

PERSISTENT HOMOLOGIES OF SIMPLE HOMEOMORPHIC POINT CLOUDS

EZRA SANTOS

ABSTRACT. Persistent homology is a recent, yet widely-used technique in topological data analysis due to its effectiveness in studying the behavior of data by efficiently detecting central holes and noise in point clouds. This paper aims to introduce simplicial homology and computing simplicial homology groups, singular homology, homotopy invariance, determining persistence and constructing barcode and persistence diagrams from the Rips complex, and finally comparing the persistence of simple homeomorphic spaces.

CONTENTS

| | |
|--|----|
| 1. Introduction | 1 |
| 2. Simplicial Homology | 2 |
| 3. Other Useful Concepts | 4 |
| 4. Persistent Homology | 5 |
| 5. Persistence of Simple Homeomorphic Shapes | 8 |
| Acknowledgments | 13 |
| References | 13 |

1. INTRODUCTION

Topological data analysis has been a powerful tool in studying data sets in various fields such as social science, ethnography, and physics. Realistically, a huge amount of large point clouds that are used in statistical studies are met with noise, such as a handful of anomalies in data that does not necessarily agree with the rest of the data set. This makes it difficult even for the human eye to assess. Algorithms from a handful of subfields have been used to tackle this problem, and one such method is the use of persistent homology. Using the basic foundations of algebraic topology, persistent homology attempts to find central holes in large point clouds, efficiently filtering out the noise.

The first section introduces simplices, which are essentially n -dimensional triangles. From there, we can glue those simplices together to form Δ -complexes and simplicial complexes. Then, we introduce the concept of simplicial homology, which is a method of studying the structure of simplicial complexes.

The second section introduces concepts that will be useful in persistent homology such as singular homology and homotopy invariance. At the end of the section, we will be able to assess the homology groups of homeomorphic spaces and any two spaces where one is deformation retractable to the other.

Date: September 13, 2023.

The third section dives into persistent homology. The Rips complex will be introduced, and several terms essential to persistent homology will be defined as well, such as filtered simplicial complexes, barcode diagrams, and persistence diagrams.

The last section compares the persistent homology of point clouds that represent homeomorphic shapes. Using Python, the persistent homology of the following pairs are evaluated: the circle and the ellipse, and the square and the rectangle.

2. SIMPLICIAL HOMOLOGY

We start with the concept of simplices, simplicial complexes, and computing homology groups of simplicial complexes.

Definition 2.1. Assigning vertices $[v_0, \dots, v_n]$ to the unit vectors along the coordinate axes of \mathbb{R}^{n+1} , we can define the standard n -simplex as

$$\Delta^n = \{(t_0, \dots, t_n) \in \mathbb{R}^{n+1} \mid \sum_i t_i = 1 \text{ and } t_i > 0 \text{ for all } i\}$$

Definition 2.2. Let $\Delta^n = [v_0, \dots, v_n]$ be an n -simplex in \mathbb{R}^n . Any $n-1$ dimensional subset of Δ^n is a face of Δ^n .

Definition 2.3. The boundary of Δ^n , namely $\delta\Delta^n$ is the union of all faces of Δ^n .

Definition 2.4. The interior of Δ^n , namely the open simplex $\mathring{\Delta}^n$, is defined as $\Delta^n - \delta\Delta^n$.

We can glue different n -simplices together in order to form Δ -complexes.

Definition 2.5. A Δ -complex on a space X is a collection of maps $\sigma_\alpha : \Delta^n \rightarrow X$, with n depending on the index α , such that:

- (1) The restriction $\sigma_\alpha|_{\mathring{\Delta}^n}$ is injective, and each point of X is in the image of such restriction.
- (2) Each restriction of σ_α to a face of Δ^n is one of the maps $\sigma_\beta : \Delta^{n-1} \rightarrow X$, where we identify Δ^{n-1} as the face of Δ^n .
- (3) A set $A \subset X$ is open if and only if $\sigma_\alpha^{-1}(A)$ is open in Δ^n for each σ_α .

A simplicial complex is a Δ -complex with simplices uniquely determined by its vertices.

We want to relate a simplicial complex to its simplices in each dimension, which will be done through simplicial homology.

Definition 2.6. Let X be a simplicial complex and $\Delta_n(X)$ the free abelian group, whose basis is given by the open simplices $e_\alpha^n \in X$ where $e_\alpha^n = (0, \dots, 1, \dots, 0)$ such that the 1 is in the n -th position. Elements of $\Delta_n(X)$, which are also called n -chains, are written as finite sums $\sum_\alpha n_\alpha e_\alpha^n$, for $n_\alpha \in \mathbb{Z}$.

To relate elements of an n -chain, we define the boundary operator.

Definition 2.7. Let $\sigma_\alpha : \Delta^n \rightarrow X$ be a mapping from an n -simplex to a topological space X . The boundary operator is defined as the homomorphism $\delta_n : \Delta^n \rightarrow \Delta^{n-1}$ such that

$$\delta_n(\sigma_\alpha) = \sum_i (-1)^i \sigma | [v_0, \dots, \hat{v}_i, \dots, v_n]$$

where \hat{v}_i denotes an omission of the i -th vertex.

Boundary operators map from an n -simplex to an $n - 1$ simplex, so applying the operator constantly yields the following chain.

$$\dots \xrightarrow{\delta_{n+2}} C_{n+1} \xrightarrow{\delta_{n+1}} C_n \xrightarrow{\delta_n} C_{n-1} \xrightarrow{\delta_{n-1}} \dots$$

In order to understand the structure of a simplicial complex through its respective chain, we introduce the concept of homology, which gives us an idea of the shape itself through its cycles and boundaries.

To find the cycles of a specific boundary operator, we find all possible linear combinations of paths such that the resulting path creates a loop. The amount of generators yield the kernel of the n -th boundary operator. For example, the 2nd boundary operator of a standard 2-simplex is isomorphic to \mathbb{Z} because the only generator is the sum of its vertices. Computing the boundaries of a shape involves finding all possible combinations of the boundary operator in the $n + 1$ dimension, hence we evaluate the image of the $n + 1$ boundary operator [1].

We use the concept of cycles and boundaries to define homology groups.

Definition 2.8. The n -th simplicial homology group is defined as the quotient group

$$H_n^\Delta(X) = \frac{Z_n^\Delta}{B_n^\Delta} = \frac{\ker \delta_n}{\text{Im } \delta_{n+1}}$$

where Z_n^Δ and B_n^Δ correspond to the cycles and boundaries of X , respectively.

Example 2.9. As an example, we will find the n -th simplicial homology group of a standard hollow triangle with ordered vertices $[v_0, v_1, v_2]$.

Proof. First, we will calculate the 0th homology group. We have three 0-simplices in a 2-simplex, and we know that every element in C_0 maps to zero. Thus, $\ker \delta_0 \cong \bigoplus_{i=1}^3 \mathbb{Z}$. To find $\text{Im } \delta_1$, let $a, b, c \in \mathbb{Z}$. We produce the following linear combination:

$$\begin{aligned} a\delta_1[v_0, v_1] + b\delta_1[v_1, v_2] - c\delta_1[v_0, v_2] &= a(v_1 - v_0) + b(v_2 - v_1) - c(v_2 - v_0) \\ &= v_0(c - a) + v_1(a - b) + v_2(b - c). \end{aligned}$$

There are two generators for $\text{Im } \delta_1$. Hence, $H_0^\Delta(X) = \frac{\bigoplus_{i=1}^3 \mathbb{Z}}{\mathbb{Z} \oplus \mathbb{Z}} \cong \mathbb{Z}$. To find $H_1^\Delta(X)$, we need to find $\ker \delta_1$. By definition, we find the kernel by setting the linear combination to zero.

$$v_0(c - a) + v_1(a - b) + v_2(b - c) = 0$$

This is only possible if $a = b = c$, thus there is only one generator. Hence, $\ker \delta_1 \cong \mathbb{Z}$. Notice that since the triangle is hollow, there are no 2-simplices. This implies that the only possible input of δ_2 is zero. Hence, $H_1(X) = \frac{\mathbb{Z}}{0} \cong \mathbb{Z}$. The homology groups of higher dimensions would yield $\frac{0}{0}$, which is 0, as 0 represents the trivial group. Hence,

$$H_n(X) = \begin{cases} \mathbb{Z} & \text{if } n = 0, 1 \\ 0 & \text{else} \end{cases}$$

□

The following propositions will be useful in studying persistent homology.

Proposition 2.10. *Corresponding to the decomposition of a simplicial complex X into its path-components X_α there is an isomorphism of $H_n^\Delta(X)$ with the direct sum $\bigoplus_\alpha H_n^\Delta(X_\alpha)$.*

Proof. Let X be a simplicial complex. Fix $n \in \mathbb{N}$. Notice that taking the n -th boundary operator of X yields the sum of all n -th boundary operators of individual n -simplices. Thus, the direct sums of $\ker(\delta_n(X_\alpha))$ and $\text{Im}(\delta_{n+1}(X_\alpha))$ are isomorphic to $\ker(\delta_n(X))$ and $\text{Im}(\delta_{n+1}(X))$, respectively. Hence, $\bigoplus_\alpha H_n^\Delta(X_\alpha) \cong H_n^\Delta(X)$. \square

Proposition 2.11. *Let X be a nonempty and path-connected simplicial complex. Then $H_0^\Delta \cong \mathbb{Z}$.*

Proof. Let X be a nonempty and path-connected simplicial complex with n 0-simplices. Fix a 0-simplex x in X . For any 0-simplex y in X , there exists a continuous function $\epsilon : I = [0, 1] \rightarrow X$ such that $\epsilon(0) = x$ and $\epsilon(1) = y$. Thus, $\delta_1(I) = y - x$. Since this holds true for all 0-simplices in X , $B_0^\Delta(X)$ has $n - 1$ generators. We know that δ_0 yields the zero map, thus $Z_0^\Delta(X) \cong \bigoplus_{i=1}^n \mathbb{Z}$. Hence, $H_0^\Delta \cong \mathbb{Z}$. \square

Proposition 2.12. *If X is a point, then $H_n(X) = 0$ for $n > 0$ and $H_0(X) \cong \mathbb{Z}$.*

Proof. Let X be a point. Since X contains one 0-simplex and no n -simplices for all $n \in \mathbb{N}$, $\ker \delta_0 \cong \mathbb{Z}$ and $\text{Im} \delta_1 = 0$. Thus, $H_0(X) \cong \mathbb{Z}$. Now fix $n > 0$. Then $\ker \delta_n = 0$ and $\text{Im} \delta_{n+1} = 0$. Hence, $H_n(X) = 0$. \square

3. OTHER USEFUL CONCEPTS

Singular homology is a way to find the homology groups of any topological space, not just simplicial complexes. Finding the homology groups of a more general space uses the same concepts of cycles and boundaries.

Definition 3.1. A singular n -simplex is defined as the map $\sigma : \Delta^n \rightarrow X$ where Δ^n is a Δ -complex.

Definition 3.2. Let $C_n(X)$ be a free abelian group whose bases are given by the set of singular n -simplices in X . Then singular n -chains are elements of $C_n(X)$ defined by the finite sum $\sum_i n_i \sigma_i$ for $n_i \in \mathbb{Z}$ and $\sigma_i : \Delta^n \rightarrow X$.

Definition 3.3. A boundary map $\delta_n : C_n(X) \rightarrow C_{n-1}(X)$ is formally defined as

$$\delta_n(\sigma) = \sum_i (-1)^i \sigma[[v_0, \dots, \hat{v}_i, \dots, v_n]$$

where \hat{v}_i denotes an omission of the i -th vertex.

Definition 3.4. The n -th singular homology group is defined as the quotient group

$$H_n(X) = \frac{Z_n(X)}{B_n(X)} = \frac{\ker \delta_n}{\text{Im} \delta_{n+1}}$$

Persistent homology only deals with simplicial complexes, but singular homology is needed as it gives us context to homeomorphic spaces. These definitions are similar to their simplicial counterparts, except that the map $\sigma : \Delta^n \rightarrow X$ implies that the image does not have to be a simplicial complex. It follows that homeomorphic spaces have isomorphic singular homology groups, but it is not necessarily so with simplicial homology groups [1].

In persistent homology, we will encounter a lot of simplicial complexes that contain hundreds of simplices. Most of the time, those simplicial complexes have trivial homology groups due to homotopy invariance.

Definition 3.5. A homotopy is defined as the family of maps $f_t : X \rightarrow Y, t \in I$, such that the function $F : X \times I \rightarrow Y$ given by $F(x, t) = f_t(x)$ is continuous.

Since I is the unit interval, then we can say that f_0 and f_1 are homotopic if there exists a homotopy f_t [1].

One can intuitively see that homotopies can be used to retract topological spaces, which is defined as follows:

Definition 3.6. Let $A \subset X$ be a subspace of a topological space X . Then A is a deformation retract of X if there exists a homotopy $F : X \times I \rightarrow X$ such that the following holds for all $x \in X$ and $a \in A$:

- (1) $F(x, 0) = x$
- (2) $F(x, 1) \in A$
- (3) $F(a, 1) = a$.

Let us, for example, consider an open-ended cylinder. This cylinder can be collapsed to a circle. Since the circle is a subspace of the cylinder, we say that the cylinder is deformation retractable to a circle.

Theorem 3.7. *If X is deformation retractable to Y , then $H_n(X) \cong H_n(Y)$.*

A proof is provided in Hatcher’s book. Using Theorem 3.7 along with Proposition 2.12 tells us that if a space is deformation retractable to a point, then that space has trivial homology.

Notice that any mapping of two simplicial complexes $f : K \rightarrow K'$ induces a homomorphism $f_* : H_n(K) \rightarrow H_n(K')$. This fact will be used when formally defining persistent homologies.

4. PERSISTENT HOMOLOGY

The goal of persistent homology is to be able to study the behavior of a large data set when put into a point cloud. This is done by creating a simplicial complex in the point cloud and studying the homology groups of that simplicial complex given a distance parameter.

For this section, we will be repeatedly using the concepts from the simplicial homology section and the propositions proven in the singular homology section. To avoid confusion, we use the notation $H_n(X)$ instead of $H_n^\Delta(X)$, as singular homology will not be involved anymore.

Suppose we have a set of points in a point cloud. Using the point cloud, we want to be able to create a simplicial complex. Thus, we will introduce the Rips complex.

Definition 4.1. Let x_α be a collection of points in Euclidean space \mathbb{E}^n . Then the Rips complex, labeled R_ϵ , is a simplicial complex such that each k -tuple of points $\{x_\alpha\}_0^k$ form a k -simplex if and only if every two points are at most ϵ apart.

For example, if three points are pairwise $\frac{\epsilon}{2}$ apart, then a 2-simplex would be formed. For different values of ϵ , different simplicial complexes could be formed, and thus the homology of the simplicial complex could change.

The goal of persistent homology is to be able to discern what the behavior of the data is when translated into a point cloud. That means a right distance parameter ϵ is necessary. The problem is what ϵ to choose: if $\epsilon = 0$, then we would get many 0-simplices, which does not give us meaningful information about the data.

However, if we choose a large ϵ , then according to Theorem 3.7 the shape itself will have trivial homology. This is why we analyze the behavior of data not from distinct individual distance parameters, but rather from the change of homology groups for all values of ϵ .

Definition 4.2. Let K be a simplicial complex with subsets $K_1 \subset K_2 \subset \dots \subset K_l = K$. Then K is a filtered simplicial complex.

Definition 4.3. Let $K_1 \subset K_2 \subset \dots \subset K_l$ be a filtered simplicial complex. Then the p -th persistent homology is defined as the pair

$$(\{H_p(K_i)\}_{1 \leq i \leq l}, \{f_{i,j}\}_{1 \leq i < j \leq l})$$

where $f : H(K_i) \rightarrow H(K_j)$ is the induced homomorphism from the inclusion map $K_i \hookrightarrow K_j$.

As a simple example, consider the following points in the Cartesian plane: $(0, 0)$, $(0, 1)$, $(-1, 3)$, and $(2, 0)$. Below is a diagram of two separate filtration steps and a table containing the 0th and 1st homology groups of all possible simplicial complexes.

| K | H_0 | H_1 |
|-------|----------------|-------|
| K_1 | \mathbb{Z}^4 | 0 |
| K_2 | \mathbb{Z}^3 | 0 |
| K_3 | \mathbb{Z}^2 | 0 |
| K_4 | \mathbb{Z} | 0 |
| K_5 | \mathbb{Z} | 0 |
| K_6 | \mathbb{Z} | 0 |

Table 1. The H_0 and H_1 homology groups of the example given above.

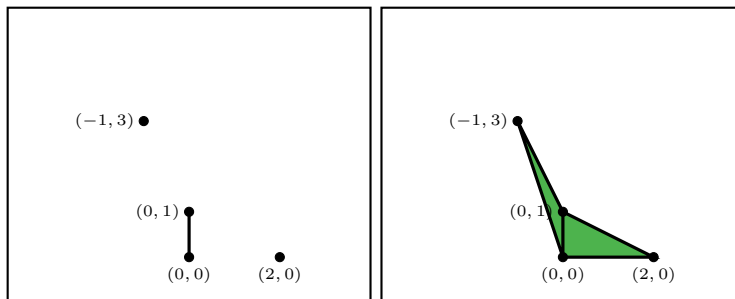


Figure 1. Visualizations of K_2 and K_5 .

As stated in Propositions 2.10 and 2.11, the zeroth homology group of the given Rips complex is the direct sum of the zeroth homology group of its components, each of which is isomorphic to \mathbb{Z} .

These new simplicial complexes are formed due to the gradual increase of the distance parameter ϵ , in accordance with the definition of the Rips complex. Additionally, notice that at any point in this filtered simplicial complex, there are no H_1 groups formed at all. This is because according to the definition of the Rips complex, once three vertices are pairwise within ϵ apart, it becomes a filled 2-simplex, which does not contribute to the first homology group.

Finding the homology groups of each simplicial complex doesn't tell us about the true structure of this specific data set. We need to find a way to determine how

each point in the data set persists over an increasing distance parameter. This is where barcodes come into play. The existence of barcodes are guaranteed because of Definition 3.4. For setting up an H_0 barcode diagram, we will let the x -axis be the filtration step (the step in which a new simplicial complex is formed) and the y -axis be the different connected components initially formed. The following shows the H_0 barcode diagram over all the filtration steps.

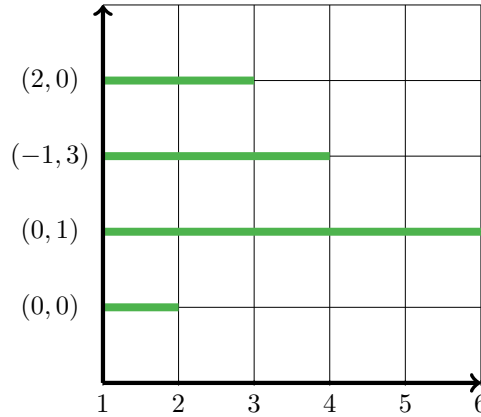


Figure 2. An H_0 barcode diagram.

A more convenient and preferred way of looking at the persistence of a point cloud is through a persistence diagram.

Definition 4.4. A persistence diagram is a space $\overline{\mathbb{R}}^2$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, such that a point (i, j) is represented by the barcode interval $[i, j)$.

A typical persistence diagram plots the birth of components against its death, and an origin that is located at $(1, 1)$ instead of $(0, 0)$. We take a point in the barcode diagram, for example $(0, 0)$, and its location in the x -axis is the filtration step in which it was born. Similarly, its location in the y -axis is the filtration step in which it dies. Plotting all four points from the example in a persistence diagram, we get the following.

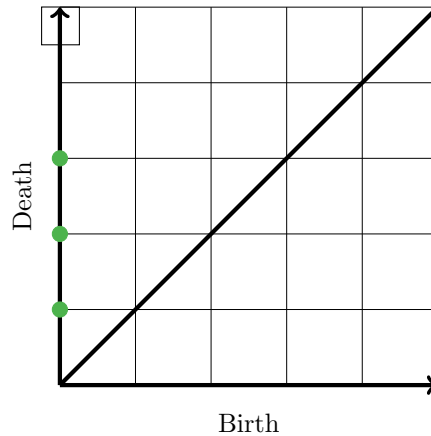


Figure 3. A persistence diagram.

The box on $(1, 6)$ is due to the fact that the point $(0, 1)$ doesn't really die at the sixth step. As ϵ increases, the connected component that is deformation retractable to $(0, 1)$ will stay the same forever, and thus the box signifies a component that will persist forever.

Indeed, this example's absence of any H_1 groups does not reveal us any central holes. Intuitively, this tells us one of two things: there are not enough data points, or this point cloud is almost similar to a single point. When H_1 groups are involved, the $y = x$ line in the persistence diagram would be useful, as points that are farthest away from this line are defined as the most persistent, and thus would signify a central hole in the point cloud [3].

It's possible to arrange these points in a way that would create a first homology group. We can imagine the points as the vertices of a rhombus. We can intuitively surmise that there will exist such an arrangement so that a specific distance parameter ϵ will create 1-simplices from adjacent vertices, and nothing else. This does not create any 2-simplices, and it counts as a one-dimensional hole, hence there exists an arrangement of four points in the Cartesian plane such that some Rips complex will have a non-zero first homology group. If this central hole persists long enough, then that H_1 would be more than just noise, but rather an important aspect of the behavior of this new point cloud.

If we can arrange a set of points and have different homology groups, it is worth discovering whether two different datasets that resemble two homeomorphic spaces will have different persistent homologies and potential causes of such differences.

5. PERSISTENCE OF SIMPLE HOMEOMORPHIC SHAPES

We have discussed in Section 3 that homeomorphic spaces have the same homology groups. Specifically, Theorem 3.7 tells us that if a topological space X is deformation retractable to a another topological space Y , then the homology groups of X and Y are the same.

In this section, we will take two data sets that resemble simple homeomorphic shapes: a circle and an ellipse. Intuitively, their diagrams should tell us that there is a central hole that persist for some time. The construction of the Rips complex and the generation of persistence diagrams was done through an original Python script that uses the MatLab and Ripser packages.

The procedure, with slight modifications of the diagrams shown in the previous section, is as follows: For the circle, a zone is made from two circles with radii 0.9 and 1.1, and then 200 random points are scattered throughout this zone. The persistence diagram won't have filtration steps in its axes, but rather the ϵ value. As a feature of the Ripser package, instead of having a box to represent infinite persistence, points will lie on a horizontal line near the top. The procedure for the ellipse is similar; the inner and outer semi-minor axes are 0.9 and 1.1, respectively. The semi-major axes are determined by a ratio modifier that is multiplied with the semi-minor axes. For this section, we will use the following ratio modifiers: 1, 3, and 10. We will be using two ellipse modifiers to surmise if eccentricity is a criterion for its change in persistence (if there are any changes at all). A GitHub repository is available here, along with potentially more examples that will not be used in this paper: <https://github.com/esan0983/phcalculations>. Refer to "ellipse.py" for this part.

The six images show the plots of the three shapes and their respective persistence diagrams.

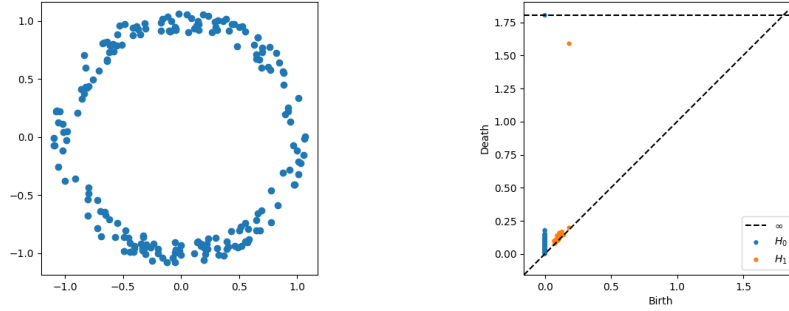


Figure 4. A 200-point plot of a circle (ratio 1) and its respective persistence diagram.

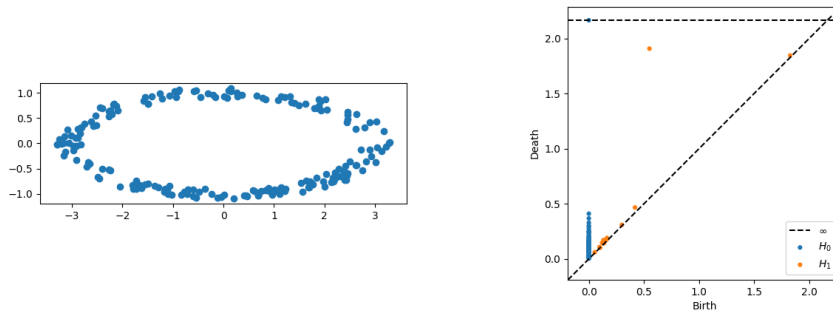


Figure 5. A 200-point plot of an ellipse with a ratio of 3 and its respective persistence diagram.

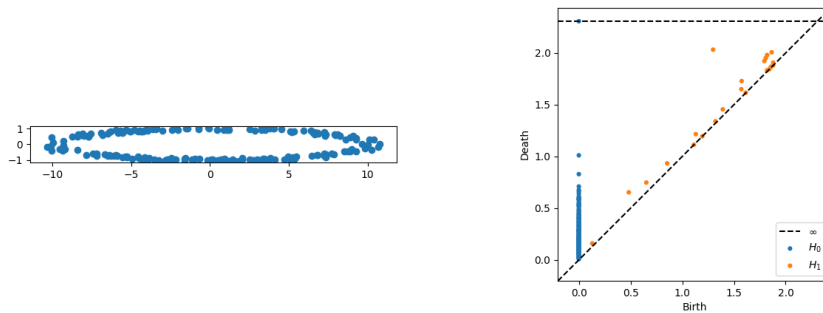


Figure 6. A 200-point plot of an ellipse with a ratio of 10 and its respective persistence diagram.

For ratio modifiers of 1 and 3 (the first four diagrams), it is clear where the central hole persists. However, while it is clear enough that there is a persistent central hole in the third case (the last two diagrams), there are significantly more holes that don't die immediately. The results for the third case overall were more chaotic than the previous ones. Running the program two more times for each shape confirms

that fact, since the persistence of the first two modifiers were almost exactly the same each time, but the set of persistence points for the third modifier abruptly changed every time. The images below show two more iterations of the code with a ratio modifier equal to 10.

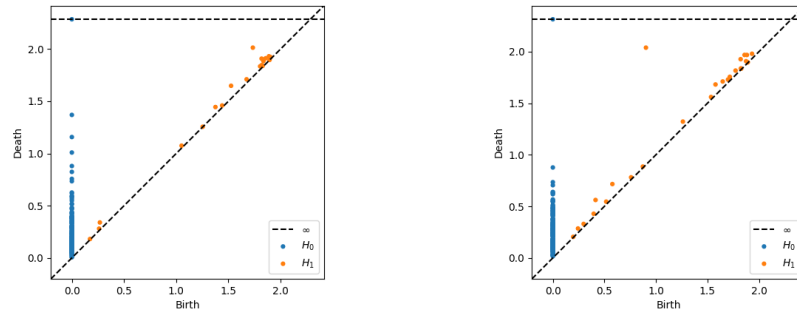
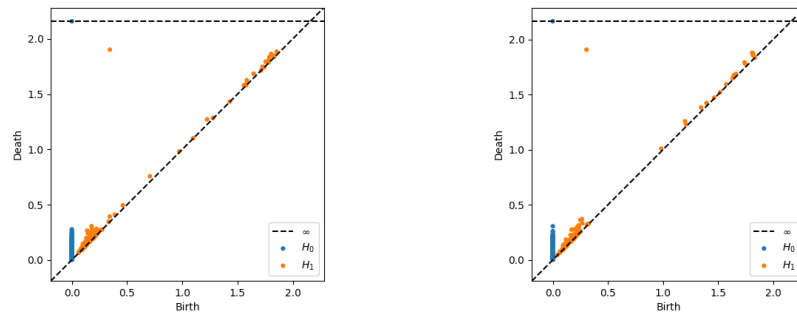


Figure 7. Two more persistence diagrams of an ellipse with a ratio of 10.

The first point of interest here is the formation of one-dimensional holes for all three shapes. For the circle, most of the holes form and die at the same ϵ value, which follows from the fact that the shape itself is radially symmetrical; a point will create lines with their neighbors almost at the same time as other points do. As the circle gets more stretched, the one-dimensional holes persist around the same time, but the times of their birth and death become more scattered.

Another point of interest is, as already mentioned, the instability of the persistence diagram as the ratio modifier increases. The human eye would be able to figure out what the central hole of an ellipse is, without regard to the eccentricity. However, persistent homology is heavily reliant on the discrete data set that is given. If a circle is stretched with an equal amount of data points, the point density would decrease, and would thus change the birth and death of one-dimensional holes.

With that in mind, the program ran three more times on the same ratio modifier of 10, but now with a sample size of 1000.



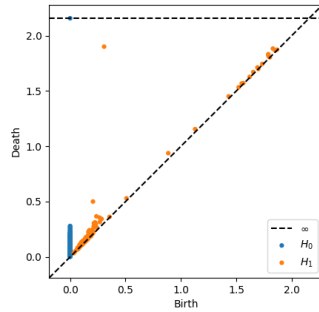


Figure 8. Three persistence diagrams of an ellipse with a ratio of 10 with a sample size of 1000.

Indeed, due to the law of large numbers, the standard deviation of the persistence of the central hole approaches zero as sample size increases. The more data points, the more continuous the shape seems, and the more likely that persistent homology will view the shape as if it has a human eye.

Nonetheless, the birth and death of the central holes are different for each shape. A modified, separate version of the code iterates through all possible ratios from one to 20, with 0.1 increments. The birth and death of the central hole follows a general trend: the higher the ratio, the later the birth and death. Refer to "eclipse_progression.py" for this part.

The GIF is included in the repository as "ellipse.gif," but four instances of varying ratios will be shown below.

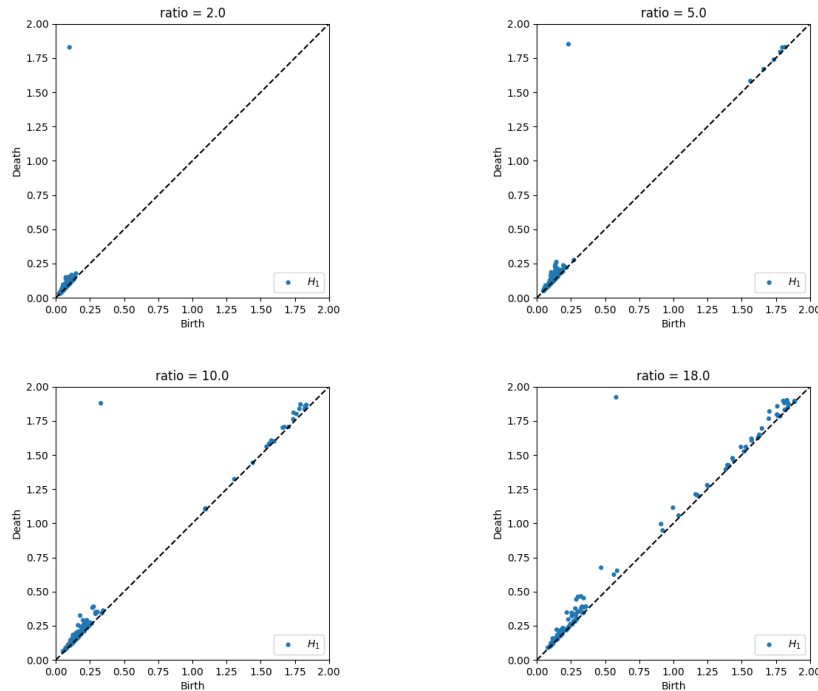


Figure 9. Four persistence diagrams of an ellipse with ratios 2, 5, 10, and 18.

This gives rise to the hypothesis that for homeomorphic shapes, as long as there is a way to determine the major and minor axes of the shape, the ratio between the two determines its overall persistence. In order to explore this even further, a similar program was made for a rectangle, with the major and minor axes being the width and height of the rectangle. Additionally, the ratios in this program will only go from one to ten. Since a significantly larger sample size is used, a subsampling algorithm that is implemented in the same package will be used, which yields a close approximation of resulting persistence diagrams [6]. As per usual, the GIF will be in the repository as "rectangle.gif," and for the program, refer to "rectangle_progression.py." The four instances will be shown below.

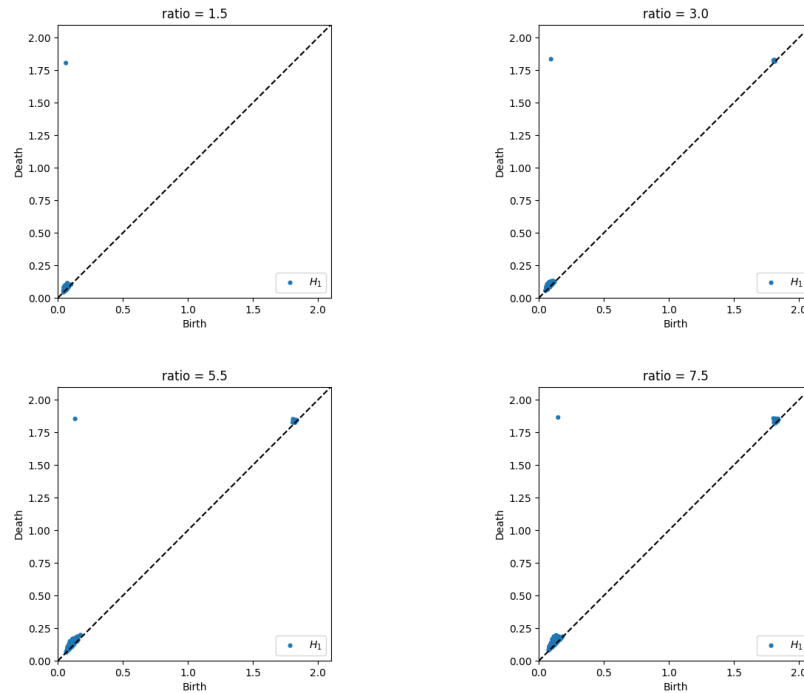


Figure 10. Four persistence diagrams of an ellipse with ratios 1.5, 3, 5.5, and 7.5.

Although a little subtler, the central hole in the persistence diagram still follows the same trend.

A question of interest that is left unanswered in this paper is if this hypothesis can be generalized to any collections of homeomorphic shapes. The evidence shown in this section worked because the examples are simple shapes, meaning that their minor and major axes can easily be determined. A potential follow up is exploring whether a trend in persistence can still be seen with general and/or random shapes, and if so, the criteria (something more reliable than ratio between axes) needed to exhibit that trend. Constructing simpler shapes around general, more random shapes and taking the persistent homology of those circumscribed simpler shapes is a potential approximation method, but it would need further verification and rigorous testing.

ACKNOWLEDGMENTS

I would like to thank my mentor, Adolfo Medina, for helping me revise the paper and making sure that I understood the foundations of algebraic topology despite my complete inexperience. I would also like to thank Adolfo's mentor, Professor Carmen Rovi, for introducing me to the beautiful subfield of persistent homology and showering us with amazing ideas for the project. I'm also grateful for our apprentice program lecturers, Professor Daniil Rudenko and Professor Laszlo Babai, for introducing us to a lot of new and beautiful math concepts, as well as the other mentors who help out on the problem sessions. Lastly, I would like to thank Professor Peter May for running the REU.

REFERENCES

- [1] Allen Hatcher. Algebraic Topology. Cambridge University Press. 2002
- [2] Robert Ghrist. Barcodes: The Persistent Topology of Data. <https://www.ams.org/journals/bull/2008-45-01/S0273-0979-07-01191-3/S0273-0979-07-01191-3.pdf>
- [3] Nina Otter, Mason Porter, Ulrike Tillman, and Heather Harrington. A roadmap for the computation of persistent homology. <https://doi.org/10.1140/epjds/s13688-017-0109-5>
- [4] Christopher Tralie, Nathaniel Saul, and Rann Bar-On. Ripser.py: A Lean Persistent Homology Library for Python. <https://joss.theoj.org/papers/10.21105/joss.00925>
- [5] Ulrich Bauer. Ripser: efficient computation of Vietoris–Rips persistence barcodes. <https://doi.org/10.1007/s41468-021-00071-5>
- [6] Nicholas Cavanna, Mahmoodreza Jahanseir, and Donald Sheehy. A Geometric Perspective on Sparse Filtrations. <https://doi.org/10.48550/arXiv.1506.03797>