

Renzo's Math 490

Introduction to Topology



Tom Babinec	Chris Best	Michael Bliss	Nikolai Brendler	Eric Fu
Adriane Fung	Tyler Klein	Alex Larson	Topcue Lee	John Madonna
Joel Mousseau	Nick Posavetz	Matt Rosenberg	Danielle Rogers	
Andrew Sardone	Justin Shaler	Smrithi Srinivasan	Pete Troyan	
Jackson Yim	Elizabeth Uible	Derek Van Farowe	Paige Warmker	
	Zheng Wu	Nina Zhang		

Winter 2007

Contents

1	Topology	9
1.1	Metric Spaces	9
1.2	Open Sets (in a metric space)	10
1.3	Closed Sets (in a metric space)	11
1.4	Topological Spaces	11
1.5	Closed Sets (Revisited)	12
1.6	Continuity	13
1.7	Homeomorphisms	14
1.8	Homeomorphism Examples	16
1.9	Theorems On Homeomorphism	18
1.10	Homeomorphisms Between Letters of Alphabet	19
1.10.1	Topological Invariants	19
1.10.2	Vertices	19
1.10.3	Holes	20
1.11	Classification of Letters	21
1.11.1	The curious case of the “Q”	22
1.12	Topological Invariants	23
1.12.1	Hausdorff Property	23
1.12.2	Compactness Property	24
1.12.3	Connectedness and Path Connectedness Properties	25
2	Making New Spaces From Old	27
2.1	Cartesian Products of Spaces	27
2.2	The Product Topology	28
2.3	Properties of Product Spaces	29

2.4	Identification Spaces	30
2.5	Group Actions and Quotient Spaces	34
3	First Topological Invariants	37
3.1	Introduction	37
3.2	Compactness	37
3.2.1	Preliminary Ideas	37
3.2.2	The Notion of Compactness	40
3.3	Some Theorems on Compactness	43
3.4	Hausdorff Spaces	47
3.5	T_1 Spaces	49
3.6	Compactification	50
3.6.1	Motivation	50
3.6.2	One-Point Compactification	50
3.6.3	Theorems	51
3.6.4	Examples	55
3.7	Connectedness	57
3.7.1	Introduction	57
3.7.2	Connectedness	58
3.7.3	Path-Connectedness	61
4	Surfaces	63
4.1	Surfaces	63
4.2	The Projective Plane	63
4.2.1	\mathbb{RP}^2 as lines in \mathbb{R}^3 or a sphere with antipodal points identified.	63
4.2.2	The Projective Plane as a Quotient Space of the Sphere	65
4.2.3	The Projective Plane as an identification space of a disc	66
4.2.4	Non-Orientability of the Projective Plane	69
4.3	Polygons	69
4.3.1	Bigons	71
4.3.2	Rectangles	72
4.3.3	Working with and simplifying polygons	74
4.4	Orientability	76
4.4.1	Definition	76

4.4.2	Applications To Common Surfaces	77
4.4.3	Conclusion	80
4.5	Euler Characteristic	80
4.5.1	Requirements	80
4.5.2	Computation	81
4.5.3	Usefulness	83
4.5.4	Use in identification polygons	83
4.6	Connected Sums	85
4.6.1	Definition	85
4.6.2	Well-definedness	85
4.6.3	Examples	87
4.6.4	$\mathbb{R}P^2 \# T = \mathbb{R}P^2 \# \mathbb{R}P^2 \# \mathbb{R}P^2$	88
4.6.5	Associativity	90
4.6.6	Effect on Euler Characteristic	90
4.7	Classification Theorem	92
4.7.1	Equivalent definitions	92
4.7.2	Proof	93
5	Homotopy and the Fundamental Group	97
5.1	Homotopy of functions	97
5.2	The Fundamental Group	100
5.2.1	Free Groups	100
5.2.2	Graphic Representation of Free Group	101
5.2.3	Presentation Of A Group	103
5.2.4	The Fundamental Group	103
5.3	Homotopy Equivalence between Spaces	105
5.3.1	Homeomorphism vs. Homotopy Equivalence	105
5.3.2	Equivalence Relation	106
5.3.3	On the usefulness of Homotopy Equivalence	106
5.3.4	Simple-Connectedness and Contractible spaces	107
5.4	Retractions	108
5.4.1	Examples of Retractions	108

5.5	Computing the Fundamental Groups of Surfaces: The Seifert-Van Kampen Theorem	110
5.5.1	Examples:	112
5.6	Covering Spaces	113
5.6.1	Lifting	117

What is this?

This is a collection of topology notes compiled by Math 490 topology students at the University of Michigan in the Winter 2007 semester. Introductory topics of point-set and algebraic topology are covered in a series of five chapters.

Foreword (for the random person stumbling upon this document)

What you are looking at, my random reader, is not a topology textbook. It is not the lecture notes of my topology class either, but rather my student's free interpretation of it. Well, I should use the word free with a little bit of caution, since they *had to* do this as their final project. These notes are organized and reflect tastes and choices of my students. I have done only a minimal amount of editing, to give a certain unity to the manuscript and to scrap out some mistakes - and so I don't claim merits for this work but to have lead my already great students through this semester long adventure discovering a little bit of topology.

Foreword (for my students)

Well, guys, here it is! You've done it all, and here is a semester worth of labor, studying, but hopefully fun as well. I hope every once in a while you might enjoy flipping through the pages of this book and reminiscing topology...and that in half a century or so you might be telling exaggerated stories to your grandchildren about this class.

A great thank to you all for a very good semester!

Chapter 1

Topology

To understand what a topological space is, there are a number of definitions and issues that we need to address first. Namely, we will discuss metric spaces, open sets, and closed sets. Once we have an idea of these terms, we will have the vocabulary to define a topology. The definition of topology will also give us a more generalized notion of the meaning of open and closed sets.

1.1 Metric Spaces

Definition 1.1.1. A *metric space* is a set X where we have a notion of distance. That is, if $x, y \in X$, then $d(x, y)$ is the “distance” between x and y . The particular distance function must satisfy the following conditions:

1. $d(x, y) \geq 0$ for all $x, y \in X$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) \leq d(x, y) + d(y, z)$

To understand this concept, it is helpful to consider a few examples of what does and does not constitute a distance function for a metric space.

Example 1.1.2. For any space X , let $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise.

This metric, called the **discrete metric**, satisfies the conditions one through four.

Example 1.1.3. The Pythagorean Theorem gives the most familiar notion of distance for points in \mathbb{R}^n . In particular, when given $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$, the distance f as

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Example 1.1.4. Suppose f and g are functions in a space $X = \{f : [0, 1] \rightarrow \mathbb{R}\}$. Does $d(f, g) = \max|f - g|$ define a metric?

Again, in order to check that $d(f, g)$ is a metric, we must check that this function satisfies the above criteria. But in this case property number 2 does not hold, as can be shown by considering two arbitrary functions at any point within the interval $[0, 1]$. If $|f(x) - g(x)| = 0$, this does not imply that $f = g$ because f and g could intersect at one, and only one, point. Therefore, $d(f, g)$ is not a metric in the given space.

1.2 Open Sets (in a metric space)

Now that we have a notion of distance, we can define what it means to be an open set in a metric space.

Definition 1.2.1. Let X be a metric space. A **ball** B of radius r around a point $x \in X$ is $B = \{y \in X | d(x, y) < r\}$.

We recognize that this ball encompasses all points whose distance is less than r from x .

Definition 1.2.2. A subset $O \subseteq X$ is **open** if for every point $x \in O$, there is a ball around x entirely contained in O .

Example 1.2.3. Let $X = [0, 1]$. The interval $(0, 1/2)$ is open in X .

Example 1.2.4. Let $X = \mathbb{R}$. The interval $[0, 1/2)$ is not open in X .

With an open set, we should be able to pick any point within the set, take an infinitesimal step in any direction within our given space, and find another point within the open set. In the first example, we can take any point $0 < x < 1/2$ and find a point to the left or right of it, within the space $[0, 1]$, that also is in the open set $(0, 1/2)$. However, this cannot be done with the second example. For instance, if we take the point within the set $[0, 1/2)$, say 0, and take an infinitesimal step to the left while staying within our given space X , we are no longer within the set $[0, 1/2)$. Therefore, this would not be an open set within \mathbb{R} .

If a set is not open, **this does not imply that it is closed**. Furthermore, there exists sets that are neither open, nor closed, and sets that are open and closed.

Lastly, open sets in spaces X have the following properties:

1. The empty set is open
2. The whole space X is open
3. The union of any collection of open sets is open
4. The intersection of any *finite* number of open sets is open.

1.3 Closed Sets (in a metric space)

While we can and will define a closed sets by using the definition of open sets, we first define it using the notion of a limit point.

Definition 1.3.1. *A point z is a **limit point** for a set A if every open set U containing z intersects A in a point other than z .*

Notice, the point z could be in A or it might not be in A . The following example will help make this clear.

Example 1.3.2. *Consider the open unit disk $D = \{(x, y) : x^2 + y^2 < 1\}$. Any point in D is a limit point of D . Take $(0, 0)$ in D . Any open set U about this point will contain other points in D . Now consider $(1, 0)$, which is not in D . This is still a limit point because any open set about $(1, 0)$ will intersect the disk D .*

The following theorem and examples will give us a useful way to define closed sets, and will also prove to be very helpful when proving that sets are open as well.

Definition 1.3.3. *A set C is a **closed** set if and only if it contains all of its limit points.*

Example 1.3.4. *The unit disk in the previous example is not closed because it does not contain all of its limit points; namely, $(1, 0)$.*

Example 1.3.5. *Let $A = \mathbb{Z}$, a subset of \mathbb{R} . This is a closed set because it does contain all of its limit points; no point is a limit point! A set that has no limit points is closed, by default, because it contains all of its limit points.*

Every intersection of closed sets is closed, and every finite union of closed sets is closed.

1.4 Topological Spaces

We now consider a more general case of spaces *without* metrics, where we can still make sense of (or rather define appropriately) the notions of open and closed sets. These spaces are called *topological spaces*.

Definition 1.4.1. *A **topological space** is a pair (X, τ) where X is a set and τ is a set of subsets of X satisfying certain axioms. τ is called a topology.*

Since this is not particularly enlightening, we must clarify what a topology is.

Definition 1.4.2. *A **topology** τ on a set X consists of subsets of X satisfying the following properties:*

1. *The empty set \emptyset and the space X are both sets in the topology.*
2. *The union of any collection of sets in τ is contained in τ .*
3. *The intersection of any finitely many sets in τ is also contained in τ .*

The following examples will help make this concept more clear.

Example 1.4.3. Consider the following set consisting of 3 points; $X = \{a, b, c\}$ and determine if the set $\tau = \{\emptyset, X, \{a\}, \{b\}\}$ satisfies the requirements for a topology.

This is, in fact, not a topology because the union of the two sets $\{a\}$ and $\{b\}$ is the set $\{a, b\}$, which is not in the set τ .

Example 1.4.4. Find all possible topologies on $X = \{a, b\}$.

1. $\emptyset, \{a, b\}$
2. $\emptyset, \{a\}, \{a, b\}$
3. $\emptyset, \{b\}, \{a, b\}$
4. $\emptyset, \{a\}, \{b\}, \{a, b\}$

The reader can check that all of these are topologies by making sure they follow the 3 properties above. The first topology in the list is a common topology and is usually called the *indiscrete topology*; it contains the empty set and the whole space X . The following examples introduce some additional common topologies:

Example 1.4.5. When X is a set and τ is a topology on X , we say that the sets in τ are open. Therefore, if X does have a metric (a notion of distance), then $\tau = \{\text{all open sets as defined with the ball above}\}$ is indeed a topology. We call this topology the **Euclidean topology**. It is also referred to as the usual or ordinary topology.

Example 1.4.6. If $Y \subseteq X$ and τ_x is a topology on X , one can define the **Induced topology** as $\tau_y = \{O \cap Y \mid O \in \tau_x\}$.

This last example gives one reason why we must only take finitely many intersections when defining a topology.

Example 1.4.7. Let $X = \mathbb{R}$ with the usual topology. Then certainly in this standard Euclidean topology, $(-1/n, 1/n)$ is an open set for any integer n . However, the infinite intersection $\bigcap_{n=1}^{\infty} (-1/n, 1/n)$ is the set containing just 0. Thus, it is a single point set, which is not open in this topology.

1.5 Closed Sets (Revisited)

As promised, we can now generalize our definition for a closed set to one in terms of open sets alone which removes the need for limit points and metrics.

Definition 1.5.1. A set C is closed if $X - C$ is open.

Now that we have a new definition of a closed set, we can prove what used to be definition 1.3.3 as a theorem: A set C is a **closed** set if and only if it contains all of its limit points.

Proof. Suppose a set A is closed. If it has no limit points, there is nothing to check as it trivially contains its limit points. Now suppose z is a limit point of A . Then if $z \in A$, it contains this limit point. So suppose for the sake of contradiction that z is a limit point and z is not in A . Now we have assumed A was closed, so its complement is open. Since z is not in A , it is in the complement of A , which is open; which means there is an open set U containing z contained in the complement of A . This contradicts that z is a limit point because a limit point is, by definition, a point such that every open set about z meets A . Conversely, if A contains all its limit points, then its complement is open. Suppose x is in the complement of A . Then it can not be a limit point (by the assumption that A contains all of its limit points). So x is not a limit point which means we can find some open set around x that doesn't meet A . This proves the complement is open, i.e. every point in the complement has an open set around it that avoids A . \square

Example 1.5.2. *Since we know the empty set is open, X must be closed.*

Example 1.5.3. *Since we know that X is open, the empty set must be closed.*

Therefore, both the empty set and X are open and closed.

1.6 Continuity

In topology a continuous function is often called a map. There are 2 different ideas we can use on the idea of continuous functions.

Calculus Style

Definition 1.6.1. *$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuous if for every $\epsilon > 0$ there exists $\delta > 0$ such that when $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \epsilon$.*

The map is continuous if for any small distance in the pre-image an equally small distance is apart in the image. That is to say the image does not “jump.”

Topology Style

In topology it is necessary to generalize down the definition of continuity, because the notion of distance does not always exist or is different than our intuitive idea of distance.

Definition 1.6.2. *A function $f : X \rightarrow Y$ is continuous if and only if the pre-image of any open set in Y is open in X .*

If for whatever reason you prefer closed sets to open sets, you can use the following equivalent definition:

Definition 1.6.3. *A function $f : X \rightarrow Y$ is continuous if and only if the pre-image of any closed set in Y is closed in X .*

Let us give one more definition and then some simple consequences:

Definition 1.6.4. Given a point x of X , we call a subset N of X a neighborhood of x if we can find an open set O such that $x \in O \subseteq N$.

1. A function $f : X \rightarrow Y$ is continuous if for any neighborhood V of Y there is a neighborhood U of X such that $f(U) \subseteq V$.
2. A composition of 2 continuous functions is continuous.

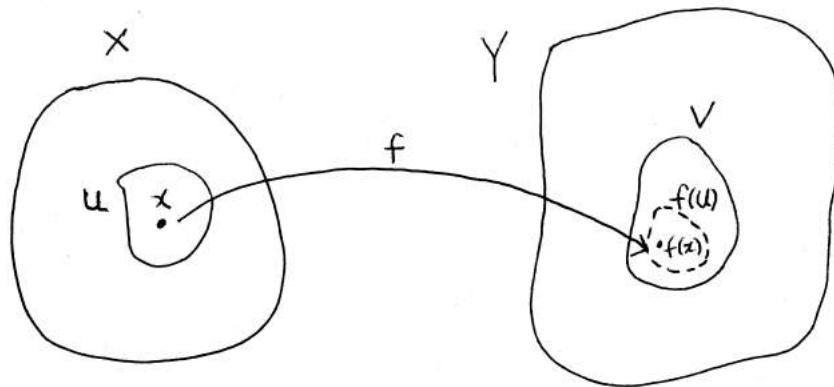


Figure 1.1: Continuity of f with neighborhoods.

1.7 Homeomorphisms

Homeomorphism is the notion of equality in topology and it is a somewhat relaxed notion of equality. For example, a classic example in topology suggests that a doughnut and coffee cup are indistinguishable to a topologist. This is because one of the geometric objects can be stretched and bent continuously from the other.

The formal definition of homeomorphism is as follows.

Definition 1.7.1. A homeomorphism is a function $f : X \rightarrow Y$ between two topological spaces X and Y that

- is a continuous bijection, and
- has a continuous inverse function f^{-1} .



Figure 1.2: Homeomorphism between a doughnut and a coffee cup.

Another equivalent definition of homeomorphism is as follows.

Definition 1.7.2. *Two topological spaces X and Y are said to be homeomorphic if there are continuous map $f : X \rightarrow Y$ and $g : Y \rightarrow X$ such that*

$$f \circ g = I_Y \quad \text{and} \quad g \circ f = I_X.$$

Moreover, the maps f and g are homeomorphisms and are inverses of each other, so we may write f^{-1} in place of g and g^{-1} in place of f .

Here, I_X and I_Y denote the identity maps.

When inspecting the definition of homeomorphism, it is noted that the map is required to be continuous. This means that points that are “close together” (or within an ϵ neighborhood, if a metric is used) in the first topological space are mapped to points that are also “close together” in the second topological space. Similar observation applies for points that are far apart.

As a final note, the homeomorphism forms an **equivalence relation** on the class of all topological spaces.

- *Reflexivity:* X is homeomorphic to X .
- *Symmetry:* If X is homeomorphic to Y , then Y is homeomorphic to X .
- *Transitivity:* If X is homeomorphic to Y , and Y is homeomorphic to Z , then X is homeomorphic to Z .

The resulting equivalence classes are called *homeomorphism classes*.

1.8 Homeomorphism Examples

Several examples of this important topological notion will now be given.

Example 1.8.1. Any open interval of \mathbb{R} is homeomorphic to any other open interval. Consider $X = (-1, 1)$ and $Y = (0, 5)$. Let $f : X \rightarrow Y$ be

$$f(x) = \frac{5}{2}(x + 1).$$

Observe that f is bijective and continuous, being the compositions of addition and multiplication. Moreover, f^{-1} exists and is continuous:

$$f^{-1}(x) = \frac{2}{5}x - 1.$$

Note that neither $[0, 1]$ nor $[0, 1)$ is homeomorphic to $(0, 1)$ as such mapping between these intervals, if constructed, will fail to be a bijection due to endpoints.

Example 1.8.2. There exists homeomorphism between a bounded and an unbounded set. Suppose

$$f(x) = \frac{1}{x}.$$

Then it follows that $(0, 1)$ and $(1, \infty)$ are homeomorphic. It is interesting that we are able to “stretch” a set to infinite length.

Example 1.8.3. Any open interval is, in fact, homeomorphic to the real line. Let $X = (-1, 1)$ and $Y = \mathbb{R}$. From the previous example it is clear that the general open set (a, b) is homeomorphic to $(-1, 1)$. Now define a continuous map $f : (-1, 1) \rightarrow \mathbb{R}$ by

$$f(x) = \tan\left(\frac{\pi x}{2}\right).$$

This continuous bijection possesses a continuous inverse $f^{-1} : \mathbb{R} \rightarrow (-1, 1)$ by

$$f^{-1}(x) = \frac{2}{\pi} \arctan(x).$$

Hence $f : (-1, 1) \rightarrow \mathbb{R}$ is a homeomorphism.

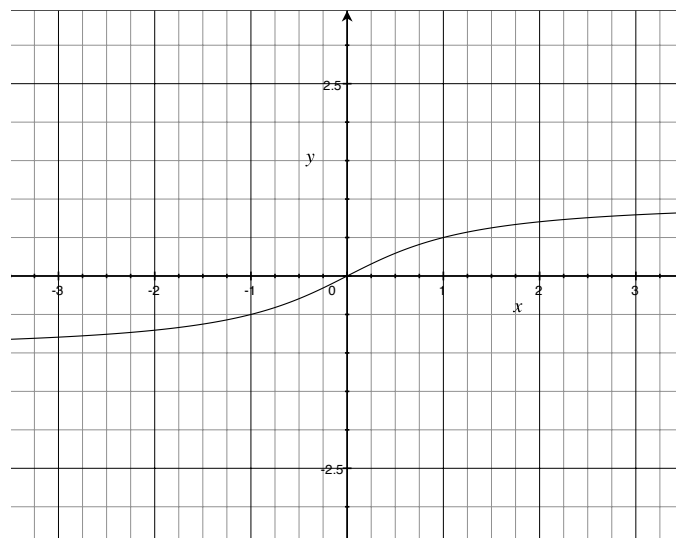
Example 1.8.4. A topologist cannot tell the difference between a circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ and a square $T = \{(x, y) \in \mathbb{R}^2 \mid |x| + |y| = 1\}$, as there is a function $f : S^1 \rightarrow T$ defined by

$$f(x, y) = \left(\frac{x}{|x| + |y|}, \frac{y}{|x| + |y|} \right)$$

which is continuous, bijective, and has a continuous inverse

$$f^{-1}(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right).$$

Both circle and square are therefore topologically identical. S^1 and T are sometimes called simple closed curves (or Jordan curves).

Figure 1.3: The graph of $f^{-1}(x)$.

Example 1.8.5. S^1 with a point removed is homeomorphic with \mathbb{R} . Without loss of generality, suppose we removed the North Pole. Then the stereographic projection is a homeomorphism between the real line and the remaining space of S^1 after a point is omitted.

Place the circle “on” the x -axis with the point omitted being directly opposite the real line. More precisely, let $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + (y - 1/2)^2 = 1/4\}$ and suppose the North Pole is $N = (0, 1)$. Using geometry, we may construct $f : S^1 \setminus N \rightarrow \mathbb{R}$ by defining

$$f(x, y) = \frac{2x}{1 - y}.$$

f is well-defined and continuous as the domain of f excludes $y = 1$, i.e. the North Pole. With the continuous inverse function

$$f^{-1}(x) = \left(\frac{4x}{x^2 + 4}, \frac{x^2 - 4}{x^2 + 4} \right),$$

we have $f \circ f^{-1} = f^{-1} \circ f = I$, hence f is a homeomorphism.

Stereographic projection is a mapping that plays pivotal roles in cartography (mapmaking) and geometry. One interesting properties of stereographic projection is that this mapping is conformal, namely it preserves the angles at which curves cross each other.

Example 1.8.6. The annulus $A = \{(x, y) \in \mathbb{R}^2 \mid 1 \leq x^2 + y^2 \leq 4\}$ is homeomorphic to the cylinder $C = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = 1, 0 \leq z \leq 1\}$ since there exists continuous function

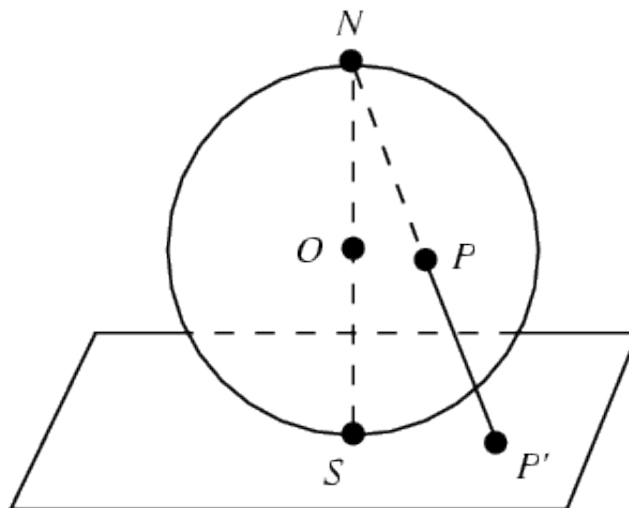


Figure 1.4: Stereographic projection of $S^2 \setminus N$ to \mathbb{R}^2 .

$f : C \rightarrow A$ and $g : A \rightarrow C$

$$f(x, y, z) = ((1 + z)x, (1 + z)y) \quad \text{and} \quad g(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}}, \sqrt{x^2 + y^2} - 1 \right)$$

such that $f \circ g = g \circ f = I$. Thus f and g are homeomorphisms.

1.9 Theorems On Homeomorphism

Theorem 1.9.1. *If $f : X \rightarrow Y$ is a homeomorphism and $g : Y \rightarrow Z$ is another homeomorphism, then the composition $g \circ f : X \rightarrow Z$ is also a homeomorphism.*

Proof. We need to establish that $g \circ f$ is continuous and has a continuous inverse for it to be a homeomorphism. The former is obvious as the composition of two continuous maps is continuous.

Since f and g are homeomorphisms by hypothesis, there exists $f^{-1} : Y \rightarrow X$ and $g^{-1} : Z \rightarrow Y$ that are continuous. By the same token, $f^{-1} \circ g^{-1}$ is continuous and

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ g^{-1} = I_Y.$$

Similarly $(f \circ g) \circ (g^{-1} \circ f^{-1}) = I_X$. Therefore $g \circ f$ is a homeomorphism. □

Various topological invariants will be discussed in Chapter 3, including connectedness, compactness, and Hausdorffness. Recall that homeomorphic spaces share *all* topological invariants. Hence the following theorem (without proof) immediately follows:

Theorem 1.9.2. *Let X and Y be homeomorphic topological spaces. If X is connected or compact or Hausdorff, then so is Y .*

1.10 Homeomorphisms Between Letters of Alphabet

Letters of the alphabet are treated as one-dimensional objects. This means that the lines and curves making up the letters have zero thickness. Thus, we cannot deform or break a single line into multiple lines. Homeomorphisms between the letters in most cases can be extended to homeomorphisms between \mathbb{R}^2 and \mathbb{R}^2 that carry one letter into another, although this is not a requirement – they can simply be a homeomorphism between points on the letters.

1.10.1 Topological Invariants

Before classifying the letters of the alphabet, it is helpful to get a feel for the topological invariants of one-dimensional objects in \mathbb{R}^2 . The number of 3-vertices, 4-vertices, (n -vertices, in fact, for $n \geq 3$), and the number of holes in the object are the topological invariants which we have identified.

1.10.2 Vertices

The first topological invariant is the number and type of vertices in an object. We think of a vertex as a point where multiple curves intersect or join together. The number of intersecting curves determines the vertex type.

Definition 1.10.1. *An n -vertex in a subset L of a topological space S is an element $v \in L$ such that there exists some neighborhood $N_0 \subseteq S$ of v where all neighborhoods $N \subseteq N_0$ of v satisfy the following properties:*

1. $N \cap L$ is connected.
2. The set formed by removing v from $N \cap L$, i.e., $\{a \in N \cap L \mid a \neq v\}$, is not connected, and is composed of exactly n disjoint sets, each of which is connected.

The definition of connectedness will be discussed much more in the course and will not be given here. Intuitively, though, a set is connected if it is all in one piece. So the definition of an n -vertex given above means that when we get close enough to an n -vertex, it looks like one piece, and if we remove the vertex from that piece, then we get n separate pieces.

We can say that for a given $n \geq 3$, the number of n -vertices is a topological invariant because homeomorphisms preserve connectedness. Thus, the connected set around a vertex must map to another connected set, and the set of n disjoint, connected pieces must map to another set of n disjoint connected pieces.

The number of 2-vertices is not a useful topological invariant. This is true because every curve has an infinite number of 2-vertices (every point on the curve not intersecting another curve is a 2-vertex). This does not help us to distinguish between classes of letters.

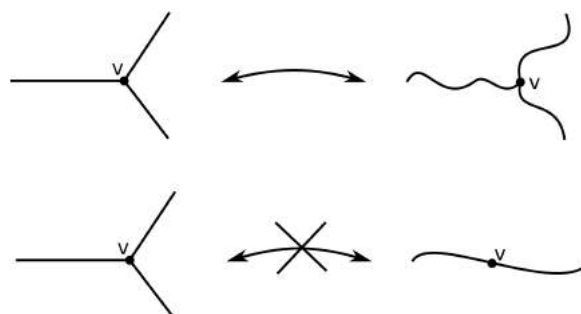


Figure 1.5: 3-vertex homeomorphism example

Three curves intersecting in a 3-vertex are homeomorphic to any other three curves intersecting in a 3-vertex. However, they are not homeomorphic to a single curve. This point is illustrated in Figure 1.5.

Similarly, any set close to a 4-vertex is homeomorphic to any other set close to a 4-vertex. In this case, “close” means a small neighborhood in which there are no other n -vertices ($n \geq 3$). Figure 1.6 shows an example of this.

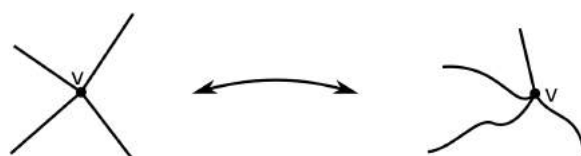


Figure 1.6: 4-vertex homeomorphism example

1.10.3 Holes

It is very clear when a one dimensional shape in two dimensional space has a hole: a shape possessing a hole closes in on itself. As a result, a topological structure with a hole cannot be continuously shrunk to a single point. Moreover, a function mapping a space with a hole to one without a hole cannot be a homeomorphism. To see this point, we note that the tearing of the circle means that nearby points on the circle can be mapped to very distant points on the line. This point, which is illustrated by figure 5, violates the continuity requirements of a homeomorphism. The number of holes is therefore an important topological invariant when classifying letters of the alphabet.

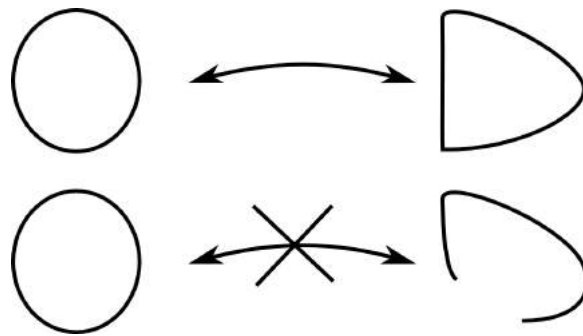


Figure 1.7: A loop is homeomorphic to any other loop (top), but not to a path with endpoints (bottom)

1.11 Classification of Letters

We now use the preceding arguments to classify the letters of the alphabet based on their topological invariants. We must use the letters printed as in the project specifications since changing the font can change the topological invariants of the letter.

0 holes, 0 three-vertices, 0 four-vertices:	C, G, I, J, L, M, N, S, U, V, W, Z
1 hole, 0 three-vertices, 0 four-vertices:	D, O
0 holes, 1 three-vertex, 0 four-vertices:	E, F, T, Y
1 hole, 1 three-vertex, 0 four-vertices:	P
0 holes, 2 three-vertices, 0 four-vertices:	H, K
1 hole, 2 three-vertices, 0 four-vertices:	A, R
2 holes, 2 three-vertices, 0 four-vertices:	B
0 holes, 0 three-vertices, 1 four-vertex:	X
1 hole, 0 three-vertices, 1 four-vertex:	Q

Table 1.1: Classification of the letters using the topological invariants

None of the letters possess a n -vertex with n greater than five, so our classification scheme is based on the number of holes, the number of 3-vertices, and the number of 4-vertices in the letter.

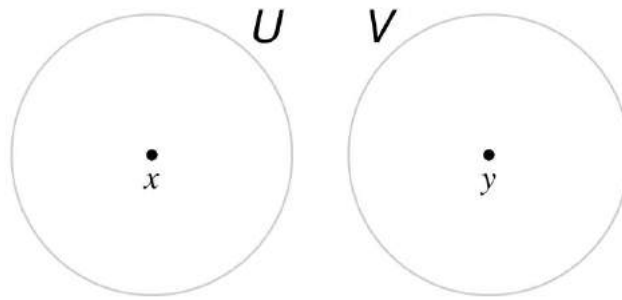
1.11.1 The curious case of the “Q”

It is interesting to consider the question of whether there exists a homeomorphism of \mathbb{R}^2 that carries \mathbb{Q} homeomorphically to \mathbb{Q} .

Well, there is not!

Intuitively, this is because such a function would need to map some points on the inside of the loop to points on the outside of the loop. This means that every point in the interior of the loop must map outside of the loop, since otherwise the image of one of the interior points would have to be on the loop in \mathbb{Q} , by continuity, which would not be an invertible operation since that point in the image would have two points in the domain mapping to it. An analogous argument for the inverse function means that points on the outside of the loop in \mathbb{Q} must all map to the inside of \mathbb{Q} , which does not allow for there to be one segment inside and one segment outside.

We *can*, however, define a homeomorphism $f : \mathbb{Q} \rightarrow \mathbb{Q}$, where these symbols represent only the subsets of \mathbb{R}^2 included in each figure. It is defined simply by mapping $q \in \mathbb{Q}$ to the same point on the loop if q is on the loop, to the lower leg in \mathbb{Q} if q is on the inner leg in \mathbb{Q} , and to the upper leg in \mathbb{Q} if q is on the outer leg in \mathbb{Q} . The mapping on the legs would preserve the ratio of

Figure 1.8: Neighborhoods U and V .

the distance along the leg from the 4-vertex versus the total length of the leg, so that it would be continuous at the 4-vertex.

1.12 Topological Invariants

A *topological invariant* of a space X is a property that depends solely on the topology of the space X . That is, a property shared by any other space that is homeomorphic to X . Intuitively, a homeomorphism between X and Y maps points in X that are “close together” to points in Y that are “close together”, and points in X not “close together” to points in Y that are not “close together”.

Below we will give a brief introduction to some of the topological invariants encountered during the course, and also give some simple examples to let the readers gain an insight on what these invariants are. (For more rigorous mathematical formulation, please refer to Chapter 3)

1.12.1 Hausdorff Property

Definition 1.12.1. If X is a topological space and S is a subset of X then a neighbourhood of S is a set V , which contains an open set U containing S . i.e. $S \subseteq U \subseteq V \subseteq X$

Definition 1.12.2. Let X be a topological space. Let $x, y \in X$. We say that x and y can be separated by neighbourhoods if there exists a neighbourhood U of x and a neighbourhood V of y such that U and V are disjoint i.e. $U \cap V = \emptyset$.

Definition 1.12.3. A space is **Hausdorff** if every two distinct points have disjoint neighbourhoods.

Example 1.12.4. A topological space is called *T1* if for any pairs of point $x, y \in X$, there is an open set O_x such that $x \in O_x$ and y isn't. Hausdorff are *T1* but the vice versa is not always true. A simple example is an infinite set endowed with the cofinite topology.

A nice property of Hausdorff spaces is that compact sets are always closed. This may fail for spaces which are non-Hausdorff. The definition of a Hausdorff space says that points can be separated by neighborhoods. It turns out that this implies something which is seemingly stronger:

in a Hausdorff space every pair of disjoint compact sets can be separated by neighborhoods. This is an example of the general rule that compact sets often behave like points.

1.12.2 Compactness Property

Definition 1.12.5. *A space is compact if every open cover has a finite subcover. (This is equivalent to “closed” and “bounded” in an Euclidean Space.)*

One of the main reasons for studying compact spaces is because they are in some ways very similar to finite sets: there are many results which are easy to show for finite sets, whose proofs carry over with minimal change to compact spaces.

Example 1.12.6. *Suppose X is a Hausdorff space, and we have a point x in X and a finite subset A of X not containing x . Then we can separate x and A by neighbourhoods: for each a in A , let $U(x)$ and $V(a)$ be disjoint neighbourhoods containing x and a , respectively. Then the intersection of all the $U(x)$ and the union of all the $V(a)$ are the required neighbourhoods of x and A .*

Note that if A is infinite, the proof fails, because the intersection of arbitrarily many neighbourhoods of x might not be a neighbourhood of x . The proof can be “rescued”, however, if A is compact: we simply take a finite subcover of the cover $V(a)$ of A . In this way, we see that in a Hausdorff space, any point can be separated by neighbourhoods from any compact set not containing it. In fact, repeating the argument shows that any two disjoint compact sets in a Hausdorff space can be separated by neighbourhoods.

Here we give some theorems related to compactness:

Theorem 1.12.7. *If f is a continuous function, then the image of a compact set under f is also compact.*

Note the above theorem implies that compactness is a topological invariant.

Theorem 1.12.8. *Any closed subset C of a compact space K is also compact.*

Given a non-compact set X , sometimes one would like to construct a homeomorphism to an open dense set in a compact space \tilde{X} . Intuitively, it is like adding some points to X to make it compact.

Theorem 1.12.9. *For any non-compact space X the one-point compactification of X is obtained by adding one extra point ∞ and defining the open sets of the new space to be the open sets of X together with the sets of the form $G \cup \infty$, where G is an open subset of X such that $X \setminus G$ is compact.*

Example 1.12.10. *Any locally compact (every point is contained in a compact neighbourhood) Hausdorff space can be turned into a compact space by adding a single point to it, by means of one-point compactification. The one-point compactification of \mathbb{R} is homeomorphic to the circle S^1 ; the one-point compactification of \mathbb{R}^2 is homeomorphic to the sphere S^2 . Using the one-point compactification, one can also easily construct compact spaces which are not Hausdorff, by starting with a non-Hausdorff space.*

1.12.3 Connectedness and Path Connectedness Properties

Definition 1.12.11. A space X is connected if it is not the union of a pair of disjoint non-empty open sets. Equivalently, a space is connected if the only sets that are simultaneously open and closed are the whole space and the empty set.

Another equivalent formal definition is:

Definition 1.12.12. A topological space X is said to be disconnected if it is the union of two disjoint nonempty open sets. Otherwise, X is said to be connected.

Example 1.12.13. The closed interval $[0, 2]$ is connected; it can, for example, be written as the union of $[0, 1)$ and $[1, 2]$, but the second set is not open in the topology of $[0, 2]$. On the other hand, the union of $[0, 1)$ and $(1, 2]$ is disconnected; both of these intervals are open in the topological space $[0, 1) \cup (1, 2]$.

A major theorem of connectedness is:

Theorem 1.12.14. Let X and Y be topological space. If $f: X \rightarrow Y$ is a continuous function, then the image of a connected subset $C \subseteq X$ under f is also connected.

Note the above theorem implies that connectedness is a topological invariant. It can also be considered a generalization of the intermediate value theorem.

Connectedness is one of the principal topological properties that is used to distinguish topological spaces. A stronger notion is that of a path-connected space, which is a space where any two points can be joined by a path.

Definition 1.12.15. The space X is said to be path-connected if for any two points $x, y \in X$ there exists a continuous function f from the unit interval $[0, 1]$ to X with $f(0) = x$ and $f(1) = y$. (This function is called a path from x to y .)

Theorem 1.12.16. A path-connected space is connected.

Note the vice-versa of the above theorem is not always true.

Example 1.12.17. Consider the space defined as the graph of the function $\sin(1/x)$ over the interval $(0, 1]$ extended by the single point $(0, 0)$. This set is then equipped with the topology induced from the Euclidean plane. It is connected but not path-connected. It is the continuous image of a locally compact space (namely, let V be the space $-1 \cup (0, 1]$, and use the map f from V to T defined by $f(-1) = (0, 0)$ and $f(x) = (x, \sin(1/x))$), but is not locally compact itself.

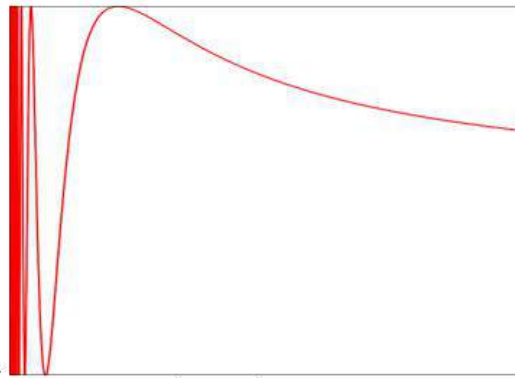


Figure 1.9: Image of the curve $\sin\left(\frac{1}{x}\right)$.

Chapter 2

Making New Spaces From Old

In Chapter 1 we learned the technical definition of the *topological spaces* - which are in some sense generalizations of *metric spaces* - that will be the object of study for the rest of these notes. Moreover, we saw how to construct and interpret functions $f : X \rightarrow Y$ that map points in one topological space X to points in another space Y . These maps, which may or may not be *continuous*, are therefore relationships between spaces and it is typically impossible to say how the *topological properties* of the domain space X and the image space Y are related. When these spaces are deformed continuously via a *homeomorphism*, however, such properties are *topological invariants*. This important result gives an approach towards determining whether or not spaces are topologically “different”. We now move on to describing methods that can be used to create new topological spaces that we will study further in future chapters. Interesting connections between old and familiar spaces will also be drawn in this chapter.

As we shall see, this subject is interesting and valuable for a number of reasons. For example, there simply is aesthetic beauty in building up complex structures and spaces from simple ones. But on a more fundamental level, many properties of the initial and final spaces are related when these techniques are applied. In particular, we consider *product spaces*, *quotient spaces* and *group actions*, and *identification spaces*.

2.1 Cartesian Products of Spaces

Let us forget about topology for a moment, take a step backwards, and consider the function $f(x, y) = xy$. When we look at this function, we typically think of its behavior as the variables x and y are varied separately. That is, for a fixed value of x we can observe the change in the function’s value as we vary y , and vice-versa. In this sense we naturally “factor” the function $f(x, y)$.

We can apply similar intuition to certain topological spaces that can be “factored” as the product $X \times Y$. Here, points in the product space $X \times Y$ are represented by Cartesian pairs (x, y) where $x \in X$ and $y \in Y$. The spaces X and Y are typically “simpler”, more familiar, or more easily visualized spaces, though this is not necessarily the case.

There exists ample examples of decomposing familiar topological spaces as product spaces. The plane \mathbb{R}^2 , for example, can be written as the product of two copies of the real line \mathbb{R} or $\mathbb{R} \times \mathbb{R}$.

The definition of Cartesian products certainly agrees with our picture of representing points as having an x-coordinate (from the first \mathbb{R} in $\mathbb{R} \times \mathbb{R}$) and a y-coordinate (from the second \mathbb{R} in $\mathbb{R} \times \mathbb{R}$). Also, the surface of a cylinder can be thought of as a circle extruded down the length of an interval. We can therefore represent the cylinder as $Cyl = S^1 \times I$, where S^1 is a circle in the (x,y) plane and I is an interval along the z axis. Likewise, the torus can be interpreted as one small circle extruded around the circumference of a second larger circle and hence written as $T = S^1 \times S^1$.

Rather exotic spaces can also be built out of familiar ones by taking products. Ultimately there are an infinitely large number of ways to combine such spaces as the disk, the sphere, the torus, and the real projective plane, but examples include the spaces $\mathbb{R}P^2 \times \mathbb{R}P^2 \times T$ and $\mathbb{R}P^2 \times \dots \times \mathbb{R}P^2$. While these may have little meaning to us as we cannot picture them in our minds or draw them on paper, they nevertheless represent real objects that can be considered.

2.2 The Product Topology

We are therefore armed with a mechanism for “combining” spaces using products. Our next step must therefore be to extend this picture to include open sets and topology. In order to make this intellectual leap, we can observe that if U and V are open sets in the component spaces X and Y , respectively, then, if the world is at all fair, $U \times V$ should be an open set in the product $X \times Y$. Making this a requirement means that, just as we can build-up arbitrary open sets in X adding open sets in a base for its topology, we can likewise build-up arbitrary open sets in $X \times Y$ as the union of open sets of the form $U \times V$. The collection of such sets $U \times V$ can be called the family \mathcal{B} , and the finite intersection of members in \mathcal{B} also lies in \mathcal{B} . The collection \mathcal{B} is therefore a *base* for a topology on $X \times Y$ that is appropriately named the *product topology*. When the set $X \times Y$ is paired with this topology, it is referred to as a *product space*.

Now that we have a notion of the product topology, it is possible to consider the relationship between open sets in a product space and those in the component spaces. As such, we define the *projections* $p_X : X \times Y \rightarrow X$ with $p_X(x, y) = x$ and $p_Y : X \times Y \rightarrow Y$ with $p_Y(x, y) = y$. This allows us to describe the product topology in terms of the projections.

Theorem 2.2.1. *When $X \times Y$ is equipped with the product topology, the projections p_X and p_Y are continuous functions and are open maps. In addition, the product topology on $X \times Y$ is the smallest topology for which the projections are continuous.*

This important result is true because open boxes $U \times V$ in the product space are, by definition, mapped to open sets U and V in the component spaces, respectively. In a reciprocal fashion, it is clear that the pre-image of open sets U, V in the component spaces are open strips $U \times Y, V \times X$ in the product space. Moreover, any given topology on a product space must necessarily contain the open sets $U \times V$ and is therefore at least as large as the product topology.

Another general result is that we can check that the continuity of a function f mapping component spaces to product spaces by composing it with the projections and checking the continuity of the resulting functions.

Theorem 2.2.2. *The function $f : Z \rightarrow X \times Y$ is continuous if and only if the composite functions $p_X f : Z \rightarrow X$ and $p_Y f : Z \rightarrow Y$ are continuous.*

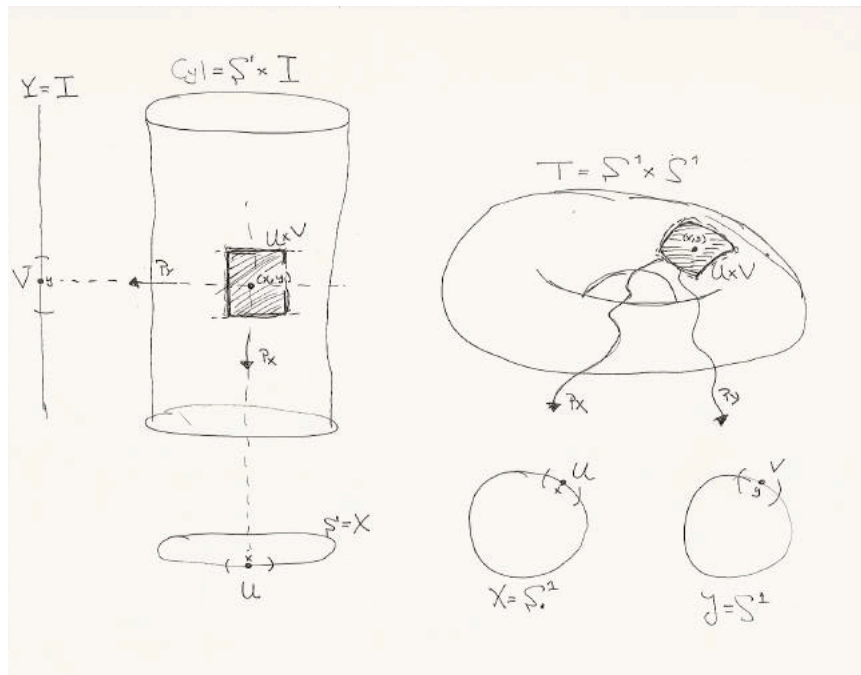


Figure 2.1: As discussed in the notes, the cylinder (left) and the torus (right) can be represented as product spaces. Notice the graphical relationship between open sets in the component spaces X and Y and those in $X \times Y$, which is crucial for developing the product topology.

These notions of product spaces and topology can be made more concrete by comparing these definitions with figure one and its examples.

2.3 Properties of Product Spaces

There are also a number of nice theorems that allow us to relate the topological properties of product spaces $X \times Y$ to those of the component spaces X and Y . While these theorems will be discussed in detail later in these notes, we nevertheless state them here.

Theorem 2.3.1. *The product space $X \times Y$ is a Hausdorff space if and only if both X and Y are Hausdorff.*

Theorem 2.3.2. *The product space $X \times Y$ is compact if and only if both X and Y are compact.*

Theorem 2.3.3. *If X and Y are connected spaces then the product space $X \times Y$ is also connected.*

It is also clear that we can extend the results shown below to products of any number of such as spaces, as in the case of $X_1 \times X_2 \times \dots \times X_n$. This is true since we can simply make the substitutions $X \mapsto X_1$ and $Y \mapsto X_2 \times \dots \times X_n$ in these theorems and make an inductive argument.

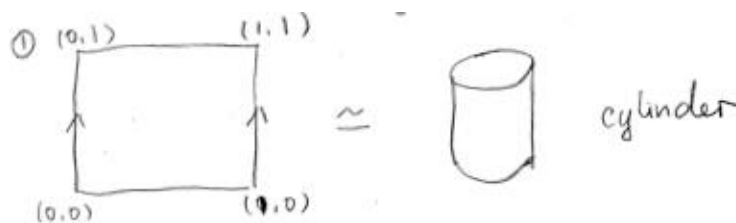
2.4 Identification Spaces

Before introducing Identification spaces, let's first define a *partition*. Let X be a topological space and let \mathcal{P} be a family of disjoint nonempty subsets of X such that $\bigcup P = X$. Such a family is usually called a partition of X . That is, partitions of X are disjoint subsets of X that together form X . For a trivial example, let elements of \mathcal{P} be the points of X so that $X = \bigcup P$.

We may form a new space Y , called an *identification space*, as follows. The points of Y are the members of \mathcal{P} and, if $\pi : X \rightarrow Y$ sends each point of X to the subset of \mathcal{P} containing it, the topology of Y is the largest for which π is continuous. Therefore, a subset O of Y is open if and only if $\pi^{-1}(O)$ is open in X . This topology is called the *identification topology* on Y .

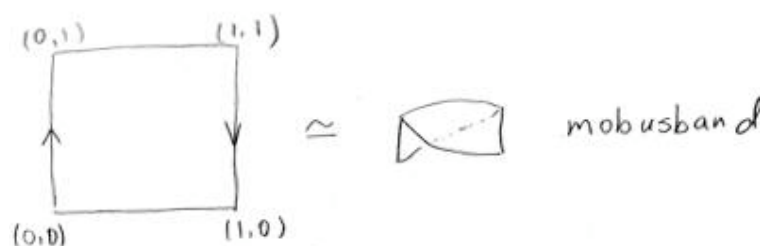
(We call Y an identification space, because all points in X that are sent to the same subset of Y have become the *same* point.)

Example 2.4.1. *Cylinder.*



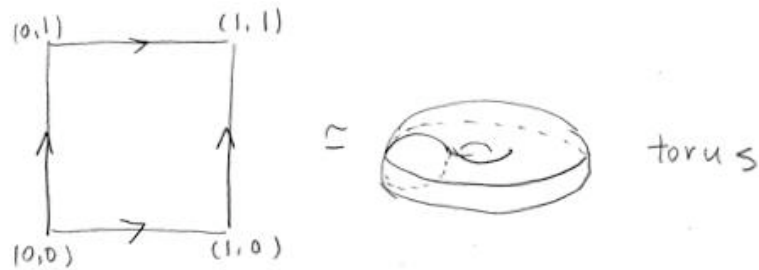
$$Y = \left[\bigcup_{a \neq 0,1} \{(a,b)\} \right] \bigcup_{0 \leq b \leq 1} \left[\bigcup \{(0,b), (1,b)\} \right].$$

Example 2.4.2. *Möbius band.*



Example 2.4.3. *Torus.*

$$Y = \left[\bigcup_{\substack{a \neq 0,1 \\ b \neq 0,1}} \{(a,b)\} \right] \bigcup_{0 < b < 1} \left[\bigcup \{(0,b), (1,b)\} \right] \bigcup_{0 < a < 1} \left[\bigcup \{(a,0), (a,1)\} \right] \bigcup \{(0,0), (0,1), (1,0), (1,1)\}.$$

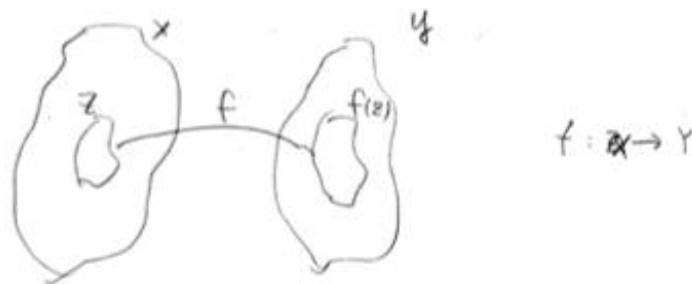


Fact: Suppose you have an identification map

$$\begin{array}{ccc}
 X & \xrightarrow{f} & Y \\
 \downarrow g \circ f & & \downarrow g \\
 & \longrightarrow & Z
 \end{array}$$

A function g is continuous if and only if $g \circ f$ is continuous.

Example 2.4.4. *Gluing 2 spaces along a function.*

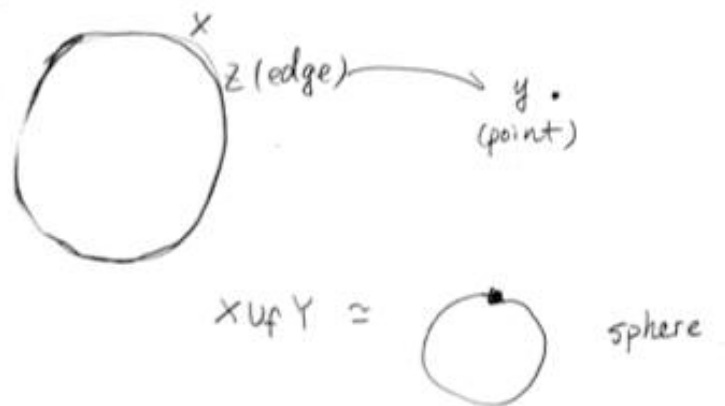


Create a partition of $X \cup Y$ as follows,

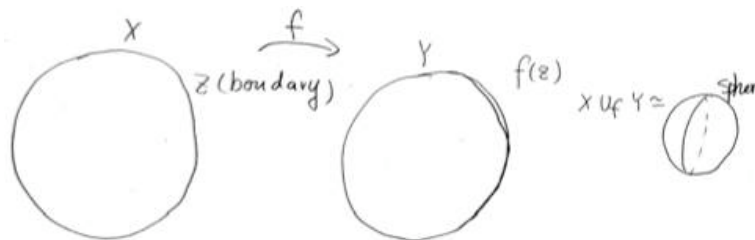
- If $x \in X \setminus Z \rightarrow \{x\}$.
- If $y \in Y \setminus f(z) \rightarrow \{y\}$.
- If $y \in f(z) \rightarrow \{x, y \mid f(x) = y\}$

The identification space obtained is denoted $X \cup_f Y$ read X glued along f to Y .

Example 2.4.5. *If we glue the boundary of the disk X to a point Y , as shown below.*



Example 2.4.6. *If we glue two disks along their boundaries, as shown below.*



Statement: A way to make a partition on a set is to define an *equivalence relation* on the set. Recall that the definition of equivalence is:

- Symmetric: $x \sim y \implies y \sim x$.
- Reflexive: $x \sim x$.
- Transitive: $x \sim y, y \sim z \implies x \sim z$.

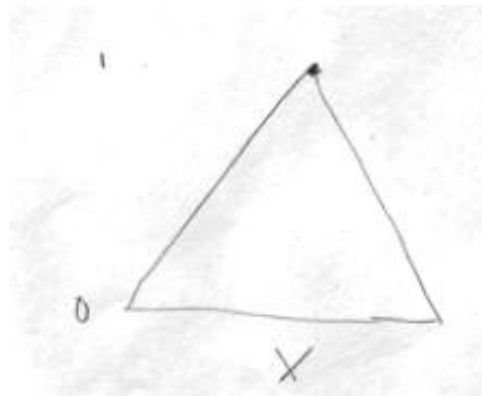
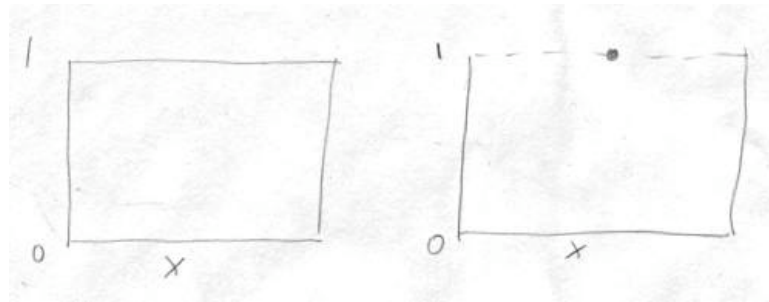
More examples on identification maps:

Example 2.4.7. *Cone*

Above we have $X \times (0, 1)$ on the left, if we identify the entire horizontal fiber $(x, 1)$ down to on point. $(x, 1 - \epsilon)$ is in a neighborhood of this point for all x , so we have the space, as below.

And from this transformation we can see why it is called as cone, the product space is compressed to one point as we approach $(x, 1)$.

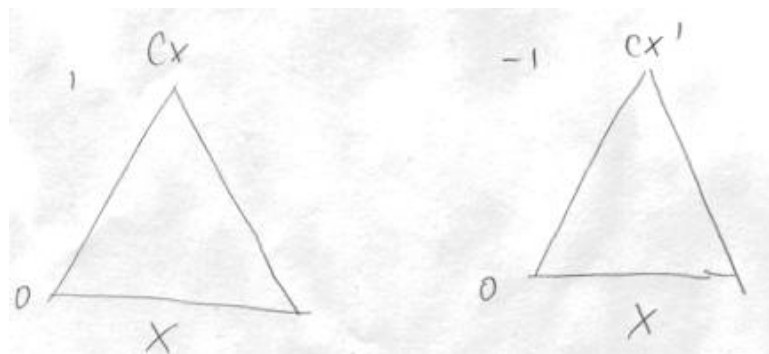
Example 2.4.8. *The suspension SX of X is the identification space obtained from $X \times (-1, 1)$ by the rule*



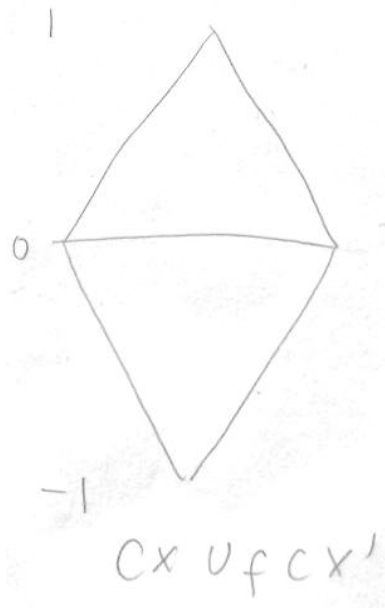
- All points of the form $(x, -1)$ are identified.
- All points of the form $(x, +1)$ are identified.
- All other points are left alone.

Then we claim that the suspension SX can be obtained by gluing two cones over X along a function.

Since the interval $(-1, 0)$ is homeomorphic to $(0, 1)$ we have two cones over each as below:



Now let's take the function $f : CX \rightarrow CX'$ taking $(x, 0)$ to $(x, 0)$ which is the identity map along the $(X, 0)$, and gluing along this function.



2.5 Group Actions and Quotient Spaces

It is also possible to create new spaces using *groups*. Before we go into the procedure, recall that a *group* is a set G equipped with a binary operation $*$ satisfying four properties.

1. G is closed under the operation $*$: if $x, y \in G$ then $x * y \in G$.
2. The operation for G is associative: if $x, y, z \in G$ then $x * (y * z) = (x * y) * z$.
3. G possesses an identity element e such that if $x \in G$, then $x * e = x$.
4. G possesses inverse elements x^{-1} such that if $x \in G$ then $x * x^{-1} = e$.

Examples of groups include the real numbers under either the usual addition or multiplication operations, as well as the integers under the addition operation.

If G is a group and X is a set or topological space, then the *group action* $G \times X \rightarrow X$ is given by $(g, x) \mapsto g \cdot x$. In addition, the group action must satisfy two additional properties.

1. The group action behaves well with respect to composition: $(gh) \cdot x = g(h \cdot x)$ for all $g, h \in G$ and $x \in X$.
2. The group action possesses an identity action: $e \cdot x = x$ for every $x \in X$.

These definitions of a group and a group action can be found in any standard algebra text.

The effect of the group action is therefore to partition a space into *orbits*, which are sets of points that can be connected by elements of the group. That is, the points x, y in the topological space X are in the same orbit if there is some $g \in G$ such that $g \cdot x = y$. Each orbit is therefore an

equivalence class under the group operation, and the space of such orbits is called a *quotient space*.

In this form, the group action and quotient spaces are somewhat abstract and it is not immediately obvious how we should use them to make interesting topological spaces. The concept of a group action is therefore best understood through examples such as the following.

Example 2.5.1. *Antipodal Action of the Sphere:* Consider the group action $\mathbb{Z}/2\mathbb{Z} \times S^2 \rightarrow S^2$, which corresponds to the quotient space $S^2/(\mathbb{Z}/2\mathbb{Z})$. The relatively uninteresting case of the zero element of $\mathbb{Z}/2\mathbb{Z}$ acting on the sphere maps each point on the sphere to itself. The one element of $\mathbb{Z}/2\mathbb{Z}$, however, identifies antipodal points on the sphere and creates lines through the origin in R^3 . But the space of real lines through the origin in R^3 is the real projective plane, as we shall see later in the notes.

Example 2.5.2. *Translational Action of the Real Line:* This can be written as the group action $\mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R}$, which corresponds to the quotient space \mathbb{R}/\mathbb{Z} . The action of \mathbb{Z} on the real line is to identify points that differ by some integer, and orbits in the real line are therefore the set of points that are separated by some integer n . This allows us to “compress” the real line to the unit interval with endpoints identified. As can be seen in the below figure, this quotient space is the circle C .

Example 2.5.3. *Translational Action of the Complex Plane:* This can be written as the group action $(\mathbb{Z} \times \mathbb{Z}) \times \mathbb{C} \rightarrow \mathbb{C}$, which corresponds to the quotient space $\mathbb{C}/(\mathbb{Z} \times \mathbb{Z})$. The action of $\mathbb{Z} \times \mathbb{Z}$ on the complex plane is to identify points that differ by some integer in both real and imaginary components. Orbits in the plane are therefore points that have both real and imaginary components differing by some integers m and n . This “compresses” the complex plane to a unit square with opposite sides identified, each with positive sense. As can be seen below, and as we shall see later in the notes, this quotient space is the torus T .

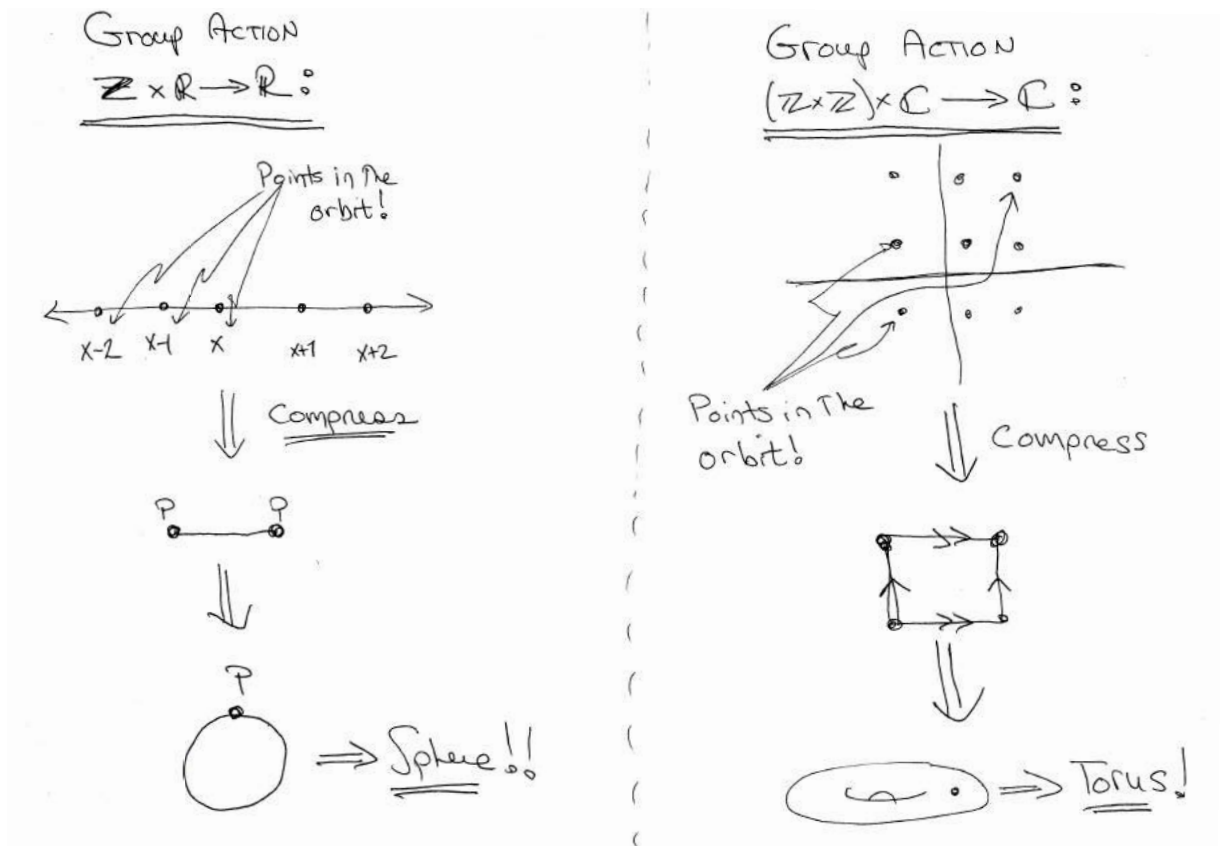


Figure 2.2: A cartoon of the effects of the translational action on the real line (left) and complex plane (right).

Chapter 3

First Topological Invariants

3.1 Introduction

Central to the study of topology is the idea of a homeomorphism between spaces. Topological invariants are special properties of spaces which are preserved under a homeomorphism. To this end, topological invariants are key in proving whether two spaces are not homeomorphic and useful as motivation toward finding homeomorphisms between spaces. Some of the fundamental topological invariants we will cover are compactness (a generalization of closed and bounded in \mathbb{R}^n to different topologies), connectedness (whether or not a space is a single piece), and Hausdorffness (whether or not there is space between points of a space).

3.2 Compactness

An intuitive way to think about the notion of compactness is to think of sets or spaces that are solid and small. In topology, compactness is a fundamental topological invariant because a space that is compact is generally easier to define and work with. Later on when the Classification Theorem is discussed, we will see that compact surfaces are special because they can be categorized as homeomorphic to one of just three different families of surfaces.

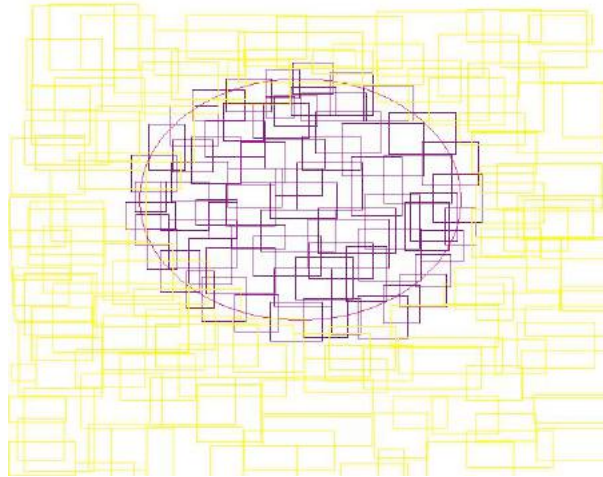
3.2.1 Preliminary Ideas

Definitions

Before formally defining what it means for a topological space to be compact, it will be helpful to introduce some useful terminology and go over some examples:

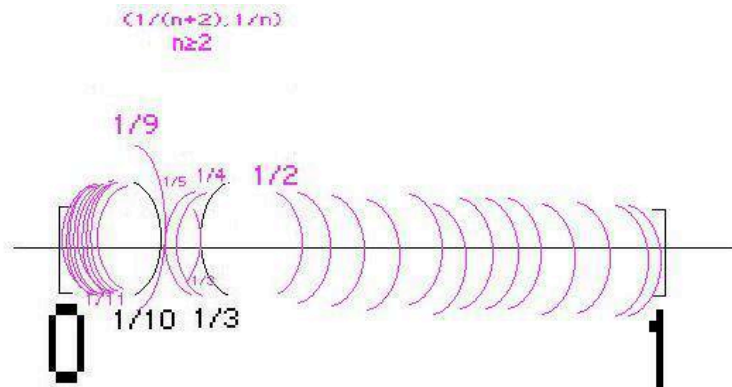
An **open cover** of a topological space X is a family of open subsets of X whose union is all of X . If \mathcal{F} is an open cover of X and \mathcal{F}' is a subfamily of \mathcal{F} , and if $\bigcup \mathcal{F}' = X$, then \mathcal{F}' is called a **subcover** of X . In the following diagram, X is represented by the ellipse, \mathcal{F} is repre-

sented by the union of the purple and yellow boxes, while \mathcal{F}' is represented by just the purple boxes:

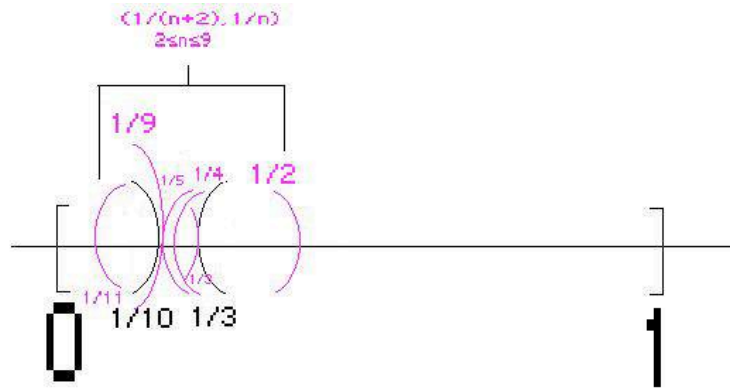


Examples of open covers and sub covers

- Let X be the closed unit interval $[0,1]$ (with the induced topology from the real line). Then $[0, 1/10)$; $(1/3, 1]$; and $(1/n+2, 1/n)$ where $n \in \mathbb{Z}$ and $n \geq 2$ is an infinite open cover of $[0,1]$.



Notice that we can take only a finite number of the open sets in the family above that still cover X , namely $[0, 1/10]$; $(1/3, 1]$; and $(1/n+2, 1/n)$ where $n \in \mathbb{Z}$ and $n \geq 2 \geq 9$



This finite number of open sets from the above family of sets therefore constitutes a finite subcover of F .

2. Open covers of the real line that do not contain a finite subcover

- $(n-1, n+1)$ with $n \in \mathbb{Z}$
- $(0, 1-1/n)$ with $n \in \mathbb{N}$
- $(-n-1, n+1)$ with $n \in \mathbb{N}$

3. An open cover of the plane

Let \mathcal{F} be the set of all open balls of radius 1 whose centers have integer coordinates, i.e., $\mathcal{F} = \{x, y \mid x^2 + y^2 \leq 1\}$. This family of open sets balls cover the plane, but if you remove any one ball, the family no longer covers the plane. Therefore this family of open sets has no proper subcover.

3.2.2 The Notion of Compactness

Now that we have an idea of what an open cover, subcover, and finite subcover of a topological space is, we can formally define what it means for a topological space to be compact.

Definition: A topological space X is compact if every open cover of X has a finite subcover.

Equivalent Definition: A topological space X is compact if for every collection of closed sets $\mathcal{C}_\alpha \in X$ such that $\bigcap \mathcal{C}_\alpha = \phi$, \exists a finite collection of closed subsets $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_n$ such that $\bigcap \mathcal{C}_n = \phi$.

The equivalent definition is true based on the first formal definition and using the fact that the complement of an open set is closed. Instead of a collection of open sets in X whose union is equal to X , we consider a collection of closed sets in X whose *intersection* is equal not to X but to ϕ , and instead of a finite subcover of open sets whose union is still equal to X , we look for a finite number of closed sets whose intersection is still equal to ϕ .

Examples of Compact Spaces

We can apply the above definition to come up with examples of compact sets in \mathbb{R}^n under different topologies:

1. Suppose \mathbb{R}^n has the discrete topology. Then every subset of \mathbb{R}^n is an open set. An example of a compact set in \mathbb{R}^n would be a set consisting of a finite number of points. This is because each point could be covered with a finite number disjoint open sets that are still in \mathbb{R}^n , so this could be a finite subcover of a family of open sets that cover the entire part of \mathbb{R}^n that contain the points. In fact, any finite subset of \mathbb{R}^n is a compact set in \mathbb{R}^n when \mathbb{R}^n has the discrete topology.
2. Suppose \mathbb{R}^n has the “stupid” (a.k.a trivial or indiscrete) topology. Then the only open sets in \mathbb{R}^n are all of \mathbb{R}^n and the empty set. Then any subset of \mathbb{R}^n is compact. This is because we have only one non empty open set!
3. Suppose \mathbb{R}^n has the finite complement topology. Then open sets are any sets with a complement that is finite. Then every open set is infinitely large and contains all but finitely many points of \mathbb{R}^n . Then closed sets of \mathbb{R}^n consist of finitely many points, so a finite subcover can always be found by covering each point one at a time with the open set whose finite complement is each of the other points in that closed set. Then the remaining sets in \mathbb{R}^n , the open sets, clearly have a finite subcover of open sets, namely themselves. Therefore, again any subset of \mathbb{R}^n with the finite complement topology is compact, including \mathbb{R}^n .
4. Note that finite subsets, under any topology, are always compact.
5. Suppose \mathbb{R}^n has the Euclidean topology. Then an example of a compact set in \mathbb{R}^1 , for example, would be $[0,1]$. This can be seen from the definition of a compact set and the example above where it is illustrated that this interval has a finite subcover.

In fact, when \mathbb{R}^n has the Euclidian topology, we have a special theorem to define all of its compact subsets.

A Generalization of the Heine-Borel Theorem

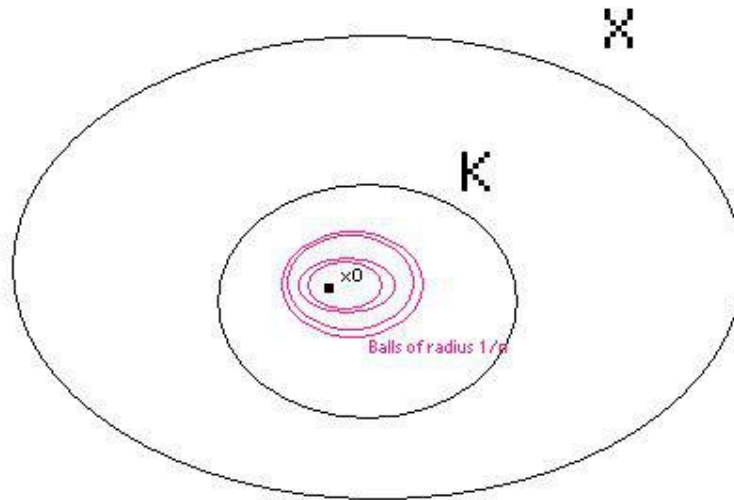
Theorem 3.2.1. *In \mathbb{R}^n with the Euclidean topology, compact sets are precisely the closed and bounded sets of \mathbb{R}^n .*

Proof:

Compact \Rightarrow Closed and Bounded:

Compact \Rightarrow Closed:

Let \mathcal{K} be a subset of \mathbb{R}^n that is compact but not closed, and let x_0 be a limit point of \mathcal{K} which is not in \mathcal{K} . Take an infinite collection of open subsets $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$ such that \mathcal{U}_1 consists of all points of distance 1 from x_0 , \mathcal{U}_2 consists of all points of distance $1/2$ from x_0 , \dots , \mathcal{U}_n consists of points of distance $1/n$ from x_0 . Then $\bigcup \mathcal{U}_n$ does not contain x_0 , but since it does contain the rest of \mathbb{R}^n , therefore it is a cover of \mathcal{K} . But notice that there is no possible way to extract a finite subcover because we chose x_0 to be a limit point.

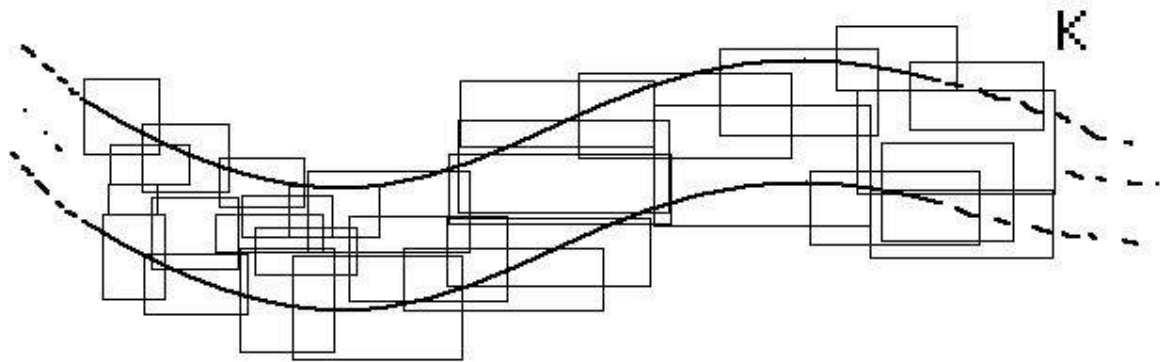


Compact \Rightarrow Bounded:

Suppose \mathcal{K} is compact but not bounded. Then for any point $x \in \mathcal{K}$, an open cover is \mathcal{U}_x where \mathcal{U}_x is a ball of radius 1 centered at x . Then $\bigcup_{x \in \mathcal{K}} \mathcal{U}_x$ covers all of \mathcal{K} . However, if \mathcal{K} is unbounded we cannot find a finite subcover of \mathcal{K} , contrary to our assumption that \mathcal{K} can be both not bounded and compact.

Closed and bounded \Rightarrow Compact:

Let \mathcal{K} be a closed and bounded subset of \mathbb{R}^n . Suppose that we can come up with an open cover of \mathcal{K} that does not admit any finite subcover, so that \mathcal{K} is not compact. Since \mathcal{K} is closed and bounded, \mathcal{K} can be split into a finite number of subsets, each of which is closed and bounded, but since \mathcal{K} does not admit a finite subcover, at least one of these subsets must not admit a finite subcover as well. Call this subset \mathcal{K}_1 . \mathcal{K}_1 can be divided into a finite number of closed and bounded subsets as well, and again, since \mathcal{K}_1 does not admit any finite subcover, at least one of the finite subsets of \mathcal{K}_1 must not admit a finite subcover. Choose one such subset and call this \mathcal{K}_2 . \mathcal{K}_2 can again similarly be subdivided in the same manner, and another closed and bounded subset that does not admit a finite subcover \mathcal{K}_3 can be chosen. This process can continue on in this manner and we get an infinite sequence of nested closed subsets such that



The above open sets cover K , but since K is not bounded, clearly none of them can be removed in such a way that their union still covers K .

the size of these sets approaches zero, and such that each of these sets does not admit a finite subcover. However, the intersection of all of these subsets is a point, which certainly can be covered by a finite subcover, a contradiction. Therefore the hypothesis must be wrong, and if K is closed and bounded then it must admit a finite subcover and therefore be compact.

The Heine-Borel Theorem states that any closed and bounded subset of the real line is a compact set. This theorem can be proven specifically. However, note that the above proof applies to \mathbb{R}^n in general, and so in particular applies to \mathbb{R}^1 as well. This therefore takes care of not only the Heine-Borel theorem, but to the equivalent of that theorem for \mathbb{R}^n in general.

3.3 Some Theorems on Compactness

Compact spaces have many interesting and useful properties.

Theorem 3.3.1. *The continuous image of a compact space is compact*

Proof: Suppose we have $f : X \rightarrow Y$ an onto continuous function with X compact. We must show Y is also compact. Take any open cover, \mathfrak{F} of Y . Since f is continuous

for each open set $O \in \mathfrak{F}$ $f^{-1}(O)$ is open in X . Define the set

$$\mathfrak{G} = \{f^{-1}(O) \mid \forall O \in \mathfrak{F}\}$$

Which is an open cover of X . X is compact, therefore \mathfrak{G} admits a finite subcover.

$$X = \bigcup_{i=1}^k f^{-1}(O_i)$$

Note $f(f^{-1}(O_i)) = O_i$ and since f is onto $Y = O_1 \cup O_2 \cup \dots \cup O_k$ and we have found a finite subcover from \mathfrak{F} . □

Notice that this automatically implies that compactness is preserved under homeomorphisms since homeomorphism are continuous bijections (with continuous inverses).

Theorem 3.3.2. *A closed subset of a compact set is also compact*

Proof: Take a compact space X with a closed subset C let \mathfrak{F} be a family of open subsets of X such that $C \subset \bigcup \mathfrak{F}$. Add the open set $U = X - C$ to \mathfrak{F} and we get a cover of X . X is compact so $\mathfrak{F} \cup U$ admits a finite subcover $O_1 \cup O_2 \cup \dots \cup O_k \cup (X - C)$. $C \subset O_1 \cup O_2 \cup \dots \cup O_k$ since $(X - C) \cap C = \emptyset$ and this new set $\{O_1, O_2, \dots, O_k\}$ is a finite subcover of C . □

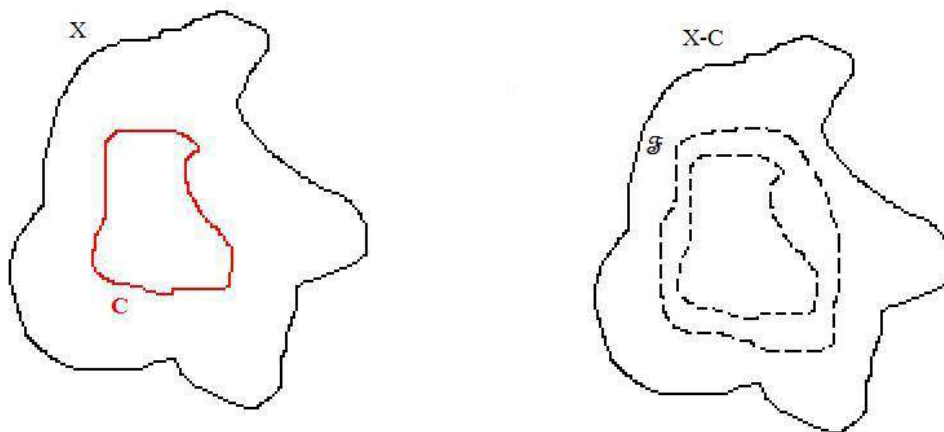


Figure 3.1: Illustration of **Theorem 2**

The basic idea behind the last proof is that since C is closed $(X - C)$ is ONE open set in X . Any open cover of X must admit a finite subcover \mathfrak{O} , if we remove one open set and restrict our attention to its complement we can still find a finite subcover. All we must do is take the original subcover \mathfrak{O} and remove the open set $(X - C)$.

Theorem 3.3.3. Bolzano-Weierstrass property: *An infinite subset of a compact space must have a limit point.*

Proof: Let X be a compact space with an infinite subset S which has no limit point. Our goal is to show that S is finite. We can find an open cover of X consisting of open neighborhoods $O(x)$ such that

$$O(x) \cap S = \begin{cases} \emptyset & \text{if } x \notin S \\ \{x\} & \text{if } x \in S \end{cases}$$

each of these $O(x)$ exist since otherwise x is a limit point of S . The open cover $\{O(x) | x \in X\}$ must admit a finite subcover, and since we defined $O(x)$ to contain no more than one point of S , S must be finite. \square

You should read this proof as $\forall x \in X$ take an open set around x , but make sure that if $x \in S$ this open set is small enough to contain only one point of S and if x is not in S make sure the open set is small enough to contain no points of S . Now together all of these open sets cover X and since X is compact we only need a finite number of these open sets, throw away all the rest. We are now left with a finite number of sets which contain AT MOST one point of S , hence S must be finite.

Now we have a convenient way to expose spaces which are not compact, look for an infinite subset which has no limit points. This is another way to discover that a compact subset of \mathbb{E}^n must be bounded, since if we let it run off to infinity it would be simple to find a subset which has no limit point.

For example take $X \in \mathbb{R}^2$; $X = \{(x, y) | -1 \leq y \leq 1\}$ a ribbon running along the x -axis shown in Figure 3.2

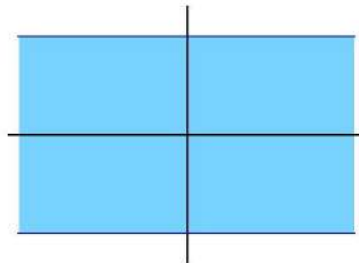


Figure 3.2: The set $X \in \mathbb{R}^2$

Of course the set $S = \{(x, 0) | x \in \mathbb{Z}\} \subset X$ and S has no limit points therefore X is not compact.

The next theorem also follows from the Bolzano-Weierstrass property

Theorem 3.3.4. *A continuous, real-valued function defined on a compact space is bounded and attains its bounds.*

Proof: Given $f : X \rightarrow \mathbb{R}$ is continuous and X is compact we know $f(X)$ is compact by Theorem 3.3.1. We also know f is bounded and $f(X)$ is closed because compact subsets of \mathbb{E}^n are closed and bounded. Since $f(X)$ is closed the supremum and infimum of $f(X)$

must exist in $f(X)$. Let these points be x_1 and x_2 in X such that $f(x_1) = \sup(f(X))$ and $f(x_2) = \inf(f(X))$. \square

Next is perhaps a more interesting property of compact spaces

Theorem 3.3.5. $X \times Y$ is compact if and only if X and Y are compact.

Before we can prove this theorem we need the following definition.

Definition 3.3.6. The **base**, β of a product space topology $X \times Y$ is all sets $U \times V$ where U is an open set in X and V is an open set in Y .

Notice that any open set in $X \times Y$ can be formed as a union of sets in β . Figure 3.3 gives you an idea what these open sets may look like.

The basic open sets are important because any open set in $X \times Y$ can be described as a union

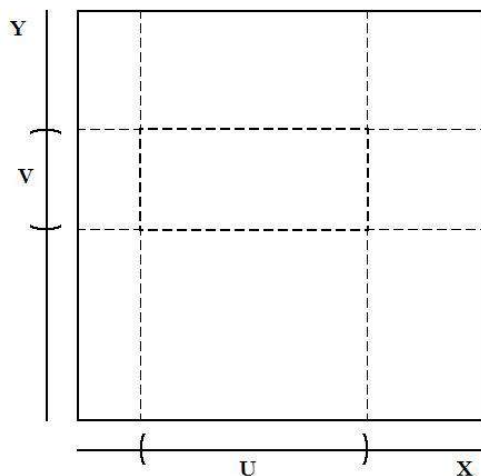


Figure 3.3: The open set $U \times V$ in $X \times Y$

of basic open sets. For this reason we can restrict our proof of compactness to covers composed of these basic open sets (see page 53 in the text).

Proof: (Theorem 3.3.5 Show $X \times Y$ compact implies X and Y are compact. We simply note that the projection functions $p_1 : X \times Y \rightarrow X$ and $p_2 : X \times Y \rightarrow Y$ are onto and continuous thus by Theorem 1 we know X and Y are compact.

Show X and Y compact implies $X \times Y$ is compact. Take any basic open cover \mathcal{O} of $X \times Y$ we must show \mathcal{O} admits a finite subcover. Select a point $x \in X$ and examine $\{x\} \times Y \subset X \times Y$. Notice $\{x\} \times Y$ is homeomorphic to Y by taking the projection p_2 of the subset. Therefore $\{x\} \times Y$ is compact and we can find a finite subcover of \mathcal{O} which contains the subset. Take the smallest subcover $U_1^x \times V_1^x, U_2^x \times V_2^x, \dots, U_n^x \times V_n^x$. Suppose we let

$$U^x = \bigcap_{i=1}^n U_i^x$$

then our subcover contains $U^x \times Y$

The family $\{U^x | \forall x \in X\}$ is an open cover of X so we take a finite subcover

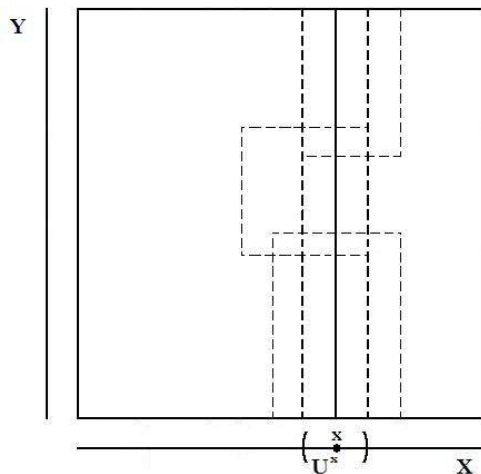


Figure 3.4: The open set $U^x \times Y$ in $X \times Y$

$U^{x_1}, U^{x_2}, \dots, U^{x_m}$ and we have

$$X \times Y = (U^{x_1} \times Y) \cup (U^{x_2} \times Y) \cup \dots \cup (U^{x_m} \times Y)$$

And each of the $U^{x_i} \times Y$ is contained in $(U_1^{x_i} \times V_1^{x_i}) \cup (U_2^{x_i} \times V_2^{x_i}) \cup \dots \cup (U_n^{x_i} \times V_n^{x_i})$ (for $V_j^{x_i}$ a basic open set) a finite number of open sets (where the $V_j^{x_i}$ are defined similarly to the $U_k^{x_i}$ to cover Y). Therefore the collection of all these open sets forms a subcover of $X \times Y$. □

This last proof of course offers us an easy way to prove that a product space is or is not compact: all we have to do is look at its components.

3.4 Hausdorff Spaces

One of the things we like to define in topology is a metric space, a space in which we can define a distance function. Distance functions are important because our favorite topologies, namely the Euclidean space, the torus, etc. are metrizable.

Often times, we would like to see if a topological space is metrizable. To find if a space is metrizable is fairly simple; all we need to do is define a metric function on the space that satisfies the axioms of a metric. However, trying to prove a space is not metrizable can be considerably difficult. Obviously, you cannot try every single function and test it against the axioms of a metric. Additionally, saying ‘I tried really hard, but couldn’t find a metric’ is also not acceptable. The property of a space being Hausdorff provides a nice, easy test to see if a space is not metrizable. If a topological space X is not Hausdorff, then it is not metrizable. Note that being Hausdorff is a necessary but not sufficient condition. All of this will be proved in a moment, but for the time being let’s get the formal definition of a Hausdorff space out of the way:

Definition 3.4.1. A topological space X is Hausdorff if given two distinct points x and $y \in X$, there exists an open set O_x and an open set $O_y \in X$ such that $x \in O_x, y \in O_y$ and $O_x \cap O_y = \emptyset$.

Now that we have a definition for a Hausdorff space, let's see some examples:

- \mathbb{R}^n with the Euclidean topology is Hausdorff, because it is metrizable.
- The discrete topology on any set is also a Hausdorff space, because we can just take O_x and O_y to be the points x and y themselves.
- We note that the stupid topology on any set is not Hausdorff. This is because there is only one non-empty open set in the entire topology, so any two points must be in the same open set.

Theorem 3.4.2. If a topological space (X, τ) is metrizable, then it is Hausdorff.

Proof:

Consider two points x and y in X , and say that $d(x, y) = a$. We can take an open ball around x of radius $\frac{a}{2}$, and an open ball with equal radius around y . This construction gives us two disjoint, open sets around any two points in X , which fulfills the axiom for a Hausdorff space.

The notion of a Hausdorff space is useful because it is a topological invariant. This means that if two topological spaces X and Y are homeomorphic, and X is Hausdorff, then Y must be Hausdorff. More formally

Theorem 3.4.3. The property of a space being Hausdorff is preserved by homeomorphism.

Proof:

Suppose we have two spaces which are homeomorphic X and Y . Assume that X is Hausdorff, we will now directly prove that Y must be Hausdorff using our homeomorphism. Start by picking two distinct points $x_1, x_2 \in X$. We now want to send these points to Y using our homeomorphism, so $f(x_1) = y_1$ and $f(x_2) = y_2$. Note that $y_1 \neq y_2$ because f is one to one. Now, let's take two open sets around x_1 and x_2 , O_1 and O_2 such that $O_1 \cap O_2 = \emptyset$. We now send O_1 and O_2 to Y using f , i.e. $f(O_1)$ and $f(O_2)$. These sets are open in Y , $y_1 \in f(O_1)$ and $y_2 \in f(O_2)$ by our homeomorphism. Finally, we note because f is an onto function, $f(O_2) \cap f(O_1) = \emptyset$. We therefore have two, disjoint open sets around any two distinct points y_1 and $y_2 \in Y$. Therefore, Y is Hausdorff.

Like many of our other favorite topological invariants (compactness, connectedness) Hausdorff spaces behave well with regards to product spaces. Specifically, if two spaces are both Hausdorff, then the product topology defined in the usual way is Hausdorff. This works the other way as well; if a product space is Hausdorff, then its two constituent spaces are Hausdorff.

Theorem 3.4.4. The product space of two topological spaces, $X \times Y$ is Hausdorff if and only if both X and Y are Hausdorff.

Proof:

Assume $X \times Y$ is Hausdorff. We know the product topology equips us with two continuous

projection functions, π_X and π_Y which maps open sets from $X \times Y$ to X and Y respectively. If we take distinct points (x_1, y_1) and $(x_2, y_2) \in X \times Y$, we have two disjoint open set $O_{x_1} \times O_{y_1}$ and $O_{x_2} \times O_{y_2}$. Now, take $\pi_X(O_{x_1} \times O_{y_1}) = O_{x_1}$ and $\pi_X(O_{x_2} \times O_{y_2}) = O_{x_2}$. We can always do this, because these open sets are just the union of the base sets, $U \times V$, of $X \times Y$. By our construction, O_{x_1} and O_{x_2} are disjoint, so X is Hausdorff. The argument for O_y is identical, so Y is Hausdorff.

Now assume X and Y are Hausdorff. Take two distinct points x_1 and $x_2 \in X$ and the open sets surrounding them, O_{x_1} and O_{x_2} . We can form disjoint open sets in $X \times Y$, just by taking $O_{x_1} \times Y$ and $O_{x_2} \times Y$. We can do the same thing for Y , and create disjoint, open sets containing the points (x, y) . Thus, $X \times Y$ is Hausdorff.

3.5 T_1 Spaces

A T_1 space is a looser condition than a Hausdorff space. This means every space that is Hausdorff is T_1 . Hausdorff spaces are sometimes called T_2 spaces. T_1 spaces are encountered less often in topology, but still have some very useful properties, the most useful probably being that points in a T_1 space are closed sets. We will start with the definition of a T_1 space.

Definition 3.5.1. *A topological space X is T_1 if given any pair of distinct points $x, y \in X$, there exists an open set O_x such that $x \in O_x$ but $y \notin O_x$*

Once again, let's come up with some examples to help us with this definition:

- We note immediately that any Hausdorff space is going to be T_1 . This is because in a Hausdorff space we can always find an open set O_y around a point y which does not intersect with O_x . Obviously, $y \notin O_x$ which fulfills the axiom for a T_1 space. Therefore, the same examples for a Hausdorff space carry over. \mathbb{R}^n , and the discrete topology are both T_1 .
- An interesting example of a space which is T_1 but is not Hausdorff is the so called finite complement topology on an infinite set S , where open sets are \emptyset, S and all subsets of S such that their complement is finite. We note that this space is T_1 because given a point $s \in S$, we can contain it in an open set $O_s \subset S$. Now, the complement of O_s must be finite and non-empty, so it must contain a point $q \notin O_s$. However, this topology is not going to be Hausdorff. No two open sets which can be disjoint in this topology!!
- If in the previous example S was a finite space, then the finite complement topology becomes the discrete topology on S , which makes points into closed sets. We would like to prove next that this is a general feature of T_1 spaces.

Theorem 3.5.2. *A topological space X is T_1 if and only if points in X are closed sets.*

Proof:

Assume X is T_1 . Take two different points x and $y \in X$. Now, take the complement of y . By the T_1 definition, we can express y^c as the union of all open sets $O_x \in X$ in which $y \notin O_x$. But the arbitrary union of open sets is open, so the complement of y is open, and y is closed.

Now, assume points are closed sets in X . Take a pair of distinct points $x, y \in X$, and consider the complement of y . This set is open, and contains x . We have thus found an open set O_x which contains x , but does not contain y , so X is T_1 .

From this proof, we can trivially prove that if a space is Hausdorff, then points are closed sets in that space. This follows from the fact that a Hausdorff space is T_1 by definition.

An interesting theorem that ties together the ideas of Hausdorffness and compactness is,

Theorem 3.5.3. *A one-to-one, onto, and continuous function from a compact space X to a Hausdorff space Y is a homeomorphism.*

Proof: Let $f : X \rightarrow Y$ be a function as described above, and take a closed subset of X , C . C is compact by Theorem 3.3.2 and $f(C)$ must be compact by Theorem 3.3.1 and thus (by theorem 3.6 on page 47 of the text), it is also closed in Y . Our function f takes closed sets to closed sets, proving f^{-1} is continuous, making f a homeomorphism. \square

I have chosen to leave out Theorem 3.6 from the book because I feel that it is not very intuitive and that its result is mainly necessary to prove this theorem.

This is a very useful result, now if we can only find a continuous bijection from a compact space to a Hausdorff space, we get a homeomorphism, (and the compactness of the Hausdorff space, as well as the Hausdorffness of the compact space) for free. Of course in practice it can be difficult to find a function satisfying these properties.

3.6 Compactification

3.6.1 Motivation

Compact spaces are often easier to work with than non-compact spaces; for instance, the classification theorem only handles compact surfaces. They're also nice because they don't ever run away to infinity. Compactification is a way to take a space X that is not compact, and to add a little bit to it in such a way that we produce a new compact space. To be more precise, we want to construct a compact space \tilde{X} such that X is homeomorphic to an open dense subset of \tilde{X} . We want it to be dense because we want to alter X as little as possible, which means we don't want to have to add very much to X in order to get to \tilde{X} .

3.6.2 One-Point Compactification

We'll assume throughout that X is both Hausdorff (meaning that distinct points can be separated by neighborhoods) and locally compact (meaning that each of its points is contained in a compact neighborhood).

Definition 3.6.1. *The **one-point compactification** \tilde{X} of a topological space X is the set $X \cup \{\infty\}$ (where ∞ is a single point called "the point at infinity") with the following topology:*

- if $\infty \notin U$, then U is open in \tilde{X} if and only if it is open in X .
- if $\infty \in U$, then U is open in \tilde{X} if and only if its complement is compact.

For example, if X is the open interval (a, b) of the real line, then \tilde{X} is a circle. All of the original open sets (open intervals on the real line) are still open, and we also get new open sets of the form $(a, c) \cup \{\infty\} \cup (d, b)$. Intuitively, we can think of this process as bringing the two ends of the interval together to meet at the point ∞ .

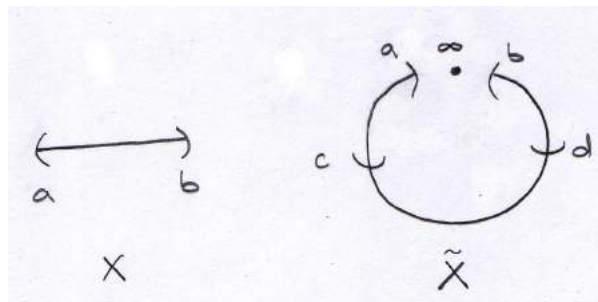


Figure 3.5: The one-point compactification of (a, b)

3.6.3 Theorems

We want to make sure that our definition does what we want it to do, and gives us something reasonable. Here are a few checks we can make:

Theorem 3.6.2. For any open set $U \subset \tilde{X}$, $U \cap X$ is open in X .

This ensures that we aren't creating "too many" new open sets.

Proof. Take an open set $U \subset \tilde{X}$. If $\infty \notin U$, then the topology we've chosen guarantees that $U \cap X = U$ is open in X .

If $\infty \in U$, then by our topology $\tilde{X} - U$ is compact. Since $\infty \in U$, the points in $\tilde{X} - U$ are all points of X . We have assumed that X is Hausdorff, so $\tilde{X} - U$ is closed in X . Thus its complement, $X - (\tilde{X} - U) = U - \infty = X \cap U$, must be open in X . \square

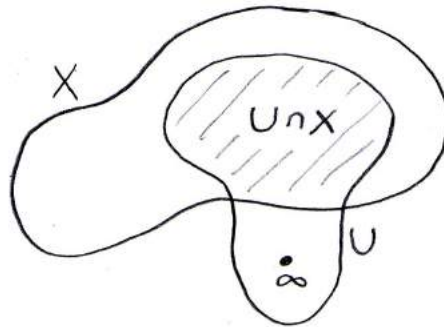


Figure 3.6: Proof of Theorem 3.6.2

Theorem 3.6.3. *If X is not compact, then X is homeomorphic to an open dense set in \tilde{X} .*

As mentioned before, this ensures that \tilde{X} is not "too large", that is, not "too much larger" than X .

Proof. First let's show that X is homeomorphic to the set $\{X\} \subset \tilde{X}$. But that's easy: construct a function that sends each point of X to the corresponding point in $\{X\}$. This function is obviously one-to-one and onto, and it is continuous (and so is its inverse) because the open sets in $\{X\}$ are exactly the open sets in X .

The set $\{X\}$ is open in \tilde{X} , because it doesn't contain ∞ and it is open in X . To show that $\{X\}$ is dense, we can simply show that it is not closed, or that ∞ is not open. (If that's the case, then $\{X\}$ isn't its own closure, and the only other option is that its closure is \tilde{X} .) If ∞ is open, then its complement, $\{X\}$, must be compact. But this would imply that X is compact, contradicting our earlier assumption. So ∞ can't be open, meaning $\{X\}$ must be dense. \square

Note that if X is already compact, then we can't prove that $\{X\} \subset \tilde{X}$ is dense.

Theorem 3.6.4. *\tilde{X} is compact.*

This had better be true, because it's the whole point of compactification!

Proof. Start with an open cover of \tilde{X} , denoted $\bigcup_{\alpha} U_{\alpha}$. We want to show that we can always extract a finite open subcover. Since we've covered \tilde{X} , at least one of these sets (call it U_{∞}) contains ∞ . By our topology, its complement $\tilde{X} - U_{\infty}$ is compact, so from the open cover $\bigcup_{\alpha} U_{\alpha}$ we can extract a finite open subcover of $\tilde{X} - U_{\infty}$, denoted $\bigcup_{\beta} U_{\beta}$. If we put U_{∞} in, then we've produced $U_{\infty} \cup \bigcup_{\beta} U_{\beta}$, which is the finite open subcover of \tilde{X} that we've been looking for. \square

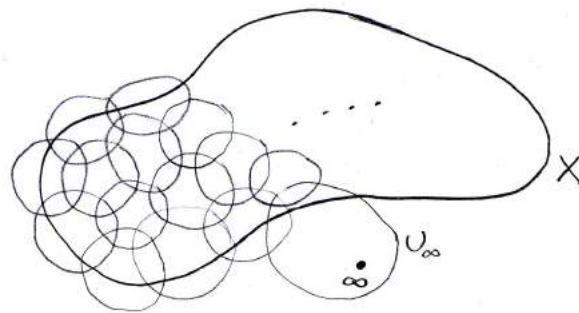


Figure 3.7: Proof of Theorem 3.6.4

Theorem 3.6.5. *If none of the components of X is compact, then \tilde{X} is connected.*

Proof. Assume that \tilde{X} isn't connected, i.e. there is some set U in \tilde{X} that is open and closed, but isn't \emptyset or \tilde{X} . Its complement, V , is also open and closed without being \emptyset or \tilde{X} . Either U or V contains ∞ ; take the one that doesn't, and call it W . W is compact because its complement is open and contains ∞ .

First let's consider the case that X is connected. We've already established that W isn't \emptyset . It can't be all of X either, because W is compact and X isn't. W is open in X because it's open in \tilde{X} and doesn't contain ∞ . It's closed in X because its complement (either $U \cap X$ or $V \cap X$) is open in X (see Theorem 3.6.2). So W is open, closed, not \emptyset , and not X , which implies that X isn't connected. This contradicts our assumption, so \tilde{X} must be connected.

But what if X isn't connected? In this case, we look at the connected components of X . Any open set including ∞ must also contain points in each of the components of X (because the complement of the open set is compact, and if the complement included an entire connected component, then that component would need to be compact, but it isn't). So W contains some points in each of the components. But this would imply that the connected components aren't connected, which is our contradiction. So again, \tilde{X} must be connected. \square

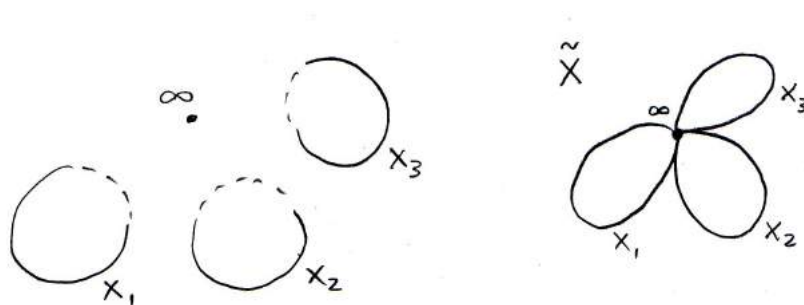


Figure 3.8: Proof of Theorem 3.6.5

Note that if X is already compact, then \tilde{X} *cannot* be connected. The set $\{\infty\}$ is open (because

it contains ∞ and its complement X is compact), closed (because its complement is open), not \emptyset , and not \tilde{X} .

Theorem 3.6.6. \tilde{X} is Hausdorff.

Proof. We want to show that for any two points $x, y \in \tilde{X}$, we can find disjoint open sets $U_x \ni x$ and $U_y \ni y$. If $x, y \neq \infty$, then we can find the appropriate U_x and U_y in X because we know that X is Hausdorff.

If $x = \infty$, then we use the fact that X is locally compact, so there is an open set $O \ni y$ contained entirely in a compact set $C \subset X$. Now consider C as a set in \tilde{X} : its complement $\tilde{X} - C$ must be open (since it contains ∞ and its complement is compact). So $U_x = \tilde{X} - C$ and $U_y = O$, and we win. \square

Theorem 3.6.7. If X is homeomorphic to Y , then \tilde{X} is homeomorphic to \tilde{Y} .

Proof. X and Y are homeomorphic, so there is some function $f : X \rightarrow Y$ that is bijective and continuous, and has a continuous inverse. We want to extend f to a function from \tilde{X} to \tilde{Y} , and show that it is also a homeomorphism. Define \tilde{f} as follows:

$$\tilde{f} = \begin{cases} f(x) & \text{if } x \neq \infty_X \\ \infty_Y & \text{if } x = \infty_X \end{cases}$$

In other words, if $x \in X$ leave $f(x)$ alone, and map the ∞ of \tilde{X} to the ∞ of \tilde{Y} . This new function is clearly bijective, because f is bijective and the extra points are mapped to each other.

To show that \tilde{f} is continuous: take an open set U in \tilde{Y} . If $\infty_Y \notin U$, then the preimage $\tilde{f}^{-1}(U) = f^{-1}(U)$ is open in \tilde{X} because $f^{-1}(U)$ is open in X . If $\infty_Y \in U$, then we know that $C = \tilde{Y} - U$ is compact, and $\tilde{f}^{-1}(U) = \tilde{f}^{-1}(\tilde{Y} - C) = \tilde{X} - f^{-1}(C)$. But $f^{-1}(C)$ must also be compact, so its complement $f^{-1}(U)$ is open in \tilde{X} .

We can use the same argument to show that \tilde{f}^{-1} is continuous. Thus \tilde{X} and \tilde{Y} are homeomorphic. \square

The converse of this theorem (if $\tilde{X} \sim \tilde{Y}$, then $X \sim Y$) is not necessarily true. For example, take $X = [0, a) \cup (a, 1]$ and $Y = [0, 1]$. In this case, $X \approx Y$, because Y is connected but X isn't. If we compactify X , the new open sets are of the form $(b, a) \cup \{\infty\} \cup (a, c)$, where $0 \leq b < a$ and $a < c \leq 1$. For \tilde{Y} , the new open sets are of the form $(d, 1) \cup \{\infty\}$, where $0 \leq d < 1$. So $\tilde{X} \sim \tilde{Y} \sim [0, 1]$, even though $X \not\approx Y$.

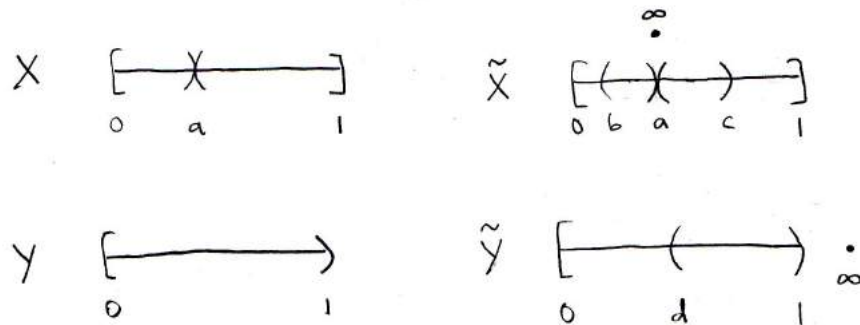


Figure 3.9: Counterexample to "if $\tilde{X} \sim \tilde{Y}$, then $X \sim Y$ "

3.6.4 Examples

Example 3.6.8 (several open intervals of the real line). *We've already seen that the one-point compactification of (a, b) is a circle. What if $X = (a, b) \cup (c, d)$? In this case, the new open sets are of the form $\{\infty\} \cup (a, a') \cup (b', b) \cup (c, c') \cup (d', d)$ (open intervals on both ends of both of the original intervals, plus the point at infinity). So \tilde{X} is homeomorphic to a figure-8, which is also called a "bouquet of two circles".*

This is generalizable to any number of open intervals: for n intervals, we get a bouquet of n circles.

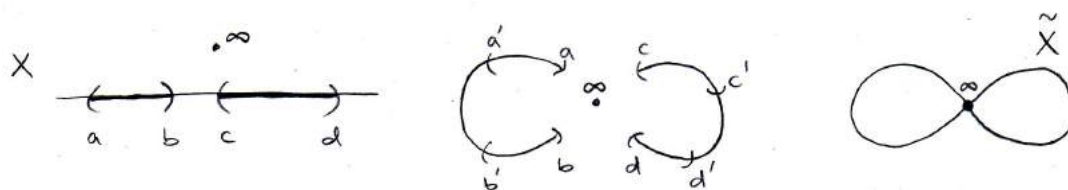


Figure 3.10: One-point compactification of $(a, b) \cup (c, d)$ (Example 3.6.8)

Example 3.6.9 (open disc in the plane). *To find the new open sets, remember that their complements must be compact, i.e. closed and bounded. So an open set containing ∞ must also contain a thin open strip around the edge of the disc. Intuitively, we bring the edges of the disc up to meet the point at infinity, and we are left with a sphere.*

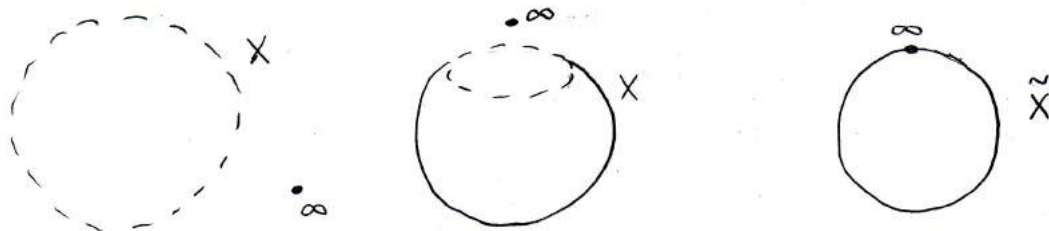


Figure 3.11: One-point compactification of an open disc (Example 3.6.9)

Example 3.6.10 (closed disc in the plane minus a point). *An open set containing ∞ must also contain a small open disc around the missing point, so its complement is a closed annulus (possibly with holes in it, if the open set is the union of several pieces). Intuitively, ∞ "fills in" the gap left by the missing point in the disc. So the one-point compactification is a closed disc.*

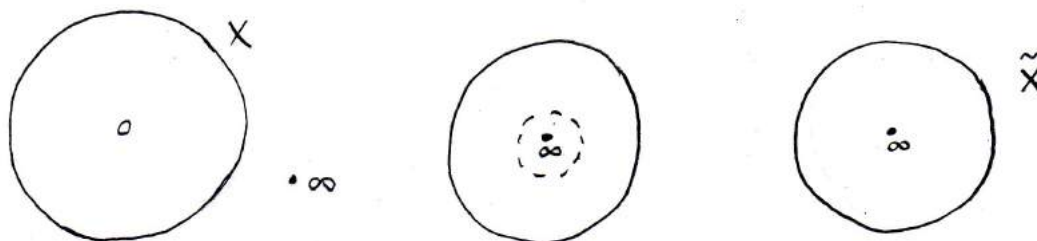


Figure 3.12: One-point compactification of a closed disc minus a point (Example 3.6.10)

Example 3.6.11 (\mathbb{R}). *The real line \mathbb{R} is homeomorphic to the open interval (a, b) . By Theorem 3.6.7, their one-point compactifications are also homeomorphic. So $\tilde{\mathbb{R}}$ is homeomorphic to a circle.*

Example 3.6.12 (\mathbb{R}^2). *The plane is homeomorphic to the open disc, so we again invoke Theorem 3.6.7 and say that $\tilde{\mathbb{R}}^2$ is homeomorphic to a sphere.*

Example 3.6.13 ($\mathbb{R}\mathbb{P}^2$ as a compactification of \mathbb{R}^2). *All of the compactifications we've talked about so far have been one-point compactifications; that is, we've added a single point to the original space X to get the new space \tilde{X} . But that isn't the only way to compactify spaces.*

For example, we sometimes think of $\mathbb{R}\mathbb{P}^2$ as a plane plus a line plus a point. In other words, we compactify the plane \mathbb{R}^2 by adding not just a point at infinity, but also a line at infinity.

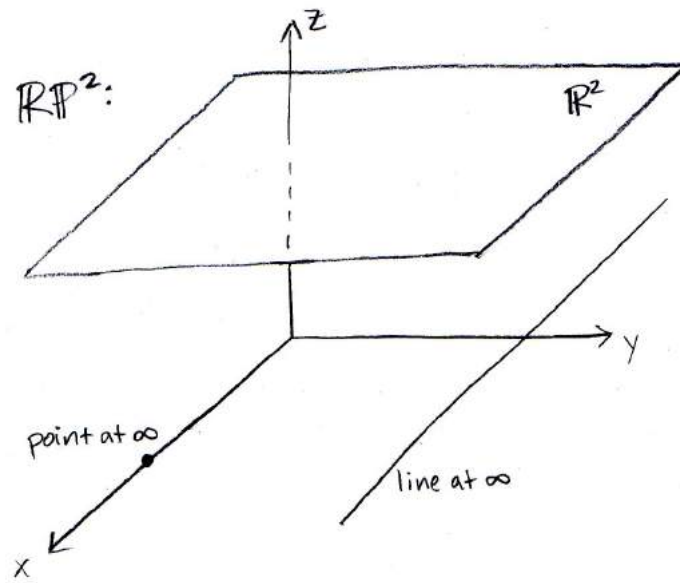


Figure 3.13: The projective plane (Example 3.6.13)

3.7 Connectedness

3.7.1 Introduction

Intuitively, a topological space is a **connected space** if it is made up of one single piece. To be a little more thorough, think of a connected space as a space which *cannot* be represented as the disjoint union of two more more nonempty open sets. Figure 3.14 gives a visual representation of a connected space X and a disconnected space Y formed by the union of its disjoint components.

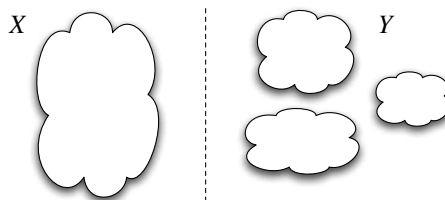


Figure 3.14: X is a connected space, Y is a disconnected space

3.7.2 Connectedness

Definition 3.7.1. A space X is **connected** if whenever it is decomposed as the union $A \cup B$ of two nonempty subsets then $\overline{A} \cap B \neq \emptyset$ or $A \cap \overline{B} \neq \emptyset$

Theorem 3.7.2. The following conditions on a space X are equivalent:

- (a) X is connected
- (b) The only subsets of X which are both open and closed are X and the empty set
- (c) X cannot be expressed as the union of two disjoint nonempty open sets
- (d) There is no onto continuous function from X to a discrete space which contains more than one point

Proof. We will show the statements' equivalence by showing that (a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a).

(a) \Rightarrow (b) Assume that X is a connected space. Let A be an open and closed subset in X such that it is not the empty set or all of X . If we define $B = X \setminus A$, it is both open and closed since it is the relative complement of A in X . Since both A and B are closed, $\overline{A} = A$ and $\overline{B} = B$, and therefore

$$\overline{A} \cap B = A \cap \overline{B} = A \cap B = \emptyset$$

But this is a contradiction since X is connected and, by definition, $\overline{A} \cap B \neq \emptyset$ or $A \cap \overline{B} \neq \emptyset$. Therefore one of A, B must be the empty set and the other be the whole space, i.e., the only subsets of X which are both open and closed are X and the empty set.

(b) \Rightarrow (c) Assume that the only subsets of X which are both open and closed are X and the empty set. Let $A \cup B = X$ where A and B are two disjoint nonempty open sets. $X \setminus A = B$, which is a closed set since it's the complement of an open set. But this means that B is both open and closed which contradicts our assumption. Therefore, X cannot be expressed as the union of two disjoint nonempty open sets.

(c) \Rightarrow (d) Assume that X cannot be expressed as the union of two disjoint nonempty open sets. Let Y be a discrete space with more than one point and let $f : X \rightarrow Y$ be an onto continuous function. Define the nonempty open sets U, V such that $Y = U \cup V$. Since f is continuous, $X = f^{-1}(U) \cup f^{-1}(V)$, but this contradicts our assumption (c). Therefore, Y can't be a discrete space with more than one point, i.e., there is no onto continuous function from X to a discrete space which contains more than one point.

(d) \Rightarrow (a) Assume that there is no onto continuous function from X to a discrete space which contains more than one point. Let's suppose that X is not connected so we can define two nonempty open sets A, B such that $X = A \cup B$ and $\overline{A} \cap B = A \cap \overline{B} = \emptyset$. Since A and B are disjoint open sets whose union equals all of X , each set is the complement of the other set's closure, i.e., $X \setminus A = \overline{B}$. Therefore we can construct a function f from X to the discrete subspace of the real line, $\{-1, 1\}$, by

$$f(x) = \begin{cases} -1, & \text{if } x \in A \\ 1, & \text{if } x \in B \end{cases}$$

But f is an onto continuous function from X to a discrete space which contains more than one point which contradicts our assumption. Therefore, X must be connected.

□

Theorem 3.7.3. *The continuous image of a connected space is connected.*

Proof. Let $f : X \rightarrow Y$ be an onto continuous function and X be a connected space. If A is a subset of Y such that A is both open and closed, then $f^{-1}(A)$ is both open and closed since f is continuous. Since $f^{-1}(A)$ is both open and closed and X is connected, $f^{-1}(A)$ must either be all of X or the empty set (condition (b) of Theorem 3.7.2). Therefore A is either the entire space Y or the empty set, i.e., Y is connected. □

Corollary 3.7.4. *If $f : X \rightarrow Y$ is a homeomorphism, then X is connected if and only if Y is connected.*

The key consequence of Corollary 3.7.4 is that **connectedness is a topological invariant**.

Theorem 3.7.5. *Let X be a topological space and let Z be a subset of X . If Z is connected and dense in X , then X is connected.*

Proof. Let $A \subset X$ such that A is both open and closed and nonempty. Since Z is dense in X we know that Z must intersect every nonempty open subset of X , and therefore $A \cap Z \neq \emptyset$. Since $A \cap Z$ is both open and closed in Z and Z is connected we deduce that $A \cap Z = Z$, i.e., $Z \subset A$. Therefore $X = \overline{Z} \subset \overline{A} = A$, which means $X = A$ and therefore X is connected (condition (b) of Theorem 3.7.2). □

Lemma 3.7.6. *Let \mathfrak{F} be a family of subsets of X whose union is all of X and A be a nonempty subset of X which is both open and closed and not equal to all of X . If each member of \mathfrak{F} is connected, then for $Z \in \mathfrak{F}$ either $Z \cap A$ is empty or all of Z .*

Proof. Assume that each member of \mathfrak{F} is connected and A is a nonempty subset of X which is both open and closed and not equal to all of X . We know that $A \cap Z$ is an open set in Z and that $Z \cap (X \setminus A) = Z \setminus A$ is an open set in Z . This means that $(A \cap Z) \cup (Z \setminus A) = Z$, yet $(A \cap Z) \cap (Z \setminus A) = \emptyset$ if $A \neq X$, i.e., Z can be expressed as the union of two disjoint nonempty open sets. This contradicts our assumption that Z , being a member of \mathfrak{F} , is connected, and therefore, $Z \cap A$ is empty or all of Z . □

Theorem 3.7.7. *Let \mathfrak{F} be a family of subsets of a space X whose union is all of X . If each member of \mathfrak{F} is connected, and if no two members of \mathfrak{F} are separated from one another in X , then X is connected.*

Proof. Let A be a subset of X which is both open and closed. Since each member of \mathfrak{F} is connected and X is the union of all the elements of \mathfrak{F} , we know via Lemma 3.7.6 that for any $Z \in \mathfrak{F}$ that $Z \cap A = \emptyset$ implies $A = \emptyset$ and $Z \cap A = Z$ implies $Z \subset A$, i.e., Z is either in A or not. But if $A \neq X$ there exists some $Z, W \in \mathfrak{F}$ such that $Z \cap A = Z$ and $W \cap A = \emptyset$, i.e., Z is in A but W is not. This is a contradiction since Z and W are members of \mathfrak{F} and no two members of \mathfrak{F} are separated. Therefore if A is nonempty it must be the entire space X , i.e., X is connected. □

Theorem 3.7.8. *If X and Y are connected spaces then the product space $X \times Y$ is connected.*

Proof. Let $x \in X$ and $p_1 : Y \rightarrow \{x\} \times Y$. The function $p_1(y) = (x, y)$ is a homeomorphism, and since Y is a connected space, by Corollary 3.7.4 the subspace $\{x\} \times Y$ of $X \times Y$ is connected.

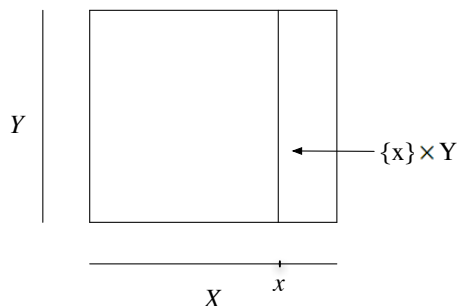


Figure 3.15: The subspace $\{x\} \times Y$ of $X \times Y$

Similarly for any point $y \in Y$, $X \times \{y\}$ is connected. Now $\{x\} \times Y$ and $X \times \{y\}$ are bound to intersect at some point (x, y) , therefore $\{x\} \times Y \cup X \times \{y\}$ is connected (apply Theorem 3.7.7).

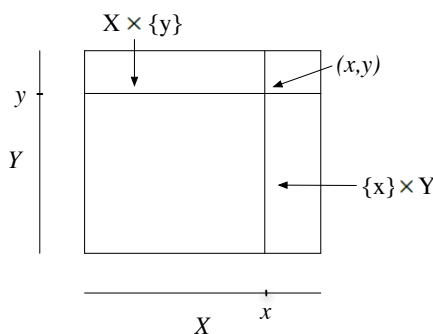


Figure 3.16: $\{x\} \times Y \cup X \times \{y\}$

We can see that $X \times Y = \bigcup_{x \in X, y \in Y} \{x\} \times Y \cup X \times \{y\}$, and for any $x_i \neq x_j$ and $y_i \neq y_j$ we have $(\{x_i\} \times Y \cup X \times \{y_i\}) \cap (\{x_j\} \times Y \cup X \times \{y_j\}) \neq \emptyset$. Therefore by applying Theorem 3.7.7 again we see that $X \times Y$ is connected. \square

A consequence of Theorem 3.7.7 is that euclidean n -space is connected since it is simply the product of a finite number of copies of the real line. We also see that given a sphere S^n in \mathbb{E}^{n+1} where $n \geq 1$, if we remove a point from S^n we're left with a space that is homeomorphic to \mathbb{E}^n . Since the closure of S^n minus a point is all of S^n , i.e., S^n is dense, by Theorem 3.7.5 we can see that S^n is a connected space for $n \geq 1$. Finally, since a torus is simply the product space $S^1 \times S^1$ we can see that it is connected.

3.7.3 Path-Connectedness

Now we'll address the topological invariant of path-connectedness, intuitively a stronger notion of connectedness where any two points in the space can be joined by a path.

Definition 3.7.9. A path in a topological space X is a continuous function $\gamma : [0, 1] \rightarrow X$. A space is **path-connected** if any two points in the space can be joined by a path.

Intuitively, $\gamma(0)$ and $\gamma(1)$ are thought of as the *beginning* and *end* points of the path respectively.

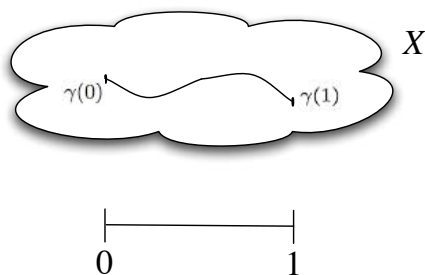


Figure 3.17: The path $\gamma : [0, 1] \rightarrow X$ on the path-connected space X

Theorem 3.7.10. A path-connected space is connected.

Proof. Assume X is a path connected space with a nonempty open and closed subset A such that $A \neq X$. Let $\gamma : [0, 1] \rightarrow X$ be a path on X such that $\gamma(0) = x$ and $\gamma(1) = y$ where $x \in X \setminus A$ and $y \in A$. Because of the continuity of γ , $\gamma^{-1}(A)$ is an open and closed subset of $[0, 1]$. But because $[0, 1]$ is a connected space, the only open and closed sets are either the empty set or $[0, 1]$, so $\gamma^{-1}(A)$ must be all of $[0, 1]$. Therefore $A = X$, i.e., X is a connected space. \square

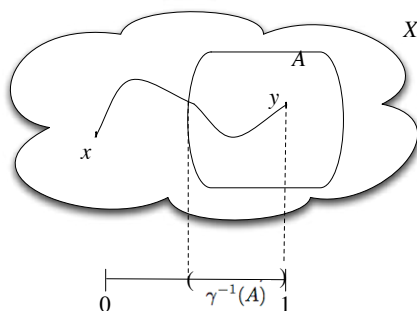


Figure 3.18: The path-connected space X

Theorem 3.7.11. A connected open subset of a euclidean space is path-connected.

Proof. Let X be a connected open subset of \mathbb{E}^n . Pick $x \in X$ and let U_x be the set of points in X that can be joined to x by a path. Let $y \in U_x$ and choose an open ball B that lies entirely in X centered at y . Then for any $z \in B$ a path can be drawn to y , and since a path can be drawn from y to x , a path can be drawn from z to x . Therefore z is path connected to x , B is contained in U_x , and U_x is open. Let the complement of U_x be V_x , which is the set of all points in X that cannot be joined to x by a path. By the same logic, V_x is open making U_x closed ($X \setminus V_x = U_x$). Since U_x is both open and closed and not empty (there exists at least one element in U_x , namely x connected to itself), then $U_x = X$. Therefore, X is path-connected. \square

Since we'll generally be working with connected open subsets of a euclidean space we'll generally also be working with path-connected spaces. Though you must note that though a path-connected space is connected, a connected space is not necessarily path-connected. Take a look at the space $X = \{B \cup C \cup \bigcup_{n=0}^{\infty} A_n\}$ where A_n is the half-line $\{x = \frac{1}{2^n} \text{ and } y > 0\}$ represented in Figure 3.19. Here X is connected since the only open and closed subsets of X are the entire

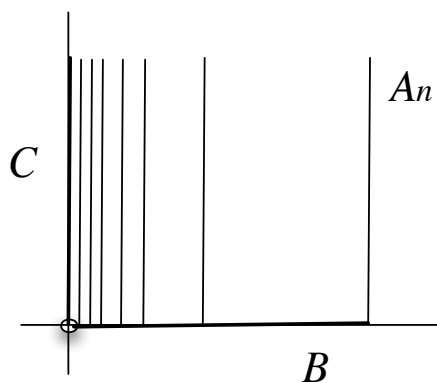


Figure 3.19: The space $X = \{B \cup C \cup \bigcup_{n=0}^{\infty} A_n\}$

space X and the empty set. Yet X is not path-connected since you can't join via a path a point on A_n or B to a point on C because the origin is not part of X .

Chapter 4

Surfaces

4.1 Surfaces

A surface is a topological space S such that every point $s \in S$ has a neighborhood homeomorphic to \mathbb{R}^2 .

4.2 The Projective Plane

As will be proven later, the study of (compact) surfaces has 3 basic surfaces as building blocks: the sphere, the torus, and the projective plane. Since essentially all surfaces can be made from these building blocks, it is important to gain a good understanding of them. The most difficult of these to grasp is probably the projective plane, \mathbb{RP}^2 , and so here we will be concerned with giving a thorough description of the projective plane in a variety of ways.

4.2.1 \mathbb{RP}^2 as lines in \mathbb{R}^3 or a sphere with antipodal points identified.

The first, and probably easiest to visualize, description of \mathbb{RP}^2 is as the space of lines through the origin in \mathbb{R}^3 . Note that this is the exact same as taking a sphere and identifying the antipodal points, as each line through the origin will intersect the sphere in exactly two points. This means that we can form a map from the space of lines through the origin to the sphere (with antipodal points identified) by sending each line to the pair of points at which the line intersects the sphere. In figure 4.1, the red line would be sent to the red points (which are just one point, since we are identifying antipodal points), and the blue lines would be sent to the blue points. We can see geometrically that this is a continuous function by noting that wiggling a line a little bit will wiggle the intersection points on the sphere a little bit as well. The one-to-one and onto conditions are also geometrically obvious, as each line intersects the sphere in exactly 2 points. The inverse is just sending a point on the sphere to the line through that point and the origin, and so these spaces are homeomorphic; that is, both are equivalent ways of describing \mathbb{RP}^2 .

The reason that this space is called the projective plane is that it is very much related to the plane, and is actually a compactification of it.

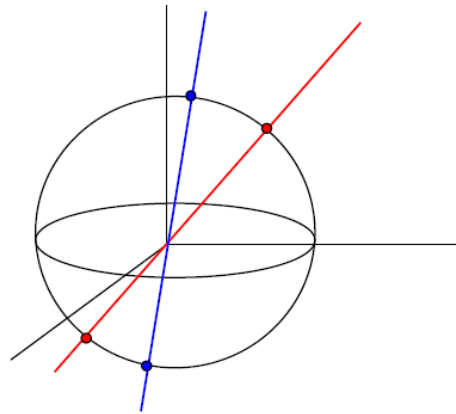


Figure 4.1: Lines through \mathbb{R}^2

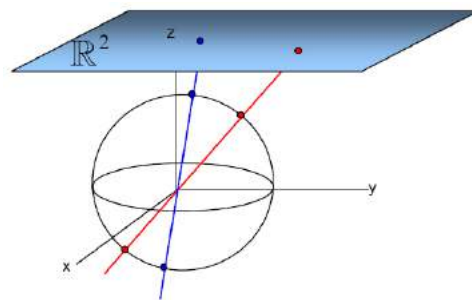


Figure 4.2: Mapping from lines in \mathbb{R}^3 to the plane

To see this, place a plane above the sphere as shown in figure 4.2. Then, note that we can define a function from a subset of $\mathbb{R}P^2$ to \mathbb{R}^2 by sending each line to its intersection with \mathbb{R}^2 . However, by doing this, we actually miss the lines that lie in the xy -plane.

To remedy this, we add another line that lies in the xy -plane, parallel to the x -axis, as shown in figure 4.3 (blue line).

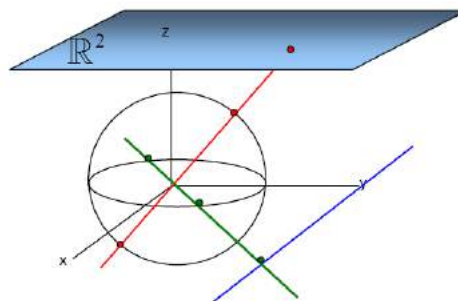


Figure 4.3: Adding a line to intersect lines in the x - y plane

Now, extend our function to send any line through in the origin in the xy -plane to the corre-

sponding intersection point on the blue line. This takes care of almost every line through the origin (i.e., every point in the projective plane), except for the x-axis, which intersects neither the plane above nor the blue line. So, we must add one more point in our image, corresponding to the x-axis.

So, now our function sends any point in \mathbb{RP}^2 (which is a line through the origin in \mathbb{R}^3) to its intersection with the plane above, the blue line, or the added point (if we are talking about the x-axis). This is another way of saying that the projective plane is equivalent to \mathbb{R}^2 plus a line plus a point. Adding the line and the point are a way of compactifying the plane.

The compactness of \mathbb{RP}^2 is one of its most important features, and we will use this in the classification theorem. For a direct proof that \mathbb{RP}^2 is compact, consider the function above from the sphere to \mathbb{RP}^2 that sends a point on the sphere to the corresponding line through that point and the origin. The continuity of this function can be seen geometrically, as we described above. But, we know that the continuous image of a compact set is compact, and therefore, since the sphere is compact, \mathbb{RP}^2 is compact as well.

4.2.2 The Projective Plane as a Quotient Space of the Sphere

If we consider the group action $\mathbb{Z}_2 \times S^2 \rightarrow S^2$ defined by

$$(1, p) \mapsto p$$

$$(-1, p) \mapsto -p$$

it is evident that the orbits are pairs of antipodal points as shown in figure 4.4.

Now, if we make the identification space whose partitions are the orbits of this group action, we see that what we are doing is identifying pairs of antipodal points of the sphere. But, this is exactly the description of the projective plane given earlier, and so we find that

$$\mathbb{RP}^2 = S^2 / \mathbb{Z}_2$$

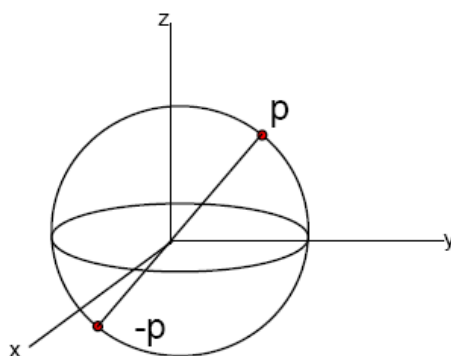


Figure 4.4: Identifying antipodal points on a sphere

4.2.3 The Projective Plane as an identification space of a disc

In constructing surfaces and proving the classification theorem, we will find that an extremely useful way of thinking about a surface is as an identification polygon. As a first example of this, we will now show that the projective plane can be described in yet another way: as a bigon with the two sides identified in opposite directions.

We want to represent the projective plane as a bigon, which is two dimensional. To do this, we start with the sphere with antipodal points identified. Then, since antipodal points are identified, one point from each pair is sufficient, and so we can ignore the entire lower hemisphere, shown in figure 4.5.

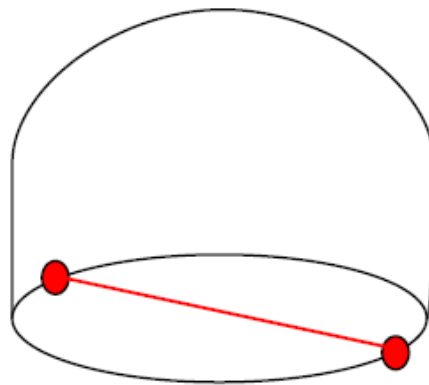


Figure 4.5: Discarding the lower hemisphere

Now, each point on the northern hemisphere represents a unique point of the projective plane, with the exception of points on the equator, which are still identified with their antipodes. This space is clearly homeomorphic to a disc with antipodal boundary points identified; the homeomorphism is to project each point (x, y, z) on the sphere to the point (x, y) directly beneath it in the plane (Figure 4.6).

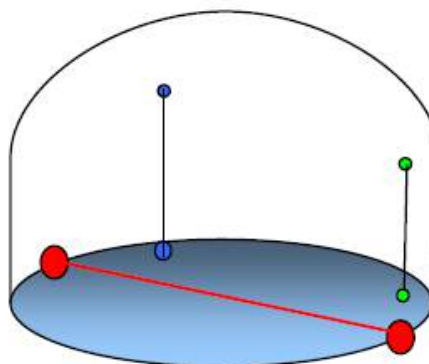


Figure 4.6: Projecting the upper hemisphere to the x - y plane

After the projection, we are left with the projective plane as shown in figure 4.7.

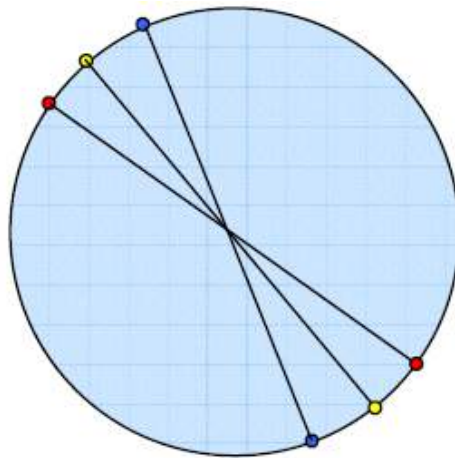


Figure 4.7: Result of the projection

This is simply a closed disc in the plane with antipodal points on its boundary circle identified. Each point represents a unique point of the projective plane (a line through the origin in \mathbb{R}^3 except for the boundary points, which are identified with their antipodes). Using the arrow notation for identification polygons, we are left with the familiar picture in figure 4.8.

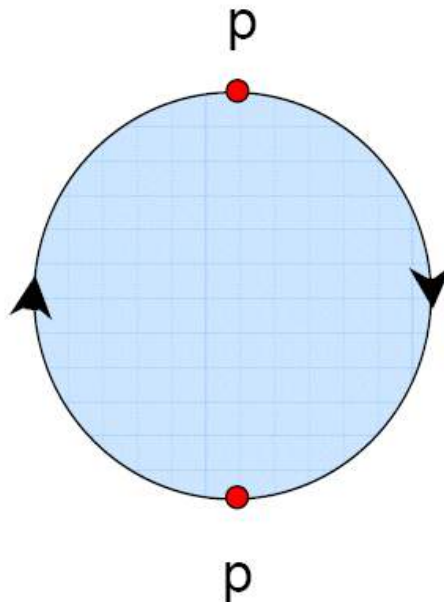


Figure 4.8: Final bigon representation

The above bigon is the representation of the projective plane that we will use the most. It will be the most useful in constructing arbitrary surfaces (through connected sums) and for the proof of the classification theorem. Another alternative, but equivalent representation is a

square with sides identified as in figure 4.9.

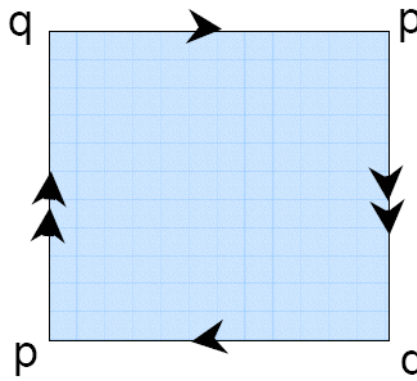


Figure 4.9: An equivalent identification polygon

Clearly identifying the sides above will give a surface, because every point will have a neighborhood that is locally a plane. This is obvious for the points on the interior of the disc, and is no harder to see for points on the boundary. For example, take the point p on the boundary of the projective plane in figure 4.10.

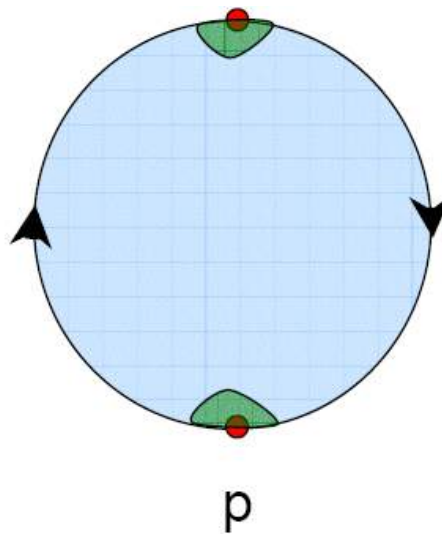


Figure 4.10: A neighborhood around point p

Because p is identified with its antipode, any neighborhood of p will look like the green region of the above picture, containing points on both sides of the polygon. Because the points on the boundary are identified, this will look locally like the plane. However, if you try to actually imagine what the entire surface would look like, it is very difficult to do, because the projective plane cannot be represented in 3D space without intersecting itself. This is because of an important property called non-orientability, to which we now turn.

4.2.4 Non-Orientability of the Projective Plane

The importance of the projective plane is that it is the most basic non-orientable, closed, compact surface, from which all other non-orientable, closed, compact surfaces can be made. (A more complete discussion of orientability will come later). A surface is non-orientable if it contains a Möbius band. A Möbius band is a rectangle with the sides identified as shown in figure 4.11.



Figure 4.11: Möbius strip as an identification space

The top and bottom sides are identified only with themselves. That the projective plane contains a Möbius band is obvious from the bigon representation (Figure 4.12).

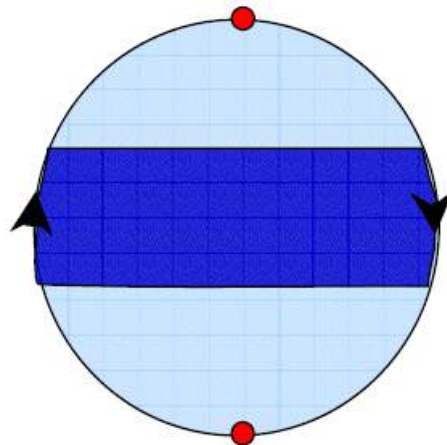


Figure 4.12: The Möbius strip in the projective plane

The dark blue strip is a Möbius band, and so the projective plane is a non-orientable surface.

4.3 Polygons

All closed, compact surfaces can be represented by a $2n$ -sided polygon. The pairwise identification of edges of known polygons can be used to build other polygons that are more complex and can represent new surfaces.

Triangulation is a process that helps with dividing up a surface. Surfaces that are divided into triangles are sometimes easier to work with because the position of the vertices determines/reveals a lot of the properties of the surface. When dividing the surface, there is a condition: any two triangles share a single edge or a single vertex or they are disjoint. An example of triangulation is shown in figure 4.13.

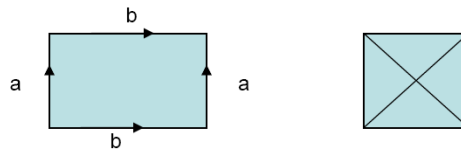


Figure 4.13: A torus divided into triangles

A final triangulated surface will be such that each edge is identified to exactly one other edge and triangles meeting at a vertex can be labeled in sequence, by identifying along an edge. In addition, any compact surface can be finitely-triangulated. This can be seen by assuming that a compact surface has an infinite triangulation. This would mean that the surface would have an infinite number of vertices, which can be arranged in a sequence with a limit point. This is impossible with an infinite triangulation; therefore, we have reached a contradiction.

For any given surface, there are many possible polygonal representations for it. As long as the conditions of triangulation are held for the triangles (or the identification polygons), they can be labeled and assembled in a variety of ways. The amount of simplification and reduction that one performs on a polygonal representation of a surface varies. Therefore many different polygons can, in effect, represent the same surface.

The polygonal representations of a surface are helpful because they are easy to work with. The process of triangulation introduces cutting and gluing, and as such we can visualize a surface and its topological equivalents more easily. It is a lot easier to draw many of the polygonal representations of a surface, as opposed to the surface itself. Using a polygonal representation also allows us to see the entire surface all at once. For all of these reasons, polygonal representations are very helpful in the analysis of surfaces.

Using polygonal representations of surfaces can also be a lot of fun. Imagine the entertainment of 2nd grade art class without the mess of cutting paper and gluing! When we are identifying edges, we don't have to actually bring together the identified edges we merely label them as such. Polygonal representations of surfaces are fun because they are like twisted puzzles you create the pieces in order to construct the mental picture of the surface.

Polygons and triangles represent identification spaces. With these you can achieve or construct new spaces from other spaces using the properties of group action. The idea behind identification spaces:

1. Start with some familiar space
2. Decide that bunches of points become the same
3. Get a new set and a new topology on it \rightarrow Identification Space

In more mathematical terms:

1. Let X be a topological space
2. Let $Y = \bigcup\{Y_i\}$ be a partition, where Y_i is a subset of X

3. Get a function:

$$X \rightarrow Y$$

$x \mapsto$ the unique subset that contains x

4. Make Y into a topological space by defining a set $U \subseteq Y$ as open only if $\Pi^{-1}(u)$ is open in X

In essence, we are gluing along a function to build the new space.

As mentioned earlier, any surface can also be constructed from our three basic building blocks: the sphere, the torus, and the projective plane. We are working towards the proof of this fact, but first, we must find the identification polygons of our two other building blocks: the sphere and the torus, just as we did for the projective plane. Once we have these polygons, we will be able to take any arbitrary identification polygon and recognize how many of each of our building blocks it contains. In doing so, we will see from the identification polygons that any surface can be built from combinations of: the sphere, the torus, and the projective plane.

4.3.1 Bigons

Let us first look at some examples of Bigons.

$\mathbb{R}P^2$

$\mathbb{R}P^2$ has a planar diagram as shown in figure 4.14.

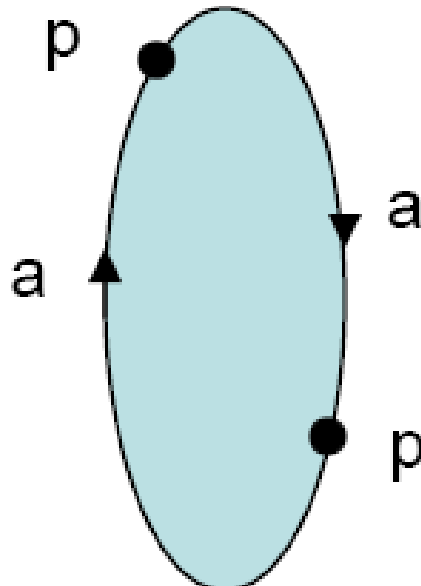


Figure 4.14: The projective plane

The projective plane doesn't exist as a set of points in \mathbb{R}^3 , but the points of the projective plane represent lines in \mathbb{R}^3 . Lines through the origin can then be considered projective points. This means that projective lines are planes through the origin. To analyze the topology of the projective plane, we can surround the origin with a sphere. The lines through the origin will cut the sphere in 2 places (called antipodal points). We can throw away the lines and keep the points. Each line is now represented by a pair of points, and since they represent the same line, we have to identify the pairs of antipodal points. Since every point in one hemisphere has a double in another, we can just work with one hemisphere. But, the points on the equator each have a double. The hemisphere can be flattened to obtain a circle with opposite points identified. Then we can obtain a bigon with the two sides identified (as seen in the planar diagram of \mathbb{RP}^2 in figure 4.14).

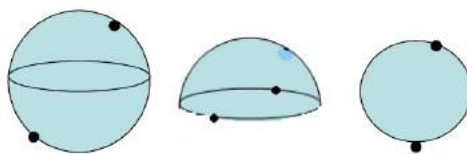


Figure 4.15: Identifying antipodal points

Thus, the projective plane is the quotient of the sphere S^2 under the action of $\mathbb{Z}/2\mathbb{Z} = \{1, -1\}$ defined by

$$1 \cdot x \mapsto x$$

$$-1 \cdot x \mapsto -x, \text{ the antipodal point on the sphere}$$

for $x \in S^2$.

Sphere

The sphere has a planar diagram as shown in figure 4.16

A sphere is the set $\{(x, y, z) : x^2 + y^2 + z^2 = r^2\}$ with the usual topology of \mathbb{R}^3 . The value r is the radius, but it is irrelevant so long as it is positive.

4.3.2 Rectangles

Another class of polygons is rectangles.

Torus

We can define a torus parametrically. A torus is the set $\{(a+b \cos \phi) \cos \theta, (a+b \cos \phi) \sin \theta, b \sin \phi) : 0 \leq \theta, \phi < 2\pi\}$ for a fixed a, b with $a > b$.

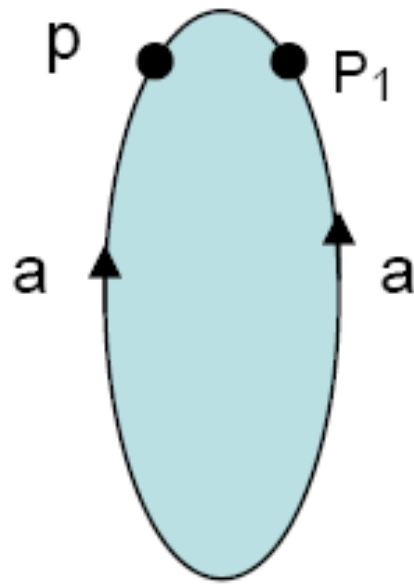


Figure 4.16: Identification polygon of a sphere

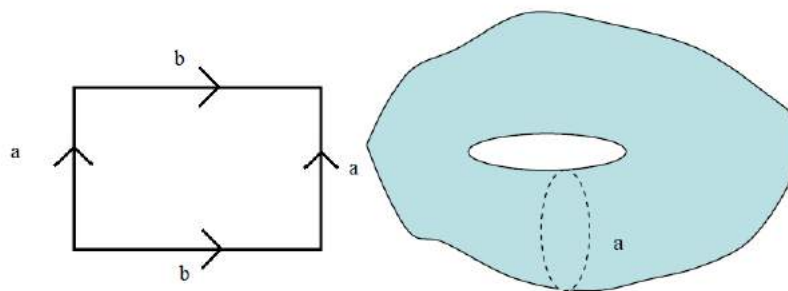


Figure 4.17: Identification polygon of a torus

Klein Bottle

The Klein Bottle can mentally be constructed by taking a cylinder, and gluing together the two circular ends with reverse orientation. This would seem to be a torus, but in order to glue the ends together in this manner, we need to bring one end around and behind the other. This presents a problem: we would have to cut inside the cylinder without passing through the walls - Not possible in \mathbb{R}^3 .

Because it is not possible in \mathbb{R}^3 , there is no subset of \mathbb{R}^3 which is homeomorphic to the Klein Bottle. This is possible in a \mathbb{R}^4 world.

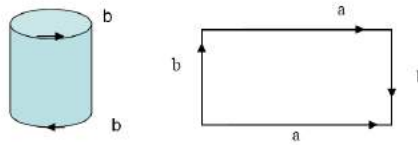


Figure 4.18: Klein Bottle

4.3.3 Working with and simplifying polygons

For any given surface there are many polygonal representations that are possible. There are many different ways to partition a surface, so there are many different representations for a surface that still maintain all of the same properties of that surface. Using several different methods of simplifying surfaces, we can give multiple polygonal representations for it.

How to simplify and work with Polygons: There are many ways that surfaces can be simplified to make working with polygons easier. These ways all represent forms of two methods that together constitute surgery:

Cutting

Cutting a surface is breaking it down into more manageable pieces. Any cut can be repaired by gluing things back together. Cutting is not a homeomorphism, unless it is followed by gluing the edges of the cut as they were before the cut. In the example of the octagon (Figure 4.22), cuts technically haven't been made because when we name two edges the same thing, we are theoretically gluing them together in an identification space.

Gluing

Gluing is the inverse of cutting, as described above. These two processes together can be described as surgery.

Surgery can be carried out in a number of ways to simplify a surface:

- Collecting twisted pairs (Figure 4.19).

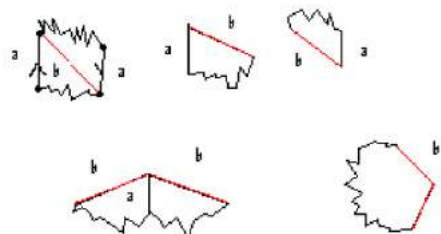


Figure 4.19: Collecting twisted pairs

- Simplifying edges (Figure 4.20).

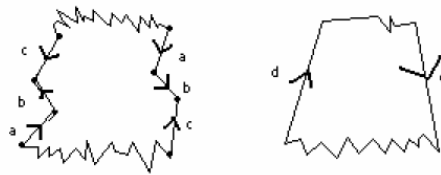


Figure 4.20: Simplifying Edges

- Eliminate adjacent pairs (Figure 4.21)

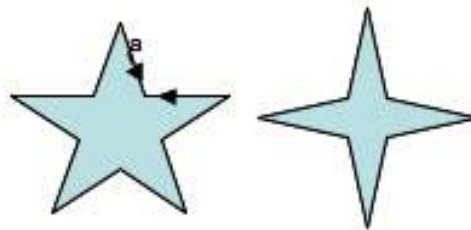


Figure 4.21: Eliminating adjacent pairs

This has eliminated two edges of the surface by canceling the adjacent pair of a's.

Some of the ways of simplifying polygons can be applied to certain octagons to achieve a double-torus: A double-torus is a genus 2 figure because it has 2 holes. The octagon in figure 4.22 is the representative polygon for the surface.

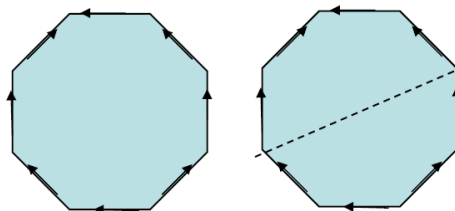


Figure 4.22: Cutting an octagon to recognize it as a genus 2 surface

The octagon is difficult to visualize as a genus 2 surface, so we simplify the octagon by cutting and gluing. We cut along the dashed line and end up with two tori, each with one hole.

We can now glue the two tori together where each hole of theirs' is located (Figure 4.23).

This process can be repeated for any $4n$ -sided polygon, where $4n=g$, to create a genus g figure.

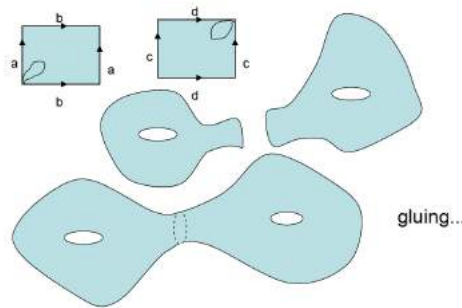


Figure 4.23: Gluing two tori



Figure 4.24: A genus 3 surface

4.4 Orientability

4.4.1 Definition

Orientability is a means of dividing compact surfaces into two classes, orientable and non-orientable and is used to help us classify surfaces. A surface is orientable if it has a top and bottom or inside and outside which are distinguishable. The more formal definition of this would be:

Def: A surface is considered **orientable** if a vector Y perpendicular to the surface at point P can be slid along the surface such that it always remains perpendicular to the surface and when it arrives back at P it points in the same direction as when it started.

Intuitively a cylinder would be an orientable surface since it has a clear inside and outside. According to the definition, as you move the vector Y along the cylinder and back around to P it points in the same direction as it originally did, verifying that a cylinder is indeed orientable.

Our intuition also tells us that a Möbius band would be a non-orientable since there is no clear top or bottom. Using the definition we see that as Y is slid along the Möbius band back to P it actually points in the opposite direction as it initially did, thus verifying that the Möbius band is non-orientable.

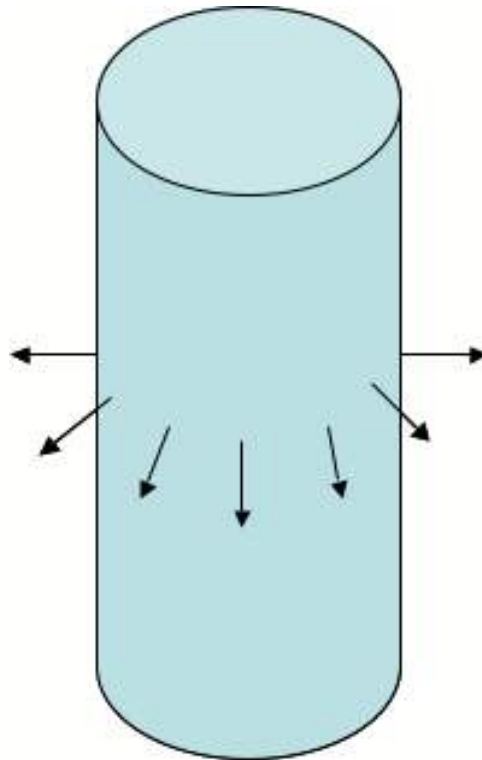


Figure 4.25: A Cylinder

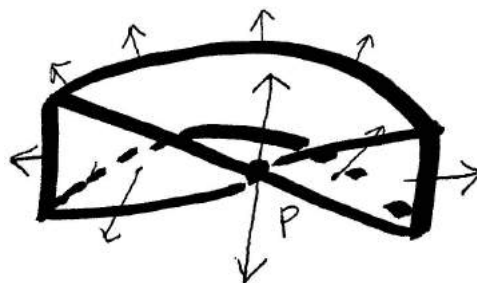


Figure 4.26: Möbius band

4.4.2 Applications To Common Surfaces

Any general surface is defined to be non-orientable if it contains at least one Möbius band and a surface is said to be orientable if it contains no Möbius bands. Since it is much easier to look at the identification polygon of more complicated surfaces it makes sense to look for a Möbius band in a surface's identification polygon rather than by looking at the surface directly. As we showed above the identification polygon of a Möbius band is as shown in figure 4.27.

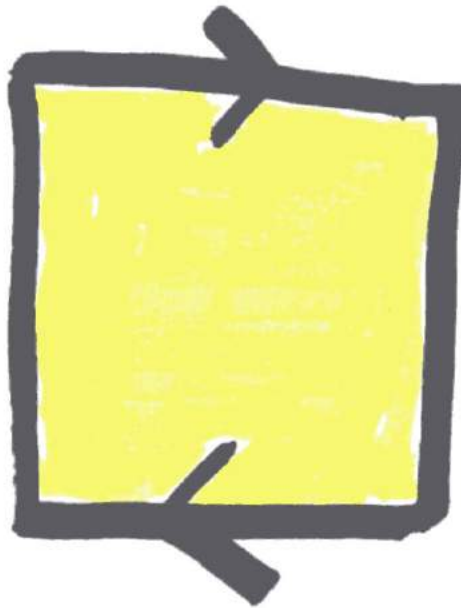


Figure 4.27: Identification polygon of a Möbius band

When looking at an identification polygon of any surface it is enough to find two sides that are attached in opposite directions.

Let's look at some of our common surfaces:

- Sphere (Figure 4.28)

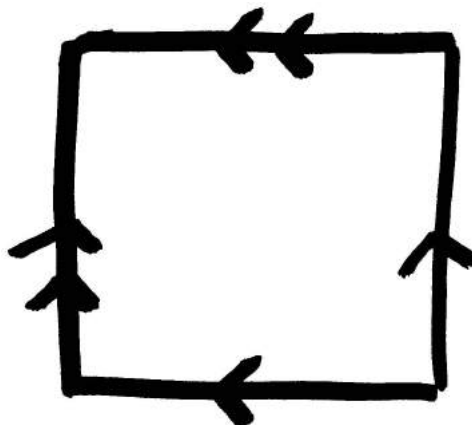


Figure 4.28: Identification polygon of a sphere

Since there is no Möbius band present, the sphere is orientable.

- Torus (Figure 4.29)

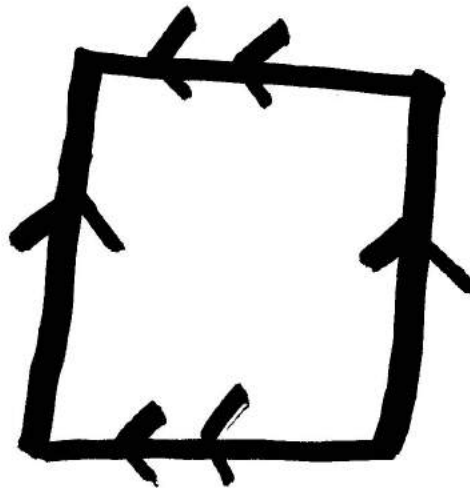


Figure 4.29: Identification polygon of a torus

Since there is no Möbius band present, the torus is orientable.

- Projective Plane (Figure 4.30)

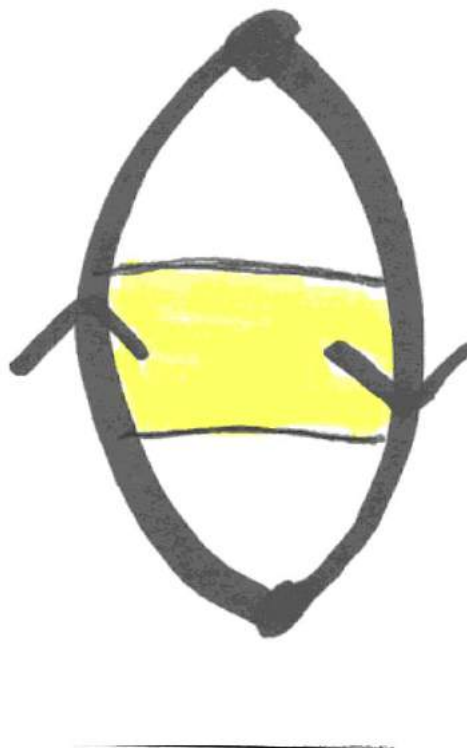


Figure 4.30: Identification polygon of a projective plane

The projective plane does contain a Möbius band as is shown by removing the disk created

by the hatched lines. Since the ends of this disks are identified oppositely, the projective plane contains a Möbius band and is thus non-orientable.

- Klein Bottle (Figure 4.31)

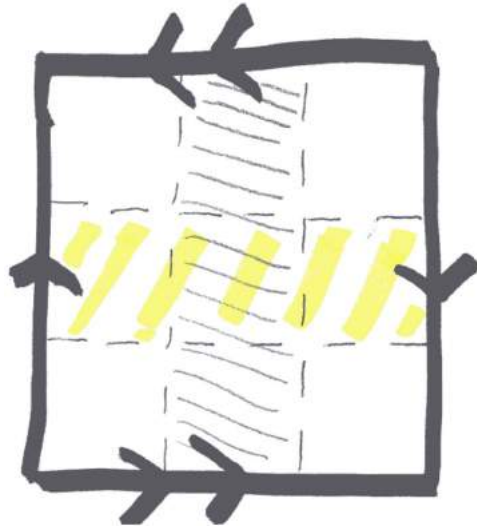


Figure 4.31: Identification polygon of a Klein Bottle

Since the Klein bottle also contains a Möbius band, it is also non-orientable.

4.4.3 Conclusion

Classifying surfaces as orientable or non-orientable is a way to distinguish surfaces from one another. Orientability allows us to split our most basic surfaces that we have used into two categories, one containing the cylinder, the sphere, and the torus and the other containing the Möbius band, the projective plane, and the Klein bottle.

4.5 Euler Characteristic

As we saw in the last section, orientability allows us to divide all compact surfaces into two distinct classes. Euler Characteristic is another way to distinguish compact surfaces from each other.

4.5.1 Requirements

The main requirements for Euler Characteristic are that the surface is compact and has a good graph on it. A good graph on **ANY** surface is a graph, a collection of vertices and edges connecting them on a surface, such that:

1. There is a vertex at every end of each edge.

2. The complement of the graph is homeomorphic to a disjoint union of disks

4.5.2 Computation

To compute the Euler Characteristic(χ) of a surface you count the number of vertices, edges and faces formed by the good graph on the surface and define:

$$\chi = \text{Number of Vertices} - \text{Number of Edges} - \text{Number of Faces}$$

or more simply:

$$\chi = V - E + F$$

It is important to show that the Euler Characteristic of a surface does not depend on the good graph that is placed on it.

Theorem 4.5.1. *The Euler Characteristic of any surface is not dependent on the good graph used to compute it.*

Proof:

There are three actions which can modify any good graph. These three actions are:

1. Add a vertex
2. Add an edge
3. Add a vertex and an edge

Given an arbitrary surface, in this case, a sphere with the good graph in figure 4.32 on it. Counting the number of vertices, edges and faces gives us 6 vertices, 7 edges and the sphere would be split into three faces. We can see that the Euler characteristic would be

$$\chi = 6 - 7 + 3 = 2$$

Adding a vertex

Adding a vertex also splits an edge and therefore increases the number of edges by one. The change in our formula from the original Euler Characteristic would be

$$\chi = 7 - 8 + 3 = 2$$

resulting in the same Euler Characteristic. Therefore adding a vertex does not change the Euler Characteristic.

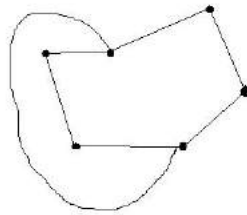


Figure 4.32: A Good Graph

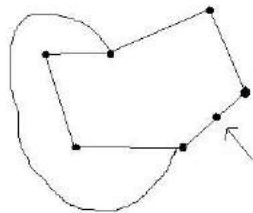


Figure 4.33: Inserting a vertex

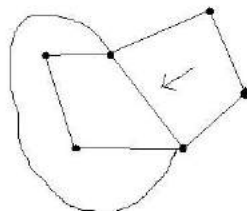


Figure 4.34: Inserting an edge

Adding an edge

Adding a new edge between two existing vertices also splits one face into two. The result-

ing equation would thus be

$$\chi = 6 - 8 + 4 = 2$$

resulting in the same Euler Characteristic. Therefore adding an edge does not change the Euler Characteristic.

Adding a vertex and an edge

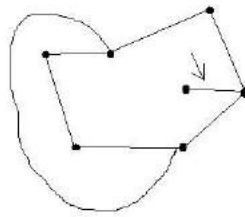


Figure 4.35: Inserting a vertex and an edge

Adding a new vertex and edge increases the number of vertices and edges but does not change the number of faces. The resulting formula would read

$$\chi = 7 - 8 + 3 = 2$$

This also results in the same Euler Characteristic. Since all three possible changes to a good graph do not change the Euler characteristic, the Euler characteristic of a surface does not depend on the good graph on it.

4.5.3 Usefulness

The Euler characteristic is a topological invariant that allows us to distinguish whether or not two surfaces are homeomorphic. Surfaces that have the same Euler characteristic are not necessarily homeomorphic, but if they are both orientable (or non-orientable), then they are. It is important to notice that an identification polygon is a good graph of a surface that has a vertex at every end of edge and there is exactly one face. Therefore, identification polygons are good graphs on a surface and can be used to compute the Euler characteristic.

4.5.4 Use in identification polygons

Given an identification polygon of a surface, it is simple to compute the Euler characteristic. Since identification polygons are good graphs on a surface, we can simply use them as graphs

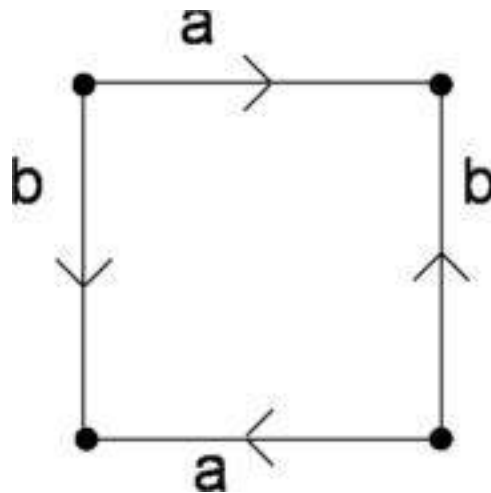


Figure 4.36: The identification polygon and good graph of a projective plane

to compute the Euler characteristic. The general procedure is to draw an edge on each side of the polygon, and a vertex at every corner. The example in figure 4.36 shows a projective plane.

Note that edges drawn on sides of the polygon which are identified are the same edge. So the number of edges in the graph will be the number of unique labels on the sides of the polygon.

Exactly one region, namely the entire polygon, is enclosed by the edges of such a graph, so the number of faces is one (Figure 4.38).

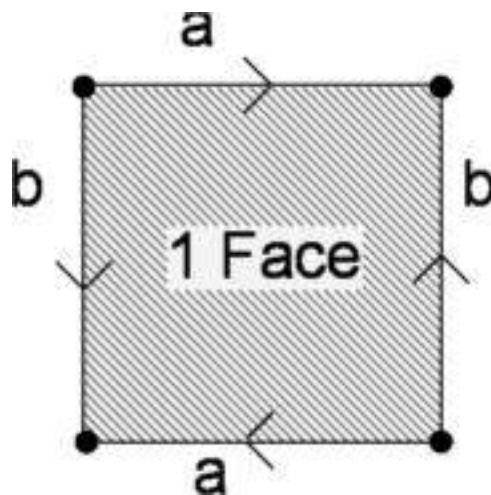


Figure 4.37: Only one face

Therefore we have only to count the number of vertices in the graph. This is only slightly more complicated than counting edges. Vertices drawn at corners which are identified are the same vertex, so we must compute the unique corners. To do this by hand, we pick an arbitrary corner and apply a label. We note the names of the sides which meet at the corner, and whether the point is at the “beginning” of the side or the “end”. Then we apply the same label to the

corresponding point at the beginning or end of the sides with the same labels, again noting any new sides which are adjacent to the point. If any corners remain unlabeled after this process, we again choose an arbitrary unlabeled corner, apply a new label, and repeat the process.

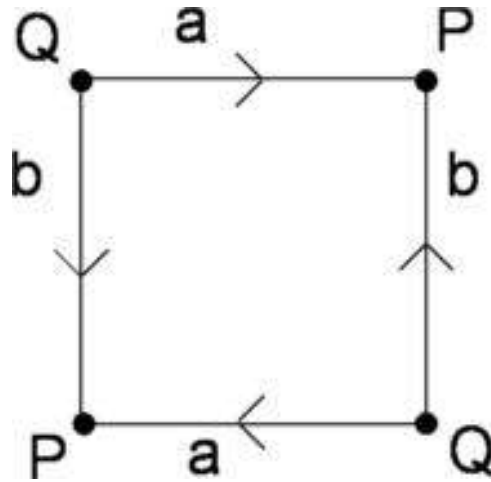


Figure 4.38: Identifying vertices

We now have the number of edges, vertices, and faces of the graph, and can compute the Euler Characteristic. In the case of the projective plane, we have $2 - 2 + 1 = 1$.

4.6 Connected Sums

4.6.1 Definition

If you are given two surfaces you can take the connected sum of the two surfaces to give one single surface. To take the connected sum you remove a disc from each surface and connect up the two resulting boundary circles by a cylinder. Connected sums are denoted by the # sign such that $A\#B$ would be the connected sum of surfaces A and B (assuming that the operation # is well-defined).

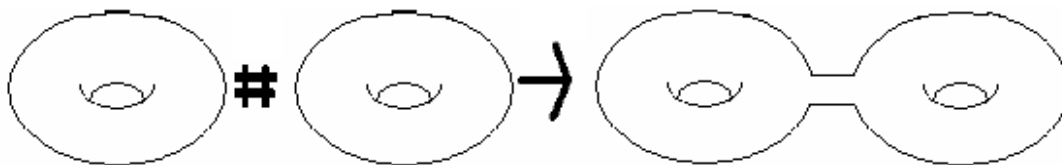


Figure 4.39: A connected sum operation on two tori

4.6.2 Well-definedness

To show that the operation of connected sums is well defined we must show that given two sets that are path connected it does not matter where you take the discs out of in each set; you will

obtain the same resulting surface (same in the sense of homeomorphic of course).

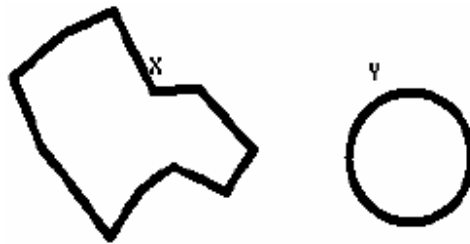


Figure 4.40: Both spaces are path connected

Given this we want to show that if we take any disc out of X and any disc out of Y and we glue them together it will be the same as if we take any other disc out of each surface and glue those together. To do this take the surface X and Y and let the disc to be taken out of Y be fixed. Then take points x_1 and $x_2 \in X$. These two points can be connected by a simple (no loops) path. And we can take a small ball around both x_1 and x_2 .

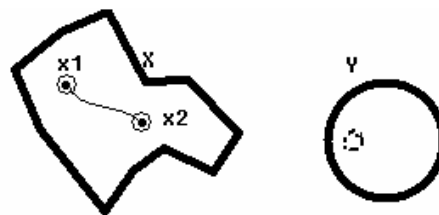


Figure 4.41: Spaces with balls removed

Now let us take a “very small amount of space around” the entire path going from x_1 to x_2 .

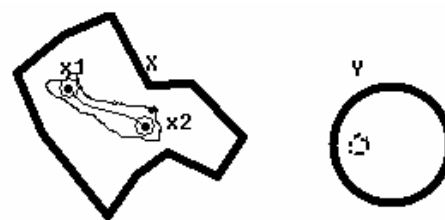


Figure 4.42: Path between resulting holes

If we observe just the closed disc we took around the two points we can relate change it to another surface because it is a homeomorphism to look like figure 4.43.

We are trying to show that taking the ball around x_1 and gluing it to the disc in Y we get the same surface as if you take a ball around x_2 and glue it to the disc in Y . So if we just have x_1 and x_2 in the portion that we took out we need to somehow show that they are equivalent in homeomorphic terms.



Figure 4.43: Area connecting x_1 and x_2

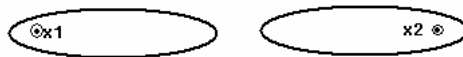


Figure 4.44: Comparing area with x_1 to area with x_2

We can show this because since all we have to do is take these in homeomorphic terms without touching the boundaries we can push and stretch the insides of the boundary. If we think in terms that this is rubber we can stretch it such that if we have a line from the boundary to the point x_1 we can stretch or loosen the boundary enough so that we have the exact position of x_2 .



Figure 4.45: Stretching disk around x_1 to a disk around x_2

If we observe in line 1 and 2 we stretch it to point x_2 such that x_1 are the same and in x_3 we loosen it. So then we can determine that gluing a disc around each point, or any disc on the surface X to the disc removed in surface Y produces the same resulting surface because they are homeomorphic.

So then if we take any disc in X and any disc in Y and glue them we can show that they produce the same resulting space because they are homeomorphic as long as X and Y are path-connected. Thus we conclude that the connected sum operation is well defined for path-connected spaces.

4.6.3 Examples

Let's look at a few examples of connected sums.

First a sphere $\#$ a sphere is yet another sphere (Figure 4.46).

A torus $\#$ another torus forms a surface with two holes. We call this a genus 2 surface (Figure 4.47).

Finally, two projective planes produce a Klein bottle (Figure 4.48).

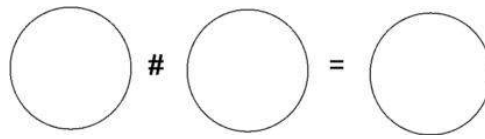


Figure 4.46: Sphere # Sphere

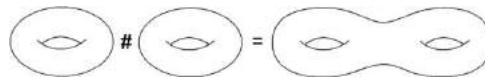


Figure 4.47: Torus # Torus = Double Torus

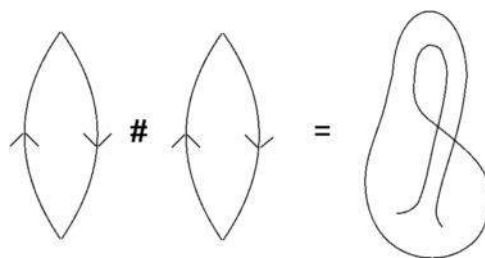


Figure 4.48: Two projective planes form a Klein bottle

4.6.4 $\mathbb{R}P^2 \# T = \mathbb{R}P^2 \# \mathbb{R}P^2 \# \mathbb{R}P^2$

First we start of with the connected sum of a projective plane and a torus, and we make a cut, removing two sides (Figure 4.49).

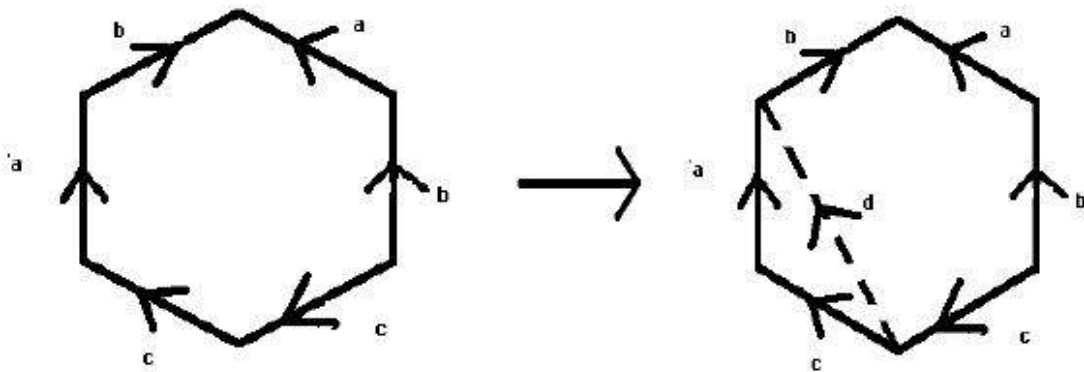


Figure 4.49: First Cut

Then we are left with two surfaces, which we glue together, and make another cut in the resulting polygon (Figure 4.50).

We glue again and observe that we are left with a projective plane minus a disc and a Klein

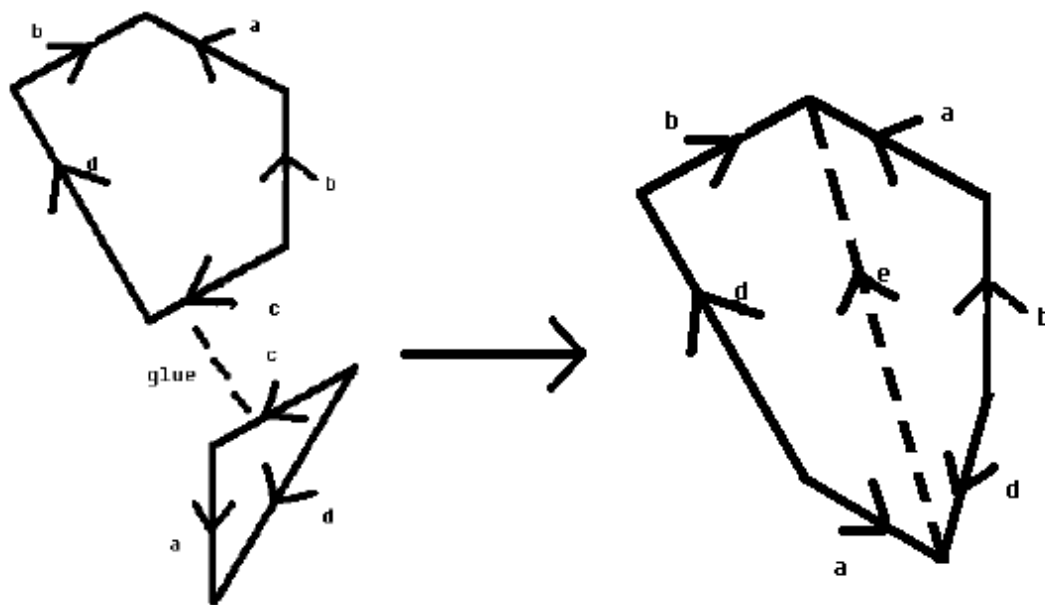


Figure 4.50: Gluing and making another cut

bottle minus a disc (Figure 4.51).

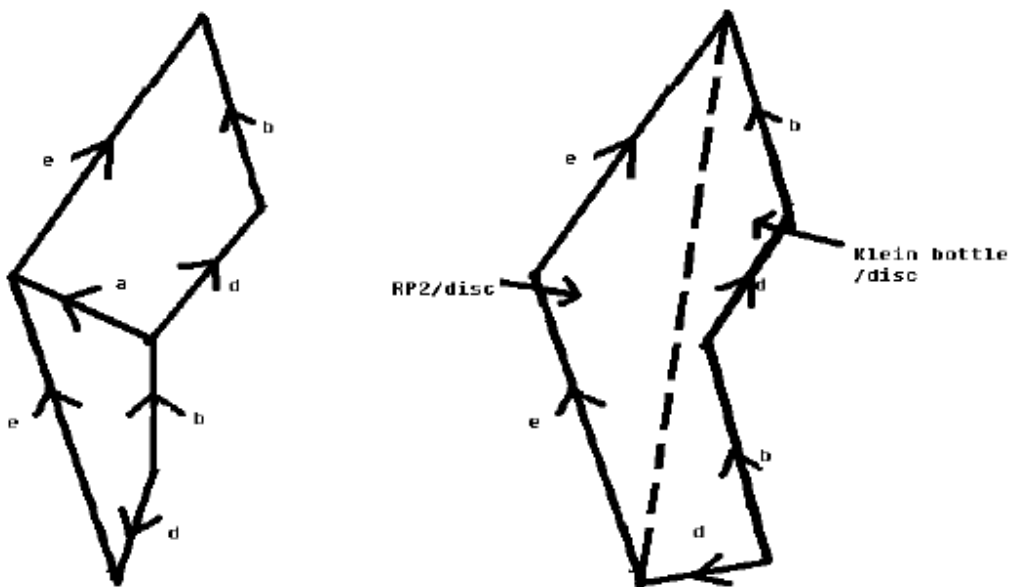


Figure 4.51: The final glue and resulting surface

4.6.5 Associativity

An important property of the connected sum operation is that it is associative. We can see this by considering three surfaces S_1, S_2, S_3 and the connected sum $(S_1 \# S_2) \# S_3$. Consider what would happen if we performed the operation in the other order: $S_1 \# (S_2 \# S_3)$. We know that where the disc is removed does not impact the result of the connected sum. So we can require that, in the first order, we connect S_3 to a disc taken from part of the surface which “used to be” S_2 , and in the second case, we connect S_1 to a part which “used to be” S_2 . Then in both cases, S_1 and S_3 were connected to part of S_2 , and the order in which they were added doesn't matter.

4.6.6 Effect on Euler Characteristic

We have studied the Euler Characteristic of surfaces, and seen that it will be a very useful property for classifying surfaces. However, to compute the Euler Characteristic of a complicated surface by drawing a good graph and counting the vertices, edges, and faces would be tedious. The connected sum gives us a simple way to construct new surfaces, so we will examine what effect the operation has on the Euler Characteristic.

Consider surfaces S_1 and S_2 with good graphs drawn on them as shown in Figure 4.52.



Figure 4.52: Two surfaces with good graphs

To determine the Euler Characteristic of their connected sum, we carefully choose the places on each to remove discs and attach the cylinder. We choose a disc which borders on exactly one vertex and is disjoint from any other part of the graph (Figure 4.53)

Next we add to the graph on the connected sum in order to restore the properties of a good graph. We add an edge around each base of the cylinder, connecting the vertex on the border to itself. Finally we add an edge from the vertex on one edge of the cylinder to the vertex on the other edge of the cylinder. By adding these edges we divide the cylinder so that the complement of the graph is a disc. The final graph is shown in figure 4.54.

Now let us compute the Euler Characteristic of the result of the operation. Let $m = \chi(S_1)$

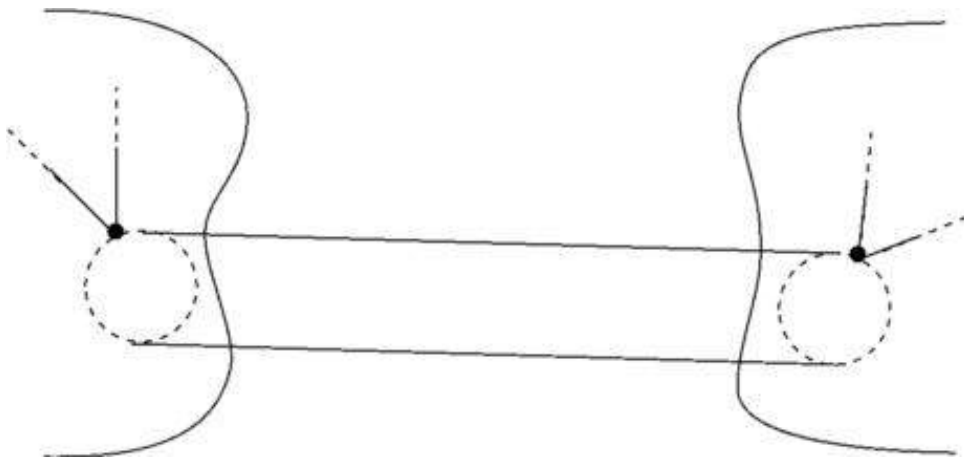


Figure 4.53: Carefully attaching a cylinder

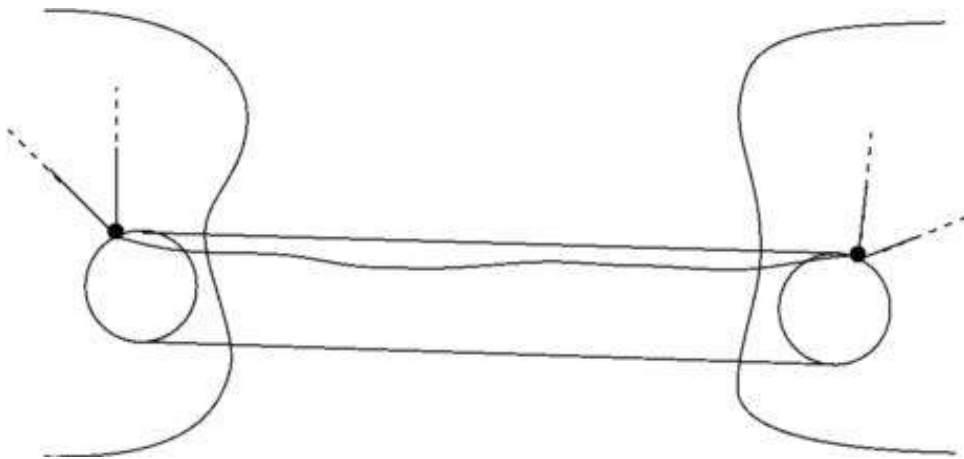


Figure 4.54: A good graph on the connected sum

and $n = \chi(S_2)$. Since we haven't removed any vertices, edges, or faces, $\chi(S_3 = S_1 \# S_2)$ will be $m+n$ plus the new vertices, edges, and faces. We did not add any vertices, However, we added three edges and a face on the cylinder, so our final Euler Characteristic is $m + n - 3 + 1 = m + n - 2$. The Euler Characteristic of the connected sum of two surfaces is two less than the sum of the Euler Characteristics of the two surfaces.

Using this result, we can conclude with the general form of the Euler Characteristic of some useful connected sums.

- Connected sum of tori

$$\chi(T) = 0.$$

$$\chi(T \# T) = 0 + 0 - 2.$$

$$\chi(\underbrace{T \# \cdots \# T}_g \text{ tori}) = 2 - 2g.$$

- Connected sum of $\mathbb{R}P^2$

$$\chi(\mathbb{R}P^2) = 1.$$

$$\chi(\mathbb{R}P^2 \# \mathbb{R}P^2) = 1 + 1 - 2 = 0.$$

$$\chi(\underbrace{\mathbb{R}P^2 \# \cdots \# \mathbb{R}P^2}_{m \mathbb{R}P^2}) = 2 - m.$$

4.7 Classification Theorem

Theorem: Any closed surface is homeomorphic either to (No two of the following are homeomorphic):

1. The Sphere
2. The sphere with a finite number of handles added
3. The sphere with a finite number of discs removed and replaced by Möbius strips.

4.7.1 Equivalent definitions

This is the classification theorem and all surfaces are homeomorphic to one of these. However, there are other ways to express the classification theorem.

- (a) Any closed surface is homeomorphic either to (No two of the following are homeomorphic):
1. The sphere
 2. A connected sum of tori
 3. A connected sum of projective planes
- (b) Any closed surface is homeomorphic either to (No two of the following are homeomorphic):
1. The sphere
 2. A connected sum of tori
 3. A connected sum of a tori and one projective plane
 4. A connected sum of a tori and two projective planes

Before showing that these are the same as the original classification theorem it is easier to show that these two are equivalent and then show that both (a) is equivalent to the classification theorem and thus (b) is also equivalent to the classification theorem.

To show that (a) and (b) are equivalent use what we found in the section about connected sums that: $\mathbb{R}P^2 \# \mathbb{R}P^2 \# \mathbb{R}P^2 \cong \mathbb{R}P^2 \# T$ (Connected sum of 3 projective planes is homeomorphic to the connected sum of a projective plane and a torus.)

If we look at this relation we can observe that we can essentially trade off a torus for two projective planes and trade two projective planes for a torus. Now in the process of going from (a) to (b) if you have at least 3 projective planes in your surface you can go about trading two

of those projective planes until you no longer have 3 projective planes in your surface. If you do so you end up with the list in (b).

Going from (b) to (a) is essentially the same where you trade off a torus one at a time for two projective planes until you have list (a).

For the equivalence between (a) and the original theorem if you note that a connected sum $S\#T$ is equivalent to attaching a cylinder to S with two discs removed. Whereas making a connected sum $S\#\mathbb{R}P^2$ means we are gluing the projective plane minus a Möbius strip, which is a Möbius strip, to S with a disc removed. Thus we find that (a) and the classification theorem are equivalent and thus (b) is also equivalent to the theorem.

4.7.2 Proof

OK, so where is all of this leading? We have a lot of content (more surfaces than your heart can desire) but still very little form (nothing to bring it all together). Thus far, we are capable of manipulating/identifying surfaces from polygons and able to create new surfaces by connect summing them. But we still don't have a complete understanding of compact surfaces. It would be nice to have some way of classifying all the possible compact surfaces.

So far we have dealt with the sphere, the torus and the projective plane. I am going to claim that you can make any surface from the connected sum of these three surfaces. In order to see if this is true, we need to show that A) there are no compact surfaces that are left out B) All these surfaces are distinct.

A) Big Assumption Let us assume that every compact closed surface can be triangulated. If we accept this fact, then by identifying all the sides of all the triangles, we can construct an identification polygon from any surface. Now all we need to establish is that all polygons can be represented as the connected sum of the sphere, tori and $\mathbb{R}P^2$'s. We will prove this by induction. We have already established that all identification bigons and quadrilaterals are well classified as connected sum of sphere, tori and $\mathbb{R}P^2$. If we take a generic polygon P , with N sides, that is well classified, we need to show P' with $N+2$ sides will be a combination of well classified polygons.

Induction Proof: Assume that a generic polygon P with N sides can be created as the connected sum of a sphere, tori and $\mathbb{R}P^2$, show that the polygon P' with $N+2$ sides is still the connected sum of P , sphere, tori and $\mathbb{R}P^2$.

(DIAGRAM BELOW!!!)

A Two adjacent sides in opposite direction collapses into an N -Gon

B Two adjacent sides in same direction can be cut to make the connected sum of an N -Gon and $\mathbb{R}P^2$.

C Two non-adjacent sides in opposite directions can be cut to result in the same as B

D Orientable polygon with all sides identified in same direction can be sliced into a connected sum of tori missing 2 discs and a cylinder

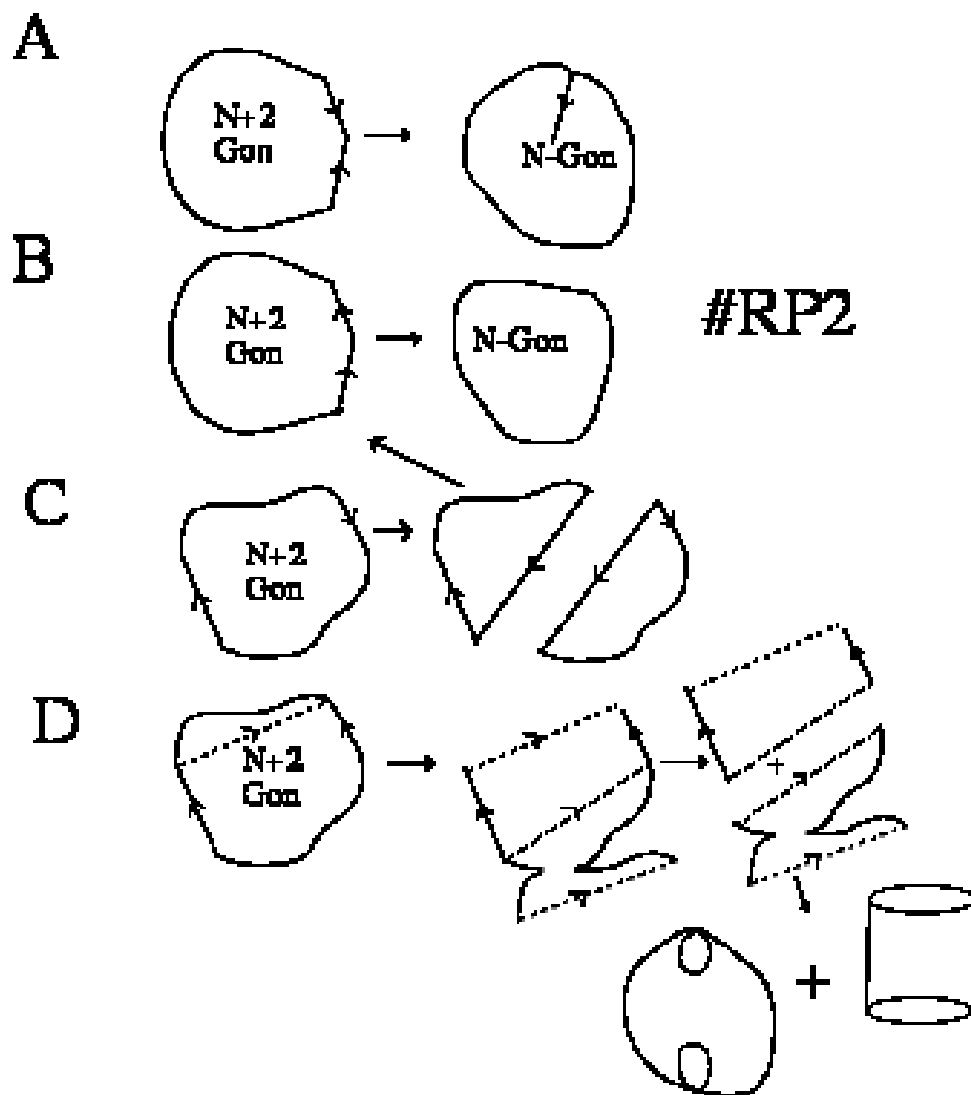


Figure 4.55: Bulky diagram Interrupting the Beauty of the Classification Theorem!

B) All surfaces are distinct We showed that the tori, $\mathbb{R}P^2$ and the Sphere are distinct in the previous section on Euler characteristic.

Ok, so we have established that any surface can be constructed from the connected sum of a sphere, tori and $\mathbb{R}P^2$, but is that the best we can do?

First of all, what does it mean to connect sum a surface P with a sphere? Since a sphere minus a disc is homeomorphic to a disc, $P\#S = P$. So connect summing the sphere with any surface is trivial.

We know that $T\#\mathbb{R}P^2 = \mathbb{R}P^2\#\mathbb{R}P^2\#\mathbb{R}P^2$. So if a surface contains a projective plane, each projective plane can "cancel" a torus to create three projective planes. Consequently, any surface containing a projective plane, can be represented purely as the connected sum of projective planes. This simplifies classification theorem to a sphere, the connected sum of $\mathbb{R}P^2$ and the

connected sum of tori.

Classification Theorem of Compact Surfaces Any closed surface is homeomorphic to

1. The sphere
2. Connected sum of tori
3. Connected sum of projective planes

The surfaces in the first two families are orientable. Since the sphere and the torus have Euler characteristics 2 and 0, respectively, it follows that the Euler characteristic of the connected sum of g tori is $2 - 2g$.

As we discussed earlier, the surfaces in the third family are nonorientable. Since the Euler characteristic of the real projective plane is 1, the Euler characteristic of the connected sum of k of them is $2 - k$.

In this way, knowing the Euler characteristic and the orientability of a surface is all that is needed to classify a closed surface.

Chapter 5

Homotopy and the Fundamental Group

5.1 Homotopy of functions

Introduction

This section introduces the notion of two functions being homotopic.

Definition 1. Let $f, g : X \rightarrow Y$ be maps. f is homotopic to g if there exists a map $F : X \times I \rightarrow Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all points $x \in X$.

The map F is called a homotopy from f to g and we write it as $f \stackrel{\sim}{\underset{F}{\rightsquigarrow}} g$.

More intuitively, if we think of the second parameter of F as “time”, then F describes a “continuous deformation” of f into g . At time 0 we have the function f , at time 1 we have the function g .

Example 5.1.1. *Enlarging a circle:*

Given $f(\theta) = (\cos(\theta), \sin(\theta))$ and $g(\theta) = (2 \cos(\theta), 2 \sin(\theta))$ for $0 \leq \theta \leq 2\pi$, using polar coordinates.

We can define the homotopy as $F(\theta, t) = ((1 + t) \cos(\theta), (1 + t) \sin(\theta))$.

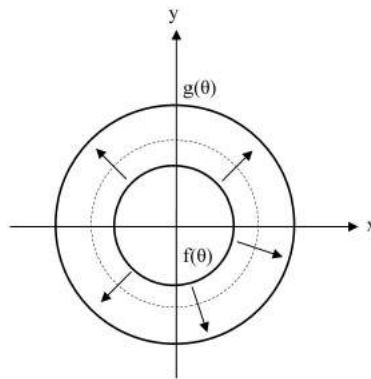


Figure 5.1: Enlarging a circle

Lemma 5.1.2. Any mapping $f : X \rightarrow \mathbb{R}$ is homotopic to the constant zero mapping. The required homotopy is $H(x, t) = (1 - t)f(x)$. f is **null-homotopic** if and only if $f \simeq$ (constant map).

Example 5.1.3. Compressing the real line interval $[-100, 100] \in \mathbb{R}$ into a point $\{0\}$.

The required homotopy would be $F(x, t) = x(1 - t)$.

Definition 2. If f and g agree on a subset A of X , we may define the homotopy from f to g with the additional property

$$F(a, t) = f(a) \quad \forall a \in A, \forall t \in I$$

In this case we say that f is homotopic to g relative to A and write $f \underset{F}{\simeq} g \text{ rel } A$.

Example 5.1.4. Homotopy between two discs

Let D_1 be a disc of radius 1, and D_2 be a disc of radius 2.

If we define the homotopy from D_2 to D_1 by mapping the overlapping points to themselves and the rest of the points to the corresponding point on the boundary of D_1 , we have obtained a homotopy $D_2 \simeq D_1 \text{ rel } D_1$.

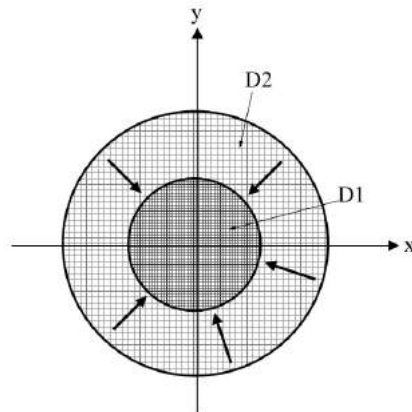


Figure 5.2: Homotopy between two discs

Lemma 5.1.5. *Suppose we have two loops $\alpha, \beta : I \rightarrow X$ based at the same point p in X . Continuously deforming α into β without moving the base point p is the same as asking for a homotopy such that $\alpha \stackrel{F}{\sim} \beta \text{ rel } \{0, 1\} \in I$.*

The homotopy is a map F from the square $I \times I$ to X which sends the bottom of the square via α , the top via β and the two vertical sides to the base point p .

By sending the two vertical sides to one point, any line linking the two vertical sides will become a loop based at p . Sliding from the bottom of the square to the top will give us a continuous family of loops starting at α and ending at β .

Figure 5.3 illustrates this situation for two loops on a torus.

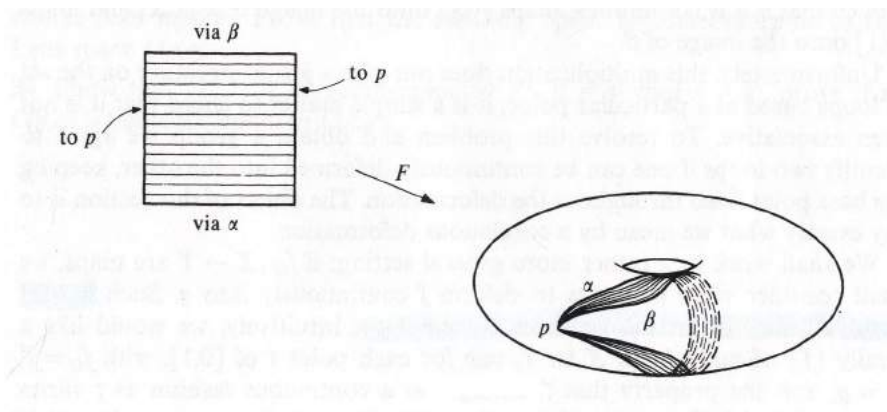


Figure 5.3: Two loops on the torus

Theorem 5.1. *Homotopy is an equivalence relation.*

Proof: We need to show that a homotopy is

- Reflexive

Proof: It is trivial to show that $f \simeq f$. The required map is $F(x, t) = f(x)$.

- Symmetric

Proof: Let $f \simeq g$, thus there exist a map $F : X \times I \rightarrow Y$ such that

$$F(x, 0) = f(x)$$

$$F(x, 1) = g(x) \text{ for all points } x \in X.$$

Define $G : X \times I \rightarrow Y$ such that $G(x, t) = F(x, 1 - t)$. Therefore,

$$G(x, 0) = g(x)$$

$$G(x, 1) = f(x)$$

which implies that $g \simeq f$.

- Transitive

Proof:

Let $f \stackrel{\sim}{\underset{F}{\simeq}} g$ and $g \stackrel{\sim}{\underset{G}{\simeq}} h$ be maps from X to Y .

$$\text{Define } H(x, t) = \begin{cases} F(x, 2t), & \text{if } t \in [0, \frac{1}{2}] \\ G(x, 2t - 1), & \text{if } t \in [\frac{1}{2}, 1] \end{cases}.$$

By simple substitution, we get the following results:

$$H(x, 0) = F(x, 2t) = f(x)$$

$$H(x, \frac{1}{2}) = F(x, 1) = G(x, 0) = g(x)$$

$$H(x, 1) = G(x, 1) = h(x)$$

Since the deformations are continuous within the two intervals of t , we can conclude that f is deformed into g from $t \in [0, \frac{1}{2}]$ and then deformed into h from $t \in [\frac{1}{2}, 1]$. Therefore $f \stackrel{\sim}{\underset{H}{\simeq}} h$.

Definition 3. Two paths f and g , mapping the interval $I = [0, 1]$ into X , are said to be **path homotopic** if

- they have the same initial point x_0 and final point x_1 .
- there is a continuous map $F : I \times I \rightarrow X$ such that

$$F(s, 0) = f(s)$$

$$F(s, 1) = g(s)$$

$$F(0, t) = x_0$$

$$F(1, t) = x_1$$

$$\forall s, t \in I.$$

We call F a path homotopy between f and g and write $f \simeq_p g$.

5.2 The Fundamental Group

5.2.1 Free Groups

The *fundamental group* of a topological space is the group formed by the multiplication of homotopy equivalence classes of loops. Knowing the fundamental group of a space can be very

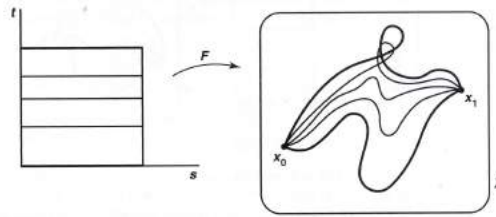


Figure 5.4: Path Homotopy

powerful because, as we’ll soon see, the fundamental group is a topological invariant and a homotopic invariant. Our goal is to devise a way to describe the fundamental group. Before we talk about the fundamental group, let’s look first at the free group and get an idea of group theory first.

Construction of the free group in n elements, denoted F_n :

First start off with n letters $a_1, a_2, a_3, \dots, a_n$

Then add $\bar{a}_1, \bar{a}_2, \bar{a}_3, \dots, \bar{a}_n$ such that $a_i \bar{a}_i = 1$.

The elements of F_n are then words comprised of the letters and compositions thereof. Each element can be expressed in only one non-trivial way. To check that it is in fact a group: Identity: $1 = \emptyset$ (\emptyset would correspond to the empty word, or the word with no letters.) Inverse: As established, the inverse for every a_i is \bar{a}_i and vice versa. Associativity $a \cdot (bc) = (ab) \cdot c$

Example - Free Group in Four Elements $F_4 = \{a, b, c, d, \bar{a}, \bar{b}, \bar{c}, \bar{d}\}$ Some examples of words in F_4 are $adc\bar{b}d\bar{c}$ or $\bar{a}cd\bar{c}bab$

Example - Free Group in One Element $F_1 = \{a, \bar{a}\}$ is pretty straightforward as far as what are the elements of the group, but it is useful in seeing that the free groups can be used to represent other groups we all know and love. It’s not always easy to figure out, but here it kind of is. If we think of a as 1 and \bar{a} as -1, and strings of these elements being addition, then F_1 is a means of describing \mathbb{Z} .

5.2.2 Graphic Representation of Free Group

The free group (with at least two generators) is not abelian by definition. Every element can be written (in a reduced form, i.e. without adjacent pairs of the form $x\bar{x}$) one and only one way. Therefore if the group were abelian, then $ab = ba$, which would be two non-trivial ways of writing the same element, contradicting the definition of the group.

In fact we can get a good picture of the group using a Cayley graph for F_2 :

As we can see, vertices of the Cayley graph are in bijection with elements of F_2 .

We can also think of the group F_n as “all possible ways of going around” n circles with one point in each identified with each other, or a bouquet of n circles. We can imagine every loop in the bouquet as corresponding to an element in F_n and define counterclockwise (arbitrarily) around a the i -th loop as a_i and clockwise around the same loop as \bar{a}_i . For example F_4 :

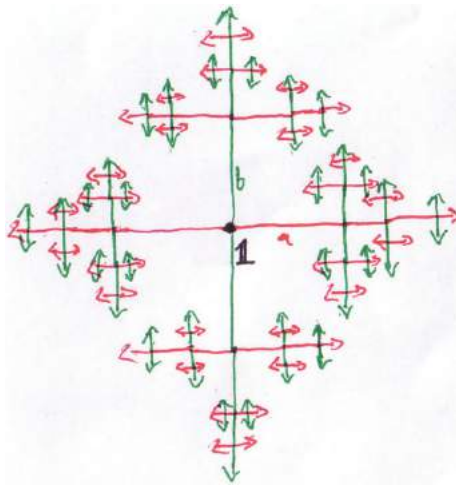


Figure 5.5: Cayley Graph for F_2

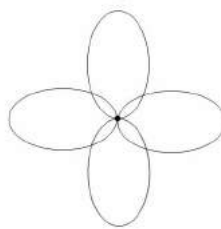


Figure 5.6: A bouquet of 4 circles

Theorem 5.2.1. *For any group G , and any choice of n elements in G , we can find a unique homomorphism h such that*

$$\begin{array}{l}
 h: F_n \longrightarrow G \\
 a_1 \longrightarrow g_1 \\
 a_2 \longrightarrow g_2 \\
 \vdots \\
 a_n \longrightarrow g_n
 \end{array}$$

Note that after a few months in a topology course, it's easy to see homomorphism and have your brain think homeomorphism. The unique homomorphism in this case need not be one-to-one (and won't be in general.)

Example - $\mathbb{Z}/5$ (The Additive Group) $F_5 = \{a, b, c, d, e, \bar{a}, \bar{b}, \bar{c}, \bar{d}, \bar{e}\}$
 Define a homomorphism, $h: F_n \longrightarrow G$. h maps:
 $a \longrightarrow 1$

$b \longrightarrow 2$
 $c \longrightarrow 3$
 $d \longrightarrow 4$
 $e \longrightarrow 5$

But of course here, for example, we would also have to have \bar{d} map to something. And if \bar{d} is the inverse of d , then clearly $\bar{d} = 2 = c$, and I think this is a good way to segue into

5.2.3 Presentation Of A Group

We present a group as:

$G = \langle \text{generators} \mid \text{relations} \rangle$ Where the generators are a set of elements that allows to produce every element in the group, and the relations are the set of all non-trivial elements identified with the identity element (the trivial examples would be things like $a_i \bar{a}_i$.)

The presentation of a free group, e.g. F_5 (to continue the example,) would be:

$$F_5 = \langle a, b, c, d, e \rangle$$

Here the group is just represented by the generators because, of course, the free group has no relations.

And if we look at the presentation of the additive group $\mathbb{Z}/5$:

$$\mathbb{Z}/5 = \langle a \mid a^5 \rangle$$

Which we know from group theory that, as 5 is a prime number, we can find a primitive root in the group that would generate the entire set (for example, $a = 3$.)

The main goal of all this is to find a useful means of determining fundamental groups. So, let h be an onto homomorphism from F_n to G . Define then the kernel of a group homomorphism to be

$K = \{r \mid h(r) = 1\}$ The kernel is a normal subgroup of G , i.e. $Ka = aK \forall a \in G$ If we quotient out the kernel from F_n we find that:

$$G \cong F_n/K = \langle a_1, a_2, \dots, a_n \mid r_1, r_2, \dots, r_m \rangle$$

Returning to the example. $\langle a \mid a^5 \rangle = \mathbb{Z}/5 \cong F_5/K = \langle a_1, a_2, a_3, a_4, a_5 \mid r_1, r_2, \dots, r_m \rangle$

Let the generator of the additive group $\mathbb{Z}/5$ be 3, and for h as defined above, let $h(a_i) = i$. The relations are then (with some algebra) $R = \{a_5, a_3 a_2, a_4 a_1, a_3^2 \bar{a}_4, a_3^3 \bar{a}_1, a_3^4 \bar{a}_2, a_3^5, \dots\}$. The list goes on, but it is easy to see that the end result could be expressed as a single element generating the whole set.

5.2.4 The Fundamental Group

After all that, we can now talk about the fundamental group of a topological space.

Theorem 5.2.2. *The set of homotopy classes of loops in topological space X based at point p forms a group under operation $\langle\alpha\rangle\cdot\langle\beta\rangle = \langle\alpha\cdot\beta\rangle$ where $\langle\alpha\rangle$ and $\langle\beta\rangle$ are the homotopy equivalence classes of α and β , and the operation consists of “concatenating loops”.*

This group is called the fundamental group and denoted $\pi_1(X, p)$, and for path connected spaces, the fundamental groups based at any two points are isomorphic to each other, so $\pi_1(X)$ would suffice notation-wise. I’ll prove this in a slightly non-rigorous way, as it’s fairly intuitive. Let p, q be two points in a path-connected space X . We can find a path, γ , betwixt p and q , and therefore we can define a loop, α , based around q in terms of a loop based around p by simply looking at $\langle\gamma^{-1}\cdot\alpha\cdot\gamma\rangle$.

We can also move from talking about the fundamental group in a space X to describe the fundamental group in a space Y (both path-connected) when there exists a continuous function $f : X \rightarrow Y$, choose $p \in X$ and $q = f(p) \in Y$. A loop α based at p in X is mapped to $f \circ \alpha$ based at q in Y . We know that homotopy equivalence and loop composition is preserved by homomorphisms, we define f_* as:

$$f_* : \pi_1(X, p) \rightarrow \pi_1(Y, f(p)) \text{ as } f_*(\langle\alpha\rangle) = \langle f \circ \alpha \rangle$$

Theorem 5.2.3. *Let X, Y, Z be path-connected topological spaces. Define functions f, g such that $X \xrightarrow{f} Y \xrightarrow{g} Z$ and $p \in X, f(p) = q \in Y, g(q) = r \in Z$. $(g \circ f)_* = g_* \circ f_*$*

Proof First off, let’s take a look at what $(g \circ f)_*$ and $g_* \circ f_*$ mean to get an idea of why this is true:

$$(g \circ f)_* : \pi_1(X, p) \rightarrow \pi_1(Z, r)$$

$$g_* \circ f_* : \pi_1(X, p) \xrightarrow{f_*} \pi_1(Y, q) \xrightarrow{g_*} \pi_1(Z, r)$$

Now to prove the theorem using algebra and the above definition of f_* , let α be a loop in X , $(g \circ f)_*(\langle\alpha\rangle) = \langle(g \circ f) \circ \alpha\rangle = \langle g \circ (f \circ \alpha)\rangle = g_*(\langle f \circ \alpha \rangle) = g_*(f_*(\langle\alpha\rangle)) = (g_* \circ f_*)(\langle\alpha\rangle)$

A special case of the above theorem would be if the function f is a homeomorphism such that $X \xrightarrow{h} Y \xrightarrow{h^{-1}} X$ and $Y \xrightarrow{h^{-1}} X \xrightarrow{h} Y$ we can see:

$$f_*^{-1} \circ f_* : \pi_1(X, p) \rightarrow \pi_1(X, p) f_*^{-1} \circ f_* = (f^{-1} \circ f)_* = (1_X)_*$$

and $f_* \circ f_*^{-1} : \pi_1(Y, f(p)) \rightarrow \pi_1(Y, f(p)) f_* \circ f_*^{-1} = (f^{-1} \circ f)_* = (1_Y)_*$

The identity here in X and Y is induced by the identity function, so f_* is clearly one-to-one and onto. Therefore, we can see that f_* is an isomorphism, that is two path-connected homeomorphic spaces have isomorphic fundamental groups. Hence, the fundamental group is a topological invariant, as if two spaces don’t have the same fundamental group, they can not be homeomorphic.

5.3 Homotopy Equivalence between Spaces

Introduction

In this section we will deal with the notion of two topological spaces being homotopy equivalent. This gives a way to talk about two spaces being “the same,” using the idea of homotopic maps, although we will see that this is not such a restrictive condition as homeomorphism.

Definition 5.3.1. *Two spaces are said to be homotopy equivalent if there are maps $f : X \rightarrow Y$ and $g : Y \rightarrow X$ such that $f \circ g \simeq I_Y$ and $g \circ f \simeq I_X$, where I is the identity map, and \simeq denotes homotopy between functions. Homotopy equivalence is denoted by \sim .*

Example 5.3.2. *Take D , the disk, and P , a point, as our two spaces. The claim is that $D \sim P$. See Figure 5.7.*

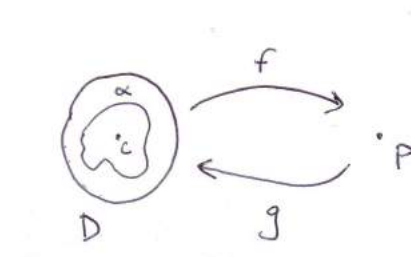


Figure 5.7: Homotopy equivalence of a disc and a point

If α is any point in D , then $f(\alpha)$ is clearly just P . Similarly, $g(P)$ is some point in D , say the center for simplicity. So, $f \circ g = P$, and therefore $f \circ g \simeq I_Y$.

For $g \circ f$, we have that $(g \circ f)(\alpha) = C$, where C is the center of the disk. Radial retraction shows that the map $g \circ f$ is homotopic to I_X .

Informally, spaces are homotopy equivalent when one can be continuously “shrunk” into the other. For example, a “hairy” sphere is homotopy equivalent to the sphere since all of the hairs can be pushed back to the surface. Since the maps between spaces are not necessarily invertible, this is possible. One could not, for example, construct maps between a surface of genus 2 and a torus to show that they are homotopy equivalent. This is because we are not able to shrink or expand the surface to get rid of one of the holes. And, as we will see below, all of the tricks to visually determine if two spaces are homeomorphic still apply.

5.3.1 Homeomorphism vs. Homotopy Equivalence

Claim 5.3.3. *If $X \cong Y$, then $X \sim Y$. In other words, if two spaces are homeomorphic, then they are homotopy equivalent.*

Proof. Since X and Y are homeomorphic, there exist bijective maps $F : X \rightarrow Y$ and $F^{-1} : Y \rightarrow X$ such that $F \circ F^{-1} = F^{-1} \circ F = I$. Hence we have exhibited maps whose compositions are in fact equal to the identity map, and so they are homotopic. Therefore $X \sim Y$. \square

Note that the converse of this claim is not necessarily true. We saw in Example 1 that $D \sim P$, but it is not true that $D \cong P$.

5.3.2 Equivalence Relation

The notion of homotopy equivalence defines an equivalence relation on the set of spaces. It is clear that $X \sim X$. This is always true because there is always the identity map from X to X .

It is also clear that $X \sim Y \Rightarrow Y \sim X$. This follows since the definition of homotopy equivalence requires that both $f \circ g$ and $g \circ f$ are homotopic to the respective identity maps.

The transitive property of the relation is perhaps easiest to see in a diagram.

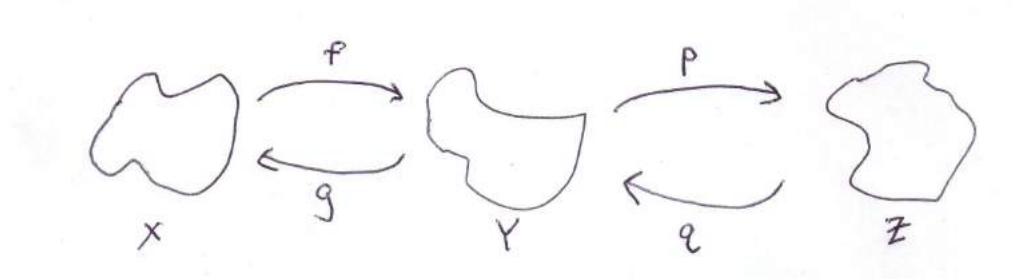


Figure 5.8: Transitivity of homotopy equivalence

We are given that $f \circ g = g \circ f \simeq I$, and that $p \circ q = q \circ p \simeq I$, so clearly the composition $(f \circ p) \circ (q \circ g) = (q \circ g) \circ (f \circ p) \simeq I$.

Therefore we can conclude that given $X \sim Y$ and $Y \sim Z$, $X \sim Z$ by using the properties of function composition.

Note that the transitivity property can be useful when trying to determine if two spaces belong to the same homotopy equivalence class. It may be easier to show that each space is homotopy equivalent to a third, simpler space, and then conclude the two are also equivalent.

5.3.3 On the usefulness of Homotopy Equivalence

One might ask why the idea of homotopy equivalence between spaces is worth studying. One of the main reasons is that it gives a powerful tool to determine the fundamental group of a space. Here is a useful theorem:

Theorem 5.3.4. *If X and Y are path-connected spaces such that $X \sim Y$, then the fundamental groups $\pi_1(X)$ and $\pi_1(Y)$ are isomorphic.*

The utility of this theorem can be seen in the next example.

Example 5.3.5. *The fundamental group of a torus minus a point:*

As usual, it is easiest to consider the torus by its identification polygon, below:

As we will see in the next section, removing a point from the polygon is equivalent to removing the entire interior, when trying to find a homotopy equivalent space. Proceeding by identifying

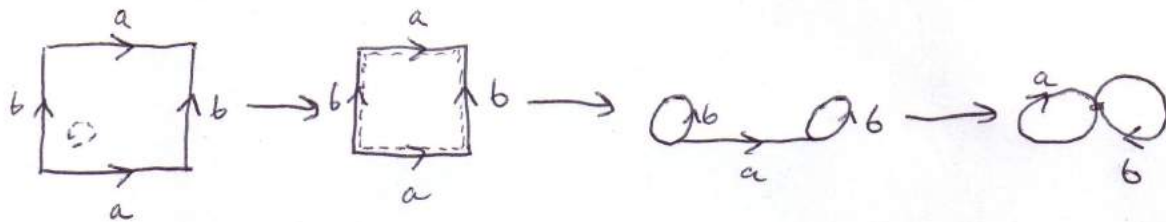


Figure 5.9: A torus missing a point

the boundary, we obtain a bouquet of 2 circles. By using the theorem, the fundamental group of the torus minus a point is therefore isomorphic to F_2 .

The theorem also shows that the fundamental group of a space is a topological invariant. The contrapositive implies that if the fundamental groups of two spaces are not isomorphic, then they are not homotopy equivalent (assuming they are path-connected), and so they are certainly not homeomorphic as seen previously.

5.3.4 Simple-Connectedness and Contractible spaces

Definition 5.3.6. A space X is said to be simply connected if $\pi_1(X)$ is the trivial group.

For example, the sphere is simply connected, since any loop on the sphere can be reduced to a point, so its fundamental group is trivial.

With the previous theorem in mind, the notion of simple-connectedness helps to tell which homotopy equivalence class a space belongs to. If a space is path-connected and homotopy equivalent to a point, then its fundamental group is trivial, and hence it is simply connected. This is the essence of a contractible space.

Definition 5.3.7. A space is said to be contractible if and only if it is homotopy equivalent to a point.

One property of a contractible space is that any two maps into a contractible space are homotopic.

Proof. First, note that 1_X , the identity map on X , is homotopic to the constant map at some point of X . This is because X is homotopy equivalent to a point, so we can pick some point in X and map it to the point, so that the constant map is “the same” as the identity map, up to homotopy.

Now suppose that there are two maps f and g from some space Z into the contractible space X . We know that $f = 1_X \circ f$. By the note above, $1_X \circ f \simeq c \circ f$, where c is the constant map at some point of X . Since the constant map “eats” whatever we compose it with, we have that $c \circ f = c \circ g \simeq 1_X \circ g = g$. \square

5.4 Retractions

Definition 5.4.1. A retraction is a homotopy equivalence between a space X and a subspace Y ($Y \subseteq X$). Note: when dealing with retractions we assume we are using the Euclidian Topology.

Formally, a homotopy is a retraction $r : X \times I \rightarrow X$ if:

- $r(x, 0) = x$
- $r(x, 1) \in Y$
- $r(y, t) = y \in Y$ for all t

What this says is that r is a retraction if at time 0, x remains unchanged in X . At time 1 ($t = 1$), points in X have been retracted into the new object Y . And finally, at any time t , and object which began in Y is still a part of the set Y ($y \in Y$)

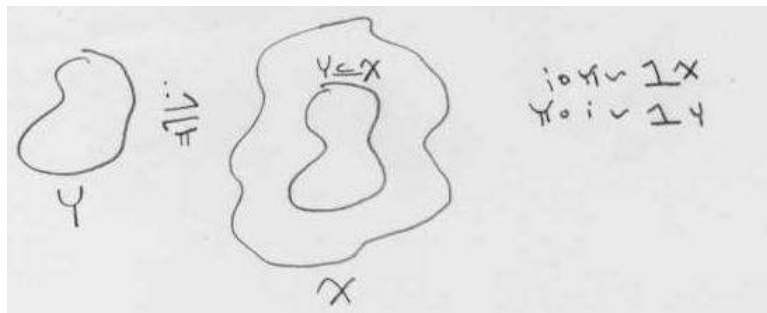


Figure 5.10: General picture of a retraction

It is important to note that retractions do not have inverses. The reasoning behind this is think of an object which retracts to a single point, as in Figure 5.11, now think of some function which expands that point into an object, can this function even exist? [hint: no]

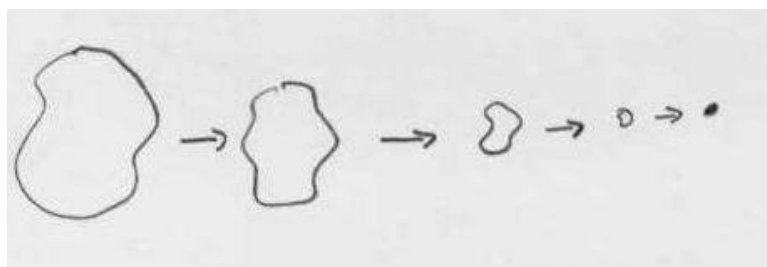


Figure 5.11: Noninvertibility of a retraction

5.4.1 Examples of Retractions

Starting with a simple 4 sided polygon (1) as shown in Figure 5.12, we can show what a retraction does to an object. Begin with the torus. If we were to remove a single point (2), we could then create a retraction on (2) to make the hole we poked in (1) larger. As t goes from 0

to 1, our polygon goes from (2) to (5). At time 1, the hole has increased to the entire polygon except for the edges. All that remains of our torus in (5) is just 2 circles joined at a point.

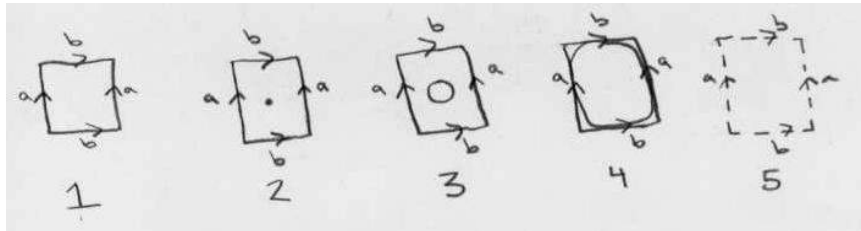


Figure 5.12: Punctured torus retraction

The question that may arise is why is it necessary for there to be a point removed from the polygon. This is because if we create a retraction that maps each point within the polygon to its nearest edge, there will always be a point that does not “know where to go”. Since our homotopy needs to be defined for all points within the set, we solve this problem by simply removing the problematic point from our polygon. Now that all points “know where to go”, we can create a function that takes them from their location at time 0, to their destination at time 1.

The next question that may arise is why can't everything be retracted into a single point. If an object is path connected, why can't we find a retraction that takes all points from their location at time 0 to a single point at time 1. The reason for this is that retractions do preserve homotopy invariants. For example, if our object is a circle, we can ‘shrink’ the circle into a smaller circle, but if we are going to make the circle into a single point we would need to “plug up” the hole it surrounds. To do this would create an entirely different object not homotopic to the circle. Figure 5.13 Shows that any retraction on a circle must preserve the hole within, no matter how small it gets.

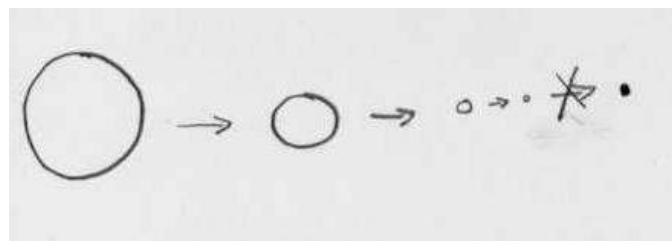


Figure 5.13: Inability to shrink a circle to a point

In addition to showing why we cannot simply plug up a hole in an object, let us show why we cannot simply ‘cut’ the object to remove the hole. Lets assume the simple case of a circle in \mathbb{R}^2 . Lets assume our retraction is as illustrated in Figure 5.14. The problem arises at the point where we ‘cut’ our circle. Does this point go to the left, to get to the bottom, or does it go to the right? For the exact same reason we must remove a point from a polygon, this point too causes us to not be able to complete a retraction as we had desired. If however, we remove that point, we are left with a line segment (not a circle) and while that is able to be retracted to a point, the circle is not.

This case of the 2d circle can be applied to any surface with a hole in it, such as a sphere (but



Figure 5.14: Attempt at retraction of a circle to a point

not a ball! A ball contains no holes, and is fully retractable to a point!)

Therefore, we now know that certain topological properties such as the number of holes an object has must be preserved when doing a retraction. Our example with the torus can be expanded to any representative polygon. Figure 5.15 shows an example of how any arbitrary n -polygon with a point removed can be retracted into n circles joined at a point.

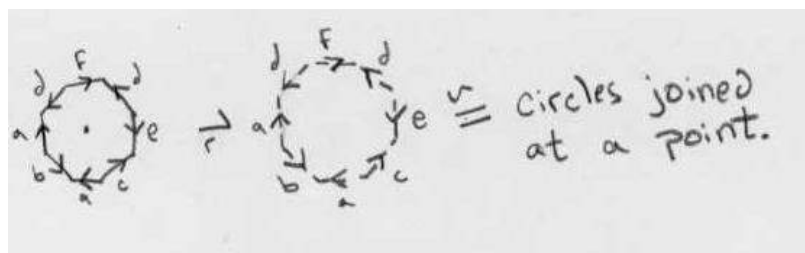


Figure 5.15: Homotopy of an n -polygon with a point removed

5.5 Computing the Fundamental Groups of Surfaces: The Seifert-Van Kampen Theorem

Introduction

The Seifert-Van Kampen Theorem is a very useful result in topology. It allows for the computation of fundamental groups of very complex topological spaces. The theorem allows one to compute the fundamental group of a topological space by breaking the space into smaller overlapping subspaces and computing the fundamental groups of these spaces. Then the fundamental group of the original space can be easily obtained. The following is the precise statement of the Seifert-Van Kampen Theorem:

Theorem 5.5.1 (Seifert-Van Kampen). *Given a space X , let X_1 and X_2 be subspaces of X such that $X_1 \cup X_2 = X$ and is open, $X_1 \cap X_2 \neq \{\phi\}$ and is open, and $X_1 \cap X_2$ is path-connected. Then $\pi_1(X)$ is just the group with generators and relations from $\pi_1(X_1)$ and $\pi_1(X_2)$ with additional relations from $\pi_1(X_1 \cap X_2)$.*

Idea Behind Proof: Let X be the following topological space:

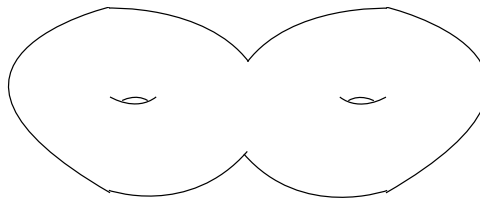


Figure 5.16: A 2-Torus

Then divide this space into two overlapping subspaces X_1 and X_2 (Figure 5.17).

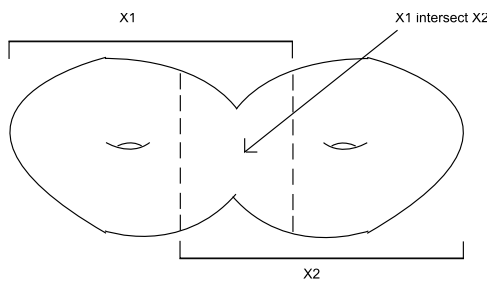


Figure 5.17: Dividing the Space

Claim 5.5.2. Any loop in X can be described by a composition of loops from X_1 and X_2 .

Idea behind proof of claim: First choose the base point x_0 somewhere in $X_1 \cap X_2$, and let l be a loop in X . If l is entirely in X_1 or entirely in X_2 , then they clearly can be described by a single loop in X_1 and X_2 respectively. Hence, let us consider loops that are not entirely in X_1 or X_2 (Figure 5.18a). Since $X_1 \cap X_2$ is path-connected, we can select a section of the loop in $X_1 \cap X_2$ that doesn't cross x_0 and homotopically map it to x_0 along paths p_1 and p_2 (Figure 5.18b). Note that this doesn't change the homotopy equivalence class of the loop. Hence, in this case we get two loops. A loop α in X_1 and a loop β in X_2 , and we see that $l = \alpha\beta$ (Figure 5.18c).

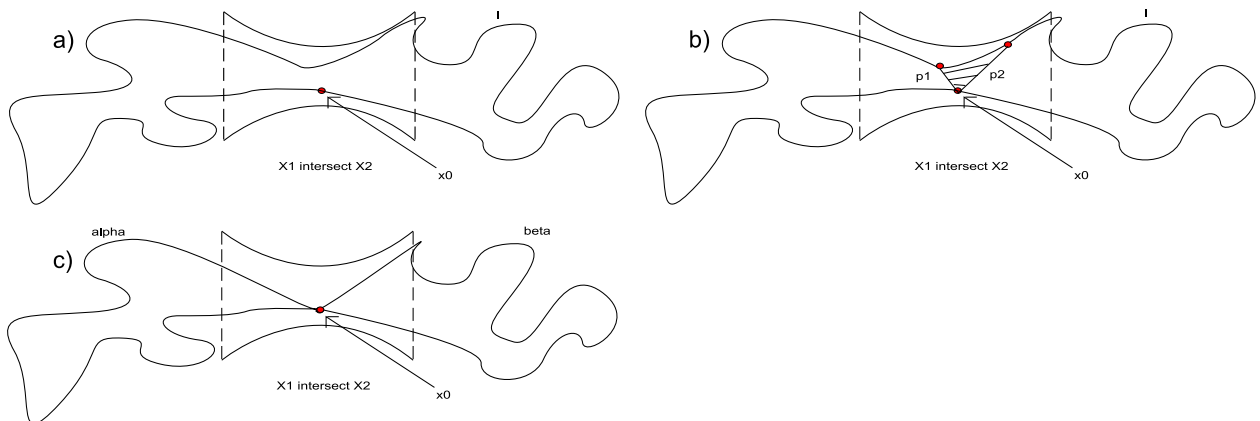


Figure 5.18: Making a Detour

It is fairly easy to see that this can be generalized to any loop in X (End of proof of claim).

Hence for any loop l in X , we can write it in the form $l = \alpha_1\alpha_2 \dots \alpha_n$, where each α_i are loops in either X_1 or X_2 . Hence the generators of $\pi_1(X)$ are just the generators from $\pi_1(X_1)$ and the generators from $\pi_1(X_2)$.

As for the relations of $\pi_1(X)$, we see that these relations can be divided into three cases:

1. The relations we get when the elements of $\pi_1(X_1)$ are equal to other elements of $\pi_1(X_1)$, which are just the relations of $\pi_1(X_1)$. We can think of these relations as just ‘words’ that are described by generators of $\pi_1(X_1)$.
2. The relations we get when the elements of $\pi_1(X_2)$ are equal to other elements of $\pi_1(X_2)$, which are just the relations of $\pi_1(X_2)$. We can think of these relations as ‘words’ that are described by generators of $\pi_1(X_2)$.
3. The relations we get when the elements of $\pi_1(X_1)$ are equal to other elements of $\pi_1(X_2)$.

All-in-all, we see that the generators and relations of $\pi_1(X)$ are just the generators and relations from $\pi_1(X_1)$ and $\pi_1(X_2)$, with additional relations from $\pi_1(X_1 \cap X_2)$.

5.5.1 Examples:

The general method for computing the fundamental groups of a space (the application of the Seifert-Van Kampen Theorem) is to first divide up your space such so that the requirements of the theorem are met. Then, the generators and relations of the smaller and simpler subspaces can be found pretty easily.

As a note on the technique for applying the Seifert-Van Kampen Theorem, it is often visually easier to represent the topological space by its identification polygon and divide up the spaces in that representation.

Example 5.5.3 (A Torus). *First we divide up the spaces as in Figure 5.19. We see that the hole in X_1 is homotopic to just the outside boundary of the identification polygon by means of deformation retract. Since all corners of the identification polygon are the same point, the boundary can be represented as a bouquet of 4 distinct loops, (loop a , \bar{a} , b , and \bar{b}). But a and \bar{a} are generated by the same generator, and b and \bar{b} are generated by the same generator. Hence a and b are our generators from X_1 and no relations. Now lets look at X_2 . X_2 is just a disk so it is homotopic to the identity. Hence, there are also no generators or relations in X_2 . To see the relations of $X_1 \cap X_2$, we consider a loop l in $X_1 \cap X_2$. The same loop l in X_1 is homotopic to the outside boundary i.e. $l = ab\bar{a}\bar{b}$, and the same loop in X_2 is homotopic to the identity. Hence $ab\bar{a}\bar{b} = 1$ and so $ab\bar{a}\bar{b}$ is our relation. Hence $\pi_1(X) = \langle a, b | ab\bar{a}\bar{b} \rangle$.*

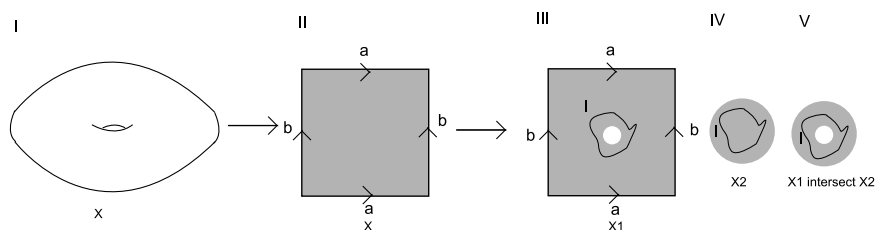


Figure 5.19: Torus

Example 5.5.4 (A Klein Bottle). *Similar to the previous example, we divide up the spaces as in Figure 5.20. We see that the hole in X_1 is homotopic to just the outside boundary of the identification space by means of deformation retract. Since all corners of the identification polygon are the same point, the boundary can be represented as a bouquet of 2 loops. The distinct loops are a and b . Hence, a and b are again our generators from X_1 and no relations. Now lets look at X_2 . X_2 is just a disk so it is homotopic to the identity. Hence, there are also no generators or relations in X_2 . To see the relations of $X_1 \cap X_2$, we consider a loop l in $X_1 \cap X_2$. The same loop l in X_1 is homotopic to the outside boundary i.e. $l = ab\bar{a}\bar{b}$, and the same loop in X_2 is homotopic to the identity. Hence $l = ab\bar{a}\bar{b} = 1$ and so $ab\bar{a}\bar{b}$ is our relation. Hence $\pi_1(X) = \langle a, b | ab\bar{a}\bar{b} \rangle$.*

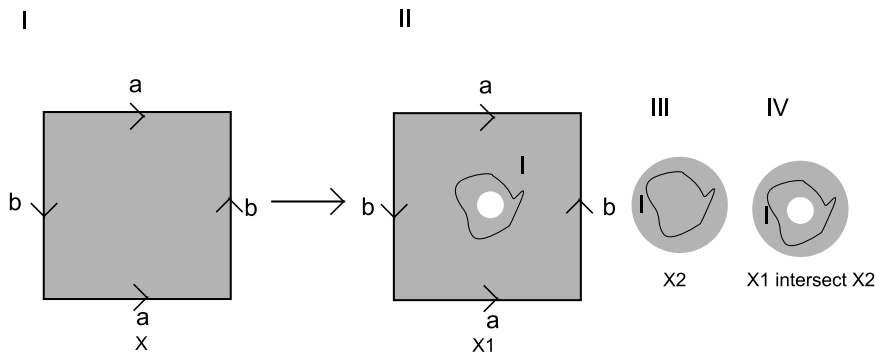


Figure 5.20: Klein Bottle

Example 5.5.5 (An n -Torus Connected to the Real Projective Plane). *Again, first we divide up the spaces as in Figure 5.21 (the identification polygon of the n -torus has $4n$ sides). We see that the hole in X_1 is homotopic to just the outside boundary of the identification polygon by means of deformation retract. Since all corners of the identification polygon are the same point, the boundary can be represented as a bouquet of $4n$ loops with half of them being inverses of the other half. Hence like in the previous examples, we only need $2n$ number of generators (a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n) and no relations. Now lets look at X_2 . X_2 is just the boundary of the projective plane. The corners are the same point so we get 2 loops, which is just the 1 distinct loop a , which means there is just the generator a and no relations. To see the relations of $X_1 \cap X_2$, we consider a loop l in $X_1 \cap X_2$. The same loop l in X_1 is homotopic to the outside boundary of X_1 i.e. $l = a_1 b_1 \bar{a}_1 \bar{b}_1 a_2 b_2 \bar{a}_2 \bar{b}_2 \dots a_n b_n \bar{a}_n \bar{b}_n$. The same loop l in X_2 is homotopic to the boundary of X_2 i.e. $l = a^2$. Hence, we get $a_1 b_1 \bar{a}_1 \bar{b}_1 a_2 b_2 \bar{a}_2 \bar{b}_2 \dots a_n b_n \bar{a}_n \bar{b}_n = a^2$. Hence $\pi_1(X) = \langle a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n | a_1 b_1 \bar{a}_1 \bar{b}_1 a_2 b_2 \bar{a}_2 \bar{b}_2 \dots a_n b_n \bar{a}_n \bar{b}_n a^2 \rangle$.*

5.6 Covering Spaces

Definition 4. A *covering map* is a surjective map

$$\pi : X \rightarrow Y$$

where for every $y \in Y$, there is a neighborhood U_y such that $\pi^{-1}(U_y)$ consists of a disjoint union of subsets $V_i \subset X$ and the restriction maps $\pi|_{V_i} : V_i \rightarrow U_y$ are all homeomorphisms.

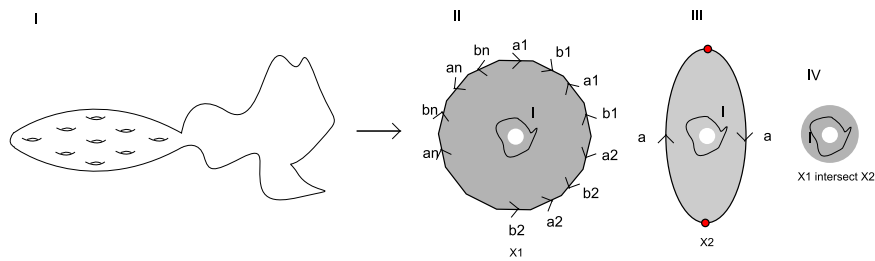


Figure 5.21: N-Torus connected to the real projective plane

Definition 5. X is a *covering space* for Y if there exists a covering map from X to Y .

This is illustrated in Figure 5.22, where each of the three open sets in X is homeomorphic to the open set in Y , and each of the points shown maps to y under π .

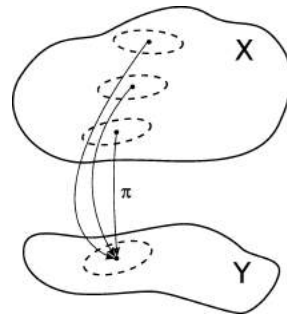


Figure 5.22: Covering map/space illustrative definition

First, we look at some simple examples to get an idea of what covering spaces and maps are.

Example 5.6.1. A simple example of a covering map from the real line to the circle is

$$\begin{aligned} \pi : \mathbb{R} &\longrightarrow S^1 \\ x &\longmapsto e^{ix} \end{aligned}$$

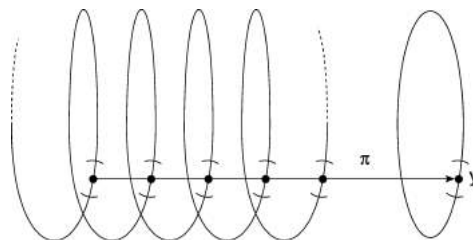


Figure 5.23: Covering map from the helix (\mathbb{R}) to the circle

This can be pictured as a projection from the infinite helix to the circle, as shown in Figure 5.23. This is a covering map, as we can see as follows. Let $y \in S^1$, and let U_y be the neighborhood around y consisting of points within $\pm\pi/2$ radians, i.e., the set $U_y = \{e^{iu} \in Y \mid \arg y - \pi/2 < u < \arg y + \pi/2\}$.

Then the preimage $\pi^{-1}(U_y)$ consists of infinitely many disjoint subsets in \mathbb{R}^2 , where each one is an interval of length π centered around the points in \mathbb{R} that map to y , namely, the points $\{\arg y + 2n\pi \mid n \in \mathbb{Z}\}$. Each of these sets is obviously homeomorphic to U_y under π .

Example 5.6.2. *Another example is*

$$\begin{aligned} \pi : \mathbb{R}^2 &\longrightarrow T \\ (x, y) &\longmapsto (x - [x], y - [y]) \end{aligned}$$

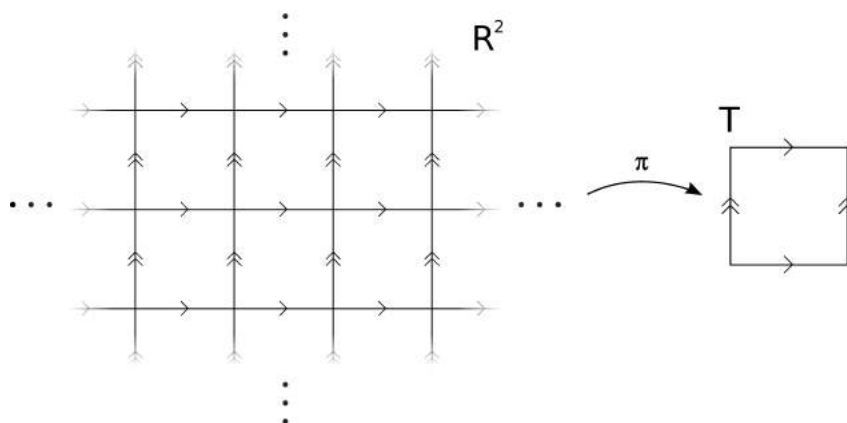


Figure 5.24: Covering map from the plane to the torus

This takes any point in the plane and maps it to the point with the same fractional part. This works as a map to the torus since the torus can be thought of as an identification space on the unit square. The operation of this map is depicted in Figure 5.24, where the arrows in \mathbb{R}^2 denote where the point is sent to under π , not that the picture is actually an identification space by itself. However, the arrows in the torus on the right are the normal identification space arrows.

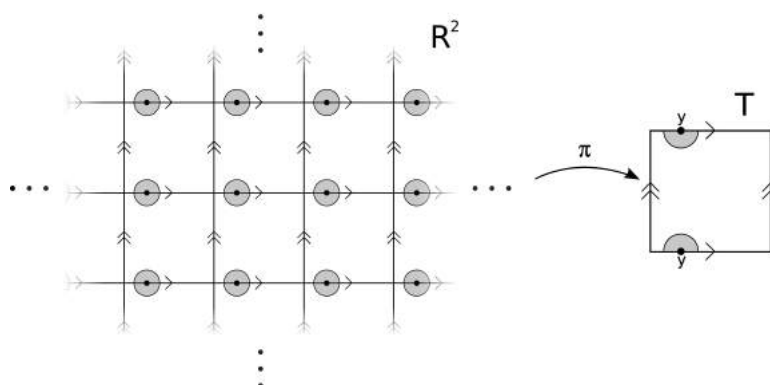


Figure 5.25: Preimage of an open disc around a point on the border of T

This is also a covering map. For $y \in T$ not on the “border”, there exists an open neighborhood of y fully contained in the torus, not crossing the border. Such an open neighborhood has a preimage which is just the neighborhood copied and shifted to by every integer coordinate pair, each of which is homeomorphic to the original neighborhood in the torus. However, it is slightly

more complicated if y is on one of the borders. Since we only need to show the properties for some open set, we can choose to look at open discs. If y is on the border, but not the corner, then there is an open disc around y which looks like two half-discs, since y is identified with the point on the opposite side of the torus. Then the preimage the open disc is a disjoint set of open discs in \mathbb{R}^2 , as shown in Figure 5.25. It is important to note that π restricted to one of the discs is a homeomorphism with the open disc in T , since although it appears to be split along the middle, it is actually still together in the identification space of the torus. Similarly, if y is the corner point, then its preimage is the set of all open discs (of the same radius) centered at all points with integer coordinates in the plane. Each disc is again homeomorphic to the original disc in T under π restricted to each disc.

Example 5.6.3 (A not-example). *There is also an interesting map to consider that appears to be a covering map, but actually isn't. Consider the map from the plane to the projective plane,*

$$\pi : \mathbb{R}^2 \longrightarrow \mathbb{R}P^2$$

which is defined as shown in Figure 5.26, which also shows the preimage of a disc.

Note that adjacent projective planes¹ are flipped so that the correct sides match. This is needed so that π is continuous.

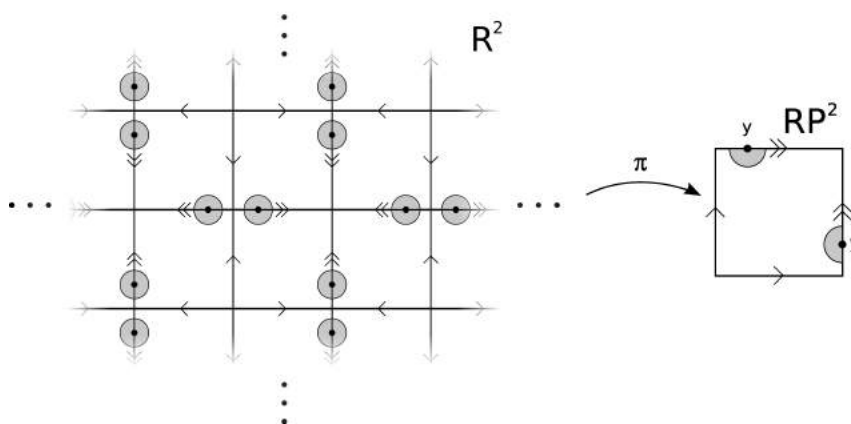


Figure 5.26: Map from the plane to the projective plane, with the preimage of a disc.

Consider the point $y \in \mathbb{R}P^2$ shown in the figure. A neighborhood of y with its corresponding preimage is shown. Note that the preimage is a disjoint set of discs which are homeomorphic to the original disc. However, the discs in the plane are not homeomorphic to the disc in the projective plane under the restriction of π to each disc. The image of each disc under the restricted map is only half of the disc, as shown in Figure 5.27, and half of an open disc is not homeomorphic to a whole open disc, since part of its boundary is closed.

Thus, this is not actually a covering map.

¹Caution: here our artist got a little confused in the drawing, and the picture is not completely correct. Up to you to figure out how to fix it!

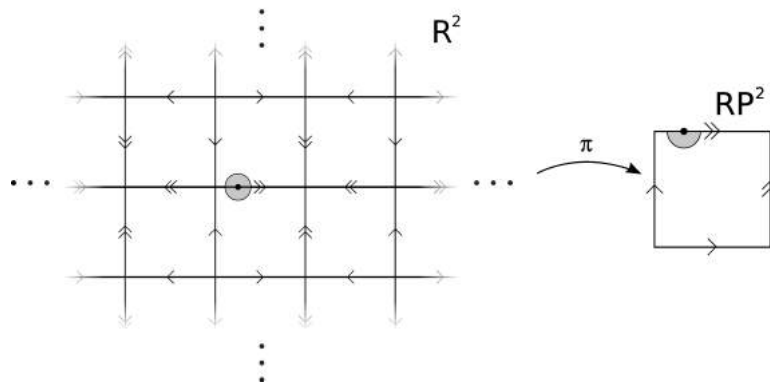


Figure 5.27: Image of one disc under π

5.6.1 Lifting

The general terms for X and Y in the above discussion are the *cover* and the *base*, respectively. So the covering map takes points from the cover to the base.

Lifting refers to taking a path γ in the base starting at point p , and getting a path $\tilde{\gamma}$ in the cover starting at a point q , where $\pi(\tilde{\gamma}) = \gamma$ and $\pi(q) = p$.

There are two basic results about lifts which are proved in pages 97–98 of Armstrong: the path-lifting lemma and the homotopy-lifting lemma.

Path-lifting lemma. Given a loop γ in Y starting at p and a point $q \in X$, where $\pi(q) = p$, there is a unique lift $\tilde{\gamma}$ in X starting at q .

The path-lifting lemma is illustrated in Figure 5.28. The main idea is that for each of the points in the cover that map to the starting point of the path in the base, there is exactly one path in the cover which is the lift of the path in the base. Note that it is unique only up to the choice of the point in the cover.

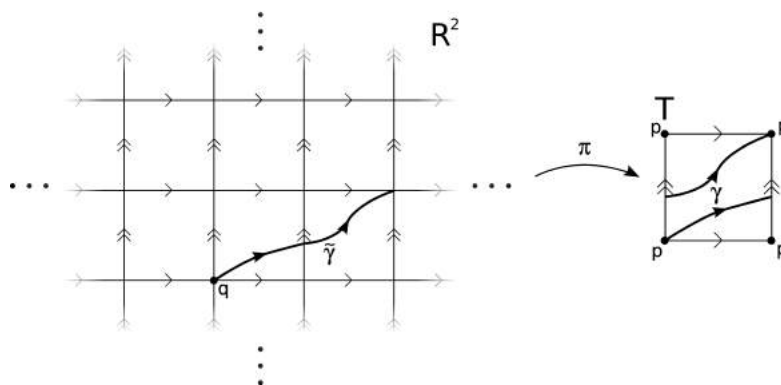


Figure 5.28: A lift of a loop γ in the torus to the cover from Figure 5.24

It is also important to note that a loop in the base does not necessarily lift to a loop in the cover, since the ending point of the loop in the cover could be one of the other points in the preimage of p .

Homotopy-lifting lemma. If $F : I \times I \rightarrow Y$ is a map such that $F(0, t) = F(1, t) = p$ for $0 \leq t \leq 1$, there is a unique map $\tilde{F} : I \times I \rightarrow X$ which satisfies $\pi \circ \tilde{F} = F$ and $\tilde{F}(0, t) = q$ for $0 \leq t \leq 1$.

The homotopy-lifting lemma is an important tool for algebraic topology. One result is that the fundamental group of the cover is a subgroup of the fundamental group of the base, namely, the subgroup of loops in the base that when lifted are still loops. There is actually a bijection between the covering spaces of X and the normal subgroups of $\pi_1(X)$. An in-depth treatment of algebraic topology is not within the scope of this document, but it is important to realize that these results are just the beginning for an entire branch of mathematics.