Basic Topology taken from [1]

1 Metric space topology

We introduce basic notions from point set topology. These notions are prerequisites for more sophisticated topological ideas—manifolds, homeomorphism, and isotopy—introduced later to study algorithms for topological data analysis. To a layman, the word topology evokes visions of "rubber-sheet topology": the idea that if you bend and stretch a sheet of rubber, it changes shape but always preserves the underlying structure of how it is connected to itself. Homeomorphisms offer a rigorous way to state that an operation preserves the topology of a domain, and isotopy offers a rigorous way to state that the domain can be deformed into a shape without ever colliding with itself.

Topology begins with a set $\mathbb{T}$ of points—perhaps the points comprising the d-dimensional Euclidean space $\mathbb{R}^d$, or perhaps the points on the surface of a volume such as a coffee mug. We suppose that there is a *metric* $d(p,q)$ that specifies the scalar *distance* between every pair of points $p, q \in \mathbb{T}$. In the Euclidean space $\mathbb{R}^d$ we choose the Euclidean distance. On the surface of the coffee mug, we could choose the Euclidean distance too; alternatively, we could choose the *geodesic distance*, namely the length of the shortest path from p to q on the mug's surface.

Let us briefly review the Euclidean metric. We write points in $\mathbb{R}^d$ as $p = (p_1, p_2, \ldots, p_d)$, where each p_i is a real-valued *coordinate*. The *Euclidean inner product* of two points $p, q \in \mathbb{R}^d$ is $\langle p, q \rangle = \sum_{i=1}^d p_i q_i$. The *Euclidean norm* of a point $p \in \mathbb{R}^d$ is $\|p\| = \langle p, p \rangle^{1/2} = (\sum_{i=1}^d p_i^2)^{1/2}$, and the *Euclidean distance* between two points $p, q \in \mathbb{R}^d$ is $d(p,q) = \|p - q\| = (\sum_{i=1}^d (p_i - q_i)^2)^{1/2}$. We also use the notation $d(\cdot,\cdot)$ to express minimum distances between point sets $P, Q \subseteq \mathbb{T}$,

$$\begin{aligned} d(p, Q) &= \inf\{d(p,q) : q \in Q\} \text{ and} \\ d(P, Q) &= \inf\{d(p,q) : p \in P, q \in Q\}. \end{aligned}$$

The heart of topology is the question of what it means for a set of points—say, a squiggle drawn on a piece of paper—to be *connected*. After all, two distinct points cannot be adjacent to each other; they can only be connected to another by an uncountably infinite bunch of intermediate points. Topologists solve that mystery with the idea of *limit points*.

Definition 1 (limit point)**.** Let $Q \subseteq \mathbb{T}$ be a point set. A point $p \in \mathbb{T}$ is a *limit point* of Q, also known as an *accumulation point* of Q, if for every real number $\epsilon > 0$, however tiny, Q contains a point $q \neq p$ such that that $d(p,q) < \epsilon$.

In other words, there is an infinite sequence of points in Q that get successively closer and closer to p—without actually being p—and get arbitrarily close. Stated succinctly, $d(p, Q \setminus \{p\}) = 0$. Observe that it doesn't matter whether $p \in Q$ or not.

Definition 2 (connected)**.** Let $Q \subseteq \mathbb{T}$ be a point set. Imagine coloring every point in Q either red or blue. Q is *disconnected* if there exists a coloring having at least one red point and at least one blue point, wherein no red point is a limit point of the blue points, and no blue point is a limit point of the red points. A disconnected point set appears at left in Figure 1. If no such coloring exists, Q is *connected*, like the point set at right in Figure 1.

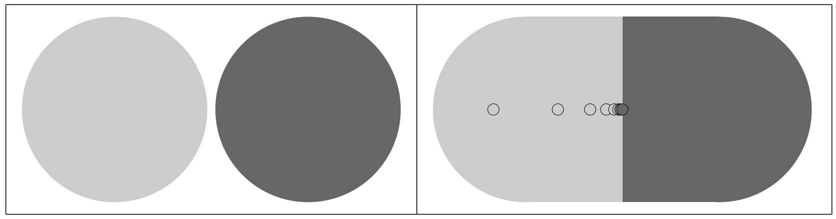


Figure 1: The disconnected point set at left can be partitioned into two connected subsets, which are colored differently here. The point set at right is connected. The dark point at its center is a limit point of the lightly colored points.

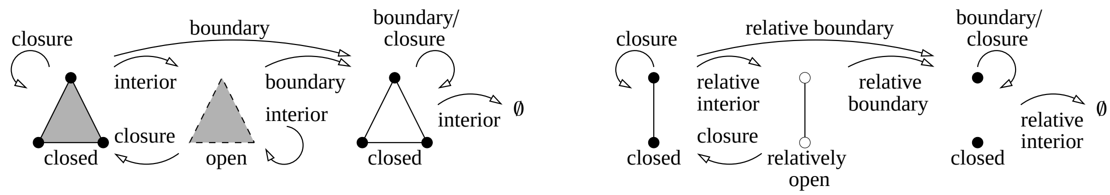


Figure 2: Closed, open, and relatively open point sets in the plane. Dashed edges and open circles indicate points missing from the point set.

We frequently distinguish between closed and open point sets. Informally, a triangle in the plane is *closed* if it contains all the points on its edges, and *open* if it excludes all the points on its edges, as illustrated in Figure 2. The idea can be formally extended to any point set.

Definition 3 (closure; closed; open)**.** The *closure* of a point set $Q \subseteq \mathbb{T}$, denoted $\mathrm{Cl}\, Q$, is the set containing every point in Q and every limit point of Q. A point set Q is *closed* if $Q = \mathrm{Cl}\, Q$, i.e. Q contains all its limit points. The *complement* of a point set Q is $\mathbb{T} \setminus Q$. A point set Q is *open* if its complement is closed, i.e. $\mathbb{T} \setminus Q = \mathrm{Cl}\,(\mathbb{T} \setminus Q)$.

For example, let $(0, 1)$ denote an *open interval* on the real number line—the set containing every $r \in \mathbb{R}$ such that $r > 0$ and $r < 1$—and let $[0, 1]$ denote a *closed interval* $(0, 1) \cup \{0\} \cup \{1\}$. The numbers zero and one are both limit points of the open interval, so $\mathrm{Cl}\,(0, 1) = [0, 1] = \mathrm{Cl}\,[0, 1]$. Therefore, $[0, 1]$ is closed and $(0, 1)$ is not. The numbers zero and one are also limit points of the complement of the closed interval, $\mathbb{R} \setminus [0, 1]$, so $(0, 1)$ is open, but $[0, 1]$ is not.

The terminology is misleading because "closed" and "open" are not opposites. In every nonempty metric space $\mathbb{T}$, there are at least two point sets that are both closed and open: $\emptyset$ and $\mathbb{T}$. The interval $(0, 1]$ on the real number line is neither open nor closed.

The definition of *open set* hides a subtlety that often misleads newcomers to point set topology: a triangle τ that is missing the points on its edges, and therefore is open in the two-dimensional metric space $\mathrm{aff}\,\tau$, is not open in the metric space $\mathbb{R}^3$. Every point in τ is a limit point of $\mathbb{R}^3 \setminus \tau$, because we can find sequences of points that approach τ from the side. In recognition of this quirk, a simplex $\sigma \subset \mathbb{R}^d$ is said to be *relatively open* if it is open relative to its affine hull.

Informally, the boundary of a point set Q is the set of points where Q meets its complement $\mathbb{T} \setminus Q$. The interior of Q contains all the other points of Q. Limit points provide formal definitions.

Definition 4 (boundary; interior)**.** The *boundary* of a point set Q in a metric space $\mathbb{T}$, denoted $\mathrm{Bd}\, Q$, is the intersection of the closures of Q and its complement; i.e. $\mathrm{Bd}\, Q = \mathrm{Cl}\, Q \cap \mathrm{Cl}\, (\mathbb{T} \setminus Q)$. The *interior* of Q, denoted $\mathrm{Int}\, Q$, is $Q \setminus \mathrm{Bd}\, Q = Q \setminus \mathrm{Cl}\, (\mathbb{T} \setminus Q)$.

For example, $\mathrm{Bd}\, [0, 1] = \{0, 1\} = \mathrm{Bd}\, (0, 1)$ and $\mathrm{Int}\, [0, 1] = (0, 1) = \mathrm{Int}\, (0, 1)$. The boundary of a triangle (closed or open) in the Euclidean plane is the union of the triangle's three edges, and its interior is an open triangle, illustrated in Figure 2. The terms *boundary* and *interior* have the same misleading subtlety as open sets: the boundary of a triangle embedded in $\mathbb{R}^3$ is the whole triangle, and its interior is the empty set. Hence the following terms.

Definition 5 (relative boundary; relative interior)**.** The *relative boundary* of a convex polyhedron $C \subset \mathbb{R}^d$ is its boundary with respect to the metric space of its affine hull—that is, $\mathrm{Cl}\, C \cap \mathrm{Cl}\, ((\mathrm{aff}\, C) \setminus C)$. The *relative interior* of C is C minus its relative boundary.

Again, we often abuse terminology by writing "boundary" for relative boundary and "interior" for relative interior. The same subtlety arises with curved ridges and surface patches, but these have fundamentally different definitions of "boundary" and "interior," which we give in Section 4.

Definition 6 (bounded; compact)**.** The *diameter* of a point set Q is $\sup_{p,q \in Q} d(p, q)$. The set Q is *bounded* if its diameter is finite, or *unbounded* if its diameter is infinite. A point set Q in a metric space is *compact* if it is closed and bounded.

As we have defined them, simplices and polyhedra are bounded. We may have unbounded polyhedra, which arise in Voronoi diagrams. Besides simplices and polyhedra, the point sets we use most are balls.

Definition 7 (Euclidean ball)**.** In $\mathbb{R}^d$, the *Euclidean d-ball* with center c and radius r, denoted $B(c, r)$, is the point set $B(c, r) = \{p \in \mathbb{R}^d : d(p, c) \leq r\}$. A 1-ball is an edge, and a 2-ball is called a *disk*. A *unit ball* is a ball with radius 1. The boundary of the d-ball is called the *Euclidean* $(d-1)$*-sphere* and denoted $S(c, r) = \{p \in \mathbb{R}^d : d(p, c) = r\}$. For example, a circle is a 1-sphere, and a layman's "sphere" in $\mathbb{R}^3$ is a 2-sphere. If we remove the boundary from a ball, we have the *open Euclidean d-ball* $B_o(c, r) = \{p \in \mathbb{R}^d : d(p, c) < r\}$.

The foregoing text introduces point set topology in terms of metric spaces. Surprisingly, it is possible to define all the same concepts without the use of a metric, point coordinates, or any scalar values at all. Section 2 discusses *topological spaces*, a mathematical abstraction for representing the topology of a point set while excluding all information that is not topologically essential.

2 Topology sans metric

In Section 1, we state that the heart of topology is to ask what it means for a set of points to be connected, and we answer that question with the concept of limit points in metric spaces. Topological spaces provide a way to describe the topology of a point set without a metric or point coordinates, so they are more abstract but more general than metric spaces. In a topological space, points are abstract entities that might have no characteristics except that they can be distinguished from one other. However, topological spaces remain founded on the concept of limit points. In place of a metric, we encode the connectedness of a point set by supplying a list of all of the open sets. This list is called a *system* of subsets of the point set. The point set and its system together describe a topological space.

Definition 8 (topological space)**.** A *topological space* is a point set $\mathbb{T}$ endowed with a *system of subsets* T, which is a set of subsets of $\mathbb{T}$ that satisfies the following conditions.

- $\emptyset, \mathbb{T} \in T$.
- For every $U \subseteq T$, the union of the subsets in U is in T.
- For every finite $U \subseteq T$, the common intersection of the subsets in U is in T.

The system T is called a *topology* on $\mathbb{T}$. The sets in T are called the *open sets* in $\mathbb{T}$. A *neighborhood* of a point $p \in \mathbb{T}$ is an open set containing p.

The axioms of Definition 8 may seem puzzling; we will not use them explicitly. Mathematicians have found them to be a simple and general set of rules from which one can derive most of the topological concepts one expects from familiarity with metric space topology.

Topological spaces may seem baffling from a computational point of view, because a point set with an interesting topology has uncountably infinitely many open sets containing uncountably infinitely many points. But from a mathematical point of view, topological spaces are attractive because they exclude information that is not topologically essential. For instance, the act of stretching a rubber sheet changes the distances between points and thereby changes the metric, but it does not change the open sets or the topology of the rubber sheet.

The charm of a pure topological space is that, for example, all 2-spheres are indistinguishable from each other, so we simply call them “the 2-sphere.”

Of course, the easiest way to define a topological space is to inherit the open sets from a metric space. For example, we can construct a topology on the d-dimensional Euclidean space $\mathbb{R}^d$ by letting T be the set of all possible open sets in $\mathbb{R}^d$. We can make the idea of “all possible open sets in $\mathbb{R}^d$” more concrete. Every open set in $\mathbb{R}^d$ is a union of a set of open d-balls, and vice versa, although sometimes requiring uncountably many d-balls. Therefore, we can let T be the set of all possible unions of open balls. In this topology, every open ball is a neighborhood of the point at its center.

In Section 1, we build the concepts of topology around the idea of limit points. Topological spaces require a different definition of limit point, but with the new definition in place, concepts that are defined in terms of limit points such as connectedness and closure extend without change to topological spaces.

Definition 9 (limit point)**.** Let $Q \subseteq \mathbb{T}$ be a point set. A point $p \in \mathbb{T}$ is a *limit point* of Q if every open set that contains p also contains a point in $Q \setminus \{p\}$.

Recall from Definition 3 that the *closure* $\mathrm{Cl}\, Q$ of a point set $Q \subseteq \mathbb{T}$ is the set containing every point in Q and every limit point of Q, and a point set Q is *closed* if $Q = \mathrm{Cl}\, Q$. It is straightforward to prove that in a topological space, the *complement* $\mathbb{T} \setminus Q$ of every open set $Q \in \mathbb{T}$ is closed, and that $\mathrm{Cl}\, Q$ is the smallest closed set containing Q.

For every point set $\mathbb{U} \subseteq \mathbb{T}$, the topology T induces a *subspace topology* on $\mathbb{U}$, namely the system of open subsets $U = \{P \cap \mathbb{U} : P \in T\}$. The point set $\mathbb{U}$ endowed with the system U is said to be a *topological subspace* of $\mathbb{T}$. The topological spaces we consider in this course are subsets of a metric space such as $\mathbb{R}^d$ that inherit its topology as a subspace topology. Examples of topological subspaces are the Euclidean d-ball $\mathbb{B}^d$, Euclidean d-sphere $\mathbb{S}^d$, open Euclidean d-ball $\mathbb{B}^d_o$, and Euclidean halfball $\mathbb{H}^d$, where

$$\begin{aligned}
\mathbb{B}^d &= \{x \in \mathbb{R}^d : \|x\| \leq 1\}, \\
\mathbb{S}^d &= \{x \in \mathbb{R}^{d+1} : \|x\| = 1\}, \\
\mathbb{B}^d_o &= \{x \in \mathbb{R}^d : \|x\| < 1\}, \\
\mathbb{H}^d &= \{x \in \mathbb{R}^d : \|x\| < 1 \text{ and } x_d \geq 0\}.
\end{aligned}$$

3 Maps, homeomorphisms, and isotopies

Two topological spaces are considered to be the same if the points that comprise them are connected the same way. For example, the boundary of a cube can be deformed into a sphere without cutting or gluing it. They have the same topology. We formalize this notion of topological equality by defining a function that maps the points of one space to points of the other and preserves how they are connected. Specifically, the function preserves limit points.

A function from one space to another that preserves limit points is called a *continuous function* or a *map*.[1] Continuity is just a step on the way to topological equivalence, because a continuous function can map many points to a single point in the target space, or map no points to a given point in the target space. If the former does not happen—that is, if the function is injective—the function is called an *embedding* of the domain into the target space. True equivalence is marked by a *homeomorphism*, a bijective function from one space to another that possesses both continuity and a continuous inverse, so that limit points are preserved in both directions.

Definition 10 (continuous function; map)**.** Let $\mathbb{T}$ and $\mathbb{U}$ be topological spaces. A function $g : \mathbb{T} \to \mathbb{U}$ is *continuous* if for every set $Q \subseteq \mathbb{T}$ and every limit point $p \in \mathbb{T}$ of Q, $g(p)$ is either a limit point of the set $g(Q)$ or in $g(Q)$. Continuous functions are also called *maps*.

Definition 11 (embedding)**.** A map $g : \mathbb{T} \to \mathbb{U}$ is an *embedding* of $\mathbb{T}$ into $\mathbb{U}$ if g is injective.

A topological space can be *embedded* into a Euclidean space by assigning coordinates to its points such that the assignment is continuous. For example, a drawing of a square is an embedding of $\mathbb{S}^1$ into $\mathbb{R}^2$. Not every topological space has an embedding into a Euclidean space, or even into a metric space—there are spaces that cannot be represented by any metric—but we will have no need for such spaces.

A homeomorphism is an embedding whose inverse is also an embedding.

Definition 12 (homeomorphism)**.** Let $\mathbb{T}$ and $\mathbb{U}$ be topological spaces. A *homeomorphism* is a bijective map $h : \mathbb{T} \to \mathbb{U}$ whose inverse is continuous too.

Two topological spaces are *homeomorphic* if there exists a homeomorphism between them.

[1]There is a small caveat with this characterization: a function g that maps a neighborhood of x to a single point $g(x)$ may be continuous, but technically $g(x)$ is not a limit point of itself, so in this sense a continuous function might not preserve all limit points. This technicality does not apply to homeomorphisms because they are bijective; homeomorphisms preserve all limit points, in both directions.

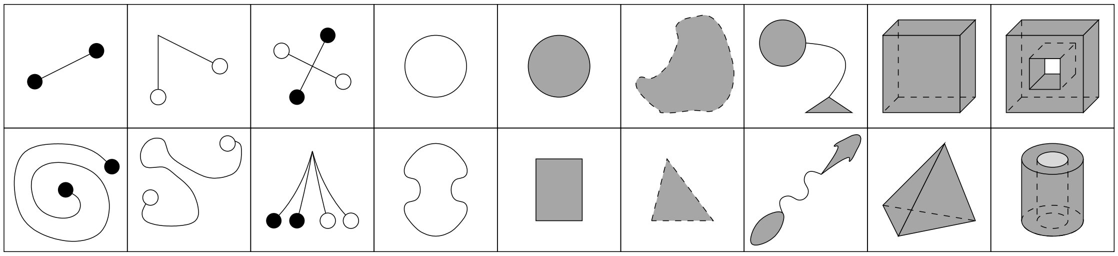


Figure 3: Each point set in this figure is homeomorphic to the point set above or below it, but not to any of the others. Open circles indicate points missing from the point set, as do the dashed edges in the point sets fourth from the right.

Homeomorphism induces an equivalence relation among topological spaces, which is why two homeomorphic topological spaces are called *topologically equivalent*. Figure 3 show pairs of topological spaces that are homeomorphic. A less obvious example is that the open d-ball $\mathbb{B}^d_o$ is homeomorphic to the Euclidean space $\mathbb{R}^d$, as demonstrated by the map $h(x) = \frac{1}{1-\|x\|}x$. The same map shows that the halfball $\mathbb{H}^d$ is homeomorphic to the Euclidean halfspace $\{x \in \mathbb{R}^d \, : \, x_d \geq 0\}$.

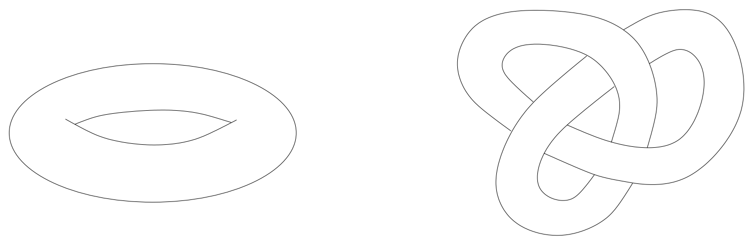
A subspace of a Euclidean space is said to be *compact* if it is bounded and closed with respect to the Euclidean space. Boundedness is a metric space concept; there is a purely topological definition of compactness, which we omit. If $\mathbb{T}$ and $\mathbb{U}$ are compact metric spaces, every bijective map from $\mathbb{T}$ to $\mathbb{U}$ has a continuous inverse. We will take advantage of this fact to prove that certain functions are homeomorphisms. When two topological spaces are subspaces of the same larger space, there is another notion of similarity that is stronger than homeomorphism, called *isotopy*. If two subspaces are isotopic, one can be continuously deformed into the other so that the deforming subspace remains always homeomorphic to its original form. For example, a cube can be continuously deformed into a ball.

Homeomorphic subspaces are not necessarily isotopic. Consider a torus embedded in $\mathbb{R}^3$, illustrated in Figure 4(a). One can embed the torus in $\mathbb{R}^3$ so that it is knotted, as shown in Figure 4(b). The knotted torus is homeomorphic to the standard, unknotted one. However, it is not possible to continuously deform one to the other while keeping it embedded in $\mathbb{R}^3$ and topologically unchanged. Any attempt to do so will cause the torus to pass through a state in which it is "self-intersecting" and not a manifold. The easiest way to recognize this fact is to look not at the topology of the tori, but at the topology of the space around them. Although the knotted and unknotted tori are homeomorphic, their complements are not. Therefore, we consider both the notion of an *isotopy*, in which a torus deforms continuously, and the notion of an *ambient isotopy*, in which not only the torus deforms; the entirety of $\mathbb{R}^3$ deforms with it.

Definition 13 (isotopy)**.** An *isotopy* connecting two spaces $\mathbb{T} \subseteq \mathbb{R}^d$ and $\mathbb{U} \subseteq \mathbb{R}^d$ is a continuous map $\xi : \mathbb{T} \times [0,1] \to \mathbb{R}^d$ where $\xi(\mathbb{T}, 0) = \mathbb{T}$, $\xi(\mathbb{T}, 1) = \mathbb{U}$, and for every $t \in [0,1]$, $\xi(\cdot, t)$ is a homeomorphism between $\mathbb{T}$ and its image $\{\xi(x,t) : x \in \mathbb{T}\}$. An *ambient isotopy* connecting $\mathbb{T}$ and $\mathbb{U}$ is a map $\xi : \mathbb{R}^d \times [0,1] \to \mathbb{R}^d$ such that $\xi(\cdot, 0)$ is the identity function on $\mathbb{R}^d$, $\xi(\mathbb{T}, 1) = \mathbb{U}$, and for each $t \in [0,1]$, $\xi(\cdot, t)$ is a homeomorphism.

For example, the map

$$\xi(x,t) = \frac{1 - (1-t)\|x\|}{1 - \|x\|}x$$



(a) (b)

Figure 4: (a) Torus. (b) Knotted torus.

sends the open d-ball $\mathbb{B}_o^d$ to itself if $t = 0$, and to the Euclidean space $\mathbb{R}^d$ if $t = 1$. Think of the parameter t as the time, so $\xi(\mathbb{B}_o^d, t)$ deforms continuously from a ball at time zero to $\mathbb{R}^d$ at time one. Hence the open d-ball and $\mathbb{R}^d$ are related by an isotopy.

Every ambient isotopy becomes an isotopy if its domain is restricted from $\mathbb{R}^d \times [0, 1]$ to $\mathbb{T} \times [0, 1]$. It is known that if two subspaces are related by an isotopy, there exists an ambient isotopy connecting them, so the two notions are equivalent.

There is another notion of similarity among topological spaces that is weaker than homeomorphism, called *homotopy equivalence*. It relates spaces that can be continuously deformed to one another but may not be homeomorphic. For example, a ball can shrink to a point, but they are not homeomorphic; there is not even a bijective function from an infinite point set to a single point. However, homotopy preserves some aspects of connectedness, such as the number of connected components and the number of holes in a space. Thus a coffee cup is homotopy equivalent to a circle, but not to a ball or a point.

To get to homotopy equivalence, we first need the concept of homotopies, which generalize isotopies so that homeomorphism is not required.

Definition 14 (homotopy)**.** Let $g : \mathbb{X} \to \mathbb{U}$ and $h : \mathbb{X} \to \mathbb{U}$ be maps. A *homotopy* is a map $H : \mathbb{X} \times [0, 1] \to \mathbb{U}$ such that $H(\cdot, 0) = g$ and $H(\cdot, 1) = h$. Two maps are *homotopic* if there is a homotopy connecting them.

For example, if $h : \mathbb{B}^3 \to \mathbb{R}^3$ is the identity map on a unit ball and $g : \mathbb{B}^3 \to \mathbb{R}^3$ maps every point in the ball to the origin, the fact that g and h are homotopic is demonstrated by the homotopy $H(x, t) = t \cdot h(x)$; hence $H(\mathbb{B}^3, t)$ deforms continuously from a point at time zero to a ball at time one. A key property of a homotopy is that, as H is continuous, at every time t the map $H(\cdot, t)$ is continuous.

It is more revealing to consider two maps that are not homotopic. Let $g : \mathbb{S}^1 \to \mathbb{S}^1$ be the identity map from the circle to itself, and let $h : \mathbb{S}^1 \to \mathbb{S}^1$ map every point on the circle to a single point $p \in \mathbb{S}^1$. Although it is easy to imagine contracting a circle to a point, that image is misleading: the map H is constrained by the requirement that every point on the circle at every time maps to a point on the circle. The circle can contract to a point only if we cut it somewhere, implying that H is not continuous.

Observe that whereas a homeomorphism is a topological relationship between two topological spaces $\mathbb{T}$ and $\mathbb{U}$, a homotopy or an isotopy (which is a special kind of homotopy) is a relationship between two maps, which indirectly establishes a relationship between two topological subspaces

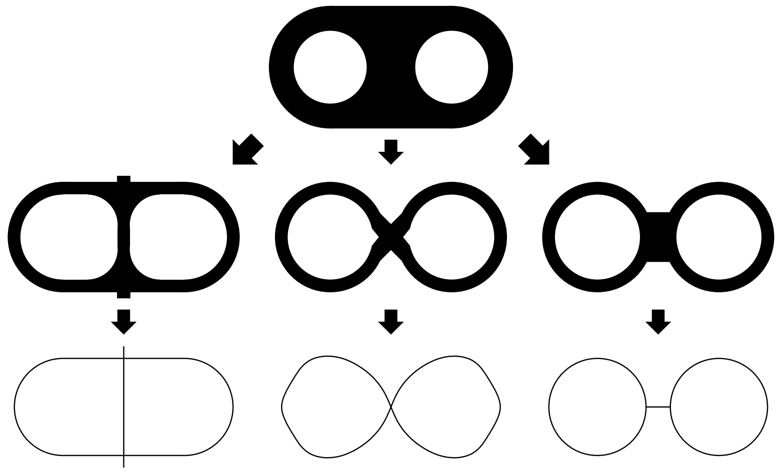


Figure 5: All seven of these point sets are homotopy equivalent, because they are all deformation retracts of the top point set.

$g(\mathbb{X}) \subseteq \mathbb{U}$ and $h(\mathbb{X}) \subseteq \mathbb{U}$. That relationship is not necessarily an equivalence class, but the following relationship is.

Definition 15 (homotopy equivalent)**.** Two topological spaces $\mathbb{T}$ and $\mathbb{U}$ are *homotopy equivalent* if there exist maps $g : \mathbb{T} \to \mathbb{U}$ and $h : \mathbb{U} \to \mathbb{T}$ such that $h \circ g$ is homotopic to the identity map $\iota_{\mathbb{T}} : \mathbb{T} \to \mathbb{T}$ and $g \circ h$ is homotopic to the identity map $\iota_{\mathbb{U}} : \mathbb{U} \to \mathbb{U}$.

Whereas homeomorphic spaces have the same dimension, homotopy equivalent spaces sometimes do not. To see that the 2-ball is homotopy equivalent to a single point p, construct a map $h : \mathbb{B}^2 \to \{p\}$ and a map $g : \{p\} \to \mathbb{B}^2$ where $g(p)$ is any point q in $\mathbb{B}^2$. Observe that $h \circ g$ is the identity map on $\{p\}$, which is trivially homotopic to itself. In the other direction, $g \circ h : \mathbb{B}^2 \to \mathbb{B}^2$ sends every point in $\mathbb{B}^2$ to q. There is a homotopy connecting $g \circ h$ to the identity map $\iota_{\mathbb{B}^2}$, namely the map $H(x, t) = (1 - t)q + tx$.

The definition of homotopy equivalent is somewhat mysterious. A useful intuition for understanding it is the fact that two spaces $\mathbb{T}$ and $\mathbb{U}$ are homotopy equivalent if and only if there exists a third space $\mathbb{X}$ such that both $\mathbb{T}$ and $\mathbb{U}$ are *deformation retracts* of $\mathbb{X}$, illustrated in Figure 5.

Definition 16 (deformation retract)**.** Let $\mathbb{T}$ be a topological space, and let $\mathbb{U} \subset \mathbb{T}$ be a subspace. A *retraction* r of $\mathbb{T}$ to $\mathbb{U}$ is a map from $\mathbb{T}$ to $\mathbb{U}$ such that $r(x) = x$ for every $x \in \mathbb{U}$. The space $\mathbb{U}$ is a *deformation retract* of $\mathbb{T}$ if the identity map on $\mathbb{T}$ can be continuously deformed to a retraction with no motion of the points already in $\mathbb{U}$: specifically, there is a homotopy $R : \mathbb{T} \times [0, 1] \to \mathbb{T}$ such that $R(\cdot, 0)$ is the identity map on $\mathbb{T}$, $R(\cdot, 1)$ is a retraction of $\mathbb{T}$ to $\mathbb{U}$, and $R(x, t) = x$ for every $x \in \mathbb{U}$ and every $t \in [0, 1]$.

If $\mathbb{U}$ is a deformation retract of $\mathbb{T}$, then $\mathbb{T}$ and $\mathbb{U}$ are homotopy equivalent. For example, any point on a line segment (open or closed) is a deformation retract of the line segment and is homotopy equivalent to it. The letter V is a deformation retract of the letter W, and also of a ball. Moreover, two spaces are homotopy equivalent if they are deformation retractions of a common space. The symbols $\varnothing$, ∞, and ⊸∘ (viewed as one-dimensional point sets) are deformation retracts

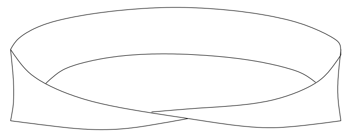


Figure 6: Möbius band.

of a double doughnut—a doughnut with two holes. Therefore, they are homotopy equivalent to each other, although none of them is a deformation retract of any of the others. They are not homotopy equivalent to X, O, ⊕, ⊙, ⊚, a ball, or a coffee cup.

4 Manifolds

A manifold is a set of points that is locally connected in a particular way. A 1-manifold has the structure of a piece of string, possibly with its ends tied in a loop, and a 2-manifold has the structure of a piece of paper or rubber sheet that has been cut and perhaps glued along its edges—a category that includes disks, spheres, tori, and Möbius bands.

Definition 17 (manifold)**.** A topological space Σ is a *k-manifold*, or simply *manifold*, if every point $x \in \Sigma$ has a neighborhood homeomorphic to $\mathbb{B}_o^k$ or $\mathbb{H}^k$. The *dimension* of Σ is k.

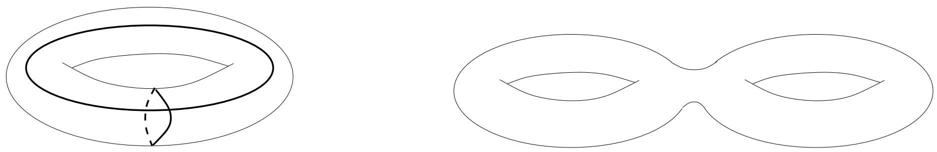
A manifold can be viewed as a purely abstract topological space, or it can be embedded into a metric space or a Euclidean space. Even without an embedding, every manifold can be partitioned into boundary and interior points. Observe that these words mean very different things for a manifold than they do for a metric space or topological space.

Definition 18 (boundary; interior)**.** The *interior* $\operatorname{Int}\Sigma$ of a manifold Σ is the set of points in Σ that have a neighborhood homeomorphic to $\mathbb{B}_o^k$. The *boundary* $\operatorname{Bd}\Sigma$ of Σ is the set of points $\Sigma \setminus \operatorname{Int}\Sigma$. The boundary $\operatorname{Bd}\Sigma$, if not empty, consists of the points that have a neighborhood homeomorphic to $\mathbb{H}^k$. If $\operatorname{Bd}\Sigma$ is the empty set, we say that Σ is *without boundary*.

A single point, a 0-ball, is a 0-manifold without boundary according to this definition. The closed disk $\mathbb{B}^2$ is a 2-manifold whose interior is the open disk $\mathbb{B}_o^2$ and whose boundary is the circle $\mathbb{S}^1$. The open disk $\mathbb{B}_o^2$ is a 2-manifold whose interior is $\mathbb{B}_o^2$ and whose boundary is the empty set. This highlights an important difference between Definitions 4 and 18 of "boundary": when $\mathbb{B}_o^2$ is viewed as a point set in the space $\mathbb{R}^2$, its boundary is $\mathbb{S}^1$ according to Definition 4; but viewed as a manifold, its boundary is empty according to Definition 18. The boundary of a manifold is *always* included in the manifold.

The open disk $\mathbb{B}_o^2$, the Euclidean space $\mathbb{R}^2$, the sphere $\mathbb{S}^2$, and the torus are all connected 2-manifolds without boundary. The first two are homeomorphic to each other, but the last two are topologically different from the others. The sphere and the torus are compact (bounded and closed with respect to $\mathbb{R}^3$) whereas $\mathbb{B}_o^2$ and $\mathbb{R}^2$ are not.

A 2-manifold Σ is *non-orientable* if, starting from a point p, one can walk on one side of Σ and end up on the opposite side of Σ upon returning to p. Otherwise, Σ is *orientable*. Spheres and balls are orientable, whereas the *Möbius band* in Figure 6 is a non-orientable 2-manifold.



(a) (b)

Figure 7: (a) Removal of the bold loops opens up the torus into a topological disk. (b) Every surface without boundary in $\mathbb{R}^3$ resembles a sphere or a conjunction of one or more tori.

A *surface* is a 2-manifold that is a subspace of $\mathbb{R}^d$. Any compact surface without boundary in $\mathbb{R}^3$ is an orientable 2-manifold. To be non-orientable, a compact surface must have a nonempty boundary (like the Möbius band) or be embedded in a 4- or higher-dimensional Euclidean space.

A surface can sometimes be disconnected by removing one or more *loops* (connected 1-manifolds without boundary) from it. The *genus* of a surface is g if $2g$ is the maximum number of loops that can be removed from the surface without disconnecting it; here the loops are permitted to intersect each other. For example, the sphere has genus zero as every loop cuts it into two balls. The torus has genus one: a circular cut around its neck and a second circular cut around its circumference, illustrated in Figure 7(a), allow it to unfold into a rectangle, which topologically is a disk. A third loop would cut it into two pieces. Figure 7(b) shows a 2-manifold without boundary of genus 2. Although a high-genus surface can have a very complex shape, all compact 2-manifolds in $\mathbb{R}^3$ that have the same genus and no boundary are homeomorphic to each other.

5 Smooth manifolds

A purely topological manifold has no geometry, but once embedded in a Euclidean space it may appear smooth or creased. Here we enrich the notion of a geometric manifold by imposing a differential structure on it. For the rest of this chapter, we are discussing only manifolds without boundary.

Consider a map $\phi : U \to W$ where U and W are open sets in $\mathbb{R}^k$ and $\mathbb{R}^d$, respectively. The map ϕ has d components, namely $\phi(x) = (\phi_1(x), \phi_2(x), \ldots, \phi_d(x))$, where $x = (x_1, x_2, \ldots, x_k)$ denotes a point in $\mathbb{R}^k$. The *Jacobian* of ϕ at x is the $d \times k$ matrix of the first-order partial derivatives

$$\begin{bmatrix} \frac{\partial\phi_1(x)}{\partial x_1} & \cdots & \frac{\partial\phi_1(x)}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial\phi_d(x)}{\partial x_1} & \cdots & \frac{\partial\phi_d(x)}{\partial x_k} \end{bmatrix}.$$

The map ϕ is *regular* if its Jacobian has rank k at every point in U. The map ϕ is C^i-continuous if the ith-order partial derivatives of ϕ are continuous.

The reader may be familiar with *parametric surfaces*, for which U is a k-dimensional *parameter space* and its image $\phi(U)$ in d-dimensional space is a parametric surface. Unfortunately, a single parametric surface cannot easily represent a manifold with a complicated topology. However, for a manifold to be smooth, it suffices that each point on the manifold has a neighborhood that looks like a smooth parametric surface.

Definition 19 (smooth manifold)**.** For any $i > 0$, a k-manifold Σ without boundary embedded in $\mathbb{R}^d$ is C^i*-smooth* if for every point $p \in \Sigma$, there exists an open set $U_p \subset \mathbb{R}^k$, a neighborhood $W_p \subset \mathbb{R}^d$ of p, and a map $\phi_p : U_p \to W_p \cap \Sigma$ such that (i) ϕ_p is C^i-continuous, (ii) ϕ_p is a homeomorphism, and (iii) ϕ_p is regular. If $k = 2$, we call Σ a C^i*-smooth surface*.

The first condition says that each map is continuously differentiable at least i times. The second condition requires each map to be bijective, ruling out "wrinkles" where multiple points in U map to a single point in W. The third condition prohibits any map from having a directional derivative of zero at any point in any direction. The first and third conditions together enforce smoothness, and imply that there is a well-defined tangent k-flat at each point in Σ. The third condition prohibits any map from having a directional derivative of zero at any point in any direction. The first and third conditions together enforce smoothness, and imply that there is a well-defined tangent k-flat at each point in Σ. The three conditions together imply that the maps ϕ_p defined in the neighborhood of each point $p \in \Sigma$ overlap smoothly. There are two extremes of smoothness. We say that Σ is C^∞-smooth if for every point $p \in \Sigma$, the partial derivatives of ϕ_p of all orders are continuous. On the other hand, Σ is *nonsmooth* if Σ is a k-manifold (therefore C^0-smooth) but not C^1-smooth.

Exercises

1. Let X be a point set, not necessarily finite, in $\mathbb{R}^d$. Prove that the following two definitions of the convex hull of X are equivalent.
   - The set of all points that are convex combinations of the points in X.
   - The intersection of all convex sets that include X.
2. In every metric space $\mathbb{T}$, the point sets $\emptyset$ and $\mathbb{T}$ are both closed and open.
   (a) Give an example of a metric space that has more than two sets that are both closed and open, and list all of those sets.
   (b) Explain the relationship between the idea of connectedness and the number of sets that are both closed and open.
3. Prove that for every subset X of a metric space, $\mathrm{Cl}\,\mathrm{Cl}\,X = \mathrm{Cl}\,X$. In other words, augmenting a set with its limit points does not give it more limit points.

References

[1] S.-W. Cheng, T. K. Dey, and J. R. Shewchuk. Delaunay Mesh Generation. CRC Press, 2012.