MAYER VIETORIS SEQUENCE FOR PERSISTENT HOMOLOGY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

YAĞMUR YILMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MATHEMATICS

JULY 2018

Approval of the thesis:

MAYER VIETORIS SEQUENCE FOR PERSISTENT HOMOLOGY

submitted by **YAĞMUR YILMAZ** in partial fulfillment of the requirements for the degree of **Master of Science in Mathematics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Yıldıray Ozan
Head of Department, **Mathematics**

Assoc. Prof. Dr. Mehmetcik Pamuk
Supervisor, **Mathematics Department, METU**

Examining Committee Members:

Assoc. Prof. Dr. Mesut Şahin
Mathematics Department, Hacettepe University

Assoc. Prof. Dr. Mehmetcik Pamuk
Mathematics Department, METU

Assoc. Prof. Dr. Mehmet Fırat Arıkan
Mathematics Department, METU

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: YAĞMUR YILMAZ

Signature :

ABSTRACT

MAYER VIETORIS SEQUENCE FOR PERSISTENT HOMOLOGY

Yılmaz, Yağmur
M.S., Department of Mathematics
Supervisor : Assoc. Prof. Dr. Mehmetcik Pamuk

July 2018, 42 pages

Persistent homology is an algebraic method for understanding topological features of discrete objects or data (finite set of points with metric defined on it). In algebraic topology, the Mayer Vietoris sequence is a powerful tool which allows one to study the homology groups of a given space in terms of simpler homology groups of its subspaces. In this thesis, we study to what extent does persistent homology benefit from Mayer Vietoris sequence.

Keywords: Persistent Homology, Mayer-Vietoris Sequence, Persistence Module

ÖZ

KARARLI HOMOLOJİ İÇİN MAYER VİETORİS DİZİSİ

Yılmaz, Yağmur
Yüksek Lisans, Matematik Bölümü
Tez Yöneticisi : Doç. Dr. Mehmetcik Pamuk

Temmuz 2018 , 42 sayfa

Kararlı homoloji ayrık nesnelerin veya verilerin (üzerinde metrik tanımlanan sonlu noktalar kümesi) topolojik özelliklerini anlamak için cebirsel bir metottur. Cebirsel topolojide Mayer Vietoris dizisi, verilen uzayı homoloji grupları daha kolay hesaplanabilir altuzaylara ayırarak verilen uzayın homoloji gruplarının hesaplanmasında kullanılan güçlü bir araçtır. Biz bu tezde, kararlı homolojinin Mayer Vietoris dizisinden ne derece yarar sağladığını araştıracağız.

Anahtar Kelimeler: Kararlı Homoloji, Mayer-Vietoris Dizisi, Kararlı Modül

To my family

ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my supervisor, Assoc. Prof. Dr. Mehmetcik Pamuk, for his invaluable guidance, encouragement, and endless support throughout my thesis.

Also, I would like to thank Neza Mramor Kosta for her support and also Prof. Dr. Mustafa Turgut Önder for many helpful conversations. I would also like to thank Assoc. Prof. Dr. Semra Pamuk and Dr. Hanife Varlı for their support during our TÜBİTAK project.

I would like to thank TÜBİTAK (1001 Project No:117F015) and Middle East Technical University (Scientific Research Project BAP-01-01-2017-003) for the financial support.

My special thanks to my mother Funda Güler, my father Mustafa Güler, and my sister Göksu Güler without whose encouragement, patience, and support, this thesis would not be possible.

Lastly, but not least, I would like to express my deepest thanks to my husband Kürşat Yılmaz for his patience and support throughout this thesis.

TABLE OF CONTENTS

CHAPTERS

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Natural data is hard to analyse and one of tool sets developed for this difficult task is persistent topology. It offers new topological descriptors (e.g. persistent homology, persistent Betti numbers, persistence diagrams) which preserve some selected geometric features through filtering functions by considering the filtration obtained from the sequence of nested level sets of the function under study by encoding the scale at which a topological feature (e.g., a connected component, a tunnel, a void) is created, and when it vanishes along the filtration.

Persistent homology is an algebraic method to determine the features of a topological space by using a scalar function or to recognize topological properties of a data set by first constructing suitable finite simplicial complex structures on the set. Persistent homology is worked in detail for the usual homology groups in general with coefficients from a field which are important for practical purposes.

The history of persistent homology goes back to 1990s. Patrizio Frosini and his collaborators [14] introduced size functions which is now known as 0-dimensional persistent homology. Then, in 1999, Vanessa Robins [21] worked on homology of sampled spaces and described the images of homomorphisms induced by inclusions same as persistent homology groups. In 2002, Edelsbrunner, Letscher and Zomordian [10], defined persistent homology groups, persistent Betti numbers, persistence diagrams and an efficient (computationally cheap) algorithm for the filtrations of alpha shapes in the case of $\mathbb{Z}/2$ coefficients. Afterwards, Zomordian and Carlsson [27] generalized this algorithm to arbitrary fields and used graded $\mathbb{F}[t]$-module structure to give an algebraic description of persistence. Moreover, they showed that the module decomposition determines all persistent Betti numbers uniquely. The relation between persistence and quiver theory was given by Carlsson and de Silva [4] which leads to zigzag persistence. In the mean time, Robert Ghrist [15], published a paper in which barcode systems, an equivalent way of representing the persistence diagrams, are introduced for persistent homology.

The stability of the persistence diagram of a function on a topological space, introduced by Cohen-Steiner, Edelsbrunner and Harer [5], is one of the essential tools used in persistent homology. Stability in this context means that, if there is a small perturbation on the function or data then the amount of change in the persistence diagrams will be small as well. In slightly more details and technical terms; measuring the distance between two functions f and g, by the L_∞-norm and between the corresponding persistence diagrams, $Dgm(f)$ and $Dgm(g)$, by the bottleneck distance, the stability result bounds the latter distance by the former, i.e.,

$$d_b(Dgm(f), Dgm(g)) \leq \|f - g\|_\infty .$$

This inequality is applied in particular to two specific problems: The first is the estimation of the homology of a closed subset of a metric space from a finite point sample, while the second problem is the comparison and classification of geometric shapes.

In 2009, after the paper of Carlsson [3], the importance of the persistent homology is revealed and it became the major tool in data analysis. This in turn leads to a new research area, which is now called topological data analysis(TDA). TDA is a field connecting different areas together such as algebraic topology, computer science, computational geometry and statistics. The aim of TDA is to develop new techniques to analyze qualitative features of data by using topological and geometric approaches. There are many promising and beneficial results in the application side of TDA in an increasing number of areas such as material science [17], [19], 3D shape analysis [22], [23], sensor networks [6], three-dimensional structure of DNA [11], robotics [2], [24], [20], proteins [26], which flaunts the importance of TDA one more time.

Topological shape analysis is one of the main problems in TDA. To have an efficient algorithm to analyze the shapes, we use persistent homology instead of classical homology. Since persistent homology reflects the changes in homology which evolves with respect to a parameter, it gives more significant results about the shape of the object. To make the algorithm provide consistent results, it is necessary to have some invariants. However, the persistence of the classical algebraic invariants of topological spaces is not yet well understood. Moreover, it is not clear yet whether the methods to determine algebraic invariants work also for persistent homology. One such method is the Mayer-Vietoris sequence which allows one to study the homology groups of a given space in terms of relatively simpler homology groups of its subspaces.

Mayer Vietoris sequence is an important tool in algebraic topology for computing homology groups of a topological space. The question we are going to address in this thesis is whether it is applicable in the persistent homology setting as well. We show that, persistent homology has a Mayer-Vietoris sequence that in general is not exact but only of order two i.e., rather than being equal, we only have $\mathrm{Im} \subseteq \ker$.

In this thesis, we show that the Mayer-Vietoris sequence is not exact if we are working with the persistent homology groups [7]. To better explain this situation, we give a concrete example for which the Mayer-Vietoris sequence is not exact. However, we show that the Mayer-Vietoris sequence is exact when we use the persistent modules instead of the persistent homology groups.

The organization of this thesis is as follows:

In Chapter 2, we introduce basic definitions and ideas related with persistent homology which will be used in the following chapters.

In Chapter 3, we construct a particular example to examine the Mayer-Vietoris sequence for persistent homology. We see that one of the problems why the Mayer-Vietoris sequence being not exact is the classes that are born and die throughout the filtration. To minimize such classes we used a perfect discrete Morse function which gives minimal number of critical cells but we have observed that the corresponding sequence is still not exact. On the other hand, we show that the Mayer-Vietoris sequence is exact if we work with persistent modules.

CHAPTER 2

PERSISTENT HOMOLOGY

2.1 Some Basic Definitions in Algebraic Topology

In this chapter, we start with briefly reviewing some basic concepts of algebraic topology and discrete Morse theory. For further details on these topics see [16], [12]. Then we introduce persistent homology in detail.

Definition 1. *The standard p-**simplex**, denoted by Δ^p, is the simplest geometric figure determined by a collection of $(p+1)$ points in some Euclidean space $\mathbb{R}^n$. Geometrically, one can think of it as the complete graph on $(p+1)$ vertices. Formally, a p-simplex in $\mathbb{R}^n$ is defined to be the set of the form*

$$\{\sum_{j=0}^{p} t_j u_j : 0 \leq t_j \leq 1 \, for \, j = 0, 1, \ldots, p \;\; and \;\; \sum_{j=0}^{p} t_j = 1\}$$

where $u_0, u_1, \ldots, u_p$ are geometrically independent points of $\mathbb{R}^n$. The points $u_0, \ldots, u_p$ are referred to as the vertices of the simplex. The non-negative integer p is referred to as the dimension of the simplex.

Definition 2. ***A finite simplicial complex** K is a collection of simplices such that whenever $\sigma \in K$ and τ is a face of σ (denoted by $\tau \leq \sigma$) then $\tau \in K$. Also, if $\sigma_1, \sigma_2 \in K$, then we have $\sigma_1 \cap \sigma_2$ is either empty or a face of both.*

Definition 3. *A linear combination of p-simplices in a simplicial complex K is called a p-chain, which is shown by $c = \sum a_i \sigma_i$, where each σ_i is a p-simplex. Throughout this thesis, we mostly work with $\mathbb{Z}/2$ coefficients, that is, $a_i \in \mathbb{Z}/2$.*

Moreover, the p-chains form a group under addition $'+'$ where the identity is the 0-chain $0 = \sum_{i=1}^{p} 0\sigma_i$, and the inverse of a chain c is c itself since $c + c = 0$ (working *with $\mathbb{Z}/2$ coefficients). This group is called the p-th chain group which is denoted by $C_p := C_p(K)$. Also, this operation is associative since addition in $\mathbb{Z}_2$ is associative. Also, C_p is abelian because addition modulo 2 is abelian.*

Definition 4. *The sum of $(p-1)$-dimensional faces of a p-simplex is called its boundary. In other words, let $\sigma = [u_0,u_1,\dots,u_p]$ be a p-simplex with the listed vertices and then **the boundary of** σ is given by*

$$\partial_p \sigma = \sum_{j=0}^{p} [u_0, \dots, \hat{u}_j, \dots, u_p].$$

Here, $\hat{u}_j$ component is omitted. If we have a p-chain $c = \sum a_i \sigma_i$, then boundary of c is defined as $\partial_p c = \sum a_i \partial_p \sigma_i$. Therefore, the boundary homomorphism can be written as $\partial_p : C_p \to C_{p-1}$.

It can be easily seen that $\partial_p(c + c^{'}) = \partial_p c + \partial_p c^{'}$. From this, we can induce that boundary operator, defined as above, gives a group homomorphism between the chain groups. So, ∂_p is called the boundary map for chains. The sequence of chain groups together with boundary homomorphisms is called the chain complex,

$$\cdots \xrightarrow{\partial_{p+2}} C_{p+1} \xrightarrow{\partial_{p+1}} C_p \xrightarrow{\partial_p} C_{p-1} \xrightarrow{\partial_{p-1}} \cdots .$$

Definition 5. *If a p-chain $c = \sum a_i \sigma_i$ has an empty boundary, i.e., $\partial c = 0$, then it is called a p-**cycle**. Since ∂ commutes with addition, p-cycle forms a subgroup of C_p which can be defined to be the kernel of the p-th boundary homomorphism denoted as $Z_p := Z_p(K) = \ker \partial_p$.*

Definition 6. *If a p-chain $c = \sum a_i \sigma_i$ is the boundary of a $(p+1)$-chain $d \in C_{p+1}$, i.e., $c = \partial d$, then it is called a p-**boundary**. Since ∂ commutes with addition, p-boundaries form a subgroup of C_p which can be defined as the image of the $(p+1)$-st boundary homomorphism. It is denoted as $B_p := B_p(K) = \operatorname{Im} \partial_{p+1}$.*

Definition 7. ***The p-th homology group** is defined to be the quotient group $H_p = Z_p/B_p$. The rank of this group is called p-th Betti number and it is denoted as $\beta_p = rank H_p$.*

2.2 Discrete Morse Theory

In our example, we are going to use discrete Morse functions to filter a given simplicial complex. In the following subsection we will quickly review the basics of discrete Morse theory, for further details see [12], [13].

Definition 8. *Let K be any finite simplicial complex and $\tau \geq \sigma$ if σ is a face of τ. We say that $f\colon K \to \mathbb{R}$ is a* ***discrete Morse function*** *if for every $\sigma^{(p)} \in K$ where p is the dimension of σ, we have*

(i) $m_1 = \#\{\tau^{(p+1)} > \sigma : f(\tau) \leq f(\sigma)\} \leq 1$

(ii) $m_2 = \#\{\gamma^{(p-1)} < \sigma : f(\gamma) \geq f(\sigma)\} \leq 1.$

Theorem 2.2.1. *The numbers m_1 and m_2 cannot be one at the same time.*

Proof. Let us assume that $m_1 = m_2 = 1$. Then there exists a $(p-1)$-cell γ and a $(p+1)$-cell τ for the p-cell $\sigma \in K$ such that $f(\gamma) \geq f(\sigma)$ and $f(\tau) \leq f(\sigma)$. Therefore, $f(\tau) \leq f(\sigma) \leq f(\gamma)$. Since K is a finite simplicial complex, there is at least one p-cell $\tilde{\sigma} \neq \sigma$ such that $\gamma^{(p-1)} < \tilde{\sigma} < \tau^{(p+1)}$. By definition 8, we get $f(\gamma) < f(\tilde{\sigma})$ and $f(\tilde{\sigma}) < f(\tau)$, this is a contradiction. Hence the numbers m_1 and m_2 cannot be one at the same time. $\square$

Definition 9. *A simplex σ is called* ***critical*** *for the discrete Morse function f, if*

(i) $\#\{\tau^{(p+1)} > \sigma : f(\tau) \leq f(\sigma)\} = 0,$

(ii) $\#\{\gamma^{(p-1)} < \sigma : f(\gamma) \geq f(\sigma)\} = 0.$

Definition 10. *Let f be a discrete Morse function on K and $t \in \mathbb{R}$.* ***The sublevel complex*** *$K(t)$ is defined as follows:*

$$K(t) = \bigcup_{f(\tau) \leq t} \bigcup_{\alpha \leq \tau} \alpha.$$

That is to say, $K(t)$ consists of all cells τ, where $f(\tau) \leq t$, together with all of their faces.

Lemma 2.2.2. *Let K be a simplicial complex with a discrete Morse function f on it and L be a subcomplex of K. Then the restriction of f is also a discrete Morse function on L such that if $\alpha \in L$ is a critical cell of f, then it is a critical cell for the restriction function.*

We note that the above lemma is used to show that the discrete Morse function which is used in the example 5 is still a discrete Morse function on the subcomplexes of the given torus.

Let m_p denote the number of critical simplices of dimension p. Let $\mathbb{F}$ be any field, and $b_p = dim H_p(K, \mathbb{F})$ the p-th Betti number with respect to $\mathbb{F}$.

Theorem 2.2.3. *[13][The Weak Morse Inequalities] [Theorem 1.7]*

For each $p = 0, 1, 2, \ldots, n$ (where n is the dimension of K)

$$m_p \geq b_p$$

and

$$m_0 - m_1 + m_2 - \cdots + (-1)^n m_n = b_0 - b_1 + b_2 - \cdots + (-1)^n b_n.$$

Theorem 2.2.4. *[13][The Strong Morse Inequalities] [Theorem 1.8]*

For each $p = 0, 1, 2, \ldots, n, n+1$,

$$m_p - m_{p-1} + \cdots + (-1)^p m_0 \geq b_p - b_{p-1} + \cdots + (-1)^p b_0.$$

Definition 11. *A discrete Morse function on K is* ***perfect*** *(with respect to $\mathbb{F}$) if the Morse inequalities are equalities.*

Note that for a perfect discrete Morse function, the number of critical cells and the Betti numbers are equal in each dimension.

Definition 12. ***A discrete gradient vector field*** *V induced by a discrete Morse function f on K is a disjoint collection of pairs $(\sigma^{(p)}, \tau^{(p+1)})$ of cells such that $\sigma < \tau$ and $f(\sigma) \geq f(\tau)$. A cell σ is critical if it is not paired with any other cell.*

Instead of working with a given discrete Morse function, it is always easier to work with the discrete gradient vector field which comes from this function. The reason is that, the discrete gradient vector field has all the information that the discrete Morse function gives us, and the use of arrows instead of numbers is more useful in many situations.

Suppose $\sigma^{(p)}$ is a non-critical simplex with $\sigma \leq \tau^{(p+1)}$ satisfying $f(\tau) \leq f(\sigma)$, then we draw an arrow from σ to τ. By theorem 2.2.1, every simplex α satisfies exactly one of the following:

(i) α is the tail of exactly one arrow.

(ii) α is the head of exactly one arrow.

(iii) α is neither the head nor the tail of an arrow.

Note that a simplex is critical if and only if it is neither the head nor the tail of an arrow.

2.3 Persistent Homology

Definition 13. *Let K be a simplicial complex and f be a height function from K to $\mathbb{R}$. For any real number a, the preimage $f^{-1}(a)$ will give a **level set** that consists of all the elements of K having the height a. Then the **sublevel set** can be defined to be all the points of K having height less than or equal to a,*

$$K_a := f^{-1}(-\infty, a] = \{x \in K : f(x) \leq a\}.$$

Definition 14. *Let K be a simplicial complex. Define a monotonic function $f\colon K \to \mathbb{R}$ by which we mean $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ. **A filtration of** f is a nested sequence of subcomplexes such as*

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K,$$

where $K_i := K_{a_i} = f^{-1}(-\infty, a_i]$ for each i and $a_1 < a_2 < \cdots < a_n$ are the function values.

Example 1. *Let K be the simplicial complex given in Figure 2.1 and let $a_1 < a_2 < \cdots < a_7$ denote different function values of f defined on K. The filtration given by this function is depicted in Figure 2.2.*

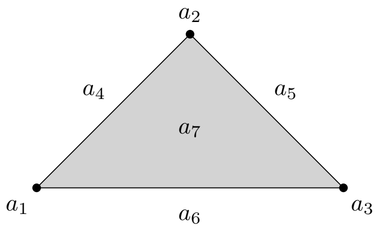


Figure 2.1: Simplicial complex K

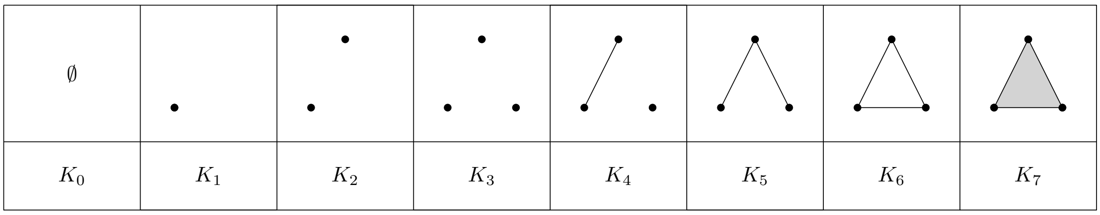


Figure 2.2: A filtration of f

Example 2. ***Lower Star Filtration*** *Let us consider the triangulation of a torus with a real valued function f defined at its vertices (see Figure 2.3) to give another example of a filtration, which in particular is called the lower star filtration. But, before our example, let us first define some preliminary notions:*

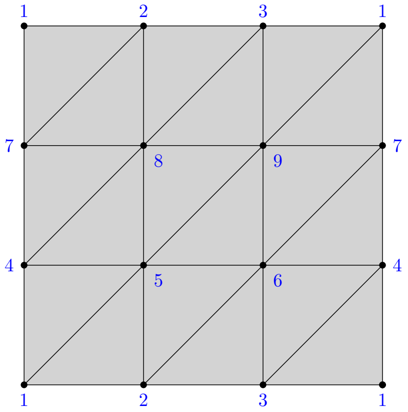


Figure 2.3: Triangulation of Torus

Let K be a simplicial complex. The ***star*** *of a simplex $\tau \in K$ is the subcomplex of K consisting of all cofaces of τ and it is denoted by*

$$St\tau = \{\sigma \in K \mid \tau \leq \sigma\}.$$

*The **link** of τ comprises of all simplices in the star of τ that are disjoint from τ denoted as*

$$Lk\tau = \{v \in St\tau \mid v \cap \tau = \emptyset\}.$$

In Figure 2.4, the stars and links (in red) of the vertices in Figure 2.3 are drawn.

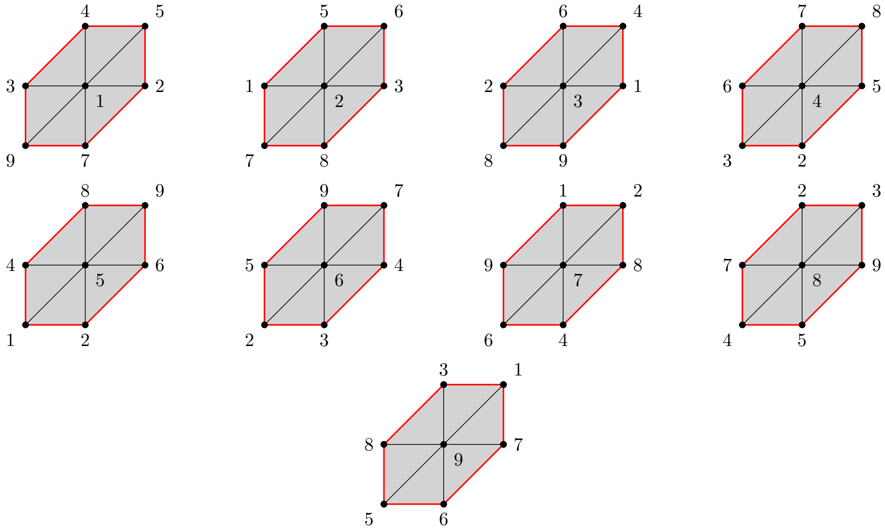


Figure 2.4: Stars of Vertices in 2.3

If we know the star and the link of a vertex, we can find the lower link and lower star of this vertex. The lower star of the vertex u_i, denoted by St_{u_i}, is the set of simplices for which $f(u_i)$ is the maximum function value,

$$St_{u_i} = \{\sigma \in Stu_i \mid x \in \sigma \Rightarrow f(x) \leq f(u_i)\}.$$

The lower link of the vertex u_i, denoted by Lk_{u_i}, is the collection of simplices in the star that do not belong to the lower star. Equivalently, it is the collection of simplices in the link whose vertices have smaller function value than u_i,

$$Lk_{u_i} = \{\sigma \in Lku_i \mid x \in \sigma \Rightarrow f(x) \leq f(u_i)\}.$$

For the triangulation of the torus given in Figure 2.3, the lower link and the lower star of vertices can be seen in Figures 2.5 and 2.6. Note that the lower stars of vertices is partition K. Furthermore, K_i is the union of the first i lower stars. When we go

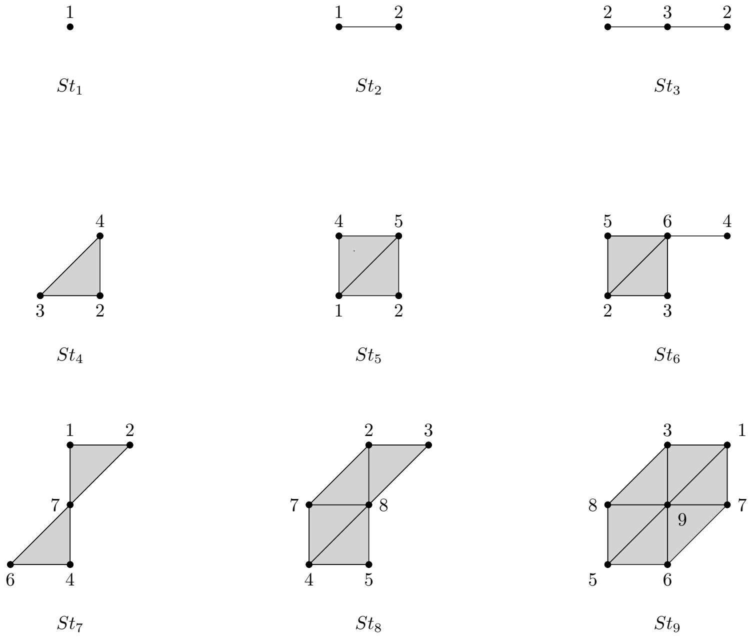


Figure 2.5: Lower Stars of Vertices in Figure 2.4

from K_{i-1} to K_i, we attach the lower star of u_i, gluing it along the lower link to the complex K_{i-1}. The nested sequence of complexes $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ is called the lower star filtration of f which is examplified in Figure 2.7

For any function defined on a simplicial complex, we can decide how many levels the filtration consists of. However, in the lower star filtration we get a filtration level for each vertex and each filtration level is obtained from the previous one. Therefore, this gives us a more consistent filtration.

We want to emphasize that instead of the sequence of complexes, our interest is concentrated on the topological evolution of the complex, which can be examined through the corresponding sequence of homology groups. Since $K_i \subseteq K_j$ for every $i \leq j$, there is an inclusion map from K_i to K_j. Hence we have an induced homomorphism denoted by

$$f_p^{i,j} \colon H_p(K_i) \to H_p(K_j),$$

and a corresponding sequence of homology groups connected by these homomorphisms,

$$\emptyset = H_p(K_0) \longrightarrow H_p(K_1) \longrightarrow \cdots \longrightarrow H_p(K_i) \overset{f_p^{i,i+1}}{\longrightarrow} H_p(K_{i+1}) \longrightarrow \cdots \longrightarrow H_p(K_n) = H_p(K)$$

one for each dimension p.

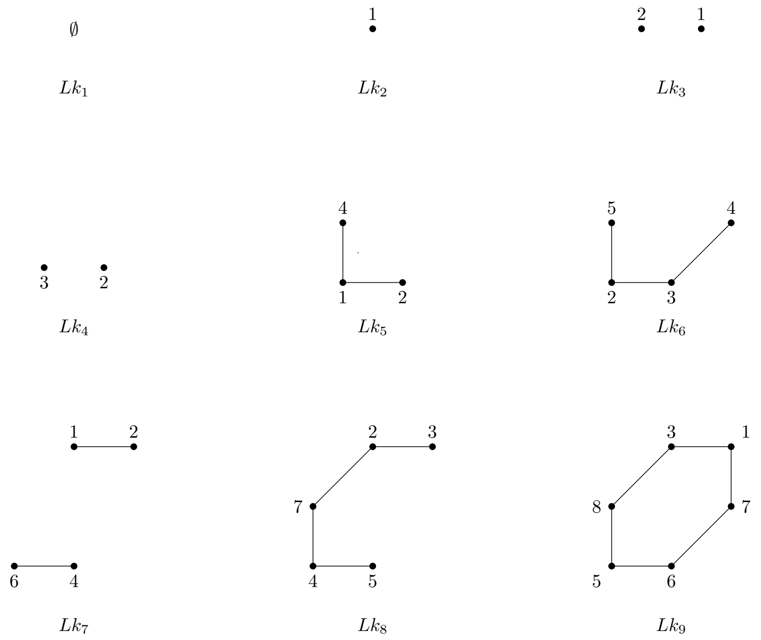


Figure 2.6: Lower Links of Vertices in Figure 2.4

Definition 15. *[9] The images of the homomorphisms induced by these inclusion maps are called* ***the p-th persistent homology groups*** *defined as*

$$H_p^{i,j} := \operatorname{Im}\{f_p^{i,j} \colon H_p(K_i) \to H_p(K_j)\} \textit{ for } 0 \le i \le j \le n.$$

Furthermore, the corresponding p***-th persistent Betti numbers*** *are defined as the ranks of these groups,*

$$\beta_p^{i,j} := \textit{rank } H_p^{i,j}.$$

By definition of the p-th persistent homology groups, any class in $H_p^{i,j}$ survive from K_i to K_j, as it were, these classes do not get quotient out by the boundaries in K_j. Therefore we can also represent $H_p^{i,j}$ as

$$H_p^{i,j} = Z_p(K_i)/(B_p(K_j) \cap Z_p(K_i)).$$

Note that, these groups exist for each dimension p and each pair of indeces $i \le j$.

Definition 16. *[9] We say that a class $\gamma \in H_p(K_i)$ is* ***born*** *at K_i if $\gamma \notin H_p^{i-1,i}$. Besides, if γ is born at K_i then it* ***dies*** *entering K_j if $f_p^{i,j-1}(\gamma) \notin H_p^{i-1,j-1}$ but $f_p^{i,j}(\gamma) \in H_p^{i-1,j}$. In other words, if γ merges with an older class, which is born before γ, at K_j then it does not live in K_j any more.*

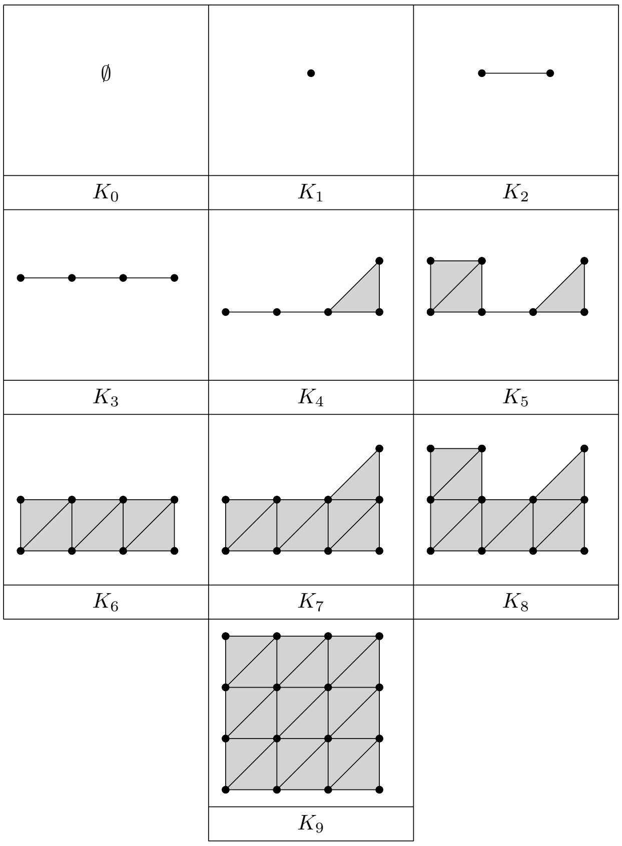


Figure 2.7: The lower star filtration of f

An explicit explanation of this definition is illustrated in Figure 2.8.

Birth and death of a class is a crucial information which is contained in the persistent homology groups. However, when a class is born, many other classes also born which are the sum of this new class and the sum of this class with the existing ones. So, it is not easy to understand what does birth and death mean. Therefore, we need a mechanism to pair births and deaths canonically.

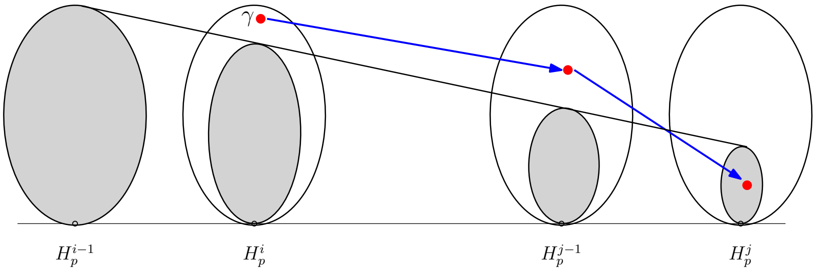


Figure 2.8: The class γ is born at K_i since it does not lie in the (shaded) image of H_p^{i-1}. Furthermore, γ dies entering K_j since this is the first time its image merges into the image of H_p^{j-1}

Elder Rule: To clarify the ambiguity when two classes merged we choose to kill the youngest one. This canonical way of killing the youngest one is called the elder rule.

Definition 17. *[9]* ***The persistence of*** *γ is the difference in function value denoted by pers$(\gamma) := a_j - a_i$, if γ is born at K_i and dies entering K_j. Also, we can define* ***the index persistence of*** *γ which is the difference in the indeces, $j - i$, instead of the persistence of γ.*

Note that if γ is born at K_i but never dies, then we call γ an **essential class** and its persistence is infinite.

Definition 18. *[9]* ***Multiplicity*** *at (i, j) is the number of p-dimensional classes born at K_i and die entering K_j. It is denoted by*

$$\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j})$$

for all $i \leq j$ and all p.

In fact, the number of classes that are born at or before K_i and die entering K_j is shown by the first difference on the right hand side, similarly, the number of classes that are born at or before K_{i-1} and die entering K_j is shown by the second difference. Drawing each point (a_i, a_j) with multiplicity $\mu_p^{i,j}$ in the extended real plane, $\bar{\mathbb{R}}^2$, we get the p-th persistence diagram of the filtration, denoted as $Dgm_p(f)$. The persistence of a class is determined by the vertical distance of a point to the diagonal (see Figure 2.10).

Indeed, $\beta_p^{k,l}$ is the number of points in the upper, left quadrant with corner point (a_k, a_l). A class that is born at K_i and dies entering K_j is counted if and only if $a_i \leq a_k$ and $a_j > a_l$. The quadrant is therefore closed along its vertical right side and open along its horizontal lower side.

Remark 1. *Suppose, we want to compute persistent homology of the sublevel sets, $K_a = f^{-1}(-\infty, a]$ using the values of f at each vertex. Its persistence diagrams are the same as for the lower star filtration of simplicial complexes,*

$$K_a = \{\sigma \in K \mid max_{\nu \in \sigma}\ f(v) \leq a\}$$

.

Example 3. *Let K be a simplicial complex and consider the given filtration of K in Figure 2.9:*

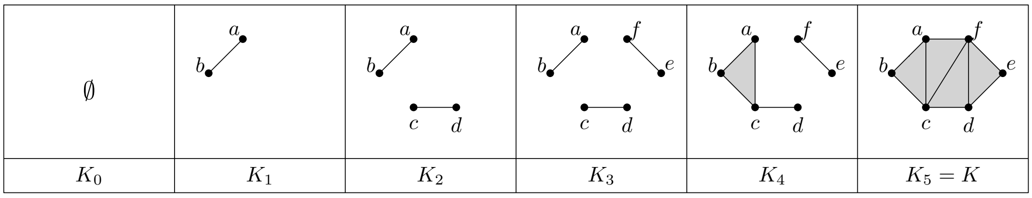


Figure 2.9: A filtration for K

Then we can draw the $zero - th$ persistence diagram according to our filtration in Figure 2.10 :

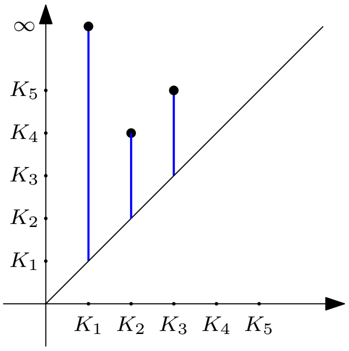


Figure 2.10: The zero-persistence diagram of K

As we know the diagram, we can get the following inferences from the filtration about $H_0(K)$:

(i) $\{a, b\}$ *is born at* K_1 *and it never dies,*

(ii) $\{c, d\}$ *is born at* K_2 *and it dies entering at* K_4*,*

(iii) $\{e, f\}$ *is born at* K_3 *and it dies entering at* K_5*.*

In Figure 2.11, one can see how to compute the persistent Betti numbers from the persistence diagrams. For example, $\beta_0^{2,4} = 1$.

Lemma 2.3.1. *[9][Fundamental Lemma Of Persistent Homology]*

The p-th persistent Betti number $\beta_p^{k,l}$ is defined for every dimension p as follows: Take a filtration $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$ and let (k, l) be a pair of indices such that $0 \leq k \leq l \leq n$. Then

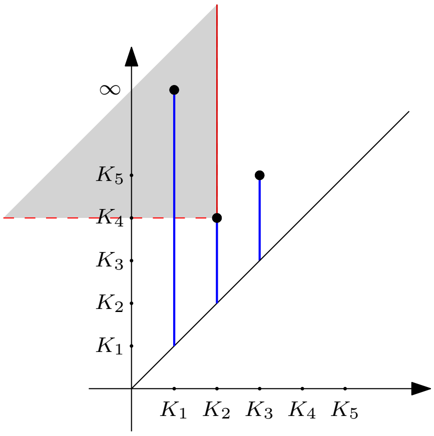


Figure 2.11: Persistent Betti numbers from persistence diagrams.

$$\beta_p^{k,l} = \sum_{i \leq k} \sum_{j > l} \mu_p^{i,j}.$$

This lemma is very important because it shows that the persistence diagram contains all information about persistent homology groups. Indeed, $\beta_p^{k,l}$ is the number of points in the upper left quadrant of the corner (k, l). A class that is born at K_i and dies entering K_j is counted for $\beta_p^{k,l}$ if and only if $i \leq k$ and $j > l$. The quadrant is therefore closed on the right and open on the bottom.

In general, there are two types of data analyzed by using persistent homology. The first type of data, that we are already working with, is topological spaces equipped with real valued function. The second one is finite metric spaces which are called point cloud. In order to analyze the second type of data by using persistent homology, we have to convert it into a global topological object: the shape of a point cloud can be studied by thickening the points at different scales of resolution and analyzing the evolution of the shape across the different resolution values. To do so, we construct a simplicial complex on it. There are many ways to do this transformation, but in this thesis we will just mention two of them, namely $\check{C}$ech complex and the Vietoris-Rips complex.

Definition 19. *[15]* ***Data*** *is represented as an unordered sequence of points in an Euclidean* n*-dimensional space* $\mathbb{R}^n$*. One type of data set for which global features are present and significant is the so-called point cloud data coming from physical objects in 3D.*

Definition 20. *[15] Given a collection of points $\{x_a\}$ in Euclidean space $\mathbb{R}^n$,* ***the $\check{C}$ech complex****, C_ϵ, is the abstract simplicial complex whose k-simplices are determined by unordered $(k+1)$-tuples of points $\{x_a\}_0^k$ whose closed $\epsilon/2$-ball neighborhoods have a point of common intersection.*

Definition 21. *[15] Given a collection of points $\{x_a\}$ in $\mathbb{R}^n$, the* ***Vietoris-Rips complex****, R_ϵ, is the abstract simplicial complex whose k-simplices correspond to unordered $(k+1)$-tuples of points $\{x_a\}_0^k$ which are pairwise within distance ϵ.*

If we have X, a subset of $\mathbb{R}^n$, instead of an abstract simplicial complex, we can construct the Vietoris-Rips complex as follows: Let d denote the chosen metric and for any $\epsilon > 0$

1. add a 0-simplex for each point in X,
2. for $x_1, x_2 \in X$, add a 1-simplex between x_1, x_2 if $d(x_1, x_2) \leq \epsilon$,
3. for $x_1, x_2, x_3 \in X$, add a 2-simplex with vertices x_1, x_2, x_3 if $d(x_1, x_2), d(x_1, x_3), d(x_2, x_3) \leq \epsilon$,
   $\vdots$

n. for $x_1, x_2, \ldots, x_n \in X$, add an n-simplex with vertices $x_1, x_2, \ldots, x_n$ if $d(x_i, x_j) \leq \epsilon$ for $1 \leq i, j \leq n$; that is, if all the points are within a distance of ϵ from each other.

We can easily see the difference between $\check{C}$ech and Vietoris Rips complex in Figure 2.12. In this figure, the Vietoris-Rips complex is a graph with a 2-simplex, while the $\check{C}$ech complex is just a graph.

Note that, since X is sitting in some metric space, and the union of these epsilon-balls forms some topological space $X(\epsilon)$ that is close in structure to X. The Nerve Theorem says that the homotopy types of $X(\epsilon)$ and C_ϵ are the same. Recall that when two topological spaces have the same homotopy type, then homology can not distinguish them. In other words, if homotopy type is too coarse for a discriminator for our data set, then persistent homology will fail us for sure. Therefore, by this theorem, if we want to learn about our point cloud, we can pick an ϵ and study the topology of

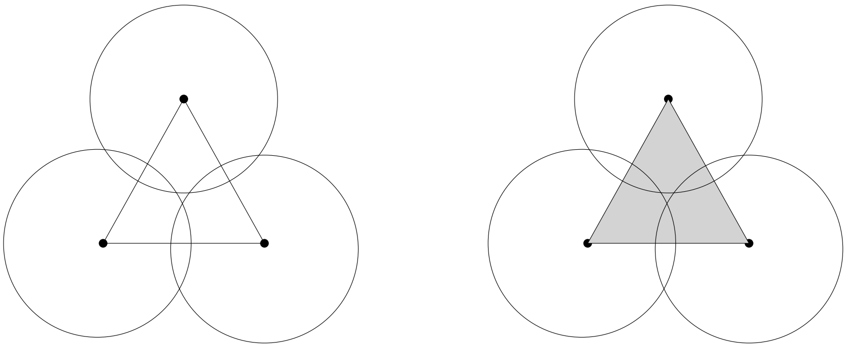

(a) $\check{C}$ech complex (b) Vietoris-Rips complex

Figure 2.12: The difference between $\check{C}$ech and Vietoris Rips complexes

the corresponding $\check{C}$ech complex C_ϵ. The reason this is called the "Nerve Theorem" is because one can generalize it to an arbitrary family of convex sets. Given some family F of convex sets, the nerve is the complex obtained by adding simplices for mutually overlapping subfamilies in the same way. The nerve theorem is actually more general, it says that with sufficient conditions on the family F being "nice," the resulting $\check{C}$ech complex has the same topological structure as F. Despite of this correspondence, it is really hard to compute $\check{C}$ech complexes. The reason is that, to determine whether there are any 5-simplices you have to check all the subsets of size 5. If the dimension of the simplices increase then the required time to find it will increase exponentially, and so computing the entire space is exponentially proportional to the size of the space. To overcome this weakness of $\check{C}$ech complexes, we use Vietoris-Rips complexes which is less time consuming.

Whether or not we are using Vietoris-Rips or $\check{C}$ech complexes to construct global object from the point cloud data sets we need to make a choice of the parameter ϵ. If we choose ϵ too small then our complex will be just a discrete set of points, if we choose ϵ too large then we will have a single high dimensional simplex. It is the reason why we do not fix the value of ϵ. To get the complex which best captures the topology of data set we start with a small ϵ and increase it slightly.

Definition 22. *A **barcode** is a representation which contains the information of persistent homology groups of a topological space X such as a persistence diagram. Barcode consists of horizontal line segments in a plane whose horizontal axis corresponds to the parameter and whose vertical axis represents an (arbitrary) ordering of homology generators.*

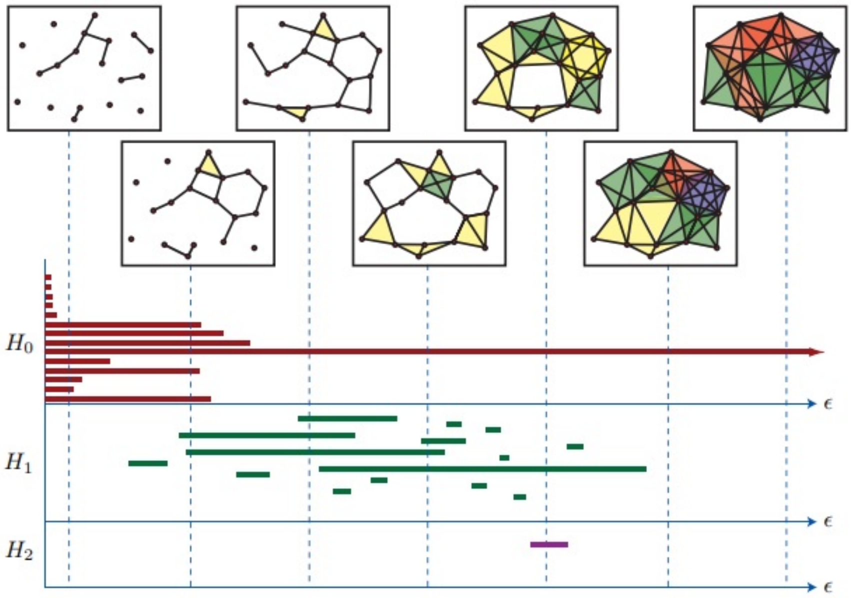


Figure 2.13: An example of a barcode of a data set [15]

2.4 Extended Persistence

Definition 23. *Let $f : \mathbb{M} \to \mathbb{R}$ be a smooth function on a d-dimensional, connected and orientable manifold $\mathbb{M}$. A point $x \in \mathbb{M}$ is critical if the total derivative of f at x is identically 0, and it is non-degenerate if the Hessian at the point is invertible. We call f as a* ***smooth Morse function*** *if the followings are satisfied:*

(i) all critical points are non-degenerate,

(ii) the critical points have different function values.

Definition 24. *[1] Let $\mathbb{M} \subseteq \mathbb{R}^3$ be a two dimensional smooth manifold which is assumed to be generic, i.e., any small perturbation does not change the embedding of $\mathbb{M}$ in $\mathbb{R}^3$. Now we define a height function as the signed distance from the plane normal to a given direction and passing through the origin. More formally, consider $\mathbb{S}^2$ as the collection of all unit vectors in $\mathbb{R}^3$, and define a function $f_u(x) =< x, u >$ which gives the height of x in the direction u. Note that, the collection of all such f*

is a 2-parameter family of height functions,

$$\textit{Height} : \mathbb{M} \times \mathbb{S}^2 \to \mathbb{R}$$

where Height$(x, u) = f_u(x)$.

Moreover, if u is a generic direction then $f_u(x)$ gives a Morse function on $\mathbb{M}$. We define the persistence of a critical point as the absolute difference in height to the paired point:

$$\textit{pers}(x) = \textit{pers}(y) = f_u(y) - f_u(x).$$

Each point $x \in \mathbb{M}$ is critical for exactly two height functions, namely for the ones in the direction of its outer and inner normals: $u = \pm n_x$. The pairs we get for the two opposite directions are the same. Hence, each point $x \in \mathbb{M}$ has unique persistence, which we use to introduce the elevation function,

$$\textit{Elevation} : \mathbb{M} \to \mathbb{R}$$

defined by Elevation$(X) =$ *pers*$(x) = f_u(y) - f_u(x)$. *We note that the elevation function is invariant under translation and rotation of $\mathbb{M}$ in $\mathbb{R}^3$.*

To define elevation function each critical points has to be paired. If we just use the sublevel sets (if we sweep the manifold from bottom to top), we cannot pair some of the critical points which are the first minimum, the last maximum and the $2g$-saddles starting with the $2g$ cycles. We are making an extension to pair this $2g + 2$ unpaired critical points. To do so, we define the superlevel sets (we sweep the manifold from top to bottom). After doing it, we pair the unpaired critical points as well.

In later work, elevation function generalized to smoothly embedded d-manifolds in $\mathbb{R}^{d+1}$.

2.4.1 Extended Filtration

Let $f_u : \mathbb{M} \to \mathbb{R}$ be a height function and $a_1 < a_2 < \cdots < a_n$ be the homological critical values of f_u. When we add the interleaved values such that $b_0 < a_1 < b_1 < a_2 < \cdots < a_n < b_n$, we get sublevel sets $M_{b_i} = f^{-1}(-\infty, b_i]$ and superlevel sets $M^{b_i} = f^{-1}[b_i, \infty)$.

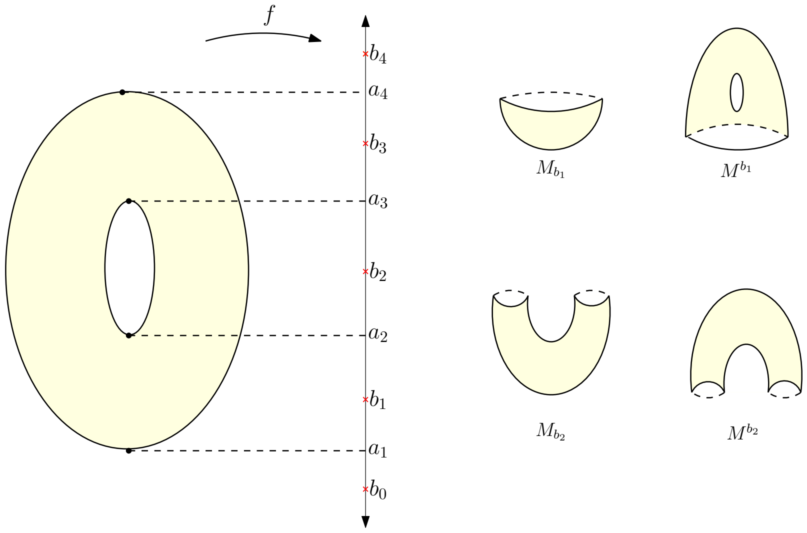


Figure 2.14: An example of sublevel and superlevel sets of torus

Now, let us briefly review Poincaré and Lefschetz duality. For further details see [16].

2.4.2 Poincaré Duality

If $\mathbb{M}$ is a d-dimensional oriented closed manifold, then

$$H_p(\mathbb{M}) \cong H^{d-p}(\mathbb{M}).$$

If a manifold is orientable, Poincaré duality holds for any coefficient group. However, if you are working with $\mathbb{Z}_2$ coefficient then the orientation has no importance. Because of the fact that any manifold is mod 2 orientable.

2.4.3 Lefschetz Duality

If $\mathbb{M}$ is an d-dimensional compact manifold with boundary $\partial\mathbb{M}$, then

$$H^{d-p}(\mathbb{M}) \cong H_p(\mathbb{M}, \partial\mathbb{M})$$

2.4.4 Extended Persistence

Let $\mathbb{M}$ be a d-dimensional closed manifold and $f : \mathbb{M} \to \mathbb{R}$ be a Morse function on it. We can imagine that $\mathbb{M}$ is smoothly embedded in $\mathbb{R}^{2d+1}$ and f maps every point to its height above some hyperplane.

We have the sequence of homology groups of connected by homomorphism,

$$\emptyset = H_p(M_{b_0}) \to H_p(M_{b_1}) \to \cdots \to H_p(M_{b_n}).$$

Applying Poincaré duality to $H_p(M_{b_n})$,

$$H_p(M_{b_n}) \cong H^{d-p}(M_{b_n})$$

and by inclusion we have the natural induced sequence

$$H_p(M_{b_n}) \cong H^{d-p}(M_{b_n}) \to H^{d-p}(M_{b_{n-1}}) \to \cdots \to H^{d-p}(M_{b_0}).$$

By applying Lefschetz duality to each cohomology group in this sequence, we get

$$H_p(M_{b_n}) \cong H_p(M_{b_n}, \partial M_{b_n}) \to H_p(M_{b_{n-1}}, \partial M_{b_{n-1}}) \to \cdots \to H_p(M_{b_0}, \partial M_{b_0}).$$

Obviously we can replace the boundaries of the sublevel sets by those of the corresponding superlevel sets, as

$$\partial M_{b_i} = \partial M^{b_i} = f^{-1}(\{b_i\}),$$

and thus receive the sequence

$$H_p(M_{b_n}, \partial M^{b_n}) \to H_p(M_{b_{n-1}}, \partial M^{b_{n-1}}) \to \cdots \to H_p(M_{b_0}, \partial M^{b_0}).$$

Finally, we apply the excision theorem to each homology group to put back in the open set interior of M^{b_i} into each pair $H_p(M_{b_i}, \partial M^{b_i})$, and end up with

$$H_p(M, M^{b_n}) \to H_p(M, M^{b_{n-1}}) \to \cdots \to H_p(M, M^{b_0}).$$

Because of $H_p(M, M^{b_n}) = H_p(M, \emptyset) = H_p(M)$, we can append this sequence to our original persistence sequence:

$$\emptyset = H_p(M_{b_0}) \to H_p(M_{b_1}) \to \cdots \to H_p(M_{b_n}) \cong H_p(M, M^{b_n}) \to \cdots \to H_p(M, M^{b_0}) = \emptyset$$

which is the extended persistence sequence we searched for.

Since we go from the trivial group to the trivial group, everything that gets born eventually dies. As a consequence, all births will be paired with corresponding deaths, as desired.

Definition 25. *[18] If we have an extended persistence sequence, then we define homomorphisms* $f_p^{i,j} : A_i \to A_j$ *for every pair* $0 \leq i \leq j \leq 2n$*, where*

$$A_i = \begin{cases} H_p(M_{b_i}) & if \quad i \leq n, \\ H_p(M, M^{b_{2n-i}}) & if \quad i \geq 2n. \end{cases}$$

Hence, p***-th extended persistent homology groups*** *can be defined as* $H_p^{i,j} = \operatorname{Im} f_p^{i,j}$.

Definition 26. *We call a homology class in the superlevel set* ***essential*** *if it lives all the way down to* b_0*, and inessential, otherwise.*

Now, let us explain the relation between the absolute homology of $\mathbb{M}^{b_i}$ and the relative homology of the pair $(\mathbb{M}, \mathbb{M}^{b_i})$ in terms of birth and death of a homology class through the extended persistence sequence.

1. A dimension p homology class of $\mathbb{M}^{b_i}$ dies at the same time a dimension $p+1$ relative homology class of $(\mathbb{M}, \mathbb{M}^{b_i})$ dies.
2. An inessential dimension p homology class of $\mathbb{M}^{b_i}$ gets born at the same time a dimension $p+1$ relative homology class of $(\mathbb{M}, \mathbb{M}^{b_i})$ gets born.
3. An inessential dimension p homology class of $\mathbb{M}^{b_i}$ gets born at the same time a dimension p relative homology class of $(\mathbb{M}, \mathbb{M}^{b_i})$ dies.

Example 4. *Consider the height function f on the genus-2 surface depicted in Figure 2.15*

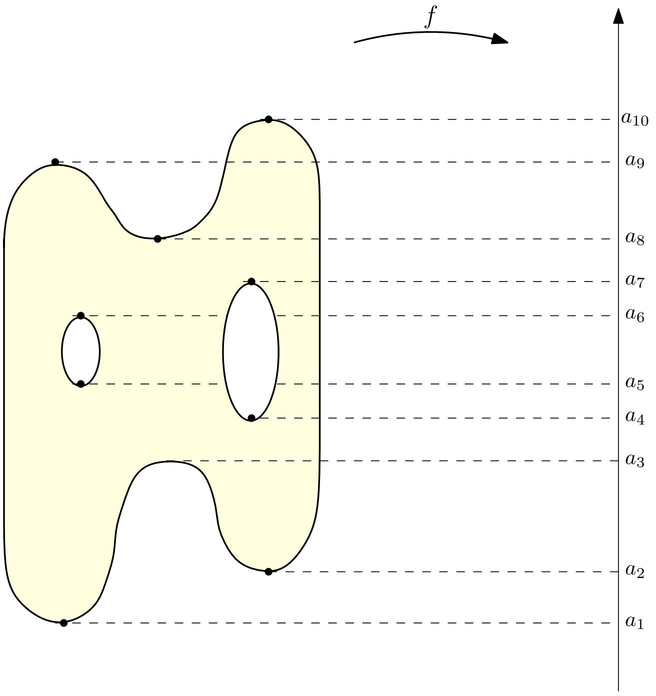


Figure 2.15: The genus-2 surface with the height function f.

First of all, we determine the birth and death of homology classes by using sublevel sets.

a_1, a_2*: give birth to classes in* H_0*,*

a_3*: kills the class in* H_0 *which is born at* a_2*,*

a_4,a_5,a_6,a_7,a_8*: give birth to classes in* H_1*,*

a_9*: kills the class in* H_1 *which is born at* a_8*,*

a_{10}*: gives birth to a class in* H_2*.*

Therefore, the classes, which are born at a_1,a_4,a_5,a_6,a_7 *and* a_{10} *are essential homology classes of genus-*2 *surface. Also* (a_2, a_3) *and* (a_8, a_9) *form pairs.*

Next, we determine the birth and death of homology classes with using superlevel sets and we will pair all homology classes of genus-2 surface

a_{10}*: kills the class in* H_0*,*

a_9: gives birth to a class in H_1,

a_8: kills the class in H_1 which is born at a_9,

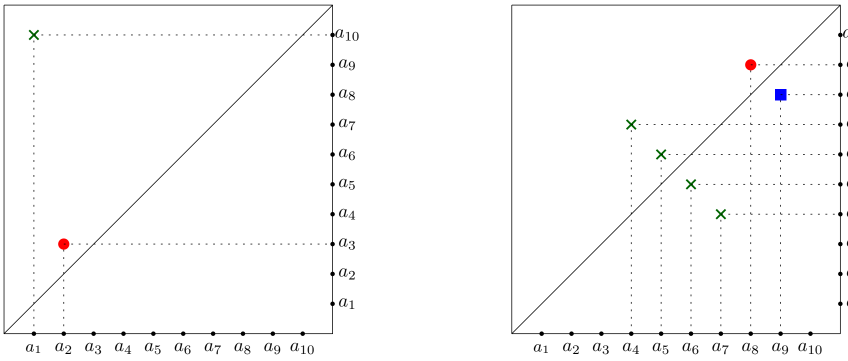
a_4,a_5,a_6,a_7: kill the classes in H_1,

a_3: gives birth to a class in H_2,

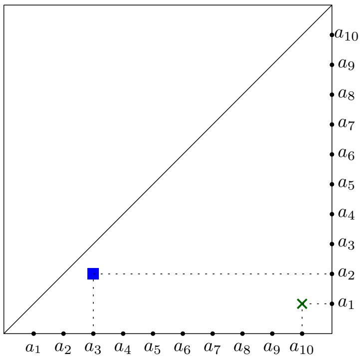
a_2: kills the class in H_2 which is born at a_3,

a_1: kills the class in H_2.

According to superlevel sets, (a_9, a_8) and (a_3, a_2) form pairs. To summarize, the pairs of critical values defining the points in the diagrams are (a_1, a_{10}), (a_2, a_3) in dimension 0, (a_4, a_7), (a_5, a_6), (a_6, a_5), (a_7, a_4), (a_8, a_9), (a_9, a_8) in dimension 1, and (a_{10}, a_1), (a_3, a_2) in dimension 2.



$0^{th}PersistenceDiagram$ $1^{st}PersistenceDiagram$



$2^{nd}PersistenceDiagram$

Figure 2.16: The 0^{th}, 1^{st}, 2^{nd} persistence diagrams of the height function in Figure 2.15

In Figure 2.16, these three diagrams contain the information of classes which are born and die going up, born going up and die coming down, and born and die coming down which are represented by different symbols. Also, there are three subdiagrams according to the location of the points in these diagrams.

1. *The collection of the points which lie above diagonal is called ordinary subdiagram and denoted as Ord.*
2. *The collection of the points which lie below diagonal is called relative subdiagram and denoted as Rel.*
3. *The collection of the points which lie on either side of diagonal is called extended subdiagram and denoted as Ext.*

2.4.5 Duality and Symmetry

Definition 27. *We call* $f\colon \mathbb{M} \to \mathbb{R}$ ***tame*** *if the homology groups of every sublevel have finite ranks and there are only finitely many values* b_i *across which the homology groups are not isomorphic.*

Theorem 2.4.1 (Persistence Duality Theorem)**.** *[9] A tame function* f *on a* d*-manifold without boundary has persistence diagrams that are reflections of each other as follows,*

$$Ord_p(f) = Rel^T_{d-p}(f)$$

$$Ext_p(f) = Ext^T_{d-p}(f)$$

$$Rel_p(f) = Ord^T_{d-p}(f).$$

Equivalently, the full p*-th persistence diagram is the reflection of the full* $(d-p)$*-th persistence diagram,* $Dgm_p(f) = Dgm^T_{d-p}(f)$*.*

We use the symbol "T" to indicate reflection across the main diagonal, mapping the point (a, b) to (b, a).

Theorem 2.4.2 (Persistence Symmetry Theorem)**.** *[9] Let f be a tame function on a d-manifold without boundary and $-f$ be its negative. Then the persistence diagrams of the two functions are reflections of each other,*

$$Ord_p(f) = Ord^R_{d-p-1}(-f)$$

$$Ext_p(f) = Ext^O_{d-p}(-f)$$

$$Rel_p(f) = Ord^R_{d-p+1}(-f).$$

We use the symbol "R" to indicate reflection across the minor diagonal, mapping the point (a, b) to $(-b, -a)$. We use the symbol "O" to indicate central reflection through the origin, mapping the point (a, b) to $(-a, -b)$.

2.5 Persistence Module

In algebraic topology, we use abelian groups as a coefficient to define homology. Many of the abelian groups are not just groups. More precisely, if it is an abelian group with respect to addition and multiplication is defined and distributive law holds we name it is a ring. In addition to be a ring, if every element has a multiplicative inverse then we call it as a field. Since the persistent homology is defined by using the homomorphisms between the homology groups, we can define persistent homology with coefficients any of the above.

Zomorodian and Carlsson [27] constructed a module structure over the polynomial ring $R = \mathbb{F}[x]$ by using the homology groups. To do so, let us remind some basic definitions of algebra: $\mathbb{F}[x]$ is a polynomial ring which consists of polynomials of x having the coefficients in $\mathbb{F}$. A subset I of R is called ideal if it is a subgroup with respect to addition and for any $x \in I, r \in R$ we have $rx, xr \in I$. An ideal is principle if it is generated by a single element. If every ideal of R is a principle then we call it principle ideal domain. With respect to the basic results in commutative algebra, $\mathbb{F}[x]$ is a principle ideal domain. The classification of finitely generated modules over principle ideal domain is very simple. Namely, we have a structure theorem which says that if M is such a finitely generated module then M can be represented as the direct sum of a finitely generated free module and a torsion module.

Definition 28. *A* ***persistence complex*** $\mathscr{C}$ *is a family of chain complexes* $\{C^i_*\}_{i\geq 0}$ *over* R*, together with chain map's* $f^i : C^i_* \to C^{i+1}_*$*, so that we have the following diagram:*

$$C^0_* \xrightarrow{f^0} C^1_* \xrightarrow{f^1} C^2_* \xrightarrow{f^2} \cdots .$$

We show a portion of the persistence complex below:

$$\begin{tikzcd}
 & \vdots \arrow[d, "\partial^i_{k+1}"] & \vdots \arrow[d, "\partial^{i+1}_{k+1}"] & \vdots \arrow[d, "\partial^{i+2}_{k+1}"] & \\
\cdots \arrow[r, "f^{i-1}"] & C^i_k \arrow[r, "f^i"] \arrow[d, "\partial^i_k"] & C^{i+1}_k \arrow[r, "f^{i+1}"] \arrow[d, "\partial^{i+1}_k"] & C^{i+2}_k \arrow[r, "f^{i+2}"] \arrow[d, "\partial^{i+2}_k"] & \cdots \\
\cdots \arrow[r, "f^{i-1}"] & C^i_{k-1} \arrow[r, "f^i"] \arrow[d, "\partial^i_{k-1}"] & C^{i+1}_{k-1} \arrow[r, "f^{i+1}"] \arrow[d, "\partial^{i+1}_{k-1}"] & C^{i+2}_{k-1} \arrow[r, "f^{i+2}"] \arrow[d, "\partial^{i+2}_{k-1}"] & \cdots \\
\cdots \arrow[r, "f^{i-1}"] & C^i_{k-2} \arrow[r, "f^i"] \arrow[d, "\partial^i_{k-2}"] & C^{i+1}_{k-2} \arrow[r, "f^{i+1}"] \arrow[d, "\partial^{i+1}_{k-2}"] & C^{i+2}_{k-2} \arrow[r, "f^{i+2}"] \arrow[d, "\partial^{i+2}_{k-2}"] & \cdots \\
 & \vdots & \vdots & \vdots &
\end{tikzcd}$$

In the above diagram, each column is a chain complex, and the chain maps f^i connect chain complexes of successively larger simplicial complexes in the filtration together.

Definition 29. *[25] The* k***-th persistence module*** $\mathcal{H}_k$ *is the family of* k*-th homology modules* H^i_k *together with module homomorphisms* $\mu^i_k : H^i_k \to H^{i+1}_k$*. A persistence module is said to be of finite type if each component module is finitely generated and there exists some integer* m *such that the maps* μ^i_k *are isomorphisms for all* $i \geq m$*.*

Definition 30. *[25] The* k*-th persistence module can be given the structure of a graded module over the polynomial ring* $R[x]$*:*

$$\mathcal{H}_k = \bigoplus_{i=0}^{\infty} H^i_k$$

where the action of x *is given by* $x.(\sum_{i=0}^{\infty} m^i) = \sum_{i=0}^{\infty} \mu^i_k(m^i)$ *for any* $m^i \in H^i_k$*.*

In other words, the action of x shifts the grading upward by 1. By this action, the persistent module makes a connection between homologies of different complexes in the filtration. Hence, it contains all of the information needed for the p-persistent k-th homology modules $H_k^{i,p}$.

Theorem 2.5.1. *[25] Suppose $\mathcal{H}_k$ is over the polynomial ring $\mathbb{F}[x]$, where $\mathbb{F}$ is a field. Then*

$$\mathcal{H}_k = (\bigoplus_i (x^{a_i})) \oplus (\bigoplus_j (x^{b_j})\ (x^{c_j}))$$

where the sums range over $1 \leq i \leq M$ and $1 \leq j \leq N$ for non-negative integers M, N and a_i, b_j, c_j are non-negative integer powers of x.

Indeed, the numbers a_i and b_j represent index in the filtration where k-dimensional homology generators are born at K_{a_i} and K_{b_j} and the number c_j represents the death of a k-dimensional homology generator which merges with an older k-dimensional homology generator at subcomplex K_{b_j}. While the free part of the above sum carries the information about essential classes which born at K_{a_i}, the torsion part gives the classes which born and die through the filtration.

CHAPTER 3

MAYER VIETORIS SEQUENCE FOR PERSISTENT HOMOLOGY

Let X be a topological space, and A and B are subspace of X such that X is the union of the interiors of A and B and $A \cap B$ is a subspace of X. Then we can compute homology groups by using the Mayer Vietoris sequence given as below:

$$\cdots \to H_{k+1}(X) \overset{\delta_k}{\to} H_k(A \cap B) \overset{\alpha_k}{\to} H_k(A) \oplus H_k(B) \overset{\beta_k}{\to} H_k(X) \to \cdots \to H_0(X) \to 0$$

where, $\delta_k([x]) = [\partial(x|_A)],\ \alpha_k([y]) = ([y],[-y])$ and $\beta_k(([z],[z'])) = [z+z']$.

The key point here is that this is a long exact sequence, i.e, $\mathrm{Im} = \ker$ for successive maps. Long exact sequences have many nice properties that make it easy to deduce, up to isomorphism, each group in the sequence. In practice, we try to choose A and B so that their homology groups can be easily computed by inspection or are already known. Then we use the exactness of the Mayer-Vietoris sequence to compute the homology of X.

Mayer Vietoris sequence is very important tool for computing homology groups of a topological space in algebraic topology. The question that we want to address in this chapter is that whether it is applicable in persistent homology as well. Persistent homology has a Mayer-Vietoris sequence that in general is not exact but only of order two i.e., rather than being equal, we only have $\mathrm{Im} \subseteq \ker$.

3.0.1 A Mayer-Vietoris Formula for Ordinary Persistent

Let A and B to be triangulable subsets of X such that $X = A \cup B$ and $Y = A \cap B$ is triangulable.

Lemma 3.0.1. *Let $u \in \mathbb{R}$. Let us endow X_u with the relative topology induced by the topology of X. Then A_u and B_u are closed sets in X_u. Moreover, $X_u = A_u \cup B_u$ and $Y_u = A_u \cap B_u = (A \cap B)_u$*

Proof. Let $A \subset X$ be a closed subset. Therefore A_u is closed in X_u since $A_u = A \cap X_u$, with A triangulable, hence closed in X (definition of subspace topology). Similarly, B_u is closed in X. $A_u = A \cap X_u$ and $B_u = B \cap X_u$ implies that

(i) $A_u \cup B_u = (A \cap X_u) \cup (B \cap X_u) = (A \cup B) \cap X_u = X \cap X_u = X_u$

(ii) $A_u \cap B_u = (A \cap X_u) \cap (B \cap X_u) = (A \cap B) \cap X_u = Y \cap X_u = Y_u$

$\square$

Let us consider the diagram below:

$$\begin{tikzcd}
 & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
\cdots \arrow[r] & H_{k+1}(X_u) \arrow[r, "\delta_k^u"] \arrow[d, "h_{k+1}"] & H_k((A \cap B)_u) \arrow[r, "\alpha_k^u"] \arrow[d, "f_k"] & H_k(A_u) \oplus H_k(B_u) \arrow[r, "\beta_k^u"] \arrow[d, "g_k"] & H_k(X_u) \arrow[r] \arrow[d, "h_k"] & \cdots \\
\cdots \arrow[r] & H_{k+1}(X_v) \arrow[r, "\delta_k^v"] \arrow[d] & H_k((A \cap B)_v) \arrow[r, "\alpha_k^v"] \arrow[d] & H_k(A_v) \oplus H_k(B_v) \arrow[r, "\beta_k^v"] \arrow[d] & H_k(X_v) \arrow[r] \arrow[d] & \cdots \\
 & \vdots & \vdots & \vdots & \vdots &
\end{tikzcd} \tag{3.1}$$

where the top horizontal line belongs to the Mayer-Vietoris sequence of the triad (X_u, A_u, B_u) and the second horizontal line belongs to the Mayer-Vietoris sequence of the triad (X_v, A_v, B_v). Also, the vertical lines are inclusion maps.

Lemma 3.0.2. *Each vertical line in above diagram is exact. Moreover, each square in the same diagram is commutative.*

Remark 2. *For every $k \in \mathbb{Z}$, and every proper point $(u, v) \in \Delta^*$,*

(i) $\operatorname{Im} f_k = H_k^{u,v}(A \cap B)$,

(ii) $\operatorname{Im} g_k = H_k^{u,v}(A) \oplus H_k^{u,v}(B)$,

(iii) $\operatorname{Im} h_k = H_k^{u,v}(X)$.

The following proposition proves that the commutativity of the squares in the above diagram induces a Mayer-Vietoris sequence of order 2 involving the k-th persistent homology groups of X, A, B, and $A \cap B$, for every integer k.

Proposition 1. *[7] Let us consider the sequence of homomorphisms of persistent homology groups*

$$\cdots \longrightarrow H_{k+1}^{u,v}(X) \xrightarrow{\delta} H_k^{u,v}(A \cap B) \xrightarrow{\alpha} H_k^{u,v}(A) \oplus H_k^{u,v}(B) \xrightarrow{\beta} H_k^{u,v}(X) \longrightarrow \cdots$$

where $\delta = \delta_k^v|_{\operatorname{Im} h_{k+1}}$, $\alpha = \alpha_k^v|_{\operatorname{Im} f_k}$, $\beta = \beta_k^v|_{\operatorname{Im} g_k}$. *Then the following statements hold:*

(i) $\operatorname{Im} \delta \subseteq \ker \alpha$,

(ii) $\operatorname{Im} \alpha \subseteq \ker \beta$,

(iii) $\operatorname{Im} \beta \subseteq \ker \delta$.

Proof. We will prove only the first case. The proofs of other cases are similar to the first one.

Let $a \in \operatorname{Im} \delta \subseteq H_k^{u,v}(A \cap B) = \operatorname{Im} f_k \subseteq H_k((A \cap B)_v)$. This implies that $a \in H_k((A \cap B)_v)$. Since $\delta = \delta_k^v|_{\operatorname{Im} h_{k+1}}$, $a \in \operatorname{Im} \delta_k^v$. So there exists an $a' \in H_{k+1}(X_v)$ such that $\delta_k^v(a') = a$. Since the third row of the above diagram is exact, $\alpha_k^v(a) = 0$ in other words $a \in \ker \alpha_k^v$. Therefore, $\alpha = \alpha_k^v$ implies $a \in \ker \alpha$ □

We have shown that $\operatorname{Im} \delta \subseteq \ker \alpha$. But $\ker \alpha \subseteq \operatorname{Im} \delta$ is not always true. Let $c \in \ker \alpha$. Since $\ker \alpha \subseteq H_k^{u,v}(A \cap B)$, $c \in H_k((A \cap B)_v)$. Moreover, $\alpha = \alpha_k^v|_{\operatorname{Im} f_k}$ implies $c \in \ker \alpha_k^v$. So, there exist a $d \in H_{k+1}(X_v)$ such that $\delta_k^v(d) = c$ since the third row of the above diagram is exact. On the other hand, $c \in H_k^{u,v}(A \cap B)$ implies $c \in H_k((A \cap B)_u)$. If $\alpha_k^u(c) \neq 0$, then we cannot be sure that there exist a $d \in H_{k+1}(X_u)$ such that $d \in H_{k+1}^{u,v}(X)$. Therefore, c may not be in the $\operatorname{Im} \delta$. Let us give further explanation in the following example.

Example 5. *For a given complex X, a filtration is defined as a nested sequence of its subcomplexes often generated by a filtering function, which in our example will be a perfect discrete Morse function.*

Consider the triangulation of the torus with a perfect discrete Morse function on it. The subspaces A and B are shown in Figure 3.1.

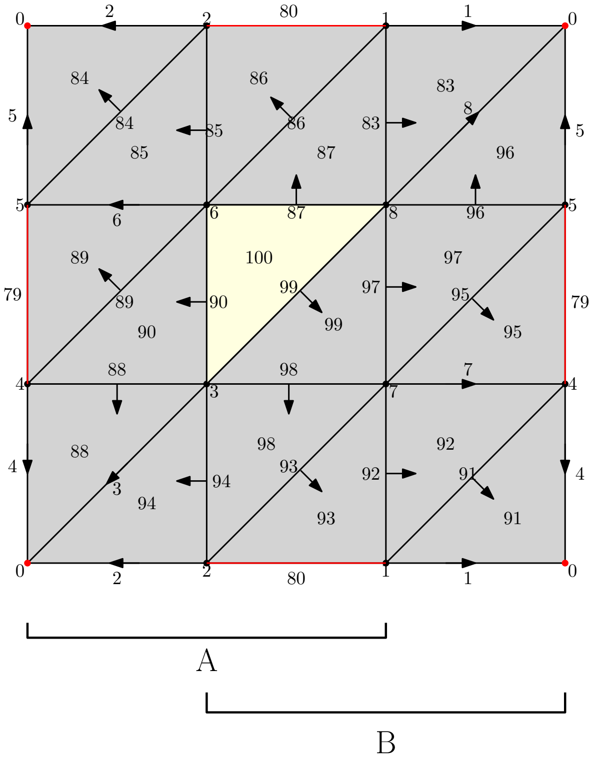


Figure 3.1: Perfect discrete Morse function on the triangulation of Torus

The perfect discrete Morse function defined as above gives a filtration for X, A, B and $A \cap B$:

$$\emptyset \subseteq X_0 \subseteq X_6 \subseteq X_8 \subseteq X_{79} \subseteq X_{95} \subseteq X_{100} = X,$$

$$\emptyset \subseteq A_0 \subseteq A_6 \subseteq A_8 \subseteq A_{79} \subseteq A_{95} \subseteq A_{100} = A,$$

$$\emptyset \subseteq B_0 \subseteq B_6 \subseteq B_8 \subseteq B_{79} \subseteq B_{95} \subseteq B_{100} = B,$$

$$\emptyset \subseteq (A \cap B)_0 \subseteq (A \cap B)_6 \subseteq \ldots \subseteq (A \cap B)_{95} \subseteq (A \cap B)_{100} = A \cap B.$$

Note that the indices in the filtration comes from the values of the given perfect discrete Morse function.

Let us consider the chain complex for the level B_{95} in Figure 3.2:

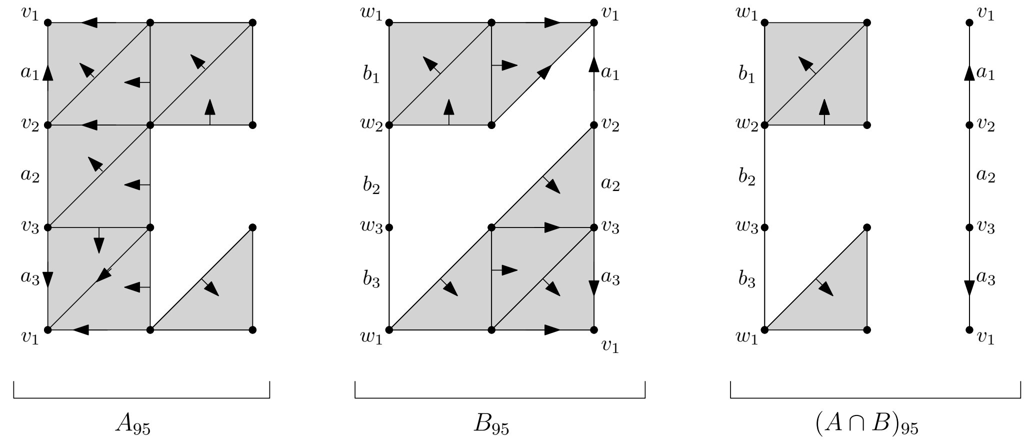


Figure 3.2: The filtration levels of A, B and $A \cap B$ which contains simplicies having the function value less than or equal to 95

$$\begin{tikzcd}
0 \arrow[r] & C_2 \arrow[r, "\partial_2"] & C_1 \arrow[r, "\partial_1"] & C_0 \arrow[r] & 0
\end{tikzcd}$$

From this sequence we can induce the followings: $\partial_1(b_1) = w_1 - w_2 \neq 0$, $\partial_1(b_2) = w_2 - w_3 \neq 0$, $\partial_1(b_3) = w_3 - w_1 \neq 0$ *but* $\partial_1(b_1 + b_2 + b_3) = 0$. *Therefore* $b_1 + b_2 + b_3$ *is a generator of* $H_1(B_{95})$*, call it* γ*. Similarly,* a_2 *is also a generator of* $H_1(B_{95})$*, call it* α*.*

After applying the same process we get:

(i) $H_1(A_{95}) = \langle \alpha \rangle$, $H_1(A_{100}) = \langle \alpha \rangle$,

(ii) $H_1(B_{95}) = \langle \alpha, \gamma \rangle$, $H_1(B_{100}) = \langle \alpha \rangle$,

(iii) $H_1((A \cap B)_{95}) = \langle \alpha, \gamma \rangle$, $H_1((A \cap B)_{100}) = \langle \alpha, \gamma \rangle$.

By the definition of persistent homology group we have,

(i) $H_1^{95,100}(A) = \langle \alpha \rangle$,

(ii) $H_1^{95,100}(B) = \langle \alpha \rangle$,

(iii) $H_1^{95,100}(A \cap B) = \langle \alpha, \gamma \rangle$.

Now consider the Mayer-Vietoris sequence for persistent homology:

$$0 \xrightarrow{\delta_1} H_1^{95,100}(A \cap B) \xrightarrow{\alpha_1} H_1^{95,100}(A) \oplus H_1^{95,100}(B) \xrightarrow{\beta_1} H_1^{95,100}(X) \longrightarrow \cdots$$

Clearly, $\operatorname{Im}\delta_1 = 0$*. Also* $\ker\alpha_1 = \langle\gamma\rangle \neq 0$*. Therefore,* $\operatorname{Im}\delta_1 \neq \ker\alpha_1$ *which implies this sequence is not exact.*

Theorem 3.0.3. *[7] For every* $u, v \in \mathbb{R}$*, with* $u \leq v$*, and for every* $k \in \mathbb{Z}$*,*

$$Ord_k^{u,v}(X) = Ord_k^{u,v}(A) + Ord_k^{u,v}(B) - Ord_k^{u,v}(Y) + rk\delta_k^v - rk\delta_k^{v,u} + rk\delta_{k-1}^u$$

with δ_k^v*,* $\delta_k^{u,v}$ *and* δ_{k-1}^u *as in above diagram.*

3.0.2 A Mayer-Vietoris Formula for Relative Persistent

By a slight modification on the idea used for ordinary persistence, we can define a Mayer-Vietoris sequence formula for relative persistence.

Theorem 3.0.4. *[8] For every* $v \leq u \in \mathbb{R}$ *and* $k \in \mathbb{Z}$*,*

$$Rel_k^{u,v}(X) = Rel_k^{u,v}(A) + Rel_k^{u,v}(B) - Rel_k^{u,v}(Y) + rk\bar{\delta}_k' - rk\bar{\delta}_k'' + rk\bar{\delta}_{k-1}$$

with $\bar{\delta}_k'$*,* $\bar{\delta}_k''$ *and* $\bar{\delta}_{k-1}$ *as:*

$$\cdots \to H_{k+1}(X, X^v) \xrightarrow{\bar{\delta}_k'} H_k(Y, Y^v) \to \cdots$$

$$\cdots \to H_k(X^v, X^u) \xrightarrow{\bar{\delta}_k''} H_{k-1}(Y^v, Y^u) \to \cdots$$

$$\cdots \to H_k(X, X^u) \xrightarrow{\bar{\delta}_{k-1}} H_{k-1}(Y, Y^u) \to \cdots$$

3.0.3 A Mayer-Vietoris Formula for Extended Persistent

We now derive Mayer-Vietoris formula for extended persistence. Since extended persistence sweep the manifold twice from top to bottom and bottom to top, it involves two different cases.

Theorem 3.0.5. *[8] For every $u, v \in \mathbb{R}$ and $k \in \mathbb{Z}$, the following statements hold:*

1. *if $u < v$, then*

$$Ext_k^{u,v}(X) = Ext_k^{u,v}(A) + Ext_k^{u,v}(B) - Ext_k^{u,v}(Y) + rk\hat{\delta}'_k - rk\hat{\delta}''_k + rk\hat{\delta}_{k-1}$$

where $\hat{\delta}'_k$, $\hat{\delta}''_k$ and $\hat{\delta}_{k-1}$ are the maps:

$$\begin{aligned} H_{k+1}(X, X^v) &\xrightarrow{\hat{\delta}'_k} H_k(Y, Y^v), \\ H_{k+1}(X,{}^{\cup}X^v_u) &\xrightarrow{\hat{\delta}''_k} H_k(Y,{}^{\cup}Y^v_u), \\ H_k(X_u) &\xrightarrow{\hat{\delta}_{k-1}} H_{k-1}(Y_u). \end{aligned}$$

2. *if $u \geq v$, then*

$$Ext_k^{u,v}(X) = Ext_k^{u,v}(A) + Ext_k^{u,v}(B) - Ext_k^{u,v}(Y) + rk\tilde{\delta}'_k - rk\tilde{\delta}''_{k-1} + rk\tilde{\delta}_{k-1}$$

where $\tilde{\delta}'_k$, $\tilde{\delta}''_{k-1}$ and $\tilde{\delta}_{k-1}$ are the maps:

$$\begin{aligned} H_{k+1}(X_u,{}^{\cap}X^v_u) &\xrightarrow{\tilde{\delta}'_k} H_k(Y_u,{}^{\cap}Y^v_u), \\ H_k({}^{\cap}X^v_u) &\xrightarrow{\tilde{\delta}''_{k-1}} H_{k-1}({}^{\cap}Y^v_u), \\ H_k(X_u) &\xrightarrow{\tilde{\delta}_{k-1}} H_{k-1}(Y_u). \end{aligned}$$

3.1 Mayer-Vietoris Sequence of Persistence Modules

Next, we show that if instead one works with the persistent modules the Mayer-Vietoris sequence for graded persistence modules is exact. Before we state the theorem, let us define

$$s^u_k \colon H_k((A \cap B)_u) \to H_k(A_u) \text{ defined by } s^u_k([y]) = [y]$$

and also

$$t^u_k \colon H_k((A \cap B)_u) \to H_k(B_u) \text{ defined by } t^u_k([y]) = [-y]$$

so that $\alpha^u_k([y]) = (s^u_k([y]), t^u_k([y]))$.

Theorem 3.1.1. *The Mayer-Vietoris sequence of graded persistence modules*

$$\cdots \to \mathcal{H}_{k+1}(X) \xrightarrow{\delta} \mathcal{H}_k(A\cap B) \xrightarrow{\alpha} \mathcal{H}_k(A)\oplus\mathcal{H}_k(B) \xrightarrow{\beta} \mathcal{H}_k(X) \to \cdots \tag{3.2}$$

where

$$\begin{aligned}\delta(m^0,m^1,\dots,m^n) &= (\delta_k^0(m^0),\delta_k^1(m^1),\dots,\delta_k^n(m^n)),\\ \alpha(y^0,y^1,\dots,y^n) &= ((s_k^0(y^0),s_k^1(y^1),\dots,s_k^n(y^n)),(t_k^0(y^0),t_k^1(y^1),\dots,t_k^n(y^n))),\\ \beta(a,b) &= (\beta_k^0(a^0,b^0),\beta_k^1(a^1,b^1),\dots,\beta_k^n(a^n,b^n)),\end{aligned}$$

for all $m^u \in H_{k+1}^u(X)$*,* $y^u \in H_k^u(A\cap B)$ *and for all* $a^u \in H_k^u(A)$ *and* $b^u \in H_k^u(B)$ *such that* $a = (a^0,a^1,\dots,a^n), b = (b^0,b^1,\dots,b^n)$*, is exact. That is,*

(i) $\operatorname{Im}\delta = \ker\alpha$*,*

(ii) $\operatorname{Im}\alpha = \ker\beta$*,*

(iii) $\operatorname{Im}\beta = \ker\delta$*.*

Proof. We prove only claim (i). Claims (ii) and (iii) can be obtained analogously.

Let $(y^0,y^1,\dots,y^n)\in\operatorname{Im}\delta$. Then there exist an element $(m^0,m^1,\dots,m^n)\in\mathcal{H}_{k+1}(X)$ such that

$$\delta(m^0,m^1,\dots,m^n) = (\delta_k^0(m^0),\delta_k^1(m^1),\dots,\delta_k^n(m^n)) = (y^0,y^1,\dots,y^n).$$

So $\delta_k^u(m^u) = y^u$ and $y^u \in \operatorname{Im}\delta_k^u = \ker\alpha_k^u$ for all u by Lemma 3.0.2. Then we get $\alpha_k^u(y^u) = (s_k^u(y^u),t_k^u(y^u)) = (0,0)$ and $s_k^u(y^u) = t_k^u(y^u) = 0$. Therefore,

$$\begin{aligned}\alpha(y^0,y^1,\dots,y^n) &= ((s_k^0(y^0),s_k^1(y^1),\dots,s_k^n(y^n)),(t_k^0(y^0),t_k^1(y^1),\dots,t_k^n(y^n)))\\ &= ((0,0,\dots,0),(0,0,\dots,0))\end{aligned}$$

and so $(y^0,y^1,\dots,y^n)\in\ker\alpha$. Hence, we get $\operatorname{Im}\delta\subseteq\ker\alpha$

Let $(c^0,c^1,\dots,c^n)\in\ker\alpha$. Then

$$\begin{aligned}\alpha(c^0, c^1, \ldots, c^n) &= ((s_k^0(c^0), s_k^1(c^1), \ldots, s_k^n(c^n)), (t_k^0(c^0), t_k^1(c^1), \ldots, t_k^n(c^n))) \\ &= ((0, 0, \ldots, 0), (0, 0, \ldots, 0)).\end{aligned}$$

By equalities of the tuples, we get $s_k^u(c^u) = 0$ and $t_k^u(c^u) = 0$ for all $c^u \in H_k^u(A \cap B)$. Thus, $c^u \in \ker \alpha_k^u = \operatorname{Im} \delta_k^u$ by Lemma 3.0.2. Then there exist an element $m^u \in H_{k+1}^u(X)$ such that $\delta_k^u(m^u) = c^u$ for all $u \geq 0$, and then we get

$$(c^0, c^1, \ldots, c^n) = (\delta_k^0(m^0), \delta_k^1(m^1), \ldots, \delta_k^n(m^n)) = \delta(m^0, m^1, \ldots, m^n).$$

Therefore, $(c^0, c^1, \ldots, c^n) \in \operatorname{Im} \delta$ which implies that $\ker \alpha \subseteq \operatorname{Im} \delta$.

□

Now, let us briefly see how the sequence that is just constructed gives an exact sequence for the Example 5. Recall that in Example 5, the reason for non-exactness is that there is a $\gamma \in H_1^{95,100}(A \cap B)$ that belongs to $\ker \alpha$ but not in $\operatorname{Im} \delta$. In terms of persistence modules γ corresponds to $(0, 0, 0, 0, \gamma, 0) \in \mathcal{H}_1(A \cap B)$ and $\alpha(0, 0, 0, 0, \gamma, 0) = ((0, 0, 0, 0, 0, 0), (0, 0, 0, 0, \gamma, 0)) \in \mathcal{H}_1(A) \bigoplus \mathcal{H}_1(B)$. Hence it is not in $\ker \alpha$. Moreover $x \cdot (0, 0, 0, 0, \gamma, 0) = (0, 0, 0, 0, 0, \gamma)$ and $\alpha(0, 0, 0, 0, 0, \gamma) = ((0, 0, 0, 0, 0, 0), (0, 0, 0, 0, 0, 0))$. But in this case $(0, 0, 0, 0, 0, \sigma) \in \mathcal{H}_2(X)$ and $\delta(0, 0, 0, 0, 0, \sigma) = (0, 0, 0, 0, 0, \gamma)$.

REFERENCES

[1] P. K. Agarwal, H. Edelsbrunner, J. Harer, and Y. Wang. Extreme elevation on a 2-manifold. *Discrete & Computational Geometry*, 36(4):553–572, 2006.

[2] S. Bhattacharya, R. Ghrist, and V. Kumar. Persistent homology for path planning in uncertain environments. *IEEE Transactions on Robotics*, 31(3):578–590, 2015.

[3] G. Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[4] G. Carlsson and V. De Silva. Zigzag persistence. *Foundations of computational mathematics*, 10(4):367–405, 2010.

[5] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007.

[6] V. De Silva and R. Ghrist. Homological sensor networks. *Notices of the American mathematical society*, 54(1), 2007.

[7] 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(5):499–527, 2011.

[8] B. Di Fabio and C. Landi. Persistent homology and partial similarity of shapes. *Pattern Recognition Letters*, 33(11):1445–1450, 2012.

[9] H. Edelsbrunner and J. Harer. *Computational topology: an introduction*. American Mathematical Soc., 2010.

[10] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 454–463. IEEE, 2000.

[11] K. Emmett, B. Schweinhart, and R. Rabadan. Multiscale topology of chromatin folding. In *Proceedings of the 9th EAI international conference on bio-inspired information and communications technologies (formerly BIONETICS)*, pages 177–180. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2016.

[12] R. Forman. Witten–morse theory for cell complexes. *Topology*, 37(5):945–979, 1998.

[13] R. Forman. A user's guide to discrete morse theory. *Sém. Lothar. Combin*, 48:35pp, 2002.

[14] P. Frosini. A distance for similarity classes of submanifolds of a euclidean space. *Bulletin of the Australian Mathematical Society*, 42(3):407–415, 1990.

[15] R. Ghrist. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75, 2008.

[16] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[17] M. Kramar, A. Goullet, L. Kondic, and K. Mischaikow. Persistence of force networks in compressed granular media. *Physical Review E*, 87(4):042207, 2013.

[18] J.-P. Litza, E.-M. Feichtner, and D. Feichtner-Kozlov. Extended persistence. 2012.

[19] T. Nakamura, Y. Hiraoka, A. Hirata, E. G. Escolar, and Y. Nishiura. Persistent homology and many-body atomic structure for medium-range order in the glass. *Nanotechnology*, 26(30):304001, 2015.

[20] F. T. Pokorny, M. Hawasly, and S. Ramamoorthy. Topological trajectory classification with filtrations of simplicial complexes and persistent homology. *The International Journal of Robotics Research*, 35(1-3):204–223, 2016.

[21] V. Robins. Towards computing homology from finite approximations. In *Topology proceedings*, volume 24, pages 503–532, 1999.

[22] P. Skraba, M. Ovsjanikov, F. Chazal, and L. Guibas. Persistence-based segmentation of deformable shapes. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 45–52. IEEE, 2010.

[23] K. Turner, S. Mukherjee, and D. M. Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[24] R. Vasudevan, A. Ames, and R. Bajcsy. Persistent homology for automatic determination of human-data based cost of bipedal walking. *Nonlinear Analysis: Hybrid Systems*, 7(1):101–115, 2013.

[25] K. G. WANG. The basic theory of persistent homology, 2012.

[26] K. Xia and G.-W. Wei. Persistent homology analysis of protein structure, flexibility, and folding. *International journal for numerical methods in biomedical engineering*, 30(8):814–844, 2014.

[27] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.