim$ does not?
(2) Let $X/\!\sim$ be a quotient space obtained from X and $Y\subset X$.
   (a) Show that there is a natural way to induce an equivalence relation on Y; and thus a quotient space $Y/\!\sim$.
   (b) Let $Y^*=\{\,[x]\in(X/\!\sim):[x]\cap Y\neq\emptyset\,\}$ be given the topology induced from $X/\!\sim$. Is Y^* homeomorphic to $Y/\!\sim$?
(3) Let $X=\mathbb{R}/\!\sim$ where $s\sim t$ if $s=t$ or $s,t\in\mathbb{Z}$ and

$$Y=\bigcup_{n=1}^{\infty}\left\{z\in\mathbb{C}:\ \left|z-\frac{1}{n}\right|=\frac{1}{n}\right\},\qquad Z=\bigcup_{n=1}^{\infty}\{\,z\in\mathbb{C}:\ |z-n|=n\,\}\,,$$

with induced topology of $\mathbb{C}=\mathbb{R}^2$. Are X, Y, Z homeomorphic?
(4) Let $A\subset X$. What can you say about $\mathrm{Int}(A)/\!\sim$ and $\mathrm{Int}(A/\!\sim)$; $\mathrm{Cl}(A)/\!\sim$ and $\mathrm{Cl}(A/\!\sim)$?

4.3.2. Quotient Topology II. Let us consider the example of the circle again. Now, we look at the circle as a subspace of $\mathbb{C}$, that is,

$$\mathbb{S}^1 = \{ z \in \mathbb{C} \ : \ |z| = 1 \} .$$

This is already a space with a chosen topology (standard). There is a special mapping, whose properties are listed in the exercise below,

$$\wp \colon \mathbb{R} \ \to \ \mathbb{S}^1 \qquad x \mapsto e^{2\pi \mathbf{i} x} .$$

EXERCISE 4.3.6. Show that the mapping $\wp$ satisfies the following:

- It is surjective;
- It is continuous;
- It is an open mapping, i.e., it takes open sets to open sets.

In fact, the third property of $\wp$ is a little strong. We indeed needs a weaker one given in the exercise below.

EXERCISE 4.3.7. A set $V \subset \mathbb{S}^1$ is open if and only if $\wp^{-1}(V) \subset \mathbb{R}$ is open. Moreover, if $\mathfrak{T}'$ is any topology on $\mathbb{S}^1$ such that $\wp \colon \mathbb{R} \to (\mathbb{S}^1, \mathfrak{T}')$ is continuous, then $\mathfrak{T}'$ is a subset of the standard topology of $\mathbb{S}^1$.

On the other hand, it is shown above that the circle can be obtained as a quotient space so there is a quotient map, $q \colon \mathbb{R} \ \to \ \mathbb{R}/\mathbb{Z} \qquad x \mapsto [x]$. The equivalence class $[x]$ can be imagined as the decimal part of x. Now, both mappings $\wp$ and q satisfy the three conditions stated in Proposition 4.4. In fact, these two mappings are related by the commutative diagram

$$\begin{tikzcd}
x \in \mathbb{R} \arrow[rr, "\wp"] \arrow[drr, "q"] & & \mathbb{S}^1 \quad e^{2\pi \mathbf{i} x} \\
& & \mathbb{R}/\mathbb{Z} \quad [x] \arrow[u, "h"]
\end{tikzcd}$$

Note that h is well-defined because if $[x_1] = [x_2]$, then $x_1 = x_2 + k$ for some $k \in \mathbb{Z}$ and so $e^{2\pi \mathbf{i} x_1} = e^{2\pi \mathbf{i}(x_2+k)}$. Indeed, h is a homeomorphism. In this way, $\mathbb{S}^1$ and $\mathbb{R}/\mathbb{Z}$ can be seen as two names for the same space; hence $\wp$ and q are two names for the same mapping.

DEFINITION 4.8. Let $(X, \mathfrak{T}_X)$ be a topological space and $f \colon X \to Y$ be a surjection. The topology $\mathfrak{T}_f$ is called the *quotient topology wrt f*, if the following conditions are satisfied:

- f is surjective (this is just stated for convenience);
- f is continuous; and
- for any topology $\mathfrak{T}'$ on Y with $f\colon X \to (Y, \mathfrak{T}')$ continuous, one must have $\mathfrak{T}' \subset \mathfrak{T}_f$.

The third condition above can always be replaced by $V \in \mathfrak{T}_Y$ if and only if $f^{-1}(V) \in \mathfrak{T}_X$.

The following exercise provides a convenient way to verify whether a function from a quotient space is continuous.

Exercise 4.3.8. As above, given $(X, \mathfrak{T}_X)$, $f\colon X \to Y$, and let $(Y, \mathfrak{T}_f)$ be the quotient space wrt f. Show that for any topological space $(Z, \mathfrak{T}_Z)$ and any mapping $g\colon (Y, \mathfrak{T}_f) \to (Z, \mathfrak{T}_Z)$, we have g is continuous if and only if $g \circ f\colon (X, \mathfrak{T}_X) \to (Z, \mathfrak{T}_Z)$ is continuous.

In fact, the above is a characterizing property of the quotient topology wrt f. That is, any topology $\mathfrak{T}_*$ on Y satisfying the above property must be $\mathfrak{T}_f$. The mathematical statement is given in the exercise.

Exercise 4.3.9. Let $\mathfrak{T}_*$ be a topology on Y satisfying the following property: For any topological space $(Z, \mathfrak{T}_Z)$ and any mapping $g\colon (Y, \mathfrak{T}_*) \to (Z, \mathfrak{T}_Z)$, we always have g is continuous if $f \circ g\colon (X, \mathfrak{T}_X) \to (Z, \mathfrak{T}_Z)$ is continuous.

Prove that $\mathfrak{T}_* \supset \mathfrak{T}_f$ (so quotient topology also has a meaning of minimality).

4.4. Examples of Spaces

A useful path to understand topology is through the recognition of examples, non-examples, and counter-examples. In this section, we will show how certain famous examples are obtained by space construction. These examples are commonly used in various fields of mathematics.

First, we will *duplicate* a space by the construction of disjoint union. Let $(X, \mathfrak{T}_X)$ be a topological space and consider the subset of $X \times \mathbb{R}$,

$$X \sqcup X \stackrel{\text{def}}{=\!=} \{ (x, 0) \in X \times \mathbb{R} \ : \ x \in X \} \cup \{ (x, 1) \in X \times \mathbb{R} \ : \ x \in X \} .$$

A typical subset of $X \sqcup X$ is of the form $(A \times \{ 0 \}) \cup (B \times \{ 1 \})$ where $A, B \subset X$. We define it to be an open set of $X \sqcup X$ if both $A, B \in \mathfrak{T}_X$.

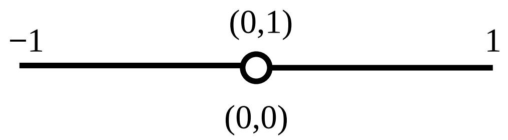
EXERCISE 4.4.1. Show that this topology of $X \sqcup X$ is exactly the induced topology from $X \times \mathbb{R}$.

Similarly, one may define the *disjoint union* $X \sqcup Y$ for any pair of topological spaces.

EXAMPLE 4.9. Define an equivalence relation $\sim$ on $[-1,1] \sqcup [-1,1]$ by $(x_1, t_1) \sim (x_2, t_2)$ if

$$(x_1, t_1) = (x_2, t_2) \qquad \text{or} \qquad x_1 = x_2 \neq 0 \,\&\, |t_1 - t_2| = 1\,.$$

In other words, we glue every pair of $x_1 = x_2 \neq 0$ and leave $(0,0), (0,1)$ unglued. The picture of this space is given below in which the middle small arcs are considered as two points.



EXAMPLE 4.10. The circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ and torus $\mathbb{T} = \mathbb{R}^2/\mathbb{Z}^2$ are defined above. In general, the n-dimensional torus is $\mathbb{T}^n \overset{\text{def}}{:=} \mathbb{R}^n/\mathbb{Z}^n$, which is homeomorphic to the product space $\underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{n \text{ times}}$.

4.4.1. Quotient on a Disk. Many interesting examples are obtained as a quotient space of a 2-dimensional disk, often viewed as a square.

EXAMPLE 4.11. The 2-dimensional sphere $\mathbb{S}^2 = \left\{ \vec{\boldsymbol{x}} \in \mathbb{R}^3 \,:\, \|\vec{\boldsymbol{x}}\| = 1 \right\}$ can be defined as a subspace of $\mathbb{R}^3$. It is also a quotient space. The method is valid for higher dimensions too.

Take the unit disk $\mathbb{D} = \{ z \in \mathbb{C} \,:\, |z| \leq 1 \}$ in $\mathbb{R}^2$ or $\mathbb{C}$. Define an equivalence relation $\sim$ on $\mathbb{D}$ by $z_1 \sim z_2$ if

$$z_1 = z_2 \quad \text{or} \quad |z_1| = 1 = |z_2|\,.$$

In other words, the whole boundary circle of $\mathbb{D}$ is identified to be one point. It can be shown that $\mathbb{D}/{\sim}$ is homeomorphic to $\mathbb{S}^2$ as suggested below.

EXERCISE 4.4.2. Define a mapping $p \colon \mathbb{D} \to \mathbb{S}^2$ such that p sends a radial line to a longitude of the sphere such as the origin is mapped to the south pole and the boundary circle is mapped to the north pole. Show that p is surjective, continuous, and open. Hence, show that $(\mathbb{D}/{\sim}) = \mathbb{S}^2$.

Exercise 4.4.3. On the disk $\mathbb{D}$, define an equivalence relation $\approx$ by $z_1 \approx z_2$ if

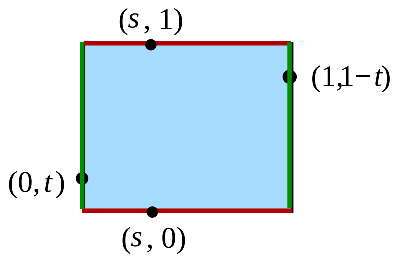
$$z_1 = z_2 \qquad \text{or} \qquad \overline{z_1} = z_2 \quad \& \quad |z_1| = 1 = |z_2| \ .$$

Show that $\mathbb{D}/\!\approx$ is also the sphere.

Example 4.12. By changing one of the identification of the torus, we obtain another surface called the Klein Bottle. This surface cannot be drawn without self-intersection in $\mathbb{R}^3$. Define an equivalence relation $\sim$ on $[0,1] \times [0,1]$ by $(s_1, t_1) \sim (s_2, t_2)$ if $(s_1, t_1) = (s_2, t_2)$ or

$$s_1 = s_2 \ \& \ |t_1 - t_2| = 1 \quad \text{or} \quad |s_1 - s_2| = 1 \ \& \ t_1 = 1 - t_2 \, .$$

Then $[0,1]^2/\sim$ is called the *Klein Bottle*. The identification is shown in the following picture.



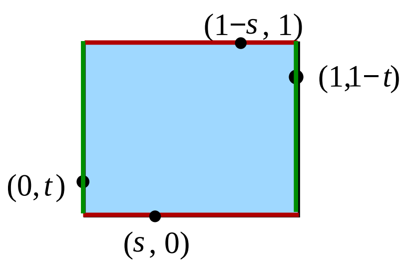
Note that after the top and bottom red edges are glued to each other. A cylinder is formed. Then if one tries to glue the left and right green edges within $\mathbb{R}^3$, one has to cut through the cylinder. Moreover, there is not a clear distinction of inside versus outside of the surface. This phenomenon is called non-orientable.

4.4.2. The Real Projective Plane. If we also change the identification of the other pair of sides, we have another non-orientable surface. This surface will be commonly seen in many branches of mathematics

Example 4.13. On $[0,1] \times [0,1]$, $(s_1, t_1) \sim (s_2, t_2)$ if $(s_1, t_1) = (s_2, t_2)$ or

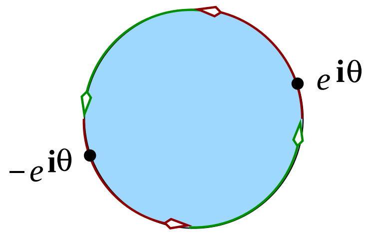
$$s_1 = 1 - s_2 \,\&\, |t_1 - t_2| = 1 \quad \text{or} \quad |s_1 - s_2| = 1 \,\&\, t_1 = 1 - t_2 \, .$$

Then $[0,1]^2/\sim$ is called the *Real Projective Plane*, denoted $\mathbb{R}\mathbf{P}^2$.



There are many ways to construct the Real Projective Plane as a quotient space. We will describe them one by one in the following.

Let $\mathbb{D} = \{ z \in \mathbb{C} : |z| \leq 1 \}$ be the unit disk. Define an equivalence relation $\sim$ on $\mathbb{D}$ by $z_1 \sim z_2$ if $z_1 = z_2$ or $|z_1| = 1 = |z_2|$ and $z_1 = -z_2$. That is, each pair of antipodal points on the boundary of $\mathbb{D}$ are identified.



It can be proved that $\mathbb{D}/{\sim}$ is also the Real Projective Plane, i.e., it is homeomorphic to the above one. The following exercise may be helpful.

Exercise 4.4.4. Let X, Y be topological spaces and $\sim$, $\approx$ are equivalence relations on X, Y respectively. If $h\colon X \to Y$ is a homeomorphism such that $x_1 \sim x_2$ if and only if $h(x_1) \approx h(x_2)$, then $X/{\sim}$ and $Y/{\approx}$ are homeomorphic.

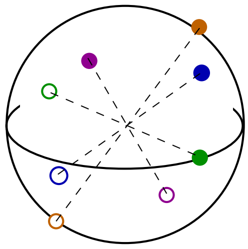
From this exercise, although $\mathbb{R}\mathbf{P}^2$ can be defined by one of the above identification, we usually prefer the second one (and the descriptons given below). It is because these descriptions work for higher dimensions. For example, $\mathbb{R}\mathbf{P}^n$ is obtained by $\mathbb{D}^n/\sim$ by identifying antipodal points on the boundary $\partial\mathbb{D}^n = \mathbb{S}^{n-1}$.

The third way to see the Real Projective Plane is the space of straight lines through the origin. In other words, the space of 1-dimensional subspace in $\mathbb{R}^3$. Define an equivalence relation $\sim$ on $\mathbb{R}^3 \setminus \{ \mathbf{0} \}$ by $\mathbf{x} \sim \mathbf{y}$ if there is $\lambda \neq 0$ such that $\mathbf{x} = \lambda\mathbf{y}$. In other words, if $\mathbf{x}$ and $\mathbf{y}$ determine the same 1-dimensional subspace in $\mathbb{R}^3$, then they are identified. Then the quotient space is also the Real Projective Plane. At this point, it may not be easy to set up a homeomorphism between this version and the pervious two versions. Readers may see why in the next definition of the Real Projective Plane.

Again, if one considers $\mathbb{R}^{n+1} \setminus \{\mathbf{0}\}$, i.e., 1-dimensional vector subspaces in $\mathbb{R}^{n+1}$, the result is the n-dimensional real projective space, $\mathbb{R}\mathbf{P}^n$. Similarly, 1-dimensional complex vector subspaces in $\mathbb{C}^{n+1}$ form the complex projective space $\mathbb{C}\mathbf{P}^n$.

In the above quotient space $(\mathbb{R}^3 \setminus \{ \mathbf{0} \})/{\sim}$, every equivalence class $[\mathbf{x}]$ is indeed a straight line passing through the origin determined by $\mathbf{x}$. Now, one may always choose in every equivalence class a representative $\mathbf{u}$ with $\|\mathbf{u}\| = 1$. In fact, there are two such representatives, a pair of antipodal points on the sphere $\mathbb{S}^2$. This leads to the fourth way of seeing the Real Projective Plane.

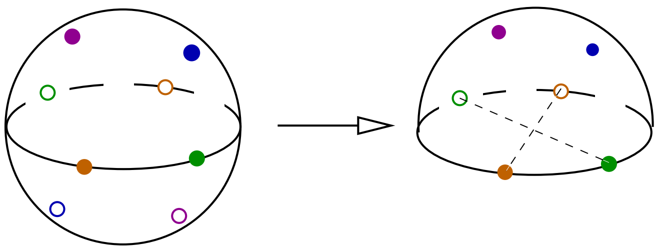
Let $\mathbb{S}^2 = \left\{ \mathbf{x} \in \mathbb{R}^3 \ : \ \|\mathbf{x}\| = 1 \right\}$ be given the standard topology. Define an equivalence relation on it by $\mathbf{x} \sim \mathbf{y}$ if $\mathbf{x} = \pm\mathbf{y}$.



Then it is easy to have a mapping

$$(\mathbb{R}^3 \setminus \{\,\mathbf{0}\,\})/\!\sim \quad \to \quad \mathbb{S}^2/\!\sim\, : \ [\mathbf{x}] \mapsto \left[\frac{\mathbf{x}}{\|\mathbf{x}\|}\right].$$

It can be shown that this is a homeomorphism and so both represent the Real Projective Plane. In fact, since on $\mathbb{S}^2/\!\sim$, two antipodal points are identified. One may take the representatives in the north hemisphere. On the open hemisphere, there is always a single representative. However, on the equator, two antipodal representatives are identified. This gives an intuition why $\mathbb{S}^2/\!\sim$ is homeomorphic to the earlier model of $\mathbb{D}/\!\sim$. Similarly, $\mathbb{R}\mathbf{P}^n$ can be defined as $\mathbb{S}^n/\sim$.



Exercise 4.4.5. Fill in the details that the above models are homeomorphic.

In fact, there is also algebraic way to see the Real Projective Plane, which we will discuss in the next section.

Exercise 4.4.6. Prove that $\mathbb{R}\mathbf{P}^1$ and $\mathbb{S}^1$ are homeomorphic.

4.5. Digression: Quotient Group

This section is written for readers who have some exposure to Abstract Algebra. It provides interesting examples of topological spaces which are commonly discussed in geometry. Beginners may skip this section.

Recall that the circle $\mathbb{S}^1$ may be seen as a quotient space of $\mathbb{R}$. Namely,

$$\mathbb{S}^1 = \mathbb{R}/\sim, \qquad \text{where } s \sim t \text{ if } (s-t) \in \mathbb{Z}.$$

In terms of Abstract Algebra, we are consider the additive group $(\mathbb{R}, +)$, in which $(\mathbb{Z}, +)$ is a normal subgroup. Then as a set, we have $\mathbb{R}/\sim$ is exactly the factor group $\mathbb{R}/\mathbb{Z}$. Furthermore, the group structure and topology structure of $\mathbb{R}$ are related. As a consequence, its quotient $\mathbb{S}^1$ also has related group and topology structures.

DEFINITION 4.14. A topological space $(X, \mathfrak{T})$ with a group structure $(X, \cdot)$ is called a *topological group* if the following two mappings are continuous,

$$(x_1, x_2) \in X \times X \mapsto x_1 \cdot x_2 \in X, \qquad x \in X \mapsto x^{-1} \in X\,.$$

In this sense, the standard $\mathbb{R}$ is a topological group.

Given a topological group $(X, \mathfrak{T})$ and let $A \subset X$ is a subgroup, which is given the induced topology $\mathfrak{T}|_A$. Then, it is clear that the mappings

$$(a_1, a_2) \in A \times A \mapsto a_1 \cdot a_2 \in A, \qquad a \in A \mapsto a^{-1} \in A$$

are continuous. This make A also a topological group, which is called a *topological subgroup* of X.

In addition, if A is a normal subgroup, we have the factor group X/A of cosets and it may be given the quotient topology. In such situation, the following two mappings

$$\begin{aligned} ([x_1], [x_2]) \in (X/A) \times (X/A) &\mapsto [x_1 \cdot x_2] \in (X/A), \\ [x] \in (X/A) &\mapsto \left[x^{-1}\right] \in (X/A) \end{aligned}$$

are well-defined and they are continuous under the quotient topology. Thus, X/A becomes a topological group also. The circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ can be seen as such space. Interestingly, the quotient group structure is exactly the multiplication inherited from $\mathbb{C}$, that is, $e^{\mathbf{i}\alpha} \cdot e^{\mathbf{i}\beta} = e^{\mathbf{i}(\alpha+\beta)}$.

Topological groups are very common in the study of topology and geometry. Many interesting examples come from matrix groups including Lie groups. The example of $\mathbb{R}$ and $\mathbb{S}^1$ can be regarded as a case of 1×1 matrix group.

EXAMPLE 4.15. Let $\mathcal{M}_n(\mathbb{R})$ be the set of all $n \times n$ matrices with real entries. Each matrix can be listed in a form of a vector in $\mathbb{R}^{n^2}$. Thus, $\mathcal{M}_n(\mathbb{R})$ is a topological space homeomorphic to $\mathbb{R}^{n^2}$. Matrix multiplication is continuous, but inverses may not be defined so it is only a semigroup.

Let $\mathrm{GL}(n,\mathbb{R}) = \{\, M \in \mathcal{M}_n(\mathbb{R}) \ : \ \det(M) \neq 0 \,\}$. It can be verified that $\mathrm{GL}(n,\mathbb{R})$ is a group with both multiplication and inverse mappings being continuous. Thus, it is a topological group.

Let $\mathrm{O}(n) = \mathrm{O}(n,\mathbb{R}) = \left\{\, Q \in \mathrm{GL}(n,\mathbb{R}) \ : \ QQ^T = Q^T Q = I \,\right\}$, that is, the group of orthogonal matrices. It is a topological subgroup of $\mathrm{GL}(n,\mathbb{R})$.

There is a natural way of regarding a smaller matrix as a larger matrix. Mathematically, define a mapping

$$\mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_{n+1}(\mathbb{R}) \ : \quad M \mapsto \begin{pmatrix} 1 & 0 \\ 0 & M \end{pmatrix}.$$

This mapping is clearly an injective homomorphism and continuous. Its restrictions also gives continuous monomorphisms $\mathrm{GL}(n,\mathbb{R}) \hookrightarrow \mathrm{GL}(n+1,\mathbb{R})$ and $\mathrm{O}(n) \hookrightarrow \mathrm{O}(n+1)$. Inductively, for all $m < n$, we have

$$\mathcal{M}_m(\mathbb{R}) \hookrightarrow \mathcal{M}_n(\mathbb{R}) \quad \text{etc.}$$

Example 4.16. Recall that $\mathbb{R}\mathbf{P}^2$ can be seen as the space of 1-dimensional vector subspaces in $\mathbb{R}^3$. This fact may be rewritten in terms of group action which is the fifth way to define $\mathbb{R}\mathbf{P}^2$. Let us first start with the sphere $\mathbb{S}^2$.

Consider $\mathrm{O}(2)$ as a topological subgroup of $\mathrm{O}(3)$. An element in the quotient topological $\mathrm{O}(3)/\mathrm{O}(2)$ is a coset of the form $Q \cdot \mathrm{O}(2)$ where $Q \in \mathrm{O}(3)$. Moreover,

$$Q_1 \cdot \mathrm{O}(2) = Q_2 \cdot \mathrm{O}(2) \text{ as elements in } \mathrm{O}(3)/\mathrm{O}(2)$$

if and only if $Q_1^{-1}Q_2 \in \mathrm{O}(2) \subset \mathrm{O}(3)$. Note that here $\mathrm{O}(2)$ is considered as a subgroup of $\mathrm{O}(3)$ and it only acts on the second and third entries. That is, $Q_1^{-1}Q_2$ is an isometry on the (x_2x_3)-plane while fixing the x_1-coordinate. Mathematically speaking, $Q_1^{-1}Q_2 \in \mathrm{O}(2)$ if and only if $Q_1(\mathbf{e}_1) = Q_2(\mathbf{e}_2)$ and

$$\left(Q_1^{-1}Q_2\right)\big|_{\{\,0\,\}\times\mathbb{R}^2} \ : \ \{\,0\,\} \times \mathbb{R}^2 \to \{\,0\,\} \times \mathbb{R}^2$$

is an isometry. Since both Q_1, Q_2 are orthogonal, the vector $Q_1(\mathbf{e}_1) = Q_2(\mathbf{e}_2)$ is of unit length, i.e., it belongs to $\mathbb{S}^2$. Therefore, we have a mapping

$$\mathrm{O}(3)/\mathrm{O}(2) \to \mathbb{S}^2 \ : \quad Q \cdot \mathrm{O}(2) \mapsto Q(\mathbf{e}_1).$$

It can be verified that this mapping is a homeomorphism. However, unlike the situation of $\mathbb{R}/\mathbb{Z}$, $\mathrm{O}(2)$ is NOT a normal subgroup of $\mathrm{O}(3)$. Therefore, $\mathrm{O}(3)/\mathrm{O}(2)$ and hence $\mathbb{S}^2$ is not a topological group.

Now, consider another subgroup $(\pm\mathrm{O}(2)) \subset \mathrm{O}(3)$ which contains all matrices of the form $\begin{pmatrix} \pm 1 & 0 \\ 0 & A \end{pmatrix}$ where A is a 2×2 orthogonal matrix. Clearly $(\pm\mathrm{O}(2))$ contains the embedded image of $\mathrm{O}(2)$. Equip the set of cosets $\mathrm{O}(3)/(\pm\mathrm{O}(2))$

with the quotient topology. Following the same argument as above, the only difference is if $Q_1^{-1}Q_2 \in (\pm\mathrm{O}(2))$ then we only have $Q_1(\mathbf{e}_1) = \pm Q_2(\mathbf{e}_1) \in \mathbb{S}^2$. Nevertheless, for the coset $Q(\pm\mathrm{O}(2))$, the set $\{ \pm Q(\mathbf{e}_1) \}$ is exactly the equivalence class representing a point in $\mathbb{R}\mathbf{P}^2$. Therefore, one can set up a homeomorphism

$$\mathrm{O}(3)/(\pm\mathrm{O}(2)) \to \mathbb{R}\mathbf{P}^2 \ : \quad Q\cdot(\pm\,\mathrm{O}(2)) \mapsto \{ \pm Q(\mathbf{e}_1) \} \,.$$

In addition, there is obviously a mapping

$$\mathrm{O}(3)/\mathrm{O}(2) \to \mathrm{O}(3)/(\pm\mathrm{O}(2)) \ : \ Q\cdot\mathrm{O}(2) \mapsto Q\cdot(\pm\mathrm{O}(2))$$

which is exactly the mapping $\mathbb{S}^2 \to \mathbb{R}\mathbf{P}^2$ taking antipodal points to the same point.

There is another way of seeing the above quotient. The normal subgroup $\{ I, -I \}$ of $\mathrm{GL}(n,\mathbb{R})$ induces a quotient group $\mathrm{O}(3)/\{\pm I\}$ and, in turns, it defines a set of cosets $\left(\mathrm{O}(3)/\{\pm I\}\right)/\mathrm{O}(2)$. Thus, we have a quotient topological space

$$\left(\mathrm{O}(3)/\{\pm I\}\right)/\mathrm{O}(2), \quad \text{which is homeomorphic to } \mathrm{O}(3)/(\pm\mathrm{O}(2))\,.$$

Example 4.17. Similarly, one can easily express $\mathbb{S}^n$ and $\mathbb{R}\mathbf{P}^n$ as quotient topological spaces $\mathrm{O}(n)/\mathrm{O}(n-1)$ and $\mathrm{O}(n)/(\pm\mathrm{O}(n-1))$ respectively. Describe this in the context of $\mathrm{O}(n)$-action on $\mathbb{R}^n$, orbit space, and isotropy group. Two vectors $\mathbf{v}, \vec{\boldsymbol{w}} \in \mathbb{R}^n$ are in the same orbit if there exists $Q \in \mathrm{O}(n)$ such that $Q(\mathbf{v}) = \mathbf{w}$. The isotropy group of an element $\mathbf{x} \in \mathbb{R}^n$ is a subgroup of $\mathrm{O}(n)$ containing all those Q with $Q(\mathbf{x}) = \mathbf{x}$. Clearly, the isotropy group of $\mathbf{e}_1$ is $\mathrm{O}(2)$. In this way, $\mathrm{O}(n)/\mathrm{O}(n-1)$ records those elements that do not fix $\mathbf{e}_1$.

Exercise 4.5.1. Furthermore, do similar thing for complex matrices and unitary matrices, and thus obtain the space $\mathbb{C}\mathbf{P}^n$.

Example 4.18. In the above example, $\mathbb{R}\mathbf{P}^n$ can be seen as the space of all 1-dimensional vector subspaces in $\mathbb{R}^n$. By taking quotient over $\mathrm{O}(n-1)$ in $\mathrm{O}(n)$, unit vectors in $\mathbb{R}^n$ are determined. But, each vector has a direction, so the 1-dimensional vector space has a direction. Further quotient over the subgroup $\{ \pm I \}$ serves the purpose of "taking the vector and its opposite direction". Thus, it gives a 1-dimensional vector space without specifying the direction.

With this understanding, one may try to look at the space of k-dimensional vector subspaces in $\mathbb{R}^n$. It is called a Grassmannian $G(n,k)$. It is not surprising that it can be obtained similarly by actions of topological groups, namely,

$$G(n,k) = \frac{\mathrm{O}(n)}{\mathrm{O}(k)\times\mathrm{O}(n-k)}\,.$$

In this quotient space, an element in the quotient of $\mathrm{O}(n-k)$ is determining an oriented k-dimensional subspace. Then $\mathrm{O}(k)$ is about all the orientations in that k-dimensional space. Thus, further quotient gives all the unoriented k-dimensional vector subspaces in $\mathbb{R}^n$.

4.6. Infinite Product

We have defined finite product spaces and mentioned that the method is not valid for infinite product. Moreover, in Section 4.3, we have given the second definition of the quotient topology wrt to a mapping. At the end of the section, characterizing properties of the quotient topology are illustrated in two exercises. It turns out that an infinite product space is better seen by similar characterizing properties.

4.6.1. Defining features. In Proposition 4.4, we have characterized the quotient topology. Let $(X, \mathfrak{T}_X)$ and $(Y, \mathfrak{T}_Y)$ be topological spaces such that

QT1: the mapping $q\colon X \to Y$ is surjective;
QT2: the mapping $q\colon (X, \mathfrak{T}_X) \to (Y, \mathfrak{T}_Y)$ is continuous;
QT3: the topology $\mathfrak{T}_Y$ on Y is the largest one to have **QT2**.

Then $\mathfrak{T}_Y$ is exactly the quotient topology $\mathfrak{T}_q$ wrt the mapping q.

There is another characterization property which provides us the convenience of checking continuity related to quotient space. In order to check the continuous of $g\colon (Y, \mathfrak{T}_Y) \to (Z, \mathfrak{T}_Z)$, it is equivalent to verify that $g \circ q\colon (X, \mathfrak{T}_X) \to (Z, \mathfrak{T}_Z)$ is continuous. This equivalence is valid and only valid when $\mathfrak{T}_Y$ is the quotient topology.

In the case of finite product, we also perform similar thing to check continuity of maps $f\colon W \to X_1 \times X_2$ by simply looking at $\pi_1 \circ f$ and $\pi_2 \circ f$. Therefore, we expect analogous characterizing properties, given as exercises below. Let $(X_1, \mathfrak{T}_1)$, $(X_2, \mathfrak{T}_2)$ be topological spaces and $\mathfrak{T}$ be a topology on $X_1 \times X_2$ with the projections $\pi_1\colon X_1 \times X_2 \to X_1$ and $\pi_2\colon X_1 \times X_2 \to X_2$.

Exercise 4.6.1. Show that if π_1 and π_2 are continuous under the topology $\mathfrak{T}$ and $\mathfrak{T}$ is the minimal one on $X_1 \times X_2$ to have both projections continuous, then $\mathfrak{T}$ is the product topology.

Likewise, there is another characterizing property.

EXERCISE 4.6.2. Assume that the $\mathfrak{T}$ on $X_1 \times X_2$ satisfies the following: for any topological space $(W, \mathfrak{T}_W)$ and any mapping $f\colon (W, \mathfrak{T}_W) \to (X_1 \times X_2, \mathfrak{T})$, f is continuous if and only if both $\pi_1 \circ f$ and $\pi_2 \circ f$ are continuous. Then $\mathfrak{T}$ is the product topology.

The characterizing properties given in Exercises 4.6.1 and 4.6.2 are very useful in product spaces. As such, we would expect a definition for an infinite product space which will keep these characterizing properties.

Let $X_\alpha, \alpha \in \mathcal{A}$ be a family of nonempty sets. The infinite product set $\prod_{\alpha\in\mathcal{A}} X_\alpha$ is the set of all functions

$$x \ : \mathcal{A} \to \bigcup_{\alpha\in\mathcal{A}} X_\alpha \qquad \text{such that for each } \alpha \in \mathcal{A},\ x(\alpha) \in X_\alpha.$$

We often denote x_α for the image $x(\alpha)$. Then, for each $\beta \in \mathcal{A}$, there is a surjective projection mapping

$$\pi_\beta \ : \ \prod_{\alpha\in\mathcal{A}} X_\alpha \to X_\beta, \qquad \pi_\beta(x) = x_\beta\,.$$

In addition, let $\mathfrak{T}_\alpha$ be a topology on X_α. Then the *product topology* $\mathfrak{T}_\prod$ is determined by the following properties.

PT1: Each π_β is surjecctive;
PT2: Each mapping $\pi_\beta \ : \ \left(\prod_{\alpha\in\mathcal{A}} X_\alpha, \mathfrak{T}_\prod\right) \to (X_\beta, \mathfrak{T}_\beta)$ is continuous;
PT3: The topology $\mathfrak{T}_\prod$ is the minimal one on $\prod_{\alpha\in\mathcal{A}} X_\alpha$ to have every π_β continuous.

From property **PT2**, if $V \in \mathfrak{T}_\beta$, then $\pi_\beta^{-1}(V) \in \mathfrak{T}_\prod$. Thus, we must have

$$\mathfrak{T}_\prod \supset \bigcup_{\beta\in\mathcal{A}} \left\{ \pi_\beta^{-1}(V) \ : \ V \in \mathfrak{T}_\beta \right\}.$$

Note that the above union is not a topology; in fact, it is not even a base. Nevertheless, it generates a topology which guarantees that each π_β is continuous. By **PT3**, it contains $\mathfrak{T}_\prod$. On the other hand, if each π_β is continuous, for each $V \in \mathfrak{T}_\beta$, $\pi_\beta^{-1}(V)$ must be open. Thus, this topology is exactly $\mathfrak{T}_\prod$.

In other words, we may form a base for the product topology, namely,

$$\mathcal{B} = \left\{ \pi_{\alpha_1}^{-1}(V_1) \cap \cdots \cap \pi_{\alpha_n}^{-1}(V_n) \ : \ V_k \in \mathfrak{T}_{\alpha_k},\ n \in \mathbb{N} \right\}.$$

Let us describe the base $\mathcal{B}$ from another point of view. A typical set in the above subbase is $\pi_\beta^{-1}(V)$ where $V \in \mathfrak{T}_\beta$. In other words, it is

$$\left(\prod_{\beta \neq \alpha \in A} X_\alpha \right) \times V \quad \overset{\text{if } \mathcal{A} = \mathbb{N}}{=\!=\!=} \quad X_1 \times \cdots X_{\beta-1} \times V \times X_{\beta+1} \times \cdots\cdots .$$

Each set in the base $\mathcal{B}$ is a finite intersection of the above sets, so it is a product of X_α except finitely many factors are replaced with open sets in corresponding spaces. Thus the base can be expressed as

$$\mathcal{B} = \left\{ \prod_{\alpha \in \mathcal{F}} U_\alpha \times \prod_{\alpha \in \mathcal{A} \setminus \mathcal{F}} X_\alpha : \ U_\alpha \in \mathfrak{T}_\alpha, \quad \text{finite } \mathcal{F} \subset \mathcal{A} \right\}.$$

In the case that $\mathcal{A} = \mathbb{N}$, each open set in $\mathcal{B}$ is of the form

$$X_1 \times \cdots X_{m_1-1} \times U_1 \times X_{m_1+1} \times \cdots X_{m_2-1} \times U_2 \times X_{m_2+1} \times \cdots\cdots ,$$

in which there are only finitely many such U's. In the case of finite product, the base $\mathcal{B}$ coincides with the previous definition.

Based on this definition, it is easy to prove the other defining property.

Exercise 4.6.3. For any topological space $(W, \mathfrak{T}_W)$ and any mapping

$$f\colon (W, \mathfrak{T}_W) \ \to \ \left(\prod_{\alpha \in \mathcal{A}} X_\alpha, \mathfrak{T}_\prod \right),$$

f is continuous if and only if each $\pi_\beta \circ f$ is continuous. Moreover, $\mathfrak{T}_\prod$ is the unique topology satisfying this property.

Example 4.19. Let $X_t = \mathbb{R}$ with the standard topology where $t \in \mathcal{A} = \mathbb{R}$. Then each element $x \in \prod_{t \in \mathcal{A}} X_t$ is indeed a function $x : \ \mathcal{A} = \mathbb{R} \to \mathbb{R} = \bigcup_{t \in \mathcal{A}} X_t$. In this situation, the infinite product space is simply the function space $\mathbb{R}^\mathbb{R}$. Similarly, we have function space $\mathbb{R}^{[a,b]}$, which contains $\mathcal{C}([a,b], \mathbb{R})$, the space of continuous functions on $[a,b]$. However, one should be careful that there can be many topologies or metrics on $\mathcal{C}([a,b], \mathbb{R})$. Only the suitable one is the same as the subspace topology induced from $\mathbb{R}^{[a,b]}$.

Exercise 4.6.4. Let $X = \mathbb{R}^\mathbb{N}$ be seen as a product space. Each element $x \in X$ is a function $x : \ \mathbb{N} \to \mathbb{R}$, i.e., a sequence of real numbers. Denote $0 \in X$ to be the constant zero function (sequence) and $x_n \in X$ the following sequence of functions (sequences)

$$x_n(k) = 0 \quad \text{for } k \leq n; \text{ and } \quad x_n(k) = 1 \quad \text{for } k > n.$$

Show that $x_n \to 0$ in the product space X.

Exercise 4.6.5. Let (X_k, d_k), $k \in \mathbb{N}$ be a countable family of metric spaces. Define a metric d on the product set $\prod_k X_k$ by

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{d_k(x_k, y_k)}{1 + d_k(x_k, y_k)} .$$

Show that the topology induced by d on $\prod_k X_k$ is exactly the product topology.

Given topological spaces $(X_\alpha, \mathfrak{T}_\alpha)$ and consider the set

$$\mathcal{B}_{\text{box}} \overset{\text{def}}{:==} \left\{ \prod_\alpha U_\alpha \ : \ U_\alpha \in \mathfrak{T}_\alpha \right\} .$$

Exercise 4.6.6. Show that $\mathcal{B}_{\text{box}}$ is a base for some topology of $\prod_\alpha X_\alpha$.

This topology defined by $\mathcal{B}_{\text{box}}$ is called the *box topology*, $\mathfrak{T}_{\text{box}}$ of the product set. Clearly, $\mathfrak{T}_{\prod} \subset \mathfrak{T}_{\text{box}}$. Also, when the index set for α is finite, they are equal.

Exercise 4.6.7. Show that the sequence in Exercise 4.6.4 does not converge in $(X, \mathfrak{T}_{\text{box}})$.

Exercise 4.6.8. For a general index set, under certain condition, the box topology and the product topology are the same. Guess the condition and justify it.

Hint. Compare the bases for box topology and product topology, what are the additional open sets in the box topology? When will these open sets become less?

CHAPTER 5

Compactness

In this chapter, one of the two ultimately important properties about topological spaces is discussed.

Perhaps, the most important property of topological spaces may be *compactness*. Many useful theorems are proved with this property. The main reason may be due to a certain sense of "finiteness" is guaranteed by this property.

In §5.1, the definition and some examples of compact spaces are introduced. Then, several useful theorems about compactness will be given in §5.2. In §5.3, the relation between compactness and Tychonoff separation axioms is discussed. In §5.4, we will discuss various concepts closely related to compactness and how they are logically connected. Finally in §5.5, locally compactness is briefly introduced with one-point compactification as a key application.

5.1. Compact Spaces and Sets

Every student will soon realize the importance of a closed interval early in an elementary course of analysis on $\mathbb{R}$. Later, when analysis is done on $\mathbb{R}^n$ of higher dimensions, sets that are closed and bounded become the attention.

Let us first obtain motivation by thinking of the good properties of a closed interval.

- If a sequence in it converges, the limit is also in it. As we have seen before, this is a matter that the set containing the sequence is closed. A closed interval has a stronger topological property.
- Some proofs on a closed interval may make use of nested intervals (see below). However, this is not the general property that we are looking for. As shown in previous sections, the fundamental reason for the method of nested intervals is that $[a, b]$ is a closed subset with finite diameter of a *complete metric space*.

- Every sequence in it must have a convergent subsequence. This is a result particularly true on closed and bounded sets. Therefore, it is likely the notion. The only shortcoming is that it involves sequences, i.e., the space must be of first countability. What about its proof? If nested intervals are not used, likely, we will need the next property below.
- We also have Heine-Borel Theorem, which is not easy to understand at the first encounter. This theorem perplexes many first time students for analysis. Usually, it is difficult to see what it means. Interestingly, this is exactly the significant notion that we are looking for.
- One may think of other good properties, such as, a continuous real-valued function on a closed interval must have maximum and minimum. This exactly requires the two most important properties in this chapter.

In a topological space $(X, \mathfrak{T})$, an *open cover* for X is collection of open sets of which the union is the whole space, i.e.,

$$\mathcal{C} = \{ U_\iota \}_{\iota \in \mathcal{I}} \subset \mathfrak{T} \quad \text{and} \quad X = \bigcup \mathcal{C} = \cup_{\iota \in \mathcal{I}} U_\iota \,.$$

A subset $\mathcal{E} \subset \mathcal{C}$, or $\{ U_\iota \}_{\iota \in \mathcal{J}}$ for $\mathcal{J} \subset \mathcal{I}$, which is also an open cover for X, i.e., $\bigcup \mathcal{E} = X$, is called a *subcover* of $\mathcal{C}$.

DEFINITION 5.1. A topological space $(X, \mathfrak{T})$ is *compact* if every open cover for X has a finite subcover. That is, for every $\mathcal{C} \subset \mathfrak{T}$ with $\bigcup \mathcal{C} = X$, there is a finite subset $\mathcal{E} \subset \mathcal{C}$ such that $\bigcup \mathcal{E} = X$.

Clearly, in this category of compact sets, one would like to include closed intervals in $\mathbb{R}$; more generally, closed and bounded sets in $\mathbb{R}^n$. Often, the following definition is more useful.

DEFINITION 5.2. A subset K in a topological space $(X, \mathfrak{T})$ (not necessarily compact itself) is *compact* if the induced space $(K, \mathfrak{T}|_K)$ is compact.

Since open sets in $\mathfrak{T}|_K$ are intersections of K with sets in $\mathfrak{T}$, it is easy to verify that K is compact if and only if for every $\mathcal{C} \subset \mathfrak{T}$ with $\bigcup \mathcal{C} \supset K$, there is a finite $\mathcal{E} \subset \mathcal{C}$ such that $\bigcup \mathcal{E} \supset K$.

EXAMPLE 5.3. As mentioned above, every closed and bounded subset in $\mathbb{R}^n$ is compact. This is exactly the content of Heine-Borel Theorem. Thus, the circle, the spheres $\mathbb{S}^n$, and the torus $\mathbb{T}$ are compact sets.

The whole Euclidean space $\mathbb{R}^n$ is not compact. Note that, first, $\mathbb{R}^n$ has a finite open cover, for example, $\{ \mathbb{R}^n \}$ has only one open set in it. Second, $\mathbb{R}^n$ also has an infinite open cover which has a finite subcover (try to give an example). However, it also has an infinite open cover that cannot be reduced to a finite subcover.

The half-closed interval $(0,1]$ is not compact. We may consider the open cover

$$\mathcal{C} = \{\, (1/n, 1] \; : \; n \in \mathbb{N} \,\} \; .$$

Any finite subcover of $\mathcal{C}$ is of the form

$$\mathcal{E} = \{\, (1/n_1, 1]\, , (1/n_2, 1]\, , \ldots , (1/n_k, 1] \,\} \; .$$

Its union $\bigcup \mathcal{E} = (1/n_m, 1] \neq (0,1]$ where $n_m = \max \{\, n_1, \ldots, n_k \,\}$.

Example 5.4. The set $\{\, 1/n \; : \; 1 \leq n \in \mathbb{Z} \,\} \cup \{\, 0 \,\}$ in $\mathbb{R}$ is compact. In fact, this is a particular case of a more general example. Let $(X, \mathfrak{T})$ a topogical space and x_n, $n \in \mathbb{N}$ be a sequence convergent to $x \in X$. Then the set $A = \{\, x_n \; : \; n \in \mathbb{N} \,\} \cup \{\, x \,\}$ is compact.

This is simply because if $\mathcal{C}$ is an open cover for A, then $x \in U_0$ for some $U_0 \in \mathcal{C}$. As $x_n \to x$, there is $N \in \mathbb{N}$ such that $\{\, x_n \; : \; n \geq N + 1 \,\} \subset U_0$. Also, for each $n \leq N$, there is an $U_n \in \mathcal{C}$ containing x_n. Thus, $\{\, U_0, U_1, \ldots, U_N \,\}$ is a finite subcover of $\mathcal{C}$.

Exercise 5.1.1.

(1) A family $\mathcal{F}$ of closed sets satisfies *finite intersection property* if every intersection of finitely many sets in $\mathcal{F}$ is nonempty. Prove that the following is equivalent to compactness: every family $\mathcal{F}$ of closed sets satisfying the finite intersection property must have $\cap \mathcal{F}$ nonempty.
(2) Show that a compact space X satisfies a property similar to the Cantor Intersection Theorem: if F_n are nonempty closed sets with $F_n \supset F_{n+1}$, then $\bigcap_{n=1}^{\infty} F_n$ is nonempty.
(3) Let $\mathcal{B}$ be a base for $\mathfrak{T}$. Assume that every open cover $\mathcal{C} \subset \mathcal{B}$ for X has a finite subcover. Prove that X is compact. *Remark*. The converse is trivially true.
Remark. The same question concerning subbase is considerably harder.
(4) Recall that a set S in a metric space (Y, d) is *bounded* iff $S \subset B(y_0, R)$ for some $y_0 \in Y$ and $R > 0$. Let (X, d) be a metric space. Prove that if $K \subset X$ is compact, it is closed and bounded.
Do you think the converse is true?
(5) Show that if a space $(X, \mathfrak{T})$ is compact and discrete then X is finite.

5.2. Compactness

Just as we have always seen in $\mathbb{R}^n$, that a subset is compact seems to be highly related to whether it is closed. This fact is partially true in general, in the sense that if the underlying topological space satisfies certain mild conditions. It turns out that the space is Hausdorff plays a key role.

EXERCISE 5.2.1.

(1) Use the indiscrete topology to create an example of a compact space X with a compact subset A which is not closed in X.
(2) Let K_α be compact subsets in a topological space $(X, \mathfrak{T})$. Prove that a finite union of K_α's is compact and, if X is Hausdorff, an arbitrary intersection of K_α's is compact.
Think about what happens to infinite union of compact sets.

THEOREM 5.5. *If $(X, \mathfrak{T})$ is a compact space and $A \subset X$ is closed, then A is a compact subset.*

Before the proof, we will also state a pseudo-converse of this theorem. These two theorems are often applied together because the requirements are commonly seen.

THEOREM 5.6. *If $(X, \mathfrak{T})$ is a Hausdorff space and $A \subset X$ is a compact set, then A is a closed subset of X.*

These two theorems may be schematically represented by the following diagram.

$$A \subset X \text{ is closed} \quad \underset{X \text{ is Hausdorff}}{\overset{X \text{ is compact}}{\rightleftarrows}} \quad A \subset X \text{ is compact}$$

The proof for the first one is simple from definition. It will be given here. The second one relates more to separation axioms so its proof will be given later.

PROOF OF THEOREM 5.5. Let $\mathcal{C} \subset \mathfrak{T}$ be an open cover for A, i.e., $\bigcup \mathcal{C} \supset A$. The natural way to use the compactness of X is to create an open cover from $\mathcal{C}$ for X. Since A is closed, $X \setminus A$ is open. Thus $\mathcal{C} \cup \{ X \setminus A \}$ is an open cover for X and it has a finite subcover $\mathcal{E}$. Then the collection $\mathcal{E} \setminus \{ X \setminus A \}$ is a finite subcover for A. □

EXERCISE 5.2.2. Let $(X, \mathfrak{T})$ be a Hausdorff space such that every proper subset of X is compact. Show that the topology is discrete. Do you think the converse is true?

THEOREM 5.7. *Let $f\colon (X, \mathfrak{T}_X) \to (Y, \mathfrak{T}_Y)$ be a continuous mapping. If X is compact, then its image $f(X) \subset Y$ is compact.*

In particular, if f is surjective, then Y is compact. From this, if f is a real-valued function, then $f(X)$ is a closed and bounded set. By boundedness, its supremum and infinmum exist; and by closedness, they both lie inside $f(X)$. Thus, they become maximum and minimum of f on X.

PROOF. Let $\mathcal{C} \subset \mathfrak{T}_Y$ be an open cover for $f(X)$, i.e., $\bigcup \mathcal{C} \supset f(X)$. Then the collection $\{ f^{-1}(V) \ : \ V \in \mathcal{C} \}$ is an open cover for X and so it has a finite subcover

$$\{ f^{-1}(V_1), f^{-1}(V_2), \ldots, f^{-1}(V_n) \} .$$

Correspondingly, $\{ V_1, V_2, \ldots, V_n \}$ is a finite subcover for $f(X)$. □

The above proof looks easy; especially when it is compared with the proof of existence of maximum and minimum in analysis. The main reason is that ε-δ arguments are now hidden in the proof. Nevertheless, this still shows the benefits of identifying the crucial topological concepts.

COROLLARY 5.8.
(1) *If X is compact, then its quotient space $X/\!\sim$ is also compact.*
(2) *If a product space $\prod_{\alpha\in\mathcal{I}} X_\alpha$ is compact, then each of its factors X_β is compact.*

Both are simple application of the above theorem by seeing that the continuous images under the quotient map or projection map. Naturally, especially about the product space, we are interested in whether the converse is true.

EXERCISE 5.2.3.
(1) Let $C(X) = \{ f\colon X \to \mathbb{R} \mid f \text{ is continuous} \}$. Prove that if X is compact, then $d(f,g) = \sup \{ |f(x) - g(x)| : \ x \in X \}$ is a metric on $C(X)$.
(2) A metric d is defined on $n \times n$ matrices by

$$d(A,B) = \left[\operatorname{tr}((A-B)(A-B)^T)\right]^{1/2} .$$

Convince yourself that the orthogonal matrices $\mathrm{O}(n)$ is compact but $\mathrm{SL}(n)$ is not, where

$$\begin{aligned} Q \in \mathrm{O}(n) \quad &\text{iff} \qquad QQ^T = Q^TQ = \text{identity}, \\ A \in \mathrm{SL}(n) \quad &\text{iff} \qquad \det(A) = 1 . \end{aligned}$$

(3) Prove that if $(X, \mathfrak{T})$ is compact and $f\colon (X, \mathfrak{T}) \to (Y, d)$ is continuous, then the image $f(X)$ is bounded.
(4) Let $f\colon X \to Y$ be continuous where X is compact and Y is Hausdorff. Show that for every closed set $F \subset X$, its projection $f(F)$ is closed.
(5) Let $p\colon X \to Y$ be a continuous surjective map such that $p(F)$ is closed for each closed set $F \subset X$ and $p^{-1}(y)$ is compact for each $y \in Y$. Show that if Y is compact, then X is compact.

Example 5.9. A quotient space $X/\!\sim$ is compact while the original space X may not be so. Consider $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, which is an obvious counter-example.

In the case of product space, given that each factor space is compact, it turns out that the product is also compact. We are going to give a proof for a finite product. The result about an arbitrary product is called Tychonoff Theorem, which is considerably more technical.

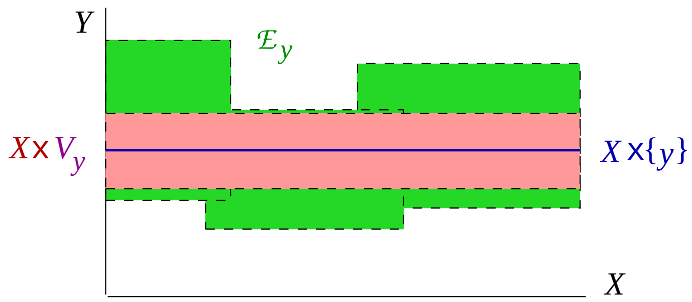
Theorem 5.10. *If $(X, \mathfrak{T}_X)$ and $(Y, \mathfrak{T}_Y)$ are compact spaces, then their product space $X \times Y$ is also compact.*

Proof. Let $\mathcal{C}$ be an open cover for $X \times Y$. For simplicity, we will assume that $\mathcal{C}$ belongs to the base $\{\, U \times V \;:\; U \in \mathfrak{T}_X, V \in \mathfrak{T}_Y \,\}$. This is valid according to an exercise in previous section. We will also discuss below how to further extend to general open cover.

Fix a $y \in Y$ and consider $X \times \{\, y \,\}$ as a subspace of $X \times Y$. The collection $\mathcal{C}$ is also an open cover for $X \times \{\, y \,\}$, which is homeomorphic to X and so is compact. Therefore, there is a finite subcover $\mathcal{E}_y$ for $X \times \{\, y \,\}$. Let

$$\mathcal{E}_y = \{\, U_1 \times V_1, U_2 \times V_2, \ldots, U_n \times V_n \,\} \qquad \text{where} \quad U_j \in \mathfrak{T}_X,\ V_j \in \mathfrak{T}_Y\,.$$

Take $V_y = \bigcap_{j=1}^n V_j$, then $X \times \{\, y \,\} \subset X \times V_y \subset \bigcup_{j=1}^n (U_j \times V_j) = \bigcup \mathcal{E}_y$ as shown in the following figure.



Perform this for each $y \in Y$, we have $\{ V_y : y \in Y \}$ an open cover for Y and so it has a finite subcover, $\{ V_{y_1}, V_{y_2}, \ldots, V_{y_m} \}$. Correspondingly, there are $\mathcal{E}_{y_j}$, $j = 1, \ldots, n$, each is a finite cover for $X \times \{ y_j \}$. Then $\mathcal{E} = \mathcal{E}_{y_1} \cup \cdots \cup \mathcal{E}_{y_m}$ is a finite cover for $X \times Y$.

In general, $\mathcal{C}$ may not belong to the base as above. Then we may work on

$$\mathcal{C}^* = \{ U \times V \ : \ U \in \mathfrak{T}_X, V \in \mathfrak{T}_Y, U \times V \subset G \text{ for some } G \in \mathcal{C} \} .$$

By the above method, a finite subcover of $\mathcal{C}^*$ can be found. Each $U_j \times V_j$ in this finite subcover is contained in an open set of $\mathcal{C}$. Therefore, we also have a finite subcover of $\mathcal{C}$. $\square$

Exercise 5.2.4. If X is compact and $H \subset X \times Y$ is closed, then its projection $\pi_Y(F)$ is closed in Y.

5.3. Compactness and Separation

Let us recall the two results that are partially converse to each other. That are Theorem 5.5 and Theorem 5.6. Schematically, we have

$$A \subset X \text{ is closed} \quad \underset{X \text{ is Hausdorff}}{\overset{X \text{ is compact}}{\rightleftarrows}} \quad A \subset X \text{ is compact}$$

We will first illustrate the power of these two results.

Theorem 5.11. *Let $(X, \mathfrak{T}_X)$ be compact and $(Y, \mathfrak{T}_Y)$ be Hausdorff. If $f \colon X \to Y$ is a continuous bijection, then f is a homeomorphism.*

Proof. It is sufficient to show that f is an open mapping, i.e., $f(U) \in \mathfrak{T}_Y$ for every open $U \in \mathfrak{T}_X$. Let $U \in \mathfrak{T}_X$, equivalently, $X \setminus U$ is a closed subset of the compact space X. By Theorem 5.5, $X \setminus U$ is compact. So, its continuous image $f(X \setminus U)$ is so. Since f is a bijection, we have $Y \setminus f(U) = f(X \setminus U)$, which is a compact subset of the Hausdorff space Y. By Theorem 5.6, $Y \setminus f(U)$ is closed, i.e., $f(U) \in \mathfrak{T}_Y$. $\square$

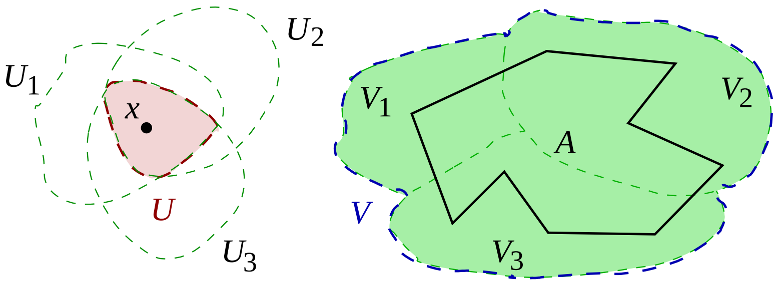
We are now ready to settle the theorem. Let us recall the statement: If $(X, \mathfrak{T})$ is Hausdorff and $A \subset X$ is compact, then A is closed in X.

Proof of Theorem 5.6. It is equivalent to establish that $X \setminus A$ is open. Precisely, let $x \in X \setminus A$, we will show that there is a $U \in \mathfrak{T}$ such that $x \in U \subset X \setminus A$. For each $a \in A$, clearly, $x \neq a$. Since X is Hausdorff, there are open sets $U_a, V_a \in \mathfrak{T}$ such that $U_a \cap V_a = \emptyset$; $x \in U_a$ and $a \in V_a$. In this way, $\{ V_a : a \in A \}$ is an open cover for A and hence it has a finite subcover, $\{ V_{a_1}, V_{a_2}, \dots, V_{a_n} \}$ and correspondingly finitely many $U_{a_1}, U_{a_2}, \dots, U_{a_n}$ such that $x \in U_{a_j}$ and $U_{a_j} \cap V_{a_j} = \emptyset$. Let

$$U = U_{a_1} \cap \cdots \cap U_{a_n} \qquad \text{and} \qquad V = V_{a_1} \cup \cdots \cup V_{a_n} .$$

Then $x \in U \in \mathfrak{T}$; $U \cap V = \emptyset$ and hence $U \subset X \setminus V \subset X \setminus A$. $\square$

The above proof can be illustrated by the following picture. This picture and the idea of the proof will be used again.



5.3.1. Separation Properties. In a Hausdorff space, two distinct points must have a reasonable separation between them. Roughly speaking, they are contained in "separated" open sets. In the above Theorem 5.6, at the end of the proof, it is indeed established that a point and a closed set are contained in "separated" open sets. Besides the Hausdorff property, there are several related weaker or stronger separation properties concerning points or closed sets. These properties are sometimes called Tychonoff Separation Axioms.

DEFINITION 5.12. A topological space $(X, \mathfrak{T})$ is:

Hausdorff or T_2 if for every $x \neq y \in X$, there are $U, V \in \mathfrak{T}$ such that $U \cap V = \emptyset$; $x \in U$ and $y \in V$.

T_1 if for every $x \neq y \in X$, there are $U, V \in \mathfrak{T}$ such that $x \in U$, $y \in V$, $x \notin V$, and $y \notin U$.

T_0 if for every $x \neq y \in X$, there exists $U \in \mathfrak{T}$ such that either $x \in U$ and $y \notin U$; or $x \notin U$ and $y \in U$.

regular if for every closed set $F \subset X$ and $x \notin F$, there are $U, V \in \mathfrak{T}$ such that $U \cap V = \emptyset$; $x \in U$ and $F \subset V$.

T_3 if it is T_1 and regular.

normal if for every closed sets $E, F \subset X$ with $E \cap F = \emptyset$, there are $U, V \in \mathfrak{T}$ such that $U \cap V = \emptyset$; $E \subset U$ and $F \subset V$.

T_4 if it is T_1 and normal.

EXERCISE 5.3.1. It can be check (by rewriting the definition) that a space is T_1 if and only if every singleton $\{ x \}$ is a closed subset.

Then, it follows easily that $T_4 \Longrightarrow T_3 \Longrightarrow T_2 \Longrightarrow T_1 \Longrightarrow T_0$.

EXERCISE 5.3.2.
(1) Let X be a Hausdorff space in which singleton is not open. Show that for any open set U and any $x \in X$, there is an open set $V \subset U$ such that $x \notin \overline{V}$.
(2) Show that an infinite set X with cofinite topology is T_1 but not T_2. What if the set is finite?
(3) Let $f, g \colon X \to Y$ be continuous mappings. What is the natural requirement for Y if you need one of the following:
    (a) The set $\{ x \in X : f(x) = y_0 \}$ is closed for every $y_0 \in Y$.
    (b) The set $\{ x \in X : f(x) = g(x) \}$ is closed.
(4) Show that X is normal if and only if for every closed sets $E, F \subset X$ with $E \cap F = \emptyset$, there exists a continuous function $f \colon X \to [0,1]$ such that $E \subset f^{-1}(0)$ and $F \subset f^{-1}(1)$.

5.3.2. Compact Hausdorff Spaces. Although there are descending implication from T_4 to T_2 in general, if the space is compact, these notions are indeed equivalent. The reasoning will make use the idea and picture in the proof of Theorem 5.6.

PROPOSITION 5.13. *A compact Hausdorff space is regular, and indeed T_3. Furthermore, it is normal, and indeed T_4.*

PROOF. Let $(X, \mathfrak{T})$ be a compact Hausdorff space. We will first show that it is regular. Let $F \subset X$ be closed and $x \notin F$. For each point $y \in F$, we have $x \neq y$. Since X is T_2, there are $U_y, V_y \in \mathfrak{T}$ such that $U_y \cap V_y = \emptyset$; $x \in U_y$ and $y \in V_y$. By Theorem 5.5, the set F is compact and so the open cover $\{ V_y : y \in F \}$ has a finite subcover $\{ V_{y_1}, \dots, V_{y_n} \}$. Now, it is easy to see that the same argument in Theorem 5.5 is applicable here. Finally, one has the required $U, V \in \mathfrak{T}$ such that $U \cap V = \emptyset$, $x \in U$ and $F \subset V$.

We may proceed one step further. Let $E, F \subset X$ be closed and $E \cap F = \emptyset$. We may apply the regular property obtained above to each point of E and the closed set F. Then the rest is similar. $\square$

EXERCISE 5.3.3. Fill in the details of the above proof.

There are two famous theorems that indicate the importance of normal spaces. One is about separation and the other about extension of continuous mappings. Both are necessary and sufficient conditions for nomality; in other words, they are only true on such spaces.

THEOREM 5.14 (Urysohn Lemma). *A topological space X is normal if and only if for every pair of closed sets $A, B \subset X$, there exists a continuous function $f: X \to [0,1]$ such that $f(A) \subset \{0\}$ and $f(B) \subset \{1\}$.*

THEOREM 5.15 (Tietz Extension Theorem). *A topological space X is normal if and only if for any closed $A \subset X$ and continuous function $f: A \to [0,1]$, there is an extension $\tilde{f}: X \to [0,1]$ such that $\tilde{f}|_A \equiv f$.*

This section is ended with some more definitions, which are included only as references.

DEFINITION 5.16. Some separation properties for a topological space X:

It is *completely regular* if for each closed $F \subset X$ and $x \notin F$, there is a continuous function $f\colon X \to [0,1]$ such that $F \subset f^{-1}(0)$ and $x \in f^{-1}(1)$.

It is $T_{3.5}$ or T_π if it is T_1 and completely regular.

It is *completely normal* if every subspace of X is normal.

It is T_5 if it is T_1 and complete normal.

It is *perfectly normal* if for every closed sets $E, F \subset X$ with $E \cap F = \emptyset$, there exists a continuous function $f\colon X \to [0,1]$ such that $E = f^{-1}(0)$ and $F = f^{-1}(1)$.

It is T_6 if it is T_1 and perfectly normal.

EXERCISE 5.3.4. (1) Let $(X, \mathfrak{T}_1)$ and $(X, \mathfrak{T}_2)$ are both compact Hausdorff spaces. Prove that if $\mathfrak{T}_1 \subset \mathfrak{T}_2$, then $\mathfrak{T}_1 = \mathfrak{T}_2$.

(2) Let Y be compact Hausdorff. For a mapping $f\colon X \to Y$, define

$$G = \{\, (x, f(x)) \in X \times Y \;:\; x \in X \,\} \,.$$

Prove that f is continuous if and only if G is a closed subset of $X \times Y$.

5.4. Locally Compactness

The Euclidean space $\mathbb{R}^n$ is not compact. Yet, it has many good properties because it is not far from one. The discussion in this section models a lot on the situation of $\mathbb{R}^n$ and $\mathbb{S}^n$.

DEFINITION 5.17. A topological space $(X, \mathfrak{T})$ is called *locally compact* if every $x \in X$ has a compact neighborhood. That is, there is a compact set $N \subset X$ such that $x \in \mathrm{Int}(N)$.

It is obvious that $\mathbb{R}^n$ is locally compact.

EXERCISE 5.4.1. (1) Is the co-finite topology locally compact?
(2) Give an example other than $\mathbb{R}^n$ that is locally compact.

REMARK . Warning: Usually, in other context of topology, we say that a space X is "locally $\star\star\star$" if for every point $x \in X$, there is a local base of "$\star\star\star$ neighborhoods". In the case of locally compactness, the definition is atypical. Obviously, if a space has local bases of compact neighborhoods everywhere, then it is locally compact. The converse is not always true. A typical example is $X = ([-1,1] \sqcup [-1,1])/\sim$ by identifying corresponding points except the origin. The points $(0,1)$ and $(0,-1)$ have compact neighborhoods, but they do not have a local base of compact neighborhoods.

EXERCISE 5.4.2. Let $(X, \mathfrak{T})$ be a Hausdorff space. Show that the following two statements are equivalent.

- Each $x \in X$ has a compact neighborhood N containing x.
- For each $x \in X$ and each neighborhood U of x, there is a compact set $K \subset U \subset X$ such that $x \in \mathrm{Int}(K)$. That is, compact neighborhoods form a local base at each $x \in X$.

One-point Compactification. A locally compact space is likely to be seen as a subspace of a compact space.

THEOREM 5.18. *Let $(X, \mathfrak{T})$ be a locally compact Hausdorff space. Then there is a compact space $(X*, \mathfrak{T}^*)$ such that*

(1) *$X^* \setminus X$ has exactly one point;*
(2) *$(X, \mathfrak{T})$ is a subspace of $(X^*, \mathfrak{T}^*)$, i.e., $\mathfrak{T} = \mathfrak{T}^*|_X$;*
(3) *the space $(X*, \mathfrak{T}^*)$ is also Hausdorff;*
(4) *if X is non-compact, then $\overline{X} = X^*$; otherwise, X is closed in X^*.*

PROOF. Let us start by defining X^* and $\mathfrak{T}^*$. Besides the four listed statements, we need to verify the topology $\mathfrak{T}^*$ and the compactness.

Take any point not in X, call it ∞; so we have $\infty \notin X$ and let $X^* = X \cup \{\infty\}$. Statement (1) is satisfied. Let

$$\mathfrak{T}^* \overset{\text{def}}{:=} \mathfrak{T} \cup \{ (X \setminus K) \cup \{\infty\} \ : \ K \subset X \text{ is compact} \} .$$

Note that since X is Hausdorff, by Theorem 5.5, $X \setminus K \in \mathfrak{T}$ whenever K is compact. Statement (2) is then clearly satisfied.

First, we will check that $\mathfrak{T}^*$ is closed under finite intersection. Clearly, if $U_1, U_2 \in \mathfrak{T}$ then $U_1 \cap U_2 \in \mathfrak{T} \subset \mathfrak{T}^*$. Also, if $K_1, K_2 \subset X$ are compact, then so is $K_1 \cap K_2$. Therefore

$$((X \setminus K_1) \cup \{\infty\}) \cap ((X \setminus K_2) \cup \{\infty\}) = (X \setminus (K_1 \cap K_2)) \cup \{\infty\} \in \mathfrak{T}^* .$$

For $U \in \mathfrak{T}$ and compact $K \subset X$, we also have

$$U \cap ((X \setminus K) \cup \{\infty\}) = U \cap (X \setminus K) \in \mathfrak{T} \subset \mathfrak{T}^* .$$

Second, observe that an arbitrary union of sets in $\mathfrak{T}^*$ is always of the form $U \cup (X \setminus K) \cup \{\infty\}$ for some $U_\alpha \in \mathfrak{T}$ and compact $K \subset X$. It is because

$$(\bigcup_\alpha U_\alpha) \cup \bigcup_\beta [(X \setminus K_\beta) \cup \{\infty\}] = V \cup (X \setminus K_{\beta_0}) \cup \{\infty\}$$

where $V = (\bigcup_\alpha U_\alpha) \cup \bigcup_{\beta \neq \beta_0} (X \setminus K_\beta) \in \mathfrak{T}$. Next, the set

$$U \cup (X \setminus K) \cup \{\infty\} = [X \setminus (K \setminus U)] \cup \{\infty\} \in \mathfrak{T}^*$$

because $K \setminus U$ is a closed subset of K and X is Hausdorff. Thus $K \setminus U$ is compact. Hence, $\mathfrak{T}^*$ is a topology on X^*.

Third, we will show that $(X^*, \mathfrak{T}^*)$ is compact. Let

$$\mathcal{C} = \{ U_\alpha \in \mathfrak{T} \}_{\alpha \in \mathcal{A}} \cup \{ (X \setminus K_\beta) \cup \{\infty\} \}_{\beta \in \mathcal{B}}$$

be an open cover for X^*. Take a compact set K_{β_0} where $\beta_0 \in \mathcal{B}$, then

$$\bigcup \mathcal{C} = \left(\bigcup_{\alpha \in \mathcal{A}} U_\alpha \right) \cup \left(\bigcup_{\beta \neq \beta_0} (X \setminus K_\beta) \right) \cup (X \setminus K_{\beta_0}) \cup \{\infty\} = X^* .$$

Then $\{ U_\alpha \in \mathfrak{T} \}_{\alpha \in \mathcal{A}} \cup \{ (X \setminus K_\beta) \}_{\beta \neq \beta_0}$ is an open cover for the compact set K_{β_0} and it has a finite subcover. Together with $(X \setminus K_{\beta_0}) \cup \{\infty\}$, a finite subcover of $\mathcal{C}$ is obtained for X^*.

Fourth, we will establish Statement (4). Note that by definition of $\mathfrak{T}^*$, X is compact if and only if $\{\infty\} = (X \setminus X) \cup \{\infty\} \in \mathfrak{T}^*$. If X is non-compact, then for any neighborhood $N = (X \setminus K) \cup \infty$ of the point ∞, $N \cap X = X \setminus K \neq \emptyset$. So, $\overline{X} = X^*$. On the other hand, if X is compact, ∞ is an isolated point in X^*.

Finally, we will show Statement (3) that $(X^*, \mathfrak{T}^*)$ is Hausdorff. The only case to handle if for a point $x \in X$ and $\infty \in X^*$. Since X is locally compact, there is a compact neighborhood K of x, that is, $x \in U \subset K$ for some $U \in \mathfrak{T}$. Then $U \cap (X \setminus K) = \emptyset$. Hence, $U \in \mathfrak{T}^*$ and $(X \setminus K) \cup \{ \infty \} \in \mathfrak{T}^*$ are the required open sets. □

Exercise 5.4.3.
(1) Show that the one-point compactification of $\mathbb{R}^n$ is homeomorphic to $\mathbb{S}^n$. Note that a possible homeomorphism is the stereographic projection.
(2) Can the real line $\mathbb{R}$ be compactified to $\mathbb{R} \cup \{ \pm\infty \}$?

5.5. Equivalences

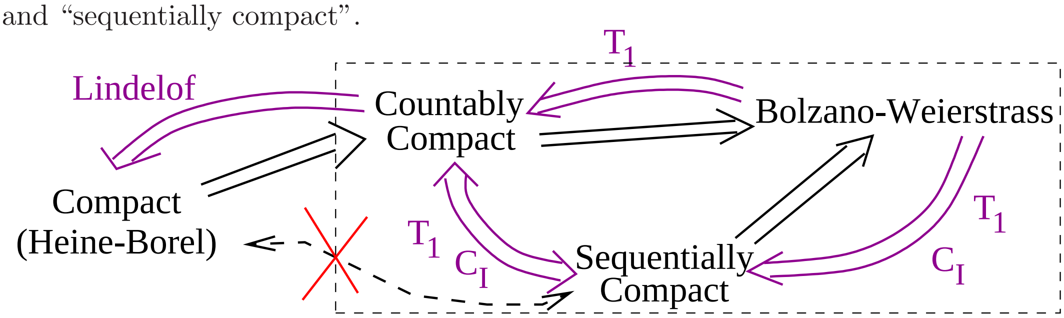
In the context of Euclidean spaces, several theorems are considered equivalent to the Heine-Borel Theorem. These are indeed properties related to compactness and they coincides with compactness under certain conditions. First, let us recall some definitions and define new properties.

Definition 5.19. A topological space $(X, \mathfrak{T})$ is said to be:

Lindelöf if every open cover for X has a countable subcover.
Countably compact if every countable open cover for X has a finite subcover.
Sequentially compact if every sequence in X has a convergent subsequence.
Having *Bolzano-Weierstrass Property* if every infinite subset $A \subset X$ has a cluster point (in X). This is called *limit point compact* by Munkres.

Theorem 5.20.
(1) *A compact topological space is always countably compact and has the Bolzano-Weierstrass property. A Lindelöf countably compact space is compact.*
(2) *A sequentially compact topological space has the Bolzano-Weierstrass property. A T_1 space of first countability with the Bolzano-Weierstrass property is sequentially compact.*
(3) *A countably compact topological space has the Bolzano-Weierstrass property. A T_1 space with the Bolzano-Weierstrass property is countably compact*
(4) *If a space is T_1 of first countability, then it is countably compact if and only if it is sequentially compact.*

The conditions for equivalences of the above properties are given in the schematic diagram below. Note that in general there is no implication between "compact" and "sequentially compact".



Example 5.21. The space $[0,1]^{\mathbb{N}}$ is an infinite product of compact spaces. Thus by Tychonoff Theorem, it is compact. However, as in Exercise 4.6.4, one may create a sequence in $[0,1]^{\mathbb{N}}$ that has no convergent subsequence. A complicated example of sequentially compact but non-compact space occurs in the order topology of uncountable ordinals.

It is trivial that compactness implies countably compactness. Moreover, A countably compact Lindelöf space is obviously compact. Statement (4) is a consequence of (2) and (3).

Sequentially compact $\Longrightarrow$ Bolzano-Weierstrass. Let $(X, \mathfrak{T})$ be a sequentially compact space and $A \subset X$ be infinite. Then it is possible to pick a *distinct* sequence $a_n \in A$. By sequentially compactness, it has a convergent subsequence a_{n_k}, $k \in \mathbb{N}$ with limit $x \in X$.

We claim that x is a cluster point of A. Take any open set $U \in \mathfrak{T}$ with $x \in U$, there is $K \in \mathbb{N}$ such that for every $k \geq K$, $a_{n_k} \in U$. Since a_{n_k}'s are distinct, there must be some $a_{n_k} \neq x$, i.e., $a_{n_k} \in A \cap (U \setminus \{x\})$. Thus, $x \in A'$. □

As we have seen before, the limit of a sequence in A is always a cluster point of A. But, a cluster point may not be always approached by a sequence. Basically, it is because a point may not have a countable local base, i.e., one cannot find a countable family of neighborhoods to represent all the neighborhoods.

Bolzano-Weierstrass $\overset{T_1, C_I}{\Longrightarrow}$ Sequentially Compact. Let x_n, $n \in \mathbb{N}$ be a sequence in X and $A = \{x_n : n \in \mathbb{N}\} \subset X$. If A is a finite set, then x_n clearly has a constant subsequence, which obviously converges. If A is infinite then by Bolzano-Weierstrass property, it has a cluster point $x \in X$. Since X is first countable, let $\{U_k : k \in \mathbb{N}\}$ be a countable local base at x.

For U_1, there is $x_{n_1} \in U_1 \setminus \{ x \}$. Then, since X is T_1, the set $\{ x_1, x_2, \ldots, x_{n_1} \}$ is closed. Thus, $(U_1 \cap U_2) \setminus \{ x_n \ : \ n \le n_1 \}$ is also a neighborhood of x, so there is $x_{n_2} \in (U_1 \cap U_2) \setminus \{ x, x_{n_1} \}$ for some $n_2 > n_1$. The process can be continued so that it is possible to pick

$$x_{n_{k+1}} \in (U_1 \cap \cdots \cap U_k \cap U_{k+1}) \setminus \{ x, x_{n_1}, \ldots, x_{n_k} \} \,, \qquad \text{for some } n_{k+1} > n_k.$$

Using the fact that $\{ U_k \ : \ k \in \mathbb{N} \}$ is a local base, one may show that $x_{n_k} \to x$ as $k \to \infty$. $\square$

Example 5.22. This example will be very useful for the understanding of the next proof. Consider the set of integers $\mathbb{Z} \subset \mathbb{R}$. It is an infinite set without any cluster point in $\mathbb{R}$. Let

$$\mathcal{C} = \{ \mathbb{R} \setminus \mathbb{Z} \} \cup \{ (n - 1/2, n + 1/2) \ : \ n \in \mathbb{Z} \} \,.$$

Then $\mathcal{C}$ is a countable open cover for $\mathbb{R}$. It clearly does not have a finite subcover. In this construction, $\mathbb{Z}$ being an infinite discrete closed subset in $\mathbb{R}$ is crucial.

Countably compact $\Longrightarrow$ Bolzano-Weierstrass. Since it is difficult to construct a cluster point for a set A simply from the assumption, we would start from the negation of Bolzano-Weierstrass property.

Let $A \subset X$ be an infinite set such that it has no cluster point in X. Without loss of generality, by taking a subset, we may assume that A is countable, say, $A = \{ a_n \ : \ n \in \mathbb{N} \}$ with $a_m \ne a_n$ for $m \ne n$. Now, we have the situation as in the example above. Since every a_n is not a cluster point of A, there exists $U_n \in \mathfrak{T}$ such that $a_n \in U_n$ and $U_n \cap A = \{ a_n \}$.

Next, we will show that $X \setminus A$ is open. Let $x \in X \setminus A$. Since x is not a cluster point of A, there is a $V \in \mathfrak{T}$ with $x \in V$ and $A \cap V = A \cap V \{ x \} = \emptyset$. This is equivalent to that $V \subset X \setminus A$. Thus, $X \setminus A$ is open.

Finally, $\{ X \setminus A \} \cup \{ U_n \ : \ n \in \mathbb{N} \}$ is a countable open cover for X which clearly does not have a finite subcover. $\square$

Bolzano-Weierstrass $\overset{T_1}{\Longrightarrow}$ Countably compact. Let $\{ U_n \in \mathfrak{T} \ : \ n \in \mathbb{N} \}$ be a countable open cover for X and we will look for a finite subcover. In order to use the given property, we try to pick $x_n \in U_n$ wisely to form an infinite set.

Let $x_1 \in X = \bigcup_{n=1}^{\infty} U_n$. Therefore, there exists n_1 (i.e., the first one) such that

$$x_1 \in U_{n_1} \qquad \text{but} \quad x_1 \notin U_1 \cup \cdots \cup U_{n_1 - 1} \,.$$

If $X \setminus (U_1 \cup \cdots \cup U_{n_1-1} \cup U_{n_1})$ is empty, then we have a finite subcover. Otherwise, pick $x_2 \in X \setminus (U_1 \cup \cdots \cup U_{n_1-1} \cup U_{n_1}) \subset \bigcup_{n>n_1} U_n$. By the construction, $x_2 \neq x_1$. Again, there exists n_2, which satisfies $n_2 > n_1$ consequently, such that

$$x_2 \in U_{n_2} \qquad \text{but} \quad x_1 \notin U_1 \cup \cdots \cup U_{n_2-1}\,.$$

Inductively, we have $x_k \in U_{n_k}$ such that $x_k \notin U_n$ for $n < n_k$ and $x_k \neq x_1, \ldots, x_{k-1}$. If the process stops in finitely many steps, then we have a finite subcover. Suppose it continues indefinitely, then we have a distinct sequence x_k constructed as above. The set $A = \{\, x_k \;:\; k \in \mathbb{N} \,\}$ is an infinite set. Take any $x \in X = \bigcup_{n=1}^{\infty} U_n$. Then $x \in U_m$ for some $m \in \mathbb{N}$ and $x \notin U_n$ for $n < m$. By ranking m among the increasing sequence n_k, we have

$$n_1 < n_2 < \cdots < n_N \leq m < n_{N+1} < n_{N+2} < \cdots\cdots\,.$$

Since $m < n_{N+1}$, by the construction process, $x_{N+1}, x_{N+2}, \ldots\ldots \notin U_m$. Therefore $U_m \cap A \subset \{\, x_1, \ldots, x_N \,\}$ is a finite set. Let

$$V = \begin{cases} U_m \setminus \{\, x_1, \ldots, x_{N-1}, x_N \,\} & \text{if } x \neq x_N, \\ U_m \setminus \{\, x_1, \ldots, x_{N-1} \,\} & \text{if } x = x_N, \text{ only occur when } m = n_N. \end{cases}$$

Since the space X is T_1, $V \in \mathfrak{T}$, $x \in V$ and $A \cap (V \setminus \{\, x \,\}) = \emptyset$. Thus, x is not a cluster point of A. Hence A is an infinite set without any cluster point. This contradicts the Bolzano-Weierstrass property. □

Exercise 5.5.1.

(1) Show that a space is countably compact if and only if every decreasing sequence of nonempty closed sets, $F_n \supset F_{n+1}$, has nonempty intersection $\bigcap_{n=1}^{\infty} F_n$.
(2) Let X has Bolzano-Weierstrass Property.
    (a) If $A \subset X$ is closed, then does A have the same property?
    (b) Does a continuous image of X also have the same property?
    (c) If X is a subspace of a Hausdorff space, is it a closed subset of it?
(3) A compact topological space often shares similar properties as a complete metric space. We have seen the Cantor Intersection Property before. Here is another one:
    Show that a contraction mapping on a compact metric space must have a fixed point. Do you think it is true for a space with other notion of compactness?

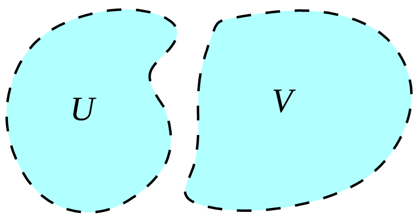
CHAPTER 6

Connectedness

Compactness discussed in the previous chapter guarantees that some sort of finiteness can be achieved. In this chapter, we discuss another important property on topological spaces, which is essential to have uniform results on the whole space. Without it, one part of a space may be very different from the other part of it. In §6.1, the concept of connectedness will be introduced, together with some examples and the standard skills of proving connectedness. In §6.2, we will discussed connected components and certain properties of connectedness. In the last section, §6.3, other related connnectivity will be explored.

6.1. Disconnected and Connected

In order to define connectedness, we start from the opposite. A topological space $(X, \mathfrak{T})$ is *disconnected* if there is a pair of nonempty open sets $U, V \in \mathfrak{T}$ such that $U \cap V = \emptyset$ and $U \cup V = X$. Intuitively, X can be separated into two nonempty pieces U and V. Such a pair of $U, V \in \mathfrak{T}$ is also called a *separation* of X.



Note that the condition of $U, V \in \mathfrak{T}$ is essential because any space can be written as $X = A \cup (X \setminus A)$ where A may not be open nor closed. Naturally, a connected space is not disconnected. But, we will give a better formulation later.

A subset $Y \subset X$ is disconnected if $(Y, \mathfrak{T}|_Y)$ is so. In such a case, one should be careful that the sets U, V are open *in* Y, but not necessarily in X.

If X is disconnected, with the above notations, then since $U = X \setminus V$ and $V \in \mathfrak{T}$, we have U and similarly V being also closed subsets of X. In other words, U and V are nontrivial (neither X nor $\emptyset$) both open and closed subsets of X.

EXAMPLE 6.1.
(1) Let $X = [0,1] \cup (2,3]$ with the induced topology of the standard $\mathbb{R}$. Then X is disconnected because $U = [0,1]$ is both *open and closed in* X. Note that $[0,1] = (-1, 1+\delta) \cap X = [0, 1.5] \cap X$ for an open interval $(-1, 1+\delta)$ and a closed interval $[0, 1.5]$. Thus, $[0,1]$ is both open and closed in X.
(2) Let $X = \left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1 \right\}$ with the induced topology of the standard $\mathbb{R}^2$. Then X is disconnected. Any branch of the hyperbola is a closed subset of $\mathbb{R}^2$, but it is both open and closed in X.

EXERCISE 6.1.1. Let $Y \subset X$ and there are disjoint $A, B \subset Y$ such that $Y = A \cup B$. If each $A \cap \overline{B} = \emptyset$ and $B \cap \overline{A} = \emptyset$, then the pair A, B is a separation of Y.

As seen above, one may only specify one sets instead of a pair. Namely, X is disconnected if there exists a both open and closed subset $U \subset X$ (i.e., U and $X \setminus U \in \mathfrak{T}$) such that neither $U = \emptyset$ nor $X \setminus U = \emptyset$. We are now ready to define a connected space by the negation of a disconnected space. The definition phrased as below is the most useful one, especially in abstract proofs.

DEFINITION 6.2. Let $(X, \mathfrak{T})$ be a topological space. It is *connected* if for each $S \subset X$ with both $S, X \setminus S \in \mathfrak{T}$, one must have $S = \emptyset$ or $S = X$. A subset $A \subset X$ is connected if the induced space $(A, \mathfrak{T}|_A)$ is so.

EXAMPLE 6.3.
(1) An indiscrete space is connected because any both open and closed subset must be trivial.
(2) A discrete space of more than one point is disconnected.
(3) As given above, $X = \left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1 \right\} \subset \mathbb{R}^2$ is disconnected. However, let

$$X_u = \left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1, x > 0 \right\}$$
$$X_\ell = \left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1, x < 0 \right\} .$$

Then both X_u and X_ℓ are connected. This is the concept of connected component which we will discussed later.

EXERCISE 6.1.2.
(1) Show that an infinite set X with the cofinite topology is connected. What if X is finite?
(2) Let X be connected and $\emptyset \neq A \subset X$. Later, we will see that if A is connected then so is $\overline{A}$. Give a counter-example of the converse.
(3) As above, if A is connected, is it necessary true that $\mathrm{Int}(A)$ is connected?
(4) Is $\mathrm{Frt}(A)$ connected if A is so? Is A connected if $\mathrm{Frt}(A)$ is so?

(5) Let X be connected and $\emptyset \neq A \subset X$. Prove that $\mathrm{Frt}(A) \neq \emptyset$.
(6) Let $A, C \subset X$ such that $C \cap A \neq \emptyset$ and $C \cap (X \setminus A) \neq \emptyset$. If C is connected then $C \cap \mathrm{Frt}(A) \neq \emptyset$.

6.1.1. Continuous Image. In the study of continuous functions on $\mathbb{R}$, it is known that the image of an interval is again an interval. This gives the idea that connectedness is preserved under continuous function. Here, in the proof, the typical use of the definition of connectedness is illustrated.

THEOREM 6.4. *Let X be a connected space. If $f\colon X \to Y$ is a continuous mapping, then its image $f(X) \subset Y$ is connected. In particular, if f is also surjective, then Y is connected.*

PROOF. Let $S \subset f(X)$ be both open and closed in $f(X)$. We are trying to show that $S = \emptyset$ or $f(X)$. Note that by taking the induced topology on $f(X)$ from Y, the mapping $f\colon X \to f(X)$ is also continuous. Thus $f^{-1}(S) \subset X$ is both open and closed in X. By the connectedness of X, $f^{-1}(S) = \emptyset$ or X. Consequently, $S = \emptyset$ or $f(X)$. □

Note that in the above proof, the continuity of $f\colon X \to f(X)$ is indeed easily verified. Nevertheless, it is still beneficial for the readers to pause and think to make sure about it.

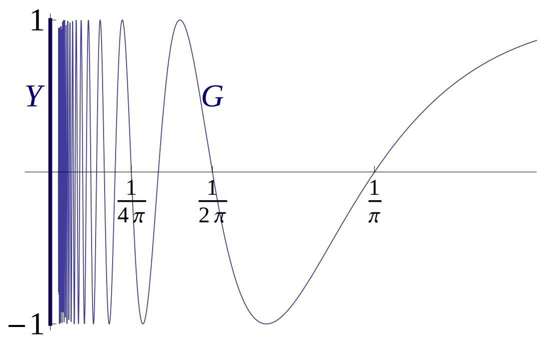
With this result, $Y = \big\{\, (0, y) \in \mathbb{R}^2 \;:\; y \in \mathbb{R} \,\big\}$ is a continuous (homeomorphic) image of $\mathbb{R}$ and so it is connected. The graph of a continuous function is also connected as seen in the exercise below.

EXERCISE 6.1.3. (1) If X is connected and $f\colon X \to Y$ is continuous, then the set $\{\, (x, f(x)) \in X \times Y \;:\; x \in X \,\}$ with the induced topology from the product space $X \times Y$ is also connected. Do you think the converse is true?
(2) Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function and $L_\alpha = \{\, \mathbf{x} \in \mathbb{R}^n \;:\; f(\mathbf{x}) = \alpha \,\}$, i.e., the level set wrt α.
  (a) If $A = L_\alpha \cup L_\beta$ with $\alpha \neq \beta$, show that A is disconnected.
  (b) Is it true that L_α is always connected?

6.1.2. A famous example. This is an example to show the strange behaviour of connectedness. One may also use a similar example to form a domain in $\mathbb{R}^2$ with a strange boundary. Let $X = Y \cup G$, which has a picture shown below, where

$$Y = \{ (0,y) \in \mathbb{R}^2 \ : \ y \in \mathbb{R} \}$$
$$G = \left\{ \left(x, \sin \frac{1}{x} \right) \in \mathbb{R}^2 \ : \ x > 0 \right\}$$

and X is given the induced topology of the standard $\mathbb{R}^2$.



This space X is a typical example of a connected space which is not path connected (to be discussed later). In order to show that X is connected, we will use the fact that both Y and G are connected, which is a consequence discussed in the previous section.

Let $S \subset X$ be both open and closed in X. Then $S \cap Y$ is both open and closed in Y. By the connectedness of Y, one must have $S \cap Y = \emptyset$ or Y. The same argument can be applied to $S \cap G$. Thus, we have the four possibilities $S = \emptyset$ or Y or G or X. We are going to show that neither Y nor G is both open and closed in X. Thus, $S = \emptyset$ or X and hence X is connected.

Since $G = X \setminus Y$, Y is both open and closed in X if and only if G is so. It is then sufficient to consider Y only. Let $U \subset \mathbb{R}^2$ be any open set containing Y. Then $U \supset A = \{ (0,y) \ : \ y \in [-1,1] \}$. Since A is compact, there is $\varepsilon > 0$ such that $U \supset (-\varepsilon, \varepsilon) \times [-1,1]$. For this ε, there is $m \in \mathbb{Z}$ such that $0 < 1/(2m\pi) < \varepsilon$ and so $G \cap U \neq \emptyset$. Therefore, if U is an open set of $\mathbb{R}^2$ such that $U \cap X \supseteq Y$, then $U \cap G \neq \emptyset$. This shows that Y cannot be open in X. Equivalently, G is not closed in X. $\square$

It should be noted that Y is indeed closed in X and thus G is open.

6.2. Components

As we have seen in a previous example, $\left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1 \right\}$ is disconnected but it is a union of two disjoint connected sets. In a way, a disconnected space is build up by pieces of connected sets.

Let X be a topological space and $x_0 \in X$.

Definition 6.5. A subset $C \subset X$ is the *connected component* of x_0 if any one of the following equivalent conditions hold.

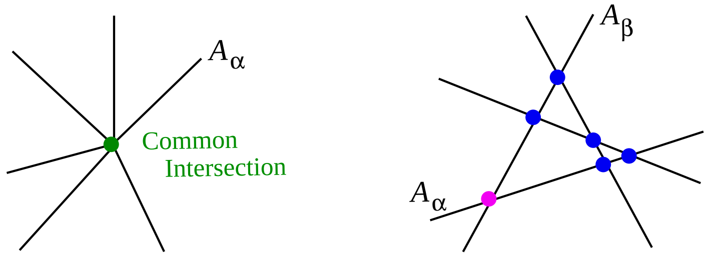
(1) C is the maximal/largest connected subset of X containing x_0; that is, if $A \subset X$ is connected and $x_0 \in A$, then $A \subset C$;
(2) $C = \bigcup \{ A \subset X \ : \ x_0 \in A \quad \text{and } A \text{ is connected} \}$;
(3) $C = [x_0]$, the equivalence class of x_0 wrt the equivalence relation $\sim$ where $x \sim y$ if there is a connected set $A \subset X$ such that $x, y \in A$.

Example 6.6. (1) In a discrete topological space, the connected component of x_0 is simply $\{ x_0 \}$.
(2) In $X = \left\{ (x,y) \in \mathbb{R}^2 \ : \ xy = 1 \right\}$; the upper and lower halves, $X_u = X \cap \{x > 0\}$ and $X_\ell = X \cap \{x < 0\}$, are the only two connected components.

Note that there are several things to consider in this definition. First, why are they equivalent? Second, the existence of C in condition (1) is actually given by condition (2). However, why is such a union in condition (2) is connected? Third, why is the relation $\sim$ given in condition (3) an equivalence relation? Interestingly, all the answers (Exercise 6.2.1) rely on the following theorem.

Theorem 6.7. *Let $A_\alpha \subset X$ be connected subspaces of X with either (i) $\bigcap_\alpha A_\alpha \neq \emptyset$ or (ii) $A_\alpha \cap A_\beta \neq \emptyset$ for each pair of indices α, β. Then $A = \bigcup_\alpha A_\alpha$ is connected.*

Pictorially, the two conditions can be represented by the diagram.



Exercise 6.2.1. Use Theorem 6.7 to show that the three definitions are valid.

Proof. Note that condition (i) implies (ii), so one only needs to prove with assuming (ii). Let $S \subset A = \bigcup_\alpha A_\alpha$ be both open and closed in A. As a consequence, for each index α, $S \cap A_\alpha$ is both open and closed in A_α. So, $S \cap A_\alpha = \emptyset$ or A_α for each α. Be careful that at this point we cannot simply conclude that $S = \bigcup(S \cap A_\alpha) = \bigcup \emptyset = \emptyset$ or $S = \bigcup(S \cap A_\alpha) = \bigcup A_\alpha = A$.

Suppose $S \cap A_\beta \neq \emptyset$ for some index β, then $S \cap A_\beta = A_\beta$ by connectedness of A_β. By condition (ii), for each α, $A_\alpha \cap S \supset A_\alpha \cap A_\beta \neq \emptyset$. Then by connectedness of A_α, one also has $S \cap A_\alpha = A_\alpha$. Thus, we have either $S \cap A_\alpha = \emptyset$ for each index α or $S \cap A_\alpha = A_\alpha$ for each index α. Hence, taking union over α, we have $S = \emptyset$ or $S = A$. □

Exercise 6.2.2.
(1) Prove the two variations of connectedness theorem:
   (a) Let A_α be a family of connected subsets in X and there is a connected subset C such that $C \cap A_\alpha \neq \emptyset$ for each α. Then $C \cup (\bigcup_\alpha A_\alpha)$ is also connected.
   (b) Let A_n be a countable family of connected subsets in X such that $A_n \cap A_{n+1} \neq \emptyset$ for all $n \in \mathbb{N}$. Then $\bigcup_n A_n$ is also connected.
(2) Let $f, g\colon (X, \mathfrak{T}_X) \to (Y, \mathfrak{T}_Y)$ be continuous functions and X is connected. Show that if there exists $x_0 \in X$ such that $f(x_0) = g(x_0)$ then $G_f \cup G_g$ is connected. Is the converse true?
(3) Let X, Y, Z be connected topological spaces and

$$f\colon (X, \mathfrak{T}_X) \to (Z, \mathfrak{T}_Z), \qquad g\colon (Y, \mathfrak{T}_Y) \to (Z, \mathfrak{T}_Z)$$

be continuous. Construct a quotient space $(X \sqcup Y)/\sim$ by $x \sim y$ if $f(x) = g(y)$. Show that if f or g is surjective, then $(X \sqcup Y)/\sim$ is connected.

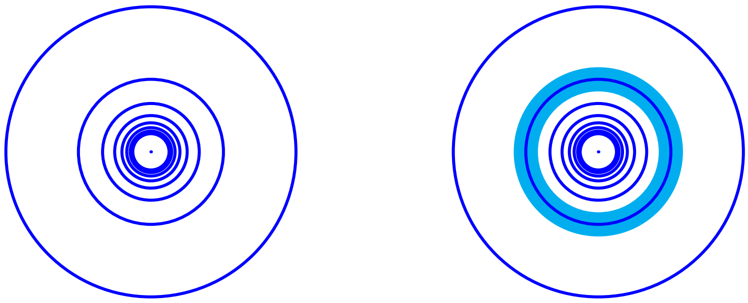
Remark. The result is intuitively obvious. Finding a clean proof may be the spirit of this exercise.
(4) Let X, Y be connected spaces and $A \subsetneq X$, $B \subsetneq Y$. Prove that $(X \times Y) \setminus (A \times B)$ is connected.

6.2.1. Open or closed? Let X be a disconnected space and C be one of its connected components, is C open or closed or both?

From the definition, $X = U \cup V$ for disjoint subsets $U, V \in \mathfrak{T} \setminus \{\emptyset, X\}$. Thus, both U and V are both open and closed. Intuitively, if either U or V is disconnected, we may further decompose it into both open and closed subsets. Therefore, it is natural (unfortunately wrong) to think that a component C is both open and closed.

EXAMPLE 6.8. Let $C_n = \{\, (x, y) \in \mathbb{R}^2 \;:\; x^2 + y^2 = \frac{1}{n^2} \,\}$, $C_\infty = \{\, (0, 0) \,\}$ and

$$X = \bigcup_{n=1}^{\infty} C_n \cup C_\infty \,.$$



Note that for each finite $n = 2, 3, \ldots$, there are $\dfrac{1}{n+1} < r_{n+1} < \dfrac{1}{n} < r_n < \dfrac{1}{n-1}$. Thus $C_n = X \cap \{\, (x, y) \;:\; r_{n+1} < x^2 + y^2 < r_n \,\}$ is open in X. Similarly, we may show that C_n is closed in X. Therefore, the components C_n are both open and closed. However, the situation for the component C_∞ is different. Any open set $\{\, (x, y) \;:\; x^2 + y^2 < \varepsilon \,\}$ containing C_∞ must also intersect some C_n for large n. Therefore, C_∞ is not open.

EXERCISE 6.2.3. (1) Show that C_∞ is closed in X.

(2) Is it true that if X has only finitely many connected components, then each one is both open and closed in X?

6.2.2. More about connectedness. Let us first consider an example of matrix topological space (in fact, a group), namely, $\mathrm{O}(n)$, the set of $n \times n$ orthogonal matrices with real entries. Recall that an $n \times n$ matrix $Q \in \mathrm{O}(n)$ if $Q^T Q = Q Q^T = I$. Clearly, $\mathrm{O}(n) \subset \mathbb{R}^{n^2}$ and it has an induced topology. The question is whether $\mathrm{O}(n)$ is connected or disconnected? Take the determinant mapping, which is continuous,

$$\det\colon\ \mathrm{O}(n) \;\to\; \mathbb{R}, \qquad Q \mapsto \det(Q)\,.$$

Note that $\det(Q^T)\det(Q) = \det(Q^T Q) = \det(I) = 1$. So, $\det(Q) = \pm 1$. Moreover, $\det(I) = 1$ and $\det(J) = -1$ where J is obtained by interchanging the first two rows and columns of I. Therefore, the image of $\mathrm{O}(n)$ under det is the discrete space $\{\, 1, -1 \,\}$, which is disconnected. By contra-positive of Theorem 6.4, $\mathrm{O}(n)$ must be disconnected. In fact, we can easily have the following general result.

PROPOSITION 6.9. *Let f be a continuous mapping on a topological space X. If its image $f(X)$ contains a discrete subset of at least two points, then X is disconnected.*

In the example of $\mathrm{O}(n)$, there is a subspace $\mathrm{SO}(n)$ which contains all orthogonal matrices of determinant 1. At this point, one should not simply conclude that $\mathrm{SO}(n)$ is connected because its image under determinant is so. One needs more information of the determinant mapping.

EXERCISE 6.2.4. Show that $\mathrm{SO}(n)$ is both open and closed in $\mathrm{O}(n)$. By this, argue that $\mathrm{SO}(n)$ is a connected component of $\mathrm{O}(n)$.

As a comparison, let us consider the set $\mathrm{U}(n)$ of $n \times n$ complex matrices such that $U \in \mathrm{U}(n)$ if $U^*U = UU^* = I$. That is the space of unitary matrices. By considering the determinant function again. One still has $|\det(U)| = 1$. So, the image is indeed $\{ z \in \mathbb{C} : |z| = 1 \}$, which is the unit circle. Thus, we cannot conclude that $\mathrm{U}(n)$ is disconnected. In fact, the connectedness of $\mathrm{U}(n)$ can be proved by below.

EXERCISE 6.2.5. Let $f \colon X \to Y$ be a mapping such that Y is having the quotient topology induced by f and is connected. Prove that if for all $y \in Y$, the subset $f^{-1}(y) \subset X$ is connected, then X is connected. Apply this result to show that $U(n)$, the unitary group, is connected.

THEOREM 6.10. *Let A be a connected subspace of X. Then every B with $A \subset B \subset \overline{A}$ is connected. In particular, $\overline{A}$ is connected.*

PROOF. Let $S \subset B$ be both open and closed in B. Thus, with the induced topology, $S \cap A$ is both open and closed in A. By connectedness of A, $S \cap A = \emptyset$ or $S \cap A = A$. First, consider that $S \cap A = \emptyset$. Thus, S is an open set and $S \subset B \setminus A$. We will argue that $x \in S \cap \overline{A}$ is a contradiction. Suppose otherwise, since $x \in \overline{A}$ and $x \in S \subset B$, $x \in \mathrm{Cl}_B(A)$. Therefore, any open neighborhood of x in B intersects A, in particular, $S \cap A$ is nonempty. Hence $S \cap A = \emptyset$ implies that $S \cap \overline{A}$ and so S is empty. Second, if $S \cap A = A$, then $S \subset B$ is a closed set containing A and so $S \supset \mathrm{Cl}_B(A) = \overline{A} \cap B = B$. To conclude, one has $S = \emptyset$ or $S = B$. □

EXERCISE 6.2.6. Prove that every connected component is closed.

EXERCISE 6.2.7. Let X be a compact Hausdorff space. If $\mathcal{F}$ is a set of closed connected subsets of X such that any two sets $F_\alpha, F_\beta \in \mathcal{F}$ satisfy $F_\alpha \subset F_\beta$ or vice versa, then $\bigcap \mathcal{F}$ is connected.

6.2.3. Two famous applications. There are two useful theorems in calculus that make use of connectedness. We give a brief description in order to highlight the importance of connectedness.

EXAMPLE 6.11. If $f\colon [a,b] \to \mathbb{R}$ is a continuous function and $y \in \mathbb{R}$ satisfies either $f(a) < y < f(b)$ or $f(b) < y < f(a)$, then there exist $x \in (a,b)$ such that $f(x) = y$. This is indeed INTERMEDIATE VALUE THEOREM.

In terms of connectedness, it can be proved as follows. Since $[a,b]$ is connected and f is continuous, $f([a,b])$ is connected and it is an interval J. Now, both $f(a), f(b) \in J$ and so either $[f(a), f(b)] \subset J$ or $[f(b), f(a)] \subset J$. Thus, y satisfying the assumption must be in J and the result follows.

EXAMPLE 6.12. If $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable on a domain Ω (i.e., open connected) and all its partial derivatives $f_j \equiv 0$, then f is a constant function.

Many students may use integration to prove this statement. That actually requires path connectedness (which is also true for a domain).

Recall that we need the Mean Value Theorem at $\mathbf{x}_0 \in \Omega$, which is only valid for points $\mathbf{x}$ with $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ for some $\delta > 0$. Thus, we already know that f is constant in a δ-ball of a point. To complete the proof, one may fix $\mathbf{x}_0 \in \Omega$ and consider $A = \{\, \mathbf{x} \in \Omega : \; f(\mathbf{x}) = f(\mathbf{x}_0) \,\}$. It can be shown that A is both open and closed. Since $\mathbf{x}_0 \in A$, it is nonempty so it must be the whole Ω.

Similar type of connectedness argument occurs very often in various studies of mathematics. Some requires further knowledge but it is good to understand this general nature through the above examples.

6.3. Other Connectivity

There are some further definitions of connectivity. Some of them require concepts in algebraic topology which we will introduce towards the end of this course. The most famous one is called simply-connected. In this section, we will only introduce two connectivity concepts which can be defined by point set topology.

6.3.1. Path connected. A topological space X is *path connected* if for every pair $x_1, x_2 \in X$, there exists a continuous path $\gamma\colon [0,1] \to X$ such that $\gamma(0) = x_1$ and $\gamma(1) = x_2$. In words, one can always join two points in X by a continous path.

EXERCISE 6.3.1. Show that a path connected space is connected.

Clearly, the converse of the exercise is not true. We have seen the example in Section 6.1.2, where

$$X = \{\, (0,y) \ :\ y \in \mathbb{R} \,\} \cup \left\{\, (x,y) \ :\ x > 0, y = \sin \frac{1}{x} \,\right\}.$$

We know that X is connected. However, if it is path connected, the function $\sin \frac{1}{x}$ must be discontinous at 0.

EXERCISE 6.3.2. Prove that there is no continuous $\gamma\colon [0,1] \to X$ such that $\gamma(0) = (0,0)$ and $\gamma(1) = \left(\frac{1}{\pi}, 0\right)$. This may be done by considering $t_n \to 0$ for $t_n \in [0,1]$.

It is clear that in $\mathbb{R}^n$, a convex set or a star-shaped set is path connected. Here, a set $A \subset \mathbb{R}^n$ is *star-shaped* if there exists $x_0 \in A$ such that for each $x \in A$, the line segment joining x_0 to x lies in A. An important fact in $\mathbb{R}^n$ is given in the following.

EXERCISE 6.3.3. Let $\Omega \subset \mathbb{R}^n$ be an open connected set. Then it is path connected. In fact, every two points in Ω can be joined a path formed by straight line segments parallel to the coordinate axes. *Hint.* Fix a point $x_0 \in \Omega$, let $S \subset \Omega$ be the set of all points $x \in \Omega$ such that x is joined to x_0 by coordinate-parallel segments. Show that S is both open and closed.

One may define an equivalence relation on X by $x_1 \sim x_2$ if there is a continuous path joining them. Each equivalence class is called a *path component*. Clearly, each path component is a subset of a connected component and they may be different from the example in Section 6.1.2. It should be noted that there is no conclusion on whether a path component is open or closed.

EXERCISE 6.3.4. Let $X = J \cup G$ where $G = \left\{\, (x,y) \ :\ x > 0, y = \sin \frac{1}{x} \,\right\}$ and $J = \{\, (0,y) \ :\ y \in [0,1] \cap \mathbb{Q} \,\}$. How many path components and connected components are there in X?

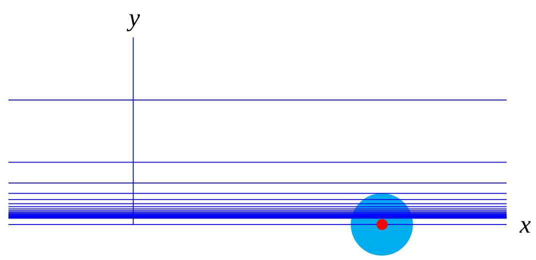
6.3.2. Locally connected. A topological space X is *locally connected* at $x \in X$ if the connected neighborhoods of x form a local base at x. That is, for every neighborhood N of x, there exists a connected neighborhood U of x such that $x \in U \subset N$. The space X is *locally connected* if it is so at every $x \in X$.

From the definition, it is clear that on a locally connected space, one may always assume that a neighborhood is connected. However, there is no implication between locally connected versus connected in any direction.

EXAMPLE 6.13. A locally connected space may not be connected. Let $X = (-1,0) \cup [2,3)$ with the standard induced topology. It is clearly locally connected but not connected.

A connected space may also not be locally connected. Let

$$F = \{\,(0,y) \;:\; y \geq 0\,\} \cup \{\,(x,0) \;:\; x \in \mathbb{R}\,\} \cup \bigcup_{n=1}^{\infty} \{\,(x,1/n) \;:\; x \in \mathbb{R}\,\}\,.$$



The space F is path connected as every point can be joined by a horizontal path to the y-axis and two points can be joined along the y-axis. However, it is not locally connected. At the point $(2,0) \in F$, take any neighborhood $U \subset F$ determined by a ball with center $(2,0)$ and radius $r < 2$, then U is disconnected.

Let X be locally connected; G be an open set and C be a connected component of G. Take any $x \in C$, then by local connectedness, there is a connected neighborhood U with $x \in U \subset G$. Since U is connected, it lies in a connected component, which must be C. Therefore, $x \in U \subset C$, i.e., $x \in \mathring{C}$. Hence, C is an open set. We thus have proved the following.

PROPOSITION 6.14. *If X is a locally connected space, then every connected component of an open subset is also open.*

EXERCISE 6.3.5. Prove the converse of the above proposition.

EXERCISE 6.3.6. Previously, we have seen that an open connected subset of $\mathbb{R}^n$ is path connected. What condition is necessary for a topological space X such that each open connected subset of X is path connected?

CHAPTER 7

Algebraic Topology

A major objective of topology study is to distinguish two abstractly defined topological spaces. To show that two spaces are the same, i.e., homeomorphic, one needs to establish a homeomorphism between them. To argue that two spaces are different, it is impossible to check the mappings between them one by one. It always is a matter of establishing a contradiction. In the study of topology, the contradiction often arises from the comparison of two algebraic objects determined by the given topological spaces. This is the key idea of algebraic topology.

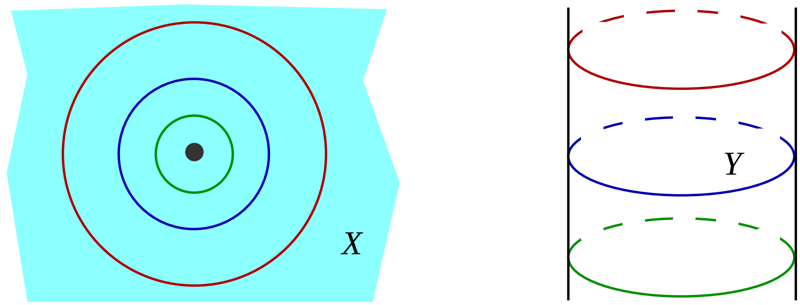
We will first discuss how invariants are exploited to distinguish different spaces in §7.1. Then, in §7.2, the concept of homotopy of maps is introduced and that it is an equivalence relation. In §7.3, the set of homotopy classes of mappings between two spaces is shown to be a simple topological invariant without mentioning its algebraic structure. Homotopy equivalences and homotopy type of a space are naturally brought into focus. Next, in §7.4, we focus our attention on homotopy of paths relative end-points. This naturally leads to the definition of fundamental group. Abundance of examples of fundamental groups are given in §7.5, with mostly intuitive arguments, so that readers will get a feeling of how it serves as an invariant. With certain familiarity of the definition, in §7.6, readers naturally sees the homotopy invariance of fundamental group. Finally, in §7.7, as an application of fundamental groups, the Brouwer's Fixed Point Theorem is proved.

7.1. Idea of Invariant

Before we start the study of algebraic topology, we will illustrate the direction and rationale of it. As mentioned above, the key is to distinguish two spaces.

7.1.1. Are they the same? Here are a series of examples to see if the given spaces are the same or different.

Example 7.1. Let $X = \mathbb{R}^2 \setminus \{ (0,0) \}$ with the standard induced topology and $Y = \mathbb{S}^1 \times (-\infty, \infty)$ with the standard product topology. Usually, X is called the *puunctured plane* while Y is called an *infinite cylinder*.



In order to show that X and Y are homeomorphic, one needs to construct a homeomorphism between them. For example, $\varphi\colon \mathbb{S}^1 \times \mathbb{R} \to \mathbb{R}^2 \setminus \{ (0,0) \}$ where

$$\varphi\left(e^{\mathbf{i}\theta}, t\right) = \left(e^t \cos\theta,\, e^t \sin\theta\right) .$$

Example 7.2. Let $X = \mathbb{R}$ and $Y = \mathbb{S}^1$, both endowed with standard topology. It is natural to expect that they are not homeomorphic. Clearly, it is impossible to examine every continuous mapping between $\mathbb{R}$ and $\mathbb{S}^1$. Thus, the method must be by contradiction. Assume that they are homeomorphic. Since $\mathbb{S}^1$ is compact and by Theorem 5.7, $\mathbb{R}$ must also be compact and it is a contradiction.

Example 7.3. Let us consider an example that both are compact, $X = [0,1]$ and $Y = \mathbb{S}^1$. We also expect that they are different but the above argument does not work. Here is one observation that may lead to a contradiction. $X \setminus \{ x_0 \}$ is disconnected except $x_0 = 0, 1$ while $Y \setminus \{ y_0 \}$ is connected for any $y_0 \in Y$.

Exercise 7.1.1. Suppose $f\colon X \to Y$ is a homeomorphism. Let $x_0 \in X$ and $f(x_0) = y_0$. Then $X \setminus \{ x_0 \}$ is homeomorphic to $Y \setminus \{ y_0 \}$.

7.1.2. General philosophy. Let us review the above examples from a general philosophical point of view. To show that two spaces are homeomorphic, basically, one has to establish the homeomorphism. On the other hand, proving that two spaces are different is by means of contradiction. In each case, we are actually setting up a function on the spaces.

- For a space X, we may assign $k(X) = \begin{cases} 1 & \text{if } X \text{ is compact} \\ -1 & \text{if } X \text{ is non-compact.} \end{cases}$. The function k satisfies that if X, Y are homeomorphic, then $k(X) = k(Y)$. Hence, from $k(\mathbb{R}) \neq k(\mathbb{S}^1)$, one concludes that $\mathbb{R}, \mathbb{S}^1$ are not homeomorphic.

- Similar, we may have $c(X)$ = number of connected components of X. Again, c satisfies that if X, Y are homeomorphic, then $c(X) = c(Y)$. However, $k([0,1]) = k(\mathbb{S}^1)$ and $c([0,1]) = c(\mathbb{S}^1)$.
- We may define $s(X) = \sup\{\, c\,(X \setminus \{x\}) \;:\; x \in X\,\}$. This function s also satisfies that if X, Y are homeomorphic, then $s(X) = s(Y)$. Now, $s([0,1]) = 2 \neq s(\mathbb{S}^1) = 1$.

Now, it can be seen that the key point is to have a function $\iota(X)$ for a topological space X such that if X, Y are homeomorphic, then $\iota(X) = \iota(Y)$. Such ι is called a *topological invariant*. Its values can be numbers (as the above examples), or polynomials, or vector spaces, etc.

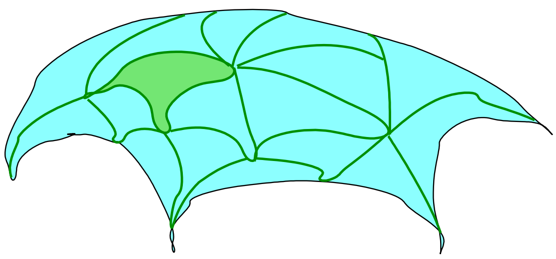
In many cases, there is an algebraic structure on a topological invariant. This provides good procedures of calculating the invariant value for a particular space. This is the reason why algebraic topology is so important.

Exercise 7.1.2. Use methods similar to the above, show that the three subspaces $\mathbb{S}^1$, $\mathbb{S}^1 \wedge \mathbb{S}^1$, and B of $\mathbb{R}^2$ are not homeomorphic, where

$$\mathbb{S}^1 \wedge \mathbb{S}^1 = \{\, z \in \mathbb{C} \;:\; |z-1| = 1 \text{ or } |z+1| = 1\,\}\,,$$
$$B = \{\, z \in \mathbb{C} \;:\; |z| = 1\,\} \cup \{\, z \in \mathbb{C} \;:\; z = \mathbf{i}y, y \in [0,1]\,\}\,.$$

7.1.3. Example: Euler Characteristic. Perhaps, the easiest algebraic invariant is called *Euler characteristic*. It can be defined for general topological spaces. For simplicity, we will focus on surfaces.

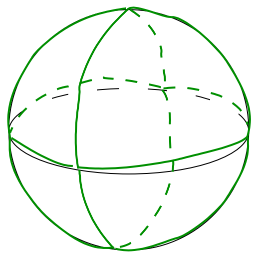
Let X be a surface. Some special subsets of X may be defined as a "triangle" in a suitable way. Each triangle has three "vertices" and three "edges". A *triangulation* Δ of X is a set of "triangles", $\Delta = \{\, T_\alpha\,\}$ satisfying certain conditions on the intersection of any two triangles, basically, $T_\alpha \cap T_\beta$ must be a common vertex or a common edge. Moreover, $X = \bigcap_\alpha T_\alpha$. Note that a surface X may have many triangulations. An illustrative picture is given below.



For a given triangulation Δ, one may "count" the following numbers, V is the number of vertices, E is the number of edges, and F is the number of triangles (faces). It is interesting that though these three numbers may changes for different triangulations, however, the number, $\chi(X) \overset{\text{def}}{:=} V - E + F$, is always the same for any triangulation. This number $\chi(X)$ is called the *Euler characteristic* of X.

REMARK . We mentioned the requirement on triangulation above. It is to govern how two triangles intersect each other. Besides topological reasons, this also guarantees that the counting will not be confused. Then, it can be proved that the sum $V - E + F$ is independent of the choice of triangulation.

EXAMPLE 7.4. Consider the triangulation of the sphere $\mathbb{S}^2$ shown below, we have $\chi(\mathbb{S}^2) = 6 - 12 + 8 = 2$.

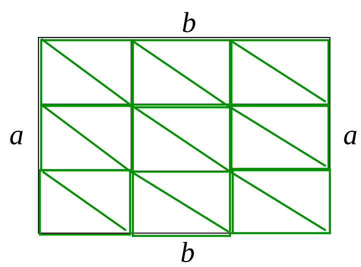


This triangulation is obtained by "blowing up" a octahedron like a ballon. Other triangulations can be obtained by other regular solid and the number of vertices, edges, and faces are shown below.

Solid	V	E	F	$\chi(\mathbb{S}^2)$
Tetrahedron	4	6	4	2
Cube	8	12	6	2
Octahedron	6	12	8	2
Dodecahedron	20	30	12	2
Icosahedron	12	30	20	2

In fact, there are many other non-regular non-symmetric triangulations on $\mathbb{S}^2$. Each gives $\chi(\mathbb{S}^2) = 2$.

EXAMPLE 7.5. The following determines a usual triangulation on the torus $\mathbb{T}$. From it, one can conclude that $\chi(\mathbb{T}) = 9 - 27 + 18 = 0$.



Interestingly, this triangulation and the counting do not depend on how the opposite sides a or b are glued. Therefore,

$$\chi(\mathbb{T}) = \chi(\mathbb{R}\mathbf{P}^2) = \chi(\mathbb{K}) = 0\,,$$

where $\mathbb{R}\mathbf{P}^2$ is the real projective plane and $\mathbb{K}$ is the Klein bottle.

PROPOSITION 7.6. *If X, Y are homeomorphic, then $\chi(X) = \chi(Y)$.*

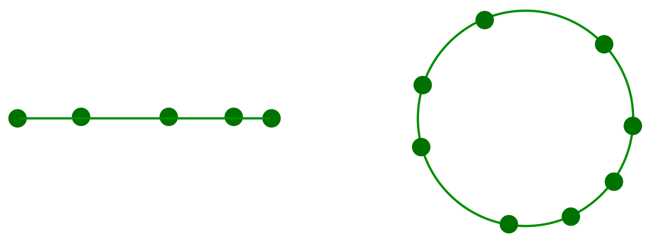
The proof is in fact not difficult, but it involves a lengthy machinery of algebraic topology, which we have not enough time to cover. From the example of torus, projective plane, and Klein bottle, the converse is not true. Furthermore, Euler characteristic is not only defined for surfaces as we will discuss below. It obeys several algebraic rules and so it is very useful. For example, here is a simple one.

PROPOSITION 7.7. *Given topological spaces X, Y and their product space $X \times Y$,*

$$\chi(X \times Y) = \chi(X) \cdot \chi(Y)\,.$$

The Euler characteristic is not only defined for surfaces. It is also defined for other dimensions and even combined objects of several dimensions. We will only give an intuitive introduction here.

EXAMPLE 7.8. On 1-dimensional topological spaces, a triangulation is roughly a division of the space into "curly intervals". In this case, there is no "faces", so $\chi(X) = V - E$. For example, one may observe the triangulations for an interval and a circle from the picture below.



Moreover, from it, one sees that $\chi([a, b]) = n-(n-1) = 1$ and $\chi(\mathbb{S}^1) = n-n = 0$. This gives another proof that $[a, b]$ and $\mathbb{S}^1$ are not homeomorphic.

The analogue of a "triangle" in a 3-dimensional space is a "curly tetrahedron". In this case, besides V, E, F, we still have T, the number of tetrahedra, and

$$\chi(X) \mathrel{:\stackrel{\text{def}}{=}} V - E + F - T\,.$$

In general, there is the concept of *k-simplex* and the cases of $k = 0, 1, 2, 3$ are points, edges, triangles, tetrahedra. Many topological spaces are made up by k-simplices of various dimensions. Again, there are requirements on their pairwise intersection. Then one may count the numbers of points, edges, triangles, etc. to get the following value for Euler characteristic,

$$\chi(X) :\stackrel{\text{def}}{=\!=} \sum_{k=0}^{n} (-1)^k N_k, \qquad N_k = \text{number of } k\text{-simplices in } X.$$

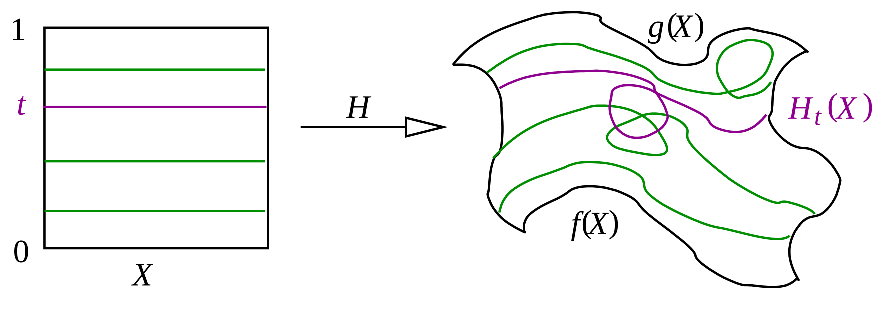
7.2. Homotopy

Very often, two homemorphic spaces are imagined as that one space can be continuously deformed to another. This indeed is not exactly correct. Nevertheless, continuous deformation is an important concept in topology. It is in some ways related to homeomorphism; and it plays a particular important role in algebraic topology.

Definition 7.9. Let X, Y be topological spaces. Two continuous mappings $f, g \colon X \to Y$ are *homotopic*, denoted $f \simeq g$ or $f \overset{H}{\simeq} g$, if there is a continuous mapping $H \colon X \times [0,1] \to Y$, called a *homotopy between* f, g such that,

$$H(x,0) = f(x) \quad \text{and} \quad H(x,1) = g(x) \qquad \text{for each } x \in X\,.$$

It is customarily to denote $H_t \colon X \to Y$, $t \in [0,1]$, the mapping defined by $H_t(x) :\stackrel{\text{def}}{=\!=} H(x,t)$ for $x \in X$. Then the above simply means that H_t is a continous family of continous mappings such that $H_0 \equiv f$ and $H_1 \equiv g$. Moreover, a homotopy is often visualized by the following picture.



In the above illustration, the space X is drawn as an interval. This is of course because of simplicity; yet it actually reflects an important case, namely, the homotopy of paths. Furthermore, the little loops and intersection of purple and green arcs demonstrate that the mapping H_t may not be one-to-one. Indeed, it may happen that $H(x_1, t_1) = H(x_2, t_2)$ for $x_1 \neq x_2$ or $t_1 \neq t_2$.

EXAMPLE 7.10. Let $R_\alpha \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the rotation of an angle α at the origin. That is,

$$R_\alpha \quad \leftrightarrow \quad \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Then for every $\alpha, \beta \in \mathbb{R}$, $R_\alpha \simeq R_\beta$ by a homotopy

$$H(\vec{\boldsymbol{x}}, t) = R_{\theta(t)}(\vec{\boldsymbol{x}}), \quad \vec{\boldsymbol{x}} \in \mathbb{R}^2,$$

where $\theta(t) = (1-t)\alpha + t\beta$ for $t \in [0,1]$. Intuitively, one may gradually increase/decrease the rotation angle to deform from R_α to R_β.

It should be noted that a rotation and a reflection in $\mathbb{R}^2$ are not homotopic. Clearly, this has to be proved by contradiction. In fact, it can be proved by considering a suitable algebraic invariant.

EXAMPLE 7.11. Consider the identity mapping id on $\mathbb{S}^1 \times \mathbb{R}$ and the mapping $f \colon \mathbb{S}^1 \times \mathbb{R} \to \mathbb{S}^1 \times \mathbb{R}$ given by

$$f\left(e^{\mathbf{i}\theta}, s\right) = \left(e^{\mathbf{i}\theta}, \frac{s}{1+|s|}\right), \quad e^{\mathbf{i}\theta} \in \mathbb{S}^1, s \in \mathbb{R}.$$

Note that the image of f is $\mathbb{S}^2 \times (-1,1)$. Its action is like compressing the infinitely long cylinder into a short cylinder. It can be seen that id $\simeq f$ by the homotopy

$$H\left(e^{\mathbf{i}\theta}, s; t\right) = \left(e^{\mathbf{i}\theta}, \frac{s}{1+t\,|s|}\right).$$

EXERCISE 7.2.1. (1) Let $\mathcal{M}$ be the set of all $n \times n$ real matrices. Any matrix $f \in \mathcal{M}$ can be seen as a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$.

(a) Show that any $f, g \in \mathcal{M}$ are homotopic.

(b) Is the homotopy between f, g above only involves mappings in $\mathcal{M}$? That is, there exists a homotopy $H \colon \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ between f, g such that for each $t \in [0,1]$, the mapping $x \mapsto H(x,t)$ also belongs to $\mathcal{M}$. We call it a homotopy through mappings in $\mathcal{M}$.

(c) Let $\mathcal{A} \subset \mathcal{M}$ be the subset of invertible matrices and $f, g \in \mathcal{A}$. Are they homotopic through mappings in $\mathcal{A}$?

(d) If $f, g \in \mathcal{P}$, the set of positive definite matrices, then there is a homotopy between f and g through mappings in $\mathcal{P}$.

(2) Let $\mathcal{M}$ be the set of all $n \times n$ real matrices. It can be given a topology induced by the standard $\mathbb{R}^{n^2}$. Show that $\mathcal{M}$ is path connected if and only if every pair of $f, g \in \mathcal{M}$ are homotopic through mappings in $\mathcal{M}$.

Among mappings between two spaces, there is always a special mapping, which is considered as the trivial element.

Definition 7.12. A mapping $\mathfrak{c}\colon X \to Y$ is a constant map at $y_0 \in Y$ if $\mathfrak{c}(x) = y_0$ for each $x \in X$. A mapping $f\colon X \to Y$ is called *null homotopic* if $f \simeq \mathfrak{c}$ for a constant map $\mathfrak{c}$ at some y_0.

Example 7.13. Any mapping $f\colon X \to \mathbb{R}^n$, $n \geq 1$ is null homotopic. In fact, the continuous mapping

$$H(x,t) = t\, f(x), \quad x \in X$$

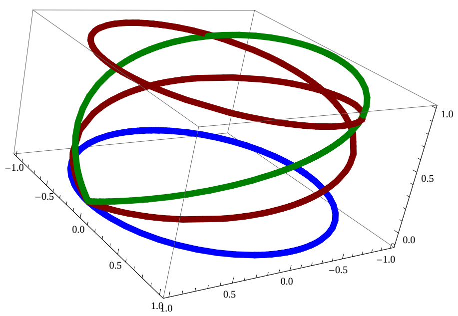
is a homotopy between f and the constant map at the origin.

Exercise 7.2.2. From the above, clearly, we have $(\star)$ that the identity mapping $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$ is homotopic to the constant map. Prove that $(\star)$ implies that any map is null homotopic. Note that this equivalence is valid for any space X, not only $\mathbb{R}^n$.

Example 7.14. Consider three mappings $f, g, h\colon \mathbb{S}^1 \to \mathbb{S}^1 \times \mathbb{R}$, where

$$\begin{aligned} f(e^{\mathbf{i}\theta}) &= \left(e^{\mathbf{i}\theta},\, 0\right) ; \\ g(e^{\mathbf{i}\theta}) &= \left(e^{\mathbf{i}\theta},\, \sin\tfrac{\theta}{2}\right) ; \\ h(e^{\mathbf{i}\theta}) &= \left(e^{\mathbf{i}(2\theta)},\, \sin\tfrac{\theta}{2}\right) . \end{aligned}$$

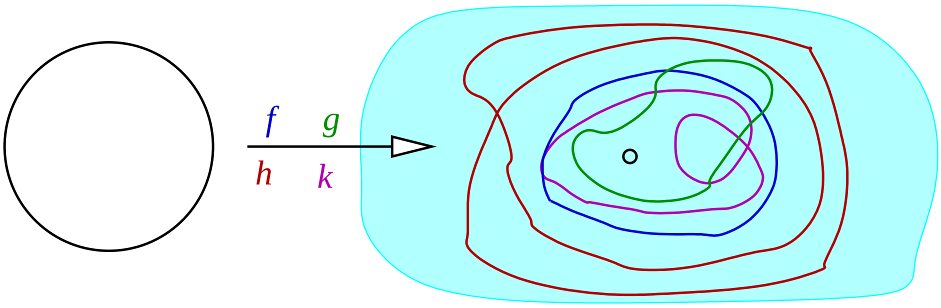
Their images in the cylinder $\mathbb{S}^1 \times \mathbb{R}$ are shown in the following picture according to their colors.



Note that if they are considered as mappings into $\mathbb{R}^3$, they are all homotopic (in fact, to a constant map). However, it is a different story when they are seen as mappings into $\mathbb{S}^1 \times \mathbb{R}$.

Exercise 7.2.3. By giving explicitly a homotopy, show that $f \simeq g$ in the above example. Intuitively, one would expect that $f, g \not\simeq h$, which may need more study later to prove.

Example 7.15. An example analogous to the above is the punctured plane, $\mathbb{R}^2 \setminus \{(0,0)\}$. Suppose four mappings f, g, h, k from $\mathbb{S}^1$ into $\mathbb{R}^2 \setminus \{(0,0)\}$ have their images shown in the following picture.



It is expected that $f \simeq g \simeq k \not\simeq h$.

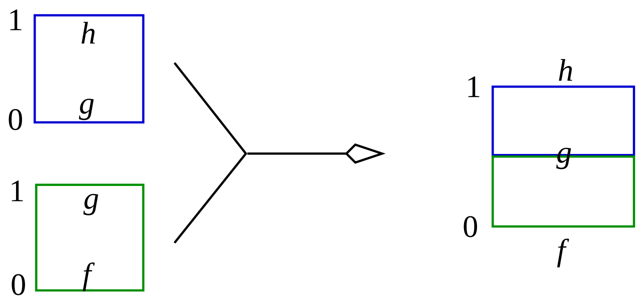
To conclude this section, we will see that homotopy of maps is an equivalence relation. More precisely, on the set of all continuous mappings between two topological spaces X, Y, the relation of homotopy is reflexive, symmetric, and transitive.

Reflexivity is obvious. For symmetry, let $f \overset{H}{\simeq} g$ and define $K\colon X \times [0,1] \to Y$ by

$$K(x,t) = H(x, 1-t)\,.$$

This essentially reverse the "time" of $[0,1]$ and gives a homotopy with $K_0 \equiv g$ and $K_1 \equiv f$.

Finally, let $f \overset{H}{\simeq} g$ and $g \overset{K}{\simeq} h$. We are usually the following schematic figure to construct the combined homotopy.



Mathematically, we define $L\colon X \times [0,1] \to Y$ by

$$L(x,t) = \begin{cases} K(x, 2t-1) & t \in \left[\frac{1}{2}, 1\right]\,; \\ H(x, 2t) & t \in \left[0, \frac{1}{2}\right]\,. \end{cases}$$

EXERCISE 7.2.4. If $f_1 \simeq g_1\colon X \to Y_1$ and $f_2 \simeq g_2\colon X \to Y_2$, show that $(f_1, f_2) \simeq (g_1, g_2)$ as mappings $X \to (Y_1 \times Y_2)$, where $(f_1, f_2)(x) = (f_1(x), f_2(x))$ and $(g_1, g_2)(x) = (g_1(x), g_2(x))$.

Here is a related version of homotopy that we will use later.

DEFINITION 7.16. Let X, Y be topological spaces, $A \subset X$. Two mappings $f, g \colon X \to Y$ are *homotopic rel* A, usually written $f \simeq g$ rel A, if there exists a homotopy $H \colon X \times [0,1] \to Y$ such that, besides $H_0 \equiv f$ and $H_1 \equiv g$,

$$H(a,t) = f(a) = g(a) \text{ for all } a \in A \text{ and for all } t \in [0,1].$$

Note that from the definition, f and g must already have the same restriction on A, i.e., $f|_A \equiv g|_A$. Moreover, there is a weaker version called *homotopy preserving* A, which only requires $H_t(A) \subset A$ for all $t \in [0,1]$.

EXERCISE 7.2.5. Let $h \colon [0,1] \to [0,1]$ be a homeomorphism such that $h(0) = 0$ and $h(1) = 1$. Show that $h \simeq \mathrm{id}_{[0,1]}$ rel $\{0,1\}$.

EXERCISE 7.2.6.

(1) Prove that both homotopy rel A and homotopy preserving A are equivalence relations on mappings.
(2) Let X, Y, Z be spaces with $A \subset X$ and $B \subset Y$. Let $f_0, f_1 \colon X \to Y$ satisfy $f_0(A) = f_1(A) \subset B$ and $g_1 \colon Y \to X$. If $f_0 \overset{F}{\simeq} f_1$ rel A and $g_0 \overset{G}{\simeq} g_1$ rel B, is it true that $g_0 \circ f_0 \simeq g_1 \circ f_1$ rel A?

7.3. Homotopy Classes and Homotopy Equivalences

7.3.1. Homotopy Classes. In the previous section, we have established the equivalence relation defined by homotopy of mappings. Let X, Y be topological spaces and $\mathcal{C}(X,Y)$ be the set of all continuous mappings from X to Y.

DEFINITION 7.17. The quotient set $\mathcal{C}(X,Y)/\simeq$ under the homotopy relation $\simeq$ is denoted $[X,Y]$ and any $[f] \in [X,Y]$ is called a *homotopy class* of f.

As we have seen above, since every continuous mapping into $\mathbb{R}^n$, $n \geq 1$, is homotopic to a constant map, $[X, \mathbb{R}^n]$ is a singleton set. This is true in particular for $[\mathbb{S}^1, \mathbb{R}^n]$. On the other hand, from the intuitive argument, the set $[\mathbb{S}^1, \mathbb{R}^2 \setminus \{(0,0)\}]$ contains at least two different elements. It is natural to expect that

$$[\mathbb{S}^1, \mathbb{R}^n] \neq [\mathbb{S}^1, \mathbb{R}^2 \setminus \{(0,0)\}] \quad \Longrightarrow \quad \mathbb{R}^n \neq \mathbb{R}^2 \setminus \{(0,0)\},$$

where the equality of two topological spaces really means homeomorphic.

In order to establish such implication, we will need the following results and apply the contra-positive.

PROPOSITION 7.18.

- *Let X, Y_1, Y_2 be topological spaces. If Y_1 is homeomorphic to Y_2, then there is a bijection between $[X, Y_1]$ and $[X, Y_2]$.*
- *Let X_1, X_2, Y be topological spaces. If X_1 is homeomorphic to X_2, then there is a bijection between $[X_1, Y]$ and $[X_2, Y]$.*

Let us try to formulate a proof for the first statement. Let $g\colon Y_1 \to Y_2$ be a homeomorphism and we would like to define a mapping $[X, Y_1] \to [X, Y_2]$. Take any $[f] \in [X, Y_1]$ where $f\colon X \to Y_1$. Clearly, we have $g \circ f\colon X \to Y_2$ and it naturally corresponds to $[g \circ f] \in [X, Y_2]$. The first thing we need is that such correspondence is well-defined, that is,

$$[f_0] = [f_1] \quad \Longrightarrow [g \circ f_0] = [g \circ f_1] \,.$$

Next, we will need to prove that the correspondence is one-to-one, that is,

$$[g \circ f_0] = [g \circ f_1] \quad \Longrightarrow [f_0] = [f_1] \,.$$

Similar things will be needed in proving the second statement. All these can be obtained from the following theorem.

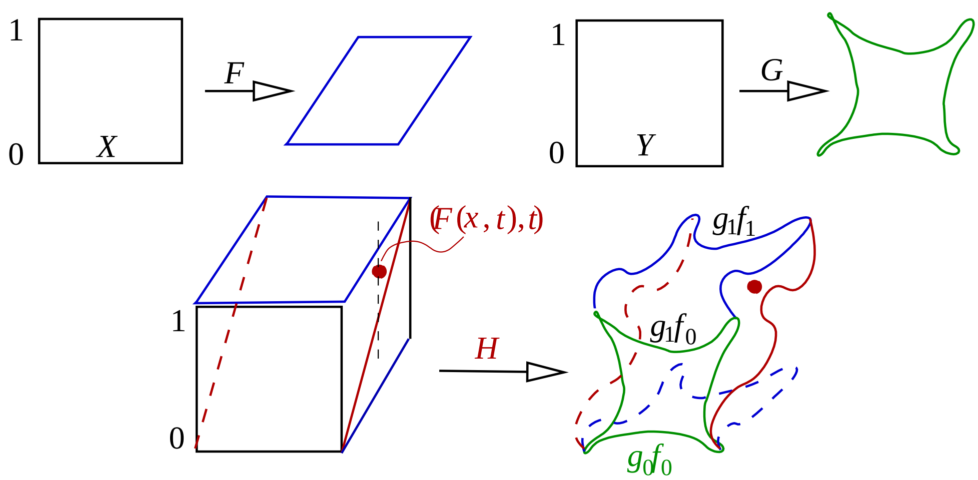
THEOREM 7.19. *Let X, Y, Z be topological spaces. If $f_0 \simeq f_1\colon X \to Y$ and $g_0 \simeq g_1\colon Y \to Z$, then*

$$(g_0 \circ f_0) \simeq (g_1 \circ f_1)\colon X \to Z \,.$$

PROOF. Suppose $f_0 \overset{F}{\simeq} f_1$ and $g_0 \overset{G}{\simeq} g_1$ where

$$(x, t) \in X \times [0, 1] \mapsto F(x, t) \in Y, \qquad (y, t) \in Y \times [0, 1] \mapsto G(y, t) \in Z \,.$$

We will define a homotopy $H\colon X \times [0, 1] \to Z$ by $H(x, t) \overset{\text{def}}{=\!=} G(F(x, t), t)$ as illustrated in the following figure.



Then it is clear that

$$H(x,0) = G(F(x,0),0) = G(f_0(x),0) = g_0 f_0(x)$$
$$H(x,1) = G(F(x,1),1) = G(f_1(x),1) = g_1 f_1(x)\,.$$

Obviously, indeed we proved that $g_0 \circ f_0 \simeq g_0 \circ f_1 \simeq g_1 \circ f_0 \simeq g_1 \circ f_1$. □

Exercise 7.3.1.

(1) Formulate and prove the analogous statement about $g_0 \circ f_0 \simeq g_1 \circ f_1$ rel A for $A \subset X$.
(2) If $f_1 \simeq g_1 \colon X \to Y_1$ and $f_2 \simeq g_2 \colon X \to Y_2$, show that $(f_1, f_2) \simeq (g_1, g_2)$ as mappings $X \to (Y_1 \times Y_2)$, where $(f_1, f_2)(x) = (f_1(x), f_2(x))$ and $(g_1, g_2)(x) = (g_1(x), g_2(x))$.
(3) Let Y be any topological space. Form the quotient space $CY = (Y \times [0,1])/{\sim}$ by the equivalence relation $\sim$ on $Y \times [0,1]$ with $(y_1, t_1) \sim (y_2, t_2)$ if $t_1 = 1 = t_2$. That is, CY is obtained by crushing the "top" $Y \times \{1\}$ to one point. Prove that any map $f \colon X \to CY$ is null homotopic.

7.3.2. Homotopy Types. Previously, we have discussed two conditions on a space X, which are proved to be equivalent,

- Every mapping $f \colon W \to X$ is null homotopic, i.e., homotopic to a constant map.
- ⋆ The identity mapping $\mathrm{id}_X \colon X \to X$ is null homotopic.

An example given before is $\mathbb{R}^n$ for $n \geq 1$.

Definition 7.20. A topological space X is called *contractible* if $\mathrm{id}_X \colon X \to X$ is null homotopic.

Example 7.21. The following are contractible spaces.

(1) The Euclidean spaces $\mathbb{R}^n$, $n \geq 1$.
(2) A convex subset of $\mathbb{R}^n$.
(3) A star-shaped subset $A \subset \mathbb{R}^n$, namely, there exists $a_0 \in A$ such that for each $a \in A$, the line segment joining a_0 to a lies in A.
(4) The most simplest example, i.e., a point $\{ x_0 \}$.

In fact, roughly speaking, in terms of algebraic topology, a contractible space is as simple as a point.

Exercise 7.3.2. Show that a contractible space is path connected.

DEFINITION 7.22. Two spaces X, Y are *of the same homotopy type* or *homotopy equivalent*, written as $X \simeq Y$, if there are continuous mappings $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$. Here, we say that f, g are *homotopy inverses* to each other.

EXAMPLE 7.23.

(1) It is clear that if X, Y are homeomorphic then they are homotopy equivalent. In that case, one exactly gets $g \circ f \equiv \mathrm{id}_X$ and $f \circ g \equiv \mathrm{id}_Y$.
(2) A contractible space X has exactly the homotopy type of a point.

EXERCISE 7.3.3. Let X be a topological space and Y be contractible. Prove that X and $X \times Y$ are homotopy equivalent. In addition, if $y_0 \in Y$ then $X \times \{y_0\} \simeq X \times Y$ where the homotopy equivalences can be chosen rel $X \times \{y_0\}$.

Now, in the light of Theorem 7.19, we indeed have the following.

PROPOSITION 7.24.

- *Let X, Y_1, Y_2 be topological spaces. If Y_1, Y_2 are of the same homotopy type, then there is a bijection between $[X, Y_1]$ and $[X, Y_2]$.*
- *Let X_1, X_2, Y be topological spaces. If X_1, X_2 are of the same homotopy type, then there is a bijection between $[X_1, Y]$ and $[X_2, Y]$.*

EXAMPLE 7.25. The product space $\mathbb{S}^1 \times \mathbb{R}$ is homotopy equivalent to $\mathbb{S}^1$ via the following mappings.

$$e^{\mathbf{i}\theta} \in \mathbb{S}^1 \xrightarrow{\quad f \quad} \left(e^{\mathbf{i}\theta}, 0\right) \in \mathbb{S}^1 \times \mathbb{R}\,;$$
$$\left(e^{\mathbf{i}\theta}, s\right) \in \mathbb{S}^1 \times \mathbb{R} \xrightarrow{\quad g \quad} e^{\mathbf{i}\theta} \in \mathbb{S}^1\,.$$

Then $g \circ f \equiv \mathrm{id}_{\mathbb{S}^1}$ while $f \circ g\left(e^{\mathbf{i}\theta}, s\right) = \left(e^{\mathbf{i}\theta}, 0\right)$. Observe that $\left(e^{\mathbf{i}\theta}, ts\right)$ gives the required homotopy between $f \circ g$ and $\mathrm{id}_{\mathbb{S}^1 \times \mathbb{R}}$, which indeed is a homotopy rel $\mathbb{S}^1 \times \{0\}$.

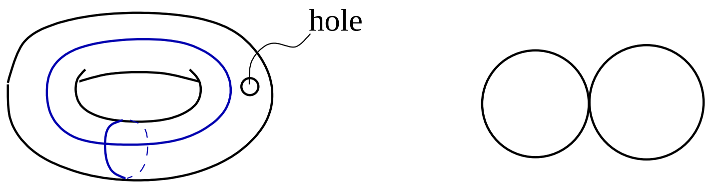
In this example, one observes that $\mathbb{S}^1 \times \{0\}$ is a special subset of $\mathbb{S}^1 \times \mathbb{R}$. It has the homotopy type of the whole space. Moreover, as mentioned above, the homotopy can be chosen to fix every point of $\mathbb{S}^1 \times \{0\}$. In such a case, $\mathbb{S}^1 \times \{0\}$ is called a strong deformation retract of $\mathbb{S}^1 \times \mathbb{R}$.

EXERCISE 7.3.4.

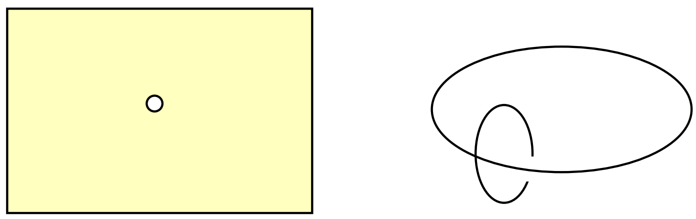
(1) Show that homotopy equivalence (homotopy type) defines an equivalence relation on all the topological spaces.

(2) Show that a space of two points, i.e., $\mathbb{S}^0 = \{-1, 1\}$ with discrete topology, is not homotopy equivalent to a one point space. In other words, $\mathbb{S}^0$ is not contractible.
(3) Consider the unit sphere $\mathbb{S}^{n-1}$ and the punctured space $\mathbb{R}^n \setminus \{0\}$. Show that they are homotopy equivalent.

EXAMPLE 7.26. The punctured torus, $\mathbb{S}^1 \times \mathbb{S}^1 \setminus \{\text{point}\}$ is homotopy equivalent to a figure-8 shape, denoted $\mathbb{S}^1 \wedge \mathbb{S}^1$. A figure-8 shape can be formed by gluing two circles at one point.



EXERCISE 7.3.5. Prove the above fact about the punctured torus by using torus as a quotient space on the square.



Finally, there are some results involving the mapping cone. These are presented in the form of exercises. Given $f\colon X \to Y$, there is a natural mapping, again denote it by f, from $X \times \{0\} \to Y$. One may define the quotient spaces (called mapping cylinder and mapping cone),

$$M_f = ((X \times [0,1]) \sqcup Y)/\sim, \quad \text{where } (x,0) \sim f(x);$$
$$C_f = ((X \times [0,1]) \sqcup Y)/\sim, \quad \text{where } (x,0) \sim f(x) \text{ and } (x_1,1) \sim (x_2,1).$$

REMARK . To understand these two objects, imagine $f\colon \mathbb{S}^1 \to \mathbb{R}$ to be the standard embedding. Then M_f is a tall hat while C_f is a wizard hat. In general, f need not to be one-to-one. In addition, if $\mathbb{D}^n$ is the closed n-dimensional unit disk and $f\colon \mathbb{S}^n \to \mathbb{D}^{n+1}$ is the standard embedding, then $C_f = \mathbb{S}^{n+1}$.

EXERCISE 7.3.6. Show that if $f, g\colon X \to Y$ are homotopic mappings, then M_f and M_g are homotopy equivalent; likewise, C_f and C_g are also homotopy equivalent.

Remark . Using this, one may prove the USELESS result: if $\mathbb{S}^n$ is contractible then so is $\mathbb{S}^{n+1}$. The converse is the USEFUL part because one may set up an induction process. Together with that $\mathbb{S}^0$ is not contractible (done above), we prove $\mathbb{S}^n$ is not contractible.

Exercise 7.3.7. Those interested may try to show if both X and Y are Hausdorff, then so are M_f and C_f.

7.4. Fundamental Group

As we have seen above, one way to understand a topological space X is to look at homotopy classes $[W, X]$ for various spaces W. Naturally, we would choose spaces W such that it is easy to investigate and there is a geometric meaning. In fact, spheres of various dimensions are good candidates.

Exercise 7.4.1. Recall the $\mathbb{S}^0$ is the discrete space $\{ -1, 1 \}$. Show that a space X has trivial $[\mathbb{S}^0, X]$ if and only if X is path connected.

Hint. If $f, g\colon \mathbb{S}^0 \to X$ are homotopic, what do you know about the two pairs of points, $f(-1), g(-1)$ and $f(1), g(1)$?

Along the same line of proof in this exercise, if X is not path connected, one sees that $[\mathbb{S}^0, X]$ still reflects the path components of X but the record is "repeated". To make it clearer, one would set up a fixed base point $x_0 \in X$. Denote $f\colon (\mathbb{S}^0, -1) \to (X, x_0)$ be a continuous mapping from $\mathbb{S}^0$ to X taking $f(-1) = x_0$. Then $[(\mathbb{S}^0, -1), (X, x_0)]$ is the set of homotopy classes rel $\{-1\}$.

Exercise 7.4.2. Show that the number of elements of $[(\mathbb{S}^0, -1), (X, x_0)]$ is exactly the number of path components of X.

Next, we would to consider $[\mathbb{S}^1, X]$. Again, it is more useful to have a fixed base point $x_0 \in X$. That corresponds to the following construction of the so-called fundamental group, $\pi_1(X, x_0)$.

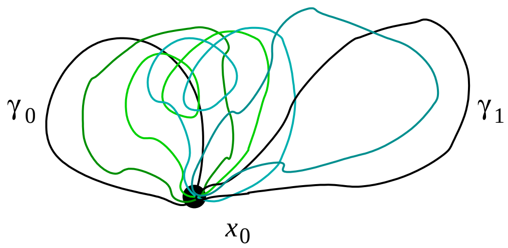
Recall that a path in X is defined as a continuous mapping γ from $[0, 1]$ into X. Throughout this section, we will consider a special kind of homotopies between two paths with the same end-points.

Definition 7.27. Let $\gamma_0, \gamma_1\colon [0, 1] \to X$ be two paths in X such that $\gamma_0(0) = \gamma_1(0)$ and $\gamma_0(1) = \gamma_1(1)$. A homotopy rel $\{ 0, 1 \}$ between the paths is also called a *path homotopy*.

Such a homotopy $H\colon [0,1]\times[0,1] \to X$ satisfies two major conditions:

(a) Homotopy: $H(s,0)=\gamma_0(s)$ and $H(s,1)=\gamma_1(s)$ for all $s\in[0,1]$;
(b) End-points fixed: $H(0,t)=\gamma_0(0)=\gamma_1(0)$ and $H(1,t)=\gamma_0(1)=\gamma_1(1)$ for all $t\in[0,1]$.

DEFINITION 7.28. Let X be a topological space and $x_0\in X$. A *loop* based at x_0 is a closed path with both end-points at x_0. A *loop homotopy* between two loops γ_0,γ_1 based at x_0 is a homotopy rel $\{0,1\}$ between them.

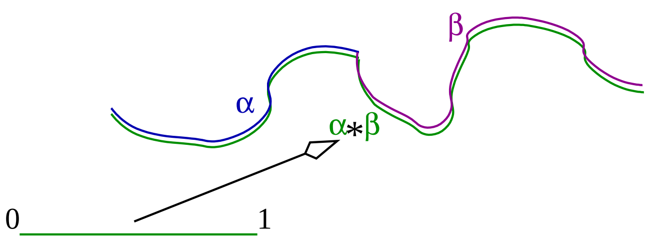


In the above, we only consider paths (loops) that have parameter in $[0,1]$ for convenience, because re-parametrization simply gives homotopic paths. The following can be proved using Exercise 7.2.5 or a modified version of Theorem 7.19.

PROPOSITION 7.29. *If $h\colon [0,1]\to[0,1]$ is a change of parameter, i.e., homeomorphism with $h(0)=0$ and $h(1)=1$; and γ is a path in X, then $h\circ\gamma\simeq\gamma$ rel $\{0,1\}$.*

7.4.1. Concatenation of paths. Let $\alpha,\beta\colon[0,1]\to X$ be paths (or loops) in X such that $\alpha(1)=\beta(0)$, i.e., the second path begins at the point where the first one terminates. There is a new path formed by first tracing α and then β, called $\alpha*\beta\colon[0,1]\to X$ defined by

$$(\alpha*\beta)(s) \mathrel{:}\overset{\text{def}}{=\!=} \begin{cases}\alpha(2s) & s\in\left[0,\frac{1}{2}\right],\\ \beta(2s-1) & s\in\left[\frac{1}{2},1\right].\end{cases}$$



The set of points in X of $(\alpha*\beta)$ is the same as the union of those of α and β. But, the above definition is to reset the parameter so that it is defined on $[0,1]$.

Now, naturally, we expect to define an algebraic operation $(\alpha, \beta) \mapsto \alpha * \beta$ on paths or loops. However, it is not exactly valid because, as mappings on $[0,1]$, $(\alpha*\beta)*\gamma \neq \alpha*(\beta*\gamma)$. It is clear that in $(\alpha*\beta)*\gamma$, α is defined on $[0,1/4]$ while it is defined on $[0,1/2]$ in $\alpha*(\beta*\gamma)$. This leads us to consider the operation on homotopy classes of paths relative end points.

Let $[\alpha]$ denote the set of paths which are homotopic rel $\{0,1\}$ to α. First, we would like to define a *product* by $[\alpha]*[\beta] = [\alpha*\beta]$ if $\alpha(1) = \beta(0)$ or if α, β are both based at x_0. The following guarantees that it is well-defined.

PROPOSITION 7.30. *If* $\alpha_0 \simeq \alpha_1$ *rel* $\{0,1\}$ *and* $\beta_0 \simeq \beta_1$ *rel* $\{0,1\}$, *then*

$$\alpha_0 * \beta_0 \simeq \alpha_1 * \beta_1 \textit{ rel } \{0,1\}.$$

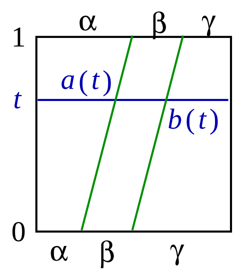
PROOF. Let F, G be the homotopy rel $\{0,1\}$ between α_0, α_1 and β_0, β_1 respectively. Obviously, the desired homotopy is obtained by combining them by a process similar to concatenation. The details are left as EXERCISE 7.4.3. □

Moreover, the product for path or loop homotopy classes is associative, that is,

$$([\alpha]*[\beta])*[\gamma] = [\alpha]*([\beta]*[\gamma]).$$

PROPOSITION 7.31. $(\alpha*\beta)*\gamma \simeq \alpha*(\beta*\gamma)$ *rel* $\{0,1\}$.

PROOF. Since the images of $(\alpha*\beta)*\gamma$ and $\alpha*(\beta*\gamma)$ are the same, the homotopy is simply changing the "speed" of the parameters. Observe from the following diagram,



and define $H\colon [0,1]\times[0,1] \to X$ with

$$H(s,t) = \begin{cases} \alpha\left(\frac{s}{a(t)}\right) & s \in [0,a(t)] \\ \beta\,(??) & s \in [a(t), b(t)] \\ \gamma\,(??) & s \in [b(t),1]. \end{cases}$$

It can be easily seen that H is the required homotopy rel $\{0,1\}$. □

EXERCISE 7.4.4. Complete the above proof by finding out $a(t), b(t)$, and those ??'s in the definition of H.

With these three propositions (7.24, 7.30, 7.31), for a topological space X with $x_0 \in X$, we define

$$\pi_1(X, x_0) \stackrel{\text{def}}{:=} \{\text{loops in } X \text{ based at } x_0 \in X\}/\simeq \quad \text{rel } \{0, 1\} .$$

For loop homotopy classes $[\alpha], [\beta] \in \pi_1(X, x_0)$, an associative product is well-defined by

$$[\alpha] * [\beta] \stackrel{\text{def}}{:=} [\alpha * \beta] .$$

In the remaining of this section, we will show that $(\pi_1(X, x_0), *)$ is a group. In this situation, we only need to exhibit an identity element and show that an inverse element exists for each element.

DEFINITION 7.32. The group $(\pi_1(X, x_0), *)$ is the *fundamental group of X at x_0*.

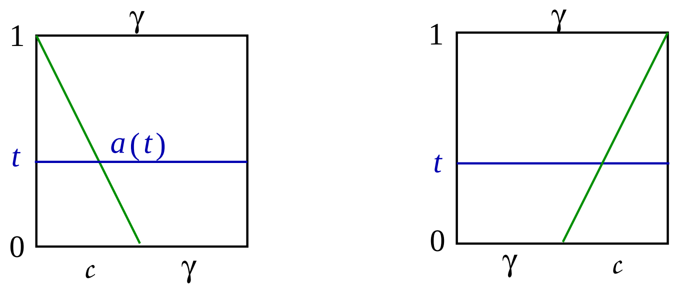
The identity element is given by the constant loop. For $x_0 \in X$, let

$$\mathfrak{c}_0 \colon [0,1] \ \to \ X, \qquad \mathfrak{c}_0(s) = x_0, \quad s \in [0,1]$$

define the continous path (loop) in X. Denote its homotopy class by $\mathbf{1}$. Similarly, for a point $x_1 \in X$ (which may also be x_0), denote a constant map $\mathfrak{c}_1$.

PROPOSITION 7.33. *For each path γ from x_0 to x_1, $\mathfrak{c}_0 * \gamma \simeq \gamma \simeq \gamma * \mathfrak{c}_1$ rel $\{0,1\}$. In particular, when $x_0 = x_1$ and α is a loop, $[\alpha] * \mathbf{1} = [\alpha] = \mathbf{1} * [\alpha]$.*

PROOF. Again, the images of the three paths are the same set in X, it is sufficient to find homotopies that essentially are changing the parameters.



The above two diagrams provide the observation to get the desired homotopies. Again, the explicit expressions of the homotopies are left as EXERCISE 7.4.5. □

Next, for a path γ in X from x_0 to $x_1 \in X$, we may define another path, which is essentially the reversed direction of γ, namely,

$$\overline{\gamma}\colon [0,1] \to X \qquad \overline{\gamma}(s) = \gamma(1-s), \quad s \in [0,1]\,.$$

If α is a loop based at x_0, its homotopy class will give the inverse element of $[\alpha]$, i.e., $[\alpha]^{-1} = [\overline{\alpha}]$.

PROPOSITION 7.34. *With the notation above, $\alpha * \overline{\alpha} \simeq \mathfrak{c}_0 \simeq \overline{\alpha} * \alpha$ rel $\{0,1\}$.*

PROOF. As $\alpha * \overline{\alpha}$ is simply tracing the whole α and then back tracing the whole loop, one can gradually trace only part the the loop α and eventually never move away from x_0. The homotopy for $\alpha * \overline{\alpha} \simeq \mathfrak{c}_0$ rel $\{0,1\}$ is given by

$$F\colon [0,1]\times[0,1] \to X \qquad F(s,t) = \begin{cases} \alpha(2st) & s \in \left[0,\frac{1}{2}\right] \\ \alpha(t(1-2s)) & s \in \left[\frac{1}{2},1\right] \end{cases}.$$

The other one for $\overline{\alpha} * \alpha \simeq \mathfrak{c}_0$ is similarly constructed. □

DEFINITION 7.35. A path connected space X is called *simply connected* or 1-connected if there exists $x_0 \in X$ such that $\pi_1(X,x_0)$ is the trivial group, i.e., it only contains the identity element $\mathbf{1} = [\mathfrak{c}_0]$.

Since fundamental groups are normally non-abelian, the trivial group is usually written as $\mathbf{1}$.

Next, we discuss the effect of the base point. More precisely, if the space is path connected, then the fundamental group is independent of the choice of the base point.

Let X be a path connected space and $x_0, x_1 \in X$. Let $\sigma\colon [0,1] \to X$ be any path joining them, i.e., $\sigma(0) = x_0$ and $\sigma(1) = x_1$. Consider a mapping $\varphi_\sigma\colon \pi_1(X,x_0) \to \pi_1(X,x_1)$ defined by the following. Let $[\alpha] \in \pi_1(X,x_0)$ be a homotopy class of a loop α based at x_0. Then $\overline{\sigma} * \alpha * \sigma$ is a loop based at x_1. Take $\varphi_\sigma([\alpha]) \overset{\text{def}}{:=} [\overline{\sigma} * \alpha * \sigma]$.

EXERCISE 7.4.6. Show that φ_σ is well defined. That is, if $\alpha_0 \simeq \alpha_1$ rel $\{0,1\}$, then $\overline{\sigma} * \alpha_0 * \sigma \simeq \overline{\sigma} * \alpha_1 * \sigma$ rel $\{0,1\}$.

With the exercise, since $\overline{\sigma}$ is a path from x_1 to x_0, it follows naturally that we also have the mapping $\varphi_{\overline{\sigma}}$ from $\pi_1(X,x_1)$ back to $\pi_1(X,x_0)$. In fact, we have $\varphi_\sigma \circ \varphi_{\overline{\sigma}} = \mathrm{id} = \varphi_{\overline{\sigma}} \circ \varphi_\sigma$. Therefore, $\varphi_{\overline{\sigma}}$ is in fact the inverse of φ_σ.

THEOREM 7.36. *The mapping φ_σ is an isomorphism from $\pi_1(X,x_0)$ to $\pi_1(X,x_1)$.*

Exercise 7.4.7. Verify that φ_σ is a homomorphism, i.e.,

$$\varphi_\sigma([\alpha]\cdot[\beta]) = \varphi_\sigma([\alpha])\cdot\varphi_\sigma([\beta])\,,$$

and other details of the proof of the above theorem. Also, show that if σ and τ are two paths from x_0 to x_1 such that $\sigma \simeq \tau$ rel $\{0,1\}$, then $\varphi_\sigma \cong \varphi_\tau$.

7.5. Useful Examples

7.5.1. Euclidean and the Simplest Ones. The Euclidean spaces $\mathbb{R}^n$, $n = 0,1,2,\ldots$ belong to the simplest topological spaces because they are contractible.

Recall that a space X is contractible if its identity map id_X is homotopic to a constant map $\mathfrak{c}$ on X. It does not matter which constant map we are taking.

Exercise 7.5.1. Let X be a path connected space, $x_1, x_2 \in X$ and $\mathfrak{c}_1, \mathfrak{c}_2$ be constant maps where $\mathfrak{c}_1(x) = x_1$ and $\mathfrak{c}_2(x) = x_2$ for all $x \in X$. Show that $\mathfrak{c}_1 \simeq \mathfrak{c}_2$.

Proposition 7.37. *If X is contractible and $x_0 \in X$, then $\pi_1(X, x_0)$ is trivial. In particular, each Euclidean space $\mathbb{R}^n$ is simply connected.*

Proof. Without loss of generality, assume $\mathfrak{c}(x) = x_0$ for all $x \in X$. Let $\mathrm{id}_X \overset{F}{\simeq} \mathfrak{c}\ : X \to X$. Then for any loop γ in X based at x_0,

$$\gamma = \mathrm{id}_X \circ \gamma \simeq \mathfrak{c}\circ\gamma\,.$$

Clearly $\mathfrak{c}\circ\gamma$ is the constant path, i.e., $[\mathfrak{c}\circ\gamma] = \mathbf{1} \in \pi_1(X, x_0)$. □

Another obvious example of contractible, and hence simply connected, space is the open disk or closed disk,

$$\{\, x \in \mathbb{R}^n \ :\ \|x\| < 1 \,\} \qquad \text{or} \qquad \{\, x \in \mathbb{R}^n \ :\ \|x\| \le 1 \,\}\,.$$

Indeed, they both are convex subsets of $\mathbb{R}^n$. Recall that a star-shaped $X \subset \mathbb{R}^n$ satisfies that there exist a point $x_0 \in X$ such that every point $x \in X$ can be joined to x_0 by a straight line. Obviously, a convex subset is of star-shaped.

Exercise 7.5.2. Show that any star-shaped subset of $\mathbb{R}^n$ is contractible and thus simply-connected.

7.5.2. The Circle and Punctured Plane. The circle $\mathbb{S}^1$ gives an important example of fundamental group because of two reasons. First, the space is simple enough so the geometry and algebraic properties are easy to visualize. Second, it has a typical structure that occurs in a more general context that naturally leads to theory of covering spaces.

PROPOSITION 7.38. *Let x_0 be a point in $\mathbb{S}^1 \subset \mathbb{R}^2 \setminus \{\mathbf{0}\}$. Then*

$$\pi_1(\mathbb{S}^1, x_0) \cong \pi_1\left(\mathbb{R}^2 \setminus \{\mathbf{0}\}\right) \cong (\mathbb{Z}, +) .$$

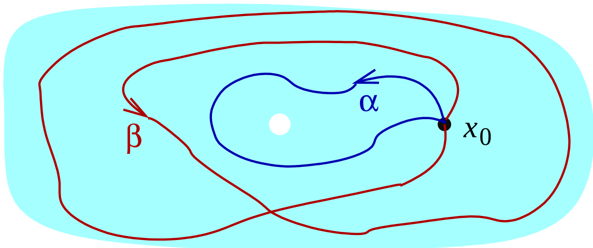
In the rest of this section, we will give the main ideas of the proof of the above fact. The first isomorphism is due to the fact that $\mathbb{S}^1$ and $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ are homotopy equivalent.

EXERCISE 7.5.3. Show that if $p: \ \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{S}^{\mathbf{1}}$ is defined by $p(\mathbf{x}) = \mathbf{x}/\left\|\mathbf{x}\right\|$, then the inclusion $\mathbb{S}^1 \hookrightarrow \mathbb{R}^2 \setminus \{\mathbf{0}\}$ and the mapping p are homotopy inverse to each other.

The second isomorphism can be described by the following. Let $[\gamma] \in \pi_1\left(\mathbb{R}^2 \setminus \{\mathbf{0}\}\right)$ be represented by a loop γ based at the point x_0. Then

$$[\gamma] \quad \mapsto \quad \text{winding number of } \gamma \ : \ \pi_1\left(\mathbb{R}^2 \setminus \{\mathbf{0}\}\right) \to \mathbb{Z}$$

is well-defined and indeed is an isomorphism. For example, in the picture below, $[\alpha] \mapsto 1$ and its backward oriented curve goes to -1; moreover, $[\beta] = [\alpha] * [\alpha] \mapsto 2$.

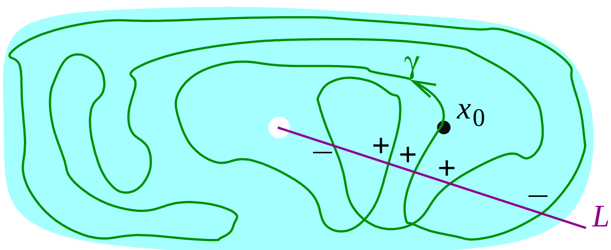


In the case that γ is a piecewise differentiable loop, we have

$$\text{Winding number of } \gamma \ = \frac{1}{2\pi} \int_\gamma \frac{dz}{z} .$$

However, when γ is simply a continuous loop, the winding number may be obtained in a topological method. For instance, let L be a half-line from $\mathbf{0}$ to infinity such that L cuts γ transversely (never tangentially). Then there are finitely many intersecting points in $L \cap \gamma$. Define

$$\text{The sign at an intersection } = \begin{cases} +1 & \text{if } \gamma \text{ cuts } L \text{ counterclockwisely} \\ -1 & \text{if } \gamma \text{ cuts } L \text{ clockwisely.} \end{cases}$$



Then the winding number is the sum of signs over all intersection points.

EXERCISE 7.5.4. Take an example of γ, try different choices of half-lines L to calculate the winding number. Convince yourself that the result is independent of the choice of L.

To come up with a proof of the independency, one may first show that the winding number will not change if the angle of L is continuously perturbed a little. And the number of intersection points may change only when L passes through a tangential intersection of γ at some angle. Then argue that two opposite intersections disappear in this situation. The similar argument may apply to a homotopy of γ.

EXERCISE 7.5.5. Convince yourself that this method of calculating winding number gives an isomorphism.

In the following, we will sketch another proof of Proposition 7.38. This is an alternative way of seeing the winding number and thus an isomorphism between $\pi_1(\mathbb{S}^1, z_0)$ and $\mathbb{Z}$ is established. It should note that the study of the fundamental group of the circle is an important revelation to the theory of covering spaces. Many books may prove the proposition at the same time of discussing covering spaces. However, due to time limitation, we will take an elementary route; though the idea behind the proof is essentially the same.

LEMMA 7.39. *Let* $\gamma\colon [0,1] \to \mathbb{C}\setminus\{0\}$ *with* $\gamma(0)=\gamma(1)=z_0$ *be a continuous loop into the punctured plane. Then there is a continuous function* $\theta\colon [0,1] \to \mathbb{R}$ *such that for all* $s\in[0,1]$,

$$\cos\theta(s) = \frac{\mathrm{Re}[\gamma(s)]}{|\gamma(s)|}, \qquad \sin\theta(s) = \frac{\mathrm{Im}[\gamma(s)]}{|\gamma(s)|}\,.$$

This function $\theta(s)$ is uniquely determined by the choice of its value $\theta(0)$. It is called a *continuous choice of the argument* of the curve γ.

SKETCH OF PROOF. First, if the image of γ lies in a half-plane of $\mathbb{C}$, then it is easy to show that such choice of $\theta(s)$ exists. In this case, one may fix the value of $\theta(0)$ according to the point $z_0\in\mathbb{C}\setminus\{0\}$. Then for the above equations of $\cos\theta(s)$ and $\sin\theta(s)$, for each $s\in[0,1]$, there exists a unique solution inside the interval $(\theta(0)-\pi, \theta(0)+\pi)$. In fact, if γ is differentiable, it is exactly given by

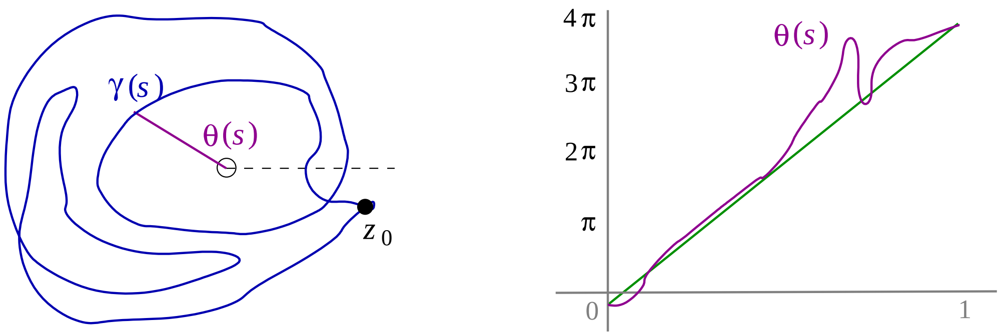
$$\theta(s) = \theta(0) + \int_0^s \frac{\gamma'(t)\,dt}{\gamma(t)}\,.$$

Second, for general γ, by compactness, one may subdivide $[0,1]$ into finitely many subintervals such that the image of γ of each subinterval lies in a half-plane. Then inductively, a continuous choice of $\theta(s)$ can be obtained. $\square$

Since $\gamma(0) = \gamma(1) = z_0$, we have

$$\cos\theta(0) = \cos\theta(1) = \frac{\mathrm{Re}[z_0]}{|z_0|}, \qquad \sin\theta(0) = \sin\theta(1) = \frac{\mathrm{Im}[z_0]}{|z_0|}.$$

It turns out that $\theta(1) = \theta(0) + 2d\pi$ for some $d \in \mathbb{Z}$. By continuity of γ and connectedness of $[0,1]$, it can be shown that d is uniquely determined by γ. An example of $d = 2$ is shown below.



Obviously, from the graph of $\theta(s)$, a homotopy between $\theta(s)$ and $\theta(0) + 2d\pi s$ rel $\{0,1\}$ is easily constructed, which in turns leads to a homotopy rel $\{0,1\}$ between $\gamma(s)$ and the d-fold circle. Thus we have

LEMMA 7.40. *For each continuous loop γ, there exists a unique $d \in \mathbb{Z}$ such that γ is homotopic to the d-fold circle relative end-point.*

Then the mapping taking the homotopy class $[\gamma] \in \pi_1(\mathbb{S}^1, z_0)$ to $d \in \mathbb{Z}$ is the desired isomorphism. Hence Proposition 7.38 is established.

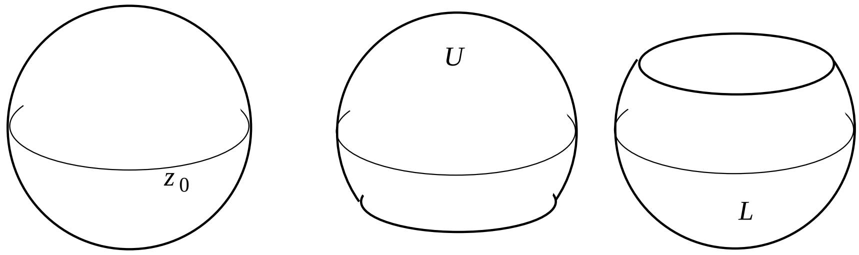
7.5.3. Spheres and Punctured Euclidean spaces. Similar to the situation of the circle and punctured plane, the n-sphere and punctured space are of the same homotopy type.

PROPOSITION 7.41. *For all $n \geq 1$, the n-sphere $\mathbb{S}^n \simeq \mathbb{R}^{n+1} \setminus \{\mathbf{0}\}$.*

However, in higher dimensional cases, the fundamental group becomes trivial.

PROPOSITION 7.42. *For $n \geq 2$, $\pi_1(\mathbb{S}^n, \mathbf{z_0}) \cong \pi_1(\mathbb{R}^{n+1} \setminus \{\mathbf{0}\}, \mathbf{z_0}) = \{\mathbf{1}\}$.*

The proof of this result makes use of the Van Kampen's Theorem, which is out of the scope of this course. An illustration of $\mathbb{S}2$ may be helpful.



Let $\mathbb{S}^2 = U \cup L$ where U, L are the upper and lower parts as shown. Then $U \cap L$ is homeomorphic to a cylinder $\mathbb{S}^1 \times [-\delta, \delta] \simeq \mathbb{S}^1$. Take a point $\mathbf{z}_0$ in the equator $\mathbb{S}^1$. In this way, Van Kampen's Theorem tells us that $\pi_1(\mathbb{S}^2, \mathbf{z}_0)$ is a "product group" of $\pi_1(U, \mathbf{z}_0)$ and $\pi_1(L, \mathbf{z}_0)$ with some correction due to $\pi_1(\mathbb{S}^1, \mathbf{z}_0)$. However, since both U and L are homeomorphic to a disk, we have $\pi_1(U, \mathbf{z}_0) = \mathbf{1} = \pi_1(L, \mathbf{z}_0)$ and hence $\pi_1(\mathbb{S}^2, \mathbf{z}_0) = \mathbf{1}$.

For higher dimensions, Van Kampen's Theorem expresses $\pi_1(\mathbb{S}^n, \mathbf{z}_0)$ in terms of $\pi_1(\mathbb{E}^n_+, \mathbf{z}_0)$, $\pi_1(\mathbb{E}^n_-, \mathbf{z}_0)$, and $\pi_1(\mathbb{S}^{n-1}, \mathbf{z}_0)$, where $\mathbb{E}^n_\pm$ denotes the upper and lower n-hemisphere. Through an inductive process on $\mathbb{S}^1 \subset \mathbb{S}^2 \subset \cdots \subset \mathbb{S}^n$, one may show that any loop in $\mathbb{S}^n$ is homotopically trivial except in $\mathbb{S}^1$.

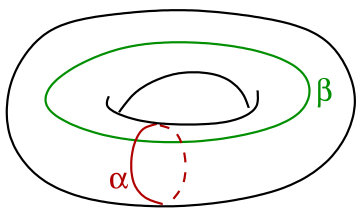
7.5.4. The Torus. Recall that the 2-dimensional torus can be seen from three perspectives, as a product of circles or as a surface of revolution in $\mathbb{R}^3$ or as a quotient space of a rectangle. We will use all of them to understand the fundamental group of the torus, whichever is convenient.

Exercise 7.5.6. Let X, Y be connected topological spaces with $x_0 \in X$ and $y_0 \in Y$. Then $\pi_1\left(X \times Y, (x_0, y_0)\right) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$, where the $\times$ on the right hand side denotes direct product of groups. *Hint.* Any loop in $X \times Y$ is of the form $(\gamma_X(s), \gamma_Y(s))$ where γ_X and γ_Y are loops in X and Y respectively.

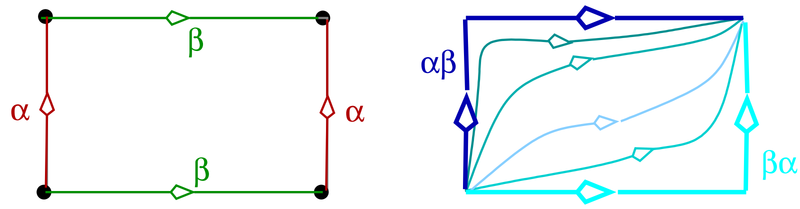
Consequently, we may apply the result of the exercise to the torus.

Proposition 7.43. $\pi_1(\mathbb{T}, *) \cong \pi_1(\mathbb{S}^1, *) \times \pi_1(\mathbb{S}^1, *) \cong (\mathbb{Z} \oplus \mathbb{Z}, +)$.

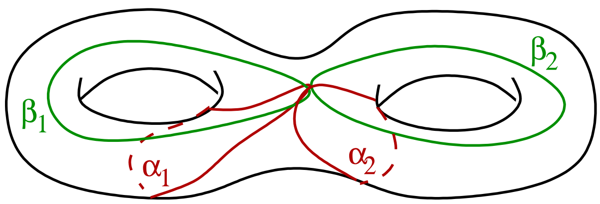
The generators $[\alpha], [\beta] \in \pi_1(\mathbb{T}, *)$ corresponding to $(1, 0), (0, 1) \in \mathbb{Z} \oplus \mathbb{Z}$ are shown in the surface of revolution below.



Note that the group $\mathbb{Z}\oplus\mathbb{Z}$ is abelian. Since $[\alpha][\beta]$ and $[\beta][\alpha]$ naturally are mapped to $(1,1)$, therefore, we should expect that $\alpha * \beta \simeq \beta * \alpha$ rel $\{0,1\}$. This can be seen in the illustration of quotient space below. Equivalently, in $\pi_1(\mathbb{T},*)$, there are two generators, α and β, but they obey a relation $\alpha\beta\alpha^{-1}\beta^{-1}=1$.

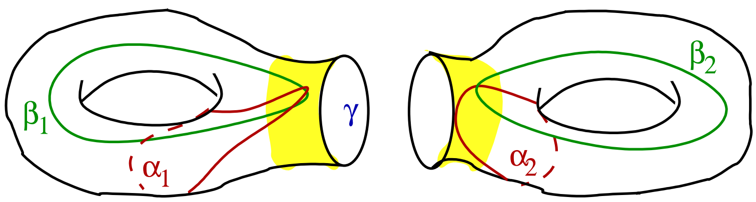


7.5.5. Surface of Genus 2. This example is intended to illustrate how the fundamental group records topological information of a space. Many steps are not rigorously enough. A surface S of genus 2 is sort of a double torus.

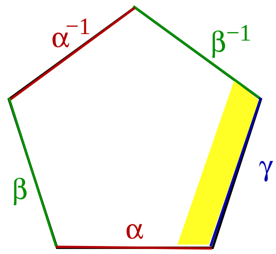


Naturally, one would expect that α_1,β_1 and α_2,β_2 are four generators for the group $\pi_1(S,x_0)$, where x_0 is the base point at the common intersection of them. However, now we do not have the same abelian property as in the situation of the torus. This can be explained by Van Kampen's Theorem again.

Let us cut the surface into two pieces, left-half and right-half, as shown in the picture below.



Unlike the torus, $\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}$ does not form the boundary of a rectangle. Instead, they are the four sides of a pentagon and the fifth side is γ, which is the cut open circle. Therefore the relation is $\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}=\gamma$.



The right-half gives a similar pentagon but the open circle is denoted by γ^{-1}. Thus the relation is $\alpha_2\beta_2\alpha_2^{-1}\beta_2^{-1} = \gamma^{-1}$. In view of Van Kampen's Theorem, the surface $S = L \cup R$ and $L \cap R$ is the yellow tube $\simeq$ cylinder $\simeq \mathbb{S}^1$. The two generators on the left and the two on the right need to satisfy a condition in the yellow tube, namely,

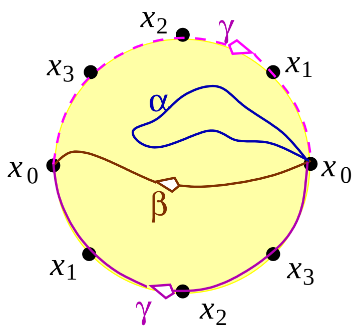
$$\alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}\alpha_2\beta_2\alpha_2^{-1}\beta_2^{-1} = 1\,.$$

Then $\pi_1(S, x_0)$ is presented by a group of four generators with the above relation.

7.5.6. Projective Plane. Recall that the 2-dimensional real projective plane can be defined as $\mathbb{R}\mathbf{P}^2 = \mathbb{D}^2/\!\sim$ where $\mathbb{D}^2 = \{\, z \in \mathbb{Z} : \ |z| \le 1 \,\}$ and

$$z_1 \sim z_2 \qquad \text{if} \qquad \begin{cases} z_1 = z_2 \\ z_1 = -z_2 \quad \text{with} \quad |z_1| = |z_2| = 1\,. \end{cases}$$

Let us observe what happens from the following picture.



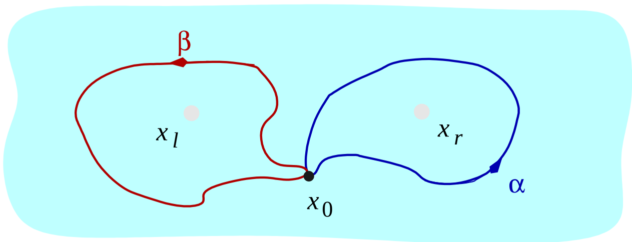
Any simple loop in $\mathbb{R}\mathbf{P}^2$ corresponds to a simple loop or a simple arc in $\mathbb{D}^2$, each illustrated by an example in the picture. A simple loop in $\mathbb{D}^2$ is null homotopic and so is the corresponding one in $\mathbb{R}\mathbf{P}^2$. An simple arc such as $\beta \in \mathbb{D}^2$ corresponds to a non-trivial loop in $\mathbb{R}\mathbf{P}^2$. Clearly, in $\mathbb{D}^2$, by "pushing downward", β is homotopic to γ on the lower half boundary. Since $\mathbb{R}\mathbf{P}^2$ is obtained by identifying the boundary of $\mathbb{D}^2$ according to that $z_1 = -z_2$. This γ in the lower half of $\mathbb{D}^2$ will become a simple arc on the upper half boundary. By "pushing upward", one sees that β is homotopic to $-\gamma$ on the upper boundary. So $\gamma \simeq -\gamma$ relative the end-point. Equivalently, $[\gamma]^2 = 1$ in the fundamental group $\pi_1(\mathbb{T}, x_0)$. With some algebraic calculations, one may prove that

PROPOSITION 7.44. $\pi_1(\mathbb{R}\mathbf{P}^2, x_0) \simeq (\mathbb{Z}/2\mathbb{Z}, +)$.

Combining the argument of the cases of the torus and the projective plane, one may also obtain that

PROPOSITION 7.45. $\pi_1(\mathbb{K}, *) \simeq (\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}, +)$ *where* $\mathbb{K}$ *is the Klein bottle.*

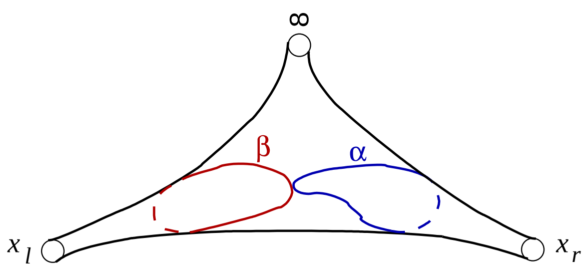
7.5.7. Two-punctured Plane. Let $X = \mathbb{R}^2 \setminus \{x_\ell, x_r\}$ be the plane with two points removed, where $x_\ell = (-1, 0)$ and $x_r = (1, 0) \in \mathbb{R}^2$. Pick a base point $x_0 \in X$ and choose two positively oriented simple closed loops α and β around x_r and x_ℓ respectively.



It is naturally imagined that α and β are generators for the fundamental group, $\pi_1(X, x_0)$, of the two-punctured plane. We will see below that $[\alpha]\,[\beta] \neq [\beta]\,[\alpha]$ in the fundamental group. In fact, there are no other relations. Thus, $\pi_1(X, x_0)$ is the free group on two generators. For those who are not familiar with algebra, this means the group consists of all possible products, including the followings,

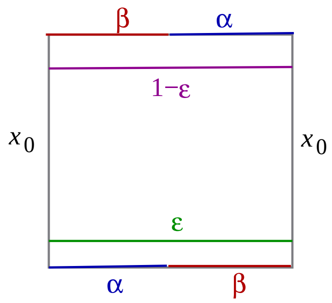
$$1, \alpha, \alpha^{-1}, \beta, \beta^{-1}, \alpha\beta, \beta^{-1}\alpha^{-1}, \beta\alpha, \alpha^{-1}\beta^{-1}, \alpha^{-1}\beta, \beta^{-1}\alpha, \alpha\beta^{-1}, \beta\alpha^{-1},$$
$$\alpha\beta\alpha, \alpha^{-1}\beta^{-1}\alpha^{-1}, \beta\alpha\beta, \beta^{-1}\alpha^{-1}\beta^{-1}, \alpha^2\beta, \beta^{-1}\alpha^{-2}, \ldots\ldots.$$

The two-punctured plane is homeomorphic to the so-called *pair of pants* as shown in the following picture.

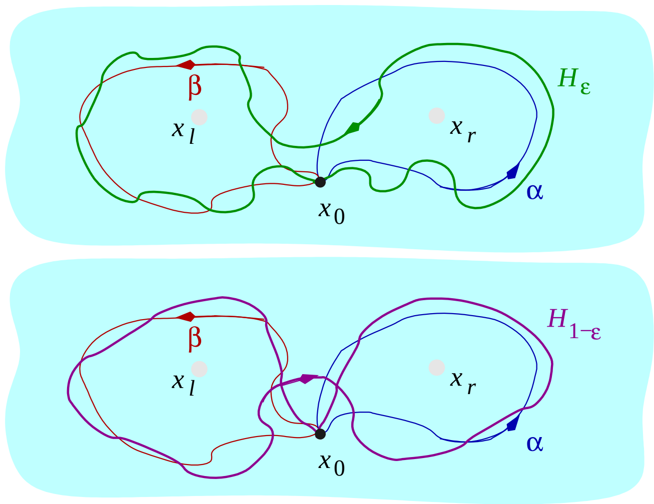


Both the two-punctured plane and the pair of pants are homotopy equivalent to a figure-8, usually denoted as $\mathbb{S}^1 \wedge \mathbb{S}^1$.

It is good to illustrate why its fundamental group $\pi_1(X, x_0)$ is not abelian. From the illustration, one also sees the importance of requiring the homotopy to fix the base point x_0. Suppose there is a homotopy H rel end-points between $\alpha * \beta$ and $\beta * \alpha$, it will be represented by the following square where slight deformation at $t = \varepsilon, 1 - \varepsilon$ are shown.



Then, the curves H_ε and $H_{1-\varepsilon}$ are shown in the pictures, which cannot be deformed to each other with the point x_0 fixed.



Note that if the starting and ending points are not fixed at x_0, the curves H_ε and $H_{1-\varepsilon}$ can be deformed to each other.

7.6. Homotopy Invariance

As it is seen in Section 7.3, a continuous function between topological spaces induces a mapping between homotopy classes involving the two spaces. This already gives us a way to distinguish two spaces. The group structure of fundamental groups provides further information for such purpose.

Let X, Y be topological spaces and $\varphi\colon X \to Y$ be a continuous function with $\varphi(x_0) = y_0$. For convenience, we usually use the notation $\varphi\colon (X, x_0) \to (Y, y_0)$ to specify such situation. More generally, if $A \subset X$ and $B \subset Y$, $\varphi\colon (X, A) \to (Y, B)$ means that $\varphi(A) \subset B$.

Given a loop γ in X with base point x_0, i.e., $\gamma\colon [0,1] \to X$ with $\gamma(0) = x_0 = \gamma(1)$. The composition $\varphi\circ\gamma\colon [0,1] \to Y$ is clearly a loop in Y with base point $y_0 \in Y$. Moreover, if two loops $\gamma_0, \gamma_1\colon [0,1] \to X$ are two homotopic rel base point x_0, by Theorem 7.19, the loops $\varphi\circ\gamma_0$ and $\varphi\circ\gamma_1$ are homotopic rel base point y_0 (Exercise 7.6.1). Therefore, given an element $[\gamma] \in \pi_1(X, x_0)$, a loop homotopy class $[\varphi\circ\gamma] \in \pi_1(Y, y_0)$ is defined, which only depends on the loop homotopy class of $[\gamma]$.

Definition 7.46. The mapping defined above is denoted

$$\varphi_\# \quad \text{or} \quad \varphi_*\colon \pi_1(X, x_0) \to \pi_2(Y, y_0)\,.$$

Not only the mapping is useful, the additional group structure is also important.

Theorem 7.47. *Let $\varphi\colon (X, x_0) \to (Y, y_0)$ be a continous function. Then the mapping $\varphi_\#\colon \pi_1(X, x_0) \to \pi_1(Y, y_0)$ is a homomorphism.*

Proof. Let α, β be loops with base point x_0 which represent $[\alpha], [\beta] \in \pi_1(X, x_0)$ respectively. It is needed to show that $\varphi_\#([\alpha]\cdot[\beta]) = \varphi_\#([\alpha])\cdot\varphi_\#([\beta])$. According to the definition of $\varphi_\#$, it is equivalent to show that

$$\varphi\circ(\alpha*\beta) \simeq (\varphi\circ\alpha)*(\varphi\circ\beta) \quad \text{rel the end point } \{y_0\}.$$

In fact, considering the definition of concatenation $*$, one has eqality above. $\square$

Exercise 7.6.2.
1. What is $\mathrm{id}_\#$ on $\pi_1(X, x_0)$ where $\mathrm{id}\colon X \to X$ is the identity map?
2. Let $x_0 \in A \subset X$. Does the inclusion map $i\colon A \to X$ induce an injective homomorphism $i_\#\colon \pi_1(A, x_0) \to \pi_1(X, x_0)$?
3. Suppose $f\colon (X, x_0) \to (Y, y_0)$ is a surjective continuous function. Is it true that $f_\#$ is also surjective?

Theorem 7.48. *If two continuous functions $\varphi \simeq \psi\colon X \to Y$ are homotopic, then they induce the same mapping, $\varphi_\# \equiv \psi_\#$ on $\pi_1(X, x_0)$.*

Proof. The proof uses Theorem 7.19 again and it is left as Exercise 7.6.3. $\square$

In principle, the above result guarantees that the homomorphism induced on the fundamental group is determined up to the homotopy of the continuous function. Furthermore, this homomorphism is particular useful because of a property called *naturality*.

THEOREM 7.49. *Let* $\varphi\colon (X, x_0) \to (Y, y_0)$ *and* $\psi\colon (Y, y_0) \to (Z, z_0)$ *be continuous functions. Then*

$$\psi_{\#} \circ \varphi_{\#} \equiv (\psi \circ \varphi)_{\#}\colon \pi_1(X, x_0) \ \to \ \pi_1(Z, z_0)\,.$$

PROOF. This is simply a consequence of $\psi \circ (\varphi \circ \alpha) = (\psi \circ \varphi) \circ \alpha$ for any loop α with base point x_0. $\square$

This above naturality result is usually represented by the following figure.

$$\begin{tikzcd}
(X, x_0) \arrow[rr, "\varphi"] \arrow[dr, "\psi\varphi"'] & & (Y, y_0) \arrow[dl, "\psi"] \\
& (Z, z_0) &
\end{tikzcd}
\qquad
\begin{tikzcd}
\pi_1(X, x_0) \arrow[rr, "\varphi_{\#}"] \arrow[dr, "\substack{\psi_{\#}\varphi_{\#} \\ (\psi\varphi)_{\#}}"'] & & \pi_1(Y, y_0) \arrow[dl, "\psi_{\#}"] \\
& \pi_1(Z, z_0) &
\end{tikzcd}$$

From Theorem 7.49, it is easy to obtain the following two useful facts.

COROLLARY 7.50.
1. *If* $\varphi\colon (X, x_0) \to (Y, y_0)$ *and* $\psi\colon (Y, y_0) \to (X, x_0)$ *are homotopy equivalences inverse to each other, then* $\psi_{\#} \equiv (\varphi_{\#})^{-1}$ *from* $\pi_1(Y, y_0)$ *to* $\pi_1(X, x_0)$.
2. *If two spaces* X, Y *are homeomorphic, then their fundamental groups are isomorphic.*

Given a pair of homotopy inverses, which satisfy $\psi \circ \varphi \simeq \mathrm{id}_X$ and $\varphi \circ \psi \simeq \mathrm{id}_Y$. However, in general $\varphi(x_0) = y_0$ but $\psi(y_0) = x_1 \neq x_0$. In this case, we have

$$\pi_1(X, x_0) \ \xrightarrow{\varphi_{\#}} \ \pi_1(Y, y_0) \ \xrightarrow{\psi_{\#}} \ \pi_1(X, x_1)\,.$$

EXERCISE 7.6.4. Explore the relation between this isomorphism $(\psi\varphi)_{\#}$ and the one in Theorem 7.36 from $\pi_1(X, x_0)$ to $\pi_1(X, x_1)$.

7.7. Brouwer Fixed Point Theorem

Let X be a topological space. A subset $A \subset X$ is called a *retract* of X if there is a continuous function $r\colon X \to A$ such that $r|_A \equiv \mathrm{id}_A$, i.e., for all $a \in A$, $r(a) = a$. Equivalently, if $i : A \hookrightarrow X$ is the inclusion map, then $r \circ i \equiv \mathrm{id}_A$.

Example 7.51.

(1) Any singleton $\{x_0\} \subset X$ is a retract of X by taking r to be the constant map on x_0. However, this is an uninteresting retract because the topology of $\{x_0\}$ does not reflect anything about the topology of X. The aim is always to look for a simple enough subset A but it still gives useful information about X.

(2) The infinite cylinder $\mathbb{S}^n \times \mathbb{R}$ has retracts $\mathbb{S}^n \times (-1, 1)$ and $\mathbb{S}^n \times \{0\}$.

Exercise 7.7.1.

(1) Show that if $B \subset A \subset X$ and A is a retract of X while B is a retract of A, then B is a retract of X.

(2) Show that the punctured torus $\mathbb{S}^1 \times \mathbb{S}^1 \setminus \{x_0\}$ has a retract of figure-8, $\mathbb{S}^1 \wedge \mathbb{S}^1$.

Proposition 7.52. *Let $A \subset X$ be a retract. Then $i_\#\colon \pi_1(A, x_0) \to \pi_1(X, x_0)$ is a monomorphism while $r_\#\colon \pi_1(X, x_0) \to \pi_1(A, x_0)$ is an epimorphism, where $x_0 \in A$.*

Proof. Be careful, it is not because i is injective and r is surjective.

Observe from the following communtative diagrams,

$$\begin{tikzcd}
(A, x_0) \arrow[rr, "i"] \arrow[dr, "r \circ i \equiv \mathrm{id}_A"'] & & (X, x_0) \arrow[dl, "r"] \\
& (A, x_0) &
\end{tikzcd}
\qquad
\begin{tikzcd}
\pi_1(A, x_0) \arrow[rr, "i_\#"] \arrow[dr, "\mathrm{id}_\# \equiv \mathrm{id}"'] & & \pi_1(X, x_0) \arrow[dl, "r_\#"] \\
& \pi_1(A, x_0) &
\end{tikzcd}$$

Since $r_\# \circ i_\# = \mathrm{id}_{\pi_1(A, x_0)}$, we have $i_\#$ injective and $r_\#$ surjective. □

Definition 7.53. Let X be a topological space and $A \subset X$ is a retract with an inclusion $i\colon A \to X$ and a retraction $r\colon X \to A$ such that $r \circ i \equiv \mathrm{id}_A$. The set A is a *deformation retract* if, in addition, $r \equiv i \circ r \simeq \mathrm{id}_X$. If further, $r \equiv i \circ r \simeq \mathrm{id}_X$ rel A, then A is a *strong deformation retract* of X.

REMARK . In the above definition of strong deformation retract, the homotopy H between r and id_X must fix every point of A at any time $t \in [0,1]$, i.e., $H_t|_A \equiv \mathrm{id}_A$. In different books, there may be all sorts of variation for this definition. For example, some may only require $H_t(A) \subset A$ or $H_t|_A$ is a homeomorphism.

PROPOSITION 7.54. *If $A \subset X$ is a deformation retract (or strong deformation retract) of X and $x_0 \in A$, then*

$$i_{\#} \colon \pi_1(A, x_0) \ \to \ \pi_1(X, x_0)$$

is an isomorphism with inverse $r_{\#}$.

PROOF. It is left as EXERCISE 7.7.2. $\square$

Let $\mathbb{S}^n = \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \ : \ \|\mathbf{x}\| = 1 \right\}$ and $\mathbb{D}^{n+1} = \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \ : \ \|\mathbf{x}\| \le 1 \right\}$ be the standard sphere and disk.

THEOREM 7.55. *The circle $\mathbb{S}^1$ is not a retract of the disk $\mathbb{D}^2$.*

REMARK . The analogous statement about $\mathbb{S}^n$ and $\mathbb{D}^{n+1}$ is true. In fact, the proof is similar but it requires a higher dimensional algebraic topological object.

PROOF. Suppose otherwise, so we have the communtative diagrams.

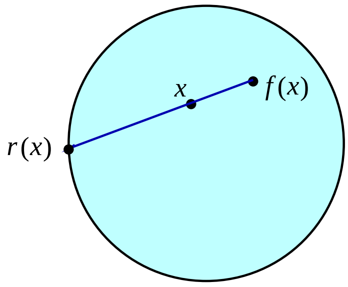
$$\begin{tikzcd}
(\mathbb{S}^1, x_0) \arrow[rr, "i"] \arrow[dr, "\mathrm{id}_{\mathbb{S}^1}"'] & & (\mathbb{D}^2, x_0) \arrow[dl, "r"] \\
& (\mathbb{S}^1, x_0) &
\end{tikzcd}
\qquad
\begin{tikzcd}
\pi_1(\mathbb{S}^1, x_0) \arrow[rr, "i_{\#}"] \arrow[dr, "\mathrm{id}_{\#}"'] & & \pi_1(\mathbb{D}^2, x_0) \arrow[dl, "r_{\#}"] \\
& \pi_1(\mathbb{S}^1, x_0) &
\end{tikzcd}$$

Note that $\pi_1(\mathbb{S}^1, x_0) \cong (\mathbb{Z}, +)$ and $\pi_1(\mathbb{D}^2, x_0)\bar{(}0, +)$. Therefore, in the second diagram, take $1 \in (\mathbb{Z}, +)\overline{\pi}_1(\mathbb{S}^1, x_0)$, one has $i_{\#}(1) = 0$ and so $r_{\#}\, i_{\#}(1) = 0$. On the other hand, $(r \circ i)_{\#}(1) = \mathrm{id}(1) = 1 \neq 0$. This leads to a contradiction. $\square$

Assuming the above Theorem 7.55 is true for general $\mathbb{S}^n$ in $\mathbb{D}^{n+1}$, we have the following important result.

THEOREM 7.56. **(Brouwer Fixed Point Theorem)** *Every continuous function $f \colon \mathbb{D}^n \ \to \ \mathbb{D}^n$ has a fixed point, i.e., $x_0 \in \mathbb{D}^n$ such that $f(x_0) = x_0$.*

Proof. Suppose otherwise, i.e., for every $x \in \mathbb{D}^n$, $f(x) \neq x$. Then for each $x \in \mathbb{D}^n$, a point $r(x)$ can be defined such that $r(x), x, f(x)$ lie on a straight line according to the following picture.



In fact, we have a continuous function $r\colon \mathbb{D}^n \to \mathbb{S}^{n-1}$. Show this by finding the explicit expression of $r(x)$, Exercise 7.7.3.

It can be easily verified that $r|_{\mathbb{S}^{n-1}} \equiv \mathrm{id}_{\mathbb{S}^{n-1}}$. Thus, it leads to the contradiction that $\mathbb{S}^{n-1}$ is a retract of $\mathbb{D}^n$. $\square$