Simplicial homology and its applications

Rubén Baeza García
Trabajo de fin de grado de Matemáticas
Universidad de Zaragoza

Director del trabajo: Miguel Ángel Marco Buzunariz
27 de junio de 2022

Resumen

La topología algebraica introduce nociones de álgebra para simplificar problemas de topología tales como saber si dos espacios topológicos pueden ser o no ser homeomorfos. En palabras de Lescheftz [7], la teoría de homología es la herramienta de mayor utilidad jamás creada en topología. Combina invariantes topológicos como el número de componentes conexas o el primer grupo fundamental abelianizado. Como el título indica, la homología simplicial y sus aplicaciones tanto teóricas como aplicadas, van a ser las protagonistas en los siguientes capítulos.

En el primero, introduciremos el *complejo simplicial*, $|K|$, una estructura formada por *símplices*, capaz de discretizar un subespacio topológico triangulable del espacio euclídeo. Su realización topológica, llamada *poliedro*, $|K|$, es homeomorfa al original. Cabe destacar la existencia de las *aplicaciones simpliciales* entre complejos simpliciales. Relacionan símplices de manera lineal y un resultado conocido como *Teorema de aproximacion simplicial* permite, valga la redundancia, aproximar aplicaciones continuas entre dos complejos $f\colon |K| \to |L|$, por medio de una aplicación simplicial cuyo dominio es un complejo más “fragmentado”, $s\colon |K^m| \to |L|$. Esta propiedad resultará clave para demostrar la invarianza bajo homotopía de los grupos de homología.

En el segundo capítulo se introduce la homología simplicial. Asociaremos una colección de grupos abelianos a un complejo simplicial K. El grupo de q-cadenas, $C_q(K)$, es el grupo libre cuyo conjunto generador es el conjunto de símplices *orientados* de dimensión q. Definiremos las *aplicaciones borde*, $\partial_q\colon C_q(K) \to C_{q-1}(K)$, que permiten construir una estructura algebraica de *complejo de cadenas*, $C(K)$. Usando el Teorema 2.1, asociaremos a cada complejo de cadenas sus grupos de homología. Se mostrará el cálculo de la homología para algunos complejos simpliciales y de este modo nos familiarizaremos con las partes libres y de torsión.
En las siguientes secciones, el objetivo será demostrar la invarianza bajo homotopía para poder afirmar que los grupos de homología son invariantes topológicos. Haremos uso de diagramas conmutativos y herramientas de álgebra homológica que relacionaremos con nuestro particular contexto.

Estos capítulos siguen como fuente principal [1], salvo en la Sección 2.4 que se usa [10] dado que evita introducir el concepto de *subdivisión estelar*. El resto de materiales se han mencionado para suplementar con visiones alternativas de la teoría o para ofrecer recursos complementarios al lector. Debemos destacar que en ambos capítulos hay ciertos resultados más cercanos al área de análisis que hemos decidido referenciar sus demostraciones en [1] por falta de espacio, pero que sin embargo son resultados clave en homología simplicial.

El último capítulo está dedicado a la *homología persistente*. Dada una nube de datos, es decir, un conjunto de puntos n-dimensionales, construiremos una *filtración*. Esta es una sucesión creciente de complejos simpliciales que se construyen siguiendo una regla de distancia euclídea. Aquí presentaremos dos: el complejo de Čech y el de Rips. Dada una filtración, consideraremos su *complejo de persistencia*, integrado por los complejos de cadenas asociados a cada parámetro

y relacionados por aplicaciones identidad. Por último, calcularemos la *homología persistente* del complejo de persistencia y mostraremos los resultados en lo que se conoce como *código de barras*. La justificación rigurosa de los *grupos de homología persistente* se basa en la construcción de un módulo graduado que garantiza la existencia de bases compatibles en los distintos grupos de homología de la filtración con el mismo orden. Esto se consigue flexibilizando la definición de grupo de cadenas a R-módulos, donde R es un DIP.

Para concluir quiero mencionar que el uso de los colores en este trabajo no tiene una función decorativa, sino que pretende guiar al lector en las demostraciones y facilitar su comprensión.

Contents

Chapter 1

Simplicial Complexes

The foundations of Simplicial Homology

Our first step is to replace complex topological spaces with simpler constructions called *simplicial complexes*. This procedure can be viewed as a discretization of infinite euclidean spaces since we will no longer work with the infinite set of points. Instead we will identify a finite set of points and work with "construction blocks" formed by subsets of it: simplices.

1.1 Simplices and simplicial complexes

Definition (Points in general position).
Let $\mathbb{E}^n$ denote the euclidean space $\mathbb{R}^n$ with the usual topology and let $v_0, v_1, \dots, v_k \in \mathbb{E}^n$, the *hyperplane spanned by k points* consists of all linear combinations $\lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_k v_k$ where all $\lambda_i \in \mathbb{E}$ and $\sum_{i=0}^{k} \lambda_i = 1$.
The points are in *general position* if any subset of them spans a strictly smaller hyperplane, or equivalently, $v_1 - v_0, v_2 - v_0, \dots, v_k - v_0$ are linearly independent.

Definition (k-simplex).
Given $k+1$ points $v_0, v_1, \dots, v_k \in \mathbb{E}^n$ in general position, we call the smallest convex set containing them a *simplex of dimension* k, or a k*-simplex* and we denote it by $(v_0, v_1, \dots, v_k)$. The points $v_0, v_1, \dots, v_k$ are called the *vertices of the simplex.*
Recall that x lies in the smallest convex set containing $v_0, v_1, \dots, v_k$ if and only if $x = \lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_k v_k$ where all $\lambda_i \geq 0$ and $\sum_{i=0}^{k} \lambda_i = 1$.

With this geometrically-intuitive definition we can visualize low-dimensional simplices:

k	Interpretation of a k-simplex
0	Point
1	Segment
2	Triangle
3	Tetrahedron

Definition (Face of a simplex).
If A and B are simplices and if the vertices of B form a subset of the vertices of A, then we say that B is a *face* of A and denote it by $B < A$.

Definition (Simplicial complex, subcomplexes and dimension).
A finite collection of simplices K in some euclidean space $\mathbb{E}^n$ is called a *simplicial complex* if whenever a simplex lies in the collection then so does each of its faces, and whenever two simplexes of the collection intersect they do so in a common face. It is natural to define a *subcomplex* of K as the subcollection of simplices which itself forms a simplicial complex, as well as defining the *dimension* of a simplicial complex to be the maximum of the dimensions of its simplices.

This definition illustrates a set of simplices where intersections can be thought of as glueing k-simplices with k-simplices. Below we find some drawn examples of sets of simplices which intersect in both ways.

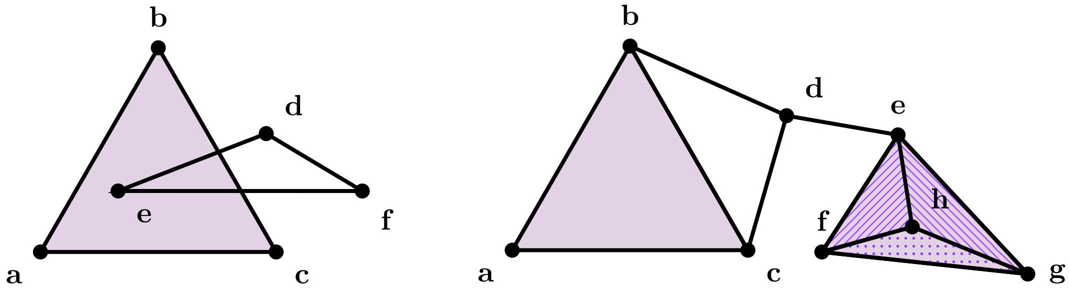


Figure 1.1: The first case is not a simplicial complex, but so the second is. Indeed it contains 8, 0-simplices; 12, 1-simplices; 5, 2-simplices and 1, 3-simplex.

Definition (Closure of a simplex).
The *closure* of a k-simplex σ, $\mathrm{Cl}(\sigma)$, is the complex consisting of σ and all its faces.

The terminology is accurate as $\mathrm{Cl}(\sigma)$ is the smallest simplicial complex containing σ. Intuitively, one can think of the closure of a 3-simplex as the result of joining every vertex of the closure 2-simplex, with a new vertex (in general position with the other three) acting as the apex of a tetrahedron. Then adding the lateral faces, new edges and the new vertex. A similar construction can be considered if we replace the closure of a 2-simplex by any simplicial complex to act as the base the *cone*.

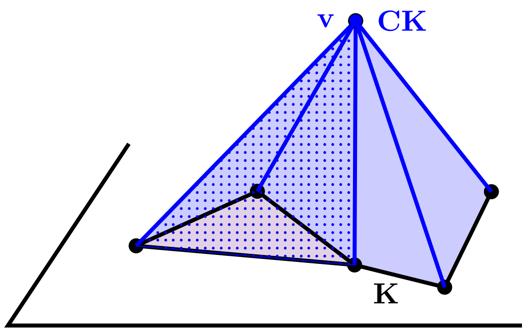


Figure 1.2: Starting from the simplicial complex K, we add the simplices colored in blue and the 3-simplex to obtain the cone CK.

Definition (Cone on a simplicial complex).
Let K be a simplicial complex in $\mathbb{E}^n$ and think of its inclusion in $\mathbb{E}^{n+1}$ as points whose last coordinate is 0. We shall now construct a new simplicial complex in $\mathbb{E}^{n+1}$ called the *cone on* K and denoted by CK.
Let $v = (0, \dots, 0, 1) \in \mathbb{E}^{n+1}$. If A is a k-simplex of K, with vertices $v_0, v_1, \dots, v_k$, the points $v_0, v_1, \dots, v_k, v$ are in general position so we get a $(k+1)$-simplex in $\mathbb{E}^{n+1}$. We call it the *join of* A *to* v. Our cone CK consists of the initial simplexes of K, the join of each of these simplices to v, and the 0-simplex v itself. We leave it to the reader to verify that the resulting set of simplices is a simplicial complex.

So far we have defined simplicial complexes in $\mathbb{E}^n$, but little have we stated about the desired topological properties. The definition below gives a topological meaning to these objects.

Definition (Polyhedron).
The union of simplexes which make up a simplicial complex is itself a subset of the euclidean space and can therefore be made into a topological space by giving it the subspace topology. This topological space is called the *polyhedron of* K and it is denoted by $|K|$.

The concept of polyhedron is key if we want to relate simplicial complexes to the topological spaces they are built from by, somehow, "triangulating" the latter.

Definition (Triangulation of a topological space).
A *triangulation of a topological space* X consists of a simplicial complex K and a homeomorphism $h\colon |K| \to X$.

For example, let K be a triangulation of the torus. It is a simplicial complex made up of 9, 0-simplices; 27, 1-simplices and 18, 2-simplices.

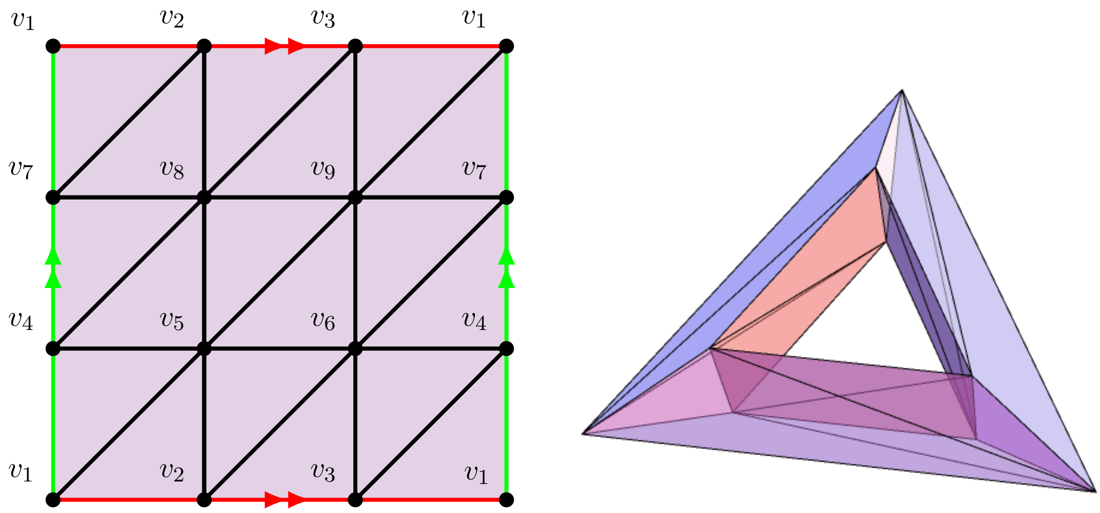


Figure 1.3: A triangulation of the torus, K, with its embedding in $\mathbb{E}^3$. Image credit to [2].

Despite having forgotten all the geometric properties from simplices such as size, we will be quite interested in subdivision processes of simplices. Thereby we shall introduce an algorithm to obtain similar simplicial complexes with more simplices but smaller in diameter.

Definition (Barycentre and barycentric coordinates).
If A is a simplex with vertices $v_0, v_1, \ldots, v_k$ then each point $x \in A$ has a unique expression of the form $x = \lambda_0 v_0 + \lambda_1 v_1 + \cdots + \lambda_k v_k$. We call $\{\lambda_i \mid i = 0, 1, \ldots, k\}$ the *barycentric coordinates* of x and the *barycentre* of the simplex A is the point $\hat{A} = \dfrac{1}{k+1}(v_0 + v_1 + \cdots + v_k)$.

Thus, the barycentre of a 0-simplex is the vertex itself, the middle point in the case of a 1-simplex and so on.

Definition (Barycentric subdivision).
Given a simplicial complex in $\mathbb{E}^n$, K, we want to create a new simplicial complex K^1 such that $|K^1| = |K|$ but whose simplexes have a smaller diametre.
The vertices of K^1 are precisely the barycentres of the simplexes of K, the vertices of K are included since they are the barycentre of their own 0-simplex.
A collection of $k+1$ such barycentres $\dot{\sigma}_0, \dots, \dot{\sigma}_k$ of $\sigma_0, \dots, \sigma_k$ simplices of K, are the vertices of a k-simplex in K^1 if $\sigma_{\tau(0)} < \sigma_{\tau(1)} < \cdots < \sigma_{\tau(k)}$ for some permutation τ of the integers $0, 1, \dots, k$.
Lastly we should check that the vertices defining simplices are in general position. It stems directly from the algorithm since $\dot{\sigma}_{\tau(i)}$ lies off the hyperplane spanned by $\dot{\sigma}_{\tau(0)}, \dots, \dot{\sigma}_{\tau(i-1)}$.
We define the m-th barycentric subdivision K^m of K inductively as the barycentric subdivision of the $(m-1)$-th barycentric subdivision, i.e. $K^m = (K^{m-1})^1$.

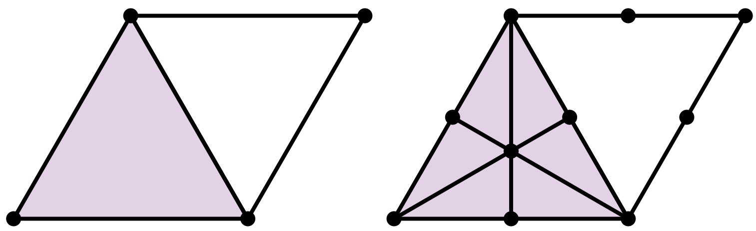


Figure 1.4: Performing the first barycentric subdivision on a simplicial complex made up of 4, 0-simplices; 5, 1-simplices and 1, 2-simplex.

As a consequence from the definition, we are replacing every q-simplex σ with a cone $C\tilde{\sigma}$ with apex $\dot{\sigma}$, where $\tilde{\sigma}$ is the simplicial complex containing the barycentric subdivisions of every $(q-1)$-simplex in $\mathrm{Cl}(\sigma)$. No two vertices of K form a 1-simplex in any of its subdivisions. In the next section, we will be required to replace a simplicial complex by another one with the same polyhedra but suficiently small simplices. In that context, the barycentric subdivision procedure will be a meaningful tool.

1.2 Maps between simplicial complexes

Given two topological spaces and their triangulations $r\colon |K| \to X$ and $s\colon |L| \to Y$. Any map $f\colon X \to Y$ induces a map $s^{-1}fr\colon |K| \to |L|$. However, we would be interested in maps that maintain properties from one simplicial complex into the other, the same way group homomorphisms do with the respective group operations.

Definition (Interior of a simplex and carrier of a point).
Let A be a simplex in $\mathbb{E}^n$ with vertices $v_0, v_1, \dots, v_k$. We define the *interior of* A to consist of those points $x \in A$ such that $x = \lambda_0 v_0 + \lambda_1 v_1 + \cdots + \lambda_k v_k$ where all $\lambda_i > 0$ and $\sum_{i=0}^{k} \lambda_i = 1$.
In addition, it is clear that any point of $x \in |K|$ lies in the interior of exactly one simplex of K which we will call the *carrier of* x. The latter is the smallest simplex of K containing x.

Definition (Simplicial map).
Let K and L be simplicial complexes. A map $s\colon |K| \to |L|$ is called *simplicial* if s takes simplices of K linearly onto simplices of L.

The condition of linearity implies that if A is a simplex of K with vertices $v_0, v_1, \ldots, v_k$ and $x = \lambda_0 v_0 + \lambda_1 v_1 + \cdots + \lambda_k v_k$ then $s(x) = \lambda_0 s(v_0) + \lambda_1 s(v_1) + \cdots + \lambda_k s(v_k)$. Hence, a simplicial map is completely determined by the image of the vertices. Besides, simplicial maps are continuous. Note that $s(A)$ may have a lower dimension than A since we do not require s to be injective. From the point of view of category theory[1], we have just created a category **Csim** whose objects are simplicial complexes and its morphisms are simplicial maps between them.

We are interested as well in finding simplicial maps which bear some resemblance with a given map between two polyhedra.

Definition (Simplicial approximation of a map).
Let $f\colon |K| \to |L|$ be a continuous map. A simplicial map $s\colon |K| \to |L|$ is said to be a *simplicial approximation* of f if $s(x)$ lies in the carrier of $f(x)$ for each $x \in |K|$.

Remark

The definition above does not state that $s(x)$ and $f(x)$ have the same carrier. Actually, it says that the carrier of $s(x)$ is a face of the carrier of $f(x)$. We will denote this situation by $s(x) \leq f(x)$.

Simplicial approximations do not always exist (see [1], page 129). However, we can prove the following:

Theorem 1.1 (Simplicial approximation theorem). *Let $f\colon |K| \to |L|$ be a continuous map between polyhedra. For m large enough, there exists a simplicial approximation $s\colon |K^m| \to |L|$ to $f\colon |K^m| \to |L|$.*

The proof of this theorem requires previous definitions and lemmas which take us longer than what we can afford in this work, but they can be found in [1]. The lemma proved below will be required in a future theorem of Chapter 2.

Lemma 1.2. *If $s\colon |K^m| \to |L|$ is a simplicial approximation of $f\colon |K| \to |L|$ and $t\colon |L| \to |M|$ is a simplicial map, then $t \circ s$ is a simplicial approximation of $t \circ f$. In particular, simplicial maps preserve the $\leq$ relation introduced in the remark.*

Proof. Assume $\sigma = (v_0, \ldots, v_k)$ is the carrier of x. Since s is a simplicial approximation of f we get $s(x) = \lambda_0 s(v_0) + \cdots + \lambda_k s(v_k) \leq f(x) = \mu_0 s(v_0) + \cdots + \mu_k s(v_k) + \mu_{k+1} w_{k+1} + \cdots + \mu_m w_m$. Applying t, $t(s(x)) = \lambda_0 t(s(v_0)) + \cdots + \lambda_k t(s(v_k))$ and $t(f(x)) = \mu_0 t(s(v_0)) + \cdots + \mu_k t(s(v_k)) + \mu_{k+1} t(w_{k+1}) + \cdots + \mu_m t(w_m)$. Hence $t(s(x)) \leq t(f(x))$, $\forall x \in |K^m|$ and $t \circ s$ is a simplicial approximation of $t \circ f$. □

We are ready to introduce *simplicial homology* in the next chapter.

[1] In [4] they follow a categorical approach of simplicial homology, treating triangulations and homology groups as functors.

Chapter 2

Simplicial Homology

A bridge between algebra and topology

The task of telling when two given topological spaces are homeomorphic is a deeply discussed topic in topology. One tool used in algebraic topology is the fundamental group: being homeomorphic implies having isomorphic fundamental groups. However, the reciprocal is not true. For example, S^3 and S^4 have the same fundamental groups, but they are not homeomorphic. In an attempt to overcome this difficulty, we introduce a new invariant based on the triangulation of a topological space: homology groups.

2.1 Associating groups to a simplicial complex

Let K be a finite simplicial complex. We are aiming to construct a group structure whose generators are precisely the q-simplices of K. Given a q-simplex, $q > 0$, there are $(q+1)!$ different orderings of its vertices. We choose to identify them up to even permutations, leaving us two canonical orderings: $(v_0, v_1, \ldots, v_q)$ and $(v_1, v_0, \ldots, v_q)$. These orderings are called *orientations* and a simplex with an explicit orientation is an *oriented simplex.* The two orientations of a simplex are said to be opposite to each other and it is expressed via their corresponding oriented simplices $(v_0, v_1, \ldots, v_q) = -(v_1, v_0, \ldots, v_q)$. For 0-simplices there is only one orientation. We have gathered enough algebraic foundations to define groups on our complex.

Definition ($C_q(K)$, q-th chain group of K)**.**
Let K be a simplicial complex. We define the *q-th chain group of K, $C_q(K)$* as the free abelian group generated by the oriented q-simplices of K, subject to the relations $\sigma + (-\sigma) = 0$ where σ is an oriented q-simplex. Elements of this group are referred to as *q-dimensional chains.*
Note that $C_q(K)$ is a free abelian group with rank equal the number of simplicial complexes in K. Besides, bear in mind that $-\lambda\sigma = \lambda(-\sigma)$ holds for all $\lambda \in \mathbb{Z}$ and σ an oriented q-simplex.

In algebra the functorial approach plays the starring role: once we have some groups we must find homomorphisms between them. One strategy to define a homomorphism between chain groups is specifying the value on each generator simplex, check that the relations are preserved and extend linearly.
Given the oriented q-simplex $(v_0, v_1, \ldots, v_q)$, there is an induced orientation on its $(q-1)$-faces. The orientation of the $(q-1)$-face $(v_0, v_1, \ldots, \hat{v}_i, \ldots, v_q)$, whose vertices are all but v_i, is defined by $(-1)^i(v_0, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_q)$. The motivation behind this choice lies in the following definition.

Definition (∂_q, boundary homomorphism).
The boundary of an oriented q-simplex is defined to be the $(q-1)$-chain determined by the sum of it $(q-1)$-dimensional faces, each taken with the orientation induced from that on the whole simplex. Explicitly, $\partial_q : C_q(K) \to C_{q-1}(K)$ via $\partial_q(v_0, v_1, \dots, v_q) = \sum_{i=0}^{q}(-1)^i(v_0, v_1, \dots, \hat{v}_i, \dots, v_q)$. Extending by linearity, $\partial_q\left(\sum_i \lambda_i\sigma_i\right) = \sum_i \lambda_i\partial_q(\sigma_i)$.
In order for it to be a group homomorphism, we should check that the relation $\sigma + (-\sigma) = 0$ is preserved, i.e., $\partial_q(-\sigma) = -\partial_q(\sigma)$. Indeed,

$$\begin{aligned}\partial_q(-\sigma) = \partial_q(v_1, v_0, \dots, v_q) &= \sum_{i=0}^{q}(-1)^i(v_1, v_0, \dots, \hat{v}_i, \dots, v_q) = \\ &= -\sum_{i=0}^{q}(-1)^i(v_0, v_1, \dots, \hat{v}_i, \dots, v_q) = -\partial_q(\sigma)\end{aligned}$$

In the special case when $q = 0$, we define the boundary of a single vertex to be 0 and set $C_{-1}(K) = 0$.

Definition ($Z_q(K)$, group of q-cycles of K and $B_q(K)$, group of bounding q-cycles of K).
Let us call the kernel of ∂_q the *group of q-cycles of K* and denote it by $Z_q(K)$ and also call the image of ∂_{q+1} the *group of bounding q-cycles of K* and denote it by $B_q(K)$.

In this case, the names of both groups are anticipating the next result: every bounding q-cycle is also a q-cycle.

Theorem 2.1. *Given any simplicial complex K, $B_q(K)$ is a subgroup of $Z_q(K)$, or equivalently $\partial_q \circ \partial_{q+1} = 0$, $\forall q$.*

Proof. Let $(v_0, v_1, \dots, v_{q+1})$ be a $q+1$-simplex of K. Now,

$$[\partial_q \circ \partial_{q+1}](v_0, v_1, \dots, v_q) =$$

$$= \partial_q\left(\sum_{i=0}^{q+1}(-1)^i(v_0, v_1, \dots, \hat{v}_i, \dots, v_{q+1})\right) = \sum_{i=0}^{q+1}(-1)^i\partial_q(v_0, v_1, \dots, \hat{v}_i, \dots, v_{q+1}) =$$

$$= \sum_{i=0}^{q+1}(-1)^i\left(\sum_{j=0}^{i}(-1)^j(v_0, v_1, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{q+1}) + \sum_{j=i+1}^{q+1}(-1)^{j-1}(v_0, v_1, \dots, \hat{v}_i, \dots, \hat{v}_j, \dots, v_{q+1})\right) =$$

$$\overset{(a)}{=} \sum_{i=0}^{q+1}(-1)^i 0 = 0$$

In (a) all the terms in pairs since each oriented $(q-1)$-simplex appears twice, once positive and once negative. □

Since chain groups are abelian, subgroups generate quotient groups. The above theorem allows us to build the core elements of simplicial homology.

Definition ($H_q(K)$, q-th homology group of K).
Given a simplicial complex K, we define its *q-th homology group* as

$$H_q(K) = \frac{Z_q(K)}{B_q(K)} = \frac{\ker \partial_q}{\operatorname{Im} \partial_{q+1}}$$

The element of $H_q(K)$ determined by a q-cycle z is called the *homology class* of z and denoted by $[z]$. Recall that any two elements $x, y \in H_q(K)$ satisfy $x = y + w$, where $w \in B_q(K)$ and they are called *homologous cycles*. Homology groups are by construction finitely generated abelian groups hence each of them is isomorphic to $\mathbb{Z}_{p_1^{a_1}} \oplus \cdots \oplus \mathbb{Z}_{p_k^{a_k}} \oplus \mathbb{Z}^n$ for some primes $p_1, \ldots, p_k$ and $a_1, \ldots, a_k, k, n \in \mathbb{Z}$. The rank of the finitely generated free abelian group $\mathbb{Z}^n$ is called the *q-th Betti number of K*, denoted by $\beta_q = n$.

At first sight, these groups could be deeemed too abstract. Indeed, that is the case for higher values of q, but for the case $q = 0$ it behaves as a counter of connected components:
Given any simplicial complex K, by the definition of ∂_0 as the zero map, $Z_0(K) = \ker \partial_0 = C_0(K)$. Notice that the latter is defined to be the free abelian group generated by the vertices of K. In the case of bounding 0-cycles, $B_0(K)$, one can observe that two vertices, say u, v connected by an edge $a = (u, v)$ generate the bounding q-cycle $\partial_1(a) = v - u$. Hence, homologous points are the ones connected by paths. In other words, when quotienting by $B_0(K)$ we are identifying points which can be reached using $1-$simplexes. We have proved this result.

Proposition 2.2. *If K is a simplicial complex, then $H_0(K)$ is a free abelian group whose rank is the number of connected components of $|K|$.*

2.2 Computing homology

Once we have introduced the theoretical construction of homology groups is time to compute them for a specific simplicial complex. For this first example we choose a triangulation of the real projective plane, we denote by P (see Figure 2.1).

We will not undertake a deep analysis[1] to show that indeed Figure 2.1 is homeomorphic to the real projective plane. Nevertheless P bears some resemblance with the latter, recalling the description of the projective plane as the unit circle with diametrically opposed points identified. $|P|$ is a connected space hence $H_0(P) \simeq \mathbb{Z}$. In addition, $C_q(P) = 0$, $\forall q > 2$ implying that $H_q(P) = 0$, $\forall q > 2$. The two remaining homology groups must be computed through calculations.
Take a 2-cycle, i.e. $c_2 \in Z_2(P)$. If c_2 has $n(0,1,3)$ as a summand, the only way to cancel out the term $n(3,0)$ in $n\partial_2(0,1,3) = n(0,1) + n(1,3) + n(3,0)$ is to include the summand $n(0,3,2)$ in c_2. Applying this argument several times leaves us with $c_2 = n(0,1,3) + n(0,3,2) + n(1,4,3) + n(3,4,5) + n(2,3,5) + n(1,2,4) + n(2,0,4) + n(0,5,4) + n(0,1,5) + n(1,2,5)$. All 2-simplices of P with coefficient n and oriented with clock-wise direction. Therefore, computing its boundary we obtain $\partial_2(c_2) = 2n(0,1) + 2n(1,2) + 2n(2,0)$, different from 0 unless $n = 0$. Thus $Z_2(P) = 0$ and $H_2(P) = 0$.

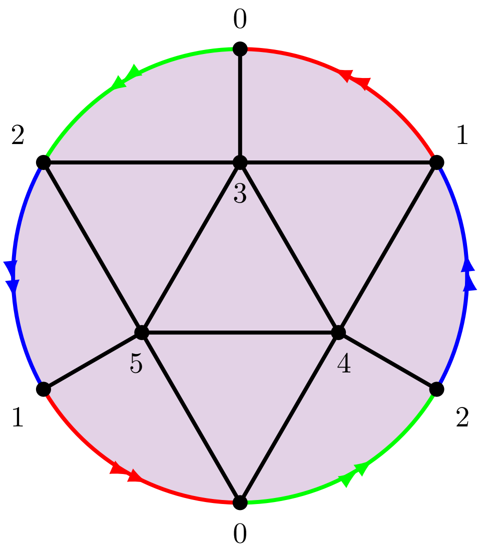


Figure 2.1: A triangulation of the real projective plane, P, made up of 6, 0-simples; 15, 1-simplices and 10, 2-simplices.

[1]The issue is discussed with all details in [6]

Let us now compute $H_1(P)$. We claim that every 1-cycle is homologous to a multiple of $c_1 = (0,1)+(1,2)+(2,0)$. Assume $c = \sum_{(i,j)\in J} \alpha_{ij}(i,j) \in Z_1(P)$, where $J = \{(i,j) \mid 0 \le i < j \le 5\}$. We will start adding to c multiples of 2-chain boundaries (remaining in the homology class of c) aiming to cancel out all the 1-chains with vertices 3, 4 or 5.

$$\begin{aligned} d = c \;+\; & \alpha_{03} && \partial_2(0,1,3) \\ +\; & (\alpha_{03}+\alpha_{13}) && \partial_2(1,4,3) \\ +\; & (\alpha_{03}+\alpha_{13}-\alpha_{34}) && \partial_2(3,4,5) \\ +\; & (\alpha_{03}+\alpha_{13}-\alpha_{34}-\alpha_{35}) && \partial_2(2,3,5) \end{aligned}$$

Hence d is homologous to c, since with each boundary summand we remove the edges $(0,3)$, $(1,3)$, $(3,4)$, $(3,5)$ respectively. Despite five edges being incident with the vertex 3 we have only removed explicitly four edges, all of them but $(2,3)$. This one has coefficient $-\alpha_{03}-\alpha_{13}-\alpha_{23}+\alpha_{35}+\alpha_{45}$ in d, which equals 0 since it is the coefficient of (3) in $\partial_1(c) = 0$.

$$\begin{aligned} e = d \;+\; & (\alpha_{03}+\alpha_{13}+\alpha_{14}) && \partial_2(1,2,4) \\ +\; & (\alpha_{03}+\alpha_{13}+\alpha_{14}+\alpha_{24}) && \partial_2(2,0,4) \\ +\; & (-\alpha_{03}-\alpha_{13}+\alpha_{34}-\alpha_{45}) && \partial_2(0,5,4) \\ +\; & (\alpha_{03}+\alpha_{13}-\alpha_{34}-\alpha_{25}-\alpha_{35}) && \partial_2(1,2,5) \\ +\; & (\alpha_{03}+\alpha_{13}-\alpha_{34}-\alpha_{15}-\alpha_{25}-\alpha_{35}) && \partial_2(0,1,5) \\ =\; & k_0(0,1)+k_1(1,2)-k_2(0,2),\ k_i \in \mathbb{Z} \end{aligned}$$

Following the pattern described before we end up with e. Observe that e is also a 1-cycle, hence $k_0 = k_1 = k_2$.
We leave it to the reader to verify $[c_1] \neq [0]$. When studying $H_2(P)$ we found out that $2c_1 \in B_1(P)$. For that reason, $H_1(P) \simeq \mathbb{Z}_2$.

Unlike in the previous example, where we computed homology[2] for a specific simplicial complex, we are now going to compute the homology groups of the closure of a $(n+1)$-simplex, Δ^{n+1}, in a more general way.

Previously, when we defined the cone of a simplex, we stated that Δ^3 is merely the cone of Δ^2. This property can be generalised: any Δ^{n+1} is the cone of some Δ^n. Therefore, it is possible to study the homology groups of Δ^{n+1} via their conic structure. First, let us introduce a preliminary result.

Lemma 2.3. *Let K be a cone, i.e. $K \simeq CL$ for some complex L whose dimension is one less than that of K, and let v be the apex of K. For $q > 0$, define $d_q : C_q(K) \to C_{q+1}(K)$ mapping a generator q-simplex $\sigma = (v_0, \ldots, v_q)$ as follows (extend by linearity):*

$$d_q(\sigma) = \begin{cases} (v, v_0, \ldots, v_q) & \text{if } \sigma \neq 0 \text{ lies in } L. \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $d_q(\sigma) = 0$ for $q > dim(L)$. Then:
1) d_q is a well-defined group homomorphism.
2) $\partial_{q+1}(d_q(\sigma)) = \sigma - d_{q-1}(\partial_q(\sigma))$.

Proof. The value of d does not depend on the ordering of the vertices, but on the orientation of the simplex. Besides, it satisfies $d_q(\sigma) + d_q(-\sigma) = 0$, hence it is a well-defined group homomorphism. Finally, let us prove 2):
If σ lies in L,

$$\partial_{q+1}(d_q(\sigma)) = \partial_{q+1}(v, v_o, \ldots, v_q) = (v_0, \ldots, v_q) + \sum_{i=0}^{q} (-1)^{i+1}(v, v_0, \ldots, \hat{v}_i, \ldots, v_q) = \sigma - d_{q-1}(\partial_q(\sigma))$$

If σ does not lie in L, assume $v_0 = v$,

[2] In [15] an algorithm for computing homology groups is presented.

$$\sigma - d_{q-1}(\partial_q(\sigma)) = (v, v_1, \ldots, v_q) - d_{q-1}[(v_1, \ldots, v_q)] - d_{q-1}\left(\sum_{i=1}^{q}(-1)^i(v_1, \ldots, \hat{v}_i, \ldots, v_q)\right) =$$
$$= \sigma - \sigma - 0 = 0 = \partial_{q+1}(d_q(\sigma))$$

□

Proposition 2.4. *Let K be a cone, then $H_0(K) \simeq \mathbb{Z}$ and $H_q(K) = 0$ for $q > 0$.*

Proof. A cone is a connected space, so $H_0(K) \simeq \mathbb{Z}$. Now, applying Lemma 2.3, take $z \in Z_q(K)$, $z = z - 0 = z - d_{q-1}(\partial_q(z)) = \partial_{q+1}(d_q(z))$, hence $Z_q(K) = B_q(K)$ for $q > 0$. □

Corollary 2.5. *$H_0(\Delta^{n+1}) \simeq \mathbb{Z}$ and $H_q(\Delta^{n+1}) = 0$, for $q > 0$ and $n \geq 0$.*

So far we have seen how to associate a simplicial complex to a topological space (by taking a triangulation) and defined homology groups that depend on that specific simplicial complex. However a topological space might be triangulated in different ways. For that reason we cannot yet associate homology groups to topological spaces. During the next sections we are going to develop the theory to prove that the homology groups obtained by two different triangulations of the same space must coincide.

2.3 Tools from homological algebra

The path we are due to follow requires a robust algebraic setup. In this section we consider the previously defined topological concepts and investigate their properties as algebraic constructions. Indeed, chain groups and boundary maps will be linked together to form a *chain complex* and simplicial maps will give rise to *chain maps*. Thus, we will develop results to be applied to the corresponding simplicial objects. This will be done at the end of the section (and throughout the next one) being able to mix two elements that had been previously discussed seperately: simplicial maps and homology groups.

Definition (Chain complex and its homology).
A *chain complex* in the category **Grp** is a (possibly infinite) sequence of groups and group homomorphisms,

$$C \equiv \qquad \ldots \xrightarrow{\partial_{q+2}} C_{q+1} \xrightarrow{\partial_{q+1}} C_q \xrightarrow{\partial_q} C_{q-1} \xrightarrow{\partial_{q-1}} \ldots$$

where $\partial_q \circ \partial_{q+1} = 0$, $\forall q$, i.e. $\mathrm{Im}(\partial_{q+1}) \subseteq \ker(\partial_q)$. The quotient $H_q(C) = \dfrac{\ker \partial_q}{\mathrm{Im}\, \partial_{q+1}}$ is known as the *q-th homology group of C*.

Definition (Chain map).
A *chain map* $\phi\colon C \to D$ is a collection of maps $\phi_q\colon C_q \to D_q$, $\forall q$ such that the following diagram commutes

$$\begin{array}{ccccccccc}
\ldots & \xrightarrow{\partial^C_{q+2}} & C_{q+1} & \xrightarrow{\partial^C_{q+1}} & C_q & \xrightarrow{\partial^C_q} & C_{q-1} & \xrightarrow{\partial^C_{q-1}} & \ldots \\
 & \circlearrowleft & \downarrow \phi_{q+1} & \circlearrowleft & \downarrow \phi_q & \circlearrowleft & \downarrow \phi_{q-1} & \circlearrowleft & \\
\ldots & \xrightarrow[\partial^D_{q+2}]{} & D_{q+1} & \xrightarrow[\partial^D_{q+1}]{} & D_q & \xrightarrow[\partial^D_q]{} & D_{q-1} & \xrightarrow[\partial^D_{q-1}]{} & \ldots
\end{array}$$

In other words, $\partial_q^D \circ \phi_q = \phi_{q-1} \circ \partial_q^C$, $\forall q$.

Since chain maps are defined on chain complexes, it would be useful to induce maps on their corresponding homology groups. The next result proves that, indeed, chain maps meet the desired property.

Proposition 2.6. *If $\phi\colon C \to D$ is a chain map, then ϕ induces a group homomorphism in the q-th homology groups*

$$\phi_{q*}\colon H_q(C) \to H_q(D), \ \text{such that} \quad \phi_{q*}(x + \operatorname{Im} \partial_{q+1}^C) = \phi_q(x) + \operatorname{Im} \partial_{q+1}^D$$

Proof. It suffices to shows that the map is well-defined, which translates into checking these inclusions:

$$\partial_q^D \left(\phi_q \left(\ker \partial_q^C\right)\right) = \phi_{q-1} \left(\partial_q^C \left(\ker \partial_q^C\right)\right) = 0, \text{ hence } \phi_q(\ker \partial_q^C) \subset \ker \partial_q^D$$

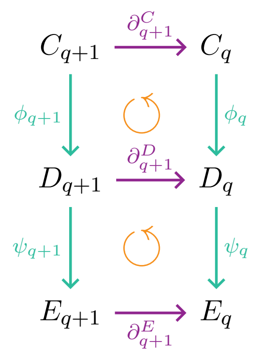
$$\phi_q \left(\operatorname{Im} \partial_{q+1}^C\right) = \phi_q \left(\partial_{q+1}^C \left(C_{q+1}\right)\right) = \partial_{q+1}^D \left(\phi_{q+1} \left(C_{q+1}\right)\right) \subset \operatorname{Im} \partial_{q+1}^D$$

In addition, we must prove that the image of a class does not depend on its representative: Let $y \in x + \operatorname{Im} \partial_{q+1}^C$, then $y = x + c$ for some $c \in \operatorname{Im} \partial_{q+1}^C$.

$$\phi_{q*}(y + \operatorname{Im} \partial_{q+1}^C) = \phi_q(y) + \operatorname{Im} \partial_{q+1}^D = \phi_q(x) + \phi_q(c) + \operatorname{Im} \partial_{q+1}^D = \phi_q(x) + \operatorname{Im} \partial_{q+1}^D \qquad \square$$

Lemma 2.7. *If $\psi\colon D \to E$ is another chain map, then $\psi \circ \phi\colon C \to E$ is a chain map and $(\psi \circ \phi)_{q*} = \psi_{q*} \circ \phi_{q*}\colon H_q(C) \to H_q(E)$.*

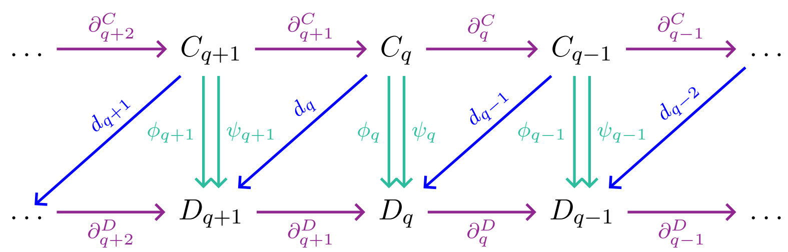
Proof. Showing that $\psi \circ \phi$ is a chain map requires verifying the commutativity of the paths. We use the commutativity of the two squares:



$$(\psi \circ \phi)_q \circ \partial_{q+1}^C = \psi_q \circ \phi_q \circ \partial_{q+1}^C = \psi_q \circ \partial_{q+1}^D \circ \phi_{q+1} =$$
$$= \partial_{q+1}^E \circ \psi_{q+1} \circ \phi_{q+1} = \partial_{q+1}^E \circ (\psi \circ \phi)_{q+1}$$

In addition,
$(\psi \circ \phi)_{q*}(x + \operatorname{Im} \partial_{q+1}^C) = (\psi \circ \phi)_q(x) + \operatorname{Im} \partial_{q+1}^E = \psi_q(\phi_q(x)) + \operatorname{Im} \partial_{q+1}^E$
$(\psi_{q*} \circ \phi_{q*})(x + \operatorname{Im} \partial_{q+1}^C) = \psi_{q*}(\phi_q(x) + \operatorname{Im} \partial_{q+1}^D) = \psi_q(\phi_q(x)) + \operatorname{Im} \partial_{q+1}^E$ $\square$

Definition (Chain homotopy).
Let $\phi, \psi\colon C \to D$ be chain maps. A *chain homotopy* $d\colon C \to D$ between ϕ and ψ is a collection of homomorphisms $d_q\colon C_q \to D_{q+1}$ such that $\phi_q - \psi_q = \partial_{q+1}^D \circ d_q + d_{q-1} \circ \partial_q^C$, $\forall q$. If such a homotopy exists, we say that ϕ and ψ are *homotopic chain maps.*



The chain homotopy relation was designed explicitly to produce this result.

Proposition 2.8. *If ϕ is chain homotopic to ψ, then $\phi_{q*} = \psi_{q*}$, $\forall q$*

Proof. Let $z \in \ker \partial_q^C$, we want to show that

$$\phi_{q*}(z + \operatorname{Im} \partial_{q+1}^C) = \psi_{q*}(z + \operatorname{Im} \partial_{q+1}^C) \iff \phi_q(z) - \psi_q(z) \in \operatorname{Im} \partial_{q+1}^D$$

Indeed,

$$\phi_q(z) - \psi_q(z) = \partial_{q+1}^D(d_q(z)) + d_{q-1}(\partial_q^C(z)) = \partial_{q+1}^D(d_q(z)) \in \operatorname{Im} \partial_{q+1}^D$$

□

Definition (Chain equivalent complexes).
Two chain complexes C and D are said to be *equivalent* if there are chain maps $\phi\colon C \to D$ and $\mu\colon D \to C$ such that $\mu \circ \phi$ and $\phi \circ \mu$ are chain homotopic to $\mathrm{Id}_C\colon C \to C$ and $\mathrm{Id}_D\colon D \to D$ respectively.

Proposition 2.9. *Chain equivalent complexes C and D have isomorphic homology groups in corresponding dimensions.*

Proof. If ϕ and μ are the chain maps required by the definition, using Lemma 2.7 we obtain

$$\begin{cases} \mathrm{Id}_{H_q(C)} = (\mu \circ \phi)_{q*} = \mu_{q*} \circ \phi_{q*}. \\ \mathrm{Id}_{H_q(D)} = (\phi \circ \mu)_{q*} = \phi_{q*} \circ \mu_{q*}. \end{cases}$$

As a direct consequence $\phi_{q*}\colon H_q(C) \to H_q(D)$ is an isomorphism for every q. □

Remark

In our simplicial setting, we will denote chain complexes induced by chain groups of a simplicial complex K and chain maps between them as follows:

$$C(K) \equiv \quad \dots \xrightarrow{\partial_{q+2}} C_{q+1}(K) \xrightarrow{\partial_{q+1}} C_q(K) \xrightarrow{\partial_q} C_{q-1}(K) \xrightarrow{\partial_{q-1}} \dots$$

$$\phi\colon C(K) \to C(L)$$

Now, as the introduction of this section suggested, we aim to use Proposition 2.6 in the scenario where chain complexes are given by chain groups of simplicial complexes related by a simplicial map $s\colon |K| \to |L|$.

Theorem 2.10. *Any simplicial map $s\colon |K| \to |L|$ induces a chain map $s\colon C(K) \to C(L)$ and homomorphisms in the homology groups $s_{q*}\colon H_q(K) \to H_q(L)$, $\forall q$.*

Proof. We shall begin by defining a chain map from $C(K)$ to $C(L)$ using s.
Let $s_q\colon C_q(K) \to C_q(L)$ be a homomorphism which acts on a generator q-simplex $\sigma = (v_0, \dots, v_q)$ as follows (extend by linearity):

$$s_q(\sigma) = \begin{cases} (s(v_0), \dots, s(v_q)) & \text{if all the vertices } s(v_0), \dots, s(v_q) \text{ are distinct.} \\ 0 & \text{otherwise.} \end{cases}$$

We are reduced to proving that this chain map commutes with the boundary operators.

The case when all the vertices $s(v_0), \dots, s(v_q)$ are distinct is clear, hence we will show the remaining one. Suppose $s(v_j) = s(v_k)$, where $j < k$. Thus, $\partial_q(s_q(\sigma)) = 0$. All terms in the sum

$$\begin{array}{ccc} C_q(K) & \xrightarrow{\partial_q} & C_{q-1}(K) \\ \downarrow{s_q} & & \downarrow{s_{q-1}} \\ C_q(L) & \xrightarrow[\partial_q]{} & C_{q-1}(L) \end{array}$$

$$s_{q-1}(\partial_q(\sigma)) = \sum_{i=0}^{q} (-1)^i s_{q-1}(v_0, \dots, \hat{v}_i, \dots, v_q)$$

vanish except $(-1)^j s_{q-1}(v_0, \dots, \hat{v}_j, \dots, v_q)$ and $(-1)^k s_{q-1}(v_0, \dots, \hat{v}_k, \dots, v_q)$ (only if v_j and v_k are the only vertices identified by s, otherwise all terms cancel out). Observe that:

$$\begin{aligned} s_{q-1}(v_0, \dots, \hat{v}_j, \dots, v_q) &= (s(v_0), \dots, \widehat{s(v_j)}, \dots, s(v_q)) = \\ &= (-1)^{k-j-1}(s(v_0), \dots, \widehat{s(v_k)}, \dots, s(v_q)) = (-1)^{k-j-1} s_{q-1}(v_0, \dots, \hat{v}_k, \dots, v_q) \end{aligned}$$

Proving our claim and Proposition 2.6 allows us to construct maps $s_{q*} \colon H_q(K) \to H_q(L)$, $\forall q$. □

2.4 Invariance of homology groups under barycentric subdivision

Our next step is proving that a simplicial complex K and its barycentric subdivision K^1 have isomorphic homology groups for each order.

This new notation will be convenient along this section. If $\sigma = (v_0, \dots, v_q)$ is a q-simplex and v is a vertex in general position with the vertices of σ, we denote by $v\sigma$ the $(q+1)$-simplex $(v, v_0, \dots, v_q)$. Besides we can extend this notation linearly to q-chains. Given $c = \sum_i \lambda_i \sigma_i$ a q-chain, vc denotes the $(q+1)$-chain $\sum_i \lambda_i (v\sigma_i)$. Recalling Lemma 2.3, one realises that we proved the following relation back then.

Lemma 2.11. *Let c be a q-chain on a simplicial complex K and v a vertex for which vc belongs to $C_{q+1}(K)$, then $\partial_{q+1}(vc) = c - v\partial_q(c)$.*

Definition (Subdivision chain maps)**.**
Let K be a simplicial complex. We define the chain map $\chi \colon C(K) \to C(K^1)$ known as *the first subdivision chain map* using an inductive procedure:

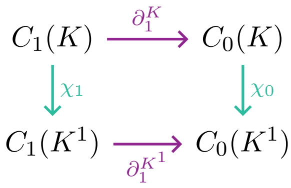
(i) Every 0-simplex of K is also a 0-simplex of K^1, hence $C_0(K)$ is a subgroup of $C_0(K^1)$ and we can take $\chi_0 \colon C_0(K) \to C_0(K^1)$ to be the inclusion map.

(ii) Now, we define the image by χ_q on a generator q-simplex σ, $\chi_q(\sigma) = \dot{\sigma}\chi_{q-1}(\partial_q(\sigma))$ where $\dot{\sigma}$ denotes the barycenter of σ. We extend by linearity to obtain $\chi_q \colon C_q(K) \to C_q(K^1)$.

For $m > 1$, the m-th subdivision map $\chi^m \colon C(K) \to C(K^m)$ is the composition of χ^{m-1}, the $(m-1)$-th subdivision chain map, with $\chi \colon C(K^{m-1}) \to C(K^m)$, the first subdivision chain map of the $(m-1)$-th barycentric subdivision K^{m-1}.

We encourage the reader to consult [10] if they need examples of how the chain maps defined in this section act on explicit simplicial complexes. One subtle detail we have not mentioned is that $\dot{\sigma}\chi_{q-1}(\partial_q(\sigma))$ is well-defined. Indeed, the barycenter of a q-simplex σ is in general position with any $(q-1)$ points which are linear combinations involving at most $(q-1)$ vertices of σ. In contrast, we have defined chain maps without proving the commutativity with boundary maps! First we will prove that $\chi \colon C(K) \to C(K^1)$ is a chain map by checking commutativity for generator simplices.

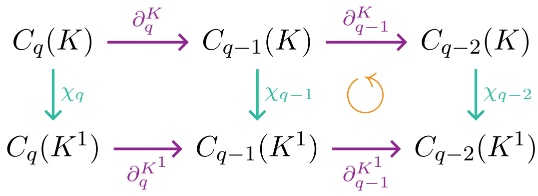
For the case $q = 1$, equalities follow, in order, from the definition of χ_1, Lemma 2.11, χ_0 being the inclusion map and $\partial_0^{K^1} \circ \partial_1^{K^1} = 0$.



$$\partial_1^{K^1}(\chi_1(\sigma)) = \partial_1^{K^1}[\dot{\sigma}\chi_0(\partial_1^K(\sigma))] = \chi_0(\partial_1^K(\sigma)) - \dot{\sigma}\partial_0^{K^1}[\chi_0(\partial_1^K(\sigma))] = \\ = \chi_0(\partial_1^K(\sigma)) - \dot{\sigma}\partial_0^{K^1}[\partial_1^{K^1}(\sigma)] = \chi_0(\partial_1^K(\sigma))$$

By induction $\partial_{q-1}^{K^1}\chi_{q-1} = \chi_{q-2}\partial_{q-1}^K$. Therefore:

$$\partial_q^{K^1}(\chi_q(\sigma)) = \partial_q^{K^1}[\dot{\sigma}\chi_{q-1}(\partial_q^K(\sigma))] = \\ = \chi_{q-1}(\partial_q^K(\sigma)) - \dot{\sigma}\partial_{q-1}^{K^1}[\chi_{q-1}(\partial_q^K(\sigma))] = \\ = \chi_{q-1}(\partial_q^K(\sigma)) - \dot{\sigma}\chi_{q-2}[\partial_{q-1}^K(\partial_q^K(\sigma))] = \\ = \chi_{q-1}(\partial_q^K(\sigma))$$



As a result, the first subdivision chain map χ is a chain map and Lemma 2.7 guarantees that χ^m is also a chain map for every m.

Definition (Standard simplicial map).
Let K be a simplicial complex. We define a *standard simplicial map* to be a simplicial map $\theta\colon K^1 \to K$ satisfying that $\theta(\dot{\sigma})$ is a vertex of σ. Observe that there might be more than one standard simplicial map between K^1 and K. We define as well a standard simplicial map $\theta\colon K^m \to K$ as the composition of m standard simplicial maps $\theta_i\colon K^i \to K^{i-1}$ for $i = 1, \dots, m$.

Proposition 2.12. *The induced chain map $\theta\colon C(K^1) \to C(K)$ is a left inverse for the first subdivision chain map $\chi\colon C(K) \to C(K^1)$. In other words, $\theta_q \circ \chi_q = Id_{C_q(K)}$ for every $q > 0$.*

Proof. Clearly $\theta_0 \circ \chi_0 = \mathrm{Id}_{C_0(K)}$. Observe that if τ is a q-simplex in K^1 then $\theta_q(\tau) = \eta\sigma$ where $\eta \in \{0, +1, -1\}$ and σ is a q-simplex of K which produces τ in its barycentric subdivision. Inductively, assume that $\theta_{q-1} \circ \chi_{q-1} = \mathrm{Id}_{C_{q-1}(K)}$ and let σ be a generator q-simplex in K^1. Since $\theta_q(\chi_q(\sigma)) = \theta_q[\dot{\sigma}\chi_{q-1}(\partial_q(\sigma))] = m\sigma$ and both θ and χ are chain maps:

$$m\partial_q^K(\sigma) = \partial_q^K(m\sigma) = \partial_q^K[\theta_q(\chi_q(\sigma))] = \theta_{q-1}[\partial_q^{K^1}(\chi_q(\sigma))] = \theta_{q-1}[\chi_{q-1}(\partial_q^K(\sigma))] = \partial_q^K(\sigma)$$

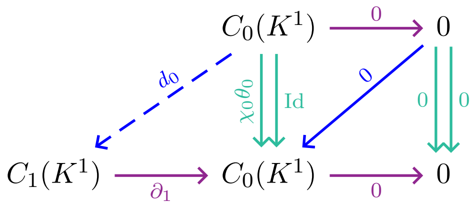
Thus, $m = 1$ and $\theta_q \circ \chi_q = \mathrm{Id}_{C_q(K)}$. □

By definition of $\chi^m\colon C(K) \to C(K^m)$ and $\theta\colon K^m \to K$, we also have that the induced chain map $\theta\colon C(K^m) \to C(K)$ is a left inverse for the m-th subdivision chain map $\chi^m\colon C(K) \to C(K^m)$.

Definition (Barycentric carrier of a simplex).
Let K be a simplicial complex with first barycentric subdivision K^1. Given a simplex σ in K^1, there exists at least one simplex in K such that σ lies in its barycentric subdivision. We will refer to the one with the biggest dimension as the *barycentric carrier of* σ and denote it by $BC(\sigma)$.

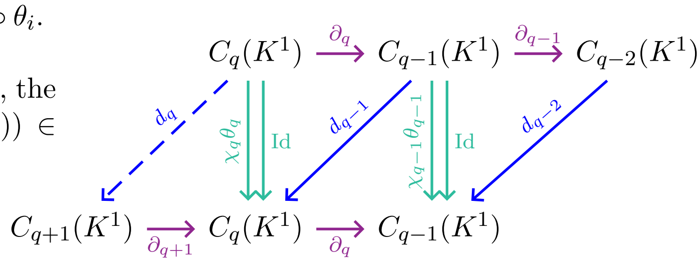
Theorem 2.13. *Barycentric subdivision does not change the homology groups of a complex.*

Proof. Our proof begins with the observation that showing a chain equivalence between $C(K)$ and $C(K^1)$ implies proving the statement above (see Proposition 2.9). According to Proposition 2.12, $\theta_q \circ \chi_q = \mathrm{Id}_{C_q(K)}$, so applying Lemma 2.7, $\theta_{q*} \circ \chi_{q*} = (\theta \circ \chi)_{q*} = (\mathrm{Id}_{C(K)})_{q*} = \mathrm{Id}_{H_q(K)}$. Let us now build a chain homotopy between $\chi \circ \theta$ and $\mathrm{Id}_{C(K^1)}$.



We begin by defining $d_0\colon C_0(K^1) \to C_1(K^1)$. Let w be a vertex in K^1 such that $\theta_0(w) = v$ where v is a vertex of some simplex σ in K whose barycenter is w. Take $d_0(w) = (v, w) \in BC(w)$ and extend by linearity to $C_0(K^1)$. Thus $\mathrm{Id}(w) - \chi_0(\theta_0(w)) = w - v = \partial_1(v, w)$, i.e. $\mathrm{Id}_{C_0(K^1)} - \chi_0 \circ \theta_0 = \partial_1 \circ d_0$, satisfying the chain homotopy condition at $q = 0$.

Now, suppose that we have defined homomorphisms $d_i\colon C_i(K^1) \to C_{i+1}(K^1)$ for $0 \le i \le q-1$, satisfying:
(1) $d_{i-1} \circ \partial_i + \partial_{i+1} \circ d_i = \mathrm{Id}_{C_i(K^1)} - \chi_i \circ \theta_i$.
(2) $d_i(\sigma)$ is always a chain in $BC(\sigma)$.
Let us show that for σ a q-simplex in K^1, the q-chain $z = \sigma - \chi_q(\theta_q(\sigma)) - d_{q-1}(\partial_q(\sigma)) \in Z_q(K^1)$.
Indeed,



$$\partial_q[\ \sigma - \chi_q(\theta_q(\sigma)) - d_{q-1}(\partial_q[\sigma])\ ] =$$

$$\partial_q[\sigma] - \partial_q[\ \chi_q(\theta_q(\sigma))\ ] - \partial_q[\ d_{q-1}(\partial_q[\sigma])\ ] \overset{(a)}{=}$$

$$\partial_q[\sigma] - \partial_q[\ \chi_q(\theta_q(\sigma))\ ] - [\ \partial_q[\sigma] - \chi_{q-1}(\theta_{q-1}(\partial_q[\sigma])) - d_{q-2}(\partial_{q-1}(\partial_q[\sigma]))\ ] \overset{(b)}{=}$$

$$-\,\partial_q[\ \chi_q(\theta_q(\sigma))\ ] + \chi_{q-1}(\theta_{q-1}(\partial_q[\sigma])) \overset{(c)}{=} 0$$

Where we have used the induction hypothesis (1), $\partial_{q-1} \circ \partial_q = 0$ and χ_q and θ_q commutativity with boundary operators in (a), (b) and (c) respectively.
Applying our induction hyphotesis (2), we conclude that $\sigma - \chi_q(\theta_q(\sigma)) - d_{q-1}(\partial_q(\sigma))$ lies in $BC(\sigma)$ as $\sigma - \chi_q(\theta_q(\sigma))$ already belongs by definition. However, barycentric subdivisions are cones implying that $\sigma - \chi_q(\theta_q(\sigma)) - d_{q-1}(\partial_q(\sigma)) = \partial_{q+1}(c) \in B_q(K^1)$ for some $c \in C_{q+1}(K^1)$. We define $d_q\colon C_q(K^1) \to C_{q+1}(K^1)$ such that $d_q(\sigma) = c$ and extend linearly. By construction of d_q we have that, $d_{q-1} \circ \partial_q + \partial_{q+1} \circ d_q = \mathrm{Id}_{C_q(K^1)} - \chi_q \circ \theta_q$ is satisfied. Moreover, $\partial_{q+1}(c)$ lying in $BC(\sigma)$ implies that also $d_q(\sigma) = c$ lies in $BC(\sigma)$.
Thus there is a chain homotopy $C(K)$ and $C(K^1)$ are chain equivalent. □

2.5 Invariance of homology groups under homotopy type

Our initial motivation was to prove that all triangulations of a topological space have isomorphic homology groups. Nevertheless, the machinery available at this point can go even further than that: two topological spaces having the same homotopy type[3] will have the same homologic structure. The invariance under homeomorphism will be just a corollary of the invariance under homotopy type.

Eventually, we are ready to induce homomorphisms between homology groups of simplicial complexes related by **any** map between their polyhedra. The proof requires prior lemmas involving a concept of "closeness" between simplicial maps.

Definition (Close simplicial maps and carrier of a simplex).
Let $s, t\colon |K| \to |L|$ be simplicial maps. We say that they are *close* simplicial maps if for every

[3] Definitions and properties related to homotopy equivalence can be found in Section 5.4 of [1]

simplex $A \in K$ there exists a simplex $B \in L$ such that both $s(A)$ and $t(A)$ are faces of B. Furthermore, we can consider the smallest simplex of L having $s(A)$ and $t(A)$ as faces. We will refer to the latter as the *carrier*[4] of A.

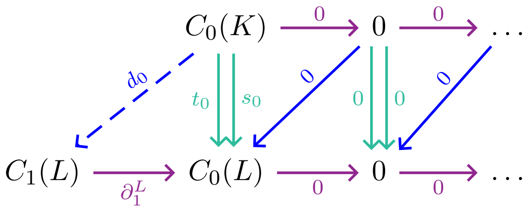
Lemma 2.14. *If $s, t\colon |K^m| \to |L|$ both simplicially approximate $f\colon |K^m| \to |L|$, then they are close simplicial maps.*

Proof. Take a q-simplex $A = (v_0, \dots, v_q)$ of K. Since simplicial approximations must map vertices of K to their images by f, $s(A) = (s(v_0), \dots, s(v_k)) = (f(v_0), \dots, f(v_k)) = (t(v_0), \dots, t(v_k)) = t(A)$ is the carrier of A. □

Lemma 2.15. *If $s, t\colon |K| \to |L|$ are close simplicial maps, then $s_{q*} = t_{q*}\colon H_q(K) \to H_q(L)$, $\forall q$.*

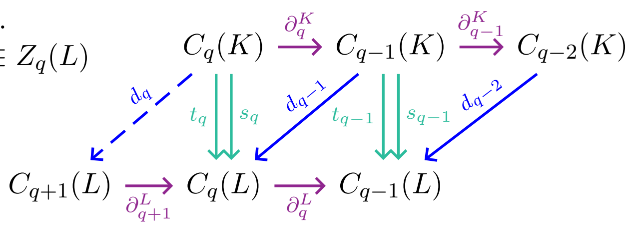
Proof. The proof consists on the construction of a chain homotopy between s and t (seen as chain maps). We will build this collection of homomorphisms in an inductive step.
Recall that simplicial maps induce chain maps as seen in Theorem 2.10. Let us start with d_0. Take $\sigma \in C_0(K)$ to be a vertex of K and define $d_0(\sigma) \in C_1(L)$ as follows:

$$d_0(\sigma) = \begin{cases} (s(\sigma), t(\sigma)) & \text{if } s(\sigma) \neq t(\sigma). \\ 0 & \text{if } s(\sigma) = t(\sigma). \end{cases}$$



Notice that $\partial_1^L \circ d_0 = t_0 - s_0\colon C_0(K) \to C_0(L)$, satisfying the chain homotopy condition[5] at $q = 0$. Besides, $d_0(\sigma)$ is a chain lying in the carrier of σ. Indeed, s and t being close implies that $s(\sigma)$ and $t(\sigma)$ are vertices of the carrier of σ.

Now, suppose that we have defined homomorphisms $d_i\colon C_i(K) \to C_{i+1}(L)$ for $0 \le i \le q-1$, satisfying:
(1) $d_{i-1} \circ \partial_i^K + \partial_{i+1}^L \circ d_i = t_i - s_i\colon C_i(K) \to C_i(L)$
(2) $d_i(\sigma)$ is always a chain in the carrier of σ.
Let us show that $t_q(\sigma) - s_q(\sigma) - d_{q-1}(\partial_q^K(\sigma)) \in Z_q(L)$ which will imply that it is also in $B_q(L)$.
Indeed,



$$\partial_q^L[\ t_q(\sigma) - s_q(\sigma) - d_{q-1}(\partial_q^K(\sigma))\ ] =$$

$$\partial_q^L[\ t_q(\sigma) - s_q(\sigma)\ ] - \partial_q^L[\ d_{q-1}(\partial_q^K(\sigma))\ ] \overset{\text{(a)}}{=}$$

$$\partial_q^L[\ t_q(\sigma) - s_q(\sigma)\ ] - [\ t_{q-1}(\partial_q^K(\sigma)) - s_{q-1}(\partial_q^K(\sigma)) - d_{q-2}(\partial_{q-1}^K(\partial_q^K(\sigma)))\ ] \overset{\text{(b)}}{=}$$

$$\partial_q^L[t_q(\sigma)] - \partial_q^L[s_q(\sigma)] - t_{q-1}(\partial_q^K(\sigma)) + s_{q-1}(\partial_q^K(\sigma)) \overset{\text{(c)}}{=} 0$$

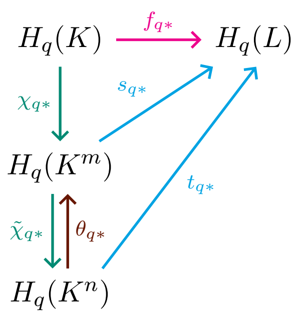
Where we have used the induction hypothesis (1), $\partial_{q-1}^K \circ \partial_q^K = 0$ and s and t commutativity with boundary operators in (a), (b) and (c) respectively.
Applying our induction hyphotesis (2), we conclude that $t_q(\sigma) - s_q(\sigma) - d_{q-1}(\partial_q^K(\sigma))$ lies in the carrier of σ. However, the latter is a cone hence, by Proposition 2.4, $t_q(\sigma) - s_q(\sigma) - d_{q-1}(\partial_q^K(\sigma)) = \partial_{q+1}^L(c) \in B_q(L)$ for some $c \in C_{q+1}(L)$. We define $d_q\colon C_q(K) \to C_{q+1}(L)$ such that $d_q(\sigma) = c$. By construction of d_q, $d_{q-1} \circ \partial_q^K + \partial_{q+1}^L \circ d_q = t_q - s_q\colon C_q(K) \to C_q(L)$ is satisfied. Moreover, $\partial_{q+1}^L(c)$ lying in the carrier of σ implies that also $d_q(\sigma) = c$ lies in the carrier of σ.

[4] Do not confuse with the carrier of a point in a simplicial complex which was defined in the first chapter.
[5] Observe in the diagram that ∂_0^K and d_{-1} are the null homomorphism, so the term $\partial_0^K \circ d_{-1}$ cancels in the sum

Thus there is a chain homotopy $d\colon C(K) \to C(L)$ between s and t. Theorem 2.8 completes the proof. □

Theorem 2.16. *Any continuous map $f\colon |K| \to |L|$ induces a homomorphism $f_{q*}\colon H_q(K) \to H_q(L)$, $\forall q$.*

Proof. So far, we just have one result which enables us to induce homomorphisms between homology groups: Theorem 2.10. Therefore, we have no choice but to use the only tool acting as a bridge between maps and simplicial maps, the simplicial approximation theorem (1.1).
Let $s\colon |K^m| \to |L|$ be a simplicial approximation of f and $\chi\colon C(K) \to C(K^m)$ the subdivision chain map.
We define f_{q*} to be $H_q(K) \xrightarrow{\chi_{q*}} H_q(K^m) \xrightarrow{s_{q*}} H_q(L)$, i.e. $f_{q*} = s_{q*} \circ \chi_{q*}\colon H_q(K) \to H_q(L)$, $\forall q$
Unfortunately, we have made a choice in that definition: the simplicial approximation s. We must show that different simplicial approximations give the same f_{q*}.



Suppose we have two simplicial approximations $s\colon |K^m| \to |L|$ and $t\colon |K^n| \to |L|$, with $n \geq m$. Let $\chi\colon C(K) \to C(K^m)$ and $\tilde{\chi}\colon C(K^m) \to C(K^n)$ be subdivision chain maps and $\theta\colon |K^n| \to |K^m|$ a standard simplicial map. We claim that $s_{q*} \circ \chi_{q*} = t_{q*} \circ \tilde{\chi}_{q*} \circ \chi_{q*}$, $\forall q$. It is easy to check that $s \circ \theta\colon |K^n| \to |L|$ simplicially approximates f (concisely summarising: $\theta(x) \leq x$ so using Lemma 1.2, $s(\theta(x)) \leq s(x) \leq f(x)$), but so does t. Thus, by Lemma 2.14, they are close simplicial maps and applying Proposition 2.7 and Lemma 2.15

$$s_{q*} \circ \theta_{q*} = (s \circ \theta)_{q*} = t_{q*}, \ \forall q$$

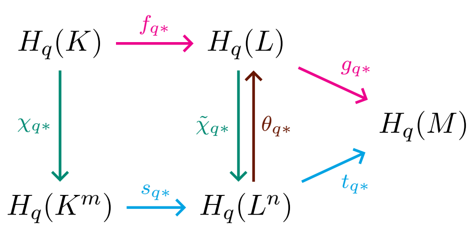
Besides, $\tilde{\chi}_{q*}$ and θ_{q*} are mutually inverse for all q which leads us to the result.

$$s_{q*} \circ \chi_{q*} = s_{q*} \circ \theta_{q*} \circ \tilde{\chi}_{q*} \circ \chi_{q*} = t_{q*} \circ \tilde{\chi}_{q*} \circ \chi_{q*}, \ \forall q.$$ □

These homology maps induced from continuous maps between polyhedra show the same functorial behaviour as showed in Lemma 2.7 for homology maps induced from chain maps.

Corollary 2.17. *If f is the identity map of $|K|$ then each $f_{q*}\colon H_q(K) \to H_q(K)$ is the identity homomorphism, and if we have two maps $|K| \xrightarrow{f} |L| \xrightarrow{g} |M|$, then $(g \circ f)_{q*} = g_{q*} \circ f_{q*}\colon H_q(K) \to H_q(M)$ for all q.*

Proof. The first part follows from construction since both maps χ_{q*} and s_{q*} will be the identity on $H_q(K)$, $\forall q$.



For the next claim, set $t\colon |L^n| \to |M|$ to be a simplicial approximation of $g\colon |L^n| \to |M|$ and $s\colon |K^m| \to |L^n|$ another one for $f\colon |K^m| \to |L^n|$. Consider the subdivision chain maps that go from the original chain complexes to the ones generated by the barycentric subdivisions acting as domains of s and t, $\chi\colon C(K) \to C(K^m)$ and $\tilde{\chi}\colon C(L) \to C(L^n)$. Also let $\theta\colon |L^n| \to |L|$ be the standard simplicial map. We claim that

$$(g \circ f)_{q*} = g_{q*} \circ f_{q*} \text{ for all } q.$$

Let us construct $(g \circ f)_{q*}$ using the definition in Theorem 2.16. First we need a simplicial approximation of $g \circ f$. Notice that, using Lemma 1.2, $t(s(x)) \leq t(f(x)) \leq g(f(x))$. Thus $t \circ s$ is

a simplicial approximation of $g \circ f$ and $(g \circ f)_{q*} = (t \circ s)_{q*} \circ \chi_{q*} = t_{q*} \circ s_{q*} \circ \chi_{q*}$.
Proceeding in the same way as in the proof of the previous theorem; $\theta \circ s\colon |K^m| \to |L|$ simplicially approximates f as $\theta(s(x)) \leq \theta(f(x)) \leq f(x)$.

As a consequence:
$g_{q*} \circ f_{q*} = t_{q*} \circ \tilde{\chi}_{q*} \circ (\theta \circ s)_{q*} \circ \chi_{q*} = t_{q*} \circ \tilde{\chi}_{q*} \circ \theta_{q*} \circ s_{q*} \circ \chi_{q*} = t_{q*} \circ s_{q*} \circ \chi_{q*} = (g \circ f)_{q*}$ □

For the last theorem we need the following lemma, which we will not prove here, but some guiding steps can be found in [1].

Lemma 2.18. *Given two homotopic maps $f, g\colon |K| \to |L|$, we can find a barycentric subdivision K^m and a sequence of simplicial maps $s_1, \dots s_n\colon |K^m| \to |L|$ such that s_1 and s_n are simplicial approximations of f and g respectively and each pair s_i, s_{i+1} are close simplicial maps.*

Theorem 2.19. *Given two homotopic maps $f, g\colon |K| \to |L|$, their induced homology maps coincide. In other words, $f_{q*} = g_{q*}\colon H_q(K) \to H_q(L)$, $\forall q$.*

Proof. Consider the subdivision chain map $\chi\colon C(K) \to C(K^m)$, by Lemmas 2.15 and 2.18

$$f_{q*} = s_{1_{q*}} \circ \chi_{q*} = \cdots = s_{n_{q*}} \circ \chi_{q*} = g_{q*}$$ □

Corollary 2.20 (Invariance of homology groups under homotopy). *If two polyhedra $|K|$ and $|L|$ have the same homotopy type, then $H_q(K) \simeq H_q(L)$, $\forall q$.*

Proof. $|K|$ and $|L|$ have the same homotopy type so there exist $f\colon |K| \to |L|$ and $g\colon |L| \to |K|$ such that $f \circ g \simeq Id_{|L|}$ and $g \circ f \simeq Id_{|K|}$. Using Corollary 2.17 and Theorem 2.19,

$$Id_{H_q(L)} = (f \circ g)_{q*} = f_{q*} \circ g_{q*} \text{ and } Id_{H_q(K)} = (g \circ f)_{q*} = g_{q*} \circ f_{q*}.$$

Therefore, f_{q*} is an isomorphim for all q. □

We can affirm that any two triangulations of the same topological space $|K|$ and $|\tilde{K}|$ coincide in their homology groups. In fact, there is an homeomorphism between them, let us say $f\colon |K| \to |\tilde{K}|$. In particular f is a homotopy equivalence with homotopy inverse f^{-1}. Now the homology groups of a topological space can be discussed without specifying which triangulation is being used to calculate them.

Let us begin with S^n. Recall that by Corollary 2.5, $H_0(\Delta^{n+1}) \simeq \mathbb{Z}$ and $H_q(\Delta^{n+1}) = 0$, for $q > 0$ and $n \geq 0$. We will denote by Σ^n the subcomplex of Δ^{n+1} formed by the simplices lying in its boundary. In other words, Σ^n contains every simplex in Δ^{n+1} but the $(n+1)$-simplex. Clearly, Σ^n is a triangulation of S^n for $n \geq 0$.

(i) For the case $\Sigma^0 \simeq S^0$, we only have two non-connected points. Therefore, $H_0(\Sigma^0) \simeq \mathbb{Z} \oplus \mathbb{Z}$ and $H_q(\Sigma^0) = 0$.

(ii) Now, $H_0(\Sigma^n) \simeq H_0(\Delta^{n+1}) \simeq \mathbb{Z}$ and $H_q(\Sigma^n) \simeq H_q(\Delta^{n+1}) \simeq 0$ for $1 < q \leq n-1$ and $n > 0$ since they have the same simplices up to dimension n and the computation of the q-th homology group does not involve simplices of dimension greater than $q+1$. Lastly, Σ^n has no $(n+1)$-simplices, $H_n(\Sigma^n) \simeq Z_n(\Sigma^n) \simeq Z_n(\Delta^{n+1}) = B_n(\Delta^{n+1})$, due to $H_n(\Delta^{n+1}) = 0$. $B_n(\Delta^{n+1}) = \partial_{n+1} C_{n+1}(\Delta^{n+1})$ is generated by the boundary of the only $n+1$ simplex in Δ^{n+1}. Hence $H_n(\Sigma^n) \simeq \mathbb{Z}$. Besides, $H_q(\Sigma^n) = 0$ for $q > n$.

Sumarising:

$$\begin{cases} H_0(S^0) \simeq \mathbb{Z} \oplus \mathbb{Z} & \\ H_q(S^0) = 0 & q > 0 \end{cases} \quad \text{and if } n > 0 \begin{cases} H_q(S^n) \simeq \mathbb{Z} & q \in \{0, n\} \\ H_q(S^n) = 0 & 0 < q < n \text{ or } q > n \end{cases}$$

Theorem 2.21. *If $m \neq n$, then S^n and S^m are not of the same homotopy type.*

Proof. $H_m(S^m)$ and $H_m(S^n)$ are isomorphic only when $m = n$. □

The relevance of homology groups as topological invariant does not limit to Theorem 2.21. Chapter 9 in [1] is devoted to further theoretical applications of homology groups such as proving the *hairy ball theorem* or the *Euler-Poincaré formula*.

Before concluding this chapter and crossing over to the more interdisciplinary applications of simplicial homology we should mention the accepted interpretation of non-trivial homology groups as “holes”. It all stems from the first homology group of a surface. Bounding 1-cycles are just, if you excuse the repetition, closed curves which bound part of the surface. In contrast, non-bounding 1-cycles (which are not the boundary of any 2-chain) correspond to closed curves which do not bound any region. An easy example of this is the torus, whose first homology group is isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$: the coordinate curves of its usual parametrization are non-bounding curves which detect the two holes in a torus (the whole inside of it and the one in the center), see Figure 2.2.
The same reasoning is extended to higher dimensions so we end up referring to non-bounding q-cycles as q-dimensional holes.

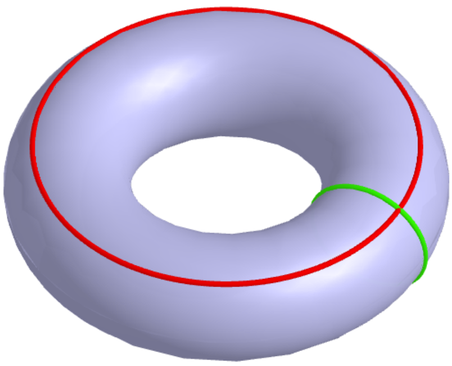


Figure 2.2: A torus with two closed curves representing non-bounding 1-cycles.

Chapter 3

Topological Data Analysis

A brief introduction into persistent homology

During the past two chapters we have been speaking about the theoretical aspects of simplicial homology: definitions, how to compute it for a simplicial complex or results we can prove with it. In this chapter, our focus changes to building simplicial complexes, which we will refer to as *filtrations*, out of clouds of points in $\mathbb{E}^n$. Afterwards an homological analysis of these filtrations will provide insight into the shape features of the original data. The natural question that quickly arises is: how should we construct such simplicial complexes? Given 13 points we could associate them with the closure of a 13-simplex or maybe just connect them with 1-simplices. Our approach is creating simplices out of points which are "close" enough.

3.1 From a point cloud to a simplicial complex

The two methods we will introduce interpret the distance condition in different ways while bearing complementary benefits.

Definition ($\mathcal{C}_\epsilon$, Čech complex).
Given $X = \{x_\alpha\}$ a finite set of points in $\mathbb{E}^n$ and $\epsilon > 0$, the *Čech complex* $\mathcal{C}_\epsilon$ is built as follows. Form closed[1] balls of radius $\frac{\epsilon}{2}$ around each point and if k balls have a non-empty intersection, its corresponding vertices will form a $(k-1)$-simplex in $\mathcal{C}_\epsilon$. Note that this construction is indeed a simplicial complex.

During this process, we might have lost information from the original data to the simplicial complex. Does this topological construction bear any resemblance with the original setting?

Theorem 3.1 (Čech theorem). *The Čech complex of a point cloud is homotopic equivalent to the union of closed $\frac{\epsilon}{2}$-balls centered on each set point. Hence, their homology groups are isomorphic.*[2]

So in fact, this simplicial complex is a topologically faithful simplicial model for the topology of a point cloud fattened by balls. On the other hand, it requires a vast number of operations for its computation, making it extremely unefficient for large sets of points. The next filtration relaxes the computational cost while weakening the topological relation.

[1]Some authors, like Afra in [12], define them as open balls, however it does not affect our theory.
[2]This theorem is also known as the Nerve theorem or Leray´s theorem whose proof can be found in [13].

Definition ($\mathcal{R}_\epsilon$, Rips complex)**.**
Given $X = \{x_\alpha\}$ a finite set of points in $\mathbb{E}^n$ and $\epsilon > 0$, the *Rips complex* $\mathcal{R}_\epsilon$ is built as follows. Form closed balls of radius $\frac{\epsilon}{2}$ around each point and if k balls have **pairwise** non-empty intersection, its corresponding vertices will form a $(k-1)$-simplex in $\mathcal{R}_\epsilon$.

We could actually approach this filtration process by obtaining a graph where two vertices are joint by an edge if their corresponding balls overlap. This way, k-simplices in $\mathcal{R}_\epsilon$ are associated with $(k+1)$-cliques. It stems directly from the definition the relation $\mathcal{C}_\epsilon \subset \mathcal{R}_\epsilon$. As promised, the Rips filtration no longer retains any topological similarity with the data cloud. For example, in Figure 3.1, $\mathcal{C}_\epsilon$ is homotopic to $S^1 \vee S^1 \vee S^1$ while $\mathcal{R}_\epsilon$ has the homotopy type of $S^1 \vee S^2$ ($X \vee Y$ denotes the wedge sum of X and Y). Using Rips complexes implies sacrificing accuracy for computability. Nevertheless, we can stablish an inclusion relation between Rips and Čech complexes which will allow us to dispense with the latter.

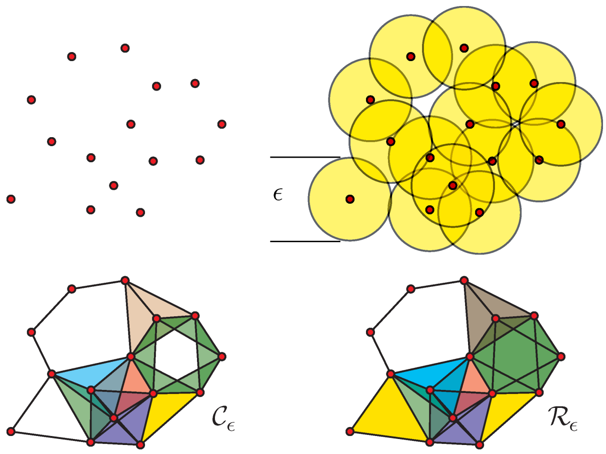


Figure 3.1: Illustration of $\mathcal{C}_\epsilon$ and $\mathcal{R}_\epsilon$ for a data cloud and some $\epsilon > 0$. Image credit to [11].

Proposition 3.2. *Let X be a set of points contained in $\mathbb{E}^d$ and fix $\epsilon > 0$, we have the following chain of inclusions.*

$$\mathcal{R}_\epsilon \subseteq \mathcal{C}_{\epsilon'} \subseteq \mathcal{R}_{\epsilon'} \text{ where } \epsilon' \geq \epsilon\sqrt{\frac{2d}{d+1}}$$

Or in general, $\mathcal{R}_\epsilon \subseteq \mathcal{C}_{\epsilon\sqrt{2}} \subseteq \mathcal{R}_{\epsilon\sqrt{2}}$.

The proof can be found in [14]. As a consequence, studying Rips complexes is enough. If one simplex appears in both $\mathcal{R}_\epsilon$ and $\mathcal{R}_{\epsilon'}$ so it does in $\mathcal{C}_{\epsilon'}$ for some $\epsilon' \geq \epsilon\sqrt{2}$.
At this point one could wonder which ϵ captures best the topology of the data cloud. Small values generate a low dimensional complex meanwhile for ϵ sufficiently large the resulting complex is the closure of a single high dimensional simplex. For example, in Figure 3.2, the data cloud has been sampled from a planar annulus. In contrast, it seems like no choice of ϵ covers that situation. By the time small holes are removed from the interior of the annulus, the inner hole is covered.
Accordingly, a new tool must be introduced to extract some significant features of the cloud data out of a sequence of filtrations of it: persistent homology.

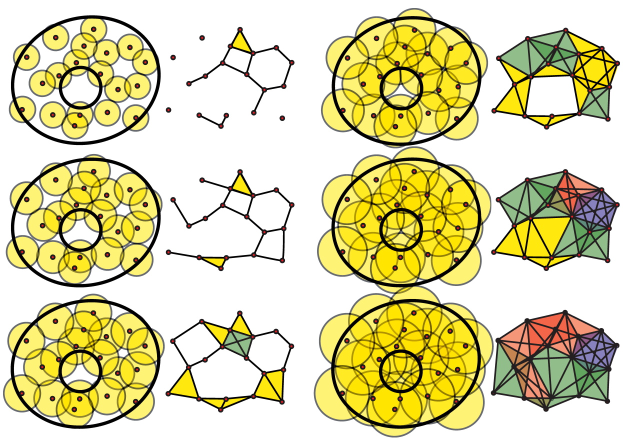


Figure 3.2: A sequence of Rips complexes for a data cloud obtained from a plannar annulus for increasing values of ϵ. Image credit to [11].

3.2 Persistent homology and barcodes

Despite being both computable and insightful, the homology of a complex associated to a point cloud at a specific ϵ is insufficient. It is a mistake to ask which value of ϵ is optimal. We are required to discern between what holes are real, hence present in a large interval of ϵ values, and what holes are mere noise, appearing in brief intervals. In this section we will introduce *persistent homology* as an algebraic mechanism to discern which non-bounding cycles endure over long intervals of ϵ.

Definition (Filtered complex and persistence complex)**.**
A *filtered complex* is an increasing sequence of simplicial complexes and a *persistence complex* is a collection of chain complexes $\mathsf{C} = \{C^i\}_{i=1}^N$ with chain maps $f^i\colon C^i \to C^{i+1}$, denoted by $\{C^i, f^i\}$.

Definition (q-th homology of a persistence complex)**.**
Given a persistence complex $\{C^i, f^i\}$, we call the collection of $\{H_q(C^i)\}$ together with the induced homology maps $f^i_{q*}\colon H_q(C^i) \to H_q(C^{i+1})$, the q-th homology of C and denote it by $H_q(\mathsf{C}) = \{H_q(C^i), f^i_{q*}\}$.

We shall relate these definitions to our context. Let $(\epsilon_i)_{i=1}^N$ be an increasing sequence of parameters and denote by $\mathcal{R} = (\mathcal{R}^i)_{i=1}^N$ the sequence of Rips complexes associated to a fixed point cloud for the sequence of parameters. $\mathcal{R}$ is indeed a filtered complex. Each $\mathcal{R}^i$ is associated to a chain complex: $C(\mathcal{R}^i)$. Besides, we can define the inclusion maps $x\colon C_q(\mathcal{R}^i) \to C_q(\mathcal{R}^{i+1})$ (not indexed for notational simplicity) since every q-chain of $\mathcal{R}^i$ is also a q-chain of $\mathcal{R}^{i+1}$. It is trivial verifying that they are chain maps. Thus, all the chain complexes obtained from the Rips filtration along with the inclusion chain maps, determine the persistence complex $\{\mathcal{R}^i, x\}$. Finally, the collection of q-th homology groups $H_q(\mathcal{R}^i)$ along with the induced homology maps $x_*\colon H_q(\mathcal{R}^i) \to H_q(\mathcal{R}^{i+1})$ are the q-th homology of $\{\mathcal{R}^i, x\}$, denoted by $H_q(\mathcal{R}) = \{H_q(\mathcal{R}^i), x_*\}$.

$$
\begin{array}{ccccccccccc}
 & & \cdots & & \cdots & & \cdots & & & & \cdots \\
 & & \downarrow x & & \downarrow x & & \downarrow x & & & & \downarrow x_* \\
\cdots & \xrightarrow{\partial_{q+2}} & C_{q+1}(\mathcal{R}^{i-1}) & \xrightarrow{\partial_{q+1}} & C_q(\mathcal{R}^{i-1}) & \xrightarrow{\partial_q} & C_{q-1}(\mathcal{R}^{i-1}) & \xrightarrow{\partial_{q-1}} & \cdots & & H_q(\mathcal{R}^{i-1}) \\
 & & \downarrow x & & \downarrow x & & \downarrow x & & & & \downarrow x_* \\
\cdots & \xrightarrow{\partial_{q+2}} & C_{q+1}(\mathcal{R}^{i}) & \xrightarrow{\partial_{q+1}} & C_q(\mathcal{R}^{i}) & \xrightarrow{\partial_q} & C_{q-1}(\mathcal{R}^{i}) & \xrightarrow{\partial_{q-1}} & \cdots & & H_q(\mathcal{R}^{i}) \\
 & & \downarrow x & & \downarrow x & & \downarrow x & & & & \downarrow x_* \\
\cdots & \xrightarrow{\partial_{q+2}} & C_{q+1}(\mathcal{R}^{i+1}) & \xrightarrow{\partial_{q+1}} & C_q(\mathcal{R}^{i+1}) & \xrightarrow{\partial_q} & C_{q-1}(\mathcal{R}^{i+1}) & \xrightarrow{\partial_{q-1}} & \cdots & & H_q(\mathcal{R}^{i+1}) \\
 & & \downarrow x & & \downarrow x & & \downarrow x & & & & \downarrow x_* \\
 & & \cdots & & \cdots & & \cdots & & & & \cdots
\end{array}
$$

Definition ((i,j)-persistent q-th homology group of a persistence complex)**.**
For $i < j$, the (i,j)-*persistent q-th homology group of a persistence complex* $\mathsf{C} = \{C^i\}_{i=1}^N$ is

$$H_q^{i \to j}(\mathsf{C}) = \frac{Z_q^i(\mathsf{C})}{B_q^j(\mathsf{C}) \cap Z_q^i(\mathsf{C})}$$

It is well-defined since both groups in the denominator are subgroups of C_q^j, so their intersection is a group contained in $Z_q^i(\mathsf{C})$, hence a subgroup.

Alternatively we could define them as the image of the induced homomorphism by x, $x_*\colon H_q(C^i) \to H_q(C^j)$ defined by $x_*([z]) = x_*(z + B_q^i(\mathsf{C})) = x(z) + B_q^j(\mathsf{C}) = z + B_q^j(\mathsf{C}) = [z]$. Hence $\operatorname{Im} x_* \simeq H_q^{i\to j}(\mathsf{C})$.

The idea is simple, we want to "project" the q-th homology classes of the i-th chain complex of C (in our setting, $C_q(\mathcal{R}^i)$ from the persistence complex $\mathcal{R}$) onto the q-th homology class of the j-th chain complex of C, as a means of tracking the lifetime of homology classes along the filtration.
There is one last obstacle preventing us from using persistent homology groups: intuitively, the computation of persistence requires compatible bases for $H_q^i(\mathsf{C})$ and $H_q^{i\to j}(\mathsf{C})$. In general this is not satisfied, but for specific cases specified later, it is possible. A new algebraic construction is required.

Definition (Graded ring and graded module).
A *graded ring* is a ring $(R, +, \cdot)$ equipped with a decomposition as direct sum of abelian groups, i.e. $R \simeq \bigoplus_i R_i,\ i \in \mathbb{Z}$ where multiplication is given by bilinear correspondences $R_i \otimes R_j \to R_{i+j}$. Elements lying in a single R_i are called *homogeneous* and have degree i. One easy example of a graded ring is the polynomial ring $R[t]$ graded non-negatively with the *standard grading*; $R_i = Rt^i,\ i \geq 0$. Elements $2t$ and t^2 are homogeneous, but not their sum.
A *graded module* M over a ring R is a R-module equipped with a direct sum decomposition $M \simeq \bigoplus_i M_i,\ i \in \mathbb{Z}$ where the action of R is given by bilinear pairings $R_i \otimes M_j \to M_{i+j}$.
A graded ring (R-module) is *non-negatively graded* if $R_i = 0$ ($M_i = 0$) for all $i < 0$.

Theorem 3.3 (Structure Theorem for PID's). *If R is a PID, then every finitely generated R-module is isomorphic to a direct sum of a finitely generated free R-module and cyclic R-modules.*

$$R^\beta \oplus \left(\bigoplus_{i=1}^m \frac{R}{r_i \cdot R} \right)$$

where $\beta \in \mathbb{Z}$ and $r_i \in R$ such that $r_i | r_{i+1}$. Similarly, every graded R-module over a graded PID R decomposes uniquely as a direct sum.

$$\left(\bigoplus_{i=1}^n \Sigma^{\alpha_i} \cdot R \right) \oplus \left(\bigoplus_{j=1}^m \Sigma^{\gamma_j} \cdot \frac{R}{r_j \cdot R} \right)$$

where r_j are homogeneous such that $rj | r_{j+1}$ $\alpha_i, \gamma_j \in \mathbb{Z}$ and Σ^α denotes an α-shift upward in grading.

This theorem allows us to think of finitely generated modules and graded modules as structures that look like vector spaces with some extra dimensions that are "finite" in size. We defined homology groups as abelian groups which are $\mathbb{Z}$-modules. Observe that there would be no contradiction in defining homology groups as finitely generated R-modules. Moreover, if R were a PID, Theorem 3.3 would give us the same type of decomposition we knew for finitely generated abelian groups.

Definition (Persistence module).
A *persistence module* $\mathcal{M}$ is a family of R-modules M^i together with homomorphisms $\varphi^i\colon M^i \to M^{i+1}$. We denote it by $\{M^i, \varphi^i\}$.

$H_q(C^i)$ can be regarded as R-modules, henceforth $H_q(\mathsf{C}) = \{H_q(C^i), f^i_{q*}\}$ are persistence modules for all q.

Suppose that we have a persistence module $\mathcal{M} = \{M^i, \varphi^i\}_{i=1}$ over a ring R. We equip the polynomial ring $R[t]$ with the standard grading and define a graded module over $R[t]$ by

$$\alpha(\mathcal{M}) = \bigoplus_{i=1}^{\infty} M^i$$

where the action of t is given by $t(m^1, m^2, m^3, \dots) = (0, \varphi^1(m^1), \varphi^2(m^2), \varphi^3(m^3), \dots)$, in other words, t shifts elements of the module up in the graduation.
Choosing a field F to be the coefficients of homology groups, we can construct the above construction for the persistence modules $H_q(\mathsf{C}, F)$ for every q. Besides, $F[x]$ is a PID allowing us to use Theorem 3.3,

$$H_q(\mathsf{C}, F) \simeq \left(\bigoplus_{i=1}^{n} x^{t_i} \cdot F[x] \right) \oplus \left(\bigoplus_{j=1}^{m} x^{r_j} \cdot \frac{F[x]}{x^{s_j} \cdot F[x]} \right)$$

From this classification we can extract all the data from the persistent homology. The free portions reveal that there are n homology generators each of them coming into existence at the parameter t_i and persisting for all the parameter values. The torsion part describes those m homology generators appearing at parameter value r_j and disappearing at $r_j + s_j$. This information can be encapsuled in time intervals. A non-vanishing generator is associated with the interval $[\epsilon_{t_i}, \infty)$ meanwhile torsion elements are associated with $[\epsilon_{r_j}, \epsilon_{r_j+s_j}]$. A graphical representation of these intervals from every persistence module in a given persistence complex is known as a *barcode*.

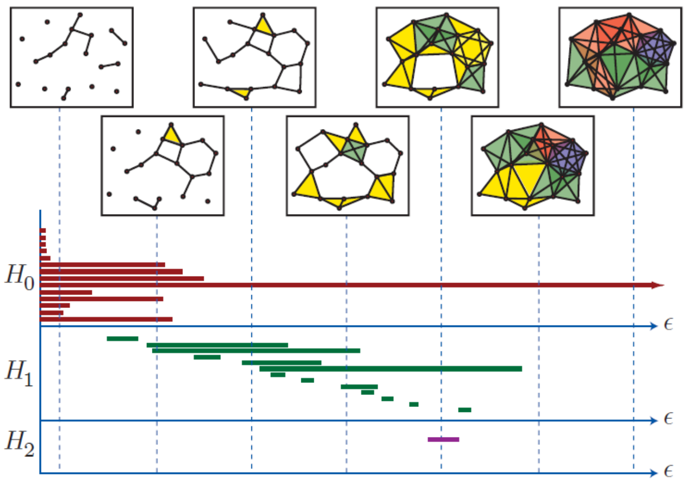


Figure 3.3: An example of a barcode for every $H_q(\mathcal{R})$, where $\mathcal{R}$ is the same filtration of Figure 3.2. The horizontal axis corresponds to the parameters and the vertical one represents an arbitrary ordering of homology generators. The point cloud likely represents a connected object with one or two significant one-dimensional holes. Image credit to [11].

How are we supposed to interpret a barcode? We would like to filter out noise from the relevant features of a data set. Therefore, long bars represent significant holes while short ones are described as "topological noise".

What is the relation between persistence modules with their associated barcode, and persistent homology groups? After all, we developed this graded $F[x]$-module structure to convey the meaning of persistent homology groups.

Theorem 3.4 (Fundamental Theorem of Persistent Homology [15], [17]). *The rank of the persistent homology group $H_q^{i\to j}(\mathsf{C}, F)$ is equal to the number of intervals in the barcode of $H_q(\mathsf{C}, F)$*

spanning the parameter interval $[\epsilon_i, \epsilon_j]$.

The existence of this theorem resides in F being a field, otherwise we could have not used the classification of graded modules over a PID (read more about this topic in [15]). As a result of the theorem, barcodes completely determined up to permutations of the bars. According to [15], the algorithms used to find the intervals of a filtration avoid computing the $F[x]$-module.

Remark

One question that might haunt readers is which line should be continued in a barcode when two different generators merge in the same homology class for the next parameter. The answer is simple, that scenario can be avoided. Theorem 3.4 and the decomposition of $H_q(\mathsf{C}, F)$ ensures that compatible bases can be chosen so that out of all the homologous cycles for the next parameter, only one homology class is not mapped to zero by x_*.

A main characteristic of persistent homology is its stability. It is stable with respect to small perturbations in the input data [17]. That is to say that given a "slightly" modified point cloud, persistent homology will output the same result.

Persistent homology is a relatively new field with undergoing research. Experts are applying topological data analysis in a wide variety of real-world issues ranging from jaw treatments [18] to money laundering investigation and many more [17].

Bibliography

[1] M.A. Armstrong, *Basic Topology*, Springer, New York, 1983.

[2] kglr, in a post from *Mathematica Stack Exchange*, available at
https://mathematica.stackexchange.com/questions/57829/torus-triangulation

[3] A. Atarés de Sus, *Classic Problems in Topology and Geometry*, available at
https://zaguan.unizar.es/record/97953/files/TAZ-TFG-2020-3256.pdf.

[4] D.L. Ferrario and R.A. Piccinini, *Simplicial Structures in Topology*, Springer, 2011.

[5] C.R.F. Maunder, *Algebraic Topology*, Cambridge University Press, Cambridge, 1980.

[6] M. Aanjaneya and M. Teillaud, *Triangulating the Real Projective Plane*, available at
http://physbam.stanford.edu/~aanjneya/mridul_files/papers/real_projective.pdf.

[7] G.E. Bredon, *Topology and Geometry*, Springer, 1993

[8] P.A. Grillet, *Abstract Algebra*, Springer, 2007

[9] P. Giblin, *Graphs, Surfaces and Homology*, Cambridge University Press, 2010

[10] F.H. Croom, *Basic concepts of algebraic topology*, Springer, New York, 1978.

[11] R. Ghrist, *Barcodes: The persistent topology of data*, Bulletin of the American Mathematical Society, Volume 45, Number 1, Pages 61-75, 2007, available at
https://www2.math.upenn.edu/~ghrist/preprints/barcodes.pdf.

[12] A. Zomorodian, *Topological Data Analysis*, AMS, 2011, available at
https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.220.7411&rep=rep1&type=pdf.

[13] J.J Rotman, *An introduction to algebraic topology*, Springer-Verlag, New York, 1988.

[14] V. de Silva and R. Ghrist, *Coverage in sensor networks via persistent homology*, available at https://www2.math.upenn.edu/~ghrist/preprints/persistent.pdf.

[15] A. Zomorodian and G. Carlsson, *Computing persistent homology*, 2005, available at
https://geometry.stanford.edu/papers/zc-cph-05/zc-cph-05.pdf.

[16] H. Edelsbrunner,D. Letscher and A. Zomorodian, *Topological Persistence and Simplification*, Discrete Comput. Geom. 28, 511–533, 2002 available at
http://www.math.uchicago.edu/~shmuel/AAT-readings/Data%20Analysis%20/Edelsbrunner-Letscher-Zomordian.pdf.

[17] N. Otter, M.A. Porter, U.Tillmann, P. Grindrod and H.A. Harrington, *A roadmap for the computation of persistent homology*, EPJ Data Science **6**, 17, 2017, available at https://epjdatascience.springeropen.com/track/pdf/10.1140/epjds/s13688-017-0109-5.pdf.

[18] H. Edelsbrunner and D. Morozov, *Persistent Homology: Theory and Practice*, 2013, available at https://pub.ist.ac.at/~edels/Papers/2012-P-11-PHTheoryPractice.pdf.

[19] J. Trzeciak, *Writing Mathematical Papers in English*, European Mathematical Society, 1995.