UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2023:51

Degree project 30 credits

Master`s Programme in Mathematics

November 2023

# Persistent Homology

## A Modern Application of Algebraic Topology in Data Analysis

Staffan Leijnse

**Department of Mathematics, Uppsala University**

Supervisor: Julian Külshammer

Examiner: Michele Del Zotto

# Persistent Homology

**A Modern Application of Algebraic Topology in Data Analysis**

Staffan Leijnse

October 13, 2023

# Contents

# 1. Introduction

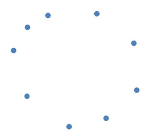
In this paper we survey some machinery for recovering the topological structure of a probability distribution on a metric space, or dense regions of that distribution, from a sample. The latest in topological data analysis can be traced back to 1992 but it wasn't until 2005 that the methodology was sufficiently developed for mainstream usage across a large number of fields, see [Per18]. In presenting the mathematics behind persistent homology some papers have taken an approach which is far too complicated for an average engineer, assuming familiarity with category theory, homology and commutative algebra. Other papers have been very focused on applications with most of the mathematical theory being neglected or hidden. This paper takes a different route and presents the required theory in a way that only requires a well developed understanding of basic linear algebra along with a basic knowledge of what topology can describe. It gives a novel and constructive proof of the main structure theorem within persistent homology, designed to only use basic mathematics and justify the intuitive way of interpreting barcodes. The mathematical approach may resemble the approach taken in [Rin16], where the goal was to classify the indecomposables of so called Dynkin Quivers. In [Oud15] the reader can find a more comprehensive, but more abstract, review of topological data analysis from a quiver representation point of view. The structure theorem presented in this paper is closely related to what is usually called the Krull-Schmidt decomposition theorem which proves existence and uniqueness of a so called Remak decomposition, found in [Gri07], p. 343. This paper ends with a presentation of this structure theorem using objects that are more standard in abstract algebra, namely graded modules.

## 1.1. The Shape of Data

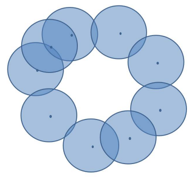
Samples, or finite point clouds, don't themselves have interesting topological properties. It is however often reasonable to assume that every point could have been slightly different, depending on what we mean by "slightly". In mathematical terms this means assuming that there is some $\epsilon$ such that every point has a ball surrounding it which is included in the space we sampled from. Let $X \subseteq \mathbb{R}^n$ be the space of possible samples and let $\mathbb{X} \subseteq X$ be a set of samples, i.e. a finite set of points. Now let $B^\epsilon(\mathbb{X})$ be the space of open $\epsilon$-balls around each point in $\mathbb{X}$. Then our assumption would be that $B(\mathbb{X}, \epsilon) \subset X$ for some $\epsilon$. Depending on our choice of $\epsilon$, $B(\mathbb{X}, \epsilon)$ could have very different shape, see Figure 1.

While the assumption that $\epsilon$-balls are part of $X$ for some $\epsilon$ is often reasonable, it is far more difficult to justify any specific choice of $\epsilon$. It is however worth noting that for any choice of $\epsilon$ the real features of $X$ will start to emerge as the sample size grows. Seen from a different perspective we can see that the real features of $X$ will emerge for a wider range of $\epsilon$ values when we have a larger sample size.

A natural question might be which is the best choice of $\epsilon$ but in the field of persistent homology, the approach is rather different. The philosophy behind persistence is to construct spaces using a range of different $\epsilon$ values and tracking which features (e.g connected components, loops etc.) persist over which parts of the $\epsilon$ range. Consider



(a) Does this sample come from a circle?



(b) Adding sufficiently large $\epsilon$ balls around each point makes $B(\mathbb{X}, \epsilon)$ into something resembling a circle.

Figure 1

the Figures 1 and 2. For sufficiently small $\epsilon$ we will get a space that is topologically the same as a point cloud. As $\epsilon$ grows we will start seeing points join together and thus the components become fewer, eventually being reduced to a single component. At some value of $\epsilon$ the loop is closed and we have a space topologically equivalent to the circle. Then for sufficiently large $\epsilon$ values the loop gets filled in and we have a space topologically equivalent to a single point. So as epsilon grows we start with a bunch of connected components which disappear until only one remains and we see a circle form and eventually disappear.

It is worth noting that there is a simple way to limit our investigation to the space of "likely" samples rather than the space of all possible samples. This involves removing points from $\mathbb{X}$ that are in sparse areas, sometimes called outliers. This creates a new $\mathbb{X}$ where the only sample points which remain are in areas where many samples have been found. We will not dive deeper into this topic but encourage the reader to think of $\mathbb{X}$ as either a set of samples from $X$ or a set of samples from $X$ with some points removed.

## 1.2. Bar Codes

Bar codes are a way of tracking and visualizing topological features which appear and disappear as $\epsilon$ grows. In the above figures we would have a bar for each topological feature, such as a connected component or a loop. Their start signifies the $\epsilon$ for which the feature appears and their end signifies the $\epsilon$ for which they disappear. This can be seen in Figure 3 taken from [LR21].

Allowing for a lack of precision one might call a barcode "a way to track the topological features of a sample surrounded by $\epsilon$-balls as $\epsilon$ is allowed to change". This paper presents the mathematics behind the creation of these barcodes in a way that (1) requires very

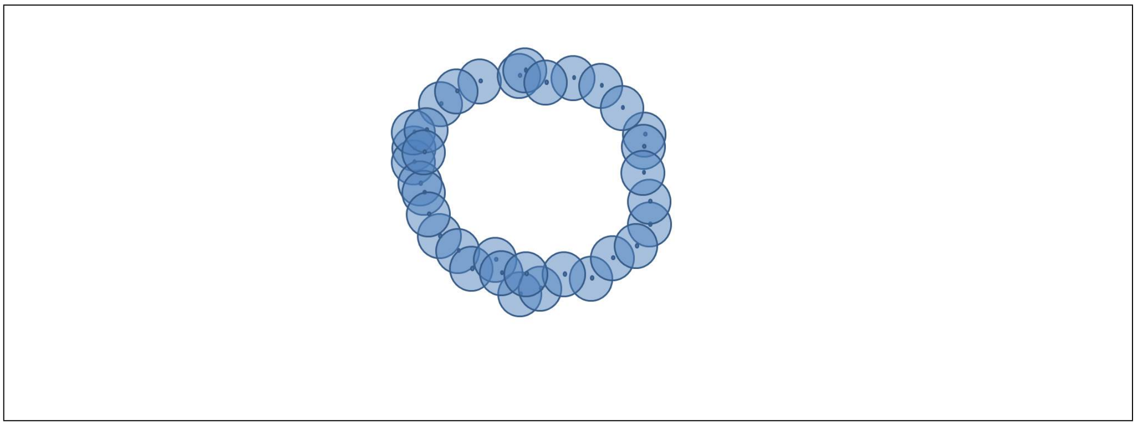


Figure 2: A smaller $\epsilon$ gives a circle as long as the sample is larger. Alternatively: with a larger sample we see the circle emerge for smaller $\epsilon$-values

little mathematical background, (2) is constructive and (3) closely tracks the intuitive way to interpret barcodes.

## 1.3. Applications

Persistent homology is one of the most popular methods within topological data analysis, see for example [TMO22]. It is used to study very general objects such as networks which touch on many fields, see [Meh19]. This makes it impossible to describe the applications and potential impact of persistent homology so instead we look at three areas as examples of applications for this methodology, namely (1) guiding robots and other machines as they figure out how to analyze and manipulate their environment, (2) processing and interpreting images and (3) understanding and diagnosing the human body.

### 1.3.1. Robotics and Machine Learning

According to [PXL22] there is a need for machine learning models which represent features in a way that can reduce data complexity while also preserving the important intrinsic information of the data. In a systematic review of persistent homology methodologies as applied to machine learning they claim that "Persistent homology provides a delicate balance between data simplification and intrinsic structure characterization, and has been applied to various areas successfully".

At a robotics lab, [Ant+21] deals with objects that are highly deformable under manipulation by a robot and consider the example of a robot hanging an apron on a hook. This requires the robot to detect and manipulate the hole formed by the straps of the apron based on visual input data. In the article they take a persistent homology approach and "prove that it is possible to detect significant topological features (such as the straps of an apron) and observe their dynamics from point clouds".

Machine learning using persistent homology has also been applied to fingerprint classification in [GGM19]. They classify fingerprints by collecting a point cloud from the

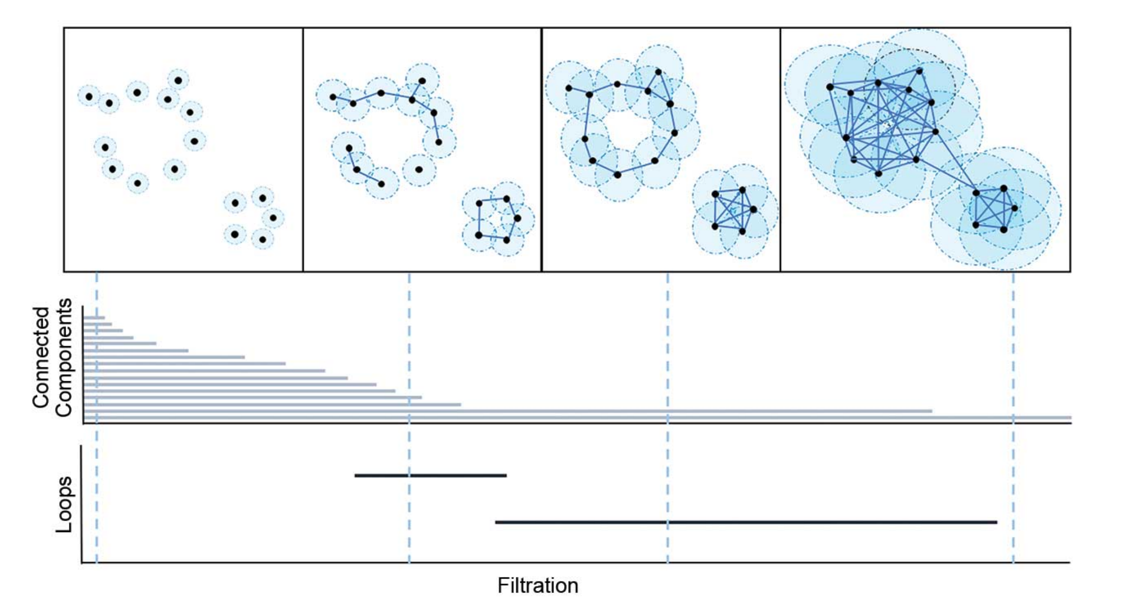



Figure 3

print and then producing the related barcodes.

### 1.3.2. Image processing

In [Chu09], researchers create a framework based on persistent homology for detecting and characterizing signals in images. According to the researchers, this technique is "general enough for dealing with noisy multivariate data" and is used to map the surface of brains. In [Woo17], researchers use persistent homology to process images in various ways such as removing noise or increasing resolution. The methodology proposed is shown to "outperform the existing state-of-the-art approaches".

### 1.3.3. Medical Data Analysis

Computer-aided diagnosis systems for cancer screenings are constantly being improved and some, such as [Qai+16], use persistent homology as a tool. In the paper the researchers "explore the degree of connectivity among nuclei using the novel idea of persistent homology profiles" which is an important step in cancer diagnosing.

Brains can be thought of as networks where different parts are connected. It is however hard to know which scale is most useful to look at. The approach in [Lee+11] is to use persistent homology as it allows researchers to capture connectivity data at all scales at the same time.

## 2. Mathematical Basis

### 2.1. Linear Algebra

In this section we present some basic definitions and results which will be useful in later sections. For a more thorough presentation, see [Axl15]. We start with equivalence relations and quotient spaces and then move on to some results in linear algebra which are not usually part of beginner courses on the subject.

**Definition 2.1.** Let $X$ be a set. By a *relation* $R$ on $X$ we mean a subset of $X \times X$ and write $x \sim_R y$ in place of $(x, y) \in R$. A relation is called an *equivalence relation* if it is (1) reflexive meaning that $x \sim_R x$ for all $x$, (2) symmetric meaning that $x \sim_R y \implies y \sim_R x$ and (3) transitive meaning that $x \sim_R y$ and $y \sim_R z$ implies that $x \sim_R z$.

When we have a relation on a set we will at times wish to extend it to be an equivalence relation. One may do this by adding relationships to $R$ to make it reflexive, symmetric and transitive. It turns out that this always generates an equivalence relation which is in fact the smallest equivalence relation containing $R$.

**Definition 2.2.** Let $R$ be a relation. By the *equivalence relation generated by* $R$ we mean the smallest equivalence relation which contains $R$ and we denote it $\hat{R}$.

**Definition 2.3.** Let $R$ be an equivalence relation on $X$. We define the *equivalence class* of $x \in X$ as $[x] = \{y \in X | x \sim_R y\}$. We define the *quotient of* $X$ *by* $R$ by $X/R := \{[x] | x \in X\}$.

**Definition 2.4.** Let $V$ be a vectorspace over a field $K$ and $W \subseteq V$ be a subspace. Let $\sim_W := \{(v_1, v_2) | v_1 - v_2 \in W\}$. We define the quotient of $V$ by $W$ as $V/W := V/ \sim_W$. By letting $[x] + [y] := [x + y]$ and $\lambda[x] := [\lambda x]$ we have a $K$-vector space structure on $V/W$.

The field $K$ will usually be either $\mathbb{R}$ or $\mathbb{Z}_2$ in this paper. The following is a basic property of vector spaces and is usually proven in basic courses on linear algebra.

**Proposition 2.5.** *$V/W$ is isomorphic to a subspace of $V$ and the dimension of $V/W$ is* $\dim(V/W) = \dim(V) - \dim(W)$.

Let $V$ and $W$ be vectorspaces over $K$ and $f : V \to W$ be a linear transformation. Another basic proposition in linear algebra is that $\operatorname{Im}(f)$ is a subspace of W. It will later be central to this paper to relate $V$ to $W/\operatorname{Im}(f)$ and we will start proving some simple results for this space. We define $\theta(f) = W/\operatorname{Im}(f)$ and present a way to simplify $\theta(f)$.

**Proposition 2.6.** *Let $f : V \to W$ be a linear transformation, $g : V \to V$ and $h : W \to W$ be invertible linear transformations. Then $\theta(f)$ is isomorphic to $\theta(hfg)$. Given a choice of bases for both spaces we get matrices $F, G$ and $H$ which correspond to $f, g$ and $h$. Then $G, H$ are invertible and $\theta(F) \cong \theta(HFG)$*

*Proof.* For any $v \in V$ we get $f(v) = fg(g^{-1}(v))$ so $\operatorname{Im}(f) = \operatorname{Im}(fg)$. This means $\theta(f) = W/\operatorname{Im}(f) = W/\operatorname{Im}(fg) = \theta(fg)$.
Now we notice that $h$ induces a map $h : W/\operatorname{Im}(f) \to W/\operatorname{Im}(hf)$ defined by $h(w + \operatorname{Im}(f)) = h(w) + \operatorname{Im}(hf)$. We check that this map is well-defined. Assume $w, w' \in W$ are such that $w + \operatorname{Im}(f) = w' + \operatorname{Im}(f)$. Then $w - w' \in \operatorname{Im}(f)$ so $\exists v \in V; w - w' = f(v)$. Now we need to check that $hw + \operatorname{Im}(hf) = hw' + \operatorname{Im}(hf))$ which is true since $hw + \operatorname{Im}(hf) - hw + \operatorname{Im}(hf) = h(w - w') + \operatorname{Im}(hf) = h(f(v)) + \operatorname{Im}(hf) = 0 + \operatorname{Im}(hf)$. Thus $h$ is well defined and it is easy to see that it is an isomorphism. So we get $\theta(f) = W/\operatorname{Im}(f) \cong W/\operatorname{Im}(hf) = \theta(hf)$. □

*Remark* 2.7. Elementary row operations can be performed by multiplying a matrix by an invertible matrix $H$ from the left and elementary column operations can be performed by multiplying by an invertible matrix $G$ on the right. Thus the above proposition says that elementary row or column operations on a matrix $A$ do not affect $\theta(A)$, up to isomorphism.

We will be dealing with particular sequences of vector spaces, for which we now prove some preliminary and very important results.

**Proposition 2.8.** *Let $U, V$ and $W$ be $K$-vectorspaces and assume a given basis, let $b : U \to V$ and $a : V \to W$ be linear transformations with matrices $B$ and $A$. Assume $AB = 0$. Let also $f : U \to U$, $g : V \to V$ and $h : W \to W$ be invertible linear transformations with matrices $F, G$ and $H$. Then*

$$\operatorname{Ker}(A)\big/_{\operatorname{Im}(B)} \cong \operatorname{Ker}(HAG^{-1})\big/_{\operatorname{Im}(GBF)} \tag{1}$$

*Note* 2.9. We prove the proposition for the linear maps $a, b$ but note that it holds for their matrices as well.

*Proof.* First we note that $\operatorname{Im}(b) \subseteq \operatorname{Ker}(a)$ since $a \circ b = 0$. In the same way as in the proof above we get $\operatorname{Ker}(a)/\operatorname{Im}(b) = \operatorname{Ker}(a)/\operatorname{Im}(b \circ f)$. Observing that $\operatorname{Ker}(a) = \operatorname{Ker}(h \circ a)$ gives us $\operatorname{Ker}(a)/\operatorname{Im}(b) = \operatorname{Ker}(h \circ a)/\operatorname{Im}(b \circ f)$. If we now prove that $\operatorname{Ker}(a)/\operatorname{Im}(a) \cong \operatorname{Ker}(ag^{-1})/\operatorname{Im}(g \circ b)$ then we are done.

We now attempt to create an isomorphism $\hat{g} : \operatorname{Ker}(a)/\operatorname{Im}(b) \to \operatorname{Ker}(a \circ g^{-1})/\operatorname{Im}(g \circ b)$. This isomorphism will be given by $g$ acting on $\operatorname{Ker}(a)/\operatorname{Im}(b)$ in the natural way. Thus let $\hat{g}$ be defined by $\hat{g}(v + \operatorname{Im}(b)) = g(v) + \operatorname{Im}(g \circ b)$.

To check that $\hat{g}$ is well defined we note that for any $v \in \operatorname{Ker}(a)$ we get $g(v) \in \operatorname{Ker}(a \circ g^{-1})$. Thus for any $v + \operatorname{Im}(b) \in \operatorname{Ker}(a)/\operatorname{Im}(b)$ we have $\hat{g}(v + \operatorname{Im}(b)) \in \operatorname{Ker}(a \circ g^{-1})/\operatorname{Im}(g \circ b)$. Also for $v, v' \in \operatorname{Ker}(a)$ such that $v + \operatorname{Im}(b) = v' + \operatorname{Im}(b)$ we have

$$\begin{aligned} \hat{g}(v + \operatorname{Im}(b)) - \hat{g}(v' + \operatorname{Im}(b)) &= gv + \operatorname{Im}(g \circ b) - g(v)' + \operatorname{Im}(g \circ b) \\ = g(v) - g(v') + \operatorname{Im}(g \circ b) &= g(v - v') + \operatorname{Im}(g \circ b) \\ &= 0 + \operatorname{Im}(g \circ b) \end{aligned}$$

since $v - v' \in \operatorname{Im}(b)$ and thus $g(v - v') \in \operatorname{Im}(g \circ b)$. This means that $\hat{g}$ is well-defined.

It is then easy to see that $\hat{g}$ is invertible with $\hat{g}^{-1}(v + \operatorname{Im}(g \circ b)) = g^{-1}v + \operatorname{Im}(b)$ in a similar way which completes the proof. □

*Remark* 2.10. This means that we may perform the following operations on the matrices $A$ and $B$ without changing $\operatorname{Ker}(A)/\operatorname{Im}(B)$:

1. Any row operation on $A$
2. Any column operation on $B$
3. Any column operation on $A$ along with the same operation, with reversed sign, on $B$, for example we may add $k$ times column $i$ to column $j$ in $A$ as long as we also add $-k$ times row $i$ to row $j$ in $B$.

**Proposition 2.11.** *Let $A$ and $B$ be matrices with entries from a field $K$ and $AB = 0$. Then we may perform row and column operations to get matrices $A'$ and $B'$ with*

$$\operatorname{Ker}(A)\big/\operatorname{Im}(B) \cong \operatorname{Ker}(A')\big/\operatorname{Im}(B')$$

*and*

$$A' = \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_m \end{bmatrix}, \tag{2}$$

*where $m, n \in \mathbb{N}$ and $0$ represents $0$-matrices of appropriate size.*

*Proof.* Through row operations on $A$ we can make it into row reduced echelon form and through column operations we may make it into $A' = HAG^{-1}$ of the form described in this proposition. According to Proposition 2.8, $\operatorname{Ker}(A')/\operatorname{Im}(B')$ will then be the same as $\operatorname{Ker}(A)/\operatorname{Im}(B)$ if we apply those row operations as column operations on $B$ to get $B' = GB$. Since $A'B' = HAG^{-1}GB = HAB = 0$ and

$$A'B' = \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} B'_{11} & B'_{12} & B'_{13} \\ B'_{21} & B'_{22} & B'_{23} \\ B'_{31} & B'_{32} & B'_{33} \end{bmatrix} = \begin{bmatrix} B'_{11} & B'_{12} & B'_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3}$$

we may conclude that $B'_{11} = B'_{12} = B'_{13} = 0$. Therefore we have

$$B' = \begin{bmatrix} 0 & 0 & 0 \\ B'_{21} & B'_{22} & B'_{23} \\ B'_{31} & B'_{32} & B'_{33} \end{bmatrix}. \tag{4}$$

Since $A'$ has all zeros on column $l$ for any $l > n$, we can perform any row operations on $B'$ which only involve rows $l > n$, as well as any column operations without changing $\operatorname{Ker}(A')/\operatorname{Im}(B')$ according to Proposition 2.8. Through these row operations we can make $B'$ into row reduced echelon form, but on the bottom right instead of top left. Through column operations we can then transform $B'$ into the form stated in this proposition. $\square$

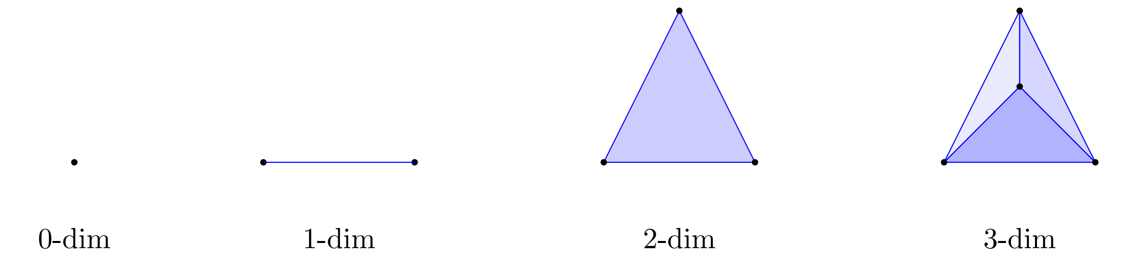


Figure 4: Simplices of dimension 0-3 containing 1-4 points

Later we will study families of $K$-vectorspaces called the chain groups $C_d$ along with linear maps $\partial_d : C_d \to C_{d-1}$ with $\partial_{d-1}\partial_d = 0$.

$$\ldots \to C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0 \tag{5}$$

We will then construct the homology groups $H_d := \mathrm{Ker}(\partial_d)/\,\mathrm{Im}(\partial_{d-1})$ and get another family of vector spaces $H_d$. In this section we have introduced the algebraic tools needed for this construction and we have proven that the $\partial$ functions can be simplified in a way that makes $H_d$ very easy to calculate using only linear algebra.

## 2.2. Simplicial Complexes

The topological properties of a space are often calculated by creating a more computation friendly space with the same properties. This is the approach which we will also take here by replacing the space of $\epsilon$-balls with a simplicial complex, i.e a set of "triangles" in various dimensions.

Let $X$ be the euclidean space $\mathbb{R}^n$ and $\mathbb{X}$ be a finite subset of $X$. Let $\epsilon$ be a fixed positive real number in this section. In future sections we will let $\epsilon$ vary and track the resulting changes in our constructions.

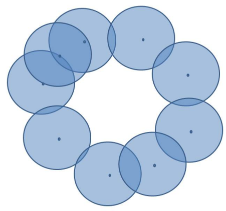
**Definition 2.12.** We define the *Delauney-Čech Cover* of $\mathbb{X}$ as $B^\epsilon(\mathbb{X})$, i.e the covering formed by putting open balls of radius $\epsilon$ around the points of $\mathbb{X}$.

A simplex is intuitively thought of as a generalization of the triangle to $d$ dimensions. See Figure 4 or, even better, the notes in [Sch19] for a more in depth exploration of simplices.

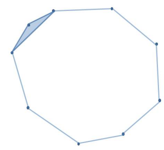
**Definition 2.13.** Let $S$ be a finite set. By an *abstract simplicial comlex* $\theta$ we mean a set of subsets of $S$ such that if $s \in \theta$ and $s' \subset s$ then $s' \in \theta$. A set $s \in \theta$ with $|s| = d+1$ is called a *d-dimensional simplex* of $\theta$. The 0-dimensional simplices are also called the *vertices* of $\theta$.

**Definition 2.14.** Given an open cover $\{U_i\}_{i\in I}$ of $\mathbb{X} = \{x_i\}_{i\in I}$, we construct a simplicial complex called the *nerve* of that covering, $N(\{U_i\})$, as follows:

For any $J \subseteq I$ let the abstract $n$-simplex $\{x_j\}_{j\in J}$ be included in $N(\{U_i\})$ iff the intersection $\bigcap_{j\in J} U_j$ is not empty.



(a) A DC-cover $B^\epsilon(\mathbb{X})$



(b) The realization of the DC-complex $DC^\epsilon(\mathbb{X})$

Figure 5: $B^\epsilon(\mathbb{X})$ is homotopy equivalent to $DC^\epsilon(\mathbb{X})$

*Note* 2.15. The nerve includes a point $x_i$ for each open set and a higher dimensional simplex whenever the corresponding open sets intersect.

**Definition 2.16.** We define the *Delauney-Čech Complex of* $\mathbb{X}$ as $DC^\epsilon(\mathbb{X}) = N(\{B^\epsilon(\{\mathbb{X}\}\})),$ i.e the nerve of the DC-covering with the same $\epsilon$.

**Definition 2.17.** Let $s = \{x_j\}_{j\in J} \subseteq \mathbb{X}$ be a simplex. By the *realization* of $s$ we mean $|s| = \{\sum \alpha_j x_j$ where $\sum \alpha_j = 1\}$. Let $\sigma = \{s_1, s_2, \ldots, s_n\}$ be a simplicial complex with $s_i \subseteq \mathbb{X}$. By the *realization* of $\sigma$ we mean $|\sigma| = \cup_{i\leq n}|s_i|$.

When the set $\mathbb{X}$ is obvious it will be omitted, as will sometimes $\epsilon$, so that $DC^\epsilon(\mathbb{X})$ is simply called the DC-complex and $B^\epsilon(\mathbb{X})$ is called the DC-cover.

We will not dive deep into the proof that the DC-cover of a finite point cloud in $\mathbb{R}^n$ is topologically the same as the realization of its DC-complex since the topology needed is elementary and also outside the scope of the algebraic approach taken in this paper. We will just outline the ideas involved. The DC-cover is a union of open balls, which are all convex sets. The intersections of convex sets are again convex sets, which are all contractible. According to the nerve theorem, originally proven in [Ale28], a cover where all intersections are contractible is homotopy equivalent to the realization of its nerve.

**Example 2.18.** In Figure 6 we see a sample, we name it $\mathbb{Y}$, of 4 different points in $\mathbb{R}^2$. First we see the DC-cover, $B^\epsilon(\mathbb{Y})$ for four different $\epsilon$-values. If we allow ourselves to abuse notation by writing $a$ instead of $\{a\}$ and $ab$ instead of $\{a, b\}$ we get the nerve of this covering as:

$$DC^{\epsilon_0}(\mathbb{Y}) = \{\{a\}, \{b\}, \{c\}, \{d\}\} = \{a, b, c, d\}$$

$$DC^{\epsilon_1}(\mathbb{Y}) = \{a, b, c, d, bc\}$$

$$DC^{\epsilon_2}(\mathbb{Y}) = \{a, b, c, d, bc, ab, ac, bd, cd\}$$

$$DC^{\epsilon_3}(\mathbb{Y}) = \{a, b, c, d, bc, ab, ac, bd, cd, abc, bcd\}$$

In Figure 6 we also see the realizations of these DC-complexes.

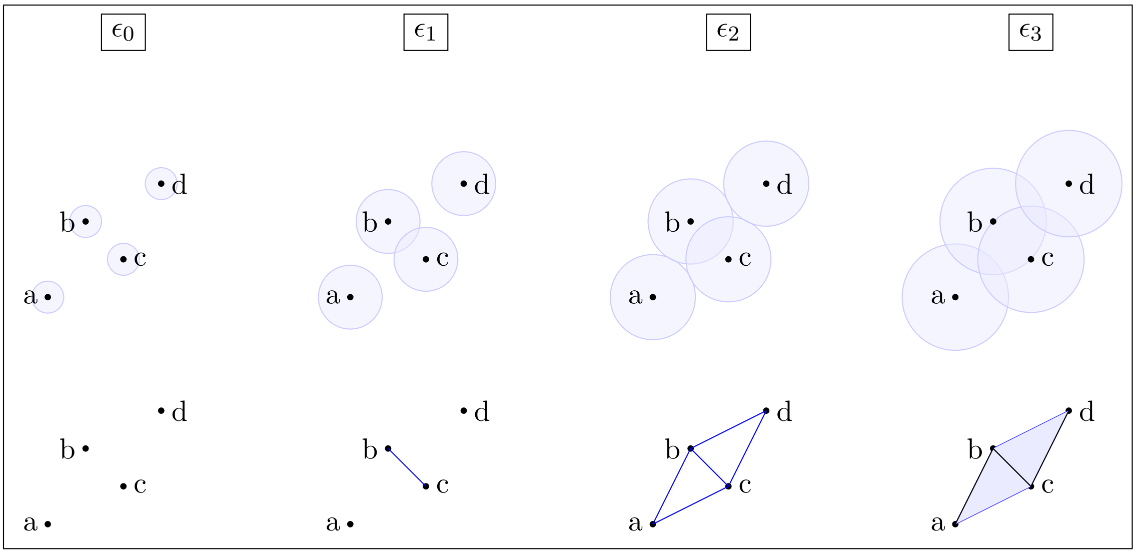


Figure 6: The DC-cover $B^\epsilon(\mathbb{Y})$ followed by the realization of the DC-complex, $|DC^\epsilon(\mathbb{Y})|$, for four different $\epsilon$-values

## 2.3. Simplicial Homology

In algebraic topology the idea is to find so called invariants of topological spaces. These invariants are constructions that are the same for all homotopy equivalent spaces. In this section we construct homology groups of simplicial complexes which are invariant. This means we may classify a family of homotopy equivalent spaces by the homology groups of their DC-complexes. After constructing such homology groups we give a simple way to calculate them using linear algebra.

**Definition 2.19.** Given a simplicial complex $S$ we define $C_d(S)$, the *d-th Chain Group* of $S$, to be the $\mathbb{Z}_2$-vector space with the $d$-simplices of $S$ as its basis. We let $C_d(S) := \{0\}$ for $d < 0$.

Since the chain group of a DC-complex will be important later, we name it $C_d^\epsilon := C_d(DC^\epsilon(\mathbb{X}))$. We also order the basis for this chain group by the $\epsilon$-value for which the basis vector appears in the underlying DC-complex. Thus let $B_d^\epsilon := \{b^\epsilon_{(d,1)}, b^\epsilon_{(d,2)}, \dots b^\epsilon_{(d,n)}\}$ be the $d$-simplices of $DC^\epsilon(\mathbb{X})$ ordered in such a way that the maximum pairwise distance between points in $b^\epsilon_{(d,m)}$ is not larger than the maximum pairwise distance between the points in $b^\epsilon_{(d,n)}$ whenever $n > m$. Then we get a basis for the vector spaces $C_d^\epsilon$ which grows as $\epsilon$ grows by adding an additional basis vector to the set $B_d^\epsilon$.

Assume that the points of $\mathbb{X}$ are ordered so that $\mathbb{X} = \{x_1, x_2, \dots, x_n\}$. Later it will be useful to consider the full set of potential simplices of a certain sample $\mathbb{X}$. Let the set of subsets of $\mathbb{X}$ be defined as $\mathbb{X}^{\mathbb{X}} = \{(x_{j_1}, x_{j_2}, \dots, x_{j_k}); 1 \leq k \leq |\mathbb{X}|, j_1 \leq j_2 \leq \dots, \leq j_k\}$.

**Definition 2.20.** For any sequence of natural numbers $j_1 < j_2 < \dots < j_n$ and any set $\{x_{j_1}, x_{j_2}, \dots x_{j_k}\}$ let $\{x_{j_1}, x_{j_2}, \dots x_{j_k}\}^{(m)} := \{x_{j_1}, x_{j_2}, \dots x_{(j_{m-1})}, x_{j_{(m+1)}}, \dots x_{j_k}\}$. In other words, when we put a superscript $m$ after the description of a simplex, this refers to that same simplex but without the $m$-th vertex in the ordering of $\mathbb{X}$.

**Definition 2.21.** Let $S$ be a simplicial complex. We define $\delta_d : C_d(S) \to C_{(d-1)}(S)$ on the basis of $C_d(S)$ as

$$\delta_d(\{x_{j_1}, \dots x_{j_d}\}) = \sum_{k \leq d} \{x_{j_1} \dots x_{j_d}\}^{(k)}.$$

$\delta_d$ is then linearly extended to the rest of $C_d(S)$. $\delta_d$ is called the $d$th *boundary operator.*

It is often notationally useful to not include the index $d$ in $\delta_d$ and regard $\delta$ as a function from the sum of all $C_d(S)$ to itself and simply call $\delta$ the boundary operator. We will however never apply $\delta$ to an object which is not contained in $C_d(S)$ for some $d$. We also often restrict $\delta$ to $C_d(S)$ for some simplicial complex $S \subseteq \mathbb{X}^{\mathbb{X}}$ without changing notation. Note that the sum in the definition of $\delta$ is well defined since $\{x_{j_1}, \dots x_{j_d}\} \in S \implies \{x_{j_1} \dots x_{j_d}\}^{(k)} \in S$ for any simplicial complex $S$.

**Proposition 2.22.** $\delta^2 = 0$

*Note* 2.23. We will later use the equivalent statement that for any $d \in \mathbb{N}, \mathrm{Im}(\delta_d) \subset \mathrm{Ker}(\delta_{(d-1)})$, the image of $\delta_d$ is a subset of the kernel of $\delta_{(d-1)}$.

*Proof.* We first prove that $\delta^2 = 0$ for all basis vectors in $C_d(S)$ which by the linearity of $\delta$ will mean that $\delta^2 = 0$ on all elements in $C_d(S)$. Assume thus that $\{x_1, \dots, x_i\} \in S$ and apply $\delta$ twice. We note that the product will be a linear combination of the type

$$\sum (\{x_1, \dots x_i\} \setminus \{x_j, x_k\}) \in C_{d-2}(S)$$

where $j, k \leq i$ and $j \neq k$. But for any such $j, k$ we get $(\{x_1, \dots x_i\} \setminus \{x_j, x_k\}) = (\{x_1, \dots x_i\} \setminus \{x_k, x_j\})$ and each appears once in the linear combination. Therefore each simplex appears $1 + 1$ times in the sum so they cancel since $1 + 1 = 0$ in $\mathbb{Z}_2$. $\square$

*Note* 2.24. The field $\mathbb{Z}_2$ was chosen to simplify the definition of $\delta$ and make this proof simple. It is not hard to do the same thing over general fields by making the sum in the definition of $\delta$ signed or by introducing a concept called orientation of simplices. This generalization complicates the calculations but gives the same algorithm and is therefore omitted.

It is natural to create a matrix representation of $\delta_d$. We let $r(s)$ and $c(s')$ be the row and column representing the basis vector $s \in C_d(\mathbb{X}^{\mathbb{X}})$ and $s' \in C_{d-1}(\mathbb{X}^{\mathbb{X}})$. The $k$-th column of $\delta_d$ will have a 1 entry in the $j$-th position iff $b_{(d-1,j)} \subseteq b_{(d,k)}$ and otherwise be 0. In other words the columns of $\delta_d$ have a 1 in the rows corresponding to a simplex inside the simplex representing the column.

**Example 2.25.** Let $\mathbb{Y}$ be the same as in Example 2.18 and let $S_i := N(B^{\epsilon_i})$ . Then $C_d(S_i)$ is the $\mathbb{Z}_2$-vectorspace over the $d$-simplices of $S_i$, the simplicial complex corresponding to $\epsilon_i$. Since $S_3 = \{a, b, c, d, ab, ac, bc, bd, cd, abc, bcd\}$ we see that $v \in C_1(S_3)$ means $v = \alpha_0 ab + \alpha_1 ac + \alpha_2 bc + \alpha_3 bd + \alpha_4 cd$ where each $\alpha_j \in \mathbb{Z}_2$.

Let $v = abc, w = abc + bcd \in C_2(S_3)$. Looking at the picture and trying to figure out what a boundary might mean we would guess that the boundary of $v$ would include $ab$, $ac$ and $bc$ while the boundary of $w$ would include $ab$, $ac$, $bd$ and $cd$. Using the definitions established above we get the boundary of $v$ as $\delta(v) = ab + ac + bc \in C_1(S_3)$ and the boundary of $w$ as $\delta(w) = \delta(abc) + \delta(bcd) = ab + ac + bc + bc + bd + cd = ab + ac + bd + cd$ since $bc + bc = (1+1)bc = 0bc$ in a $\mathbb{Z}_2$-vector space.

We check that $\delta^2(v) = 0$.

$$\begin{aligned}\delta^2(v) &= \delta(ab + ac + bc) = \delta(ab) + \delta(ac) + \delta(bc)\\ &= a + b + a + c + b + c = (1+1)(a+b+c) = 0\end{aligned}$$

We know that $\delta_d$ has a matrix representation which does not depend on $\epsilon$ but we may restrict $\delta_d$ to $C_d(S_3)$. We then get the matrices

$$\delta_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \delta_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \delta_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}.$$

**Definition 2.26.** A *d-cycle* is an element of $C_d(S)$ in the kernel of $\delta_d$. A *d-boundary* is an element of $C_d(S)$ in the image of $\delta_{(d+1)}$. We use $Z_d(S)$ to denote the subspace consisting of cycles in $C_d(S)$ and $B_d(S)$ to denote the subspace of boundaries in $C_d(S)$.

**Definition 2.27.** The *d-th homology group* of a simplicial complex $S$ is defined as the quotient (of vector spaces)

$$H_d(S) = {Z_d(S)} \Big/ {B_d(S)}, \tag{6}$$

the space of cycles divided by the space of boundaries.

*Remark* 2.28. Both $Z_d$ and $B_d$ are subspaces of $C_d$ and $B_d \subseteq Z_d$ by Proposition 2.22. The quotient (6) is thus well defined.

When we construct our homology groups $H_d$ from a DC-complex, they will depend on $\epsilon$. Since this is the construction we will study we define $H_d^\epsilon := H_d(DC^\epsilon(\mathbb{X}))$

**Definition 2.29.** The $d$-th Betty number of a space is $\beta_d = \dim(H_d)$. When the space is dependent on $\epsilon$ we write $\beta_d^\epsilon = \dim(H_d^\epsilon)$.

It is worth noting that $\dim(H_d) = \dim(Z_d) - \dim(B_d)$. We can simplify $\delta$ using Proposition 2.11 and get $\dim(Z_d)$ as the number of zero columns in $\delta_d$ and $\dim(B_d)$ as the number of non-zero columns in $\delta_{d+1}$.

**Example 2.30.** Let $\mathbb{Y}$ be the same as in Example 2.18 and 2.25. We will now calculate the homology groups $H_d^\epsilon$ with $\epsilon = \epsilon_3$. Starting with $d = 1$ we get $H_1 = \mathrm{Im}(\delta_2) / \operatorname{Ker} \delta_1$. According to Proposition 2.8 we can perform (1) row operations on $\delta_2$, (2) column operations on $\delta_1$ and (3) column operations on $\delta_2$ along with the corresponding row operations on $\delta_1$ without changing $H_1$. We do these operations to get $\delta_2$ and $\delta_1$ to the form of Proposition 2.11. We get

$$\delta_2' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \delta_1' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$H_1 \simeq \mathrm{Ker}(\delta_1')/\operatorname{Im}\delta_2'$. $\mathrm{Ker}(\delta_1') = \mathrm{Im}(\delta_2') = \langle (1,0,0,0,0)^T, (0,1,0,0,0)^T \rangle$. Thus $H_1 = \{0\}$ which can be seen by the fact that there are no loops in Figure 6 for $\epsilon = \epsilon_3$. We get $\beta_1 = 0$.

If we instead let $\epsilon = \epsilon_2$ then $S_2 = \{0\}$ and therefore $\delta_2 = 0$ meaning that $H_1(S_2) = \langle (1,0,0,0,0)^T, (0,1,0,0,0)^T \rangle$ and $\beta_1 = 2$.

## 2.4. Functoriality

In the last section we fixed $\epsilon$ and created objects which depended on our choice of $\epsilon$. In this section we will allow $\epsilon$ to vary and create maps between the objects we get for different values of $\epsilon$. It is through these maps that we can connect the homology of $B^{\epsilon_j}(\mathbb{X})$ to the homology of $B^{\epsilon_k}(\mathbb{X})$ and track the changes in homological features that happen as $\epsilon$ increases. These maps will be shown to 'respect' the $\delta$ function defined above which is a feature called 'fuctoriality'. This functoriality feature will allow us to connect one loop in $B^{\epsilon_j}(\mathbb{X})$ to a loop in $B^{\epsilon_k}(\mathbb{X})$ and ask if they are in some sense the same loop.

For $\epsilon_j \leq \epsilon_k$ we can see that $B^{\epsilon_j}(\mathbb{X}) \subseteq B^{\epsilon_k}(\mathbb{X})$ and therefore $DC^{\epsilon_j}(\mathbb{X}) \subseteq DC^{\epsilon_k}(\mathbb{X})$. This also means that for any $d \in \mathbb{N}$ we get $C_d^{\epsilon_j} \subseteq C_d^{\epsilon_k}$ since the basis for $C_d^{\epsilon_j}$ is a subset of the basis for $C_d^{\epsilon_k}$. Now for any $d \in \mathbb{N}$ we may form the inclusion $\imath_{\epsilon_j,\epsilon_k} : C_d^{\epsilon_j} \to C_d^{\epsilon_k}$ by mapping $c \in C_d^{\epsilon_j}$ to $c \in C_d^{\epsilon_k}$ . To simplify notation we often call this map $\imath$ when the spaces are obvious.

We now assert that the boundary map $\delta$ from the previous section is functorial with respect to the inclusion map $\imath$, which is easily seen straight from the definitions.

$$\begin{tikzcd} C_d^{\epsilon_j} \arrow[r, "\imath"'] \arrow[d, "\delta"'] & C_d^{\epsilon_k} \arrow[d, "\delta"] \\ C_{d-1}^{\epsilon_j} \arrow[r, "\imath"'] & C_{d-1}^{\epsilon_k} \end{tikzcd}$$

**Proposition 2.31.** $\imath\delta = \delta\imath$

The inclusion map $\imath : C_d^{\epsilon_j} \to C_d^{\epsilon_k}, \epsilon_j \leq \epsilon_k$ can be restricted to the $d$-cycles $Z_d^{\epsilon_j}$ and will then be a map $\imath : Z_d^{\epsilon_j} \to Z_d^{\epsilon_k}$ since any $d$-cycle in $C_d^{\epsilon_j}$ will also be an $d$-cycle in $C_d^{\epsilon_k}$. We may therefore create a map $\hat{\imath} : H_d^{\epsilon_j} \to H_d^{\epsilon_k}$ by $z + B_d^{\epsilon_j} \mapsto z + B_d^{\epsilon_k}$. The map $\hat{\imath}$ is no longer neccessarily an inclusion since $B_d^{\epsilon_k}$ may be a bigger space than $B_d^{\epsilon_j}$ and therefore $z + B_d^{\epsilon_k}$ may be zero even if $z + B_d^{\epsilon_j}$ is not.

**Example 2.32.** Assume that a $d$-dimensional hole appears at $\epsilon_a$ and is filled in at $\epsilon_b$. Let $z \in Z_d^{\epsilon_a}$ be a loop around this hole. Then

$$\hat{\imath}_{\epsilon_a,\epsilon_r}(z + B_d^{\epsilon_a}) = \begin{cases} z + B_d^{\epsilon_r} \neq 0 & \text{if } \epsilon_a \leq \epsilon_r \leq \epsilon_b, \\ 0 & \text{otherwise} \\ . \end{cases}$$

If the hole never closed we could think of $\epsilon_b$ as infinity and the result still holds. If $d = 0$ then the loop around this 0-dimensional hole would be the same as a connected component. If $x_i$ is a point in $\mathbb{X}$ then this will represent a connected component of $B^\epsilon(\mathbb{X})$ for all $\epsilon$. For sufficiently large $\epsilon$ all $x_i$ would be identified as the same connected component but $x_i$ would still represent a non-zero element of $H_d^\epsilon$. We would therefore get $\epsilon_a = 0$ and $\epsilon_b = \infty$ since all points are representatives of some connected component for all $\epsilon$.

# 3. Persistent Homology

This section contains a presentation of persistence as well as the main results of this paper. The first subsection explains the concept of $P$-persistent objects which is going to be the structure upon which the later construction rests. In the following subsection we narrow our view to $\mathbb{N}$-persistent vector spaces and prove that these can be represented by barcodes.

## 3.1. Definitions

The abstract mathematical view of a $P$-persistent object is as a functor from a category formed by the partially ordered sets to another category. This paper takes a different and more accessible route, closely related to representation theory.

Perhaps the most intuitive way of viewing persistent objects is to simply draw a picture of an oriented graph and on each vertex we place an object, such as a vector space, as shown in Figures 7 and 8. For the remainder of this presentation we will limit ourselves to $P$-persistent vector spaces and we will only briefly consider cases other than $P = \mathbb{N}$.

**Definition 3.1.** A partially ordered set $P$ is a set along with a binary relation $\leq$ which is reflexive, antisymmetric, and transitive. I.e for any $r, s, t \in P$:

1. $r \leq r$ (reflexivity);
2. If $r \leq s$ and $s \leq r$ then $r = s$ (antisymmetry);
3. If $r \leq s$ and $s \leq t$ then $r \leq t$ (transitivity).

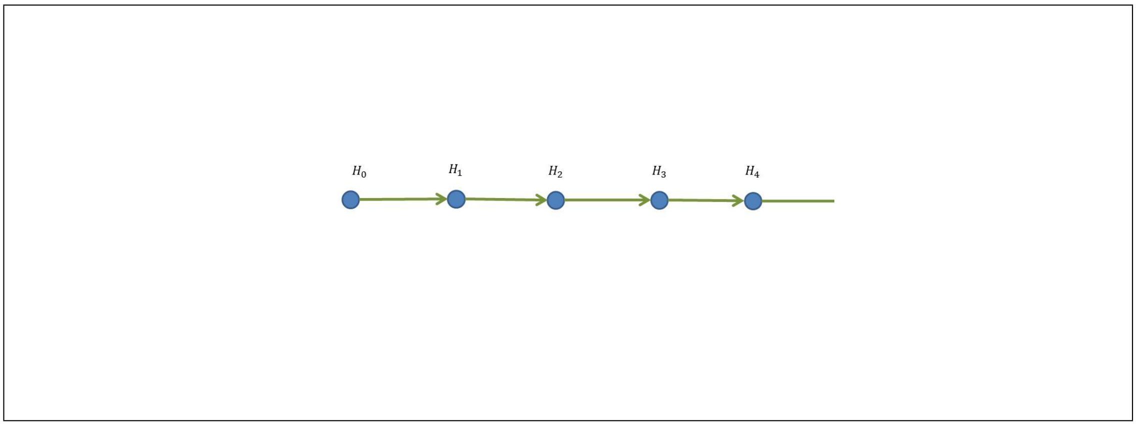



Figure 7: An $\mathbb{N}$-persistent object with vector spaces $H_i$ as objects

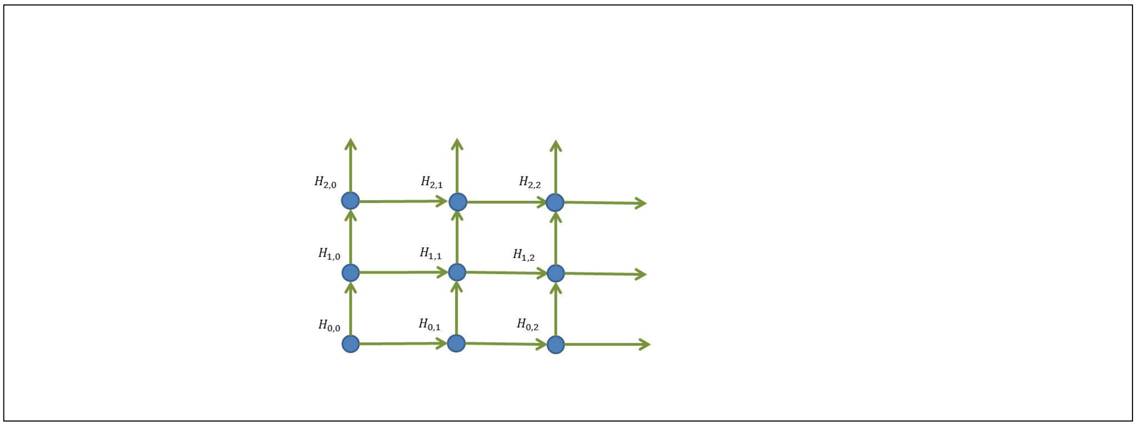



Figure 8: An $\mathbb{N}^2$-persistent object

**Example 3.2.** $\mathbb{N}$ and $\mathbb{R}$ are examples with the normal $\leq$ but we may also make $\mathbb{N}^n$ and $\mathbb{R}^n$ into partially ordered sets by imposing the so called product order. Let $x = (x_1, x_2, \ldots, x_n)$ be a vector in $\mathbb{R}^n$ and define $\leq$ by $x \leq x'$ iff $x_i \leq x'_i$ for all $i$. Then $\mathbb{R}^n$ is a partially ordered set with binary relation $\leq$. The same construction works for $\mathbb{N}^n$ and if $n = 2$ we get the relation corresponding to Figure 8.

**Definition 3.3.** Let $P$ be a partially ordered set. By a *$P$-persistent vector space* we mean:

1. A set of vector spaces $\{V_a\}$, one for each $a \in P$,
2. Linear transformations $\phi_{(a,b)} : V_a \to V_b$ whenever $a \leq b$ in $P$, such that $\phi_{(b,c)} \circ \phi_{(a,b)} = \phi_{(a,c)}$

A $P$-persistent vector space will be denoted $(\{V_a\}, \phi)$ even though $\phi$ is actually a set of functions.

**Example 3.4.** Let $(\{V_\epsilon\}, \phi)$ be an $\mathbb{R}$-persistent vector space. Since $\mathbb{R}$ is not a computer friendly space we wish to convert our $\mathbb{R}$-persistent vector space into a $\mathbb{N}$-persistent vector space. To do this we create a map $\epsilon_{(\cdot)} : \mathbb{N} \to \mathbb{R}$ to be any order preserving map, i.e $a \leq b \implies \epsilon_a \leq \epsilon_b$. Now, for any $a \leq b \in \mathbb{N}$ we get vector spaces $V_{\epsilon_a}, V_{\epsilon_b} \in \{V_\epsilon\}$ and a linear transformation $\phi_{(a,b)} : V_{\epsilon_a} \to V_{\epsilon_b}$ which fulfills the condition in the definition. Therefore $(\{V_{\epsilon_a}\}, \phi)$ is an $\mathbb{N}$-persistent vector space. Our function $\epsilon_{(\cdot)}$ can be thought of as a function of the sort $\epsilon_a = Ca$ where $C \in \mathbb{R}$ is a small positive constant.

**Definition 3.5.** Let $(\{V_a\}, \phi)$ and $(\{W_a\}, \psi)$ be two $P$-persistent vector spaces. By an *homomorphism* $T : (\{V_a\}, \phi) \to (\{W_a\}, \psi)$ we mean a set of linear transformations $T = \{T_a\}$ where $T_a : V_a \to W_a$ such that $\psi_{(a,b)} \circ T_a = T_b \circ \phi_{(a,b)}$ for all $a \leq b \in P$. In other words, the following diagram should commute for all $a \leq b \in P$.

$$\begin{tikzcd} V_a \arrow[r, "\phi_{(a,b)}"'] \arrow[d, "T_a"'] & V_b \arrow[d, "T_b"] \\ W_a \arrow[r, "\psi_{(a,b)}"'] & W_b \end{tikzcd}$$

By an *isomorphism* of $P$-persistent vector spaces we mean a homomorphism $T$ where each $T_a$ has an inverse $T_a^{-1}$.

*Note* 3.6. If $\{T_a\}$ is an isomorphism then so is $\{T_a^{-1}\}$. To see this start with $w_a \in W_a$ and apply the fact that $\{T_a\}$ is a homomorphism so $\psi_{(a,b)} \circ T_a = T_b \circ \phi_{(a,b)}$ to $T_a^{-1}(w_a)$. This gives $\psi_{(a,b)} T_a(T_a^{-1}(w_a)) = T_b(\phi_{(a,b)} T_a^{-1}(w_a))$. Now apply $T_b^{-1}$ to both sides and cancel terms to get $T_b^{-1}\psi_{(a,b)} = \psi_{(a,b)} T_a^{-1}(w_a)$.

**Example 3.7.** For any $d \in \mathbb{N}, \epsilon \in \mathbb{R}_+$ we have a $\mathbb{Z}_2$-vector space $C_d^\epsilon$. We define the inclusion map $\imath : C_d^{\epsilon_a} \to C_d^{\epsilon_b}$ whenever $\epsilon_a \leq \epsilon_b$ as in the previous section. Then $(\{C_d^\epsilon\}, \imath)$ is an $\mathbb{R}_+$-persistent $\mathbb{Z}_2$-vector space. Let $\delta_\epsilon$ be defined as the boundary operator from the previous section. Then $\delta$ is a homomorphism from $(\{C_{d+1}^\epsilon\}, \imath)$ to $(\{C_d^\epsilon\}, \imath)$ due to the functoriality described in Theorem 2.31.

**Definition 3.8.** By the *direct sum* of two $P$-persistent vector spaces, $(\{V_a\}, \phi) \bigoplus (\{W_a\}, \psi)$ we mean the $P$-persistent vector space $(\{V_a \bigoplus W_a\}, \phi \bigoplus \psi)$.

**Lemma 3.9.** *Let $(\{V_a\}, \phi)$ and $(\{W_a\}, \psi)$ be two $\mathbb{N}$-persistent $K$-vector spaces. Then $(\{V_a\}, \phi) \bigoplus (\{W_a\}, \psi) = (\{W_a\}, \psi) \bigoplus (\{V_a\}, \phi)$.*

*Proof.* Let $T_a : V_a \bigoplus W_a \to W_a \bigoplus V_a$ be defined by $T_a(v, w) = (w, v)$. $T_a$ is clearly bijective for any $a \in \mathbb{N}$ so all we have to prove is that $\{T_a\}$ is a homomorphism of $\mathbb{N}$-persistent $K$-vector spaces. For any $a, b \in \mathbb{N}, v \in V_a, w \in W_a$ we need to get the same result whether we apply $T_a$ first and then $(\phi \oplus \psi)_{(a,b)}$ or if we apply $(\phi \oplus \psi)_{(a,b)}$ first and then $T_b$.

$$(\psi \oplus \phi)_{(a,b)}(T_a(v, w)) = (\psi \oplus \phi)_{(a,b)}(w, v) = (\psi_{(a,b)}(w), \phi_{(a,b)}(v))$$
$$T_b((\phi \oplus \psi)_{(a,b)}(v, w)) = T_b(\phi_{(a,b)(v), \psi_{(a,b)}(w)}) = (\psi_{(a,b)}(w), \phi_{(a,b)}(v))$$

Since the results are the same, $\{T_a\}$ is a homomorphism. □

**Example 3.10.** Let $U(\alpha, \beta)$ be defined as the $\mathbb{N}$-persistent $\mathbb{Z}_2$-vector space $(\{U(\alpha, \beta)_a\}, \hat{\imath})$ where $U(\alpha, \beta)_a = \mathbb{Z}_2$ for $\alpha \leq a \leq \beta$ and $U_a = \{0\}$ otherwise. Let $\hat{\imath}_{(a,b)} = \mathrm{Id}$ if $b \leq \beta$ and $\hat{\imath}_{(a,b)} = 0$ otherwise. Then $U(\alpha, \beta)$ is called a bar starting at $\alpha$ and ending at $\beta$. A barcode is a finite sum of bars.

**Example 3.11.** Let $U = (\{U_a\}, \phi) := U(2, 3) \bigoplus U(1, 4)$ as defined in the above example. We now calculate $\phi_{(a,a+1)}$ for a few values of $a \in \mathbb{N}$. $\phi_{(1,2)} : U_1 = \{0\} \bigoplus \mathbb{Z}_2 \to \mathbb{Z}_2 \bigoplus \mathbb{Z}_2 = U_2$ with $\phi_{(1,2)}(0, \lambda) = (0, \lambda)$. We also get

$$\begin{aligned} \phi_{(2,3)}(\lambda_1, \lambda_2) &= (\lambda_1, \lambda_2) \\ \phi_{(3,4)}(\lambda_1, \lambda_2) &= (0, \lambda_2) \\ \phi_{(4,5)}(0, \lambda_2) &= (0, 0). \end{aligned}$$

**Definition 3.12.** Let $(\{V_a\}, \phi), (\{W_a\}, \psi)$ be $P$-persistent vector spaces. We say that $(\{V_a\}, \phi)$ is a *subspace* of $(\{W_a\}, \psi)$ if for all $a \in P, V_a \subseteq W_a$ and $\phi(v) = \psi(v)$ for all $v \in V_a$. The *quotient* $(\{W_a\}, \psi)/(\{V_a\}, \phi)$ is defined as the $P$-persistent vector space $(\{W_a/V_a\}, \psi)$ where $\psi_{(a,b)}(w + V_a) = \psi_{(a,b)}(w) + V_b$.

*Note* 3.13. It is easy to see that this is well-defined because $\psi_{(a,b)}(w) + V_b - \psi_{(a,b)}(w') + V_b = \psi_{(a,b)}(w - w') + V_b$.

We can present the homology groups $H_d^\epsilon$ from previous sections as an $\mathbb{R}$-persistent vector space $\{H_d^\epsilon, \hat{\imath}\}$. Using the transformation from $\mathbb{R}$-persistent to $\mathbb{N}$-persistent we can instead present them as an $\mathbb{N}$-persistent vector space $\{H_d^{\epsilon_a}, \hat{\imath}\}$. We have also developed the tools required to calculate the vector spaces $H_d^{\epsilon_a}$ for each $a \in \mathbb{N}$ by reducing the matrices corresponding to the boundary operator $\delta$. What we still lack is a way to rewrite this as a barcode. We do this in the next section and follow closely the intuitive interpretation of a bar, namely as the persistence of a topological feature.

### 3.2. Classification of $\mathbb{N}$-persistent $K$-vector spaces

Let $K$ be a field. We start this section by constructing simple $\mathbb{N}$-persistent $K$-vector spaces called bar codes. We then prove that the $\mathbb{N}$-persistent $K$-vector space $(\{H_d^{\epsilon(a)}\}, \hat{\imath})$ defined in previous sections is isomorphic to a bar code in a unique way. This means that given a sequence of $\epsilon$-values and a sample, there is exactly one way to construct a bar code, with the caveat that the individual bars may be written in different orders.

The definition of a bar looks complicated but can be easily understood as something intuitive. A bar from $\alpha$ to $\beta$ is a sequence of one dimensional vector spaces $K$ over $K$ starting from $\alpha$ and ending at $\beta$. Outside of this range the bar is just $\{0\}$ which is a zero-dimensional vector space over $K$.

**Definition 3.14.** Let $\alpha \in \mathbb{N}$, $\beta \in \mathbb{N} \cup \{\infty\}$. For any $a \in \mathbb{N}$ define

$$U(\alpha, \beta)_a = \begin{cases} K & \text{if } \alpha \leq a \leq \beta, \\ \{0\} & \text{otherwise,} \end{cases}$$

where $K$ is the $K$-vector space over itself and $\{0\}$ is the trivial $K$-vector space. Define linear transformations $\phi_{a,b} : U(\alpha,\beta)_a \to U(\alpha,\beta)_b$ as

$$\phi_{a,b}(k) = \begin{cases} k & \text{if } \alpha \leq a \leq b \leq \beta, \\ 0 & \text{otherwise.} \end{cases}$$

The case $\beta = \infty$ is handled in the natural way so that $a \leq \infty$ for all $a \in \mathbb{N}$. $(\{U(\alpha,\beta)_a\}, \phi)$ is then an $\mathbb{N}$-persistent vector space and is called a *bar* starting at $\alpha$ and ending at $\beta$. To make the notation easier we define $U(\alpha,\beta) := (\{U(\alpha,\beta)_a, \phi\})$.

**Definition 3.15.** By a *barcode* we mean a direct sum of bars. A barcode is called *finite* if it can be written as a finite sum of bars.

*Note* 3.16. Any finite barcode can be written as

$$U = \bigoplus_{i=0}^{I} U(\alpha_i, \beta_i),$$

for some sequence of $(\alpha_i, \beta_i) \in \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$, $I \in \mathbb{N}$. Also note that

$$U_a = (\bigoplus_{i=0}^{I} U(\alpha_i, \beta_i))_a = \bigoplus_{i=0}^{I} U(\alpha_i, \beta_i)_a.$$

Next we wish to prove that barcodes are isomorphic if and only if they have the same bars, possibly in different order. The "only if" part of that statement is surprisingly complicated and to solve it we create an invariant, i.e. a construction that doesn't change as we transform the barcode in certain ways.

**Definition 3.17.** Let $U$ be a finite barcode. We define the *rank matrix* of $U$ as the $\mathbb{N} \times \mathbb{N}$ matrix $\mathrm{R}(U)$ where $\mathrm{R}(U)_{(a,b)} = \mathrm{Rank}(\phi_{(a,b)})$ when $a \leq b \in \mathbb{N}$ and $\mathrm{R}(U)_{(a,b)} = 0$ otherwise.

We will now prove that the sequence $(\alpha_i, \beta_i)$ uniquely determines, and is uniquely determined by, $\mathrm{R}(U)$ up to the ordering of factors.

**Lemma 3.18.** *Let*

$$U = \bigoplus_{i=0}^{I} U(\alpha_i, \beta_i)$$

*and let* $I(a,b) = |\{i; (\alpha_i, \beta_i) = (a,b)\}|$. *Then*

1. *For* $a \leq b \in \mathbb{N}$ *we get* $\mathrm{R}(U)_{(a,b)} = |\{i; \alpha_i \leq a \leq b \leq \beta_i\}|$.
2. *For* $a \leq b \in \mathbb{N}$ *we get* $I(a,b) = (\mathrm{R}(U)_{(a,b)} - \mathrm{R}(U)_{(a-1,b)}) - (\mathrm{R}(U)_{(a,b+1)} - \mathrm{R}(U)_{(a-1,b+1)})$.
3. *For* $a \in \mathbb{N}$ *we get* $I(a,\infty) = \inf\{\mathrm{R}(U)_{(a,c)}; c \geq a\} - \inf\{\mathrm{R}(U)_{(a-1,c)}; c \geq a\}$.

*Note* 3.19. The exact formulas are not important for this paper. The only important thing is that formulas exist so that the sequence $(\alpha_i, \beta_i)$ can be used to calculate $\mathrm{R}(U)$ and vice versa.

*Proof.*

1. $\phi_{(a,b)} : U_a \to U_b$ is the identity on any $U(\alpha_i, \beta_i)$ such that $\alpha_i \leq a \leq b \leq \beta_i$ and the zero map otherwise. In other words $(\phi_{(a,b)})_{|U(\alpha_i,\beta_i)_a}$ is the identity map from $K$ to $K$ if $\alpha_i \leq a \leq b \leq \beta_i$ and the zero map from $K$ to $\{0\}$ otherwise. Thus each $i$ such that $\alpha_i \leq a \leq b \leq \beta_i$ adds 1 to the rank of $\phi_{(a,b)}$. Therefore $R(U)_{(a,b)} = |\{i; \alpha_i \leq a \leq b \leq \beta_i\}|$.

2. The idea of this proof is that the first paranthesis, $(\mathrm{R}(U)_{(a,b)} - \mathrm{R}(U)_{(a-1,b)})$, is the number of bars that start exactly at $a$ and reach to $b$ or further. The second paranthesis, $(\mathrm{R}(U)_{(a,b+1)} - \mathrm{R}(U)_{(a-1,b+1)})$, is the number of bars that start exactly at $a$ and reach to $b+1$ or further. The difference between these two numbers is the number of bars that start exactly at $a$ and end exactly at $b$. The formal proof looks like this:

$$\begin{aligned}\mathrm{R}(U)_{(a,b)} - \mathrm{R}(U)_{(a-1,b)} &= |\{i; \alpha_i \leq a \leq b \leq \beta_i\}| - |\{i; \alpha_i \leq a-1 \leq b \leq \beta_i\}| \\ &= |\{i; \alpha_i = a \leq b \leq \beta_i\}|\end{aligned}$$

   Therefore:

$$\begin{aligned}&(\mathrm{R}(U)_{(a,b)} - \mathrm{R}(U)_{(a-1,b)}) - (\mathrm{R}(U)_{(a,b+1)} - \mathrm{R}(U)_{(a-1,b+1)}) \\ &= |\{i; \alpha_i = a \leq b \leq \beta_i\}| - |\{i; \alpha_i = a \leq b+1 \leq \beta_i\}| \\ &= |\{i; \alpha_i = a \leq b = \beta_i\}| = I(a,b)\end{aligned}$$

3. First, $\inf\{\mathrm{R}(U)_{(a,c)}; c \geq a\} = |\{i; \forall c \geq a, \alpha_i \leq a \leq c \leq \beta_i\}| = |\{i; \alpha_i \leq a, \beta_i = \infty\}|$. Therefore $\inf\{\mathrm{R}(U)_{(a,c)}; c \geq a\} - \inf\{\mathrm{R}(U)_{(a-1,c)}; c \geq a\} = |\{i; \alpha_i = a, \beta_i = \infty\}|$.

$\square$

**Lemma 3.20.** *Two finite barcodes*

$$U = \bigoplus_{i=0}^{I} U(\alpha_i, \beta_i), V = \bigoplus_{j=0}^{J} U(\hat{\alpha}_j, \hat{\beta}_j),$$

*are isomorphic iff* $\mathrm{R}(U) = \mathrm{R}(V)$.

*Proof.* If $\mathrm{R}(U) = \mathrm{R}(V)$ then $U$ and $V$ have the same bars by Lemma 3.18, possibly in different order. To prove that $U \cong V$ we only need to apply Lemma 3.9 a finite number of times to rewrite $V$ as a sum in the same way as $U$.

Now assume that $F : U \to V$ is an isomorphism. We only need to show that $\mathrm{R}(U)_{(a,b)} = \mathrm{R}(V)_{(a,b)}$ for $a \leq b \in \mathbb{N}$, since $\mathrm{R}(U)_{(a,b)} = \mathrm{R}(V)_{(a,b)} = 0$ when $a > b$. Since $F$ is a homomorphism we get $F_b \circ \phi_{(a,b)} = \psi_{(a,b)} \circ F_a$ where we use $\phi$ and $\psi$ to denote the linear transformation from the definition of a barcode for $U$ and $V$ respectively.

Since $F$ is an isomorphism of $\mathbb{N}$-persistent $K$-vector spaces we know that $F_a$ and $F_b$ are vector space isomorphisms, meaning they are of full rank. Since we also know $\mathrm{Rank}(F_b \circ \phi_{(a,b)}) = \mathrm{Rank}(\psi_{(a,b)} \circ F_a)$ we can deduce that $\mathrm{Rank}(\phi_{(a,b)}) = \mathrm{Rank}(\psi_{(a,b)})$ and therefore $\mathrm{R}(U)_{(a,b)} = \mathrm{R}(V)_{(a,b)}$. $\square$

Now the final lemma is just an application of the above lemmas. This is a version of the Krull-Remak-Schmidt theorem which is adapted to the constructions of this paper.

**Lemma 3.21.** *Two finite barcodes*

$$U = \bigoplus_{i=0}^{I} U(\alpha_i, \beta_i), V = \bigoplus_{j=0}^{J} U(\hat{\alpha}_j, \hat{\beta}_j),$$

*are isomorphic iff the sequences* $(\alpha_i, \beta_i)$ *and* $(\hat{\alpha}_j, \hat{\beta}_j)$ *are identical, except possibly in different order.*

*Note* 3.22. This is the same as saying that two finite barcodes are the same iff they have the same bars, possibly in different order.

*Proof.* If the sequences $(\alpha_i, \beta_i)$ and $(\hat{\alpha}_j, \hat{\beta}_j)$ are the same except possibly in different order, then $\mathrm{R}(U) = \mathrm{R}(V)$ by Lemma 3.18 and thus $U \cong V$ by Lemma 3.20.

If on the other hand $U \cong V$ then $\mathrm{R}(U) = \mathrm{R}(V)$ by Lemma 3.20 meaning that the sequences $(\alpha_i, \beta_i)$ and $(\hat{\alpha}_j, \hat{\beta}_j)$ are the same, except possibly in different order, by Lemma 3.18. $\square$

**Definition 3.23.** Let $(\{V_a\}, \phi)$ be an $\mathbb{N}$-persistent vector space, let $\alpha \in \mathbb{N}$ and $v \in V_\alpha$. Then the $K$-span of $v$ is $\langle v \rangle_K = \{kv; k \in K\} \subseteq V_\alpha$. By the $\phi$-span of $v$ we mean the $\mathbb{N}$-persistent vector space $\langle v \rangle_\phi = (\{\langle \phi_{\alpha,a}(v) \rangle_K\}, \phi) \subseteq (\{V_a\}, \phi)$ where $\phi_{(\alpha,a)}(v) := 0$ if $a < \alpha$.

The following is stated without proof since it is easy to see.

**Lemma 3.24.** *Let* $v \in V_\alpha$ *and let* $\beta \in \mathbb{N} \cup \{\infty\}$ *be the largest number such that* $\phi_{(\alpha,a)}(v) \neq 0$ *for all* $a \leq \beta$. *Then* $\langle v \rangle_\phi \cong U(\alpha, \beta)$, *the bar starting at* $\alpha$ *and ending at* $\beta$.

We know that each $H_d^{\epsilon_a}$ is a finite dimensional vector space but there are infinitely many such vector spaces. It does however turn out that $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ is finitely generated in the sense that there is a finite part of the structure which generates everything after it. We give spaces with these properties the name "tame" and then prove that $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ is tame.

**Definition 3.25.** Let $V = (\{V_a\}, \phi)$ be an $\mathbb{N}$-persistent $K$-vector space. $V$ is called *tame* if each $V_a$ is finite dimensional and there exists $N \in \mathbb{N}$ such that $V_a = \phi_{(N,a)}(V_N)$ for all $a \geq N$.

**Lemma 3.26.** $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ *is tame.*

*Proof.* We first prove that $DC^{\epsilon_a} = DC^{\epsilon_b} \implies (\hat{\imath} : H_d^{\epsilon_a} \to H_d^{\epsilon_b}) = \mathrm{Id}$. We then prove that there is an $N \in \mathbb{N}$ such that for all $b \geq a \geq N$ we have $DC^{\epsilon_b} = DC^{\epsilon_a}$, finishing the proof since each $H_d^{\epsilon_\alpha}$ is finite dimensional.

If $DC^{\epsilon_a} = DC^{\epsilon_b}$, then for all $d$ we get : $C_d^{\epsilon_a} = C_d^{\epsilon_b}$ and $\delta_d(C_d^{\epsilon_a}) = \delta_d(C_d^{\epsilon_b})$. This means that $B_d^{\epsilon_a} = B_d^{\epsilon_b}$ and $Z_d^{\epsilon_a} = Z_d^{\epsilon_b}$, which gives $H_d^{\epsilon_a} = H_d^{\epsilon_b}$. Since $DC^{\epsilon_a} = DC^{\epsilon_b}$, we also get $(\imath : DC^{\epsilon_a} = DC^{\epsilon_b}) = \mathrm{Id}$, which finally gives $(\imath : H_d^{\epsilon_a} \to H_d^{\epsilon_b}) = \mathrm{Id}$.

To prove that $\exists N \in \mathbb{N} : \forall b \geq a \geq N; DC^{\epsilon_a} = DC^{\epsilon_b}$ we first note that (1) if $a \leq b$, then $DC^{\epsilon_a} \subset DC^{\epsilon_b}$. We also note that $\forall a, DC^{\epsilon_a} \subset 2^{\mathbb{X}} = \{\text{subsets of } \mathbb{X}\}$. Since $|2^{\mathbb{X}}| = 2^{|\mathbb{X}|}$, $|DC^{\epsilon_a}|$ is an increasing sequence of natural numbers with an upper bound and thus has a maximum. This together with (1) finishes the proof. $\square$

We are now going to find a vector in $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ which has a span that can be separated from $H_d^{\epsilon_a}$ so that $(\{H_d^{\epsilon_a}\}, \hat{\imath}) = \langle v \rangle_{\hat{\imath}} \bigoplus (\{H_d^{\epsilon_a}\}, \hat{\imath}) / \langle v \rangle_{\hat{\imath}}$. For this to work we need some form of unique factorization over the entire structure so we have something specific to factor out. Each of the $H_d^{\epsilon_a}$ is a vector space so unique factorization is obtained simply through a choice of basis but just because $b_1$ and $b_2$ are basis vectors for $H_d^{\epsilon_a}$ that does not mean $\hat{\imath}(b_1), \hat{\imath}(b_2)$ are basis vectors of $H_d^{\epsilon_a+1}$. In fact, it is possible that $\hat{\imath}(b_1) = 0$ or $\hat{\imath}(b_1) = \hat{\imath}(b_2) \neq 0$. To get the unique factorization that we need, we establish a way to construct a basis for the entire structure and use it to prove the next lemma, which is the main problem in this section. In other contexts we would say that we are proving that the so called "splitting lemma" can be applied to $\langle b_0 \rangle$.

**Lemma 3.27.** *(The Splitting Lemma) Let $V = (\{V_a\}, \phi)$ be a tame $\mathbb{N}$-persistent $K$-vector space. Let $\alpha \in \mathbb{N}$ be the smallest number such that $V_\alpha \neq \{0\}$ and let $b_0 \in V_\alpha$, $b_0 \neq 0$. Then $V/\langle b_0 \rangle_\phi$ is also tame and*

$$V \cong \langle b_0 \rangle_\phi \bigoplus V/\langle b_0 \rangle_\phi.$$

*Proof.* Let $\beta \in \mathbb{N} \cup \{\infty\}$ be the largest number such that $\phi_{(\alpha,a)}(b_0) \neq 0$ for all $a \leq \beta$. Now extend $\{b_0\}$ to an ordered basis $B_\alpha := \{b_0, b_1, \ldots\}$ for $V_\alpha$. Now for any $a$ such that $\alpha \leq a \leq \beta$ assume that we have bases $B_\alpha, B_{\alpha+1}, \ldots B_a$ for $V_\alpha, V_{\alpha+1} \ldots V_a$. Now form an ordered basis $B_{a+1}$ by first including $\phi_{(a,a+1)}(B_a)$ in order and then extending that set into an ordered basis for $V_{a+1}$.

Now for any $a$ such that $\alpha \leq a \leq \beta$ the first basis vector in $B_a$ is $\phi_{(\alpha,a)}(b_0)$. This means that any vector in $V_\alpha$ has a uniquely determined $\phi_{(\alpha,a)}(b_0)$ coefficient. In other words any vector $v \in V_a$ can be written in the basis $B_a$ as $v = \lambda_0 v_0 + \lambda_1 v_1 + \ldots + \lambda_n v_n$ where $\lambda_0$ is the $\phi(b_0)$-coefficient.

Now for any $a \leq \beta$ let $T_a : V_a \to \langle b_0 \rangle_\phi \bigoplus V_r/\langle b_0 \rangle_\phi$ be defined by $T_a(v) := (\lambda_0, v + \langle \phi_{(\alpha,a)}(b_0) \rangle_K$. For $a > \beta$ let $T_a(v) := (0, v)$. We will now show that $T$ is an $\mathbb{N}$-persistent isomorphism.

It is a well defined function due to our choice of basis since any $v$ has a uniquely determined $\lambda_0$. To see that it is an $\mathbb{N}$-persistent homomorphism we need $\phi_{(a,b)} \circ T_a =$

$T_b \circ \phi_{(a,b)}$ for all $a \leq b$. Now

$$\begin{aligned}\phi_{(a,b)} \circ T_a(v) &= \phi_{(a,b)}(\lambda_0, v + \langle \phi_{(\alpha,a)}(b_0) \rangle_K)) \\ &= (\lambda_0, v + \langle \phi_{(\alpha,b)}(b_0) \rangle_K) \\ &= T_b \circ \phi_{(a,b)}\end{aligned}$$

due to the linearity of $\phi$. Surjectivity is obvious from the construction. For injectivity assume that $T(v) = T(v_0 + v_1) = T(w_0 + w_1) = T(w)$ where $v_0$ and $w_0$ are the $\phi(b_0)$ coefficients of $v$ and $w$ multiplied by $\phi(b_0)$. We get $v_0 = w_0$ due to our choice of basis which also means $v_1 + \langle \phi_{(\alpha,a)}(b_0) \rangle_K = w_1 + \langle \phi_{(\alpha,a)}(b_0) \rangle$. We get $v_1 - w_1 \in \langle \phi_{(\alpha,a)}(b_0) \rangle$ but due to our choice of basis they can't differ by a factor of $\hat{\imath}_{(\alpha,a)}(b_0)$ other than 0. This means $v_1 = w_1$.

To see that $V/\langle b_0 \rangle_\phi$ is tame we let $N \in \mathbb{N}$ be such that $V_a = \phi_{(N,a)}(V_N)$ for all $a \geq N$. If $(\langle b_0 \rangle)_N = \{0\}$ then it is easy to see that $(V/\langle b_0 \rangle_\phi)_a = \phi_{(N,a)}((V/\langle b_0 \rangle_\phi)_N)$. If on the other hand $(\langle b_0 \rangle)_N \cong K$ then so is $(\langle b_0 \rangle)_N$ for any $a \geq N$ meaning the right hand side and the left hand side are again the same. □

We are now ready for the major result. The persistent homology group we obtained in earlier chapters can be decomposed into a barcode in a unique way.

**Theorem 3.28.** *Fix an $i \in \mathbb{N}$ and a sequence of real numbers $\{\epsilon_a\}$, such that $\epsilon_a \leq \epsilon_b$ whenever $a \leq b$. Then there exists a unique (up to ordering) sequence of pairs $\{(\alpha_j, \beta_j)\}, j \leq J \in \mathbb{N}$ such that $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ is isomorphic to*

$$\bigoplus_{j=0}^{J} U(\alpha_j, \beta_j),$$

*where $\forall j, \alpha_j \in \mathbb{N}, \beta_j \in \mathbb{N} \cup \{\infty\}, 0 \leq \alpha_j \leq \beta_j$.*

*Proof.* Lemma 3.27 allows us to factor out one basis vector $b_0$ from the first non-zero $H_d^{\epsilon_\alpha}$. This reduces the dimension of the first $H_f^{\epsilon_\alpha}$ with dimension bigger than 0 as well as any $H_d^{\epsilon_a}$ with $\hat{\imath}_{(\alpha,a)}(b_0) \neq 0$. Repeat while there is some $H_d^{\epsilon_\alpha} \neq \{0\}$. Since each $H_d^{\epsilon_\alpha}$ is finite dimensional and, by Lemma 3.26, there exists some $N \in \mathbb{N}$ such that $H_d^{\epsilon_a} = \hat{\imath}_{(N,a)}(H_d^{\epsilon_N})$ this process will end. Lemma 3.24 tells that the resulting sum is of the form of the theorem and Lemma 3.21 tells us that the barcode is unique, up to ordering of factors. □

We have not only proven the existence and uniqueness of this decomposition but have also derived a way to calculate and conceptualize it. Take an element $v$ from the first non-zero homology group and see for which $\epsilon$-values $\hat{\imath}(v)$ remains non-zero. The start and end of this element is the start and end of our first bar. Now identify that element with 0 by quotienting out its span and repeat the process to get the second bar. Continue until all features are gone.

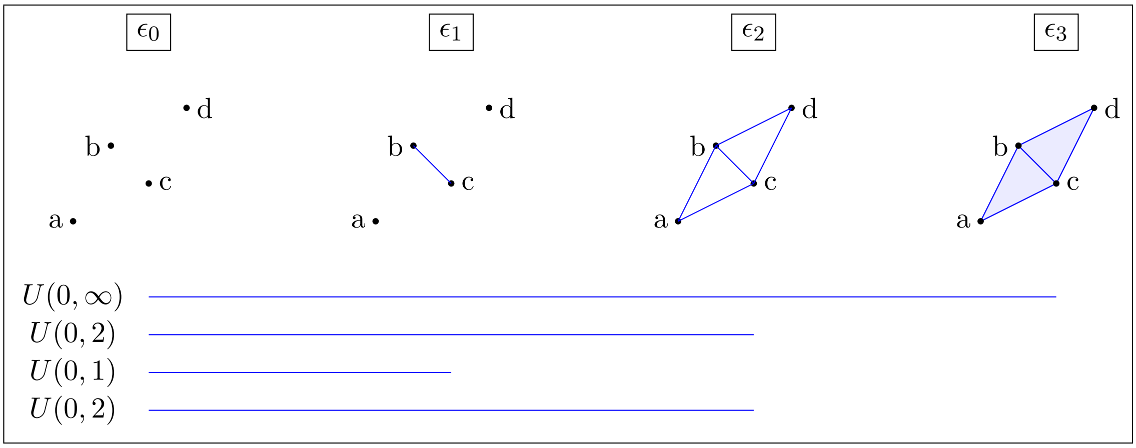


Figure 9: The realization of $DC^{\epsilon}(\mathbb{Y})$ for four different $\epsilon$-values along with the 0-dimensional barcode representing how connected components are created and disappear as $\epsilon$ grows

**Example 3.29.** We will now calculate the barcodes of dimension 0 from Example 2.18 seen in Figure 9. Start by picking a feature of $H_0^{\epsilon_0}$, we choose the point $a$, and track how long that feature persists. $a$ will never be the boundary of a higher dimensional simplicial complex so we start by adding $\langle a \rangle_{\hat{\imath}} \cong U(0,\infty)$ to the barcode of $H_0^{\epsilon}$. Now split off $\langle a \rangle_{\hat{\imath}}$ and consider $H_0^{\epsilon}/\langle a \rangle_{\hat{\imath}}$, that is the same space but we identify $a$ with 0. Pick another 0-dimensional feature, say $b$. When $\epsilon = \epsilon_2$ we get $b$ connected to $a$ and therefore $b$ becomes 0 in $H_0^{\epsilon_2}/\langle a \rangle_{\hat{\imath}}$. Thus $b$ gives rise to the bar $U(0,2)$. Picking $c$ next gives us the bar $U(0,1)$ and lastly $d$ gives us the bar $U(0,2)$. Thus we conclude $H_0^{\epsilon} \cong U(0,\infty) \bigoplus U(0,2) \bigoplus U(0,1) \bigoplus U(0,2)$. Picking features in a different order would have given the same bars, except possibly in a different order.

Now we look at the barcode and try to figure out which samples may give rise to such a barcode. First we see 4 bars, meaning the sample has 4 components which are distinct at $\epsilon_0$. The bar $U(0,\infty)$ is always present and simply signifies that for any $\epsilon$ there is at least one connected component. $U(0,1)$ tells us that a connection between components is formed at $\epsilon_1$. The two $U(0,2)$ bars signify that two more connections are formed at $\epsilon_2$ meaning that the sample is complex is then fully connected.

# 4. Abstract Approach

In this chapter we present a more general, but mathematically more abstract, picture of why the algorithm presented in earlier chapters works. Here we assume familiarity with the basic theory surrounding groups, rings and modules as well as categories. For the rest of the paper $R$ will denote a ring, $K$ will denote a field and $M$ will denote a module. The first section introduces graded rings and modules and proves some basic results. The second section proves an equivalence of categories between the category of $\mathbb{N}$-persistent $R$-modules introduced in earlier chapters and the category of $\mathbb{N}$-graded $R[t]$-modules.

The third section provides an equivalent description of bar codes as $K[t]$-modules and proves the main theorem of the section, that any finitely generated $\mathbb{N}$-graded $K[t]$-module is isomorphic to a uniquely determined finite barcode.

## 4.1. Graded Rings and Modules

We will assume familiarity with rings and modules as well as their homomorphisms, as described in books such as [Gri07], and start by defining graded rings and modules. We will also look closer at $\mathbb{N}$-graded $K[t]$-modules which will later be shown to be equivalent to $\mathbb{N}$-persistent $K$-vector spaces.

**Definition 4.1.** A *non-negatively graded ring*, or $\mathbb{N}$*-graded ring*, $R$ is a ring together with a a set of subgroups of the additive group of $R$, $\{R_a\}_{a\in\mathbb{N}}$, such that $R = \bigoplus R_a$, the direct sum of abelian groups, and $R_aR_b \subseteq R_{a+b}$ for all $a, b \geq 0$.

Note that the decomposition is part of the definition of the graded ring. That means that given an $\mathbb{N}$-graded ring $R$ we already have a decomposition of the ring and may refer to the elements in $R_a$. There are two common ways to write the elements of a graded ring $r \in R = \bigoplus R_a$. One way is as a vector $r = (r_0, r_1, \ldots)$ and another is as a sum $\sum r_a$ depending on the ring in question. Either way we are assuming that only a finite number of the $r_a$ are non-zero.

**Example 4.2.** The set of polynomials with coefficients in $\mathbb{R}$, $\mathbb{R}[t]$, is a ring and it can be written as $\mathbb{R}[t] = \bigoplus \mathbb{R}t^a$. For this to be a graded ring we need $\mathbb{R}t^a\mathbb{R}t^b \subseteq \mathbb{R}t^{a+b}$, which is true since $\mathbb{R}t^a\mathbb{R}t^b = \mathbb{R}t^{a+b}$. Examples of elements in $\mathbb{R}[t]$ are $2 - 3t^2$ or $t + t^3$. Written as vectors these elements would be $(2, 0, -3t^2, 0, 0, \ldots)$ and $(0, t, 0, t^3, 0, 0, \ldots)$, respectively. Viewed as an $\mathbb{R}$-vector space, $\mathbb{R}[t]$ has a basis consisting of $1, t, t^2, \ldots$ and we can write the examples above in that basis as the coordinate vectors $(2, 0, -3, 0, 0, \ldots)$ and $(0, 1, 0, 1, 0, 0, \ldots)$ respectively. Multiplication of elements becomes unintuitive when written in this form but there is a simple way to view the action of multiplication by $t$. $t(2, 0, -3, 0, 0, \ldots) = (0, 2, 0, -3, 0, \ldots)$ and $t(0, 1, 0, 1, 0, 0, \ldots) = (0, 0, 1, 0, 1, 0, \ldots)$. In other words, $t()$ acts on $\mathbb{R}[t]$ by a "rightward shift" of the elements in $\mathbb{R}[t]$.

**Definition 4.3.** Let $r = (r_0, r_1, r_2, \ldots)$ be an element of a graded ring $R = \bigoplus R_a$. We call $r$ *homogeneous* of *degree* $a$ if $r_a \neq 0$ and $r_b = 0$ for all $b \neq a$. We write $\deg(r) = a$ if $r$ is homogeneous of degree $a$ but let $\deg(r)$ be undefined if $r$ is not homogeneuous.

*Note* 4.4. $R_a$ is the set of elements of degree $a$. Since $R_aR_b \subseteq R_{a+b}$ we can see that if $r$ and $r'$ are homogeneous then $rr'$ is either zero or $rr'$ is also homogeneous and $\deg(rr') = \deg(r) + \deg(r')$.

**Definition 4.5.** Let $R$ be an $\mathbb{N}$-graded ring and $r = \sum r_a \in R$. Then $r_a \neq 0$ are called the *homogeneous components* of $r$. The homogeneous component of least degree is called the *lowest homogeneous component* of $r$.

**Example 4.6.** The homogeneous elements of $\mathbb{R}[t]$ are the elements of the form $r = kt^a$, where $k \in \mathbb{R}$ and $a \in \mathbb{N}$. Multiplying two elements of degree $a$ and $b$, $r = kt^a$ and $s = lt^b$

gives an element of degree $a + b$ because $rs = klt^{i+j}$. The homogeneous components of $r = 2 - t^2$ are 2 and $-t^2$ with the lowest homogeneous component 2 of degree 0. The homogeneous components of $s = t + t^3$ are $t$ and $t^3$ with lowest homogeneous component $t$ of degree 1. We see that the lowest homogeneous component of $rs$ is $2t$ of degree $0 + 1$.

**Definition 4.7.** Let $R$ and $S$ be $\mathbb{N}$-graded rings. An ($\mathbb{N}$-graded ring) homomorphism $F : R \to S$ is a ring homomorphism where $F(r_a) \in S_a$ for every homogeneous element $r_a \in R_a$. A homomorphism is called an *isomorphism* if it has an inverse.

*Note* 4.8. The inverse of a homomorphism of $\mathbb{N}$-graded rings is also a homomorphism since $F^{-1}(S_a) \subseteq R_a$.

**Example 4.9.** Let $t()$ be the function $t() : K[t] \to K[t]$ which multiplies by $t$. This fulfills the requirement of a (non-unital) ring homomorphism since $rt(p(t)) = rtp(t) = t(rp(t)$ for any $p(t) \in K[t]$. This is however not a homomorphism of $\mathbb{N}$-graded rings since $\deg(1) = 0$ but $(t(1)) = \deg(t) = 1$.

**Proposition 4.10.** *Let $R, S$ be $\mathbb{N}$-graded rings and $F$ be a homomorphism $F : R \to S$. Then there exists a unique family of group homomorphisms $f_i : R_i \to S_i$ such that $F(r) = \sum f_a(r_a)$ for any $r \in R$.*

*Proof.* Since $F(R_a) \subseteq S_a$ by definition we can define $f_a : R_a \to S_a$ as the restriction of $F$ to $R_a$. We note that $f_a$ is not a homomorphism of rings because $R_a$ is not a ring but it is a homomorphism of abelian groups. Any $r \in R$ can be written as $r = \sum r_a$ where $r_a \in R_a$ so $F(r) = F(\sum r_a) = \sum F(r_a) = \sum f_a(r_a)$ since $F$ is a ring homomorphism. $\square$

**Definition 4.11.** Let $R$ be an $\mathbb{N}$-graded ring. An $\mathbb{N}$-graded $R$-module $M$ is an $R$-module together with a set of subgroups of $M = \bigoplus M_a$ such that $R_a M_b \subseteq M_{a+b}$ for all $a, b \in \mathbb{N}$. An element in $M_a$ is said to be *homogeneous* of degree $a$

**Example 4.12.** Let $K$ be a field and $M := \bigoplus_{a \in \mathbb{N}} K$. To make $M$ a $K[t]$-module we define multiplication by $k \in K$ on an element $m = (m_0, m_1, \ldots) = (m_a)_{a \in \mathbb{N}}$ by $km = (km_a)_{a \in \mathbb{N}}$ and multiplication by $t$ by a rightward shift in $M$. That is $t(m_0, m_1, m_2, \ldots) = (0, m_0, m_1, m_2, \ldots)$. For $kt^a \in Kt^a$ and $m_b \in M_b$ we get $kt^a m_b = t^a(km_b) \in M_{a+b}$ so $M$ is an $\mathbb{N}$-graded $K[t]$-module. It is worth noting that $M$ is a $K$-vector space with an infinite basis $B = \{b_0, b_1, \ldots\}$ where $b_a$ is the vector with 1 in the $a$-th place and 0 everywhere else. As a $K[t]$-module it however has a basis consisting of a single element $b_0$ since $m = (\sum\limits_{a \in \mathbb{N}} m_a t^a) b_0$. It is also worth noting that $b_0$ is the only $b_a$ to give a basis for $M$ since $K[t]b_a \cap M_{a-1} = \{0\}$.

**Definition 4.13.** Let $M, N$ be $\mathbb{N}$-graded $R$-modules. An ($\mathbb{N}$-graded module) *homomorphism* $F : M \to N$ is a module homomorphism such that $F(m_a) \in N_a$ for every homogeneous element $m_a \in M_a$. A homomorphism is called an *isomorphism* if it has an inverse.

Just like for rings we can define the graded components and the least graded components of an $m \in M$ for an $\mathbb{N}$-graded $R$-module $M$.

**Proposition 4.14.** *Let $M, N$ be $\mathbb{N}$-graded $R$-modules and let $F : M \to N$ be a homomorphism. Then there exists a unique family of group homomorphisms $f_a : M_a \to N_a$ such that $F(m) = \sum f_a(m_a)$ for any $m \in M$.*

*Proof.* Entirely analogous to the proof of Proposition 4.10. $\square$

**Example 4.15.** Let $M = \mathbb{R}^n$, the standard $n$-dimensional $\mathbb{R}$-vector space. Let $T$ be an $n \times n$ matrix and assume that $M$ is an $\mathbb{N}$-graded $\mathbb{R}[t]$-module with $(\sum r_a t^a)v := \sum r_a T^a(v)$. For this to work $T$ can't have any non-zero eigenvalues. To see this we let $v \in M$ be an eigenvector of $T$ and let the lowest homogeneous component of $v$ be $v_a$ with degree $a$. Since the lowest homogeneous component of $\lambda v$ is $\lambda v_a$ of degree $a$ but the lowest homogeneous component of $Tv$ has degree at least $a + 1$ we can determine that $\lambda = 0$.

Submodules and quotient modules are defined exactly as for ungraded modules. There are however some properties that are worth noting. If $N = \bigoplus N_a \subseteq M = \bigoplus M_a$ is a submodule then $N_a \subseteq M_a$ for all $a \in \mathbb{N}$. We define $\imath : N \to M$ as the inclusion and note that this is an $\mathbb{N}$-graded module homomorphism. According to Proposition 4.14 the inclusion $\imath : N \to M$ can be decomposed into $\imath_a : N_a \to M_a$. The same is true for the quotient $M/N$ and the projection $\pi : M \to M/N$ with $\pi(m) = m + N$. This projection decomposes into $\pi_a : M_a \to (M/N)_a = M_a/N_a$.

**Definition 4.16.** Let $M$ be an $\mathbb{N}$-graded $R$-module and $m \in M$. Then the *span* of $m$ is written $\langle m \rangle := Rm = \{rm; r \in R\}$.

**Example 4.17.** Let $m_a \in M_a$ be a homogeneous element of degree $a$ in an $R[t]$-module $M$. Then $\langle m_i \rangle = R[t]m_a = Rm_a \bigoplus Rtm_a \bigoplus Rt^2 m_a \bigoplus \ldots$. This is an $\mathbb{N}$-graded $R[t]$-module, call it $N$. Then $N_b = \{0\}$ for $b < a$ and $N_b = Rt^{b-a}m_a$ for $b \geq a$. If $R$ is a field, say $R = K$, and $t^{b-a}m_a \neq 0$ then $N_b \cong K$. Thus the span of a homogeneous element $m_a$ in a $K[t]$-module is a sequence of one-dimensional $K$-vector spaces starting at $a$ and ending whenever, if ever, $t^{b-a}m_a = 0$.

**Example 4.18.** Let $M = K[t]$ be regarded as a $K[t]$-module. Then $N = t^2 K[t]$ is a submodule consisting of all polynomials of degree at least 2 with coefficients in $K$. $(M/N)_a = M_a/N_a \cong K$ for $0 \leq a \leq 1$ and $(M/N)_a \cong \{0\}$ otherwise. So $M/N = K[t]/(t^2 K[t])$ is isomorphic to $\{\lambda_0 + \lambda_1 t; \lambda_0, \lambda_1 \in K\}$ as a $K[t]$-module with $(k_0 + k_1 t + k_2 t^2 + \ldots)(\lambda_0 + \lambda_1 t) = k_0\lambda_0 + (k_0\lambda_1 + k_1\lambda_0)t$. For $m = \sum m_i = \sum \lambda_i t^i$ we get $\pi(m) = \lambda_0 + \lambda_1 t + t^2 K[t]$.

**Definition 4.19.** An $\mathbb{N}$-graded $R$-module $M$ is called *finitely* generated if there are $b^{(0)}, b^{(1)}, \ldots, b^{(N)} \in M$ such that $M = \{ \sum_{n=0}^{N} r^{(n)} b^{(n)}; r^{(n)} \in R\}$

**Proposition 4.20.** *Let $M$ be a finitely generated $\mathbb{N}$-graded $R$-module. Then there is a finite generating set $b^{(0)}, b^{(1)}, \ldots, b^{(N)} \in M$ where each $b^{(i)}$ is homogeneous.*

*Proof.* The homogeneous parts of a generating set will also be a generating set as each element can be written as the sum of its homogeneous parts. These can be chosen to be finite since each generator is required to be a finite sum and the generating set can be chosen to be finite. □

We finish this part by proving a version of the so called "splitting lemma". This mirrors Lemma 3.27 from the previous section. It states that under certain conditions we may separate out a submodule $N$ as a summand of $M$ and be left with $M/N$. For a vector space $V$ it is well known that these conditions always hold so we may write $V = W \bigoplus V/W$ whenever $W$ is a subspace of $V$. In the case of a $\mathbb{N}$-graded $R$-modules we however need a projection onto $N$ which fulfills certain conditions.

**Lemma 4.21.** *(The Splitting Lemma) Let $M$ be a $\mathbb{N}$-graded $R$-module, let $N$ be a submodule of $M$ and let $\imath : N \to M$ be the inclusion $\imath(n) = n$ . If there exists a homomorphism $\psi : M \to N$ such that $\psi \circ \imath = Id_N$ then $M \cong N \bigoplus M/N$.*

*Note* 4.22. We will use this for $N = \langle v \rangle = Rv$. What the lemma says in that case is that if we can give elements in $M$ a $v$-coordinate in a way that respects the action of $R$, then we may split off $\langle v \rangle$.

*Proof.* Let $\phi : M \to N \bigoplus M/N$ be defined by $\phi(m) = (\psi(m), m + N)$. We can easily see that $\phi$ is a module homomorphism since $\psi$ and the projection $\pi : M \to M/N$ are homomorphisms. We will however check that $\phi$ satisfies the condition of a graded module homomorphism that $\phi(m_a) \in (N \bigoplus M/N)_a$ for any $m_a \in M_a$. Since $\psi$ is a homomorphism we get $\psi(m_a) \in N_a$ which means that $\psi(m_a)$ is either 0 or it is homogeneous of degree $a$. The same is true for $\pi(m_a)$ meaning that $\phi(m_a)$ is either homogeneous of degree $a$ or it is 0 which means $\phi(m_a) \in (N \bigoplus M/N)_a$.

To show that $\phi$ is injective we just need to show that $\mathrm{Ker}(\phi) = 0$. If $\phi(m) = 0$ then $\psi(m) = 0$ and $m + N = 0$. $m + N = 0$ means that $m \in N$ . If $m \in N$ then $m = \imath(m)$ and $\psi(m) = \psi(\imath(m)) = m$ so $\psi(m) = 0 \Rightarrow m = 0$ for any $m \in N$.

To show surjectivity we let $(n, m + N) \in N \bigoplus M/N$ and find $m'$ such that $\phi(m') = (n, m+N)$. Intuitively we may think of $\imath\psi(m)$ as the projection of $m$ onto $N$ so $m - \imath\psi(m)$ should be the orthogonal part of $m$. Thus pick any representative $m$ for $m + N$ and let $m' = n + (m - \imath\psi(m))$. Now $\phi(m') = (\psi(n + m - \imath\psi(m)), n + m - \imath\psi(m) + N)$. We check the first component and get $\psi(n+m-\imath\psi(m)) = \psi(n)+\psi(m)-\psi(m) = n$ since $\psi(n) = n$ and $\psi \circ \imath = \mathrm{Id}$. As for the second component we get $n + m - \imath\psi(m) + N = m + N$ since $n \in N$ and $\imath\psi(m) \in N$. Thus $\phi(m') = (n, m + N)$. □

**Example 4.23.** Let $K$ be a field and let $M = K[t] \bigoplus K[t]$ be a graded module over the graded ring $R = K[t]$ with ring action defined by $p(t)(q(t), r(t)) = (p(t)q(t), p(t)r(t))$. Let $N = \langle (1, 0) \rangle = K[t](1, 0)$. $N$ is then a submodule of $M$ and there is a natural inclusion $\imath : N \to M$ defined by $\imath((p(t), 0) = (p(t), 0)$. Let $\psi : M \to N$ be defined by $\psi(p(t), q(t)) = (p(t), 0)$. It is easy to check that $\psi$ is a homomorphism and that $\psi \circ \imath = \mathrm{Id}_N$. Thus, by the splitting lemma, $M = N \bigoplus M/N$.

**Example 4.24.** Now let $M$ be as in the previous example but let $N = \langle(t,0)\rangle = K[t](t,0)$. $N$ is still a submodule and we may still form $M/N$ and also define the inclusion $\imath$. There is however no $\psi$ fulfilling the conditions in the splitting lemma. If such a $\psi$ existed we would need $t\psi(1,0) = t0 = 0$, since $\psi$ needs to preserve the degree and $N$ has no element of degree 1, but also $t\psi(1,0) = \psi(t,0) = (t,0)$, since $\psi$ needs to be a module homomorphism and $\psi \circ \imath = \mathrm{Id}$. This is a contradiction so no such $\psi$ can exist. We will later see that, for $K[t]$-modules, when $v$ is homogeneous of the lowest degree in $M$ there exists a $\psi$ which fulfills the conditions in the lemma.

## 4.2. Correspondence between persistent modules and graded modules

In this section we establish an equivalence between the category of $\mathbb{N}$-persistent $R$-modules and the category of $\mathbb{N}$-graded modules over $R[t]$. We then prove that the $\mathbb{N}$-persistent $K$-vector spaces we defined as "tame" in Section 3.2 correspond to finitely generated $K[t]$-modules.

We start by establishing an equivalence of categories, $\theta$ between the category of $\mathbb{N}$-persistent $R$-modules and the category of $\mathbb{N}$-graded $R[t]$-modules. To do this we first need to associate to every $\mathbb{N}$-persistent $R$-module $M$ an $\mathbb{N}$-graded $R[t]$-module $\theta(M)$. Then we prove that $\theta$ defines a functor and define an inverse functor $\theta^{-1}$ such that $\theta \circ \theta^{-1} = \mathrm{Id}$, the identity functor on the category of $\mathbb{N}$-graded $R[t]$-modules, and $\theta^{-1} \circ \theta = \mathrm{Id}$, the identity functor on the category of $\mathbb{N}$-persistent $R$-modules.

We let $R$ be a ring and $(\{M_a\}, \phi)$ be an $\mathbb{N}$-persistent $R$-module as defined in Section 3.1. That is a set of $R$-modules $\{M_a\}_{a\in\mathbb{N}}$ along with module homomorphisms $\phi_{(a,b)} : M_a \to M_b$ such that $\phi_{(a,b)} \circ \phi_{(b,c)} = \phi_{(a,c)}$. Recall also that a homomorphism $\xi : (\{M_a\}, \phi) \to (\{N_a\}, \psi)$ is a set of module homomorphisms $\xi_a : M_a \to N_a$ such that $\xi_b \circ \phi_{(a.b)} = \psi_{(a,b)} \circ \xi_a$.

To create an $R[t]$-module associated to $(\{M_a\}, \phi)$ we need to define (1) the group of the module along with its decomposition into homogeneous components, (2) the action of $R[t]$ on its elements, (3) the image of homomorphisms between $\mathbb{N}$-persistent $R$-modules, i.e $\theta(\xi)$, where $\xi : (\{M_a\}, \phi) \to (\{N_a\}, \psi)$ is a homomorphism of $\mathbb{N}$-persistent $R$-modules.

1. For $\alpha \in M_a$ let $\deg(\alpha) = a$ and let the group of the module $\theta((\{M_a\}, \phi))$ be $\bigoplus M_a$.
2. The action of $r \in R$ on $m \in \bigoplus M_a$ is defined as $r(\sum m_a) = \sum r(m_a)$. The action of $t$ is defined $t(\sum m_a) = \sum \phi_{(a,a+1)}(m_a)$. This is then extended linearly to $R[t]$.
3. We need a homomorphism $\theta(\xi) : \bigoplus M_a \to \bigoplus N_a$. Let $\theta(\xi)_a : M_a \to N_a$ be given by $\theta(\xi)_a(m_a) = \xi(m_a)$. Now let $\theta(\xi)(\sum m_a) = \sum \theta(\xi)_a(m_a)$.

It is relatively easy, but very cumbersome, to check that $\theta$ is a functor so this proof is moved to Appendix A and simply outline the interesting parts here. We need $\theta(\xi)$ to be an $\mathbb{N}$-graded $R[t]$-module homomorphism. That $\theta(\xi)(M_a) \subseteq N_a$ is straightforward from the definition and $\theta(\xi)$ commutes with $r(t) \in R[t]$ since multiplication by $r$ commutes with $\xi$ by definition and multiplication by $t$ correspond to application of $\phi$ or $\psi$ which also commutes with $\xi$ by definition. We also need $\theta(\xi \circ \psi) = \theta(\xi) \circ \theta(\psi)$ which is straightforward from the definition.

To prove that $\theta$ is an equivalence of categories we create a inverse functor $\theta^{-1}$ which becomes the identity when composed with $\theta$ on both the left and right. Defining $\theta^{-1}$ requires (1) for a given $\mathbb{N}$-graded $R[t]$-module $M$ we need to define the groups of an $\mathbb{N}$-persistent $R$-module $\theta^{-1}(M) = \{M_a\}$, (2) define maps $\phi_{(a,b)} : M_a \to M_b$ whenever $a \leq b$ so that their compositions satisfy $\phi_{(a,b)} \circ \phi_{(b,c)} = \phi_{(a,c)}$ and finally (3) $\theta^{-1}(F)$ for any homomorphism $F : M \to N$ where $M$, $N$ are $\mathbb{N}$-graded $R[t]$-modules.

1. For $M = \bigoplus M_a$ let the set of $R$-modules of the $\mathbb{N}$-persistent $R$-module $\theta^{-1}(M)$ be given by $\{M_a\}$.
2. Let $\phi_{(a,b)} := t^{b-a}()$, the operator that multiplies by $t^{b-a}$.
3. Note first that Proposition 4.14 allows us to decompose $F$ into uniquely determined $f_a : M_a \to N_a$. This means that we may define $\theta^{-1}(F) : \theta^{-1}(M) \to \theta^{-1}(N)$ by defining each $(\theta^{-1}(F))_a(\alpha_a) = \theta^{-1}(f_a(\alpha_a)) \in (\theta^{-1}(N))_a$ for $\alpha_a \in M_a$.

Note that $\theta^{-1}(f_a)$ is an $R$-module homomorphism because $f_a$ is. The requirement that $F$ is an $\mathbb{N}$-graded $R[t]$-module homomorphism means that $F$ commutes with multiplication by $t$ which translates into $\theta^{-1}(F)$ commuting with $\phi$. This makes $\theta^{-1}(F)$ an $\mathbb{N}$-persistent $R$-module homomorphisms. Thus $\theta^{-1}$ is a functor. Proving that the compositions $\theta \circ \theta^{-1}$ and $\theta^{-1} \circ \theta$ are identities is simply a matter of checking that $\theta\theta^{-1}$ and $\theta^{-1}\theta$ of various objects give the same object. This is very easy and intuitive, but quite technical and extremely cumbersome, so it is omitted.

There is no known decomposition of general $\mathbb{N}$-graded $R[t]$-modules, but there is a decomposition of the types of $R[t]$-modules we can obtain by the process described in earlier sections, $\theta((\{H_d^{\epsilon_a}\}, \hat{\imath}))$. We defined such persistent modules as tame in earlier sections and proved that $(\{H_d^{\epsilon_a}\}, \hat{\imath})$ was tame. We restate the definition of tame and then prove that tame $\mathbb{N}$-persistent vector spaces are the ones that get mapped to finitely generated $F[t]$-modules by $\theta$.

We remind ourselves that an $\mathbb{N}$-persistent $R$-module $(\{M_a\}, \phi)$ is called *tame* if (1) $M_a$ is finitely generated for any $a$ and (2) $M_a = \phi_{(a-1,a)}(M_{a-1})$ for sufficiently large $a$.

**Lemma 4.25.** *Let $(\{M_a\}, \phi)$ be a tame $\mathbb{N}$-persistent $R$-module. Then $\theta((\{M_a\}), \phi)$ is a finitely generated $\mathbb{N}$-graded $R[t]$-module.*

*Proof.* By definition we see that $\theta((\{M_a\}, \phi))$ is an $\mathbb{N}$-graded $R[t]$-module so all we need to prove is that it is finitely generated. Since $(\{M_a\}, \phi)$ is tame we can find a number $A$ such that $\phi_{(a,a+1)}$ is surjective for any $a \geq A$. Since each $M_a$ is finitely generated we may choose a finite generating set for all $M_a$ with $a \leq A$. This constitutes a set of generators not only for each $M_a$ with $a \leq A$ but also for $M_a$ with $a \geq A$ since for any $a \geq A$, $M_a \subseteq t^{a-A}(M_A)$. □

We have now seen that instead of working with persistent modules we could have defined everything as $\mathbb{N}$-graded $K[t]$-modules. We have also seen that our algorithm would then have produced finitely generated $K[t]$-modules.

## 4.3. $\mathbb{N}$-graded $K[t]$-modules

We start this section by defining the simplest building blocks of $\mathbb{N}$-graded $K[t]$-modules where $K$ is a field. We call these bars and barcodes just like the persistent $K$-vector spaces we defined in Section 3.2. We then check that these correspond to the bars and barcodes defined earlier, meaning that our abuse of notation is somewhat justified. We finish by providing a method to rewrite any finitely generated $K[t]$-modules into a barcode in a way that is consistent with the intuitive interpretation of a bar as the persistence of a feature.

**Definition 4.26.** By a *bar* starting at $\alpha \in \mathbb{N}$ and ending at $\beta \in \mathbb{N}$ we mean the $\mathbb{N}$-graded $K[t]$-module

$$V(\alpha, \beta) := \{\sum_{a=\alpha}^{\beta} \lambda_a t^a; \lambda_a \in K\}$$

with the natural multiplication by elements in $K[t]$ except $t^a = 0$ in $V(\alpha, \beta)$ whenever $a > \beta$. A *barcode* is a finite direct sum of bars.

*Note* 4.27. We may also write $V(\alpha, \eta) = t^\alpha K[t]/(t^{\beta+1}K[t])$ when $\beta \in \mathbb{N}$ and $V(\alpha, \infty) = t^\alpha K[t]$ when $\beta = \infty$.

**Example 4.28.** Let $V := V(1, 4) \bigoplus V(2, 3)$. Then an element $v \in V$ can be written as

$$v = (\sum_{a=1}^{4} \lambda_a t^a, \sum_{b=2}^{3} \mu_b t^b).$$

$V$ is generated by $b_0 = (t, 0)$ and $b_1 = (0, t^2)$ since

$$v = (\sum_{a=1}^{4} \lambda_a t^a, \sum_{b=2}^{3} \mu_b t^b) = (\sum_{a=1}^{4} \lambda_a t^{a-1})b_0 + (\sum_{b=2}^{3} \mu_b t^{a-2})b_1.$$

We don't however have a unique representation of elements in $V$ as a linear combination of $\lambda b_0 + \mu b_1$ such that $\lambda, \mu \in K[t]$ since for example $0 = t^4 b_0 = t^4 b_1$. In fact no basis can exist over $K[t]$ since any non-empty set of elements in $V$ will be linearly dependent but we may still write $V = \langle b_0 \rangle \bigoplus \langle b_1 \rangle$.

**Example 4.29.** Let $V := V(1, \infty) \bigoplus V(2, \infty) = tK[t] \bigoplus t^2 K[t]$. Then $V$ has a basis, namely $\{b_0, b_1\}$ from the previous example. Also we see that $V = \langle b_0 \rangle \bigoplus \langle b_1 \rangle$.

We now prove that the bars defined above correspond to the bars defined in Section 3.2.

**Proposition 4.30.** *Let $U(\alpha, \beta)$ be an $\mathbb{N}$-persistent $K$-vector space bar. Then $\theta(U(\alpha, \beta)) \cong V(\alpha, \beta)$, an $\mathbb{N}$-graded $K[t]$-module bar.*

*Proof.*

$$\theta(U(\alpha,\beta)) = \bigoplus_{a=0}^{\alpha-1}\{0\} \bigoplus_{a=\alpha}^{\beta} K \bigoplus_{a=\beta+1}^{\infty} \{0\}$$

The action of $t$ on $\theta(U(\alpha,\beta))$ is a "rightward shift" in the sum except on the elements in the $\beta$-th component of the sum where $t$ is the zero map. In other words, we may write elements in $\theta(U(\alpha,\beta))$ as $\sum_{a=\alpha}^{\beta} \lambda_a t^a$, where $\lambda_a \in K$ and $t^{\beta+1} = 0$ which is the same as $V(\alpha,\beta)$. □

A barcode has sums in two "directions", $a$ for the grading of the module, and $i$ for the number in the sum of bars. We now define an invariant of barcodes $\mathrm{R}(V)$ which we later use to prove that barcodes are isomorphic iff they have the same bars.

**Definition 4.31.** Let $V$ be a finite barcode. We define the *rank matrix* of V as $\mathrm{R}(V) \in \mathbb{N}\times\mathbb{N}$ where $\mathrm{R}(V)_{(a,b)} = \dim(t^{b-a}(V_a))$ when $a \le b$ and $\mathrm{R}(V)_{(a,b)} = 0$ otherwise.

*Note* 4.32. If we view $t_{(a,b)}(.)$ as a linear transformation $t_{(a,b)}(.) : V_a \to V_b$ then $\mathrm{R}(V)_{(a,b)}$ is the rank of $t_{(a,b)}(.)$.

We now provide formulas to calculate the rank matrix from the bars of a barcode, as well as formulas to calculate the bars from the rank matrix. The idea that such formulas would exist has been raised as a question in other papers such as the bachelor thesis [Sco16], where the author proves the existence for a smaller partially ordered set, namely $\{0,1,2\} \subseteq \mathbb{N}$. The actual formulas presented below are not important for this paper, the existence of such formulas however is very important as it proves that the rank matrix uniquely determines the bars and vice versa.

**Proposition 4.33.** *Let*

$$V = \bigoplus_{i=0}^{I} V(\alpha_i,\beta_i)$$

*be a finite barcode and let $I(a,b)$ be the number of $i$ such that $(\alpha_i,\beta_i) = (a,b)$. Then*

1. *For $a \in \mathbb{N}$ we get $\dim(V_a) = |\{i; \alpha_i \le a \le \beta_i\}|$.*
2. *For $a \le b \in \mathbb{N}$ we get $\mathrm{R}(V)_{(a,b)} = |\{i; \alpha_i \le a \le b \le \beta_i\}|$.*
3. *For $a \le b \in \mathbb{N}$ we get $I(a,b) = (\mathrm{R}(V)_{(a,b)} - \mathrm{R}(V)_{(a-1,b)} - (\mathrm{R}(V)_{(a,b+1)} - \mathrm{R}(V)_{(a-1,b+1)})$.*
4. *For $a \in \mathbb{N}$ we get $I(a,\infty) = \inf\{\mathrm{R}(V)_{(a,c)}; c \ge a\} - \inf\{\mathrm{R}(V)_{(a-1,c)}; c \ge a\}$.*

*Proof.* 1. Because $(M \bigoplus N)_a = M_a \bigoplus N_a$ we get

$$V_a = (\bigoplus_{i=0}^{I} V(\alpha_i,\beta_i))_a = \bigoplus_{i=0}^{I} V(\alpha_i,\beta_i)_a$$

Now $V(\alpha_i,\beta_i)_a = K$ iff $\alpha_i \le a \le \beta_i$ and otherwise $V(\alpha_i,\beta_i)_a = \{0\}$ so the sum $\bigoplus_{i=0}^{I} V(\alpha_i,\beta_i)_a$ has one $K$ for each $V(\alpha_i,\beta_i)$ such that $\alpha_i \le a \le \beta_i$. In other words $V_a \cong K^n$ where $n$ is the number of $i$ such that $\alpha_i \le a \le \beta_i$.

2. Let $T$ be the linear map $T : V_a \to t^{b-a}V_a \subseteq V_b$ which multiplies by $t^{b-a}$. We wish to calculate $\dim(T(V_a))$. By the rank nullity theorem we get $\dim(V_a) = \dim(\mathrm{Ker}(T)) + \dim(T(V_a))$, so $\dim(T(V_a)) = \dim(V_a) - \dim(\mathrm{Ker}(T))$. For any non-zero $v \in V(\alpha_i,\beta_i)_a$ we get $Tv = 0$ iff $\beta_i < a + b - a = b$. This means that $\mathrm{Ker}(T)$ has one $K$ for each $V(\alpha_i,\beta_i)$ such that $\alpha_i \le a$ and $\beta_i < b$ and therefore $\dim(\mathrm{Ker}(T)) = |\{i; \alpha_i \le a, b > \beta_i\}|$. Now $\dim(V_a) - \dim(\mathrm{Ker}(T)) = |\{i; \alpha_i \le a \le \beta_i\}| - |\{i; \alpha_i \le a \le \beta_i, b > \beta_i\}| = |\{i; \alpha_i \le a \le b \le \beta_i\}|$.

3. This is identical to the proof of Statement 2 in Lemma 3.18.

4. First, $\inf\{\mathrm{R}(U)_{(a,c)}; c \ge a\} = |\{i; \forall c \ge a, \alpha_i \le a \le c \le \beta_i\}| = |\{i; \alpha_i \le a, \beta_i = \infty\}|$. Therefore $\inf\{\mathrm{R}(U)_{(a,c)}; c \ge a\} - \inf\{\mathrm{R}(U)_{(a-1,c)}; c \ge a\} = |\{i; \alpha_i = a, \beta_i = \infty\}|$. ☐

**Lemma 4.34.** *Two finite barcodes*

$$V = \bigoplus_{i=0}^{I} V(\alpha_i,\beta_i), W = \bigoplus_{j=0}^{J} V(\hat{\alpha}_j,\hat{\beta}_j),$$

*are isomorphic iff the sequences $(\alpha_i,\beta_i)$ and $(\hat{\alpha}_j,\hat{\beta}_j)$ are identical, except possibly in different order.*

*Proof.* If the sequences $(\alpha_i,\beta_i)$ and $(\hat{\alpha}_j,\hat{\beta}_j)$ are the same then we $V \cong W$ since $M \bigoplus N \cong N \bigoplus M$ for any $\mathbb{N}$-graded modules $M$ and $N$ so we may reorder the bars from $V$ to the same order as in $W$ since these are both finite sums.

To prove the converse we first need the fact that, for any isomorphism $F : V \to W$ and any $c \in \mathbb{N}$, $f_c : V_c \to W_c$ as described in Proposition 4.14 is a $K$-vector space isomorphism. To see this we first note that $V_c$ and $W_c$ are $K$-vector spaces by definition and $f_c$ is linear because $F$ is a homomorphism. Furthermore $f_c$ is a bijection because $F$ is so that $F^{-1}F(v_c) = v_c$ for any $v_c \in V_c$ and so there exists a $f_c^{-1}$, namely the restriction of $F^{-1}$ to $W_c$.

Now assume that $F : V \to W$ is an isomorphism and let $a \le b \in \mathbb{N}$. If we prove that $\mathrm{R}(V)_{(a,b)} = \mathrm{R}(W)_{(a,b)}$ then the sequences $(\alpha_i,\beta_i)$ and $(\hat{\alpha}_j,\hat{\beta}_j)$ are identical, except possibly in different order, by Proposition 4.33. Now

$$\begin{aligned}\mathrm{R}(V)_{(a,b)} &= \dim(t^{b-a}V_a) = \dim(f_b(t^{b-a}V_a)) \\ &= \dim(t^{b-a}(f_a(V_a)) = \dim(t^{b-a}(W_a)) = \mathrm{R}(W)_{(a,b)}\end{aligned}$$

where we only use the facts that $F$ is a homomorphism and that $f_a$ as well as $f_b$ are vector space isomorphisms. ☐

Next we prove that certain cyclic modules, that is modules of the type $K[t]v, v \in M_a$, are isomorphic to bars.

**Lemma 4.35.** *Let $M$ be an $\mathbb{N}$-graded $K[t]$-module, let $v \in M_\alpha$ and let $\beta \in \mathbb{N} \cup \{\infty\}$ be the largest number such that $t^{\beta-\alpha}v \neq 0$. Then $K[t]v \cong V(\alpha, \beta)$.*

*Proof.* Let $\phi_\alpha : V(\alpha,\beta)_\alpha \to (K[t]v)_\alpha$ be defined by $\phi_\alpha(1) = v$ and extend linearly to $V(\alpha,\beta)_\alpha$. Then $\phi_\alpha$ is a vector space isomorphism between one dimensional $K$-vector spaces. Now for any $a \in \mathbb{N}$ such that $\alpha \leq a \leq \beta$ extend $\phi$ linearly with regard to $t$ such that $\phi_a(1_{V(\alpha,\beta)_a}) = \phi_a(t^{a-\alpha}(1_{V(\alpha,\beta)_\alpha})) = t^{a-\alpha}\phi_\alpha(1_{V(\alpha,\beta)_\alpha}) = t^{a-\alpha}v$. Now extend $\phi_a$ linearly with regard to $K$ to make $\phi_a : V(\alpha,\beta)_a \to (K[t]v)_a$ a vector space isomorphism of one dimensional $K$-vector spaces. For any $a$ such that $a < \alpha$ or $a > \beta$ both spaces are $\{0\}$ so let $\phi_a(0) = 0$. This makes $\phi_a$ a vector space isomorphism for any $a \in \mathbb{N}$, meaning that $\phi$ is bijective, and $\phi$ is an $\mathbb{N}$-graded module homomorphism by definition. □

We now wish to pick a vector $v$ and then use the splitting lemma to rewrite $M \cong K[t]v \bigoplus M/(K[t]v)$. If we can prove that this is possible and that $M/(K[t]v)$ is finitely generated with a smaller generating set then we can keep going until we get $M$ as a direct sum of cyclic modules $K[t]v_\alpha$. These cyclic modules have been proven to each be isomorphic to a bar so $M$ would be isomorphic to a barcode, and barcodes have been proven to be isomorphic iff they are the same up to ordering of summands. To apply the splitting lemma we however need a homomorphism $\psi : M \to K[t]v$ such that $\psi \circ \imath = \mathrm{Id}$.

**Lemma 4.36.** *Let $M$ be a finitely generated $\mathbb{N}$-graded $K[t]$-module and let $v$ be a homogeneous element of minimal degree in $M$. Then $M \cong K[t]v \bigoplus M/(K[t]v)$.*

*Proof.* Let $\alpha$ be the smallest number such that $M_\alpha \neq 0$. Then $\deg(v) = \alpha$. Construct a basis $B_\alpha$ for $M_\alpha$ as a $K$-vector space using $v$ as the first basis vector so that $B_\alpha = \{b_\alpha^{(0)}, b_\alpha^{(1)}, \dots b_a^{(i_\alpha)}\}$ with $v = b_\alpha^{(0)}$. Now any vector $m\alpha \in M_\alpha$ can be written as $m_\alpha = \sum \lambda_j b_a^{(k)}$ in a unique way with its $v$ component first. Let $\psi_\alpha(m_\alpha) = \lambda_0 b_\alpha^{(0)} = \lambda_0 v$, that is the $v$ component of $m_\alpha$ in the basis $B_\alpha$. Then $\psi_\alpha : M_\alpha \to \langle v \rangle_a$ is a linear map.

In order to create a homomorphism $\psi$ we will need $t\psi(m_\alpha) = \psi(tm_\alpha) = \sum \lambda_j t b_a^{(j)}$. So if $m_\alpha$ has $v$ component $\lambda_0$ then we need $tm_\alpha$ to have $tv$ component $\lambda_0$. This works as long as the basis, $B_{a+1}$ for $M_{a+1}$ is an extension of $tB_a$. Therefore for any $a > \alpha$ pick a basis, $B_a$, for $M_a$ extending $tB_{a-1}$ and let $\psi_a(m_a)$ be the $t^{a-\alpha}v$ component of $m_a$. For $a < \alpha$ we get $\psi_a : \{0\} \to \{0\}$ so $\psi_a = 0$. For $a$ such that $t^{a-\alpha}v = 0$ we get $\psi_a : M_a \to \{0\}$ so we let $\psi_a = 0$. $\psi$ is a homomorphism by construction and $\psi(\imath(n)) = n$ for any $n \in \langle v \rangle$ so $\psi$ fulfills the conditions in the splitting lemma (Lemma 4.21). Thus $M \cong K[t]v \bigoplus M/(K[t]v)$ □

Now we are ready to put everything together to prove the main theorem, a version of the Krull-Remak-Schmidt theorem, that every finitely generated $\mathbb{N}$-graded $K[t]$-module is isomorphic to a finite barcode. The proof also demonstrates how such a barcode can be obtained, namely (1) select a vector $v$ of minimal degree, (2) obtain the first bar going from $\deg(v)$ to the largest number $\beta \in \mathbb{N}$ such that $t^\beta v \neq 0$ and (3) quotient out $\langle v \rangle$. Repeat for $M/\langle v \rangle$.

**Theorem 4.37.** *Let $M$ be a finitely generated $\mathbb{N}$-graded $K[t]$-module. Then there exists a unique (up to ordering of bars) and finite barcode $V$ such that $M \cong V$*

*Proof.* If $M$ is isomorphic to a finite barcode then Lemma 4.34 proves that there exists a unique (up to ordering of bars) and finite barcode $V$ such that $M \cong V$ so it is enough to prove that $M$ is isomorphic to a finite barcode.

Since $M$ is finitely generated, by Proposition 4.20, $M$ has a finite generating set $B$ of homogeneous elements. Pick any element $b_0 \in B$ of least degree. By Lemma 4.36 we get $M \cong \langle b_0 \rangle \bigoplus M/\langle b_0 \rangle$. Then $M/\langle b_0 \rangle$ is generated by $\pi(B \setminus \{b_0\})$ in $M/\langle b_0 \rangle$ which is a strictly smaller set of graded elements. Thus we may do the same thing for $M/\langle b_0 \rangle$. This process ends so $M$ can be written as a direct sum of cyclic modules, which are isomorphic to bars according to Lemma 4.35. Thus $M \cong V(\alpha_0, \beta_0) \bigoplus V(\alpha_1, \beta_1) \bigoplus \ldots \bigoplus V(\alpha_I, \beta_I)$ for some sequence of $(\alpha_i, \beta_i)_{i \leq I}$. $\square$

# A. Category Equivalence

This section fills out the details of the proof sketched in Section 4.2 that there is a correspondence between $\mathbb{N}$-persistent $K$-vector spaces and $\mathbb{N}$-graded $K[t]$-modules for a given field $K$.

An $\mathbb{N}$-persistent $K$-vector space $(\{V_a\}, \phi)$ is given by a set of $K$-vector spaces $\{V_a\}_{a \in \mathbb{N}}$ and a linear function for each $a \leq b \in \mathbb{N}$, $\phi_{(a,b)} : V_a \to V_b$ such that $\phi_{(b,c)} \circ \phi_{(a,b)} = \phi_{(a,c)}$. A homomorphism $F : (\{V_a\}, \phi) \to (\{W_a\}, \psi)$ is a set of linear transformations $F_a : V_a \to W_a$ such that $F_b \circ \phi_{(a,b)} = \psi_{(a,b)} \circ F_a$.

We defined the $\mathbb{N}$-graded $K[t]$-module $\theta(\{V_a\}, \phi)$ by letting the group of the module be defined by $M = \bigoplus_{a \in \mathbb{N}} V_a$. The action of $k \in K$ on $m = (v_0, v_1, \ldots)$ is defined by $km := (kv_0, kv_1, \ldots)$ and the action of $t$ is defined by $tm := (0, \phi_{(0,1)}(v_0), \phi_{(1,2)}(v_1), \ldots)$. This is then linearly extended to $p(t) = \sum \lambda_a t^a \in K[t]$ so that $p(t)m := \sum \lambda_a t^a m$. The homogeneous elements of degree $a$ are elements $m = (v_0, v_1, \ldots) \in M$ such that $v_a \neq 0$ while $v_b = 0$ for all $b \neq a$. We define $\theta(F) : \bigoplus V_a \to \bigoplus W_a$ by

$$\theta(F)(v_0, v_1, \ldots) := (F_0(v_0), F_1(v_1), \ldots).$$

To prove that $\theta$ is a functor we need to prove that for any $\mathbb{N}$-persistent $K$-vector spaces $(\{U_a\}, \phi), (\{V_a\}, \psi), (\{W_a\}, \xi)$ and any $\mathbb{N}$-persistent $K$-vector space homomorphisms $F : (\{U_a\}, \phi) \to (\{V_a\}, \psi), G : (\{V_a\}, \psi) \to (\{W_a\}, \xi)$ we have

1. $\theta((\{V_a\}, \phi))$ is a $\mathbb{N}$-graded $K[t]$-module,
2. $\theta(F)$ is a $\mathbb{N}$-graded $K[t]$-module homomorphism,
3. $\theta(\mathrm{Id}) = \mathrm{Id}$,
4. $\theta(G) \circ \theta(F) = \theta(G \circ F)$.

*Proof.* 1. Since $K[t]_b = Kt^b$ we need to prove $Kt^b\theta((\{V_a\},\phi))_c \subseteq \theta((\{V_a\},\phi))_{(b+c)}$. We note that $\theta((\{V_a\},\phi))_c = V_c$ by definition so that

$$\begin{aligned} Kt^b\theta((\{V_a\},\phi))_c &= Kt^bV_c = K\{t^bv_c; v_c \in V_c\} = K\{\psi_{(c,b+c)}(v_c); v_c \in V_c\} \\ &= K\psi_{(c,b+c)}(V_c) \subseteq V_{(b+c)} \end{aligned}$$

since $(\{V_a\},\psi)$ is an $\mathbb{N}$-persistent $K$-vector space.

2. In order for $\theta(F)$ to be a module homomorphism we need $p(t)\theta(F)(v) = \theta(F)(p(t)v)$ for any $p(t) \in K[t]$. We get

$$\begin{aligned} p(t)\theta(F)(v) &= \\ \sum \lambda_a t^a(F_0(v_0), F_1(v_1), \ldots)) &= \\ \sum \lambda_a(0,0,\ldots,0,\phi_{(0,a)}(F_0(v_0)), \phi_{(1,a+1)}F_1(v_1)\ldots) &= \\ \sum(0,0,\ldots,0,\lambda_a\phi_{(0,a)}(F_0(v_0)), \lambda_a\phi_{(1,a+1)}(F_1(v_1))\ldots) &= \end{aligned}$$

where the vectors have $a$ zeroes at the beginning. Now since $F$ is a homomorphism of $\mathbb{N}$-persistent $K$-vector spaces we get

$$\begin{aligned} \sum(0,0,\ldots,0,\lambda_a\phi_{(0,a)}(F_0(v_0)), \lambda_a\phi_{(1,a+1)}(F_1(v_1))\ldots) &= \\ \sum \theta(F)(\lambda_a t^a v)) &= \\ \theta(F)\sum(\lambda_a t^a v)) &= \\ \theta(F)(p(t)v) & \end{aligned}$$

3. We get $\theta(\mathrm{Id})(v_0, v_1, \ldots) = (\mathrm{Id}_0(v_0), \mathrm{Id}_1(v_1), \ldots) = (v_0, v_1, \ldots)$ and thus $\theta(\mathrm{Id})(m) = m$ so $\theta(\mathrm{Id}) = \mathrm{Id}$

4. We check $\theta(G \circ F)(v) = \theta(G) \circ \theta(F)(v)$.

$$\begin{aligned} \theta(G \circ F)(v_0, v_1, \ldots) &= ((G \circ F)_0(v_0), (G \circ F)_1(v_1), \ldots) \\ &= (G_0 \circ F_0(v_0), G_1 \circ F_1(v_1), \ldots) \\ &= (G_0(F_0(v_0), G_1(F_1(v_1), \ldots) \\ &= \theta(G)(\theta(F)(v)) = \theta(G) \circ \theta(F)(v) \end{aligned}$$

where we only use that $F_a$ and $G_a$ are linear transformations so that $(G_a \circ F_a)(v_a) = G_a(F_a(v_a))$.

□

We now define $\theta^{-1}$ from the category of $\mathbb{N}$-graded $K[t]$-modules to $\mathbb{N}$-persistent $K$-vector spaces. Let $M = \sum V_a$ be a $\mathbb{N}$-graded $K[t]$-module. Then each $V_a$ is a $K$-module and thus a $K$-vector space. Now let $\theta^{-1}(M) = (\{V_a\}, t(.))$ where $t(.)_{(a,b)} : V_a \to V_b$ is the linear transformation given by multiplication by $t^{b-a}$. We know from Proposition 4.14 that any homomorphism $F : M = \bigoplus V_a \to N = \bigoplus W_a$ has an associated family of linear transformations $f_a : V_a \to W_a$ such that for any $m = (v_0, v_1, \ldots)$ we get $F(m) = (f_0(v_0), f_1(v_1), \ldots)$. Thus we let $\theta^{-1}(F) : (\{V_a\}, t(.)) \to (\{W_a\}, t(.))$ be defined by $\theta^{-1}(F)_a = f_a : V_a \to W_a$. To prove that $\theta^{-1}$ is a functor we need to prove that for any $\mathbb{N}$-graded $K[t]$-modules $M = \bigoplus V_a, N = \bigoplus W_a, U = \bigoplus U_a$ and any homomorphisms $F : M \to N, G : N \to U$ we have

1. $\theta^{-1}(M)$ is an $\mathbb{N}$-persistent $K$-vector space,
2. $\theta^{-1}(F)$ is an $\mathbb{N}$-persistent $K$-vector space homomorphism,
3. $\theta^{-1}(\mathrm{Id}) = \mathrm{Id}$,
4. $\theta^{-1}(G) \circ \theta^{-1}(F) = \theta^{-1}(G \circ F)$.

*Proof.*

1. It is clear that $\theta^{-1}(M) = (\{V_a\}, t(.))$ has a set of vector spaces and linear transformations $t(.)_{(a,b)}$ whenever $a \leq b$ but we need to check that $t(.)_{(b,c)} \circ t(.)_{(a,b)} = t(.)_{(a,c)}$. But $t(.)_{(b,c)} \circ t(.)_{(a,b)}$ is multiplication by $t^{b-a}$ followed by multiplication by $t^{c-b}$ which is the same as multiplication by $t^{c-a}$ which is $t(.)_{(a,c)}$.

2. We need to prove that for any $a, b \in \mathbb{N}, t_{(a,b)}(.) \circ \theta^{-1}(F)_a = \theta^{-1}(F)_b \circ t_{(a,b)}(.)$. Since $\theta^{-1}(F)_a = f_a$ and $\theta^{-1}(F)_b = f_b$ and $t_{(a,b)}(.)$ is the same as multiplication by $t^{b-a}$ we need to prove $t^{b-a} f_a(v_a) = f_b(t^{b-a} v_a)$ for any $v_a \in V_a$. This is true by definition since $F$ is an $\mathbb{N}$-graded $K[t]$-module homomorphism.

3. If $\mathrm{Id} : M \to M$ then $\mathrm{Id}_a : V_a \to V_a$ is also the identitiy. Thus $\theta^{-1}(\mathrm{Id}) : (\{V_a\}, t(.)) \to (\{V_a\}, t(.))$ is defined by $\theta^{-1}(\mathrm{Id})_a = \mathrm{Id}_a : V_a \to V_a$ which is the identity on each $V_a$.

4. For any $a \in \mathbb{N}$ we get $\theta^{-1}(G \circ F)_a : V_a \to U_a$ and $\theta^{-1}(G \circ F)_a(v_a) = (G \circ F)_a(v_a) = g_a \circ f_a(v_a) = g_a(f_a(v_a)) = \theta^{-1}(G)_a \circ \theta^{-1}(F)_a(v_a)$. Since homomorphisms between persistent vector spaces are fully defined by their action on each $V_a$ and the LHS acts the same as the RHS for any $a \in \mathbb{N}$ this means that they are the same.

□

We now need to check that $\theta\theta^{-1}$ is the identity in the category of $\mathbb{N}$-graded $K[t]$-modules and $\theta^{-1}\theta$ is the identity on the category of $\mathbb{N}$-persistent $K$-vector spaces.

$\theta\theta^{-1}(M) = \theta(\{V_a\}, t(.)) = \bigoplus V_a = M$. The action of elements from $K$ on $\theta\theta^{-1}(M)$ is the same by definition and so is the action of $t$. This makes the action of all elements in $K[t]$ the same. Thus $M = \theta\theta^{-1}(M)$.

$\theta^{-1}\theta(\{V_a\}, \phi)) = \theta^{-1}(\bigoplus V_a) = (\{V_a\}, t(.))$ where $t(.)$ is defined by applying $\phi$ and thus $\theta^{-1}\theta(\{V_a\}, \phi)) = (\{V_a\}, \phi)$.

# References


[Ale28] Paul Alexandroff. “Über den allgemeinen Dimensionsbegriff und seine Beziehungen zur elementaren geometrischen Anschauung. [In Mathematische Annalen/]”. In: *Mathematische Annalen* (1928).

[Gri07] Pierre Antoine Grillet. *Abstract Algebra, second edition*. Springer New York, NY, 2007.

[Chu09] Kim PT Chung MK Bubenik P. “Persistence diagrams of cortical surface data.” In: *Inf Process Med Imaging.* 21 (2009), pp. 386–97.

[Lee+11] Hyekyoung Lee et al. “Discriminative persistent homology of brain networks”. In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro.* 2011, pp. 841–844. DOI: 10.1109/ISBI.2011.5872535.

[Axl15] Sheldon Axler. *Linear Algebra Done Right.* Springer, 2015.

[Oud15] Steve Oudot. *Persistence Theory: From Quiver Representations to Data Analysis.* American Mathematical Society, 2015.

[Qai+16] Talha Qaiser et al. “Persistent Homology for Fast Tumor Segmentation in Whole Slide Histology Images”. In: *Procedia Computer Science* 90 (2016). 20th Conference on Medical Image Understanding and Analysis (MIUA 2016), pp. 119–124. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2016.07.033. URL: https://www.sciencedirect.com/science/article/pii/S1877050916312133.

[Rin16] Claus Michael Ringel. “Representation theory of Dynkin quivers. Three contributions”. In: *Frontiers of Mathematics in China* 11 (4 2016), pp. 765–814.

[Sco16] Simon Scott. *Introduction to Representation Theory of Quivers. [On https://uu.diva-portal.org/smash/get/diva2:972601/FULLTEXT01.pdf/].* B.Sc. Thesis. Sept. 2016.

[Woo17] Jong Chul Ye Woong Bae Jaejun Yoo. *Beyond Deep Residual Learning for Image Restoration: Persistent Homology-Guided Manifold Simplification.* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2017.

[Per18] Jose A. Perea. *A Brief History of Persistence.* 2018. arXiv: 1809.03624 [math.AT].

[GGM19] Noah Giansiracusa, Robert Giansiracusa, and Chul Moon. “Persistent Homology Machine Learning for Fingerprint Classification”. In: Dec. 2019, pp. 1219–1226. DOI: 10.1109/ICMLA.2019.00201.

[Meh19] Ahmed Fatmaoui Mehmet Aktas Esra Akbas. “Persistence homology of networks: methods and applications”. In: *Applied Network Science* 4 (2019).

[Sch19] Jonathan Schneider. “Geometry of Simplexes [On http://homepages.math.uic.edu/jschnei3/Writing/Simplexes/]”. In: (2019).

[Ant+21] Rika Antonova et al. “Sequential Topological Representations for Predictive Models of Deformable Objects”. In: *Proceedings of the 3rd Conference on Learning for Dynamics and Control*. Ed. by Ali Jadbabaie et al. Vol. 144. Proceedings of Machine Learning Research. PMLR, July 2021, pp. 348–360. URL: https://proceedings.mlr.press/v144/antonova21a.html.

[LR21] Nicole Lazar and Hyunnam Ryu. “The Shape of Things: Topological Data Analysis. [On https://chance.amstat.org/2021/04/topological-data-analysis/]”. In: *Chance* (Apr. 2021).

[PXL22] Chi Seng Pun, Kelin Xia, and Si Xian Lee. *Persistent-Homology-based Machine Learning and its Applications – A Survey*. 2022.

[TMO22] Renata Turkes, Guido F Montufar, and Nina Otter. “On the Effectiveness of Persistent Homology”. In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo et al. Vol. 35. Curran Associates, Inc., 2022, pp. 35432–35448. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/e637029c42aa593850eeebf46616444d-Paper-Conference.pdf.