NTNU
Norwegian University of
Science and Technology

Topology and Data

Øyvind Brekke

Master of Science in Physics and Mathematics
Submission date: June 2010
Supervisor: Nils A. Baas, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Learn to use, and describe the theory behind Plex, which is a software package for computing persistent homology of finite simplicial complexes, often generated from point cloud data. In particular, the theory behind persistent homology is to be included.

Learn to use, and describe the theory behind Mapper, which is a computational method for extracting simple descriptions of high dimensional data sets in the form of simplicial complexes.

Run tests on point clouds generated from known topological spaces.

There is also an opportunity to have a look at biological data, if this presents itself as resonable thing to do.

Assignment given: 25. January 2010
Supervisor: Nils A. Baas, MATH

Preface

This is my master's thesis(course code TMA4900), written the spring 2010 at the Norwegian University of Science(NTNU). I want to thank professor Nils A. Baas for his support and advice. I also want to thank my brother, Birger Brekke, who has been a companion throughout this writing process, while working on his own thesis on the same subject. Finally, I want to thank Till Tantau as author and developer of TikZ, which is a graphical tool package for LaTeX, it has been invaluable in creating the figures and diagrams in this thesis.

Øyvind Brekke, June 21, 2010

Abstract

Today there is an immense production of data, and the need for better methods to analyze data is ever increasing. Topology has many features and good ideas which seem favourable in analyzing certain datasets where statistics is starting to have problems. For example, we see this in datasets originating from microarray experiments. However, topological methods cannot be directly applied on finite point sets coming from such data, or atleast it will not say anything interesting. So, we have to modify the data sets in some way such that we can work on them with the topological machinery. This way of applying topology may be viewed as a kind of discrete version of topology. In this thesis we present some ways to construct simplicial complexes from a finite point cloud, in an attempt to model the underlying space. Together with simplicial homology and persistent homology and barcodes, we obtain a tool to uncover topological features in finite point clouds. This theory is tested with a Java software package called JPlex, which is an implementation of these ideas. Lastly, a method called Mapper is covered. This is also a method for creating simplicial complexes from a finite point cloud. However, Mapper is mostly used to create low dimensional simplicial complexes that can be easily visualized, and structures are then detected this way. An implementation of the Mapper method is also tested on a self made data set.

Contents

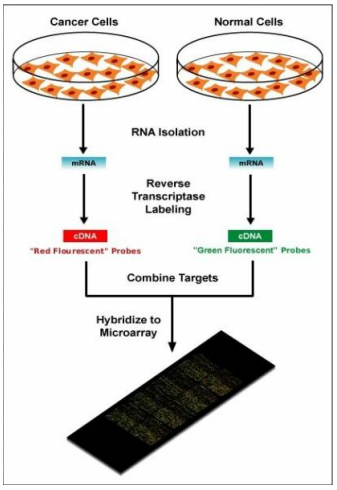


Figure 1: A Microarray experiment.

1 Introduction

In the present day, data of varoius kinds are being produced as never before. These might come from new experimental methods, in the study of different kinds of networks like Facebook, scanning of geometrical objects, results from psychological questionaire responses and so on. The challenge is that this production of data is developing faster than our capabilities to deal with them. This is were topology might be in a position to contribute to the greater good. Let us take a look at a new experimental method using microarrays and how topology might have favourable attributes to deal with challenges present in analyzing these types of data sets.

A microarray is a tool for analyzing genes. All biological cells contain a full set of chromosomes and identical genes, but only a small part of these are active. This means that different cells produce different types and amounts of messenger RNA(mRNA), which are the blueprints for making proteins. This process is termed *gene expression*. The microarrays work by exploiting the fact that a given mRNA molecule is able to bind, *hybridize*, to the DNA template from which it originated. So a microarray is a small membrane or glass slide containing an array of selected DNA samples, called probes. These may number in the hundreds or thousands. The microarray is then exposed to mRNA from sample cells alone or together with mRNA from control cells. This will then produce different light intensities or color variations on the array, depending on which of the metods above is used, and relative gene expression levels are expressed at the specific probe spots for specific genes.

With this new experimental method, several challenges have arised.

Qualitative differentiations. Often one want to try and distuingish gene expres-

sions in different types of specimens. Say we have breast cancer patients with or without some type of mutation, and that the types of tumors are classified beforehand by some other method, for instance biopsy. Then, if there are differences in the gene expressions of these two classes of tumors, this can be used for diagnostics.

High-dimensionality. A challenge when dealing with microarray data is the high-dimensionality. Microarrays enable scientists to check thousands of genes simultaniously, but the samples may only vary in the order of tens to hundreds, this is due to costs and sample availability. Statistical analysis has been the tool of choice for dealing with these types of data. The problem is that standard statistical methodology works best in the opposite case, when the samples are in the thousands and the dimension is in the tens or hundreds.

Relative quantities. As mentioned microarray analysis can only say something about the relative gene expressions in a sample. DNA and mRNA consists of strands with nucleotide base pairs, and in a microarray experiment the more correct pairs there are between a DNA probe strand and an mRNA strand, the stronger the bonds are. After an experiment the amount of pairs on a specific gene probe depends on experiment conditions like temperature under hybridization, washing after hybridization and amount of mRNA available. This means that it is the relative differences between quantities that are important and not the quantities on their own.

No theoretical backed up metric. In contrast to physics, there is no theoretical backed up metric to rely on to measure distance in these data sets. Metrics are mostly computed with regard to all genes that are registered at some high enough value, or with regard to genes that show some meaningfull variation across the sample set. However, often there are relatively few genes that separate different specimen classes, and cluster analysis, which creates groups of cluster with respect to these metrics, will often not pay enough attention to the effects of these genes. In short one can not really trust the metrics that much here.

Noise. There may be a lot of noise in microarray experiments. Some mRNAs may cross-hybridize to other probes which are supposed to detect other mRNA, there might be cases where probes are based on information incorrectly associated with the gene they are supposed to detect, and there are measurement difficulties as well.

Let us see what topology has to offer. Topology only study the properties of geometric objects that does not depend on coordinates. When in a metric space, only relative positions are important, which is the case with data sets from microarrays.

Topology is the branch of mathematics that deals with qualitatively differentiation. More specifically, we will in this paper focus on homology, which is a crude measure of topological properties. Through the lenses of homology a circle and a square are not dis-

tuingishable. This is because homology is invariant under continuous deformations. This means that because a circle and a square can be deformed to one another without any tearing they are not distuingished. Whereas the letter 'B' and a circle is distuingished because the letter 'B' cannot be deformed to a circle without ripping the letter apart and visa versa. This line of thought is then applied for higher dimensional surfaces as well.

Because homology is invariant under continuous deformations, it follows that homology is not as sensitive to the metrics chosen as other methods might be. As an example, let us say that two points are connected by an edge if they are within some distance of each other. Because of the invariance properties of homology, the actual length of the edge does not really matter, only that it exists in the first place! This also makes homological methods less affected by noise. The natural follow up question might be how to decide if there is an edge or not, but before we look into that let us adress the high-dimensionality.

To adress the high-dimensionality, let us just go ahead and see how one can deal with such data sets from a topological point of view. From microarray data sets we get a matrix which consists of genes along the rows and samples in the columns. The entries in this matrix are numbers which represent relative gene expression levels. The column vectors can then be viewed as vectors in some metric space, for instance in euclidean space, and we call them *point clouds*. Looking at the rows as vectors is also possible, giving other studying possibilities.

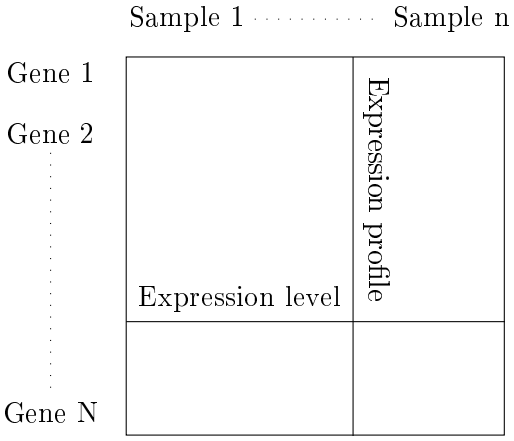


Figure 2: Matrix consisting of relative gene expression levels.

The dimensional reduction methods that are often used are various clustering methods, these methods detect connectedness in the set. Say we are given a point cloud in a metric space. Then we can decide that some points are close enough to be connected and makes a cluster. Hence, we get a collection of clusters. Furthermore, there might be connections between such clusters and clusters of clusters, and so on. What it does not capture is higher dimensional connectivity data and structures. For instance, if we

have points sampled from a sphere, which has a void in it. Can we extract this kind of information?

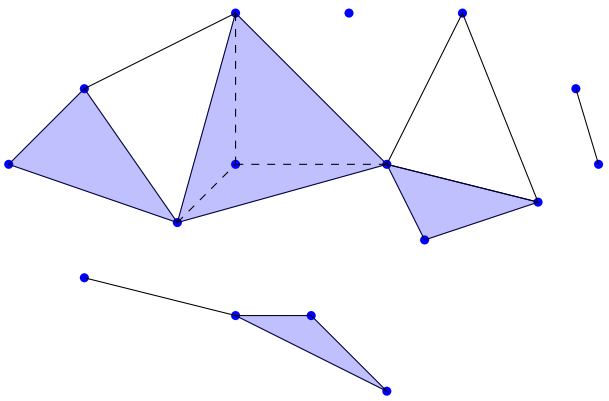


Figure 3: A simplicial complex.

A central idea is to use what we call simplicial complexes, which are built from p-simplices, a 0-simplex is a point, a 1-simplex is an edge, a 2-simplex is a triangle, a 3-simplex is a tetrahedron, and a p-simplex is a higher dimensional analog for a triangle. So what happens is that some points are used as scaffolds for creating edges, like in clustering, but then the edges can be used as scaffolds for building triangels, the triangles as scaffold for building tetrahedrons and so on. In this way we hope to create a good model of the underlying space, with similar topological features. These constructions can then be analysed with homology to detect global higher-dimensional features. An example is given in Figure 4.

When constructing a simplicial complex, we have to decide when points may be connected by an edge and when these may be used for creating a triangle and so on. This will often depend on a distance ϵ, as in Figure . Which epsilon should one choose? In Figure , if we choose a too small ϵ, then no points will be connected, too large an ϵ and all points will be connected. Remember that we generally do not know what the underlying structure is, so how can we decide which ϵ gives the correct structure? Moreover, there is a good chance that such an ϵ does not even exist or it exists in a very small neighborhood. A way to deal with this is persistent homology. It is possible to construct simplicial complexes such that they increase in size with respect to some ϵ, like in Figure , such that a simplicial complex constructed with a low ϵ is a *subset* of a simplicial complex constructed with a higher ϵ. Hence, if we start with some small ϵ and increase it in steps we get a sequence of simplicial complexes, where topological features appear and dissappear. Then if features persist over a long time it might indicate that these are actual features of the underlying space while short lived features are most likely noise. This will be be illustrated by sets of intervals that we will call *Barcodes*.

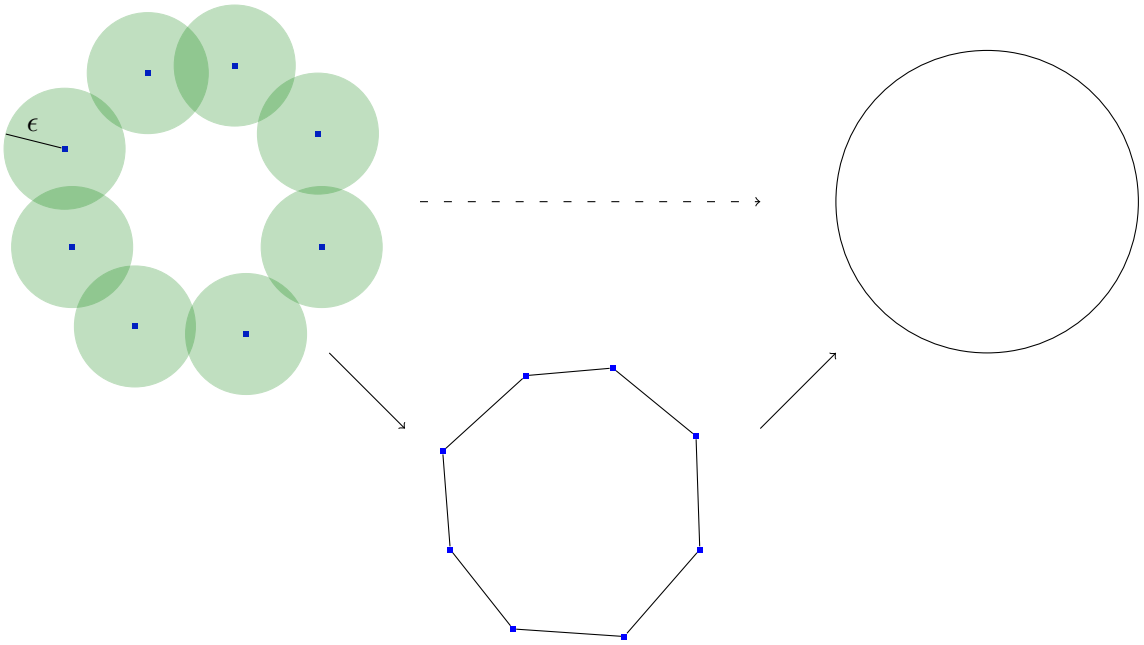


Figure 4: The point-cloud on the left seems to have been sampled from a circle. One way to build a simplicial complex is to draw edges between the points who are closer to each other than some distance ϵ under the metric in question. From this we get a simplicial complex and something even more resembling a circle. Through the lenses of homology this is exactly the same as a circle or a loop. So from points that are 0-dimensional objects we have inferred a 2-dimensional property.

To conclude the introduction, it seems that topology has favourable features to analyze such types of large data sets, like those coming from microarray experiments. This way of applying topology may be viewed as a discrete version of topology, where ideas in topology are being applied on discrete sets in some way. We will look into simplicial homology and present some simplicial constructions which tries to approximate topological spaces from points sampled from them. Then we will introduce persistent homology, barcodes, and an algorithm for computing these barcodes from a point cloud. Furthermore, we will run some tests on points sampled from the unit torus and sphere with a program called JPlex which uses this algorithm. We will also introduce a method called Mapper, which is yet another method for creating simplicial complexes. Mapper does mainly produce low dimensional simplicial complexes that are easily visualizable for us humans, and hence exploits our own innate ability to detect features and structures. We will also test the Mapper method on a simple data set. From now on it is assumed that the reader has knowledge about general topology. For this topic, one may for example consult [20].

2 Simplicial homology

The main material in this introduction on simplicial homology has been gathered from [10], [30]. The reader is assumed to have knowledge of basic abstract algebra. For this topic, consult for example [1].

In the introduction we motivated the use of creating simplicial complexes on data sets to approximate the homology of a underlying space. We will now introduce simplicial homology. The homology theory which is applicable to all topological spaces is called singular homology, which can be viewed as the "correct" answer, but we will not define singular homology here. Now, we would like to have some automatic way to compute the homology of spaces via computer computations, but singular homology is not suited for this and computational methods are done by hand. However, the computation of simplicial homology can be made combinatorial, which makes it available for computer computations. The underlying topological spaces will naturally not be simplicial complexes but, if the underlying space is homotopy equivalent to a simplicial complex, then their homology will be the same. If such a homotopy equivalence does not exist, the constructed simplicial complex will hopefully still approximate the space in some way and tell us something about it.

2.1 Homological algebra

Since we are going to introduce a homology theory we introduce some basic homological algebra.

Definition 2.1. A *chain complex* (A_*, d_*) is a sequence of abelian groups or modules,

$$\cdots \longrightarrow A_{p+2} \xrightarrow{d_{p+2}} A_{p+1} \xrightarrow{d_{p+1}} A_p \xrightarrow{d_p} A_{p-1} \xrightarrow{d_{p-1}} A_{p-2} \longrightarrow \cdots$$

connected by homomorphisms $d_p : A_p \to A_{p-1}$ such that the composition of any two consecutive maps is zero, i.e. $d_{p-1} \circ d_p = 0$ for all n. These maps are called *boundary operators*.

Definition 2.2. If $C = (A_*, d_*)$ is a chain complex, then we define the p-th homology of this chain complex to be the quotient

$$H_p(C) = \frac{\ker(d_p)}{\mathrm{im(d_{p-1})}}.$$

This is ok since the property $d \circ d = 0$ gives that $\mathrm{im(d_{p+1})} \subseteq \mathrm{ker(d_p)}$.

Definition 2.3. If (A_*, d_*^A) and (B_*, d_*^B) are two chain complexes then a chain map $\phi : (A_*, d_*^A) \to (B_*, d_*^B)$ is given by a sequence of homomorphisms $\phi^p : (A_*, d_*^A) \to (B_*, d_*^B)$ such that $d_p^B \circ \phi_p = \phi_{p-1} \circ d_p^A$.

2.2 Simplicial complexes

Simplicial homology is based on computing homology of simplicial complexes so we start with defining simplices and simplicial complexes.

Definition 2.4. Let x and y be points in $\mathbb{R}^n$. Define the *segment* from x to y to be $\{(1-t)x + ty | 0 \leq t \leq 1\}$. A subset $C \subseteq \mathbb{R}^n$ is *convex* if for all pairs (x, y) in C the segment from x to y lies entirely in C. Note that an arbitrary intersection of convex sets is convex. Let $A \subseteq \mathbb{R}^n$, and define the *convex hull* of A to be the intersection of all convex sets in $\mathbb{R}^n$ which contain A.

Definition 2.5. A p-simplex σ in $\mathbb{R}^n$ is the convex hull of a set of points $K = \{v_0, v_1, ..., v_p\}$ such that $\{v_1 - v_0, v_2 - v_0, ..., v_p - v_0\}$ are p linearly independent vectors in $\mathbb{R}^n$. We call $\{v_0, v_1, ..., v_p\}$ the vertices of σ.

If we had chosen any other ordering $\sigma_{i_0}, \sigma_{i_1}, ..., \sigma_{i_p}$ of the vertices, then $\{\sigma_{i_1} - \sigma_{i_0}, \sigma_{i_1}\sigma_{i_0}, ..., \sigma_{i_p} - \sigma_{i_0}\}$ would still be p-linearly independent vectors in $\mathbb{R}^n$.

Definition 2.6. Let σ be a p-simplex defined by $K = \{v_0, v_1, ..., v_p\} \subset \mathbb{R}^n$. A simplex τ defined by $L \subseteq K$ is called a *face* of σ, and if $|\tau| = (p-n)$, then τ is called a $(p-n)$-face of σ.

Definition 2.7. If a p-simplex σ is defined by the set $S = \{e_0, e_1, ..., e_p\}$ consisting of the unit vectors in $\mathbb{R}^n$, then σ is called the standard p-simplex.

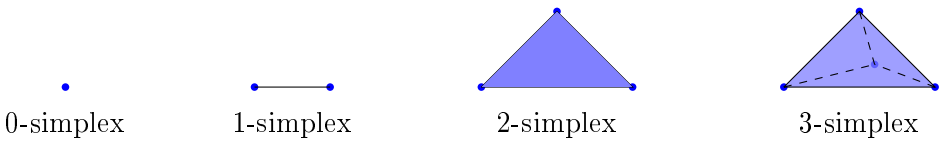


Figure 5: Examples of simplices.

Definition 2.8. A *simplicial complex* S in $\mathbb{R}^n$ is a set of simplices s.t.

1. If $\sigma \in S$, then every face of σ is in S.
2. If $\sigma, \tau \in S$ and $\sigma \cap \tau \neq \emptyset$, then $\sigma \cap \tau$ is a simplex whose vertices are also vertices of both σ and τ.

See Figure 3 again.

Definition 2.9. Let S be a simplicial complex. The *support* $|S| = \bigcup_{\sigma \in S} \sigma$ of S in $\mathbb{R}^n$ is the union of the simplices which belong to it.

Definition 2.10. Let S and T be two simplicial complexes. A map $f : S \to T$ is called a simplicial map if whenever a set of vertices $\{v_0, ..., v_n\}$ of S span a simplex in S, then $\{f(v_0), ..., f(v_n)\}$ span a simplex in T.

Definition 2.11. A *subcomplex* of a simplicial complex S is a simplicial complex $L \subseteq S$.

Definition 2.12. A filtration of a simplicial complex S is a nested sequence of subcomplexes

$$\emptyset = S^0 \subseteq S^1 \subseteq \cdots \subseteq S^m = S.$$

We call a simplicial complex S with a filtration for a *filtered complex*.

This was the geometric definition of a simplicial complex. However, we can define simplicial complexes without any geometry involved. This definition is purely combinatorial, and hence easily stored and manipulated in a computer. So it is this type that is used during applications, however, as we will see, we can still think of them intuitively as the geometric ones.

Definition 2.13. An *abstract simplicial complex* is a pair (K, Σ), where Σ is a collection of nonempty finite subsets of K, called *(abstract) simplices*, such that:

1. For all $v \in K$, $\{v\} \in \Sigma$. The sets $\{v\}$ are called the vertices of K.
2. If $\tau \subseteq \sigma \in \Sigma$, then $\tau \in \Sigma$.

Moreover, we call (K, Σ) a *finite abstract simplicial complex* if $|K|$ is finite. A simplex σ is a p-simplex if $|\sigma| = p + 1$ and if $|\sigma| = 1$ then σ is called an *edge*. If $\tau \subseteq \sigma$, then τ is a *face* of σ.

We can relate this combinatorial definition to the geometric one.

Definition 2.14. Let S be a simplicial complex and let K be the vertex set of S, i.e. all the vertices in S. Furthermore, let Σ be the collection of all subsets $\{v_{i_0}, v_{i_1}, ..., v_{i_n}\}$ of K such that the vertices $\{v_{i_0}, v_{i_1}, ..., v_{i_n}\}$ span a simplex of S. The collection Σ is called the *vertex scheme* of S.

Definition 2.15. If the abstract simplicial complex (K, Σ) is isomorphic with the vertex scheme of the simplicial complex S, then S is said to be a geometric realization of (K, Σ).

Definition 2.16. Given a finite abstract simplicial complex (K, Σ) we can create a geometric realization $|(K, \Sigma)|$ of it in the following way. Let $\phi : K \to \{1, 2, ..., n\}$ be a bijection. We then define $|(K, \Sigma)|$ to be the subspace of $\mathbb{R}^n$ given by the union $\cup_{\sigma \in \Sigma} c(\sigma)$, where $c(\sigma)$ is the convex hull of the set $\{e_{\phi_s}\}$, and where e_i are the unit vectors $\mathbb{R}^n$. The space $|(K, \Sigma)|$ is a simplicial complex and we will call it the *standard geometric realization* of (K, Σ).

Example 2.17. Let ϕ be the map that maps $\phi(a) = 1$, $\phi(b) = 2$, $\phi(c) = 3$ and $\phi(d) = 4$. Let the abstract simplicial complexes be given by $K = \{a, b, c\}$, $\Sigma = \{\{a, b\}, \{a\}, \{b\}, \{c\}\}$, $\Sigma' = \{\{\Sigma\, \{a, c\}, \{b, c\}\}$, and $\Sigma'' = \{\{\Sigma', \{a, b, c\}\}$. Then we get the following showed in Figure 6. If $K = \{a, b, c, d\}$ and $\Sigma''' = \{\{\Sigma''\}, \{a, d\}, \{d\}\}$, then we get the following showed in Figure 7.

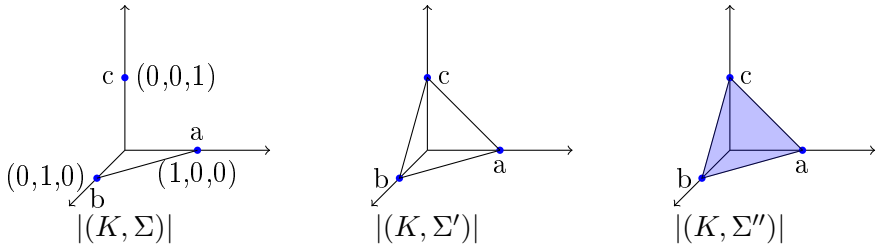


Figure 6: See Example 2.17 for an explanation.

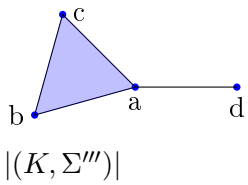


Figure 7: See Example 2.17 for an explanation.

Theorem 2.18. *All geometric realizations of an abstract simplicial complex are homotopy equivalent.*

Hence, all geometric realizations of an abstract simplicial complex are viewed as the same in eyes of homology.

Finally, abstract subcomplexes and filtrations are defined in a similar manner as we did earlier for simplicial complexes. Hence, due to the similarity we will often just say a complex to refer to both the geometric and abstract version.

2.3 Simplicial homology

To define a homology on simplicial complexes, we create chain complexes from the simplicial complexes. We create our abelian groups in the chain complex as follows.

Definition 2.19. Let S be a simplicial complex. An *orientation* of a p-simplex $\sigma = \{v_0, v_1, ..., v_p\} \in K$ is an equivalence class of orderings of the vertices of σ, where

$$(v_0, v_1, ..., v_p) \simeq (v_{\tau(0)}, v_{\tau(1)}, ..., v_{\tau(p)})$$

are equivalent if the parity of the permutation τ is even. A simplex with an ordering is called an *oriented simplex*, and is denoted by $[\sigma]$.

Definition 2.20. The *p-th chain group* $C_p(S)$ of a simplicial complex S, is the free abelian group on the oriented p-simplices, where $[\sigma] = -[\tau]$ if $\sigma = \tau$ and σ and τ have different orientation. An element $c \in C_p(S)$ is called a *p-chain* and is of the form $c = \sum n_q[\sigma_q]$ with $\sigma_q \in S$ and coefficients $n_q \in \mathbb{Z}$.

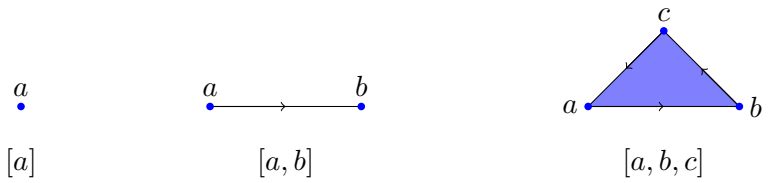


Figure 8: Oriented simplices.

Next we need a boundary operator $\partial_p : C_p \to C_{p-1}$ so that $\partial_{p-1} \circ \partial_p = 0$. Let us consider a basis element $\sigma = [v_0, ..., v_p] \in C_p$. If we do a sum over all the $(p-1)$-faces of σ in a way such that if we do it twice we end up with the empty set, then this is our desired operator.

Definition 2.21. Let S be a simplicial complex and let $\sigma = [v_0, v_1, ..., v_p] \in S$. The boundary homomorphism $\partial_p : C_p(S) \to C_{p-1}(S)$ is given by

$$\partial_p \sigma = \sum_{i=0}^{n} (-1)^i [v_0, v_1, ..., \hat{v}_i, ..., v_n],$$

where $[v_0, v_1, ..., \hat{v}_i, ..., v_n]$ denotes that v_i is removed. This is well defined; that is, ∂_k is the same for every ordering in the same orientation. Also $\partial \circ \partial = 0$.

We are now able to define the homology of a simplicial complex.

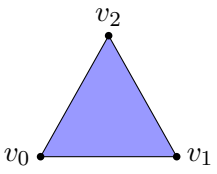
Definition 2.22. Elements in $Z_p = \ker(\partial_p)$ are called *cycles*.

Definition 2.23. Elements in $B_p = \mathrm{im}(\partial_{\mathrm{p}+1})$ are called *boundaries*.

Definition 2.24. Let S be a simplicial complex. The p-th homology group of S is defined to be the quotient

$$H_p(S) = \frac{Z_p}{B_p}.$$

Example 2.25. The following triangle T is a simplicial complex consisting of one 2-



simplex and its faces.

Let us check if $\partial \circ \partial = 0$. Since $\partial_0([v_i]) = 0$ we have that $\partial_0 \circ \partial_1 = 0$. Let $z \in \mathbb{Z}$, then for $\partial_1 \circ \partial_2$ we have that:

$$\begin{aligned} \partial_1 \circ \partial_2(z[v_0, v_1, v_2]) &= \partial_1(z[v_1, v_2] - z[v_0, v_2] + z[v_0, v_1]) \\ &= z([v_2] - [v_1] - [v_2] + [v_0] + [v_1] - [v_0]) \\ &= 0. \end{aligned}$$

Since there are no p-chains for $p > 2$, we get that $\partial_p \circ \partial_{p+1} = 0$ for all p.

Next let us compute $H_0(T)$. Let $a, b, c, z_i \in \mathbb{Z}$.

$$\partial_1(a[v_0, v_1] + b[v_1, v_2] + c[v_0, v_2]) = (a - b)[v_1] - (a + c)[v_0] + (b + c)[v_2].$$

A 0-chain $z_0[v_0] + z_1[v_1] + z_2[v_2]$ is in $im(\partial_1)$ iff $z_0 = a - b$, $z_1 = a + c$ and $z_2 = b + c$ for some $a, b, c \in \mathbb{Z}$. This gives two degrees of freedom, i.e. $\text{im}(\partial_1) \simeq \mathbb{Z} \oplus \mathbb{Z}$. Since $\partial_0[v_i] = 0$, $\ker(\partial_0) = \{a[v_0] + b[v_1] + c[v_2] | a, b, c \in \mathbb{Z}\} \simeq \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$. This gives that

$$H_0(T) = \frac{\mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}}{\mathbb{Z} \oplus \mathbb{Z}} \cong \mathbb{Z}.$$

Now, we would like to use this definition of homology on other topological spaces than simplicial complexes. The idea is to turn a given space into a simplicial complex in a way that does not change the homology of the space. This is not always possible, but if it is the resulting simplicial complex we can then handle. Let us first recall the following.

Theorem 2.26. *If X and Y are two topological spaces and $f : X \to Y$ is a homotopy equivalence, then X and Y induce the same homology in any homology theory.*

Particularly remember than any homeomorphism between two spaces is a homotopy equivalence.

Definition 2.27. A *triangulation* of a topological space X is a homeomorphism $T : |S| \to X$ from the support of S to X. If such a triangulation of X exists, we say that X is triangulable.

Definition 2.28. If $T : |S| \to X$ is a triangulation of X, then we define the p-th homology group of X with respect to T as

$$H_p^T(X) = H_p(S). \tag{1}$$

To define the simplicial homology $H^{simp}(X)$ of a triangulable topological space X, we would have to consider refinements of triangulations and taking a direct sum, we will not do that. However we can think of $H^{simp}(X)$ as the same as $H_p^T(X)$, where T is any triangulation of X, since the following theorem will ensure us that they are isomorphic. It also ensures us that $H_p^T(X)$ and $H_p^{sing}(X)$ are the same isomorphically, where $H_p^{sing}(X)$ denotes singular homology.

Theorem 2.29. *If $T : |S| \to X$ is any triangulation of a topological space X then there are natural isomorphisms*

$$H_p^{simp}(X) \cong H_p^{sing}(X) \cong H_p^T(X),$$

where $H_p^{simp}(X)$ denotes p-th simplicial homology group and $H_p^{sing}(X)$ denotes the p-th singular homology group.

Remark 2.30. *Note that a complex does not necessarily need to be a triangulation of a topological space X to have the same homology. It is enough that it is a triangulation of a topological space Y that is homotopy equivalent to X.*

As a consequence of Theorem 2.18, the homology of geometric realizations of an abstract simplicial complex are all the same. The "abstract simplicial homology" is constructed exactly in the same manner as for the simplicial homology, just exchange the p-simplex from a simplicial complex with its corresponding abstract p-simplex giving us the same thing. Computation of the simplicial homology does then become purely combinatorial.

2.4 Alternative coefficients and betti numbers

Earlier we created chain groups $C_p(S)$ with coefficients in $\mathbb{Z}$. We can however view these groups as a $\mathbb{Z}$-modules, and as a result the homology groups $H_p(S)$ also become $\mathbb{Z}$-modules. In this case we say that the homology groups have *ground ring of coefficients* in $\mathbb{Z}$. Now, we can change this ground ring of coefficients. So, if the ring is a Principal Ideal Domain(PID) D, we will obtain homology groups that can be viewed as D-modules.

Definition 2.31. Let S be a simplicial complex, and let D be a PID. Then we let $H_p(S; D)$ denote the homology group of S with coefficients in D.

The structure of such finitely generated D-modules is given by the following theorem.

Theorem 2.32. *If D is a PID and M is a finitely generated D-module, then*

$$M \cong \left(D^{\mathcal{B}}\right) \oplus \left(\bigoplus_{i=1}^{m} \frac{D}{d_i D}\right),$$

where $d_i \in D$ such that $d_i \mid d_{i+1}$, and $\mathcal{B} \in \mathbb{Z}$. The left part in the above expression is called the free part while the right part is called the torsion part. If $D = F$ is a field, then the torsion part dissappears as a field does not have any non-trivial ideals, that is, any element in F generates F. Hence, if M is an F-module it has the form

$$M \cong D^{\mathcal{B}}.$$

Definition 2.33. Let S be a simplicial complex, and let D be a PID. Then

$$H_p(S; D) \cong (D^r) \oplus \left(\bigoplus_{i=1}^{m} \frac{D}{d_i D}\right)$$

and the p-th *betti number* of S with coefficients in D is r. The context will make clear what space and ground coefficient ring is used, and we will denote the p-th betti number by $\mathcal{B}_p$.

Remark 2.34. *Singular homology groups over any topological space also have this structure, and betti numbers is defined in the same way.*

The torsion part is what recognizes torsion in a space. If the space is without torsion, then the space is fully described by its betti numbers which then is the same for any ring. We may also mention that subcomplexes of $\mathbb{S}^3$ do not have torsion. We can understand the p-th Betti number $\mathcal{B}_k$ as a measure of the p-dimensional connectivity of a space. For a better intuitive notion, let us see what this means in dimensions two and three. In two dimensions $\mathcal{B}_0$ is the number of connected components while $\mathcal{B}_1$ is the number of holes and all higher betti numbers are zero. In three dimensions $\mathcal{B}_0$ is again the number of connected components, $\mathcal{B}_1$ is the number of tunnels or handles, and $\mathcal{B}_2$ is the number of voids.

Example 2.35. In Figure 9 we have three(triangulable) spaces. The figure to the left is two annuluses connected together, the figure in the middle is a torus, and the figure to the right is $\mathbb{S}^2$.

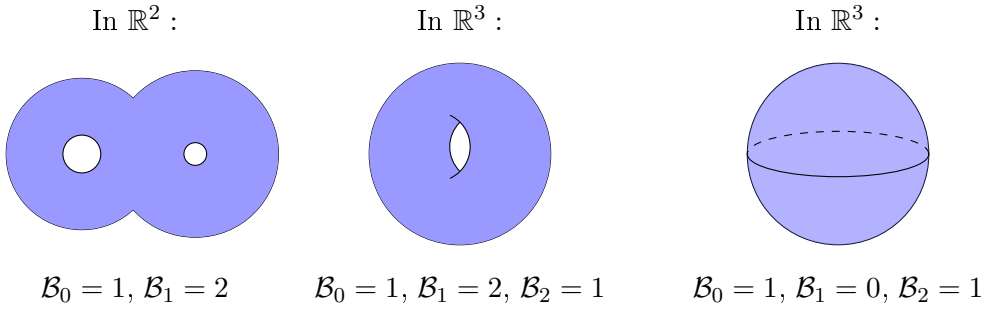


Figure 9: Betti numbers for different topological spaces.

Later on when we introduce persistent homology and barcodes, it is these betti numbers we will use to track features and see how they change and persist over time for a filtered complex.

We may add that even though computing homology over fields does not detect torsion in a space directly, it can be detected indirectly. That is, if we compute homology over a space with two different fields as coefficients and the results are different, then the space has torsion. Let us give an example.

Example 2.36. Let K be the Klein bottle(which is triangulable). The homology of the Klein bottle is as follows.

$$H_0^{simp}(K) = \mathbb{Z}$$
$$H_0^{simp}(K) = \mathbb{Z} \times \mathbb{Z}_2$$
$$H_0^{simp}(K) = 0$$

Now if we compute homology over the fields $\mathbb{Z}_2$ and $\mathbb{Z}_3$, we get the betti numbers,

$$\mathcal{B}_p(K, \mathbb{Z}_2) = \{1, 2, 1\}$$
$$\mathcal{B}_p(K, \mathbb{Z}_3) = \{1, 1, 0\}$$

for $p = 0, 1, 2$ respectively. We see that with coefficients in $\mathbb{Z}_2$ the Klein bottle is misidentified as the torus, but we understand that it is not the torus as the betti numbers over $\mathbb{Z}_3$ gives a different result.

3 The nerve

Material in this section is gathered from [3], [24], and [29].

In this section we will introduce an important construction called the nerve complex of a topological space X, which is construcuted from an open cover of X. It is an abstract simplicial complex, and we will assert that there is a continuous map between the space X and the nerve complex, and furthermore that with a certain type covers it is a homotopy equivalence. This assures us that they induce the same homology. The nerve complex is the basic idea from which we will construct simplicial complexes from point clouds in a hope to create a good model of the underlying space.

In this section let X be a a topological space, and let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be a finite open cover of X.

Definition 3.1. A *p-simplex* σ of $\mathcal{U}$ is an ordered collection of $p+1$ sets $\{U_{\alpha_0}, ..., U_{\alpha_q}\}$ where $U_{\alpha_i} \in \mathcal{U}$ such that $\bigcap_{i=0}^{q} U_{\alpha_i} \neq \emptyset$. This intersection is called the *support* of σ and is denoted $|\sigma|$.

Definition 3.2. The *Nerve of* $\mathcal{U}$ is the collection of p-simplices in $\mathcal{U}$, and we denote it by $\mathcal{N}(\mathcal{U})$.

Definition 3.3. $\mathcal{N}(\mathcal{U})$ is an abstract simplicial complex. And we will call the standard geometric realization of $\mathcal{N}(\mathcal{U})$ for the *nerve complex* of $\mathcal{U}$.

We want to find a connection between X and $\mathcal{N}(\mathcal{U})$, and we will do this via the following construction.

Definition 3.4. Let $\mathcal{U} = \{U_{\alpha \in A}\}$ be a finite open cover of X. Then let

- $\triangle[A]$ denote the standard simplex with vertex set A.
- $\triangle[S] \subseteq \triangle[A]$, where $\emptyset \neq S \subseteq A$ denote the face of $\triangle[A]$ with vertex set S.
- $X[S] = \bigcap_{s \in S} U_s \subseteq X$.

By the *Mayer-Vietoris blowup* of X associated to $\mathcal{U}$, we mean the subspace

$$\mathcal{M}(X,\mathcal{U}) = \bigcup_{\emptyset \neq S \subseteq A} \triangle[S] \times X[S] \subseteq \triangle[A] \times X.$$

We note that there are two natural projection maps $f : \mathcal{M}(X,\mathcal{U}) \to X$ and $g : \mathcal{M}(X,\mathcal{U}) \to \triangle[A]$. Let us first adress the map f.

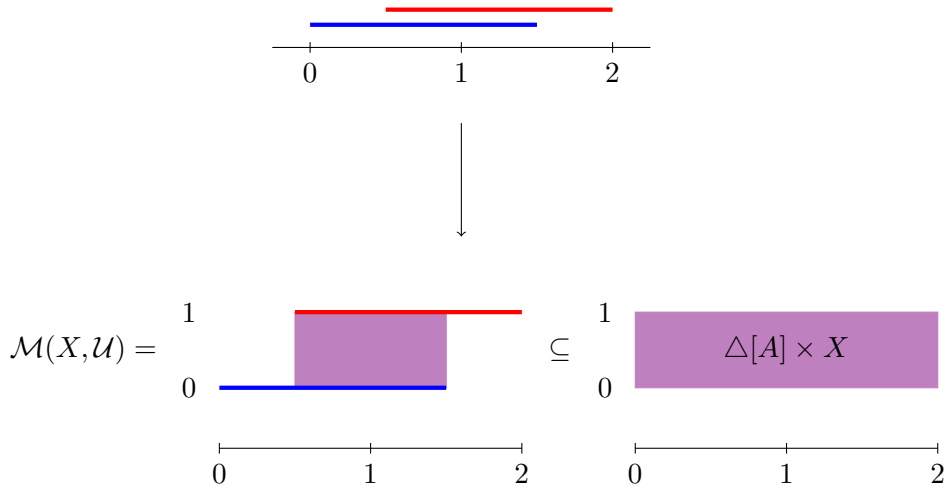


Figure 10: A covering of $X = [0,2]$ by $\mathcal{U} = \{[0,3/2),(1/2,2]\}]\}$ with corresponding Mayer-Vietoris blowup construction $\mathcal{M}(\mathcal{U})$.

Theorem 3.5. *The projection map $f : \mathcal{M}(X,\mathcal{U}) \to X$ is a homotopy equivalence when X has the homotopy type of a finite simplicial complex and the covering consists of open sets. Furthermore, one can obtain an explicit homotopy inverse $\rho : X \to \mathcal{M}(X,\mathcal{U})$ of f.*

Let us see how we can express this homotopy inverse ρ, for that we need the following definition.

Definition 3.6. A *partition of unity subordinate to the finite open covering $\mathcal{U}$* is a family of real valued functions $\{\phi_\alpha\}_{\alpha \in A}$ such that

- $0 \leq \phi_\alpha(x) \leq 1$ for all $\alpha \in A$ and $x \in X$.
- $\sum_{\alpha \in A} \phi_\alpha(x) = 1$ for all $x \in X$.
- The closure of the set $\{x \in X \mid \phi_\alpha(x) \geq 0\}$ is contained in the open set U_α.

Now, assume we are given a partition of unity $\{\phi_\alpha(x)\}_{\alpha \in A}$. Then let $\mathcal{T}(x) \subset A$ be the set of all α so that $x \in U_\alpha$. Next define $\rho : X \to \mathcal{M}(X,\mathcal{U})$ by

$$\rho(x) = (\phi_{\alpha_0}, \phi_{\alpha_1}, ..., \phi_{\alpha_l}) \times \{x\},$$

where $\{\alpha_0(x), \alpha_1(x), ..., \alpha_l(x)\}$ is an enumeration of the set $\mathcal{T}(x)$. The map ρ is easily checked to be continuous. An example is given in Figure 11.

Moving on to the projection map $g : \mathcal{M}(X,\mathcal{U}) \to \triangle[A]$. Let us first define the following.

Definition 3.7. A cover $\mathcal{U}$ such that all the sets $X[S]$ are either empty or contractible is called a *good* cover.

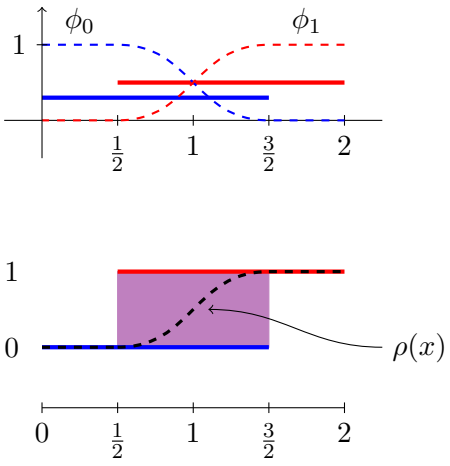


Figure 11: The top figure shows a covering of $X = [0, 2]$ by $\mathcal{U} = \{[0, 3/2), (1/2, 2]\}]\}$ with a partition of unity $\{\phi_0, \phi_1\}$ subordinate to the covering. The bottom figure shows the Mayer-Vietoris blowup construction with the map $\rho(x) : X \to \mathcal{M}(X,\mathcal{U})$ constructed via the partition of unity. We see that $(f \circ \rho) \simeq id_X$ and $(\rho \circ f) \simeq \mathcal{M}(X,\mathcal{U})$, whre $f : \mathcal{M}(X,\mathcal{U}) \to X$ is the projection map. Hence, f is a homotopy equivalence and $X \simeq \mathcal{M}(X,\mathcal{U})$.

Now note that $im(g) = \mathcal{N}(\mathcal{U})$, an illustration is given in Figure 12, and consider the following theorem.

Theorem 3.8. *The projection map $g : \mathcal{M}(X,\mathcal{U}) \to \triangle[A]$ is a homotopy equivalence onto its image, i.e. the nerve complex, if the cover $\mathcal{U}$ is a good cover.*

We now observe that $g \circ \rho : X \to \mathcal{N}(\mathcal{U})$ gives us an explicit map from X to $\mathcal{N}(\mathcal{U})$. Furthermore if we have a good covering, then $g \circ \phi$ is a homotopy equivalence. This gives us the following lemma.

Lemma 3.9. *(Nerve lemma) Let $\mathcal{U}$ be a good cover of a topological space X, then the geomtric realization of $\mathcal{N}(\mathcal{U})$ is homotopy equivalent to X.*

As a consequence $\mathcal{N}(\mathcal{U})$ and X are the same in view of homology theories.

4 Simplicial complex constructions

There are several ways to construct simplicial complexes from point clouds so to approximate the underlying topological space. We will present some important constructions here. Resources have been gathered from [5], [3], [11], [7], and [8].

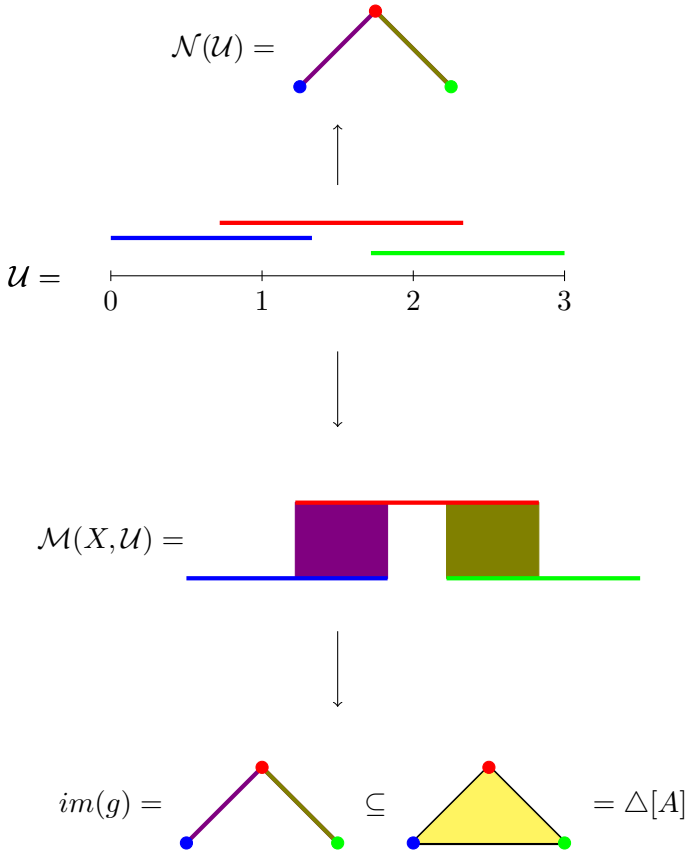


Figure 12: Example of a good cover $\mathcal{U}$ such that g is a homotopy equivalence onto its image, furthermore we see that the image of g is the same as the nerve complex of $\mathcal{U}$.

In this section let X be the underlying topological space that we want to approximate and let $Z \subset X$ be a finite set of sample points from X and assume Z is embedded in some metric space.

4.1 The Čech and Vietoris-Rips complexes

We start by introducing the Čech complex.

Definition 4.1. Given some real number $\epsilon > 0$. Let $\beta_\epsilon(Z) = \{B_\epsilon(z)\}_{z \in Z}$ be the set of open balls under the given metric. The *Čech complex* of Z is then defined to be the nerve of $\beta_\epsilon(Z)$ and will be denoted by $\check{C}(Z, \epsilon)$.

This is a very useful construction. The set $\bigcup_{z \in Z} B_\epsilon(z)$ is a good cover of an open set in X, hence the nerve lemma (Lemma 3.9) tells us that $\check{C}(Z, \epsilon)$ is homotopy equivalent to $\bigcup_{z \in Z} B_\epsilon(z)$. Also, consider the following theorem.

Theorem 4.2. *Let $X = M$ be a compact Riemannian manifold. Then there exists an $e > 0$ such that the nerve of $\beta_\epsilon(Z) = \{B_\epsilon(z)\}_{z \in Z}$ is homotopic equivalent to M whenever*

$\epsilon < e$. Moreover, for all $\epsilon \leq e$ there exists a subset $Z \subset M$ such that $\check{C}(Z, \epsilon)$ is homotopy equivalent to M as well.

As we know homotopy equivalent spaces induce the same homology. The Ĉech complex is however computationally expensive. This is because to check if k+1 vertices forms a p-simplex we have will have to check if the interesction of k+1 ϵ-balls is non-empty. And all of these high dimensional simplexes has to be stored. To avoid this we can construct the following complex.

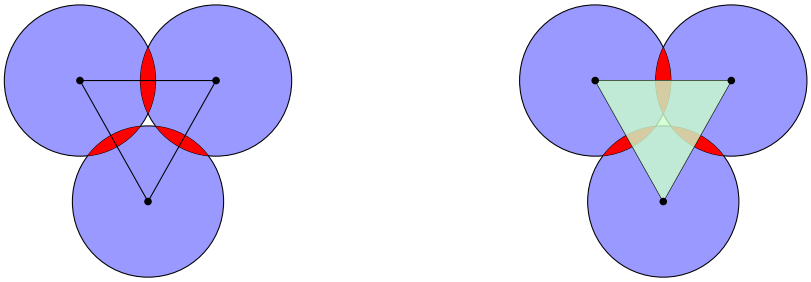
Definition 4.3. Given some $\epsilon > 0$. Let the *Vietoris-Rips complex* for Z denoted by $VR(Z, \epsilon)$ be the complex whose vertex set is Z, and where $\{z_0, z_1, ..., z_k\}$ spans a p-simplex iff $d(z_i, z_j) \leq \epsilon$ for all $0 \leq i, j \leq k$.

In this construction we only need to calculate the distance between pairs of vertices and store this. From this it is straight forward to generate the whole complex.

For the Ĉech and the Vietoris-Rips complex we have inclusions

$$\check{C}(Z, \epsilon) \subseteq VR(Z, 2\epsilon) \subseteq \check{C}(Z, 2\epsilon),$$

which is easy to check. It is possible to find tighter bounds and which also depends on the dimension, but this is enough to make the following point. This bounds gives us some control over the Vietoris Rips complex. Any feature that persist under the inclusion $VR(Z, \epsilon) \hookrightarrow VR(Z, 2\epsilon)$ has to be a topological feature of $C(Z, 2\epsilon)$. This idea of features persisting as such ϵ varies is the key idea in persistent homology, which we will treat later on. Figure 13 illustrates the difference between the two complexes.



$\check{C}(Z, \epsilon)$ $VR(Z, 2\epsilon)$

Figure 13: Illutration of the difference between the Čech complex and the Vietoris-Rips complex.

4.2 Motivating witness complexes

The Čech and Vietoris-Rips complexes are theoretically very nice, but they are using the whole sample set Z as vertex set. This gives us very high dimensional complexes

and makes them computationally very expensive. It would be nice if we could construct complexes with fewer vertices, but somehow still use the information given by the rest of the points in Z, so to still give us a good approximation of the underlying space. An idea is that the points not used as vertices can be used as "witnesses" for the existence of a potential simplex. An illustration of this is in figure 14. To make this idea more tangible we start by introducing Voronoi decomposition of a metric space.

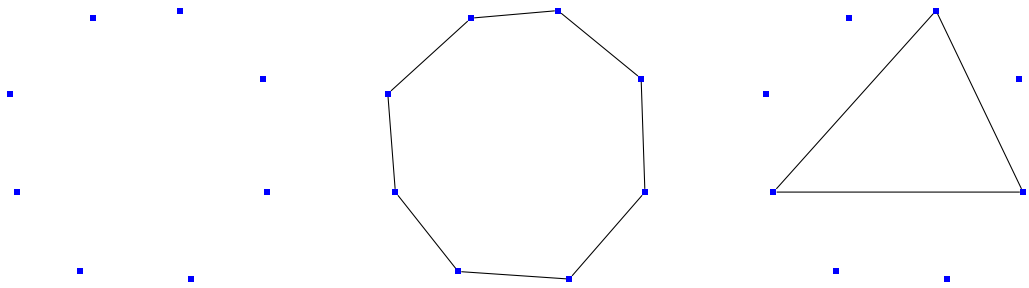


Figure 14: Here we have a set of points lying on a circle. Both figures on the right manage to replicate the same homology of a circle, but the one on the far right does it with fewer points. If we use a bit of imagination we may be able to view the points, not used in the construction, as witnesses for the existence of the edges.

Definition 4.4. Let Y be any metric space. Let $\mathcal{L} \subset Y$ be a finite subset called, the landmark points of Y. We will discuss the selection of these later. Let $l \in \mathcal{L}$ and define the *Voronoi cells* associated to l by

$$V_l = \{y \in Y | d(y,l) \leq d(y,l')\} \text{ for all } l' \in \mathcal{L}.$$

This is called the Voronoi-decomposition of Y with respect to $\mathcal{L}$. The Voronoi cells form a covering of Y, and we call the nerve of this covering for the *Delaunay complex*, denoted by $Del(Y, \mathcal{L})$. An illustration of a Voronoi-decomposition is given in Figure 15.

In the Delaunay complex the points $\{l_0, ..., l_n\} \subseteq \mathcal{L}$ span a p-simplex only if their corresponding Voronoi cells have a point in common, i.e. a "witness". Hence, the Delaunay complex creates very small and low dimensional complexes. If $X \subset \mathbb{R}^n$ is a submanifold of euclidean space and $\mathcal{L} \subset X$ is sampled finely enough then by the nerve lemma (Lemma 3.9) the complex $Del(X, \mathcal{L})$ is homotopy equivalent to X. However, as we want to study finite sets, the Delaunay complex does not seem to give us much. When given a finite point set, the probability that there is a point equidistant between two landmark points equals zero and the Delaunay complex will consist of just the vertices. However, we can make this idea of a witness work with some modification. This leads us to the witness complexes. They can be viewed as approximations to the Delaunay complex.

4.3 Witness complexes

Maybe the most natural modification of the Delaunay complex is to allow points to "move around a little bit"; so to weaken the criteria that witnesses have to be equidistant

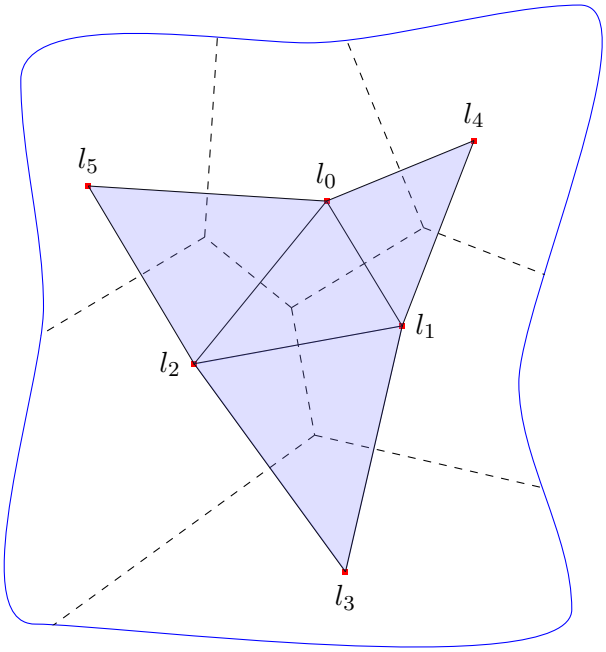


Figure 15: Voronoi-decomposition of a subspace in $\mathbb{R}^2$ with corresponding Delaunay complex.

between landmark points.

Let $\mathcal{L} \subset Z$ be a set of landmark points and let ϵ be some real number greater than zero.

Definition 4.5. For every point $z \in Z \setminus \{l_0, ..., l_p\}$ let m_z denote the minimum distance from z to any point in $\mathcal{L}$. We say that z is an *ϵ-strong witness* for $\{l_0, ..., l_p\}$ iff $d(z, l_i) \leq m_z + \epsilon$ for all $0 \leq i \leq p$. If this holds for $\epsilon = 0$, then z is called a strong witness.

Note that this also allows the remaining landmark points to be witnesses.

Definition 4.6. The *strong witness complex* $W^s(Z, \mathcal{L}, \epsilon)$ is the complex with vertex set $\mathcal{L}$, where $\{l_0, l_1, ..., l_p\} \subseteq \mathcal{L}$ spans a p-simplex iff $\{l_0, ..., l_i\}$ have an ϵ-strong witness for all $i = 0, 1, ..., p$.

Definition 4.7. Let $\{l_0, l_1, ..., l_p\} \subseteq \mathcal{L}$ span a p-simplex iff all the pairs $\{l_i, l_j\}$ has a ϵ-strong witness. This is a Vietoris-Rips version of the strong witness complex, and we denote it by $W^s_{VR}(Z, \mathcal{L}, \epsilon)$.

Remark 4.8. *$W^s(Z, \mathcal{L}, 0)$ is the same complex as $Del(Z, \mathcal{L})$. This supports the notion of the strong witness complex being an approximation of the Delaunay complex.*

Next up is the ϵ-weak witness. It might be less apparent why this witness type gives us an approximation of the delaunay complex, but we will argue that it does.

Definition 4.9. Let $z \in Z \setminus \{l_0, ..., l_p\}$. We say that z is a *ϵ-weak witness* for $\{l_0, ..., l_p\}$ iff $d(z, l) + \epsilon \geq d(z, l_i)$ for all $l \in \mathcal{L} \setminus \{l_0, ..., l_p\}$ and for all $i = 0, 1, ..., p$. If $\epsilon = 0$ then we just call z a weak witness.

Again this allows landmark points to be witnesses.

Definition 4.10. The *weak witness complex* $W^w(Z, \mathcal{L}, \epsilon)$ is the complex with vertex set $\mathcal{L}$, and where $\{l_0, ..., l_p\} \subseteq \mathcal{L}$ spans a p-simplex iff all $\tau \subseteq \{l_0, ..., l_p\}$ have an ϵ-weak witness.

Definition 4.11. Let $\{l_0, l_1, ..., l_p\} \subseteq \mathcal{L}$ span a p-simplex iff all the pairs $\{l_i, l_j\}$ have a ϵ-weak witness. This is a Vietoris-Rips version of the weak witness complex and we denote it by $W^w_{VR}(Z, \mathcal{L}, \epsilon)$.

Definition 4.12. If $\epsilon = 0$ then we will just write $W^w(Z, \mathcal{L})$ and $W^w_{VR}(Z, \mathcal{L})$ for the two weak witness complexes.

Remark 4.13. *The weak witness complex $W^w(Z, \mathcal{L})$ is often called the* strict witness complex.

First we can note that the Delaunay complex can be viewed as the strict witness complex with the additional requirement that the witnesses are equidistant from the landmarks. Secondly we have a theorem telling us that they are the same in $\mathbb{R}^n$.

Theorem 4.14. *Suppose $\mathcal{L} \subset \mathbb{R}^n$ is a finite collection of points. Then $\{l_0, l_1, ..., l_p\} \subset \mathcal{L}$ has a strong witness iff $\{l_0, l_1, ..., l_p\}$ and all its subsets have a weak witness.*

Figure 16 is a diagram illustating that it is plausible that X is homotopy equivalent to $W^w(Z, \mathcal{L})$.

$$X \hookleftarrow \mathrm{Del(X, \mathcal{L})} \hookleftarrow W^w(X, \mathcal{L}) \hookleftarrow W^w(Z, \mathcal{L})$$

Figure 16: Here $X \subset \mathbb{R}^n$ is a submanifold of euclidean space, Z is some reasonable sample of X and $\mathcal{L} \subset \mathcal{Z}$ is some well distributed set of landmark points. Then by the nerve lemma (Lemma 3.9) it is possible that X is homotopy equivalent to $\mathrm{Del(X, \mathcal{L})}$. Next, we may have that $\mathrm{Del(X, \mathcal{L})} = \mathrm{W^w(X, \mathcal{L})}$ which would be similar to theorem 4.14. Finally, it is not unreasonable that Z contains weak witnesses that does the same job as the weak witnesses in X, hence we may have that $W^w(X, \mathcal{L}) = W^w(Z, \mathcal{L})$. Together this makes it atleast plausible that X is homotopy equivalent to $W^w(Z, \mathcal{L})$.

We add one more witness type construction.

Definition 4.15. Let $z \in Z \setminus \{l_0, ..., l_p\}$. We say that z is an ϵ-witness for $\{l_0, ..., l_p\}$ iff $d(z, l_i) \leq \epsilon$ for all $0 \leq i \leq p$. Then $SW(Z, \mathcal{L}, \epsilon)$ will the complex with vertex set L, and where $\{l_0, ..., l_p\}$ spans a p-simplex off all $\tau \subset \{l_0, ..., l_p\}$ have an ϵ-witness. We will call this the simple witness complex.

The motivation for the Simple witness complex is the following relation.

$$W(Z, \mathcal{L}, \epsilon) \subseteq VR(\mathcal{L}, 2\epsilon) \subseteq W(Z, \mathcal{L}, 2\epsilon).$$

Remark 4.16. *The Weak witness complex, Vietoris-Rips weak witness complex, and the Simple witness complex may also be viewed as instances of a group of complex constructions called Lazy witness complexes, where these three are the important ones. The Simple witness complex is just a name given in this paper. See [8].*

In the following figures 17, 18 and 19 we are in $\mathbb{R}^2$, the l_i's, with red dots correspond to landmark points while the $z_i's$, with blue dots represent the rest of the points in the set. The intention is to illustrate the differences between the different complex constructions.

What is good about the witness complexes is that we will be able to control the size of them by choosing fewer or more landmark points, the number of higher dimensional simplices is small given the number of vertices and they are said to give relieable topological estimates when tested empirically. The drawback is that they do not possess the same theoretical tractability that the Ĉech and Vietoris-Rips complexes has.

4.4 Choosing landmark points

There are two ways that are being are used to choose landmark points, the random method and the maxmin method. The maxmin method works as follows. Given a metrix space (X, d), first choose a landmark, $l_1 \in X$ by random. Then, inductively, if $l_1, l_2, ..., l_{i-1}$ have been chosen, let $l_i \in X \setminus \{l_1, l_2, ..., l_{i-1}\}$ be the data point which maximises the function

$$z \mapsto min\{d(z, l_1), d(z, l_2), ..., d(z, l_{i-1})\}.$$

The maxmin method will generally pick out more evenly placed landmarks than the random method, but it has a tendency to pick extremal points as landmarks which is often just noise. The random method will most likely pick points be located in high density areas and not pick extremal points. At any rate, both methods are said to work well.

4.5 Combinatorial delaunay triangulation

We introduce one more complex construction, the combinatorial Delaunay triangulation(CDT). Given Z with a metric, we will then create a new metric on Z. This is done by making a weighted graph from Z. Let $z_0, z_1 \in Z$, then $[z_0 z_1]$ is an edge in the graph

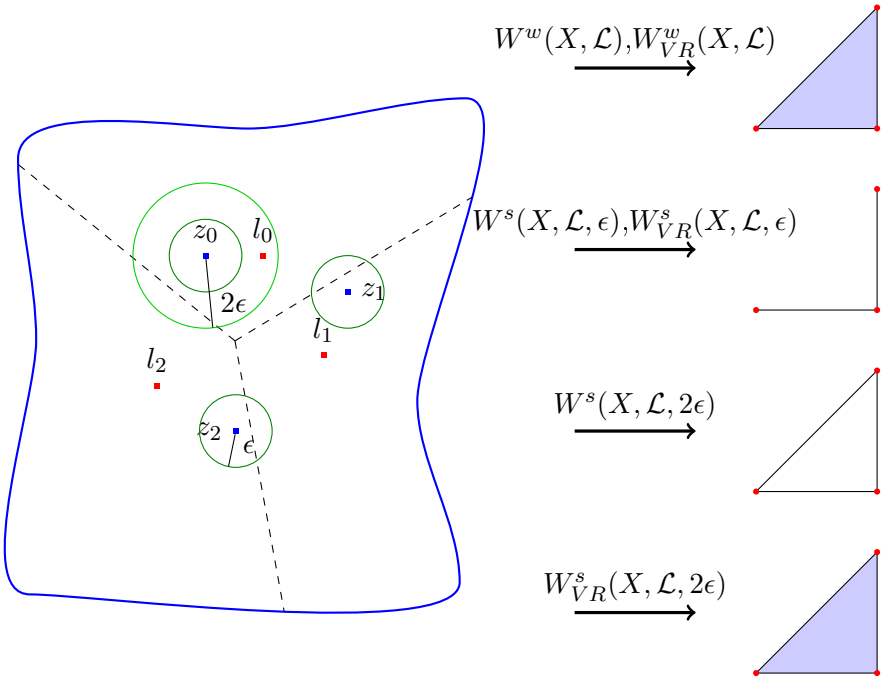


Figure 17: There are no points equidistant between the landmark points. Hence the Delaunay complex consists of just the vertices. Every pair $\{l_i, l_j\}$ of landmark points have a weak witness, and since $\{l_0, l_1, l_2\}$ are the only landmark points the triple also has a weak witness. This results in a 2-simplex for the two weak witness types. For some $\epsilon > 0$ the points z_1 and z_2 becomes ϵ-strong witnesses for the pairs $\{l_0, l_1\}$ and $\{l_1, l_2\}$ respectively, resulting in a simplicial complex with two connected edges for the ϵ-strong witness types. If we increase to 2ϵ we get a an ϵ-strong witness ("2ϵ-strong witness") for $\{l_0, l_2\}$ as well

.

iff z_1 is one of the k nearest neighbors of z_0 and visa versa. Alternatively, we can say that $[z_0 z_1]$ is an edge iff $d(z_0, z_1) < \epsilon$ for some real number $\epsilon > 0$. Then we assign the weight of each edge $[z_0 z_1]$ as the distance $d(z_0, z_1)$. This generates a weighted graph G. G defines a new metric d_G by considering the path of least weight(shortest path). We can now define the CDT.

Definition 4.17.

$$CDT((X,\mathcal{L}) = W^w((Z, d_G), \mathcal{L})$$
$$CDT_{VR}((X,\mathcal{L}) = W^w_{VR}((Z, d_G), \mathcal{L})$$

where d_G is as bove.

The advantage of this complex is that due to the construction of the graph G it copes better to curvature in the sample set. That is, the CDT creates better models of spaces with a lot of curvature. For example, consider the letter "U". What is the distance

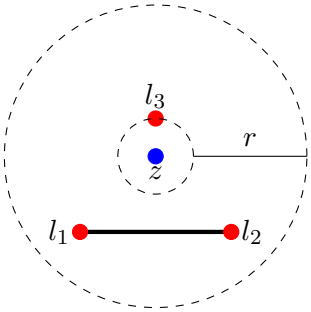


Figure 18: In this figure z is not an ϵ-weak witness for the edge between l_1 and l_2 if $\epsilon = 0$, due that l_3 is closest to z. The landmark point l_3 is however a witness for the edge. If ϵ is increased enough, say to r, then z becomes a weak witness for the edge as well.

between the two top most points? If we have to move along the "U", then this will be much longer than the actual distance on the paper. This is what the CDT takes into account.

5 Some algebra

Material in this section is gathered from [1], [18], and [2].

In this section we will go through the algebraic theory we will need when we introduce persistent homology in the next section.

We will always assume that R is a commutative ring with identity. Hence we will not talk about left or right modules as they coincide, and just say modules.

5.1 Graded rings and modules

We start by introducing the algebraic structures that we will be needing.

Definition 5.1. A ring R is called *graded* if there exists a family of subgroups $\{R_n\}_{n\in\mathbb{Z}}$ of R such that

1. $R = \bigoplus_n R_n$ (as abelian groups).
2. $R_n \cdot R_m \subseteq R_{n+m}$ for all n, m.

A graded ring is called *non-negatively graded* or $\mathbb{N}$-graded if $R_n = 0$ for all $n \leq 0$. A non-zero element $x \in R_n$ is called a *homogeneous* element of R of degree n.

Definition 5.2. 1. Any ring can be made into a graded ring by setting $R_0 = R$ and $R_n = 0$ for all $n \neq 0$. This is called the *trivial grading*.

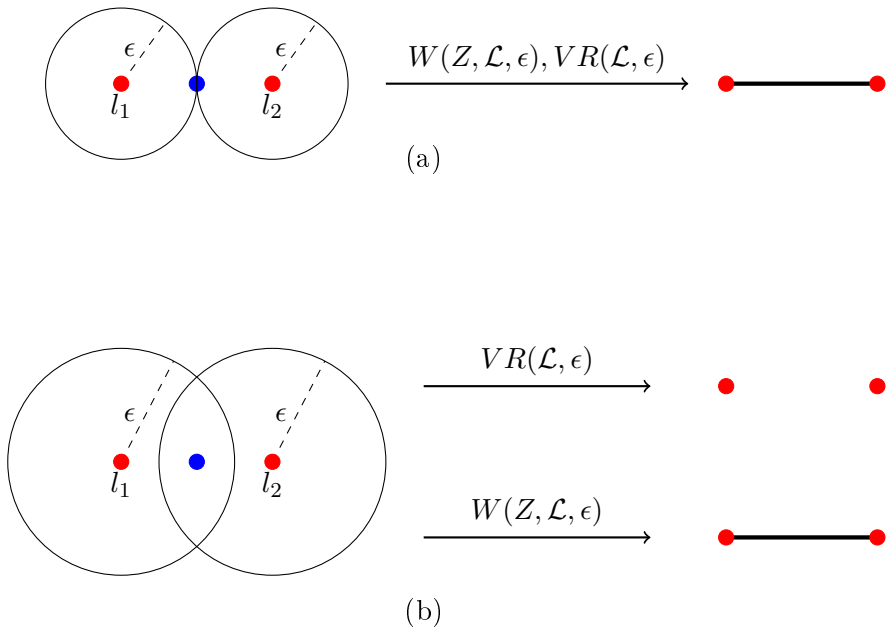


Figure 19: Motivating the relation between the ϵ-witness complex and the Vietoris-Rips complex. (a) The two complexes are the same, but if the blue witness point is perturbed slightly then $W(Z, \mathcal{L}, \epsilon) \leq VR(\mathcal{L}, \epsilon)$. (b) $W(Z, \mathcal{L}, \epsilon)$ is potentiallly larger than $VR(\mathcal{L}, \epsilon)$ since it can have a witness, as in this example.

2. Let $S = R[x_1, ..., x_d]$ be the polynomial ring over R with d indeterminates, and let $m = \{m_1, m_2, ..., m_d\}$ and $X^m = \{x_1^{m_1}, x_2^{m_2}, ..., x_d^{m_d}\}$. Then we can assign a grading to S by setting

$$S_n = \{ \sum_{m \in \mathbb{N}^d} r_m x^m \mid r_m \in R \text{ and } m_1 + \cdots + m_d = n\}.$$

This is called the *standard grading* on the polynomial ring $R[x_1, ..., x_d]$, and it is a nonnegative grading.

Remark 5.3. *Particularly we have that the standard grading on* $R[t]$ *is* $R[t]_n = \{rt^n \mid r \in R\}$. *If we set* $1 = (1, 0, 0, ...)$, $t = (0, 1, 0, ...)$, $t^2 = (0, 0, 1, 0, ...)$ *etc., then we can write* $a_0 + a_1 t + a_2 t^2 + \cdots a_n t^n \in R[t]$ *as* $(a_0, a_1, a_2..., a_n)$.

Definition 5.4. Let $R = \bigoplus_n R_n$ be a graded ring. A subring S of R is called a *graded subring* of R if $S = \sum_n (R_n \cap S)$. Equivalently, a subring S is graded if for every element $f \in S$, all the homogeneous components of f regarded as an element in R, are in S.

Example 5.5. Let $R[t]$ be the polynomial ring with the standard grading. Now $R[t] = \{a_0 + a_1 t + a_2 t^2 + \cdots\} = \{(a_0, a_1, a_2, ...)\}$ where $a_i \in R$. Then $t^n R[t] = \{t^n(a_0 + a_1 t + a_2 t^2 + \cdots)\} = \{(0, ..., 0, a_0, a_1, a_2, ...)\}$, with zeros in the n first positions. Consequently

we have that

$$(t^n R[t])_k = \begin{cases} 0 & \text{if } 0 \leq k \leq n \\ t^k R & \text{if } k \geq n \end{cases}$$

where $t^k R = \{t^k r \mid r \in R\}$.

Definition 5.6. Let R be a graded ring and M an R-module. We say that M is a graded R-module if there exists a family of subgroups $\{M_n\}_{n\in\mathbb{Z}}$ of M such that

1. $M = \bigoplus M_n$ (as abelian groups).
2. $R_n \cdot M_m \subseteq M_{n+m}$ for all n, m.

A nonzero element $x \in M_n$ is called a *homogeneous* element of M of degree n.

Example 5.7. 1. If R is a graded ring, then R is a graded R-module.

2. If $\{M_\lambda\}$ is a family of graded R-modules, then $\bigoplus_\lambda M_\lambda$ is a graded R- module.

Remark 5.8. *Recall that if $\{R_i\}_{i=1}^n$ is a finite family of rings, then the notions of direct sum and direct product coincide. Hence we will freely change between the two.*

Definition 5.9. Let M be a graded R-module, and let $n \in \mathbb{N}$. Then we can define a new graded R-module $M(n)$ by setting $M(n)_k = M_{n+k}$. We say that M is shifted by n.

Example 5.10. Let $R[t]$ be the polynomial ring with the standard grading and we can view it as a graded $R[t]$-module. We then have that $R[t](n)_k = (t^n R[t])_k$.

Next, let N be a submodule of a graded R-module M. We can assign a grading to N by setting $N_i = N \cap M_i$. We give the following definition.

Definition 5.11. A submodule N, of a graded R-module M, is said to be a *graded submodule* of M if its grading is defined by $N_i = N \cap M_i$.

Example 5.12. Equip $R[t]$ with the standard grading and let $M = t^{\alpha_1} R[t] \oplus t^{\alpha_2} R[t] \oplus \cdots \oplus t^{\alpha_n} R[t]$, then $N = t^{\beta_1}(t^{\alpha_1} R[t]) \oplus t^{\beta_2}(t^{\alpha_2} R[t]) \oplus \cdots \oplus t^{\beta_n}(t^{\alpha_n} R[t])$ is a graded submodule of M.

Proposition 5.13. *Let R be a graded ring, M a graded R-module and N a graded submodule of M. Then M/N is a graded R-module, where*

$$(M/N)_n = (M_n + N)/N = \{m \in M_n\}.$$

Proof. Proof is found in [18].

Definition 5.14. The graded R-module M/N is called the *graded quotient module* of M modulo N.

Example 5.15. Equip $R[t]$ with the standard grading then $t^\alpha R[t]$ is a graded R-submodule of $R[t]$, and we have that

$$\frac{R[t]}{t^\alpha R[t]} = \{a_0[1], +a_1[t] + a_2[t]^2 + \cdots\},$$

where the equivalence classes $[t^n] = [t^m]$ iff $m - n = r\alpha$, for $n, m, r \in \mathbb{Z}$.

Definition 5.16. Let R be a graded ring and M, N graded R-modules. Let $f : M \to N$ be an R-homomorphism. Then f is said to be *graded of degree* d if $f(M_n) \subseteq N_{n+d}$ for all n.

Example 5.17. Let R be a graded ring, M a graded R-module, and $r \in R_d$. Define $\mu_r : M \mapsto M$ by $\mu_r m = rm$ for all $m \in M$. Then μ_r is a graded R-homomorphism of degree d.

Definition 5.18. Let M and N be two graded R-modules. We say that M and N are *isomorphic as graded R-modules* if there exists a degree zero isomorphism from M to N and we denote this by $M \cong_0 N$.

Remark 5.19. *Notice that*

$$\bigoplus_{i=1}^{n} R[t] \cong t^{\alpha_1} R[t] \oplus t^{\alpha_2} R[t] \oplus \cdots \oplus t^{\alpha_n} R[t],$$

but that

$$\bigoplus_{i=1}^{n} R[t] \ncong_0 t^{\alpha_1} R[t] \oplus t^{\alpha_2} R[t] \oplus \cdots \oplus t^{\alpha_n} R[t],$$

unless $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Proposition 5.20. *Let M, N be graded R-modules. If $f : M \to N$ is a graded R-homomorphism, then*

- *$ker(f)$ is a graded submodule of M, and*
- *$im(f)$ is a graded submodule of N.*

Proposition 5.21. *If $C = \{C_*, \partial_*\}$ is a chain complex of graded R-modules with graded maps. Then the homology R-modules $H_i(C)$ are graded for all i.*

Theorem 5.22. *(fundamental theorem of [graded] R-homomorphisms). Let M and N be graded R-modules and let $f : M \to N$ be a degree 0 R-homomorphism. Then $M/ker(f) \cong_0 f(M)$.*
Proof. This is proved in exactly the same manner as for the normal fundamental theorem of R-homomorphisms by adding the fact that f is of degree 0.

Theorem 5.23. *Let A and B be graded R-submodules of graded R-modules M and N, respectively. Then*

$$\frac{M \times N}{A \times B} \cong_0 \frac{M}{A} \times \frac{N}{B}.$$

Proof. Define a mapping

$$f : M \times N \to \frac{M}{A} \times \frac{N}{B}$$

by $f(m,n) = (m+A, n+B), m \in M, n \in N$, and f is clearly onto and of degree 0. Now

$$\begin{aligned} ker(f) &= \{(m,n) \mid (m+A, n+B) = (0+A, 0+B)\} \\ &= \{(m,n) \mid m \in A, n \in B\} \\ &A \times B. \end{aligned}$$

Hence, by theorem 5.22 we have

$$\frac{M \times N}{A \times B} \cong_0 \frac{M}{A} \times \frac{N}{B}.$$

Definition 5.24. A finite sequence $x_1, .., x_n$ of elements of an R-module M is called linearly dependent if, for any $a_1, ..., a_n \in R, \sum_{i=1}^{n} a_i x_i = 0$ implies $a_1 = a_2 = \cdots = a_n = 0$. A finite sequence is called *linearly independent* if it is not linearly independent.

Now recall the following for non-graded modules.

Definition 5.25. A subset B of an R-module M is called a basis if

1. M is generated by B.
2. B is a linearly independent set.

Definition 5.26. Suppose R is a ring with identity, and let $M = R^n$ be an R-module. Now let $e_i = (0, ..., 1, 0, ..., 0)$, be the n-tuple where all the components are zero except the ith component which is 1. Then $\{e_1, e_2, ..., e_n\}$ is a basis for R^n, and we call it the *standard basis* for R^n.

Definition 5.27. An R-module is called a *free* R-module if M admits a basis.

Theorem 5.28. *Let M be a free R-module with a basis $\{e_1, ..., e_n\}$. Then $M \cong R^n$.*

Lemma 5.29. *Let R be a PID, and let M be a free R-module with a basis of m elements. Then any submodule N of M is also free with a basis consisting of n elements, where $n \leq m$.*

For graded modules we have the following.

Theorem 5.30. *If M is a graded free R-module, then there exists a unique set of integers $\{n_1, n_2, ..., n_k\}$ such that $M \cong_0 \bigoplus_{i=1}^{k} R(n_i)$.*

Proof. See [18].

Example 5.31. Let $R[t]$ be graded with the standard grading, and let M be a graded free $R[t]$-module. Then

$$M \cong_0 t^{\alpha_1}R[t] \oplus t^{\alpha_2}R[t] \oplus \cdots \oplus t^{\alpha_n}R[t], \text{ where } \alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{N}.$$

Moreover, it has a basis consisting of homogeneous elements $\{e_1, e_2, ..., e_n\}$ where $e_i = (0, ..., t^{\alpha_i}, 0, ..., 0)$ with all zeros except at position i. We will call this the *standard basis* of a graded free $R[t]$-module.

Lemma 5.32. *Let R be a graded PID, and let M be a graded free R-module with a basis of m elements. Then any graded submodule N of M is also free with a basis consisting of n elements, where $n \leq m$.*
Proof. See [1] for non-graded case.

Remark 5.33. *Particularly recall that if F is a field, then $F[t]$ is a PID.*

Finally, the structure for all finitely generated graded modules over a graded PID is given by the following theorem.

Theorem 5.34. *Every finitely generated graded module over a graded PID D can be uniquely written as*

$$\left(\bigoplus_{i=1}^{n} D(\alpha_i)\right) \oplus \left(\bigoplus_{j=1}^{m} \frac{D}{d_j D}(\gamma_j)\right) \text{ where } d_j | d_{j+1}.$$

Particularly, if $D = F[t]$, then the module can be written as

$$\left(\bigoplus_{i=1}^{n} t^{\alpha_i}F[t]\right) \oplus \left(\bigoplus_{j=1}^{m} t^{\gamma_j}\frac{F[t]}{t^{\beta_j}F[t]}\right) \text{ where } t^{\mathcal{B}_j} | t^{\mathcal{B}_{j+1}}.$$

Proof. The proof of this theorem is the same as for the non-graded case, one just have to make sure that all the isomorphisms in the proof are graded isomorphisms.

5.2 Matrix representations

When we introduce persistent homology in the next section, we will create chain complexes consisting of graded free $F[t]$-modules, and the maps will be graded $F[t]$-homomorphisms of degree 0. To compute the homology modules we will express the boundary maps as matrices. These matrices will depend on the bases we choose for the graded free chain modules. So, in this section we will show how to obtain such matrix representations and how such matrices with respect to different bases are related. We will assume that all polynomials $R[t]$ are given the standard grading.

Notation 5.35. *Let*

$$(a_1, a_2, ..., a_m) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

Definition 5.36. Let M be a finitely generated free module over R, and let $\mathcal{B} = \{e_1, e_2, ..., e_m\}$ be any basis for M. If $g \in M$ we can express it as a unique linear combination of the basis elements, i.e. $g = \sum_{i=1}^{m} b_i e_i where b_i \in R$. Furthermore, define $g_{\mathcal{B}} = (b_1, b_2, ..., b_m)$ to be the element g *relative* to the basis $\mathcal{B}$. We will call the collection $M_{\mathcal{B}} = \{g_{\mathcal{B}} \mid g \in M\}$ for the free R-module M relative to the basis $\mathcal{B}$, and note that $M_{\mathcal{B}} = R^m$.

Notation 5.37. *If $\mathcal{B} = \{e_1, e_2, ..., e_m\}$ is a basis for a finitely generated free R-module M, then we will write $[\mathcal{B}] = [e_1, e_2, ..., e_m]$, where the e_i's expand in the columns, so it is an $m \times m$ matrix.*

Remark 5.38. *If $g \in M$ and $\mathcal{B}$ is a basis for M, then $[\mathcal{B}]g_{\mathcal{B}} = g$.*

Example 5.39. Let $M = t^{\alpha_1}R[t] \oplus t^{\alpha_2}R[t]$ be a graded free $R[t]$-module, and let $\mathcal{B} = \{(t^{\alpha_1}, 0), (0, t^{\alpha_2})\}$ and $\mathcal{C} = \{(t^{\alpha_1}, t^{\alpha_2}), (0, t^{\alpha_2})\}$ be two bases on M.

Now, suppose that $g = (t^{\alpha_1}, t^{\alpha_2}) \in M$. Then,

$$g = 1 \cdot (t^{\alpha_1}, 0) + 1 \cdot (0, t^{\alpha_2}),$$

which gives us that $g_{\mathcal{B}} = (1, 1)$. Also, we have that

$$g = 1 \cdot (t^{\alpha_1}, t^{\alpha_2}) + 0 \cdot (0, t^{\alpha_2}),$$

so $g_{\mathcal{C}} = (1, 0)$.

Next, suppose that $h = (a_1 t^{\beta_1}, a_2 t^{\beta_2}) \in M$, where $a_1, a_2 \in R$ and $\mathcal{B}_i \in \mathbb{N}$ and $\mathcal{A}_i \leq \mathcal{B}_i$, for $i = 1, 2$. Then

$$h = a_1 t^{\beta_1 - \alpha_1} \cdot (t^{\alpha_1}, 0) + a_2 t^{\beta_2 - \alpha_2} \cdot (0, t^{\alpha_2}),$$

so $h_{\mathcal{B}} = (a_1 t^{\beta_1 - \alpha_1}, a_2 t^{\beta_2 - \alpha_2})$. Also, we have that

$$h = a_1 t^{\beta_1 - \alpha_1} \cdot (t^{\alpha_1}, t^{\alpha_2}) + (a_2 t^{\beta_2 - \alpha_2} - a_1 t^{\beta_1 - \alpha_1}) \cdot (0, t^{\alpha_2}),$$

so $h_{\mathcal{B}} = (a_1 t^{\beta_1 - \alpha_1}, a_2 t^{\beta_2 - \alpha_2} - a_1 t^{\beta_1 - \alpha_1})$.

Definition 5.40. Let M and N be finitely generated free R-modules with bases $\mathcal{B} = (e_1, e_2, ..., e_m)$ and $\mathcal{C} = (f_1, f_2, ..., f_n)$ respectively. Let $\phi : N \to M$ be an R-homomorphism given by

$$\phi(f_j) = \sum_{i=1}^{m} a_{ij} e_i,$$

where $j = 1, ..., n$, and $a_{ij} \in R$. Then the matrix $A = (a_{ij})$ is called the matrix representation of ϕ relative to the bases $\mathcal{B}$ and $\mathcal{C}$. Furthermore we can define a R-homomorphism $\mu_A : N_{\mathcal{C}} \to M_{\mathcal{B}}$ by setting $\mu_A(g_{\mathcal{C}}) = A g_{\mathcal{C}}$.

Remark 5.41. *With the above definition we have that* $\phi(g) = [\mathcal{B}]Ag_{\mathcal{C}}$*, and also that*

$$g \in ker(\phi) \Leftrightarrow g_{\mathcal{C}} \in ker(\mu_A) = \text{ null space of } A, \text{and}$$
$$f \in im(\phi) \Leftrightarrow f_{\mathcal{B}} \in im(\mu_A) = \text{ column space } A.$$

Definition 5.42. If ϕ is as in Definition 5.40, then the dimension of $im(\phi)$ is called the rank of ϕ.

Definition 5.43. Let $\mathcal{B} = \{e_1, e_2, ..., e_m\}$ and $\mathcal{B}' = \{e'_1, e'_2, ..., e'_m\}$ be two ordered bases of M. Then we can write each e'_j as a linear combination of e_j's and visa versa. So we can write

$$e'_j = \sum_{i=1}^{m} p_{ij} e_i,$$

where $j = 1, ..., m$, and $p_{ij} \in R$. The $m \times m$ matrix $P = (p_{ij})$ is called the *matrix of transformation from* $\mathcal{B}'$ *to* $\mathcal{B}$.

Remark 5.44. *If* P *is a matrix transformation from* $\mathcal{B}'$ *to* $\mathcal{B}$ *and if* Q *is a matrix transformation from* $\mathcal{B}$ *to* $\mathcal{B}'$*, then* $PQ = I = QP$*, which also implies that* P *and* Q *are invertible matrices.*

Definition 5.45. Let M and N be two finitely generated free modules over R, and let $\phi : N \to M$ be an R-homomorphism. Then the matrix representation of ϕ relative to the standard bases of M and N will be called the *standard matrix representation.* This goes for the graded case as well.

Theorem 5.46. *Let* M *and* N *be finitely generated free modules over* R*. Let* $A = (a_{ij})$ *be a matrix representing an* R*-homomorphism* $\phi : N \to M$ *relative to bases* $\mathcal{B} = \{e_1, e_2, ..., e_m\}$ *and* $\mathcal{C} = \{f_1, f_2, ..., f_n\}$ *of* M *and* N*, respectively. Then we have the following*

1. *The matrix of* ϕ *relative to a new pair of bases* $\mathcal{B}' = \{e'_1, e'_2, ..., e'_m\}$ *and* $\mathcal{C}' = \{f'_1, f'_2, ..., f'_n\}$ *of* M *and* N*, respectively, is* $P^{-1}AQ$*, where* P *and* Q *are the matrices of transformation from* $\mathcal{B}'$ *to* $\mathcal{B}$ *and* $\mathcal{C}'$ *to* $\mathcal{C}$*, respectively.*

2. *For any pair of invertible matrices* $P \in R^{m \times m}$ *and* $Q \in R^{n \times n}$*, there exists bases* $\mathcal{B}'$ *of* M *and* $\mathcal{C}'$ *of* N*, such that the matrix representation of* ϕ *relative to* $\mathcal{B}'$ *and* $\mathcal{C}'$ *is* $P^{-1}AQ$*.*

Remark 5.47. *Suppose conditions are as in the theorem above, and that* $g \in N$*. Then we have the following*

$$\begin{aligned} \phi(g) &= [\mathcal{B}]Ag_{\mathcal{C}} \\ &= [\mathcal{B}']A'g_{\mathcal{C}'} \\ &= [\mathcal{B}']P^{-1}AQg_{\mathcal{C}'}. \end{aligned}$$

Particularly $im(\phi) = \phi(N) = [\mathcal{B}]AN_{\mathcal{C}} = [\mathcal{B}']P^{-1}AQR^n$.

Example 5.48. Suppose $M = R[t] \oplus R[t] \oplus tR[t]$ and $N = tR[t] \oplus tR[t] \oplus t^3R[t]$ are two graded free $R[t]$-modules. Furthermore let $\mathcal{B} = \{e_1, e_2, te_3\}$ and $\mathcal{C} = \{te_1, te_2, t^3e_3\}$ be bases for M and N, respectively, where $\{e_1, e_2, e_3\}$ denotes the standard basis for R^3. Now, suppose we have an $R[t]$-homomorphism $\phi : N \to M$ which is given by $\phi(a, b, c) = (a + c, a + b, b + c)$. Then

$$\begin{aligned}
\phi(te_1) &= \phi(t, 0, 0) = (t, t, 0) = t \cdot e_1 + t \cdot e_2 + 0 \cdot te_3 \\
\phi(te_2) &= \phi(0, t, 0) = (0, t, t) = 0 \cdot e_1 + t \cdot e_2 + 1 \cdot te_3 \\
\phi(t^3e_3) &= \phi(0, 0, t^3) = (t^3, 0, t^3) = t^3 \cdot e_1 + 0 \cdot e_2 + t^2 \cdot te_3.
\end{aligned}$$

Hence, the matrix representation of ϕ with respect to $\mathcal{B}$ and $\mathcal{C}$ is

$$A = \left[\begin{array}{c|ccc} & te_1 & te_2 & t^3e_3 \\ \hline e_1 & t & 0 & t^3 \\ e_2 & t & t & 0 \\ te_3 & 0 & 1 & t^2 \end{array}\right].$$

Now, suppose that $g = (at^2, bt^4, ct^3)$. We then have that

$$g = at \cdot (t, 0, 0) + bt^3 \cdot (0, t, 0) + c \cdot (0, 0, t^3),$$

so $g_{\mathcal{C}} = (at, bt^3, c)$. This means that

$$[\mathcal{B}]Ag_{\mathcal{C}} = \begin{bmatrix} 1 & 0 & t^3 \\ 0 & 1 & 0 \\ 0 & 0 & t \end{bmatrix} \begin{bmatrix} t & 0 & t^3 \\ t & t & 0 \\ 0 & 1 & t^2 \end{bmatrix} \begin{bmatrix} at \\ bt^3 \\ c \end{bmatrix}.$$

Hence

$$[\mathcal{B}]Ag_{\mathcal{C}} = (at^2 + ct^3, at^2 + bt^4, bt^4 + ct^3) = \phi(g).$$

Next, suppose that $\mathcal{B}' = \mathcal{B}$ and $\mathcal{C}' = \{te_1, te_2, t^3e_3 - t^3e_2 - t^3e_1\}$ are alternative bases for M and N. Then

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is the matrix transformation from $\mathcal{B}'$ to $\mathcal{B}$, and

$$Q = \begin{bmatrix} 1 & 0 & -t^2 \\ 0 & 1 & -t^2 \\ 0 & 0 & 1 \end{bmatrix}$$

is the matrix transformation from $\mathcal{C}'$ to $\mathcal{C}$. Now

$$g = (at + ct^2) \cdot (t, 0, 0) + (bt^3 + ct^2) \cdot (0, t, 0) + c \cdot (-t^2, -t^2, t^3),$$

so $g_{\mathcal{C}'} = (at + ct^2, bt^3 + ct^2, c)$. Finally, this gives us that

$$[\mathcal{B}']P^{-1}AQg_{\mathcal{C}'} = (at^2 + ct^3, at^2 + bt^4, bt^4 + ct^3) = \phi(g) = [\mathcal{B}]Ag_{\mathcal{C}}.$$

5.3 Basis changes by elementary matrix operations

We now know how to obtain matrix representations of our alleged boundary maps. Next, via elementary matrix operations we can put the matrix representations in a form called *Smith normal form*, from which it will be easy to deduce the homology modules. Each elementary matrix operation corresponds to multiplying the matrix on the left or right with an invertible matrix. So, in this section we will go through how to keep track of the basis changes as we put our matrix representation into Smith normal form.

Let D be a PID, and specifically recall that $F[t]$, where F is a field, is a PID.

Definition 5.49. Let A be an $m \times n$ matrix over D. The following three types of operations are called elementary row (column) operations.

1. Swap row i (column i) with row j (column j).

2. Multiply a row (column) with a nonzero invertible element $q \in D$.

3. Add q times row j (column j) to row i (column i), where $i \neq j$ and $q \in D$.

Notation 5.50. *Let e_{ij} be the $n \times n$ matrix with zeros everywhere except at position $i \times j$ where it is* 1. *Also let I denote the identity matrix.*

These operations can be identified with multiplying A with the following invertible matrices.

Theorem 5.51. *Let A be an $m \times n$ matrix over D.*

1. *Let $E_{ij} = I - e_{ii} - e_{jj} + e_{ij} + e_{ji}$, where $i \neq j$. Then $E_{ij}A(AE_{ij})$ amounts to doing elementary row (column) operation of type 1. Also note that $E_{ij}^{-1} = E_{ij}$.*

2. *Let $L_i(q) = I + (q-1)e_{ii}$, where $q \in D$ is invertible. Then $L_i(q)A(AL_i(q))$ amounts to doing elementary row (column) operation of type 2. Also note that $L_i(q)^{-1} = I(q^{-1} - 1)e_{ii}$.*

3. *If $M_{ij}(q) = I + qe_{ij}$, where $i \neq j$ and $q \in R$. Then $M_{ij}(q)A(AM_{ji}(q))$ amounts to doing elementary row (column) operation of type 3. Also note that $M_{ij}(q)^{-1} = I - qe_{ij}$.*

Definition 5.52. Two $m \times n$ matrices A and B over D are said to be equivalent if there exists an invertible matrix $P \in R^{m \times m}$ and $Q \in R^{n \times n}$ such that $B = PAQ$.

Theorem 5.53. *If A is a an $m \times n$ matrix over a PID. Then A is equivalent to a matrix that has the diagonal form*

$$\begin{bmatrix} a_1 & & & & & & \\ & a_2 & & & & & \\ & & \ddots & & & & \\ & & & a_k & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

where $a_i \neq 0$ *and* $a_1 \mid a_2 \mid \cdots \mid a_n$. *The nonzero diagonal elements* $a_1, ..., a_n$ *are called the* invariant factors *of* A.

Suppose now that M and N are finitely generated free D-modules and that $\phi : N \to M$ is a D-homomorphism. Let $A = (a_{ij})$ be the representation of ϕ relative to some bases $\mathcal{B} = \{e_1, e_2, ..., e_m\}$ and $\mathcal{C} = \{f_1, f_2, ..., f_n\}$ for M and N, respectively. Then $A \in D^{m \times n}$, so we can reduce it to the Smith normal form. Now, all of the elementary matrices, E_{ij}, $L_i(q)$, and $M_{ij}(q)$, are invertible, consequently any composition of them is invertible, and reducing A will result in a new representation matrix for ϕ relative to some new bases. So, let us explore what the basis changes are when we do elementary row and column operations.

We start with elementary row operations. Multiplication on the left by some invertible matrix amounts to a change in the basis for M from $\mathcal{B}$ to some new basis $\mathcal{B}'$.

1. Swap row i with row j. This is the same as the multpiplication $E_{ij}A$. The matrix of transformation from $\mathcal{B}'$ to $\mathcal{B}$ is then $E_{ij}^{-1} = E_{ij} = I - e_{ii} - e_{jj} + e_{ji}$. Consequently we have the following change in basis:

$$\begin{aligned} e'_i &= e_j \\ e'_j &= e_i \\ e'_l &= e_l \text{ for all } l \neq i, j. \end{aligned}$$

2. Multiply row i with some invertible element $q \in D$. This is the same as the multiplication $L_i(q)A$. The matrix of transformation from $\mathcal{B}'$ to $\mathcal{B}$ is then $L_i(q)^{-1} = I(q^{-1} - 1)e_{ii}$. Consequently we have the following change in basis:

$$\begin{aligned} e'_i &= q^{-1}e_i \\ e'_l &= e_l \text{ for all } l \neq i. \end{aligned}$$

3. Adding q times row j to row i, $q \in D$. This is the same as the multiplication $M_{ij}(q)A$. The matrix of transformation from $\mathcal{B}'$ to $\mathcal{B}$ is then $M_{ij}(q)^{-1} = I - qe_{ij}$. Consequently we have the following change in basis:

$$\begin{aligned} e'_j &= e_j - qe_i \\ e'_l &= e_l \text{ for all } l \neq j. \end{aligned} \tag{2}$$

Going about in the same fashion we will get the following basis changes for column operations.

1. Swap column i with column j. This gives the following change in basis:

$$\begin{aligned} f_i' &= f_j \\ f_j' &= f_i \\ f_l' &= f_l \text{ for all } l \neq i, j. \end{aligned}$$

2. Multiply column i with some invertible element $q \in D$. This gives the following change in basis:

$$\begin{aligned} f_i' &= q f_i \\ f_l' &= f_l \text{ for all } l \neq i. \end{aligned}$$

3. Adding q times column j to column i, $q \in D$. This gives the following change in basis:

$$\begin{aligned} f_i' &= f_i + q f_j \\ f_l' &= f_l \text{ for all } l \neq j. \end{aligned} \tag{3}$$

What if $D = F[t]$ is a graded ring over a field F, and M and N are finitely generated graded free $F[t]$-modules? What will elementary matrix operations do to the degree and homogeneity of basis elements? Let us explore.

1. Swapping rows (columns). This clearly does not change any degree.

2. Multiplying a row (column) by an invertible element. We notice that the only invertible elements in $F[t]$ are the elements $q \in F_0 = F$. Consequently $\deg(e_i') = \deg(q^{-1}e_i') = \deg(e_i)$ [$\deg(f_i') = \deg(qf_i') = \deg(f_i)$].

3. Adding q times row j (column i) to row i (column j), $q \in F[t]$. Then the change in basis is $e_j' = e_j - qe_i$ [$f_i' = f_i + qf_j$]. Now, if

$$\begin{aligned} deg(q) &= deg(e_j) - deg(e_i) \\ [deg(q) &= deg(f_i) - deg(f_j)], \end{aligned} \tag{4}$$

   then $\deg(e_j') = \deg(e_j)$ [$\deg(f_i)$ - $\deg(f_i)$]. And this is the type of type 3 elementary operations that we need in order to eliminate rows (columns) in reducing a matrix.

Any composition of such operations will naturally not alter anything as well. We conclude that reducing the representation matrix with such elementary matrix operations will not alter the degree or homogeinity of the basis elements.

5.4 Computing graded homology modules

In this section we will show how to compute the graded homology modules we have been talking about. Now, we will assume that we know what the cycle modules are, as this will be part of an inductive process in the next section.

Let F be a field and assign the standard grading to $F[t]$. Now, suppose we have a chain complex (C_n, ∂_n) of graded free $F[t]$-modules, with degree 0 $F[t]$-homomorphisms as boundary maps. We know by lemma 5.32 that all graded submodules of a graded free module is free. Hence, let us suppose that for some k we have the following,

$$ker(\partial_k) = t^{\gamma_1}F[t] \oplus t^{\gamma_2}F[t] \oplus \cdots \oplus t^{\gamma_n}F[t], \text{ and}$$
$$C_{k+1} = t^{\alpha_1}F[t] \oplus t^{\alpha_2}F[t] \oplus \cdots \oplus t^{\alpha_m}F[t],$$

where $\alpha_i \geq \gamma_i$ for all i, j. Now, we want to compute $H_k = ker(\partial_k)/\text{im}(\partial_{k+1})$, so we need to find $im(\partial_{k+1})$. Let e_i and f_i be the standard bases for F^m and F^n, respectively. Then we have that

$$\mathcal{B} = \{t^{\gamma_j} f_j\}_{j=1}^n = \{0, ..., t^{\gamma_j}, 0, ..., 0\}_{j=1}^n \text{ is a basis for } Z_k \text{ and}$$
$$\mathcal{C} = \{t^{\alpha_i} e_i\}_{i=1}^m = \{0, ..., t^{\alpha_i}, 0, ..., 0\}_{i=1}^n \text{ is a basis for } C_{k+1}$$

are bases for $ker(\partial_k)$ and C_{k+1}. Next we apply our boundary operator on the basis elements of C_{k+1}, which gives us that

$$\partial_{k+1}(t^{\alpha_i} e_i) = t^{\alpha_i}(a_{1i}, a_{2i}, ..., a_{ni}),$$

for some $a_{ji} \in F[t]_0 = F$, and this is necessarily so since ∂_{k+1} is of degree 0. We continue by expressing the result as a sum of basis elements of Z_k, so

$$\begin{aligned} t^{\alpha_i}(a_{1i}, a_{2i}, ..., a_{ni}) &= t^{\alpha_i} \sum_{j=1}^{n} a_{ji} t^{-\gamma_j}(t^{\gamma_j} f_j) \\ &= \sum_{j=1}^{n} a_{ji} t^{\alpha_i - \gamma_j}(t^{\gamma_j} f_j). \end{aligned}$$

The resulting matrix representation is then

$$M_{k+1} = (t^{\alpha_i - \gamma_j} a_{ji}),$$

where $a_{ji} \in F$.

Next, by doing elementary row and column operations we can put M_{k+1} into Smith normal form, which we know is equivalent to multpiplying M_{k+1} on the right and left by appropriate invertible matrices P^{-1} and Q. From these we can deduce the new basis elements and their degrees as we have discussed. Now, note that since $a_{ji} \in F$, these elements have inverses, hence these will all dissappear in the Smith normal form as they

are easily removed by type 2 elementary operations. The image of ∂_{k+1}, remembering our notation for this, will be the following.

$$\begin{aligned} im(\partial_{k+1}) &= \partial_{k+1}(C_{k+1}) = [\mathcal{B}]M_{k+1}[C_{k+1}]_{\mathcal{C}} \\ &= [\{t^{\gamma_j} f_j\}]PM_{k+1}QF[t]^m \end{aligned}$$

$$= [\{t^{\gamma_{j_l}} f_{j_l}\}] \begin{bmatrix} t^{\alpha_{i_1}-\gamma_{j_1}} & & & & & & & \\ & t^{\alpha_{i_2}-\gamma_{j_2}} & & & & & & \\ & & \ddots & & & & & \\ & & & t^{\alpha_{i_r}-\gamma_{j_r}} & & & & \\ & & & & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & \end{bmatrix} F[t]^m$$

$$= (t^{\alpha_{i_1}-\gamma_{j_1}})t^{\gamma_{j_1}}F[t] \oplus (t^{\alpha_{i_2}-\gamma_{j_2}})t^{\gamma_{j_2}}F[t] \oplus \cdots \oplus (t^{\alpha_{i_r}-\gamma_{j_r}})t^{\gamma_{j_r}}F[t].$$

Where r is the rank of ∂_{k+1}, and $\{i_l\}$ and $\{j_l\}$ are some permutations of the sets $\{1, ..., m\}$ and $\{1, ..., n\}$, respectively. Next we note that $(t^{\alpha_{i_l}-\gamma_{j_l}})t^{\gamma_{j_l}}F[t]$ is a graded submodule of $t^{\gamma_{i_l}}F[t]$, and hence by Theorem 5.23 we have that

$$\begin{aligned} H_k &= ker(\partial_k)/im(\partial_{k+1}) \\ &= \frac{t^{\gamma_{j_1}}F[t] \oplus t^{\gamma_{j_2}}F[t] \oplus \cdots \oplus t^{\gamma_{j_n}}F[t]}{(t^{\alpha_{i_1}-\gamma_{j_1}})t^{\gamma_{j_1}}F[t] \oplus (t^{\alpha_{i_2}-\gamma_{j_2}})t^{\gamma_{j_2}}F[t] \oplus \cdots \oplus (t^{\alpha_{i_r}-\gamma_{j_r}})t^{\gamma_{j_r}}F[t]} \\ &\cong_0 t^{\gamma_{j_1}}\frac{F[t]}{(t^{\alpha_{i_1}-\gamma_{j_1}})F[t]} \oplus \cdots \oplus t^{\gamma_{j_r}}\frac{F[t]}{(t^{\alpha_{i_r}-\gamma_{j_r}})F[t]} \oplus t^{\gamma_{j_{r+1}}}F[t] \oplus \cdots \oplus t^{\gamma_{j_m}}F[t]. \end{aligned}$$

6 Persistence

Material in this section is gathered from citeulike:3148973, [3], [9], and [11].

Our simplicial complex constructions presented earlier, from the Ĉech complex to the witness complexes, all induce a sequence of increasing complexes as the parameter ϵ increases. That is, they induce a filtered complex. In Figure 20 we have an example of this. Here points have been sampled from a space with two holes, like a double annulus, and from this Ĉech complexes for various choices of ϵ have been constructed. Now, homology will view the double annulus as one connected component with two loops, that is $\mathcal{B}_0 = 1$ and $\mathcal{B}_1 = 2$. However none of the corresponding Ĉech complexes for the various values for ϵ, displayed here, manage to obtain the same homology as the the original space. Furthermore, if we do not know what the underlying space the samples come from, which will be the reality of things, we may not even have any idea of how to begin choosing an ϵ. Hence, instead of searching for one correct choice for ϵ, what we will do is

consider whole filtered complexes and see if we can single out features that persist longer than others.

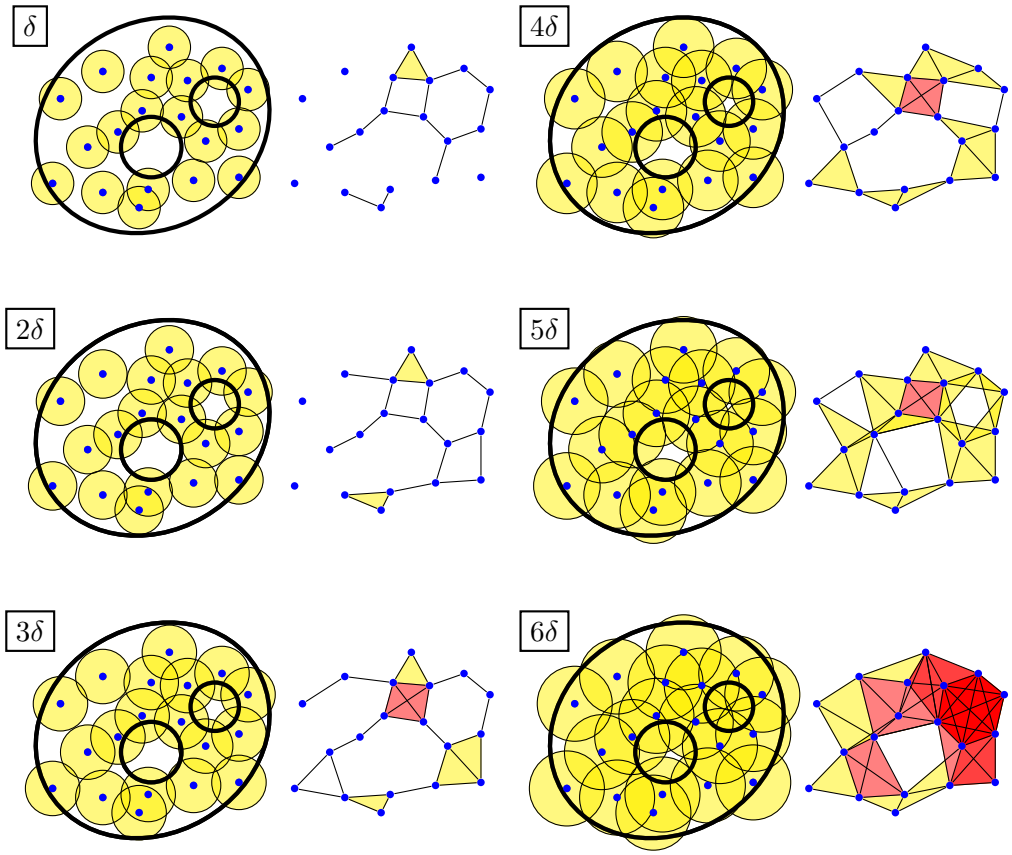


Figure 20: A filtered complex created by the Ĉech complex construction. Yellow indicating 2-simplices and darker red indicating higher dimensional simplices.

Let us look at the exmaple in Figure 20. At $\epsilon = \delta$ we have two loops. The biggest one lasts through all values up to $\epsilon = 6\delta$ where it is closed up. The other loop that is created at $\epsilon = \delta$, only lasts until $\epsilon = 3\delta$. This might indicate that the first loop is more likely to be an actual feature of the data set, which it actually is, while the other loop is less likely to be such a feature, which it actually is not.

The types of features we will be measuring are those that are detected by betti numbers, that is the number of connected components, number of loops, number of voids, and so on. What we will do is create what we will call *Barcodes*. The barcodes for Figure 20 is shown in Figure 21. Each interval represents a feature , which in this case is connected components or loops. The start of the intervals indicates the time of birth of such a feature, and the endpoint the time of death of such a feature. Counting the

number of blue and red lines in Figure 21, at some value for ϵ, gives us $\mathcal{B}_0$ and $\mathcal{B}_1$ for that value of ϵ, respectively

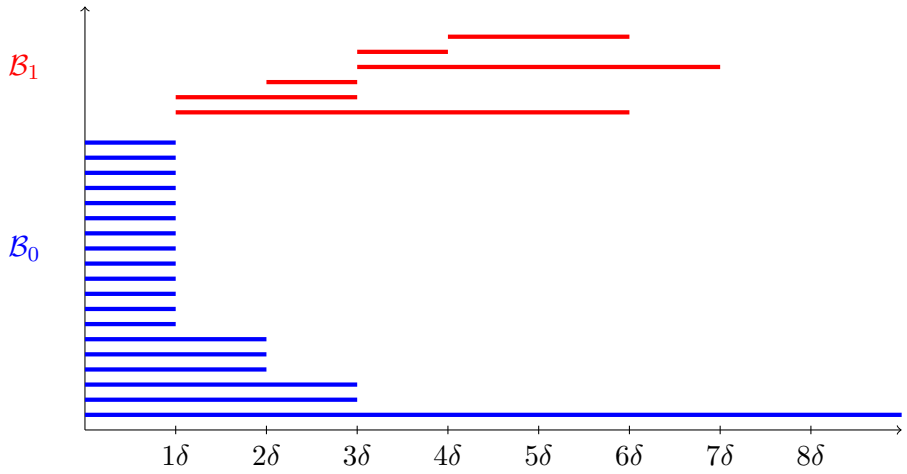


Figure 21: Betti 0 and 1 barcodes for the filtered complex in figure 20.

In this section we will create a homology structure that contains information from all the complexes in such a filtered complex. It will be a graded $F[t]$-module, and we will relate these modules to sets of intervals which will be the barcodes. Finally, we introduce an algorithm for computing these barcodes from a filtered complex.

6.1 Persistent homology and structure

In the following we assume that R is a commutative ring with identity.

Definition 6.1. A *persistence complex* $\mathcal{C}$ is a family of chain complexes $\{C_*^i\}_{i\geq 0}$ over R, together with chain maps $f^i : C_*^i \to C_*^{i+1}$, which gives us the following diagram

$$C_*^0 \xrightarrow{f^0} C_*^1 \xrightarrow{f_1} C_*^2 \xrightarrow{f_2} \cdots .$$

If we extend in the vertical direction we get the diagram in Figure 22.

$$\begin{array}{ccccccc}
 & \downarrow{\scriptstyle \partial_3} & & \downarrow{\scriptstyle \partial_3} & & \downarrow{\scriptstyle \partial_3} & \\
\cdots & \xrightarrow{f^2} & C_2^2 & \xrightarrow{f^1} & C_2^1 & \xrightarrow{f^0} & C_2^0 \\
 & \downarrow{\scriptstyle \partial_2} & & \downarrow{\scriptstyle \partial_2} & & \downarrow{\scriptstyle \partial_2} & \\
\cdots & \xrightarrow{f^2} & C_1^2 & \xrightarrow{f^1} & C_1^1 & \xrightarrow{f^0} & C_1^0 \\
 & \downarrow{\scriptstyle \partial_1} & & \downarrow{\scriptstyle \partial_1} & & \downarrow{\scriptstyle \partial_1} & \\
\cdots & \xrightarrow{f^2} & C_0^2 & \xrightarrow{f^1} & C_0^1 & \xrightarrow{f^0} & C_0^0
\end{array}$$

Figure 22: A persistence complex expanded in the vertical direction.

If we have a filtered complex K and form chain complexes from the subsets $\{K^i\}_{0\leq i\leq m}$, then we get a persistence complex where the f^i's are the induced inclusion maps. That is, we have the following commutative diagram, shown in Figure 23, where C_* is the functor taking complexes to chain complexes.

$$\begin{array}{ccc}
K^i & \xrightarrow{C_*} & C_*^i \\
{\scriptstyle \iota_i}\downarrow & \circ & \downarrow{\scriptstyle f^i = C_*(\iota_i)} \\
K^{i+1} & \xrightarrow[C_*]{} & C_*^{i+1}
\end{array}$$

Figure 23: Chain complexes formed from a filtered complex.

Definition 6.2. A *persistence module* $\mathcal{M}$ is a family of R-modules $\{M^i\}$, together with R-homomorphisms $\phi : M^i \to M^{i+1}$.

For example, the homology of a persistence complex is a persistence module, where the map ϕ^i just maps a homology class to the one that contains it. Also, if we have a persistence complex $\mathcal{C}$ as above, the chain groups $\{C_k^i\}_i$ with the chain maps $\{f^i\}_i$ become a persistence module.

Definition 6.3. A persistence complex $\{C_*^i, f^i\}$ is of *finite type* if each component complex, i.e., each $\{C_j^i\} \in \{C_*^i\}$ is a finitely generated R-module, and if the maps f^i are isomorphisms for $i \geq m$ for some $m \in \mathbb{N}$.

Similarly we define.

Definition 6.4. A persistence module $\{M^i, \phi^i\}$ is of finite type if each M^i is a finitely generated R-module, and if the maps ϕ^i are isomorphisms for $i \geq m$ for some $m \in \mathbb{N}$.

For example, if we form a persistence module $\mathcal{C}$ from a filtered complex $K = \{K^i\}$ it will be of finite type. Then, we can turn $\mathcal{C}$ into a persistence module of finite type by taking the homology of the chain complexes in $\mathcal{C}$. That is, we have the following commutative diagram shown in Figure 24.

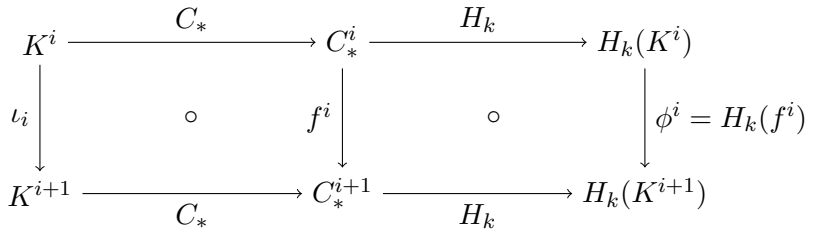


Figure 24: A persistence module induced by a filtered complex.

Definition 6.5. Let $\mathcal{M} = \{M^i, \phi^i\}$ be a persistence module over a ring R. Form the polynomial ring $R[t]$ and equip it with the standard grading. We can now form a graded $R[t]$-module by setting

$$\alpha(\mathcal{M}) = \bigoplus_{i=0}^{\infty} M^i,$$

where the action of multiplying by t is given by

$$t(m^0, m^1, m^2, ...) = (0, \phi^0(m^0), \phi^1(m^1), \phi^2(m^2), ...),$$

that is, a shift upward in gradation.

Theorem 6.6. *The correspondence α defines an equivalence of categories between the category of persistence R-modules of finite type, and the category of finitely generated $\mathbb{N}$-graded modules over $R[t]$.*

There does not exist any simple classifications of graded $R[t]$-modules where R is not a field, for example $\mathbb{Z}$. However, if $R = F$ is a field, then we know that the structure is

$$\Big[\bigoplus_{j=1}^{n} t^{\gamma_j} F[t]\Big] \oplus \Big[\bigoplus_{i=1}^{n} t^{\alpha_i} \frac{F[t]}{(t^{n_i})F[t]}\Big].$$

So, let us see if we can make some sense of things if we start with a filtered complex, and a field F as ground ring of coefficents.

Definition 6.7. Let $\mathcal{C} = \{C_*, f^i\}$ be a persistence complex generated by a filtered complex K, then the $\{C_k^i, f^i\}$ are persistence modules where the C_k^i's are free F-modules generated by the k-simplices $\sigma \in K^i$. Let the σ's form the standard basis for C_k^i. Then, $\{C_k^i, f^i\}$ corresponds to a graded free $F[t]$-module. So let us define $C_k = \alpha(\{C_k^i, f^i\})$.

Next, suppose that $C_{k+1} = t^{\alpha_1}F[t] \oplus \cdots \oplus t^{\alpha_m}F[t]$, where each $\alpha_i \in \mathbb{N}$ is the filtration time at which some k-simplex was added. Moreover, let $\{\sigma'_1, ..., \sigma'_m\}$ form the standard basis for $C_{k+1}^M = F^m$, where M is the maximum filtration time in the filtered complex so that this is the biggest complex in the filtered complex. This will induce the standard basis $\{\sigma_i = t^{\alpha_i}\sigma'_i\}_{i=1}^m$ for C_{k+1}. We call the σ_i for *graded simplices*. Also, the boundary maps $\partial_{k+1}^i : C_{k+1}^i \to C_k^i$, induce a boundary map $\partial_{k+1} : C_{k+1} \to C_k$, and applying it on the basis elements we have that

$$\partial_{k+1}(\sigma_i) = \partial_{k+1}(t^{\alpha_i}\sigma'_i) = t^{\alpha_i}\partial_{k+1}^i(\sigma'_i).$$

Clearly $\partial \circ \partial = 0$, and we see that ∂ is a graded $F[t]$-homomorphism of degree 0. Hence, we have a chain complex

$$\cdots \to C_{k+1} \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \to \cdots .$$

Definition 6.8. Given the above chain complex, we then define the k-th persistent homology $F[t]$-module to be

$$H_k^{pers} \frac{ker\partial_k}{im\partial_{k+1}}.$$

If we assume as we did earlier, in the end of Section 5, that

$$C_{k+1} = t^{\alpha_1}F[t] \oplus t^{\alpha_2}F[t] \oplus \cdots \oplus t^{\alpha_m}F[t] \text{ and}$$
$$Z_k = t^{\gamma_1}F[t] \oplus t^{\gamma_2}F[t] \oplus \cdots \oplus t^{\gamma_n}F[t]$$

then,

$$H_k = t^{\gamma_{j_1}} \frac{F[t]}{(t^{\alpha_{i_1}-\gamma_{j_1}})F[t]} \oplus \cdots \oplus t^{\gamma_{j_r}} \frac{F[t]}{(t^{\alpha_{i_r}-\gamma_{j_r}})F[t]} \oplus t^{\gamma_{j_{r+1}}}F[t] \oplus \cdots \oplus t^{\gamma_{j_m}}F[t].$$

We interpret this as follows. The γ_{j_l}'s tell us when a k-cycle is created with the birth of some k-simplex in $K^{\gamma_{j_l}}$, which in turn generates a $t^{\gamma_{j_l}}F[t]$ part in the cycle group Z_k. The α_{i_l}'s tell us when a k-cycle becomes the boundary of some $(k+1)$-chain with the birth of some $(k+1)$-simplex in $K^{\alpha_{i_l}}$, which turns the corresponding part in the cycle group into a torsion part in the homology $F[t]$-module. Basically this means that the γ_{j_l}'s tell us when a generating k-cycle occurs, and the differences $(\alpha_{i_l} - \gamma_{j_l})$ tell us how long until they become a boundary. So, it seems that to each such homology group we can relate it to a set of intervals.

Definition 6.9. A $\mathcal{P}$-interval is an ordered pair (i, j) with $0 \leq i \leq j \in \mathbb{Z}^\infty = \mathbb{Z} \cup \{\infty\}$.

We can associate a graded $F[t]$-module to a set $\mathcal{S}$ of $\mathcal{P}$-intervals via a bijection Q.

Definition 6.10. Define the bijection Q from the set of $\mathcal{P}$-intervals to the set of finitely generated graded $F[t]$-modules by

$$Q(i,j) = t^i \frac{F[t]}{t^{j-i}F[t]},$$

where $Q(i,\infty) = t^i F[t]$. Furthermore, if $\mathcal{S} = \{(i_1,j_1),(i_2,j_2),...,(i_n,j_n)\}$ is a set of $\mathcal{P}$-intervals, define

$$Q(\mathcal{S}) = \bigoplus_{l=1}^{n} Q(i_l,j_l).$$

Corollary 6.11. *The correspondence $\mathcal{S} \to Q(\mathcal{S})$ defines a bijection between the finite sets of $\mathcal{P}$-intervals and the finitely generated graded modules over the graded ring $F[t]$. Consequently, the isomorphism classes of persistence modules of finite type over F are in bijective correspondence with the finite sets of $\mathcal{P}$-intervals.*

Hence, we have our barcodes.

6.2 Algorithm.

In this section we present an algorithm for computing barcodes from a filtered complex. We will use the filtration in Figure 25 as a running example, and we will do the computations over $\mathbb{Z}_2$ coefficients. Table 1 shows the simplices in Figure 25 sorted according to their degree.

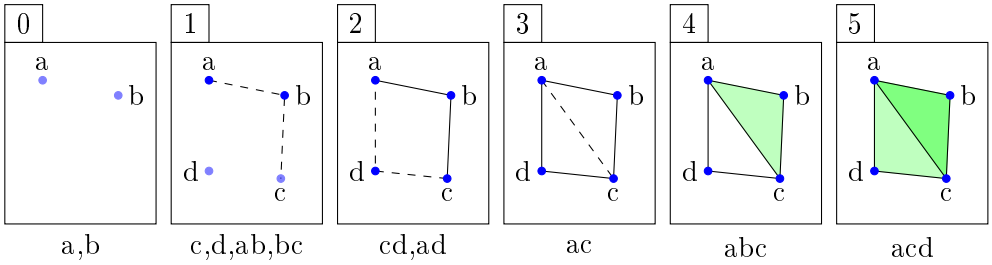


Figure 25: A filtered complex. Newly added simplices are dashed or semitransparent.

Table 1: Simplices in Fig 25 sorted according to degree.

a	b	c	d	ab	bc	cd	ad	ac	abc	acd
0	0	1	1	1	1	2	2	3	4	5

Throughout this section we let $\{e_j\}$ and $\{\hat{e}_i\}$ represent homogeneous bases for C_k and C_{k-1}, respectively. Since we use $\mathbb{Z}_2$ coefficients we deduce from what we have done earlier that the matrix representation M_k of ∂_k has the following property

$$\deg M_k(i,j) = \deg e_j - \deg \hat{e}_i,$$

where $M_k(i,j)$ denotes the element at position (i,j). In our example ∂_1 represented relative to the standard bases $\{a,b,c,d\} \subset C_0$ and $\{ab, bc, cd, ad, ac\} \subset C_1$, will be

$$M_1 = \left[\begin{array}{c|ccccc} & ab & bc & cd & ad & ac \\ \hline d & 0 & 0 & t & t & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right].$$

To compute the homologies we need to represent $\partial_k : C_k \to C_{k-1}$ by a homogeneous basis for C_k and Z_{k-1}, reduce it to Smith normal form, and read off the description for H_k as we have discussed. We will do this inductively in increasing dimension. For $k = 0$ we have that $\partial = 0$, hence $Z_0 = C_0$, hence we can use the standard basis for C_0 to represent Z_0. Now, suppose we have a representation of M_k relative to the standard bases for C_k and Z_{k-1}. Let $\{e_j\}$ be the standard basis for C_k and $\{\hat{e}_i\}$ be a homogeneous basis for Z_{k-1}. First sort the basis elements $\{\hat{e}_i\}$ in reverse order and $\{e_j\}$ in increasing order. This is done in our matrix representation M_1. Then we reduce the matrix to column echelon form, i.e.

1. All zero columns are at the right of the matrix.
2. Moving from left to right. The leading element, called the pivot, of each nonzero column after the first, is below the previous pivot element in the previous row. A row (column) with a pivot is called a pivot-row (column).

$$\left[\begin{array}{ccccccc} \mathbf{*} & 0 & & & & 0 \\ & \mathbf{*} & 0 & \cdots & & \\ * & * & 0 & & & \vdots \\ & * & \mathbf{*} & 0 & \cdots & \\ * & & * & 0 & \cdots & 0 \end{array}\right].$$

Figure 26: The column-echelon form. An * indicates a nonzero element and one of bold type * indicates a pivot.

An illustration is given Figure 26, and in our example we get that

$$\tilde{M}_1 = \left[\begin{array}{c|ccccc} & cd & bc & ab & z_1 & z_2 \\ \hline d & \mathbf{t} & 0 & 0 & 0 & 0 \\ c & t & \mathbf{1} & 0 & 0 & 0 \\ b & 0 & t & \mathbf{t} & 0 & 0 \\ a & 0 & 0 & t & 0 & 0 \end{array}\right],$$

where $z_1 = ad - cd - t \cdot bc - t \cdot ab$ and $z_2 = ac - t^2 \cdot bc - t^2 \cdot ab$ forms a homogeneous basis for Z_1. The way we achieve this column echelon form is via elementary column

operations of type 1 and type 3.

Column reduction algorithm:

1. From the left moving to the right in the columns, call the first column with leading entry at some row index j, a pivot column, making row j a pivot row.
2. Again, moving from the left to right, we pick out nonpivot columns. Then moving from top to bottom in each column, remove elements in pivot rows with the corresponding pivot columns by using type 3 column operations. This we can do since we have ordered our basis for C_k in increasing order, furthermore, they will not alter the degree of the basis elements. We continue removing elements in the column until the column is all zeros, or if we arrive at a new pivot element giving us a new pivot column.
3. Finally we use swaps to put the matrix in wanted column echelon form.

Let us apply the algortihm on our example. First we find the initial pivot columns, and we have marked them below.

$$M_1 = \left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & ad & ac \\ \hline d & 0 & 0 & t & t & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right].$$

Then, following the above recipe we get

$$\left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & ad & ac \\ \hline d & 0 & 0 & t & t & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right] \xrightarrow[(1)]{\substack{\text{(col 4)}\to \\ \text{(col 4)} \\ -1\cdot\text{(col 3)}}} \left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & \substack{ad \\ -cd} & ac \\ \hline d & 0 & 0 & t & 0 & 0 \\ c & 0 & 1 & t & t & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right] \xrightarrow[(2)]{\substack{\text{(col 4)}\to \\ \text{(col 4)} \\ -t\cdot\text{(col 2)}}}$$

$$\left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & \substack{ad-cd \\ -t\cdot bc} & ac \\ \hline d & 0 & 0 & t & 0 & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & t^2 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right] \xrightarrow[(3)]{\substack{\text{(col 4)}\to \\ \text{(col 4)} \\ -t\cdot\text{(col 1)}}} \left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & ad-cd-t\cdot bc-t\cdot ab & ac \\ \hline d & 0 & 0 & t & 0 & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & 0 & t^3 \end{array}\right] \xrightarrow[(4)]{\substack{\text{(col 5)}\to \\ \text{(col 5)} \\ -t^2\cdot\text{(col 2)}}}$$

$$\left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & ad-cd-t\cdot bc-t\cdot ab & ac-t^2\cdot bc \\ \hline d & 0 & 0 & t & 0 & 0 \\ c & 0 & 1 & t & 0 & 0 \\ b & t & t & 0 & 0 & t^3 \\ a & t & 0 & 0 & 0 & t^3 \end{array}\right] \xrightarrow[(5)]{\substack{\text{(col 5)}\to \\ \text{(col 5)} \\ -t^2\cdot\text{(col 1)}}} \left[\begin{array}{c|ccccc} & \mathbf{ab} & \mathbf{bc} & \mathbf{cd} & ad-cd-t\cdot bc-t\cdot ab & ac-t^2\cdot bc-t^2\cdot ab \\ \hline d & 0 & 0 & t & 0 & 0 \\ c & 0 & 1 & t & 0 & 0 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[(6)]{\text{col (3)}\leftrightarrow\text{col (1)}} \left[\begin{array}{c|ccccc} & \mathbf{cd} & \mathbf{bc} & \mathbf{ab} & \begin{smallmatrix}ad-cd\\-t\cdot bc-t\cdot ab\end{smallmatrix} & \begin{smallmatrix}ac-t^2\cdot bc\\-t^2\cdot ab\end{smallmatrix} \\ \hline d & t & 0 & 0 & 0 & 0 \\ c & t & 1 & 0 & 0 & 0 \\ b & 0 & t & t & 0 & 0 \\ a & 0 & 0 & t & 0 & 0 \end{array}\right]$$

Lemma 6.12. *The pivots in column-echelon form are the same as in Smith normal form. Moreover, the degree of the basis elements on pivot rows is the same in both forms.*

Proof. Because of our sort of the basis elements for Z_{k-1}, the degree of row basis elements $\hat{e}_i$ is decreasing moving from the top to bottom row. This implies that in a column deg $M_k(i,j) \leq$ deg $M_k(i',j)$, where $i \leq i'$. Therefore we can eliminate nonzero elements below pivots, and we know that such operations will not alter the pivot elements or the degrees of the new basis elements.

The corollary of this is that we can read of the description for the homology directly from the column echelon form without the need for any row operations.

Corollary 6.13. *Let $\tilde{M}_k$ be the column-echelon form for ∂_k relative to bases $\{e_j\}$ and $\{\hat{e}_i\}$ for C_k and Z_{k-1}, respectively. If row i has pivot $\tilde{M}(i,j) = t^n$, it contributes*

$$t^{deg\ \hat{e}_i}\frac{F[t]}{t^n}$$

to the description of H_{k-1}. If row i does not have a pivot element, then it contributes

$$t^{deg\ \hat{e}_i}F[t]$$

to the description of H_{k-1}.

In our example with $\tilde{M}_1$ we get

$$H_0 = F[t] \oplus \frac{F[t]}{tF[t]} \oplus t\frac{F[t]}{tF[t]}.$$

Finally, we want to represent ∂_{k+1} relative to the basis we have computed for Z_k. Let M_k and M_{k+1} be the standard matrix representations. Since the domain of ∂_k and codomain of ∂_{k+1} are the same, column operations done on M_k induce row operations on M_{k+1}. Type 1 and type 3 column operations on M_k will do the following.

- Swap column i with column j in $M_k \Leftrightarrow$ swap row i with row j in M_{k+1}.
- Replace column i with column i + q·column j in $M_k \Leftrightarrow e_i \to e_i + qe_j$ in $C_k \Leftrightarrow e'_i \to e'_i + qe'_j$ in $C_{k+1} \Leftrightarrow$ replace row j with row j - q·row i in M_{k+1}.

In our example we have

$$M_2 = \left[\begin{array}{c|cc} & abc & acd \\ \hline ac & t & t^2 \\ ad & 0 & t^3 \\ cd & 0 & t^3 \\ bc & t^3 & 0 \\ ab & t^3 & 0 \end{array}\right].$$

The previous column operations in M_1 then becomes the following row operations in M_2.

$$\left[\begin{array}{c|cc} & abc & acd \\ \hline ac & t & t^2 \\ ad & 0 & t^3 \\ cd & 0 & t^3 \\ bc & t^3 & 0 \\ ab & t^3 & 0 \end{array}\right] \xrightarrow{(1)} \left[\begin{array}{c|cc} & abc & acd \\ \hline ac & t & t^2 \\ \substack{ad\\-cd} & 0 & t^3 \\ cd & \mathbf{0} & \mathbf{0} \\ bc & t^3 & 0 \\ ab & t^3 & 0 \end{array}\right] \xrightarrow{(2)}$$

$$\left[\begin{array}{c|cc} & abc & acd \\ \hline ac & t & t^2 \\ \substack{ad-cd\\-t\cdot bc} & 0 & t^3 \\ cd & 0 & 0 \\ bc & \boldsymbol{t^3} & \boldsymbol{t^4} \\ ab & t^3 & 0 \end{array}\right] \xrightarrow{(3)} \left[\begin{array}{c|cc} & abc & acd \\ \hline ac & t & t^2 \\ \substack{ad-cd\\-t\cdot bc-t\cdot ab} & 0 & t^3 \\ cd & 0 & 0 \\ bc & t^3 & t^4 \\ ab & \boldsymbol{t^3} & \boldsymbol{t^4} \end{array}\right] \xrightarrow{(4)}$$

$$\left[\begin{array}{c|cc} & abc & acd \\ \hline \substack{ac\\-t^2\cdot bc} & t & t^2 \\ \substack{ad-cd\\-t\cdot bc-t\cdot ab} & 0 & t^3 \\ cd & 0 & 0 \\ bc & \mathbf{0} & \mathbf{0} \\ ab & t^3 & t^4 \end{array}\right] \xrightarrow{(5)} \left[\begin{array}{c|cc} & abc & acd \\ \hline \substack{ac-t^2\cdot bc\\t^2\cdot ab} & t & t^2 \\ \substack{ad-cd\\-t\cdot bc-t\cdot ab} & 0 & t^3 \\ cd & 0 & 0 \\ bc & 0 & 0 \\ ab & \mathbf{0} & \mathbf{0} \end{array}\right]$$

$$\xrightarrow{(6)} \left[\begin{array}{c|cc} & abc & acd \\ \hline \substack{ac-t^2\cdot bc\\t^2\cdot ab} & t & t^2 \\ \substack{ad-cd\\-t\cdot bc-t\cdot ab} & 0 & t^3 \\ ab & \mathbf{0} & \mathbf{0} \\ bc & 0 & 0 \\ cd & \mathbf{0} & \mathbf{0} \end{array}\right]$$

However, we will see that we can obtain this representation in a very easy manner. Let us notice a few things. We have that $\partial_k \partial_{k+1} = 0$ which implies that we must have $M_k M_{k+1} = 0$, and this relationship will be unchanged during such paired elementary operations as above. Then, if M_k is reduced to column echelon form $\tilde{M}_k$, we must have that the induced matrix $\breve{M}_{k+1}$ from M_{k+1} must be of the form as illustrated in Figure 27.

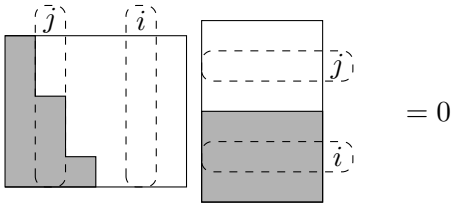


Figure 27: Since $\partial_k \partial_{k+1} = 0$ then we must have that $M_k M_{k+1} = 0$. When M_k is put into column echelon form corresponding row operations must zero out rows in M_{k+1} that correspond to pivot columns in M_k.

Notice that for each pivot column j in $\tilde{M}_k$, row j in $\breve{M}_{k+1}$ is all zeros. This means that when we do type 3 column operations to remove all other pivot elements in pivot row j, then the corresponding row operations on M_{k+1} will eventually zero out row j in M_{k+1}, as we see in our running example. Furthermore, these operations will not affect row i in M_{k+1}. Thus, we can easily obtain the representation of M_{k+1} relative to the standard basis for C_{k+1} and the basis computed for Z_k.

Lemma 6.14. *To represent ∂_{k+1} relative to the standard basis for C_{k+1} and the basis computed for Z_k, simply delete rows in M_{k+1} that correspond to pivot columns in $\tilde{M}_k$.*

Consequently, we do not need row operations at all. In our running example we get the following. Removing the three bottom rows in M_2 corresponding to pivot columns in $\tilde{M}_1$ we have

$$\breve{M}_2 = \left[\begin{array}{c|cc} & abc & acd \\ \hline z_2 & t & t^2 \\ z_1 & 0 & t^3 \end{array}\right].$$

Here ad and ac are replaced with their corresponding basis elements $z_1 = ad - cd - t \cdot bc - t \cdot ab$ and $z_2 = ac - t^2 \cdot bc - t^2 \cdot ab$.

6.3 Pseudocode.

From our previous section we have seen that if we sort our basis, we only need to do column operations. Furthermore, by just eliminating rows corresponding to pivot columns in the previous dimension, we get the desired basis change. Due to this we do not need matrix representations, instead we will represent the boundary operators with boundary chains corresponding to the boundary columns. To keep track of all our moves, we will create an array T.

First of all we give a total ordering of all the graded simplices. We order the simplices first by increasing degree, then by dimension, and then it does not matter, so we assign arbitrarily. If m is the number of simplices then T will have m slots and slot i is denoted by $T[i]$. A simplex is added accordingly to its index in the total ordering. A σ_i can be

marked if its boundary corresponds to a non-pivot column. Finally, if the boundary d of a simplex σ_j corresponds to a pivot-column, it is stored in T[i], where i is the index of the highest indexed row basis element in d.

Table 2: A table for the array T for the filtration in Fig. 25.

	0	1	2	3	4	5	6	7	8	9	10
Simplex	a	b	c	d	ab	bc	cd	ad	ac	abc	acd
(Degree)	0	0	1	1	1	1	2	2	3	4	5
Marked	x	x	x	x				x	x		
Pivot column		$\substack{b \\ -a}$	$\substack{c \\ -b}$	$\substack{d \\ -c}$				ad	ac		

The actual algorithm is as follows.

ComputeIntervals(K)
for $k = 0$ to $\dim(K)$ **do**
  $L_k = \emptyset$
  for $j = 0$ to $m - 1$ **do**
    $d =$ RemovePivotRows(σ_j)
    if $d = \emptyset$ **then**
      Mark σ_j
    else
      $i =$ maxindex d; $k = \dim(\sigma_j)$
      Store d in $T[i]$
      $L_k = L_k \cup$ (deg σ_i,deg σ_j)
    end if
  end for
  for $j = 0$ to $m - 1$ **do**
    if σ_j is marked and $T[j]$ is has no pivot column entry **then**
      $k =$ dim σ_j; $L_k = L_k \cup$(deg σ_j,∞)
    end if
  end for
end for

where RemovePivotRows is

chain RemovePivotRows(σ)
k = dim σ; $d = \partial_k \sigma$
Remove unmarked terms in d
while $d \neq \emptyset$ **do**
  $i =$ maxindex d
  if $T[i]$ is empty **then**

```
        break
    end if
    Let q be the coefficient of σ_i in T[i]
    Let b be the coefficient of σ_i in d
    d = d − bq^{-1}T[i]
end while
return d
```

RemovePivotRows. RemovePivotRows takes as input a simplex σ. The boundary column d of σ is computed, and all terms in d corresponding to marked simplices are removed. Then, by drawing on the pivot columns stored in T, it does gauss eliminations resulting in either a empty column or a new pivot column. The empty column or pivot column is then returned.

ComputeIntervals. We traverse through the simplices according to their index. Suppose we have arrived at simplex σ_j. The first thing the algorithm does is to check, via RemovePivotRows, whether σ_i's boundary column corresponds to a zero column or a pivot column. If a pivot-column is returned we store it in the appropriate slot $T[i]$. This will give us a $\mathcal{P}$-interval $(\deg \sigma_i, \deg \sigma_j)$. When we have traversed through all the simplices we go through the filtration once more, in search for infinite $\mathcal{P}$-intervals, i.e. marked simplices with no pivot column entry.

Running time. The algorithm has the same running time as Gaussian elimination over field, i.e. it is $\mathcal{O}(m^3)$, worst case.

In our example we wil get the following $\mathcal{P}$-intervals. $L_0 = \{(0,\infty),(0,1),(1,1),(1,2)\}$ and $L_1 = \{(2,5),(3,4)\}$.

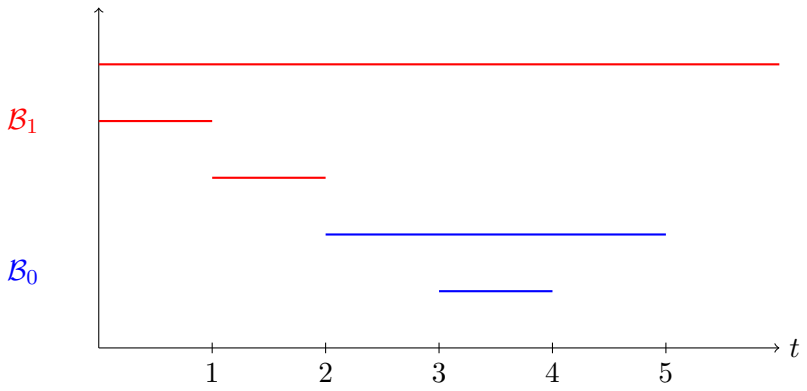


Figure 28: Barcodes for the filtered complex in figure 25.

7 JPlex tests

JPlex is a Java software package for computing the persistent homology of filtered complexes using the algortihm we have presented. Furthermore, it makes use of some of our complex constructions to create filtered complexes. They are, the Vietoris-Rips complex, Strong witness complex, Weak witness complex, Vietoris-Rips weak witness complex, and the Simple witness complex. The authors are Harlan Sexton and Mikael Vejdemo Johansson. JPlex can be used together with Beanshell or MATLAB. In this paper JPlex has been used together with MATLAB.

7.1 Sphere

In this test 500 points have been sampled from the unit sphere by using uniformly distributed pseudorandom numbers to generate spherical coordinates of the unit sphere. From these, 20 landmark points have been chosen by random. Then the above mentioned complex constructions have been used to create filtered complexes, and barcodes have been computed. Simplices are also constricted to dimension 3, as this is only what we need to compute betti numbers up to dimension 2. Typical barcodes obtained is shown in figures 29 to 33. The correct betti numbers for the sphere is $\mathcal{B}(B_0) = 1$, $\mathcal{B}(B_1) = 2$, and $\mathcal{B}(B_2) = 1$. For all constructions we present typical ϵ values for when the correct betti numbers are achieved. We also state the total number of simplices in the corresponding complexes for these ϵ values. All the results presented for the different complex constructions are obtained by the same set of sample and landmark points, except for the Vietoris-Rips complex. For this complex it was necessary to generate a new set of 20 points.

First of all we notice that the barcodes for the Simple witness and Vietoris-Rips constructions are not as convincing as the others. They manage, but not necessarily always, to obtain the correct betti numbers. But again, the barcodes are not that convincing. Furthermore they create, as we will see, a lot more simplices before they obtain the correct betti numbers. This results in a much longer computation time. The Vietoris-Rips complex and the Simple complex in figures 29 and 30 obtain the correct betti numbers at $\epsilon \approx 1.65$ and $\epsilon \approx 1.65$, resulting in 950 and 1257 simplices, respecitvely. They also give very similar barcodes. That is, $W(Z, \mathcal{L}, \epsilon)$ is very similar to $VR(\mathcal{L}, 2\epsilon)$. This makes sense. Recall figure 38. If two landmark points are a bit closer to each other than 2ϵ, then there is great amount of potential witness points (498 for an edge). Hence, there is a good chance that there is one in the intersection of the two ϵ-balls around each point.

We observe that the two weak witness type constructions, Vietoris-Rips weak witness and Weak witness, almost get it right right from the start at about $\epsilon \approx 0,05$. This is not surprising due to the weak witness definition allowing a point to be a witness right from the start. That is, $(p+1)$ landmarks spans a p-simplex if they are the $(p+1)$ closest points to another point. They also show almost identical barcodes. Altough the starting points for all the edges is the same for both, all barcodes in the Vietoris-Rips version

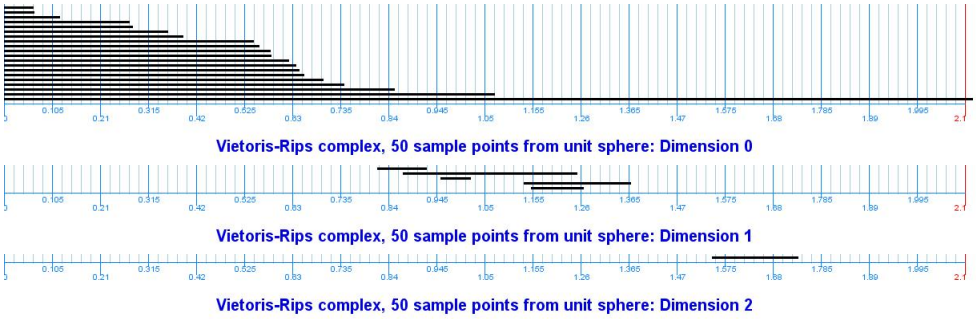


Figure 29: Barcodes for the Vietoris-Rips complex with sample points from a sphere.

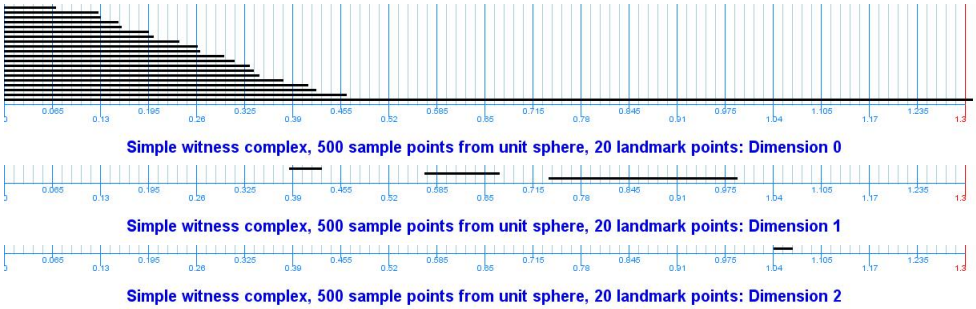


Figure 30: Barcodes for the Simple witness complex with sample points from a sphere.

ends quicker due to the Vietoris-Rips complex is a filled in version of the Weak witness complex. Note that the scales are different for the two barcode sets. The number of simplices at achieved correct betti numbers for the Vietoris-Rips weak witness complex and the Weak witness complex is 348 and 324, respectively. The Vietoris-Rips weak version constructs as expected more simplices. However it is much quicker. It did not matter much for this test, but for a greater number of sample or landmark points it is much quicker, without actually having done any extensive research on this except a few tests. But it figures as the Vietoris-Rips weak version only need to compute distances between two points.

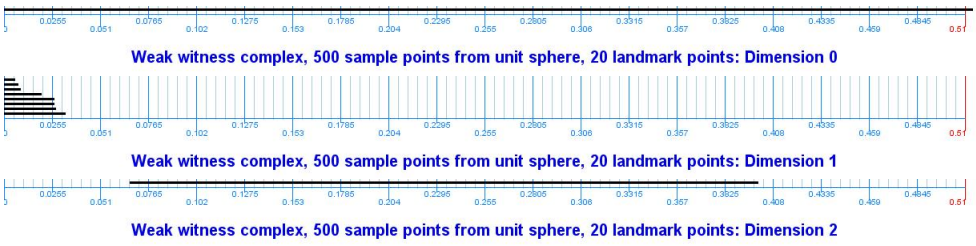


Figure 31: Barcodes for the Weak witness complex with sample points from a sphere.

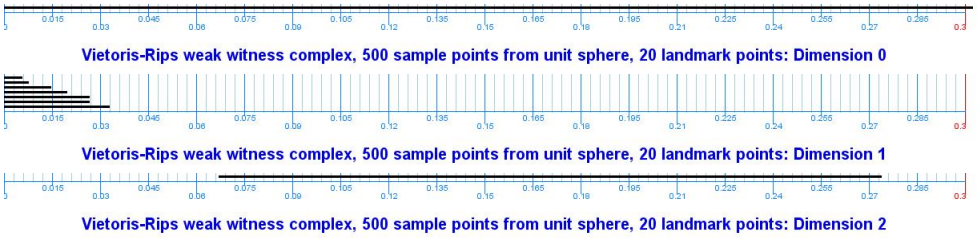


Figure 32: Barcodes for the Vietoris-Rips weak witness complex with sample points from a sphere

The strong witness complex obtain correct betti numbers at $\epsilon \approx 0.15$ with 228 simplices. This is less than the weak witness types, which makes sense as the strong witness is a stricter witness type. Recall figure 17.

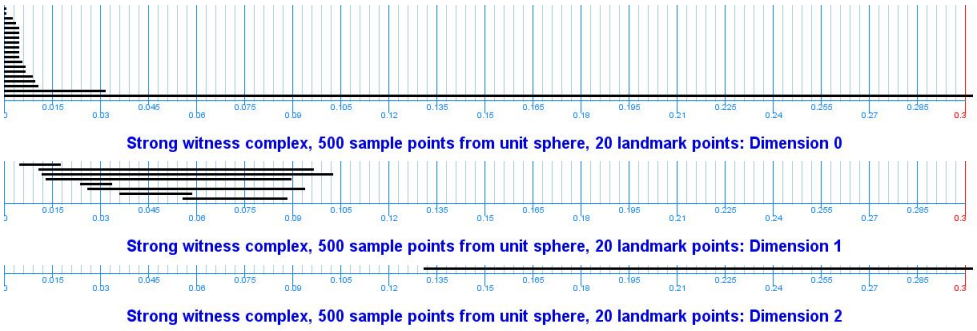


Figure 33: Barcodes for the Strong witness complex with sample points from a sphere

7.2 Torus

In this test 100^2 points have been sampled from a 100×100 grid on the 2-dimensional unit torus in $\mathbb{R}^4$, and then noise has been added. From these, 50 landmark points have been chosen randomly. Then the witness complex constructions have been used to create filtered complexes, and barcodes have been computed. Simplices are constricted to dimension 3, as this is only what we need to compute betti numbers up to dimension 2. Typical barcodes obtained is shown in figures 34 to 38. The correct betti numbers for the sphere is $\mathcal{B}(B_0) = 1$, $\mathcal{B}(B_1) = 0$, and $\mathcal{B}(B_2) = 1$. For all constructions we present typical ϵ values for when the correct betti numbers are achieved. We also state the total number of simplices in the corresponding complexes for these ϵ values. All the results presented for the different complexes are obtained by the same sets of sample and landmark points, except for the Vietoris-Rips complex. For this complex it was necessary to generate a new set of points, so 7^2 points were generated.

The Simple witness and Vietoris-Rips construction numbers are, respectively, $\epsilon \approx$

FAIL and $\epsilon \approx 1.36$. The Vietoris-Rips does a better job on this one. I do not know what the explanation for this might be, but maybe it has something to do with the witness points making for a more prololonged continuous like conversion from vertices to edges, to 2-simplices, and to 3-simplices. While with the Vietoris-Rips the conversions are more abrupt and gathered. Looking at the barcodes for the other witness constructions we see that they look much better.

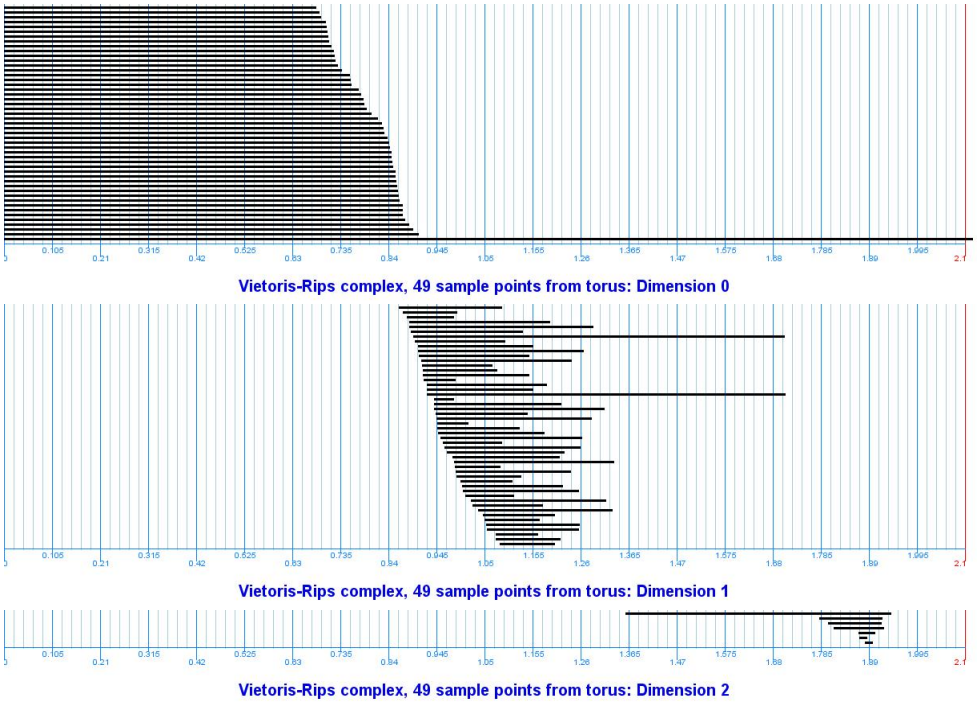


Figure 34: Barcodes for the Vietoris-Rips complex with sample points from a torus.

The weak complexes get it right practically right from the start. The valid interval for the Vietoris-Rips weak witness complex is much shorter than for the Weak witness complex. Which again is as expected as the the Vietoris-Rips weak complex is a filled in version of the Weak witness complex. The Vietoris-Rips weak witness construction seems to get into some trouble at $\epsilon \approx 0.07$ when the filling og the simplices with the Vietoris-Rips part goes a skew. The numbers are $\epsilon \approx 0.003$ for both, and 449 simplices for the Weak witness complex and 475 for the Vietoris-Rips weak witness complex.

The Stong witness complex does a very good job in this one and the numbers are $\epsilon \approx 0.05$ and 73, with the same explanations as for the sphere test.

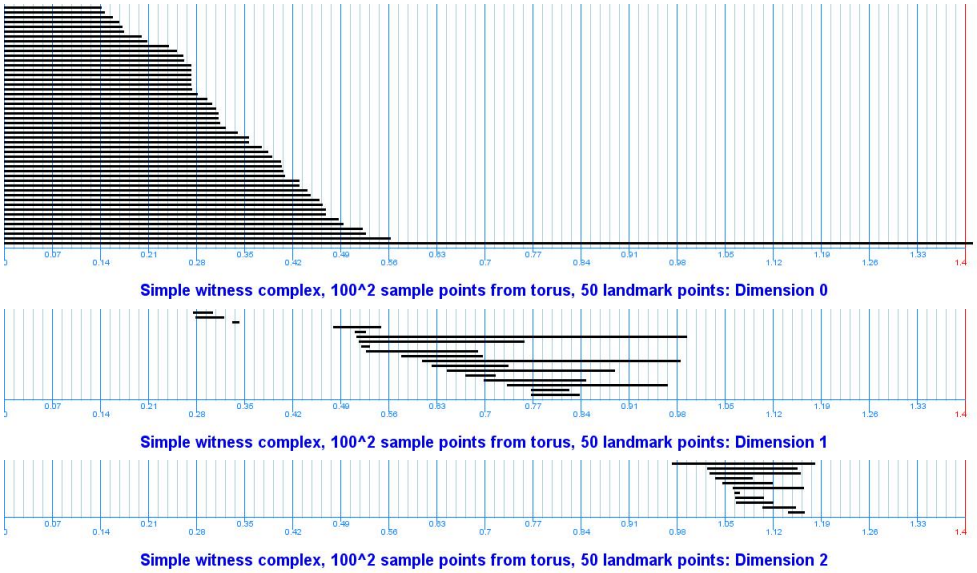


Figure 35: Barcodes for the Simple witness complex with sample points from a torus.

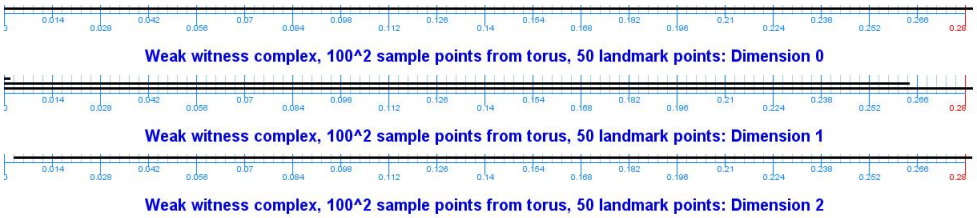


Figure 36: Barcodes for the Weak witness complex with sample points from a torus.

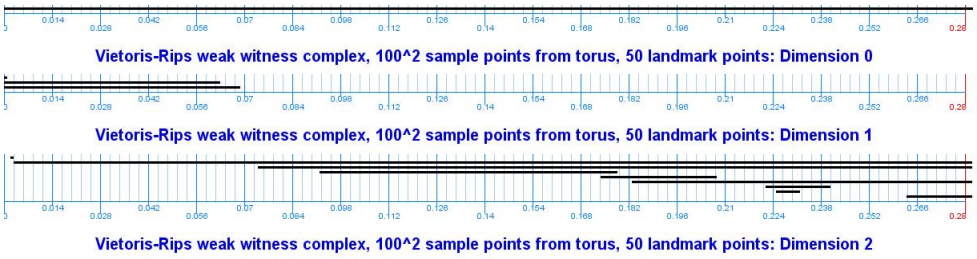


Figure 37: Barcodes for the Vietoris-Rips weak witness complex with sample points from a torus.

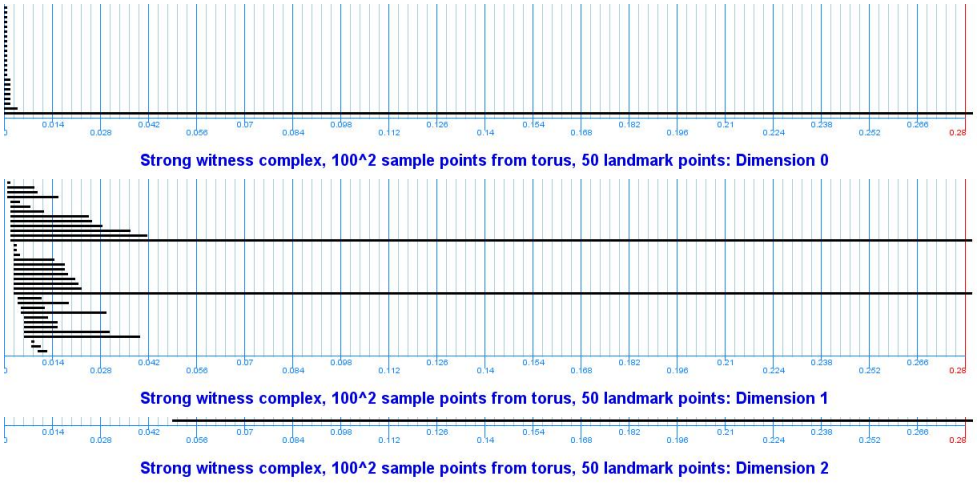


Figure 38: Barcodes for the Strong witness complex with sample points from a torus.

8 Mapper

Material in this section is gathered from [3] and [24].

Up to this point we have used homology, specifically betti numbers and barcodes, as indicators of intrinsic geometric features of data sets. However, humans are very good at recognizing patterns in dimensions less than 3. It seems then like a good idea to develop methods that tap into this potential. The simplest way to visualize a space is by seeing its component structure, hence any clustering algorithm can be regarded as a visualization method. The method we will introduce now is called *Mapper* and is based on topological ideas with the nerve complex construction at the core.

8.1 Topological idea

Recall from the nerve section that we obtained a map $g \circ \rho : X \to \mathcal{N}(\mathcal{U})$ where $\mathcal{U}$ was a cover for a topological space X. Now whether or not $g \circ \rho$ is a homotopy equivalence we still have a continuous map giving us a kind of coordinization of the space X. An illustration is given in Figure 39. An ordinary coordinization provide a map to Euclidean space of some dimension, and can provide useful insight into the space under investigation. We have now instead obtained a map from our space to a complex. We then say that this complex represent our space, and if it is of low dimension it may be easily visualized.

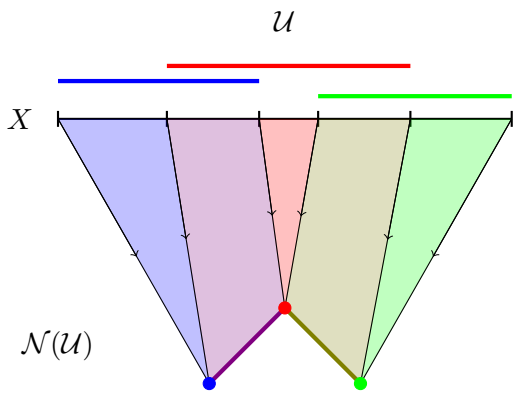


Figure 39: Illustrating the values of a coordinization map from an interval X with covering $\mathcal{U}$, to the corresponding nerve complex of $\mathcal{U}$.

We will now introduce a new variant of the nerve construction which produces a more sensitive target for our coordinization map.

Definition 8.1. Let X be any topological space, and let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be a covering of X. For each $\alpha \in A$ we can write $U_\alpha = \bigcup_{i \in I}(U_{\alpha,i})$ where $\{U_{\alpha,i}\}_{i \in I}$ are the path connected

components of U_α. The set $\hat{\mathcal{U}} = \{U_{\alpha,i}\}_{\alpha\in A, i\in I}$ then forms a new covering of X, and we define a new nerve complex

$$\mathcal{N}^{\pi_0}(\mathcal{U}) = \mathcal{N}(\hat{\mathcal{U}}).$$

We illustrate the difference with an example.

Example 8.2. Let $X = \mathbb{S}^1$, and let $\mathcal{U}$ be a covering of X given by the sets $U_0 = \{(x,y) \mid y < 0\}$, $U_1 = \{(x,y) \mid y > 0\}$, and $U_2 = \{(x,y) \mid y \neq \pm 1\}$. Note that U_0 and U_1 have one connected component each, while U_2 has two connected components. We observe that $\mathcal{N}^{\pi_0}(\mathcal{U})$ is homeomorphic to X while $\mathcal{N}(\mathcal{U})$ is not. This is illustrated in Figure 40.

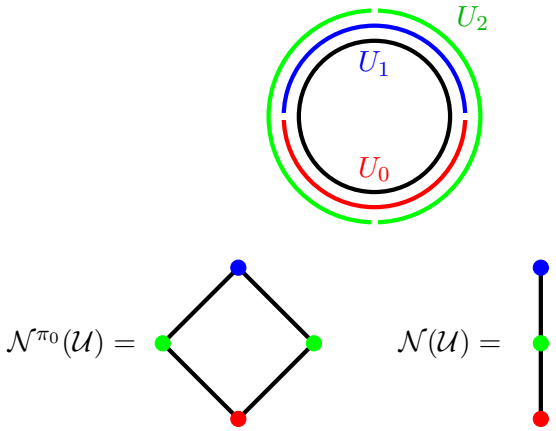


Figure 40: See Example 8.2.

Now, we need a way to construct coverings of spaces. Earlier we have done this by covering the space with open ϵ-balls. Now we introduce a new method. Suppose we are given a continuous map $\rho : X \to Z$ from our topological space X to a metric space Z. The map ρ is called the *reference map* or the *filter*, and the metric space is called the *parameter space*. Suppose that $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ is a open or closed covering of Z, then the set $\rho^*\mathcal{U} = \{\rho^{-1}(U_\alpha)\}_{\alpha\in A}$ is a covering of X, for which we can compute $\mathcal{N}^{\pi_0}(\rho^*\mathcal{U})$.

Typical parameter spaces are $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{S}^1$.

Example 8.3. Let $Z = \mathbb{R}$ then we can construct the covering $\mathcal{U}(R,e)$ which consists of all intervals $[kR-e, (k+1)R+e]$ where $k \in \mathbb{N}$. This gives us two parameter choices, and as long as $e \leq R/2$ there will never be more than two intervals intersecting, restricting the resulting nerve complex to dimension 1. Product of such intervals give coverings of $\mathbb{R}^n$.

Example 8.4. Let $Z = \mathbb{S}^1$, $N \geq 2$ in $\mathbb{N}$, and $\epsilon \geq 0$ in $\mathbb{R}$. Then a covering of $\mathbb{S}^1$ is $\mathcal{U}[N,\epsilon] = \{U_j\}_{0\leq j\leq N}$ where

$$U_j = \{(cos(x), sin(x)) \mid x \in [\frac{2\pi j}{N} - \epsilon, \frac{2\pi j}{N} + \epsilon]\},$$

whenever $\epsilon \geq \frac{\pi}{N}$.

We give an example of the whole topological idea.

Example 8.5. Let $X = \mathcal{S}^1$, $Z = [0,1]$, and let $\mathcal{U} = \mathcal{U}(\frac{1}{3}, \frac{1}{9})$ be a covering of Z. Further suppose we have a map $\rho : X \to Z$ given by $\rho(x,y) = y$. Then we have the following illustrated in in Figure 41.

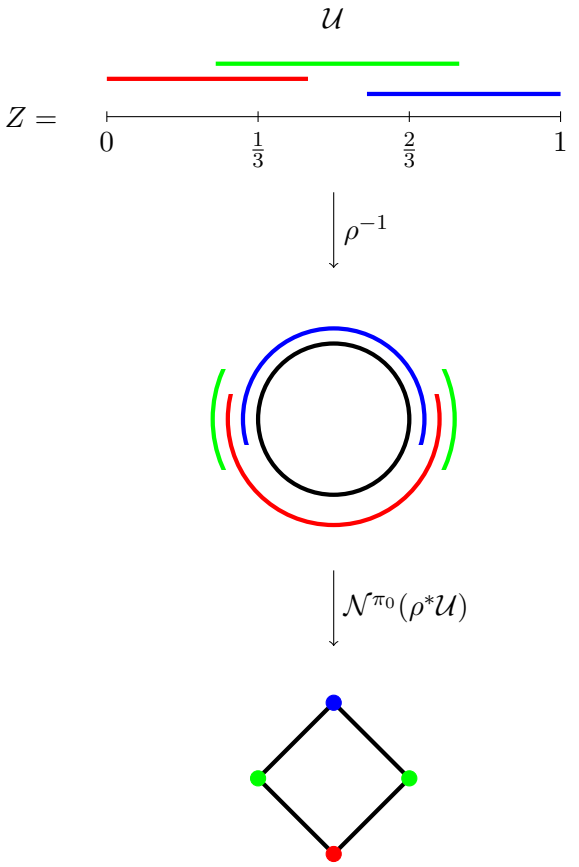


Figure 41: See Example 8.5.

8.2 Statistical version

We want to apply the above idea to point clouds. We can still make a reference map $\rho : X \to Z$ from a point cloud X to a parameter space Z, and $\rho^*\mathcal{U} = \{\rho^{-1}(U_\alpha)\}_{\alpha \in A}$ is still a covering of X. However it is meaningless to split up this covering into its path connected components, which is just points. The idea is instead to replace the notion of path connected components with the notion of clusters of points. So, instead of finding the path connected components of $\rho^{-1}(U_\alpha)$, we find the clusters of $\rho^{-1}(U_\alpha)$. We can then write

$$X_\alpha = \rho^{-1}(U_\alpha) \ = \bigcup_{c \in C} (X_{\alpha,c}),$$

where $\{X_{\alpha,c}\}_{c\in C}$ is the set of clusters in $\rho^{-1}(U_\alpha)\}$ found by the clustering algorithm. Analogous to the topological case we can now define a new covering

$$\widehat{\rho^*\mathcal{U}} = \{X_{\alpha,c}\}_{\alpha\in A, c\in C}.$$

for which can compute $\mathcal{N}(\widehat{\rho^*\mathcal{U}})$.

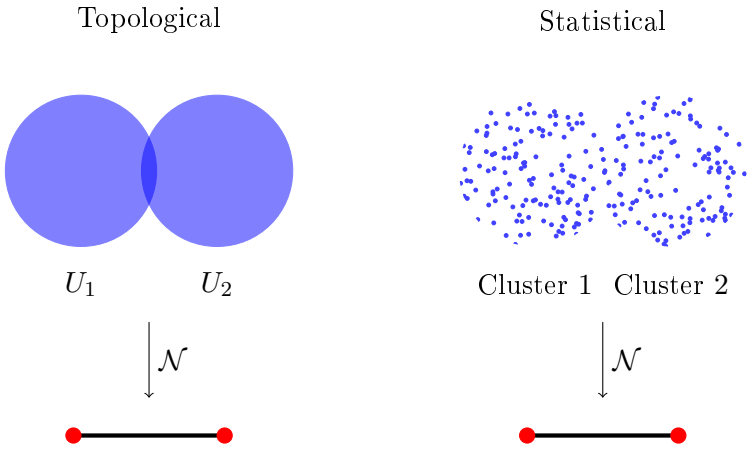


Figure 42: Moving from the notion of connected components to the notion of clusters.

Any clustering method will do, but the method we will be using is called single linkage clustering, and works as follows. Set a paramenter ϵ. The clusters $\{U_{\alpha,c}\}$ of X_α will be equivalence classes under the equivalence relation $\sim_\epsilon$ defined by $x \sim_\epsilon y$ iff $d(x,y) \leq \epsilon$. Note that this is exactly the same way we decide whether or not there is an edge between two points in $VR(X,\epsilon)$, and hence the number of connected components in $VR(X,\epsilon)$ is the same as the number of clusters found single linkage clustering. This will enable us to take advantage of persistent homology with this clustering algorithm to find good choices for ϵ. But we come back to that later in the end of this section.

Let us recap our method.

1. Define a reference map $\rho : X \to Z$ from a point cloud X to a metric space Z.
2. Create a covering $\mathcal{U} = \{U_\alpha\}_{\alpha\in A}$ of Z.
3. Construct the subsets $X_\alpha = \rho^{-1}(U_\alpha)$.
4. Use some clustering scheme to find the clusters $\{X_{\alpha,c}\}_{c\in C}$ of X_α.
5. Compute $\mathcal{N}(\widehat{\rho^*\mathcal{U}}) = \mathcal{N}(\{X_{\alpha,c}\}_{\alpha\in A, c\in C})$.

An example is given in Figure 43.

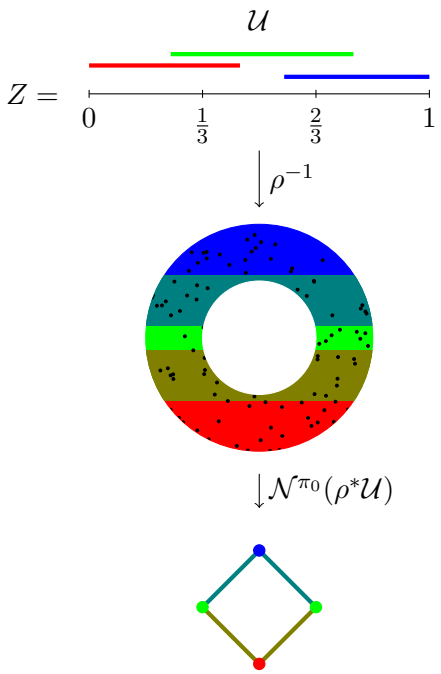


Figure 43: Here X are points sampled from an annulus in $\mathbb{R}^2$ with outer diameter equal to 1 and inner diameter equal to $1/2$. The parameter space is $Z = [0, 1]$, and the covering for Z is $\mathcal{U} = \mathcal{U}[1/3, 1/9]$. The reference map $\rho : X \to Z$ is given by $f(x, y) = y$.

8.3 Multiresolution, topological idea

We would like to produce a multiresolution, or multiscale structure, so that we can try to distinguish between true features and artifacts in the data set. The intuition is that features appearing at many different resolutions or scales are more likely to be actual features of the set than those appearing at fewer resolutions or scales. We start by giving the following definition.

Definition 8.6. Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ and $\mathcal{V} = \{V_\beta\}_{\beta \in B}$ be two coverings of a topological space X. A *map of coverings* from $\mathcal{U}$ to $\mathcal{V}$ is a set map $\theta : A \to B$ so that for all $\alpha \in A$, we have $U_\alpha \subseteq V_{\theta(\alpha)}$.

Moreover, given a map of coverings $\theta : A \to B$ from $\mathcal{U}$ to $\mathcal{V}$ we have an induced map of complexes $\mathcal{N}(\theta) : \mathcal{N}(\mathcal{U}) \to \mathcal{N}(\mathcal{V})$.

Example 8.7. Let $\mathcal{U}[R, e]$ be the covering of $\mathbb{R}$ as defined earlier. Then the identity map $i : \mathbb{Z} \to \mathbb{Z}$ yields a map of coverings $\mathcal{U}[R, e] \to \mathcal{U}[R, e']$ whenever $e \leq e'$, which also induces a map of complexes between the corresponding nerve complexes. Figure 44 is an example with two coverings $\mathcal{U}(1, 1/4)$ and $\mathcal{U}(1, 1/3)$.

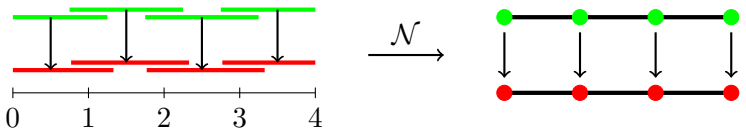


Figure 44: See Example 8.7.

Example 8.8. The map of integers $k \to \lfloor \frac{k}{2} \rfloor$ yields a map of coverings from $\mathcal{U}[R, e] \to \mathcal{U}[2R, e]$, which further induces a map of complexes between the corresponding nerve complexes. Figure 45 is an example with two coverings $\mathcal{U}(1, 1/4)$ and $\mathcal{U}(2, 1/4)$.

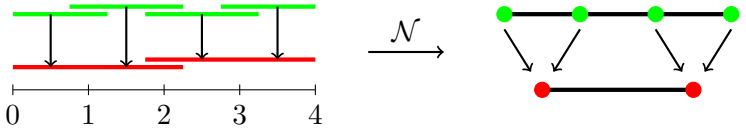


Figure 45: See Example 8.8.

These two examples can naturally be generalized to higher dimensions.

Let us now consider a topological space X, together with a reference map $\rho : X \to Z$ from X to a metric space Z, coverings $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ and $\mathcal{V} = \{V_\beta\}_{\beta \in B}$, and a map of coverings $\theta : A \to B$ from $\mathcal{U}$ to $\mathcal{V}$ on Z. Now if $U_\alpha \subseteq V_{\theta(\alpha)}$, then $\rho^{-1}(U_\alpha) \subseteq \rho^{-1}(V_{\theta(\alpha)})$, consequently each path connected component of $\rho^{-1}(U_\alpha)$ is a subset of exactly one path connected component of $\rho^{-1}(V_{\theta(\alpha)})$. Hence we obtain a map of coverings

$$\widehat{\rho^*\mathcal{U}} = \{\widehat{\rho^{-1}(U_\alpha)}\}_{\alpha \in A} \to \widehat{\rho^*\mathcal{V}} = \{\widehat{\rho^{-1}(V_\beta)}\}_{\beta \in B},$$

which induces a map of complexes

$$\mathcal{N}^{\pi_0}(\rho^*\mathcal{U}) \xrightarrow{\mathcal{N}(\theta)} \mathcal{N}^{\pi_0}(\rho^*\mathcal{V}).$$

Consequently, if we have a family of coverings $\{\mathcal{U}_i\}_{i=0}^N$, and maps of coverings $\{\theta_i : \mathcal{U}_i \to \mathcal{U}_{i+1}\}_{i=0}^N$, there is an induced diagram

$$\mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_0) \xrightarrow{\mathcal{N}^{\pi_0}(\theta_0)} \mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_1) \xrightarrow{\mathcal{N}^{\pi_0}(\theta_1)} \cdots \xrightarrow{\mathcal{N}^{\pi_0}(\theta_{N-1})} \mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_N).$$

We see that moving to the left in the diagram results in a finer resolution, while moving to the right results in a coarser resolution.

Example 8.9. Let X be the shape at the top in Figure 46. The parameter space is $Z = [0, 8]$ with coverings $\mathcal{U}(1, 1/4)$, $\mathcal{U}(2, 1/4)$, and $\mathcal{U}(4, 1/4)$, and let the maps of coverings be the one induced by the integer map $k \to \lfloor \frac{k}{2} \rfloor$ as in the Example 8.8. Looking at Figure 46 we notice that finer coverings gives a complex that captures finer features of the space X. Furthermore we see which vertices of a finer resolution corresponds to

which vertices of a coarser resolution. It is just like zooming in to discover finer details, but with the greater danger of being deluded by artifacts, atleast when we get to the statistical version.

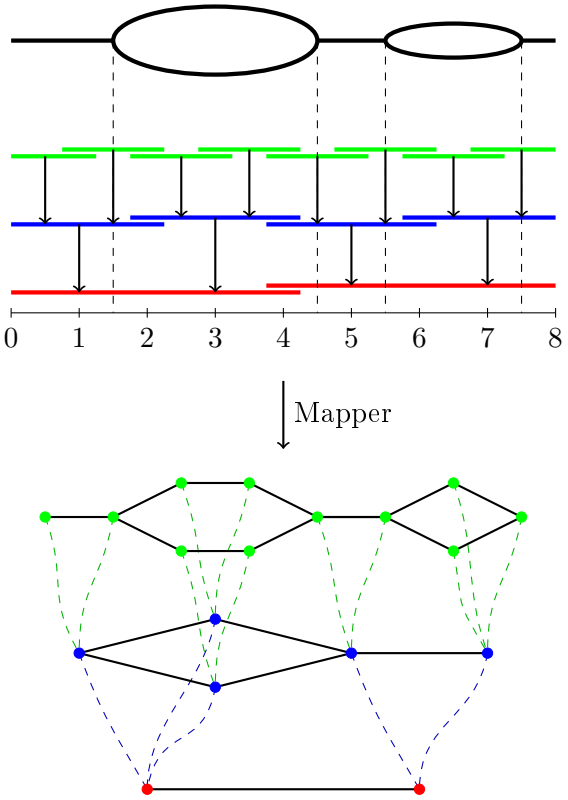


Figure 46: Figure for example 8.9.

8.4 Multiresolution, statistical version

We want to apply the above technique to point clouds. So we assume the same conditions as above, but let X be a point cloud. To apply the same idea to point clouds there is one thing we need to be sure of. In the topological version above we knew that each connected component of $\rho^{-1}(U_\alpha)$ was included in exactly one connected component of $\rho^{-1}(V_{\theta(\alpha)})$. Similarily if X is now a point cloud, we need to know that each cluster of $\rho^{-1}(V_\alpha)$ is included in exactly one cluster of $\rho^{-1}(V_{\theta(\alpha)})$. An illustration is given in Figure 47.

We arrive at the following definiton for clustering algorithms satisfying this property.

Definition 8.10. Suppose X and Y are point clouds such that $X \subset Y$, and that we have an inclusion map $\iota : X \to Y$. A clustering algorithm is said to be *functorial* if

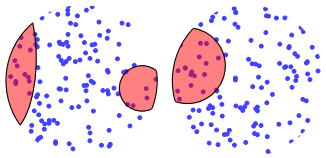


Figure 47: Suppose the clustering algorithm partitions the above point cloud into the two obvious clusters. Next, we apply the algorithm to the subset of the point cloud denoted by the red areas. Then the clustering algorithm must not partiton the subset point cloud in such a way that we get a cluster containing points both from the left and right clusters found earlier. In other words, for instance, the two red areas close to each other should not then become a cluster.

the inlcusion of each cluster constructed in X is included in exactly one of the clusters constructed in Y. That is, the diagram in Figure 48 commutes.

$$\begin{array}{ccc} X & \longrightarrow & \text{Clusters of } X \\ {\scriptstyle \iota}\downarrow & \circ & \downarrow \text{Inclusion of clusters} \\ Y & \longrightarrow & \text{Clusters of } Y \end{array}$$

Figure 48: See definition 8.10.

Let us recap. We are given a point cloud X together with a reference map $\rho : X \to Z$ from X to a metric space Z, coverings $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ and $\mathcal{V} = \{V_\beta\}_{\beta \in B}$ of Z, and a map of coverings $\theta : A \to B$. We then obtain inclusions $\rho^{-1}(U_\alpha) \subseteq \rho^{-1}(V_{\theta(\alpha)})$. By applying a functorial clustering algorithm we find $\widehat{\rho^*\mathcal{U}}$ and $\widehat{\rho^*\mathcal{V}}$, where now each element(cluster) in $\widehat{\rho^*\mathcal{U}}$ can be included in exactly one element(cluster) in $\widehat{\rho^*\mathcal{V}}$. Hence we obtain a simplical map

$$\mathcal{N}^{\pi_0}(\rho^*\mathcal{U}) \xrightarrow{\mathcal{N}(\theta)} \mathcal{N}^{\pi_0}(\rho^*\mathcal{V}).$$

Moreover, if we again have a family of coverings $\{\mathcal{U}_i\}_{i=0}^N$, and maps of coverings $\{\theta_i : \mathcal{U}_i \to \mathcal{U}_{i+1}\}_{i=0}^N$, there is an induced diagram

$$\mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_0) \xrightarrow{\mathcal{N}^{\pi_0}(\theta_0)} \mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_1) \xrightarrow{\mathcal{N}^{\pi_0}(\theta_1)} \cdots \xrightarrow{\mathcal{N}^{\pi_0}(\theta_{N-1})} \mathcal{N}^{\pi_0}(\rho^*\mathcal{U}_N).$$

Where moving to the left in the diagram results in a finer resolution, while moving to the right results in a coarser resolution.

8.5 Filters

The big question is ofcourse how to find appropriate and good reference maps. These may be some user defined functions given someone who has a certain knowledge about

the data sets in question. However, there are some maps frequently used in the statistical field of analyzing point clouds, which are natural to use. They are functions that carry interesting geometric features about data sets in general.

Density. Any density estimator applied to the point cloud can provide useful information about the data set. Often it is exactly this information one is looking for. For example, letting $\epsilon > 0$, we can estimate density by using a Gaussian kernel as:

$$f_\epsilon = C_\epsilon \sum_y exp(\frac{-d(x,y)^2}{\epsilon}),$$

where $x, y \in X$, C_ϵ is a constant such that $\int f_\epsilon(x)dx = 1$, and ϵ works as a smoothing parameter.

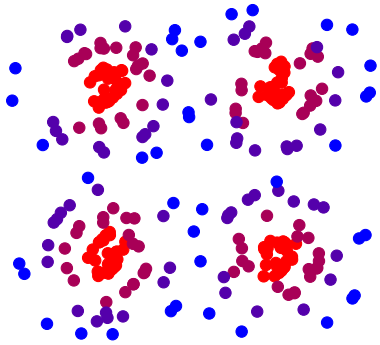


Figure 49: Points in high density areas in red, and points in low density areas in blue.

Eccentricity. Maps in this family is based on the notion of measuring the distance points have to some notion of the center in the data set. The set does not necessarily actually have a center, but a point minimizing such a function can be thought as being one. For example one can use the function

$$E_p(X) = (\frac{\sum_{y \in X} d(x,y)^p}{N})^{1/p},$$

where $x, y \in X$.

Mapper used with such reference maps is good at recognizing qualitative properties of certain shapes, e.g. mapper should easily be able to recover the shape in Figure 51.

Graph Laplacians. This family of functions originate from considering a Laplacian operator on graphs. In particular, their eigenfunctions produce functions on the vertex set of the graphs. These eigenfunctions can produce useful filters. See [15] for details.

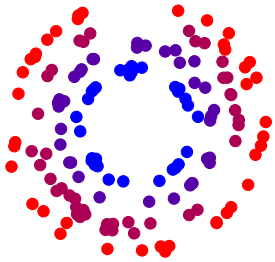


Figure 50: Illustration of eccentricity. Red points are points with high eccentricity, while blue points are points with lower eccentricity.

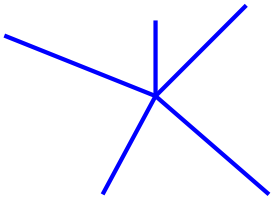


Figure 51: Mapper used with a eccentricity function as reference map, should easily be able to recover this shape.

8.6 Scale space

In our construction we need to specify an ϵ for the clustering method. Finding good choices for ϵ is generally difficult. Furthermore, it is likely that we want to specify different choices for ϵ for different regions. In this section we will present a systematic way to help choosing ϵ as described above, based on the single linkage clustering method.

Recall that single linkage clustering produces the same edges as $VR(X, \epsilon)$, and hence they have the same number of connected components(clusters). We can then obtain interesting information about the behaviour of the clusters under all values of ϵ by computing the barcodes for $\mathcal{B}_0$ for $VR(X, \epsilon)$.

Let $E(X) \subset [0, \infty)$ be the collection of all endpoints of the intervals occuring in the barcode. This gives us a finite set $E(X) = [e_1, e_2, ..., e_t]$ where $e_i < e_{i+1}$. Now, let $e_i \leq \eta < \eta' < e_{i+1}$, then the inclusion $VR(X, \eta) \hookrightarrow VR(X, \eta')$ is an isomorphism, which induce a bijection on the sets of connected components. Each interval $(e_i, e_{i+1}]$ is therefore called a *stability interval*.

Definition 8.11. Let X be a point cloud, $\rho : X \to Z$ a reference map to a metric space, and $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ a covering of Z. We define a complex $SS = SS(X, \rho, \mathcal{U})$ as follows. The vertices of SS are pairs (α, I), where $\alpha \in A$, and where I is a stability interval for the point cloud $X_\alpha = \rho^{-1}(U_\alpha)$. The set $\{(\alpha_0, I_0), (\alpha_1, I_1), ..., (\alpha_p, I_p)\}$ forms a p-simplex iff

1. $U_{\alpha_0} \cap \cdots \cap U_{\alpha_p} \neq \emptyset$.

2. $I_0 \cap \cdots \cap I_p \neq \emptyset$.

Now, the vertex map $(\alpha, I) \to \alpha$ induce a map of complexes $p : SS \to \mathcal{N}(\mathcal{U})$. A *scale choice for X and $\mathcal{U}$*, is a simplicial map $s : \mathcal{N}(\mathcal{U}) \to SS$ such that $p \circ s = Id_{\mathcal{N}(\mathcal{U})}$.

So, the choice of a scale map s will, in principle, give us a choice for the scale parameter for different regions α. Say we choose a scale map s for X and $\mathcal{U}$, and that we set $s(\alpha) = (\alpha, I_\alpha)$. Now for each α we may choose an $\epsilon_\alpha \in I_\alpha$, and it does not matter which one we choose, as long as we just choose one. Then we have chosen scale parameters varying with α and we can create a complex with these scale parameter choices as input for the single linkage clustering method. Specifically, we create a complex with vertex sets (α, c) where c is a cluster obtained by applying single linkage clustering to $\rho_{-1}(U_\alpha)$ with scale parameter choices ϵ_α.

Example 8.12. Let X be some point cloud, and let our parameter space Z be an interval with a covering $\mathcal{U}$ as illustrated in Figure 52.

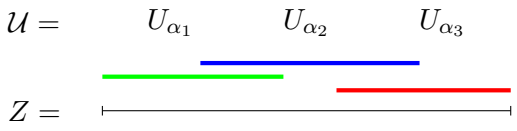


Figure 52: Parameter space Z with covering $\mathcal{U}$.

Now suppose that we have a map $\rho : X \to Z$, and that computing the $\mathcal{B}_0$ barcodes for the Vietoris-Rips complexes on the sets $\rho^{-1}(U_{\alpha_i})$ yields the following stability intervals marked in red, green and blue in Figure 53

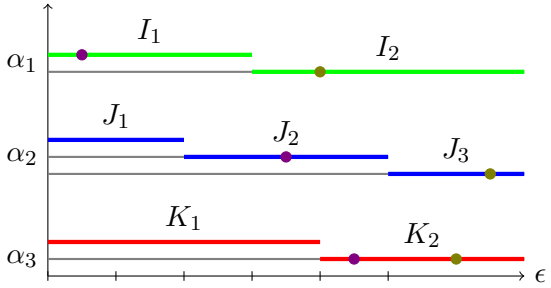


Figure 53: Supposed barcodes.

We can now create the simplicial complex $SS(X, \rho, \mathcal{U})$, and we can look for scale choices for X and $\mathcal{U}$. In Figure 54 are two simplicial maps that are possible scale choices, and where the colored dots in the barcodes in Figure 53 are possible ϵ choices corresponding to these scale choices.

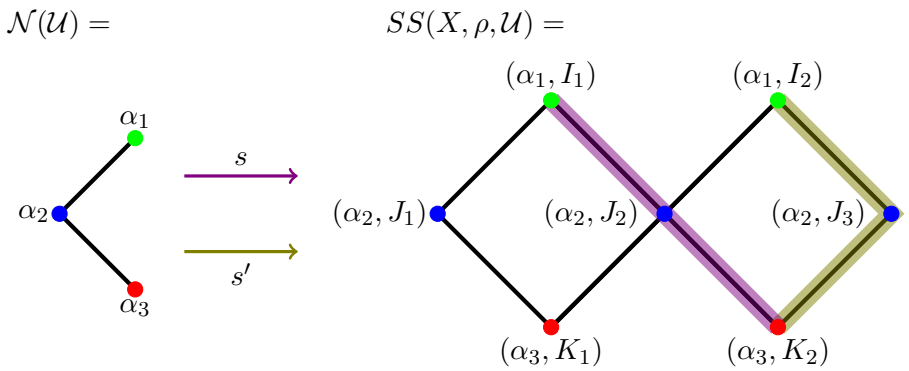


Figure 54: Two possible scale choices.

Note that even though for example $I_1 \cap K_2 = \emptyset$, there is a scale choice s where we can pick an ϵ in each I_1 and K_2. This is due to the fact that $I_1 \cap J_2 \neq \emptyset$ and $K_2 \cap J_2 \neq \emptyset$. So J_2 connects the two stability intervals together and ensures continuity in the choices for ϵ. Hence, the difference in choices for ϵ are allowed to be quite different in different regions, as long as this continuity is preserved.

Let us review what the method offers.

- There is a kind of continuity in the choice of the scale parameter varying with α. Likely the data sets will be of such a nature that continuously varying the ϵ as above is the natural way to go. For example, we most likely want to choose small epsilons in high density areas and bigger epsilons in less dense areas.

- Stability intervals have a notion of length, hence it is possible to compare different choices of s. The intuitive idea is again that scale choices that are stable over long ranges of parameter choices is better than the more unstable ones. For example, we can sum up all the stability intervals corresponding to a scale choice s, maybe even with some weights, and make comparisons.

8.7 Mapper on population data

In this test an implementantion of the Mapper method in MATLAB is used, and the author is Gurjeet Singh. The results have been visualized using an open source graph visualization software called Graphviz.

We are given samples from an unknown population of people, which we want to know more about. The total amount of 3150 samples have been taken from this population. For each sample two measurements have been taken.

- The current radius of their eyes, that is the current size of their eyes, for which it is believed to indicate how much experience an individual is taking in at the moment.

- The curvature of the mouth, indicating anything from a smiling to a neutral to a sad person.

Furthermore, in this test we assume, for obvious reasons, that a persons eyes or smile cannot change instantaniously to another form, but has to transform in some continuous way. So no instant jumping from a smiling mouth to a sad mouth without some transitory fase. Now, the eyes and mouth give us two parameters for each individual, and hence we can map each individual onto the plane in $\mathbb{R}^2$ giving us a point cloud $X \subset \mathbb{R}^2$. This is illustratet in Figure 55. The data set obtained from the 3150 samples is shown in Figure 56. The color of each point indicates the value of the filter map on that point.

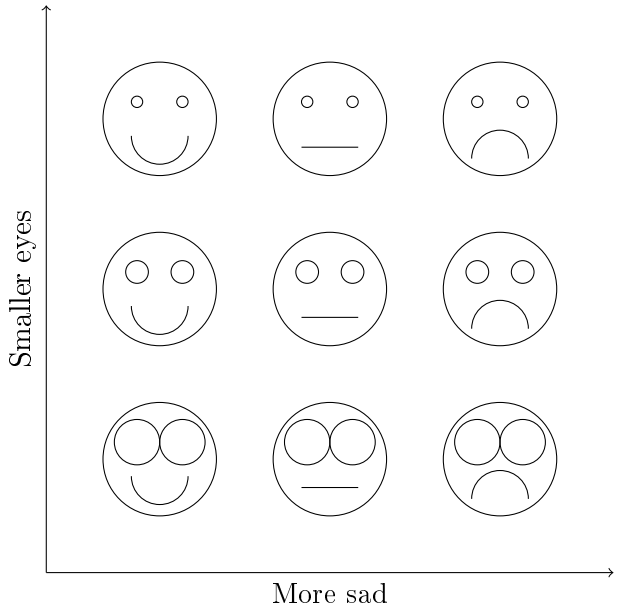


Figure 55: Individuals mapped to the plane according to the size of their eyes and curvature of mouth.

The filter map chosen for this test is a kernel density function $\rho : X \to \mathbb{R}$, as introduced earlier. The Mapper implementation sets the parameter space to be $Z = [min(\rho), max(\rho)]$. A covering of this space is then created by first defining subintervals of length $(max(\rho) - min(\rho)) * resolution$, where the resolution is an input parameter for Mapper. Then another input parameter specifies the percentage overlap between the intervals, giving us a covering of Z. So, let us see if we can detect any structure with the Mapper method.

In Figure 57 the results for different resolutions are shown. The percentage of overlap is 25 percent. As we can see the finer the resolution the finer structures we obtain. The main structure of this data set seems to be best represented by the resolution in Figure 57(c), i.e. three connected components. Resolutions in Figure 57(a) and 57(b) seems

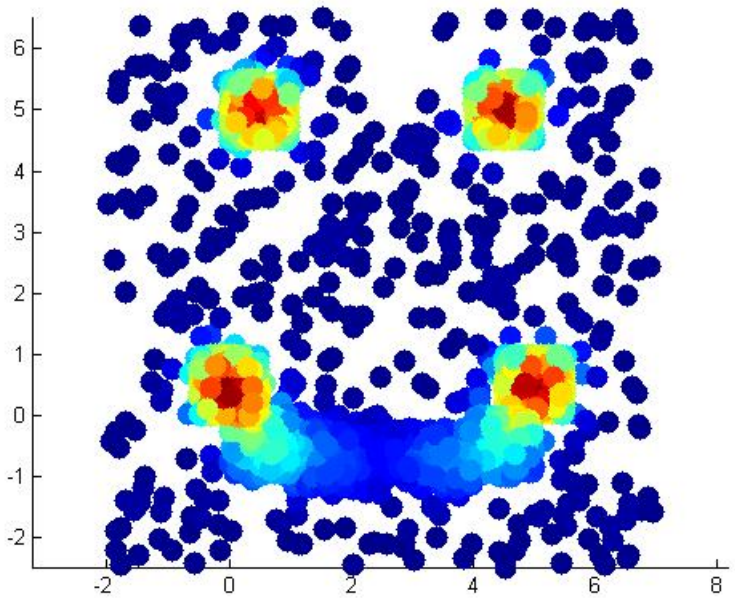
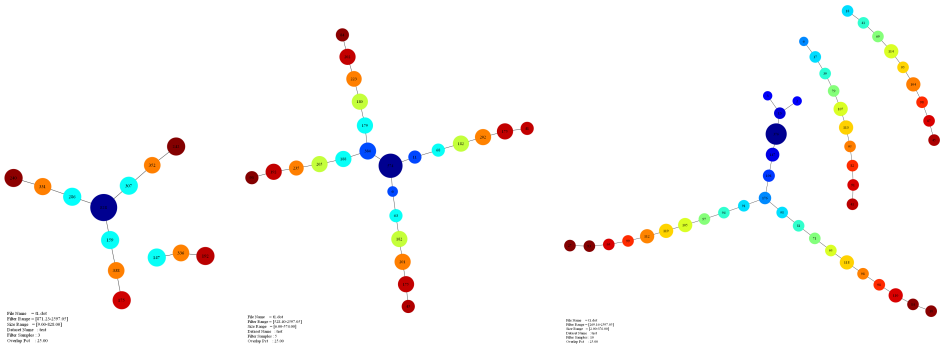
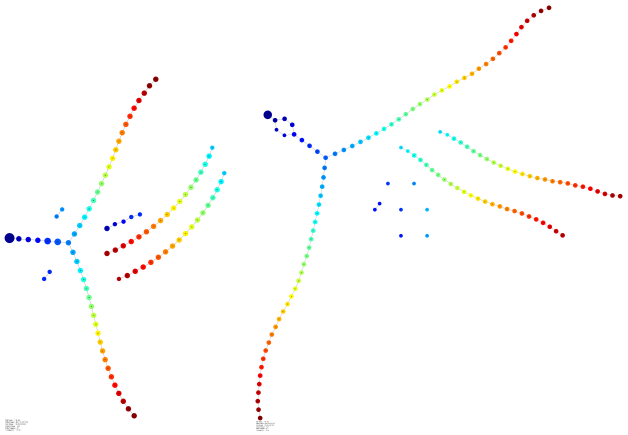


Figure 56: Population sample mapped to the plane. Red indicates high filter value, while blue indicates low filter value, that is high and low density estimates. Each sample point is placed accordingly to the mapping illustrated in Figure 55.

to be too coarse compared to the others, while Figures 57(d) and 57(e) starts to show properties that seem likely to be artifacts.



(a) Resolution = 1/3 (b) Resolution = 1/5 (c) Resolution = 1/10



(d) Resolution = 1/20 (e) Resolution = 1/30

Figure 57: The Mapper method has been applied to the data set in Figure 56. The filter map is a kernel density estimator. The percentage overlap was chosen to be 25 percent, and Figures 57(a) to 57(e) shows the results for different resolutions. The size of the nodes indicates the number of data points that belongs to this node, and the color indicates the average value of the filter map on the points belonging to this node. Again red indicates a high value and blue a low value.

These results may indicate a few properties about this population and we use Figure 57(c) as a reference. For the largest component with three endpoints(or five if we count the two small blue ones) we have that

- the points corresponding to the flare with the bluest nodes are tracked down to be points all over the plane, which explains the the low density represented by the blue color. This is about 20 percent of the data set.
- the middle node that connects the three flares consists of individuals with a neutral mouth and big eyes. This indicates individuals having a shock, as a lot of experience is taken in, but they have not yet grasped the consequences, which explains the neutral mouth.
- the two other flares, starting from the node in the middle of them, represents individuals that goes from big eyes to medium sized eyes, and from neutral mouth to either very happy or very sad. With increasingly higher density the further we move out in these two flares. They are about 25 percent each of the data set.

For the two other connected components we have that

- the points representing individuals in these components are all individuals with very small eyes, and they are either happy or sad. They are about 15 percent each of the data set.

It seems that about 50 percent of the individuals in this test moves between being very sad and very happy, after having some kind of shock, and that most of the time they resides in either of these states osr, after having a shock, about 25 percent go slightly mad and behaves noisy. Furthermore it seems that 30 percent of the population just stay either very happy or very sad, which makes sense due to their small eyes, as they do not take in much experience and hence does not experience much that alters their state of being.

References

[1] P. B. Bhattacharya, S. K. Jain, and S. R. Nagpaul. *Basic Abstract Algebra.* Cambridge University Press, Cambridge, second edition, 1994.

[2] N. Bourbaki. *Eléments de Mathématique - Algèbre (Chapitres 1 à 3).*

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

[4] G. Carlsson and F. Memoli. Persistent clustering and a theorem of j. kleinberg. 2008.

[5] G. Carlsson and V. D. Silva. Topological approximation by small simplicial complexes. Technical report, MISCHAIKOW, AND T. WANNER, 2003.

[6] G. Carlsson, A. Zomorodian, A. Collins, and L. Guibas. Persistence barcodes for shapes. In *SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 124–135, New York, NY, USA, 2004. ACM.

[7] V. de Silva. A weak definition of delaunay triangulation, 2003.

[8] V. de Silva and G. Carlsson. Topological estimation using witness complexes. In M. Alexa and S. Rusinkiewicz, editors, *Eurographics Symposium on Point-Based Graphics*. The Eurographics Association, 2004.

[9] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, July 2002.

[10] J. W. Frances Kirwan. *An Introduction to Intersection Homology Theory*. Chapman & Hall/CRC, 2006.

[11] R. Ghrist. Barcodes: The persistent topology of data. Technical report, 2007.

[12] A. Hatcher. *Algebraic Topology*. Cambridge University Press, November 2001.

[13] A. Krowne and N. Egge. Planetmath: a collaborative, web-based mathematical encyclopedia, 2001.

[14] S. Kwon, Y.-H. Chu, H.-S. Yi, and C. Han. Dna microarray data analysis for cancer classification based on stepwise discriminant analysis and bayesian decision theory. 2001.

[15] S. Lafon and A. B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1393–1403, 2006.

[16] J. M. Lee. *Introduction to Topological Manifolds (Graduate Texts in Mathematics)*. Springer, May 2000.

[17] I. Madsen and J. Tornehave. *From Calculus to Cohomology*. Cambridge University Press, Cambridge, UK, 1997.

[18] T. Marley. Graded rings and modules, 1993.

[19] G. McLachlan. Classification of microarray gene expression data. 2003.

[20] J. Munkres. *Topology*. Prentice Hall, December 1999.

[21] A. I. Saeed. Introduction to microarray analysis and mev. 2005.

[22] B. W. Silverman. *Density estimation: for statistics and data analysis*. London, 1986.

[23] R. Simon, M. D. Radmacher, K. Dobbin, and L. M. McShane. Pitfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification. *J. Natl. Cancer Inst.*, 95(1):14–18, 2003.

[24] G. Singh, F. Memoli, and G. Carlsson. Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. pages 91–100, Prague, Czech Republic, 2007. Eurographics Association.

[25] G. K. Smyth, Y. H. Yang, and T. Speed. Statistical issues in cdna microarray data analysis. pages 111–136. 2003.

[26] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, December 2000.

[27] J. W. Vick. *Homology Theory: An Introduction to Algebraic Topology*. Springer, 2nd edition, January 1994.

[28] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete Comput. Geom*, 33:249–274, 2005.

[29] A. Zomorodian and G. Carlsson. Localized homology. *Computational Geometry*, 41(3):126–148, November 2008.

[30] A. J. Zomorodian. Computing and comprehending topology: Persistence and hierarchical morse complexes, 2001.