ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Detecting Singularities: A Topological Approach

Bachelor Thesis

Anna Walder

14.10.2024

Advisor: Dr. Sara Kališnik Hintz
Department of Mathematics, ETH Zürich

Abstract

In this thesis we present TARDIS (Topological Algorithm for Robust DIscovery of Singularities), a novel framework for detecting singularities within data sets. The methodology leverages persistent homology, a popular invariant in topological data analysis (TDA), to measure intrinsic dimensional changes and identify singular points. In part, persistent intrinsic dimension (PID) and Euclidicity are two metrics within TARDIS that quantify the local geometric and topological properties of data. We demonstrate the effectiveness and limitations of TARDIS through a series of experiments.

Acknowledgements

I am very grateful for my advisor, Dr. Sara Kališnik Hintz for her support with this thesis. Thanks to her advice, comments and input I could learn so much about topological data analysis and about writing a thesis. Having Sara as a mentor has been an incredibly enriching experience.

Contents

Acronyms and Abbreviations

w.r.t.	with respect to
TDA	Topological Data Analysis
Point cloud	finite metric space
PVS	Persistent vector space
PID	Persistent Intrinsic Dimension
PLH	Persistent Local Homology
point cloud	finite metric space

Chapter 1

Introduction

Many machine learning algorithms assume that the underlying data lies on or close to an unknown manifold of low intrinsic dimension. This is called the *manifold hypothesis*. Manifolds are rigid topological spaces that resemble Euclidean spaces locally. Each point has a neighborhood homeomorphic to some $\mathbb{R}^n$ for some n. This n is referred to as the intrinsic dimension of the point. The intrinsic dimension of points on a manifold is constant. For some data sets the manifold hypothesis is justified by prior knowledge about the data, for example, if the data entries are highly correlated. One such example are black and white natural pictures. For complex data the manifold hypothesis might fail.

In many such cases we can model the underlying space as a union of several manifolds with varying intrinsic dimension, called a stratified space. In a stratified space points that are on a manifold are called regular, points that do not have such a neighborhood are singularities. An example of a stratified space is the pinched torus which is illustrated in Figure 1.1. Here any neighborhood of the pinched point p differs from a neighborhood of any other point, so this is clearly not a manifold.

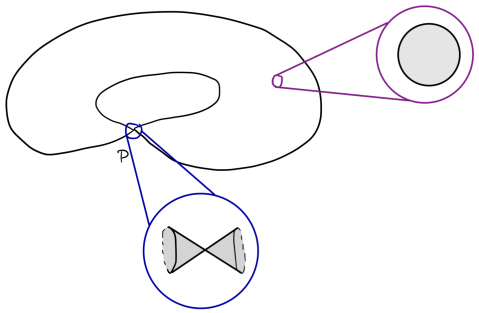


Figure 1.1: Pinched torus with two types of neighborhoods highlighted.

In this report we discuss the TARDIS algorithm [VRR23], which is designed to detect singularities in some given data set. This algorithm measures how similar the neighborhood of some data point is to a neighborhood of a regular point. This is captured by the Euclidicity score, where a low score indicates a regular point. To measure this similarity the algorithm relies on persistent homology, a popular method in topological data analysis (TDA).

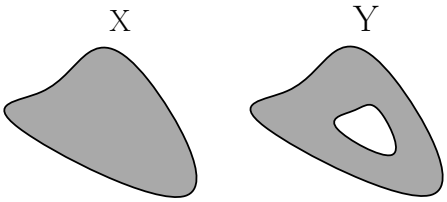


Figure 1.2: Homology captures the difference between these two spaces.

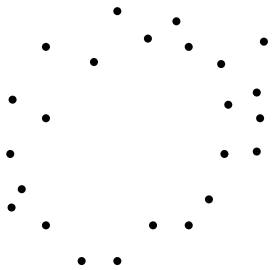


Figure 1.3: A finite metric space, that looks like a circle.

Persistent homology is based on a measure from algebraic topology called homology which provides a way to distinguish between topological spaces by identifying and counting the number of holes, voids, and connected components. For example, in Figure 1.2 both X and Y have each exactly one connected component, but while Y has one hole, X has none. The number of occurrences of each feature is an invariant of this space. Since these numbers do not agree for X and Y, we deduce that X and Y are different spaces. Persistent homology adapts this idea for finite metric spaces.

To see why we can not apply homology directly, consider the finite metric space in Figure 1.3. Here, we as humans see that these points lie along a circle, but how can we capture this with homology? If we apply homology we just get the number of connected components! This is due to the limited structure of such spaces. One way to overcome this issue is to thicken up the points, or mathematically speaking, to cover the points with balls. This gives us more structure. But how large should these covering balls be? If we choose them too small, as in Figure 1.4, we do not gain much. If they are too big, as in Figure 1.7, we cover up the circle.

Persistent homology tackles this problem by considering all possible radii instead of choosing one. By letting the radius vary it is possible to measure how long a feature, such as a connected component or a hole persists. In Figure 1.4 we have 23 connected components and no holes. While the radius of the balls increases, the components merge together until we finally have just one component as in Figure 1.7.

Not all features we detect in this progress are real features. For example, in Figure 1.5 we get two small circles. From an intuitive point of view it is

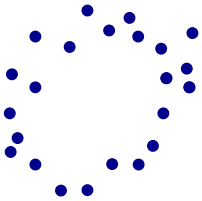


Figure 1.4: 23 connected components

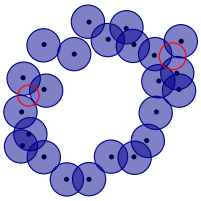


Figure 1.5: Two small circles

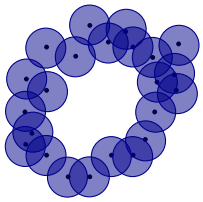


Figure 1.6: One big circle

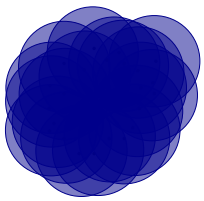


Figure 1.7: No holes visible

Figure 1.8: Different choices of radii on a point cloud.

clear that these two circles are due to noise. Instead we want to label the hole captured in Figure 1.6 as the real hole. To distinguish the importance of a circle we track how long it persists. This means we make note of how big the radii of the balls are, as soon as a new circle appears and how big the radii need to be for it to vanish. We see that the two holes in Figure 1.5 will disappear if we increase the radius just a little bit. This small difference in the two tracked radii indicates that this is not a circle of interest. On the other hand, the hole in Figure 1.6 will survive for much longer, indicating that this is a real hole.

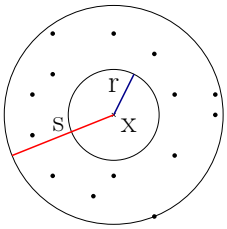


Figure 1.9: Annulus with outer radius s and inner radius r.

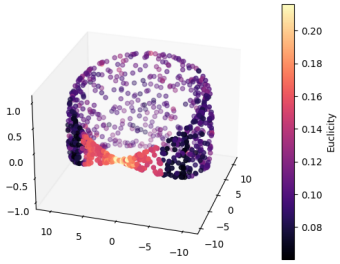


Figure 1.10: Euclidicity scores of the pinched torus.

The TARDIS algorithm computes persistent homology of some point x for various annuli centered at x. An example annulus is illustrated in Figure 1.9. The average of the persistent homology computed for these annuli is compared to persistent homology of an annulus on some manifold. This means that a feature that only appears on some of the annuli is less important than a feature that appears on every annulus. For example, consider again the pinched torus. A point close to the pinched point p has different persistent homology on an annulus with a small enough inner and outer radius than on an annulus that includes p. This means that by averaging over persistent homology computed for annuli of different sizes provides a way to see how

close some point is to a singularity. Thus, by varying the outer radius s and the inner radius r we characterize the topological features not only by the time they appear or vanish but also for which choices of radii they persist. The key observation here is that neighborhoods of regular points differ from neighborhoods of points close to a singularity. The more these two neighborhoods deviate the bigger is the Euclidicity score for this point, indicating a singular region. For example, consider an annulus around the singular point of the pinched torus. Each annulus has two connected components. An annulus in $\mathbb{R}^2$ has one connected component. This difference is captured by a high Euclidicity score for the singular point. This example highlights why a correct dimension estimate is crucial; if we compare the annuli of a regular point on the pinched torus to annuli on $\mathbb{R}^3$ the Euclidicity score is high as well, since these two annulus have different first homology.

In a series of experiments we can observe the effect of different choices of parameters on the TARDIS algorithm. We discuss a data driven way to find a minimal inner and a maximal outer radius for this annuli and compare this method to taking fixed values for the radii. The experiments show that Euclidicity is expressive and does highlight the singularities one sees in Figure 1.10. By comparing different numbers of annuli per point the experiments show that fewer annuli make the computation faster while the singularities are still easy to detect.

Chapter 2

Simplicial Complexes And Homology

In this chapter we define simplicial complexes, a special type of topological spaces. To classify simplicial complexes we define homology. Informally speaking homology counts the number of n-dimensional holes. Finally with local homology we will define neighborhoods in the context of simplicial complexes.

We follow Hatcher's *Algebraic Topology* [Hat02] and *Topological pattern recognition for point cloud data* by Carlsson [Car14].

2.1 Simplicial Complexes

Simplicial complexes are combinatorial objects built by gluing vertices, edges, triangles, tetrahedrons etc. along common edges. We start by defining simplices as convex hulls of points in general positions.

Definition 2.1 *Let $S = \{x_0, x_1, \ldots, x_n\} \subset \mathbb{R}^k$. We say that S is in **general position** if S is not contained in any affine hyperplane of $\mathbb{R}^k$ of dimension less than n.*

Definition 2.2 *If $S = \{x_0, x_1, \ldots, x_n\} \subset \mathbb{R}^k$ is in general position, then the **simplex spanned by S** is the convex hull $\sigma = \sigma(S)$ of S in $\mathbb{R}^k$. The dimension of the simplex is n so we can also refer to it as the **n-simplex**. For $k \leq n$ a **k-face** (or, simply a **face**) of σ is a k-simplex that is the convex hull of a nonempty subset of S.*

We call a 0-simplex a point, a 1-simplex an edge and a 2-simplex a triangle.

Definition 2.3 *A **geometric simplicial complex** is a finite collection X of simplices in a Euclidean space, that satisfies the following two restrictions.*

- *X contains every face of each simplex in X.*
- *For any two simplices $\sigma, \tau \in X$, their intersection $\sigma \cap \tau$ is either emtpy or a face of both σ and τ.*

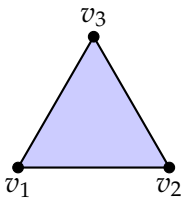


Figure 2.1: An example of a simplicial complex.

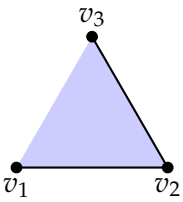


Figure 2.2: A collection of simplices that does not form a simplex.

Remark 2.4 *We often write $\sigma = x_0x_1 \cdots x_n$ or $\sigma = [x_0, x_1, \cdots x_n]$ for the simplex spanned by the vertices $\{x_0, x_1, \cdots, x_n\}$.*

It is also possible to regard simplicial complexes as a purely combinatorial object.

Definition 2.5 *An **abstract simplicial complex** X is the pair $X = (V(X), \Sigma(X))$. $V(X)$ denotes a finite set, the vertices of X. $\Sigma(X)$ is a subset of the family of all non-empty subsets called the simplices of $V(X)$, satisfying that if*

$$\sigma \in \Sigma(X) \text{ and } \emptyset \neq \tau \subset \sigma \Rightarrow \tau \in \Sigma(X).$$

To every geometric simplicial complex we can assign an abstract simlipicial complex and vice versa. One can think of this as two different points of view on the same object. This gives the freedom to consider a simplicial complex as a combinatorial or as a geometric object depending on the context.

Definition 2.6 *A geometric simplicial complex X in $\mathbb{R}^d$ is called a **geometric realization** of an abstract simplicial complex X' if and only if there is an embedding $e\colon V(X') \to \mathbb{R}^d$ that takes every k-simplex $[x_0, x_1, \cdots, x_k] \in X'$ to a k-simplex in X that is the convex hull of $e(x_0), e(x_1), \cdots, e(x_k)$.*

Remark 2.7 *To get an abstract simplicial X' complex from a geometric simplicial complex X simply make a list $V(X')$ of all the vertices in X, and a list $\Sigma(X')$ of all the faces in X.*

Example 2.8 *Let $X = (V(X), \Sigma(X))$ be given by $V(X) = \{v_1, v_2, v_3\}$ and $\Sigma(X) = \{v_1, v_2, v_3, v_1v_2, v_2v_3, v_1v_3, v_1v_2v_3\}$. A geometric realization of this simplex is drawn in Figure 2.1. Note that this is also an example of a 2-simplex. v_1v_2 is an example of a face of $v_1v_2v_3$. Now consider $X' = (V(X'), \Sigma(X'))$ given by $V(X') = \{v_1, v_2, v_3\}$ and $\Sigma(X') = \{v_1, v_2, v_3, v_1v_2, v_2v_3, v_1v_2v_3\}$. This is not a simplicial complex see Figure 2.2.*

In order to describe a simplicial complex locally we need a notion of a neighborhood for simplices.

Definition 2.9 *For a simplical complex $X = (V(X), \Sigma(X))$ and $x \in V(X)$, we define the **link of x** $\mathrm{Lk}(x)$ to be the union of simplices $\sigma \in \Sigma(X)$ with $x \notin \sigma$ and $x \cup \sigma \in \Sigma(X)$. The **open star of x** $\mathrm{st}(x)$ as the collection of of simplices $\in \Sigma(X)$ that contain x. We call the closure of $\mathrm{st}(x)$ the **star of x** and write $\mathrm{St}(x)$.*

Intuitively we can think of the open star as an open neighborhood of x. The open star is not necessarily a simplicial complex since it is not downward closed. Observe that the star is contractible. Furthermore, we could also write $\mathrm{St}(x) = \mathrm{st}(x) \cup \mathrm{Lk}(x)$, meaning we could think of the link as the boundary of the open star.

Example 2.10 *The open star of the yellow vertex in Figure 2.3 is marked in blue. We see that this is not a simplicial complex, since not all the faces of the blue triangles are included. The link of this vertex is drawn in Figure 2.4. This are exactly the faces that are needed to make the open star a simplicial complex.*

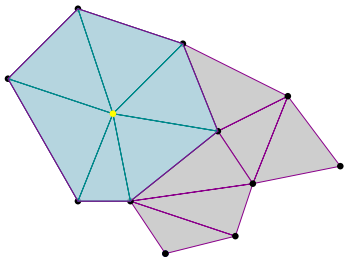


Figure 2.3: Open star of the yellow vertex, (including the yellow vertex).

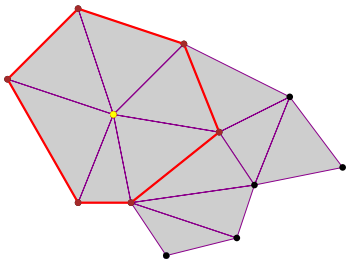


Figure 2.4: Link of the yellow vertex

2.2 Simplicial Homology

Simplicial homology allows to study and classify topological spaces based on their structural features such as loops and voids. To do so, we consider n-simplices and investigate whether they form cycles. The cycles that are not a boundary are features of the space.

Definition 2.11 *Let k be a field and S a finite set. We denote the k-span of S by $V_k(S)$ and call it the **free k-vector space on the set S**.*

Definition 2.12 *Let k be a field and X a simplicial complex. We denote by $C_n(X)$ the free k-vector space on the set of n-simplices. We call $\xi \in C_n(X)$ **n-chain**.*

A basis for $C_n(X)$ is given by the set of all n-simplices. Each chain can be written as a formal sum of n-simplices $\xi = \sum a_k \sigma_k$, where $a_k \in k$ for some field k. Addition and scalar multiplication is defined in the obvious way.

From now on we set $k = \mathbb{Z}_2$ unless stated otherwise. This makes some computations easier, in particular, this allows to forget signs.

Example 2.13 *In Figure 2.1 $C_1(X)$ is given by* $\text{span}\{v_1v_2, v_2v_3, v_1v_3\}$. *$C_0(X)$ is given by* $\text{span}\{v_1, v_2, v_3\}$. *An example of a 1-chain is $v_1v_2 + v_2v_3$.*

Definition 2.14 *Given an n-simplex σ, we define* ***the boundary $\partial_n\sigma$*** *as the union of its (n-1)-dimensional faces, i.e. if $\sigma = [x_0, x_1, \cdots, x_n]$ then*

$$\partial_n\sigma = \sum_{k=0}^{n}[x_0, x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_n].$$

For an n-chain $\xi = \sum a_k\sigma_k$, the boundary is the sum of the boundaries of its simplices, i.e. $\partial_n\xi = \sum a_k(\partial_n\sigma_k)$. The boundary takes a n-chain to a (n-1)-chain, we see that $\partial_n(\xi + \xi') = \partial_n(\xi) + \partial_n(\xi')$ and thus $\partial_n \colon C_n(X) \to C_{n-1}(X)$ is a homomorphism, called the ***boundary map****. We can represent this map with a matrix, called the* ***boundary matrix****. This matrix will be useful to compute persistence.*

Example 2.15 *Consider again Example 2.13, here the boundary matrix is given by*

$$\partial_1 = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \end{array} \begin{array}{c} \begin{array}{ccc} v_1v_2 & v_1v_3 & v_2v_3 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}$$

Definition 2.16 *We call the sequence*

$$\ldots \longrightarrow C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} \ldots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0$$

a ***chain complex****. We define the* ***n-th homology group*** *of this chain complex to be the quotient group $H_n = \ker\partial_n / \operatorname{Im}\partial_{n+1}$. Elements of $\ker\partial_n$ are called* ***cycles*** *and elements of $\operatorname{Im}\partial_{n+1}$ are* ***boundaries****. The rank of H_n is called the* ***n-th Betti number*** *of X.*

2.2.1 Relative Homology Groups

Sometimes it is useful to disregard a certain amount of structure. To do so we consider quotient groups.

Definition 2.17 *Given a simplicial complex X and a subcomplex $A \subset X$, we define the* ***relative homology group*** *$C_n(X, A)$ as the quotient $C_n(X)/C_n(A)$.*

Since the boundary map $\partial : C_n(X) \to C_{n-1}(X)$ takes $C_n(A)$ to $C_{n-1}(A)$ it induces a quotient boundary map and we get a relative quotient chain:

$$\ldots \longrightarrow C_{n+1}(X, A) \xrightarrow{\partial_{n+1}} \ldots \xrightarrow{\partial_2} C_1(X, A) \xrightarrow{\partial_1} C_0(X, A) \xrightarrow{\partial_0} 0$$

This chain complex allows us to define relative homology groups.

Definition 2.18 *The **relative homology groups** $H_n(X, A) = \ker \partial / Im\ \partial$, where ∂ is the boundary map of the chain complex above.*

2.3 Singular Homology

Simplicial homology is quite intuitive. Unfortunately, it is not defined for general topological spaces. An extension is called singular homology, a more flexible and robust framework. Here, the building blocks are singular simplices which are continuous maps from standard simplices into the space under study. These singular simplices can be of any shape.

For an introduction into singular homology we refer to Hatcher [Hat02, Section 2.1].

One important result from singular homology is the excision theorem.

Theorem 2.19 *For subspaces $A, B \subset X$ whose interior cover X, the inclusions $(B, A \cap B) \hookrightarrow (X, A)$ induces isomorphisms $H_n(B, A \cap B) \to H_n(X, A)$ for all n.*

The proof of this theorem can be found in Hatcher [Hat02, Theorem 2.20].

Chapter 3

Persistent Homology

In this chapter we introduce Vietoris-Rips and Čech complexes as well as filtrations in general. The main result of this chapter is the decomposition theorem which classifies all finitely generated persistent vector spaces.
This chapter is mainly based on Hatcher's *Algebraic Topology* [Hat02], Carlsson's *Topological pattern recognition for point cloud data* [Car14] and the book *Computational Topology for Data Analysis* [DW22].

3.1 Point Cloud Filtrations

To apply homology to a point cloud we need to transform the data. One option is to cover the points with balls. To get a simplicial complex we regard the points as vertices and connect them whenever their covering balls intersect. The nerve theorem ensures that this construction maintains the structure of the space.

Definition 3.1 *Given a finite collection of sets $\mathcal{U} = \{U_a\}_{a\in A}$, we define the **nerve** of $\mathcal{U}$ to be the simplical complex $\mathcal{N}(\mathcal{U})$ whose vertex set is the index set A. $\{a_0, a_1, \cdots, a_k\} \subset A$ spans a k-simplex in $\mathcal{N}(\mathcal{U})$ if and only if $U_{a_0} \cap U_{a_1} \cap \cdots \cap U_{a_k} \neq \emptyset$.*

Figure 3.1: Covering $\mathcal{U}$ of some space.

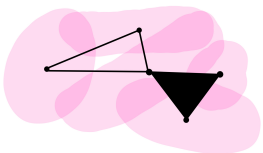


Figure 3.2: The nerve $\mathcal{N}(\mathcal{U})$ of this covering.

Example 3.2 *Consider the covering $\mathcal{U}$ in Figure 3.1. The nerve $\mathcal{N}(\mathcal{U})$ in Figure 3.2 is obtained by adding a vertex on each set $U_i \in \mathcal{U}$ and an edge between two vertices*

whenever the corresponding sets overlap, and add a triangle whenever three sets intersect.

The nerve theorem ensures that the simplicial complex captures all the information of the covering. This is a standard result which is beyond the scope of this thesis, but the interested reader is referred to [DW22, Theorem 2.1].

Theorem 3.3 (Nerve Theorem) *Given a finite (open or closed) cover $\mathcal{U}$ of a metric space X, the underlying space $|\mathcal{N}(\mathcal{U})|$ is homotopy equivalent to X if every non-empty intersection $\bigcap_{i\in I} U_{a_i}$ of cover elements is contractible.*

Applying homology to a point cloud directly yields trivial n-th homology groups for $n > 0$. To give more structure to the point cloud we 'thicken up' the points, and observe the union of this thickened points.

Definition 3.4 *For a metric space (X, d_X) and $P = \{p_1, \cdots, p_k\} \subset X$ a finite subset, the **Čech complex** at time t, $\check{C}(P, t)$, is a simplicial complex given by the nerve of the set $\{B(p_i, t)\}_i$, where $B(p_i, t) = \{x \in X \mid d(p_i, x) \leq t\}$.*

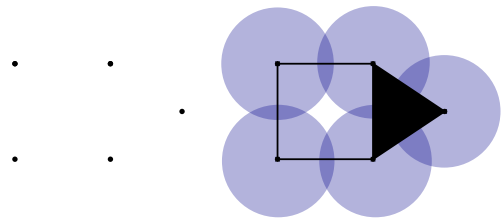


Figure 3.3: Left: point cloud. Right: Čech complex at the time t= 1.2.

Example 3.5 *In Figure 3.3 we have a point cloud on the left and the Čech complex on this point cloud at time t= 1.2 on the right. On the Čech complex we have a non-trivial 1st homology group, giving us a global understanding of the data points.*

Remark 3.6 *Since balls are convex, every intersection of a collection of balls is convex, and therefore contractible. This implies that the nerve theorem always holds for Čech complexes.*

Instead of computing intersections of balls the Vietoris-Rips complex only computes the pairwise distances of the points, making it more computationally feasible.

Definition 3.7 *Given a finite metric space (X, d_X), the **Vietoris-Rips complex** at time t, $VR(X, t)$ is a simplical complex, where the n-simplex $(x_0, x_1, \cdots, x_n) \in VR(X, t) \Leftrightarrow \forall\, 1 \leq i, j \leq n \colon d_X(x_i, x_j) \leq 2t$.*

The Čech complex and the Vietoris-Rips complex are closely connected as the next proposition shows.

Proposition 3.8 *Let P be a finite subset of a metric space (X,d). Then,*

$$\check{C}(P,t) \subset VR(P,t) \subset \check{C}(P,2t).$$

Proof *First inclusion:* If there is a point x in the intersection $\bigcap_{i=1}^{k} B(p_i,t)$, we have that for every pair (i,j) $1 \leq i,j \leq k$ the distances $d(p_i,p_j)$ are at most $2t$. It follows that for every simplex $[p_1,\cdots,p_k] \in \check{C}(P,t)$ is also in $VR(P,t)$.
Second inclusion: Let $[p_1,\cdots,p_k] \in VR(P,t)$. By definition of the VR complex we have that for every $1 \leq i \leq k$ $d(p_i,p_1) \leq 2t$. Then $\emptyset \neq p_1 \in \bigcap_{i=1}^{k} B(p_i,2t)$ and we conclude that $[p_1,\cdots,p_k]$ is a simplex in $\check{C}(P,2t)$. □

This inclusions together with Theorem 3.3 gives us a theoretical guarantee for the Vietoris-Rips complex as well.
We can use these complexes to approximate the underlying space. The question is what parameter to choose. Instead of choosing one t we simply consider all possible choices of $t \in \mathbb{R}_+$, this yields a filtration.

Definition 3.9 *Given a topological space X, a **filtration** is a nested sequence of subspaces*

$$\mathcal{F}\colon \emptyset = X_0 \subset X_1 \subset X_2 \subset \cdots \subset X_n = X.$$

Definition 3.10 *If X is a simplicial complex in the above definition, we call the filtration $\mathcal{F} = \mathcal{F}(K)$ a **simplicial filtration**, it consist of a nested sequence of closed subcomplexes*

$$\mathcal{F}\colon \emptyset = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_n = K.$$

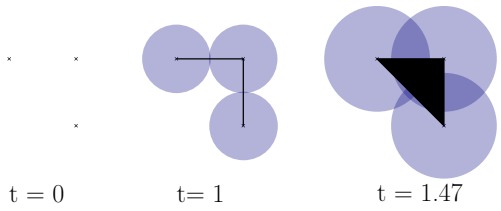


Figure 3.4: Simplicial filtration induced by the Čech complex.

Example 3.11 *The Čech and the Vietoris-Rips complexes induce simplicial filtrations; by letting the time t vary from 0 to ∞, we get a filtration where only at finitly many times the complex changes. For any point cloud P and different times $t_1 < t_2$ we have by definition that $\mathcal{F}(P,t_1) \subset \mathcal{F}(P,t_2)$, where $\mathcal{F}$ is either the Čech or the Vietoris-Rips complex. An example of a Čech filtration is illustrated in Figure 3.4, there are three different Čech complexes for different times t. Here we have that if $t \geq 1.46$ the complex does not change anymore. In Figure 3.5 there are the first four steps of the filtration induced by the Vietoris-Rips complex. We see in both examples*

that there are only finitely many changes of the simplicial complex since there are finitely many points in a point cloud.

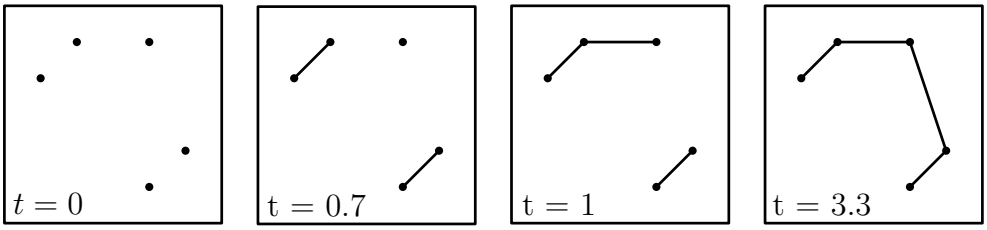


Figure 3.5: 4 steps of a Vietoris-Rips filtration.

Not every simplicial filtration is induced by the Čech or the Vietoris-Rips complex as the next example shows.

Example 3.12 *An example of a simplicial filtration with 4 steps is given in Figure 3.6. We have that* $K_1 = \{a, b\}$, $K_2 = K_1 \cup \{c, ab\}$ *and so on. We see from the pictures that* $K_i \subset K_{i+1}$ *for* $1 \leq i \leq 3$. *This is an example for a simplicial filtration that is not induced by the Čech complex or by the Vietoris-Rips complex.*

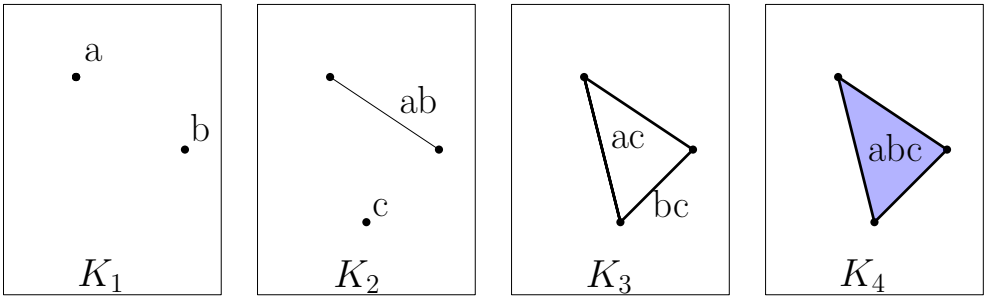


Figure 3.6: Simplicial filtration with 4 steps.

Definition 3.13 *Let* X *be any set and* $\rho\colon X \to \mathbb{R}_+$. *We call such a pair* (X, ρ) *an* $\mathbb{R}_+$***-filtered set***.

An $\mathbb{R}_+$-filtered set also induces a filtration as the following example shows.

Example 3.14 *Consider the simplicial complex* $X = (V(X), \Sigma(X))$ *with* $\Sigma(X) = \{a, b, c, ab, bc, ac, abc\}$. *If we assign to each face the time it appears in the filtration shown in Figure 3.6, we get an* $\mathbb{R}_+$*-filtered set. For example,* $\rho(a) = 1$, $\rho(ab) = 2$, $\rho(c) = 2$, $\rho(bc) = 3$, $\rho(abc) = 4$.
In K_3 *we have a circle which is not a boundary, the circle becomes a boundary in* K_4, *the lifespan of this circle is time of death-time of birth = 1 step.*

3.2 Persistent Vector Spaces

In this section we introduce persistent vector spaces. We state and prove the decomposition theorem, which classifies all finitely presented persistent vector spaces up to isomorphisms.

Definition 3.15 *Let $r, r' \in R$. By a **persistence vector space (PVS)** over a field k we refer to a family of k-vector spaces $\{V_r\}_{r\in\mathbb{R}}$, together with linear maps $L_V(r,r')\colon V_r \to V_{r'}$ where $r \leq r'$ and $L_V(r',r'') \circ L_V(r,r') = L_V(r,r'')$ where $r \leq r' \leq r''$. We define sub-PVS and quotient spaces of PVS in the obvious way.*

Remark 3.16 *Sometimes persistent vector spaces are referred to as persistent modules.*

Definition 3.17 *A linear transformation f between two PVS $\{V_r\}_{r\in\mathbb{R}}$ and $\{W_r\}_{r\in\mathbb{R}}$, is a family of transformations $f_r\colon V_r \to W_r$ which make the following diagram commute:*

$$\begin{array}{ccc} V_r & \xrightarrow{L_V(r,r')} & V_{r'} \\ \downarrow{\scriptstyle f_r} & & \downarrow{\scriptstyle f_{r'}} \\ W_r & \xrightarrow{L_W(r,r')} & W_{r'} \end{array}$$

Example 3.18 *Given an $\mathbb{R}_+$-filtered set (X,ρ), we get an induced PVS $\{W(X,\rho)_r\}$ where*

$$\{W(X,\rho)_r\} = span\{x \in X\colon \rho(x) \leq r\} \subset V_k(X).$$

Definition 3.19 *We call a PVS **free** if it is isomorphic to one of the form $\{V(X,\rho)_r\}$ for some $\mathbb{R}_+$-filtered set (X,ρ), it is **finitely generated** if X is finite.*

Remark 3.20 *Any linear combination $\sum_x a_x x \in V_k(X)$ lies in $W(X,\rho)_r$ iff $a_x = 0\,\forall x$ with $\rho(x) > r$.*

Definition 3.21 *A PVS is **finitely presented** if it is isomorphic to a PVS of the form $\{W_r\}/im(f)$, where $f\colon \{V_r\} \to \{W_r\}$ is a linear transformation between finitly generated free PVS.*

Example 3.22 *For $a, b \in R$ with $a \leq b$ we define the **interval PVS** as the family of k-vector spaces*

$$P(a,b)_r = \begin{cases} k, & r \in [a,b) \\ 0, & r \notin [a,b) \end{cases}$$

together with linear maps $L(r,r') = id_k$ for $r, r' \in [a,b)$ and zero maps otherwise. $P(a,b)$ is an example of a finitely presented PVS.

Observe that for a finitely generated free PVS $\{V(X,\rho)_r\}$ we have that $V_k(X,\rho)_r = V_k(X)$ for a sufficiently large r. Thus any linear map $f\colon \{V(Y,\sigma)_r\} \to \{V(X,\rho)_r\}$ induces a linear map $f_\infty\colon V_k(Y) \to V_k(X)$. Using the basis $\{x\}_{x\in X}$ of $V_k(X)$ and $\{y\}_{y\in Y}$ of $V_k(Y)$ we can determine the (X,Y)-matrix $A(f) = [a_{xy}]$ with $a_{xy} \in k$.

Example 3.23 *Consider again Figure 3.6. In Example 3.14 we established that this filtration is induced by the $\mathbb{R}_+$-filtered set (X,ρ). Together with inclusion maps we see that this is an example of a free PVS.*

We establish some properties of this matrix, which will come by handy later on.

Proposition 3.24 *The (X,Y)-matrix $A(f)$ has the property that $[a_{xy}] = 0$ whenever $\rho(x) > \sigma(y)$. Conversely, every (ρ,σ)-adapted (X,Y)-matrix A uniquely determines a linear transformation of persistence vector spaces*

$$f_A\colon \{V_k(Y,\sigma)_r\} \to \{V_k(X,\rho)_r\}.$$

The correspondences $f \to A(f)$ and $A \to f_A$ are inverses to each other.

Proof The basis vector y lies in $V_k(Y,\sigma)_{\sigma(y)}$. By definition we have

$$f(y) = \textstyle\sum_{x\in X} a_{xy}x.$$

From Remark 3.20 we know that $\sum_{x\in X} a_{xy}x \in V_k(X,\rho)_{\sigma(y)}$ iff $\forall\, x, y$ with $\rho(x) > \sigma(y)$ we have $a_{xy} = 0$. □

Definition 3.25 *Let (X,ρ) and (Y,σ) be two $\mathbb{R}_+$-filtered sets. We call the (X,Y)-matrix with the property from Proposition 3.24 (ρ,σ)-**adapted**.*

Corollary 3.26 *Let (X,ρ) and (Y,σ) be two $\mathbb{R}_+$-filtered sets and $A = [a_{xy}]$ be a (ρ,σ)-adapted matrix. Then A determines a PVS via the correspondence*

$$A \xrightarrow{\theta} V_k(X,\rho)/im(f_A).$$

$\theta(A)$ is a finitely presented PVS. Conversely, any finitely presented PVS is isomorphic to one of the form $\theta(A)$ for some (ρ,σ)-adapted matrix A.

Remark 3.27 *We write $\theta(A)$ to refer to the quotient space $V_k(X,\rho)/im(f_A)$.*

Definition 3.28 *Let (X,ρ) be an $\mathbb{R}_+$-filtered set. We identify the group of automorphism on $V_k(X,\rho)$ with the group of all invertible (ρ,ρ)-adapted (X,X)-matrices.*

Proposition 3.29 *Let (X,ρ) and (Y,σ) be two $\mathbb{R}_+$-filtered sets, let A be a (ρ,σ)-adapted (X,Y)-matrix. Let B be a (ρ,ρ)-adapted (X,X)-matrix and C a (σ,σ)-adapted (Y,Y)-matrix. Then BAC is also (ρ,σ)-adapted and $\theta(A) \cong \theta(BAC)$.*

Proof Let $A = [a_{xy}]$ and $B = [b_{\tilde{x},x}]$. We first show that the matrix $BA = [d_{xy}]$ is (ρ, σ)-adapted. Note that

$$d_{xy} = \sum_{x \in X} b_{\tilde{x},x} a_{xy} = \sum_{\substack{x \in X \\ \rho(x) \leq \rho(\tilde{x}) \leq \sigma(y)}} b_{\tilde{x},x} a_{x,y}.$$

Where the second equality follows since B, A are adapted matrices and thus $\rho(\tilde{x}) > \rho(x)$ implies that $[b_{\tilde{x},x}] = 0$ and $\rho(x) > \sigma(y)$ implies that $A = [a_{xy}] = 0$. We see that for $\rho(x) > \sigma(y)$ we have $d_{xy} = 0$, thus BA is (ρ, σ)-adapted. Analogously we see that BAC is (ρ, σ)-adapted as well.
The isomorphism follows directly from the linearity of f_{BAC} and since the equivalence classes $[x]$ are preserved. □

The matrices B and C in the above proposition will play an important part in the decomposition theorem and in the computation algorithm. They preserve the (ρ, σ)-adaptedness in a given matrix A.

Definition 3.30 *Let (X, ρ) and (Y, σ) be two $\mathbb{R}_+$-filtered sets and A be a (ρ, σ)-adapted (X, Y)-matrix. We denote the rows of A by $r(x)$ and the columns by $c(y)$. We define an **adapted row operation** to be a an operation that adds a multiple of r(x) to $r(\tilde{x})$ where $\rho(x) \geq \rho(\tilde{x})$. Similarly, we define an **adapted column operation** as an operation that adds a multiple of $c(y)$ to $c(\tilde{y})$ where $\sigma(y) \leq \sigma(\tilde{y})$.*

Before turning our attention to the main result of this section we introduce a special type of PVS; the interval PVS. This space is easy to visualize as we'll discuss in the next section.

Example 3.31 *Let $a, b \in \mathbb{R}_+$ and $a < b$. Let $X = \{x\}$ together with the filtration $\rho(x) = a$ and $Y = \{y\}$ together with the filtration $\sigma(y) = b$. (X, ρ) and (Y, σ) are $\mathbb{R}_+$-filtered sets. Consider the matrix $A = (1)$ which maps y to x. Note that A is (ρ, σ)-adapted matrix, since we have $a < b$ and therefore y appears after x is already present. $V_k(X)_r$ is defined by:*

$$V_k(X)_r = \begin{cases} \emptyset & \text{if } r < a \\ k & \text{if } r \geq a \end{cases}$$

The image of the map f_A induced by A is given by:

$$im(f_A)_r = \begin{cases} 0 & \text{if } r < b \\ k & \text{if } r \geq b \end{cases}$$

We combine this together and get that:

$$\theta(A)_r = \begin{cases} \emptyset & \text{if } r \notin [a,b) \\ k & \text{if } r \in [a,b) \end{cases}$$

Thus we can conclude that $\theta(A) \cong P(a,b)$.

We are now in a position to state the most important result from this section.

Theorem 3.32 (Decomposition Theorem) *Every finitely presented PVS over some field k is isomorphic to a finite direct sum of the form*

$$P(a_1,b_1) \oplus P(a_2,b_2) \oplus \cdots \oplus P(a_n,b_n)$$

for some $a_i \in [0,\infty)$, $b_i \in [0,\infty]$ *and* $a_i < b_i$ *for all i. Furthermore this decomposition is unique up to isomorphisms.*

Proof *Uniqueness:*
Let $\{V_r\} \cong \bigoplus_{i\in I} P(a_i,b_i)$ and $\{V_r\} \cong \bigoplus_{j\in J} P(c_j,d_j)$ be two decompositions where I and J are finite sets. Let $a_{\min}$ and $c_{\min}$ denote the smallest values of a_i and c_j respectively. We can characterize $a_{\min}$ as $\min\{r \in \mathbb{R} \colon V_r \neq 0\}$, from this we conclude that $a_{\min} = c_{\min}$. Set $b_{\min} := \min\{b_i \colon a_i = a_{\min}\}$ and analogously $d_{\min} := \min\{d_i \colon c_i = c_{\min}\}$. $b_{\min}$ and $d_{\min}$ are characterized by $\min\{r' \colon \ker(L(r,r')) \neq 0\}$, which implies $b_{\min} = d_{\min}$.
$P(a_{\min},b_{\min}) = P(c_{\min},d_{\min})$ appears in both decompositions. For each decomposition, consider the sum of all the occurrences of $P(a_{\min},b_{\min})$, these are both sub-PVS of $\{V_r\}$. This sub-PVS are isomorphic to the kernel $\{W_r\}$ of the map.

$$\operatorname{im}(L(a_{min},r)) \xrightarrow{L(r,b_{min})|\operatorname{im}(L(a_{min},r))} V_{bmin}.$$

This implies that the number of summands of the form $P(a_{min},b_{min})$ in the two decompositions is the same. Set $I' = \{i \in I \colon a_i = a_{min}\}$ and $J' = \{j \in J \colon c_j = c_{min}\}$. If we form the quotients, and get the decomposition

$$\{\{V_r\}/\{W_r\}\} \cong \bigoplus_{i\in I\setminus I'} P(a_i,b_i) \text{ and } \{\{V_r\}/\{W_r\}\} \cong \bigoplus_{j\in J\setminus J'} P(c_i,d_i).$$

Repeating this process on the obtained quotient space gives us the desired result.

Existence:
Consider a (ρ,σ)-adapted (X,Y)-matrix A where every row and every column has at most one non-zero element which is equal to 1. We denote these non-zero pairs by $\{(x_1,y_1),\cdots,(x_n,y_n)\}$. Then,

$$\theta(A) = \bigoplus_{x\in X} V_k(x,\rho)/im(f_A) \cong$$
$$\bigoplus_{i\in\{1,\cdots,n\}} V_k(x_i,\rho)/im(f_A) \oplus \bigoplus_{x\in X\setminus\{x_1,\cdots,x_n\}} V_k(x,\rho)/im(f_A) \cong$$
$$\bigoplus_{i\in\{1,\cdots,n\}} P(\rho(x_i),\sigma(y_i)) \oplus \bigoplus_{x\in X\setminus\{x_1,\cdots,x_n\}} P(\rho(x),\infty).$$

Where the last equation follows since $V_k(x_i,\rho)/im(f_A)$ only depends on the image of y_i. For a general matrix A with rows and/or columns with more than one non-zero element, we'll show that there are matrices B and C as in Proposition 3.29 such that BAC has the property, that there is no more than one non-zero element, which is equal to 1 in each row and each column. If we find such matrices we are done by Proposition 3.29.
Finding such matrices is equivalent to transforming A by adapted row and adapted column operations. First we take a $y \in Y$ which minimizes $\sigma(y)$ and with $c(y) \neq 0$. Next, find a x which maximizes $\rho(x)$ and for which $a_{xy} \neq 0$. By the way we chose x, we can add multiples of $r(x)$ to all the other rows in $c(y)$ so we have just one non-zero element $[a_{xy}]$ left. The way we choose y allows us to add multiples of $c(y)$ to zero out all the other entries in $r(X)$, except for $[a_{xy}]$. By multiplying r(x) with $1/a_{xy}$ we create a row and a column where there is exactly one non-zero element equal to 1. Observe that we only used adapted operations. We denote by ρ', σ' the restriction of ρ (or σ) to $X\setminus\{x\}$ and $Y\setminus\{y\}$ respectively. Now delete r(x) and c(y) and repeat the whole process inductively with the (ρ',σ')-adapted $(X\setminus\{x\}, Y\setminus\{y\})$-matrix. Each operation on this smaller matrix can be interpreted as an operation on the original matrix. Since X and Y are finite, we eventually get a matrix A with at most one non-zero entry in each row and in each column. □

3.3 Persistent Homology

Due to the decomposition theorem we have a standardized way of representing a finitely presented PVS with the so-called barcodes or persistent diagrams.

Definition 3.33 *Let $\{V_k\} \cong \bigoplus_{i\in I} P(a_i,b_i)$ be the interval decomposition of dimension p of a given PVS $\{V_k\}$. The **barcode** of $\{V_k\}$ is the multiset of intervals $[a_i,b_i) \subset (\mathbb{R}\cup\infty)$ for $i \in \{1,\cdots,n\}$.*

Definition 3.34 *Let $\{V_k\} \cong \bigoplus_{i\in I} P(a_i,b_i)$ be the interval decomposition of dimension p of a given PVS $\{V_k\}$. The collection of points $\{(a_i,b_i)\} \subset (\mathbb{R}\cup\infty)^2$ with proper multiplicity as well as the points on the diagonal Δ: $\{(x,x)\} \subset (\mathbb{R}\cup\infty)^2$ with infinite multiplicity constitute the **persistence diagram** $Dgm_p(\{V_k\})$ of the PVS $\{V_k\}$.*

Remark 3.35 *Both the barcode as well as the persistence diagram, encode the same information about a given PVS.*

Note that the 0 persistence points on the diagonal are mostly there for technical purposes. Usually we consider points close to the diagonal as noise, whereas points further away represent real features. Typically the x-axis correspond to the birth time whereas the y-axis corresponds to the death time. We observe that there are no points bellow the diagonal, otherwise something would die before its birth.

Example 3.36 *In Figure 3.7 the point cloud seems to be sampled from two cycles. The persistence diagram in Figure 3.8 contains the information about the zeroth and first homology of this point cloud. Since the data is close to each other we quickly have connected components vanishing. We see this in the diagram on the bottom left. For some time we have two connected components until they merge at time 5.5. For the first homology group we have two off-diagonal points. They correspond to the two cycles in the point cloud. The points close to the diagonal are due to noise.*

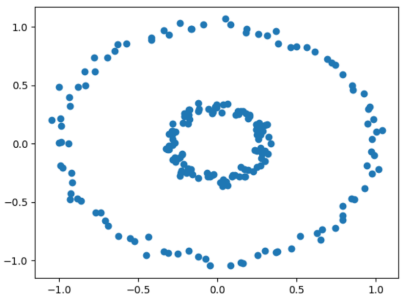


Figure 3.7: Point cloud.

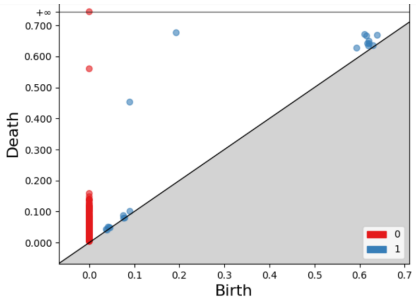


Figure 3.8: Persistence diagram displaying 0th and 1st homology group of the point cloud.

Definition 3.37 *The **p-th persistent homology** $\mathbf{PH_p}$ is the collection of all the points on the persistence diagram Dgm_p.*

3.4 Distance Metric on Persistence Diagrams

Given two data sets with the respective persistent diagrams, we want to have some notion of similarity of persistence diagrams, to evaluate how different some diagrams (and thus the underlying data sets) are. This is important since we want that a small perturbation in the data should still result in similar persistence diagrams.

Definition 3.38 *Let X, Y be finite metric spaces with respective filtrations $\mathcal{F}$ and $\mathcal{G}$. Let $\Pi = \{\pi \colon Dgm_p(\mathcal{F}) \to Dgm_p(\mathcal{G}) \colon \pi \text{ is bijective}\}$ be the set of all matchings between the corresponding persistence diagrams. Then the **bottleneck distance** is given by:*

$$d_b(Dgm_p(\mathcal{F}), Dgm_p(\mathcal{G})) = \inf_{\pi \in \Pi} \sup_{x \in Dgm_p(\mathcal{F})} \|x - \pi(x)\|_\infty.$$

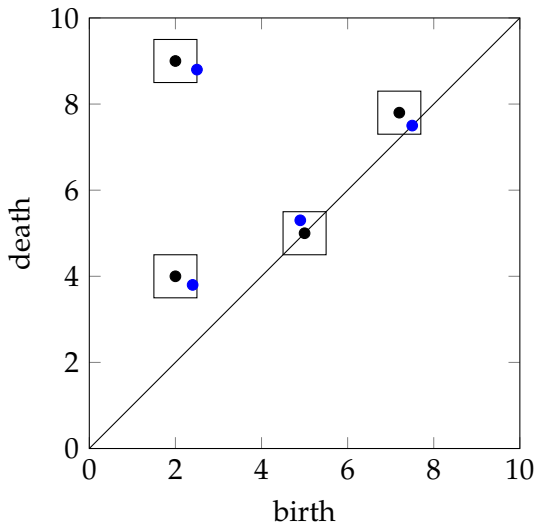


Figure 3.9: The black points are from a different filtration than the blue ones. The rectangles correspond to balls in the infinity norm.

The reason we take points on the diagonal with infinite multiplicity is to ensure we always have bijections. Another reason is that we want to find optimal maps between the points. We illustrate the idea of the bottleneck distance in Figure 3.9, here the blue points are from a different filtration than the black points. Note that the rectangular boxes correspond to balls in the infinity norm. This example also demonstrates why the diagonal allows us to get optimal distances; without it the bottleneck distance here would be way bigger: If we can not match points to the diagonal we need to map the blue point with coordinates $(4.9, 5.3)$ to the black point at $(7.2, 7.8)$, but this means that the radius of the balls would become five times larger than they are now!

As a measure of similarity for point clouds that live in the same ambient space we use the Hausdorff distance.

Definition 3.39 *The **Hausdorff distance** of two non-empty subsets $A, B \subset X$ of a metric space is $d_H(A, B) := \inf\{\varepsilon \geq 0 \colon A \subset B_\varepsilon, B \subset A_\varepsilon\}$ where $A_\varepsilon = \bigcup_{a \in A}\{x \in X \colon d(x, a) \leq \varepsilon\}$ is the ε-thickening of A in X.*

One can think of the Hausdorff distance of two point clouds as the minimum (this is indeed a minimum since the sets are compact) ε one needs, such that the balls of radius ε centred at the points of one point cloud cover the other point cloud as well.

Example 3.40 *In Figure 3.10 we have a green space X and a red space Y. Both spaces are totally bounded. To see that the Hausdorff distance between this spaces*

equal to 1 *is we first note that* $X \subset Y_1$ *and* $Y \subset X_1$*, showing* $d_H(X, Y) \leq 1$*. On the other hand, we have for any smaller thickening* $\varepsilon < 1$ *that* Y *is no longer included in* X_ε*.*

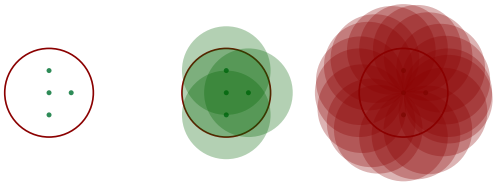


Figure 3.10: Left: Two subsets of $\mathbb{R}^2$. Middle: 1-thickening of the green space Right: 1-thickening of red space.

Chazal et al. showed in *Gromov-Hausdorff Stable Signatures for Shapes using Persistence* [CCSG+09] a more general version of the following stability theorem of persistence diagrams. For this theorem we need that we deal with totally bounded spaces, that is a space which can be covered by finitely many balls where the radius of the balls is chosen arbitrary.

Theorem 3.41 *Let* X, Y *be totally bounded metric spaces. Then* $\forall p \geq 0$ *we have*

$$d_b(Dgm_p(\mathcal{F}(X), Dgm_p(\mathcal{F}(Y)) \leq 2d_H(X, Y),$$

where $\mathcal{F}$ *is either the Vietoris-Rips or the Čech filtration.*

Chapter 4

Manifolds and Stratified Spaces

In this chapter we define manifolds, regular points and singularities. To have a less rigid space that allows for changes of dimension within the space we introduce stratified spaces. This spaces are built with manifolds of various dimensions.
The first section is based on Hatcher's *Algebraic Topology* [Hat02, Chapter 3.3] and on the book *Introduction to Smooth Manifolds* by Lee [Lee13]. Section 2.2 is based on Banagl's *Topological Invariants of Stratified Spaces* [Ban07, Chapter 4] and on Friedman's *Singular Intersection Homology* [Fri20, Chapter 2].

4.1 Manifolds

Definition 4.1 *A point in some topological space that has an open neighborhood homeomorphic to $\mathbb{R}^n$ is called* ***regular****. Points that are not regular are called* ***singularities****. An* ***n-dimensional topological manifold*** *or* ***n-manifold*** *for short, is a Hausdorff space M in which each point is regular for some fixed dimension n.*

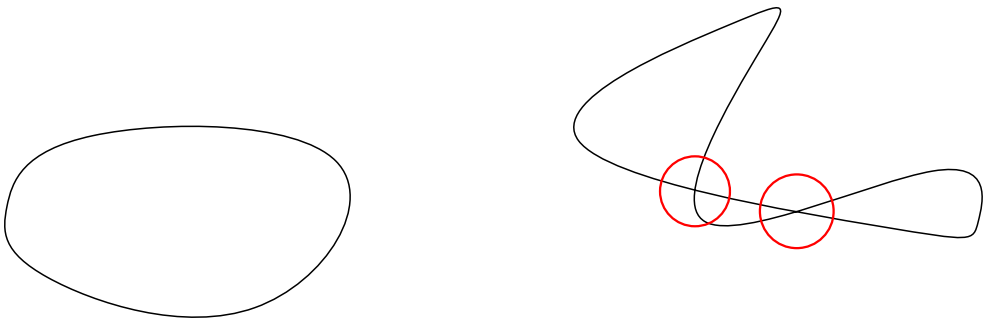


Figure 4.1: Example of a 1-manifold and a nonexample.

Example 4.2 *On the left of Figure 4.1 there is an example of a 1-manifold on the left and a nonexample on the right. The two marked points are singular points since no neighborhood of this points resemble* $\mathbb{R}$.

4.2 Stratified Spaces

Manifolds are very well-behaved spaces, but this also imposes serious restrictions. We want to turn our attention to a more general framework which also allows singularities as well as different intrinsic dimensions in one space.

Definition 4.3 *A* 0*-dimensional* ***(topological) stratified pseudomanifold*** *is a countable set of points with discrete topology.*
For $n > 0$*, an n-dimensional (topological) stratified pseudomanifold* X_n *is an n-dimensional filtered space of closed subsets*

$$X = X_n \supset X_{n-1} \supset X_{n-2} \supset \cdots \supset X_{-1} = \emptyset$$

such that:

1. $\forall i \leq n$ *: Each connected component of* $X_i \setminus X_{i-1}$ *is an i-dimensional manifold.*
2. $X_n \setminus X_{n-1}$ *is dense in* X*.*
3. *Local normal triviality: For each point* $x \in X_i \setminus X_{i-1}$*, there exists*
   - *an open neighborhood* U_x *of* x *in* X*.*
   - *a compact stratified pseudomanifold* L *of dimension* $m := n - i - 1$ *with stratification* $L = L_m \supset L_{m-1} \supset \cdots \supset L_0 \supset L_{-1} = \emptyset$*.*
   - *a homeomorphism* $\Phi \colon U_x \to \mathbb{R}^i \times c^\circ L$*. Here we denote by* c° *the open cone given by* $c^\circ L := L \times (0,1]/L \times \{0\}$*.*
4. *We call* $X_n \setminus X_{n-1}$ *the top stratum. For* $i \neq n$ *we call* $X_i \setminus X_{i-1}$ *a (i-dimensional) stratum. An n-dimensional stratification is called classical if* $X_{n-1} = X_{n-2}$*.*

We refer to L as the link. The link is only interesting for points on $X_i \setminus X_{i-1}$ for $i < n$. If $i = n$ we have that the dimension of the link $m = 2 - 2 - 1 = -1$ and thus we only have $L_{-1} = \emptyset$. One can think of the link as the boundary of a neighborhood of a point which has a lower intrinsic dimension than the space itself.

Example 4.4 *Let* X *be an n-dimensional manifold. Then* $X = X_n$ *and* $X_{n-1} = X_{n-2} \cdots = X_{-1} = \emptyset$*.* $X_n \setminus X_{n-1} = X_n$ *which is trivially dense in* X_n*. Intuitively the link should the empty set since there are only points on* X_n*. By a calculation of the links dimension* $m = 2 - 2 - 1 = -1$ *we verify that the link is indeed the empty set.*

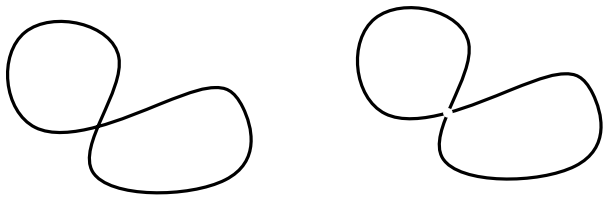


Figure 4.2: Left: Example of a 1-dimensional stratified pseudomanifold. Right: Top stratum.

Example 4.5 *Consider the squiggly 8-figure in Figure 4.2. The two loops on the right form the top stratum, and the removed point is the 0-dimensional stratum. Obviously the two loops $X_1 \setminus X_0$ are dense in the whole figure X_1. Furthermore, each point in $X_1 \setminus X_0$ has a neighborhood homeomorphic to $\mathbb{R}^1$. Here L is the empty set since $m = 1 - 1 - 1 = -1$.*

Figure 4.3: Left: Torus with two meridians highlighted. Right: Pinched torus

Example 4.6 *Given a torus T one obtains a pinched torus $\tilde{T}$ by compressing one meridian to a single point. This object has a croissant like shape and is illustrated in Figure 4.3. The pinched torus is 2-dimensional stratified pseudomanifold. The stratification is given by*

$$X_2 = \textit{pinched torus;}$$
$$X_1 = X_0 = \textit{pinched point.}$$

Clearly $X_2 \setminus X_1$ is dense in X_2. $X_1 \setminus X_0 = \emptyset$, which aligns with our intuition that there is no 1-dimensional manifold in this object.
For the pinched point $p \in X_0 \setminus X_{-1} = X_0$, there exists a neighborhood U_p and a $2 - 0 - 1 = 1$-dimensional compact stratification L. L_1 is given by $S^1 \sqcup S^1$. The open cone $c^\circ S^1$ is homeomorphic to D^2, this can be achieved by projection. The open cone $c^\circ S^1 \sqcup S^1$ is illustrated in Figure 4.4, one can think of two disks D^2 that are glued together at the middle point and afterwards are pulled in opposite directions. We observe that this is exactly the neighborhood of p.
Note that this is an example of a classical stratification.

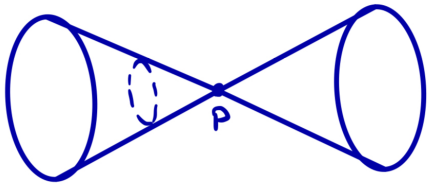


Figure 4.4: The open cone of $S^1 \sqcup S^1$.

4.3 Local Homology

We want to localize homology, so we shift from the usual global point of view to a local viewpoint. This shift will help us in the following chapter to detect singularities and to quantify local dimensions. Furthermore, we introduce the star of a vertex to get a notion of neighborhood, in the context of simplicial complexes.

We adapt the definition of relative homology from Chapter 2 a bit in order to obtain local homology groups.

Definition 4.7 *If we take $A = X\backslash\{x\}$ in the definition of relative homology, we get the* ***local homology group*** $H_n(X, X\backslash\{x\})$.

In what follows we want to motivate the use of the link $\mathrm{Lk}(x)$ as a tool to study the local behaviour of a space. For this we need a long exact reduced sequence, for an introduction see [Hat02][Section 2.1].
If we apply Theorem 2.19 to a topological space X, $x \in X$, with $A = X\backslash\{x\}$ and $B = \mathrm{St}(x)$ we get:

$$H_n(X, X\backslash\{x\}) \cong H_n(\mathrm{St}(x), \mathrm{St}(x)\backslash\{x\}).$$

We consider the long exact reduced sequence for the pair $(\mathrm{St}(x), \mathrm{St}(x)\backslash\{x\})$:

$$\cdots \to \widetilde{H}_n(\mathrm{St}(x)) \to H_n(\mathrm{St}(x), \mathrm{St}(x)\backslash\{x\}) \to \widetilde{H}_{n-1}(\mathrm{St}(x)\backslash\{x\}) \to \widetilde{H}_{n-1}(\mathrm{St}(x)) \to \cdots$$

Since the star $St(x)$ is contractible we get $\forall n\colon \widetilde{H}_n(\mathrm{St}(x)) = 0$, thus we can update the sequence:

$$\cdots \to 0 \to H_n(\mathrm{St}(x), \mathrm{St}(x)\backslash\{x\}) \to \widetilde{H}_{n-1}(\mathrm{St}(x)\backslash\{x\}) \to 0 \to \cdots$$

By exactness of the sequence we have

$$\forall n\colon H_n(\mathrm{St}(x), \mathrm{St}(x)\backslash\{x\}) \cong \widetilde{H}_{n-1}(\mathrm{St}(x)\backslash\{x\}).$$

Observe that $\mathrm{St}(x)\backslash x$ deformation retracts to $\mathrm{Lk}(x)$. Combining everything together we get

$$H_n(X, X\backslash\{x\}) \cong \widetilde{H}_{n-1}(\mathrm{St}(x)\backslash\{x\}) \cong \widetilde{H}_{n-1}(\mathrm{Lk}(x)).$$

We see that $\mathrm{Lk}(x)$ already contains all the information about the homology of $X\backslash x$.

Proposition 4.8 *The dimension of an n-manifold M is characterized by the local homology group:*

$$H_i(M, M \setminus \{x\}) \neq 0 \Leftrightarrow i = n \quad x \in M \tag{4.1}$$

Proof Since an n-manifold is locally homeomorphic to $\mathbb{R}^n$ we have that $H_i(M, M\backslash\{x\}) \cong H_i(\mathbb{R}^n, \mathbb{R}^n\backslash\{0\})$. By Theorem 2.19 and a long exact sequence we get $H_i(\mathbb{R}^n, \mathbb{R}^n\backslash\{0\}) \cong \widetilde{H}_{i-1}(\mathbb{R}^n\backslash\{0\})$. $\mathbb{R}^n\backslash\{0\}$ is contractible and thus $\widetilde{H}_{i-1}(\mathbb{R}^n\backslash\{0\}) \cong \widetilde{H}_{i-1}(S^{n-i})$ which is nonzero iff $i = n$. □

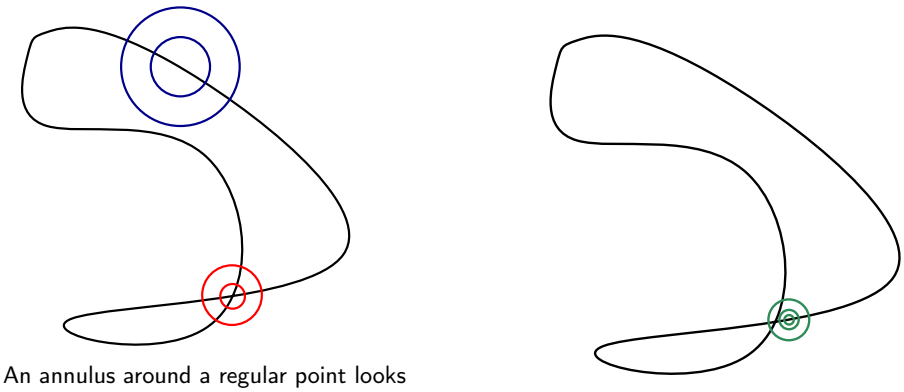
Chapter 5

TARDIS

In this chapter we discuss the TARDIS (Topological Alogorithm for Robust DIscovery of Singularities) framework suggested by von Rohrscheidt and Rieck in *Topological Singularity Detection at Multiple Scales* [VRR23]. This algorithm detects singularities in some given data X set by comparing neighborhoods of some point $x \in X$ with the neighborhood of a point on a manifold. The neighborhoods we consider are annuli with changing inner and outer radius.

5.1 Persistent Local Homology

In the context of stratified spaces, we have seen that there are changes of dimension within the space. If we consider homology we get global information about our space. In the context of manifolds we need a more local tool if we deal with stratified spaces. One fundamental observation that



(a) An annulus around a regular point looks different from an annulus around a singular point.

(b) Change of outer radius changes neighborhood of points close to a singularity.

Figure 5.1: An annulus captures different information depending on the point and the radii.

we will make use of is that an annulus around a point differs depending on the dimension of this area as well as the regularity: In Figure 5.1a we see that any small enough annulus (in red) around a singular point has four connected components, whereas any small enough annulus around a regular point (in blue) has two connected components.

5.1.1 Persistent Intrinsic Dimension

Definition 5.1 *For a metric space (X, d) and $x \in X$, we denote the **intrinsic annulus** of x with respect to the radii r, s by $A_r^s(x) := \{y \in X : r \leq d(y, x) \leq s\}$.*

Definition 5.2 *To get a **tri-filtration** by choosing a set of tuples (r, s) of inner and outer radius. For each such tuple we apply a filtration $\mathcal{F}$ on the intrinsic annulus A_r^s. We denote this tri-filtration by $\mathcal{F}(A_r^s(x), t)$.*

By applying homology to a tri-filtration we get a framework to describe a space locally. The advantage of considering different annuli is that we can get a notion of proximity to a singularity. Consider again the singular point in Figure 5.1b, a point close to this singularity has for a small outer radius an annulus homeomorphic to two intervals. As we increase the outer radius we include at some radius R the singular point, this neighborhood is not anymore homeomorphic to two intervals. This hints that that our point is not far away from the singularity.

Definition 5.3 *We define the **i-th persistent local homology (PLH)** of some x as*

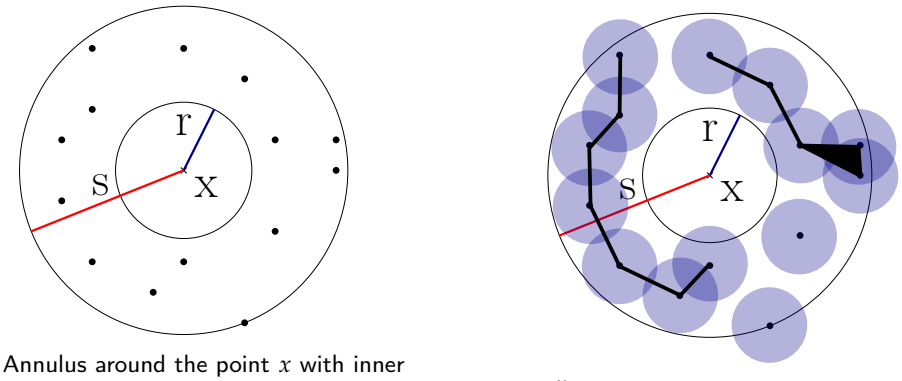
$$PLH_i(x; \mathcal{F}) := PH_i(\mathcal{F}(A_r^s(x), \bullet)).$$

For some fixed r and s and a filtration $\mathcal{F}$, such as the Vietoris-Rips or the Čech filtration.

In other words, the i-th PLH of a point x is the collection of the i-th homology groups of a filtration on any intrinsic annulus centered at x.

Example 5.4 *Consider the intrinsic annulus around x in Figure 5.2a. We choose as filtration the Čech filtration and apply it to the points on the annulus, one such step is shown in Figure 5.2b. We track the persistent homology for this filtration and repeat this process over several annuli with different inner and outer radii to obtain the tri-filtration $\check{C}(A_\bullet^\bullet(x), \bullet)$.*

Theorem 5.5 (Stability of PLH) *Given a finite metric space X and $x \in X$. Let $A_r^s(x)$ and $A_{r'}^{s'}(x)$ be two intrinsic annuli with $|r - r'| \leq \varepsilon_1$ and $|s - s'| \leq \varepsilon_2$. Furthermore, let Dgm_i, Dgm_i' denote the persistence diagrams corresponding to $PH_i(\mathcal{F}(A_r^s(x), \bullet))$ and $PH_i(\mathcal{F}(A_{r'}^{s'}(x), \bullet))$, where $\mathcal{F}$ is either the Vietoris-Rips or the Čech filtration. Then $\frac{1}{2} d_B(Dgm_i, Dgm_i') \leq \max\{\varepsilon_1, \varepsilon_2\}$.*



(a) Annulus around the point x with inner radius r and outer radius s.

(b) Čech complex on annulus for some t.

Figure 5.2: Example of an annulus around x with Čech filtration.

Proof Set $\varepsilon = \max\{\varepsilon_1, \varepsilon_2\}$, by assumption we have that each annulus is contained in an ε-thickening of the the other annulus:

$$A_r^s(x) \subset (A_{r'}^{s'}(x))_\varepsilon \text{ and } A_{r'}^{s'}(x) \subset (A_r^s(x))_\varepsilon.$$

By definition of the Hausdorff distance it follows $d_H(A_r^s(x), A_{r'}^{s'}(x)) \leq \varepsilon$. Let $\delta > 0$ be arbitrary. Since each annulus $A_r^s(x)$ can be covered by a finite union of balls $B_\delta(\cdot)$, $A_r^s(x)$ is totally bounded. Thus we can use Theorem 3.41 and conclude:

$$\tfrac{1}{2} d_b(Dgm_i, Dgm_i') \leq d_H(A_r^s(x), A_{r'}^{s'}(x)) \leq \varepsilon. \qquad \square$$

Definition 5.6 *For a point cloud $X \subset \mathbb{R}^N$, we define the **persistent intrinsic dimension (PID)** of $x \in X$ at scale ε as*

$$i_x(\varepsilon) := \max\{i \in \mathbb{N} \mid \exists r < s < \varepsilon \text{ s.t. } PH_{i-1}(\mathcal{F}(A_r^s(x), \bullet)) \text{ is not empty}\}.$$

Intuitively PID gives us the highest-dimensional feature that persists in this annuli up to a certain scale ε.

Example 5.7 *If we consider a line embedded in $\mathbb{R}^3$ and some fixed ε, every annulus around a point on this line will have two connected components. Therefore we have a non trivial 0-th homology group but higher homology groups are not present. We thus have that the PID of this line is 1.*

We would like to generalize this example: For an n-manifold we would expect that the PID of any point is equal to n.

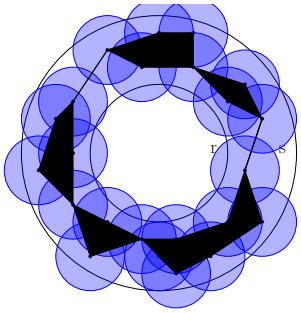


Figure 5.3: The Čech-complex on the data points of the 2-dimensional annulus is homotopic to S^1 for enough data points.

Conjecture 1 *Let $M \subset \mathbb{R}^N$ be an n-dimensional, compact, smooth manifold and $X = \{x_1, \cdots, x_S\}$ a collection of random uniform samples from M, that is a sample where every point is equally likely to be drawn. For S large enough and using the Čech filtration we have that there exists $\varepsilon_1, \varepsilon_2 > 0$ such that for all $x \in X$ and for all $\varepsilon \in (\varepsilon_1, \varepsilon_2)$ we have $i_x(\varepsilon) = n$. Moreover ε_1 can be chosen arbitrary small by increasing S (the number of samples).*

Proof We only give the idea of the proof, since a formal proof involves a lot of technical statements.
$\underline{i_X(\varepsilon) \leq n}$: Let $x \in X$ be any point, since M is a manifold there is a neighborhood of x, $U_x \cong \mathbb{R}^n$. Since M is smooth we can assume (by possibly shrinking it) that U_x is arbitrarily close to being flat. Therefore there exists some $\varepsilon_2 > 0$ such that for all $r, s < \varepsilon_2$ we have that $A_{r,s}(x) \subset U_x$, so $\forall n \geq n_0, \forall t$ we have $\Phi_i(\check{\mathrm{C}}(A_{r,s}(x), t)) = 0$. Since this is true $\forall t$ it follows $\Phi_i(\check{\mathrm{C}}(A_{r,s}(x), \bullet)) = 0$, by definition $i_X(\varepsilon) \leq n$.
$\underline{i_x(\varepsilon) \geq n}$: For enough samples, or equivalently S large enough $\exists t_0$ such that $\check{\mathrm{C}}(A_r^s(x), t_0) \simeq S^{n-1}$. An example of this is drawn in Figure 5.3 for $n = 2$. We conclude that $H_{n-1}(\check{\mathrm{C}}(A_r^s(x), t_0))$ is not trivial and thus $PH_{n-1}(\check{\mathrm{C}}(A_r^s(x), \bullet))$ is not empty, again by definition, $i_x(\varepsilon) \geq n$.
By increasing the sample size we can decrease ε_2 arbitrarily. □

For a compact stratified space, we have that $X_i \setminus X_{i-1}$ is a i-dimensional manifold and thus by performing the formalism of Conjecture 1 to every stratum, we expect that PID captures the right dimension of stratified spaces as well.

5.1.2 Euclidicity

Now that we are equipped with a local dimension measure, we want to measure how much some neighborhood deviates from being 'Euclidean'. As we already observed in Chapter 3, we know that a neighborhood of a singular point looks different from a neighborhood of a regular point. To find singularities in a data set, we take some local dimension estimation such as

PID and compare the neighborhoods of the data points to the neighborhoods of points sampled from a Euclidean model space.

Definition 5.8 *We define the **Euclidean annulus** as*

$$\mathcal{EA}_r^s(x) := \textit{random uniformly distributed samples of } \{y \in \mathbb{R}^n : r \leq d(x,y) \leq s\}.$$

The cardinality of the sample coincides with the number of points of the annulus we want to consider.

Analogously to Definition 5.3 we can define the persistent local homology of this Euclidean annulus:

Definition 5.9 *We define the **i-th persistent local homology of an Euclidean model space** of some x as*

$$PLH_i^{\mathcal{E}}(x;\mathcal{F}) := PH_i(\mathcal{F}(\mathcal{EA}_\bullet^\bullet(x),\bullet)).$$

From now on we focus on $\mathcal{F} = VR$ and use the notation $PLH_i^{\mathcal{E}}(x) := PLH_i^{\mathcal{E}}(x;VR)$.

We observe that $PLH_i^{\mathcal{E}}(x)$ depends on the random sample. Thus we rather consider $PLH_i^{\mathcal{E}}(x)$ to be a sample of random variables $\boldsymbol{PLH}_i^{\mathcal{E}}(x)$. Now that we have a reference annulus we can measure how much the annuli sampled from data sets resemble this Euclidean annulus.

Definition 5.10 *Let $D(\cdot,\cdot)$ be a distance measure for a persistence modules, such as the interleaving distance. The **Euclidicity** of x, denoted by $\mathfrak{E}(x)$ is defined by*

$$\mathfrak{E}(x) := \mathbb{E}[D(PLH_{n-1}(x), \boldsymbol{PLH}_i^{\mathcal{E}}(x))].$$

Intuitively the Euclidicity is the expected distance of the persistent homology on some annuli around the point x and annuli which are sampled from a manifold. If x has a neighborhood similar to a Euclidean space we have a small distance between the two persistent modules. On the other hand if x is a singularity, one expects that any neighborhood is quite different from an Euclidean neighborhood, and thus the Euclidicity of such an X would be high.
To actually calculate $\mathfrak{E}$ one needs to make several decisions such as the range of the radii of the annuli or a distance measure. We will discuss one possible implementation.
The idea is to choose a grid Γ of possible radii r and s, and compute for each $(r,s) \in \Gamma$ the bottleneck distance between the PLH of the Vietoris-Rips complex on the sampled points and the PLH of the Euclidean reference space. Afterwards the Euclidicity is set as the average of these bottleneck distances:

$$\mathfrak{E}(x) \approx \frac{1}{|\Gamma|} \sum_{(r,s)\in\Gamma} d_B(PH_i(\mathcal{F}(A_r^s(x),\bullet)), PH_i(\mathcal{F}(\mathcal{EA}_r^s(x),\bullet)) \tag{5.1}$$

This grid needs an upper and a lower bound for the inner and the outer radius as well as a step size. There are two methods: One idea is to hard-code values for the minimal and maximal outer and inner radius $(r_{min}, r_{max}, s_{min}, s_{max})$, and choose a step size. The paper discusses a data driven approach: Given a point $x \in X$ from the data set, one sets the maximal outer radius s_{max} to the distance of the k-nearest neighbor, the minimal inner radius r_{min} to the smallest distance to any neighbor; and the maximum inner radius r_{max} and the minimum outer radius s_{min} to the $\lfloor \frac{k}{3} \rfloor$th nearest neighbor. The choice of k depends on the intrinsic dimension of the space as well as the density of the points.

Algorithm 1 Initializing inner and outer radius with the k-nearest neighbor method

Require: Data set $X = \{x_1, \cdots, x_n\}$, a point $x \in X$ a distance metric d, list L of size n, $k < n$

1: **for** $x_i \in X$ **do**
2: $\quad L[i] = d(x, x_i)$
3: **end for**
4: **sort L**
5: $s_{max} = L[k]$
6: $r_{min} = L[1]$
7: $s_{min} = L[\lfloor \frac{k}{3} \rfloor]$
8: $r_{max} = L[\lfloor \frac{k}{3} \rfloor]$

5.2 Experiments

In this section we demonstrate the two different approaches and show the influence of different parameter choices. For the experiments we use the code provided by von Rohrscheid and Rieck on Github (https://github.com/aidos-lab/TARDIS/tree/main). The script used to create the plots can be found in the Appendix 6.

5.2.1 Parameter

The code allows for several choices of parameters, the following list contains the ones we will focus on, in the different experiments:

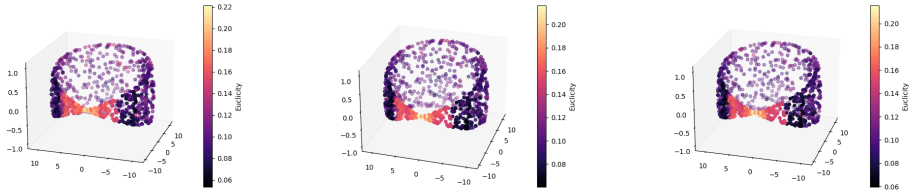
- -k : Number of neighbors to compute the k nearest neighbor.
- -q : Number of query points; for this points the Euclidicity is calculated.
- - -seed : Seed for random sampling, for all experiments this is set to 12, for reproducibility purposes.

- -r : Minimal inner radius.
- -R : Maximal inner radius.
- -s : Minimal outer radius.
- -S : Maximal outer radius.
- - -num-steps : Number of steps in the grid to go from the inner to the outer radius.
- -d : This value should be the known or estimated intrinsic dimension. It is an upper bound for the PID calculation.

We discuss the effect of the k nearest neighbor method compared to hard-coded inner and outer radii, as well as different choices of k, step sizes and estimated dimension.

5.2.2 Step Size

We start by fixing some choices for the radii and comparing different choices of step sizes. For the radii of the pinched torus we choose the values that are also used in the example on Github (https://github.com/aidos-lab/TARDIS/tree/main): r = 0.05, R = 0.45, s = 0.2, S = 0.6. For the radii of the wedged sphere we choose some smaller values: r = 0.05, R = 0.25, s = 0.2, S = 0.4.
We plotted the Euclidicity score of the points for the step size num-steps $\in \{5, 10, 20\}$. It turns out that visually the pinched point is easy to spot for all step sizes, see Figure 5.4 and Figure 5.5.



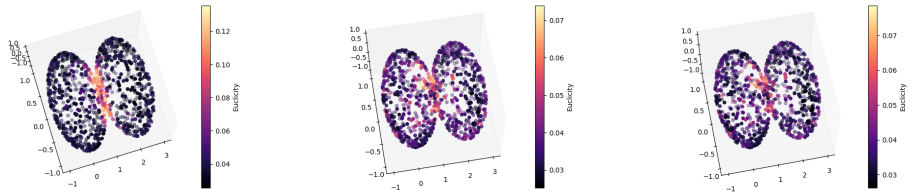
(a) Number of steps = 5 **(b)** Number of steps = 10 **(c)** Number of steps = 20

Figure 5.4: Effect of different step choices on Euclidicity score of the pinched torus.

This robustness could be used to make the computation faster by choosing a smaller step size.

5.2.3 k Nearest Neighbor vs Fixed Choice

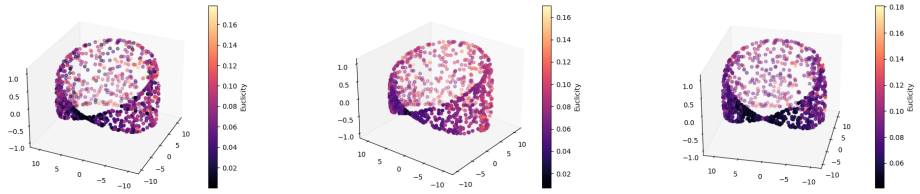
Let d be a metric. The k nearest neighbor method calculates the distance of some point $x \in X$ to all other points, and orders them in increasing order



(a) Number of steps = 5 **(b)** Number of steps = 10 **(c)** Number of steps = 20

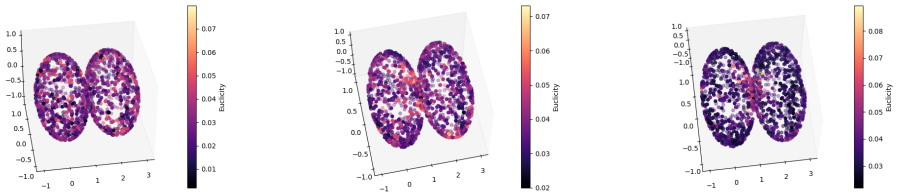
Figure 5.5: Effect of different step choices on Euclidicity score of the wedged spheres.

$\{x_1, \cdots, x_m\}$. x_k is the k nearest neighbor of x. We use $d(x, x_k)$ to initialize s_{max}, we set $s_{min} = r_{max}$ to $d(\lfloor \frac{k}{3} \rfloor, x)$ and r_{min} to $d(x, x_1)$.
For the pinched torus the k nearest neighbor method behaves unexpectedly; the singular regions have a lower Euclidicity score than the regular regions. This is contrary to the values we get for fixed values (compare Figure 5.6b to Figure 5.4b). We can still detect the singular region but not by looking for high but for small values. The reason for this behaviour is not yet clear.
We always used a step size equal to 10 for the pinched torus and equal to 5 for the wedged sphere, this is due to the observation in the previous section. In Figure 5.6a, 5.6b and 5.6c the torus is plotted for $k \in \{5, 20, 50\}$.
For the wedged sphere we tried $k \in \{5, 50, 100\}$. For the wedged sphere (see



(a) k = 5 **(b)** k = 20 **(c)** k = 50

Figure 5.6: K nearest neighbor for the pinched torus.



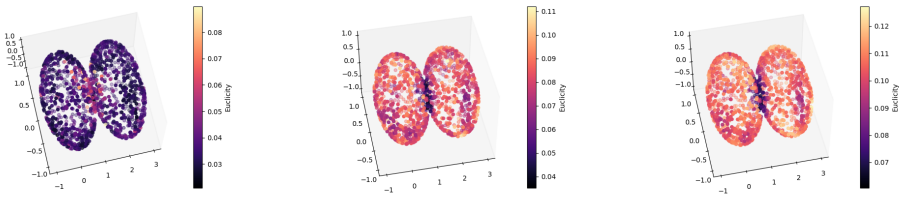
(a) k = 5 **(b)** k = 50 **(c)** k = 100

Figure 5.7: K nearest neighbor for the wedged sphere.

Figure 5.7) we get the expected result: The points around the gluing point have a higher Euclidicity score than points further away. We see for both objects that a higher value of k yields more contrast.

5.2.4 Dimension Estimate

If the estimated dimension d is wrong this implementation of the Euclidicity score is not always robust. In the following experiments we fixed the following parameters: k = 100, num-steps = 10, q = 1500.
We let the estimated dimension d vary from $2-4$. As one can observe in Figure 5.8, the Euclidicity score for $d = 3$ and $d = 4$ becomes inverted. If the estimated dimension is off, the Euclidicity score is hard to interpret. A possible explanation for the example of the wedged sphere is that the link of the gluing point is given by $S^1 \sqcup S^1$ which is possibly closer to a 2D-sphere (the reference annulus) than the link of a regular point, which is just a circle.



(a) d = 2 **(b)** d = 3 **(c)** d = 4

Figure 5.8: Effect of different estimated intrinsic dimensions.

Chapter 6

Appendix

Listing 6.1: Script to plot the Euclidicity score

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Read the data
# change this to the appropriate path if needed
# Pinched Torus
data = pd.read_csv(r'..\output\Pinched_torus.txt',
                   sep=",",
                   skiprows=1,
                   header=None,
                   encoding='utf-16')
# Wedged Spheres
#data = pd.read_csv(r'..\output\Wedged_spheres_2D.txt',
                   sep=",",
                   skiprows=1,
                   header=None,
                   encoding='utf-16')

data.columns = ['x0', 'x1', 'x2', 'Euclicity', 'PID']

data = data.iloc[:, :4]

# Display the first few rows of the dataframe
data.head()
print(data.head())

fig = plt.figure()
```

```
ax = fig.add_subplot(projection='3d')

#use this for Euclicity
sc = ax.scatter(data['x0'],
                data['x1'],
                data['x2'],
                c=data['Euclicity'],
                cmap='magma')

plt.colorbar(sc, ax=ax, label='Euclicity')

ax.grid(False)

plt.show()
```

Listing 6.2: Script to plot persistent intrinsic dimension

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Read the data
data = pd.read_csv(r'..\output\Pinched_torus.txt',
                sep=",",
                skiprows=1,
                header=None,
                encoding='utf-16')
#data = pd.read_csv(r'..\output\Wedged_spheres_2D.txt',
                sep=",",
                skiprows=1,
                header=None,
                encoding='utf-16')
data.columns = ['x0', 'x1', 'x2', 'Euclicity', 'PID']

data = data.iloc[:, [0, 1, 2, 4]]
# Display the first few rows
data.head()
print(data.head())

fig = plt.figure()
ax = fig.add_subplot(projection='3d')

#use this for PID
sc = ax.scatter(data['x1'],
            data['x0'],
```

```
                data['x2'],
                c=data['PID'],
                cmap='cividis')

    plt.colorbar(sc, ax=ax, label='PID')

    ax.grid(False)

    plt.show()
```

Bibliography

[Ban07] Markus Banagl. *Topological Invariants of Stratified Spaces*. Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2007.

[Car14] Gunnar Carlsson. Topological pattern recognition for point cloud data. *Acta Numerica*, 23:289–368, 2014.

[CCSG+09] Frédéric Chazal, David Cohen-Steiner, Leonidas J. Guibas, Facundo Mémoli, and Steve Y. Oudot. Gromov-hausdorff stable signatures for shapes using persistence. *Computer Graphics Forum*, 28(5):1393–1403, 2009.

[DW22] Tamal Krishna Dey and Yusu Wang. *Computational Topology for Data Analysis*. Cambridge University Press, 2022.

[Fri20] Greg Friedman. *Singular Intersection Homology*. New Mathematical Monographs. Cambridge University Press, 2020.

[Hat02] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002. Accessed on April 29, 2024.

[Lee13] John M. Lee. *Introduction to Smooth Manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition, 2013.

[VRR23] Julius Von Rohrscheidt and Bastian Rieck. Topological singularity detection at multiple scales. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 35175–35197. PMLR, 23–29 Jul 2023.