Analysis of Simplicial Complexes
Juttmann, D.

D. Juttmann

Analysis of Simplicial Complexes

Bachelor Thesis

Supervisor: Prof. Dr. S.J. Edixhoven

Date Bachelor Exam: 2016–06–22

Mathematical Institute, Leiden University

Contents

Introduction

In this thesis we review some results from an article by M. Robinson [5], and take a more in depth look at the theory behind his work. The first chapter will introduce two types of topological spaces we will work with. Starting from a combinatorial concept called abstract simplicial complexes, we build simplicial complexes, as well as a finite version of these spaces.

In the second chapter we take a look at sheaves on abstract simplicial complexes. After some general theory of sheaves, we show how sheaves on simplicial complexes can be characterised. These results are then applied to study interference in wireless networks, which are represented by abstract simplicial complexes. These networks are assumed to transmit data between nodes on a single channel, which may cause interference at certain points in the network if there are multiple nodes sending information at the same time. We show how the language of sheaves can be used to describe which nodes can communicate simultaneously without creating interference.

The third chapter takes a look at the theory of persistent homology for simplicial complexes. While homology can provide information about the structure of a space, with persistent homology we can analyse a whole sequence (filtration) of nested simplicial complexes. Its purpose is to find topological features that persist throughout the filtration, i.e. that give the most relevant information about its general structure.

The homology of such filtrations will be a sequence of modules and homomorphisms between them, which are called persistence modules. We will prove a theorem that decomposes these persistence modules in simpler parts, which provides invariants that completely describe the homology of the filtration. This will be achieved through the theory of graded modules, with the help of a generalised version of the structure theorem of finitely generated modules over a PID.

Finally some experiments are done with the software Perseus for computing persistent homology. We generate random point clouds in $\mathbb{R}^{n+1}$ that resemble an $\mathbb{S}^n$, and show that Perseus can find the major n-th homology features of the data that agree with the n-th homology of the $\mathbb{S}^n$.

1 Abstract Simplicial Complexes

The spaces we will study are based on abstract simplicial complexes, which are a purely combinatorial construct.

Definition 1. *A (finite)* abstract simplicial complex *on a finite set N is a subset $X \subseteq \mathcal{P}(N)$ that is closed under taking subsets, i.e. for any $x \in X$ and $y \subseteq x$, we also have $y \in X$. The elements of an abstract simplicial complex are called* cells.

To understand its relation with topology, we take a quick look at simplices.

Definition 2. *The* convex hull *of a set $S \subseteq \mathbb{R}^n$ is the smallest convex subset of $\mathbb{R}^n$ containing S, notation $\mathrm{Conv}(S)$. For $k \in \mathbb{N}$, a k*-simplex *is the convex hull of $k+1$ vectors $\{v_0, \dots, v_k\} \subseteq \mathbb{R}^n$ such that $\{v_0 - v_1, \dots, v_0 - v_k\}$ are linearly independent. Any proper subset of these vectors also generates a simplex, which is called a* face *of the larger simplex.*

These simplices, with their induced topology from $\mathbb{R}^k$, can be used to construct larger topological spaces called simplicial complexes. This is achieved by gluing simplices together at their faces.

To illustrate the following definition, we will explain how an abstract simplicial complex can be interpreted as actual simplicial complex. Each cell represents a simplex, and a cell that is a subset of a larger cell, will represent a simplex that is the face of a larger simplex. This simplicial complex is called the realisation of the abstract simplicial complex.

Definition 3. *Let X be an abstract simplicial complex on set N, where without loss of generality we take $N = \{e_1, \dots, e_n\}$ the standard basis of $\mathbb{R}^n$. The* realization $|X|$ *of this abstract simplicial complex is the topological subspace of $\mathbb{R}^n$ defined as*

$$|X| = \bigcup_{x \in X} \mathrm{Conv}(x)$$

The simplex representing a cell $x \in X$ is simply the convex hull of x. Note that for every point $p \in |X|$, there exists a unique smallest set $x \in X$ such that $p \in \mathrm{Conv}(x)$: namely the set $x = \bigcap\{y \in X : p \in \mathrm{Conv}(y)\}$. This gives a map $q : |X| \to X \setminus \{\emptyset\}$, which is in fact surjective: for a non-empty $x \in X$, the set $\mathrm{Conv}(x) \setminus \bigcup\{\mathrm{Conv}(y) : y \subset x\}$ is non-empty, and its elements will be mapped by q to x. Therefore, a quotient space can be defined.

Definition 4. *Let X be a simplicial complex. The finite topological space X_{f} induced by X is the space $X \setminus \{\emptyset\}$ with the quotient topology from q.*

The space X_{f} can also be defined in a different way. Since the space is finite, each $x \in X_{\mathrm{f}}$ has a smallest open neighbourhood, namely the intersection of all its (finitely many!) open neighbourhoods. The collection of these minimal neighbourhoods automatically generates the topology of X_{f}.

Definition 5. *Let X be an abstract simplicial complex, and $x \in X \setminus \{\emptyset\}$, then the* star *of x is the set $\mathrm{Star}(x) = \{y \in X : x \subseteq y\}$. More generally, for non-empty subsets $Y \subseteq X \setminus \{\emptyset\}$ we also define $\mathrm{Star}(Y) = \bigcup_{x \in Y} \mathrm{Star}(x)$.*

The smallest neighbourhood of a point $x \in X_{\mathrm{f}}$ turns out to be $\mathrm{Star}(x)$.

2 Sheaves

If X is a topological space, then it can be seen as a category where its open sets are the objects, and the inclusion maps $\subseteq$ are the morphisms. Sheaves can be defined using contravariant functors from X to another category $\mathcal{C}$. Typical choices for $\mathcal{C}$ are the category of sets, or the category of abelian groups.

Definition 6. *Let X be a topological space. A* presheaf *$\mathcal{F}$ on X over a category $\mathcal{C}$ is a contravariant functor $X \to \mathcal{C}$:*

$$U \mapsto \mathcal{F}(U)$$
$$(U \subseteq V) \mapsto (r_{U,V} : A(V) \to A(U))$$

that is to say, for all $U \subseteq V \subseteq W$ open, $\mathcal{F}$ has the properties

$$r_{U,U} = \mathrm{id}_{\mathcal{F}(U)}$$
$$r_{U,W} = r_{U,V} \circ r_{V,W}$$

The maps $r_{\bullet,\bullet}$ are called restriction maps*. For open sets $U \subseteq X$, the elements of $\mathcal{F}(U)$ are called* sections *of U.*

Furthermore we will write $r_{U,V}(s) =: s|_U$ for restrictions of sections.

Definition 7. *Let X be a topological space and $\mathcal{C}$ the category of sets or abelian groups. A* sheaf *$\mathcal{F}$ on X over $\mathcal{C}$ is a presheaf on X over $\mathcal{C}$ with the following two additional properties:*

- *Let $U \subseteq X$ open, $(U_i)_{i\in I}$ an open cover of U. If $s, s' \in \mathcal{F}(U)$ are such that $\forall i \in I : s|_U = s'|_U$, then $s = s'$.*
- *Let $U \subseteq X$ open, $(U_i)_{i\in I}$ an open cover of U. If an $(s_i)_{i\in I} \in \prod_{i\in I} \mathcal{F}(U_i)$ has the property that $\forall i, j \in I : s_i|_{U_i\cap U_j} = s_j|_{U_i\cap U_j}$, then there exists an $s \in \mathcal{F}(U)$ such that $\forall i \in I : s|_{U_i} = s_i$.*

For both presheaves and sheaves we can define what morphisms are.

Definition 8. *Let $\mathcal{F}, \mathcal{G}$ (pre)sheaves on the space X. Then a* morphism of sheaves *$\phi : \mathcal{F} \to \mathcal{G}$ is a collection of maps $\phi_U : \mathcal{F}(U) \to \mathcal{G}(U)$ for all $U \subseteq X$ open, such that for all open sets $U \subseteq V$ the following diagram commutes:*

$$\begin{array}{ccc} \mathcal{F}(V) & \xrightarrow{\phi_U} & \mathcal{G}(V) \\ r_{U,V} \downarrow & & r_{U,V} \downarrow \\ \mathcal{F}(U) & \xrightarrow{\phi_V} & \mathcal{G}(U) \end{array}$$

2.1 The Étalé space

Given a sheaf $\mathcal{F}$ on X, it can also be realised in a different way. First of all we define the stalks and germs of a (pre)sheaf.

Definition 9. *Let $\mathcal{F}$ be a presheaf of sets on X, and $x \in X$. Then the* stalk *of $\mathcal{F}$ at x is the set*

$$\mathcal{F}_x = \Big(\bigsqcup_{U \ni x} \mathcal{F}(U) \Big) / \sim$$

where for sections $s \in \mathcal{F}(U), t \in \mathcal{F}(V)$ we have $s \sim t$ if there exists a neighbourhood $W \subseteq U \cap V$ of x such that $s|_W = t|_W$. The equivalence classes are called the germs *at x, with an equivalence class of a section s being called the germ of s at x, denoted s_x.*

With the stalks of $\mathcal{F}$ we can now create a topological space $\text{Spé}(\mathcal{F}) = \bigsqcup_{x \in X} \mathcal{F}_x$, with the following topology (see also [1] p. 5-6). Given any $U \subseteq X$ open and $s \in \mathcal{F}(U)$, then the set $\{s_x : x \in U\}$ is open. The topology generated by all these opens is the topology on $\text{Spé}(\mathcal{F})$. Note that there exists a natural continuous projection back to X defined by $\pi : \text{Spé}(\mathcal{F}) \to X : s_x \mapsto x$. This étalé space gives rise to a new sheaf $\mathcal{F}'$ on X:

$$\begin{aligned} U &\mapsto \{s : X \to F : \pi \circ s = \text{id}_X, s \text{ continuous}\} \\ (U \subseteq V) &\mapsto (s \mapsto s|_U) \end{aligned}$$

This construction always produces a sheaf, and $\mathcal{F}'$ is in fact the same sheaf as $\mathcal{F}$.

2.1.1 Stalk maps

Given a map of sheaves $\phi : \mathcal{F} \to \mathcal{G}$ on a space X, then for all $x \in X$ it induces maps on the stalks $\phi_x : \mathcal{F}_x \to \mathcal{G}_x$ that for any $s \in \mathcal{F}(U)$ maps $s_x \mapsto (\phi_U(s))_x$. This is a well defined map: for $s \in \mathcal{F}(U), t \in \mathcal{F}(V)$ with $s_x = t_x$, there exists a neighbourhood $W \in U \cap V$ of x with $s|_W = t|_W$. Because ϕ is a sheaf morphism, we find

$$\phi_U(s)|_W = \phi_U(s|_W) = \phi_U(t|_W) = \phi_U(t)|_W$$

and thus $(\phi_U(s))_x = (\phi_U(t))_x$. This shows the definition is independent of representatives.

2.2 Sheaves on Abstract Simplicial Complexes

Let X an abstract simplicial complex and $\mathcal{F}$ a presheaf on its corresponding finite space X_f. Then there is a simple way of describing the stalk at any point $x \in X_\text{f}$. Any section of an open neighbourhood of x can be restricted to the smallest neighbourhood $\text{Star}(x)$ of x. Then it follows that for each germ at x we can use its representative from $\mathcal{F}(\text{Star}(x))$ as a canonical representative. In essence, we can say that $\mathcal{F}_x = \mathcal{F}(\text{Star}(x))$.

Suppose now that we have two points $x \subseteq y$ in X_f. Then $\text{Star}(y) \subseteq \text{Star}(x)$, and with the above interpretation of $\mathcal{F}_x$, this inclusion gives rise to a map between stalks $\mathcal{F}_x \to \mathcal{F}_y$, which is simply the restriction function $r_{\text{Star}(y),\text{Star}(x)}$.

Definition 10. *For X an abstract simplicial complex and $\mathcal{F}$ a presheaf on X_f, The map $\mathcal{F}_{y \rightsquigarrow x} : \mathcal{F}_x \to \mathcal{F}_y = r_{\text{Star}(y),\text{Star}(x)}$ is called the* specialization function *induced by $\mathcal{F}$ and the specialization $y \rightsquigarrow x$.*

Because of the functoriality of the map $\big(\text{Star}(y) \subseteq \text{Star}(x)\big) \mapsto r_{\text{Star}(y),\text{Star}(x)}$, it follows immediately that the specialization functions have the following properties:

Lemma 1. *The specialization functions have the properties*

$$\forall x \in X_{\mathrm{f}} : \mathcal{F}_{x \rightsquigarrow x} = \mathrm{id}_{\mathcal{F}_x}$$
$$\forall x, y, z \in X_{\mathrm{f}}, x \subseteq y \subseteq z : \mathcal{F}_{z \rightsquigarrow y} \circ \mathcal{F}_{y \rightsquigarrow x} = \mathcal{F}_{z \rightsquigarrow x}$$

From just these stalks and the specialization functions, it is possible to create a new sheaf. In case the initial presheaf was already a sheaf, the new sheaf will be the same as the original. Otherwise the process creates the associated sheaf. This works as follows. Given for a sheaf $\mathcal{F}$ on X_{f} the stalks $\{\mathcal{F}_x\}_{x \in X_{\mathrm{f}}}$ and specialization functions $\{\mathcal{F}_{y \rightsquigarrow x}\}_{x \subseteq y \in X_{\mathrm{f}}}$, we create a new sheaf $\mathcal{F}'$ on X:

$$U \mapsto \big\{(s(x) \in \mathcal{F}_x)_{x \in U} : \forall x, y \in U, x \subseteq y : \mathcal{F}_{y \rightsquigarrow x}(s(x)) = s(y)\big\} \qquad (1)$$
$$(U \subseteq V) \mapsto \big(r_{U,V} : (s(x) \in \mathcal{F}_x)_{x \in V} \mapsto (s(x) \in \mathcal{F}_x)_{x \in U}\big)$$

The restriction maps here are simply restrictions of tuples.

Theorem 1. *Let $X, \mathcal{F}, \mathcal{F}'$ as above, where $\mathcal{F}$ is a sheaf. Then the map $\phi : \mathcal{F} \to \mathcal{F}'$ defined by*

$$\phi_U : \mathcal{F}(U) \to \mathcal{F}'(U) : s \mapsto (s_x)_{x \in U}$$

is a sheaf isomorphism.

Proof. First of all, note that this is in fact a morphism of sheaves, because for any $s \in \mathcal{F}(U)$ and $W \subseteq U$ we have

$$\phi_U(s|_W) = ((s|_W)_x)_{x \in W} = (s_x)_{x \in W} = (s_x)_{x \in U}|_W = \phi_U(s)|_W$$

Now it suffices to prove that the stalk maps induced by ϕ are all isomorphisms [7]. Therefore let $x \in X_{\mathrm{f}}$, and consider the stalk map ϕ_x which does the following with $s_x \in \mathcal{F}_x$ for some $s \in \mathcal{F}(\mathrm{Star}(x))$:

$$\phi_x(s_x) = (\phi_{\mathrm{Star}(x)}(s_x))_x = ((s_y)_{y \in \mathrm{Star}(x)})_x$$

This map is certainly injective: For $s, t \in \mathcal{F}(\mathrm{Star}(x)), s \neq t$, the elements $(s_y)_{y \in \mathrm{Star}(x)}$ and $(t_y)_{y \in \mathrm{Star}(x)}$ will not be in the same germ of $\mathcal{F}'$. Furthermore, it is also surjective: let $(s(y) \in \mathcal{F}_y)_{y \in \mathrm{Star}(x)}$ a section in $\mathcal{F}'(\mathrm{Star}(x))$. This tuple contains $s(x) = t_x$ for some $t \in \mathcal{F}(\mathrm{Star}(x))$ as one of its elements. By definition of the specialization functions, the other coefficients of the tuple must be of the form

$$s(y) = \mathcal{F}_{y \rightsquigarrow x}(s(x)) = \mathcal{F}_{y \rightsquigarrow x}(t_x) = t_y$$

Hence this section is of the form $(t_y)_{y \in \mathrm{Star}(x)}$, and therefore its germ at x is the image of t_x under ϕ_x. This shows that each germ is in the image of ϕ_x, concluding the proof. □

2.2.1 Cellular sheaves

Given a space X_{f} induced by some abstract simplicial complex X, the question arises whether any collection of stalks and specialization functions can be turned into a sheaf. While the process described by (1) in the previous section produces a sheaf, this new sheaf does not necessarily have the same stalks and specialization functions. Theorem 2 will show that, for the new sheaf to have the same stalks and specialization functions, it is required that the specialization functions (still denoted with $\mathcal{F}_{y \rightsquigarrow x}$) have the two properties listed in lemma 1.

In fact we can see X_{f} in another way as a category, where its cells are the objects and the inclusion maps the morphisms. Then the conditions from lemma 1 can be summarised with the condition that the map $T : (x \subseteq y) \mapsto \mathcal{F}_{y \rightsquigarrow x}$ must be functorial.

Definition 11. *Let X be an abstract simplicial complex on a set N. A* cellular sheaf *on X_{f} is a covariant functor $T : X_{\mathrm{f}} \to \mathbf{Set}$ where*

- *for $x \in X$, $T(x)$ is called the* stalk *of T at x;*
- *for $x, y \in X, x \subseteq y$ the map $T(x \subseteq y) =: T_{y \rightsquigarrow x}$ is called the* specialization function *for the specialization $y \rightsquigarrow x$.*

Theorem 2. *Let T be a cellular sheaf on X_{f}. The sheaf $\mathcal{T}$ on X_{f} induced by T through the process described in theorem 1 has the same stalks and specialization functions as T.*

Proof. we show the maps $\phi_x : T(x) \to \mathcal{T}_x : t \mapsto \big((T_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(x)}\big)_x$ are bijections that commute with the specialization functions. This map is well defined: the coordinates of the tuples $(T_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(x)}$ specialize to each other as required, because of the functoriality of T.

First of all, the map is injective. If $t, t' \in T(x)$ are two different elements from the stalk of T at x, then the tuples $(\mathcal{T}_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(x)}$ and $(\mathcal{T}_{y \rightsquigarrow x}(t'))_{y \in \mathrm{Star}(x)}$ are not the same, since their coordinates at x are t and t' respectively. Since they cannot be restricted to a smaller neighbourhood of x, their germs are not the same either, so ϕ_x is injective.

Next, let $\big((t(y))_{y \in \mathrm{Star}(x)}\big)_x$ be any germ of $\mathcal{T}$ at x. By definition of $\mathcal{T}$, this is of the form $\big((T_{y \rightsquigarrow x}(t(x)))_{y \in \mathrm{Star}(x)}\big)_x$, which shows that it is the image of $t(x) \in T_x$ under ϕ_x. Hence ϕ_x is surjective.

Finally, note that any specialization $z \rightsquigarrow x$ commutes with this map: for $t \in T(x)$ we find

$$\begin{aligned}
\mathcal{T}_{z \rightsquigarrow x}(\phi_x(t)) &= r_{\mathrm{Star}(z), \mathrm{Star}(x)}\big((T_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(x)}\big) \\
&= \big((T_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(x)}\big)_z \\
&= \big((T_{y \rightsquigarrow x}(t))_{y \in \mathrm{Star}(z)}\big)_z \\
&= \big((T_{y \rightsquigarrow z} \circ T_{z \rightsquigarrow x}(t))_{y \in \mathrm{Star}(z)}\big)_z \\
&= \phi_x(T_{z \rightsquigarrow x}(t))
\end{aligned}$$

□

2.3 The Link Complex

In this section we will focus on some of the results from [5]. Here we look at networks of wireless communication nodes. The nodes of such networks can communicate with each other only if they are within reach. Hence we can model such a wireless network as a graph: its nodes are the vertices of the graph, and an edge between a pair of vertices exists whenever the pair of nodes can communicate.

Definition 12. *Let (N, E) be a graph with nodes N and edges E. Then the* link complex *of (N, E) is the clique complex, i.e. the abstract simplicial complex $X = \{x \in \mathcal{P}(N) : \forall n, m \in x : (n, m) \in E\}$.*

Note that the link complex X contains still all connectivity information of the original graph. The singletons in X correspond to the nodes and cells with two elements are the edges.

We will now assume that all nodes in the wireless network are communicating on the same channel. This means that if two nodes are sending data at the same time, interference will occur at a third node that can receive information from both.

Definition 13. *Let X be the link complex of a network with nodes in N. A subset $S \subseteq N$ is called* interference free*, if for all $c, c' \in S, c \neq c'$ there is no $d \in N$ such that $\{c, d\} \in X$ and $\{c', d\} \in X$.*

That is to say, transmitting from all nodes in S causes no interference at any node in the network.

2.3.1 The transmission sheaf

To study interference on networks, we consider sheaves on the finite spaces induced by link complexes.

Definition 14. *Let X be the link complex of a network with nodes in N. The* transmission sheaf *is the cellular sheaf T of sets on X_{f} defined as*

$$c \mapsto \{n \in N : c \cup \{n\} \in X\} \cup \{\bot\}$$
$$(c \subseteq d) \mapsto \left(T_{d \rightsquigarrow c} : T(c) \to T(d) : \begin{cases} n \mapsto n & \text{if } n \in T(d) \\ n \mapsto \bot & \text{if } n \notin T(d) \\ \bot \mapsto \bot \end{cases} \right)$$

where $\bot$ is a 'silence' symbol that is not an element of N.

Since this transmission sheaf is a cellular sheaf, it can be extended to a sheaf on X_{f}. We can now look at sections over open sets $U \subseteq X$. They are tuples indexed by U with coordinates in the stalks of T, or more precisely functions $s : U \to N \cup \{\bot\}$, such that for all $c, d \in U$

- $s(c) \in T(c)$;
- if $c \subseteq d$, then $T_{d \rightsquigarrow c}(s(c)) = s(d)$

For sections on all of X (global sections), we have a nice interpretation. In the following two theorems, we will prove such sections give all possible lists of nodes in N that may transmit simultaneously without causing interference at any other node in the network. A section simply encodes for each clique of nodes in N which node is sending data to the nodes of that clique. The symbol $\bot$ indicates that the clique is not receiving data from any node.

Theorem 3. *Let X be a link complex and s a global section of the transmission sheaf T on X_{f}, then the set $S = \mathrm{img}(s) \setminus \{\bot\}$ is interference free.*

Proof. We first show that if $n \in S$, then $s(\{n\}) = n$. Because $n \in \operatorname{img}(s)$, we know that there exist $c \in X$ with $s(c) = n$. Since s is a section, we find that $n \in T(c)$, which means that $c \cup \{n\} \in X_{\mathrm{f}}$. Now from $c \subseteq c \cup \{n\}$ it follows that $T_{c\cup\{n\}\rightsquigarrow c}(n) = s(c \cup \{n\})$. Because $n \in T(c \cup \{n\})$, we find $T_{c\cup\{n\}\rightsquigarrow c}(n) = n$, therefore $s(c \cup \{n\}) = n$. Now considering that

$$T_{c\cup\{n\}\rightsquigarrow\{n\}}(s(\{n\})) = s(c \cup \{n\}) = n$$

and that $T^{-1}_{c\cup\{n\}\rightsquigarrow\{n\}}(\{n\}) = \{n\}$ we conclude that $s(\{n\}) = n$.

Now, if S has only one element, then there cannot be any interference. Therefore suppose $\#S \geq 2$, and let $n, n' \in S$ two distinct nodes. Suppose there is some $k \in S$ where interference occurs from n and n', that is to say $\{n, k\} \in X_{\mathrm{f}}$ and $\{n', k\} \in X_{\mathrm{f}}$. From the first part of the proof, we know that $s(n) = n$ and $s(n') = n'$. Then we find

$$\begin{aligned} s(\{n,k\}) &= T_{\{n,k\}\rightsquigarrow\{n\}}(s\{n\}) = T_{\{n,k\}\rightsquigarrow\{n\}}(n) = n \\ s(\{n',k\}) &= n' \end{aligned}$$

finally note that $T_{\{n,k\}\rightsquigarrow\{k\}}(s(\{k\})) = s(\{n, k\}) = n$. That implies $s(\{k\}) = n$, but by replacing n with n' we can similarly argue that $s(\{k\}) = n'$. This is a contradiction, and we conclude that there cannot be a node where interference occurs. Hence S is interference free. □

Theorem 4. *Let X be a link complex in N with transmission sheaf T, and $S \subseteq N$ interference free, then there is a unique section s such that* $\operatorname{img}(s) \setminus \{\bot\} = S$.

Proof. Note that for each $c \in X_{\mathrm{f}}$, there exists at most one $n \in S$ such that $c \cup \{n\} \in X$: if there were another such $n' \in S$, the nodes in c would get interference from n and n'. This means we can take $s : X \to N \cup \{\bot\}$ to be the map

$$c \mapsto \begin{cases} n & \text{if } c \cup \{n\} \in X \text{ and } n \in S \\ \bot & \text{otherwise} \end{cases}$$

This map satisfies the requirements for a global section:

- In either case $s(c) \in T(c)$ by definition of T.
- Let $c \subseteq d$, then we consider the two cases. If $s(c) = n$ then $s(d)$ must be either n or $\bot$, since $d \cup \{n\} \in X$ implies $c \cup \{n\} \in X$. If $n \in T(d)$ we get $d \cup \{n\} \in X$, and therefore $s(d) = n$, leading to $T_{d\rightsquigarrow c}(s(c)) = n = s(d)$. If $n \notin T(d)$ we get $s(d) =\bot$, and also $T_{d\rightsquigarrow c}(s(c)) =\bot= s(d)$.
  In case $s(c) =\bot$, then also $s(d) =\bot$, and $T_{d\rightsquigarrow c}(s(c)) =\bot= s(d)$ follows.

Finally, suppose s, s' are two different global sections, then there exists a $c \in X_{\mathrm{f}}$ with $s(c) \neq s'(c)$. Without loss of generality we can assume $s(c) \neq\bot$, which means $s(c) \in \operatorname{img}(s) \setminus \{\bot\}$. However, if $s(c) \in \operatorname{img}(s') \setminus \{\bot\}$, then $s'(\{s(c)\}) = s(c)$, and this leads to

$$\begin{aligned} s'(\{s(c)\} \cup c) &= T_{\{s(c)\}\cup c\rightsquigarrow\{s(c)\}}(s'(\{s(c)\})) \\ &= T_{\{s(c)\}\cup c\rightsquigarrow\{s(c)\}}(s(c)) = s(c) \\ s'(\{s(c)\} \cup c) &= T_{\{s(c)\}\cup c\rightsquigarrow c}(s'(c)) \in \{s'(c), \bot\} \end{aligned}$$

Since $s(c) \notin \{s'(c), \bot\}$, this is a contradiction. We find that $s(c) \notin \operatorname{img}(s')\setminus\{\bot\}$, thus $\operatorname{img}(s) \setminus \{\bot\} \neq \operatorname{img}(s') \setminus \{\bot\}$. This means that the section inducing S is unique. □

3 Persistent Homology

In this section we will take a closer look at the homology of filtrations of simplicial complexes, and how persistent homology can be used to analyse them.

3.1 Cellular homology

The theory of persistent homology is based on the Homology groups of topological spaces. There is a general theory of homology, but since we are interested in homology of simplicial complexes, we will restrict ourselves to cellular homology. Cellular homology of simplicial complexes is in fact exactly the same as the standard homology, as is shown in [2] theorem 2.27 and Lemma 2.49. However, cellular homology is much easier to compute in practice.

From now on we work with simplicial complexes on a set of points that has a total ordering.

Definition 15. *Let X be an abstract simplicial complex and R be a ring. For $k \in \mathbb{N}$, the* k-th simplicial chain module *of $|X|$ with coefficients in R is the free R-module*

$$\Delta_k(X,R) = \bigoplus_{\sigma \in X : |\sigma| = k+1} R$$

where the generators are the k-cells of X, and its elements are written as $\sum_i r_i \sigma_i$ with $r_i \in R$ and $\sigma_i \in X$.

There exist natural maps $\Delta_k(X,R) \to \Delta_{k-1}(X,R)$ called boundary maps.

Definition 16. *Let X be an abstract simplicial complex on the set of points $\{v_1, \dots, v_n\}$. The* boundary map $\partial_k : \Delta_k(X,R) \to \Delta_{k-1}(X,R)$ *is the R-module morphism defined by the images of basis elements $\sigma = \{v_{n_1}, \dots, v_{n_k}\} \in X$, where $n_1 < \ldots < n_k$:*

$$\partial_k(\sigma) = \sum_{i=1}^{k} (-1)^i (\sigma \setminus \{v_{n_i}\})$$

These boundary maps have the property that $\partial_k \circ \partial_{k+1} = 0$ ([2] Lemma 2.1), in other words $\mathrm{img}(\partial_{k+1}) \subseteq \ker(\partial_k)$.

Definition 17. *For an abstract simplicial complex X and boundary maps ∂_k, the* k-th simplicial homology *of $|X|$ is defined as $H_k(|X|,R) = \ker(\partial_k)/\mathrm{img}\partial_{k+1}$.*

The assignment of k-th homology to a space is functorial, as mentioned in [2] proposition 2.9. This means that for a simplicial complex $|X|$ and a subcomplex $|Y| \subseteq |X|$, the inclusion map f induces a map $f_* : H_k(|Y|,R) \to H_k(|X|,R)$ between the k-th homology of the two spaces. In fact we can apply this to longer filtrations $X_0 \subseteq X_1 \subseteq X_2 \subseteq \ldots$ to obtain chains of maps on the k-th homology of these spaces. Next, we will take a closer look at persistent homology, which studies such chains.

3.2 Persistence Modules

The k-th homology for a sequence of nested spaces manifests itself as a sequence of R-modules $H_k^0 \to H_k^1 \to H_k^2 \to \ldots$. We will call such chains persistence modules.

Definition 18. *A* persistence module (M, t) *over a ring* R *is a sequence of* R*-modules* $M = (M_i)_{i \in \mathbb{N}}$ *and homomorphisms* $t = (t_i : M_i \to M_{i+1})_{i \in \mathbb{N}}$*. It is called* finite *if for all* $i \in \mathbb{N}$ *the* M_i *are finitely generated and there exists an* $n \in \mathbb{N}$ *such that for all* $i \geq n$ *the maps* t_i *are all isomorphims.*

In the category of persistence modules over a ring R, we have the following morphisms:

Definition 19. *Let* (M, t) *and* (N, s) *two persistence modules over a ring* R*. A* morphism of persistence modules $\phi : (M, t) \to (N, s)$ *is a sequence of* R*-module homomorphisms* $(\phi_i : M_i \to N_i)_{i \in \mathbb{N}}$ *such that the following diagram commutes:*

$$\begin{array}{ccccccc} M_0 & \xrightarrow{t_0} & M_1 & \xrightarrow{t_1} & M_2 & \xrightarrow{t_2} & \dots \\ \downarrow \phi_0 & & \downarrow \phi_1 & & \downarrow \phi_2 & & \\ N_0 & \xrightarrow{s_0} & N_1 & \xrightarrow{s_1} & N_2 & \xrightarrow{s_2} & \dots \end{array}$$

For a persistence module that represents the homology of a filtration of simplicial complexes, the elements of the R-modules represent in some way topological features in the filtration. Furthermore, their pre-images and images within the persistence module indicate how these features persist within the filtration. The longer a feature persists, the more representative it is of the general structure of the filtration. However, the persistence of arbitrary elements tells us little about the exact shape of the space. A central idea behind computing persistent homology is to break down persistence modules into a direct sum of simpler parts. A direct sum of persistence modules can be defined as follows.

Definition 20. *The direct sum of two persistence modules* (M, t) *and* (N, s) *is the persistence module*

$$(M_0 \oplus N_0) \stackrel{s_0 \oplus t_0}{\longrightarrow} (M_1 \oplus N_1) \stackrel{s_1 \oplus t_1}{\longrightarrow} \dots$$

Especially in the case where our ring is a field, it turns out that we can write a persistence module in a nice way as direct sum of indecomposable persistence modules. The indecomposable persistence modules can be characterised by persistence intervals.

Definition 21. *A* persistence interval *is a pair* $\langle m, n \rangle \in \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$*, with* $m < n$*.*

Such a persistence interval $\langle m, n \rangle$ represents the following indecomposable persistence module (M, t) over a field F. For all $i \in \mathbb{N}$ the modules M_i are F if $i \in [m, n)$, and zero otherwise; the t_i are identity maps if $M_i = M_{i+1} = F$, and zero otherwise. This leads us to the the theorem that provides a decomposition for persistence modules.

Theorem 5. *Let* (M, t) *be a finite persistence module over the field* F*. Then* (M, t) *can be decomposed into a direct sum of indecomposable persistence modules. The multiplicity of each indecomposable is uniquely determined.*

The set of persistence intervals corresponding to these indecomposable parts now describes the topological features of the filtration without redundancy. For each complex in the filtration they provide a basis for the homology of the space. The proof of theorem 5 requires some theory of graded rings and modules, which we will cover in the next section. We will show an equivalence between graded modules and persistence modules. The main idea is that for a persistence module (M, t) over F the maps t_i induce an endomorphism on the F-vector space $\bigoplus_{i\in\mathbb{N}} M_i$, namely the map $\bigoplus_{i\in\mathbb{N}} t_i$. This gives the space an $F[t]$ module structure, which is in fact a graded structure. Theorem 5 will then follow from a similar structure theorem for graded modules.

3.3 Graded Rings and Modules

Definition 22. *A* graded ring *is a commutative ring R together with a direct sum decomposition of abelian groups*

$$R = R_0 \oplus R_1 \oplus R_2 \oplus \dots$$

such that for all $i, j \in \mathbb{N}$ the inclusion $R_i R_j \subseteq R_{i+j}$ holds. The R_i are called the homogeneous parts *of R and their elements are called the* homogeneous elements *of R. For $r \neq 0$ homogeneous, its* degree $\deg(r)$ *is the integer such that $r \in R_{\deg(r)}$.*

A typical example of a graded ring is the polynomial ring over a field F, since we can write $F[t] = F \oplus Ft \oplus Ft^2 \oplus \dots$.

Definition 23. *Let $R = R_0 \oplus R_1 \oplus \dots$ be a graded ring. A* graded R-module *is an R-module M together with a direct sum decomposition of abelian groups*

$$M = \bigoplus_{i\in\mathbb{Z}} M_i$$

such that for all $i \in \mathbb{N}, j \in \mathbb{Z}$ the inclusion $R_i M_j \subseteq M_{i+j}$ holds. Again, the M_i are called the homogeneous parts of M. An element of a homogeneous part is called homogeneous and if non-zero, its degree is the corresponding index of the homogeneous part. Any element $m \in M$ can be written as sum of its homogeneous components *$m = \sum_i m_i$, with $m_i \in M_i$.*

Any graded ring $R = \bigoplus_{i\in\mathbb{N}} R_i$ is automatically also a graded R-module by taking for negative indices i simply $R_i = 0$. Note that it is possible to have different gradings on a module. For example, it is possible to shift the grading upwards or downwards, illustrated by the following notation.

Notation 1. *For $M = \bigoplus_{i\in\mathbb{Z}} M_i$ a graded R-module and $n \in \mathbb{Z}$, the grade shift $\Sigma^n M$ is the graded module with homogeneous parts $(\Sigma^n M)_i = M_{i-n}$.*

Definition 24. *The graded direct sum of two graded R-modules $A = \bigoplus_{i\in\mathbb{Z}} A_i$ and $B = \bigoplus_{i\in\mathbb{Z}} B_i$ is the graded module*

$$A \oplus B = \bigoplus_{i\in\mathbb{Z}} (A_i \oplus B_i)$$

If we forget the grading, this coincides with the usual direct sum of modules. Next we look at graded submodules and morphisms.

Definition 25. *Let M be a graded R-module. A submodule $N \subseteq M$ is called a* graded submodule *if $N = \bigoplus_{i\in\mathbb{Z}} (M_i \cap N)$.*

Definition 26. *Let M, N be two graded R-modules. An R-module homomorphism $f : M \to N$ is a* graded homomorphism *if* $\forall i \in \mathbb{Z} : f(M_i) \subseteq N_i$.

Lemma 2. *The kernel of a graded module homomorphism $f : M \to N$ is a graded submodule of M.*

Proof. The map f can be split in its grade components $f_i : M_i \to N_i$ such that $f = \bigoplus_{i\in\mathbb{Z}} f_i$. For $m = \sum_i m_i \in M$ we have $f(m) = \sum_i f(m_i) = \sum_i f_i(m_i)$. Thus $f(m)$ is only zero when all $f_i(m_i)$ are zero. Therefore, it follows that $\ker(f) = \bigoplus_{i\in\mathbb{Z}} \ker(f_i) = \bigoplus_{i\in\mathbb{Z}} \ker(f) \cap M_i$. □

Lemma 3. *Let M be a graded R-module and $N \subseteq M$ be a graded submodule. Then the quotient M/N is also graded, with its homogeneous parts being $(M/N)_i = (M_i + N)/N \cong M_i/N_i$.*

Proof. Note first that we can give $\bigoplus_{i\in\mathbb{Z}}((M_i+N)/N)$ a graded R-module structure: multiplication of $[m] \in (M_i + N)/N$ with a homogeneous element $r \in R_j$ is defined as $r \cdot [m] = [rm] \in (M_{i+j} + N)/N$. This is independent of representatives, because for $[m] = [m'] \in (M_i + N)/N$ we find $r[m] - r[m'] = [r(m - m')]$ and this is zero, since $m - m' \in N$.

Now the embedding maps $(M_i + N)/N \subseteq M/N$ induce the R-module homomorphism $f : \bigoplus_{i\in\mathbb{Z}}((M_i + N)/N) \to M/N : \sum_i [m_i] \mapsto [\sum_i m_i]$. This map is surjective: Let $[m] \in M/N$ with $m = \sum_i m_i$ the homogeneous parts of m, then $[m]$ has at least the element $\sum_i [m_i]$ in its pre-image. Also f is injective. Let $\sum_i [m_i] \in \bigoplus_{i\in\mathbb{Z}}((M_i+N)/N)$, where we can take homogeneous representatives $m_i \in M_i$. Now suppose that $f(\sum_i [m_i]) = [\sum_i m_i] = 0$. Then $\sum_i m_i \in N$, and since the m_i are homogeneous, it follows that $\forall i \in \mathbb{Z} : m_i \in N_i$. This implies that $\sum_i [m_i] = 0$.

Since f defines an R-module isomorphism, the grading on $\bigoplus_{i\in\mathbb{Z}}((M_i + N)/N)$ induces a grading on M/N, with the i-th homogeneous part being $(M/N)_i = (M_i + N)/N$. Finally, one of the isomorphism theorems (e.g. [6] theorem 8.1.4) gives us $(M_i + N)/N \cong M_i/(M_i \cap N)$, and because N is a graded submodule, this is equal to $M_i/(M_i \cap N) = M_i/N_i$. □

3.3.1 Structure of graded modules

Similar to the structure theorem for finitely generated modules of PIDs, there exists a version of this theorem for graded modules. Here we will prove a specific result for finitely generated graded modules over a polynomial ring $F[t]$, where F is a field. The graded module will be written as the direct sum of summands that come in two types: $\Sigma^n F[t]$ and $\Sigma^n F[t]/(t^r)$. We will call such modules indecomposables – if they are written as direct sum $M \cong N \oplus N'$ of two graded modules, one of those must be zero for the following reason. The homogeneous part M_n is isomorphic to F and cannot be decomposed as an F-module. Thus it must also be the homogeneous part of degree n of either N or N'. But since M_n generates the entire $F[t]$-module M, it also generates one of the modules N or N' entirely and the other must be zero.

Theorem 6. *Let F be a field and M a finitely generated graded module over $F[t]$. Then M can be written up to isomorphism as a direct sum of graded* *modules*

$$M \cong \Big(\bigoplus_i \Sigma^{a_i} F[t] \Big) \oplus \Big(\bigoplus_j \Sigma^{b_j} F[t]/(t^{r_j}) \Big) \tag{2}$$

with the $a_i, b_j \in \mathbb{Z}$ and $r_j \in \mathbb{N}_{>0}$ uniquely determined up to re-indexing.

Proof. Let $m_1, \dots, m_n$ be a finite set of non-zero homogeneous $F[t]$-module generators of M, with degrees $d_1, \dots, d_n$. Such a set exists, since we can take a finite set of generators, and then split up every element into its non-zero homogeneous parts. Then we can define the following graded homomorphism:

$$f : \bigoplus_{i=1}^{n} \Sigma^{d_i} F[t] \to M : e_i \mapsto m_i$$

where $(e_1, \dots, e_n)$ is the standard basis of $\bigoplus_{i=1}^{n} \Sigma^{d_i} F[t]$, the elements having degrees $d_i, \dots, d_n$ respectively. Since $F[t]$ is a PID and $\ker(f)$ is a graded submodule of a finitely generated $F[t]$-module, $\ker(f)$ is also finitely generated. Therefore we can again take a set of homogeneous non-zero generators $k_1, \dots, k_\ell$ for $\ker(f)$, which have degrees $d'_1, \dots d'_\ell$. This allows us to define the graded homomorphism

$$g : \bigoplus_{j=1}^{\ell} \Sigma^{d'_i} F[t] \to \bigoplus_{i=1}^{n} \Sigma^{d_i} F[t] : e'_j \mapsto k_j$$

with $(e'_1, \dots, e'_\ell)$ being the standard basis of $\bigoplus_{i=1}^{\ell} \Sigma^{d'_i} F[t]$, the elements having degrees $d'_1 \dots, d'_\ell$. As a map between free modules, we can now represent g as a $n \times \ell$ matrix $G = (g_{i,j})$, such that $k_j = \sum_{i=1}^{n} g_{i,j} e_i$. Since the k_j are homogeneous, its n components $g_{i,j} e_i$ are all homogeneous and of degree $\deg(k_j)$ if they are non-zero. This also means that all entries of the matrix G are homogeneous, more precisely $g_{i,j} \in F[t]_{d'_j - d_i}$. Through basis changes that maintain homogeneity of the basis elements, we will now make G diagonal.

We can use any of the following row or column operations:

- swap two rows/columns;
- multiply a row/column by a unit;
- add to a row/column a multiple of another row/column, where the multiple is a homogeneous element of suitable degree to preserve the homogeneity of the matrix.

The first two operations do not affect the homogeneity of the basis elements. Operations of the third type will take the following forms. Suppose that R us an $n \times n$ matrix representing such a row operation, i.e. a basis transformation in the codomain of g, and C an $\ell \times \ell$ matrix representing a column operation, i.e. a basis transformation in the domain. They will be of the form

$$R = \begin{pmatrix} 1 & & & \\ & \ddots & a & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & b & \ddots & \\ & & & 1 \end{pmatrix}$$

i.e. matrices that have ones on the diagonal, together with exactly one other non-zero homogeneous entry elsewhere. Then the original map $g(v) = Gv$ gives us a new map $Rg(v) = (RGC)(C^{-1}v)$ on the transformed bases with matrix RGC. Thus the columns of R and C^{-1} describe the new bases on the codomain and domain respectively in terms of the old bases. Note that C^{-1} is the same as C, but with b replaced with $-b$.

Suppose the entry a of R sits at index (i, i'). Since left multiplication with R maintains homogeneity of G, and because for all $0 < j \leq \ell$ the value $g_{i,j} + ag_{i',j}$ must be homogeneous, it follows that $\deg(a) = (d'_j - d_i) - (d'_j - d_{i'}) = d_{i'} - d_i$. Furthermore, R changes the basis of the codomain $(e_1, \dots, e_n)$ by replacing $e_{i'}$ with $e_{i'} + ae_i$. Considering the degree of a, this new basis vector is again homogeneous of degree $d_{i'}$.

Similarly, suppose the entry b of C sits at index (j, j'), Since right multiplication with C also maintains homogeneity of G, and because for all $0 < i \leq n$ the value $g_{i,j'} + bg_{i,j}$ must be homogeneous, it follows that $\deg(b) = (d'_{j'} - d_i) - (d'_j - d_i) = d'_{j'} - d'_j$. Also C^{-1} changes the basis of the domain $(e'_1, \dots, e'_\ell)$ by replacing $e'_{j'}$ with $e'_{j'} - be'_j$. Given the degree of b, this new basis vector is also homogeneous of degree $d'_{j'}$.

This shows that the third type of row/column operation also maintains homogeneity of the bases. Now the standard diagonalization process ([3] IV.26) can be applied to obtain the Smith normal form. The result is a diagonal homogeneous matrix $H = (h_{i,j})$, and since we can multiply by units, we can assume the elements on the diagonal to be either t^n, $n \in \mathbb{N}$ or zero. Using this basis, we can now write M as the cokernel of f:

$$M \cong \operatorname{img}(f)/\ker(f) = \bigoplus_{i=1}^{\min(n,\ell)} \Sigma^{\deg(c_i)} F[t]/(h_{i,i})$$

where $c_1, \dots, c_\ell$ are the degrees of the new basis vectors of the domain of H. Any entries $h_{i,i}$ that are zero contribute to the free part of (2), whereas the non-zero entries lead to quotients $F[t]/(t^n)$. In case $h_{i,i}$ is constant, the quotient is zero and can be omitted.

Uniqueness. Now suppose that we have two finite direct sums of indecomposables that are isomorphic:

$$M_1 = \bigoplus_{\substack{b \in \mathbb{Z} \\ r \in \mathbb{N}_{>0} \cup \{\infty\}}} M(b,r)^{m_1(b,r)} \xrightarrow[f]{\sim} \bigoplus_{\substack{b \in \mathbb{Z} \\ r \in \mathbb{N}_{>0} \cup \{\infty\}}} M(b,r)^{m_2(b,r)} = M_2 \qquad (3)$$

where we use the notation $M(b, r) = \Sigma^b F[t]/(t^r)$ and $M(b, \infty) = \Sigma^b F[t]$, and m_1 and m_2 are functions $\mathbb{Z} \times (\mathbb{N}_{>0} \cup \{\infty\}) \to \mathbb{N}$ with a finite support that encode the multiplicities of each indecomposable. Note that if we forget the grading of any decomposition, we get a decomposition that follows the standard structure theorem for finitely graded modules over a PID. Since such decompositions are unique, the number of non-zero summands in (3) must be equal on both sides. We will now prove that $m_1 = m_2$ with induction on the number of non-zero summands.

In the base case there are no non-zero modules in (3), so $m_1 = m_2 = 0$. Therefore, suppose there is at least one non-zero summand. For M_1 there exists a minimal index $i_{\min}$ such that the homogeneous part $M_{1,i_{\min}}$ is non-zero. This index is also minimal for M_2, since f is an isomorphism of graded modules and $f(M_{1,i_{\min}}) = M_{2,i_{\min}}$. Furthermore we have for all $n \in \mathbb{N}$ the following maps that multiply by t^n.

$$\begin{array}{ccc} M_{1,i_{\min}} & \xrightarrow{\cdot t^n} & M_{1,i_{\min}+n} \\ \downarrow f & \circlearrowleft & \downarrow f \\ M_{2,i_{\min}} & \xrightarrow{\cdot t^n} & M_{2,i_{\min}+n} \end{array}$$

Now there are two options. If all the maps $\cdot t^n$ are injective, then the graded submodule of M_1 generated by $M_{1,i_{\min}}$ is $M(i_{\min}, \infty)^{m_1(i_{\min},\infty)}$, and it is isomorphic through f to $M(i_{\min}, \infty)^{m_2(i_{\min},\infty)}$. Note also that $m_1(i_{\min}, \infty) = m_2(i_{\min}, \infty) = \dim_F(M_{1,i_{\min}})$. If not all maps $\cdot t^n$ are injective, there exists a minimal $n_{\min} \in \mathbb{N}$ such that $\ker(\cdot t^{n_{\min}}) \subseteq M_{1,i_{\min}}$ is not zero. Then the graded submodule of M_1 generated by this kernel is $M(i_{\min}, n_{\min})^{m_1(i_{\min},n_{\min})}$. Using the same process we find the submodule $M(i_{\min}, n_{\min})^{m_2(i_{\min},n_{\min})} \subseteq M_2$, which is again through f isomorphic with the submodule of M_1. Again, $m_1(i_{\min}, n_{\min}) = m_1(i_{\min}, n_{\min}) = \dim_F(\ker(\cdot t^{n_{\min}}))$.

In either case we find that for some pair (m, b) the functions m_1 and m_2 coincide, and that the corresponding submodules $M(b, r)^{m_1(b,r)} \subseteq M_1$ and $M(b, r)^{m_2(b,r)} \subseteq M_2$ are isomorphic through f. If we then quotient out these submodules, we are left with two isomorphic modules with a decomposition with strictly fewer non-zero summands. By the induction hypothesis these two direct sums have the same multiplicity functions. Thus we conclude that also for the original sums in (3) the function m_1 and m_2 are equal. □

3.3.2 Example of diagonalization

Let F be a field. The process of diagonalizing a homogeneous $F[t]$-valued matrix on homogeneous basis, is best illustrated with an example. Suppose that we have a 3×2 matrix G that is written on bases (e'_1, e'_2) of degrees $(3, 4)$ for the domain, and (e_1, e_2, e_3) of degrees $(1, 2, 3)$ for the codomain,

$$G = \begin{pmatrix} t^2 & 4t^3 \\ 2t & 3t^2 \\ 1 & 2t \end{pmatrix}$$

The following steps can be taken to obtain a homogeneous diagonal form. First the gcd of all entries is moved to the top left corner of the matrix by switching the first and third row. This corresponds to the new basis (e_3, e_2, e_1) of the codomain.

$$G_1 = \begin{pmatrix} 1 & 2t \\ 2t & 3t^2 \\ t^2 & 4t^3 \end{pmatrix}$$

Now the first column can be cleared by subtracting the first row $2t$ times from the second, and t^2 times from the third.

$$G_2 = \begin{pmatrix} 1 & 0 & 0 \\ -2t & 1 & 0 \\ -t^2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2t \\ 2t & 3t^2 \\ t^2 & 4t^3 \end{pmatrix} = \begin{pmatrix} 1 & 2t \\ 0 & -t^2 \\ 0 & 2t^3 \end{pmatrix}$$

The 3×3 matrix tells us the corresponding change of the basis of the codomain: the new basis will be $(-t^2 e_1 - 2te_2 + e_3, e_2, e_1)$. Next we subtract the first column $2t$ times from the second.

$$G_3 = \begin{pmatrix} 1 & 2t \\ 0 & -t^2 \\ 0 & 2t^3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -t^2 \\ 0 & 2t^3 \end{pmatrix}$$

Here the inverse of the 2×2 matrix gives the new basis for the domain, namely $(e'_1, 2te_1 + e'_2)$. At this point, the first row and column are clear, and the problem has been reduced to a smaller 2×1 matrix. Continuing the same process, we only need one more step to reach a diagonal form.

$$G_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 2t & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -t^2 \\ 0 & 2t^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & t^2 \\ 0 & 0 \end{pmatrix}$$

The final homogeneous bases are $(e'_1, 2te_1 + e'_2)$ for the domain, and $(-t^2 e_1 - 2te_2 + e_3, 2te_1 - e_2, e_1)$ for the codomain.

3.3.3 Graded modules and persistence modules

Finally we show how the structure theorem for graded modules translates to theorem 5. For this we use the following relation between graded modules and persistence modules.

Lemma 4. *Let R be a ring. Then there is an equivalence of categories between the category of persistence modules over R and the category of non-negatively*[1] *graded $R[t]$-modules given by the functor G:*

$$(M_0 \xrightarrow{t_0} M_1 \xrightarrow{t_1} M_2 \xrightarrow{t_2} \ldots) \mapsto \bigoplus_{i \in \mathbb{N}} M_i$$
$$\big(\phi : (M,t) \to (N,s)\big) \mapsto \bigoplus_{i \in \mathbb{N}} \phi_i$$

where for any $m \in M_i$ we define $t \cdot m = t_i(m)$ in the graded module $\bigoplus_{i \in \mathbb{N}} M_i$. The inverse functor G^{-1} is given by

$$\bigoplus_{i \in \mathbb{N}} M_i \mapsto (M_0 \xrightarrow{t_0} M_1 \xrightarrow{t_1} M_2 \xrightarrow{t_2} \ldots)$$
$$\Big(\phi : \bigoplus_{i \in \mathbb{N}} M_i \to \bigoplus_{i \in \mathbb{N}} N_i\Big) \mapsto (\phi|_{M_i})_{i \in \mathbb{N}}$$

with the t_i defined as the maps $m \mapsto tm$.

All requirements for functoriality are easily checked, so the proof will be omitted. Theorem 5 is now a simple corollary to this lemma and theorem 6.

Proof theorem 5. Let (M,t) any finite persistence module. Then G from lemma 4 will map it to a graded module that is finitely generated: we can simply take

[1] In the grading all subgroups with negative index are zero.

generators for all the M_i in the chain before it stabilizes. This allows us to apply theorem 6 and obtain a unique decomposition. Finally note that G^{-1} has the following effect on the indecomposable parts:

$$\Sigma^a F[t] \mapsto \Big(\underbrace{0 \to \ldots \to 0}_{a} \to F \overset{\text{id}}{\to} F \overset{\text{id}}{\to} \ldots\Big)$$
$$\Sigma^b F[t]/(t^r) \mapsto \Big(\underbrace{0 \to \ldots \to 0}_{b} \to \underbrace{F \overset{\text{id}}{\to} \ldots \overset{\text{id}}{\to} F}_{r} \to 0 \to \ldots\Big)$$

Thus $M_0 \to M_1 \to \ldots$ decomposes into chains of the desired form. This decomposition is unique by uniqueness of the decomposition of the corresponding graded module. □

3.4 Examples of Persistent Homology with Perseus

Several programs exist for the computation of persistent homology. Perseus [4] is one such program, that can take filtrations of simplicial complexes and compute the persistence intervals for the homology of the filtration. Because the persistence intervals are only well defined for homology with coefficients in a field, Perseus takes its coefficients in $\mathbb{F}_2$.

3.4.1 A small example

We start with a simple example of a filtration of simplicial complexes to illustrate how Perseus works. The sample filtration in figure 1 will be used. Note that for compatibility with Perseus, we start at time 1.

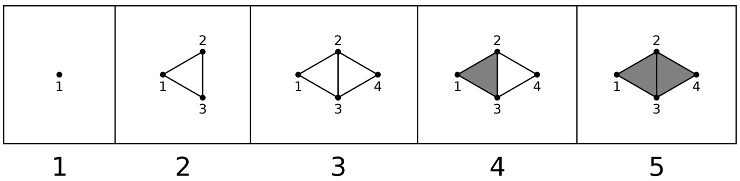

Figure 1: A small example of a filtration of simplicial complexes.

Since the example in figure 1 is a filtration, if a simplex appears at time t, it remains also in the complex at any time greater than t. The input for Perseus is therefore a file that lists for all occurring simplices what the earliest time is that they appear. Apart from first line, which is a separate parameter for Perseus, each line lists a simplex. The first number is its dimension, next are its vertices indexed by natural numbers, and finally its birth time.

```
1
0 1 1
0 2 2
0 3 2
0 4 3
1 1 2 2
1 1 3 2
```

```
1 2 3 2
1 2 4 3
1 3 4 3
2 1 2 3 4
2 2 3 4 5
```

The output given by Perseus is a list of persistence intervals for each homology up to at most the top dimension of the filtration of complexes. Each line shows two numbers, the first being the start of the interval and the second where it ends (i.e. the index where the persistence module is zero again). A second value of -1 indicates an interval that persists indefinitely. Our small example gives the following outputs for the 0-th homology

```
1 -1
```

and the 1-st homology

```
2 4
3 5
```

For this particular filtration the 0-th Homology is not very interesting. The complex is at every stage connected, thus there is only one generator that starts at time 1 and remains indefinitely. The 1-st homology however gives us the following results. At times 1 and 5 the space is contractible, so 1-st homology is trivial. At time 2 and 4 we have a space with one hole, and the homology of these spaces is generated by a cycle of 1-simplices around these holes. For time 2 the cycle $(1,2)+(2,3)+(3,1)$ is a generator, for time 4 we have the generator $(2,3)+(3,4)+(4,2)$. At time 3 we have a space with two holes, which happens to be generated by both the generators from times 2 and 4. Since we look at homology with coefficients in $\mathbb{F}_2$, we get the following chain of $\mathbb{F}_2$ vector spaces as persistence module:

$$H_1 : \qquad 0 \to \mathbb{F}_2 \xrightarrow{f} \mathbb{F}_2^2 \xrightarrow{g} \mathbb{F}_2 \to 0$$

We can give matrices for f and g on the bases

$$\begin{aligned} ((1,2) + (2,3) + (3,1)) &\text{ at } t = 2 \\ ((1,2) + (2,3) + (3,1)\,,\ (2,3) + (3,4) + (4,2)) &\text{ at } t = 3 \\ ((2,3) + (3,4) + (4,2)) &\text{ at } t = 4 \end{aligned}$$

which are $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0 & 1\end{smallmatrix}\right)$ respectively. Because of the well chosen generators, we can immediately see that it decomposes into two chains generated by cycles $(1,2) + (2,3) + (3,1)$ and $(2,3) + (3,4) + (4,2)$.

$$\begin{array}{ccccccccc} 0 & \to & \mathbb{F}_2 & \to & \mathbb{F}_2 & \to & 0 & \to & 0 \\ \oplus & & \oplus & & \oplus & & \oplus & & \oplus \\ 0 & \to & 0 & \to & \mathbb{F}_2 & \to & \mathbb{F}_2 & \to & 0 \end{array}$$

The output of Perseus tells us there should be a chain starting at time 2 that disappears at time 4, as well as one that starts at time 3 and disappears at time 5. This corresponds with the above analysis of the filtration.

3.4.2 Larger filtrations

For a more substantial experiment, we will take a look at some filtrations of simplicial complexes based on random point data. We will use Perseus to analyse sets of points in $\mathbb{R}^{n+1}$ that are clustered around the surface of an n-sphere. The data is generated by first uniformly sampling points on the surface of the standard n-sphere of radius 1. Then, to simulate an imperfect approximation of the sphere, we scale the points to a radius that is normally distributed for a given mean radius and standard deviation. This provides a point cloud that should roughly resemble the surface of the sphere, but with irregularities that may throw off simple analysis of the shape.

To analyse the shape of the point cloud, a filtration of Vietoris-Rips complexes will be generated from it. To create a Vietoris-Rips complex of a point cloud for a given distance threshold d, one first creates the graph on the points that connects any two points that have a distance of at most d between each other. The clique complex of this graph is then the corresponding Vietoris-Rips complex. This process can be applied for a sequence of increasing distances d, which results in a filtration of simplicial complexes.

Unlike in the first example, we will let Perseus generate a filtration of simplicial complexes directly from the point data. Instead of a threshold distance d, Perseus assigns to each point a radius r of a ball centred around that point, and then connects a pair of points in the graph if they have overlapping balls. This is equivalent to the Vietoris-Rips complex with distance d if we set $r = d/2$ for all points. An example of a cloud of 100 points that approximates an $\mathbb{S}^1$ of radius 1 is shown in figure 2. The distances of the points to the origin were sampled from a normal distribution with $\mu = 1$ and $\sigma = 0.2$.

We let the radii of the disks grow from 0 to 0.4 with 40 steps of size 0.01. The output from Perseus for the 1-st Homology is then as follows:

```
9 10
10 11
10 12
10 13
12 14
15 16
15 16
13 17
15 18
17 22
23 24
18 -1
```

The intervals show that most 1-cycles found are very short lived. However there is one cycle starting at $t = 18$ that persists indefinitely, making it the only significant feature of the 1-st persistent homology. This seems to correspond well with the approximate structure of an $\mathbb{S}^1$, which also has one single generator for its 1-st homology.

Finally, similar results can be achieved for higher dimensions. A similar experiment with 200 points normally distributed around an $\mathbb{S}^3$ with standard deviation

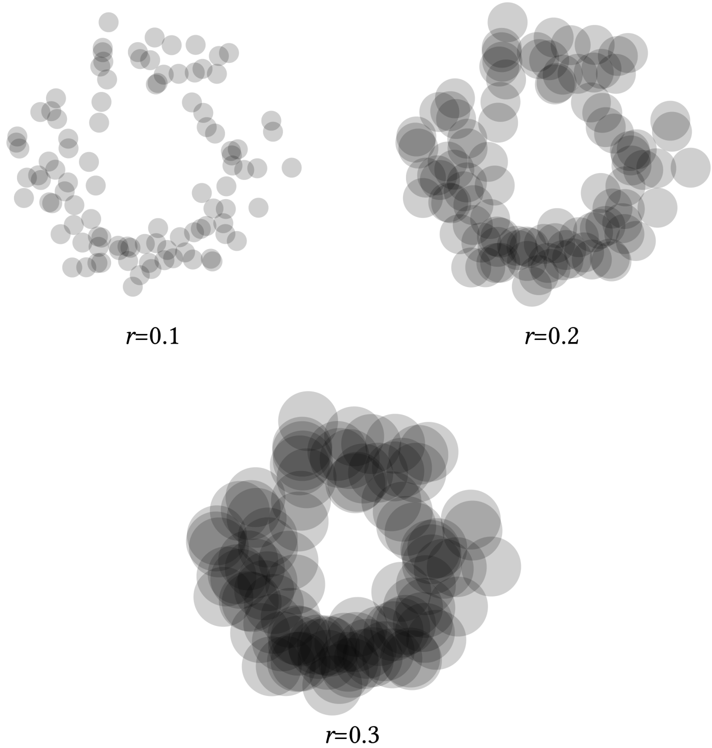

Figure 2: Example for points around an $\mathbb{S}^1$ with disks of different radii r.

0.1 gives the following output for its 3-rd Homology. Here 60 steps of size 0.01 were taken from $r = 0$ to $r = 0.6$.

```
52 53
52 -1
```

Besides one short lived 3-cycle only one significant generator is found. This again corresponds with the 3-rd homology of the $\mathbb{S}^3$ only being generated by a single 3-cycle.

Unfortunately for higher dimensions Perseus's system demands become very high. When using few points on the n-sphere or small radii, no n-cycles will be found in the n-th homology. Using too many points or radii that are too large, Perseus can easily run out of memory.

References

[1] Bredon, G. E. (1997). *Sheaf Theory (Second Edition)*. Springer New York.

[2] Hatcher, A. (2002). *Algebraic Topology*. Cambridge University Press.

[3] Mac Duffee, C.C. (1933). *The Theory of Matrices*. Springer Berlin Heidelberg.

[4] Nanda, V. *Perseus, the Persistent Homology Software*. `http://www.sas.upenn.edu/~vnanda/perseus`, accessed 2016-04-11.

[5] Robinson, M. (2013). Analyzing wireless communication network vulnerability with homological invariants. arXiv:1311.1532v1 [cs.NI].

[6] Shult, E., Surowski, D. (2015). *Algebra: A Teaching and Source Book*. Springer International Publishing.

[7] Stacks Project. Tag 007S, accessed 2016-06-10.