Part I

A Geometric Introduction to Topology

1

Basic point set topology

1.1 Topology in $\mathbb{R}^n$

Topology is the branch of geometry that studies "geometrical objects" under the equivalence relation of homeomorphism. A *homeomorphism* is a function $f: X \to Y$ which is a bijection (so it has an inverse $f^{-1}: Y \to X$) with both f and f^{-1} being continuous. One of the prime aims of this chapter will be to enhance our understanding of the concept of continuity and the equivalence relation of homeomorphism. We will also discuss more precisely the "geometrical objects" in which we are interested (called *topological spaces*), but our viewpoint will primarily be to understand more familiar spaces better (such as surfaces) rather than to explore the full generalities of topological spaces. In fact, all of the spaces we will be interested in exist as subspaces of some Euclidean space $\mathbb{R}^n$. Thus our first priority will be to understand continuity and homeomorphism for maps $f: X \to Y$, where $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$. We will use bold face $\boldsymbol{x}$ to denote points in $\mathbb{R}^k$.

One of the methods of mathematics is to abstract central ideas from many examples and then study the abstract concept by itself. Although it often seems to the student that such an abstraction is hard to relate to in that we are frequently disregarding important information of the particular examples we have in mind, the technique has been very successful in mathematics. Frequently, the success is rooted in the following idea: knowing less about something limits the avenues of approach available in studying it and this makes it easier to prove theorems (if they are true). Of course, the measure of the success of the abstracted idea and the definitions it suggests is frequently whether the facts we can prove are useful back in the specific situations which led us to abstract the idea in the first place. Some of the most important contributions to mathematics have been made by those who have figured out good definitions. This is difficult for the student to appreciate since definitions are usually presented as if they came from some supreme being. It is more likely that they have evolved through many wrong guesses and that what is presented is what has survived the test of time.

It is also quite possible that definitions and concepts which seem so right now (or at least after a lot of study) will end up being modified at a later stage.

We now recall from calculus the definition of continuity for a function $f: X \to Y$, where X and Y are subsets of Euclidean spaces.

Definition 1.1.1. f is *continuous at* $\boldsymbol{x} \in X$ if, given $\epsilon > 0$, there is a $\delta > 0$ so that $d(\boldsymbol{x}, \boldsymbol{y}) < \delta$ implies that $d(f(\boldsymbol{x}), f(\boldsymbol{y})) < \epsilon$. Here d indicates the Euclidean *distance function* $d((x_1, \ldots, x_k), (y_1, \ldots, y_k)) = ((x_1 - y_1)^2 + \cdots + (x_k - y_k)^2))^{1/2}$. We say that f is *continuous* if it is continuous at $\boldsymbol{x}$ for all $\boldsymbol{x} \in X$.

It will be convenient to have a slight reformulation of this definition. For $\boldsymbol{z} \in \mathbb{R}^k$, we define the *ball* of radius r about $\boldsymbol{z}$ to be the set $B(\boldsymbol{z}, r) = \{\boldsymbol{y} \in \mathbb{R}^k : d(\boldsymbol{z}, \boldsymbol{y}) < r\}$ If C is a subset of $\mathbb{R}^k$ and $\boldsymbol{z} \in C$, then we will frequently be interested in the intersection $C \cap B(\boldsymbol{z}, r)$, which just consists of those points of C which are within distance r of $\boldsymbol{z}$. We denote by $B_C(\boldsymbol{x}, r) = C \cap B(x, r) = \{\boldsymbol{y} \in C : d(\boldsymbol{y}, \boldsymbol{x}) < r\}$. Our reformulation is given in the following definition.

Definition 1.1.2. $f: X \to Y$ is *continuous at* $\boldsymbol{x} \in X$ if given $\epsilon > 0$, there is a $\delta > 0$ so that $B_X(\boldsymbol{x}, \delta) \subset f^{-1}(B_Y(f(\boldsymbol{x}), \epsilon))$. f is *continuous* if it is continuous at $\boldsymbol{x}$ for all $\boldsymbol{x} \in X$.

Exercise 1.1.1. Show that the reformulation Definition 1.1.2 is equivalent to the original Definition 1.1.1. This requires showing that, if f is continuous in Definition 1.1.1, then it is also continuous in Definition 1.1.2, and vice versa.

We reformulate in words what Definition 1.1.2 requires. It says that a function is continuous at $\boldsymbol{x}$ if, when we look at the set of points in X that are sent to a ball of radius ϵ about $f(\boldsymbol{x})$, no matter what $\epsilon > 0$ is given to us, then this set always contains the intersection of a ball of some radius $\delta > 0$ about $\boldsymbol{x}$ with X. This definition leads naturally to the concept of an open set.

Definition 1.1.3. A set $U \subset \mathbb{R}^k$ is *open* if given any $\boldsymbol{y} \in U$, then there is a number $r > 0$ so that $B(\boldsymbol{y}, r) \subset U$. If X is a subset of $\mathbb{R}^k$ and $U \subset X$, then we say that U is *open in* X if given $\boldsymbol{y} \in U$, then there is a number $r > 0$ so that $B_X(\boldsymbol{y}, r) \subset U$.

In other words, U is an open set in X if it contains all of the points in X that are close enough to any one of its points. What our second definition is saying in terms of open sets is that $f^{-1}(B_Y(\boldsymbol{y}, \epsilon))$ satisfies the definition of an open set in X containing $\boldsymbol{x}$; that is, all of the points in X close enough to $\boldsymbol{x}$ are in it. Before we reformulate the definition of continuity entirely in terms of open sets, we look at a few examples of open sets.

Example 1.1.1. $\mathbb{R}^n$ is an open set in $\mathbb{R}^n$. Here there is little to check, for given $\boldsymbol{x} \in \mathbb{R}^n$, we just note that $B(\boldsymbol{x}, r) \subset \mathbb{R}^n$, no matter what $r > 0$ is.

Example 1.1.2. Note that a ball $B(\boldsymbol{x}, r) \subset \mathbb{R}^n$ is open in $\mathbb{R}^n$. If $\boldsymbol{y} \in B(\boldsymbol{x}, r)$, then if $r' = r - d(\boldsymbol{y}, \boldsymbol{x})$, then $B(\boldsymbol{y}, r') \subset B(\boldsymbol{x}, r)$. To see this, we use the triangle inequality for the distance function: $d(\boldsymbol{z}, \boldsymbol{y}) < r'$ implies that

$$d(\boldsymbol{z}, \boldsymbol{x}) \leq d(\boldsymbol{z}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{x}) < r' + d(\boldsymbol{y}, \boldsymbol{x}) = r.$$

Figure 1.1 illustrates this for the plane.

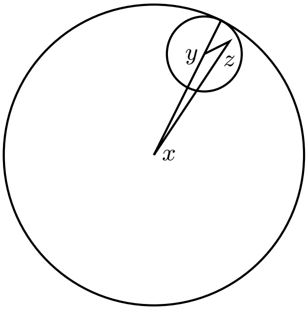


Figure 1.1. Balls are open.

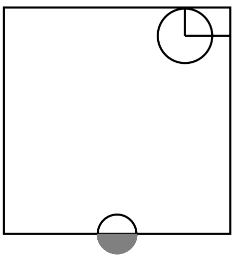


Figure 1.2. Open and closed rectangles.

Example 1.1.3. The inside of a rectangle $R \subset \mathbb{R}^2$, given by $a < x < b$, $c < y < d$, is open. Suppose (x, y) is a point inside of R. Then let $r = \min(b-x, x-a, d-y, y-c)$. Then if $(u, v) \in B((x, y), r)$, we have $|u-x| < r$, $|v-y| < r$, which implies that $a < u < b$, $c < v < d$, so $(u, v) \in R$. However, if the perimeter is included, the rectangle with perimeter is no longer open. For if we take any point on the perimeter, then any ball about the point will contain some point outside the rectangle. We illustrate this in Figure 1.2.

Example 1.1.4. The right half plane, consisting of those points in the plane with first coordinate positive, is open. For given such a point (x, y) with $x > 0$, then if $r = x$, the ball of radius r about (x, y) is still contained in the right half plane. For any $(u, v) \in B((x, y), r)$ satisfies $|u - x| < r$ and so $x - u < x$, which implies $u > 0$.

Example 1.1.5. An interval (a, b) in the line, considered as a subset of the plane (lying on the x-axis), is not open. Any ball about a point in it would have to contain some point with positive y-coordinate, so it would not be contained in (a, b). Note, however, that it is open in the line, because, if $x \in (a, b)$ and $r = \min(b - x, x - a)$, then the intersection of the ball of radius r about x with the line is contained in (a, b). Of course, the line itself is not open in the plane. Thus we have to be careful in dealing with the concept of being open in X, where

X is some subset of a Euclidean space, since a set which is open in X need not be open in the whole space.

Exercise 1.1.2. Determine whether the following subsets of the plane are open. Justify your answers.

(a) $A = \{(x, y) : x \geq 0\}$,
(b) $B = \{(x, y) : x = 0\}$,
(c) $C = \{(x, y) : x > 0 \text{ and } y < 5\}$,
(d) $D = \{(x, y) : xy < 1 \text{ and } x \geq 0\}$,
(e) $E = \{(x, y) : 0 \leq x < 5\}$.

Note that all of these sets are contained in A. Which ones are open in A?

We now give another reformulation for what it means for a function to be continuous in terms of the concept of an open set. This is the definition that has proved to be most useful to topology.

Definition 1.1.4. $f : X \to Y$ is *continuous* if the inverse image of an open set in Y is an open set in X. Symbolically, if U is an open set in Y, then $f^{-1}(U)$ is an open set in X.

Note that this definition is not local (i.e. it is not defining continuity at one point) but is global (defining continuity of the whole function). We verify that this definition is equivalent to Definition 1.1.2. Suppose f is continuous under Definition 1.1.2 and U is an open set in Y. We have to show that $f^{-1}(U)$ is open in X. Let $\boldsymbol{x}$ be a point in $f^{-1}(U)$. We need to find a ball about $\boldsymbol{x}$ so that the intersection of this ball with X is contained in $f^{-1}(U)$. Now $\boldsymbol{x} \in f^{-1}(U)$ implies that $f(\boldsymbol{x}) \in U$, and U open in Y means that there is a number $\epsilon > 0$ so that $B_Y(f(\boldsymbol{x}), \epsilon) \subset U$. But Definition 1.1.2 implies that there is a number $\delta > 0$ so that $B_X(\boldsymbol{x}, \delta) \subset f^{-1}(B_Y(f(\boldsymbol{x}), \epsilon)) \subset f^{-1}(U)$, which means that $f^{-1}(U)$ is open in X; hence f is continuous using Definition 1.1.4.

Suppose that f is continuous under Definition 1.1.4 and $\boldsymbol{x} \in X$. Let $\epsilon > 0$ be given. We noted above that a ball is open in $\mathbb{R}^k$ and the same proof shows that the intersection of a ball with Y is open in Y. Since $B_Y(f(\boldsymbol{x}), \epsilon)$ is open in Y, Definition 1.1.4 implies that $f^{-1}(B_Y(f(\boldsymbol{x}), \epsilon))$ is open in X. But the definition of an open set then implies that there is $\delta > 0$ so that $B_X(\boldsymbol{x}, \delta) \subset f^{-1}(B_Y(f(\boldsymbol{x}), \epsilon))$; hence f is continuous by Definition 1.1.2.

Before continuing with our development of continuity, we recall from calculus some functions which were proved to be continuous there. It is shown in calculus that any differentiable function is continuous. This includes polynomials, various trigonometric and exponential functions, and rational functions. Certain constructions with continuous functions, such as taking sums, products, and quotients (where defined), are shown to give back continuous functions. Other important examples are inclusions of one Euclidean space in another and projections onto Euclidean spaces (e.g. $P(x, y, z) = (x, z)$). Also, compositions of continuous functions are shown to be continuous. We re-prove this latter fact with the open-set definition.

Suppose $f: X \to Y$ and $g: Y \to Z$ are continuous. We want to show that the composition $gf: X \to Z$ is continuous. Let U be an open set in Z. Since g is continuous, $g^{-1}(U)$ is open in Y; since f is continuous, $f^{-1}(g^{-1}(U))$ is open in X. But $f^{-1}(g^{-1}(U)) = (gf)^{-1}(U)$, so we have shown that gf is continuous. Note that in this proof we have not really used that X, Y, Z are contained in some Euclidean spaces and that we have our particular definition of what it means for a subset of Euclidean space to be open. All we really are using in the proof is that in each of X, Y, Z, there is some notion of an open set and the continuous functions are those that have inverse images of open sets being open. Thus the proof would show that even in much more general circumstances, compositions of continuous functions are continuous. We pursue this in the next section.

1.2 Open sets and topological spaces

The notion of an open set plays a basic role in topology. We investigate the properties of open sets in X, where X is a subset of some $\mathbb{R}^n$. First note that the empty set is open since there is nothing to prove, there being no points in it around which we have to have balls. Also, note that X itself is open in X since given any point in X and any ball about it, then the intersection of the ball with X is contained in X. This says nothing about whether X is open in $\mathbb{R}^n$.

Next suppose that $\{U_i\}$ is a collection of open sets in X, where i belongs to some indexing set I. Then we claim that the union of all of the U_i is open in X. For suppose $\boldsymbol{x}$ is a point in the union, then there must be some i with $\boldsymbol{x} \in U_i$. Since U_i is open in X, there is a ball about $\boldsymbol{x}$ with the intersection of this ball with X contained in U_i, hence contained in the union of all of the U_i.

We now consider intersections of open sets. It is not the case that arbitrary intersections of open sets have to be open. For example, if we take our sets to be balls of decreasing radii about a point $\boldsymbol{x}$, where the radii approach 0, then the intersection would just be $\{\boldsymbol{x}\}$ and this point is not an open set in X. However, if we only take the intersection of a finite number of open sets in X, then we claim that this finite intersection is open in X. Let $U_1, \ldots, U_p$ be open sets in X, and suppose $\boldsymbol{x}$ is in their intersection. Then for each i, $i = 1, \ldots, p$, there is a radius $r_i > 0$ so that the intersection of X with the ball of radius r_i about $\boldsymbol{x}$ is contained in U_i. Let r be the minimum of the r_i (we are using the finiteness of the indexing set to know that there is a minimum). Then the ball of radius r is contained in each of the balls of radius r_i, and so its intersection with X is contained in the intersection of the U_i. Hence the intersection is open.

The properties that we just verified about the open sets in X turn out to be the crucial ones when studying the concept of continuity in Euclidean space, and so the natural thing mathematicians do in such a situation is to abstract these important properties and then study them alone. This leads to the definition of a topological space.

Definition 1.2.1. Let X be a set, and let $\mathcal{T} = \{U_i \colon i \in I\}$ be a collection of subsets of X. Then $\mathcal{T}$ is called a *topology* on X, and the sets U_i are called the *open sets* in the topology, if they satisfy the following three properties:

(1) the empty set and X are open sets;
(2) the union of any collection of open sets is open;
(3) the intersection of any finite number of open sets is open.

If $\mathcal{T}$ is a topology on X, then $(X, \mathcal{T})$, or just X itself if $\mathcal{T}$ is made clear by the context, is called a *topological space.*

Our discussion above shows that if X is contained in $\mathbb{R}^n$ and we define the open sets as we have, then X with this collection of open sets is a topological space. This will be referred to as the "standard" or "usual" topology on subsets of $\mathbb{R}^n$ and is the one intended if no topology is explicitly mentioned. Note that Definition 1.1.4 makes sense in any topological space. We use it to define the notion of continuity in a general topological space. Our proof above that the composition of continuous functions is continuous goes through in this more general framework. As we said before, the spaces that we are primarily interested in are those that get their topology from being subsets of some Euclidean space. Nevertheless, it is frequently useful to use the notation of a general topological space and to give more general proofs even though we are dealing with a very special case. We will also use quotient space descriptions of subsets of $\mathbb{R}^n$, which will require us to use topologies more generally defined than those of $\mathbb{R}^n$ and its subsets.

One of the important properties of $\mathbb{R}^n$ and its subsets as topological spaces is that the topology is defined in terms of the Euclidean distance function. A special class of topological spaces are the metric spaces, where the open sets are defined in terms of a distance function.

Definition 1.2.2. Let X be a set and $d\colon X \to \mathbb{R}$ a function. d is called a *metric* on X if it satisfies the following properties:

(1) $d(x, y) \geq 0$ and $= 0$ iff $x = y$;
(2) $d(x, y) = d(y, x)$;
(3) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

The metric d then determines a topology on X, which we denote by $\mathcal{T}_d$, by saying a set U is open if given $x \in U$, there is a ball $B_d(x, r) = \{y \in X : d(x, y) < r\}$ contained in U. $(X, \mathcal{T}_d)$ (or more simply denoted (X, d)) is then called a *metric space.*

To verify that the definition of a topology on a metric space does indeed satisfy the three requirements for a topology is left as an exercise. The proof is essentially our proof that Euclidean space satisfied those conditions. Also, it is easy to verify that the usual distance function in $\mathbb{R}^n$ satisfies the conditions of a metric.

From the point of view of some forms of geometry, the particular distance function used is very important. From the point of view of topology, the important idea is not the distance function itself, but rather the open sets that it determines. Different metrics on a set can determine the same open sets. For

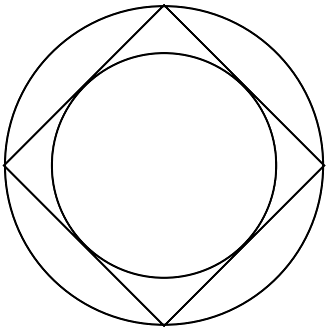


Figure 1.3. Comparing balls.

an example of this, let us consider the plane. Let d denote the usual Euclidean metric in the plane and let $d'((x,y),(u,v)) = |x-u| + |y-v|$. We will leave it as an exercise to verify that d' is a metric. We will use a subscript to indicate the metric being used when determining balls and open sets. As illustrated in Figure 1.3, balls in the metric d' look like diamonds. We show that these two metrics determine the same open sets. Since the open sets are determined by the balls and each type of ball is open, it is enough to show that if $B_d(\boldsymbol{z}, r)$ is a ball about $\boldsymbol{z}$, then there is a number r' so that $B_{d'}(\boldsymbol{z}, r') \subset B_d(\boldsymbol{z}, r)$, and conversely, that each ball $B'_d(\boldsymbol{z}, r')$ contains a ball $B_d(\boldsymbol{z}, r)$. First suppose that we are given a radius r for a ball $B_d(\boldsymbol{z}, r)$. We need to find a radius r' so that $B_{d'}(\boldsymbol{z}, r') \subset B_d(\boldsymbol{z}, r)$. Note that we want $|x_1 - u_1| + |x_2 - u_2| < r'$ to imply that $(x_1 - u_1)^2 + (x_2 - u_2)^2 < r^2$. But if $r' = r$, then this will be true as can be seen by squaring the first inequality. For the other way, given a ball $B_{d'}(\boldsymbol{z}, r')$, we need to find a ball $B_d(\boldsymbol{z}, r)$ within it. Here $r = r'/2$ will work: $(z_1 - u_1)^2 + (z_2 - u_2)^2 < (r')^2/4$ implies that $|z_1 - u_1| < r'/2, |z_2 - u_2| < r'/2$, and so $d'(\boldsymbol{z}, \boldsymbol{u}) < r'$. As Figure 1.3 suggests, we could actually take $r = r'/\sqrt{2}$. This figure shows the inclusions $B_d(\boldsymbol{z}, r/\sqrt{2}) \subset B_{d'}(\boldsymbol{z}, r) \subset B_d(\boldsymbol{z}, r)$.

From the topological point of view, the best value of r given r' is not really of much importance; it is just the existence of an appropriate r. The existence can be seen geometrically.

Exercise 1.2.1. Verify that the definition of an open set for a metric space satisfies the requirements for a topology.

Exercise 1.2.2. Verify that d' is a metric.

We give two examples of a metric space besides the usual topology on a subset of $\mathbb{R}^n$. For the first example, we take as a set $X = \mathbb{R}^n$, but define a metric d by $d(\boldsymbol{x}, \boldsymbol{y}) = 1$ if $\boldsymbol{x} \neq \boldsymbol{y}$, and $d(\boldsymbol{x}, \boldsymbol{x}) = 0$. It is straightforward to check that this satisfies the conditions for a metric. Then a ball $B\left(\boldsymbol{x}, \frac{1}{2}\right) = \{\boldsymbol{x}\}$, so one point sets are open. Hence every set, being a union of one-point sets, will be open. The topology on a set X where all sets are open is called the *discrete topology.*

The next example is of no special importance to us here, but similar constructions are very important in analysis. The points in our space will be continuous functions defined on the interval $[0, 1]$. We can then define the distance between two such functions to be $d(f, g) = \int_0^1 |f(x) - g(x)| \, dx$. We leave it as an exercise to check that this satisfies the definition of a metric.

Exercise 1.2.3. Show that the above definition of the distance between two functions does satisfy the three properties required of a metric. This depends on the fact, which you may assume in your argument, that the integral of a nonnegative continuous function is positive unless the function is identically 0.

We give an example of a topological space which is not a metric space. To define a topology on a set, we have to give a collection of subsets of the set (which we will call open sets) and then verify that they satisfy the three properties required of open sets in a topology. The simplest example of a nonmetric space is to take any set X with more than one point and define the open sets by saying that the only open sets are ϕ and X. This topology is called the *indiscrete topology* on X. For a slightly more complicated example, we will take our set to be the set with three points $\{a, b, c\}$ and then define the following sets to be open: $\phi, \{a, b\}, \{a, b, c\}$. We may verify that this collection of open sets does satisfy the three required properties: the empty set and the whole space are open, unions of open sets are open, and finite intersections of open sets are open. Of course, this is just one of many possible topologies on the three-point set. In order to get a better feeling for the requirements of a topology, we will leave it as an exercise to find some more topologies on this set.

Exercise 1.2.4.

(a) Find five different topologies for the set $\{a, b, c\}$.

(b) Find all the possible topologies on the set $\{a, b, c\}$.

How do we know that the topology that we put on $\{a, b, c\}$ does not arise from some metric? The answer lies in a separation property that any metric space possesses and our topology does not. Given any two distinct points x, y in a metric space, there is some distance $r = d(x, y)$ between them. Then the ball of radius $r/2$ about x does not intersect the ball of radius $r/2$ about y and vice versa. Hence there are two disjoint open sets, one of which contains x and the other y. But this is not true for the points a and b in the topology given above, since every open set which contains b also contains a. The same argument shows that the indiscrete topology on any set X with at least two points does not come from a metric. A topological space X is called *Hausdorff* if given $x, y \in X$ there are disjoint open sets U_x, U_y with $x \in U_x,\ y \in U_y$. The argument above says a metric space is Hausdorff, and our examples are shown not to arise from a metric since they are not Hausdorff.

We look at some specific examples of continuous functions. The inclusion of a subset B of A into A will always be continuous, where $A \subset \mathbb{R}^n$. For if $B \subset A, i : B \to A$ is the inclusion, and if U is an open set in A, then $i^{-1}(U) = B \cap U$. We need to see why $B \cap U$ is open in B if U is open in A. Let $\boldsymbol{x} \in B \cap U$. Then U open in A means that there is a ball $B(\boldsymbol{x}, r)$ with $B(\boldsymbol{x}, r) \cap A \subset U$.

Since $B(\boldsymbol{x}, r) \cap B \subset B \cap U$, $B \cap U$ is open in B. Note that this proof would work equally as well in any metric space as long as we use the same metric for the subset. In a general topological space, we have to specify how we get the topology on the subset from the topology on the original set.

Definition 1.2.3. Suppose A is a topological space and $B \subset A$. A set $V \subset B$ is open in the *subspace topology* on B iff V is the intersection of B with an open set in the whole space A; that is, V is open in B iff $V = U \cap B$, where U is open in A.

It is straightforward to show that an inclusion map is continuous when the subset has the subspace topology. From now on, we will assume that a subset is given the subspace topology unless otherwise stated. The topology on a subset of $\mathbb{R}^n$ coming from using the usual metric is a special case of the subspace topology.

Exercise 1.2.5. For $X \subset \mathbb{R}^n$, show that the usual topology on X is the same as the subspace topology.

Here is another useful construction for continuous functions. Suppose that $f : A \to B$ is continuous and C is a subset of B which contains the image of f. Then we may regard f as a function from A to C. This function, which we denote by f_C, is still continuous when C is given the subspace topology. For if we take an open set V of C, it will have the form $V = U \cap C$, where U is open in B. Then $f_C^{-1}(V) = f^{-1}(U)$ is open since U is open and f is continuous.

Putting these last two constructions together and using the fact that compositions of continuous functions are continuous shows that if we start with a function f from $\mathbb{R}^n$ to $\mathbb{R}^m$ which we already know is continuous, such as a polynomial, and then restrict the function to a subset A and restrict the range to a subset B which contains $f(A)$, then this new function with restricted domain and range will be continuous.

For many constructions involving continuous functions, it is more convenient to work with the concept of closed sets rather than open sets.

Definition 1.2.4. A set $C \subset X$ is said to be *closed* if its complement $X \backslash C$ is open.

From their definition, the closed sets are completely determined by the open sets and vice versa. From the three properties that the open sets satisfy, we can deduce three properties that the closed sets must satisfy:

(1) the empty set and X are closed sets;
(2) the intersection of any collection of closed sets is closed;
(3) the union of any finite number of closed sets is closed.

Critical for verifying these properties from the properties of open sets are DeMorgan's laws regarding complements:

(1) $X \backslash \cup_i A_i = \cap_i (X \backslash A_i)$;
(2) $X \backslash \cap_i A_i = \cup_i (X \backslash A_i)$.

First, the empty set and the whole space X will be closed since their complements (X and the empty set) are open. Second, any intersection of closed sets will be closed since the complement of the intersection will be the union of the individual complements, and thus will be open since the union of open sets is open. Finally, any finite union of closed sets will be closed since the complement of the finite union will be the intersection of the individual complements and so will be open since the finite intersection of open sets is open. It is possible to define a topology in terms of the concept of closed sets and work with closed sets instead of open sets. The most familiar example of a closed set is the closed interval $[a, b]$. We leave it as an exercise to show that it is closed.

Exercise 1.2.6. Show that $[a, b]$ is a closed set in $\mathbb{R}$. Show that a rectangle (including the perimeter) is a closed set in $\mathbb{R}^2$.

Exercise 1.2.7. Show that $[a, b)$ is neither open nor closed in $\mathbb{R}$.

We now prove a couple of useful propositions involving the concept of closed sets. Each proposition follows from corresponding statements involving open sets by taking complements.

Proposition 1.2.1. *$f : A \to B$ is continuous iff the inverse images of closed sets are closed.*

Proof. Suppose f is continuous and C is a closed subset in B. Then $B \backslash C$ is open and $f^{-1}(C) = A \backslash f^{-1}(B \backslash C)$ is closed since it is the complement of an open set in A. The converse follows similarly and is left as an exercise. □

Exercise 1.2.8. Complete the proof above by proving the converse.

Proposition 1.2.2. *If $A \subset X$ has the subspace topology, then $D \subset A$ is closed in A iff $D = A \cap E$, where E is closed in X.*

Proof. By the definition of the subspace topology, the open sets in A are the intersections of A with the open sets in X. What we are trying to prove here is a similar statement for closed sets. Suppose D is closed in A. Then $D = A \backslash F$, where F is open in A. Then $F = A \cap G$, where G is open in X. Hence, if $E = X \backslash G$, then E is closed in X and

$$D = A \backslash F = A \backslash (A \cap G) = A \cap (X \backslash G) = A \cap E.$$

The converse is left as an exercise. □

Exercise 1.2.9. Complete the proof above by proving the converse.

Exercise 1.2.10. Suppose A is a closed subset of X. Then $D \subset A$ is closed in A (with the subspace topology) iff D is closed in X.

Definition 1.2.5. The *closure* of a set $A \subset X$, denoted $\bar{A}$, is the intersection of all closed sets containing A. The *interior* of A, denoted $\operatorname{int} A$, is the union of all open sets contained in A. A point in $\operatorname{int} A$ is called an *interior point* of A. The *boundary* of A, denoted $\operatorname{Bd} A$, is $\bar{A} \cap \overline{X \backslash A}$. A point in $\operatorname{Bd} A$ is called a *boundary point* of A.

Exercise 1.2.11. Show that $\bar{A}$ is closed and $\operatorname{int} A$ is open.

To find $\bar{A}$ in examples, it is useful to have another characterization. Note that a point x is *not* in $\bar{A}$ exactly when there is a closed set C containing A which does not contain x. But this means that $X \backslash C$ is an open set containing x which is disjoint from A, or, equivalently, is contained in $\mathrm{int}(X \backslash A)$. Thus $\bar{A}$ consists of points of A and points not in A that have the property that every open set about them intersects A. Since points of A also have that property, points of $\bar{A}$ can be characterized in that every open set about them intersects A nontrivially. The description of $X \backslash \bar{A}$ above can also be rephrased as saying $X \backslash \bar{A} = \mathrm{int}(X \backslash A)$. Using the definition of $\mathrm{Bd}\, A$ and the reformulation of $\bar{A}$, we can characterize points of $\mathrm{Bd}\, A$ as those points where every open set intersects both A and $X \backslash A$.

As an example, we determine $\bar{A}$, $\mathrm{int}\, A$, and $\mathrm{Bd}\, A$ for $A = \{(x, y) \colon x > y > 0\}$. First note that this set is open since it is the intersection of the two open sets, $A_1 = \{(x, y) \colon x - y > 0\}$ and $A_2 = \{(x, y) \colon y > 0\}$. The sets A_1 and A_2 are open since they are the inverse images of $(0, \infty)$ under the continuous functions $x - y$ and y, respectively. Thus $\mathrm{int}\, A = A$. The closure is found from A by adding the rays $x = y$ and $y = 0$ within the first quadrant. These points are in the closure since every open ball about a point in them will intersect A. The set $B = \{(x, y) \colon x \geq y \geq 0\}$ is closed since it is the intersection of two closed sets, $B_1 = \{x, y) \colon x - y \geq 0\}$ and $B_2 = \{(x, y) \colon y \geq 0\}$. These sets are closed since they are the inverse images of $[0, \infty)$ under the continuous functions $x - y$ and y, respectively. Thus $\bar{A} = B$. The set $X \backslash A$ is closed since its complement is open. Thus $\overline{X \backslash A} = X \backslash A$. Hence $\mathrm{Bd}\, A = \bar{A} \cap \overline{X \backslash A} = \{(x, y) \colon x \geq 0, x = y\} \cup \{(x, y) \colon x \geq 0, y = 0\}$.

Exercise 1.2.12. Find $\bar{A}$, $\mathrm{int}\, A$, and $\mathrm{Bd}\, A$ for the following sets A in $\mathbb{R}^2$:

(a) $\{(x, y) \colon x \geq 0,\ y \neq 0\}$;
(b) $\{(x, y) \colon x \in \mathbb{Q},\ y > 0\}$;
(c) $\{(x, y) \colon x^2 + y^2 < 1\}$.

Exercise 1.2.13. Show that $\bar{A} = \mathrm{Int}\, A \cup \mathrm{Bd}\, A$ and $\mathrm{Int}\, A \cap \mathrm{Bd}\, A = \emptyset$.

We will now prove a piecing lemma, which is very useful in verifying that certain functions which are constructed by piecing together continuous functions are themselves continuous.

Lemma 1.2.3 (Piecing lemma). *Suppose $X = A \cup B$, where A and B are closed subsets of X. Let $f : X \to Y$ be a function so that the restrictions of f to A and B (given the subspace topology) are each continuous (another way of saying this is that the compositions of f with the inclusions of A and B into X give continuous functions). Then f is continuous.*

Proof. Let $C \subset Y$ be closed. Our hypothesis then says that $A \cap f^{-1}(C)$ is closed in A and $B \cap f^{-1}(C)$ is closed in B. But Exercise 1.2.10 then says that these two sets are in fact closed in X since A and B are assumed to be closed subsets of X. Then $f^{-1}(C) = (A \cap f^{-1}(C)) \cup (B \cap f^{-1}(C))$ is closed since it is the union of two closed sets. □

Exercise 1.2.14. Prove the analog of Lemma 1.2.3 where the word closed is replaced by the word open. Give an example to show that the conclusion that f is continuous is not true without some hypothesis on the sets A, B.

We will give many examples of continuous functions in the next section constructed by piecing together continuous functions defined on closed subsets. We state the definition of a homeomorphism and give the relevant version of the piecing lemma for homeomorphisms.

Definition 1.2.6. A *homeomorphism* is a bijection (1–1 and onto) between topological spaces so that the map and its inverse are both continuous. If $f : X \to Y$ is a homeomorphism, then we will say X is *homeomorphic* to Y, denoted $X \simeq Y$.

Homeomorphism gives an equivalence relation on topological spaces, as it satisfies the three conditions of an equivalence relation: (1) reflexivity—the identity $1_X : X \to X$ has continuous inverse 1_X; (2) symmetry—if $f : X \to Y$ has continuous inverse $g : Y \to X$, then g has f as its continuous inverse; (3) transitivity—if $f : X \to Y$ has continuous inverse f^{-1}, and $g : Y \to Z$ has continuous inverse g^{-1}, then $gf : X \to Z$ has continuous inverse $f^{-1}g^{-1}$. A topologist looks at homeomorphic spaces as being essentially the same. One of the fundamental problems of topology is to decide when two topological spaces are homeomorphic. One technique for solving this problem (more successful in showing that spaces are not homeomorphic than in showing that they are homeomorphic) is to find properties of spaces which are preserved by homeomorphisms. We will study two such properties in this chapter, compactness and connectedness. Later we will study an invariant that is associated to any topological space called the fundamental group of the space. It has the property that homeomorphic topological spaces have isomorphic fundamental groups, and thus it may be used to distinguish between topological spaces up to homeomorphism.

We state our lemma for piecing together homeomorphisms. It follows from the piecing lemma in a straightforward manner, and we leave the proof as an exercise.

Lemma 1.2.4 (Piecing lemma for homeomorphisms). *Suppose that $X = A \cup B, Y = C \cup D$, where A, B are closed in X, and C, D are closed in Y. Let $f : A \to C$ and $g : B \to D$ be homeomorphisms, and suppose that the restrictions of f and g to the intersection $A \cap B$ agree as maps into Y. Define $h : X \to Y$ by $h|A = f$ and $h|B = g$ (or we could start with h and define f and g just by restricting them to A and B). If h is a bijection (this just requires that the only points that are in the image of both f and g are the points in the image of $A \cap B$), then h is a homeomorphism.*

Exercise 1.2.15. Prove the piecing lemma for homeomorphisms.

1.3 Geometric constructions of planar homeomorphisms

We now look at some geometric constructions which give continuous functions and homeomorphisms. For simplicity, we will restrict our domain space to the plane, although these constructions have analogues for other $\mathbb{R}^n$.

Our first example is a rotation. If a point in the plane is given by $r(\cos\theta, \sin\theta)$, then a rotation by an angle ϕ sends this to $r(\cos(\theta+\phi), \sin(\theta+\phi))$. One way of seeing that this is continuous is to note that distances between points are unchanged by this map. A map between metric spaces which leaves the distance between any two points unchanged is continuous; we leave this as an exercise.

Exercise 1.3.1.

(a) Show that any map from $\mathbb{R}^2$ to $\mathbb{R}^2$ which leaves distances between points unchanged (i.e. $d(f(\boldsymbol{x}), f(\boldsymbol{y})) = d(\boldsymbol{x}, \boldsymbol{y})$) is continuous.

(b) Generalize this to show that $f\colon (X, d) \to (Y, d')$ with $d'(f(x), f(y)) \leq Kd(x, y), K > 0$, is continuous.

That a rotation does in fact preserve distances can be checked using trigonometric formulas and the distance formula in the plane. Another way of seeing that a rotation by ϕ is continuous is to note that it is given by a linear map, $\boldsymbol{x} \to A\boldsymbol{x}$, where $\boldsymbol{x}$ represents a point in the plane as a column vector and A is the 2×2 matrix

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

For a rotation, A is an orthogonal matrix, which means that it preserves the Euclidean inner product between vectors, and hence preserves distances between points. Multiplication by any matrix can be shown to give a continuous map. This is usually shown indirectly in advanced calculus courses by noting that a linear map is differentiable (it gives its own derivative) and that differentiable maps are continuous. It could also be shown directly using part (b) of Exercise 1.3.1 and the inequality $|A\boldsymbol{x} - A\boldsymbol{y}| \leq \|A\|\,|\boldsymbol{x} - \boldsymbol{y}|$ shown in linear algebra. Note that a rotation is reversible; after rotating a point by an angle θ, we can get back to our original point by rotating by an angle $-\theta$. From the matrix point of view, the matrix A is invertible. Either way may be used to show that rotation represents a homeomorphism from the plane to itself.

Another familiar geometric operation which gives a continuous map (and a homeomorphism) is a translation, $T_{\boldsymbol{v}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{v}$. This is seen to be continuous either directly from the definition or by the fact that it preserves distances between points. Its inverse is translation by $-\boldsymbol{v}$, and so it gives a homeomorphism.

Of course, we could rotate about some other point besides the origin. This also preserves distances and so can be shown to give a homeomorphism. Note that a rotation by angle ϕ about the point $\boldsymbol{x}$ is the composition of a translation by $-\boldsymbol{x}$ to send $\boldsymbol{x}$ to the origin, then a rotation of angle ϕ about the origin, and finally a translation by $\boldsymbol{x}$ to send the origin back to $\boldsymbol{x}$. A composition of homeomorphisms will give a homeomorphism, since a composition of continuous maps is continuous and the inverse of gf, given that g and f have inverses, is $f^{-1}g^{-1}$.

Another geometric construction which gives a homeomorphism is a reflection through a line. That this gives a homeomorphism follows from the fact that it is its own inverse and that it preserves distances between points. Alternatively, reflections through lines passing through the origin are given by multiplication by orthogonal matrices, and other reflections are conjugate to these using translations which move the line to one passing through the origin.

We may reinterpret the equivalence relation of congruence of triangles frequently studied in high school in terms of these three types of homeomorphisms: translations, rotations, and reflections. Suppose two triangles T_1, T_2 are congruent. Then they have corresponding sides A_1, B_1, C_1 and A_2, B_2, C_2, which are of the same length, and the angles between corresponding sides are the same. Let $\boldsymbol{v}_1$ be the vertex between A_1 and B_1 and $\boldsymbol{v}_2$ the vertex between A_2 and B_2. First translate the plane so that the vertex $\boldsymbol{v}_2$ is sent to $\boldsymbol{v}_1$. Now rotate about $\boldsymbol{v}_1$ so that the side A_2 lies along the side A_1. Now either the two triangles will agree or we can get from shifted triangle T_2 to T_1 by reflecting through the line going through side A_1. Thus two triangles are congruent if we can get from one to the other by a composition of translations, rotations, and reflections. Note that each type of map used above preserves distances between points. A map from the plane to itself which preserves distances between points is called a *rigid motion* or an *isometry*. In general, the term *isometry* is used for a map between metric spaces which preserves distance between points and their images.

It can be shown that any rigid motion of the plane is just a composition of translations, rotations, and reflections. We outline this argument. Starting with a rigid motion f, we get a new rigid motion g from f by translating by $-f(\mathbf{0}) : g = T_{-f(\mathbf{0})}f$. Then $g(\mathbf{0}) = \mathbf{0}$. Now we use the relation of the dot product with the distance function $\langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle = d(\boldsymbol{x},\boldsymbol{y})^2$ to show that $g(\mathbf{0}) = \mathbf{0}$ and $d(g(\boldsymbol{x}), g(\boldsymbol{y})) = d(\boldsymbol{x},\boldsymbol{y})$ implies that $\langle g(\boldsymbol{x}), g(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$. Thus g will send unit vectors to unit vectors and orthogonal vectors to orthogonal vectors. In particular, $\boldsymbol{q}_1 = g(\boldsymbol{e}_1), \boldsymbol{q}_2 = g(\boldsymbol{e}_2)$ are orthogonal unit vectors. If Q denotes multiplication by the orthogonal matrix $\begin{pmatrix}\boldsymbol{q}_1 & \boldsymbol{q}_2\end{pmatrix}$ with column vectors $\boldsymbol{q}_1, \boldsymbol{q}_2$, then Q is a rotation or reflection, and $h = Q^{-1}g$ is a rigid motion which preserves $\mathbf{0}, \boldsymbol{e}_1, \boldsymbol{e}_2$. Then h can be shown to be the identity by using the relation $\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{e}_1\rangle \boldsymbol{e}_1 + \langle \boldsymbol{v}_2, \boldsymbol{e}_2\rangle \boldsymbol{e}_2$.

Exercise 1.3.2. Fill in the details of the argument sketched above to show that a rigid motion in the plane is the composition of rotations, reflections, and translations.

Another familiar geometric relation is the similarity of triangles. If two triangles are similar, their angles will correspond exactly, but corresponding side lengths need not be equal but only have to have some common ratio k. If T_1 and T_2 are similar, we may use a rigid motion to align them so that sides A_1 and A_2 lie on the same line, as do the sides B_1 and B_2. Then the shifted T_2 will be sent to T_1 by a map that takes a line through $\boldsymbol{v}_1$ and sends the line to itself by

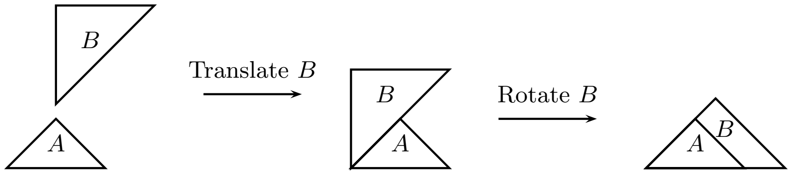


Figure 1.4. Similarity transformation.

shrinking or expanding along the line by a factor of k (in terms of the distance to $\boldsymbol{v}_1$). This last map may be described as a composition of a translation of $\boldsymbol{v}_1$ to the origin, multiplication of a vector by k, and then a translation of the origin back to $\boldsymbol{v}_1$. The multiplication by k gives a continuous map, and its inverse is given by multiplication by $1/k$, so it gives a homeomorphism. We illustrate the first three steps in a similarity in the Figure 1.4. In this figure, no reflection was necessary as part of the rigid motion.

We have seen that congruences and similarities are both examples of homeomorphisms. In geometry, a triangle and a rectangle are distinguished from one another by the number of sides, and two triangles, although possibly not similar, still are seen to have the same "shape". We will see below that the inside of a triangle and the inside of a rectangle are in fact homeomorphic. Thus what is meant when one says that two triangles have the same shape and a triangle and a rectangle do not? It means that we are looking at the triangle and rectangle through "affine linear eyes".

There is a standard triangle $\Delta(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2)$ with vertices $\boldsymbol{e}_0 = \boldsymbol{0}$, $\boldsymbol{e}_1$, $\boldsymbol{e}_2$. Each point in it can be expressed as $(\lambda_1, \lambda_2) = \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2$, with $\lambda_1, \lambda_2 \geq 0$ and $0 \leq \lambda_1 + \lambda_2 \leq 1$. We define $\lambda_0 = 1 - \lambda_1 - \lambda_2$, and then we can write $(\lambda_1, \lambda_2) = \lambda_0\boldsymbol{e}_0 + \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2$, where $\lambda_0, \lambda_1, \lambda_2 \geq 0$ and $\lambda_0 + \lambda_1 + \lambda_2 = 1$. Now suppose we have another triangle with vertices $\boldsymbol{e}_0, \boldsymbol{v}_1, \boldsymbol{v}_2$, where $\boldsymbol{v}_1, \boldsymbol{v}_2$ are linearly independent. If $V = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{pmatrix}$, then multiplication by V is a linear transformation which gives a homeomorphism between $\Delta(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2)$ and $\Delta(\boldsymbol{e}_0, \boldsymbol{v}_1, \boldsymbol{v}_2)$. If three points $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ satisfy the property that $\boldsymbol{v}_1 = \boldsymbol{a}_1 - \boldsymbol{a}_0$, $\boldsymbol{v}_2 = \boldsymbol{a}_2 - \boldsymbol{a}_0$ are linearly independent, then we say that $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ are *affinely independent*. This is equivalent to $\lambda_1\boldsymbol{a}_0 + \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 = 0$, $\lambda_0 + \lambda_1 + \lambda_2 = 0$ implying $\lambda_0 = \lambda_1 = \lambda_2 = 0$. If $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ are affinely independent, then there is a triangle $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$ with vertices $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$. Translation by $\boldsymbol{a}_0$ gives a homeomorphism between $\Delta(\boldsymbol{e}_0, \boldsymbol{v}_1, \boldsymbol{v}_2)$ and $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$, where $\boldsymbol{v}_1 = \boldsymbol{a}_1 - \boldsymbol{a}_0$, $\boldsymbol{v}_2 = \boldsymbol{a}_2 - \boldsymbol{a}_0$. The composition of multiplication by V and the translation then gives a map, called an *affine linear* map, which is a homeomorphism between the standard triangle $\Delta(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2)$ and $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$. This affine linear map A has the property that $\lambda_0\boldsymbol{e}_0 + \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2$ is sent to $\lambda_0\boldsymbol{a}_0 + \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2$. In particular, this means that the triangle $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$ is characterized as the points $\lambda_0\boldsymbol{a}_0 + \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2$ where $\lambda_i \geq 0$, $\lambda_0 + \lambda_1 + \lambda_2 = 1$. If $\Delta(\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2)$ is another triangle with affinely independent vertices $\boldsymbol{b}_0, \boldsymbol{b}_1$, $\boldsymbol{b}_2$, then there is an affine linear map B sending the standard triangle to it. Then $C = BA^{-1}$ gives an affine linear map sending $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$ to $\Delta(\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2)$. Thus any two triangles in the plane are homeomorphic via a canonical affine linear map, and the image of a triangle under an affine linear map will be another triangle. In particular, there is no affine linear map sending a triangle to a rectangle. Affine linear maps from one triangle $\Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$ to another triangle $\Delta(\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2)$ are determined completely by the map on the vertices $\boldsymbol{a}_i \to \boldsymbol{b}_i$ and the affine linearity condition $\sum \lambda_i \boldsymbol{a}_i \to \sum \lambda_i \boldsymbol{a}_i$.

Exercise 1.3.3.

(a) Show that $\boldsymbol{a}_1 - \boldsymbol{a}_0,\ \boldsymbol{a}_2 - \boldsymbol{a}_0$ are linearly independent iff $\lambda_0 \boldsymbol{a}_0 + \lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 = \mathbf{0},\ \lambda_0 + \lambda_1 + \lambda_2 = 0$ implies $\lambda_0 = \lambda_1 = \lambda_2 = 0$.

(b) Show that if $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ are affinely independent, then $\lambda_1 \boldsymbol{a}_0 + \lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 = \mu_1 \boldsymbol{a}_0 + \mu_1 \boldsymbol{a}_1 + \mu_2 \boldsymbol{a}_2$ with $\sum \lambda_i = \sum \mu_i = 1$ implies $\mu_i = \lambda_i, i = 0, 1, 2$.

(c) Show that any finite composition of translations and linear maps in the plane can be written as a single composition TL, where T is a translation and L is a linear map.

(d) Show that any composition M of translations and linear maps satisfies $M(\sum_{i=1}^k \lambda_i \boldsymbol{a}_i) = \sum_{i=1}^k \lambda_i M(\boldsymbol{a}_i)$ when $\sum_{i=1}^k \lambda_i = 1$. Conversely, show that if M satisfies this condition for any three affinely independent points, then M is a composition of a translation and a linear map, so is an affine linear map.

(e) Show that an affine linear map sending $\boldsymbol{a}_i$ to $\boldsymbol{b}_i$ will always send a line segment $\overline{\boldsymbol{a}_0\boldsymbol{a}_1}$ to the line segment $\overline{\boldsymbol{b}_0\boldsymbol{b}_1}$ via $(1-t)\boldsymbol{a}_0 + t\boldsymbol{a}_1 \to (1-t)\boldsymbol{b}_0 + t\boldsymbol{b}_1,\ 0 \le t \le 1$.

Triangles and rectangles are not equivalent under invertible affine linear maps. A triangle and a rectangle are homeomorphic, however. Moreover, the homeomorphism may be taken to be "piecewise linear". If $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_3$ are the vertices of the triangle and $\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ are the vertices of the rectangle, then we can divide the rectangle into two triangles $B_1 = \Delta(\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2), B_2 = \Delta(\boldsymbol{b}_0, \boldsymbol{b}_2, \boldsymbol{b}_3)$ by introducing the edge $\overline{\boldsymbol{b}_0\boldsymbol{b}_2}$ (see Figure 1.5). We can also introduce a vertex $\boldsymbol{a}_2$ in the triangle at the midpoint of $\overline{\boldsymbol{a}_1\boldsymbol{a}_3}$ and then an edge $\overline{\boldsymbol{a}_0\boldsymbol{a}_2}$. Now the triangle is divided into two triangles, $A_1 = \Delta(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2)$ and $A_2 = \Delta(\boldsymbol{a}_0, \boldsymbol{a}_2, \boldsymbol{a}_3)$. The

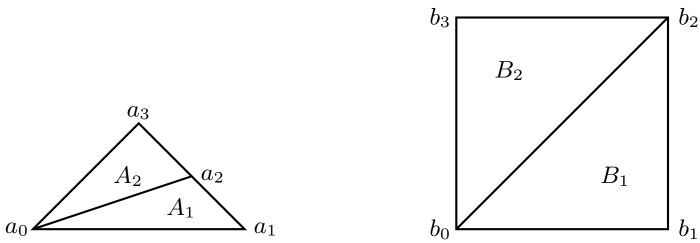


Figure 1.5. PL homeomorphism between a triangle and a rectangle.

map sending $\boldsymbol{a}_i$ to $\boldsymbol{b}_i$ can be extended affine linearly on triangles to give maps sending A_i to B_i. Figure 1.5 shows how the triangle and square are subdivided. This defines a homeomorphism between the triangle and the rectangle. That it is a homeomorphism follows from the piecing lemma for homeomorphisms. Note that on the triangles A_1, A_2, the map is affine linear $(\sum_i \lambda_i \boldsymbol{a}_i \to \sum_i \lambda_i \boldsymbol{b}_i)$. Our homeomorphism is an example of a *piecewise linear (PL) homeomorphism* of planar regions—the domain and range are divided into triangles, and the homeomorphism is an affine linear homeomorphism on each triangle.

We can generalize the argument above to show that any two convex polygonal regions in the plane are homeomorphic. By a polygonal region R, we mean a region that is bounded by a *closed polygonal path*; that is, $f([0,n])$, where $f\colon [0,n] \to \mathbb{R}^2$ with f affine linear on $[i, i+1]$, $f(i) = \boldsymbol{x}_i$, $\boldsymbol{x}_0 = \boldsymbol{x}_n$ and $f(a) = f(b)$ implies $a = b$ or $\{a,b\} = \{0,n\}$. The region R is called *convex* if R lies on one side of each line $\overline{\boldsymbol{x}_i \boldsymbol{x}_{i+1}}$ or, equivalently, line segments joining two points of R are in R. The region R bounded by P is then given by the union of line segments joining points in P. The idea of the proof that two convex polygonal regions are homeomorphic is to divide each region into the same number of triangles and then send the triangles to each other consistently. Our argument above with a triangle and a rectangle is the simplest case of this procedure.

Exercise 1.3.4.

(a) Construct a PL homeomorphism between a square and a hexagon.

(b) Show that any two convex polygonal regions are homeomorphic via a PL homeomorphism.

So far all of our examples of homeomorphisms have been piecewise linear. Here is an example of one that is not. The unit disk $D^2 = \{(x,y)\colon x^2 + y^2 \le 1\}$ is homeomorphic to the square $S = \{(x,y)\colon |x| \le 1, |y| \le 1\}$ (hence to any convex polygonal region). The homeomorphism may be described geometrically as follows. Each ray from the origin intersects D^2 and S in a line segment. The intersection with D^2 is sent linearly to the intersection with S.

We can verify that this is a homeomorphism by deriving a formula for it. This is somewhat tedious, however, so we will give a geometrical explanation, leaving the verification based on this as an exercise. We describe some corresponding open sets from our construction. Given a point $\boldsymbol{x}$ inside the disk which is not the center, we get $f(\boldsymbol{x})$ by first forming the circle about the center on which $\boldsymbol{x}$ lies, then forming the square which circumscribes this circle, and then sending $\boldsymbol{x}$ to the the point $f(\boldsymbol{x})$ on the intersection of the perimeter of this square and the ray through $\boldsymbol{x}$. The region between two circles is then sent to the region between the corresponding squares. The basic open sets inside the circle are given by the region between two circles, which lie between two lines of angles $\theta = \theta_1,\ \theta = \theta_2$, as well as disks about the center. For the inside of a square, the basic open sets are given by regions between two smaller squares, again limited by the same two radial lines, as well as small squares about the center. Our map gives a correspondence between these basic open sets about $\boldsymbol{x}$ and $f(\boldsymbol{x})$ as pictured in Figure 1.6. At the center, a small disk about the center corresponds to a small

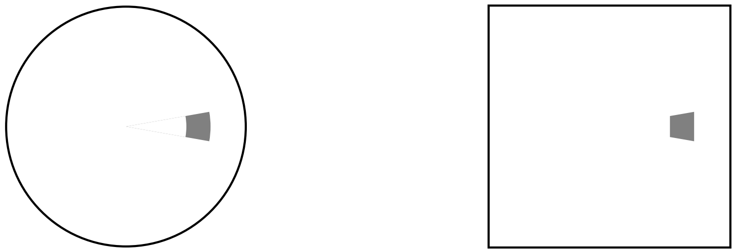


Figure 1.6. Basic open sets for disk and square.

rectangle about the center. From these facts, we can verify that the map is a homeomorphism.

Exercise 1.3.5. Use the geometrical facts cited above to verify that our construction gives a homeomorphism. You will need to use the fact that any open set about a point contains one of the basic open sets as described above.

Note that this homeomorphism sends the boundary circle to the perimeter of the square. In fact, if the homeomorphism of the circle to the perimeter of the square is given by $\boldsymbol{x} \to f(\boldsymbol{x})$, then our homeomorphism is just $t\boldsymbol{x} \to tf(\boldsymbol{x}),\ 0 \leq t \leq 1$. We are using the convexity of each region to realize the region as the "cone" on its boundary and extending the homeomorphisms of boundaries by "coning".

This same idea could be used to give a homeomorphism between the unit disk and the inside of an ellipse, for example.

Exercise 1.3.6. Write down a formula for a homeomorphism between the unit disk D^2 and the ellipse $E = \{(x, y)\colon x^2 + y^2/4 \leq 1\}$, and check whether it satisfies $f(tx, ty) = tf(x, y),\ 0 \leq t \leq 1$.

In the exercise above and the preceding example, there is a common idea. We take two subspaces in the plane A, B and points $p \notin A,\ q \notin B$. Then we form spaces pA, qB from taking the line segments joining p to points of A and line segments joining q to points of B. The set A is chosen so that each point in pA lies on a unique line segment from p to a unique point of A (and similarly for qB). In the case of the inside of the circle and the inside of the square, A is the circle and B is the square. For the disk and the inside of the ellipse, A is the circle and B is the ellipse. In both cases, $p = q = \mathbf{0}$. Then we take a homeomorphism $f\colon A \to B$, and then get a homeomorphism $F\colon pA \to qB$ by sending $(1 - t)p + ta$ to $(1 - t)q + tf(a)$. That F turns out to be a homeomorphism depends on pA, qB having the appropriate types of corresponding basic open sets. This can be rephrased in terms of the notion of a quotient topology, which we will study in Section 1.7. The construction of F from f is called *coning*.

We have seen many examples of different regions in the plane that turn out to be homeomorphic. Each of the regions so far has been homeomorphic to a disk. An important problem of topology is to characterize all regions in the plane that are homeomorphic to the disk. The homeomorphism would send the circle to a homeomorphic image—this is called a *simple closed curve*. Thus, a region R homeomorphic to a disk would have to be "bounded" by a simple closed curve.

The Jordan curve theorem and the Schönflies theorem combine to say that, if C is a simple closed curve in the plane, then it "bounds" a region R, and the homeomorphism $f : S^1 \to C$ extends to a homeomorphism between the unit disk D and R. The Jordan curve theorem says that the complement of the curve separates into two open connected pieces, one of which is bounded and the other of which is unbounded. It says the curve is the boundary of each piece. The Schönflies theorem then says that the bounded piece is homeomorphic to a disk and the unbounded piece is homeomorphic to the complement of a closed disk. We discuss connectedness in Section 1.6 and have a project to prove both theorems in the polygonal case in Section 1.8. A full proof of the Jordan curve theorem and it's generalization, the Jordan separation theorem, is given in terms of homology in Section 6.14 (see Theorems 6.14.2 and 6.14.6). The full proof of the Schönflies theorem can be found in [22]. A proof of the generalization of the Schönflies theorem to higher dimensions for locally flat embeddings is given in [5] based on the proof by Morton Brown [6].

A natural question would be to ask for examples of regions in the plane that are not homeomorphic to a disk. A simple example would be an annulus (see Figure 1.7), which is the region enclosed between two circles. There are two ways of seeing that this is not homeomorphic to a disk. One way is to compare their boundaries. The annulus has two circles as boundary and the disk has one. Of course, we have to understand why one circle is not homeomorphic to two circles (this can be based on the concept of connectedness, which we will study later) and why a homeomorphism between the annulus and the disk must restrict to a homeomorphism between their boundaries. A justification of the last fact actually leads us to the other reason that they are not homeomorphic. This involves the ideas surrounding the fundamental group of a space. Intuitively speaking, there is a circle (the middle circle) in the annulus which cannot be deformed continuously to a point, but every circle in the disk may be deformed

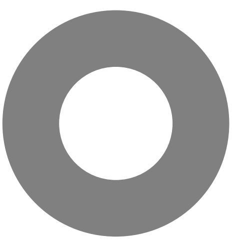


Figure 1.7. Annulus.

to a point (just contract the whole disk to its center and see what happens to the circle). This idea is responsible for a large number of applications and is pursued in Chapter 3. The classification of regions in the plane up to homeomorphism is a special case of the classification of surfaces with boundary. This latter topic is pursued in Chapter 2.

1.4 Compactness

We now discuss the property of compactness. We will discuss this in the context of a general topological space, but will specialize to metric spaces or subspaces of $\mathbb{R}^n$ on occasion.

Definition 1.4.1. Let X be a topological space. A subset $A \subset X$ is said to be *compact* if whenever A is contained in a union of open sets U_i (called an *open cover* of A), then A is contained in the union of a finite subcollection of these open sets (called a *finite subcover*).

This can be rephrased in terms of the open sets of A in the subspace topology by saying that whenever A is written as the union of a collection of open sets in A, then it may be written as the union of a finite number of these open sets.

One of the prime reasons that compactness is important as a topological concept is that it is preserved by continuous maps.

Proposition 1.4.1. *Let $f : X \to Y$ be continuous and X compact. Then the image set $f(X)$ is compact.*

Proof. Let $\mathcal{V} = \{V_i\}$ be an open cover of $f(X)$. Then $\mathcal{U} = \{U_i\} = \{f^{-1}(V_i)\}$ is an open cover of X. Since X is compact, there is a finite subcover $U_{i(1)}, \ldots, U_{i(k)}$ of X. Then the corresponding open sets $V_{i(1)}, \ldots, V_{i(k)}$ give a finite subcover of $f(X)$. □

In particular, this implies that if two sets are homeomorphic, then either both are compact or both are not compact. A property that is invariant under homeomorphisms is called a *topological invariant.* Thus compactness is a topological invariant.

Let us look at some examples.

Example 1.4.1. The real line $\mathbb{R}$ is not compact since it can be written as the union of intervals $U_k = (-k, k)$ where k ranges over the integers, and it cannot be written as a union of a finite subcollection of these open sets. The same idea will show that, for a subset of $\mathbb{R}$ to be compact, it must be bounded (i.e. contained in a large interval). For if it is not, then we can use the collection $\{(-k, k)\}$ to cover the subset, and it cannot be contained in any finite subcollection of these. We leave it as an exercise to generalize this to subsets of metric spaces.

Exercise 1.4.1. A set A of a metric space is said to be *bounded* if it is contained in some ball $B(x, r)$. Show that a subset of a metric space which is compact must be bounded.

Example 1.4.2. A finite set $A = \{a_1, \ldots, a_k\} \subset X$ is compact. For if it is contained in a union of open sets U_i, then there must be some set $U_{i(j)}$ in the collection which contains a_j. Thus $U_{i(1)}, \ldots, U_{i(k)}$ gives a finite subcover of A.

Exercise 1.4.2. Show that a finite union of compact sets is compact.

Exercise 1.4.3. Decide whether or not the following subsets of $\mathbb{R}$ are compact:

(a) $A = \{1/n : n \in \mathbb{N}\}$;
(b) $B = \{0\} \cup A$;
(c) $(0, 1]$.

We have seen that $\mathbb{R}$ is not compact, but $\mathbb{R}$ is closed as a subset of itself. Thus a closed set does not have to be compact. A compact set does not have to be closed in a general topological space, either. For example, the two-point space, where the only open sets are the empty set and the space itself, has either of its points as a compact subset, but that point is not a closed set with this topology. However, if we are dealing with subsets of Euclidean space and the standard topology, then compact sets are closed. We will give a proof in the more general situation of a metric space.

Proposition 1.4.2. *In a metric space, compact sets are closed.*

Proof. Let X be a metric space and A a compact subset of X. To show that A is closed, we have to show that its complement is open. Let $x \in X \backslash A$; we need to find a ball about x that does not intersect A. Let y be a point of A. Then we can find disjoint balls $B(y, r(y))$ and $B(x, r(x))$. The union of the $B(y, r(y))$ over all y in A will contain A; since A is compact, there is a finite subcollection of these balls which covers A. Then the intersection of the corresponding balls about x will be an open set about x which does not intersect the union of the subcollection, and hence does not intersect A. □

The crucial property of a metric space X which we used here was that a metric space is Hausdorff.

Definition 1.4.2. X is called *Hausdorff* if given $x, y \in X$ then there are disjoint open sets U_x, U_y with $x \in U_x, y \in U_y$.

Exercise 1.4.4. Show that in a Hausdorff space, compact sets are closed. (Hint: In a general topological space, a set U will be open if given $x \in U$, then there is an open set $U(x)$ with $x \in U(x) \subset U$; for then we can write U as the union of the sets $U(x)$ as x ranges over the points of U, and the union of open sets is open. With this criterion for a set to be open, the proof in the metric case can be modified to prove the result.)

The next proposition allows us to deduce that certain sets are compact by knowing that they are closed subsets of a compact set.

Proposition 1.4.3. *Let X be compact and let A be closed in X. Then A is compact.*

Proof. Suppose that $\mathcal{U} = \{U_i\}$ is a collection of open sets of X whose union contains A. Then the U_i together with $X \backslash A$ is a collection of open sets whose union is X, and so some finite subcollection will contain X. Since no points of A are contained in $X \backslash A$, then the U_i in this subcollection will contain A. □

We combine the propositions connecting compact and closed sets to prove the following very useful proposition that certain bijections between sets are homeomorphisms.

Proposition 1.4.4. *Let $f : X \to Y$ be a bijection (i.e. 1–1 and onto). Suppose that f is continuous, X is compact, and Y is Hausdorff. Then f is a homeomorphism.*

Proof. Since f is a bijection, it has an inverse $f^{-1} : Y \to X$. To see that f is a homeomorphism, we need to see that f^{-1} is continuous. We use here the characterization of a continuous function as one which has the inverse image of a closed set being closed. Let C be a closed set in X. Then X compact implies that C is compact. But $(f^{-1})^{-1}(C) = f(C)$ is the image of a compact set, and so is compact. In a Hausdorff space, a compact set is closed, so $f(C)$ is closed as required. □

In the proof above, we showed that if $f : X \to Y$ is continuous, X is compact, and Y is Hausdorff, then f sends closed sets to closed sets. A map which sends closed sets to closed sets is called a *closed map*. When f is invertible, then f^{-1} being continuous is the same thing as f being a closed map.

This proposition would no longer be true if we removed the hypothesis that X is compact. For example, consider the function $f : [0, 1) \to S^1$ given by $f(x) = (\cos 2\pi x,\ \sin 2\pi x)$.

Exercise 1.4.5. Show that the function f defined above is a bijection that is continuous but is not a homeomorphism. (Hint: Consider the open set $\left[0, \frac{1}{2}\right) \subset [0, 1)$ and its image.)

We begin studying some basic compact sets in the reals. We first show that a closed interval $[a, b]$ is compact in the usual topology of the line. This proof is based on the least upper bound property of the real numbers, which we now review. A subset $A \subset \mathbb{R}$ is said to have an *upper bound* u if $a \leq u$ for all $a \in A$. u is called the *least upper bound* of A if it is a upper bound and it is less than or equal to any other upper bound. The *least upper bound property* of the real numbers asserts that any nonempty subset of the reals with an upper bound has a least upper bound. This property does not hold for the rationals; for example, the set of rational numbers with square less than 2 has an upper bound, but does not have a least upper bound. As a subset of the reals, the least upper bound would be $\sqrt{2}$. We can think of the reals as being formed from the rationals by adding to the rationals all the least upper bounds of subsets of the rationals that are not already in the rationals.

Theorem 1.4.5. *The closed interval $[a, b]$ is compact.*

Proof. Suppose that we have an open cover $\mathcal{U} = \{U_i\}$ of $[a, b]$. Consider the set $A = \{x \in [a, b]\colon [a, x]$ has a finite subcover$\}$. We intend to show that $A = [a, b]$. First note that A is not empty since some U_i contains a, and thus must contain some interval $[a, b_1]$, for $b_1 > a$. Since b is an upper bound for A, the set A must have a least upper bound, which we will call u. We want to show that $u = b$ and that $b \in A$. Suppose first that $u < b$. Since $u \in [a, b]$, there must be some element of the cover, which we will call $U_{i(u)}$, which contains u. Now $U_{i(u)}$ contains some interval $[u_1, u_2]$, where $a < u_1 < u < u_2 < b$. Since u is the least upper bound for A, there must be an element $a_1 \in A$ with $u_1 < a_1 \leq u$ (if not, then u_1 would be a smaller upper bound, contradicting the choice of u as the least upper bound). But then $[a, a_1]$ is covered by a finite number of the U_i and thus so is $[a, u_2]$ (just use those that cover $[a, a_1]$ together with $U_{i(u)}$). But this contradicts u being an upper bound for A, since now $u_2 \in A$. Thus the least upper bound must be b. Now choose an element $U_{i(b)}$ of the cover which contains b, and choose u_1 with $[u_1, b] \subset U_{i(b)}$. Then b being the least upper bound for A implies that there is an element $a_1 \in A$ with $u_1 < a_1 \leq b$. But $[a, a_1]$ is covered by a finite number of the U_i and $[a_1, b]$ is contained in $U_{i(b)}$, so $[a, b]$ is contained in a finite number of the elements of the cover, showing that it is compact. □

As a corollary, we can now characterize the compact sets in the line.

Corollary 1.4.6. *$A \subset \mathbb{R}$ is compact iff it is closed and bounded.*

Proof. If it is compact, then it must be bounded by Exercise 1.4.1 and closed by Proposition 1.4.2. Conversely, suppose that it is closed and bounded. Since it is bounded, it is contained in some closed interval $[a, b]$. Since it is closed as a subset of the line, it will also be closed in $[a, b]$. But this makes it a closed subset of a compact space, and so it is compact. □

Exercise 1.4.6. Analogous to the definition of least upper bound is that of greatest lower bound. Give a definition of greatest lower bound for a set $A \subset \mathbb{R}$ and use the least upper bound property to show that a set with a lower bound must have a greatest lower bound.

Exercise 1.4.7. Give an example of a closed, bounded subset A of a metric space X that is not compact. (Hint: Consider the metric space X itself to be a bounded noncompact subset of $\mathbb{R}$ and $A = X$.)

For $\mathbb{R}$ we extract an important property of a closed bounded set.

Proposition 1.4.7. *A compact subset A of $\mathbb{R}$ has a largest element M and a smallest element m; that is, $m \leq a \leq M$ for all $a \in A$.*

Proof. We show that it has a largest element; the proof for a smallest element is analogous. Since A is compact, it is bounded, and so has a least upper bound u. We claim that $u \in A$, and hence u will be the largest element of A. Suppose that u is not in A; then we claim that A could not be closed. For every interval about u has to contain an element of A in order for u to be the least upper bound of A. But this means that the complement of A is not open; hence A is not closed. □

Now we give an application of this to analysis.

Proposition 1.4.8. *Let* $f : X \to \mathbb{R}$ *be continuous and* X *compact. Then* f *assumes a maximum (and minimum) on* X; *that is, there are* $x, y \in X$ *with* $f(x) \leq f(z) \leq f(y)$ *for all* $z \in X$.

Proof. To say that f assumes a maximum just means that $f(X)$ has a largest element. But X compact and f continuous means that $f(X)$ is compact and so has a largest element. □

When X is a closed interval, this is the familiar theorem from calculus that a continuous function assumes a maximum and a minimum on a closed interval.

1.5 The product topology and compactness in $\mathbb{R}^n$

We wish to generalize our characterization of compact sets in $\mathbb{R}$ to show that a subset of $\mathbb{R}^n$ is compact iff it is closed and bounded. The only missing ingredient from our proof above is knowing that a cube $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ is compact. This can be proved inductively if we can show that the product of compact sets in a product of Euclidean spaces is compact. To do this most efficiently, we need to discuss the notion of a product topology on the product $X \times Y$ of two topological spaces.

Suppose that X and Y are topological spaces and consider their product $X \times Y = \{(x, y) : x \in X,\ y \in Y\}$. We will define a topology on $X \times Y$ by saying that a set $W \subset X \times Y$ is open if given any $(x, y) \in W$, then there are open sets U in X and V in Y so that $(x, y) \in U \times V \subset W$. In particular, products of open sets will be open, and the general open set will be a union of products of open sets. It is not difficult to verify that this definition of open sets does satisfy the three requirements for a topology, which is called the *product topology.*

Exercise 1.5.1. Verify that open sets as defined above satisfy the three properties required of a topology.

Now the product topology in the plane is not defined in exactly the same way as the usual metric topology, but it does give the same topology; that is, it gives the same collection of open sets. To see this, first note that if a set W is open in the plane in the usual metric topology, and (x, y) is a point of A, then there is a small ball about (x, y) that is contained in W. But inside this ball we can find a rectangle that is a product of intervals which contains (x, y). Hence W is open in the product topology. Conversely, suppose W is open in the product topology, and $(x, y) \in W$. Then there is a product $U \times V$ of open sets (which we may choose to be intervals) with $(x, y) \in U \times V \subset W$. Then the rectangle $U \times V$ is contained in W. But then we can find a ball contained in the rectangle and containing (x, y), so W is open in the metric topology (see Figure 1.3). Inductively, a similar argument shows that the metric topology on $\mathbb{R}^n$ arises as the inductive product of n copies of $\mathbb{R}$ using the product topology.

Thus to show that a product of closed intervals is compact in $\mathbb{R}^n$, it suffices to show that the product of compact sets is compact in the product topology. We first need a preliminary lemma on product spaces.

Proposition 1.5.1. *Suppose X and Y are topological spaces and let $X \times Y$ have the product topology. Then the inclusions $i_x : Y \to X \times Y, i_x(y) = (x, y)$, $i_y : X \to X \times Y$, $i_y(x) = (x, y)$, are continuous. Moreover, each projection $p_X : X \times Y \to X$, $p_X(x, y) = x$, $p_Y : X \times Y \to Y$, $p_Y(x, y) = y$, is continuous. In particular, the map $X \to X \times \{y\}$ given from i_y by restricting the range, and $Y \to \{x\} \times Y$ given from i_x similarly, are homeomorphisms, where $X \times \{y\}$ and $\{x\} \times Y$ are given the subspace topology.*

Proof. We first show that i_x is continuous; the proof is analogous for i_y. Let W be an open set in the product topology on $X \times Y$, and let $y \in i_x^{-1}(W)$. Then $(x, y) \in W$, so there are open sets U, V with $(x, y) \in U \times V \subset W$. Then $y \in V \subset i_x^{-1}(W)$, so $i_x^{-1}(W)$ is open (using the hint in Exercise 1.4.4). We now show that p_X is continuous; the proof for p_Y is analogous. Let U be an open set in X. Then $p_X^{-1}(U) = U \times Y$, which is an open set in the product topology. Finally, note that i_x and p_Y are inverses to one another (when properly restricted) and so give homeomorphisms between Y and $\{x\} \times Y$; similarly, i_y and p_X give homeomorphisms between X and $X \times \{y\}$. □

We now show that the product of compact spaces is compact.

Theorem 1.5.2 (Tychanoff). *Suppose X and Y are compact. Then the product $X \times Y$ is compact.*

Proof. Let $\mathcal{W} = \{W_i\}$ be an open cover of the product. Fix $x \in X$ and consider the set $\{x\} \times Y$. It is homeomorphic to Y, so it is compact. Thus there are a finite number of the W_i, which we will denote by $W_{i_{x,1}}, \ldots, W_{i_{x,k}}$, which cover $\{x\} \times Y$. Let $W_x = W_{i_{x,1}} \cup \cdots \cup W_{i_{x,k}}$. Then for each $y \in Y$, select an open set

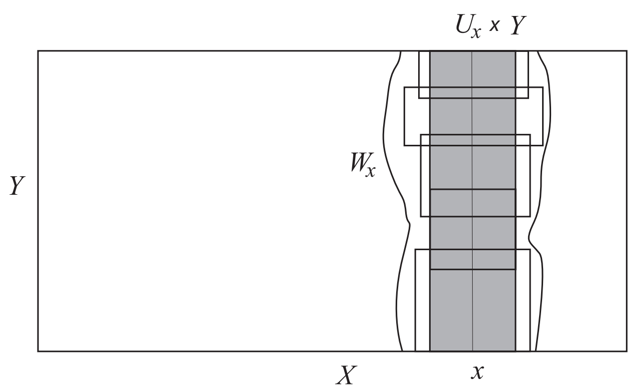


Figure 1.8. A tube $U_x \times Y \subset W_x$.

$U_y \times V_y$ with $(x, y) \in U_y \times V_y \subset W_x$. Then this gives an open cover of $\{x\} \times Y$, and so there is a finite subcover $U_{y_1} \times V_{y_1}, \ldots, U_{y_p} \times V_{y_p}$. Let $U_x = \bigcap_{j=1}^{p} U_{y_j}$. Note that $\{x\} \times Y \subset U_x \times Y \subset W_x$. The set $U_x \times Y$ is sometimes called a *tube* about $\{x\} \times Y$ inside W_x. This is illustrated in Figure 1.8. As x varies over X, the sets U_x give an open cover of X and so there is a finite subcover $U_{x(1)}, \ldots, U_{x(r)}$. Then $U_{x(1)} \times Y, \ldots, U_{x(r)} \times Y$ will cover $X \times Y$, and so will the corresponding $W_{x(i)}$. But each $W_{x(i)}$ is the union of a finite number of sets in our original cover, and so we will get a covering by a finite number of sets in our original cover. □

The Tychanoff theorem holds for infinite products as well, and it is closely related to the axiom of choice. See [24] for a discussion and proof in this context.

Now we are ready to characterize the compact sets in $\mathbb{R}^n$.

Theorem 1.5.3 (Heine–Borel). *A subset of $\mathbb{R}^n$ is compact iff it is closed and bounded.*

Proof. We showed that compact implies closed and bounded in a metric space. Suppose A is closed and bounded. Then A will be a closed subset of some large cube (which is compact) and hence will be compact. □

We now wish to introduce another form of compactness, sequential compactness, and show that it is equivalent to compactness in a metric space. In general, these concepts are not equivalent but counterexamples are rather sophisticated. In the course of doing so, we will also introduce the concept of the Lebesgue number of a cover, and show that compact metric spaces have Lebesgue numbers, a fact which will be very useful to us in Chapter 3.

Definition 1.5.1. A sequence in X is a function $s : \mathbb{N} \to X$, where $\mathbb{N}$ denotes the natural numbers. We usually denote $s(n)$ by s_n and the sequence by $\{s_n\}$. A subsequence s' of a sequence s is a sequence formed by taking the composition $s' = sj$, where $j : \mathbb{N} \to \mathbb{N}$ is order preserving ($a < b$ implies $j(a) < j(b)$). It is usually denoted by s_{n_i} where $n_i = j(i)$. A sequence is said to converge to x if given an open set U about x, there is a natural number N so that $n > N$ implies $s_n \in U$.

Definition 1.5.2. X is called *sequentially compact* if every sequence in X has a convergent subsequence.

We wish to give a criterion for a sequence to have a subsequence which converges to x. If a subsequence converges to x, then the definition of convergence implies that for any open set U containing x, there are an infinite number of values of n so that $s_n \in U$. We show the converse is true in a metric space.

Proposition 1.5.4. *Suppose X is a metric space and $x \in X$. If $\{s_n\}$ is a sequence so that for any ball about x, the ball contains an infinite number of the s_n (this means that there are an infinite number of n so that s_n is in the ball), then there is a subsequence which converges to x.*

Proof. Choose n_1 so that s_{n_1} is contained in the ball of radius 1 about x. Since there are an infinite number of the s_n in the ball of radius $\frac{1}{2}$ about x, we can find n_2 so that $n_2 > n_1$ and $s_{n_2} \in B\left(x, \frac{1}{2}\right)$. Inductively, we then use the same idea to choose $n_1 < n_2 < n_3 < \cdots$ so that $s_{n_j} \in B(x, 1/j)$. This will give us our convergent subsequence. We leave the details as an exercise. □

Exercise 1.5.2. Fill in the details in the proof above.

Proposition 1.5.5. *In a metric space, compactness implies sequential compactness.*

Proof. We prove the contrapositive. Suppose X is not sequentially compact and s_n is a sequence with no convergent subsequence. If there are only a finite number of distinct values s_n, then some value must be repeated infinitely often and we can use this to get a constant, hence convergent, subsequence. Thus we may assume that there are an infinite number of distinct values s_n. For each $x \in X$, there is no subsequence which converges to x. By the criterion above, there is an open set U_x about x which contains only a finite number of the s_n. But a covering of X by these balls, one for each x, can have no finite subcover, since a finite subcover would have to contain only a finite number of the values of the sequence (which are infinite in number), and hence could not contain all of X. □

The proof above does not need the full strength of the metric space hypothesis, just the existence for each x of a sequence of open sets U_n with $U_{i+1} \subset U_i$ about x so that any open set about x contains some U_i. This property is called first countability and is pursued in Exercises 1.9.39–1.9.41 at the end of the chapter.

Exercise 1.5.3. Show that if $\{s_n : n \in \mathbb{N}\}$ is finite, then the sequence has a convergent subsequence.

We now show that in a metric space, sequential compactness implies compactness. To prove this, we introduce the concept of the Lebesgue number of a cover. Let A be a subset of the metric space (X, d). Consider $D_A = \{d(a_1, a_2) : a_1, a_2 \in A\}$. If D_A is bounded from above, define $d_A = \sup D_A$. We will call d_A the *diameter* of the set A.

Definition 1.5.3. A covering $\mathcal{U} = \{U_i\}$ of a metric space is said to have *Lebesgue number* $\delta > 0$ if every set $A \subset X$ of diameter less than δ is contained in some element of the covering.

Proposition 1.5.6. *Let X be a metric space which is sequentially compact. Then every open covering of X has a Lebesgue number.*

Proof. We prove the contrapositive: if there is an open cover with no Lebesgue number, then there is a sequence with no convergent subsequence. Let $\mathcal{U} = \{U_i\}$ be an open cover with no Lebesgue number. Then there is a sequence of sets $\{A_n\}$ with the diameter of A_n less than $1/n$ which are not contained in any element of the cover. Choose $a_n \in A_n$. Then we claim that $\{a_n\}$ is a sequence with no convergent subsequence. Suppose there were a subsequence $\{a_{n_k}\}$ which converges to a point x, and choose an element U_p of the cover containing x.

Choose m large enough so that $B(x, 1/m) \subset U_p$, and choose $k_1 \geq 2m$ so that if $k \geq k_1$, $a_{n_k} \in B(x, 1/2m)$. Then if $a \in A_{n_k}$, $d(a,x) \leq d(a, a_{n_k}) + d(a_{n_k}, x) < 1/2m + 1/2m = 1/m$. But this means $A_{n_k} \subset U_p$, which is a contradiction. □

Proposition 1.5.7. *In a metric space, sequential compactness implies compactness.*

Proof. A metric space is *totally bounded* if given $\epsilon > 0$, we can cover X by a finite number of balls of radius ϵ. We first show that X sequentially compact implies that it is totally bounded. We show this by proving the contrapositive. Suppose X cannot be covered by a finite number of balls of radius ϵ. Let $x_1 \in X$. Since $B(x_1, \epsilon)$ does not cover X, choose $x_2 \notin B(x_1, \epsilon)$. Since $B(x_1, \epsilon) \cup B(x_2, \epsilon)$ does not cover X, we may choose $x_2 \notin B(x_1, \epsilon) \cup B(x_2, \epsilon)$. Inductively, we can choose a sequence $\{x_n\}$ in this manner with $x_{n+1} \notin \bigcup_{k=1}^{n} B(x_k, \epsilon)$. Since $d(x_n, x_k) \geq \epsilon$ for $k < n$, any ball of diameter ϵ can contain at most one x_n, so the sequence can have no convergent subsequence.

Now suppose X is sequentially compact and $\mathcal{U} = \{U_i\}$ is an open cover. By Proposition 1.5.6 we can find a Lebesgue number δ for this cover. By the above argument, there is a cover of X by a finite number of balls of radius $\delta/3$. But each such ball will be of diameter less than δ, so it will lie in an element of our original cover, $B(x_k, \delta/3) \subset U_{i(k)}, k = 1, \ldots, n$. Then $U_{i(1)}, \ldots, U_{i(n)}$ give a finite subcover of our original cover. □

Definition 1.5.4. Let $(X, d_X), (Y, d_Y)$ be metric spaces. Then $f : X \to Y$ is said to be *uniformly continuous* if given $\epsilon > 0$ there exists a $\delta > 0$ such that for $x_1, x_2 \in X, d_X(x_1, x_2) < \delta$ implies $d_Y(f(x_1), f(x_2)) < \epsilon$.

Exercise 1.5.4. Show that f uniformly continuous implies f is continuous, but construct an example to show that the converse does not hold.

Exercise 1.5.5. Let $f : X \to Y$ be a continuous map of the compact metric space (X, d_X) to the metric space (Y, d_Y). Show that f is uniformly continuous. (Hint: Use the Lebesgue number of the covering $\{f^{-1}(B(y, \epsilon/2))\}_{y \in Y}$ of X.)

1.6 Connectedness

We next want to discuss the concept of connectedness. The definition is given in terms of its negation, as it is easier to say what we mean by a space not being connected.

Definition 1.6.1. A topological space X is called *separated* if it is the union of two disjoint, nonempty open sets. A subset $A \subset X$ is *separated* if A is separated as a topological space, using the subspace topology. A set is called *connected* if it is not separated.

Exercise 1.6.1. Show that a space X is connected iff the only subsets of X which are both open and closed are $\emptyset$ and X.

We rephrase the conditions for a subset $A \subset X$ to be separated or connected in terms of open sets in X.

Proposition 1.6.1.

(a) *$A \subset X$ is separated iff there are two open sets $U, V \subset X$ so that $A \subset U \cup V$, $U \cap V \cap A = \emptyset$, $U \cap A \neq \emptyset, V \cap A \neq \emptyset$.*

(b) *A set $A \subset X$ is connected iff whenever U, V are open sets in X so that $U \cap V \cap A = \phi, A \subset U \cup V$, then $A \subset U$ or $A \subset V$.*

Proof. We only prove (a), leaving (b) as an exercise. Suppose A is separated. Then there are two disjoint nonempty sets U', V' which are open in A so that $A = U' \cup V'$. Since U', V' are open in A, there are open sets U, V in X with $U' = U \cap A$, $V' = V \cap A$. Since U' and V' are disjoint, we have $U \cap V \cap A = \emptyset$. Since $A = U' \cup V'$, we have $A \subset U \cup V$. This proves one direction. For the other direction, given U, V with $A \subset U \cup V$, $U \cap V \cap A = \emptyset, U \cap A \neq \emptyset$, $V \cap A \neq \emptyset$, then defining $U' = U \cap A$, $V' = V \cap A$ gives two nonempty sets U', V' which are open in A and show that A is separated. □

Exercise 1.6.2. Deduce (b) from (a).

Example 1.6.1. We use Proposition 1.6.1 to describe some examples of separated sets. The first example we give is the union of two points in the line $X = \{0, 1\}$. To see that this is separated, we choose $U = (-0.1, 0.1), V = (0.9, 1.1)$. A similar example would be to let $Y = [0, 1] \cup [2, 3]$. Then we could choose $U = (-0.1, 1.1), V = (1.9, 3.1)$. Our final example may be somewhat less intutive. The rationals $\mathbb{Q}$ in the line are separated. Here we can choose $U = (-\infty, \sqrt{2}), V = (\sqrt{2}, \infty)$. We will show that the $\mathbb{R}$ itself is connected, so the missing irrational numbers were crucial in separating the rational ones. Note that the openness condition in the definition is crucial. For example, you cannot get an interval being separated by dividing it into two pieces, say $[0, 2] = [0, 1] \cup (1, 2]$. The problem is that to get an open set U about $[0, 1]$ you have to include points greater than 1 and so it will not be disjoint with an open set about $(1, 2]$.

We first investigate connectedness for subsets of the line. Consider the following property:

(*) If $x, y \in A \subset \mathbb{R}$, then the interval $[x, y] \subset A$.

Proposition 1.6.2. *Any connected set in the line satisfies (*) or, equivalently, any set that does not satisfy (*) is separated.*

Proof. If A does not satisfy (*), then there are points x, y, z with $x < y < z$ and $x, z \in A$ and $y \notin A$. But then A is separated by the two open sets $(-\infty, y)$ and (y, ∞). □

What are the sets that satisfy (*)? The next proposition says that they are just the intervals, rays, and $\mathbb{R}$.

Proposition 1.6.3. *A set $A \subset \mathbb{R}$ satisfies (*) iff it is an interval, a ray, or $\mathbb{R}$.*

Proof. It is straightforward to see that an interval, ray, or $\mathbb{R}$ satisfies (*). Suppose A satisfies (*). There are a number of cases to consider; we will only consider one of the cases and leave the completion of the proof as an exercise. We consider the case where A is bounded both from above and below. Let a be the greatest lower bound and b the least upper bound of A. This implies $A \subset [a, b]$. We will show that $(a, b) \subset A$. Let c be a point in (a, b). Since a is the greatest lower bound, there is an element $e \in A$ with $a \leq e < c$. Similarly, b being the least upper bound implies that there is an element $f \in A$ with $c < f \leq b$. But (*) implies that $[e, f] \subset A$ and so $c \in A$. Hence $(a, b) \subset A$. But $A \subset [a, b]$, so there are four possibilities for $A\colon (a, b), [a, b), (a, b], [a, b]$, each of which is an interval. The other cases one has to consider are when A is not bounded on one side or the other or both. □

Exercise 1.6.3. Complete the proof of the proposition by considering the other cases.

Our previous two propositions say that the only possibilities for connected sets in $\mathbb{R}$ are intervals, rays, and $\mathbb{R}$. We now show that they are connected.

Proposition 1.6.4. *Any interval, ray, or* $\mathbb{R}$ *is connected.*

Proof. We will just give the proof for a closed interval $[a, b]$, and leave the other cases for the reader. In Proposition 1.6.1 we re-expressed the condition of connectivity by saying that a set is connected if, whenever it is contained in the union of two open sets U, V with $U \cap V \cap A = \emptyset$, then it is entirely contained in one of the two sets. Suppose that $[a, b]$ is contained in the union of two open sets U, V with $U \cap V \cap [a, b] = \emptyset$. Assume that $a \in U$. To show that $[a, b]$ is connected, we must show that $[a, b] \subset U$. Analogously to the proof that $[a, b]$ is compact, we form the set $A = \{x \in [a, b]\colon [a, x] \subset U\}$. Since U is open, we see that A contains some $x > a$. Since A is bounded, it must have a least upper bound u. We first claim that $u \in U$. If not, then $u \in V$ and so there is $u_1 < u$ with $[u_1, u] \subset V$ since V is open. But u being the least upper bound of A means that there is $c \in A$ with $u_1 < c \leq u$. But then $c \in U \cap V \cap A$, which is a contradiction. If $u \neq b$, we can find an interval $[u_1, u_2] \subset U$, with $u_1 < u < u_2$, and so $[a, u_2] \subset U$, contradicting the choice of u as an upper bound. Thus we must have $[a, b] \subset U$, and so $[a, b]$ is connected. □

Exercise 1.6.4. Show that $\mathbb{R}$ is connected.

The three preceding propositions together yield the following theorem.

Theorem 1.6.5. *The connected sets in* $\mathbb{R}$ *are intervals, rays, and* $\mathbb{R}$.

Here is a useful proposition about connectedness, which could be used to show that $\mathbb{R}$ is connected, knowing that $[a, b]$ is connected.

Proposition 1.6.6. *Suppose that* A_i *is a collection of connected subsets of a topological space* X *so that they all have at least one point* a *in common. Then the union* $A = \cup_i A_i$ *is connected.*

Proof. Suppose $A \subset U \cup V$, where $U \cap V \cap A = \emptyset$, and suppose further that $a \in U$. Then we have to show that $A \subset U$. But each A_i being connected will imply that $A_i \subset U$, so $A \subset U$. □

Exercise 1.6.5. Use the proposition above to deduce that $\mathbb{R}$ is connected from the fact that a closed interval is connected.

Unfortunately, there is no nice characterization of connected subsets of other Euclidean spaces as there is for compact subsets, although the above proposition is very useful in recognizing connected sets.

We prove that connectedness is preserved under continuous maps, and hence gives another topological invariant for a space up to homeomorphism.

Proposition 1.6.7. *The continuous image of a connected space is connected.*

Proof. Suppose $f: X \to Y$ is continuous and X is connected. Suppose $f(X) \subset U \cup V$, where $U \cap V \cap f(X) = \emptyset$ and U, V are open. Then $X \subset f^{-1}(U) \cup f^{-1}(V)$, and $X \cap f^{-1}(U) \cap f^{-1}(V) = \emptyset$. Now f continuous implies that $f^{-1}(U)$ and $f^{-1}(V)$ are open, and so X connected means that X is contained in one of these, say $f^{-1}(U)$. Hence $f(X) \subset U$, and so $f(X)$ is connected. □

Since the continuous image of a connected set is connected, so is a homeomorphic image. Hence connectedness is also a topological invariant. This fact could be used to show, for example, that two disjoint intervals could not be homeomorphic to one interval.

A somewhat more intuitive property than connectedness is path connectedness.

Definition 1.6.2. A space X is called *path connected* if, given $x, y \in X$, there is a continuous map $f: [0,1] \to X$ (called a *path* in X) with $f(0) = x$ and $f(1) = y$. We say that the path connects x to y.

There is an equivalence relation generated by this definition as follows: we say $x \sim y$ if there is a path connecting x to y. The constant path at x shows $x \sim x$. That $x \sim y$ implies $y \sim x$ can be seen by composing a path connecting x to y with a self homeomorphism of $[0,1]$ which is order reversing; usually one uses the linear map $\alpha(t) = 1 - t$, but any order reversing homeomorphism will work. That $x \sim y, y \sim z$ implies $x \sim z$ involves reparametrizing the paths and lying them end on end. Geometrically, we just traverse the path connecting x to y and then traverse the path from y to z. However, to get a parametrized path from the two paths involves reparametrizing them so that their domains fit together nicely. For example, we can compose f with $\alpha(t) = 2t$, so $f\alpha(0) = f(0) = x, f\alpha(\frac{1}{2}) = f(1) = y$. Then we could similarly reparametrize g with an affine linear map $\beta: [\frac{1}{2}, 1] \to [0,1]$ and define the path connecting x to z by making it $f\alpha$ on $[0, \frac{1}{2}]$ and $g\beta$ on $[\frac{1}{2}, 1]$. We leave the details as an exercise.

Exercise 1.6.6. Show that the relation $x \sim y$ as defined above is an equivalence relation.

The equivalence classes under this equivalence relation are called the *path components* in X. For example, if $X = [0,1] \cup [2,3]$, then the intervals $[0,1]$ and $[2,3]$ would be the path components. A set is path connected iff it has only one path component.

We show that path connectedness is preserved by continuous maps, hence, by homeomorphisms, so is a topological invariant.

Proposition 1.6.8. *Suppose X is path connected and $f: X \to Y$ is a continuous map. Then $f(X)$ is path connected.*

Proof. Let $u = f(x), v = f(y)$ be points of $f(X)$. Since X is path connected, there is a path α connecting x and y. Then $f\alpha$ is a path connecting u and v. □

The basic relationship between the two forms of connectivity is given by the following proposition.

Proposition 1.6.9. *If X is path connected, then X is connected.*

Proof. Pick a point $x \in X$, and for each point $y \in X$, choose a path connecting x to y. The images of these paths are all connected since they are images of connected sets under continuous maps, and each of them contains x. Their union (as we let y range over all of the points of X) is all of X, and so we get that X is connected by applying Proposition 1.6.6. □

It is not the case that a connected set has to be path connected. Here is an example of a set in the plane, called the *topologist's sine curve*, which is connected but is not path connected. Our set is based on the $\sin 1/x$ curve. Figure 1.9 shows a global and a local view (near a point on the y-axis) of its graph. It is the union of two sets, A and B. Here A is just the graph of $\sin 1/x$, where $0 < x \leq 1$, and B is the segment along the y-axis where the y-coordinate ranges from -1 to 1. To see that $A \cup B$ is connected, the idea is that if it were contained in a union $U \cup V$ of open sets with no points in both U and V, then since A and B are each connected (being the images of connected sets under continuous maps), each would have to lie entirely in one of the sets. Suppose that $B \subset U$. Then since U is open, we can show that at least one point of A must also lie in U. Since A is connected, then all of A must also lie in U and so $A \cup B$ lies in U. That $A \cup B$ is not path connected is based on the idea that there can be no path connecting a point of A to a point of B. The basic idea is to use the fact that a small ball about a point in B will intersect A in an infinite number of disjoint arcs, and to show that for $A \cup B$ to be path connected, we would have to be able to connect points in different arcs while staying in such a ball, which is impossible. Verification of the details are left as Exercises 1.9.44–1.9.46 at the end of the chapter.

We now consider some examples of path connected, hence connected, sets in Euclidean spaces.

Example 1.6.2. As our first example, note that $\mathbb{R}^n$ is path connected. We can take a straight line path connecting any two points $\boldsymbol{x}, \boldsymbol{y}, f(t) = (1-t)\boldsymbol{x} + t\boldsymbol{y}$. By analogous reasoning, any convex set (a set where straight lines joining any two

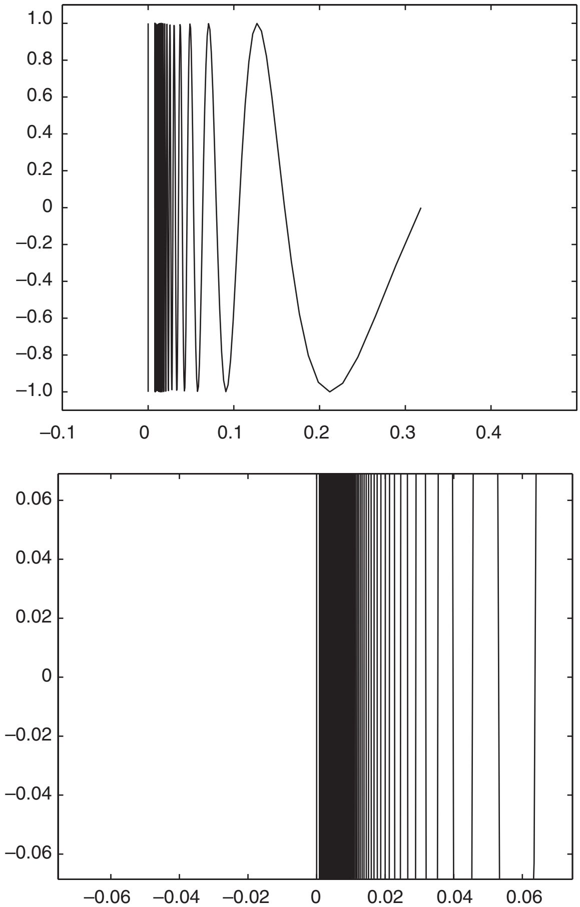


Figure 1.9. The topologist's sine curve—two views.

points in the set lie in the set) is path connected. This contains balls and cubes, for example.

Example 1.6.3. The unit sphere $S^n \subset \mathbb{R}^{n+1}$ is path connected, $n \geq 1$. The best way to see this is to show that $\mathbb{R}^{n+1}\backslash\{\mathbf{0}\}$ is path connected, and then show that there is a continuous map from $\mathbb{R}^{n+1}\backslash\{\mathbf{0}\}$ onto S^n. To see that $\mathbb{R}^{n+1}\backslash\{\mathbf{0}\}$ is path connected, note that if $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{n+1}$ and the straight line joining them does not pass through $\mathbf{0}$, then it may be used to give a path connecting them as before. If it does pass through $\mathbf{0}$, then choose a point $\boldsymbol{z}$ that is not on this line (here we use $n \geq 1$). Then the straight line from $\boldsymbol{x}$ to $\boldsymbol{z}$ together with the straight line from $\boldsymbol{z}$ to $\boldsymbol{y}$ may be used to give a path from $\boldsymbol{x}$ to $\boldsymbol{y}$. We can get a continuous map from $\mathbb{R}^{n+1}\backslash\{\mathbf{0}\}$ onto S^n by projecting along lines through the origin. Precisely, this map is given by the formula, $P(\boldsymbol{x}) = \boldsymbol{x}/|\boldsymbol{x}|$, where $|\boldsymbol{x}|$ denotes the length of $\boldsymbol{x}$.

Exercise 1.6.7. Show that a union of path connected sets with a point in common is path connected. (Hint: Let z be the common point. Then show that given x, y in the union, we can find a path that joins them by using paths in individual path connected spaces that join x to z and join z to y.)

Although a connected set need not be path connected, here is a situation where that is true.

Proposition 1.6.10. *Let A be an open subset of $\mathbb{R}^n$. If A is connected, then A is path connected.*

Proof. We show that A has only one path component, hence is path connected. Note that each path component P is open in A, since each point has a ball about it contained in A and each point of the ball can be connected to the center by a straight line path. If A had more than one path component, let P_1 be a path component and P_2 be the union of the other path components. Then P_1, P_2 give a separation of A into two disjoint, nonempty open sets, a contradiction. □

We conclude this section on connectedness by proving a version of the intermediate value theorem.

Proposition 1.6.11 (Intermediate value theorem). *Suppose that $f : X \to \mathbb{R}$ is a continuous function and X is connected. Let $c = f(x_1)$ and $d = f(x_2)$ and suppose that $c < e < d$. Then there is $x_3 \in X$ with $f(x_3) = e$.*

Exercise 1.6.8. Prove the intermediate value theorem using the fact that $f(X)$ is connected and our characterization of connected sets in the line.

This theorem is encountered in calculus when X is a closed interval $[a, b]$. In this context, it says that a continuous function must assume every value between $f(a)$ and $f(b)$. Another way of stating this is to say that the closed interval with end points $f(a), f(b)$ is a subset of $f([a, b])$. By combining compactness and connectedness, we can describe completely what the image of a closed interval under a continuous map to the reals must be. Since it must be connected, it has to be an interval, a ray, or all of the reals. Since it must be compact, the only possible choice is a closed interval. The end points of this interval will be the minimal value and the maximal value of the function. We state this as a proposition.

Proposition 1.6.12. *If $f : [a, b] \to \mathbb{R}$ is continuous, then $f([a, b]) = [m, M]$ where m, M are the minimal and maximal values of the function.*

Exercise 1.6.9. Show that the letter T is not homeomorphic to the letter O. (Hint: Consider what happens when a point is removed from each letter and the corresponding connectivity properties.)

Exercise 1.6.10. Show that S^1 is not homeomorphic to $\mathbb{R}$ by showing $S^1 \backslash \{x\}$ is not homeomorphic to $\mathbb{R} \backslash \{y\}$.

Exercise 1.6.11. Show that two disjoint concentric circles in the plane are not homeomorphic to one circle.

1.7 Quotient spaces

We discuss the notion of a quotient space, which is also called an identification space. We will be using quotient spaces extensively in Chapter 2 when we study surfaces.

Definition 1.7.1. Suppose X, Y are topological spaces, and we have a surjective map $q: X \to Y$. Then we say Y has the *quotient topology with respect to* (X, q) if $U \subset Y$ is open iff $q^{-1}(U) \subset X$ is open. Y is then called a *quotient space* of X and q is called a *quotient map.*

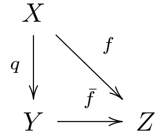
A simple example of a quotient map is the map $q: \mathbb{R} \to S^1$ where $q(t) = \mathrm{e}^{2\pi \mathrm{i} t}$. The arcs in the circle which provide a basis for its topology have as their inverse images the unions of disjoint intervals in the reals.

Note that the map q is continuous when Y has the quotient topology. For whenever $U \subset Y$ is open, the definition of the quotient topology requires that $q^{-1}(U)$ has to be open.

Exercise 1.7.1. Suppose $q: X \to Y$ and Y has the quotient topology with respect to (X, q). Show $C \subset Y$ is closed iff $q^{-1}(C) \subset X$ is closed.

Quotient spaces often arise by starting with some known space X and then forming Y from X by identifying certain points of X, usually by means of an equivalence relation we put on points of X. The map q then sends a point $x \in X$ to the equivalence class of all points that are identified with x. In this context, Y is sometimes called an *identification space* and the quotient map q is called an *identification map.* The equivalence class containing x is denoted by $[x]$ and the map sending a point to its equivalence class is denoted by $q(x) = [x]$. A simple, but quite important, example comes from starting with $X = [0, 1]$, and then making 0 equivalent to 1 the only nontrivial equivalence relation. The quotient space then can be imagined geometrically by taking a piece of string and then joining the end points to get a circle up to homeomorphism for the quotient space Y.

Suppose $f: X \to Z$ is a continuous function and $Y = X/\sim$ is formed from X by identifying points in X within the same equivalence class, $q: X \to Y, q(x) = [x]$. Then f induces a map $\bar{f}: Y \to Z$ if whenever x_1 is equivalent to x_2 then $f(x_1) = f(x_2)$; that is, identified points are sent to the same point by f. We define $\bar{f}$ by $\bar{f}([x]) = f(x)$. This is well defined because, if we choose $[x_1] = [x_2]$, then $x_1 \sim x_2$ and $f(x_1) = f(x_2)$. We are defining $\bar{f}$ by $\bar{f}q(x) = \bar{f}([x]) = f(x)$. We call $\bar{f}$ *the map induced by* f. The quotient topology is set up so that f continuous implies $\bar{f}$ is continuous. For if U is an open set in Z, then to check that $\bar{f}$ is continuous, we verify that $\bar{f}^{-1}(U)$ is open in Y. To check this, we use the quotient map $q: X \to Y, q(x) = [x]$. Then $\bar{f}^{-1}(U)$ is open in Y iff $q^{-1}(\bar{f}^{-1}(U))$ is open in X. Since $\bar{f}q = f$, the condition is that $f^{-1}(U)$ is open, which is true since f is continuous.



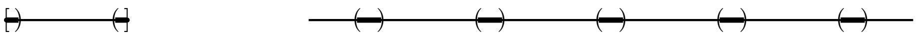


Figure 1.10. Saturated open sets $q^{-1}(U)$ about $[0]$ for $[0,1]$ and $\mathbb{R}$.

When a quotient space is formed by identifying points, it is difficult to picture the equivalence classes directly and the open sets in the quotient space. What we can do, however, is picture their inverse images within the space X. The sets $q^{-1}(U)$ are open sets that are *saturated* with respect to the equivalence relation. This means that if $x \in q^{-1}(U)$ and $x \sim y$, then $y \in q^{-1}(U)$. For a simple example, consider $X = [0,1]$ with the only nontrivial equivalence being $0 \sim 1$. Then the basis for the topology of $X/\sim$ will have inverse images being open intervals in $(0,1)$ and also sets of the form $[0,a) \cup (b,1]$ for $0 < a < b < 1$. The last set is a saturated open set that contains the equivalence class $\{0,1\}$. A related example uses $X' = \mathbb{R}$ and forms the quotient space using the equivalence relation $x \sim x+n, n \in \mathbb{Z}$. A typical equivalence class is $\{\ldots, x-2, x-1, x, x+1, x+2, \ldots\}$. A basic open set U about this point will have inverse image $q^{-1}(U) = \cup_{n\in\mathbb{Z}}(x+n-\epsilon, x+n+\epsilon)$, where $\epsilon < \frac{1}{2}$. This is just an interval about x together with all of its translates by integers. See Figure 1.10.

We prove some elementary propositions about quotient spaces. The first proposition formalizes our last observation about induced maps.

Proposition 1.7.1. *Let Y be a quotient space of X with quotient map $q\colon X \to Y$. Let $g\colon Y \to Z$ be a map. Then g is continuous iff the composition gq is continuous.*

Proof. If g is continuous, then the composition is continuous since the composition of continuous functions is continuous. Conversely, suppose the composition is continuous and $U \subset Z$ is an open set. Look at $g^{-1}(U)$. To see that it is open, we have to show that $q^{-1}(g^{-1}(U))$ is open. But $q^{-1}(g^{-1}(U)) = (gq)^{-1}(U)$, so it is open since gq is continuous. □

Proposition 1.7.2. *Suppose Y is a quotient space with respect to (X,q) and Y' is a quotient space with respect to (X',q'). Let $f\colon X \to X'$, $\bar{f}\colon Y \to Y'$ be maps with $q'f = \bar{f}q$. We also could express this by saying that the following diagram is commutative.*

$$\begin{array}{ccc} X & \xrightarrow{f} & X' \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle q'} \\ Y & \xrightarrow{\bar{f}} & Y' \end{array}$$

Then $\bar{f}$ is continuous if f is continuous.

Proof. To show $\bar{f}$ is continuous, we need to show that $\bar{f}q$ is continuous, by Proposition 1.7.1. But $q'f = \bar{f}q$ and f, q' continuous imply $q'f$ is continuous. □

Proposition 1.4.4 has a nice application for quotient spaces.

Proposition 1.7.3. *Suppose $f : X \to Y$ is a surjective continuous map, X is compact and Y is Hausdorff. Define an equivalence relation on X by saying $u \sim v$ iff $f(u) = f(v)$; the equivalence classes are the inverse images $f^{-1}(y)$. Then the induced map $\bar{f} : X/\sim \to Y$ is a homeomorphism.*

Proof. Proposition 1.7.1 implies that $\bar{f}$ is continuous. It is a bijection since we are identifying points in X which map to the same point. Since $X/\sim$ is the continuous image of the compact space X by the quotient map $q : X \to X/\sim$, we have that $X/\sim$ is compact. Then Proposition 1.4.4 implies that $\bar{f}$ is a homeomorphism. □

We now apply these propositions to the quotient spaces $Y = [0, 1]/\sim$ and $Y' = \mathbb{R}/\sim$. Consider the map $f : [0, 1] \to S^1$ given by $f(t) = (\cos 2\pi t, \sin 2\pi t)$. This is a continuous surjection and the only nontrivial inverse image is $f^{-1}\{(1, 0)\} = \{0, 1\}$. Thus if we form the quotient space Y from the interval $X = [0, 1]$ by identifying 0 with 1, then Proposition 1.7.3 implies that the induced map $\bar{f}$ is a homeomorphism.

We could instead start with $X' = \mathbb{R}$ and identify x with $x + n, n \in \mathbb{Z}$ to form the quotient space Y'. We claim that Y' is also homeomorphic to the circle. We start with the same map p, now considered as a map from the reals. It determines a map $\bar{p} : Y' \to S^1$ by $\bar{p}[t] = (\cos 2\pi t, \sin 2\pi t)$. This is well defined since $(\cos 2\pi(t + n), \sin 2\pi(t + n)) = (\cos 2\pi t, \sin 2\pi t)$ and is continuous, by Proposition 1.7.1. Note that it is onto since both q and p are. It is also 1–1, since $\bar{p}[t] = \bar{p}[t']$ implies $(\cos 2\pi t, \sin 2\pi t) = (\cos 2\pi t', \sin 2\pi t')$. But this only happens if $t = t' + n$ for some integer n; hence $[t] = [t']$. To see that $\bar{p}$ is in fact a homeomorphism, we can no longer use Proposition 1.7.3 since $\mathbb{R}$ is not compact. We need to see that its inverse $\bar{p}^{-1}$ is continuous. But this is equivalent to $(\bar{p}^{-1})^{-1}(U) = \bar{p}(U)$ being open when U is open; that is, $\bar{p}$ sends open sets to open sets. Since $\bar{p}(U) = pq^{-1}(U)$, this condition is equivalent to p sending saturated open sets to open sets. But p is an *open map*; that is, it sends open sets to open sets. Hence $\bar{p}$ is a homeomorphism from $\mathbb{R}/\sim$ to S^1.

We state, for future use, the principle used in the last example.

Proposition 1.7.4. *Suppose $f : X \to Y$ is a surjective continuous map. Define an equivalence relation on X by saying $u \sim v$ iff $f(u) = f(v)$; the equivalence classes are the inverse images $f^{-1}(y)$. Then the induced map $\bar{f} : X/\sim \to Y$ is a homeomorphism exactly when p sends saturated open sets $q^{-1}(U)$ to open sets. In particular, it is a homeomorphism if p is an open map.*

Since each of Y, Y' is homeomorphic to S^1, they are homeomorphic to each other. We now show this more directly. Let $q : X \to Y$, $q' : X' \to Y'$ be the identification maps. Define $i : X \to X'$ by inclusion. Since $[i(0)] = [0] = [1] = [i(1)]$, i induces a map $\bar{i} : Y \to Y'$ defined by $\bar{i}([x]) = [i(x)]$. Thus we have a commutative diagram (i.e. $\bar{i}q = q'i$) by definition. Thus $\bar{i}$ is continuous since i is. Next note that $\bar{i}$ is 1–1 since $\bar{i}q$ is except for $0, 1$, and $[0] = [1]$ in Y. $\bar{i}$ maps onto Y' since any $[y] \in Y'$ is represented by a y between 0 and 1. We leave it as

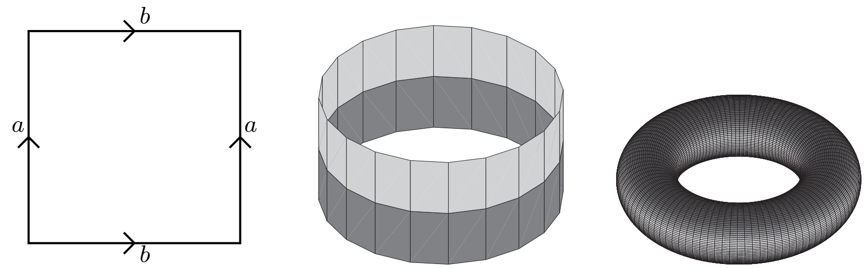


Figure 1.11. Cylinder and torus as quotient spaces of the square.

an exercise to construct an inverse for $\bar{i}$ and to prove it is continuous.

$$\begin{array}{ccccc} X = [0,1] & \xrightarrow{i} & X' = \mathbb{R} & & \\ \downarrow q & & \downarrow q' & \searrow p & \\ Y & \xrightarrow{\bar{i}} & Y' & \xrightarrow{\bar{p}} & S^1 \end{array}$$

Exercise 1.7.2. Construct an inverse for $\bar{i}$ and show that it is continuous. (Hint: Consider the discontinuous function from X' to X defined by sending x to $x - [x]$, where $[x]$ denotes the greatest integer in x; i.e. the unique integer satisfying $[x] \leq x < [x] + 1$.)

Consider the product of the circle with itself. This space is called a *torus* and will be studied in more depth in Chapter 2. From our description of the circle as a quotient space, we may give a description of $S^1 \times S^1$ as a quotient space. We take the product $\mathbb{R} \times \mathbb{R}$ and make the following identifications: $(s,t) \sim (s+m, t+n)$, where $m, n \in \mathbb{Z}$. An alternate description would be to take $[0,1] \times [0,1]$ and identify $(0,t)$ with $(1,t)$ and $(s,0)$ with $(s,1)$. A pictorial description is given in Figure 1.11. It is supposed to indicate that we identify the edges labeled a and the edges labeled b. Geometrically, we can think of gluing the edges labeled a together to form a cylinder (the b edges becoming circles) and then gluing the two circles together to get a torus.

Exercise 1.7.3. Describe basic open sets in the quotient space $[0,1] \times [0,1]/(0,t) \sim (1,t), (s,0) \sim (s,1)$ about each of the points $[(0,0)]$, $[(\frac{1}{2},0)]$, $[(0,\frac{1}{2})]$, and $[(\frac{1}{2},\frac{1}{2})]$. Describe the inverse image $q^{-1}(U)$ of each of these basic open sets.

Exercise 1.7.4. Show that the quotient space formed from a square by identifying all of the points in the bottom edge of the square to each other is homeomorphic to a triangle. (Hint: Start with the map from the rectangle to the triangle preserving y-levels and sending the bottom edge of the rectangle to the bottom vertex of the triangle. See Figure 1.12, where the bottom line that is to be collapsed to a point is thickened, as is the image point.)

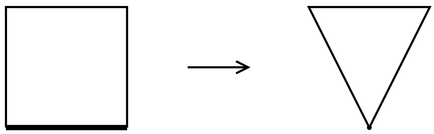


Figure 1.12. Triangle as a quotient space of the square.

We now discuss quotient spaces that are formed from two disjoint sets by identifying certain points in one of the sets with points in the other by means of a function. Suppose A and B are disjoint topological spaces. Then the union of A and B can be regarded as a topological space by saying a set is open iff it is the union of an open set in A with an open set in B. We will denote the union with this topology as $A \bigsqcup B$, and call it the disjoint union. Frequently, we will perform this construction when A and B are not disjoint. In this case we will regard them as disjoint by distinguishing points by saying the point comes from A or it comes from B. This is the reason for our terminology "disjoint union"—we want to emphasize that we are regarding the two sets as disjoint. Now suppose K is a closed subset of B and f is a homeomorphism from K onto a closed subset $f(K)$ of A. Then we may form the quotient space $A \cup_f B = (A \bigsqcup B)/x \sim f(x), x \in K \subset B$, formed from the disjoint union by identifying $x \in K$ with $f(x) \in f(K)$.

Proposition 1.7.5. *Let $g : A \cup_f B \to C$ be a map induced from continuous functions $g_A : A \to C$, $g_B : B \to C$ with $g_A f = g_B|K$. That is, if $x \in A \subset A \cup_f B$, then $g(x) = g_A(x)$, and if $x \in B, g(x) = g_B(x)$. Then g is continuous.*

Proof. To show g is continuous, we have to show that the composition $gq : A \bigsqcup B \to A \cup_f B \to C$ is continuous. But the topology on the disjoint union is just the union of the topologies on A and B. Since the restriction of this composition to A, B is just g_A, g_B, respectively, it is continuous. □

Proposition 1.7.6. *Let $A_i \cup_{f_i} B_i = A_i \bigsqcup B_i / x \sim f_i(x)$, $x \in K_i \subset B_i$ be the quotient space of $A_i \cup B_i$ coming from identifying $x \in K_i \subset B_i$ with $f(x) \in A_i$ via a homeomorphism $f_i : K_i \to f(K_i)$, $i = 1, 2$. Suppose there are homeomorphisms $F_A : A_1 \to A_2, F_B : B_1 \to B_2$ with $F_B(K_1) = K_2$ and $F_A f_1 = f_2 F_B$. Then the map $F : A_1 \cup_{f_1} B_1 \to A_2 \cup_{f_2} B_2$ given by $F(x) = F_A(x)$ if $x \in A_1$ and $F(x) = F_B(x)$ if $x \in B_1$ is a homeomorphism.*

Exercise 1.7.5. Prove Proposition 1.7.6.

Given a topological space X and closed subsets A, B with $A \cup B = X$, we can regard X as a quotient space of $A \bigsqcup B$ using $\mathrm{id} : A \cap B \subset B \to A \cap B \subset A, \mathrm{id}(x) = x$. For the inclusion maps give a map $q : A \bigsqcup B \to X$; this induces $A \cup_{\mathrm{id}} B \to X$, which is a bijection. To see that it is a homeomorphism just requires showing X has the quotient topology. A set C in X is closed iff $C \cap A$ and $C \cap B$ are closed since $C = (C \cap A) \cup (C \cap B)$, and A and B are assumed closed. The quotient topology on X from $(A \bigsqcup B, q)$ comes from requiring C to be closed iff $q^{-1}(C)$ is closed in $A \cup B$; that is, $C \cap A$ and $C \cap B$ are closed. Thus X does have the quotient topology and so $\bar{q}$ is a homeomorphism.

Suppose now we have homeomorphisms $h_A : A \to A', h_B : B \to B'$. Then Proposition 1.7.6 implies $X = A \cup_{id} B \simeq A' \cup_f B'$, where $f : h_B(A \cap B) \to h_A(A \cap B)$ is $f(x) = h_A h_B^{-1}(x)$. We will use this in situations where we can choose A', B' to be particularly nice spaces such as disks or rectangles.

As an example, consider the annulus $X = \{(x_1, x_2) : 1 \leq x_1^2 + x_2^2 \leq 2\}$. We can first break X up into $A = X \cap \{(x_1, x_2) : x_2 \leq 0\}$ and $B = X \cap \{(x_1, x_2) : x_2 \geq 0\}$. We will give a number of different descriptions of the annulus as a quotient space (see Figure 1.13). The variety of descriptions given below illustrate that a space may arise as a quotient space in many different ways. The simplest description comes from using $f : [-1, 1] \times [1, 2] \to X, f(s, t) = (t \cos \pi s, t \sin \pi s)$. The first coordinate s is used to wrap the interval around the circle (giving the angle up to a factor of π), and the second coordinate t measures the distance from the origin. This map sends $(-1, t)$ and $(1, t)$ to the same point $(-t, 0)$ and is otherwise 1–1. Thus f induces a homeomorphism between the quotient space $Q_1 = [-1, 1] \times [1, 2]/(-1, t) \sim (1, t)$ and the annulus X. We could replace the interval $[1, 2]$ by the homeomorphic interval $[-1, 1]$ and thus identify Q_1, and hence X, to the quotient space $Q_2 = [-1, 1] \times [-1, 1]/(-1, t) \sim (1, t)$. We will think of this as the standard description of the annulus as a quotient of the square $D^1 \times D^1$, where we are identifying the left-hand boundary interval to the right hand boundary interval. We depict this identification and the corresponding image on the annulus by labeling the identified edges with the letter a.

We now split the interval $[-1, 1]$ into two intervals $[-1, 0]$ and $[0, 1]$ and think of it as a quotient of the disjoint union by identifying the two copies of 0. Using

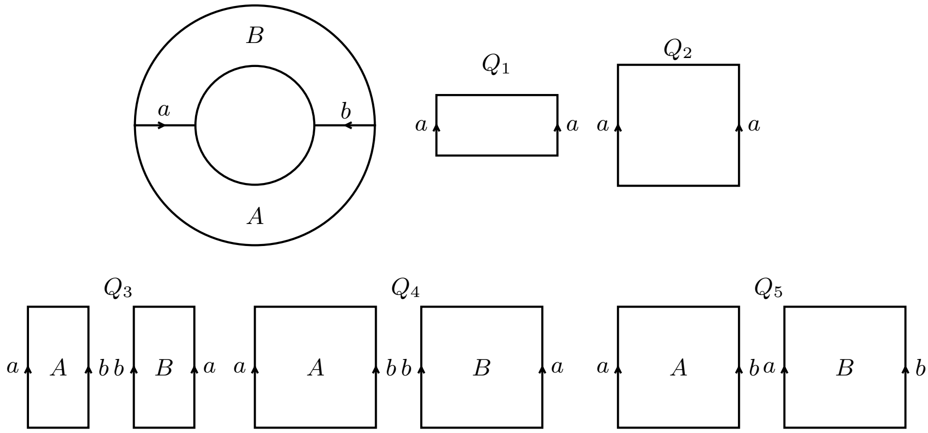


Figure 1.13. Expressing the annulus as a quotient space.

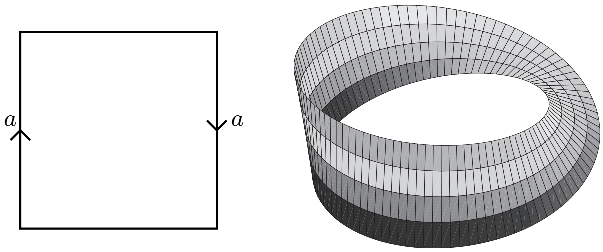


Figure 1.14. Möbius band.

this, the inclusion gives maps of the disjoint union $[-1,0] \times D^1 \bigsqcup [0,1] \times D^1$ to $D^1 \times D^1$. This map then induces a homeomorphism between the quotient space $Q_3 = [-1,0] \times D^1 \cup_f [0,1] \times D^1$ and Q_2, where $f(-1,t) = (1,t), f(0,t) = (0,t)$. This identification of the two copies of $0 \times D^1$ is labeled with b, as is its image in the annulus. Now by identifying $[-1,0]$ and $[0,1]$ with D^1 using the order preserving affine linear maps, we can re-express Q_3 as the quotient $Q_4 = D^1 \times D^1 \cup_g D^1 \times D^1$, where $g(-1,t) = g(1,t),\ g(1,t) = (-1,t)$. As another description, form a quotient $Q_5 = D^1 \times D^1 \cup_h D^1 \times D^1$, where $h(-1,t) = (-1,t), h(1,t) = (1,t)$; that is, h is the identity on the identified edges. The homeomorphism $F\colon Q_5 \to Q_4$ is induced from the map that sends the left-hand copy of $D^1 \times D^1$ to itself via the identity, and sends the right-hand copy of $D^1 \times D^1$ to itself via $(s,t) \to (-s,t)$. We use Proposition 1.7.6 to see that this induces a homeomorphism from Q_5 to Q_4.

Descriptions such as the last ones will be very useful to us in Chapter 2, where we study surfaces. We will decompose a surface into a number of pieces, each of which is homeomorphic to D^2 or $D^1 \times D^1$ and then think of the surface as a quotient space of the disjoint union of these nice pieces. The structure of the surface will be contained in the pieces involved and how they are glued together.

Here is another example. The Möbius band B is formed from a rectangular strip by identifying the ends after making a half twist as in Figure 1.14. More formally, $B = D^1 \times D^1/(-1,t) \sim (1,-t)$. We might also write this as a quotient space formed from two rectangles by splitting $D^1 = [-1,0] \cup [0,1]$ to form $Q_1' = [-1,0] \times D^1 \cup_k [0,1] \times D^1$, with $k(-1,t) = (1,-t), k(0,t) = (0,t)$. By identifying $[-1,0]$ and $[0,1]$ with D^1, we can re-express this as a quotient space $D^1 \times D^1 \cup_p D^1 \times D^1$, with $p(-1,t) = (1,t), p(1,t) = (-1,-t)$.

Exercise 1.7.6. Consider the space X formed from two copies of $R = D^1 \times D^1$ by identifying $\{-1,1\} \times D^1$ to itself via d with $d(1,y) = (1,-y)$ and $d(-1,y) = (-1,-y)$; that is, $X = R \cup_d R$. Construct a homeomorphism between X and the annulus.

Exercise 1.7.7. Suppose $X = A \cup_g B, Y = A \cup_{g'} B$, where $g, g'\colon K \subset B \to g(K),\ g'(K) \subset A$ are homeomorphisms. Suppose $(g')^{-1} g\colon K \to K = h|K$, where

$h: B \to B$ is a homeomorphism. Show that the identity on A and h on B piece together to give a homeomorphism from X to Y.

Exercise 1.7.8. Show that the Möbius band can also be described as a quotient space $D^1 \times D^1 \cup_f D^1 \times D^1$, with $f(-1,t) = (-1,t),\ f(1,t) = (1,-t)$.

Exercise 1.7.9. Identify all points in the lower half of the circle to each other. Show that the resulting quotient space $S^1/\sim$ is homeomorphic to S^1. (Hint: Find a continuous map from the circle to the circle which sends the lower half of the circle to a point and is 1–1 elsewhere.)

Exercise 1.7.10. Put an equivalence relation on the unit disk by making all points on the boundary circle equivalent to each other. Show that the resulting quotient space $D^2/\sim$ is homeomorphic to S^2. (Hint: Send diameters to great circles with the origin going to the south pole and the boundary circle going to the north pole.)

1.8 The Jordan curve theorem and the Schönflies theorem

In this section we outline proofs of the Jordan curve theorem and the Schönflies theorem for polygonal curves. The section is essentially in the form of a project to fill in the details of the outline to prove these results. The proofs of these theorems in the polygonal case will provide us with many opportunities to apply the concepts from the chapter in justifying geometric steps in the argument.

We start by carefully stating these theorems in their general versions.

Definition 1.8.1. A *simple closed curve* in the plane is a function $f: S^1 \to \mathbb{R}^2$ which is a homeomorphism onto its image. The image $C = f(S^1) \subset \mathbb{R}^2$ is sometimes also called a simple closed curve when the parametrization is not important. Alternatively, a simple closed curve in the plane can be given as a map $f: [a,b] \to \mathbb{R}^2$, with $f(x) = f(y)$ for $x \neq y$ iff $\{x,y\} = \{a,b\}$ so that when the quotient space $[a,b]/a \sim b$ is identified with S^1, the induced map $\bar{f}$ is a homeomorphism onto its image.

Theorem 1.8.1 (Jordan curve theorem). *Let $C = f(S^1)$ be a simple closed curve in the plane. Then $\mathbb{R}^2 \backslash C$ is the disjoint union of two open sets A, B so that each is path connected. Moreover, one of these sets A is bounded and the other B is unbounded. Also, C is the boundary of each of these sets.*

Theorem 1.8.2 (Schönflies theorem). *Let $C = f(S^1)$ be a simple closed curve in the plane and $\mathbb{R}^2 \backslash C = A \cup B$ as given by the Jordan curve theorem, with A bounded. Then there is a homeomorphism of the plane to itself which sends the open unit disk to A and the closed unit disk to $A \cup C$.*

The Jordan curve theorem was first stated as a theorem by Camille Jordan (1838–1932) in his Cours d'Analyse in the late nineteenth century. His original proof was very complicated and was found to have gaps, which required considerable effort to fill in. Modern proofs use homology theory, where the separation

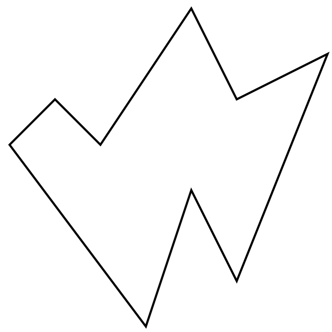


Figure 1.15. A polygonal simple closed curve.

part of the theorem is expressed by saying that $H_0(\mathbb{R}^2 \backslash C)$ is the free abelian group on two generators. H_0 measures the path components of a space; the two generators correspond to A and B. The difficulty, in general, has to do with the very wild nature a simple closed curve may have. If the curve is restricted somewhat, then the theorem becomes much easier. The Schönflies theorem was proved in 1908.

In this section we will only look at the case of a polygonal simple closed curve, which is the image of a map $p : [0, n] \to \mathbb{R}^n$ where, on each subinterval $[k, k+1]$, the map is an affine linear map onto a line segment L_k determined by the points (called *vertices*) $p(k) = \boldsymbol{v}_k$ and $p(k+1) = \boldsymbol{v}_{k+1}$. We assume that $p(0) = p(n)$ but $p(a) \neq p(b)$ if $a \neq b$ unless $\{a, b\} = \{0, n\}$. Note that the quotient space $[0, n]/0 \sim n$ is homeomorphic to S^1 and p determines a map $\bar{p} : S^1 \to \mathbb{R}^2$ as in the original definition of a simple closed curve. Figure 1.15 shows an example of a polygonal simple closed curve and the bounded region which it bounds. We will assume that adjacent segments in C do not lie on the same line.

We give an outline of the proofs of these theorems, giving the major steps with illustrations when appropriate.

Step 1. Show that both theorems are unaffected by composing $f : S^1 \to \mathbb{R}^2$ with an affine linear homeomorphism (sends lines to lines) $H \colon R^2 \to \mathbb{R}^2$. Use this to show that we can reduce the theorems to the case that no segment in the polygonal curve is horizontal, which we will assume from now on.

Step 2. There are two types of points in C, *edge points* in $p(k, k+1)$, and vertices, which are the points $p(k)$. The vertices can be divided into two types, *regular vertices* and *special vertices.* The special vertices are those which are a local maxima or local minima for the y-coordinate on C, and the regular vertices are the others. Figure 1.16 shows neighborhoods of each type of point, and smaller *regular neighborhoods* within these which consist of nearby parallel line segments. Show that such neighborhoods exist for each type of point in C.

Step 3. Consider a horizontal line at height y_0. Suppose it intersects C in k points (not counting any special vertices). Show that there is a number ϵ so that horizontal curves at height between $y_0 - \epsilon$ and $y_0 + \epsilon$ intersect C in l points besides

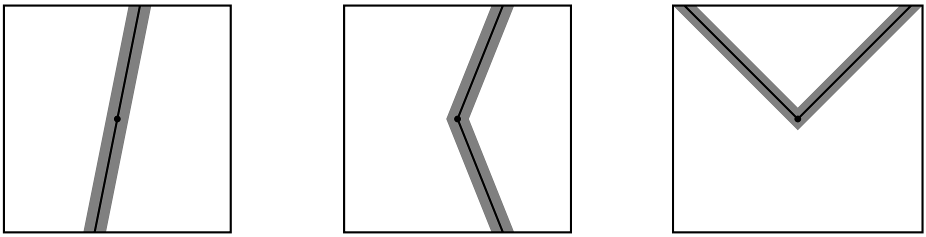


Figure 1.16. Nice neighborhoods.

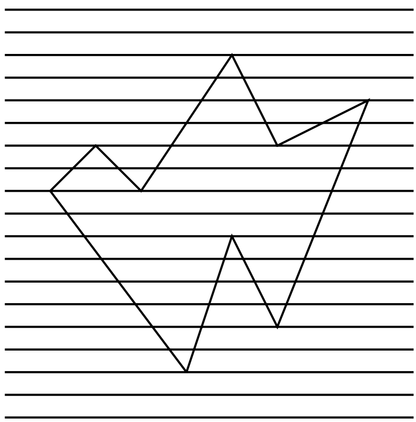


Figure 1.17. How lines intersect C.

special vertices, where $k \equiv l \mod 2$. It is necessary to consider line segments in C which are missed at height y_0 as well as special vertices at height y_0. Use this fact to show that the function that sends y to the number of points of C mod 2 that are not special vertices at height y is a continuous function from $\mathbb{R}$ to $\{0, 1\}$. Show that the horizontal line at height y intersects C in an even number of points which are not special vertices. See Figure 1.17 for an illustration.

Step 4. For each $(x, y) \notin C$, define $I(x, y)$ to be 0 if there are an even number of points of C (not counting special vertices) at height y to the left of (x, y), and equal to 1 when there are an odd number of such points. Show that I is continuous, and that the sets $A = I^{-1}(\{1\})$ and $B = I^{-1}(\{0\})$ are disjoint open sets with $\mathbb{R}^2 \backslash C = A \cup B$.

Step 5. Figure 1.18 shows a regular neighborhood of the curve C consisting of parallel polygonal curves near C. Show that C has such a regular neighborhood $N(C)$. Show that that $N(C)$ is homeomorphic to the annulus $S^1 \times [\frac{1}{2}, \frac{3}{2}] \subset \mathbb{R}^2$ enclosed between the circles of radii $\frac{1}{2}$ and $\frac{3}{2}$ with C corresponding to $S^1 \times \{1\}$.

Figure 1.18. A regular neighborhood.

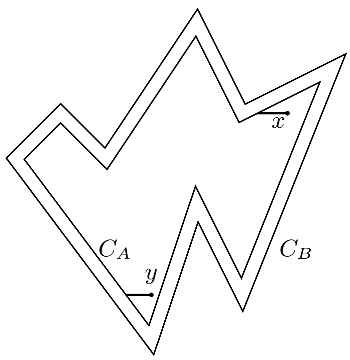


Figure 1.19. Using C_A to connect $x, y \in A$.

In particular, show that $N(C)\backslash C$ consists of two sets which are path connected but that $N(C)\backslash C$ is not path connected.

Step 6. Show that $\operatorname{Bd} N(C) = C_A \cup C_B$, where $C_A \subset A$ and $C_B \subset B$ are parallel polygonal curves to C. Use the curves C_A and C_B to show that each of A and B are path connected. See Figure 1.19 for a motivating example of such a path connecting two points $x, y \in A$ that uses C_A. Use $N(C)$ to show that $\bar{A} = A \cup C,\ \bar{B} = B \cup C$.

Step 7. Use the fact that C is compact to show that $\bar{A}$ is compact and $\bar{B}$ is not compact.

These steps then complete the proof of the polygonal version of the Jordan curve theorem. We now outline an approach to proving the polygonal Schönflies theorem. Our starting point is the setup from the polygonal Jordan curve theorem above.

For our proof of the Schönflies theorem, we first need to modify C slightly before we give our argument, but in a way that does not change the validity of the Schönflies theorem.

Step 1. Show that there is a homeomorphism $h : \mathbb{R}^2 \to \mathbb{R}^2$ which is the identity outside a regular neighborhood of C so that the special vertices all occur at

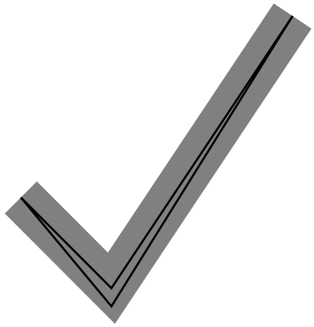


Figure 1.20. Moving a vertex.

different y-values. The idea is depicted in Figure 1.20, where we push the vertex vertically, keep the boundary of the part of the regular neighborhood of two adjacent edges that come together at the vertex fixed, and extend this to a PL map of the regular neighborhood. This allows an extension via the identity outside of the regular neighborhood to get a homeomorphism of $\mathbb{R}^2$ which displaces the vertex slightly. Show that this modification does not affect the validity of the Schönflies theorem. Figure 1.20 shows the original piece of the curve and the displaced piece in a regular neighborhood.

Step 2. Use compactness to show that there is the minimal value m assumed by C and the maximal value M assumed by C and the A lies between the $y = m$ and $y = M$. Moreover, show that there is a minimal special vertex at height m and a maximal special vertex at height M. Show that there are an even number of special vertices, half of which are local minima and half local maxima.

Step 3. Show that if there are just two special vertices, then A is homeomorphic to a triangle, and that the homeomorphism can be chosen to fix pointwise a small subtriangle at the bottom of A and is the identity outside a large rectangle containing A. Use induction on the number V of vertices, with starting point $V = 3$. Your homeomorphism should be expressible as a composition of homeomorphisms which are the identity outside a small neighborhood of a triangle which is being worked on. At each step a triangle is added or removed from A where two of its sides are on C and the third side is not. The interior of the triangle will lie entirely in A or entirely in B. The argument should show the existence of such triangles. The requirement that there are no horizontal lines occurring in any intermediate steps may require working on two adjacent triangles in a single reduction step. As a hint, we illustrate an example of a complete reduction of such a region to a triangle in Figure 1.21. The dotted lines show intermediate triangles being used and the numbering shows new edges in C as it is homeomorphed to the bottom triangle.

Step 4. The general argument is by induction on the ordered pairs (V, S), where V denotes the total number of vertices and S is the number of special vertices. The ordering is lexicographic ordering: $(V_1, S_1) < (V_2, S_2)$ iff (1) $V_1 < V_2$, or (2) $V_1 = V_2$ and $S_1 < S_2$. The starting point for the induction is $(3, 2)$

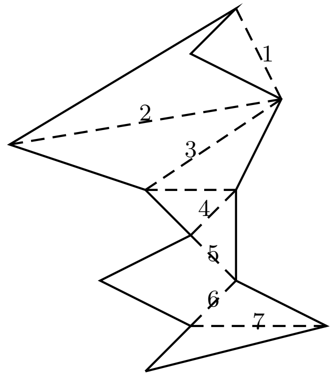


Figure 1.21. Homeomorphing A to a triangle.

and the way it works is to either keep the total number of vertices the same and reduce the number of special vertices by 2, or reduce the number of total vertices. Consider the lowest special vertex (minimal y-value) v_{m} which is a local maximum. There are a number of cases to consider. A useful concept to look at is the position of the first two vertices on the segments moving downward from v_{m} and look for ways to homeomorph $\mathbb{R}^2$ to simplify the image of C so that it has one fewer regular vertex. All of your homeomorphisms should fix a small triangle at the bottom of A and the region outside a large rectangle containing A. In fact, at each step the homeomorphism should just fix everything outside a region near a triangle on which you are working. Some steps may require composing a couple of these as well as introducing new vertices to avoid horizontal edges. Figure 1.22 illustrates how C is deformed to a C' with a single local maximum and local minimum. Intermediate triangles being used are indicated with dotted lines.

1.9 Supplementary exercises

Definition 1.9.1. A collection $\mathcal{B} = \{B_i : i \in I\}$ of subsets of X is called a *basis* and its elements are called *basis elements* if the following properties are satisfied:

(a) every $x \in X$ is contained in some B_i;

(b) if $x \in B_i \cap B_j, B_i, B_j \in \mathcal{B}$, then there is a basis element B_k with $x \in B_k \subset B_i \cap B_j$.

The topology $\mathcal{T}_{\mathcal{B}}$ determined by the basis $\mathcal{B}$ is defined as follows: a set $U \subset X$ is open if, for every $x \in U$, there is a basis element B_i with $x \in B_i \subset U$.

The first six problems concern the concept of a basis and the topology which it determines.

Exercise 1.9.1. Verify that $\mathcal{T}_{\mathcal{B}}$ satisfies the three properties required of a topology.

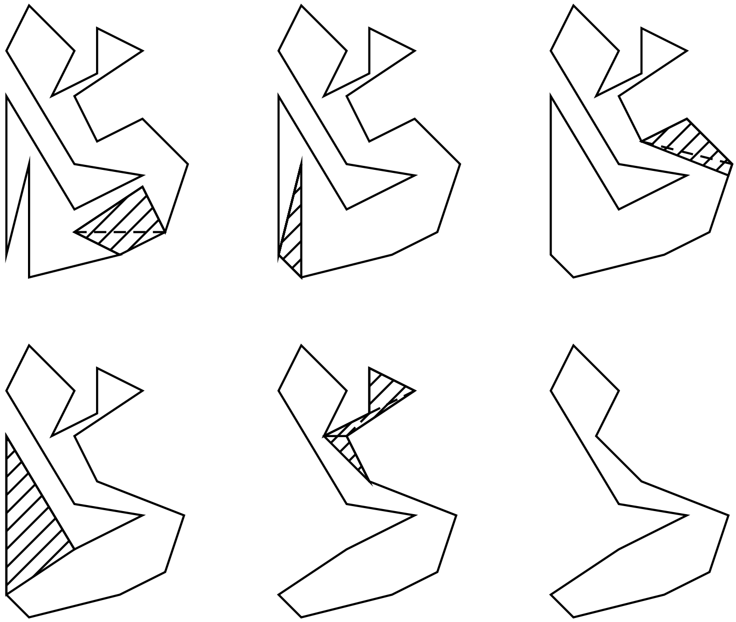


Figure 1.22. Removing excess special vertices.

Exercise 1.9.2. Show that the set of balls $\{B(x,r),\ x \in X,\ r > 0\}$ is a basis for the topology of a metric space X; that is, show that it is a basis and the topology it determines is the metric topology.

Exercise 1.9.3. Show that any open set in $\mathcal{T}_{\mathcal{B}}$ is a union of basis elements. (Hint: For $x \in U$, choose a basis element $B_{i(x)}$ with $x \in B_{i(x)} \subset U$.)

Exercise 1.9.4. Suppose the topology for Y is determined by a basis. Show that $f : X \to Y$ is continuous iff, for each basis element B of Y, $f^{-1}(B)$ is open.

Exercise 1.9.5. Show that the open intervals give a basis for the topology of $\mathbb{R}$.

Exercise 1.9.6. Combine the last two exercises to show that a map to $\mathbb{R}$ is continuous iff, for each open interval $I \subset \mathbb{R}$, we have $f^{-1}(I)$ is open. Formulate and prove an analogous statement for maps to a metric space in terms of balls.

Exercise 1.9.7. Let X be a metric space with metric d and let A be a subset of X. Show that the metric topology on A given by d is the same as the subspace topology.

Exercise 1.9.8. Suppose $A \subset B \subset X$, and B has the subspace topology. Show that, if A is open in B and B is open in X, then A is open in X.

Exercise 1.9.9. If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$, let $d(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1|$. Which of the three properties of a metric does d satisfy?

Exercise 1.9.10. On $\mathbb{R}^2 \backslash \{\mathbf{0}\}$, write each point in polar coordinates as (r, θ), where $0 \leq \theta < 2\pi$. Define $d((r_1, \theta_1), (r_2, \theta_2)) = |r_1 - r_2| + |\theta_1 - \theta_2|$. Show that this gives a metric on $\mathbb{R}^2 \backslash \{\mathbf{0}\}$ but the topology formed is not the usual topology.

Definition 1.9.2. A point x is called a *limit point* of a set A if every open set U containing x intersects $A \backslash \{x\}$ in a nonempty set; that is, $U \cap (A - \{x\}) \neq \emptyset$ for U open, $x \in U$. Denote the limit points of A by A'.

Exercise 1.9.11. Show that a set A is closed iff it contains all of its limit points.

Exercise 1.9.12. Find the limit points of the following subsets of $\mathbb{R}$: $(a)(0,1)$; $(b)\mathbb{Q}$, the rationals; (c) $\{1/n \colon n \in \mathbb{N}\}$.

Exercise 1.9.13. Let X be a metric space and $A \subset X$. Show that x is a limit point of A iff every ball $B(x, r)$ contains infinitely many points of A.

Exercise 1.9.14. Show that, in Hausdorff space X with subset A, x is a limit point of A iff every open set containing x contains infinitely many points of A.

Exercise 1.9.15.

(a) Show that if C is a closed set containing A, then $\bar{A} \subset C$.
(b) Show that if V is an open set contained in A, then $V \subset \operatorname{int} A$.

Exercise 1.9.16. Show that A is closed iff $A = \bar{A}$, and A is open iff $A = \operatorname{int} A$.

Exercise 1.9.17. Show that $\bar{A} = A \cup A'$.

Exercise 1.9.18. Show that $\overline{B(z,r)} = \{x \colon d(z,x) \leq r\}$ and $\operatorname{Bd} B(z,r) = \{x \colon d(z,x) = r\}$.

Definition 1.9.3. A topological space X is called *limit point compact* if every infinite set has a limit point.

Exercise 1.9.19. Show that a compact space is limit point compact.

Exercise 1.9.20. Show that if X is a metric space, then X is limit point compact iff it is sequentially compact iff it is compact.

Exercise 1.9.21. Show that if $f \colon X \to Y$ is continuous at x and x_n is a sequence converging to x, then $f(x_n)$ converges to $f(x)$.

Exercise 1.9.22. Show that a map $f \colon (X, d) \to (Y, d')$ between metric spaces is continuous at x iff for every sequence x_n which converges to x, the sequence $f(x_n)$ converges to $f(x)$.

Exercise 1.9.23. Show that, in a Hausdorff space, the limit of a sequence is well defined; that is, if x_n converges to x and to y, then $x = y$.

Exercise 1.9.24. Show that, in a Hausdorff space, if x_n converges to x, then x is the only limit point of the set of all of values $\{x_n \colon n \in N\}$. Is the converse true? Give a proof or counterexample.

Exercise 1.9.25. Show that a finite set in a Hausdorff space is closed.

Definition 1.9.4. A Hausdorff space is called *regular* if, given $x \in X$ and a closed set C with $x \notin C$, then there are disjoint open sets U and V with $x \in U$ and $C \subset V$. A Hausdorff space is called *normal* if, whenever C, D are disjoint closed subsets, then there are disjoint open sets U, V with $C \subset U, D \subset V$.

Exercise 1.9.26. Show that a compact Hausdorff space is regular. (Hint: Use the fact that a closed subset of a compact space is compact. Get an open cover of the closed, hence compact, set C where there is an open set V_y for each point $y \in C$ with $y \in V_y$ and an open set U_y so that $x \in U_y$ with $U_y \cap V_y = \emptyset$.)

Exercise 1.9.27. Suppose $x \in U \subset X$, where X is regular and U is open. Show that there is an open set V with $x \in V \subset \bar{V} \subset U$. (Hint: Consider the point x and the disjoint closed set $X \backslash U$.)

Exercise 1.9.28. Show that a compact Hausdorff space is normal. (Hint: Apply the conclusion of Exercise 1.9.26 to pairs x, D, where $x \in C$ and use the compactness of C.)

Exercise 1.9.29. Suppose $C \subset U \subset X$, where X is normal, C is closed and U is open. Then show that there is an open set V with $C \subset V \subset \bar{V} \subset U$. (Hint: C and $X \backslash U = D$ are disjoint closed sets.)

Exercise 1.9.30. Show that a metric space is normal. (Hint: If C, D are disjoint closed subsets, then cover C by balls $B(c, r(c))$ disjoint from D and cover D by balls $B(d, r(d))$ disjoint from C. Then show that $U = \cup_{c \in C} B(c, r(c)/2)$ and $V = \cup_{d \in D} B(d, r(d)/2)$ are disjoint open sets containing C and D.)

Definition 1.9.5. Let C be a subset of a metric space X. For each $x \in X$, define the *distance from x to C* by $d(x, C) = \inf\{d(x, y) \colon y \in C\}$.

Exercise 1.9.31. Show that $\{x \colon d(x, C) = 0\} = \bar{C}$. Show that if C is closed and $x \notin C$, then $d(x, C) > 0$.

Exercise 1.9.32. Show that $C_\epsilon = \{y \colon d(y, C) < \epsilon\}$, where $\epsilon > 0$, is an open set containing C.

Exercise 1.9.33. Show that the function $f : X \to \mathbb{R}$ given by $f(x) = d(x, C)$ is continuous. (Hint: Look at the inverse image of an interval.)

Exercise 1.9.34. Suppose C, D are disjoint closed sets in a metric space X. Use the notion of the distance from a point to a set to define $d(C, D)$ in terms of $d(x, D)$ for $x \in C$. Give an example to show that this distance could be 0. Show that if X is compact, then the distance must be positive.

The next two exercises give versions of Urysohn's lemma and the Tietze extension theorem for metric spaces.

Exercise 1.9.35. Suppose the (X, d) is a metric space and A, B are disjoint closed sets. Show that the function

$$f(x) = \frac{d(x, A) - d(x, B)}{d(x, A) + d(x, B)}$$

is a continuous real-valued function $f : X \to [-1, 1]$ with $f^{-1}\{-1\} = A$, $f^{-1}\{1\} = B$. The existence of a function from X to $[-1, 1]$ which is -1 on A and 1 on B is called Urysohn's Lemma and holds in the more general situation of a normal space.

Exercise 1.9.36. Suppose X is a metric space and $C \subset X$ is a closed subset, with a continuous function $f : C \to \mathbb{R}$. This exercise leads you through a proof that there is a continuous extension $F : X \to \mathbb{R}$. This result is called the Tietze extension theorem and holds in the more general situation of a normal space.

(a) Reduce to the case where f is bounded by considering the composition of f with a homeomorphism from $\mathbb{R}$ to $(-1, 1)$.

(b) Because of (a), we assume from now on that $f : X \to [-M, M]$. We inductively define continuous maps from X to $[-M, M]$ which give better and better approximations to f on C. Let $A_1 = f^{-1}([-M, -M/3])$, and $B_1 = f^{-1}([M/3, M])$. Show that A_1, B_1 are disjoint closed subsets of X. Apply Exercise 1.9.35 to show that there is a continuous map $g_1 : X \to [-M/3, M/3]$ with $A_1 \subset g_1^{-1}\{-M/3\}$, and $B_1 \subset g_1^{-1}\{M/3\}$. Show that $|f(x) - g_1(x)| \le 2M/3$ on C.

(c) Repeat the construction in (b) applied to $h_1 = f - g_1$ defined on C to construct $g_2 : X \to [-2M/9, 2M/9]$ so that

$$h_1^{-1}\left(\left[\frac{-2M}{3}, \frac{-2M}{9}\right]\right) \subset g_2^{-1}\left(\left\{\frac{-2M}{9}\right\}\right),$$
$$h_1^{-1}\left(\left[\frac{2M}{9}, \frac{2M}{3}\right]\right) \subset g_2^{-1}\left(\left\{\frac{2M}{9}\right\}\right)$$

and $|f(x) - g_1(x) - g_2(x)| \le 4M/9$.

(d) Use induction to construct a sequence of maps $g_n : X \to [-2^{n-1}M/3^n, 2^{n-1}M/3^n]$ so that if $h_n(x) = f(x) - g_1(x) - \cdots - g_n(x)$, then $|h_n(x)| \le 2^n M/3^n$ on C and $|g_n(x)| \le 2^{n-1}M/3^n$.

(e) Define $g(x) = \sum_{n=1}^{\infty} g_i(x)$. Show that $g(x)$ converges uniformly to a continuous function $g : X \to [-M, M]$ which is an extension of f; that is, $g(x) = f(x)$ for $x \in C$.

Exercise 1.9.37. Urysohn's lemma states that for a normal space X with disjoint closed sets A, B, there is a continuous function $f : X \to [-1, 1]$ with $A \subset f^{-1}\{-1\}$ and $B \subset f^{-1}\{1\}$. The Tietze extension theorem states that for a normal space X with a closed subset C and a continuous function $g : C \to \mathbb{R}$, there is continuous function $g : X \to \mathbb{R}$ with $f(x) = g(x)$ for $x \in C$. Show that Urysohn's lemma is equivalent to the Tietze extension theorem. (Hint: Use the argument in Exercise 1.9.36 to show that Urysohn's lemma implies the Tietze extension theorem.)

Definition 1.9.6. A set is *countable* if it can be put in 1–1 correspondence with the natural numbers $\mathbb{N}$ or is finite. For example, the rationals $\mathbb{Q}$ and n-tuples of rationals $\mathbb{Q}^n = \{(r_1, r_2, \ldots, r_n) : r_i \in \mathbb{Q}\}$ are countable. A space X is called *first* *countable* if, for each $x \in X$, there is a countable basis of open sets containing x; that is, there is a collection $\{B_n : n \in \mathbb{N}\}$ of open sets containing x so that, if U is an open set containing x, then there is a set $B_k \subset U$. This is called a *neighborhood basis.* A space X is called *second countable* if there is a countable basis for the topology of X. A space X is called *countably compact* if every countable open cover has a finite subcover. A metric space is called *separable* if there is a countable set $\{x_n : n \in \mathbb{N}\}$ so that every open set contains at least one x_n (we say $\{x_n\}$ is *dense* in X and $\{x_n\}$ is a *countable dense subset*).

Exercise 1.9.38. Show that $\mathbb{Q}$ is dense in $\mathbb{R}$, and hence $\mathbb{R}$ is separable.

Exercise 1.9.39. Show that a metric space is first countable.

Exercise 1.9.40. Show that, if X is first countable and x is a limit point of C, then there is a sequence $x_i \in C$ which converges to x.

Exercise 1.9.41. Show that if X is first countable, Hausdorff, and compact, then X is sequentially compact. (Hint: Adapt the proof given in Section 1.5 for metric spaces.)

Exercise 1.9.42. Show that a metric space is separable iff it is second countable. (Hint: If it is separable, use balls about the countable dense subset to get a countable basis. A countable number of countable sets is still countable. If it is second countable, select a countable set by choosing one point from each basis element.)

Exercise 1.9.43. Show that compactness implies countable compactness, and that the converse holds in a second countable space. (Hint: In a second countable space show that for each open covering $\{U_i\}_{i \in I}$ there is a covering by basis elements so that each basis element is contained in an element of the given covering $\{U_i\}$.)

Consider the space $X = A \cup B$, where $A = \{(x, \sin 1/x) \colon 0 < x \leq 1\}$ and $B = \{(0, y) \colon -1 \leq y \leq 1\}$. The space X is called the *topologist's sine curve.* The next three problems show that X is connected but not path connected. See Figure 1.9.

Exercise 1.9.44. Show that $A \cup B = \bar{A}$.

Exercise 1.9.45. Show that the closure of a connected set is connected, thus implying that $A \cup B$ is connected.

Exercise 1.9.46. We show here that $A \cup B$ is not path connected. Suppose $A \cup B$ were path connected. Let $f : [0, 1] \to A \cup B$ be a path with $f(0) = (0, 0)$ and $f(1) = (1, \sin 1)$. Consider the set $S = \{t \colon f([0, t]) \in B\}$. Let $p : \mathbb{R}^2 \to \mathbb{R}$ be $p(x, y) = x$.

(a) Show that S is nonempty, is bounded from above, and has a least upper bound $u < 1$.

(b) Show that $u \in S$.

(c) Show that there is a neighborhood $N_{f(u)}$ of $f(u)$ consisting of an infinite number of separated arcs, and a neighborhood N_u of u with $f(N_u) \subset N_{f(u)}$.

(d) Show that there is $u_1 > u,\ u_1 \in N_u$ with $pf(u_1) > 0$.

(e) Show that there are disjoint open sets $U, V \subset N_{f(u)}$ with $f(u) \in U,\ f(u_1) \in V$, and $U \cup V = N_{f(u)}$; that is, $N_{f(u)}$ is separated by U, V.

(f) By looking at $f|[u, u_1]$, arrive at a contradiction.

Definition 1.9.7. A topological space X is called *locally path connected at* x if for each open set V containing x, there is a path connected open set U with $x \in U \subset V$. It is called *locally path connected* if it is locally path connected at each $x \in X$. It is called *locally connected at* x if, for each open set V containing x, there is a connected open set U with $x \in U \subset V$. It is called *locally connected* if it is locally connected for each $x \in X$.

Exercise 1.9.47. Show that a locally path connected space is locally connected.

Exercise 1.9.48. Show that the topologist's sine curve is not locally connected.

Exercise 1.9.49. Show that an open set in $\mathbb{R}^n$ is locally path connected.

Exercise 1.9.50. Show that the path components of a locally path connected space are open sets.

Exercise 1.9.51. Show that if X is locally path connected and connected, then X is path connected. (Hint: Modify the proof that a connected open set in $\mathbb{R}^n$ is path connected.)

Exercise 1.9.52. Give an example of a path connected space which is not locally path connected.

Exercise 1.9.53. Define an equivalence relation on X by $x \sim y$ if there is a connected set containing both x and y. The equivalence classes are called the *components* of X.

(a) Verify that this is an equivalence relation.

(b) Show that each component is connected, that any two components are equal or disjoint, and that the union of the components is X.

(c) Show that any connected subset of X intersects at most one component and is a subset of that component.

(d) Show that each path component is contained in a component, and that a component is a disjoint union of path components.

(e) Show that a component is a closed set.

Exercise 1.9.54.

(a) Show that a space is locally connected iff, for each open set U, each component of U is open in X.

(b) Show that a space is locally path connected iff, for each open set U, each path component of U is open in X.

Exercise 1.9.55. A collection $\mathcal{D}$ of subsets of X is said to satisfy the *finite intersection property* (F.I.P.) if for every finite subcollection $\{D_1, \ldots, D_k\}$ of $\mathcal{D}$, the intersection $D_1 \cap \cdots \cap D_k \neq \emptyset$. Show that X is compact iff for every collection $\mathcal{D}$ of closed sets satisfying the F.I.P., the intersection of all of the elements of $\mathcal{D}$ is nonempty. (Hint: If not, consider the covering of X by the complements $\{X \setminus D_i\}$.)

Exercise 1.9.56. Let (X, d) be a compact metric space, and $f : X \to X$ continuous. $x \in X$ is called a *fixed point* of f if $f(x) = x$. f is called a *contraction* if there is a number $a < 1$ such that $d(f(x), f(y)) \leq ad(x, y)$ for all $x, y \in X$. Show that a contraction has a unique fixed point. This result is known as the *contraction mapping principle* and plays a key role in analysis. (Hint: Consider $\cap f^n(X)$. where f^n denotes the n-fold composition of f with itself and use the finite intersection property from the previous exercise.)

Definition 1.9.8. X is called *locally compact at x* if there is an open set U and a compact set C with $x \in U \subset C$. It is called *locally compact* if it is locally compact at each $x \in X$.

Exercise 1.9.57. Show that a compact space is locally compact.

Exercise 1.9.58. Show that $\mathbb{R}^n$ is locally compact.

Exercise 1.9.59. Show that if X is Hausdorff and locally compact at x, then there is an open set U containing x so that $\bar{U}$ is compact.

Exercise 1.9.60. Suppose X is a locally compact Hausdorff space. Show that if $x \in U \subset X$, U open, then there exists an open set V containing x such that $\bar{V}$ is compact and $\bar{V} \subset U$. (Hint: Use the preceding exercise and the argument of Exercise 1.9.27.)

The following exercise leads through the construction of the one-point compactification of a locally compact Hausdorff space.

Exercise 1.9.61. Suppose that X is a locally compact Hausdoff space. Form a new space X^+, called the *one-point compactification* of X as $X^+ = X \cup \{p\}$, the disjoint union of X and an added point p. A set $U \subset X^+$ is called open if (1) U is an open set in X, or (2) $p \in U$ and $X^+ \setminus U$ is a compact set in X.

(a) Show that this definition of open set satisfies the three properties of a topology.

(b) Show that the subspace topology on $X \subset X^+$ is the same as its usual topology.

(c) Show that X^+ is a compact Hausdorff space.

Exercise 1.9.62. Show that if X is a locally compact Hausdorff space and Y is a compact Hausdorff space with $Y \setminus \{y_0\} \simeq X$, then $Y \simeq X^+$.

Exercise 1.9.63.

(a) Show that the one-point compactification of $\mathbb{R}$ is homeomorphic to S^1.

(b) Show that the one-point compactification of $\mathbb{R}^2$ is homeomorphic to S^2. (Hint: Use projection from the point $p = (0, 0, 1)$ to get a homeomorphism from $S^2 \setminus \{p\}$ to $\mathbb{R}^2$.)

(c) Show that the one-point compactification of $\mathbb{R}^n$ is homeomorphic to S^n.

Exercise 1.9.64. Show that, if X is compact, the one-point compactification of X is $X^+ = X \bigsqcup \{p\}$, where the set $\{p\}$ is an open set in the disjoint union.

Exercise 1.9.65. Consider the space $\mathbb{R}^\infty$, which is the product of a countably infinite number of copies of $\mathbb{R}$. This is given the product topology with basis the sets which are product of a finite number of intervals with a product of copies of the reals:

$$(a_1, b_1) \times \cdots \times (a_n, b_n) \times \mathbb{R} \times \cdots \times \mathbb{R} \times \cdots .$$

Show that this space is not locally compact. (Hint: Look at basic open sets and show that they do not have compact closure.)

Exercise 1.9.66.

(a) Show that the following two subsets of $\mathbb{R}$ are not homeomorphic: $A = \{1/n : n \in \mathbb{N}\} \cup \{0\},\ B = \mathbb{N}$.

(b) Show that B^+ is homeomorphic to A.

Exercise 1.9.67. Show that ∞ is not homeomorphic to O (Hint: Consider where the crossing point of ∞ could go under a homeomorphism.)

The next four exercises concern the homeomorphism type of the letters of the alphabet. In each case, assume the letter is written as given below in the sans serif style, with no adornments:

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Exercise 1.9.68. Show that the letter X is not homeomorphic to the letter Y, but that the letter Y is homeomorphic to the letter T.

Exercise 1.9.69. Construct a homeomorphism between the letter D and the letter O.

Exercise 1.9.70. Prove that the letter A is not homeomorphic to the letter B.

Exercise 1.9.71. Group the letters of the alphabet into equivalence classes so that equivalent letters are homeomorphic and nonequivalent letters are not homeomorphic.

Exercise 1.9.72. Prove that if an open set $U \subset \mathbb{R}^2$ is path connected, then any two points in U can actually be connected by a polygonal path in U.

Exercise 1.9.73. For each of the following subsets of $\mathbb{R}^2$ indicate which of the following properties it possesses, namely, (i) compact; (ii) connected; (iii) path connected; (iv) open; (v) closed:

(a) $A = \{x_1, x_2) : x_1 \geq 0, 4 < x_2 \leq 8\}$;
(b) $B = \{(x_1, x_2) : x_1^2 + x_2^2 = 25\}$;
(c) $C = A \cap B$;
(d) $D = \{(x_1, x_2) : x_1^2 + x_2^2 < 1\}$;
(e) $E = \bar{D}$.

Exercise 1.9.74. Show that the torus is not homeomorphic to an open set in $\mathbb{R}^2$. (Hint: Use the properties of compactness and connectedness.)

Exercise 1.9.75. Show that if $A \subset S^1$ with $A \neq S^1$, then A is not homeomorphic to S^1.

Exercise 1.9.76. Put an equivalence relation $\sim$ on $\mathbb{R}^2$ by saying that two points are equivalent if they both lie on the circle of radius r about the origin. Show that $\mathbb{R}^2/\sim$, with the quotient topology, is homeomorphic to $[0, \infty)$.

Exercise 1.9.77. Identify all points in the lower hemisphere of the sphere S^2. Show that the resulting quotient space $S^2/\sim$ is homeomorphic to S^2.

Exercise 1.9.78. Identify points on the boundary circles of an annulus A between the circles of radius 1 and radius 2 that lie on the same ray from the origin. Show that the resulting quotient space $A/\sim$ is homeomorphic to the torus T^2.

Exercise 1.9.79. Identify the points on the outer circle of an annulus A to one-point and the points on the inner circle to a (different) point. Show that $A/\sim$ is homeomorphic to S^2. Describe what space you would get if you identified the points on both circles to a single point.

Exercise 1.9.80. Consider the quotient space X formed from two copies of $D^1 \times D^1$ using $f : \{-1, 1\} \times D^1 \to \{-1, 1\} \times D^1$ by $f(-1, x) = (1, -x)$, $f(1, y) = (-1, y)$, $X = (D^1 \times D^1) \cup_f (D^1 \times D^1)$. Decide whether X is homeomorphic to the annulus or the Möbius band, and prove your assertion.

Exercise 1.9.81. Show that the upper hemisphere of S^2 is homeomorphic to D^2 and similarly for the lower hemisphere. Use this to show S^2 is homeomorphic to $D^2 \cup_g D^2$ for $g : S^1 \to S^1$ and determine g.

Exercise 1.9.82. Show that $D^2 \cup_f D^2$, where $f : K \to K$ is $f(x) = x$, and $K = \{(x_1, x_2) \in S^1 : x_1 \geq 0\}$ is homeomorphic to D^2. (Hint: First choose a homeomorphism h from D^2 to $D^1 \times D^1$ where $h(K) = \{(x_1, x_2) \in D^1 \times D^1 : x_1 = 1\}$. Use this to get a homeomorphism $D^2 \cup_f D^2 \simeq (D^1 \times D^1) \cup_g (D^1 \times D^1)$, where $g : h(K) \to h(K)$ is $g(x) = x$.)

Exercise 1.9.83. Construct a homeomorphism between a square and a diamond.

Exercise 1.9.84. Construct a homeomorphism between the two regions in Figure 1.23.

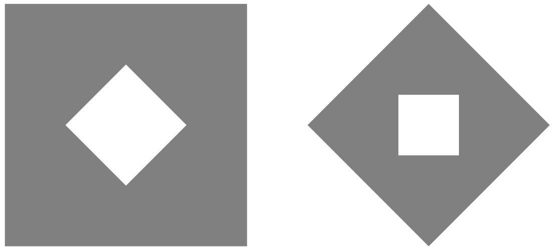


Figure 1.23. Annular regions.

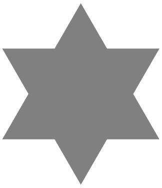


Figure 1.24. Star.

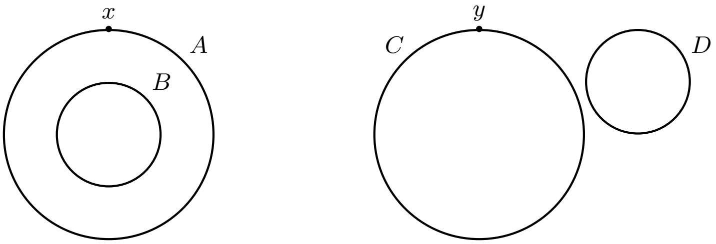


Figure 1.25. Two pairs of circles.

Exercise 1.9.85. Show that any two rectangles in the plane are homeomorphic.

Exercise 1.9.86. Construct a homeomorphism between the inside of a square and the star in Figure 1.24.

Exercise 1.9.87. Show that there is no homeomorphism of the plane to itself which sends the unit circle to itself and sends $(0, 0)$ to $(2, 0)$.

Exercise 1.9.88. Construct an example of a simple closed curve in the plane where a horizontal line intersects the curve in an infinite number of points but the curve contains no horizontal line segments.

Exercise 1.9.89. Show that the complement of two disjoint polygonal simple closed curves in the plane consists of three disjoint open, path connected sets.

The next three problems concern Figure 1.25.

Exercise 1.9.90. Show that there is a homeomorphism sending $A \cup B$ to $C \cup D$.

Exercise 1.9.91. Show that any homeomorphism sending $A \cup B$ to $C \cup D$ which sends $x \in A$ to $y \in C$ must send A homeomorphically to C and B homeomorphically to D.

Exercise 1.9.92. Show that there does not exist a homeomorphism of the plane sending $A \cup B$ to $C \cup D$. (Hint: Consider the regions bounded by A and C.)

Exercise 1.9.93. Consider a polygonal path P that is not closed and does not intersect itself. Show that $\mathbb{R}^2 \setminus P$ is path connected using a polygonal path. (Hint: Use induction on the number of segments in the path.)

Exercise 1.9.94. Show directly that the triangle with vertices $(0,0), (1,0), (0,1)$ separates the plane into two nonempty disjoint open path connected sets, one of which is bounded and the other not.

Exercise 1.9.95. Consider the shaded region in Figure 1.26 (which is not homeomorphic to a disk). Analyze how the region changes as we move upward past the special vertices.

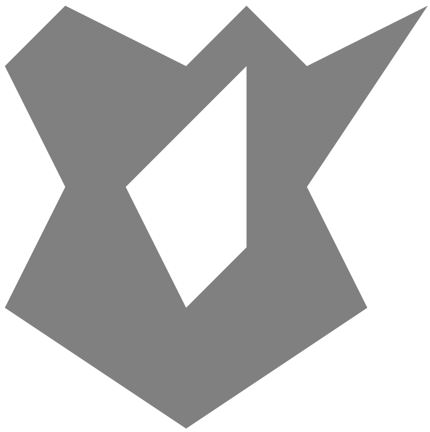


Figure 1.26. A polygonal annular region.

Figure 1.27. A curvy disk.

Exercise 1.9.96. Show that the region in Figure 1.26 is homeomorphic to the region R enclosed between the squares $[-1, 1] \times [-1, 1]$ and $[-2, 2] \times [-2, 2]$. Do this both by a direct argument and by breaking the each region into two regions to which we can apply the polygonal Schönflies theorem.

Exercise 1.9.97. Describe a homeomorphism between the region in Figure 1.27 and a disk.