3-2018

Introduction to Computational Topology Using Simplicial Persistent Homology

Jason Turner
Union College - Schenectady, NY

Brenda Johnson
Union College - Schenectady, NY

Ellen Gasparovic
Union College - Schenectady, NY

Introduction to Computational Topology
Using Simplicial Persistent Homology

By

Jason Turner

* * * * * * * * *

Submitted in partial fulfillment
of the requirements for
Honors in the Department of Mathematics

UNION COLLEGE
June 2018

Abstract

TURNER, JASON Introduction to Computational Topology Using Simplicial Persistent Homology, and its usage in Analyzing BuckyBall® Arrangements. Department of Mathematics, June 2018.

ADVISORS: Johnson, Brenda and Gasparovic, Ellen

The human mind has a natural talent for finding patterns and shapes in nature where there are none, such as constellations among the stars. Persistent homology serves as a mathematical tool for accomplishing the same task in a more formal setting, taking in a cloud of individual points and assembling them into a coherent continuous image. We present an introduction to computational topology as well as persistent homology, and use them to analyze configurations of BuckyBalls®, small magnetic balls commonly used as desk toys.

Preface

For hundreds and thousands of years, humans have found patterns in everything around them: gods and heavenly figures in the stars, faces in the sand dunes on Mars, shapes in the clouds, and much more.

In more recent times, there have been attempts using machine learning and other algorithms [10] to replicate this innate human ability. Persistent homology, based in algebraic topology, is another such attempt. The tools in persistent homology have been used in image recognition [3] as well as the analysis of a variety of other systems, such as the human gait [6].

Our motivation for this text is to build a repertoire of ideas and examples from topology to introductory ideas in persistent homology. After establishing this background, we utilize the analytic tools provided by persistent homology to analyze arrangements of BuckyBalls®, which are small magnetic spheres commonly used as a desk toy. We use persistent homology to not only obtain information about the geometry of arrangements of BuckyBalls®, but also their energy profile.

In this thesis, we present the following:

(i) A brief introduction to topology (Sections 1.1 - 1.4),

(ii) Geometric and abstract simplicial complexes (Section 2.1),

(iii) Čech and Vietoris-Rips complexes (Section 2.2),

(iv) Homology groups of a simplicial complex (Section 3.1), including calculating the homology groups of surfaces using labeled diagrams (Section 3.2),

(v) Exact sequences of homology groups (Section 4.1), including calculating the homology of surfaces using them (Section 4.2),

(vi) Persistent homology groups and persistence diagrams, including their stability (Sections 5.1 and 5.2), and

(vii) An analysis of the arrangement of simulated BuckyBall® systems using persistent homology (Sections 6.2 and 6.3).

Ideally, the reader has already taken an undergraduate course in algebra and analysis, although all necessary material from these areas is provided in Section 1.1.

Throughout the text, we strive to provide meaningful examples of most, if not all, ideas and concepts presented. We hope the reader is able to, with minimal external effort, understand and appreciate the material covered.

Should the reader be interested in further studies, we recommend [1] as an introductory text for topology, [9] as an in-depth reference for algebraic topology, and [4] as introductory reading for computational topology and topological data analysis. Our discussions throughout the first five chapters of this text are based on these three texts, and many proofs presented here are similar to those found in them.

Contents

Chapter 1: Topology

Topology (along with geometry) may be considered to be the “third Musketeer” of pure mathematics, alongside algebra and analysis. In recent years, it has gained popularity in applied mathematics as well, finding use in data analysis among other fields.

Section 1.1 includes most, if not all, background information from set theory, algebra, and analysis required to understand the rest of this text. It also introduces the reader to our notation, which generally agrees with the standard rigorous notation.

Sections 1.2 - 1.4 serve as an introduction to topology, including ideas such as open and closed sets, the Heine-Borel Theorem, the subspace, product, and quotient topologies, as well as continuous functions and homeomorphisms.

Section 1.1: Preliminaries

In this section, we review a selection of ideas from set theory, algebra, and analysis that will be of great use in our discussions throughout the text. For the sake of brevity, many proofs are left to the reader.

Ideas from Set Theory

Lemma 1.1. Let A, B, and C be sets. The following relations hold:

(i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

(ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

(iii) $A \times (B \cup C) = (A \times B) \cup (A \times C)$,

(iv) $A \times (B \cup C) = (A \times B) \cup (A \times C)$,

(v) $A \times (B \cap C) = (A \times B) \cap (A \times C)$,

(vi) $A \times (B - C) = (A \times B) - (A \times C)$,

(vii) $A - (B \cup C) = (A - B) \cup (A - C)$, and

(viii) $A - (B \cap C) = (A - B) \cup (A - C)$.

Lemma 1.2. Let $f : X \to Y$ be a function between sets X and Y. Let $A,\ B \subseteq X$ and $V,\ W \subseteq Y$. The following relations hold:

(i) $f(A \cup B) = f(A) \cup f(B)$,

(ii) $f(A \cap B) \subset f(A) \cap f(B)$,

(iii) $f(A) - f(B) \subset f(A - B)$,

(iv) $f^{-1}(V \cup W) = f^{-1}(V) \cup f^{-1}(W)$,

(v) $f^{-1}(V \cap W) = f^{-1}(V) \cap f^{-1}(W)$, and

(vi) $f^{-1}(V - W) = f^{-1}(V) - f^{-1}(W)$.

The Union Lemma gives the criterion for the union of a collection C of subsets A_i of a set X to equal X itself, i.e., $\bigcup\limits_{A_i \in C} A_i = X$. It will prove itself to be extremely useful throughout our introduction to topology (Sections 1.2 - 1.4), where we rely on it for a number of proofs.

Lemma 1.3. THE UNION LEMMA. Let X be a set and C be a collection of subsets of X. For each element x of X, let A_x be a set in C which contains x. Then $\bigcup\limits_{x \in X} A_x$ is X itself.

Proof. We prove the Union Lemma via double containment.

The union $\bigcup\limits_{x \in X} A_x$ is clearly a subset of X, as each A_x is a subset of X.

Let y be an arbitrary element of X. By hypothesis, there exists an $A_y \in C$ such that $y \in A_y$. As A_y is contained in $\bigcup\limits_{x \in X} A_x$, X is a subset of $\bigcup\limits_{x \in X} A_x$. ■

Example 1.1. Consider the collection of sets given by $C = \{\{n-1,\ n,\ n+1\} \mid n \in \mathbb{Z}\}$. By the Union Lemma, $\mathbb{Z}$ is the union of all such three-element sets.

Ideas from Algebra

Of particular relevance to our studies are abelian groups, which we shall write additively, i.e., we write the group operation of any abelian group to be $+$.

As such, 0 denotes the identity element, $-g$ denotes the additive inverse of g, and, for a positive integer n, ng denotes the n-fold sum $g + g + \cdots + g$.

Definition 1.4. An abelian group G is **free** if there exists a subset of elements $\left\{g_\alpha\right\}$ of G, called a **basis**, such that each element g of G may be expressed uniquely as a finite sum

$$g = \sum n_\alpha \, g_\alpha$$

where n_α is an integer. The number of elements in such a basis is called the **rank** of G.

If each element g in G may be written as such a finite sum, but not necessarily uniquely, then we say that $\left\{g_\alpha\right\}$ **generates** G.

Example 1.2. Consider the group $(\mathbb{Z}, +)$. All of $\mathbb{Z}$ is generated by 1, and thus $\mathbb{Z}$ is a free abelian group of rank 1.

It is also possible to construct a free abelian group using a set S in the following manner: The **free abelian group** G **generated by** S is the set of all functions $\phi : S \to \mathbb{Z}$ such that $\phi(x) \neq 0$ for only finitely many values of x, and we add two such functions by adding their values. Given $x \in S$, there exists a characteristic function ϕ_x for x defined by setting

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

The functions $\left\{\phi_x \mid x \in S\right\}$ form a basis for G, for each function $\phi \in G$ may be written uniquely as the finite sum

$$\phi = \sum n_x \phi_x,$$

where $n_x = \phi(x)$ and the summation extends over all x for which $\phi(x) \neq 0$. This differs from previously defined free abelian group in that it erases all "extra" properties of the elements in S.

Example 1.3. Consider the free abelian group $\left\{ a + b\sqrt{2} \mid a,\ b \in \mathbb{Z} \right\}$. We may also consider this group as a ring by imposing the natural multiplicative structure.

However, the free abelian group generated by the elements 1 and $\sqrt{2}$ is the group $\{(a,\ b) \mid a,\ b \in \mathbb{Z}\}$ with addition as its operation. The aforementioned possible multiplicative structure has been eliminated.

In the same vein that linear functions defined on the basis of a vector space may be extended to all elements in that vector space, homomorphisms defined on the basis of a free abelian group may be extended uniquely to homomorphisms of the entire group.

Suppose G is an abelian group and that $\left\{G_\alpha\right\}$ is a collection of subgroups of G. If each $g \in G$ may be written uniquely as a finite sum $g = \sum g_\alpha$ where $g_\alpha \in G_\alpha$ for each α, then G is said to be the **internal direct sum** of the groups G_α, and we write

$$G = \oplus_\alpha G_\alpha.$$

This should appear similar to free abelian groups, for if G is free abelian with $\left\{g_\alpha\right\}$ as its basis, then G is the direct product of the cyclic subgroups generated by each G_α.

The **external direct sum** of a family of abelian groups $\left\{G_\alpha\right\}$ is a group G consisting of all tuples (g_α) such that $g_\alpha = 0_{G_\alpha}$ for all but finitely many values of α.

We use the following lemma in our argument that internal and external direct sums are actually one in the same. However, the lemma may be ignored and the statement may be simply accepted by the reader without any loss of understanding.

Lemma 1.5. Let G be an abelian group. If G is the internal direct sum of the subgroups $\{G_\alpha\}$, then there are homomophisms

$$j_\beta : G_\beta \to G \qquad \text{and} \qquad \pi_\beta : G \to G_\beta$$

such that $\pi_\beta \circ j_\alpha$ is the zero homomorphism if $\alpha \neq \beta$ and the identity homomorphism if $\alpha = \beta$.

Conversely, suppose $\{G_\alpha\}$ is a family of abelian groups, ad there are homomorphisms j_β and π_β as above. Then j_β is a monomorphism. Furthermore, if the groups $j_\alpha(G_\alpha)$ generate G, then G is their internal direct sum.

Proof. Suppose $G = \oplus G_\alpha$. We define j_β to be the inclusion homomorphism. To define π_β, write $g = \sum g_\alpha$, where $g_\alpha \in G_\alpha$ for each α; and let $\pi_\beta(g) = g_\beta$. Uniqueness of the representation of g shows π_β is a well-defined homomorphism.

Consider the converse. As $\pi_\alpha \circ j_\alpha$ is the identity, j_α is injective (and π_α is surjective). If the groups $j_\alpha(G_\alpha)$ generate G, every element of G can be written as a finite sum $\sum j_\alpha(g_\alpha)$, by hypothesis. To show this representation is unique, suppose

$$\sum j_\alpha(g_\alpha) = \sum j_\alpha(g'_\alpha).$$

Applying π_β, we see that $g_\beta = g'_\beta$. ■

We now argue that internal and external direct sums are identical; suppose G is the external direct sum of the groups $\{G_\alpha\}$. Then for each β, we define $\pi_\beta : G \to G_\beta$ to be the projection onto the β^{th} factor. And we define $j_\beta : G_\beta \to G$ by letting it carry the element $g \in G_\beta$ to the tuple (g_α), where $g_\alpha = 0_{G_\alpha}$ for all α different from β, and $g_\beta = g$. Then $\pi_\beta \circ j_\alpha = 0$ for $\alpha \neq \beta$, and $\pi_\alpha \circ j_\alpha$ is the identity. It follows that G is the internal direct sum of the groups $G'_\alpha = j_\alpha(G_\alpha)$, where G'_α is isomorphic to G_α.

We typically denote the internal and external direct sums using

$$G = G_1 \oplus \cdots \oplus G_n \quad \text{and} \quad G = \oplus G_\alpha$$

relying on the context to clarify what is meant (if it is important). For example, we may express that G is of rank four by writing $G \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus Z$.

A subgroup G_1 of G is a **direct summand** in G if there exists a subgroup G_2 of G such that $G = G_1 \oplus G_2$. If H_1 is a subgroup of G_1 and H_2 is a subgroup of G_2, then

$$\frac{G}{H_1 \oplus H_2} \cong \frac{G_1}{H_1} \oplus \frac{G_2}{H_2} \quad \text{and} \quad \frac{G}{G_1} \cong G_2.$$

It is well known[1] and extremely useful that all finitely generated abelian groups G may be expressed as the direct sum

$$G \cong (\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}) \oplus \left(\mathbb{Z}_{a_1} \oplus \cdots \oplus \mathbb{Z}_{a_s}\right)$$

where each a_i is a power of a prime, and $\mathbb{Z}_{a_i}$ denotes $\mathbb{Z}$ modulo a_i.

Ideas from Analysis

In this text, we are primarily concerned with **Euclidean N-space**, denoted $\mathbb{R}^N$, and subsets thereof. The standard means for measuring distance in this space is via the **Euclidean distance formula** which is defined as follows: For points $p = (p_1, \ldots, p_N)$ and $q = (q_1, \ldots, q_N)$, the distance between p and q is

$$d(p, q) = \sqrt{\left(p_1 - q_1\right)^2 + \cdots + \left(p_N - q_N\right)^2}.$$

We commonly denote the distance between the point p and the origin $\mathcal{O}$ by $\|p\|$, regardless of the dimension of the Euclidean space it is in. Note that the Euclidean distance satisfies the following conditions:

(i) For all p, q in $\mathbb{R}^N$, $d(p, q) \geq 0$, and $d(p, q) = 0$ if and only if $p = q$.

(ii) For all p, q in $\mathbb{R}^N$, $d(p, q) = d(q, p)$.

[1] See *the fundamental theorem of finitely generated abelian groups*.

(iii) For all p, q, r in $\mathbb{R}^N$, $d(p, r) \leq d(p, q) + d(q, r)$ (known as the *Triangle Inequality*).

A set $A \subset \mathbb{R}^N$ is **bounded** if there exists a real number B such that $\|x\| \leq B$ for all x in A.

Example 1.4. The unit disc centered at the point $(3, -7)$ in $\mathbb{R}^2$ is bounded. For this, $B \geq 8$.

A set $A \subset \mathbb{R}^N$ is **convex** if for every pair of points p, q in A, the line segment PQ adjoining them lies entirely in A.

Example 1.5. Every regular polygon in $\mathbb{R}^2$ is convex. The five-point star, on the other hand, is not convex.

Definition 1.6. The N**-sphere**, denoted $S^N(p, r)$, is the set of all points a distance r away from the point $p = (p_1, \dots, p_{N+1})$ in $\mathbb{R}^{N+1}$, i.e.,

$$S^N(p, r) = \left\{x \in \mathbb{R}^{N+1} \mid d(p, x) = r\right\}.$$

Definition 1.7. The **closed** N**-ball**, denoted $\overline{\mathrm{B}^N}(p, r)$ is the set of all points of distance r or less away from $p = (p_1, \dots, p_N)$ in $\mathbb{R}^N$, i.e.,

$$\overline{\mathrm{B}^N}(p, r) = \left\{x \in \mathbb{R}^N \mid d(x, p) \leq r\right\}.$$

Definition 1.8. Similarly, the **open** N**-ball**, denoted $B^N(p, r)$, is the set of all points of distance less than r away from $p = (p_1, \dots, p_N)$ in $\mathbb{R}^N$, i.e.,

$$B^N(p, r) = \left\{x \in \mathbb{R}^N \mid d(x, p) < r\right\}.$$

Example 1.6. The 0-sphere of radius r centered about the point p in $\mathbb{R}^0$, denoted $S^0(p,\ r)$, is made up of the points $p-r$ and $p+r$. The closed 1-ball $\overline{\mathrm{B}^1}(p,\ r)$ is the closed interval $[p-r,\ p+r]$, while the open 1-ball $B^1(p,\ r)$ is the open interval $(p-r,\ p+r)$. Notice that $S^0(p,\ r)$ is a sort of "boundary" of $B^1(p,\ r)$ and $\overline{\mathrm{B}^1}(p,\ r)$.

The 2-sphere $S^2(p,\ r)$ is the sphere centered about the point p of radius r in $\mathbb{R}^3$. The closed 3-ball $\overline{\mathrm{B}^3}(p,\ r)$ is the ball centered at the point p of radius r in $\mathbb{R}^3$, while the open 3-ball $B^3(p,\ r)$ is all of the points in $\overline{\mathrm{B}^3}(p,\ r)$ excluding those in $S^2(p,\ r)$.

Definition 1.9. A **sequence** in a non-empty set X is a mapping $s:\ \mathbb{Z}^+ \to X$, usually expressed as a list

$$\langle s(1),\ s(2),\ \dots\rangle = \langle s_1,\ s_2,\ \dots\rangle \text{ or } \langle s_n\rangle.$$

A sequence $\langle s_n\rangle$ in $\mathbb{R}^N$ is said to **converge** to a point $s \in \mathbb{R}^N$ if, for all $\epsilon > 0$, there exists a natural number $M \in \mathbb{N}$ such that for all $m \geq M$,

$$d(s_m,\ s) < \epsilon.$$

Example 1.7. Let $\epsilon > 0$ and consider the sequence $\left\langle \frac{1}{\sqrt{m}} \right\rangle$ in $\mathbb{R}$. This sequence converges to 0. Let $M > \frac{1}{\epsilon^2}$ and $m \geq M$. Then

$$\begin{aligned} \left\| \frac{1}{\sqrt{m}} - 0 \right\| &= \frac{1}{\sqrt{m}} \\ &\leq \frac{1}{\sqrt{M}} \\ &< \epsilon \end{aligned}$$

as desired.

The **infimum** of a set $A \subset \mathbb{R}$ is the greatest lower bound of A, while the **supremum** of A is its least upper bound.

Example 1.8. Let $A = (0,\ 5) \cup (\pi,\ 7]$ The infimum of A is 0, while the supremum of A is 7.

The **diameter** of a set A of points in Euclidean space is the supremum over the distances between its points, denoted $\operatorname{diam}(A)$.

Example 1.9. Consider the set $A = \{(5,\ 2),\ (1,\ -3),\ (0,\ 1)\}$ in $\mathbb{R}^2$. The maximum distance between any two points in this set is $\sqrt{26}$ (between the first and third points), so $\operatorname{diam}(A) = \sqrt{26}$.

Section 1.2: Introduction to Topological Spaces

The field of topology is entirely centered around objects known as *topological spaces* and *continuous* functions between them. In this section, we will rigorously define topological spaces, as well as give concise ways to discuss their structure. In later sections, we will define the continuity of function between them and discuss their interactions with various topological properties.

We will also introduce a selection of ideas from topology that will be called upon to further our discussion about *simplicial complexes* throughout Chapter 2.

Definition of a Topology and Open Sets

Definition 1.10. Let X be a set. A **topology** $\mathcal{T}$ on X is a collection of subsets of X, each called an **open set**, such that:

(i) The empty set and X are open sets,

(ii) The intersection of finitely many open sets is an open set, and

(iii) The union of any collection of open sets is an open set.

The set X together with a topology $\mathcal{T}$ on X is called a **topological space**.

Although a topological space is made up of two things, a set X and a collection $\mathcal{T}$ of subsets of X which form a topology on X, we tend to refer to the set X as a topological space and leave it implicitly understood that there is a topology on X. In the case we discuss two topologies on X, we will denote the collections of subsets of X which form the topology differently.

Example 1.10. Consider the set $X = \{a, b, c, d\}$ and the following collections of subsets of X:

$$\mathcal{T}_1 = \{\varnothing, \{a\}, X\}$$
$$\mathcal{T}_2 = \{\varnothing, \{b\}, \{b, c\}, X\}$$

We show that $\mathcal{T}_1$ and $\mathcal{T}_2$ form topologies on X as follows:

(i) Each collection contains $\varnothing$ and X, i.e., $\varnothing$ and X are open sets.

(ii) In $\mathcal{T}_1$:

$$\varnothing \cap \{a\} = \varnothing, \qquad \varnothing \cap X = \varnothing, \qquad \{a\} \cap X = \{a\}.$$

In $\mathcal{T}_2$:

$$\begin{aligned} &\varnothing \cap \{b\} = \varnothing, &&\varnothing \cap \{b, c\} = \varnothing, &&\varnothing \cap X = \varnothing, \\ &\{b\} \cap X = \{b\}, &&\{b, c\} \cap X = \{b\}, &&\{b\} \cap \{b, c\} = \{b\}. \end{aligned}$$

Thus, the intersection of a finite number of sets in $\mathcal{T}_1$ are in $\mathcal{T}_1$, and similarly for $\mathcal{T}_2$, i.e., the intersection of finitely many open sets is an open set.

(iii) The third condition is left as an exercise to the reader.

It is also apparent from the previous example that one may define different topologies on the same set X. Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be topologies on X. If $\mathcal{T}_1 \subseteq \mathcal{T}_2$, then $\mathcal{T}_2$ is said to be **finer** than $\mathcal{T}_1$ and $\mathcal{T}_1$ is said to be **coarser** than $\mathcal{T}_2$. Furthermore, if $\mathcal{T}_2$ is finer than $\mathcal{T}_1$ but is not equal to $\mathcal{T}_1$, then $\mathcal{T}_2$ is **strictly finer** than $\mathcal{T}_1$. **Strictly coarser** is defined similarly.

The **trivial topology** on a set X consists of the open sets $\varnothing$ and X itself, and is the coarsest topology on X. The **discrete topology** on X, on the other hand, defines all subsets of X as open and is thus the finest topology on X.

The following proposition allows us to determine whether or not a set is open in a topological space, which proves useful when comparing topologies which are defined differently. To simplify future discussion we utilize the following definition:

Let X be a topological space and $x \in X$. An open set U containing x is said to be a **neighborhood of** x.

Proposition 1.11. Let X be a topological space and let $A \subseteq X$. A is open in X if and only if for each $x \in A$, there is a neighborhood U of x such that $x \in U \subseteq A$.

Proof. First, suppose A is open in X and $x \in A$. If we let $U = A$, then U is a neighborhood of x for which $x \in U \subseteq A$.

Now suppose that for every $x \in A$ there exists a neighborhood U_x of x such that $x \in U_x \subseteq A$. By the Union Lemma (Lemma 1.3), A is exactly equal to the union of all such U_x. As A is the union of open sets, it is therefore open. ■

Basis for a Topology

To describe a finite-dimensional vector space, one typically defines a vector basis (which is much shorter and more efficient than trying to describe every vector in the space!) In the same vein, it is possible to generate a topology utilizing a smaller collection of open sets, called a *basis*, and their unions.

Definition 1.12. Let X be a set and $\mathcal{B}$ be a collection of subsets of X. We say $\mathcal{B}$ is a **basis (for a topology on** X) if the following statements hold:

(i) For each $x \in X$, there is a B in $\mathcal{B}$ such that $x \in B$, and

(ii) If $B_1,\ B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

The sets in $\mathcal{B}$ are referred to as **basis elements**.

Example 1.11. Consider the real numbers, and the collection of open intervals $\mathcal{B} = \{(a,\ b) \subset \mathbb{R} \mid a < b\}$ in $\mathbb{R}$. We will show this is a basis for a topology on $\mathbb{R}$:

(i) Every point in $\mathbb{R}$ is contained in an open interval, and thus in a set in $\mathcal{B}$.

(ii) If two open intervals intersect, they do so in an open interval. Hence, any point in the intersection of two sets in $\mathcal{B}$ is contained in another set in $\mathcal{B}$ within the intersection.

Example 1.12. Let X be a set and $\mathcal{B} = \{\{x\} \mid x \in X\}$. We will show this is a basis for a topology on X:

(i) Every $x \in X$ lies in the set $\{x\}$ in $\mathcal{B}$.

(ii) Every pair of distinct sets in $\mathcal{B}$ is disjoint, so the second condition is vacuously satisfied.

A basis $\mathcal{B}$ generates a topology $\mathcal{T}$ on X in the following way: The open sets in $\mathcal{T}$ are the empty set and every set that is equal to a union of basis elements in $\mathcal{B}$. To prove this, we utilize the following Lemma:

Lemma 1.13. Let $\mathcal{B}$ be a basis. Assume that $B_1,\ B_2,\ \ldots,\ B_n \in \mathcal{B}$ and that $x \in \bigcap_{i=1}^{n} B_i$. Then there exists $B' \in \mathcal{B}$ such that $x \in B' \subset \bigcap_{i=1}^{n} B_i$.

Proof. We show this via induction on n. The $n = 2$ condition holds by the second condition in the definition of a basis (Definition 1.12).

Assume the result is true for $n - 1$. Let $B_1,\ \ldots,\ B_n$ be sets in $\mathcal{B}$ and $x \in \bigcap_{i=1}^{n} B_i$. Clearly $x \in \bigcap_{i=1}^{n-1} B_i$, and the induction hypothesis implies that there exists $B^* \in \mathcal{B}$ such that $x \in B^* \subset \bigcap_{i=1}^{n-1} B_i$.

Thus $x \in B^* \cap B_n$. By the second condition in the definition of a basis, there exists $B' \in \mathcal{B}$ such that $x \in B' \subset B^* \cap B_n$. Since $B^* \subset \bigcap_{i=1}^{n-1} B_i$ as well, it follows that $x \in B' \subset \bigcap_{i=1}^{n} B_i$. Thus, if the result holds for $n - 1$, it holds for n, and the result follows by induction. ∎

Proposition 1.14. The topology $\mathcal{T}$ generated by a basis $\mathcal{B}$ is a topology.

Proof. We will show that the topology $\mathcal{T}$ generated by a basis $\mathcal{B}$ satisfies the three conditions for a topology (Definition 1.10):

(i) The empty set is in $\mathcal{T}$ by definition, and since every point in X is contained in some basis element, X is the union of all basis elements and is therefore in $\mathcal{T}$ as well.

(ii) Let $V = \displaystyle\bigcap_{i=1}^{n} U_i$, where each U_i is in $\mathcal{T}$. If any U_i is empty, then so is V and V is open. Thus, assume each U_i is a union of basis elements. We will show that V is a union of basis elements as well. Let $x \in V$. Then $x \in U_i$ for each U_i.

Since each U_i is a union of basis elements, there exists a basis element B_i such that $x \in B_i \subset U_i$ for each i. Then $x \in \displaystyle\bigcap_{i=1}^{n} B_i$ and, by Lemma 1.13, there exists a basis element B_x such that $x \in B_x \subset \displaystyle\bigcap_{i=1}^{n} B_i \subset V$. It follows from the Union Lemma (Lemma 1.3) that $V = \displaystyle\bigcup_{x \in V} B_x$, and therefore V is a union of basis elements. Thus, a finite intersection of open sets is open in $\mathcal{T}$.

(iii) Let $V = \bigcup U_\alpha$ where each U_α is either an empty set or a union of basis elements. If each U_α is empty, then so is V; on the other hand if some U_α is non-empty, then V is the union of basis elements since it is the union of basis elements making up each U_α. Thus, an arbitrary union of open sets is open in $\mathcal{T}$. ■

Note that each basis element is itself an open set in the topology generated by the basis.

Example 1.13. Below are several topologies on $\mathbb{R}$, which would be difficult to describe without using a basis:

(i) The **standard topology on** $\mathbb{R}$ is generated by the aforementioned basis $\mathcal{B} = \{(a, b) \subset \mathbb{R} \mid a < b\}$.

(ii) The **lower limit topology** is that which is generated by the basis $\mathcal{B} = \{[a, b) \subset \mathbb{R} \mid a < b\}$. Note that the intervals $[a, b)$ and (a, b) are both open in the lower limit topology, as (a, b) is the union of basis elements $B_i = [a - \frac{1}{i}, b)$ where $i = 1, 2, \ldots$

(iii) The **upper limit topology** is that which is generated by the basis $\mathcal{B} = \{(a,\ b] \subset \mathbb{R} \mid a < b\}$. Note that $(a,\ b)$ is open in the upper limit topology for a similar reason.

The following proposition is reminiscent of Proposition 1.11, and is also useful in determining whether or not a set is open by investigating its elements. In Proposition 1.11, a set was open if and only if each point in that set was in an open neighborhood contained within the set. In the following proposition, we show that a set is open if and only if each point in that set is in a basis element contained entirely in that set.

Proposition 1.15. Let X be a set and $\mathcal{B}$ be a basis for a topology on X. Then U is open in the topology generated by $\mathcal{B}$ if and only if for each $x \in U$, there exists a basis element $B_x \in \mathcal{B}$ such that $x \in B_x \subset U$.

Proof. Suppose U is an open set in the topology generated by $\mathcal{B}$ and that $x \in U$. Since U is a union of basis elements, there is at least one basis element B' making up that union that contains x. Clearly then, $x \in B' \subset U$.

Now suppose that $U \subset X$ is such that for each $x \in U$ there exists a $B_x \in \mathcal{B}$ such that $x \in B_x \subset U$. By the Union Lemma (Lemma 1.3), $U = \bigcup_{x \in U} B_x$, and therefore U is a union of basis elements. Thus, U is an open set in the topology generated by $\mathcal{B}$. ∎

It is possible that two bases will generate the same topology. We will consider two unique bases for topologies on $\mathbb{R}^2$, and show that they generate the same topology. Recall that the open 2-ball of radius r centered at x is defined as $B^2(x,\ r) = \left\{y \in \mathbb{R}^2 \mid d(x,\ y) < r\right\}$. We first consider the collection of all such open balls. Before we prove that this collection is a basis and generates a topology, we prove the following supporting lemma.

Lemma 1.16. Let $y \in \mathbb{R}^2$ and $r > 0$. Then for every $x \in B^2(y,\ r)$ there exists an $\epsilon > 0$ such that $B^2(x,\ \epsilon) \subset B^2(y,\ r)$.

Proof. Suppose $x \in B^2(y,\ r)$, i.e., $d(x,\ y) < r$. Choose ϵ such that $0 < \epsilon < r - d(x,\ y)$. We claim that $B^2(x,\ \epsilon) \subset B^2(y,\ r)$. Suppose $z \in B^2(x,\ \epsilon)$. Then we have

$$\begin{aligned} d(y,\ z) &< d(y,\ x) + d(x,\ z) \\ &< d(y,\ x) + \epsilon \\ &< d(y,\ x) + r - d(x,\ y) \\ &< r. \end{aligned}$$

■

Proposition 1.17. The collection

$$\mathcal{B} = \{B^2(x,\ \epsilon) \mid x \in \mathbb{R}^2,\ \epsilon > 0\}$$

is a basis for a topology on $\mathbb{R}^2$.

Proof. Since each $x \in \mathbb{R}^2$ is contained in the basis element $B^2(x,\ 1)$, the first condition for a basis is satisfied.

Suppose $x \in B^2\left(p,\ r_1\right) \cap B^2\left(q,\ r_2\right)$. By Lemma 1.16, there exist $\epsilon_1,\ \epsilon_2 > 0$ such that $B^2\left(x,\ \epsilon_1\right) \subset B^2\left(p,\ r_1\right)$ and $B^2\left(x,\ \epsilon_2\right) \subset B^2\left(q,\ r_2\right)$. Let $\epsilon = \min\{\epsilon_1,\ \epsilon_2\}$. Then

$$B^2(x,\ \epsilon) \subset B^2\left(x,\ \epsilon_1\right) \cap B^2\left(x,\ \epsilon_2\right) B^2\left(p,\ r_1\right) \cap B^2\left(q,\ r_2\right),$$

and it follows that $\mathcal{B}$ satisfies the conditions for a basis. ■

Note, the collection of open N-balls generates the **standard topology on** $\mathbb{R}^N$. In $\mathbb{R}^2$, open rectangles and open half-planes are also open in the standard topology.

Before we assert that the collection of open rectangles in $\mathbb{R}^2$ is a basis that generates the standard topology, we prove the following useful proposition. This proposition is reminiscent of Proposition 1.11 and it will likewise continue to be useful when introducing new constructions for topologies in later sections.

Proposition 1.18. Let X be a set with topology $\mathcal{T}$, and let $\mathcal{C}$ be a collection of open sets in X. If, for each open set U in X and for each $x \in U$, there is an open set V in $\mathcal{C}$ such that $x \in V \subset U$, then $\mathcal{C}$ is a basis that generates the topology $\mathcal{T}$.

Proof. First, we check that $\mathcal{C}$ is a basis. Let $x \in X$. Since X itself is an open set, there is an open set $V \in \mathcal{C}$ such that $x \in V \subset X$. Therefore every point in X is contained in some open set V in the collection $\mathcal{C}$.

Suppose now that $x \in V_1 \cap V_2$ for two open sets $V_1,\ V_2 \in \mathcal{C}$. As V_1 and V_2 are open, $V_1 \cap V_2$ is open and, by our hypothesis, there must be an open set $V_3 \in \mathcal{C}$ such that $x \in V_3 \subset V_1 \cap V_2$. Hence, $\mathcal{C}$ is a basis.

We now must check that the topology $\mathcal{T}'$ generated by $\mathcal{C}$ coincides with $\mathcal{T}$. Suppose U is open in $\mathcal{T}$. Then, by the hypothesis, for every $x \in U$ there is an open set $V \in \mathcal{C}$ such that $x \in V \subset U$. By Proposition 1.15, U is open in $\mathcal{T}'$, i.e., $\mathcal{T} \subset \mathcal{T}'$. Now suppose W is open in $\mathcal{T}'$. Then W is the union of open sets in $\mathcal{C}$, all of which are open in $\mathcal{T}$, i.e., W is open in $\mathcal{T}$ and $\mathcal{T}' \subset \mathcal{T}$. Thus, the result holds. ■

It is straightforward to utilize this to show that the collection of open rectangles in $\mathbb{R}^2$ is a basis that generates the standard topology.

Proposition 1.19. On the plane $\mathbb{R}^2$, let

$$\mathcal{B} = \{(a,\ b) \times (c,\ d) \subset \mathbb{R}^2 \mid a < b,\ c < d\}.$$

Then $\mathcal{B}$ is a basis, and the topology $\mathcal{T}'$ generated by $\mathcal{B}$ is the standard topology on $\mathbb{R}^2$.

The proof of this is left to the reader.

Closed Sets

It is now time to introduce *closed sets*, the complementary concept to open sets.

Definition 1.20. A subset A of a topological space X is **closed** if the set $X - A$ is open.

Example 1.14. Consider $\mathbb{R}$ equipped with the standard topology. Since $(0, 1)$ is open, $(-\infty, 0] \cup [1, \infty)$ is closed.

It is not difficult to show that all closed intervals $[a, b]$ and single-point sets $\{c\}$ are closed in the standard topology on $\mathbb{R}$.

To shorten the statement of the following proposition, we define the following: Let $[a, b]$ and $[c, d]$ be bounded closed intervals in $\mathbb{R}$. Then $[a, b] \times [c, d] \subset \mathbb{R}^2$ is called a **closed rectangle**.

Proposition 1.21. Closed balls and closed rectangles are closed sets in the standard topology on $\mathbb{R}^2$.

Proof. The proof for closed balls is left to the reader.

Let $A = [a, b] \times [c, d]$ be a closed rectangle in $\mathbb{R}^2$. $\mathbb{R}^2 - A$ may be expressed as the union of four open half-planes: $\{(x, y) \mid x < a\}$, $\{(x, y) \mid x > b\}$, $\{(x, y) \mid y < c\}$, and $\{(x, y) \mid y > d\}$. Since each of these half-planes is an open set (which can be shown by utilizing Proposition 1.11), $\mathbb{R}^2 - A$ is an open set and thus A is closed. ■

By definition, a set C is closed if its complement is open. On the other hand, a set U is open if its complement is closed, which may be shown in a brief argument: Consider a set U in a topology X whose complement, $X - U$ is closed. As $X - U$ is closed, $X - (X - U) = U$ is open.

Unlike a door or a window in your favorite internet browser, a set may be both open and closed, typically referred to as being *clopen*. We illustrate this fact in the following examples.

Example 1.15. Consider a topological space X. By definition, both $\varnothing$ and X are open in X. Notice that $X - \varnothing = X$ and $X - X = \varnothing$ are both open in X, meaning that $\varnothing$ and X are closed in X as well.

Example 1.16. Consider a set X equipped with the discrete topology. Every subset A of X is an open set. The compliment of A, being another subset of X, is also open in X. Therefore, every subset A of X is also closed in X, i.e., every subset of X is clopen in the discrete topology.

Although we typically will refrain from doing so here, it is possible to define a topological space X by denoting which subsets of X are closed in the topology. The following theorem illustrates this.

Theorem 1.22. Let X be a topological space. The following statements about the collection of closed sets in X hold:

(i) $\varnothing$ and X are closed,

(ii) The intersection of any collection of closed sets is a closed set, and

(iii) The union of finitely many closed sets is a closed set.

Proof. (i) X is closed as $\varnothing = X - X$ is open. Similarly, $\varnothing$ is closed as $X = X - \varnothing$ is open.

The proofs of (ii) and (iii) are left to the reader. ■

In the beginning of this section, we very briefly mentioned *topological properties*, in reference to their interactions with *continuous* functions. Topological properties lend us an understanding of the space as a whole, without necessarily needing to define its open sets or a basis which generates it. A noteworthy topological property is being *Hausdorff*, and it will be mentioned again several times throughout this text.

Definition 1.23. A topological space X is **Hausdorff** if for every pair of distinct points x and y in X, there exist disjoint neighborhoods U and V of x and y, respectively.

Example 1.17. The standard topology on $\mathbb{R}$ is, in fact, Hausdorff. For any two distinct real numbers a and b (with $a < b$), we may construct disjoint open intervals which contain them:

$$U = \left(a - 1,\ \frac{a+b}{2}\right) \text{ and } V = \left(\frac{a+b}{2},\ b + 1\right).$$

Example 1.18. Every set X with the discrete topology is Hausdorff: given distinct points x and y in X, the sets $\{x\}$ and $\{y\}$ are disjoint neighborhoods of x and y, respectively.

The following proposition is an example of the useful properties of Hausdorff spaces:

Proposition 1.24. If X is a Hausdorff space, then every single-point subset of X is closed.

Proof. Let x be an arbitrary element of X. We will show $X - \{x\}$ is open.

Let $y \in X - \{x\}$ be arbitrary. As X is Hausdorff, there are disjoint neighborhoods U and V containing x and y, respectively. It follows that $x \notin V$, and therefore $V \subset X - \{x\}$. Since every $y \in X - \{x\}$ is in a neighborhood contained in $X - \{x\}$, it follows from Proposition 1.11 that $X - \{x\}$ is open. ■

Relating Subsets to Open and Closed Sets

The ideas of the *interior*, *closure*, *limit points*, and *boundary* of an arbitrary subset of a topological space allow us to construct related open and closed sets. These concepts will prove fruitful in our later discussion on *simplicial complexes* in the next chapter.

Definition 1.25. Let A be an arbitrary subset of a topological space X. The *interior* of A, denoted $\text{Int}(A)$ is the union of all open sets contained in A. The **closure** of A, denoted $\overline{A}$, is the intersection of all closed sets containing A. Clearly, the interior of A is open while its closure is closed, and we have that $\text{Int}(A) \subseteq A \subseteq \overline{A}$.

We obtain the following properties of the interior and closure directly from their definitions:

(i) If U is an open set in X and $U \subseteq A$, then $U \subset \text{Int}(A)$.

(ii) If C is closed in X and $A \subseteq C$, then $\overline{A} \subseteq C$.

(iii) If $A \subseteq B$, then $\text{Int}(A) \subseteq \text{Int}(B)$ and $\overline{A} \subseteq \overline{B}$.

(iv) A is open if and only if $A = \text{Int}(A)$.

(v) A is closed if and only if $A = \overline{A}$.

Example 1.19. Consider $A = [0,\ 1)$ as a subset of $\mathbb{R}$ with the standard topology. Then $\operatorname{Int}(A) = (0,\ 1)$ and $\overline{A} = [0,\ 1]$.

Now consider A as a subset of $\mathbb{R}$ with the discrete topology. Then $\operatorname{Int}(A) = \overline{A} = [0,\ 1)$.

The following propositions provide easy ways of determining whether a point is in the interior or closure of a given set A. For brevity, their proofs are left to the reader.

Proposition 1.26. Let X be a topological space, A a subset of X, and $y \in X$. Then y is in the interior of A if and only if there exists an open set U such that $y \in U \subseteq A$.

Proposition 1.27. Let X be a topological space, A a subset of X, and $y \in X$. Then y is in the closure of A if and only if every open set U containing y interests A.

Limit points ease the discussion on the interior and closure of A, as well as provide a useful way of determining if a set is closed or not. Before we discuss limit points however, we must define the convergence of a sequence in a topological space. It should be reminiscent of the definition of convergence of a sequence from analysis.

Definition 1.28. A sequence $\langle x_n \rangle$ in a topological space X **converges to** $x \in X$ if, for every neighborhood U of x, there exists a natural number $M \in \mathbb{N}$ such that for all $n \geq M$, $x_n \in U$. x is the **limit** of $\langle x_n \rangle$, and we write $\langle x_n \rangle \to x$.

Definition 1.29. Let A be a subset of a topological space X. A point x in X is a **limit point** of A if every neighborhood of x intersects A in a point other than x. Alternatively, x is a limit point of A if there exists a sequence $\langle a_n \rangle$ entirely in A which converges to x.

Example 1.20. Consider the set $A = [0,\ 1]$ in $\mathbb{R}$ with the standard topology. Every point in A is also a limit point of A. For an arbitrary x in A, every neighborhood of x may be expressed as the union of open intervals, one of which is $(x - \epsilon,\ x + \epsilon)$ for some $\epsilon > 0$. This set has a non-empty intersection with A that includes at least one point other than x.

Now consider $A = (0,\ 1]$ in $\mathbb{R}$ with the discrete topology. 0 is a limit point of A as the sequence $\left\langle \frac{1}{n} \right\rangle$ is entirely in A and converges to 0.

This example shows that a limit point of A may or may not actually be contained in A.

The following proposition relates limit points and the closure of a set.

Proposition 1.30. Let A be a subset of a topological space X, and let A' denote the set of limit points of A. Then $\overline{A} = A \cup A'$.

Proof. We proceed by double containment.

First note that $A \subseteq \overline{A}$, so it suffices to show that $A' \subseteq \overline{A}$. Let $x \in A'$. Then every neighborhood of x intersects A. By Proposition 1.27, x is in the closure of A.

Now suppose that x is in the closure of A. Either $x \in A$ or $x \in \overline{A} - A$. In the former case, the result follows immediately. In the latter case, Proposition 1.27 implies that every open set containing x intersects A. Since $x \notin A$, such an intersection must contain points other than x, and the result follows. ■

Along with the basic properties of closure, Proposition 1.30 implies the following corollary.

Corollary 1.31. A subset A of a topological space is closed if and only if it contains all of its limit points.

The notions of closure and interior lend themselves naturally to defining the *boundary* of a set.

Definition 1.32. The **boundary** of a subset A of a topological space X, denoted $\mathrm{Bd}\,(A)$, is the set $\overline{A} - \mathrm{Int}\,(A)$.

Example 1.21. Consider the set $A = [0,\ 1]$ in $\mathbb{R}$ with the standard topology. The boundary of A consists of the points 0 and 1.

Now consider $A = (0,\ 1]$ in $\mathbb{R}$ with the discrete topology. Its boundary is empty.

The following are notable properties of the boundary of A whose proofs rely primarily on the definition of boundary and ideas from set theory:

(i) The boundary of A consists exactly of the points of x of X whose neighborhoods intersect both A and $X - A$.

(ii) $\mathrm{Bd}(A)$ is closed.

(iii) $\mathrm{Bd}(A) \subseteq A$ if and only if A is closed.

(iv) $\mathrm{Bd}(A) \cap A = \varnothing$ if and only if A is open.

(v) $\mathrm{Bd}(A) = \varnothing$ if and only if A is both closed and open.

Compact Sets and the Heine-Borel Theorem

To aid in our later discussion of *simplicial complexes*, which are topological spaces that "live in" $\mathbb{R}^N$, we introduce the idea of *compact* sets as well as one of the most fundamental theorems regarding compact sets in $\mathbb{R}^N$, the *Heine-Borel Theorem*. The hurried reader may simply take the results that follow and press onward, although the idea of compactness is important enough to our work to warrant study.

Consider a subset K of a topological space X. An **open cover** of K is a collection $\mathcal{C}$ of open subsets of X whose union contains K, and is said to "cover" K. A subcollection $\mathcal{D}$ of $\mathcal{C}$ which covers K is said to be a **subcover** of K. $\mathcal{D}$ is sometimes referred to as a "subcover of $\mathcal{C}$" at the risk of misguiding the reader.

Definition 1.33. A subset K of a topological space X is said to be **compact** if every open cover of K has a finite subcover.

Example 1.22. The open interval $A = (0,\ 1)$ in $\mathbb{R}$ equipped with the usual topology is not compact. Consider the open cover $\mathcal{C} = \left\{ \left(\frac{1}{n},\ 1\right) \mid n \in \mathbb{N} \right\}$ of A. Any finite subcollection would have some maximal N, and thus would no longer cover A.

Consider the closed interval $[0,\ 1]$ in $\mathbb{R}$ with the standard topology. The collection

$$\mathcal{D} = \left\{ \left(\frac{1}{n},\ 2\right) \mid n \in \mathbb{N} \right\} \cup \left(-1,\ \frac{1}{10}\right)$$

has a finite subcover, namely $\left\{ \left(-1,\ \frac{1}{10}\right),\ \left(\frac{1}{11},\ 2\right) \right\}$. However, this is not sufficient to claim that $[0,\ 1]$ is compact in $\mathbb{R}$.

As evident by the above example, it can be difficult to prove that a particular set is compact in a topological space. The Heine-Borel theorem make it much easier to show that a given set is compact in $\mathbb{R}^N$.

Lemma 1.34. Let $\mathbb{R}$ be equipped with the standard topology. If $A \subset \mathbb{R}$ is closed and bounded, then A contains its supremum and infimum.

Proof. We prove here that the infinum of A is in A, the proof that the supremum of A is in A is similar.

Let a be the infimum of A. If $a \in A$, we are done. By Proposition 1.30, it suffices to show that a is a limit point of A.

Let $r > 0$ and consider the open interval $(a - r,\ a + r)$. As a is the infimum of A, this interval must intersect A, or else $a + \frac{r}{s}$ would be a greater lower bound of A. Thus, a is a limit point of A. ■

Theorem 1.35. THE HEINE-BOREL THEOREM. Consider a non-empty subset K of $\mathbb{R}$ equipped with the standard topology. K is compact if and only if K is bounded and closed.

Proof. Let K be compact in $\mathbb{R}$ equipped with the standard topology.

Fix some $x_0 \in \mathbb{R}$, and let U_n denote the open ball of radius n centered on x_0. Consider the open cover $\left\{U_n \mid n \in \mathbb{N}\right\}$ of K. As K is compact, there is some finite subcollection $\left\{U_1,\ \dots,\ U_n\right\}$ which still covers K. But $\bigcup_{i=1}^{n} U_i = U_n = B^1\left(x_0,\ n\right)$, and thus K is bounded.

Let $U = \mathbb{R} - K$ and consider $x \in U$. By Proposition 1.11, it suffices to produce an $\epsilon > 0$ such that $B^1\left(x,\ \epsilon\right) \subset \mathbb{R} - K$.

For each point p in K, let $U_p = \left\{q \in \mathbb{R} \mid |q - x| > \frac{|p-x|}{2}\right\}$. As $\mathbb{R} - U_p$ is a closed ball, U_p is open. Also note that each p in K is in U_p as well, and thus $\left\{U_p \mid p \in K\right\}$ is an open cover of K.

As K is compact, there exists a finite subcover $\left\{U_{p_1},\ \dots,\ U_{p_n}\right\}$. Let $\epsilon = \min\left\{\frac{\left|p_1-x\right|}{2},\ \dots,\ \frac{\left|p_n-x\right|}{2}\right\} > 0$. We will show $B(x,\ \epsilon) \cap K = \varnothing$ via contradiction. Therefore, $\mathbb{R} - K$ is closed, meaning that K is open.

Suppose $q \in B^1(x,\ \epsilon) \cap K$. As $a \in B(x,\ \epsilon)$, $|x - q| < \epsilon$. Since $q \in K$, q is in some U_{p_i} for $i = 1,\ \ldots,\ n$, which implies that $|q,\ x| > \frac{|c-p_i|}{2} \geq \epsilon$, a contradiction.

Conversely, let $K \subseteq \mathbb{R}$ be closed and bounded. Let a be the infimum of K and b be the supremum of K. By Lemma 1.34, $a,\ b \in K$.

Define, for all x in $\mathbb{R}$,

$$K_x = K \cap (-\infty,\ x],$$

and let C be an arbitrary open cover for K. We will show C has a finite subcover.

Let $G = \left\{x \in \mathbb{R} \mid K_x \text{ is covered by a finite subcover of } C\right\}$. Note, a is in G as $K_a = \{a\}$ must be covered by one element of C. Also observe that $x \in G$ implies that $y \in G$ for all $y < x$, as $K_y \subset k_x$ and K_x is covered.

We will show that G is not bounded above, as that implies b must be in G, i.e., $K_b = K$ has a finite subcover.

Assume by way of contradiction that G is bounded above, and let m be its supremum. We have two cases:

Case 1. Let $m \in K$. Then there exists a $U_0 \in C$ that contains m. By 1.11, there exists an $\epsilon > 0$ such that $(m - \epsilon,\ m + \epsilon) \subseteq U_0$.

As m is the supremum of G, $m - \epsilon \in G$ and thus $K_{m-\epsilon}$ has a finite subcover $\left\{U_1,\ \ldots,\ U_n\right\}$. Hence, $U_0 \cup \bigcup_{i=1}^n U_i$ is a finite subcover of $K_{m+\frac{\epsilon}{2}}$, which implies $m + \frac{\epsilon}{2} \in G$. This contradicts the fact that m is an upper bound of G.

Case 2. Let $m \notin K$. Then $m \in \mathbb{R} - K$, which is open. Thus, there exists an $\epsilon > 0$ such that $(m - \epsilon,\ m + \epsilon)$ is contained in $\mathbb{R} - K$. Then:

$$\begin{aligned}
K_{m+\frac{\epsilon}{2}} &= K \cap \left(-\infty,\ m + \frac{\epsilon}{2}\right] \\
&= \left(K \cap \left(-\infty,\ m - \frac{\epsilon}{2}\right]\right) \cup \left(K \cap \left(m - \frac{\epsilon}{2},\ m + \frac{\epsilon}{2}\right)\right) \\
&= K_{m-\frac{\epsilon}{2}} \cup \varnothing
\end{aligned}$$

But $m - \frac{\epsilon}{2} \in G$, so a finite subcover of C covers $K_{m-\frac{\epsilon}{2}} = K_{m+\frac{\epsilon}{2}}$, so $m + \frac{\epsilon}{2} \in G$, contradicting the assumption that m bounds G from above.

Therefore, G is not bounded above. Thus, $b \in G$, so $K \subset K_b$ has a finite subcover. ■

The Heine-Borel Theorem also generalizes to $\mathbb{R}^N$, the proof of which we will not cover here.

Section 1.3: The Subspace, Product, and Quotient Topologies

Given particular topological spaces, it is possible to create a multitude of new ones. The *subspace topology* evolves from a subset of a topological space, the *product topology* is derived from the product of two topological spaces, and the *quotient topology* is analogous to gluing spaces together or collapsing them.

In this section, we rigorously define such topologies and discuss their properties and interpretations. However, we are primarily concerned with only the subspace and product topologies, and our discussion on the quotient topology may be skimmed without any risk of not understanding our later discussions.

The Subspace Topology

Definition 1.36. Let X be a topological space and let $Y \subset X$. The **subspace topology** on Y is defined as $\mathcal{T}_Y = \{U \cap Y \mid U \text{ is open in } X\}$ and, with this topology, Y is called a **subspace** of X. We say that $V \subset Y$ is **open in** Y if V is an open set in the subspace topology on Y.

In essence, the open sets of the subspace topology on Y are the intersections of Y with all open sets of X. We show that the subspace topology on Y is in fact a topology.

(i) $\varnothing$ and Y are both open in Y, since $\varnothing = \varnothing \cap Y$ and $Y = X \cap Y$.

(ii) Let $V_1, \dots, V_n$ be open in Y. Then for each i there exists a set U_i, which is open in X, such that $V_i = U_i \cap Y$. Hence,

$$\begin{aligned} V_1 \cap \cdots \cap V_n &= (U_1 \cap Y) \cap \cdots \cap (U_n \cap Y) \\ &= (U_1 \cap \cdots \cap U_n) \cap Y. \end{aligned}$$

Since $U_1 \cap \cdots \cap U_n$ is open in X, it follows that $V_1 \cap \cdots \cap V_n$ is open in Y (and thus finite intersections of sets are open in Y.)

(iii) Suppose $\{V_\alpha\}$ is a collection of open sets in Y. Then for each α, there exists an open set U_α in X such that $V_\alpha = U_\alpha \cap Y$. Therefore,

$$\bigcup V_\alpha = \bigcup (U_\alpha \cap Y) = \left(\bigcup U_\alpha\right) \cap Y.$$

As $\bigcup U_\alpha$, it follows that $\bigcup V_\alpha$ is open in Y (and thus arbitrary unions of sets are open in Y).

Example 1.23. Consider $\mathbb{R}$ with the standard topology and $I = [0,\ 1]$ as a subset of $\mathbb{R}$. With the subspace topology, the open sets in I are all open intervals in $\mathbb{R}$ intersected with $[0,\ 1]$. The open sets in this subspace fall into three categories (as well as finite intersections and arbitrary unions of these sets):

(i) The empty set and the closed interval $[0,\ 1]$.

(ii) Open intervals of the form $(a,\ b)$ where $0 < a < b < 1$.

(iii) Half-open intervals of the form $[0,\ a)$ or $(a,\ 1]$, where $0 < a < 1$.

Note that a set may be open in the subspace topology on I, even though it may be not open in $\mathbb{R}$.

Example 1.24. The subspace topology that $\mathbb{Z}$ inherits from the standard topology on $\mathbb{R}$ is in fact the discrete topology. Since open intervals in $\mathbb{R}$ are open, and each integer is contained in an open interval which contains no other point, the single-point sets containing each integer are open in the subspace topology on $\mathbb{Z}$. Arbitrary unions of these sets are open as well, meaning that every subset of $\mathbb{Z}$ is open in $\mathbb{Z}$.

In general, for any subset Y of $\mathbb{R}^N$, the **standard topology** on Y is the topology that Y inherits as a subspace of $\mathbb{R}^N$ with the standard topology.

Naturally, for a subspace Y of a topological space X, a set $C \subset Y$ is closed in Y if C is closed in the subspace topology on Y. The following propositions allow us to "translate" between X and Y, respectively.

Proposition 1.37. Let X be a topological space, and let $Y \subset X$ have the subspace topology. Then $C \subset Y$ is closed in Y if and only if $C = D \cap Y$ for some closed set D in X.

Proof. Let $C \subset Y$ be closed. As $Y - C$ is open in the subspace topology on Y, there exists some open set $W \subset X$ such that $Y - C = W \cap Y$. Thus, $X - W$ is closed in X. We will show that $C = (X - W) \cap Y$ via double containment.

Let $x \in C$. So $x \in Y$ and $x \notin Y - C$. Hence, $x \notin W \cap Y$, but $x \in Y$ so $x \notin W$. Therefore, $x \in X - W$, which implies $x \in (X - W) \cap Y$.

Let $x \in (X - W) \cap Y$, i.e., $x \in X - W$ and $x \in Y$. As $x \notin W$, $x \notin W \cap Y$ which implies $x \notin Y - C$. However, as $x \in Y$, x must be in C. Thus, the result holds.

Conversely, let $C = D \cap Y$ for some closed $D \subset X$. Then $X - D$ is open in X and $(X - D) \cap Y$ is open in Y. We will show $(X - D) \cap Y = Y - C$, or equivalently $Y - ((X - D) \cap Y) = C$ via double containment.

Let $x \in Y - ((X - D) \cap Y)$. So $x \in Y$ and $x \notin (X - D)$. The latter implies that $x \in D$. Therefore, $x \in D \cap Y = C$. Let $x \in C$. So $x \in D \cap Y$. Thus, $x \notin (X - D)$ and $x \in Y - ((X - D) \cap Y)$. ■

Proposition 1.38. Let X be a topological space and $\mathcal{B}$ a basis for the topology on X. If $Y \subset X$, then the collection

$$\mathcal{B}_Y = \{B \cap Y \mid B \in \mathcal{B}\}$$

is a basis for the topology on Y.

Proof. Note that $\mathcal{B}_Y$ is a collection of open sets in the subspace topology on Y. Let W be an open set in the subspace topology on Y and let $y \in W$ be arbitrary. Then $W = U \cap Y$, where U is open in X. There exists a basis element $B \in \mathcal{B}$ such that $y \in B \subset U$. Thus, $y \in B \cap Y \subset U \cap Y = W$. Since $B \cap Y \in \mathcal{B}_Y$, it now follows from Theorem 1.18 that $\mathcal{B}_Y$ is a basis for the subspace topology on Y. ■

Example 1.25. Consider the circle $S^1 \subset \mathbb{R}^2$ (its center and radius are irrelevant) with the standard topology. As open balls form a basis for the standard topology on $\mathbb{R}^2$, their intersection with S^1 form a basis for the standard topology on S^1. These intersections are of the form of open intervals on the circle, consisting of all points between two angles in the circle.

The Product Topology

Unfortunately, constructing the topology on the product of two topological spaces is not as straightforward as constructing the subspace topology. For topological spaces X and Y, we cannot define the collection $\mathcal{C}$ of open sets in $X \times Y$ to be the product of open sets in X and Y. However, we can use $\mathcal{C}$ as a basis to generate such a topology.

Definition 1.39. Let X and Y be topological spaces and $X \times Y$ be their product. The **product topology** on $X \times Y$ is the topology generated by the basis

$$\mathcal{B} = \{U \times V \mid U \text{ is open in } X \text{ and } V \text{ is open in } Y\}.$$

We verify that this is in fact a basis for a topology on $X \times Y$.

Proposition 1.40. The collection $\mathcal{B}$ is a basis for a topology on $X \times Y$.

Proof. Every point (x, y) is in $X \times Y$, and $X \times Y \in \mathcal{B}$, satisfying the first condition for a basis.

Now assume $(x, y) \in (U_1 \times V_1) \cap (U_2 \times V_2)$ where U_1, U_2 are open in X and V_1 and V_2 are open sets in Y. Let $U_3 = U_1 \cap U_2$ and $V_3 = V_1 \cap V_2$. Then U_3 is open in X and V_3 is open in Y, and hence $U_3 \times V_3 \in \mathcal{B}$. Also,

$$U_3 \times V_3 = (U_1 \cap U_2) \times (V_1 \cap V_2) = (U_1 \times V_1) \cap (U_2 \times V_2),$$

and thus $(x, y) \in U_3 \times V_3 \subset (U_1 \times V_1) \cap (U_2 \times V_2)$, satisfying the second condition for a basis. ∎

Example 1.26. Consider $\mathbb{R}$ with the standard topology. Then the product topology on $\mathbb{R} \times \mathbb{R}$ is generated by the collection of open rectangles $(a, b) \times (c, d) \subset \mathbb{R} \times \mathbb{R}$. As noted in the previous section, this is the standard topology on $\mathbb{R}^2$.

As with open sets, products of closed sets are closed sets in the product topology. However, this does not describe all closed sets in the product space.

The basis described in the definition of product space $X \times Y$ (Definition 1.39) is very, very large. It consists of all possible products of all open sets in X with all open sets in Y. The following proposition provides a much smaller basis for $X \times Y$ by utilizing the bases for X and Y individually.

Proposition 1.41. If $\mathcal{C}$ is a basis for X and $\mathcal{D}$ is a basis for Y, then

$$\mathcal{E} = \{C \times D \mid C \in \mathcal{C} \text{ and } D \in \mathcal{D}\}$$

is a basis that generates the product topology on $X \times Y$.

Proof. Each set $C \times D \in \mathcal{E}$ is an open set in the product topology; therefore by Proposition 1.18, it suffices to show that for every open set W in $X \times Y$ and every point $(x, y) \in W$, there is a set $C \times D$ in $\mathcal{E}$ such that $(x, y) \in C \times D \subset W$.

Since W is open in $X \times Y$, we know that there are open sets U in X and V in Y such that $(x, y) \in U \times V \subset W$. So $x \in U$ and $y \in V$. Since U is open in X, there is a basis element $C \in \mathcal{C}$ such that $x \in C \subset U$. Similarly, since V is open in Y, there is a basis element $D \in \mathcal{D}$ such that such that $y \in D \subset V$. Thus, $(x, y) \in C \times D \subset U \times V \subset W$.

By Proposition 1.18, it follows that $\mathcal{E} = \{C \times D \mid C \in \mathcal{C} \text{ and } D \in \mathcal{D}\}$ is a basis for the product topology on $X \times Y$. ■

When picturing the product space, we can imagine a copy of the second space at each point in the first space, or vice versa. In other words, we can “drag” one space along another, and all points obtained by this process are the points in the product space.

Example 1.27. Let S^1 be the circle and $I = [0, 1]$. By the above analogy, we can imagine the product space $S^1 \times I$ as the space obtained by dragging I along the circle, obtaining an annulus.

Example 1.28. The product space $S^1 \times S^1$ is the torus. This is evident when utilizing the aforementioned analogy: imagine dragging a circle in a circular path perpendicular to its face. By Proposition 1.41, the basis of $S^1 \times S^1$ consists of rectangular patches.

Now consider two topological spaces, X and Y, as well as two sets $A \subset X$ and $B \subset Y$. We now have two natural ways to topologize $A \times B$: 1) as a subspace of $X \times Y$, and 2) as the product of subspaces $A \subset X$ and $B \subset Y$. However, these two approaches result in the same topology.

Before we prove that these two topologies are in fact equivalent, we prove the following lemma.

Lemma 1.42. For $A, C \subset X$ and $B, D \subset Y$, $(C \times D) \cap (A \times B) = (C \cap A) \times (D \cap B)$.

Proof. We proceed by double containment. Let $(x, y) \in (C \times D) \cap (A \times B)$. So $(x, y) \in (C \times D)$, i.e., $x \in C$ and $y \in D$. Similarly, $x \in A$ and $y \in B$. Thus, as $x \in C \cap A$ and $y \in D \cap B$, $(x, y) \in (C \cap A) \times (D \cap B)$.

Let $(x, y) \in (C \cap A) \times (D \cap B)$. So $x \in C$ and $x \in A$. Similarly, $y \in D$ and $y \in B$. Thus, $(x, y) \in C \times D$ and $(x, y) \in A \times B$ and the result holds. ■

Theorem 1.43. Let X and Y be topological spaces, and assume $A \subset X$ and $B \subset Y$. Then the topology on $A \times B$ as a subspace of the product $X \times Y$ is the same as the product topology on

$A \times B$, where A has the subspace topology inherited from X and B has the subspace topology inherited from Y.

Proof. Let $\mathcal{T}_1$ be the topology inherited by $A \times B$ as a subspace of the product space $X \times Y$ and let $\mathcal{T}_2$ be the topology inherited by $A \times B$ as a product space of the subspaces $A \subset X$ and $B \subset Y$. We will show $\mathcal{T}_1 = \mathcal{T}_2$ via double containment.

Let U be open in $\mathcal{T}_1$. We will show that U can be written as the product of open sets in A and B, each equipped with the subspace topology, i.e., U is open in $\mathcal{T}_2$. As U is open in $\mathcal{T}_1$, it may be written as the intersection of an open set $C \times D \subset X \times Y$ and $A \times B$, i.e. $U = (C \times D) \cap (A \times B)$. Note that C is open in X and D is open in Y. By Lemma 1.42, $U = (C \cap A) \times (D \cap B)$. Under the subspace topology, $C \cap A$ is open in A and $D \cap B$ is open in B. Thus, $\mathcal{T}_1 \subset \mathcal{T}_2$.

Let U be open in $\mathcal{T}_2$. We will show U can be written as the intersection of an open set in $X \times Y$ and $A \times B$, i.e. U is open in $\mathcal{T}_1$. As U is open in $\mathcal{T}_2$, it may be written at the product of open sets in A and B, i.e. $U = C \times D$ where C is open in A with the subspace topology and D is open in B with the subspace topology. As C is open in A with the subspace topology, $C = P \cap A$ where P is open in X. Similarly, $D = Q \cap B$ where Q is open in Y. By Lemma 1.42, $U = C \times D = (P \cap A) \times (Q \cap B) = (P \times Q) \cap (A \times B)$ As P is open in X and Q is open in Y, $P \times Q$ is open in $X \times Y$. Thus, we have expressed U as the intersection of an open set in $X \times Y$ (namely $P \times Q$) with $A \times B$. ■

We may extend this definition to a product $X_1 \times \cdots \times X_n$ of n topological spaces. The collection

$$\mathcal{B} = \left\{ U_1 \times \cdots \times U_n \mid U_i \text{ is open in } X_i \text{ for each } i = 1, \ \ldots, \ n \right\}$$

is a basis for $X_1 \times \cdots \times X_n$. Also note that this product space is "associative" in the following sense: Given three topological spaces X, Y, and Z, the following product spaces are *topolog-*

ically equivalent, i.e., they are indistinguishable as far as topology is concerned:

$$X \times Y \times Z \cong (X \times Y) \times Z \cong X \times (Y \times Z).$$

Example 1.29. As mentioned previously, the standard topology on $\mathbb{R}^2$ is the same as the product topology on $\mathbb{R}\times\mathbb{R}$. This result generalizes to Euclidean spaces of arbitrary finite dimension, i.e., the standard topology on $\mathbb{R}^N$ is the same as the product topology that results from taking the product of N copies of $\mathbb{R}$.

The Quotient Topology

The quotient topology allows us to construct a topological model that mimics the process of gluing together or collapsing parts of one or more topological spaces. For example, we may glue the opposite edges of a rectangle together to form a torus. We now define the quotient topology.

Definition 1.44. Let X be a topological space and A a set (not necessarily a subset of X). Let $p : X \to A$ be a surjective map. Define a subset U of A to be open in A if and only if $p^{-1}(U)$ is open in X. The resulting collection of open sets in A is called the **quotient topology induced by** p, and the function p is called the **quotient map**. The topological space A is called the **quotient space**.

We verify that the quotient space is in fact a topological space.

Proposition 1.45. Let $p : X \to A$ be a quotient map. The quotient topology induced by p is a topology.

Proof. (i) The set $p^{-1}(\varnothing) = \varnothing$ is open in X. The set $p^{-1}(A) = X$ is open in X as well. Thus, $\varnothing$ and A are open in the quotient topology.

(ii) Suppose each of the sets U_i, $i = 1, ..., n$, is open in the quotient topology on A. Then $p^{-1}\left(\bigcap U_i\right) = \bigcap p^{-1}\left(U_i\right)$, which is a finite intersection of open sets in X, and therefore is open in X. Hence, $\bigcap U_i$ is open in the quotient topology, and it follows that the finite intersection of open sets in the quotient topology is an open set in the quotient topology.

(iii) Suppose each of the sets in the collection $\{U_\beta\}_{\beta\in B}$ is open in the quotient topology on A. Then $p^{-1}\left(\bigcup\left(U_\beta\right)\right) = \bigcup p^{-1}\left(U_\beta\right)$, which is a union of open sets in X, and therefore is open in X. Thus, $\bigcup\left(U_\beta\right)$ is open in the quotient topology, implying that the arbitrary union of open sets in the quotient topology is an open set in the quotient topology.

Hence, the quotient topology is a topology on A. ■

Example 1.30. Consider $\mathbb{R}$ with the standard topology, and define the quotient map

$$p: \mathbb{R} \to \{a,\ b,\ c\} \text{ by } p(x) = \begin{cases} a, & \text{if } x < 0, \\ b, & \text{if } x = 0, \\ c, & \text{if } x > 0. \end{cases}$$

The quotient topology on A has the following open sets:

$$\{a,\ b,\ c\}, \qquad \{a,\ c\}, \qquad \{a\}, \qquad \{c\}, \qquad \varnothing.$$

The definition of a quotient space does not particularly seem to relate to the idea of gluing parts of spaces together or collapsing them. However, these notions arise when we consider quotient spaces defined on partitions of X, which are of great interest.

Let X^* be a collection of mutually disjoint subsets of X whose union is X, and let $p: X \to X^*$ be the surjective map which takes each point in X to the corresponding element of X^* that contains it. Then p induces a quotient topology on X^*, which we think about in the following way: Let S be a subset of X in X^*. Then p identifies all points in S to the same element of X^*, thereby "collapsing" S to a single point in the quotient space. A set U of points in X^* is open in the quotient topology on X^* exactly when the union of the subsets of X, corresponding to the points in U, is an open subset of X.

Example 1.31. Let $X = [0,\ 1]$ and consider the partition X^* that is made up of single point sets $\{x\}$ for all $0 < x < 1$ and the two-point set $D = \{0,\ 1\}$. In the quotient topology on X^*, we think of D as a single point, i.e., we glue the endpoints of X together and form S^1.

A subset of X^* which does not contain D, a collection of single-point sets, is open in X^* exactly when the union of those single point sets is open in $(0,\ 1)$. A subset of X^* that contains D is open in X^* when the union of all sets making up the subsets is an open set inside $[0,\ 1]$. This open subset must contain 0 and 1, and thus must contain the intervals $[0,\ a)$ and $(b,\ 1]$, which are open in the subspace topology on $[0,\ 1]$.

Section 1.4: Continuous Functions and Homeomorphisms

Now that we have discussed a wide variety of topological spaces, it is time to introduce the idea of *continuous* functions, and *homeomorphisms*. A topology on a set defines a notion of proximity and continuous functions preserve this proximity. A homeomorphism is a special type of continuous mapping that allows us to define the notion of *topological equivalence*.

Continuous Functions

We begin by stating the definition of a continuous function from a typical calculus or real analysis course, and then state the definition of a continuous function from a topological perspective.

Definition 1.46. A function $f : \mathbb{R} \to \mathbb{R}$ is **continuous** if, for every $x_0 \in \mathbb{R}$ and every $\epsilon > 0$, there exists a $\delta > 0$ such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. This is the $\epsilon - \delta$ **definition of continuity**.

Let X and Y be topological spaces. A function $f : X \to Y$ is **continuous** if $f^{-1}(V)$ is open in X for every open set V in Y. This is the **open set definition of continuity**.

We may translate the $\epsilon - \delta$ definition of continuity into more topological concepts. Let X

and Y be topological spaces. A function $f : X \to Y$ is continuous if, for every $x \in X$ and every open set U containing $f(x)$, there exists a neighborhood V of x such that $f(V) \subset U$. (U is playing the role of the ϵ interval, while V is playing the role of the δ interval.) This translation is actually equivalent to the open set definition, as shown in the following proposition.

Proposition 1.47. A function $f : X \to Y$ is continuous in the open set definition of continuity if and only if, for every $x \in X$ and every open set U containing $f(x)$, there exists a neighborhood V of x such that $f(V) \subset U$.

Proof. First, suppose that the open set definition holds for functions $f : X \to Y$. Let $x \in X$ and an open set $U \subset Y$ containing $f(x)$ be given. Set $V = f^{-1}(U)$. It follows that $x \in V$ and that V is open in X since f is continuous by the open set definition. Clearly $f(V) \subset U$.

Now assume that for every $x \in X$ and every open set U containing $f(x)$, there exists a neighborhood V of x such that $f(V) \subset U$. We show that $f^{-1}(W)$ is open in X for every open set W in Y. Thus, let W be an arbitrary open set in Y. To show that $f^{-1}(W)$ is open in X, choose an arbitrary $x \in f^{-1}(W)$. It follows that $f(x) \in W$, and therefore there exists a neighborhood V_x of x in X such that $f\left(V_x\right) \subset W$, or equivalently $V_x \subset f^{-1}(W)$. Thus, for an arbitrary $x \in f^{-1}(W)$ there exists an open set V_x such that $x \in V_x \subset f^{-1}(W)$. Proposition 1.11 implies that $f^{-1}(W)$ is open in X. ■

Example 1.32. The following families of functions are continuous over the domains for which they are defined:

(i) Rational functions;

(ii) Trigonometric functions and their inverses;

(iii) Exponential functions, and;

(iv) Logarithmic functions.

There are a variety of nice ways to check if a function is continuous, as illustrated in the following proposition.

Proposition 1.48. Let X and Y be topological spaces and let $\mathcal{B}$ be a basis for the topology on Y. Then $f : X \to Y$ is continuous if and only if $f^{-1}(B)$ is open in X for every $B \in \mathcal{B}$

Proof. Suppose $f : X \to Y$ is continuous. Then $f^{-1}(V)$ is open in X for every V open in Y. Since every basis element is open in Y, it follows that $f^{-1}(B)$ is open in X for all $B \in \mathcal{B}$.

Now suppose $f^{-1}(B)$ is open in X for every $B \in \mathcal{B}$. Let V be an open set in Y. Then V is a union of basis elements, say $V = \bigcup B_\alpha$. Thus,

$$f^{-1}(V) = f^{-1}\left(\bigcup B_\alpha\right) = \bigcup f^{-1}(B_\alpha).$$

By assumption, each set $f^{-1}\left(B_\alpha\right)$ is open in X, and thus so is their union. Hence, f is continuous. ■

Proposition 1.49. Let X and Y be topological spaces. A function $f : X \to Y$ is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set $C \subset Y$.

A proof of this is left to the reader.

Continuous functions also preserve several properties, such as the convergence of sequences.

Proposition 1.50. Assume that $f : X \to Y$ is continuous. If a sequence $\langle x_n \rangle$ in X converges to a point x, then the sequence $\left\langle f\left(x_n\right)\right\rangle$ in Y converges to $f(x)$.

Proof. Let U be an arbitrary neighborhood of $f(x)$ in Y. Since f is continuous, $f^{-1}(U)$ is open in X. Furthermore, $f(x) \in U$ implies that $x \in f^{-1}(U)$. The sequence $\langle x_n \rangle$ converges to x; thus, there exists a natural number $M \in \mathbb{N}$ such that $x_n \in f^{-1}(U)$ for all $n \geq M$. It follows that $f\left(x_n\right) \in U$ for all $n \geq N$, and therefore the sequence $\left\langle f\left(x_n\right)\right\rangle$ converges to $f(x)$. ■

The following proposition and lemma provide convenient ways of creating continuous functions from our few examples.

Proposition 1.51. Let $f : X \to Y$ and $g : Y \to Z$ be continuous. Then the composition function $g \circ f : X \to Z$ is continuous.

Proof. Suppose that $f : X \to Y$ and $g : Y \to Z$ are continuous, and let U be an open set in Z. Then $(g \circ f)^{-1}(U) = f^{-1}\left(g^{-1}(U)\right)$. Since g is continuous, $g^{-1}(U)$ is open in Y, and since f is continuous, $f^{-1}\left(g^{-1}(U)\right)$ is open in X. Hence, by Definition 1.46, $f \circ g$ is continuous. ■

Lemma 1.52. THE PASTING LEMMA. Let X be a topological space and let A and B be closed subsets of X such that $A \cup B = X$. Assume that $f : A \to Y$ and $g : B \to Y$ are continuous and $f(x) = g(x)$ for all x in $A \cap B$. Then $h : X \to Y$ defined by

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

is a continuous function.

Proof. By Proposition 1.49, it suffices to show that if C is closed in Y, then $h^{-1}(C)$ is closed in X. Thus, suppose C is closed in Y. Note that $h^{-1}(C) = f^{-1}(C) \cup g^{-1}(C)$. Since f is continuous, it follows by Proposition 1.49 that $f^{-1}(C)$ is closed in A. Proposition 1.37 then implies that $f^{-1}(C) = D \cap A$ where D is closed in X. Now, D and A are both closed in X, and $f^{-1}(C) = D \cap A$; therefore $f^{-1}(C)$ is closed in X. Similarly, $g^{-1}(C)$ is closed in X. Thus, $h^{-1}(C)$ is the union of two closed sets in X, and is therefore closed in X as well. Hence, h is continuous. ■

The Pasting Lemma proves especially useful, as it allows us to construct continuous piecewise functions.

Homeomorphisms

Homeomorphisms provide the most fundamental sense of topological equivalence: they preserve all the properties of a topology, and therefore give a bijective correspondence between points and between open sets in two topological spaces.

Definition 1.53. Let X and Y be topological spaces, and let $f : X \to Y$ be a bijection with inverse $f^{-1} : Y \to X$. If both f and f^{-1} are continuous functions, then f is said to be a

homeomorphism. If there exists a homeomorphism between X and Y, we say that X and Y are **homeomorphic** or **topologically equivalent**, and we denote this by $X \cong Y$.

We take a brief detour and note that the following three facts imply that topological equivalence is indeed an equivalence relation:

(i) The function $id : X \to X$, defined by $id(x) = x$, is a homeomorphism.

(ii) If $f : X \to Y$ is a homeomorphism, then so is $f^{-1} : X \to Y$.

(iii) If $f : X \to Y$ and $g : Y \to Z$ are homemorphisms, then so is $g \circ f : X \to Z$.

Also, note that saying that the inverse f^{-1} of a bijective function f is continuous is equivalent to saying that the image of every open set under f is open. Similarly, saying that a bijection f is continuous is equivalent to saying that the image of every open set under f^{-1} is open.

Example 1.33. The function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$ is a bijection with inverse given by $f^{-1}(x) = \sqrt[3]{x}$. Since f and f^{-1} are both rational functions on $\mathbb{R}$, they are both continuous, and thus f is a homeomorphism.

Example 1.34. The following subsets of $\mathbb{R}$ are homeomorphic to $\mathbb{R}$ itself (with arbitrary real scalars $a < b$):

(i) Open intervals: $(a,\ b)$, $(-\infty,\ a)$, $(a,\ \infty)$, $\mathbb{R}$.

(ii) Closed, bounded intervals: $[a,\ b]$.

(iii) Half-open intervals and closed, unbounded intervals: $[a,\ b)$, $(a,\ b]$, $(-\infty,\ a]$, $[a,\ \infty)$.

We provide some examples of homeomorphisms which imply these equivalences below:

(i) For open, bounded intervals $(a,\ b)$, the function

$$f(x) = \left(\frac{a+b}{\pi}\right) \arctan(x) + \left(\frac{b+a}{2}\right)$$

with inverse

$$f^{-1}(x) = \tan\left(\left(\frac{\pi}{b-a}\right)\left(x - \frac{b+a}{2}\right)\right)$$

suffices. For open, unbounded intervals $(a,\ \infty)$, the function $g(x) = \log(x-a)$ with inverse $g^{-1}(x) = e^x + a$ suffices.

(ii) For closed, bounded intervals, the function

$$h(x) = \left(\frac{a+b}{2}\right)\sin(x) + \frac{a+b}{2}$$

with inverse

$$h^{-1}(x) = \arcsin\left(\frac{2}{a+b}x - 1\right)$$

suffices.

Any property of a topological space preserved by homeomorphism is considered a **topological property**. Below, we prove our previous statement that being Hausdorff is a topological property.

Proposition 1.54. If $f : X \to Y$ is a homeomorphism and X is Hausdorff, then Y is Hausdorff.

Proof. Suppose that X is Hausdorff and $f : X \to Y$ is a homeomorphism. Let $x,\ y \in Y$ be distinct. Then $f^{-1}(x),\ f^{-1}(y) \in X$ are distinct as well. Thus, there exist disjoint open sets U and V containing $f^{-1}(x)$ and $f^{-1}(y)$, respectively. It follows that $f(U)$ and $f(V)$ are disjoint sets containing x and y respectively. Therefore, Y is Hausdorff. ■

We have discussed showing that two spaces are topologically equivalent by producing a homeomorphism between them. It is sometimes easier to prove that two topological spaces are not equivalent by showing that they do not share some topological property, e.g., one is Hausdorff while the other is not.

Example 1.35. A topological space is said to be **discrete** if all of its subsets are open. Being discrete is a topological property.

Clearly, $\mathbb{R}$ equipped with the discrete topology is discrete and $\mathbb{R}$ equipped with the trivial topology is not discrete. Therefore, they cannot be homeomorphic.

Example 1.36. Another topological property is being **connected**. A topological space X is connected if there does not exist a pair of disjoint non-empty open sets whose union is X.

$\mathbb{R}$ equipped with the standard topology is connected (see [1]). The interval $[0, 1)$ equipped with the discrete topology is not connected.

Chapter 2: Simplicial Complexes

In topological data analysis, *simplicial complexes* serve as the main data structure used to represent topological spaces. They allow us to capture many important properties about a space while remaining geometrically simple and allowing efficient computations.

Section 2.1 gives a rigorous introduction to *simplices*, *geometric simplicial complexes* and their usual topology, *simplicial maps* between geometric simplicial complexes, as well as *abstract simplicial complexes* and their *geometric realizations*.

Section 2.2 discusses *Čech complexes* and *Vietoris-Rips complexes*, which are simplicial complexes constructed from point clouds in $\mathbb{R}^N$. These objects are foundational to several areas of topological data analysis, and serve as the primary objects of interest in our discussion of *persistent homology* in Chapter 5.

Section 2.1: Simplices and Simplicial Complexes

Before we state rigorous definitions of simplices and simplicial complexes, we shall provide more intuitive descriptions to assist the reader in their understanding. We intend the reader to consult these descriptions, along with the accompanying figures, in order to gain a deeper appreciation for these objects.

In essence, an *n*-simplex is a generalization of the triangle found in geometry. A 2-simplex is exactly a triangle, while a 0-simplex is a point, a 1-simplex is an edge, a 3-simplex is a tetrahedron, and so forth.

The *proper faces* of an *n*-simplex are all of the lower-dimension simplices contained in it.

For example, the proper faces of a 3-simplex are the 2-dimensional faces of the tetrahedron, along with its edges and vertices.

A *geometric simplicial complex* is a collection of *n*-simplices whose intersections with one another are faces of each. Figure 2.1 shows a geometric simplicial complex with 0-, 1-, and 2-simplices.

An *abstract simplicial complex* is a purely set theoretic description of geometric simplicial complexes, and we will show using the Geometric Realization Theorem (Theorem 2.15) that they are indeed analogous.

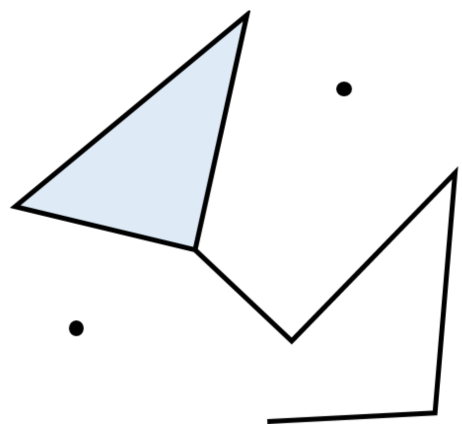


Figure 2.1: *A geometric simplicial complex. The* 0*-simplices are the vertices of, the* 1*-simplices are the edges, and the* 2*-simpliex is the triangle.*

Simplices

To rigorously define a simplex, we must first introduce one main idea from analytic geometry of Euclidean space: *geometric independence*.

We refer to points and vectors in $\mathbb{R}^N$ somewhat interchangeably, using each term to bring up associations with different ideas.

A given set of points $\{a_0, \dots, a_n\}$ of $\mathbb{R}^N$ is said to be **geometrically independent** if, for any real scalars t_i, the equations

$$\sum_{i=0}^{n} t_i = 0 \text{ and } \sum_{i=0}^{n} t_i a_i = \mathcal{O},$$

where $\mathcal{O}$ is the origin in $\mathbb{R}^N$, imply that $t_0 = t_1 = \cdots = t_n = 0$.

It is not difficult to show that the set $\{a_0, \ldots, a_n\}$ is geometrically independent if and only if the vectors

$$a_1 - a_0, \ \ldots, \ a_n - a_0$$

are linearly independent in the sense of ordinary linear algebra.

Hence, sets containing one point, two distinct points, three non-collinear points, four non-coplanar points (and so on) form geometrically independent sets in $\mathbb{R}^N$.

Definition 2.1. Let $\{a_0, \ldots, a_n\}$ be a geometrically independent set in $\mathbb{R}^N$. The ***n*-simplex** σ spanned by $a_0, \ldots, a_n$ is the set of all points $x \in \mathbb{R}^N$ such that

$$x = \sum_{i=0}^{n} t_i\, a_i, \text{ where } \sum_{i=0}^{n} t_i = 1$$

and $t_i \geq 0$ for all i.

The scalars t_i are uniquely determined by x and are called the **barycentric coordinates** of the point x of σ with respect to $a_0, \ldots, a_n$. It is common to refer to the barycentric coordinates of a given point x as $t_i(x)$.

Example 2.1. A 0-simplex is spanned by a single point a_0, and thus consists of just a point.

A 1-simplex is spanned by two points a_0 and a_1, and thus is a line segment with these points as endpoints. More specifically, it is all points x of the form

$$x = t\, a_0 + (1 - t)\, a_1$$

with $0 \leq t \leq 1$, which should remind one of the parametric equation for a line segment.

A 2-simplex is spanned by three points a_0, a_1 and a_2, and is thus a triangle with these points as vertices. It is made up of the point a_0 along with all points x of the form

$$x = t_0\, a_0 + \left(1 - t_0\right) \left[\left(\frac{t_1}{1 - t_0}\right) a_1 + \left(\frac{t_2}{1 - t_0}\right) a_2\right]$$

where $0 \leq t_0 < 1$, $0 \leq t_1,\ t_2 \leq 1$, and $t_0 + t_1 + t_2 = 1$. The expression in square brackets represents the line segment between a_1 and a_2, as when $t_0 = 0$, $t_2 = 1 - t_1$ which is reminiscent of the 1-simplex discussed earlier. Non-zero values of t_0 result in all points of the 2-simplex not on the line segment joining a_1 and a_2.

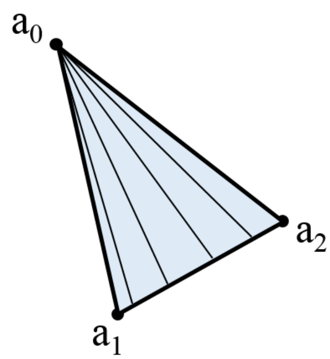


Figure 2.2: *A 2-simplex with vertices $a_0,\ a_1,\ a_2$. The points along the edge a_1a_2 are given by the expression in square brackets in the equation for a 2-simplex. The points along the line segments emanating from a_0 are given by a fixed value of t_0 and varying t_1 and t_2.*

The **vertices** of a simplex σ are the points $a_0,\ \ldots,\ a_n$ which span it, and the number n is called the **dimension** of σ.

Any simplex spanned by a subset of $\{a_0,\ \ldots,\ a_n\}$ is called a **face** of σ. Specifically, the face of σ spanned by $a_1,\ \ldots,\ a_n$ is called the **face opposite the vertex** a_0.

The faces of σ different from σ itself are called the **proper faces** of σ. Their union is called the **boundary** of σ and is denoted $\mathrm{Bd}(\sigma)$ while the **interior** of σ is defined by the equation $\mathrm{Int}(\sigma) = \sigma - \mathrm{Bd}(\sigma)$. The interior of σ is also sometimes called an **open simplex**.

Example 2.2. The boundary of the 2-simplex in Figure 2.2 is the union of the edges $a_0a_1,\ a_0a_2$, and a_1a_2. Note that for each of these edges, at least one of $t_0,\ t_1$, and t_2 are zero. Its interior are all points not on these edges, i.e., when none of $t_0,\ t_1$, and t_2 are 0. Its proper faces are the vertices $a_0,\ a_1$, and a_2 as well as the edges $a_0a_1,\ a_0a_2$, and a_1a_2.

The statements in Example 2.2 regarding the barycentric coordinates generalize to all simplices: the boundary of a simplex σ consists of all points x with at least one barycentric coordinate $t_i(x)$ equal to zero, while its interior consists of all points x whose barycentric coordinates $t_i(x)$ are all non-zero.

It then follows that every point $x \in \sigma$ is in the interior of some face s of σ.

The following are some properties of simplices, the proofs of which are left to the interested reader:

(i) The barycentric coordinates $t_i(x)$ of x with respect to $a_0, \ldots, a_n$ are continuous functions of x.

(ii) Any n-simplex σ is equal to the union of all line segments joining a_0 to points of the simplex s spanned by $a_1, \ldots, a_n$. Two such line segments intersect only in the point a_0.

(iii) Any n-simplex σ is a compact, convex set in $\mathbb{R}^N$, which equals the intersection of all convex sets in $\mathbb{R}^N$ containing $a_0, \ldots, a_n$.

(iv) Given a simplex σ, there is one and only one geometrically independent set of points spanning σ.

(v) The interior of any simplex σ is convex and is open in the plane P; its closure is σ. Additionally, $\operatorname{Int}(\sigma)$ equals the union of all open line segments joining a_0 to points of $\operatorname{Int}(s)$, where s is the face of σ opposite a.

(vi) There is a homeomorphism of any n-simplex σ with the unit ball B^n that carries $\operatorname{Bd}(\sigma)$ onto the unit sphere S^{n-1}.

Geometric Simplicial Complexes

We are now ready to introduce *geometric simplicial complexes*, their properties, and the standard topology defined on them.

Definition 2.2. A **geometric simplicial complex** K in $\mathbb{R}^N$ is a collection of simplices in $\mathbb{R}^N$ such that:

(i) Every face of a simplex of K is in K, and

(ii) The intersection of any two simplices of K is a face of each of them.

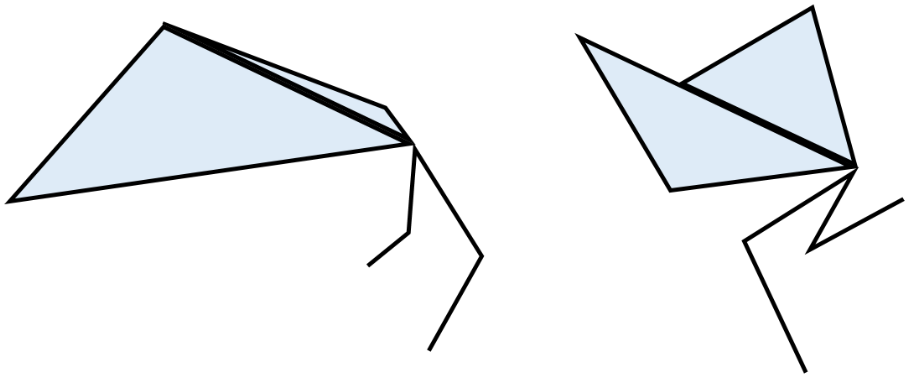


Figure 2.3: *The figure on the left is indeed a geometric simplicial complex, as the intersection of any two simplices is a face of each of them. For instance, the two* 2*-simplices intersect at an edge of each. The figure on the right is not a geometric simplicial complex, as the two* 2*-simplices intersect at a line segment that is not an edge of one of them.*

An alternative condition to (ii) is the following: Every pair of distinct simplices in K have disjoint interiors. We verify this by the following arguments:

We first show that if σ and τ in a geometric simplicial complex K share an interior point, then $\sigma = \tau$; and that if every pair of distinct simplices in a collection K of simplices have disjoint interiors, then their intersection is a face of each of them.

First, let x be a point in the interior of σ and τ, and let $s = \sigma \cap \tau$. If s were a proper face of σ, then x would be in its boundary, contradicting the fact that the interior and boundary are disjoint sets. Thus, $s = \sigma$ and it may be similarly shown that $s = \tau$.

Alternatively, let σ and τ have disjoint interiors and $\sigma \cap \tau$ be non-empty. We show that this intersection is a face σ' of σ that is spanned by the vertices $b_0, \ldots, b_m$ of σ which are in τ via double containment.

Clearly, $\sigma' \subseteq \sigma \cap \tau$ as $\sigma \cap \tau$ is convex and contains $b_0, \ldots, b_m$. Let $x \in \sigma \cap \tau$. Then $x \in \mathrm{Int}(s) \cap \mathrm{Int}(t)$ for some face s of σ and some face t of τ.

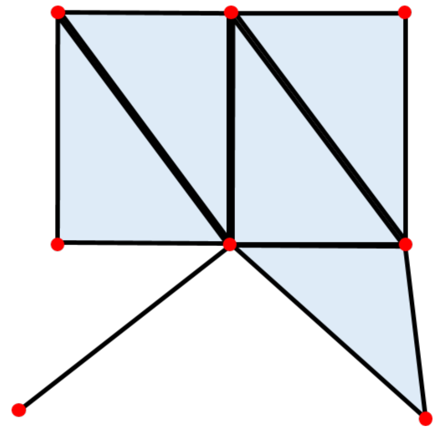


Figure 2.4: *The* 0*-skeleton (red) of a simplicial complex.*

As every pair of distinct simplices in K have disjoint interiors, $s = t$. Hence, the vertices of s lie in τ, so that by definition they are elements of the set $\{b_0, \ldots, b_m\}$. Then s is a face of σ', so $x \in s \subseteq \sigma'$ as desired.

A subcollection L of K that contains all faces of its elements L is a geometric simplicial complex in its own right, called a **subcomplex of** K.

One notable subcomplex of K is the collection of all simplices of K of dimension at most p, called a p**-skeleton** of K and denoted $K^{(p)}$. The points of the collection $K^{(0)}$ are called the **vertices** of K, as shown in Figure 2.4.

By viewing the simplices of a geometric simplicial complex K as subspaces of $\mathbb{R}^N$, we may define a topology on K that is typically finer than the subspace topology and has many useful properties. Before we do this, however, we must introducing the notion of *embedding* one topological space into another.

Definition 2.3. Let X and Y be topological spaces. An **embedding of** X **in** Y is a function $f : X \to Y$ that maps X homeomorphically to the subspace $f(X)$ in Y. We think of this as placing a copy of X in Y.

Definition 2.4. Let $|K|$ be the subset of $\mathbb{R}^N$ that is the union of the simplices of K.

Let K be a simplicial complex in $\mathbb{R}^N$, and $|K|$ be the union of its simplices. Consider each simplex of K with its natural topology as a subspace of $\mathbb{R}^N$. We may equip $|K|$ with a topology by declaring a subset A of $|K|$ to be closed in $|K|$ if and only if $A \cap \sigma$ is closed in σ

(considered as a subspace of $\mathbb{R}^N$, for each σ in K). This indeed defines a topology on $|K|$, for this collection of sets is closed under finite unions and arbitrary intersections.

The set $|K|$ equipped with this topology is called the **underlying space of** K, the **polytope of** K, or (typically for finite simplicial complexes) a **polyhedron**.

The polytope $|K|$ is usually finer than the topology K would inherit as a subspace of $\mathbb{R}^N$ by the following logic, since if A is closed in K equipped with the subspace topology, then $A = B \cap K$ for some closed set B in $\mathbb{R}^N$ (Proposition 1.37). Additionally, $B \cap \sigma$ is closed in σ for each simplex σ of K, meaning $B \cap K = A$ is closed in $|K|$ as well.

Example 2.3. Let K be the collection of all simplices in $\mathbb{R}$ of the form $[m,\ m+1]$, where m is a non-zero integer, along with all simplices of the form $\left[\frac{1}{n+1},\ \frac{1}{n}\right]$ for a positive integer n, along with all faces of these simplices.

The following are some simplices within this geometric simplicial complex:

$$[-4,\ -3]\,,\quad [-3,\ -2]\,,\quad [-2,\ -1]\,,\quad [-1,\ 0]\,,\quad [1,\ 2]\,,\quad [2,\ 3]\,,\quad [3,\ 4]\,,$$

$$\left[\frac{1}{9},\ \frac{1}{8}\right],\quad \left[\frac{1}{8},\ \frac{1}{7}\right],\quad \left[\frac{1}{7},\ \frac{1}{6}\right],\quad \left[\frac{1}{6},\ \frac{1}{5}\right],\quad \left[\frac{1}{5},\ \frac{1}{4}\right],\quad \left[\frac{1}{4},\ \frac{1}{3}\right],\quad \left[\frac{1}{3},\ \frac{1}{2}\right].$$

Note that K is equal to $\mathbb{R}$ as a set, but $|K|$ is not equivalent to $\mathbb{R}$ as a topological space. For instance, the set of points $1/n$ is closed in $|K|$ but not in $\mathbb{R}$ (since it does not contain its limit point, 0).

Example 2.4. Let K be a collection of 1-simplices $\sigma_1,\ \sigma_2,\ \ldots$ and their vertices, where σ_i is the 1-simplex in $\mathbb{R}^2$ having vertices at the origin and $(1,\ 1/i)$. K is indeed a geometric simplicial complex.

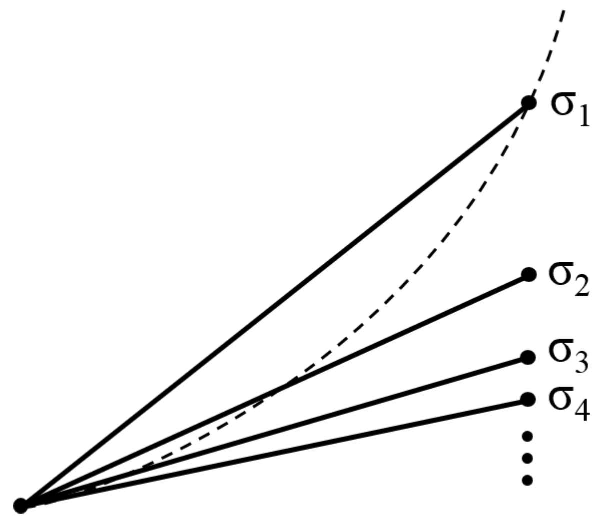


Figure 2.5: *The geometric simplicial complex consisting of simplices σ_i with a common endpoint at the origin and other endpoint* $(1,\ 1/i)$.

The intersection of K with the open parabolic arc $\left\{(x,\ x^2) \mid x > 0\right\}$ is closed in $|K|$, because its intersection with each simplex σ_i is a single point.

The intersection is not closed in K equipped with the subspace topology, however, because it has the origin as a limit point.

There are, however, situations in which $|K|$ is the same as K equipped with the subspace topology, such as when K contains finitely many simplices:

Let A be closed in $|K|$. Then $A \cap \sigma$ is closed in σ for each simplex σ of K, and hence is closed in $\mathbb{R}^N$ as A is the union of finitely many sets $A \cap \sigma$.

The following lemma provides an open set definition for the polytope of K, while the subsequent proposition provides a stronger condition for the equivalence of the topology of $|K|$ and K considered as a subspace.

Lemma 2.5. A set U is open in $|K|$ if and only if $U \cap \sigma$ is open in each simplex σ of K.

Proof. Let U be open in $|K|$. Thus, $|K| - U$ is closed in $|K|$ and $(|K| - U) \cap \sigma = \sigma - (U \cap \sigma)$ is closed in each simplex σ in K. Hence, $U \cap \sigma$ is open for each σ in K, as desired.

Conversely, let $U \cap \sigma$ be open in each simplex $\sigma \in K$. Then $\sigma - (\sigma \cap U)$ is closed in each σ. But $\sigma - (\sigma \cap U) = (|K| - U) \cap \sigma$, so $|K| - U$ is closed in $|K|$. Thus, $|K| - (|K| - U) = U$ is open in $|K|$. ■

Proposition 2.6. Let K be a complex in $\mathbb{R}^N$. The topology induced on K as a subspace of $\mathbb{R}^N$ is the same as the topology of the polytope of K if and only if each point x of $|K|$ lies in an open set of $\mathbb{R}^N$ that intersects only finitely many simplices of K.

Proof. We will prove the forward implication via its contrapositive.

Let $x \in K$ be such that for all open sets $U \subset \mathbb{R}^N$ that contain x, $U \cap \sigma$ is non-empty for infinitely many simplices σ in K. We will produce a set that is closed in $|K|$, but not closed in the subspace topology on K.

Let U_ϵ denote the open ball of radius ϵ centered at x. Consider the sequence $U_1 \supset U_{1/2} \supset U_{1/3} \supset \dots$ in $\mathbb{R}^N$ which all contain x.

For each U_i, choose an x_i such that $x \neq x_i$ and each x_i is in a unique simplex σ_i of K. Let $\langle x_i \rangle$ denote the sequence of all such x_i.

The set made up of points in $\langle x_i \rangle$ is closed in $|K|$ as it intersects each simplex σ of $|K|$ at a single point, meaning that it is closed in each simplex σ.

However, the set made up of points in the sequence is not closed in K with the subspace topology, as it does not contain its limit point x (Corollary 1.31).

Conversely, let each point x of K lie in an open set U of $\mathbb{R}^N$ that intersects only finitely many simplices of K, $|K|$ be the polytope of K, and $\mathcal{S}_K$ be K equipped with the subspace topology. We already know that $\mathcal{S}_K \subseteq |K|$, and thus it suffices to show that $|K| \subseteq \mathcal{S}_K$.

Let U be open in $|K|$. By Lemma 2.5, $U \cap \sigma$ is open in each simplex σ of K. As U intersects only finitely many σ, we consider an arbitrary simplex τ of K that intersects U.

As $U \cap \tau$ is open in τ, there exists an open ball $B(x, \epsilon)$ for each $x \in U \cap \tau$ such that $B(x, \epsilon) \subseteq U \cap \tau$. The union of all such balls is open in $\mathcal{S}_K$, and equals U (by the Union Lemma). ■

Lemmas 2.7, 2.8, and 2.9 are on the topological properties of polyhedra, may be skimmed and accepted without losing any appreciation for our later discussions.

Lemma 2.7. If L is a subcomplex of K, then $|L|$ is a subspace of $|K|$ that is closed in $|K|$

when treated as a subset of $|K|$. In other words, $|L|$ is a **closed subspace** of $|K|$.

Specifically, if $\sigma \in K$, then σ is a closed subspace of $|K|$.

Proof. Suppose A is closed in $|L|$. If σ is a simplex of K, then $\sigma \cap |L|$ is the union of those faces s_i of σ that belong to L. Since A is closed in $|L|$, the set $A \cap s_i$ is closed in s_i and hence closed in σ. Since $A \cap \sigma$ is the finite union of the sets $A \cap s_i$, it is closed in σ. We conclude that A is closed in $|K|$.

Conversely, if B is closed in $|K|$, then $B \cap \sigma$ is closed in σ for each $\sigma \in K$, and in particular for each $\sigma \in L$. Hence, $B \cap |L|$ is closed in $|L|$. ■

Lemma 2.8. A map $f : |K| \to X$ is continuous if and only if the restriction of f to σ, denoted $f\big|_{\sigma}$, is continuous for each $\sigma \in K$.

Proof. If f is continuous, then so is $f\big|_{\sigma}$ since σ is a subspace of K.

Conversely, suppose each map $f\big|_{\sigma}$ is continuous. If C is a closed set of X, then $f^{-1}(C) \cap \sigma = \left(f\big|_{\sigma}\right)^{-1}(C)$, which is closed in σ by continuity of $f\big|_{\sigma}$. Thus, $f^{-1}(C)$ is closed in $|K|$ by definition. ■

We may further generalize our previous notion of barycentric coordinates from just simplicies to geometric simplicial complexes.

If x is a point of the geometric simplicial complex K, then x is interior to precisely one simplex of K, whose vertices are $a_0, \ldots, a_n$. Then,

$$x = \sum_{i=0}^{n} t_i \, a_i,$$

where $t_i > 0$ for each i and $\sum t_i = 1$.

For an arbitrary vertex v of K, we define the **barycentric coordinate** $t_v(x)$ of x with respect to v by setting $t_v(x) = 0$ if v is not one of the vertices a_i, and $t_v(x) = t_i$ if $v = a_i$.

For a fixed v, the function $t_v(x)$ is continuous (in the polytope $|K|$) when restricted to a fixed simplex σ of K, since it is either identically zero on σ or equals the barycentric coordinate of

x with respect to the vertex v of σ in the sense formerly defined. By Lemma 2.8, $t_v(x)$ is continuous on σ.

Lemma 2.9. $|K|$ is Hausdorff.

Proof. Given $x_0 \neq x_1$, there is at least one vertex v such that $t_v(x_0) \neq t_v(x_1)$.

Choose r between these two numbers; then the sets $\{x \mid t_v(x) < r\}$ and $\{x \mid t_v(x) > r\}$ are the required disjoint open sets. ■

Lemma 2.10. If K is finite, then $|K|$ is compact. Conversely, if a subset A of $|K|$ is compact, then $A \subset |K_0|$ for some finite subcomplex K_0 of K.

Proof. Note that by the Heine-Borel Theorem, as each simplex σ is a closed and bounded set, they are compact. If K is finite, then $|K|$ is a finite union of compact subspaces σ, and hence is compact.

Now suppose A is compact and A does not lie in the polytope of any finite subcomplex of K. Choose a point $x_s \in A \cap \mathrm{Int}(s)$ whenever this set is non-empty. Then the set $B = \{x_s\}$ is infinite. Furthermore, every subset of B is closed, since its intersection with any simplex of σ is finite. Being closed and discrete, B has no limit point, contrary to the fact that every infinite subset of a compact space has a limit point. ■

The following three subspaces of the polytope on K are often useful when studying local properties of $|K|$:

(1) Let v be a vertex of K. The **star of v in K**, denoted $\mathrm{St}(v)$ or $\mathrm{St}(v, K)$, is the union of the interiors of those simplices of K that have v as a vertex.

   $\mathrm{St}(v)$ is open in $|K|$, since it consists of all points x of $|K|$ such that $t_v(x) > 0$. Its complement, the union of all simplices of K that do not have v as a vertex, is the polytope of a subcomplex of K.

(2) The closure of $\mathrm{St}(v)$, denoted $\overline{\mathrm{St}}(v)$, is the **closed star of v in K**. It is the union of all simplices of K having v as a vertex, and is the polytope of a subcomplex of K.

(3) The set $\overline{\mathrm{St}}(v) - \mathrm{St}(v)$ is called the **link of v in K** and is denoted $\mathrm{Lk}(v)$. It is also the polytope of a subcomplex of K, the intersection of $\overline{\mathrm{St}}(v)$ and the complement of $\mathrm{St}(v)$.

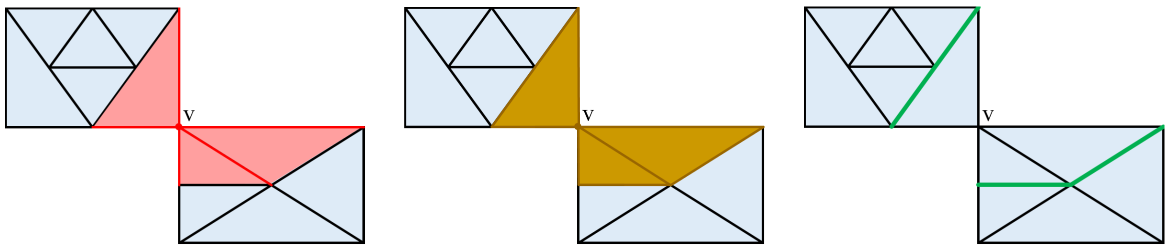


Figure 2.6: *(Left) The open star of v (red). (Middle) The closed star of v (brown). (Right) The link of v (green).*

Note that the star and closed star of v in K are path connected, meaning that every two points in $\mathrm{St}(v)$ and $\overline{\mathrm{St}}(v)$ may be connected by a path contained entirely in them. The link of v in K, however, may or may not be connected.

Simplicial Maps

We are now ready to introduce the idea of *simplicial maps*, which map one geometric simplicial complex into another by mapping each simplex of the former to a simplex of the latter. We will later use these to construct geometric simplicial complexes from abstract simplicial complexes.

Lemma 2.11. Let K and L be geometric simplicial complexes, and let $f : K^{(0)} \to L^{(0)}$ be a map. Suppose that whenever the vertices $v_0, \ldots, v_n$ of K span a simplex of K, the points $f(v_0), \ldots, f(v_n)$ are the vertices of a simplex of L. Then f can be extended to a continuous map $g : |K| \to |L|$ such that

$$x = \sum_{i=0}^{n} t_i v_i \quad \Rightarrow \quad g(x) = \sum_{i=0}^{n} t_i f(v_i).$$

We call g the **(linear) simplicial map** induced by the vertex map f.

Proof. Note that although the vertices $f(v_0), \ldots, f(v_n)$ of L are not necessarily distinct, still they span a simplex τ of L, by hypothesis. When we "collect terms" in the expression for $g(x)$, it is still true that the coefficients are non-negative and their sum is 1; thus $g(x)$ is a point of τ. Hence, g maps the n-simplex σ spanned by $v_0, \ldots, v_n$ continuously to the simplex τ whose vertex set is $\left\{f\left(v_0\right), \ldots, f\left(v_n\right)\right\}$.

The map g is continuous as a map of σ into τ, and hence as a map of σ into $|L|$. Then by Lemma 2.8, g is continuous as a map of $|K|$ into $|L|$. ■

Many of the properties of the simplicial map are those that are retained from the original function f when extended to the continuous map g. In fact, simplicial maps and geometric simplicial complexes are somewhat analogous to continuous maps and topological spaces.

Suppose $g: |K| \to |L|$ and $h: |L| \to |M|$ are simplicial maps. By definition, $x = \sum t_i v_i$ (where v_i are distinct vertices of $\sigma \in K$), so $g(x) = \sum t_i g\left(v_i\right)$.

In fact, this formula holds true even if not all v_i are not distinct, so long as $\left\{v_0, \ldots, v_n\right\}$ is the vertex set of a simplex of K. For example, suppose

$$x = \sum_{i=0}^{n} t_i v_i,$$

where $t_i \geq 0$ for all i and $\sum t_i = 1$; and suppose that $v_0 = v_1$ and the vertices $v_1, \ldots, v_n$ are distinct. We write

$$x = \left(t_0 + t_1\right) v_0 + t_2 v_2 + \cdots + t_n v_n;$$

then by definition

$$\begin{aligned} g(x) &= \left(t_0 + t_1\right) g\left(v_0\right) + t_2 g\left(v_2\right) + \cdots + t_n g\left(v_n\right) \\ &= \sum_{i=0}^{n} t_i g\left(v_i\right). \end{aligned}$$

In the present case, even though the vertices $g\left(v_0\right), \ldots, g\left(v_n\right)$ of L are not necessarily distinct, the following formula holds:

$$h(g(x)) = h\left(\sum t_i g\left(v_i\right)\right) = \sum t_i h\left(g\left(v_i\right)\right).$$

Therefore $h \circ g$ is a simplicial map; in other words, the composition of simplicial maps is simplicial.

The analogy between simplicial maps and continuous functions continues to the idea of a *simplicial homeomorphism*.

Lemma 2.12. Suppose $f : K^{(0)} \to L^{(0)}$ is a bijective correspondence such that the vertices $v_0, \ldots, v_n$ of K span a simplex of K if and only if $f\left(v_0\right), \ldots, f\left(v_n\right)$ span a simplex of L. Then the induced simplicial map $g : |K| \to |L|$ is a homeomorphism, called a **simplicial homeomorphism** or an **isomorphism** of K with L.

Proof. Each simplex σ of K is mapped by g onto a simplex τ of L of the same dimension as σ. We need only show that the linear map $h : \tau \to \sigma$ induced by the vertex correspondence f^{-1} is the inverse of the map $g : \sigma \to \tau$. And for that we note that if $x = \sum t_i v_i$, then $g(x) = \sum t_i f\left(v_i\right)$ by definition; whence

$$\begin{aligned} h(g(x)) &= h\left(\sum t_i f\left(v_i\right)\right) = \sum t_i f^{-1}\left(f\left(v_i\right)\right) \\ &= \sum t_i v_i = x. \end{aligned}$$

■

Corollary 2.13. Let Δ^N denote the complex consisting of an N-simplex and its faces. If K is a finite complex, then K is isomorphic to a subcomplex of Δ^N for some N.

Proof. Let $v_0, \ldots, v_n$ be the vertices of K. Choose $a_0, \ldots, a_N$ to be the geometrically independent points in $\mathbb{R}^N$, and let Δ^N consist of the N simplex they span, along with its faces. The vertex map $f\left(v_i\right) = a_i$ induces an isomorphism of K with a subcomplex of Δ^N. ■

In essence, a sufficiently large N-simplex is chosen such that every vertex of K can be mapped to a vertex of the N-simplex.

Abstract Simplicial Complexes

Abstract simplicial complexes are set theoretic objects greatly resembling geometric simplicial complexes. Many terms used to describe geometric simplicial complexes are also used when discussing abstract simplicial complexes for reasons that will become apparent after discussing the Geometric Realization Theorem.

Definition 2.14. An **abstract simplicial complex** is a collection $\mathcal{S}$ of finite non-empty sets, such that if A is an element of $\mathcal{S}$, then so is every non-empty subset of A.

Example 2.5. The collections

$$\mathcal{S}_1 = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\},$$

$$\mathcal{S}_2 = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}\},$$

are both abstract simplicial complexes.

The element A of $\mathcal{S}$ is called a **simplex** of $\mathcal{S}$; its dimension is one less than its order. Each non-empty subset of A is called a **face of** A.

The **dimension of** $\mathcal{S}$ is the largest dimension of one of its simplices (or infinite if there is no such largest dimension). The **vertex set** V **of** $\mathcal{S}$ is the union of the one-point elements of $\mathcal{S}$; and we make no distinction between the vertex $v \in V$ and the 0-simplex $\{v\} \in \mathcal{S}$.

Example 2.6. Recall $\mathcal{S}_1$ and $\mathcal{S}_2$ from Example 2.4. These collections both have the same vertex set, namely $\{a, b, c\}$.

A subcollection of $\mathcal{S}$ that is itself an abstract simplicial complex is called a **subcomplex of** $\mathcal{S}$. Two abstract complexes $\mathcal{S}$ and $\mathcal{T}$ are **isomorphic** if there is a bijective correspondence f mapping the vertex set of $\mathcal{S}$ to the vertex set of $\mathcal{T}$ such that $\left\{a_0, \ldots, a_n\right\} \in \mathcal{S}$ if and only if $\left\{f(a_0), \ldots, f(a_n)\right\} \in \mathcal{T}$.

For a geometric simplicial complex K and its vertex set V, let $\mathcal{K}$ be the collection of all subsets $\{a_0, \ldots, a_n\}$ of V such that the vertices $a_0, \ldots, a_n$ span a simplex of K. The collection

$\mathcal{K}$ is called the **vertex scheme** of K. It is an important example of an abstract simplicial complex, as it lends itself nicely to the proof of the Geometric Realization Theorem.

The Geometric Realization Theorem may be considered to be the "Rosetta Stone" of abstract and geometric simplicial complexes, allowing us to translate between them.

Theorem 2.15. GEOMETRIC REALIZATION THEOREM.

(1) Every abstract simplicial complex $\mathcal{S}$ is isomorphic to the vertex scheme of some geometric simplicial complex K.

(2) Two simplicial complexes are (linearly) isomorphic if and only if their vertex schemes are isomorphic as abstract simplicial complexes.

Proof. (1) Let Δ^N be the collection of all simplices in $\mathbb{R}^N$ spanned by finite subsets of the standard basis $\left\{e_1, \ldots, e_N\right\}$ for $\mathbb{R}^N$ as a vector space, which is equivalent to the definition used in the proof of Corollary 2.13.

Then Δ^N is a simplicial complex; if σ and τ are two simplices of Δ^N, then their combined vertex set is geometrically independent and spans a simplex of Δ^N.

Now let $\mathcal{S}$ be an abstract complex with vertex set V. Choose a sufficiently large N such that there is an injective function $f : V \to \left\{e_1, \ldots, e_N\right\}$. For instance, let N be the order of V.

We specify a subcomplex K of Δ^N by the condition that for each abstract simplex $\{a_0, \ldots, a_n\} \in \mathcal{S}$, the geometric simplex spanned by $f(a_0), \ldots, f(a_n)$ is to be in K.

It is immediate that K is a simplicial complex and $\mathcal{S}$ is isomorphic to the vertex scheme of K; f is the required correspondence between their vertex sets.

(2) Follows from Lemma 2.12. ■

Using the Geometric Realization Theorem, we are able to define the **geometric realization** of an abstract simplicial complex $\mathcal{S}$ as the vertex scheme of the geometric simplicial complex K with which it is isomorphic. It is unique up to a linear isomorphism.

There is an alternative method of constructing a geometric realization of an abstract simplicial complex: a *labeling of vertices*.

Given a finite complex L, a **labeling of the vertices of** L is a surjective function f mapping the vertex set of L to the **set of labels**.

There is an abstract simplex $\mathcal{S}$ corresponding to this labeling whose vertices are the labels and whose simplices consist of all sets of the form $\{f(v_0), \ldots, f(v_m)\}$, where $v_0, \ldots, v_n$ span a simplex of L.

Let K be a geometric realization of $\mathcal{S}$. Then the vertex map of $L^{(0)}$ onto $K^{(0)}$ derived from f extends to a surjective simplicial map $g : |L| \to |K|$. Then K is the **complex derived from the labeled complex** L and g is the **associated pasting map**.

Example 2.7. Suppose we wish to construct a geometric simplicial complex K whose polytope is homeomorphic to a cylinder. One way to do so is to draw K as a collection of 6 simplices as shown below:

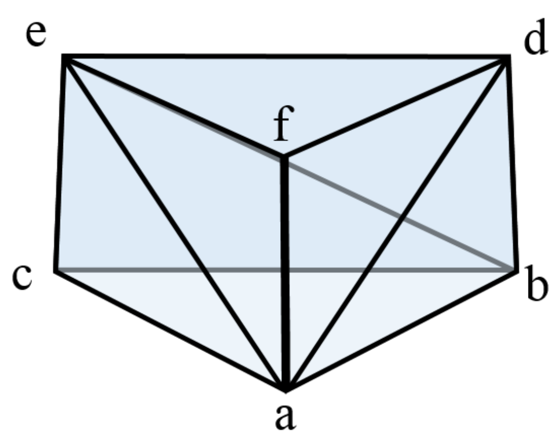


Figure 2.7: *A simplicial complex with the polytope of a cylinder. Note, the interiors of triangles abc and def are not included.*

We may instead draw the diagram shown below:

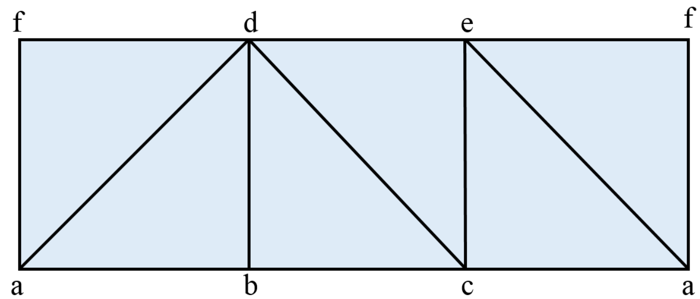


Figure 2.8: *A simplicial diagram with a labeling of vertices that, when pasted, give a simplicial complex with a polytope of a cylinder.*

This diagram consists of two things:

(i) A complex L whose underlying space is a rectangle.

(ii) A particular labeling of vertices of L.

We consider it as short-hand for an abstract simplicial complex $\mathcal{S}$ whose vertex set consists of the letters $a,\ b,\ c,\ d,\ e$, and f and whose simplices are the sets

$$\{a,\ f,\ d\},\ \{a,\ b,\ d\},\ \{b,\ c,\ d\},\ \{c,\ d,\ e\},\ \{a,\ c,\ e\},\ \{a,\ e,\ f\},$$

along with their non-empty subsets.

The abstract simplicial complex L is isomorphic to the vertex scheme of the geometric simplicial complex K, so it specifies the same simplicial complex up to isomorphism, i.e., K is a geometric realization of L.

Example 2.8. Recall the abstract simplicial complexes from Example 2.5:

$$\mathcal{S}_1 = \{\{a\},\ \{b\},\ \{c\},\ \{a,\ b\},\ \{a,\ c\},\ \{b,\ c\},\ \{a,\ b,\ c\}\},$$

$$\mathcal{S}_2 = \{\{a\},\ \{b\},\ \{c\},\ \{a,\ b\},\ \{a,\ c\}\}.$$

We may construct similar diagrams to the previous example to construct their geometric realizations:

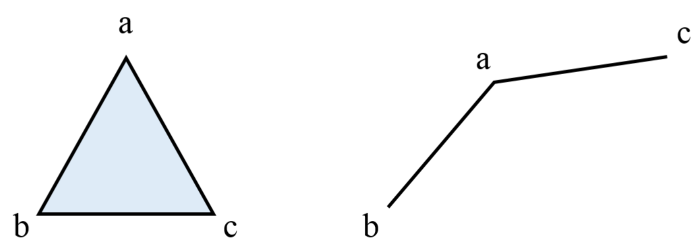


Figure 2.9: *The geometric realizations of $\mathcal{S}_1$ (left) and $\mathcal{S}_2$ (right).*

These diagrams also serve as the geometric realizations of $\mathcal{S}_1$ and $\mathcal{S}_2$, as each of their vertices has a unique label.

Due to the notion of a geometric realization of an abstract simplicial complex, any geometric simplicial complex may be considered as an abstract simplicial complex and vice versa. Henceforth, we distinguish between the two only when necessary, and simply refer to them as simplicial complexes.

Section 2.2: Čech and Vietoris-Rips Complexes

Čech and *Vietoris-Rips complexes* are simplicial complexes constructed from point clouds in Euclidean N-space commonly used in topological data analysis. We begin our discussion with two important results, and then specifically discuss these two types of complexes.

Helly's Theorem

Helly's Theorem describes the intersection patterns of finite collections of convex sets in $\mathbb{R}^N$. Our particular interest in this theorem will become more clear in our future discussion of *nerves*. Consider the following examples.

Example 2.9. Consider three closed intervals in $\mathbb{R}$ that have non-empty pairwise intersections.

For instance, let $A = [1, 5]$, $B = [3, 7]$, and $C = [0, 3]$. Note that $A \cap B = [3, 5]$, $A \cap C = [1, 3]$, and $B \cap C = \{3\}$. Also notice that $A \cap B \cap C = \{3\}$.

This non-empty common intersection will hold for any number greater than 2 of closed intervals with non-empty pairwise intersection in $\mathbb{R}$. To see why this is true, consider the endpoints bounding each closed interval and their presence in the various combinations of each intersection. A proof is included in [4].

Helly's Theorem generalizes this result to higher dimensions.

Theorem 2.16. HELLY'S THEOREM. Let F be a finite collection of closed, convex sets in $\mathbb{R}^N$. Every $N+1$ of the sets in F have a non-empty common intersection if and only if all sets in F have a non-empty common intersection.

Proof. We prove only the forward implication and proceed by induction over the dimension N and the number of sets n in F.

As stated in Example 2.9 and shown in [4], given $N = 1$ and for all n closed, convex sets, the statement holds. Specifically, the statement holds for $n = N + 1$.

Now suppose we have a minimal counterexample consisting of $n > N + 1$ closed, convex sets in $\mathbb{R}^N$. Specifically, this counterexample is minimal in n for a given N. We denote these n sets by $X_1,\ X_2,\ \ldots,\ X_n$. As $X_1,\ X_2,\ \ldots,\ X_n$ is a minimal counterexample, the set $Y_n = \bigcap_{i=1}^{n-1} X_i$ is non-empty and disjoint from X_n.

Specifically, by assumption all n of the sets do not have a non-empty common intersection as $X_1,\ X_2,\ \ldots,\ X_n$ are a counterexample to the statement. However, as they are the minimal counterexample, the sets $X_1,\ X_2,\ \ldots,\ X_{n-1}$ do indeed have a non-empty common intersection, which we have denoted Y_n. Thus, as all n of the sets do not have a non-empty common intersection, Y_n and X_n must be disjoint.

As Y_n and X_n are also closed and convex, we may find an $(N-1)$-dimensional plane h that separates and is disjoint from both sets. Let F' be the collection of sets $Z_i = X_i \cap h$ for $1 \le i \le n-1$. Each Z_i is a non-empty, closed, convex set in $\mathbb{R}^{N-1}$ as any N of the first $n-1$ sets X_i have a common intersection with each other and with X_n. This also implies that any N of the sets in F' have a non-empty common intersection.

By minimality of the counterexample, the statement implies that $\bigcap F'$ is non-empty, i.e.,

$$\bigcap F' = \bigcap_{i=1}^{n-1} \left(X_i \cap h\right) = Y_n \cap h.$$

Thus, Y_n and h are not disjoint, a contradiction. ∎

Example 2.10. Consider the sets $A,\ B,\ C$, and D shown in Figure 2.10.

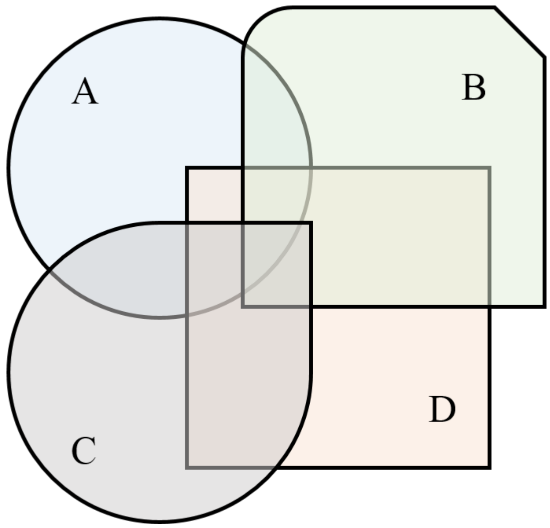


Figure 2.10: *Every three of four sets $A,\ B,\ C$, and D in $\mathbb{R}^2$ have non-empty common intersection. All four indeed have a non-empty common intersection, as implied by Helly's Theorem.*

Every three of four sets $A,\ B,\ C$, and D in Figure 2.10 have non-empty common intersection. All four indeed have a non-empty common intersection, as implied by Helly's Theorem.

Homotopy Type

Homotopy type is a notion of equivalence between topological spaces that is weaker than topological equivalence. It is useful when we desire to "simplify" a topological space in a rigorous way.

Definition 2.17. Let $X,\ Y$ be topological spaces and $f,\ g :\ X \to Y$ be continuous maps between them. A **homotopy** between f and g is a continuous map $H :\ X \times [0,\ 1] \to Y$ such that $H\,(x,\ 0) = f(x)$ and $H\,(x,\ 1) = g(x)$ for all $x \in X$.

The homotopy H may be thought of as a time-series of functions $f_t :\ X \to Y$ such that $f_t(x) = H(x,\ t)$, with $f_0 = f$ and $f_1 = g$.

Example 2.11. Consider the functions $f,\ g :\ \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$ and $g(x) = 3\,x^3$, which are both continuous functions.

The function $H : \mathbb{R} \times [0,\ 1] \to \mathbb{R}$ by $H(x,\ t) = (1 + 2\,t)\ x^{2+t}$ is a homotopy between f and g as $H(x,\ 0) = x^2 = f(x)$ and $H(x,\ 1) = 3\,x^3 = g(x)$.

Homotopies define an equivalence relation on functions, and we write $f \simeq g$ if there is a homotopy between them.

Definition 2.18. Two topological spaces X and Y are **homotopy equivalent**, or have the same **homotopy type**, if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$, and we write $X \simeq Y$, which is indeed an equivalence relation. The functions f and g are referred to as **homotopy equivalences** or **homotopy inverses** of one another.

The condition $Y \subseteq X$ is a specific case of homotopy equivalence. In this case, Y is a **retract** of X if there exists a continuous map $r : X \to Y$ such that $r(y) = y$ for all $y \in Y$; r is called a **retraction**. If there is a homotopy between r and id_X, then Y is a **deformation retract** and r is a **deformation retraction**. If Y is a single point, then X has the homotopy type of a point and is said to be **contractible**.

Example 2.12. Any closed path $\mathcal{P}$ in Euclidean space equipped with the subspace topology is contractible by the following argument:

Note that $\mathcal{P}$ is homeomorphic to the closed interval $I = [0,\ 1]$ in $\mathbb{R}$ with the usual topology. Consider the closed set $P = \{0\}$ in $\mathbb{R}$ and the zero function $f : I \to P$.

The function $H : I \times [0,\ 1] \to P$ by $H(x,\ t) = t\,x$ is a homotopy between f and the identity function id_I, as it is a continuous function with the property $H(x,\ 0) = 0 = f(x)$ and $H(x,\ 1) = x = \mathrm{id}_I(x)$. Thus, I is contractible.

As I is contactible and is homeomorphic to $\mathcal{P}$, $\mathcal{P}$ is contractible as well.

Example 2.13. Consider the unit disc $\overline{\mathrm{B}^2}\,(\mathcal{O},\ 1)$. Using a similar homotopy as in Example 2.12, we find that it, too, is contractible.

Nerves and the Nerve Theorem

We are now ready to define *nerves*, which will lead us to the Nerve Theorem and complete the groundwork for our discussion on Čech and Vietoris-Rips complexes.

Definition 2.19. Let F be a finite collection of sets in $\mathbb{R}^N$ (which may or may not be convex). The **nerve** of F consists of all non-empty subcollections X of F whose sets have a non-empty common intersection, i.e.,

$$\mathrm{Nrv}\,(F) = \left\{ X \subseteq F \;\middle|\; \bigcap X \neq \varnothing \right\}.$$

Example 2.14. Recall the sets A, B, C, and D shown in Figure 2.10. The nerve of $F = \{A,\ B,\ C,\ D\}$ is

$$\begin{aligned}\mathrm{Nrv}\,(F) = \{&\{A\},\ \{B\},\ \{C\},\ \{D\},\ \{A,\ B\},\ \{A,\ C\},\ \{A,\ D\},\\ &\{B,\ C\},\ \{B,\ D\},\ \{C,\ D\},\ \{A,\ B,\ C\},\ \{A,\ B,\ D\},\\ &\{A,\ C,\ D\},\ \{B,\ C,\ D\},\ \{A,\ B,\ C,\ D\}\}.\end{aligned}$$

Note that all non-empty subsets of the elements of $\mathrm{Nrv}\,(F)$ are also in $\mathrm{Nrv}\,(F)$, implying that it is an abstract simplicial complex. This in fact holds for all nerves and, by the Geometric Realization Theorem, we may thus represent the nerve as a geometric simplicial complex.

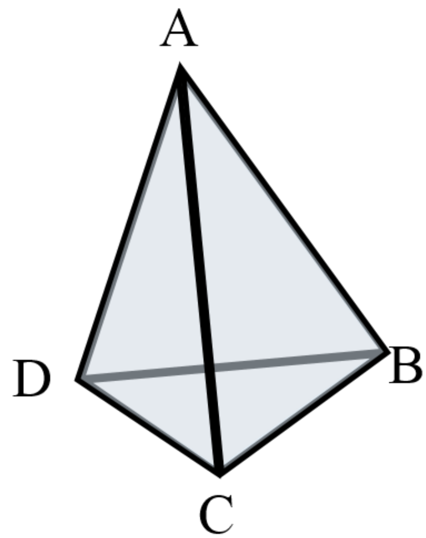


Figure 2.11: *The geometric realization of Nrv* (F) *is a tetrahedron, including its interior.*

As the nerve of a collection of sets may be represented as a geometric simplicial complex, it is sensible to discuss its topology and, more importantly for our purposes, its homotopy type. The Nerve Theorem greatly simplifies our understanding of the homotopy type of a given collection by equating it to that of the collection itself.

Theorem 2.20. NERVE THEOREM. Let F be a finite collection of closed, convex sets in Euclidean space. Then the nerve of F and the union of the sets in F have the same homotopy type.

A proof of the Nerve Theorem is explicitly omitted in [4], and although it may be found in [8], it should be left to only the most interested of readers.

Note that Helly's Theorem imposes a constraint on the structure of the nerve: if the sets live in $\mathbb{R}^N$, then a subcollection of $k \geq N+1$ sets cannot have all $\begin{pmatrix} k \\ N+1 \end{pmatrix}$ N-simplices in the nerve without having the entire k-simplex in the nerve.

Constructing Čech and Vietoris-Rips Complexes

We have now laid all necessary groundwork for our discussion of *Čech* and *Vietoris-Rips complexes*. We describe their construction and prove the Čech and Vietoris-Rips Lemma, which describes the relationship between them.

Definition 2.21. Let $r > 0$ be fixed, and consider a finite set S of points in $\mathbb{R}^N$. Let $\overline{\mathrm{B}^N}(x,\ r)$ denote the closed ball of radius r about the point x. The **Čech complex** of S and r is isomorphic to the nerve of this collection of balls, i.e.,

$$\check{\mathrm{C}}\mathrm{ech}\,(r) = \left\{ \sigma \subseteq S \;\middle|\; \bigcap_{x\in\sigma} \overline{\mathrm{B}^N}(x,\ r) \neq \varnothing \right\}.$$

Example 2.15. Below are figures depicting the construction of the geometric realization of a Čech complex for a fixed set of points S and an increasing radius.

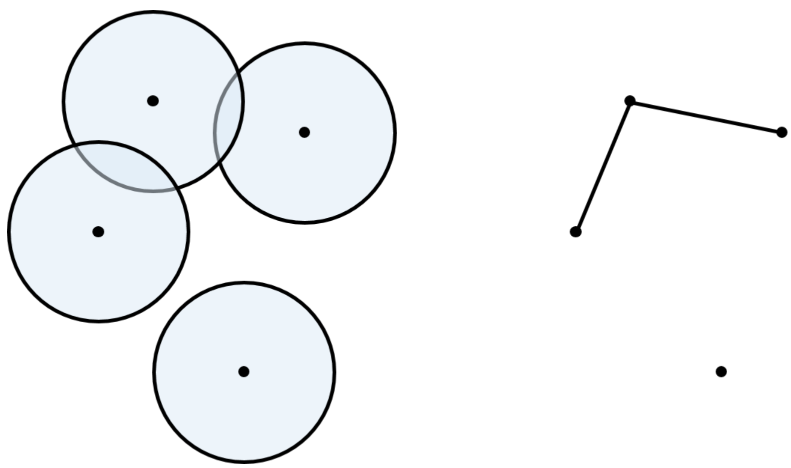


Figure 2.12: *(Left) A set of points contained in circles of a radius r. (Right) The geometric realization of the Čech complex of the set.*

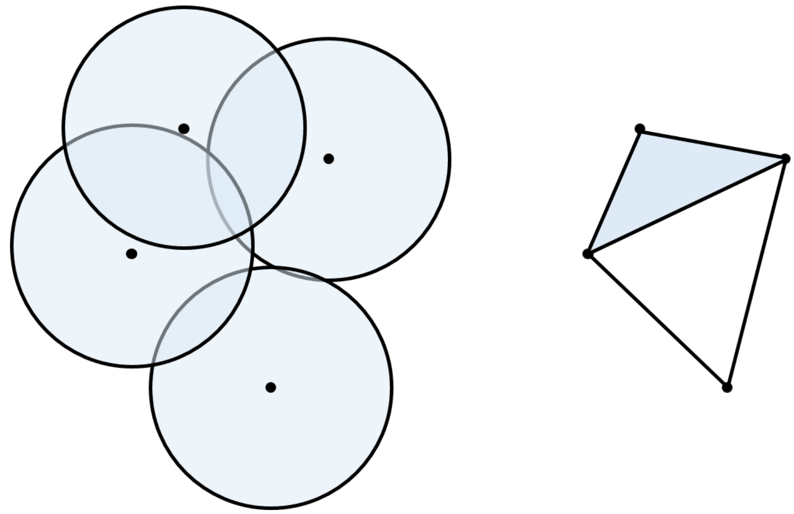


Figure 2.13: *Using a slightly larger radius, we obtain a different Čech complex for the same points. The* 2*-simplex indicates that the upper-most three points are closer to one another than to the bottom-most point.*

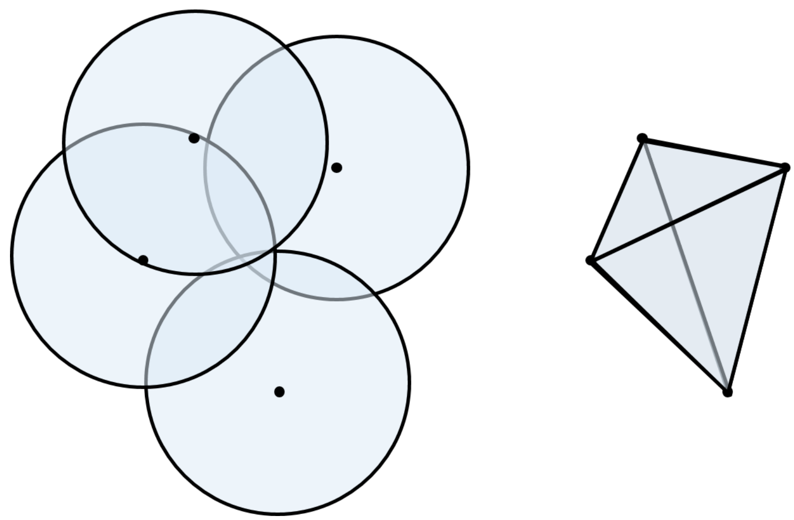


Figure 2.14: *Using an even larger radius, we obtain a tetrahedron, including its interior. In essence, we have erased the information about the relative positions of the points.*

Notice that as we increase the radius of each ball, the geometric realization of the Čech complex may change drastically. At the middle stage, depicted in Figure 2.13, we obtain a 2-simplex and a hole bounded by 1-simplices. By the last stage, depicted in Figure 2.14, we have closed the hole and filled in the 3-simplex that the four vertices span. However note that for every $r_0 \leq r$, $\check{\text{C}}\text{ech}\left(r_0\right) \subseteq \check{\text{C}}\text{ech}\,(r)$.

From this, we may also see that for a given set of points in $\mathbb{R}^N$, the Čech complex may not have a geometric realization in $\mathbb{R}^N$. Specifically, our set of points S is contained in $\mathbb{R}^2$, but the geometric realization of the final Čech complex is a tetrahedron, which lives in $\mathbb{R}^3$.

Consider a subset σ of points in S. Let the **miniball** of σ, which we denote $\mathrm{mb}(\sigma)$, be the smallest closed ball that contains σ, which is unique. Note that the diameter of the set containing the upper-most three points is less than twice the radius of the balls in both Figures 2.13 and 2.14. This indeed holds for all subsets of points in all Čech complexes, i.e., the radius of $\mathrm{mb}(\sigma)$ is less than or equal to r if and only if $\sigma \in \check{\mathrm{C}}\mathrm{ech}(r)$.

Definition 2.22. Similarly to the Čech complex, we utilize a set S of points in $\mathbb{R}^N$ and a fixed $r > 0$ to construct the **Vietoris-Rips complex**. The Vietoris-Rips complex consists of all subsets of diameter at most $2r$, i.e.,

$$\text{Vietoris-Rips}(r) = \left\{ \sigma \subset S \;\middle|\; \mathrm{diam}(\sigma) \leq 2r \right\}.$$

This implies that whenever a boundary of a simplex is in the Vietoris-Rips complex, then its interior is as well. For example, consider the case of a 2-simplex spanned by $\{a, b, c\}$ whose boundary is included in the complex for a given r. This means that the distance between any two of a, b, c is less than $2r$, and thus the diameter of $\{a, b, c\}$ is less than $2r$.

Example 2.16. Below is a figure depicting the construction of the geometric realization of a Vietoris-Rips complex for the same set of points S as in Example 2.14.

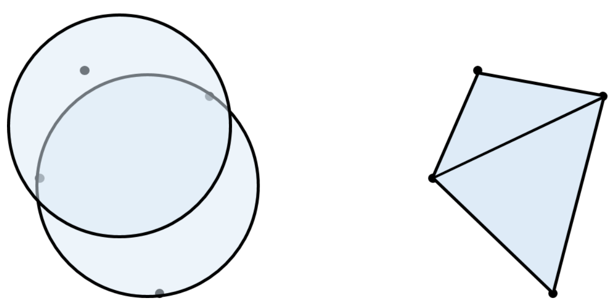


Figure 2.15: *Unlike the Čech complex featured in Figure 2.13, the Vietoris-Rips complex of this radius contains two* 2*-simplices.*

From this example, we see that the edges and vertices are identical in both the Čech and Vietoris-Rips complexes of a given set and fixed radius, and further $\check{\mathrm{C}}\mathrm{ech}(r) \subseteq \text{Vietoris-Rips}(r)$ for any given set of points and any fixed radius r.

The Čech and Vietoris-Rips Lemma relates Čech and Vietoris-Rips complexes, and provides credence to our preference for the latter for computational purposes.

Lemma 2.23. ČECH AND VIETORIS-RIPS LEMMA. Let S be a finite set of points in some Euclidean space and $r \geq 0$. Then Vietoris-Rips $(r) \subseteq$ Čech $\left(\sqrt{2}\, r\right)$.

Proof. A simplex is **regular** if all its edges have the same length. A representation for dimension N is the **standard N-simplex**, denoted ∇^N (not to be confused with the aforementioned Δ^N), which is the simplex spanned by the endpoints of the unit vectors in $\mathbb{R}^{N+1}$. Note that each edge of ∇^N has length $\sqrt{2}$. We will utilize the fact that we may always express a simplicial complex as a subcomplex of ∇^N for a sufficiently large N (see [4]), along with the interplay between the length of the edges of ∇^N and the definition of the Čech and Vietoris-Rips complexes to prove the result.

The **barycenter** of a simplex is the arithmetic mean position of its vertices. By symmetry, the barycenter of ∇^N is the point z whose $N + 1$ coordinates are all $\frac{1}{N+1}$. Thus, $\|z\| = 1/\sqrt{N+1}$.

The barycenter is also the center of the smallest N-sphere that passes through the vertices of ∇^N. Let r_N be the radius of that sphere, and note that $r_N = \sqrt{1 - \|z\|^2} = \sqrt{\frac{N}{N+1}}$. As the dimension goes to infinity, r_N approaches 1 from below.

Note that the smallest diameter of such an N-sphere occurs when $N = 1$, as $2\, r_1 = 2\sqrt{\frac{1}{2}} = \sqrt{2}$ and r_N increases with N. Thus, any set of $N + 1$ or fewer vertices for which the same N-ball of diameter $2\, r_N$ is the miniball has a pair of points at distance $\sqrt{2}$ or larger.

It follows that every simplex of diameter $\sqrt{2}$ or less belongs to Čech $\left(r_N\right)$. Hence, Vietoris-Rips $(r) \subseteq$ Čech $\left(\sqrt{2}\, r\, r_N\right)$, and we also have Čech $\left(\sqrt{2}\, r\, r_N\right) \subseteq$ Čech $\left(\sqrt{2}\, r\right)$ as $r_N \leq 1$ for all N.

Since we may always express a simplicial complex as a subcomplex of ∇^N for a sufficiently large N, the result holds. ■

By definition, the Vietoris-Rips complex of a set of vertices is easier to compute than a Čech complex for the same set. The Čech and Vietoris-Rips Lemma shows that the Vietoris-

Rips complex contains a portion of the information contained in the Čech complex, and that amount of information is sufficient for our purposes.

Chapter 3: Homology Groups of a Simplicial Complex

The driving force behind the development of *homology* was the observation that two topological spaces can be distinguished by examining their holes. For instance, the disk and the annulus differ by a single hole in their center.

In Section 3.1, we introduce *homology groups* in non-negative dimensions. These groups provide information about the space by counting the number of holes indirectly through what surrounds them. We discuss a modification to homology groups, called *reduced homology groups*, which are identical except in the 0^{th} dimension.

In Section 3.2, we utilize the diagrams introduced in Example 2.7 to calculate the homology groups of a torus and a Klein bottle.

Section 3.1: Homology Groups

The construction of homology groups necessitates the ability to determine the presence or lack of a "hole". To accomplish this, we introduce the idea of a *boundary operator*, which in turn requires the notion of *orientation* in the context of simplices.

Homology in Dimensions Greater Than 0

Let σ be a simplex. Define two orderings of its vertex set to be equivalent if they differ from one another by an even permutation. If $\dim(\sigma) > 0$, the orderings of the vertices of σ then fall into two equivalence classes. Each of these classes is called an **orientation** of σ. An **oriented simplex** is a simplex σ together with an orientation of σ.

If $v_0, \ldots, v_p$ are geometrically independent, we shall use the symbol

$$v_0 \ldots v_p$$

to denote the simplex they span and

$$\left[v_0, \ldots, v_p\right]$$

to denote the oriented simplex with the particular ordering $\left(v_0, \ldots, v_p\right)$.

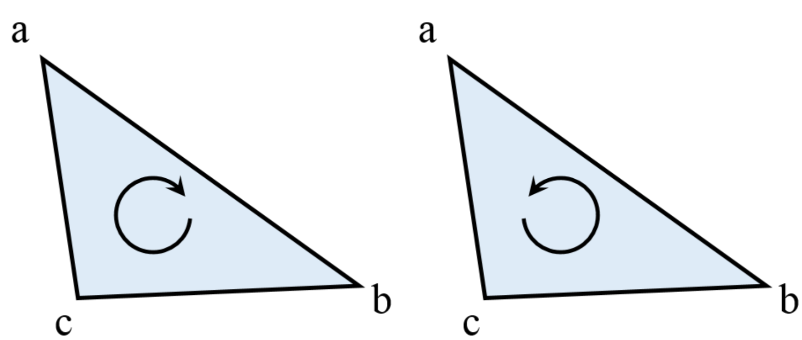


Figure 3.1: *A* 2*-simplex with its two possible orientations,* $[a, b, c]$ *(left) and* $[a, c, b]$ *(right).*

Definition 3.1. Let K be a simplicial complex. A ***p*-chain** on K is a function c from the set of oriented p-simplices of K to the integers, such that:

(i) $c(\sigma) = -c(\sigma')$ if σ and σ' represent opposite orientations of the same simplex, and

(ii) $c(\sigma) = 0$ for all but finitely many oriented p-simplices σ.

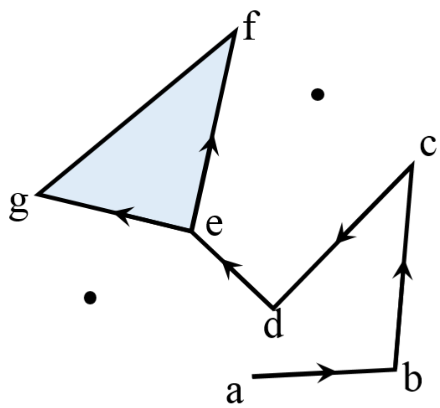


Figure 3.2: *A simplicial complex with the* 1*-chain* $[a,\ b]+[b,\ c]+[c,\ d]+[d,\ e]+[e,\ f]+[e,\ g]$ *indicated by arrows.*

For an oriented p-simplex σ, we also use σ to denote the p-chain that takes σ to 1, σ' to -1, and all other p-simplices to 0, as shown in Figure 3.2.

The p-chains together with p-chain addition form an abelian group called the **group of oriented p-chains of K**, denoted $C_p(K)$.

If $p < 0$ or $p > \dim(K)$, then $C_p(K)$ is the trivial group, as there are no simplices of dimension less than 0 or greater than the dimension of the simplicial complex which contains them.

We say a chain c is **carried by a subcomplex L of K** if c has value 0 on every simplex of K that is not in L.

Lemma 3.2. As might be expected, $C_p(K)$ is free abelian. Orient each p-simplex σ, and consider each as a p-chain consisting only of σ. These p-chains form a basis for $C_p(K)$.

Proof. Once all p-simplices of K are oriented arbitrarily, each p-chain may be written uniquely as a finite linear combination

$$c = \sum n_i\,\sigma_i$$

of the corresponding elementary chains σ_i.

The chain c assigns the value n_i to the oriented p-simplex σ_i, the value $-n_i$ to the opposite orientation of σ_i, and the value 0 to all oriented p-simplices not appearing in the summation. ■

Example 3.1. Consider the simplicial complex in Figure 3.2, which we now denote K, and orient the edge between f and g as $[f,\ g]$. The 1-chains $[a,\ b]$, $[b,\ c]$, $[c,\ d]$, $[d,\ e]$, $[e,\ f]$, $[e,\ g]$, and $[f,\ g]$ form a basis for $C_1(K)$.

The group $C_0(K)$ has a natural basis as any 0-simplex has only one orientation. For $p > 0$, $C_p(K)$ has no natural basis, and the simplices must be oriented arbitrarily in order to obtain a basis.

Corollary 3.3. Any function f from the oriented p-simplices of K to an abelian group G extends uniquely to a homomorphism $C_p(K) \to G$, provided that $f(-\sigma) = -f(\sigma)$ for all oriented p-simplices σ.

A proof of this is included in [9].

Definition 3.4. We now define a homomorphism

$$\partial_p : C_p(K) \to C_{p-1}(K)$$

called the **boundary operator**. For an oriented simplex $\sigma = \left[v_0,\ \dots,\ v_p\right]$ with $p > 0$, we define

$$\partial_p\sigma = \partial_p\left[v_0,\ \dots,\ v_p\right] = \sum_{i=0}^{p}(-1)^i\left[v_0,\ \dots,\ \hat{v}_i,\ \dots,\ v_p\right],$$

where $\hat{v}_i$ means that the vertex v_i is to be deleted from the array. Since $C_p(K)$ is trivial for $p < 0$, the operator ∂_p is the trivial homomorphism for $p \leq 0$.

To ensure ∂_p is well defined and that $\partial_p(-\sigma) = -\partial_p(\sigma)$, it suffices to show the summation on the right-hand side changes sign if we exchange two adjacent vertices in the array $\left[v_0,\ \dots,\ v_p\right]$. We compare the expressions for

$$\partial_p\left[v_0,\ \dots,\ v_j,\ v_{j+1},\ \dots,\ v_p\right] \text{ and } \partial_p\left[v_0,\ \dots,\ v_{j+1},\ v_j,\ \dots,\ v_p\right].$$

For $i \neq j,\ j+1$, the i^{th} terms in these two expressions differ precisely by a sign; the terms are identical except that v_j and v_{j+1} have been interchanged.

For $i = j,\ j+1$, the terms in $\partial_p \left[v_0,\ \dots,\ v_j,\ v_{j+1},\ \dots,\ v_p\right]$ are

$$(-1)^j \left[\dots,\ v_{j-1},\ \hat{v}_j,\ v_{j+1},\ v_{j+2},\ \dots\right] + (-1)^{j+1} \left[\dots,\ v_{j-1},\ v_j,\ \hat{v}_{j+1},\ v_{j+2},\ \dots\right]$$

while the terms in $\partial_p \left[v_0,\ \dots,\ v_{j+1},\ v_j,\ \dots,\ v_p\right]$ are

$$(-1)^j \left[\dots,\ v_{j-1},\ v_j,\ \hat{v}_{j+1},\ v_{j+2},\ \dots\right] + (-1)^{j+1} \left[\dots,\ v_{j-1},\ \hat{v}_j,\ v_{j+1},\ v_{j+2},\ \dots\right]$$

which only differ by a sign.

Example 3.2. Consider the oriented 2-simplex $[a,\ b,\ c]$, as shown in Figure 3.1. Its boundary is

$$\partial_2 [a,\ b,\ c] = [b,\ c] - [a,\ c] + [a,\ b] = [a,\ b] + [b,\ c] + [c,\ a]\,,$$

which is the orientation we would expect its boundary to inherit.

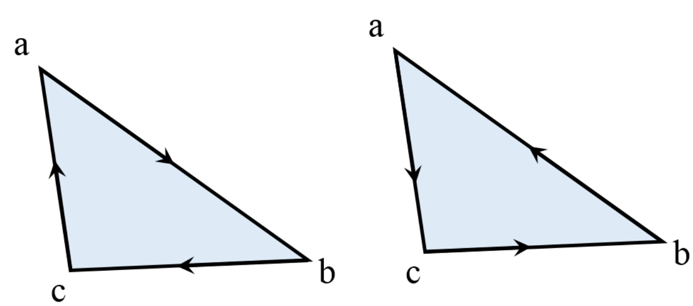


Figure 3.3: *The boundaries of* $[a,\ b,\ c]$ *(left) and* $[a,\ c,\ b]$ *(right) inherit the orientation we would expect from the orientation of the* 2*-simplices.*

One of the most crucial facts about the boundary operator is the fact that $\partial_{p-1} \circ \partial_p = 0$ for all p, which is shown by the following computation:

$$\begin{aligned}\partial_{p-1}\,\partial_p \left[v_0,\ \dots,\ v_p\right] &= \sum_{i=0}^{p} (-1)^i\,\partial_{p-1} \left[v_0,\ \dots,\ \hat{v}_i,\ \dots,\ v_p\right] \\ &= \sum_{j<i} (-1)^i\,(-1)^j \left[\dots,\ \hat{v}_j,\ \dots,\ \hat{v}_i,\ \dots\right] \\ &\quad + \sum_{j>i} (-1)^i\,(-1)^{j-1} \left[\dots,\ \hat{v}_i,\ \dots,\ \hat{v}_j,\ \dots\right].\end{aligned}$$

The terms of these two summations cancel in pairs.

Example 3.3. Recall from Example 3.1 that the boundary of $[a, b, c]$ is $[a, b] + [b, c] + [c, a]$. Applying the boundary operator once more (and excluding the square brackets for the intermediate 0-simplices), we obtain

$$\partial_1 ([a, b] + [b, c] + [c, a]) = b - a + c - b + a - c = 0.$$

Definition 3.5. The kernel of $\partial_p : C_p(K) \to C_{p-1}(K)$ is called the group of ***p*-cycles** and denoted $Z_p(K)$.

Example 3.4. Consider the 1-chain $[a, b]+[b, c]+[c, a]$, which is a connected path beginning and ending at a. From Example 3.2, we know that its boundary is 0, meaning that it is in the kernel of ∂_1. Therefore, it is a 1-cycle.

Definition 3.6. The image of $\partial_{p+1} : C_{p+1}(K) \to C_p(K)$ is called the group of ***p*-boundaries** and is denoted $B_p(K)$.

Example 3.5. Recall that the boundary of $[a, b, c]$ is

$$\partial_2 [a, b, c] = [a, b] + [b, c] + [c, a] .$$

Therefore $[a, b] + [b, c] + [c, a]$ is in the image of ∂_2 and is thus a 1-boundary.

Note that each boundary of a $p + 1$ chain is automatically a p-cycle, i.e., $B_p(K) \subseteq Z_p(K)$.

Definition 3.7. We define the p^{th} **homology group** as

$$H_p(K) = Z_p(K) / B_p(K),$$

which depends only on the polytope $|K|$.

The elements of homology groups are equivalence classes with a representative p-cycle c. All other elements of this equivalence class are of the form $c + \partial d$, where d is a $p + 1$-chain. That is, all elements of an equivalence class differ by a boundary, and we say that any two cycles that differ by a boundary are **homologous**.

Example 3.6. Recall the simplicial complex K from Examples 2.12 and 2.15, shown below:

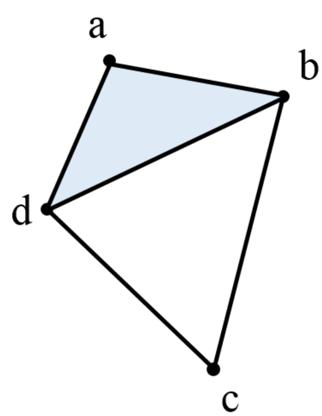


Figure 3.4: *A simplicial complex with three* 1*-cycles:* $[a,\ b]+[b,\ d]+[d,\ a]$, $[b,\ c]+[c,\ d]+[d,\ b]$*, and* $[a,\ b]+[b,\ c]+[c,\ d]+[d,\ a]$*. The last two of these differ by the boundary of* $[a,\ b,\ d]$*, and are thus homologous.*

The simplicial complex K contains three 1-cycles:

$$x = [a,\ b] + [b,\ d] + [d,\ a]\,,$$

$$y = [b,\ c] + [c,\ d] + [d,\ b]\,, \text{ and}$$

$$z = [a,\ b] + [b,\ c] + [c,\ d] + [d,\ a]\,.$$

Note that $x+y = z$, and thus every 1-cycle is of the form $m\,x+n\,y$. Therefore, $Z_1\,(K) \cong \mathbb{Z}\oplus\mathbb{Z}$.

The simplicial complex K also contains only one 1-boundary, namely x. Therefore, $B_1\,(K) \cong \mathbb{Z}$.

Thus, $H_1\,(K) \cong Z_1\,(K)\,/\,B_1\,(K) \cong \mathbb{Z}$. Note that the rank of $H_1\,(K)$ is the same as the number of 1-dimensional "holes" in K, namely the hole bounded by y.

Example 3.5 illustrates the information we may collect from the homology groups of a simplicial complex: the number of holes in each dimension. For clarity, we say that a hole is dimension p when it is bounded by a p-chain. To ease our discussion about these holes, we define the *Betti numbers* of a space below.

Definition 3.8. Let K be a simplicial complex with homology groups $H_p(K)$. The p^{th} **Betti number** β_p is the rank of the p^{th} homology group, and is the total number of p-dimensional holes.

Zero-Dimensional and Reduced Homology

Continuing our observation, the rank of the zeroth-dimensional homology group should be a count of the number of 0-dimensional holes. Upon inspection, however, it counts the number of connected components!

To amend this, we introduce the *augmentation map*, which is a homomorphism used in place of ∂_0 to define the *reduced homology groups*. The reduced homology groups are identical to the aforementioned homology groups in all non-negative dimensions, with the exception of the zeroth dimension where their rank is one less than the number of connected components.

We begin our discussion with a brief foray into the zeroth homology group.

Proposition 3.9. Let K be a complex. Then the group $H_0(K)$ is free abelian. If $\{v_\alpha\}$ is a collection consisting of one vertex from each component of $|K|$, then the homology classes of the chains v_α form a basis for $H_0(K)$.

The proof of this proposition is very detailed, contains many moving parts, and fails to lend valuable insight for our future discussions. The reader may simply accept the results of Proposition 3.9 and proceed.

Proof. This proof was obtained from in [9].

Step 1. If v and w are vertices of K, let us define $v \sim w$ if there is a sequence

$$a_0, \ \ldots, \ a_n$$

of vertices of K such that $v = a_0$ and $w = a_n$, and $a_i\, a_{i+1}$ is a 1-simplex of K for each i. This relation is clearly an equivalence relation. Given v, define

$$C_v = \bigcup \{\mathrm{St}\,(w) \mid w \sim v\}.$$

We show that the sets C_v are components of $|K|$.

Note first that C_v is open because it is the union of open sets. Furthermore, $C_v = C_{v'}$ if $v \sim v'$.

Second, we show that C_v is connected, in fact, path connected. Given v, let $w \sim v$ and let x be a point of $\mathrm{St}\,(w)$. Choose a sequence $a_0, \ldots, a_n$ of vertices of K, as before. Then the broken line path with successive vertices $a_0, \ldots, a_n, x$ lies in C_v. Since $a_i \sim v$ by definition, so that $\mathrm{St}\left(a_i\right) \subset C_v$, and in particular the line segment $a_i\, a_{i+1}$ lies in C_v. Similarly, the line segment $a_n\, x$ lies in $\mathrm{St}\left(a_n\right)$, which is contained in C_v. Hence, C_v is path connected.

Third, we show distinct sets C_v and $C_{v'}$ are disjoint. Suppose x is a point of their intersection. Then $x \in \mathrm{St}\,(w)$ for some w equivalent to v, and $x \in \mathrm{St}\,(w')$ for some w' equivalent to v'. Since x has positive barycentric coordinates with respect to both w and w', some simplex of K has w and w' as vertices. Then $w\,w'$ must be a 1-simplex of K, so $w \sim w'$. It follows that $v \sim v'$, so that the two sets C_v and $C_{v'}$ are the same.

Being connected, open, and disjoint, the sets C_v are necessarily the components of $|K|$. Note that each is the space of a subcomplex of K; each simplex of K, being connected, lies entirely in one component of $|K|$.

Step 2. Now we prove the proposition. Let $\{v_\alpha\}$ be a collection of vertices containing one vertex v_α from each component C_α of $|K|$. Given a vertex w of K, it belongs to some component of K, say C_α. By hypothesis, $w \sim v_\alpha$, so there is a sequence $a_0, \ldots, a_n$ of vertices of K, as before, leading from v_α to w. The 1-chain

$$\left[a_0,\, a_1\right] + \left[a_1,\, a_2\right] + \cdots + \left[a_{n-1},\, a_n\right]$$

has as its boundary the 0-chain $a_n - a_0 = w - v_\alpha$. Thus, the 0-chain w is homologous to the 0-chain v_α. We conclude that every chain in K is homologous to a linear combination of the elementary 0-chains v_α.

We now show that no non-trivial chain of the form $c = \sum n_\alpha v_\alpha$ bounds. Suppose $c = \partial d$ for some 1-chain d. Since each 1-simplex of K lies in a unique component of $|K|$, we can write $d = \sum d_\alpha$, where d_α consists of those terms of d that are carried by C_α. Since $\partial d = \sum \partial d_\alpha$ and ∂d_α is carried by C_α, we conclude that $\partial d_\alpha = n_\alpha v_\alpha$. It follows that $n_\alpha = 0$ for each α. Let $\epsilon : C_0(K) \to \mathbb{Z}$ be the homomorphism defined by setting $\epsilon(v) = 1$ for each vertex v of K. Then $\epsilon(\partial[v, w]) = \epsilon(w - v) = 1 - 1 = 0$ for any elementary 1-chain $[v, w]$. As a result, $\epsilon(\partial d) = 0$ for every 1-chain d. In particular, $0 = \epsilon(\partial d) = \epsilon\left(n_\alpha v_\alpha\right) = n_\alpha$. ■

Step 1 creates a collection of subsets of K that are the components of K. Specifically, $v \sim v'$ if there is a path of 1-simplices entirely in K joining them and each unique C_v is an open connected component of K and all C_v are disjoint from each other. A drawing makes this especially apparent.

Step 2 establishes that any two vertices in a C_v are homologous, every chain in K can be expressed as a linear combination of 0-chains determined by a set consisting of one vertex from each C_α, and that no non-trivial 0-chain bounds, which is analogous to showing that the linear combinations are unique.

Recall that from our observations, the 0$^{\text{th}}$ Betti number would count the number of 0-dimensional holes, which would be one less than the number of connected components. We now introduce the *augmentation map* and define *reduced holomology groups*, which allow us to count the number of holes in all dimensions $p > 0$ and count the number of connected components in dimension $p = 0$.

Definition 3.10. Let $\epsilon : C_0(K) \to \mathbb{Z}$ be the surjective homomorphism defined by $\epsilon(v) = 1$ for each vertex v of K. Then if c is a 0-chain, $\epsilon(c)$ equals the sum of the values of c on the vertices of K. The map ϵ is called the **augmentation map** for $C_0(K)$. We have just noted that $\epsilon(\partial d) = 0$ if d is a 1-chain. We define the **reduced homology group** of K in dimension 0, denoted $\tilde{H}_0(K)$, by the equation

$$\tilde{H}_0(K) = \ker(\epsilon)/\mathrm{Im}\left(\partial_1\right).$$

If $p > 0$, we let $\tilde{H}_p(K)$ denote the usual group $H_p(K)$.

We similarly define the **reduced Betti numbers** $\tilde{\beta}_p$ as the rank of the reduced homology groups.

The following theorem confirms that the zeroth reduced homology group does indeed count the number of connected components, which is one greater than the number of zero-dimensional holes.

Theorem 3.11. The group $\tilde{H}_0(K)$ is free abelian and

$$\tilde{H}_0(K) \oplus \mathbb{Z} \cong H_0(K).$$

Thus, $\tilde{H}_0(K)$ vanishes if $|K|$ is connected. If $|K|$ is not connected, let $\{v_\alpha\}$ consist of one vertex from each component of $|K|$; let α_0 be a fixed index. Then the homology classes of the chains $v_\alpha - v_{\alpha_0}$, for $\alpha \neq \alpha_0$, form a basis for $\tilde{H}_0(K)$.

Proof. Given a 0-chain c, it is homologous to a 0-chain of the form $c' = \sum n_\alpha v_\alpha$, where the chain c' bounds if and only if $n_\alpha = 0$ for all α. Now if $c \in \ker(\epsilon)$, then $\epsilon(c) = \epsilon(c') = \epsilon(\sum n_\alpha v_\alpha) = \sum n_\alpha = 0$. If $|K|$ has only one component, this implies that $c' = 0$. If $|K|$ has more than one component, it implies that c' is a linear combination of the 0-chains $v_\alpha - v_{\alpha_0}$. ■

Section 3.2: Calculating Homology Groups of Surfaces Using Labeled Simplicial Complexes

We now utilize labeled diagrams, akin to those introduced in Example 2.7, to compute the homology groups of well-known surfaces.

Preliminaries

For each of these surfaces, we will be using alternate labelings of the same diagram, namely the simplicial complex L shown below:

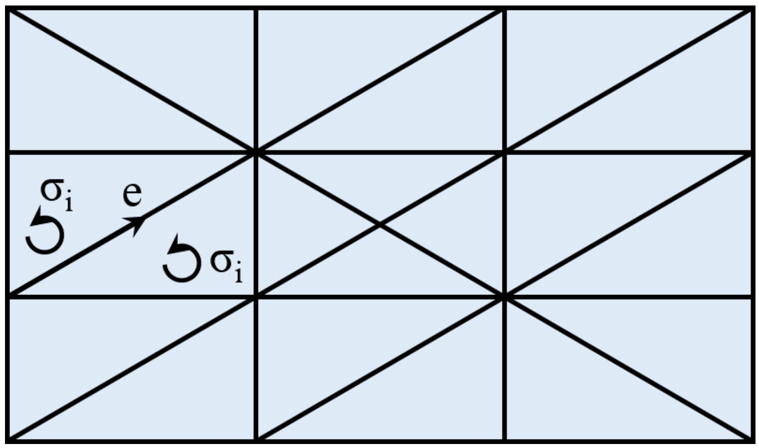


Figure 3.5: *The diagram L which we will be using throughout this section. Each of its* 2-*simplices are oriented counterclockwise, while its* 1-*simplices are oriented arbitrarily.*

Let each of the 2-simplices in L be oriented counterclockwise and the 1-simplices be oriented arbitrarily. We denote the complex whose polytope is the boundary of the rectangle by $\mathrm{Bd}(L)$.

Lemma 3.12. (1) Every 1-cycle of L is homologous to a 1-cycle carried by $\mathrm{Bd}(L)$.

(2) If d is a 2-chain of L and if ∂d is carried by $\mathrm{Bd}(L)$, then d is a multiple of the chain $\sum \sigma_i$.

Proof. We give a partial proof of (1) along with the final result. The technique used is sometimes called "pushing a p-chain off of p-simplices", and involves arguing that a given p-chain is homologous to a p-chain carried by some subcomplex, and proceeding until the desired subcomplex is reached.

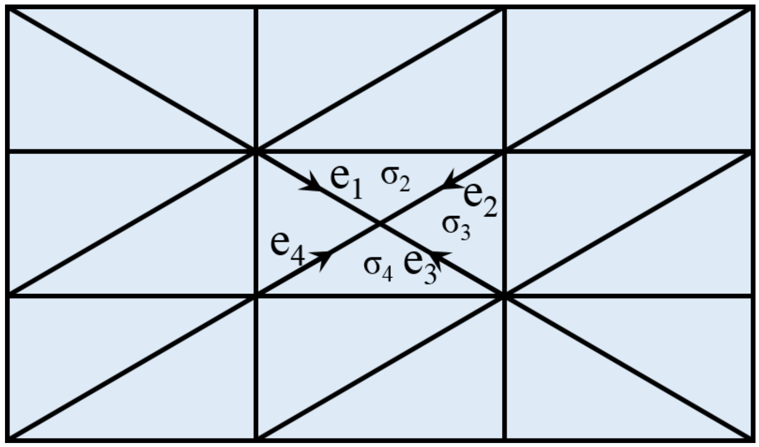


Figure 3.6

Consider an arbitrary 1-chain c of L, and let a be the value of c on e_4, as shown in Figure 3.6. By direct computation, the chain

$$c_1 = c + \partial\left(a\,\sigma_4\right)$$

has value 0 on e_4. By modifying c by a boundary, we have essentially "pushed it off" of e_4. Note also that if c is a cycle, then c_1 is a cycle as well. This holds for all subsequent steps of this process.

Let b be the value of c_1 on e_3. Then the chain

$$c_2 = c_1 + \partial\left(b\,\sigma_3\right)$$

has value 0 on e_3, and still has value 0 on e_4 as e_4 does not appear in $\partial\sigma_3$.

We may continue this process to show that any chain on L is homologous to a chain on the complex L' shown below:

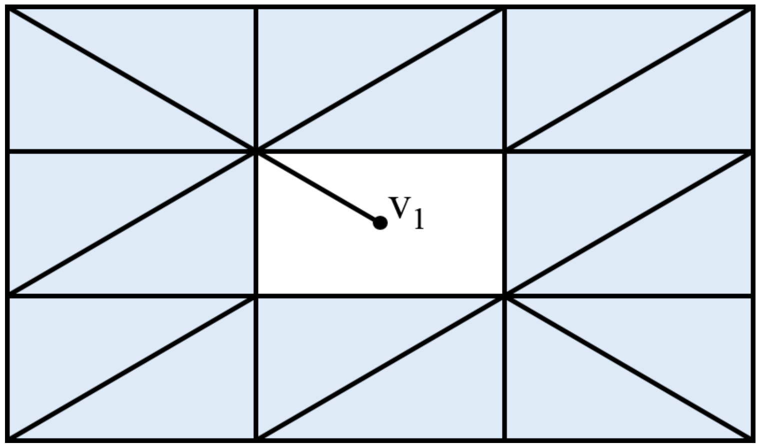


Figure 3.7

We may repeat this process using all remaining 2-simplices of L' to show that all p-chains of L are homologous to p-chains carried by the complex L'' shown below:

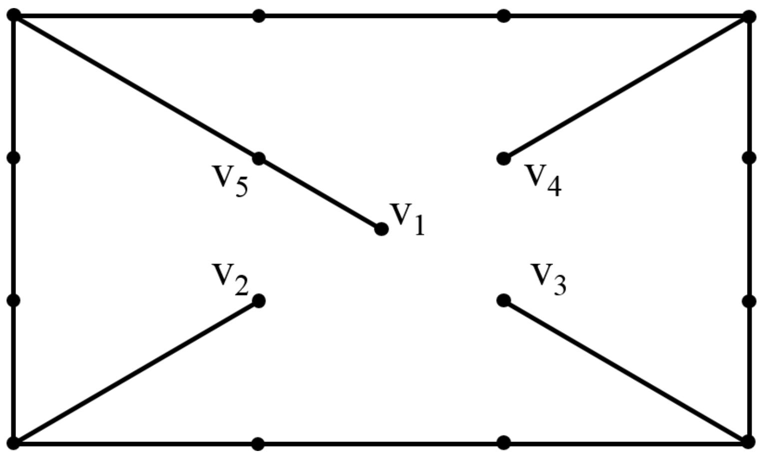


Figure 3.8

In the case where the original 1-chain c is a cycle, it must be carried by $\mathrm{Bd}\,(L)$, for otherwise c_2 would have a non-zero coefficient on one or more of the vertices $v_1, \ldots, v_5$.

To prove (2), first consider a 2-chain d of L. If σ_i and σ_j in d have an edge e in common, then that edge cannot be on $\mathrm{Bd}\,(L)$ and ∂d must have a value 0 on e. It follows that d must have the same value on σ_i as it does on σ_j. Continuing this process, we see that d has the same value on every oriented 2-simplex σ_i, i.e., d is a multiple of the chain $\sum \sigma_i$. ■

We are now ready to compute the homology of some common surfaces.

Homology Groups of the Torus

We may represent the torus T as the polytope of the labeled rectangle L shown below. Orient each 2-simplex counterclockwise and let γ denote their sum. In addition, let

$$w_1 = [a,\ b] + [b,\ c] + [c,\ a]\,,$$

$$z_1 = [a,\ d] + [d,\ e] + [e,\ a]\,.$$

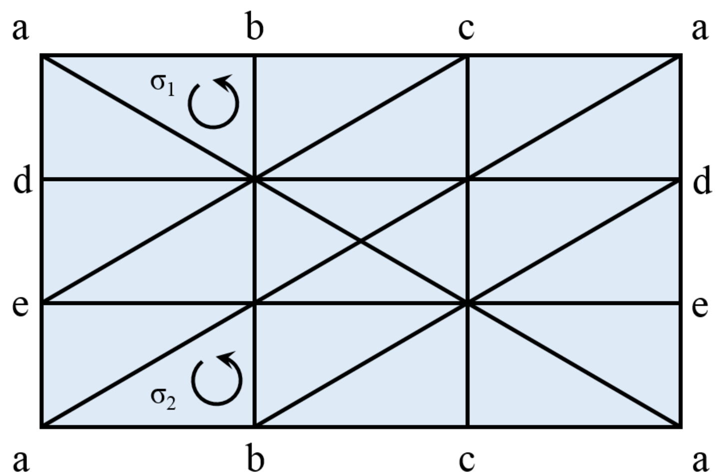


Figure 3.9: *The polytope of this labeled rectangle is the torus.*

We will show that

$$H_1(T) \cong \mathbb{Z} \oplus \mathbb{Z}, \qquad H_2(T) \cong \mathbb{Z},$$

where w_1 and z_1 generate $H_1(T)$ and γ generates $H_2(T)$. Figure 3.10 shows where w_1 and z_1 are located on the torus.

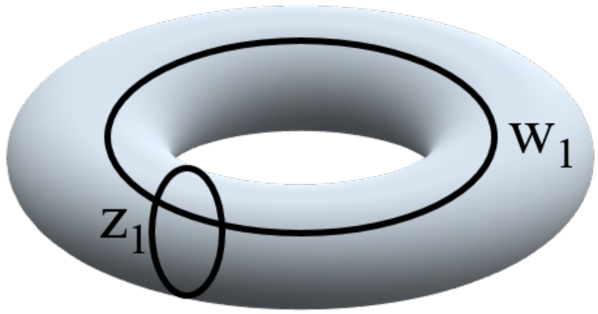


Figure 3.10: *The cycles w_1 and z_1 surround the central hole and the tube, respectively, on the torus following the identification of vertices of L.*

Let $g : |L| \to |T|$ be the pasting map; let $A = g\,(|\mathrm{Bd}\,(L)|)$. Then A is homeomorphic to a space that is the union of two circles with a point in common. (Such a space is called a **wedge of two circles**.) Orient the 1-simplices of T arbitrarily.

As g makes identifications only among simplices of $\mathrm{Bd}\,(L)$, the arguments we gave earlier in proving Lemma 3.12 apply verbatim to prove the following:

(i) Every 1-cycle of T is homologous to a 1-cycle carried by A, and

(ii) If d is a 2-chain of T and if ∂d is carried by A, then d is a multiple of γ.

In the complex T, two further results hold:

(iii) If c is a 1-cycle of T carried by A, then c is of the form $n\,w_1 + m\,z_1$, and

(iv) $\partial\gamma = 0$.

The proof of (iii) is given by Lemma 3.12 and the fact that A is just the 1-dimensional complex formed by z_1 and w_1 and pictured in Figure 3.10.

With regards to (iv), it is clear that $\partial\gamma$ has a value 0 on every 1-simplex of T not in A. We may check directly that it also has value 0 on each 1-simplex in A. For example, the chain $[a,\ b]$ appears in the expression for $\partial\sigma_1$ with value -1 and in the expression for $\partial\sigma_2$ with value $+1$, so that $\partial\gamma$ has value 0 on $[a,\ b]$.

Using results (i - iv), we can compute the homology of T.

Every 1-cycle of T is homologous to a 1-cycle of the form $c = n\,w_1 + m\,z_1$, by (i) and (iii). Such a cycle bounds only if it is trivial: For if $c = \partial d$ for some d, then (ii) applies to show that $d = p\gamma$ for some p; since $\partial\gamma = 0$ by (iv), we have $c = \partial d = 0$. We conclude that

$$H_1(T) \cong \mathbb{Z} \oplus \mathbb{Z},$$

and the 1-cycles w_1 and z_1 generate $H_1(T)$.

To compute $H_2(T)$, note that by (ii) any 2-cycle d of T must be of the form $p\gamma$ for some p. Each such 2-chain is in fact a cycle, by (iv), and there are no 3-chains for it to bound. We conclude that

$$H_2(T) \cong \mathbb{Z},$$

and this group has as generator the 2-cycle γ.

It is not difficult to see that $H_0(T) \cong \mathbb{Z}$, as the torus consists of a single connected component.

Homology Groups of a Klein Bottle

We may represent the Klein bottle $\mathcal{K}$ as the polytope of the labeled rectangle S shown below. Orient each 2-simplex counterclockwise and let γ denote their sum. In addition, let

$$w_1 = [a,\ b] + [b,\ c] + [c,\ a]$$

$$z_1 = [a,\ d] + [d,\ e] + [e,\ a]\,.$$

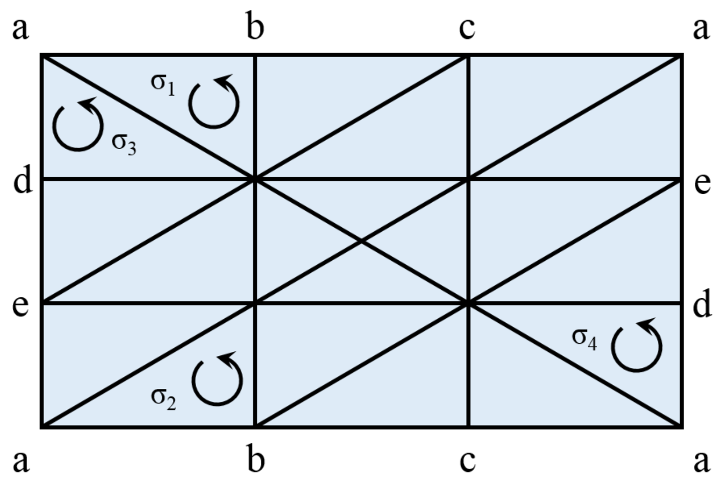


Figure 3.11: *The polytope of this labeled rectangle is the Klein bottle.*

We will show that

$$H_1(\mathcal{K}) \cong \mathbb{Z} \oplus \mathbb{Z}_2, \qquad H_2(\mathcal{K}) = 0,$$

where w_1 and z_1 generate $H_1(\mathcal{K})$, specifically w_1 generates $\mathbb{Z}$ and z_1 generates $\mathbb{Z}_2$. Figure 3.10 shows where w_1 and z_1 are located on the Klein bottle.

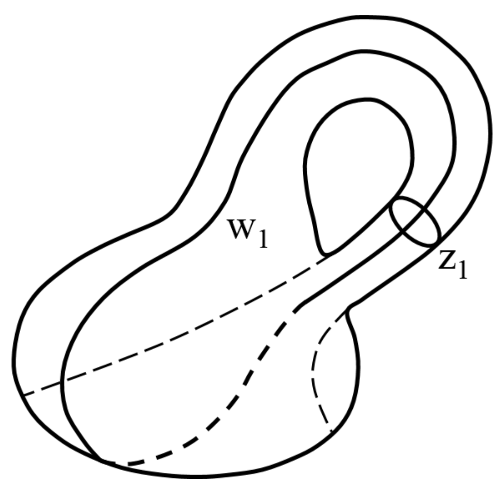


Figure 3.12: *The cycle w_1 goes through the Klein Bottle, while z_1 does not.*

Let $g : |S| \to \mathcal{K}$ be the pasting map. Let $A = g(|\mathrm{Bd}(S)|)$; as before, it is the wedge of two circles. Orient the 2-simplices of S as before; let γ be their sum. Orient the 1-simplices of A arbitrarily.

Note that (i) and (ii) from the calculations of the torus hold; neither involve particular identifications on the boundary. Because A is the wedge of two circles, (iii) holds as well. Part (iv) differs from the previous example, since we now have $\partial\gamma = 2\,z_1$.

This equation follows by direct computation. For example, $[a,\ b]$ appears in $\partial\sigma_1$ with coefficient -1 and in $\partial\sigma_2$ with coefficient $+1$, while $[a,\ d]$ appears in both $\partial\sigma_3$ and $\partial\sigma_4$ with coefficient $+1$.

Putting these facts together, we compute the homology of $\mathcal{K}$: as follows.

Every 1-cycle of $\mathcal{K}$ is homologous to a cycle of the form $c = n\,w_1 + m\,z_1$, by (i) and (iii). If $c = \partial d$ for some d, then $d = p\gamma$ by (ii); whence $\partial d = 2p\,z_1$. Thus, $n\,w_1 + m\,z_1$ bounds if and only if m is even and n is zero. We conclude that

$$H_1(\mathcal{K}) \cong \mathbb{Z} \oplus \mathbb{Z}_2.$$

The cycle z_1 represents the generator of $\mathbb{Z}_2$, and w_1 represents a generator $\mathbb{Z}$.

To compute $H_2(\mathcal{K})$, note that any 2-cycle d of S must be of the form $p\,\gamma$ by (ii); since $p\,\gamma$ is not a cycle, by (iv), we have

$$H_2(\mathcal{K}) = 0.$$

As the Klein bottle has a single connected component, $H_0(\mathcal{K}) \cong \mathbb{Z}$.

Chapter 4: Exact Sequences of Homology Groups

In Section 3.2, we utilized labeled diagrams and clever arguments to calculate the homology groups of various surfaces. *Long exact sequences of homology groups* provide another way to calculate the homology groups of surfaces, and primarily rely on ideas from algebra and previously calculated homology groups.

In Section 4.1, we begin our exploration of long exact sequences by examining *short exact sequences of vector spaces* as well as *long exact sequences* of collections of vector spaces, called *chain complexes*.

In Section 4.2, we utilize long exact sequences to calculate the homology of the sphere S^d for all dimensions d as well as the homology groups of the torus.

Section 4.1: Exact Sequences

Exact sequences are used in a wide variety of areas of mathematics. We only require exact sequences of homology groups for our purposes, although we introduce them as exact sequences of free abelian groups. Many terms we use in this more general context are the same as those used for exact sequences of homology groups, easing the transition between the two.

Maps Between Free Abelian Groups

Let $f : U \to V$ be a homomorphism between free abelian groups U and V. We define the kernel and image of f as usual:

(i) $\ker(f) = \{u \in U \mid f(u) = 0_V\}$, and

(ii) $\mathrm{Im}(f) = \{v \in V \mid f(u) = v \text{ for some } u \in U\}$.

Definition 4.1. If we have three free abelian groups and two homomorphisms, $f : U \to V$ and $g : V \to W$, then the sequence $U \xrightarrow{f} V \xrightarrow{g} W$ is **exact** at V if $\mathrm{Im}(f) = \ker(g)$, which implies $g \circ f = 0$.

More generally, if $0 \xrightarrow{a} U \xrightarrow{b} V$ is a sequence, then exactness at U is equivalent to injectivity of $U \xrightarrow{b} V$, as $\mathrm{Im}(a) = 0 = \ker(b)$. Similarly, if $V \xrightarrow{c} W \xrightarrow{d} 0$, then exactness at W is equivalent to surjectivity of $V \xrightarrow{c} W$, as $\ker(d) = W = \mathrm{Im}(c)$.

A **short exact sequence** is a sequence of length 5,

$$0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0,$$

that starts and ends with the trivial vector space and is exact at U, V, and W. Hence, by the above statement, f is injective and g is surjective. In this situation, it is always true that $V \cong U \oplus W$.

Edelsbrunner introduces exact sequences in [4] using exact sequences of vector spaces with linear transformations between them. The statement that $V \cong U \oplus W$ when $0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0$ holds when U, V, W are vector spaces or free abelian groups, but may fail to hold when U, V, W are arbitrary abelian groups.

Definition 4.2. We now begin to move toward using long exact sequences for homology groups by considering a sequence of free abelian groups with homomorphisms between them, $\mathcal{U} = (U_p, u_p)$ with $u_p : U_p \to U_{p-1}$.

If $u_p \circ u_{p+1} = 0$ for every p, then $\mathcal{U}$ is a **chain complex** and the u_p its **boundary maps**. These boundary maps should remind the reader of the boundary operator ∂_p for p-chains.

In the same way that we defined cycle, boundary, and homology groups for simplicial complexes, we now define them for chain complexes:

(i) Cycle groups: $Z_p(\mathcal{U}) = \ker(u_p)$;

(ii) Boundary groups: $B_p(\mathcal{U}) = \mathrm{Im}(u_{p+1})$, and;

(iii) Homology groups: $H_p(\mathcal{U}) = Z_p(\mathcal{U}) / B_p(\mathcal{U})$.

Definition 4.3. Let $\mathcal{V} = (V_p, v_p)$ be another chain complex. A **chain map** is a sequence of homomorphisms $\phi_p : U_p \to V_p$, one for each dimension p, that commute with the boundary maps (specifically, $v_p \circ \phi_p = \phi_{p-1} \circ u_p$ for every p).

Commutativity between chain maps and boundary maps guarantees that cycles go to cycles and boundaries go to boundaries, i.e., $\phi_p(Z_p(\mathcal{U})) \subseteq Z_p(\mathcal{V})$ and $\phi_p(B_p(\mathcal{U})) \subseteq B_p(\mathcal{V})$.

Thus, the chain map induces a map on homology $(\phi_p)_* : H_p(\mathcal{U}) \to H_p(\mathcal{V})$ for every dimension p.

Let $\mathcal{W} = (W_p, w_p)$ be a third chain complex and the sequence of $\psi_p : V_p \to W_p$ a second chain map. The sequence $\mathcal{U} \to \mathcal{V} \to \mathcal{W}$ is **exact at** $\mathcal{V}$ if $\ker(\psi_p) = \mathrm{Im}(\phi_p)$ for every p.

A **short exact sequence** of chain complexes is a sequence of length 5,

$$0 \to \mathcal{U} \xrightarrow{\phi} \mathcal{V} \xrightarrow{\psi} \mathcal{W} \to 0,$$

that begins and ends with the trivial chain complex and is exact at $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$.

Equivalently, there is a short exact sequence of vector spaces $0 \to U_p \to V_p \to W_p \to 0$ for each dimension p. This implies that each ϕ_p is injective, each ψ_p is surjective, and $V_p \cong U_p \oplus W_p$ for all p.

Long Exact Sequences of Homology Groups

We may adapt our established framework of exact sequences of free abelian groups to exact sequences of homology groups, and some results may be carried over without any additional work. We may now begin our discussion on the method for constructing long exact sequences of homology groups from short exact sequences of chain complexes.

Lemma 4.4. SNAKE LEMMA. Let $0 \to \mathcal{U} \xrightarrow{\phi} \mathcal{V} \xrightarrow{\psi} \mathcal{W} \to 0$ be a short exact sequence of chain complexes. There is a well-defined map $D\colon H_p(\mathcal{W}) \to H_{p-1}(\mathcal{U})$, called the **connecting homomorphism** such that

$$\cdots \to H_p(\mathcal{U}) \to H_p(\mathcal{V}) \to H_p(\mathcal{W}) \xrightarrow{D} H_{p-1}(\mathcal{U}) \to \dots$$

is a **long exact sequence**, i.e., an infinite sequence which is exact at each entry.

Proof. Other than the connecting homomorphism D, the maps in the long exact sequence are induced by the chain maps. We construct D and the proof that the sequence is exact is omitted here but may be found in [5].

$$\begin{tikzcd}
 & & V_{p+1} \arrow[r, "\psi_{p+1}"] \arrow[d] \arrow[dr, phantom, "\square_3"] & W_{p+1} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & U_p \arrow[r, "\phi_p"] \arrow[d] \arrow[dr, phantom, "\square_2"] & V_p \arrow[r, "\psi_p"] \arrow[d] \arrow[dr, phantom, "\square_0"] & W_p \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & U_{p-1} \arrow[r, "\phi_{p-1}"] \arrow[d] \arrow[dr, phantom, "\square_1"] & V_{p-1} \arrow[r, "\psi_{p-1}"] \arrow[d] & W_{p-1} \arrow[r] & 0 \\
0 \arrow[r] & U_{p-2} \arrow[r, "\phi_{p-2}"] & V_{p-2} & &
\end{tikzcd}$$

For brevity, we omit the subscripts on the boundary maps u, v, w from the text of the proof. The squares in each diagram are commutative squares, e.g., $\square_0$ implies that $w\psi_p = \psi_{p-1}v$ as a map from V_p to W_{p-1} in the diagram above.

Step 1: Define γ.

$$\begin{tikzcd} & \beta \arrow[r, "\psi_p"] \arrow[d] \arrow[dr, phantom, "\square_0"] & \alpha \arrow[d] \\ \gamma \arrow[r, "\phi_{p-1}"] \arrow[d] \arrow[dr, phantom, "\square_1"] & v(\beta) \arrow[r, "\psi_{p-1}"] \arrow[d] & 0 \\ 0 \arrow[r, "\phi_{p-2}"] & 0 & \end{tikzcd}$$

Let $\alpha \in W_p$ be a cycle representing a class in $H_p(\mathcal{W})$. Exactness at W_p implies ψ_p is surjective, which means there exists a chain $\beta \in V_p$ with $\psi_p(\beta) = \alpha$.

Since α is a cycle and thus has zero boundary, the boundary of β lies in the kernel of the chain map ψ_p by commutativity of $\square_0$, i.e., $v(\beta) \in \ker(\psi_{p-1})$.

Exactness at V_{p-1} means that $\operatorname{Im}(\phi_{p-1}) = \ker(\psi_{p-1})$, which implies that there exists a chain $\gamma \in U_{p-1}$ whose image under the first chain map is the boundary of β, i.e., $\phi_{p-1}(\gamma) = v(\beta)$.

Step 2: γ is a cycle. We continue to utilize the diagram included in Step 1.

By commutativity of $\square_1$ and the composition of boundary maps being zero, specifically $v\,v = 0$, we have $\phi_{p-2}\,u(\gamma) = 0$.

However, by exactness at U_{p-2}, ϕ_{p-2} is injective and thus $u(\gamma) = 0$.

Therefore, γ is a cycle and thus represents a class in $H_{p-1}(\mathcal{U})$. Additionally, this class is the image of the class represented by α under the connecting homomorphism D.

D goes left, from α to β, then down to $v(\beta)$, and then left again to γ. We may draw this as a snake cutting through the diagram, hence the name of the Snake Lemma.

It suffices to show that our answer does not depend on our choices of α and β.

Step 3: Choice of β.

$$\begin{tikzcd}
\mu \arrow[r, "\phi_p"] \arrow[d] \arrow[dr, phantom, "\square_2"] & \beta,\ \beta_0 \arrow[r, "\psi_p"] \arrow[d] \arrow[dr, phantom, "\square_0"] & \alpha \arrow[d] \\
\gamma,\ \gamma_0 \arrow[r, "\phi_{p-1}"] & \upsilon\left(\beta\right),\ \upsilon\left(\beta_0\right) \arrow[r, "\psi_{p-1}"] & 0
\end{tikzcd}$$

Let β_0 be another choice for β, i.e., β_0 is a chain in V_p with $\psi_p\left(\beta\right) = \alpha$. By our efforts in Step 1, we know there is an element γ_0 of U_{p-1} such that $\phi_{p-1}\left(\gamma_0\right) = \upsilon\left(\beta_0\right)$.

The fact that $\psi_p(\beta) = \psi_p(\beta_0) = \alpha$ along with exactness at V_p implies $\beta - \beta_0 \in \ker\left(\psi_p\right) = \operatorname{Im}\left(\phi_p\right)$. Thus, there exists a chain $\mu \in U_p$ with $\phi_p\left(\mu\right) = \beta - \beta_0$.

By commutativity of $\square_2$,

$$\upsilon\left(\beta\right) - \upsilon\left(\beta_0\right) = \upsilon\,\phi_p\left(\mu\right) = \phi_{p-1}\,u\left(\mu\right) = \phi_{p-1}\left(\gamma\right) - \phi_{p-1}\left(\gamma_0\right).$$

By exactness at V_{p-1}, ϕ_{p-1} is injective, and thus $u\left(\mu\right) = \gamma - \gamma_0$. Therefore, γ and γ_0 differ by the boundary $u\left(\mu\right)$ and thus represent the same homology class.

Step 4: Choice of α.

$$\begin{tikzcd}
 & \varrho \arrow[r, "\psi_{p+1}"] \arrow[d] \arrow[dr, phantom, "\square_3"] & \nu \arrow[d] \\
\mu' \arrow[r, "\phi_p"] \arrow[d] \arrow[dr, phantom, "\square_2"] & \upsilon\left(\varrho\right),\ \beta,\ \beta_0 \arrow[r, "\psi_p"] \arrow[d] \arrow[dr, phantom, "\square_0"] & \alpha,\ \alpha_0 \arrow[d] \\
\gamma,\ \gamma_0 \arrow[r, "\phi_{p-1}"] & 0,\ \upsilon\left(\beta\right),\ \upsilon\left(\beta_0\right) \arrow[r, "\psi_{p-1}"] & 0
\end{tikzcd}$$

Let α_0 be a different choice for α, i.e., α_0 is a cycle in W_p representing the same homology class in $H_p\left(\mathcal{W}\right)$ as α. We define β_0 and γ_0 in the same manner we defined β and γ using α, i.e., $\psi_p\left(\beta_0\right) = \alpha_0$ and $\phi_{p-1}\left(\gamma_0\right) = \upsilon\left(\beta_0\right)$.

Since α and α_0 are in the same homology class in $H_p\left(\mathcal{W}\right)$, there exists a chain $\nu \in W_{p+1}$ such that $w\left(\nu\right) = \alpha - \alpha_0$.

By exactness at W_{p+1}, ψ_{p+1} is surjective and thus there exists a chain $\varrho \in V_{p+1}$ with $\psi_{p+1}(\varrho) = \nu$.

By commutativity of $\square_3$, $\upsilon(\varrho)$ and $\beta - \beta_0$ both map to $\alpha - \alpha_0$. This implies that their difference lies in $\ker(\psi) = \mathrm{Im}(\phi)$ and there is a chain $\mu' \in U_p$ with $\phi_p(\mu') = \upsilon(\varrho) + \beta_0 - \beta$.

Using the commutativity of $\square_2$ and the fact that $\upsilon\upsilon = 0$, we see that $\phi_{p-1}\, u(\mu') = \upsilon\left(\beta_0 - \beta\right)$.

The injectivity of ϕ_{p-1} from the exactness at U_{p-1} implies that the preimage of $\upsilon\left(\beta - \beta_0\right)$ is $\gamma - \gamma_0$ and hence $u(\mu') = \gamma - \gamma_0$.

Hence, γ and γ_0 differ by a boundary and represent the same homology class, as required.

■

The Mayer-Vietoris Sequence

Given two topological spaces, the *Mayer-Vietoris sequence* relates their homology to the homology of their union and intersection. In other words, we can use it to compute the homology of an unknown space using the homology of two subspaces whose union is the desired space.

Theorem 4.5. MAYER-VIETORIS SEQUENCE THEOREM. Let K be a simplicial complex and K', K'' subcomplexes such that $K = K' \cup K''$. Let $A = K' \cap K''$. There there exists a long exact sequence

$$\ldots H_p(A) \to H_p(K') \oplus H_p(K'') \to H_p(K) \to H_{p-1}(A) \to \ldots$$

and similarly for the reduced homology groups.

Proof. Note that $C_p(A)$ is a subgroup of both $C_p(K')$ and $C_p(K'')$ on the level of chains.

By forming the direct sums $C_p(K') \oplus C_p(K'')$ for all dimensions p, we obtain a chain complex $C(K') \oplus C(K'')$ whose boundary maps are those ∂_p for the chain groups, define component-wise.

Within this direct sum, there are two copies of $C_p(A)$, which we address in the following manner:

Let $i' : A \to K'$ and $i'' : A \to K''$ be inclusion mappings of A into K' and K'' as simplicial complexes, and let $j' : K' \to K$ and $j'' : K'' \to K$ be the inclusion mappings of K' and K'' into K.

Set $i(a) = (i'(a), i''(a))$ and $j(x, y) = j'(x) - j''(y)$, which are indeed chain maps as each simplex is mapped to a simplex of equal or less dimension. These create a short exact sequence of chain complexes

$$0 \to C(A) \xrightarrow{i} C\left(K'\right) \oplus C\left(K''\right) \xrightarrow{j} C(K) \to 0.$$

The long exact sequence of homology groups implied by the Snake Lemma is the **Mayer-Vietoris sequence**, and may be adapted to the reduced homology sequence as well. ■

Consider the maps induced on homology groups by i and j, namely $i^* : H_{p-1}(A) \to H_{p-1}(K') \oplus H_{p-1}(K'')$ and $j^* : H_p(K') \oplus H_p(K'') \to H_p(K)$. Exactness of the Mayer-Vietoris sequence at $H_p(K)$ implies that $H_p(K) \cong \ker(i^*) \oplus \mathrm{Im}(j^*)$ in the case where the homology groups are free abelian.

Thus, there are two types of homology classes in K; those in $\ker(i^*)$ and $\mathrm{Im}(j^*)$:

- A homology class in $\ker(i^*)$ corresponds to a $(p-1)$-dimensional cycle $\gamma \in A$ that bounds both in K' and K''.

  Let α' and α'' be the p-chains in K' and K'', respectively, that γ bounds. If we write $\gamma = \partial\alpha' = \partial\alpha''$, then $\alpha = \alpha' - \alpha''$ is a cycle in K that represents a homology class in $\ker(i^*)$.

- A homology class in $\mathrm{Im}(j^*)$ is one that lives in K', K'', or both.

To be thorough, we work through the construction of the connecting homomorphism D in this context, which is of the form of the construction in the proof of the Snake Lemma.

Consider a homology class in $H_p(K)$ and define one in $H_{p-1}(A)$ as follows:

Let this class be represented by a p-cycle α of K. As before, there exists β, a p-chain in $C_p(K') \oplus C_p(K'')$ such that $j(\beta) = \alpha$.

There are several such β, which we obtain by writing $\alpha = \alpha' + \alpha''$, with $\alpha' \in K'$, $\alpha'' \in K''$, and setting $\beta = (\alpha',\ \alpha'')$.

Any two such decompositions of α give different β, but any two of these β differ by something in A.

Now consider the boundary of β, $\partial\beta = (\partial\alpha',\ \partial\alpha'')$. As α is a cycle, $\partial\alpha' = -\partial\alpha''$ and lies in A. Thus, the cycle γ in the construction of D is $\partial\alpha'$.

In Section 3.2, we utilized labeled diagrams to calculated specific homology groups of surfaces. The Mayer-Vietoris sequence allows us to, given knowledge of the homology groups of some spaces, quickly calculate the homology groups of other spaces.

We note that if two spaces are homeomorphic or of the same homotopy type, then their homology groups will be isomorphic. The interested reader may consult Chapter 2 of [9].

Homology Groups of the Sphere S^d

We use the Mayer-Vietoris sequence to compute the Betti numbers, and thus the homology groups, of S^d, specifically

$$\tilde{\beta}_p\left(S^d\right) = \begin{cases} 1 & \text{if } p = d, \\ 0 & \text{if } p \neq d. \end{cases}$$

We begin by writing S^d as the union between its upper and lower hemispheres, $S^d = U \cup L$. Each U and L are homeomorphic to balls of dimension $d - 1$ and intersect in a sphere of dimension $d - 1$.

Using the Mayer-Vietoris sequence, we may compute the homology of S^d inductively:

$$\cdots \to \tilde{H}_p\left(S^{d-1}\right) \to \tilde{H}_p(U) \oplus \tilde{H}_p(L) \to \tilde{H}_p\left(S^d\right) \to \tilde{H}_{p-1}\left(S^{d-1}\right) \to \dots$$

We base our induction with the observation that S^0 consists of two points, so $\tilde{H}_0\left(S^0\right) \cong \mathbb{Z}$ and $\tilde{H}_p\left(S^0\right) = 0$ for all $p > 0$.

For any d, the sequence decomposes into parts of the form

$$0 \oplus 0 \xrightarrow{a} \tilde{H}_p\left(S^d\right) \xrightarrow{b} \tilde{H}_{p-1}\left(S^{d-1}\right) \xrightarrow{c} 0 \oplus 0,$$

where $0 \oplus 0$ is the zero element in the direct sum of the homology groups of the two hemispheres.

By exactness of the sequence, $\mathrm{Im}\,(a) = \ker(b)$, and thus b must be injective. Furthermore, as $\mathrm{Im}\,(b) = \ker(c)$, b must be surjective as well. Therefore, $\tilde{H}_p\left(S^d\right) \cong \tilde{H}_{p-1}\left(S^{d-1}\right)$, and thus the result holds.

Notice that the generator of $\tilde{H}_p\left(S^d\right)$ consists of two chains, one from each hemisphere, whose boundary is the generating cycle of $\tilde{H}_{p-1}\left(S^{d-1}\right)$.

Homology Groups of the Torus

Recall from Section 3.2 that the Betti numbers of the torus T are

$$\beta_p(T) = \begin{cases} 1 & \text{if } p = 0; \\ 2 & \text{if } p = 1; \\ 1 & \text{if } p = 2; \\ 0 & \text{if } p > 2. \end{cases}$$

Consider the torus as the union of two cylinders C_1 and C_2, whose intersection is made up of two disjoint copies of S^1, which we denote SS. Note also that each of these cylinders is homotopic to S^1, and thus have the same homology groups (see [4]). Thus, the Mayer-Vietoris sequence is of the form

$$\cdots \to H_p(SS) \to H_p\left(S^1\right) \oplus H_p\left(S^1\right) \to H_p(T) \to H_{p-1}(SS) \to \dots$$

Using known sequences, we observe that the result holds for $p > 2$. We write the portion of the sequence relevant for $p = 0,\ 1,\ 2$ below, where we have utilized our consideration of T

as the union of two cylinders ,

$$H_2\left(S^1\right) \oplus H_2\left(S^1\right) \to H_2(T) \xrightarrow{a} H_1(SS) \xrightarrow{b} H_1\left(S^1\right) \oplus H_1\left(S^1\right) \xrightarrow{c} H_1(T) \xrightarrow{d}$$

$$H_0(SS) \to H_0\left(S^1\right) \oplus H_0\left(S^1\right) \to H_0(T) \to H_{-1}(SS).$$

Using known homology groups, we obtain

$$0 \oplus 0 \to H_2(T) \xrightarrow{a} \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{b} \mathbb{Z} \oplus \mathbb{Z} \xrightarrow{c} H_1(T) \xrightarrow{d} \mathbb{Z} \oplus \mathbb{Z} \to \mathbb{Z} \oplus \mathbb{Z} \to H_0(T) \to 0.$$

As T is connected, we find that $H_0(T) = \mathbb{Z}$, and thus the result for $p = 0$ holds.

Recall the map $i: A \to K' \oplus K''$ by $i(a) = (i'(a), i''(a))$ defined in the proof of the Mayer-Vietoris Sequence Theorem, which induces the homomorphisms b in the above sequence.

Let the cycles α' and α'' generate the homology groups of SS. As they are in each of the cylinders C_1 and C_2, we find that $i(\alpha') \simeq i(\alpha'') \simeq (\alpha', \alpha'') \simeq (\alpha', \alpha')$. Therefore, the kernel of b is isomorphic to $\mathbb{Z}$.

By exactness of the sequence, $\ker(a) = H_2(T)$ and $\operatorname{Im}(a) = \ker(b) \cong \mathbb{Z}$. Therefore, as a is a homomorphism, $H_2(T) \cong \mathbb{Z}$ as desired.

Recall the map $j: K' \oplus K'' \to K$ defined in the proof of the Mayer-Vietoris Sequence Theorem, which induces the homomorphism c in the above sequence, and note that i also induces the homomorphism d. Also recall that the two types of homology classes in K are those in $\ker(i^*)$ and $\operatorname{Im}(j^*)$.

These observations imply that the homology classes in $H_1(T)$ are the direct sum of $\ker(c)$ and $\operatorname{Im}(d)$, and thus $H_2(T) \cong \mathbb{Z} \oplus \mathbb{Z}$ as desired.

Chapter 5: Persistent Homology

Persistent homology utilizes the geometry of a topological space, represented by a finite simplicial complex, to measure the significance of its topological features.

Section 5.1 includes the definition for *filtrations*, which serve as the foundation for *persistent homology groups*.

Section 5.2 introduces *persistence diagrams*, which are plots that encode all of the information contained in the persistent homology groups. It also includes a brief discussion on the stability of persistence homology.

Section 5.1: Persistent Homology Groups

Before we are able to discuss persistent homology groups, we introduce filtrations, which in turn require the notion of *level sets* and *sublevel sets*:

Definition 5.1. Let K be a simplicial complex and $f : K \to \mathbb{R}$. The function f is said to be **monotonically increasing along chains of faces** if $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ.

Example 5.1. Consider the simplicial complex K shown below, whose simplices σ are labeled with the value $f(\sigma)$ that $f : K \to \mathbb{R}$ maps them to.

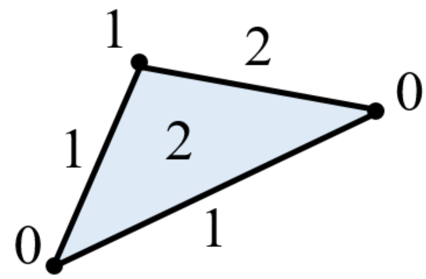


Figure 5.1

The function f is monotonically increasing along chains of faces, as $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ.

Definition 5.2. The preimage $f^{-1}(a)$ of each real number a is called a **level set**, which consists of all simplices that f maps to a.

The **sublevel set**, denoted $K(a)$, consists of all simplices which f maps to at most a, i.e., $f^{-1}(\infty, a]$.

Example 5.2. Consider the following simplicial complex K, with each simplex labeled with the value that a monotonically increasing function f assigns it.

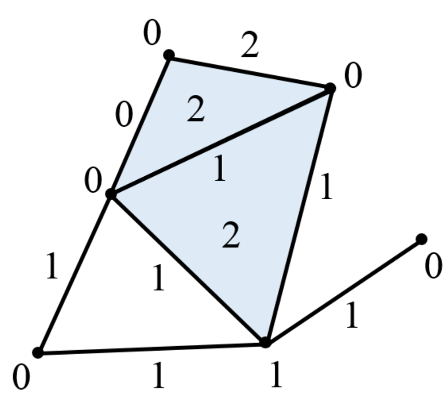


Figure 5.2

The following figure shows the sublevel sets $K(0)$, $K(1)$, and $K(2)$ of K.

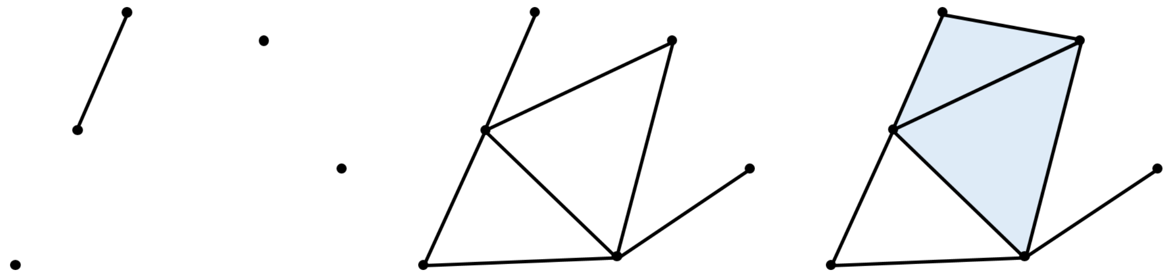


Figure 5.3: *The sublevel sets $K(0)$ (left), $K(1)$ (middle), and $K(2)$ (right).*

Observe that the sublevel set $K(a) = f^{-1}(\infty,\ a]$ is a subcomplex of K for every $a \in \mathbb{R}$, and more specifically $K(a)$ is a subcomplex of $K(b)$ whenever $a \leq b$.

This observation holds for all sublevel sets of all simplicial complexes, and is the result of the monotonicity of f.

We are now ready to define *filtrations*, which serve as the foundation of our exploration into persistent homology:

Definition 5.3. Let m be the number of simplices in K, and $f : K \to \mathbb{R}$ be a monotonically increasing along chains of faces. We obtain $n + 1 \leq m + 1$ different subcomplexes, which we arrange as an increasing sequence

$$\varnothing = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K.$$

Specifically, if $a_1 < a_2 < \cdots < a_n$ are the function values of the simplices in K and $a_0 = -\infty$, then $K_i = K\left(a_i\right)$ for each i.

This sequence of complexes is called a **filtration** of f, and we think of it as a construction by adding collections of simplices until the final simplicial complex is achieved.

Example 5.3. We may consider of Čech and Vietoris-Rips complexes as filtrations of the complex $\lim\limits_{r\to\infty}$ Čech$(r) = \lim\limits_{r\to\infty}$ Vietoris-Rips(r), which we denote K. By the Vietoris-Rips Lemma, the definition of each filtration is similar, and rigorously define $f : K \to \mathbb{R}$ for Vietoris-Rips complexes for the sake of brevity.

Recall that we include a simplex σ in Vietoris-Rips(r) if $\operatorname{diam}(\sigma) \leq 2r$. We use this to define $f(\sigma) = \frac{\text{diam}(\sigma)}{2}$ for each simplex σ of K. The function is indeed monotonically increasing on chains of faces, as for any sets S and T of points, $\text{diam}(S) \le \text{diam}(T)$ whenever $S \subseteq T$.

Persistent homology is concerned with the topological evolution of these filtrations, as expressed by the corresponding sequence of homology groups.

For every $i \le j$, we have an inclusion map from the underlying space of K_i to the underlying space of K_j. Similarly to the functions i and j from our construction of the Mayer-Vietoris sequence in the previous chapter, these inclusion mappings induce a chain map, and thus an induced homomorphism $f_p^{i,j} : H_p(K_i) \to H_p(K_j)$ for each dimension p.

The filtration thus corresponds to a sequence of homology groups connected by homomorphisms

$$0 = H_p(K_0) \to H_p(K_1) \to \cdots \to H_p(K_n) = H_p(K),$$

for each dimension p.

As we go from K_{i-1} to K_i, we gain new homology classes and lose some when they become trivial or merge with each other. We collect the classes that are born at or before a given threshold and die after another threshold in groups:

Definition 5.4. The p^{th}**-persistent homology groups** are the images of the homomorphisms induced by inclusion, $H_p^{i,j} = \text{Im}\left(f_p^{i,j}\right)$ for $0 \le i \le j \le n$.

The corresponding p^{th}**-persistent Betti numbers** are the ranks of these groups. Reduced persistent homology groups and reduced persistent Betti numbers are defined similarly.

The persistent homology groups consist of the homology classes of K_i that are still alive at K_j, i.e., $H_p^{i,j} = Z_p(K_i) / \left(B_p(K_j) \cap Z_p(K_i)\right)$. There is such a group for each dimension p and each index $i \le j$.

Example 5.4. Consider the filtration made up of the sublevel sets $K(0)$, $K(1)$, and $K(2)$ from Example 5.2, whose vertices we label in the figure below

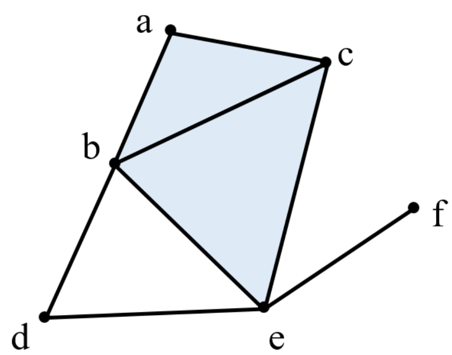


Figure 5.4

We calculate $H_0^{i,j}$ for $i = 0, 1$ and $j = 1, 2$.

[$i = 0, j = 1, 2$.] As $K(0)$ has 5 vertices, $Z_0(K(0)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$. Although the vertices $\{a, b, c, d, f\}$ generate this group, it is more useful to note that the 0-chains

$$\{a, a-b, a-c, a-d, a-f\}$$

also generate $Z_0(K(0))$.

Four of these five 0-cycles, specifically $a-b$, $a-c$, $a-d$, and $a-f$, are boundaries of 1-chains in $K(1)$. Thus, $B_0(K(1)) \cap Z_0(K(0)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$. Therefore, $H_0^{0,1} \cong \mathbb{Z}$.

Similarly, $B_0(K(2)) \cap Z_0(K(0)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$, and thus $H_0^{0,2} \cong \mathbb{Z}$.

[$i = 1, j = 1, 2$.] As $K(1)$ has 6 vertices, $Z_0(K(1)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$. As with $Z_0(K(0))$, it is useful to note that

$$\{a, a-b, a-c, a-d, a-e, a-f\}$$

generate $Z_0(K(1))$.

Five of these six 0-chains are boundaries of 1-chains in $K(1)$ as well, specifically $a-b$, $a-c$, $a-d$, $a-e$, and $a-f$. Thus, $B_0(K(1)) \cap Z_0(K(1)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$, and in turn $H_0^{1,1} \cong \mathbb{Z}$.

Similarly, $B_0(K(2)) \cap Z_0(K(1)) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$, and thus $H_0^{1,2} \cong \mathbb{Z}$.

Consider a class γ in $H_p(K_i)$. It is **born at** K_i if $\gamma \notin H_p^{i-1,i}$. It **dies entering** K_j if it merges with an existing class as we go from K_{j-1} to K_j, i.e., $f_p^{i,\,j-1}(\gamma) \notin H_p^{i-1,\,j-1}$ and $f_p^{i,\,j}(\gamma) \in H_p^{i-1,\,j}$.

Rule 5.5. PERSISTENT HOMOLOGY ELDER RULE. If γ is born at K_i and dies entering K_j, then the difference in function value is the **persistence** of γ, $\operatorname{pers}(\gamma) = a_j - a_i$. We sometimes denote the persistence instead by the difference in index, which we appropriately call the **index persistence** of the class.

If γ is born at K_i and never dies, then we set its (index) persistence to infinity.

Section 5.2: Persistence Diagrams

Persistence diagrams are used to represent the persistent homology groups of a filtration by plotting when each homology class is born and when each homology class dies on the plane $\mathbb{R}^2$. *Barcodes* encode this same information on a bar plot.

We represent a collection of persistent Betti numbers by drawing points in two dimensions. Let $\mu_p^{i,\,j}$ be the number of p-dimensional classes born at K_i and dying entering K_j. We have

$$\mu_p^{i,\,j} = \left(\beta_p^{i,\,j-1} - \beta_p^{i,\,j}\right) - \left(\beta_p^{i-1,\,j-1} - \beta_p^{i-1,\,j}\right),$$

for all $i < j$ and all p.

The first difference on the right-hand side counts the classes that are born at or before K_i and die entering K_j. The second difference counts the classes that are born at or before K_{i-1} and die entering K_j.

Example 5.5. Recall the persistent homology groups we calculated in Example 5.4. We may use these groups to calculate $\mu_0^{1,\,2}$

$$\begin{aligned}\mu_0^{1,\,2} &= \left(\beta_0^{1,\,1} - \beta_0^{1,\,2}\right) - \left(\beta_0^{0,\,1} - \beta_0^{0,\,2}\right)\\ &= (1-1) - (1-1)\\ &= 0.\end{aligned}$$

This is in line with the interpretation of each difference, as all but a single 0-cycle die entering K_2, which never dies.

Definition 5.6. Recall that $a_1 < a_2 < \cdots < a_n$ are the function values of the simplices in K, with $a_0 = -\infty$. Drawing each point (a_i, a_j) with multiplicity ${\mu_p}^{i,j}$, we obtain the p^{th} **persistence diagram of the filtration**, denoted $\mathrm{Dgm}_p(f)$. The value a_i represents the function value at which a class γ is born, while the value a_j represents the function value which γ dies entering.

Note that the multiplicity with which we draw a point does not manifest in the physical representation of the diagram, but is considered to be contained within the diagram.

The persistence of γ is represented by the vertical distance of the point from the diagonal. Since multiplicities are defined only for $i < j$, all points lie on or above the diagonal. For computational convenience, we consider the diagonal to have uncountably infinite many points.

Example 5.6. Below is the persistence diagram in dimension 1 for the filtration presented in Example 5.2. Note that the complex contains a 1-cycle which is born at 1 but never dies, so it does not appear in the persistence diagram.

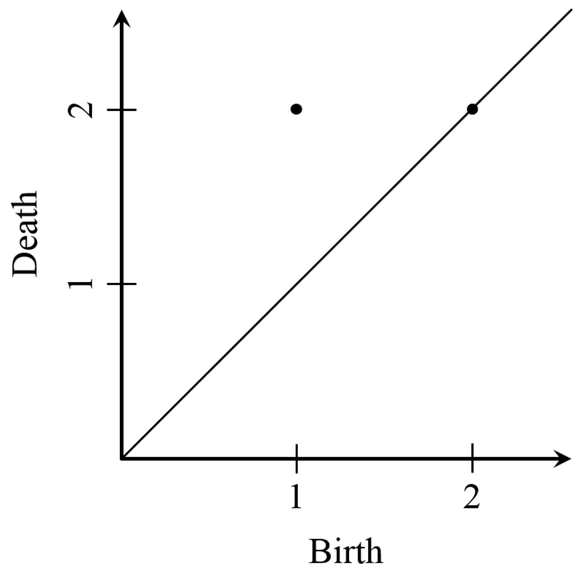


Figure 5.5: *The persistence diagram in dimension 1 of the filtration presented in Example 5.1. Each point on the diagram represents the birth and death of a homology class in the filtration. Note that the complex contains a* 1*-cycle which is born at* 1 *but never dies, so it does not appear in the diagram.*

Alternatively, we may represent a persistence diagram using a *barcode*, which utilizes line segments instead of points to represent persistence.

Definition 5.7. Consider a p-cycle γ with persistence $\text{pers}(\gamma) = a_j - a_i$, which may be infinite. In the p^{th} **barcode of the filtration**, we represent the persistence of each cycle using a line segment of length $\text{pers}(\gamma)$.

Example 5.7. Below is the barcode in dimension 1 for the filtration presented in Example 5.1. Note that the complex contains a 1-cycle which is born and dies at 2, and thus is not represented in the barcode.

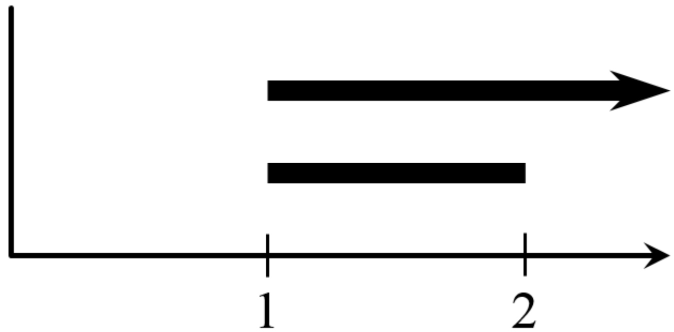


Figure 5.6: *The barcode in dimension 1 of the filtration presented in Example 5.1. Each bar in the barcode represents the birth and death of a homology class in the filtration. Note that the complex contains a* 1*-cycle which is born and dies at* 2*, so it does not appear in the barcode.*

The Fundamental Theorem of Persistent Homology states that the persistence diagram actually encodes all of the information about the persistent homology groups when using modulo 2 coefficients.

Theorem 5.8. FUNDAMENTAL THEOREM OF PERSISTENT HOMOLOGY. Let $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ be a filtration. For every pair of indices $0 \leq k \leq l \leq n$, and every dimension p, the p^{th} persistent Betti number is $\beta_p^{k,\,l} = \sum_{i \leq k} \sum_{j > l} \mu_p^{i,\,j}$.

A proof of the Fundamental Theorem of Persistent Homology may be found in [11].

Bottleneck Stability of Persistence Diagrams

Persistent homology is often used, in conjunction with Vietoris-Rips complexes, to analyze data encoded as a point cloud. As with any data analysis tool, we desire that persistence diagrams (and thus barcodes) for similar point clouds are similar. To rigorously define this stability, we introduce the *bottleneck distance* between two persistence diagrams and state the *Stability Theorem for Filtrations*.

Recall that persistence diagrams consist of finitely many points on and above the diagonal representing the birth and death of cycles, and uncountably many points along the diagonal.

Let X and Y be two arbitrary persistence diagrams, not necessarily corresponding to the same filtration nor the same simplicial complex. For example, X and Y may be persistent diagrams for different filtrations of the same simplicial complex K, or they may diagrams for filtrations of different simplicial complexes K_1 and K_2. We will consider bijections $\eta : X \to Y$, and will record the supremum of the distances between corresponding points for each.

It is noteworthy that the mappings $\eta : X \to Y$ are easy to construct for any two persistence diagrams X and Y, as X and Y contain finitely many points off of the diagonal but infinitely many points on the diagonal. Any points off the diagonal of X may be mapped to points on the diagonal of Y, and vice versa.

For example, consider a point $a_X = \left(x_1, x_2\right)$ in X representing the birth and death of a cycle in the filtration for X, as well as the point $a_Y = \left(y_1, y_2\right)$ in Y representing the cycle that η maps a_X to, i.e., $\eta\left(a_X\right) = a_Y$. Specifically, each a_X and a_Y represent the birth and death of a cycle in the filtrations of X and Y respectively. The bijection η creates a correspondence between the cycles whose birth and death are represented by a_X and a_Y, and thus refer to a_Y as the cycle that η maps a_X to. We measure the distance between a_X and a_Y as $\left\|a_X - a_Y\right\|_\infty = \max\left\{\left|x_1 - y_1\right|, \left|x_2 - y_2\right|\right\}$.

Definition 5.9. The **bottleneck distance** between the diagrams X and Y, denoted $W_\infty(X, Y)$ is the infinum over all bijections $\eta : X \to Y$ of the supremum of all such distances, i.e.,

$$W_\infty(X, Y) = \inf_{\eta : X \to Y} \sup_{a_x \in X} \left\|a_X - \eta\left(a_X\right)\right\|_\infty.$$

Example 5.8. Consider the superposition of persistence diagrams X (black) and Y (red) shown below.

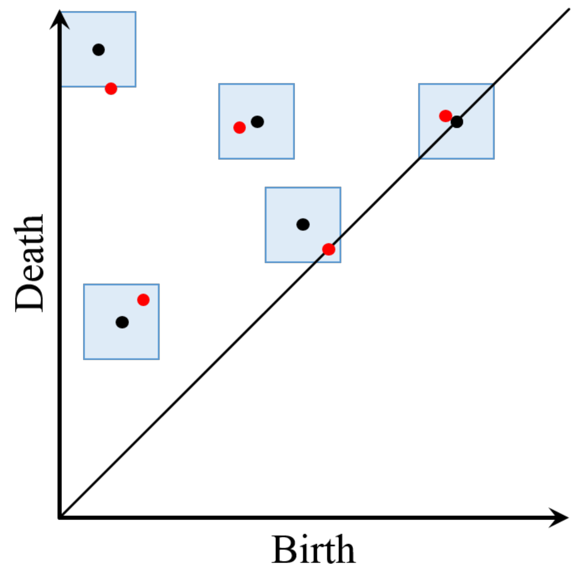


Figure 5.7

Each square has side length twice the bottleneck distance, and are centered at the points of X. Note that they also contain all points of Y. This holds for any two persistence diagrams.

Note also that the bottleneck distance satisfies the same properties as the Euclidean distance, specifically:

(i) It is non-negative for any two persistence diagrams $X,\ Y$, and zero when $X = Y$,

(ii) The bottleneck distance between X and Y is equal to the bottleneck distance between Y and X, and

(iii) It obeys the triangle inequality.

We now lay the groundwork for the Stability Theorem for Filtrations: Let K be a simplicial complex and consider two monotonic functions $f,\ g : K \to \mathbb{R}$. The straight-line homotopy $f_t = (1-t)\ f + t\,g$ is in fact monotonic and yields a persistence diagram for each dimension p and each $t \in [0,\ 1]$.

Fix a dimension p and consider the family of persistence diagrams in $\mathbb{R}^2 \times [0, 1]$. We may obtain a three-dimensional representation of the evolution of the persistent homology from $f_0 = f$ to $f_1 = g$ by drawing t as a third coordinate axis.

In this representation, each off-diagonal point is of the form $x(t) = \left(f_t(\sigma), f_t(\tau), t\right)$, where σ and τ are simplices in K. Note that adding σ in the construction of K represents the birth of a class in the p-dimensional homology group, while adding τ represents its death.

There are finitely many values $t_0, t_1, \ldots, t_n$ of t at which the pairing of simplices changes, and within each interval $\left(t_i, t_{i+1}\right)$ the pairing σ, τ is constant. This gives rise to a line segment of points $x(t)$ between the planes $t = t_i$ and $t = t_{i+1}$. There are two notable behaviors of this line segment:

(i) If the endpoint at t_{i+1} is off the diagonal, then there is another unique line segment that begins at that point. If this second line segment corresponds to the same simplex pair, then it continues on the same straight line. If it does not correspond to the same simplex pair, then it makes a turn relative to the first line segment.

(ii) If the endpoint at t_{i+1} is on the diagonal, then the line segment ends there.

We call the collection of polygonal paths formed by these line segments **vineyards**, and each polygonal path is called a **vine**.

The final ingredient necessary for the Stability Theorem for Filtrations is the L_∞-*distance*, which is a distance measure between two functions, and reminiscent of the definition of the bottleneck distance.

Definition 5.10. The L_∞**-distance** between two functions $f, g : K \to \mathbb{R}$, denoted $\|f - g\|_\infty$, is defined as $\max\{|f(\sigma) - g(\sigma)| \mid \sigma \in K\}$.

Now, let τ be another simplex in K and consider the derivative of each line segment in each vine in the vineyard

$$x(t) = (1 - t)\,(f(\sigma), f(\tau), 0) + t\,(g(\sigma), g(\tau), 1),$$

which is

$$\frac{\partial x}{\partial t} = (g(\sigma) - f(\sigma),\ g(\tau) - f(\tau),\ 1) = \left(\frac{\partial x}{\partial t}[\sigma],\ \frac{\partial x}{\partial t}[\tau],\ 1\right),$$

where we have used $\frac{\partial x}{\partial t}[\sigma]$ as short-hand for $g(\sigma) - f(\sigma)$ and similarly for τ.

By projecting the endpoints of this line segment back into $\mathbb{R}^2$, we obtain two points

$$\begin{aligned} x_i &= \left(1 - t_i\right)(f(\sigma),\ f(\tau)) + t_i\ (g(\sigma),\ g(\tau)) &&= \left(x_{i,\,1},\ x_{i,\,2}\right) \\ x_{i+1} &= \left(1 - t_{i+1}\right)(f(\sigma),\ f(\tau)) + t_{i+1}\ (g(\sigma),\ g(\tau)) &&= \left(x_{i+1,\,1},\ x_{i+1,\,2}\right) \end{aligned}$$

such that $\left\|x_i - x_{i+1}\right\|_\infty$ is $t_{i+1} - t_i$ times greater than the larger of the differences between f and g at σ and τ. Specifically,

$$\begin{aligned} \left\|x_i - x_{i+1}\right\|_\infty &= \max\left\{\left|x_{i,\,1} - x_{i+1,\,1}\right|,\ \left|x_{i,\,2} - x_{i+1,\,2}\right|\right\} \\ &= \left(t_{i+1} - t_i\right)\ \max\left\{\frac{\partial x}{\partial t}[\sigma],\ \frac{\partial x}{\partial t}[\tau]\right\}. \end{aligned}$$

Let v be the simplex in K such that $\|f - g\|_\infty = |f(v) - g(v)|$. This serves as an upper bound for the slope of any line segment in the vineyard when projected into $\mathbb{R}^2$, and thus bounds the distance $\left\|x_i - x_{i+1}\right\|_\infty$ between the projected endpoints of any vine. We may now write the Stability Theorem for Filtrations:

Theorem 5.11. STABILITY THEOREM FOR FILTRATIONS. Let K be a simplicial complex and $f,\ g:\ K \to \mathbb{R}$ be two monotonic functions. For each dimension p, the bottleneck distance between the diagrams $X = \mathrm{Dgm}_p(f)$ and $Y = \mathrm{Dgm}_p(g)$ is bounded from above by the L_∞-distance between f and g, i.e.,

$$W_\infty(X,\ Y) \leq \|f - g\|_\infty.$$

Note that there exists an analogous theorem regarding the stability of barcodes.

Chapter 6: Persistent Homology of 2-D Configurations of BuckyBalls®

In this chapter, we will use persistent homology and barcodes to analyze configurations of BuckyBalls®. This system serves as an excellent example of how one utilizes the ideas we have introduced throughout this thesis in a real-world application.

We begin in Section 6.1 by introducing the basic physical principles which guide BuckyBall® interactions.

In Section 6.2, we introduce two representations for configurations as point clouds: the $\mathbb{R}^2$ *representation* and the *potential representation*.

We further explore these representations in Section 6.3 by analyzing symmetric arrangements of BuckyBalls®.

Section 6.1: Introduction to BuckyBalls®

BuckyBalls® are small spherical neodymium magnets often used as desk toys. As discussed by Mann and Monello (henceforth M & M) in [7], they arrange themselves into a wide variety of minimum-energy configurations, as shown in Figure 6.1.

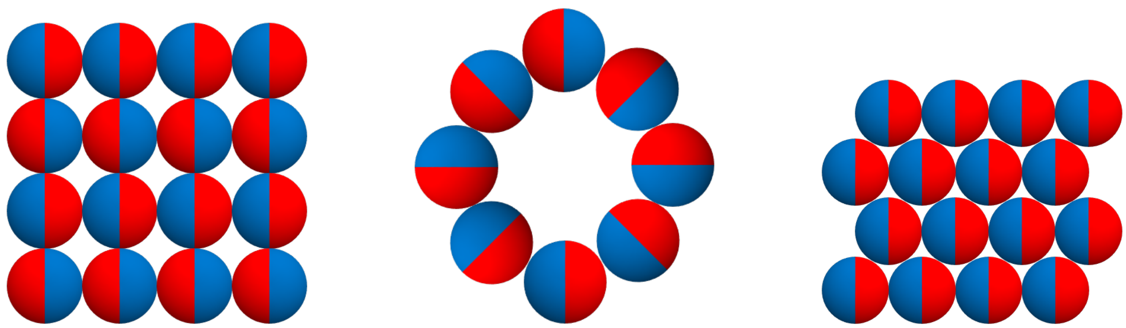


Figure 6.1: *Stable symmetric configurations of BuckyBalls®, whose magnetic poles have been colored for added clarity. A square lattice (left), a ring (middle), and a hexagonal lattice (right).*

The most prominent question that M & M pose is on the two-dimensional arrangements that these balls tend to form on their own. To answer this, they first studied the energetic favorability of highly symmetric configurations. From there, they created computer simulations of random arrangements of balls settling into stable arrangements, and used statistical analysis on the final structures.

We instead shall use persistent homology to study these structures, and extract their homological features.

Magnetic Dipoles of BuckyBalls®

In [7], M & M assumed each ball to be uniformly magnetized, with dipole moment **m**. and assigned to each four coordinates to describe their position and orientation in space: x and y Cartesian coordinates, polar angle θ, and azimuthal angle ϕ. These angular coordinates correspond to the direction of the ball's dipole moment, as shown in Figure 6.2.

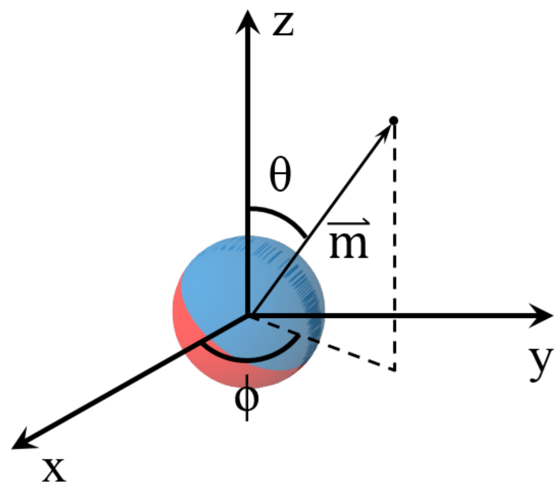


Figure 6.2: *A BuckyBall® with labeled coordinates and dipole moment.*

Following non-dimensionalization, each ball is of radius 0.5 and the potential between two BuckyBalls® is given by

$$\Phi_{12} = \frac{2}{{r_{12}}^3} \left(\hat{\mathbf{m}}_1 \cdot \hat{\mathbf{m}}_2 - 3 \left(\hat{\mathbf{m}}_1 \cdot \hat{\mathbf{r}}_{12}\right) \left(\hat{\mathbf{m}}_2 \cdot \hat{\mathbf{r}}_{12}\right)\right), \tag{6.1}$$

where $\mathbf{r}_{12}$ is the displacement vector between the two balls, and $\mathbf{m}_1$ and $\mathbf{m}_2$ are their respective dipole moments. We may write the direction of these dipole moments as

$$\hat{\mathbf{m}}_i = \sin\theta_i \, \cos\phi_i \, \hat{\mathbf{x}} + \sin\theta_i \, \sin\phi_i \, \hat{\mathbf{y}} + \cos\theta_i \, \hat{\mathbf{z}}.$$

Section 6.2: Representing BuckyBall® Arrangements as Point Clouds

By using persistent homology, we wish to gain insight into the arrangement of a system of BuckyBalls®, whether they have reached some energy minimum, and the geometry of that energy minimum. In order to use persistent homology, we represent the BuckyBalls® as a point cloud, and construct a filtration using Vietoris-Rips complexes. However, we have come to an impasse: how do we represent arrangements of BuckyBalls® as a point cloud?

The $\mathbb{R}^2$ Representation

A preliminary answer would be to represent each ball by its Cartesian coordinates in $\mathbb{R}^2$ and measure distance between them with the Euclidean distance, which we aptly refer to as the $\mathbb{R}^2$ **representation**. Indeed, this does capture many of the geometric aspects of an arrangement of balls: gaps and holes, as well as the balls' relative proximity. Consider the arrangement in Figure 6.3 and its associated barcode in Figure 6.4 using this representation, obtained using JavaPlex [2].

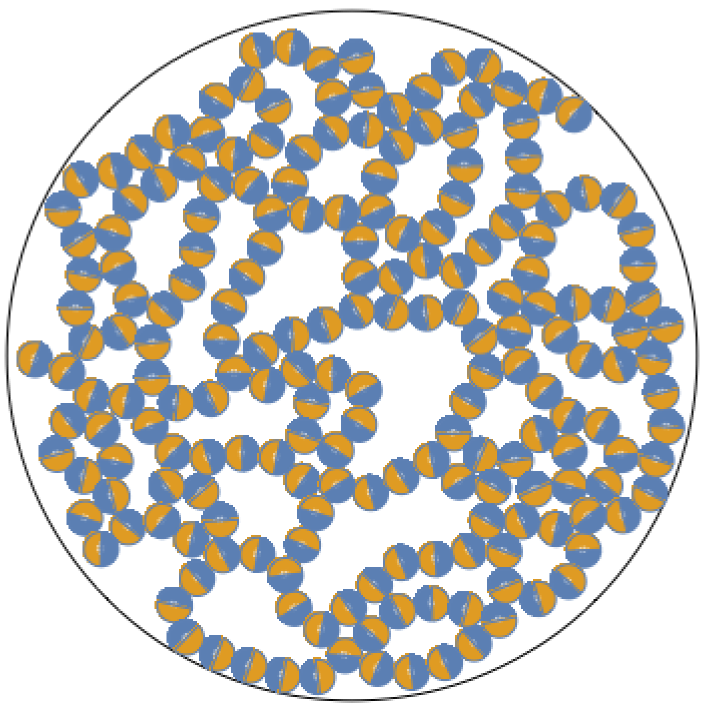


Figure 6.3: *An energy-minimized 200 ball arrangement.*

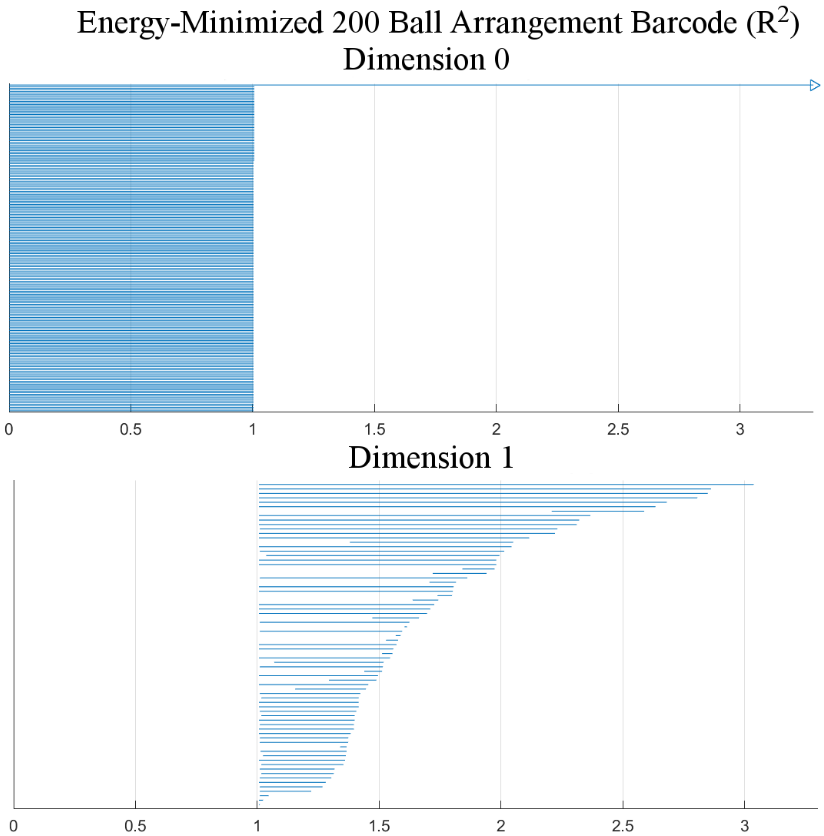


Figure 6.4: *The barcode of an energy-minimized 200 ball arrangement shown in Figure 6.3 using the* $\mathbb{R}^2$ *representation of the system. We have omitted higher-dimensional barcodes, as dimensions* 2 *through* 5 *were found to be trivial.*

This indeed captures an amount of the geometry of the arrangement, as expected. For instance, all but one of the holes in dimension 0 die at 1, which implies that the center of each ball is a distance 1 away from each other. Thus, as the radius of each ball is 0.5, each ball must be touching at least one other ball. The interested reader can also find which bar in the 1 dimensional barcode represent which hole in the arrangement.

However, the $\mathbb{R}^2$ representation fails to capture much of any information regarding the potential of the arrangement. For instance, consider the following two arrangements of balls:

Figure 6.5

The arrangement of balls on the left in Figure 6.5 would have the least possible potential energy in the system, while the arrangement on the right would have the most. The behavior of these arrangements would be radically different, but the $\mathbb{R}^2$ representation fails to capture this.

The Potential Representation

Before we are able to discuss the *potential representation*, we must first discuss an alternative way to list data to construct Vietoris-Rips complexes: a *distance matrix*.

Consider a finite set of points $\{x_1, x_2, \ldots, x_n\}$. A **distance matrix** for this set of points is of the form

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$

where r_{ij} is the distance between x_i and x_j.

In JavaPlex, such a distance matrix need not be symmetric, square, non-negative, nor satisfy the triangle inequality. The first three of these cause computational errors in JavaPlex [2], while failing to satisfy the triangle inequality does not lead to any issues when constructing a Vietoris-Rips complex.

Recall Equation 6.1 which gives the potential energy between any two dipoles, and Figure 6.5 which illustrates the least and greatest potential interactions, with dimensional potential of -4 and 4, respectively.

Therefore, by adding 4 to the potential of each interaction, we may define a “distance” between any two balls with which we may build a Vietoris-Rips complex. This notion of distance yields the **potential representation** of the arrangement, and we denote this notion of distance as the **potential distance**.

Unlike the $\mathbb{R}^2$ representation, the potential representation is far less intuitive. For this reason, we investigate this representation using highly symmetric arrangements.

Section 6.3: Homology of Symmetric Arrangements

N-Ball Chains

The standard N-ball chain consists of N BuckyBalls® arranged North-South in the same manner as those on the left in Figure 6.5. In the $\mathbb{R}^2$ representation, the N-ball chain has trivial homology in all dimensions besides 0, where its barcode resembles that in Figure 6.4.

To analyze this arrangement in the potential representation, consider the potential distance between any two adjacent balls B_i and B_{i+1}. The non-dimensionalized potential energy between these two is -4, so their potential distance is 0. As this holds for all pairs of adjacent balls in the N-chain, the N-chain must be homologous to a point, i.e., has trivial homology in all dimensions greater than 0 and a single persistent homology class in dimension 0 for all potential distances.

We now consider the case when the magnetic dipoles are not aligned properly, e.g., $\phi = \pi/2$ for all balls in an N-chain on the x-axis. In the $\mathbb{R}^2$ representation, such an arrangement has the same persistent homology. In the potential representation, however, we obtain nontrivial homology in dimensions greater than 0 and different persistent homology in dimension 0.

Consider the arrangement shown below, which we shall call the **perpendicular chain**.

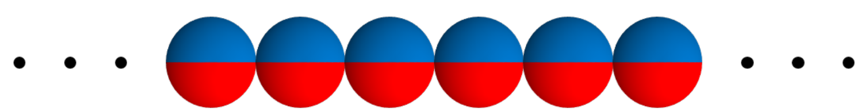
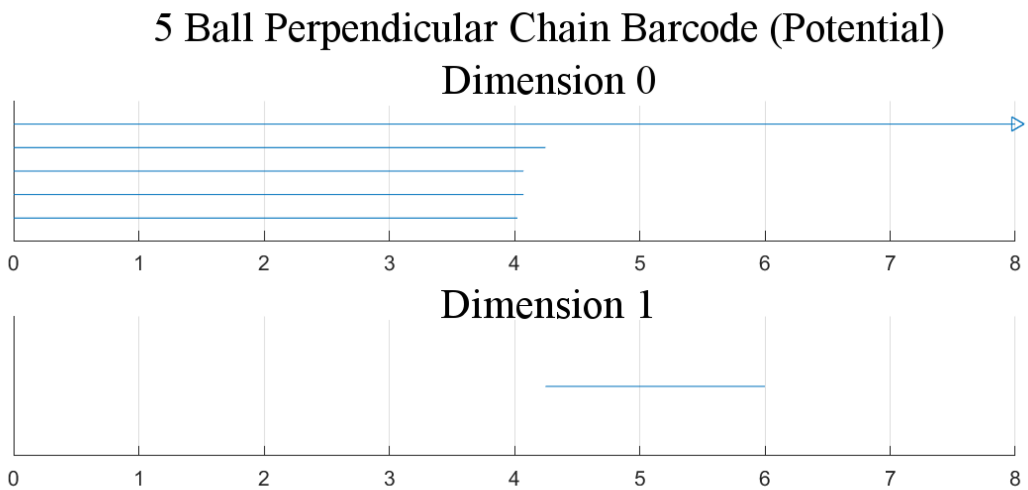
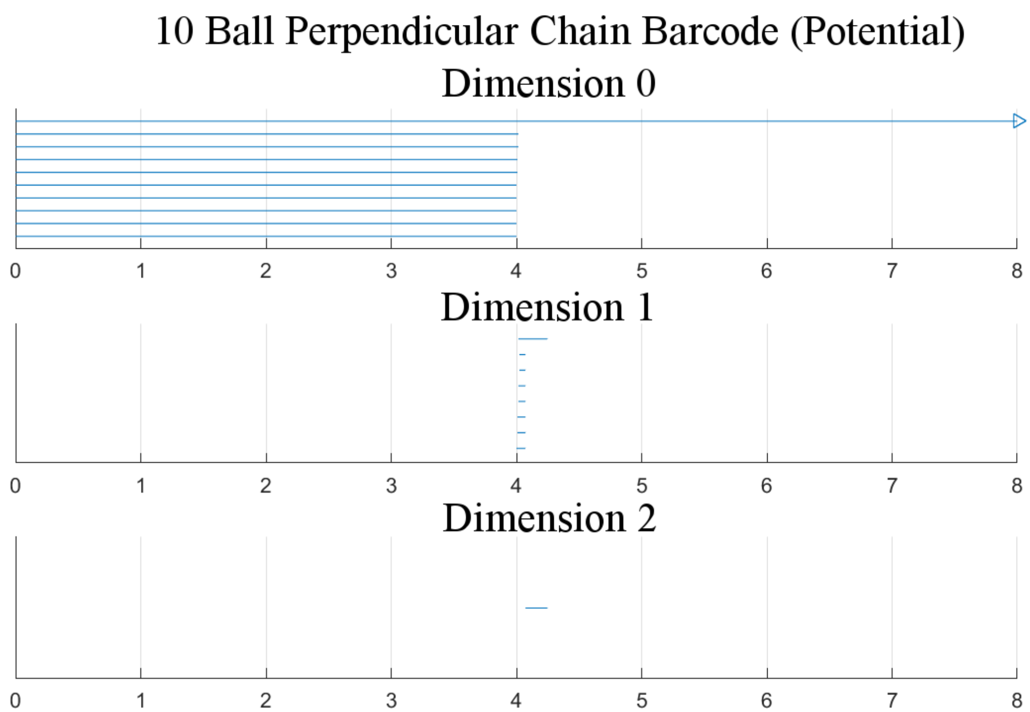
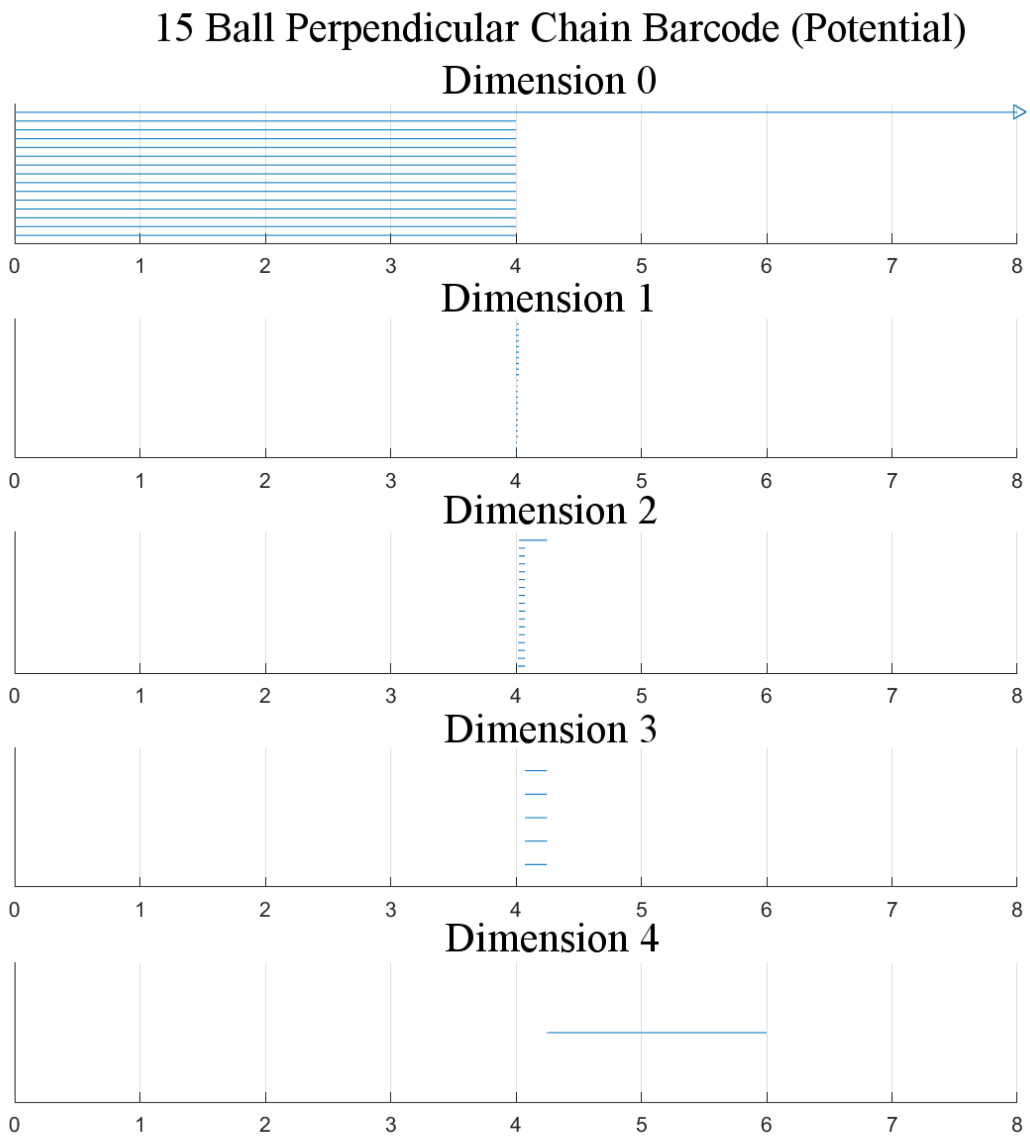


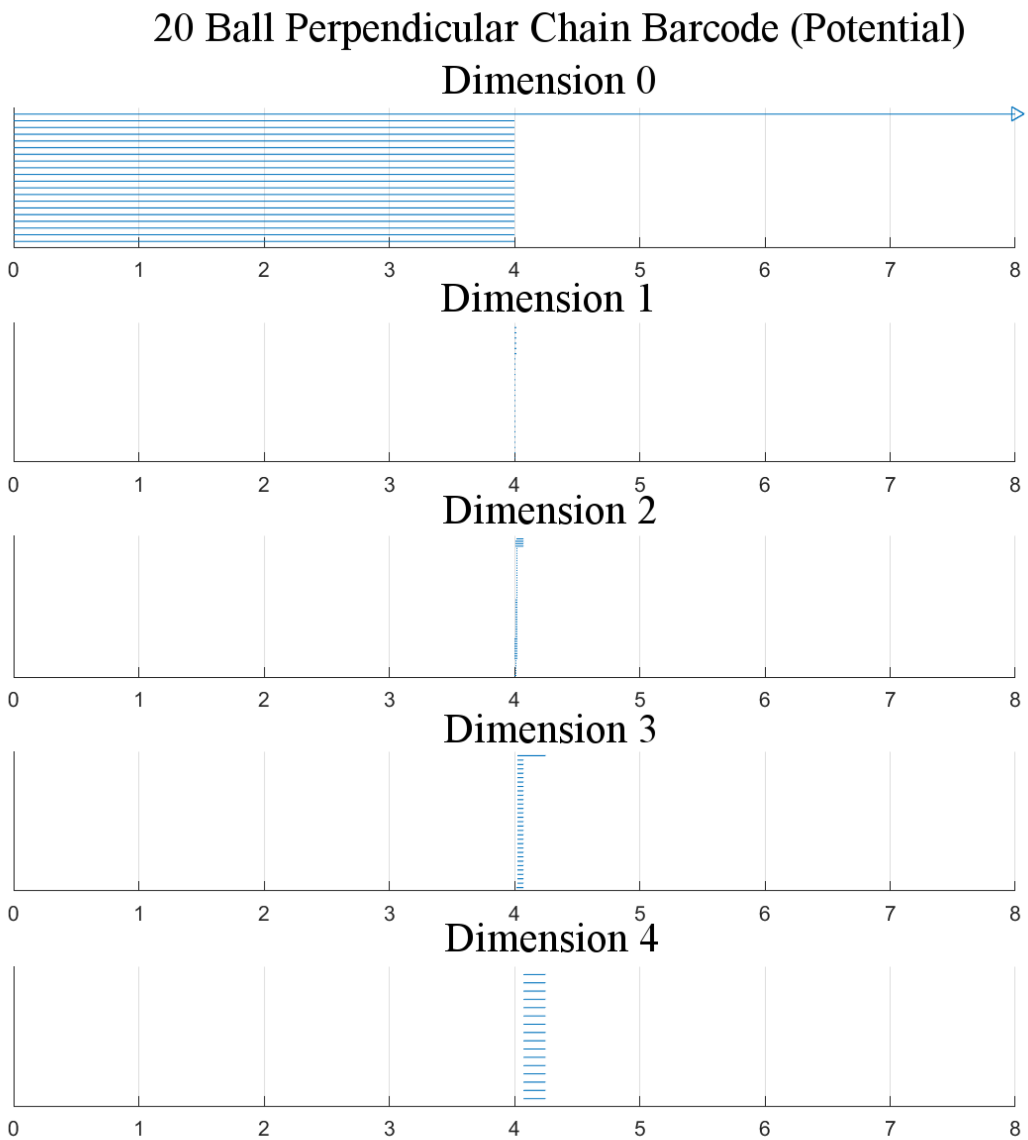
Figure 6.6

The homology of this structure depends on the the number of balls, as evident in the barcodes shown below. Note, the dimensions for each arrangement in which homology is trivial have been excluded.


5 Ball Perpendicular Chain Barcode (Potential)
Dimension 0
0 1 2 3 4 5 6 7 8
Dimension 1
0 1 2 3 4 5 6 7 8


10 Ball Perpendicular Chain Barcode (Potential)
Dimension 0
0 1 2 3 4 5 6 7 8
Dimension 1
0 1 2 3 4 5 6 7 8
Dimension 2
0 1 2 3 4 5 6 7 8


15 Ball Perpendicular Chain Barcode (Potential)
Dimension 0
Dimension 1
Dimension 2
Dimension 3
Dimension 4
0
1
2
3
4
5
6
7
8



We note the following patterns in the barcodes as the number of balls increases:

(i) In dimension 0, all but one interval die at a potential distance of 4. Note that this value corresponds to balls separated by some large distance, i.e., weaker interactions, as would be expected by increasing the number of balls in the perpendicular chain. We note that as we add balls to the system, the endpoint of each finite bar approaches 4 from above. This corresponds to the most attractive interaction that each ball experiences.

(ii) In all dimensions greater than 0, the finite bars approach a length of 0 at potential distance

4. In addition, with the exception of one or two finite bars in each dimension, all bars are of similar length (specifically within 0.15).

From these observations, we conjecture the meaning of bars in each dimension:

(i) The length of each finite bar in dimension 0 corresponds to the minimal potential distance interaction a ball experiences in the system, i.e., its most attractive interaction.

(ii) Finite bars in all dimensions n greater than zero correspond to minimal potential interactions between combinations of $4n$ balls. For example, the finite bar in dimension 1 of the 5-ball perpendicular chain corresponds to the interaction between balls $1,\ 2,\ 4$, and 5.

Furthermore, we believe that the number of bars is related to the number of symmetries in the system. For example, there are only 3 distinguishable balls in the 5-ball perpendicular chain, as 1 and 5 as well as 2 and 4 cannot be distinguished by their interactions with other balls in the system.

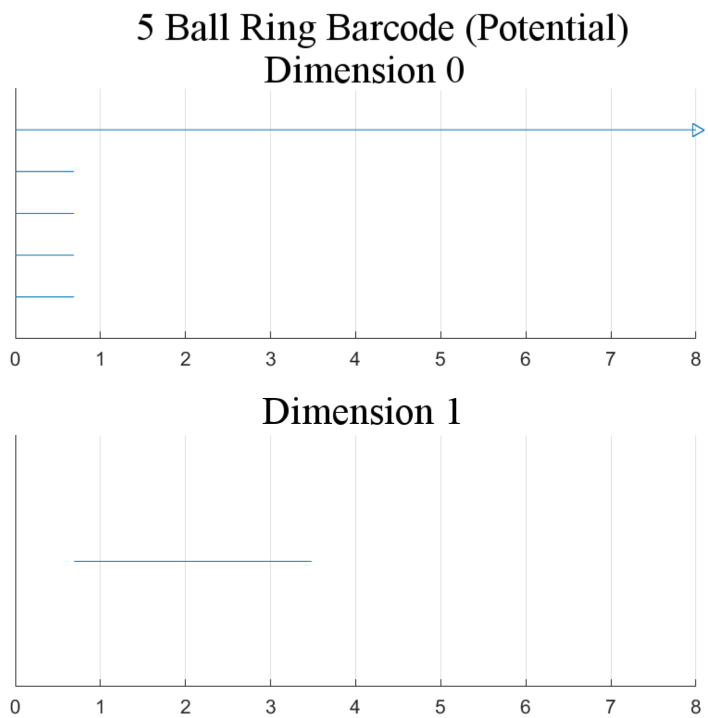
N-Ball Rings

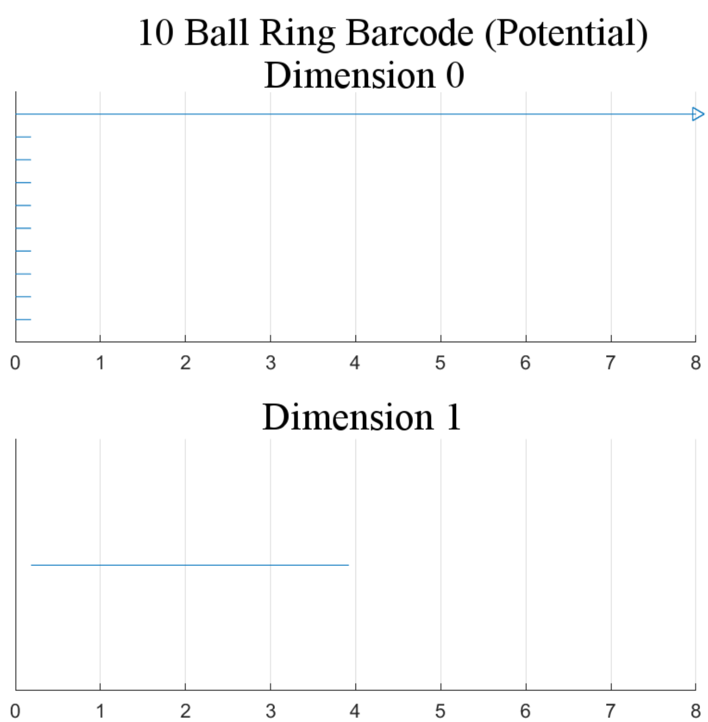
The standard N-ball ring consists of N BuckyBalls® arranged in a ring facing North-South, as shown in the middle image of Figure 6.1. The position and dipole moment of ball i is given by

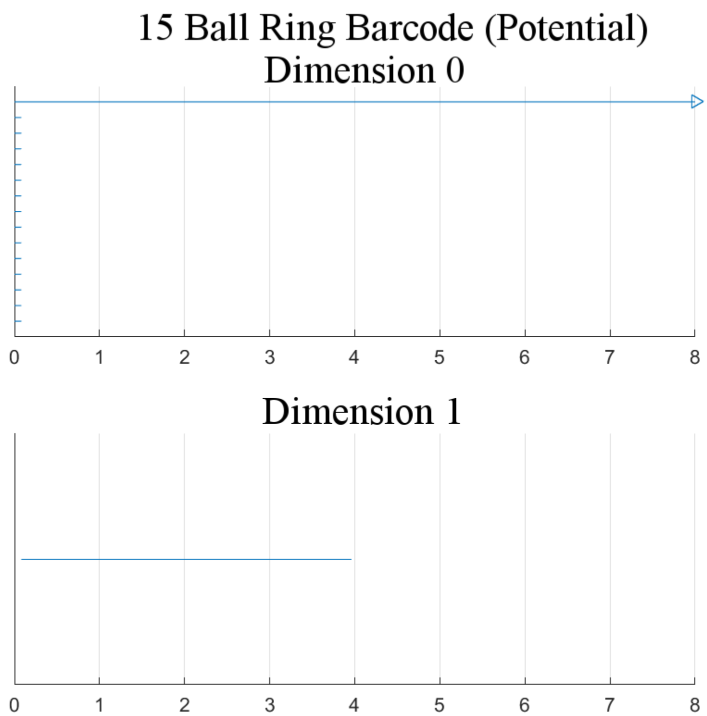
$$\mathbf{r}_i = \frac{1}{2\,\sin(\pi/N)}\left[\cos\left(\frac{2\,\pi\,i}{N}\right)\hat{\mathbf{x}} + \sin\left(\frac{2\,\pi\,i}{N}\right)\hat{\mathbf{y}}\right],$$

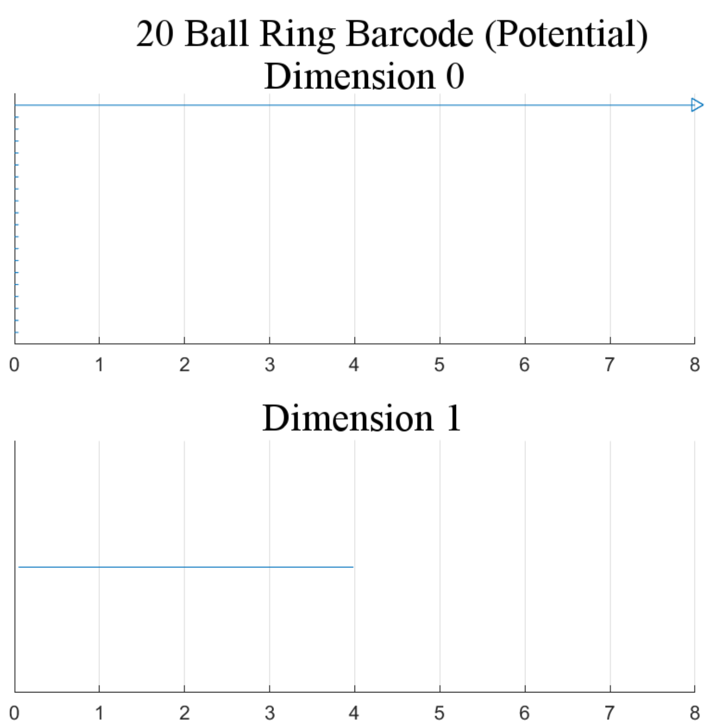
$$\hat{\mathbf{m}}_i = -\sin\left(\frac{2\,\pi\,i}{N}\right)\hat{\mathbf{x}} + \cos\left(\frac{2\,\pi\,i}{N}\right)\hat{\mathbf{y}}. \tag{6.2}$$

In the $\mathbb{R}^2$ representation, the homology of an N-ball ring is the same of that of S^1 with diameter $1/\sin(\pi/N)$. In the potential representation, we find similar homology. Note, the dimensions for each arrangement in which homology is trivial have been excluded.


5 Ball Ring Barcode (Potential)
Dimension 0
0 1 2 3 4 5 6 7 8
Dimension 1
0 1 2 3 4 5 6 7 8


10 Ball Ring Barcode (Potential)
Dimension 0
0 1 2 3 4 5 6 7 8
Dimension 1
0 1 2 3 4 5 6 7 8





In fact, as the number of balls increases, we find that the persistent homology approaches that of a circle with diameter 4, which is the potential distance between antipodal balls in the ring.

Note that as the number of balls increases, the N-ball ring locally approaches the N-ball chain. This is reflected in the 0-dimensional homology, which becomes trivial as the number of balls increases.

In addition, recall that we conjecture that the number of bars in higher dimensions is related to the number of symmetries in the system. In the N-ball ring, there is only one distinguishable ball, implying that the system is highly symmetric.

Chapter 7: Bibliography

[1] C. Adams and R. Franzosa, *Introduction to Topology: Pure and Applied*, Pearson Education, Inc., 2007.

[2] H. Adams, A. Tausz, and M. Vejdemo-Johansson, *JavaPlex: A research software package for persistent (co)homology*, Proceedings of ICMS 2014 (Han Hong and Chee Yap, eds.), Lecture Notes in Computer Science 8592, 2014, pp. 129–136.

[3] B. Di Fabio and C. Landi, *A mayer–vietoris formula for persistent homology with an application to shape recognition in the presence of occlusions*, Foundations of Computational Mathematics **11** (2011), no. 5, 499.

[4] H. Edelsbrunner and J. Harer, *Computational Topology: An Introduction*, American Mathematical Society, 2010.

[5] I. Evenor, *Homological Algebra: Five and Snake Lemma*, accesible from https://www.tau.ac.il/ itzhakevenor/homological-algebra.pdf, August 2014.

[6] J. Lamar-León, E. B. García-Reyes, and R. Gonzalez-Diaz, *Human gait identification using persistent homology*, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (Berlin, Heidelberg), Springer Berlin Heidelberg, 2012, pp. 244–251.

[7] N. Mann and J. A. Monello, *Two-Dimensional Stable Configurations of BuckyBalls®*, contact mannn@union.edu for the document.

[8] M. C. McCord, *Homotopy type comparison of a space with complexes associated with its open covers*, Proceedings of the American Mathematical Society, vol. 18, Lecture Notes in Computer Science 8592, no. 4, August 1967, pp. 705 – 708.

[9] J. R. Munkres, *Elements of Algebraic Topology*, Benjamin/Cummings Publishing Company, Inc., Redwood City, CA, 1984.

[10] M. N. Nasser, *Pattern recognition and machine learning*, Journal of Electronic Imaging **16** (2007).

[11] A. Zomorodian and G. Carlsson, *Computing persistent homology*, Discrete & Computational Geometry **33** (2005), no. 2, 249–274.