

UNIVERSITY OF QUEENSLAND

HONOURS THESIS

Persistence Modules over \mathbb{R}

Author:

Mitchell Riley

Supervisor:

A/Prof Benjamin Burton

November 11, 2014

Persistence is to the character of man as carbon is to steel.

NAPOLEON HILL

Acknowledgements

I would like to express my gratitude to my supervisor Ben Burton for putting up with my frequent changes in thesis topic and direction. I apologise that this computational topology project turned out to have little to do with computation or topology!

I also owe a debt to my fellow honours students, without whom this year would not have been nearly as much fun.

Contents

1	Introduction	1
2	Background Material	2
2.1	Free modules	2
2.2	Graded rings and modules	2
2.3	Structure theorem	3
2.4	Simplicial complexes	4
2.5	Simplicial homology	5
2.6	Singular homology	8
2.7	Filtrations	9
2.8	Multisets	11
3	Standard Persistent Homology	12
3.1	Persistent homology groups	12
3.2	Persistence modules	13
3.3	Algorithms	15
4	Persistence Modules	16
4.1	Interval modules	17
4.2	Interval decompositions	18
4.3	Quiver calculations	20
5	Persistence Diagrams	22
5.1	Abstract r-measures	22
5.2	The persistence measure	23
5.3	Equivalence	25
5.4	q-tame persistence modules	28
6	Stability	30
6.1	Interleaving distance	30
6.2	Bottleneck distance	32
6.3	Interpolation lemma	36
6.4	Box lemma	38
6.5	The stability theorem	39
7	Category Theory Primer	43
7.1	Functors and natural transformations	44
7.2	Limits and colimits	45
8	The Observable Category	47
8.1	Construction	47
8.2	Structure and stability in Ob	52
A	Implementation in Haskell	57

1 Introduction

Topological data analysis is an emerging field of applied algebraic topology that provides noise-resistant methods of interpreting real-world data. The cornerstone of topological data analysis is *persistent homology*, a technique introduced by Zomorodian and Carlsson [21] which relates topological features of a data set over different levels of detail. Each feature can be assigned an interval that describes the levels of detail at which it appears. The main insight is that any important feature of the data set will appear over a large interval.

The algebraic object underlying persistent homology is the *persistence module*, which is a diagram of vector spaces indexed by a partially ordered set. In practice, this indexing set is a finite subset of the natural numbers. Section 3 shows that persistence modules of this sort have a simple structure, and always decompose into a sum of interval modules. Each interval is interpreted as a topological feature.

This set of intervals is known as the *persistence diagram* associated with the persistence module. One of the most important results on persistence diagrams is the stability theorem, which states that small changes in the initial data lead to small changes in the persistence diagram. This fact justifies the persistence diagram's use as an invariant of real-world data.

In proving this theorem, we need to pass to persistence modules indexed over the real line. Here we face a problem: not every persistence module decomposes as a sum of interval modules. Sections 4 and 5 of this thesis are dedicated to working around this issue and finding an appropriate definition for persistence diagrams in this setting. Here we follow arguments made by Chazal et al. [9]. With this foundation in place, we then prove the stability theorem in Section 6.

Finally, Sections 7 and 8 follow a recent paper by Chazal, Crawley-Boevey and de Silva [6] to show that, in a very precise sense, the difficulties we face above are all localised to the behaviour of persistence modules over infinitesimally small intervals. We define the category of persistence modules and relate it to the much better behaved ‘observable’ category. This allows us to prove a strong structure theorem for persistence modules over the real numbers.

The main contribution of this thesis is a self-contained treatment of persistence modules over the real line, collecting results from many different sources. A number of proofs have been expanded or simplified from those that appear in the literature.

2 Background Material

We begin with a review of the material necessary to understand standard persistent homology. We use \mathbf{N} , \mathbf{Z} and \mathbf{R} to denote the natural numbers, integers and real numbers respectively, as we reserve blackboard bold to later denote persistence modules.

2.1 Free modules

Definition 2.1. A *module over a ring R* , or *R -module*, is an abelian group M equipped with an action of R satisfying the following properties. Let 1_R denote the multiplicative identity of the ring, and rx the action of the ring element r on the module element x . For any $r, s \in R$ and $x, y \in M$:

- $r(x + y) = rx + ry$;
- $(r + s)x = rx + sx$;
- $(rs)x = r(sx)$;
- $1_R x = x$.

Example 2.2. If we take R to be a field, the above conditions are the conditions of a vector space, thus, a vector space over F is precisely an F -module.

Example 2.3. Any abelian group M is a \mathbf{Z} -module via the action $nx = \underbrace{x + \cdots + x}_{n \text{ times}}$.

Abelian groups are often referred to as \mathbf{Z} -modules for this reason.

Definition 2.4. The *free R -module on a set S* is the module formed by taking every formal linear combination of elements of S with coefficients in R . These linear combinations are given an R -module structure as follows. Let $a_i, b_i \in R$ and $s_i \in S$, then:

$$\begin{aligned} (a_1 s_1 + a_2 s_2 + \dots) + (b_1 s_1 + b_2 s_2 + \dots) &= (a_1 + b_1) s_1 + (a_2 + b_2) s_2 + \dots; \\ a(a_1 s_1 + a_2 s_2 + \dots) &= (aa_1) s_1 + (aa_2) s_2 + \dots \end{aligned}$$

The *free abelian group on S* is the free \mathbf{Z} -module on S .

2.2 Graded rings and modules

Definition 2.5. A *graded ring* is a ring with a decomposition into abelian groups:

$$R = \bigoplus_{n \in \mathbf{N}} R_n$$

such that multiplication satisfies $R_n R_m \subseteq R_{n+m}$. Elements of any R_n are called *homogeneous elements of degree n* and for any $r \in R_n$ we write $\deg r = n$.

Example 2.6. The prototypical example of a graded ring is the polynomial ring $R[t]$, which decomposes as:

$$R[t] = \bigoplus_{n \in \mathbf{N}} t^n \cdot R.$$

Given a module over a graded ring, we can have the module structure respect the grading of the ring.

Definition 2.7. A *graded module* over a graded ring R is a module M with a decomposition:

$$M = \bigoplus_{n \in \mathbf{N}} M_n$$

where the action of R satisfies $R_n M_m \subseteq M_{n+m}$.

We will see that the structure of standard persistent homology is equivalent to a graded module over a particular ring.

2.3 Structure theorem

Recall that a *principal ideal domain* (PID) is an integral domain in which every ideal is generated by one element. Such ideals are called *principal* ideals.

Example 2.8. The integers \mathbf{Z} are a simple example of a PID. Given any ideal, we can find the greatest common divisor of all its elements. This element must generate the whole ideal.

For an example of a ring that is *not* a PID, consider the ring $\mathbf{Z}[x]$. The ideal $\langle x, 2 \rangle$ cannot be generated by one element.

Example 2.9. For our purposes, an important example of a PID is the polynomial ring $\mathbf{k}[x]$, where \mathbf{k} is a field.

Finitely generated modules over a PID can be uniquely decomposed using a generalisation of the structure theorem for finitely generated abelian groups [13].

Theorem 2.10. *If R is a PID, every finitely generated module over R is isomorphic to one of the form*

$$R^\beta \oplus \left(\bigoplus_{i=1}^m R / \langle d_i \rangle \right)$$

where $d_i \in R$, $\beta \in \mathbf{N}$, and $d_i \mid d_{i+1}$.

When $R = \mathbf{Z}$, we recover the structure theorem for finitely generated abelian groups. Given a graded module, we can say more:

Theorem 2.11. *Every graded module over a PID R is isomorphic to one of the form*

$$\left(\bigoplus_{i=1}^n \Sigma^{\alpha_i} R \right) \oplus \left(\bigoplus_{i=1}^m \Sigma^{\gamma_i} R / \langle d_i \rangle \right)$$

where $d_i \in R$ are homogeneous elements with $d_i | d_{i+1}$, $\alpha_i, \gamma_j \in \mathbf{N}$. The symbol Σ^k denotes a formal upward shift of k steps in the grading, i.e., $\Sigma^k R$ is the graded ring with $(\Sigma^k R)_n = R_{n-k}$.

2.4 Simplicial complexes

To begin with, we restrict our focus to a class of topological spaces that have a particularly simple description: simplicial complexes. Intuitively, these are spaces constructed of vertices, edges, triangles, tetrahedra, and so on, with these shapes glued together along their ‘faces’.

Definition 2.12. An *abstract simplicial complex* $K = (V, S)$ is a set V together with a collection S of subsets of V called *simplices* such that for every $v \in V$, $\{v\} \in S$, and if $\tau \subseteq \sigma \in S$, then also $\tau \in S$.

The subsets $\{v\}$ of V are called the *vertices* of K . A simplex $\sigma \in S$ is of dimension k if σ has $k + 1$ elements. If $\tau \subset \sigma$, we say τ is a *face* of σ .

An *oriented simplex* is a simplex together with a choice of orientation for that simplex. Any ordering of the vertices determines an orientation, and two orientations are considered equivalent if they differ by an even permutation. We write $[\omega]$ for an oriented simplex, and use the ordered sequence $[v_1, v_2, \dots]$ to denote the oriented simplex with vertices $\{v_1, v_2, \dots\}$ and the given orientation.

Example 2.13. We will use the following simplicial complex as a running example. Let $K = (V, S)$ with:

$$\begin{aligned} V &= \{a, b, c, d\} \\ S &= \{\emptyset, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, d\}, \{d, c\}, \{a, b, d\}\} \end{aligned}$$

Abstract simplicial complexes are useful in practice because of their simple structure and combinatorial properties. In theoretical settings, it is more convenient to consider general topological spaces.

We can associate a topological space to an abstract simplicial complex by pairing each k -dimensional subset in S with the convex hull of $k + 1$ independent points in \mathbf{R}^d for some $d \geq k$. Such a space is called a *geometric realisation*. With this pairing we recover the notions of shape we are familiar with: 1-simplices correspond to edges, 2-simplices to triangles and so on. In such a realisation we require that the nontrivial intersection of any two simplices is one of their common faces.

Example 2.14. Figure 1 shows one realisation of our abstract simplicial complex K .

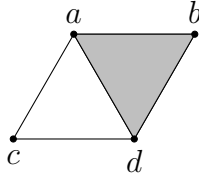


Figure 1: The simplicial complex K

Given any abstract simplicial complex K , we can always construct a geometric realisation.

Theorem 2.15. *Every abstract simplicial complex has a geometric realisation.*

Proof. Let n be the number of vertices in the simplicial complex K . Choose a one-to-one correspondence between the vertices of K and the unit vectors $\{e_1, e_2, \dots, e_n\}$ in \mathbf{R}^n . For each simplex in K , construct the convex hull of the corresponding points in \mathbf{R}^n . By construction, these convex hulls intersect either on a common face or not at all. Our geometric realisation is the union of these convex hulls, inheriting the subspace topology of \mathbf{R}^n . \square

We will refer to an abstract simplicial complex and its realisation interchangeably.

2.5 Simplicial homology

Homology is a procedure that, given a topological space, produces a sequence of abelian groups, H_k , that represent the ‘holes’ in the space. This set of groups is a topological invariant; any two homeomorphic spaces have the same homology groups. There are a number of different types of homology. We begin by introducing simplicial homology, which can only be applied to simplicial complexes.

The k -th homology group, $H_k(K)$, describes the k -dimensional holes in K . For example, the ‘hole’ in the middle of a circle is a 1-dimensional hole, so $H_1(\mathbb{S}^1)$ has rank 1. A 0-dimensional hole is interpreted as a gap between two components, so H_0 counts the path-connected components of K .

Example 2.16. Consider the n -sphere \mathbb{S}^n . The homology groups are as follows:

$$H_k(\mathbb{S}^n) = \begin{cases} \mathbf{Z} & k = 0 \text{ or } n; \\ 0 & \text{otherwise.} \end{cases}$$

The non-trivial H_0 represents the single connected component of \mathbb{S}^n , and non-trivial H_n the n -dimensional hole.

Example 2.17. Consider the standard embedding of a torus \mathbb{T} in \mathbf{R}^3 . The torus has a single component, two 1-dimensional holes (one circle through the middle,

one running around the loop), and one 2-dimensional hole (the void inside the torus). The corresponding homology groups are:

$$H_k(\mathbb{T}) = \begin{cases} \mathbf{Z} & k = 0 \\ \mathbf{Z} \oplus \mathbf{Z} & k = 1 \\ \mathbf{Z} & k = 2 \\ 0 & \text{otherwise.} \end{cases}$$

Definition 2.18. The k -th chain group of K , denoted C_k , is the free abelian group with basis the set of oriented simplices of K , where we let $[\sigma] = -[\tau]$ if σ and τ are the same simplex with the opposite orientation.

Elements of C_k are known as k -chains, and have the form $c = \sum_i n_i [\sigma_i]$, with σ_i k -simplices in K .

Let σ be a k -simplex with vertices $[v_0, \dots, v_k]$. Given such a k -simplex, we would like to calculate its boundary, $\partial_k \sigma$. The faces of σ are the $(k-1)$ -simplices with vertex sets of the form $[v_0, \dots, \hat{v}_i, \dots, v_k]$, where \hat{v}_i indicates that v_i is deleted from the sequence. We might expect the boundary to be the chain given by the sum of the faces, but a better choice is the following

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

When the dimension is unambiguous, we write ∂ for ∂_k .

Example 2.19. For generic simplices of dimension 0,1,2:

$$\begin{aligned} \partial[v_0] &= 0 \\ \partial[v_0, v_1] &= [v_1] - [v_0] \\ \partial[v_0, v_1, v_2] &= [v_1, v_2] - [v_0, v_2] + [v_0, v_1] \end{aligned}$$

Intuitively, we make this definition so that orientations are kept consistent as we travel around the boundary of a simplex.

The boundary operator ∂_k can be turned into a homomorphism $C_k \rightarrow C_{k-1}$ by extending it linearly on the basis elements of C_k . Together, these groups and maps form what is known as a *chain complex*, C_* .

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

The boundary operator allows us to define two important subgroups of each C_k : the *cycle group* $Z_k = \ker \partial_k$ and the *boundary group* $B_k = \text{im } \partial_{k+1}$. These groups match with our intuition of what a cycle and boundary should be; a cycle generalises a closed loop or surface, and a boundary is a cycle that bounds a higher dimensional shape.

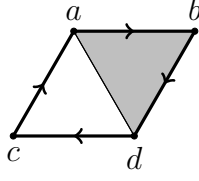


Figure 2: K with the cycle $[a, b] + [b, d] - [c, d] - [a, c]$ highlighted

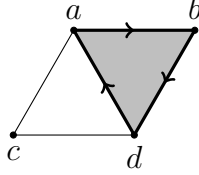


Figure 3: K with the boundary $\partial[a, b, d]$ highlighted

Example 2.20. Returning to our earlier example, the chain $[a, b] + [b, d] - [c, d] - [a, c]$ is an element of Z_1 , as:

$$\begin{aligned} \partial([a, b] + [b, d] - [c, d] - [a, c]) &= \partial[a, b] + \partial[b, d] - \partial[c, d] - \partial[a, c] \\ &= ([b] - [a]) + ([d] - [b]) - ([d] - [c]) - ([c] - [a]) \\ &= 0 \end{aligned}$$

We represent an oriented simplex $[x, y]$ as an arrow pointing $x \rightarrow y$. This chain is shown in Figure 2.

The chain $\partial[a, b, d] = [b, d] - [a, d] + [a, b]$ is an element of B_1 . This chain is shown in Figure 3. In fact, this chain is a generator for the group B_1 , as $[a, b, d]$ is the only 2-simplex. Note the consistent orientation of the edges.

Proposition 2.21. For any k , $\partial_{k-1} \circ \partial_k = 0$.

Proof. By definition,

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

Then,

$$\begin{aligned} \partial_{k-1} \partial_k[v_0, \dots, v_k] &= \sum_{j < i} (-1)^i (-1)^j [v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_k] \\ &\quad + \sum_{j > i} (-1)^i (-1)^{j-1} [v_0, \dots, \hat{v}_i, \dots, \hat{v}_j, \dots, v_k] \\ &= \sum_{j < i} (-1)^i (-1)^j [v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_k] \\ &\quad - \sum_{j < i} (-1)^i (-1)^j [v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_k] \\ &= 0 \end{aligned}$$

where in the second sum we swap the roles of i and j . □

This Proposition implies the inclusion $B_k \subseteq Z_k$.

Definition 2.22. The k -th simplicial homology group of K is $H_k = Z_k/B_k$. Elements of H_k are equivalence classes of ‘homologous’ cycles.

Example 2.23. Let us calculate H_1 of our complex K . This is easiest when we choose compatible bases for Z_1 and B_1 . Let σ and τ be the cycles:

$$\begin{aligned}\sigma &= [a, b] + [b, d] - [a, d] \\ \tau &= [a, d] - [c, d] - [a, c]\end{aligned}$$

It is not too hard to see that the cycle and boundary groups are

$$\begin{aligned}Z_1 &= \langle \sigma, \tau \rangle \\ B_1 &= \langle \sigma \rangle\end{aligned}$$

and therefore,

$$H_1 = Z_1/B_1 = \langle \tau \rangle \cong \mathbf{Z}.$$

This is the answer we expect, as our simplicial complex K has one 1-dimensional hole by inspection.

Definition 2.24. The k -th Betti number of K , β_k , is the rank of the free submodule of H_k .

2.6 Singular homology

We would like to define a notion of homology that applies to all topological spaces, not just simplicial complexes. Rather than considering spaces constructed from simplices, we will consider arbitrary continuous maps from simplices into our space.

Definition 2.25. The *standard k -simplex* Δ^k is the convex hull of the standard unit vectors $\{e_1, e_2, \dots, e_{k+1}\}$ in \mathbf{R}^{k+1} .

Definition 2.26. A *singular k -simplex* in a topological space X is a continuous map $\sigma : \Delta^k \rightarrow X$.

The image of σ may not resemble a simplex at all; all that is required is that σ be continuous. Now let C_k be the free abelian group with basis the set of singular k -simplices in X . As before, elements of C_k are called (singular) k -chains and have the form $c = \sum_i n_i \sigma_i$.

We define our boundary map $C_{k+1} \rightarrow C_k$ using a formula analogous to that in simplicial homology:

$$\partial_k(\sigma) = \sum_i (-1)^i \sigma|_{[v_0, \dots, \hat{v}_i, \dots, v_k]}.$$

Here we implicitly identify $[v_0, \dots, \hat{v}_i, \dots, v_k]$ with Δ^{k-1} , so $\sigma|_{[v_0, \dots, \hat{v}_i, \dots, v_k]}$ becomes a map $\Delta^{k-1} \rightarrow X$, i.e., a singular $(k-1)$ -simplex.

Definition 2.27. The k th singular homology group of K is $H_k = Z_k/B_k = \ker \partial_k / \operatorname{im} \partial_{k+1}$.

Because each C_k is so tremendously large, it is not clear that H_k defined in this way gives a useful invariant. However, in the case of simplicial complexes we have the following theorem.

Theorem 2.28. *The simplicial homology of a simplicial complex is isomorphic to the singular homology of its geometric realisation.*

Proof. See Hatcher [16]. □

Given a continuous map between topological spaces, there should be some relationship between the homology groups.

Theorem 2.29. *A continuous map $f : X \rightarrow Y$ induces a homomorphism $f_* : H_k(X) \rightarrow H_k(Y)$ for all k .*

Proof. The map $f : X \rightarrow Y$ induces a map $f_\# : C_k(X) \rightarrow C_k(Y)$ as follows. Given any k -simplex $\sigma : \Delta^k \rightarrow X$, we can compose with f to get a k -simplex $f_\#(\sigma) = f\sigma : \Delta^k \rightarrow Y$. Extending this map linearly via $f_\#(\sum_i n_i \sigma_i) = \sum_i n_i f_\#(\sigma_i)$ gives us the required map $C_k(X) \rightarrow C_k(Y)$.

This map $f_\#$ satisfies $f_\# \partial = \partial f_\#$, as:

$$\begin{aligned} f_\# \partial(\sigma) &= f_\# \left(\sum_i (-1)^i \sigma|_{[v_0, \dots, \hat{v}_i, \dots, v_k]} \right) \\ &= \sum_i (-1)^i f\sigma|_{[v_0, \dots, \hat{v}_i, \dots, v_k]} = \partial f_\#(\sigma). \end{aligned}$$

We have that $f_\#$ sends boundaries to boundaries as $\partial f_\#(\sigma) = f_\# \partial(\sigma)$. It also sends cycles to cycles, as $\partial \sigma = 0$ implies $\partial f_\#(\sigma) = f_\# \partial(\sigma) = 0$.

Therefore $f_\#$ induces a homomorphism $f_* : H_k(X) \rightarrow H_k(Y)$. □

Earlier, when we defined the chain groups using free abelian groups, we implicitly chose our *ground ring* R to be the integers. We can generalise this to any ring R by defining the chain groups to be free R -modules. Our homology groups H_k then also become R -modules, and we write $H_k(X; R)$ for homology with coefficients in R . When R is a field, the torsion disappears and the homology group H_k is a vector space. The homology group is completely described by the integer β_k , which may depend on the chosen field.

2.7 Filtrations

In its most general sense, a filtration of some mathematical structure S is a collection of subsets $\{S_i\}$ with index set I , such that $i \leq j \in I$ implies $S_i \subseteq S_j$. For persistent homology we are interested in filtrations of topological spaces.

Definition 2.30. Given a subset $A \subset \mathbf{R}$, a *filtration* of a topological space X is a family of subspaces $\{X_\alpha\}_{\alpha \in A}$ that are nested with respect to inclusion, that is, for all $\alpha \leq \alpha'$, $X_\alpha \subseteq X_{\alpha'}$. We use X^* to denote a filtration of X .

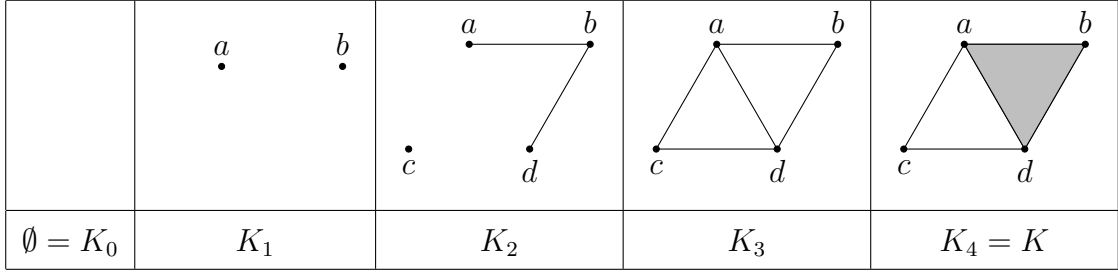


Figure 4: A filtration of K

An important type of filtration is one formed by sublevel sets of a real-valued function on a space. Given a function $f : X \rightarrow \mathbf{R}$, let $X_a = f^{-1}((-\infty, a])$ be the sublevel set. It is clear that for $a \leq b$, $X_a \subseteq X_b$. This gives us a filtration $X^f = \{X_a\}_{a \in \mathbf{R}}$.

For simplicial complexes, it is more natural to consider filtrations where each intermediate space is also a simplicial complex.

Definition 2.31. A *filtration* of a simplicial complex $K = (V, S)$ is a function $\rho : S \rightarrow \mathbf{R}$, such that for any face $\tau \subseteq \sigma$, $\rho(\tau) \leq \rho(\sigma)$. This monotonicity guarantees that for all $a \in \mathbf{R}$, $K_a = \rho^{-1}((-\infty, a])$ is a valid simplicial complex. Let $\{\rho_i\}_{i=1 \dots n}$ be the values ρ takes on the simplices of K . We then have the following sequence of subcomplexes:

$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K$$

Example 2.32. Figure 4 shows one possible filtration of our space K .

A practical source of filtered complexes comes from point cloud data.

Definition 2.33. Let (X, d_X) be a metric space. The *Vietoris-Rips complex* of X with parameter α is the simplicial complex with vertex set X and simplices given by subsets of X that are mutually no more than α distance apart. In other words,

$$\mathbf{VR}(X, \alpha) = \{\sigma \subseteq X \mid d_X(x_i, x_j) \leq \alpha \text{ for all } x_i, x_j \in \sigma\}.$$

Example 2.34. The simplest example is a set of points $P \subset \mathbf{R}^n$ endowed with the ordinary Euclidean metric. Figure 5 shows a set of points $P \subset \mathbf{R}^2$. A ball is placed around each point. On the right hand side we note the subsets of P in which the ball around any point contains the entire subset. These subsets become our simplices.

Clearly, if a simplex $\sigma \in \mathbf{VR}(X, \alpha')$ then $\sigma \in \mathbf{VR}(X, \alpha)$ for any $\alpha \geq \alpha'$. Again, we get a filtration $\{\mathbf{VR}(X, \alpha)\}_{\alpha \in \mathbf{R}}$.

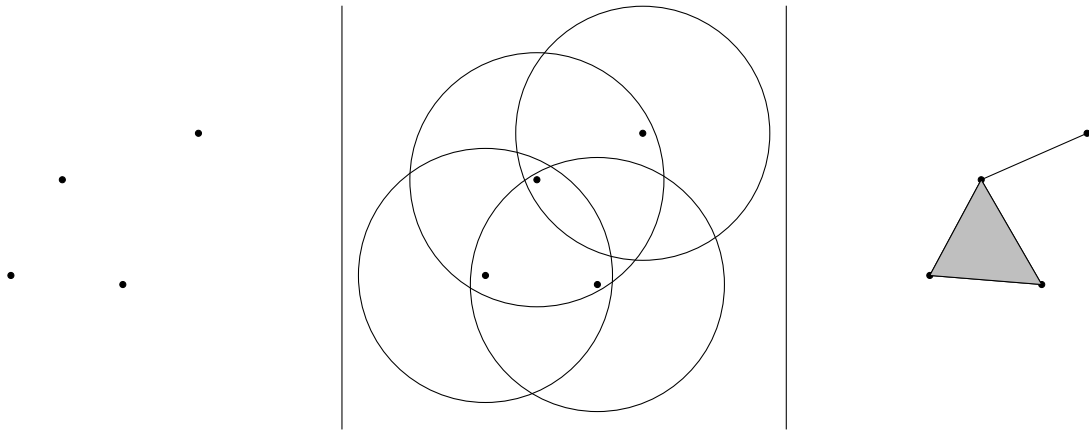


Figure 5: The Vietoris-Rips complex of a set of points for a particular α

2.8 Multisets

One of the main objects we will define in this work is the *persistence diagram*, a multiset of points in the plane. Informally, a multiset is a set in which elements are permitted to appear more than once.

Definition 2.35. A *multiset* is a pair $\mathbf{A} = (A, m)$, where A is an ordinary set and $m : A \rightarrow \mathbf{N}$ is the *multiplicity function*. For any $a \in A$, $m(a) \in \mathbf{N}$ is the number of times a appears in \mathbf{A} .

It will be convenient to think of a multiset \mathbf{A} as an ordinary set A , where we artificially distinguish elements with multiplicity $m(a) > 1$ by pretending each is labelled a_1, a_2, \dots , and so on. In what follows, we will often take this perspective without comment.

3 Standard Persistent Homology

We begin by describing persistent homology over a field for finite simplicial complexes. These conditions allow a complete characterisation of the persistent homology of a filtration via simple combinatorial objects called barcodes. In practice, because our data is finite, we always deal with finite simplicial complexes. Algorithms exist to compute the barcode of a filtration of a simplicial complex.

3.1 Persistent homology groups

A filtration is a way of gradually building up a topological space. We can extend our definition of homology to apply to each intermediate space in a filtration. Because our filtration is finite, without loss of generality we consider filtrations indexed by the natural numbers.

First, fix a dimension n and consider all homology to be done at this dimension. Let K be a filtration and let H_i denote the homology of the space K_i .

If we wish to use a filtration to find topological features of a space, we might consider, for example, the sequence of Betti numbers produced by moving up the filtration. Unfortunately, from these alone there is no way to distinguish true features of the space from the noise caused by our construction of the filtration.

In any filtration there are inclusion maps $\iota_i^j : K_i \rightarrow K_j$ for any $i \leq j$. By Theorem 2.6, each induces a homomorphism $\eta_i^j : H_i \rightarrow H_j$. When we move to the right in this sequence, we could gain homology classes or lose some as they become trivial or merge together.

The idea behind persistence is that any significant topological feature of the data should exist along a large interval of the filtration. This leads us to define the persistent homology groups of our filtration, which will represent homology classes in each space that still exist some number of spaces later in the filtration.

Definition 3.1 (Persistent homology). Let K_i be a filtration. The *persistent homology groups of K* , denoted H_i^j , are the images of the induced homomorphisms η_i^j .

These groups consist of the homology classes of K_i that are still “alive” in K_j . An alternative definition is to factor out the cycle group of K_i by the boundary group of K_j :

$$H_i^j = Z_i / (B_j \cap Z_i).$$

Both B_j and Z_i are subgroups of C_j , so the intersection $B_j \cap Z_i$ is well-defined.

We can also track the lifetime of an individual homology class. Let $\gamma \in H_i$. We say γ is *born at K_i* if it first appears in K_i , i.e., $\gamma \notin H_{i-1}^i$. The class *dies entering K_j* if it merges with an older class when we move from K_{j-1} to K_j . That is,

1. $\eta_i^{j-1}(\gamma) \notin H_{i-1}^{j-1}$ (γ in K_{j-1} is not the image of a class in K_{i-1}); but,
2. $\eta_i^j(\gamma) \in H_{i-1}^j$ (γ in K_j is the image of a class in K_{i-1}).

The *lifetime* of a homology class γ is the half-open interval $[i, j)$, where γ is born at K_i and dies entering K_j . If γ is born at K_i but then never dies, its lifetime is $[i, \infty)$.

Example 3.2. Consider the filtration given earlier as Figure 4. The homology class $\sigma = [a, b] + [b, d] - [a, d]$ is born at stage K_3 , as it is not the image of a homology class in any earlier K_i . It dies one step later in K_4 , as the class becomes trivial. The lifetime of σ is therefore $[3, 4)$.

3.2 Persistence modules

At this point it is not clear whether we can get a succinct description of the persistent homology of a filtration. In their foundational paper, Zomorodian and Carlsson [21] showed that a description exists for finite simplicial complexes when the homology coefficient ring is a field.

Intuitively, computing persistent homology requires us to find compatible bases for H_i and H_j . We begin by constructing one large structure that contains the homology groups for all steps in the filtration.

Definition 3.3. A *persistence module* \mathcal{M} is a family of R -modules $\{M_i\}_{i \geq 0}$ together with module maps $\varphi_i : M_i \rightarrow M_{i+1}$.

The set of homology groups H_i form a persistence module, with the maps φ_i induced by the inclusion maps $K_i \rightarrow K_{i+1}$.

Given a persistence module, we can construct a single graded module that contains the same data. Let $\mathcal{M} = \{M_i, \varphi_i\}$ be a persistence module over the ring R . Now define the graded module over $R[t]$:

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

where the action of R is just the action on each component, and the action of t is

$$t \cdot (m_0, m_1, m_2, \dots) = (0, \varphi_0(m_0), \varphi_1(m_1), \varphi_2(m_2), \dots).$$

In homology terms, multiplication by t moves a homology class up the filtration by one. If some homology class γ exists in K_i , then $t \cdot \gamma$ is an element of K_{i+1} .

Definition 3.4. A graded module $R[t]$ is of *finite type* if each M_n is finite dimensional over R .

Because our filtration is finite, its homology is also finite and therefore the graded $R[t]$ module will be of finite type.

Theorem 3.5 (Correspondence). *The map α is a one-to-one correspondence between the set of persistence modules over R and the set of graded modules of finite type over $R[t]$.*

Proof. We construct an inverse to α in the obvious way. If G is a graded module of finite type over $R[t]$, let M_i be the submodule of G with grading i considered as a R -module and set $\varphi_i(a) = t \cdot a \in M_{i+1}$. \square

From this we learn there is little hope for a simple classification of persistence modules when the ground ring is not a field. Even in the simple case of $R = \mathbf{Z}$, $\mathbf{Z}[t]$ -modules are known to be extremely difficult to classify.

When R is a field \mathbf{k} , the graded ring $\mathbf{k}[t]$ is a PID, so we can classify the persistence modules using Theorem 2.11, the structure theorem for graded modules over PIDs. The only graded ideals of $\mathbf{k}[t]$ are homogeneous of the form $\langle t^n \rangle$, so

$$\alpha(\mathcal{M}) \cong \left(\bigoplus_{i=1}^n \Sigma^{\alpha_i} \mathbf{k}[t] \right) \oplus \left(\bigoplus_{i=1}^m \Sigma^{\beta_i} \mathbf{k}[t] / \langle t^{n_i} \rangle \right).$$

Remember that Σ^k represents an upward shift of k steps in the grading.

This decomposition holds the description for persistence that we seek. In the first sum, the free part, each summand gives a class γ that first appears at time α_i , and $t \cdot \gamma, t^2 \cdot \gamma, \dots$ remain nonzero for all steps in the filtration after that. This corresponds to a persistence interval of $[\alpha_i, \infty)$.

In the second sum, the torsion part, a class τ first appears at time β_i , and exists as $t \cdot \tau, t^2 \cdot \tau, \dots, t^{n_i-1} \cdot \tau$. After this, t^{n_i} becomes zero, so our homology class is mapped to zero at time $\beta_i + n_i$. This gives us an interval $[\beta_i, \beta_i + n_i)$.

These sets of intervals are known as *barcodes*. More precisely:

Definition 3.6. A *barcode* is a finite multiset of ordered pairs (i, j) , with $0 \leq i < j$, and $i, j \in \mathbb{Z} \cup \{\infty\}$.

The barcode exactly determines the structure of a $\mathbf{k}[t]$ -module, so we can restate the correspondence as follows:

Corollary 3.7. *There is a bijective correspondence between barcodes and persistence modules of finite type over a field.*

Example 3.8. Let us return to the filtration shown in Figure 4 and compute the barcode for the 0-dimensional homology. The space K_1 introduces two new components, so two new homology classes. Both classes are mapped into the same component in K_2 , so one of the classes dies entering K_2 . An additional class is born in K_2 , which then joins the rest of the space in K_3 . The barcode is therefore $\{[1, 2), [1, \infty), [2, 3)\}$.

Similarly, two 1-dimensional homology classes are created in K_3 . In K_4 , one of these becomes trivial. The barcode is then $\{[3, 4), [3, \infty)\}$.

Barcodes can be represented pictorially by simply plotting each interval $[p, q)$ as a point (p, q) in the plane. This set of points must lie above the line $x = y$. Features with short lifetimes will correspond to points that lie close to the diagonal.

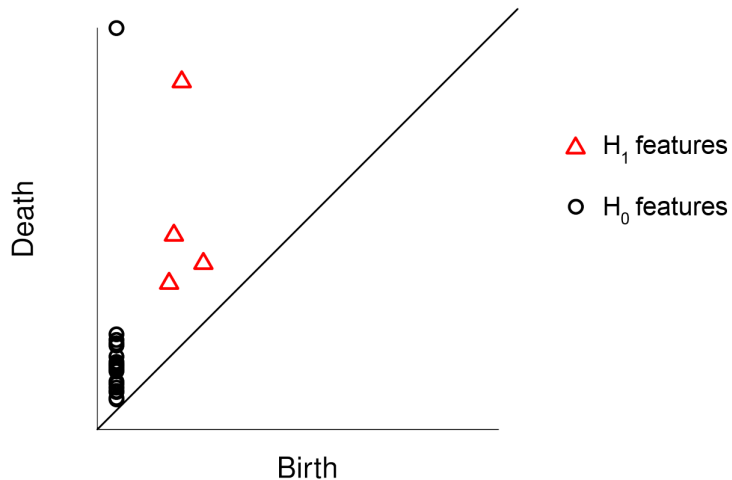


Figure 6: A sample persistence diagram

3.3 Algorithms

Zomorodian and Carlsson also provide an algorithm to systematically calculate the barcode of an input filtration. Appendix A contains a Haskell implementation of the algorithm, completed as part of this project. Haskell was chosen as it is well suited for defining mathematical objects such as simplices, groups, vector spaces, and so on.

The algorithm has since been improved in a number of ways. Most notably, calculating persistent cohomology [12] has been shown to have much better performance in practice. Poincaré duality implies that over $\mathbb{Z}/2\mathbb{Z}$, the barcodes produced this way are the same as would be produced by persistent homology.

4 Persistence Modules

In its first incarnation, persistent homology was an algorithm that given a filtration, directly produced a barcode. The previous section shows there is an intermediate object, the persistence module, that captures all the important algebraic data contained in the filtration. In this section we give a more general definition of persistence modules and derive some of their basic properties.

As before, we fix a field \mathbf{k} with all vector spaces over this field.

Definition 4.1. Let \mathbf{P} be a partially ordered set. A *persistence module* \mathbb{V} over \mathbf{P} is a collection of vector spaces $\{V_t \mid t \in \mathbf{P}\}$ and linear maps $\{v_s^t \mid s, t \in \mathbf{P}, s \leq t\}$ such that for all $r \leq s \leq t$, we have $v_r^t = v_s^t \circ v_r^s$.

These linear maps are often called the *structure maps* of \mathbb{V} . We recover the definition of the previous section when we set $P = \mathbf{N}$, the natural numbers. In what follows, we will mainly be interested in persistence modules over the real numbers \mathbf{R} . If K is a filtration, we will write $H(K)$ for the persistence module induced by K .

Definition 4.2. If \mathbb{V} is a persistence module over \mathbf{T} , and $\mathbf{P} \subset \mathbf{T}$, then the *restriction of \mathbb{V} to \mathbf{P}* , written $\mathbb{V}_{\mathbf{P}}$, is the persistence module obtained by forgetting the spaces and maps that lie outside \mathbf{P} . To be specific, $\mathbb{V}_{\mathbf{P}}$ is the persistence module with spaces and structure maps

$$\begin{aligned} (V_{\mathbf{P}})_t &= V_t \\ (v_{\mathbf{P}})_s^t &= v_s^t \end{aligned}$$

for $s, t \in \mathbf{P}$.

As with any algebraic object, we also require a notion of a mapping between persistence modules.

Definition 4.3. A *homomorphism* $\Phi : \mathbb{U} \rightarrow \mathbb{V}$ between two persistence modules over \mathbf{P} is a collection of linear maps $\{\phi_t : U_t \rightarrow V_t \mid t \in \mathbf{P}\}$ such that the following diagram commutes for all $s \leq t$.

$$\begin{array}{ccc} U_s & \xrightarrow{u_s^t} & U_t \\ \phi_s \downarrow & & \downarrow \phi_t \\ V_s & \xrightarrow{v_s^t} & V_t \end{array}$$

We compose morphisms in the obvious way. Suppose we have $\Phi : \mathbb{U} \rightarrow \mathbb{V}$ and $\Psi : \mathbb{V} \rightarrow \mathbb{W}$, where $\Phi = \{\phi_t\}$ and $\Psi = \{\psi_t\}$. Then composition $\Psi\Phi$ is given

by $\Psi\Phi = \{\psi_t\phi_t\}$. To see that the relevant diagram commutes, notice that

$$\begin{array}{ccc} U_s & \xrightarrow{u_s^t} & U_t \\ \phi_s \downarrow & & \downarrow \phi_t \\ V_s & \xrightarrow{v_s^t} & V_t \\ \psi_s \downarrow & & \downarrow \psi_t \\ W_s & \xrightarrow{w_s^t} & W_t \end{array}$$

commutes, as both the top and bottom squares do.

The *identity* and *zero* homomorphisms are just the identity and zero maps at each point. The *image*, *kernel* and *cokernel* of a homomorphism are also defined pointwise with the canonical structure maps. For example, the kernel of Φ is given by:

$$\begin{aligned} \text{Ker}(\Phi)_t &= \ker(\phi_t) \\ \ker(\Phi)_s^t &= u_s^t|_{\ker(\phi_t)}. \end{aligned}$$

For convenience, we also define $\text{Hom}(\mathbb{U}, \mathbb{V}) = \{\text{homomorphisms } \mathbb{U} \rightarrow \mathbb{V}\}$ and $\text{End}(\mathbb{V}) = \text{Hom}(\mathbb{V}, \mathbb{V})$.

4.1 Interval modules

The simplest examples of persistence modules are the interval modules. These will correspond to points in our barcode.

Definition 4.4. Let $J \subseteq \mathbf{R}$ be an interval. Then \mathbb{I}^J is defined to be the persistence module with spaces

$$I_t = \begin{cases} \mathbf{k} & \text{if } t \in J; \\ 0 & \text{otherwise} \end{cases}$$

and maps

$$i_s^t = \begin{cases} 1 & \text{if } t, s \in J; \\ 0 & \text{otherwise.} \end{cases}$$

Over \mathbf{R} , providing the endpoints is not enough to determine the interval as the interval could be either closed or open on each end. To simplify the notation, we describe the endpoints of intervals using *decorated real numbers*: real numbers with a superscript $+$ or $-$ annotation. For finite intervals, these decorated numbers are interpreted as follows:

$$\begin{aligned} (p^-, q^-) &= [p, q); \\ (p^-, q^+) &= [p, q]; \\ (p^+, q^-) &= (p, q); \\ (p^+, q^+) &= (p, q]. \end{aligned}$$

We also use the natural ordering

$$p^- < p < p^+ < q^- < q < q^+$$

whenever $p < q$. If the annotation is unknown or unimportant, an asterisk is used. The expression (p^*, q^*) with $p^* < q^*$ thus represents an arbitrary interval.

4.2 Interval decompositions

We would like to be able to analyse persistence modules by decomposing them into interval modules. There is a natural definition for the sum of two persistence modules.

Definition 4.5. The *direct sum* $\mathbb{W} = \mathbb{U} \oplus \mathbb{V}$ of two persistence modules is the persistence module with spaces and structure maps:

$$\begin{aligned} W_t &= U_t \oplus V_t \\ w_s^t &= u_s^t \oplus v_s^t. \end{aligned}$$

This definition generalises in the obvious way to direct sums with arbitrary summands.

We say that a persistence module \mathbb{W} is *indecomposable* if the only decompositions of the form $\mathbb{W} = \mathbb{U} \oplus \mathbb{V}$ are $\mathbb{W} \oplus 0$ and $0 \oplus \mathbb{W}$.

Proposition 4.6. *Interval modules are indecomposable.*

Proof. Let \mathbb{I} be an interval module with decomposition $\mathbb{I} = \mathbb{U} \oplus \mathbb{V}$. Then the projection maps onto \mathbb{U} and \mathbb{V} are idempotent elements of $\text{End}(\mathbb{I})$. Any endomorphism of \mathbb{I} acts on each I_t by scalar multiplication and, as the structure maps are commutative, the scalar must be the same for each t . Therefore, $\text{End}(\mathbb{I}) = \mathbf{k}$, and the only idempotents are 0 and 1. Our projection maps are therefore either 0 or 1, so we must have one map 0 and the other map 1 as required. \square

If our persistence module is decomposable into interval modules, the following theorem guarantees that the decomposition is unique up to reordering. The theorem is the analogue of the Krull-Schmidt theorem [2] for group decompositions, which has since been generalised to many algebraic objects for which ‘decomposition’ makes sense.

Theorem 4.7 (Krull-Schmidt-Azumaya). *Suppose a persistence module \mathbb{V} is decomposable into interval modules in two different ways:*

$$\mathbb{V} = \bigoplus_{l \in L} \mathbb{I}^{J_l} = \bigoplus_{m \in M} \mathbb{I}^{K_m}.$$

Then $|L| = |M|$, and there exists a permutation π such that $J_l = K_{\pi(l)}$ for all $l \in L$.

Proof. Here we specialise an argument given by Barot [3] for decompositions with finitely many summands. Let $\Phi : \mathbb{V} \rightarrow \mathbb{W}$ be an isomorphism, and let $\mathbb{V} = \mathbb{V}_1 \oplus \cdots \oplus \mathbb{V}_n$ and $\mathbb{W} = \mathbb{W}_1 \oplus \cdots \oplus \mathbb{W}_m$ be interval decompositions.

We can write Φ as a matrix $\Phi = (\Phi_{ji})$ where each Φ_{ji} is a homomorphism $\mathbb{V}_i \rightarrow \mathbb{W}_j$. Similarly, we can write $\Phi^{-1} = \Psi$ as a matrix (Ψ_{ji}) . Consider:

$$\mathbf{1}_{\mathbb{V}_1} = (\Psi\Phi)_{11} = \sum_{l=1}^j \Psi_{1l}\Phi_{l1}.$$

As the left-hand side is nonzero, one of the summands $\Psi_{1l}\Phi_{l1}$ must be nonzero. Without loss of generality, assume $\Psi_{11}\Phi_{11}$ is nonzero and therefore invertible. Because \mathbb{V}_1 and \mathbb{W}_1 are indecomposable, Ψ_{11} and Φ_{11} are onto and hence are themselves invertible.

Now we construct two change of basis matrices:

$$\alpha = \begin{pmatrix} \mathbf{1}_{\mathbb{W}_1} & 0 & \cdots & 0 \\ -\Phi_{21}\Phi_{11}^{-1} & \mathbf{1}_{\mathbb{W}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\Phi_{m1}\Phi_{11}^{-1} & 0 & \cdots & \mathbf{1}_{\mathbb{W}_m} \end{pmatrix} \quad \beta = \begin{pmatrix} \mathbf{1}_{\mathbb{V}_1} & -\Phi_{11}^{-1}\Phi_{12} & \cdots & -\Phi_{11}^{-1}\Phi_{1n} \\ 0 & \mathbf{1}_{\mathbb{V}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1}_{\mathbb{V}_m} \end{pmatrix}$$

and set $\Phi' = \alpha\Phi\beta$. Then $\Phi' : \mathbb{V} \rightarrow \mathbb{W}$ is another isomorphism of the form $\Phi' = \begin{pmatrix} \Phi_{11} & 0 \\ 0 & \Lambda \end{pmatrix}$ with $\Lambda : \bigoplus_{i=2}^n \mathbb{V}_i \rightarrow \bigoplus_{j=2}^m \mathbb{W}_j$ an isomorphism of the remaining summands. The result follows by induction.

The infinite case is much harder, and we refer the reader to Azumaya [2]. To apply this theorem, the indecomposable modules must satisfy a certain condition: that $\text{End}(\mathbb{I})$ be a ‘local’ ring. As $\text{End}(\mathbb{I}) = \mathbf{k}$ is a field and all fields are local, this condition is satisfied. In the finite case above, this property is required to conclude that one of $\Psi_{1l}\Phi_{l1}$ is invertible. \square

We saw in the previous section that if our indexing set is a finite subset of \mathbf{R} and the vector spaces are pointwise finite dimensional then the persistence module decomposes into interval modules. A number of researchers have independently shown that we can drop the restriction on the dimension of the vector spaces:

Theorem 4.8 (Auslander [1]). *Let \mathbf{T} be a finite, totally ordered set. Then any persistence module over \mathbf{T} decomposes as a sum of interval modules.*

The theorem is false for arbitrary persistence modules. The following module provides a counterexample.

Example 4.9 (Webb [9]). Let \mathbb{V} be the persistence module over the non-positive integers with spaces

$$\begin{aligned} V_0 &= \{\text{sequences of real numbers } (x_1, x_2, \dots)\} \\ V_{-n} &= \{\text{sequences with } x_1 = \cdots = x_n = 0\} \text{ for } n \geq 1 \end{aligned}$$

and structure maps the natural inclusions v_{-m}^{-n} for $m \geq n$.

First note that $\dim(V_0)$ is uncountable. Any vector in a vector space must be able to be written as a linear combination of finitely many basis elements. It can be proven that no countable basis exists for sequences of real numbers.

Now, because the structure maps are all injective, the intervals in an interval decomposition must all be of the form $[-n, 0]$. Since $\dim(V_{-n}/V_{-n-1}) = 1$, each $[-n, 0]$ occurs with multiplicity 1. This implies that $\mathbb{V} \cong \bigoplus_{n \geq 0} \mathbb{I}[-n, 0]$, which contradicts the fact that $\dim(V_0)$ is uncountable.

4.3 Quiver calculations

In this section, we introduce some notation that will simplify future calculations. We then prove some basic properties of persistence modules over finite sets using this tool.

A persistence module \mathbb{V} over a finite set $a_1 < a_2 < \dots < a_n$ can be thought of as a diagram of vector spaces:

$$V_{a_1} \rightarrow V_{a_2} \rightarrow \dots \rightarrow V_{a_n}$$

Diagrams of this sort are also known as ‘quiver representations of type A_n ’. Theorem 4.8 states that such a persistence module decomposes as a sum of interval modules. Because we are over a finite set, we can represent such interval modules pictorially. We use filled circles to represent indices where the module is \mathbf{k} and open circles for where the module is 0.

Consider the finite set $a < b < c$. There are precisely six interval modules over $\{a, b, c\}$:

$$\begin{aligned} \mathbb{I}[a, a] &= \bullet_a - \circ_b - \circ_c & \mathbb{I}[a, b] &= \bullet_a - \bullet_b - \circ_c & \mathbb{I}[a, c] &= \bullet_a - \bullet_b - \bullet_c \\ \mathbb{I}[b, b] &= \circ_a - \bullet_b - \circ_c & \mathbb{I}[b, c] &= \circ_a - \bullet_b - \bullet_c \\ \mathbb{I}[c, c] &= \circ_a - \circ_b - \bullet_c \end{aligned}$$

Definition 4.10. Let \mathbb{V} be a persistence module over \mathbf{R} and $\mathbf{T} \subset \mathbf{R}$ a finite subset. Given an arbitrary interval module

$$\mathbb{I} = \circ_r - \dots - \circ - \bullet_a - \dots - \bullet_b - \circ - \dots - \circ_s$$

on \mathbf{T} , we write

$$\langle \mathbb{I} \mid \mathbb{V}_{\mathbf{T}} \rangle \text{ or } \langle \circ_r - \dots - \circ - \bullet_a - \dots - \bullet_b - \circ - \dots - \circ_s \mid \mathbb{V} \rangle \text{ or } \langle [a, b] \mid \mathbb{V}_{\mathbf{T}} \rangle$$

for the multiplicity of \mathbb{I} in the interval decomposition of $\mathbb{V}_{\mathbf{T}}$. We will omit $\mathbb{V}_{\mathbf{T}}$ when the persistence module and indexing sets are clear.

Example 4.11. Consider one of the structure maps $v_s^t : V_s \rightarrow V_t$. Then:

$$\langle \bullet_s - \bullet_t \mid \mathbb{V} \rangle = \text{rank}(v_s^t).$$

If our persistence module is the direct sum of simpler modules, the multiplicity of an interval is simply the sum of the multiplicities in these simpler modules.

Proposition 4.12. *Let \mathbb{V} be written as a direct sum*

$$\mathbb{V} = \bigoplus_{i \in I} \mathbb{V}^i.$$

Then

$$\langle [a, b] \mid \mathbb{V}_{\mathbf{T}} \rangle = \sum_{i \in I} \langle [a, b] \mid \mathbb{V}_{\mathbf{T}}^i \rangle$$

for any interval module $[a, b]$ in a finite indexing set \mathbf{T} .

Proof. Individually decompose the $\mathbb{V}_{\mathbf{T}}^i$ into interval modules as in Theorem 4.8. The sum of all these interval modules is a decomposition of $\mathbb{V}_{\mathbf{T}}$. The multiplicity of $[a, b]$ in $\mathbb{V}_{\mathbf{T}}$ is then the sum of the multiplicities in $\mathbb{V}_{\mathbf{T}}^i$. \square

The final property of the multiplicities we discuss here is the so called ‘restriction principle’. This allows us to relate the multiplicities of different finite restrictions of the same module.

Proposition 4.13. *Let \mathbf{S} and \mathbf{T} be finite index sets with $\mathbf{S} \subset \mathbf{T}$. For any interval module \mathbb{I} over \mathbf{S} ,*

$$\langle \mathbb{I} \mid \mathbb{V}_{\mathbf{S}} \rangle = \sum_{\mathbb{J}} \langle \mathbb{J} \mid \mathbb{V}_{\mathbf{T}} \rangle$$

where \mathbb{J} runs over all interval modules in \mathbf{T} that restrict to \mathbb{I} over \mathbf{S} .

Example 4.14. Consider the finite indexing set $a < p < b < q < c$. Then:

$$\langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \rangle = \langle \circ_a \text{---} \circ_p \text{---} \bullet_b \text{---} \bullet_q \text{---} \bullet_c \rangle + \langle \circ_a \text{---} \bullet_p \text{---} \bullet_b \text{---} \bullet_q \text{---} \bullet_c \rangle$$

Proof. Consider an interval decomposition of $\mathbb{V}_{\mathbf{T}}$. This restricts to an interval decomposition of $\mathbb{V}_{\mathbf{S}}$. Instances of \mathbb{I} in this decomposition appear precisely when the interval \mathbb{J} restricts to \mathbb{I} over \mathbf{S} . \square

5 Persistence Diagrams

In this section we define the diagram of a persistence module over \mathbf{R} , generalising the barcode of standard persistent homology. If a persistence module \mathbb{V} is decomposable into interval modules, we can use the obvious definition: the multiset in \mathbf{R}^2 with a point at (p, q) for each interval $\mathbb{I}(p^*, q^*)$. When \mathbb{V} does *not* decompose into interval modules, we will have to work a little harder.

The strategy is as follows. To every persistence module over \mathbf{R} we associate an ‘r-measure’: a function that assigns every rectangle in \mathbf{R}^2 an integer measure. We then show there is a bijection between the appropriate classes of r-measures and multisets in \mathbf{R}^2 .

5.1 Abstract r-measures

First we define r-measures and derive some of their basic properties.

Definition 5.1. Let \mathcal{D} be a subset of \mathbf{R}^2 . We define $\text{Rect}(\mathcal{D})$ to be the set of closed rectangles lying inside \mathcal{D} , i.e.,

$$\text{Rect}(\mathcal{D}) = \{[a, b] \times [c, d] \subset \mathcal{D} \mid a < b \text{ and } c < d\}$$

An *r-measure* on \mathcal{D} is a function $\mu : \text{Rect}(\mathcal{D}) \rightarrow \{0, 1, 2, \dots\} \cup \{\infty\}$ that is additive under horizontal and vertical splits:

$$\begin{aligned} \mu([a, b] \times [c, d]) &= \mu([a, p] \times [c, d]) + \mu([p, b] \times [c, d]) \\ \mu([a, b] \times [c, d]) &= \mu([a, b] \times [c, q]) + \mu([a, b] \times [q, d]) \end{aligned}$$

for $a < p < b$ and $c < q < d$. Figure 7 shows how to interpret these equations pictorially.

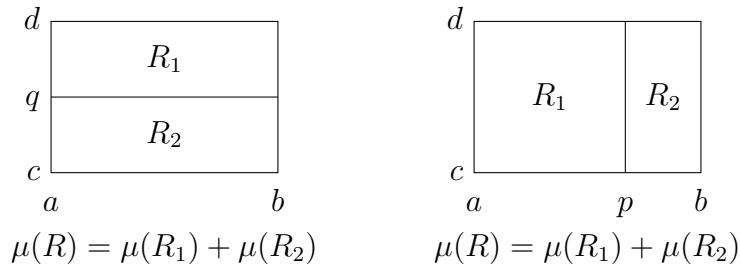


Figure 7: Horizontal and vertical splitting for r-measures

We will see that r-measures have many of the features of ordinary measures when restricted appropriately to rectangles in the plane.

Proposition 5.2 (Finite additivity). *Let μ be an r-measure on \mathcal{D} . If a rectangle R can be written as a union $R = R_1 \cup \dots \cup R_k$ of rectangles with disjoint interiors, then $\mu(R) = \mu(R_1) + \dots + \mu(R_k)$.*

Proof. Let $R = [a, b] \times [c, d]$. By the vertical splitting property, the statement holds for any decomposition $R_i = [a_i, a_{i+1}] \times [c, d]$ with $a = a_1 < a_2 < \dots < a_n = b$. Similarly, by the horizontal splitting property, the statement holds for $R_i = [a, b] \times [c_j, c_{j+1}]$ with $c = c_1 < c_2 < \dots < c_m = d$.

By performing both types of splits, it follows that finite additivity holds for any ‘product’ decomposition of the form $R_{ij} = [a_i, a_{i+1}] \times [c_j, c_{j+1}]$. Now for an arbitrary decomposition $R = R_1 \cup \dots \cup R_k$, decompose each component R_i into the above form, so the entire decomposition also has the form of a ‘product’. The result follows. \square

Proposition 5.3 (Monotonicity). *If $R \subseteq S$ then $\mu(R) \leq \mu(S)$.*

Proof. This follows from the above proposition, as S must have a decomposition containing R as a component. This can always be done with at most nine rectangles using the coordinates of R as the locations of the splits, as suggested by Figure 8. Then:

$$\begin{aligned} \mu(S) &= \mu(R) + \mu(R_1) + \dots \\ &\geq \mu(R). \end{aligned}$$

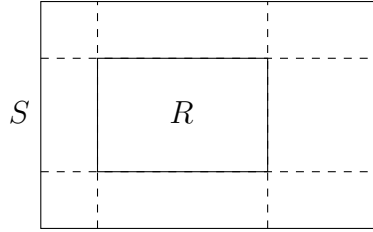


Figure 8: Splitting used in the proof of monotonicity

\square

Proposition 5.4 (Subadditivity). *If $R \subseteq R_1 \cup \dots \cup R_n$ then $\mu(R) \leq \mu(R_1) + \dots + \mu(R_n)$.*

Proof. Let $\{a_i\}$ be the x -coordinates of the corners of every rectangle, and $\{c_j\}$ the y -coordinates. Subdivide every rectangle into tiles of the form $[a_i, a_{i+1}] \times [c_j, c_{j+1}]$. Every tile that comprises R must belong to some R_i and the result follows by additivity. \square

5.2 The persistence measure

From any persistence module \mathbb{V} , we construct an r-measure that ‘counts’ the points of the barcode of \mathbb{V} that fall inside each rectangle. We can use this r-measure to construct a persistence diagram for *any* persistence module, not only those that decompose into intervals.

Let \mathcal{H} be the half-plane in \mathbf{R}^2 given by:

$$\mathcal{H} = \{(p, q) \mid p \leq q\}$$

Definition 5.5. The *persistence measure* of \mathbb{V} is the function

$$\mu_{\mathbb{V}}(R) = \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle,$$

for any $R = [a, b] \times [c, d]$ with $a < b \leq c < d$.

Proposition 5.6. $\mu_{\mathbb{V}}$ is a r -measure on \mathcal{H} .

Proof. We must show that $\mu_{\mathbb{V}}$ is additive under horizontal and vertical splitting. For horizontal splitting, let $a < p < b \leq c < d$. Then, by the restriction principle,

$$\begin{aligned} \mu_{\mathbb{V}}([a, b] \times [c, d]) &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \langle \circ_a \text{---} \bullet_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle + \langle \circ_a \text{---} \circ_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \langle \circ_a \text{---} \bullet_p \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle + \langle \text{---} \circ_p \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \mu_{\mathbb{V}}([a, p] \times [c, d]) + \mu_{\mathbb{V}}([p, b] \times [c, d]). \end{aligned}$$

For vertical splitting, let $a < b \leq c < q < d$, then:

$$\begin{aligned} \mu_{\mathbb{V}}([a, b] \times [c, d]) &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_q \text{---} \circ_d \mid \mathbb{V} \rangle + \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \bullet_q \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \langle \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_q \text{---} \mid \mathbb{V} \rangle + \langle \circ_a \text{---} \bullet_b \text{---} \bullet_q \text{---} \circ_d \mid \mathbb{V} \rangle \\ &= \mu_{\mathbb{V}}([a, b] \times [c, q]) + \mu_{\mathbb{V}}([a, b] \times [q, d]). \end{aligned}$$

□

Let us consider the simplest case: when \mathbb{V} is an interval module \mathbb{I}^J for some $J = (p^*, q^*)$. Let $R = [a, b] \times [c, d]$ be a rectangle, then $\mu_{\mathbb{V}}(R)$ is determined as follows. The restriction of \mathbb{I}^J to the indexing set $\mathbf{T} = \{a, b, c, d\}$ is clearly either an interval or zero. Therefore, $\mu_{\mathbb{V}}(R)$ is equal to either 0 or 1. $\mu_{\mathbb{V}}(R) = 1$ precisely when $\mathbb{I}_{\mathbf{T}}^J = \circ_a \text{---} \bullet_b \text{---} \bullet_c \text{---} \circ_d$, i.e., $[b, c] \subseteq J \subseteq (a, d)$.

Using the notation for extended real numbers, this condition is equivalent to $a^+ \leq p^* \leq b^- \leq c^+ \leq q^* \leq d^-$. This relationship is easiest to explain pictorially. We represent an interval (p^*, q^*) as the point $(p, q) \in \mathbf{R}^2$, with a tick to indicate the annotations. For example, the point (p^+, q^+) has a tick towards the upper right. Figure 9 demonstrates this representation.

The intervals (p^*, q^*) that are ‘detected’ by the rectangle R are precisely the intervals whose point (p, q) and tick mark are contained in R . We abuse notation and say $(p^*, q^*) \in R$ when the interval $J = (p^*, q^*)$ satisfies $[b, c] \subseteq J \subseteq (a, d)$ as above. The points shown in Figure 9 are examples of points ‘contained’ in a rectangle R .

When \mathbb{V} is decomposable, $\mu_{\mathbb{V}}$ is easy to describe.

Proposition 5.7. Let $V = \bigoplus \mathbb{I}^{l_i}$ be a decomposable persistence module with intervals $\{l_i\}$. Then

$$\mu_{\mathbb{V}}(R) = |\{l_i \mid l_i \in R\}|$$

Proof. This follows from the discussion above and Proposition 4.12. □

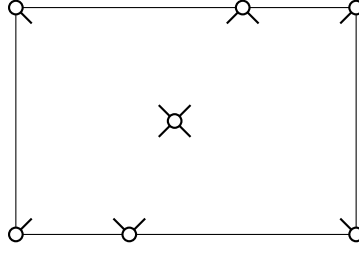


Figure 9: Some decorated points ‘detected’ by R

5.3 Equivalence

We now have a mechanism to pass from persistence modules to r-measures. In this section we show that every finite r-measure uniquely determines a multiset of decorated points.

Definition 5.8. The r -interior of a region $\mathcal{D} \subseteq \mathbf{R}^2$ is the set of decorated points:

$$\mathcal{D}^\blacksquare = \{(p^*, q^*) \mid \text{there exists } R \in \text{Rect}(\mathcal{D}) \text{ with } (p^*, q^*) \in R\}.$$

The r-interior of a region is all the points that are ‘detected’ by an r-measure on that region. For example, the upper half plane $\mathcal{H} = \{(p, q) \mid p \leq q\}$ has r-interior:

$$\mathcal{H}^\blacksquare = \{(p^*, q^*) \mid p < q\} \cup \{(p^-, p^+) \mid p \in \mathbf{R}\}.$$

Theorem 5.9 (Equivalence). *Let $\mathcal{D} \subseteq \mathbf{R}^2$. There is a bijection between*

- *r-measures on \mathcal{D} such that $\mu(R) < \infty$ for every $R \in \text{Rect}(\mathcal{D})$; and*
- *multisets of decorated points \mathbf{A} in \mathcal{D}^\blacksquare such that $\text{card}(\mathbf{A}|_R) < \infty$ for every $R \in \text{Rect}(\mathcal{D})$,*

where the measure corresponding to a multiset \mathbf{A} is given by the formula

$$\mu(R) = \text{card}(\mathbf{A}|_R).$$

The proof of this theorem occupies the remainder of this section.

Proof. One direction of the bijection is easy. Given a multiset \mathbf{A} , we define μ as above. We need only verify that the formula above defines an r-measure. Given any horizontal or vertical split $R = R_1 \cup R_2$, any decorated point $(p^*, q^*) \in R$ belongs to exactly one of R_1 and R_2 . Therefore:

$$\mu(R) = \text{card}(\mathbf{A}|_R) = \text{card}(\mathbf{A}|_{R_1}) + \text{card}(\mathbf{A}|_{R_2}) = \mu(R_1) + \mu(R_2).$$

Now for the reverse direction. For an r-measure μ , let \mathbf{A} be the multiset in \mathcal{D}^\blacksquare given by the multiplicity function:

$$m(p^*, q^*) = \min\{\mu(R) \mid R \in \text{Rect}(\mathcal{D}), (p^*, q^*) \in R\}.$$

This is well defined by the well ordering principle, as each $\mu(R)$ is a non-negative integer.

We must prove that \mathbf{A} satisfies the formula in the statement of the theorem and show that it is the unique such multiset. From above, we know that the multiset \mathbf{A} defines an r -measure ν via:

$$\nu(R) = \text{card}(\mathbf{A}|_R).$$

We now show that $\mu = \nu$ by induction on $\mu(R)$, which is finite by the conditions of the theorem.

In the base case, consider all R such that $\mu(R) = 0$. Let (p^*, q^*) be a point contained in one such R . Then:

$$0 \leq m(p^*, q^*) \leq \mu(R) = 0.$$

Therefore $m(p^*, q^*) = 0$ inside R , and $\nu(R) = \mu(R) = 0$.

Now assume that $\mu(R) = \nu(R)$ for every R with $\mu(R) < k$. We must now show that for any rectangle R_0 with $\mu(R_0) = k$, also $\nu(R_0) = k$. Split R_0 into four equal quadrants Q_1, Q_2, Q_3, Q_4 . By finite additivity, we have:

$$\begin{aligned}\mu(R) &= \mu(Q_1) + \mu(Q_2) + \mu(Q_3) + \mu(Q_4) \\ \nu(R) &= \nu(Q_1) + \nu(Q_2) + \nu(Q_3) + \nu(Q_4).\end{aligned}$$

If more than one of $\mu(Q_i)$ is non-zero, then $\mu(Q_i) < k$ for all i , and the result follows by induction. In the remaining case, there is a single quadrant with measure k and the remainder have measure 0. Let R_1 be this distinguished quadrant.

We now repeat the argument with R_1 . Either we find the quadrants satisfy $\mu(Q_i) < k$ or there is another quadrant with $\mu(Q_i) = k$. The only unhandled case is when we find an infinite sequence of rectangles $R_0 \supset R_1 \supset \dots$ with $\mu(R_i) = k$ for all i . The diameters of the quadrants tend to zero, so the intersection $\bigcap R_i$ contains a single point (r, s) .

Therefore, the only contribution to $\nu(R)$ must come from decorated points that belong to all the R_i , i.e., decorated points of the form (r^*, s^*) . Assume for now that (r, s) lies on the interior of every R_i . Divide each R_i into quadrants $R_i^{++}, R_i^{-+}, R_i^{+-}, R_i^{--}$ sharing (r, s) as a common corner, and such that (r^+, s^+) lies in R_i^{++} for all i , (r^-, s^+) lies in R_i^{-+} , and so on.

The result follows from the following claim: Let (ε_i) and (η_i) be non-increasing sequences of positive real numbers that converge to 0. Then:

$$m(r^+, s^+) = \lim_{i \rightarrow \infty} \mu([r, r + \varepsilon_i] \times [s, s + \eta_i]).$$

To see that this is true, let R be a rectangle that contains (r^+, s^+) . Then for i sufficiently large, $R_i \subseteq R$. By monotonicity, $\mu(R_i)$ must eventually stabilise to a finite limit. Then:

$$m(r^+, s^+) \leq \min \mu(R_i) = \lim_{i \rightarrow \infty} \mu(R_i) \leq \mu(R).$$

Now, taking the minimum over all R , the RHS becomes $m(r^+, s^+)$ and we have the equality we require.

Returning to the proof of the theorem, we have $m(r^+, s^+) = \lim_{i \rightarrow \infty} \mu(R_i^{++})$, with the sequence eventually stabilising at this value. We also have the obvious equivalent statements for the remaining decorated points (r^*, s^*) .

For sufficiently large i , we therefore have:

$$\begin{aligned} \nu(R_0) &= m(r^+, s^+) + m(r^+, s^-) + m(r^-, s^+) + m(r^-, s^-) \\ &= \mu(R_i^{++}) + \mu(R_i^{-+}) + \mu(R_i^{+-}) + \mu(R_i^{--}) = \mu(R_i) = k. \end{aligned}$$

If (r, s) does not lie in the interior of every R_i then beyond some N , the point (r, s) lies on the boundary of R_i for all $i > N$. We then have fewer cases to consider: if (r, s) is on an edge then only two decorated points of the form (r^*, s^*) lie in every R_i . If (r, s) is on a corner then only one decorated point needs to be considered. In either case, the above argument goes through.

By induction, we have $\mu(R) = \nu(R)$ for every rectangle R . All that remains is to show uniqueness. Let \mathbf{B} be some other multiset with associated r-measure ν' such that $\mu = \nu'$. We must show that $\mathbf{A} = \mathbf{B}$.

Let $m'(p^*, q^*)$ be the multiplicity function for \mathbf{B} , and consider any point $(p^*, q^*) \in \mathcal{D}^\blacksquare$. Let R be a rectangle with (p^*, q^*) at a corner. Because $\nu(R) = \nu'(R) = \mu(R) < \infty$, we can choose a rectangle such that (p^*, q^*) is the only point with positive multiplicity in R . Then:

$$m(p^*, q^*) = \nu(R) = \mu(R) = \nu'(R) = m'(p^*, q^*).$$

Therefore $m = m'$ for all $(p^*, q^*) \in \mathcal{D}^\blacksquare$ and $\mathbf{A} = \mathbf{B}$. □

The theorem leads to the following definitions.

Definition 5.10. The *decorated diagram* $\mathbf{Dgm}(\mathbb{V})$ of a persistence module \mathbb{V} is the unique multiset of decorated points such that:

$$\mu_{\mathbb{V}}(R) = \text{card}(\mathbf{Dgm}(\mathbb{V})|_R)$$

If we forget the topology of the ‘intervals’ we have the *undecorated diagram* $\mathbf{dgm}(\mathbb{V})$. This is a multiset of ordinary points in \mathcal{D} defined by:

$$\mathbf{dgm}(\mathbb{V}) = \{(p, q) \mid (p^*, q^*) \in \mathbf{Dgm}(\mathbb{V})\} \cap \text{int } \mathcal{D}$$

where ‘int \mathcal{D} ’ denotes the ordinary interior of \mathcal{D} .

Proposition 5.7 implies that for any persistence module \mathbb{V} that decomposes into interval modules, $\mathbf{Dgm}(\mathbb{V})$ is precisely the set of intervals in the decomposition.

5.4 q-tame persistence modules

The stability theorem of later sections requires a condition on the structure maps of the persistence modules. This condition is quite natural and is satisfied in practical applications, as demonstrated by the theorems of this section.

Definition 5.11. A persistence module is *q-tame* if for any $s < t$, the structure map v_s^t has finite rank.

Such modules are called ‘q-tame’ because, for any infinite quadrant $[-\infty, b] \times [c, +\infty]$ we have:

$$\mu_{\mathbb{V}}([-\infty, b] \times [c, +\infty]) = \langle \bullet_b \text{---} \bullet_c \mid \mathbb{V} \rangle = \text{rank } v_b^c < \infty.$$

Note that q-tameness does not mean there are no limit points in the persistence diagram, as we may have limit points on the diagonal $\Delta = \{(x, x) \mid x \in \mathbf{R}\}$. By limit point, we mean a point such that every neighbourhood of that point contains infinitely many other points.

Example 5.12. The module

$$\mathbb{V} = \bigoplus_{n=1}^{\infty} \mathbb{I}[0, \tfrac{1}{n}]$$

is q-tame with $\text{dgm}(\mathbb{V}) = \{(0, \frac{1}{n}) \mid n \geq 1\}$. This has an limit point at $(0, 0)$.

We now give some examples of how q-tame modules occur in the wild. Our original motivation for studying persistence modules was to understand persistent homology in the setting of finite simplicial complexes. The results in the remainder of this work apply as a result of the following theorem.

Theorem 5.13. *Let X be the geometric realisation of a finite simplicial complex, and $f : X \rightarrow \mathbb{R}$ a continuous function. Then the persistence module $H(X^f)$ induced by the sublevelset filtration of X is q-tame.*

Proof. Let X^a denote the sublevel set $X^a = f^{-1}((-\infty, a])$. We must show that, for any $a < b$, $H(X^a) \rightarrow H(X^b)$ has finite rank.

Begin with any triangulation of X and repeatedly subdivide the simplices of X until no simplex intersects both $A = f^{-1}(a)$ and $B = f^{-1}(b)$. To see that this is always possible, note that because X is compact and both A and B are closed, A and B are themselves compact. Because A and B are compact subsets of X , the function $d(x, B)$ attains a minimum on A . This minimum ε must be nonzero, as A and B are disjoint. We therefore subdivide the simplices until the length of every edge is less than $\varepsilon/2$.

Let Y be the union of every closed simplex that meets X^a . Then Y is a finite simplicial complex, and we have inclusions $X^a \subseteq Y \subseteq X^b$.

The map $H(X^a) \rightarrow H(X^b)$ therefore factorises as $H(X^a) \rightarrow H(Y) \rightarrow H(X^b)$. Because Y is a finite simplicial complex, $H(Y)$ is finite dimensional and $H(X^a) \rightarrow H(X^b)$ has finite rank. \square

The Vietoris-Rips complex of a finite set must be a finite simplicial complex, and any simplicial filtration of such a complex leads to a q-tame persistence module. The following theorem shows that a much larger class of metric spaces also lead to a q-tame persistence modules. [7]

Definition 5.14. A metric space is *totally bounded* if, for any $\varepsilon > 0$, the space can be covered with a finite number of open balls of radius ε .

Any compact metric space is totally bounded. For any ε , construct the cover with an open ball of radius ε at every point. By compactness there is a finite subcover which is precisely the cover we require.

Theorem 5.15. *Let (X, d_X) be a totally bounded metric space. Then the persistence module $H(\mathbf{VR}(X, \alpha))$ is q-tame.*

Proof. (Outline) Here we provide a sketch of the proof as the full details are not relevant to the rest of this work. Again we must show that, for any $a < b$, $H(\mathbf{VR}(X, a)) \rightarrow H(\mathbf{VR}(X, b))$ has finite rank. Let $\varepsilon = (b - a)/2$. Because X is totally bounded, there is a finite subset $Y \subset X$ such that for any $x \in X$, there is a $y \in Y$ with $d_X(x, y) < \varepsilon/2$. Let C be some (multivalued) matching between the elements of X and Y so that $d_Y(x, C(x)) \leq \varepsilon/2$.

Let $\sigma \in \mathbf{VR}(X, a)$, and $\tau \subseteq C(\sigma)$ any finite subset. For any $y, y' \in \tau$ there exist $x, x' \in \sigma$ such that $y \in C(x), y' \in C(x')$ and therefore:

$$d_Y(y, y') \leq d_X(x, x') + \varepsilon \leq a + \varepsilon.$$

It follows that $\tau \in \mathbf{VR}(Y, a + \varepsilon)$. Symmetrically, for any $\tau \in \mathbf{VR}(Y, a + \varepsilon)$, $\sigma' \subseteq C(\tau) \in \mathbf{VR}(X, a + 2\varepsilon)$.

After a little work, this implies the factorisation:

$$H(\mathbf{VR}(X, a)) \rightarrow H(\mathbf{VR}(Y, a + \varepsilon)) \rightarrow H(\mathbf{VR}(X, a + 2\varepsilon)) = H(\mathbf{VR}(X, b)).$$

Again, the vector space $H(\mathbf{VR}(Y, a + \varepsilon))$ is finite dimensional so the structure maps all have finite rank. \square

6 Stability

With the necessary foundations all in place, we can now prove the stability theorem. For the persistence diagram of a data set to be a useful invariant, we would hope that small changes in the data lead to small changes in the persistence diagram. The stability theorem shows that this is the case, once we have pinned down the correct definitions for a ‘small change’ in data and a ‘small change’ in persistence diagram.

Theorem 6.1 (Stability). *Let $f, g : X \rightarrow \mathbf{R}$ be functions on a topological space X and let $\mathbb{U} = H(X^f), \mathbb{V} = H(X^g)$ be the sublevel persistence modules. If \mathbb{U} and \mathbb{V} are q -tame, then:*

$$d_b(\mathrm{dgm}(\mathbb{U}), \mathrm{dgm}(\mathbb{V})) \leq \|f - g\|_\infty$$

where d_b denotes the ‘bottleneck distance’ between two diagrams, yet to be defined, and $\|\cdot\|_\infty$ denotes the supremum norm.

In this section we will prove the stability theorem after first proving two preliminary results: the box lemma and interpolation lemma.

The stability theorem was first proven in a much more restricted setting by Cohen-Steiner, Edelsbrunner and Harer [10]. They required that X be a triangulable space and that f and g be ‘tame’ functions, i.e., functions with only finitely many homological critical values. Such restrictions are lifted here, following arguments by Chazal et al. [8] that were later simplified by the same authors [9].

6.1 Interleaving distance

Our first step is to translate the distance $\|f - g\|_\infty$ into an algebraic statement about persistence modules. Let us suppose $\|f - g\|_\infty < \delta$. Then for any $x \in X$ we have $f(x) < g(x) + \delta$ and $g(x) < f(x) + \delta$, which implies the following inclusion of sublevel sets:

$$\begin{aligned} f^{-1}((-\infty, a]) &\subseteq g^{-1}((-\infty, a + \delta]) \subseteq f^{-1}((-\infty, a + 2\delta]) \\ g^{-1}((-\infty, a]) &\subseteq f^{-1}((-\infty, a + \delta]) \subseteq g^{-1}((-\infty, a + 2\delta]). \end{aligned}$$

These induce maps $U_a \rightarrow V_{a+\delta} \rightarrow U_{a+2\delta}$ and $V_a \rightarrow U_{a+\delta} \rightarrow V_{a+2\delta}$ for all $a \in \mathbf{R}$. What we have is *almost* an isomorphism of persistence modules, but where the maps in each direction shift the index by δ .

Definition 6.2. A *homomorphism of degree δ* from \mathbb{U} to \mathbb{V} is a collection of maps $\phi_t : U_t \rightarrow V_{t+\delta}$ such that:

$$\begin{array}{ccc} U_s & \xrightarrow{u_s^t} & U_t \\ \phi_s \downarrow & & \downarrow \phi_t \\ V_{s+\delta} & \xrightarrow{v_{s+\delta}^{t+\delta}} & V_{t+\delta} \end{array}$$

30

commutes for all $s \leq t$. We write $\text{Hom}^\delta(\mathbb{U}, \mathbb{V})$ for the collection of all such homomorphisms.

As an example, the ordinary structure maps $v_t^{t+\delta}$ form a homomorphism of degree δ . We write this homomorphism as $1_{\mathbb{V}}^\delta$.

When we have two mutually ‘inverse’ shift maps as above, we say that \mathbb{U} and \mathbb{V} are δ -interleaved. More precisely:

Definition 6.3. Two persistence modules \mathbb{U} and \mathbb{V} are δ -interleaved if there are maps $\Phi \in \text{Hom}^\delta(\mathbb{U}, \mathbb{V})$ and $\Psi \in \text{Hom}^\delta(\mathbb{V}, \mathbb{U})$ such that

$$\Psi\Phi = 1_{\mathbb{U}}^{2\delta}, \quad \Phi\Psi = 1_{\mathbb{V}}^{2\delta}$$

Expanding the maps involved, these equations state that for every $t \in \mathbf{R}$, the following two diagrams commute:

$$\begin{array}{ccc} U_t & \xrightarrow{\quad} & U_{t+2\delta} \\ & \searrow \quad \nearrow & \\ & V_{t+\delta} & \end{array}, \quad \begin{array}{ccc} & U_{t+\delta} & \\ \nearrow & & \searrow \\ V_t & \xrightarrow{\quad} & V_{t+2\delta} \end{array}$$

Note that a 0-interleaving is precisely an isomorphism.

If two modules \mathbb{U} and \mathbb{V} are δ -interleaved, they are certainly $(\delta+\epsilon)$ -interleaved for any $\epsilon > 0$. We can construct the required maps easily:

$$\begin{aligned} \Phi' &= \Phi 1_{\mathbb{U}}^\epsilon \\ \Psi' &= \Psi 1_{\mathbb{V}}^\epsilon. \end{aligned}$$

Then

$$\Phi'\Psi' = \Phi 1_{\mathbb{U}}^\epsilon \Psi 1_{\mathbb{V}}^\epsilon = 1_{\mathbb{V}}^\epsilon \Phi \Psi 1_{\mathbb{V}}^\epsilon = 1_{\mathbb{V}}^\epsilon 1_{\mathbb{V}}^{2\delta} 1_{\mathbb{V}}^\epsilon = 1_{\mathbb{V}}^{2(\delta+\epsilon)}$$

with the $\Psi'\Phi'$ direction identical.

This suggests that we can measure the distance between two persistence modules by finding the smallest δ for which the modules are δ -interleaved.

Definition 6.4. The *interleaving distance* between two persistence modules is:

$$d_i(\mathbb{U}, \mathbb{V}) = \inf\{\delta \mid \mathbb{U}, \mathbb{V} \text{ are } \delta\text{-interleaved}\}.$$

In the case of sublevel persistence modules, we have the inequality

$$d_i(\mathbb{U}, \mathbb{V}) \leq \|f - g\|_\infty$$

by the discussion above.

The inf in this definition is not necessarily attained. For example, the two interval modules $\mathbb{I}[p, q)$ and $\mathbb{I}[p, q]$ have interleaving distance 0, but the two modules are clearly not isomorphic.

Proposition 6.5. *The interleaving distance satisfies the triangle inequality:*

$$d_i(\mathbb{U}, \mathbb{W}) \leq d_i(\mathbb{U}, \mathbb{V}) + d_i(\mathbb{V}, \mathbb{W})$$

for any three persistence modules $\mathbb{U}, \mathbb{V}, \mathbb{W}$.

Proof. Let (Φ_1, Ψ_1) be a δ_1 -interleaving between \mathbb{U} and \mathbb{V} , and (Φ_2, Ψ_2) a δ_2 -interleaving between \mathbb{V} and \mathbb{W} . The compositions $\Phi = \Phi_2\Phi_1$ and $\Psi = \Psi_1\Psi_2$ yield a $\delta = (\delta_1 + \delta_2)$ -interleaving between \mathbb{U} and \mathbb{W} . To confirm this is a valid interleaving, we calculate:

$$\begin{aligned}\Psi\Phi &= \Psi_1\Psi_2\Phi_2\Phi_1 = \Psi_1 1_{\mathbb{V}}^{2\delta_2}\Phi_1 = \Psi_1\Phi_1 1_{\mathbb{U}}^{2\delta_2} = 1_{\mathbb{U}}^{2\delta_1} 1_{\mathbb{U}}^{2\delta_2} = 1_{\mathbb{U}}^{2\delta} \\ \Phi\Psi &= \Phi_2\Phi_1\Psi_1\Psi_2 = \Phi_2 1_{\mathbb{V}}^{2\delta_1}\Psi_2 = \Phi_2\Psi_2 1_{\mathbb{W}}^{2\delta_1} = 1_{\mathbb{W}}^{2\delta_2} 1_{\mathbb{W}}^{2\delta_1} = 1_{\mathbb{W}}^{2\delta}.\end{aligned}$$

Now, taking the infimum over δ_1 and δ_2 we have our result. \square

6.2 Bottleneck distance

We now define the distance on the other side of the stability theorem, the ‘bottleneck distance’. The intuition here is that two diagrams are close if we can match the points of the two diagrams such that the distance between paired points is small. There is some subtlety when we have points close to the diagonal Δ . We allow such points to be ‘matched’ to the diagonal so that the appearance of small features does not prevent a matching of the remaining points.

Again we work in the open half plane given by:

$$\mathcal{H} = \{(p, q) \mid p < q\}.$$

The distance between points is calculated using the ℓ^∞ -metric. In \mathbf{R}^2 , this distance is

$$d^\infty((p, q), (r, s)) = \max(|p - r|, |q - s|)$$

for two points and

$$d^\infty((p, q), \Delta) = \frac{1}{2}(q - p)$$

for a point and the diagonal. This use of ℓ^∞ is not arbitrary, as the following propositions show.

Proposition 6.6. *Let $\mathbb{U} = \mathbb{I}(p^*, q^*)$ and $\mathbb{V} = \mathbb{I}(r^*, s^*)$ be interval modules. Then:*

$$d_i(\mathbb{U}, \mathbb{V}) \leq d^\infty((p, q), (r, s))$$

Proof. We must show that for any $\delta > d^\infty((p, q), (r, s))$, the modules \mathbb{U} and \mathbb{V} are δ -interleaved. There is an obvious candidate interleaving: set the interleaving maps to 1 when both the domain and codomain are \mathbf{k} and 0 otherwise.

First we show that this definition gives us valid module homomorphisms of degree δ . This is equivalent to verifying that the diagram

$$\begin{array}{ccc} U_t & \longrightarrow & U_{t'} \\ \downarrow & & \downarrow \\ V_{t+\delta} & \longrightarrow & V_{t'+\delta} \end{array}$$

commutes. Because the spaces and maps are all either 1 or 0, the only possible obstruction is some diagram of the form:

$$\begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \downarrow & & \downarrow \\ \circ & \longrightarrow & \bullet \end{array} \quad \text{or} \quad \begin{array}{ccc} \bullet & \longrightarrow & \circ \\ \downarrow & & \downarrow \\ \bullet & \longrightarrow & \bullet \end{array}$$

In the first case, we must have $p^* < t$ and $t + \delta < r^*$, but $\delta > r - p$ so this is impossible. In the second case, we have $q^* < t'$ and $t' + \delta < s^*$, but $\delta > s - q$. Therefore, $\Phi : \mathbb{U} \rightarrow \mathbb{V}$ is a module homomorphism of degree δ . By symmetry, $\Psi : \mathbb{V} \rightarrow \mathbb{U}$ is also a valid homomorphism.

Finally, we must show that Φ and Ψ satisfy $\Psi\Phi = 1_{\mathbb{U}}^{2\delta}$ and $\Phi\Psi = 1_{\mathbb{V}}^{2\delta}$. For the first identity, we have the following diagram.

$$\begin{array}{ccc} U_t & \longrightarrow & U_{t+2\delta} \\ & \searrow & \nearrow \\ & V_{t+\delta} & \end{array}$$

This can only fail to commute in the following case:

$$\begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ & \searrow & \nearrow \\ & \circ & \end{array}$$

The top row implies that $p^* < t < t + 2\delta < q^*$. Because $\delta > r - p$ and $\delta > q - s$, we have:

$$r^* < (p + \delta)^* < t + \delta < (q - \delta)^* < s^*.$$

Therefore $t + \delta$ lies in the interval (r^*, s^*) and the above situation is impossible.

The other identity follows by symmetry. We conclude that \mathbb{U} and \mathbb{V} are interleaved for all $\delta > d^\infty((p, q), (r, s))$, and therefore $d_i(\mathbb{U}, \mathbb{V}) \leq d^\infty((p, q), (r, s))$. \square

Proposition 6.7. *Let $\mathbb{U} = \mathbb{I}(p^*, q^*)$ be an interval module, and 0 denote the zero persistence module. Then:*

$$d_i(\mathbb{U}, 0) = d^\infty((p, q), \Delta).$$

Proof. Because one of the modules is the zero module, all interleaving maps must be the zero map. The only condition that could fail is $\Psi\Phi = 1_{\mathbb{U}}^{2\delta}$, i.e., $0 = 1_{\mathbb{U}}^{2\delta}$. This holds when $\delta > \frac{1}{2}(q - p)$ and fails when $\delta < \frac{1}{2}(q - p)$. \square

As we can match points to the diagonal, we weaken our notion of bijection to the following.

Definition 6.8. A *partial matching* between two multisets A and B is a subset M of $A \times B$ such that:

- for every $\alpha \in A$ there is at most one $\beta \in B$ such that $(\alpha, \beta) \in M$; and,
- for every $\beta \in B$ there is at most one $\alpha \in A$ such that $(\alpha, \beta) \in M$.

Definition 6.9. A partial matching is a δ -*matching* when:

- for every $(\alpha, \beta) \in M$, $d^\infty(\alpha, \beta) \leq \delta$;
- if $\alpha \in A$ is unmatched then $d^\infty(\alpha, \Delta) \leq \delta$;
- if $\beta \in B$ is unmatched then $d^\infty(\beta, \Delta) \leq \delta$.

We now define the bottleneck distance in a similar manner to the interleaving distance.

Definition 6.10. The *bottleneck distance* between two multisets in \mathcal{H} is:

$$d_b(A, B) = \inf\{\delta \mid \text{there exists a } \delta\text{-matching between } A \text{ and } B\}.$$

Proposition 6.11. *The interleaving distance satisfies the triangle inequality*

$$d_b(A, C) \leq d_b(A, B) + d_b(B, C)$$

for any three multisets A, B, C .

Proof. Let M_1 be a δ_1 -matching between A and B , and M_2 a δ_2 matching between B and C . We define a $\delta = (\delta_1 + \delta_2)$ -matching between A and C by:

$$M = \{(\alpha, \gamma) \mid \text{there exists } \beta \in B \text{ such that } (\alpha, \beta) \in M_1, (\beta, \gamma) \in M_2\}.$$

If $(\alpha, \gamma) \in M$, then there is a β linking α and γ , so $d^\infty(\alpha, \gamma) \leq d^\infty(\alpha, \beta) + d^\infty(\beta, \gamma) \leq \delta_1 + \delta_2 = \delta$.

If α is unmatched in M , there are two possibilities. If α is unmatched in M_1 then $d^\infty(\alpha, \Delta) \leq \delta_1 \leq \delta$. If α is matched with β in M_1 then β must be unmatched in M_2 , in which case:

$$d^\infty(\alpha, \Delta) \leq d^\infty(\alpha, \beta) + d^\infty(\beta, \Delta) \leq \delta_1 + \delta_2 = \delta.$$

The same argument shows that if γ is unmatched in M then $d^\infty(\gamma, \Delta) \leq \delta$. Therefore M is a valid matching. The proposition follows by taking the infimum over all matchings M_1 and M_2 . \square

When we work in the realm of q-tame persistence modules, we can show that the inf is attained.

Theorem 6.12. *Let \mathbf{A} and \mathbf{B} be locally finite multisets in \mathcal{H} . If for every $\eta > \delta$ there exists an η -matching between \mathbf{A} and \mathbf{B} then there exists a δ -matching.*

Proof. Let \mathbf{M}_n be a $(\delta + \frac{1}{n})$ -matching for all $n \geq 1$. We will construct a δ -matching \mathbf{M} as the ‘limit’ of this sequence. Let $\chi, \chi_n : \mathbf{A} \times \mathbf{B} \rightarrow \{0, 1\}$ denote the indicator functions of \mathbf{M} and \mathbf{M}_n .

We construct χ as follows. Because \mathbf{A} and \mathbf{B} are locally finite, they must be countable. Fix an enumeration (a_ℓ, b_ℓ) of $\mathbf{A} \times \mathbf{B}$. We inductively define a descending sequence:

$$\mathbf{N} = \mathbf{N}_0 \supseteq \mathbf{N}_1 \supseteq \cdots \supseteq \mathbf{N}_\ell \supseteq \cdots$$

such that $\chi_n(a_\ell, b_\ell)$ has the same value for all $n \in \mathbf{N}_\ell$. The descending nature of the sequence means that $\chi_n(a_\ell, b_\ell)$ also agrees for all χ_n in $\mathbf{N}_{\ell'}$ with $\ell' > \ell$. We can then define $\chi(a_\ell, b_\ell)$ to be this common value.

Given $\mathbf{N}_{\ell-1}$, consider the two sets

$$\{n \in \mathbf{N}_{\ell-1} \mid \chi_n(a_\ell, b_\ell) = 0\} \text{ and } \{n \in \mathbf{N}_{\ell-1} \mid \chi_n(a_\ell, b_\ell) = 1\}.$$

At least one of these sets infinite, and we define \mathbf{N}_ℓ to be one such infinite set. This construction guarantees that we never run out of ‘choices’ for the next set of χ_n .

We now must prove that this χ is a valid δ -matching between \mathbf{A} and \mathbf{B} . First, note that for any finite subset $F \in \mathbf{A} \times \mathbf{B}$, there exists an χ_N beyond which χ agrees with χ_n on F for all $n \geq N$. Indeed, choose N to be the largest index ℓ of all pairs in F . By the definition of χ , the claim holds.

We can now check the conditions for a δ -matching.

- For $a \in \mathbf{A}$, there is at most one $b \in \mathbf{B}$ with $\chi(a, b) = 1$.

Proof. Suppose a is matched with both b and b' . By the claim above, there must be an n such that $\chi_n(a, b) = \chi_n(a, b') = 1$, which contradicts the fact that χ_n is a valid partial matching. \square

- For $b \in \mathbf{B}$, there is at most one $a \in \mathbf{A}$ with $\chi(a, b) = 1$.

Proof. Follows symmetrically. \square

- For $a \in \mathbf{A}$, if $d^\infty(a, \Delta) > \delta$, there is a $b \in \mathbf{B}$ with $\chi(a, b) = 1$.

Proof. Choose N such that $d^\infty(a, \Delta) > \delta + \frac{1}{N}$. The set:

$$F = \{b \in \mathbf{B} \mid d^\infty(a, b) \leq \delta + \frac{1}{N}\}$$

is finite, as \mathbf{B} is locally finite and this set lies away from the diagonal. By the claim above, there an ℓ for which $\chi(a, b) = \chi_n(a, b)$ for all $b \in F$ and $n \in \mathbf{N}_\ell$. Also, if $n \geq N$ then $\chi_n(a, b) = 1$ for some $b \in F$, as \mathbf{M}_n is a $(\delta + \frac{1}{n})$ -matching. Therefore $\chi(a, b) = \chi_n(a, b) = 1$ for some $b \in \mathbf{B}$. \square

- For $b \in \mathbf{B}$, if $d^\infty(b, \Delta) > \delta$, there is a $a \in \mathbf{A}$ with $\chi(a, b) = 1$.

Proof. Follows symmetrically. \square

- If $\chi(a, b) = 1$, then $d^\infty(a, b) \leq \delta$.

Proof. There exist infinitely many $n \in \mathbf{N}$ such that $\chi_n(a, b) = 1$. For all such n , the equation $d^\infty(a, b) \leq \delta + \frac{1}{n}$ holds. Because n becomes arbitrarily large, the result follows. \square

This completes the proof of the theorem. \square

6.3 Interpolation lemma

The first main ingredient of the stability theorem is the interpolation lemma, which proves that for any interleaved pair of persistence modules, there is a 1-parameter family of modules that interpolates between them. We begin by giving an alternative characterisation of an interleaving between two modules.

The plane \mathbf{R}^2 has a partial order on it defined by

$$(p_1, q_1) \leq (p_2, q_2) \text{ whenever } p_1 \leq p_2 \text{ and } q_1 \leq q_2$$

For any real number x , we define shifted diagonal Δ_x to be:

$$\Delta_x = \{(p, q) \mid q - p = 2x\} \subset \mathbf{R}^2.$$

Any persistence modules over \mathbf{R} can be then considered as a persistence module over Δ_x , by identifying each $t \in \mathbf{R}$ with the point $(t - x, t + x) \in \Delta_x$.

Proposition 6.13. *Two persistence modules \mathbb{U}, \mathbb{V} are $|y - x|$ -interleaved if and only if there is a persistence module \mathbb{W} over $\Delta_x \cup \Delta_y$ such that $\mathbb{W}_{\Delta_x} = \mathbb{U}$ and $\mathbb{W}_{\Delta_y} = \mathbb{V}$.*

Proof. Let $\delta = |y - x|$, and let us assume $x < y$. In addition to the maps within each of Δ_x and Δ_y , we also have maps between the two lines. Let $\Phi = \{\phi_t\}$ be the vertical maps from Δ_x to Δ_y and $\Psi = \{\psi_t\}$ the horizontal maps from Δ_y to Δ_x . Then ϕ_t and ψ_t are maps:

$$\begin{aligned} \phi_t : U_t = W_{t-x, t+x} &\rightarrow W_{t-x, t+2y-x} = V_{t+\delta} \\ \psi_t : V_t = W_{t-y, t+y} &\rightarrow W_{t+y-2x, t+y} = U_{t+\delta}. \end{aligned}$$

The composition law for \mathbb{W} implies that $\Psi\Phi = 1_{\mathbb{U}}^{2\delta}$ and $\Phi\Psi = 1_{\mathbb{V}}^{2\delta}$ as required.

The module \mathbb{W} does not contain more information than the interleaving, as any map w_S^T can be factored into the form:

$$\begin{aligned} w_S^T &= u_s^t && \text{if } S, T \in \Delta_x \\ w_S^T &= v_s^t && \text{if } S, T \in \Delta_y \\ w_S^T &= v_{s+\delta}^t \phi_s && \text{if } S \in \Delta_x \text{ and } T \in \Delta_y \\ w_S^T &= u_{s+\delta}^t \psi_s && \text{if } S \in \Delta_y \text{ and } T \in \Delta_x. \end{aligned}$$

\square

Lemma 6.14 (Interpolation lemma). *Let \mathbb{U} and \mathbb{V} be δ -interleaved modules. Then there exists a family of persistence modules $\{\mathbb{U}_x \mid x \in [0, \delta]\}$ such that $\mathbb{U}_0 = \mathbb{U}$, $\mathbb{U}_\delta = \mathbb{V}$ and $\mathbb{U}_x, \mathbb{U}_y$ are $|y - x|$ interleaved for all $x, y \in [0, \delta]$.*

Proof. In view of the previous proposition, this is equivalent to showing that there exists a persistence module $\overline{\mathbb{W}}$ over the strip $\Delta_{[0, \delta]}$ where $\overline{\mathbb{W}}|_{\Delta_0} = \mathbb{U}$ and $\overline{\mathbb{W}}|_{\Delta_\delta} = \mathbb{V}$. Then, any restriction to $\Delta_x \cup \Delta_y$ provides the required interleavings for $x, y \in [0, \delta]$.

To simplify the proof, we begin by scaling and translating the plane so that the interval $[0, \delta]$ is replaced with $[-1, 1]$. The δ -interleaving between \mathbb{U} and \mathbb{V} gives us a persistence module \mathbb{W} over $\Delta_{-1} \cup \Delta_1$.

We wish to construct two persistence modules over the strip $\Delta_{[-1, 1]}$ and a map Ω between them. The persistence module $\overline{\mathbb{W}}$ will be the image of this map.

We first construct four persistence modules over all of \mathbf{R}^2 as follows.

$$\begin{array}{lll} \mathbb{A} = \mathbb{U}[p-1] & \text{given by } A_{(p,q)} = U_{p-1} & \text{and } a_{(p,q)}^{(r,s)} = u_{p-1}^{r-1} \\ \mathbb{B} = \mathbb{V}[q-1] & \text{given by } B_{(p,q)} = V_{q-1} & \text{and } b_{(p,q)}^{(r,s)} = v_{q-1}^{s-1} \\ \mathbb{C} = \mathbb{U}[q+1] & \text{given by } C_{(p,q)} = U_{p+1} & \text{and } c_{(p,q)}^{(r,s)} = u_{q+1}^{s+1} \\ \mathbb{D} = \mathbb{V}[p+1] & \text{given by } D_{(p,q)} = V_{q+1} & \text{and } d_{(p,q)}^{(r,s)} = v_{p+1}^{r+1} \end{array}$$

We now define four module maps by abuse of notation:

$$\begin{array}{lll} 1_{\mathbb{U}} : \mathbb{A} \rightarrow \mathbb{C} & \text{defined at } (p, q) \text{ to be} & u_{p-1}^{q+1} : U_{p-1} \rightarrow U_{q+1} \\ \Phi : \mathbb{A} \rightarrow \mathbb{D} & \text{defined at } (p, q) \text{ to be} & \phi_{p-1} : U_{p-1} \rightarrow V_{p+1} \\ \Psi : \mathbb{B} \rightarrow \mathbb{C} & \text{defined at } (p, q) \text{ to be} & \psi_{q-1} : V_{q-1} \rightarrow U_{q+1} \\ 1_{\mathbb{V}} : \mathbb{B} \rightarrow \mathbb{D} & \text{defined at } (p, q) \text{ to be} & v_{q-1}^{p+1} : V_{q-1} \rightarrow V_{p+1}. \end{array}$$

The maps Φ and Ψ are the interleaving maps and are defined on all of \mathbf{R}^2 . The map $1_{\mathbb{U}}$ is defined whenever $p-1 \leq q+1$, and $1_{\mathbb{V}}$ is defined when $q-1 \leq p+1$. It follows that all four maps are defined when $-2 \leq q-p \leq 2$. This is precisely the region $\Delta_{[-1, 1]}$.

Now let $\Omega \in \text{Hom}(\mathbb{A} \oplus \mathbb{B}, \mathbb{C} \oplus \mathbb{D})$ be the map given by the following matrix:

$$\Omega = \begin{pmatrix} 1_{\mathbb{U}} & \Psi \\ \Phi & 1_{\mathbb{V}} \end{pmatrix}.$$

We claim that $\overline{\mathbb{W}} = \text{im}(\Omega)$ is the persistence module over $\Delta_{[-1, 1]}$ that we require. The only thing to verify is that $\mathbb{U} \cong \overline{\mathbb{W}}|_{\Delta_{-1}}$ and $\mathbb{V} \cong \overline{\mathbb{W}}|_{\Delta_1}$.

On $\Delta_{-1} = \{(t+1, t-1)\}$ we have:

$$\begin{aligned} (\mathbb{A} \oplus \mathbb{B})_t &= U_t \oplus V_{t-2} \\ (\mathbb{C} \oplus \mathbb{D})_t &= U_t \oplus V_{t+2} \end{aligned}$$

and

$$\omega_t = \begin{pmatrix} u_t^t & \psi_{t-2} \\ \phi_t & v_{t-2}^{t+2} \end{pmatrix}.$$

Note that the map v_{t-2}^{t+2} can be factorized using the interleaving maps as $\phi_t \psi_{t-2}$. We then have:

$$\omega_t = \begin{pmatrix} 1 & \psi_{t-2} \\ \phi_t & v_{t-2}^{t+2} \end{pmatrix} = \begin{pmatrix} 1 & \psi_{t-2} \\ \phi_t & \phi_t \psi_{t-2} \end{pmatrix} = \begin{pmatrix} 1 \\ \phi_t \end{pmatrix} \begin{pmatrix} 1 & \psi_{t-2} \end{pmatrix},$$

which, when written as a diagram of modules, takes the form

$$\mathbb{U} \oplus \mathbb{V}[t-2] \xrightarrow{\Omega_1} \mathbb{U} \xrightarrow{\Omega_2} \mathbb{U} \oplus \mathbb{V}[t+2],$$

where

$$\Omega_1(U_t \oplus V_t) = U_t + \Psi(V_t) \quad \text{and} \quad \Omega_2(U_t) = U_t \oplus \Phi(U_t).$$

Note that Ω_1 is surjective and Ω_2 is injective for every t . Therefore $\mathbb{U} \cong \text{im}(\Omega|_{\Delta_{-1}})$.

The argument for $\mathbb{V} \cong \overline{\mathbb{W}}|_{\Delta_1}$ is almost identical. On $\Delta_{-1} = \{(t-1, t+1)\}$ the modules are

$$\begin{aligned} (\mathbb{A} \oplus \mathbb{B})_t &= U_{t-2} \oplus V_t \\ (\mathbb{C} \oplus \mathbb{D})_t &= U_{t+2} \oplus V_t \end{aligned}$$

and Ω is

$$\omega_t = \begin{pmatrix} u_{t-2}^{t+2} & \psi_t \\ \phi_{t-2} & v_t^t \end{pmatrix} = \begin{pmatrix} \psi_t \phi_{t-2} & \psi_t \\ \phi_{t-2} & 1 \end{pmatrix} = \begin{pmatrix} \psi_t \\ 1 \end{pmatrix} \begin{pmatrix} \phi_{t-2} & 1 \end{pmatrix}$$

This is equivalent to

$$\mathbb{U}[t-2] \oplus \mathbb{V} \xrightarrow{\Omega_3} \mathbb{V} \xrightarrow{\Omega_4} \mathbb{U}[t+2] \oplus \mathbb{V}$$

where

$$\Omega_3(U_t \oplus V_t) = \Phi(U_t) + V_t \quad \text{and} \quad \Omega_4(V_t) = \Phi(V_t) \oplus V_t$$

Again we have Ω_3 surjective and Ω_4 injective, so $\mathbb{V} \cong \text{im}(\Omega|_{\Delta_1})$. \square

6.4 Box lemma

The second ingredient of the stability theorem is the box lemma, which allows us to relate the persistence diagrams of two interleaved persistence modules locally.

Let $R = [a, b] \times [c, d]$ be a rectangle in \mathbf{R}^2 . The δ -thickening of R , written R^δ , is the rectangle:

$$R^\delta = [a - \delta, b + \delta] \times [c - \delta, d + \delta].$$

We will also refer to the thickening of a single point. If $\alpha = (p, q)$, then:

$$\alpha^\delta = [p - \delta, p + \delta] \times [q - \delta, q + \delta].$$

Lemma 6.15 (Box lemma). *Let \mathbb{U}, \mathbb{V} be δ -interleaved persistence modules. Let R be a rectangle. Then $\mu_{\mathbb{U}}(R) \leq \mu_{\mathbb{V}}(R^\delta)$ and $\mu_{\mathbb{V}}(R) \leq \mu_{\mathbb{U}}(R^\delta)$.*

Proof. Let $R = [a, b] \times [c, d]$ and $R^\delta = [A, B] \times [C, D]$. We have the following persistence modules:

$$\mathbb{U}_{a,b,c,d} : U_a \rightarrow U_b \rightarrow U_c \rightarrow U_d$$

and

$$\mathbb{U}_{A,B,C,D} : V_A \rightarrow V_B \rightarrow V_C \rightarrow V_D$$

by restricting to the appropriate finite sets. The interleaving maps allow us to combine these persistence modules:

$$\mathbb{W} : V_A \xrightarrow{\Psi_A} U_a \longrightarrow U_b \xrightarrow{\Phi_b} V_B \longrightarrow V_C \xrightarrow{\Psi_C} U_c \longrightarrow U_d \xrightarrow{\Phi_d} V_D$$

where $\Phi : \mathbb{U} \rightarrow \mathbb{V}$ and $\Psi : \mathbb{V} \rightarrow \mathbb{U}$ are the interleaving maps. We now calculate:

$$\begin{aligned} \mu_{\mathbb{V}}(R^\delta) &= \langle \circ_A \text{-----} \bullet_B \text{---} \bullet_C \text{-----} \circ_D \mid \mathbb{V} \rangle \\ &= \langle \circ_A \text{-----} \bullet_B \text{---} \bullet_C \text{-----} \circ_D \mid \mathbb{W} \rangle \\ &= \langle \circ_A \text{---} \circ_a \text{---} \bullet_b \text{---} \bullet_B \text{---} \bullet_C \text{---} \bullet_c \text{---} \circ_d \text{---} \circ_D \mid \mathbb{W} \rangle + \text{other terms} \\ &\geq \langle \circ_A \text{---} \circ_a \text{---} \bullet_b \text{---} \bullet_B \text{---} \bullet_C \text{---} \bullet_c \text{---} \circ_d \text{---} \circ_D \mid \mathbb{W} \rangle \\ &= \langle \text{---} \circ_a \text{---} \bullet_b \text{-----} \bullet_c \text{---} \circ_d \text{---} \mid \mathbb{W} \rangle \\ &= \langle \text{---} \circ_a \text{---} \bullet_b \text{-----} \bullet_c \text{---} \circ_d \text{---} \mid \mathbb{U} \rangle \\ &= \mu_{\mathbb{U}}(R). \end{aligned}$$

The other inequality follows by symmetry. □

6.5 The stability theorem

Finally, we combine the above results to prove the stability theorem. Recall the statement:

Theorem 6.16 (Stability). *Let $f, g : X \rightarrow \mathbf{R}$ be functions on a topological space X and let $\mathbb{U} = H(X^f), \mathbb{V} = H(X^g)$ be the sublevel persistence modules. If \mathbb{U} and \mathbb{V} are q -tame, then:*

$$d_b(\text{dgm}(\mathbb{U}), \text{dgm}(\mathbb{V})) \leq \|f - g\|_\infty.$$

The first step is to apply the inequality $d_i(\mathbb{U}, \mathbb{V}) \leq \|f - g\|_\infty$. The stability theorem can be proven by showing that:

$$d_b(\text{dgm}(\mathbb{U}), \text{dgm}(\mathbb{V})) \leq d_i(\mathbb{U}, \mathbb{V}).$$

In other words,

Proposition 6.17. *Let \mathbb{U}, \mathbb{V} be q -tame persistence modules that are η -interleaved for all $\eta > \delta$. Then there exists a δ -matching between $\mathbf{dgm}(\mathbb{U})$ and $\mathbf{dgm}(\mathbb{V})$.*

In view of Theorem 6.12, this is a consequence of the following easier proposition:

Proposition 6.18. *Let \mathbb{U}, \mathbb{V} be q -tame persistence modules that are δ -interleaved. Then there exists a δ -matching between $\mathbf{dgm}(\mathbb{U})$ and $\mathbf{dgm}(\mathbb{V})$.*

Proposition 6.17 follows, because if there is a η -matching for all $\eta > \delta$, then there is a δ -matching.

The interpolation lemma states that there exists a family of persistence modules $\{\mathbb{U}_x \mid x \in [0, \delta]\}$ such that \mathbb{U}_x and \mathbb{U}_y are $|y-x|$ -interleaved for all $x, y \in [0, \delta]$. Let $\{\mu_x\}$ be the associated r-measures.

First we show that the ‘Hausdorff distance’ between $\mathbf{dgm}(\mu_x)$ and $\mathbf{dgm}(\mu_y)$ is always less than $|y-x|$. Intuitively, the Hausdorff distance is similar to the bottleneck distance but we are permitted to match multiple points from one diagram to a single point in the other.

Lemma 6.19. *Let $A = \mathbf{dgm}(\mu_x)$, $B = \mathbf{dgm}(\mu_y)$ and $\eta = |y-x|$. Then:*

- *If $a \in A$ and $d^\infty(a, \Delta) > \eta$ then there exists $b \in B$ with $d^\infty(a, b) \leq \eta$; and,*
- *If $b \in B$ and $d^\infty(b, \Delta) > \eta$ then there exists $a \in A$ with $d^\infty(a, b) \leq \eta$.*

Proof. Let $a \in A$ be a point such that $d^\infty(a, \Delta) > \eta$. Choose an ε such that $\eta + \varepsilon < d^\infty(a, \Delta)$.

By the box inequality,

$$1 \leq \mu_x(a^\varepsilon) \leq \mu_y(a^{\eta+\varepsilon})$$

and we have at least one point of B in $a^{\eta+\varepsilon}$. Because ε was arbitrary and B is locally finite, there must be at least one point b of B in a^η . It follows that $d^\infty(a, b) \leq \eta$.

The second statement follows symmetrically. \square

We now prove the theorem when $\mathbf{dgm}(\mu_x)$ is finite for all x . Once we have done this, we can finish off the proof of the main theorem using a compactness argument. For convenience we will write $A_x = \mathbf{dgm}(\mu_x)$.

Lemma 6.20. *If A_x is finite for all x , then the theorem holds.*

Proof. We claim that for every $x \in [0, \delta]$, there exists a number $\rho(x) > 0$ such that A_x, A_y are $|y-x|$ -matched whenever $|y-x| < \rho(x)$.

Let a_1, \dots, a_k be the distinct points of A_x , with n_1, \dots, n_k their multiplicities. Define $\rho(x)$ to be:

$$\rho(x) = \min \begin{cases} \frac{1}{2}d^\infty(a_i, \Delta) & \text{for all } i \\ \frac{1}{2}d^\infty(a_i, a_j) & \text{for all } i, j. \end{cases}$$

We must show that if $|y - x| < \rho(x)$ then A_x, A_y are $|y - x|$ -matched. Let $\eta = |y - x|$. By Δ^η , we mean the thickening:

$$\begin{aligned}\Delta^\eta &= \{\alpha \in \mathbf{R}^2 \mid d^\infty(\alpha, \Delta) \leq \eta\} \\ &= \{(p, q) \in \mathbf{R}^2 \mid q - p \leq 2\eta\}.\end{aligned}$$

By the previous lemma, A_x and A_y have Hausdorff distance less than η . It follows that A_y is contained in the closed set:

$$\Delta^\eta \cup \alpha_1^\eta \cup \dots \cup \alpha_k^\eta.$$

The definition of $\rho(x)$ is chosen so that all of these sets are disjoint.

Choose $\epsilon > 0$ so that $2\eta + \epsilon < 2\rho(x)$. For every box a_i^η , the box lemma implies:

$$n_i = \mu_x(a_i^\epsilon) \leq \mu_y(a_i^{\eta+\epsilon}) \leq \mu_x(a_i^{2\eta+\epsilon}) = n_i.$$

We can therefore match the n_i copies of a_i with the points of A_y that are in the square a_i^η , for all i . The only unmatched points are the ones in Δ^η , so this defines a valid partial matching.

This shows that A_0 and $A_\delta = B$ are δ -matched by the following argument.

The triangle inequality for matchings states that if A_0, A_x are x -matched, and A_x, A_y are $|y - x|$ -matched, then A_0, A_y are y -matched. We must show that we can continue matching in this way until A_0 and A_δ are δ -matched.

Let $m = \sup\{x \in [0, \delta] \mid A_0 \text{ and } A_x \text{ are } x\text{-matched}\}$. First we show that this sup is attained. Choose an m' , such that $m - m' < \rho(m)$. Then A_0 and $A_{m'}$ are m' -matched by the definition of m , and $A_{m'}$ and A_m are $|m' - m|$ -matched by the argument above. Therefore A_0 and A_m are m -matched.

Finally, we must have $m = \delta$. If $m < \delta$ then there exists a m'' with $m < m'' < \delta$ and $m'' - m < \rho(m)$. Again we have that A_0 and $A_{m''}$ are m'' -matched, contradicting the definition of m .

Therefore A_0 and A_δ are δ -matched. \square

Theorem 6.21. *The stability theorem holds, without assuming each A_x is finite.*

Proof. We prove this using a similar compactness argument to Theorem 6.12. Let $\{\mathcal{H}_n\}$ be an increasing sequence of open subsets of \mathcal{H} such that the union of these sets is \mathcal{H} and each \mathcal{H}_n has compact closure.

Because each A_x is locally finite, we have that $A_x \cap \mathcal{H}_n$ is finite for all x and n . By the previous lemma, for each n we have a δ -matching M_n between $A_0 \cap \mathcal{H}_n$ and $A_\delta \cap \mathcal{H}_n$.

We now reuse the construction and arguments of Theorem 6.12 to get a limit multiset M . We must now verify that this M gives a δ -matching. Firstly, every matched pair is separated by at most δ , as this is true for every M_n . Secondly, every $a \in A_0$ is matched with at most one $b \in A_\delta$ and vice versa, for the same reason as before.

All that remains to be checked is that every $a \in A_0$ with $d^\infty(a, \Delta) > \delta$ is matched. Consider the thickening a^δ for some such a . This square is compact

and contained in \mathcal{H} , so must be contained in \mathcal{H}_n for n sufficiently large. Therefore a is matched in \mathbf{M}_n for n sufficiently large. Because \mathbf{A}_δ is locally finite, α has only finitely many points within δ distance from a . As in the earlier theorem, there exists an n beyond which the matchings \mathbf{M}_n all agree on this finite set. Because the \mathbf{M}_n are all valid matchings, we conclude that a is matched with some b in \mathbf{M} .

By symmetry, any $b \in \mathbf{A}_\delta$ with $d^\infty(b, \Delta) > \delta$ is matched in \mathbf{M} , so \mathbf{M} is the required matching. \square

7 Category Theory Primer

Like much of algebraic topology, the theory of persistence modules has been reinterpreted using the language of category theory. This section introduces the main concepts used in category theory and shows how many of the constructions we've seen fit into this viewpoint.

In the following chapter we will use these techniques to prove a strong structure theorem for q-tame persistence modules.

Definition 7.1. A *category* \mathcal{C} consists of:

- a class $\text{Ob}(\mathcal{C})$ of objects;
- for every $a, b \in \text{Ob}(\mathcal{C})$, a class $\text{Hom}(a, b)$ of morphisms; and,
- for every $a, b, c \in \text{Ob}(\mathcal{C})$, a binary operation

$$\text{Hom}(b, c) \times \text{Hom}(a, b) \rightarrow \text{Hom}(a, c), \quad (g, f) \mapsto g \circ f \text{ or } gf$$

such that:

- morphism composition is *associative*: if $f : a \rightarrow b$, $g : b \rightarrow c$ and $h : c \rightarrow d$ then $f \circ (g \circ h) = (f \circ g) \circ h$; and,
- every object has an *identity* morphism: for every $a \in \text{Ob}(\mathcal{C})$, there exists a morphism $1_a \in \text{Hom}(a, a)$ such that for every $f : x \rightarrow a$, $1_a \circ f = f$ and for every $g : a \rightarrow y$, $g \circ 1_a = g$.

Many familiar mathematical objects form a category. We have the categories:

- **Set**, of sets and functions;
- **Top**, of topological spaces and continuous maps;
- **Grp**, of groups and homomorphisms; and,
- **Vect_k**, of **k**-vector spaces and linear maps.

It is routine to check that these classes satisfy the conditions for a category above.

The morphisms in a category are often functions with some added conditions, but this is not always the case. Given any poset P , we can consider P as a category by taking the elements of P as our objects, and including precisely one morphism $s \rightarrow t$ for each $s \leq t$.

We can apply category theory to the study of persistence because, given a indexing poset P , the class of persistence modules together with module homomorphisms forms a category **Pers_P**.

A *subcategory* \mathcal{S} of a category \mathcal{C} is a category with objects and morphisms subcollections of the objects and morphisms of \mathcal{C} . We require that these objects and morphisms form a valid category.

A subcategory \mathcal{S} is *full* if for every $a, b \in \text{Ob}(\mathcal{S})$, we have $\text{Hom}_{\mathcal{S}}(a, b) = \text{Hom}_{\mathcal{C}}(a, b)$. In other words, \mathcal{S} includes all the morphisms between a and b that exist in \mathcal{C} .

7.1 Functors and natural transformations

The next fundamental concept in category theory is the notion of a *functor* between two categories.

Definition 7.2. Given two categories \mathcal{C}, \mathcal{D} , a *functor* $F : \mathcal{C} \rightarrow \mathcal{D}$ is a mapping that assigns:

- to every object $c \in \mathcal{C}$ an object $F(c) \in \mathcal{D}$; and,
- to every morphism $f : x \rightarrow y$ in \mathcal{C} , a morphism $F(f) : F(x) \rightarrow F(y)$ in \mathcal{D}

such that:

$$\begin{aligned} F(1_a) &= 1_{F(a)} \\ F(g \circ f) &= F(g) \circ F(f). \end{aligned}$$

The most important example of a functor we have seen so far is homology. If the ground ring is a field \mathbf{k} , then $H_n(-) : \mathbf{Top} \rightarrow \mathbf{Vect}_{\mathbf{k}}$ is a functor assigning to each space X the n -dimensional homology of X , and to each continuous map $f : X \rightarrow Y$ the linear map given by Theorem 2.6.

In fact, persistence modules themselves can be thought of as functors. Let P be a poset. Then persistence modules over P are precisely functors $\mathbb{V} : P \rightarrow \mathbf{Vect}_{\mathbf{k}}$, where we consider P as a category using the construction earlier.

Moving up the abstraction tower, we can also define *natural transformations* between two functors.

Definition 7.3. Given two functors $F, G : \mathcal{C} \rightarrow \mathcal{D}$, a *natural transformation* $\tau : F \Rightarrow G$ is a function that assigns to every object $c \in \mathcal{C}$, a morphism $\tau_c : F(c) \rightarrow G(c)$ in \mathcal{D} , such that for every $f : c \rightarrow c'$ in \mathcal{C} , the diagram

$$\begin{array}{ccc} F(c) & \xrightarrow{\tau_c} & G(c) \\ F(f) \downarrow & & \downarrow G(f) \\ F(c') & \xrightarrow{\tau_{c'}} & G(c') \end{array}$$

commutes.

If we choose two categories \mathcal{C}, \mathcal{D} , we can form the *functor category* $[\mathcal{C}, \mathcal{D}]$ which has as its objects the class of all functors $\mathcal{C} \rightarrow \mathcal{D}$, and morphisms the natural transformations between them.

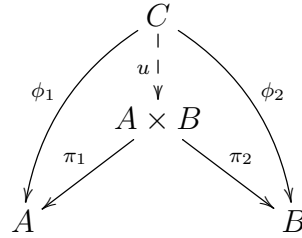
A module homomorphism $\mathbb{U} \rightarrow \mathbb{V}$ is precisely a natural transformation between the modules, when we consider the modules as functors. We can therefore make the identification $\mathbf{Pers}_P = [P, \mathbf{Vect}_{\mathbf{k}}]$.

7.2 Limits and colimits

The utility of category theory becomes apparent when we consider the ‘universal’ objects that often appear in mathematics. A common template for new constructions is to find the simplest object that satisfies certain properties.

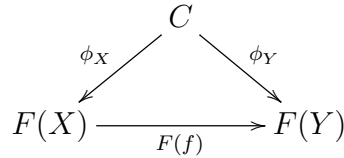
Consider the ‘direct product’ operation on sets. In the case of sets, we know the direct product $A \times B$ is the collection of all ordered pairs (a, b) with $a \in A, b \in B$. We could instead characterise the direct product as the ‘universal’ object with maps to the two components.

Definition 7.4. The direct product of two sets A and B is the unique set $A \times B$ with two maps $\pi_1 : A \times B \rightarrow A$ and $\pi_2 : A \times B \rightarrow B$ such that whenever some other set C exists with maps $\phi_1 : C \rightarrow A$, $\phi_2 : C \rightarrow B$, these maps factor through a unique map $u : C \rightarrow A \times B$.

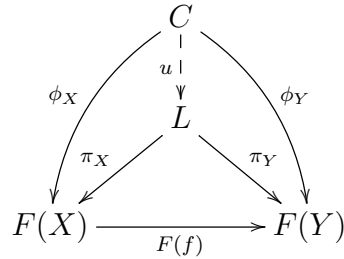


A *limit* in category theory is the generalisation of this idea.

Definition 7.5. Let $F : \mathcal{J} \rightarrow \mathcal{C}$ be a functor. A *cone to F* is an object C of \mathcal{C} together with a family of morphisms $\phi_X : C \rightarrow F(X)$ for every $X \in \mathcal{J}$, such that the following diagram commutes for every $f : X \rightarrow Y$ in \mathcal{J} .



The *limit* of F is a cone (L, π_X) to F such that for any other cone (C, ϕ_X) to F , there is a unique morphism $u : C \rightarrow L$ such that $\phi_X = \pi_X \circ u$ for all $X \in \mathcal{J}$. In other words, there exists a morphism u making the following diagram commute:



A limit is *small* if in the indexing category \mathcal{J} , the collection of objects and morphisms forms a set and not a proper class.

We recover the above example when we set \mathcal{J} to the discrete category $\{1, 2\}$, \mathcal{C} to the category **Set**, and F to the functor assigning $F(1) = A$ and $F(2) = B$.

A *colimit* is the categorical dual of a limit, which we can define by reversing the arrows in the definition above.

Definition 7.6. Let $F : \mathcal{J} \rightarrow \mathcal{C}$ be a functor. A *cocone from F* is an object C of \mathcal{C} together with a family of morphisms $\phi_X : F(X) \rightarrow C$ for every $X \in \mathcal{J}$, such that the following diagram commutes for every $f : X \rightarrow Y$ in \mathcal{J} .

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ & \searrow \phi_X & \swarrow \phi_Y \\ & C & \end{array}$$

The *colimit* of F is a cocone (L, π_X) from F such that for any other cocone (C, ϕ_X) from F , there is a unique morphism $u : L \rightarrow C$ such that $\phi_X = u \circ \pi_X$ for all $X \in \mathcal{J}$. In other words, there exists a morphism u making the following diagram commute:

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ & \searrow \pi_X & \swarrow \pi_Y \\ & L & \\ & \downarrow u & \\ & C & \end{array}$$

ϕ_X (curved arrow from $F(X)$ to C) ϕ_Y (curved arrow from $F(Y)$ to C)

If we choose $\mathcal{J} = \{1, 2\}$, $\mathcal{C} = \mathbf{Set}$, and F the functor $F(1) = A$ and $F(2) = B$ as before, the colimit of F is the disjoint union $A \coprod B$. The maps π_1, π_2 are then the inclusions of A and B into $A \coprod B$.

Limits and colimits do not necessarily always exist. A category is called *complete* if it contains all small limits, and *cocomplete* if it contains all small colimits. The categories **Set** and **Vect_k** are both complete and cocomplete.

8 The Observable Category

We have seen throughout the previous sections of this work that persistence modules over \mathbf{R} are quite badly behaved. Not every persistence module decomposes into interval modules. We restrict our attention to q-tame persistence modules, but even those do not always decompose into interval modules.

We then defined the persistence diagram as an invariant for q-tame persistence modules, but it is not a complete invariant: two non-isomorphic q-tame persistence modules can have the same persistence diagram.

In a recent paper, Chazal et al. [6] show that these deficiencies vanish completely if we make a small adjustment to the category we are working in. Philosophically, the important data carried by persistence modules is how information persists over intervals, so features that exist over short intervals are seen as unimportant. Taken to the extreme we have what we refer to as *ephemeral features*; features that exist only over a single point.

In this section we will construct the *observable category* of persistence modules \mathbf{Ob} , where such ephemeral information is ignored by design. We then show that this category has all the properties we could want: every q-tame module decomposes into a sum of interval modules and the persistence diagram is a complete invariant.

We again restrict our attention to persistence modules over \mathbf{R} . To simplify the notation, we will write \mathbf{Pers} to mean $\mathbf{Pers}_{\mathbf{R}}$.

8.1 Construction

First, a more formal definition of the persistence modules we are eliminating.

Definition 8.1. An *ephemeral module* is one where the structure maps v_s^t are 0 for all $s < t$.

The collection of all ephemeral modules and morphisms forms a category \mathbf{Eph} , which is a full subcategory of \mathbf{Pers} . We now define a weaker notion of isomorphism, where two modules are considered equivalent if they only differ by ephemeral modules.

Definition 8.2. A *weak isomorphism* is a module homomorphism Φ such that both $\ker \Phi$ and $\operatorname{coker} \Phi$ are ephemeral modules. Recall that, for persistence modules, kernels and cokernels are done pointwise.

For example, consider the module map $\Phi : \mathbb{I}[p, q] \rightarrow \mathbb{I}[p, q]$, given by $\phi_t = 1$ where both domain and codomain are \mathbf{k} . Then Φ is a weak isomorphism as $\ker \Phi = \mathbb{I}[q, q]$ and $\operatorname{coker} \Phi = 0$, both ephemeral.

We now wish to construct the ‘quotient category’ $\mathbf{Pers}/\mathbf{Eph}$, which we refer to as the *observable category* \mathbf{Ob} . By quotient category, we mean that we have a quotient functor $\pi : \mathbf{Pers} \rightarrow \mathbf{Ob}$ satisfying the following universal property: π sends all weak isomorphisms of \mathbf{Pers} to actual isomorphisms in \mathbf{Ob} , and any

other functor $F : \mathbf{Pers} \rightarrow \mathcal{C}$ that also sends weak isomorphisms to isomorphisms factors uniquely through π .

In problems of this sort there is often no explicit construction of the quotient category. In this case, however the objects and morphisms of \mathbf{Ob} are easy to describe. We will give the construction of \mathbf{Ob} first and then verify that it satisfies the universal property.

Ordinary module morphisms in \mathbf{Pers} are collections of linear maps $\phi_t : U_t \rightarrow V_t$. In what follows, it will be more useful to think of them as maps $\phi_s^t : U_s \rightarrow V_t$, such that $\phi_s^t = v_q^t \phi_p^q u_s^p$ whenever $s \leq p \leq q \leq t$. We can convert readily between the two representations by setting $\phi_t = \phi_t^t$ in one direction, and $\phi_s^t = \phi_t u_s^t = v_s^t \phi_s$ in the other. With this in mind:

Definition 8.3. An *observable morphism* or *ob-morphism*, $\Phi^\circ : \mathbb{U} \dashrightarrow \mathbb{V}$, is a collection of linear maps $\phi_s^t : U_s \rightarrow V_t$ for all $s < t$, such that $\phi_s^t = v_q^t \phi_p^q u_s^p$ whenever $s \leq p < q \leq t$. Note that, in contrast with ordinary morphisms, s must be strictly less than t .

If $\Phi^\circ : \mathbb{U} \dashrightarrow \mathbb{V}$ and $\Psi^\circ : \mathbb{V} \dashrightarrow \mathbb{W}$ are ob-morphisms, the composition $\Psi^\circ \Phi^\circ$ is given by $(\Psi^\circ \Phi^\circ)_s^t = \psi_u^t \phi_s^u$ for any $s < u < t$. The condition above guarantees that this is well defined.

Definition 8.4. The *observable category* \mathbf{Ob} is the collection of persistence modules together with ob-morphisms. The quotient functor $\pi : \mathbf{Pers} \rightarrow \mathbf{Ob}$ maps every object to itself, and every morphism $\Phi = \{\phi_s^t \mid s \leq t\}$ to $\pi(\Phi) = \Phi^\circ = \{\phi_s^t \mid s < t\}$ by forgetting each ϕ_t^t .

In the remainder of this section we show that \mathbf{Ob} satisfies the universal property. Theorem 8.5 shows that π sends weak isomorphisms to isomorphisms, and Theorem 8.9 shows that other functors of this sort must factor uniquely through π .

Theorem 8.5. *If $\Phi : \mathbb{V} \rightarrow \mathbb{W}$ is a weak isomorphism then $\Phi^\circ = \pi(\Phi)$ is invertible in \mathbf{Ob} and is therefore an isomorphism.*

Proof. We construct an inverse $\Psi^\circ = \{\psi_s^t\}$ as follows. For a given $s < t$, choose an intermediate $s < u < t$.

Because $\text{coker } \Phi$ is ephemeral, the composition of $w_s^u : W_s \rightarrow W_u$ with the natural map $W_u \rightarrow \text{coker } \phi_u$ is zero. Otherwise, we would have some subspace of $h \subset V_s$ such that $w_s^u \phi_s(h) = \phi_u v_s^u(h) \subset \text{coker } \phi_u$, which is a contradiction. Therefore, w_s^u factors as a map $\omega_s^u : W_s \rightarrow \text{im } \phi_u$ followed by the inclusion of $\text{im } \phi_u$ into W_u .

Also, $\ker \Phi$ is ephemeral, so the composition of the inclusion $\ker \phi_u \rightarrow V_u$ and $v_u^t : V_u \rightarrow V_t$ is zero. We now get an induced map $\tau_u^t : \text{im } \phi_u \rightarrow V_t$, so that the composition of this map with the natural map $V_u \rightarrow \text{im } \phi_u$ is v_u^t .

We now define our inverse map to be $\psi_s^t = \tau_u^t \omega_s^u$. This gives a valid inverse for ϕ , as each map only ‘forgets’ information in the kernel or cokernel, both of which are ephemeral. \square

Given an arbitrary persistence module \mathbb{V} , we now construct a ‘simpler’ persistence module $\overline{\mathbb{V}}$ that captures the non-ephemeral information in \mathbb{V} .

Definition 8.6. Let \mathbb{V} be a persistence module. We define $\overline{\mathbb{V}}$ to be the persistence module with spaces

$$\overline{V}_t = \text{colim} \{V_s \mid s < t\}$$

and structure maps given by the universal property of colimits. These colimits exist as \mathbf{Vect}_k is cocomplete.

This definition is a little opaque, so we unpack it as follows. The collection $\{V_s \mid s < t\}$ forms a diagram indexed by all real numbers less than t which we informally represent as follows:

$$\dots \longrightarrow V_{s_1} \longrightarrow \dots \longrightarrow V_{s_2} \longrightarrow \dots < V_t$$

\overline{V}_t is then the colimit of this diagram, i.e. the vector space and maps so that

$$\begin{array}{ccccccc} \dots & \longrightarrow & V_{s_1} & \longrightarrow & \dots & \longrightarrow & V_{s_2} & \longrightarrow & \dots & < & V_t \\ & & & & & & \searrow & & \searrow & & \downarrow \\ & & & & & & & & & & \overline{V}_t \end{array}$$

is the initial such vector space that makes the diagram commute. If we choose some u with $t < u$, then there is a collection of maps $\{V_{t_i} \rightarrow \overline{V}_u \mid t_i < u\}$. In particular, this includes maps from $\{V_{s_i} \rightarrow \overline{V}_u \mid s_i < t\}$! Our diagram looks like the following:

$$\begin{array}{ccccccccccc} \dots & \longrightarrow & V_{s_1} & \longrightarrow & \dots & \longrightarrow & V_{s_2} & \longrightarrow & \dots & \longrightarrow & V_{t_1} & \longrightarrow & \dots & \longrightarrow & V_{t_2} & \longrightarrow & \dots \\ & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & & \searrow \\ & & & & & & & & & & \overline{V}_t & \xrightarrow{\quad \overline{v}_t^u \quad} & & & & & \overline{V}_u \end{array}$$

By the universal property these maps $\{V_{s_i} \rightarrow \overline{V}_u\}$ must factor uniquely through \overline{V}_t , and we define the structure map \overline{v}_t^u to be this map.

Example 8.7. If \mathbb{V} is any interval with endpoints (p, q) , then $\overline{\mathbb{V}} = \mathbb{I}(p, q]$. Looking at \overline{V}_p , this is the colimit of the diagram of maps $\{V_s \mid s < p\}$, all of which are 0. The space \overline{V}_q is the colimit of $\{V_s \mid s < q\}$, which are constant k approaching q .

For any ob-morphism $\phi^\circ : \mathbb{V} \dashrightarrow \mathbb{W}$, we can generate maps $\overline{V}_t \rightarrow \overline{W}_t$ as suggested by the following diagram:

$$\begin{array}{ccccccc} \dots & \longrightarrow & V_{s_1} & \longrightarrow & \dots & \longrightarrow & V_{s_2} & \longrightarrow & \dots & < & V_t \\ & \searrow & & \searrow & & \searrow & & \searrow & & & \\ \dots & \longrightarrow & W_{s_1} & \longrightarrow & \dots & \longrightarrow & W_{s_2} & \longrightarrow & \dots & < & W_t \\ & & & & & & \searrow & & \searrow & & \downarrow \\ & & & & & & & & & & \overline{W}_t \end{array}$$

By the universal property, the vertical paths to \overline{W}_t must factor uniquely through a map $\overline{V}_t \rightarrow \overline{W}_t$, which is the map we require. We can therefore think of ‘bar’ as a functor $\mathbf{Ob} \rightarrow \mathbf{Pers}$.

The universal property also generates, for each t , a map $\overline{V}_t \rightarrow V_t$. It is straightforward but tedious to verify that this is a module homomorphism $n^\mathbb{V} : \overline{\mathbb{V}} \rightarrow \mathbb{V}$.

Proposition 8.8. *Given any persistence module \mathbb{V} , the morphism $n^\mathbb{V} : \overline{\mathbb{V}} \rightarrow \mathbb{V}$ is a weak isomorphism.*

Proof. As $n^\mathbb{V}$ is a module morphism, for any $s < t$ we have a commutative diagram

$$\begin{array}{ccc} \overline{V}_s & \xrightarrow{\overline{v}_s^t} & \overline{V}_t \\ n_s^\mathbb{V} \downarrow & \nearrow & \downarrow n_t^\mathbb{V} \\ V_s & \xrightarrow{v_s^t} & V_t \end{array}$$

where the diagonal map is given by the universal property of \overline{V}_t . The bottom triangle commutes by the definition of $n^\mathbb{V}$ and the top triangle commutes by the definition of \overline{v}_s^t .

Therefore, the map \overline{v}_s^t sends $\ker(n_t^\mathbb{V})$ to zero, which means $\ker(n^\mathbb{V})$ is ephemeral. Similarly, v_s^t sends V_s to $\text{im}(n_t^\mathbb{V})$, so therefore to zero in $\text{coker}(n_t^\mathbb{V})$, and $\text{coker}(n^\mathbb{V})$ is ephemeral. \square

With the groundwork in place, we can now prove the universal property of π .

Theorem 8.9. *Let $F : \mathbf{Pers} \rightarrow \mathcal{C}$ be a functor that takes weak isomorphisms to actual isomorphisms. Then there is a unique functor $G : \mathbf{Ob} \rightarrow \mathcal{C}$ such that $F = G\pi$.*

Proof. Because \mathbf{Ob} has the same objects as \mathbf{Pers} , the functor G is uniquely determined on objects. We now consider morphisms.

First consider the diagram:

$$\begin{array}{ccc} \overline{\mathbb{V}} & \xrightarrow{\overline{\Phi}} & \overline{\mathbb{W}} \\ n^\mathbb{V} \downarrow & & \downarrow n^\mathbb{W} \\ \mathbb{V} & \xrightarrow{\Phi} & \mathbb{W} \end{array}$$

We wish to show that this diagram commutes. This is equivalent to showing that the diagram of vector spaces

$$\begin{array}{ccc} \overline{V}_t & \xrightarrow{\overline{\phi}_t} & \overline{W}_t \\ n_t^\mathbb{V} \downarrow & & \downarrow n_t^\mathbb{W} \\ V_t & \xrightarrow{\phi_t} & W_t \end{array}$$

50

commutes for all t . Consider any space V_s with $s < t$. As Φ is a module map, we have that the following compositions of maps are equal:

$$V_s \xrightarrow{\phi_s} W_s \xrightarrow{w_s^t} W_t = V_s \xrightarrow{v_s^t} V_t \xrightarrow{\phi_t} W_t$$

Now by the universal property of colimits, the maps v_s^t and w_s^t factor through \overline{V}_t and \overline{W}_t respectively:

$$V_s \xrightarrow{\phi_s} W_s \rightarrow \overline{W}_t \xrightarrow{n_t^{\mathbb{W}}} W_t = V_s \rightarrow \overline{V}_t \xrightarrow{n_t^{\mathbb{V}}} V_t \xrightarrow{\phi_t} W_t$$

Now, by the definition of $\overline{\Phi}$, we can rewrite the left hand side to:

$$V_s \rightarrow \overline{V}_t \xrightarrow{\overline{\phi}_t} \overline{W}_t \xrightarrow{n_t^{\mathbb{W}}} W_t = V_s \rightarrow \overline{V}_t \xrightarrow{n_t^{\mathbb{V}}} V_t \xrightarrow{\phi_t} W_t$$

This holds for *every* $s < t$, so by the uniqueness of the map $\overline{V}_t \rightarrow W_t$, we must have that $n_t^{\mathbb{W}} \overline{\phi}_t = \phi_t n_t^{\mathbb{V}}$ as required.

Now consider an ob-morphism $\Phi^\circ : \mathbb{V} \dashrightarrow \mathbb{W}$. We have a commutative diagram in **Ob**:

$$\begin{array}{ccc} \overline{\mathbb{V}} & \xrightarrow{\pi(\overline{\Phi})} & \overline{\mathbb{W}} \\ \pi(n^{\mathbb{V}}) \downarrow & & \downarrow \pi(n^{\mathbb{W}}) \\ \mathbb{V} & \xrightarrow{\Phi^\circ} & \mathbb{W} \end{array}$$

Recall that both $n^{\mathbb{V}}$ and $n^{\mathbb{W}}$ are weak isomorphisms, so $\pi(n^{\mathbb{V}})$ and $\pi(n^{\mathbb{W}})$ are both isomorphisms. If there is a functor G with $F = G\pi$, applying G to the diagram above forces us to define

$$G(\Phi^\circ) = F(n^{\mathbb{W}})F(\overline{\Phi})F(n^{\mathbb{V}})^{-1}$$

for every ob-morphism Φ° . This is a valid functor, as if $\Psi^\circ : \mathbb{W} \dashrightarrow \mathbb{U}$ then:

$$\begin{aligned} G(\Psi^\circ \Phi^\circ) &= F(n^{\mathbb{U}})F(\overline{\Psi^\circ \Phi^\circ})F(n^{\mathbb{V}})^{-1} \\ &= F(n^{\mathbb{U}})F(\overline{\Psi^\circ} \overline{\Phi^\circ})F(n^{\mathbb{V}})^{-1} \\ &= F(n^{\mathbb{U}})F(\overline{\Psi^\circ})F(\overline{\Phi^\circ})F(n^{\mathbb{V}})^{-1} \\ &= F(n^{\mathbb{U}})F(\overline{\Psi^\circ})F(n^{\mathbb{W}})^{-1}F(n^{\mathbb{W}})F(\overline{\Phi^\circ})F(n^{\mathbb{V}})^{-1} \\ &= G(\Psi^\circ)G(\Phi^\circ). \end{aligned}$$

Now suppose $\Phi^\circ = \pi(\Phi)$. Returning to the first diagram we considered, we have that:

$$\begin{array}{ccc} \overline{\mathbb{V}} & \xrightarrow{\overline{\Phi}} & \overline{\mathbb{W}} \\ n^{\mathbb{V}} \downarrow & & \downarrow n^{\mathbb{W}} \\ \mathbb{V} & \xrightarrow{\Phi} & \mathbb{W} \end{array}$$

commutes. Applying F ,

$$F(\Phi)F(n^{\mathbb{V}}) = F(n^{\mathbb{W}})F(\overline{\Phi}),$$

or in other words,

$$\begin{aligned} F(\Phi) &= F(n^{\mathbb{W}})F(\overline{\Phi})F(n^{\mathbb{V}})^{-1} \\ &= G(\Phi^\circ) = G(\pi(\Phi)). \end{aligned}$$

Therefore $F = G\pi$ on morphisms and π satisfies the universal property. \square

The above construction is an example of *localisation*. There is a well developed theory for quotient categories \mathcal{C}/\mathcal{A} when the category \mathcal{C} is an *abelian category*, and \mathcal{A} is a full subcategory satisfying some technical conditions. Here we give a quick overview of the concepts involved.

An abelian category is one that shares many useful properties with \mathbf{Ab} , the category of abelian groups. In particular, in an abelian category every morphism has a kernel and cokernel, and each pair of objects has a product and coproduct. If \mathcal{C} is a small category and \mathcal{D} is an abelian category, any functor category $[\mathcal{C}, \mathcal{D}]$ is also abelian, with the relevant operations all performed pointwise.

Because $\mathbf{Vect}_{\mathbf{k}}$ is abelian, $\mathbf{Pers} = [\mathbf{R}, \mathbf{Vect}_{\mathbf{k}}]$ is also abelian. This retroactively justifies the algebraic manipulations involving kernels and cokernels that have appeared throughout previous sections of this work.

The theory of localisation shows that when these conditions are satisfied, not only does the quotient \mathcal{C}/\mathcal{A} always exist, but it is also itself an abelian category. Therefore, we do not lose any algebraic power when working in \mathbf{Ob} .

A full treatment of localisation can be found in Popescu [18].

8.2 Structure and stability in \mathbf{Ob}

We begin by quoting some decomposition results for persistence modules over \mathbf{R} , as given by Crawley-Boevey [6, 11].

Definition 8.10. A sequence *stabilises* if beyond some point the terms in the sequence are all equal. Let \mathbb{V} be a persistence module over \mathbf{R} .

- i) \mathbb{V} has the *descending chain condition on images* if for all $t \geq s_1 > s_2 > \dots$, the chain

$$V_t \supseteq \text{im}(v_{s_1}^t) \supseteq \text{im}(v_{s_2}^t) \supseteq \dots$$

stabilises.

- ii) Given $s \leq t$, the space V_s has the *descending chain condition on t -bounded kernels* if for all $t < \dots < r_2 < r_1$ the chain

$$V_s \supseteq \ker(v_s^{r_1}) \supseteq \ker(v_s^{r_2}) \supseteq \dots$$

stabilises.

- iii) \mathbb{V} has the *descending chain condition on sufficient bounded kernels* if for all t and nonzero $v \in V_t$, there exists $s \leq t$ such that $v \in \text{im}(v_s^t)$ and V_s has the descending chain condition on t -bounded kernels.

Note that any q-tame persistence module must have the descending chain condition on images. For all $s < t$, $\text{rank } v_s^t$ is always finite, so $\text{im}(v_s^t)$ must be finite dimensional. The well ordering principle implies that the chain must then stabilise. For a similar reason, in a q-tame persistence module, V_s has the descending chain condition on t -bounded kernels for every $s < t$.

However, q-tame persistence modules need not satisfy the descending chain condition on sufficient bounded kernels. Consider the module

$$\mathbb{V} = \bigoplus_{n=1}^{\infty} \mathbb{I}[0, \frac{1}{n}]$$

seen earlier. Because the spaces with $t < 0$ are all 0, any nonzero $v \in V_0$ is not in the image of any v_s^0 with $s < 0$. In the definition we are therefore forced to take $s = 0$. But V_0 does not satisfy the descending chain condition on 0-bounded kernels. To see this, take $r_n = \frac{1}{n}$, then:

$$V_0 \supseteq \ker(v_0^1) \supseteq \ker(v_0^{\frac{1}{2}}) \supseteq \ker(v_0^{\frac{1}{3}}) \supseteq \dots$$

does not stabilise.

When the conditions do hold, we can use the following theorem.

Theorem 8.11 (Crawley-Boevey [11]). *Any persistence module with the descending chain condition on images and on sufficient bounded kernels is a direct sum of interval modules.*

The strategy now is to show that every q-tame persistence module is weakly isomorphic to one that satisfies the above conditions. When we pass through to **Ob**, this becomes a true isomorphism.

Definition 8.12. The *radical* of a persistence module \mathbb{V} is the submodule $\text{rad } \mathbb{V}$ defined by:

$$(\text{rad } \mathbb{V})_t = \sum_{s < t} \text{im}(v_s^t).$$

By construction, $\text{rad } \mathbb{V}$ satisfies the descending chain condition on sufficient bounded kernels, as we have eliminated the $v \in V_t$ such that $v \notin \text{im}(v_s^t)$ for any $s < t$. Therefore $\text{rad } \mathbb{V}$ satisfies the conditions of the above theorem and decomposes as a direct sum of interval modules.

Proposition 8.13. *The natural inclusion $I = (i_t) : \text{rad } \mathbb{V} \rightarrow \mathbb{V}$ is a weak isomorphism.*

Proof. Firstly, $\ker I = 0$ as I is an inclusion, so $\ker I$ is ephemeral. Let $v \notin \text{im}(v_s^t)$ for any $s < t$, so that $v + \text{im}(i_t) \in \text{coker}(i_t)$. For any $u > t$, $v_t^u(v) \in \text{im}(i_t)$ so v is mapped to 0 in $\text{coker}(i_u)$. We conclude that $\text{coker } I$ is also ephemeral. \square

Therefore, when we pass to **Ob** we have the following structure theorem.

Theorem 8.14. *Any q -tame persistence module is isomorphic in **Ob** to a direct sum of interval modules.*

We can transplant the stability result of early sections into **Ob**, in view of the following propositions.

Proposition 8.15. *The interleaving distance between \mathbb{U} and \mathbb{V} is invariant under ob-isomorphisms.*

Proof. Suppose \mathbb{U} and \mathbb{U}' are ob-isomorphic. Then $d_i(\mathbb{U}, \mathbb{U}') = 0$, as we can take the ob-isomorphism maps $\{\phi_t^{t+\varepsilon}\}$ as an interleaving for all $\varepsilon > 0$. From the triangle inequality it follows that $d_i(\mathbb{U}, \mathbb{V}) = d_i(\mathbb{U}', \mathbb{V})$. \square

Proposition 8.16. *The undecorated persistence diagram is invariant under ob-isomorphisms.*

Proof. Let $\Phi : \mathbb{V} \dashrightarrow \mathbb{W}$ be an ob-isomorphism with inverse $\Psi : \mathbb{W} \dashrightarrow \mathbb{V}$. Recall that the undecorated diagram is given by the multiplicity function

$$m(p, q) = \min\{\mu(R) \mid R \in \text{Rect}(\mathcal{D}), (p, q) \in R\}.$$

We show that $m_{\mathbb{V}}(p, q) = m_{\mathbb{W}}(p, q)$ for all (p, q) . Let $R = [a, b] \times [c, d]$ be the rectangle that attains the minimum in the definition of $m_{\mathbb{V}}(p, q)$ and select a rectangle $R' = [a', b'] \times [c', d']$ lying inside R so $(p, q) \in R' \subset R$. Following the same argument as the box lemma, we have:

$$\begin{aligned} \mu_{\mathbb{V}}(R) &= \langle \circ_A \text{-----} \bullet_B \text{---} \bullet_C \text{-----} \circ_D \mid \mathbb{V} \rangle \\ &\geq \langle \circ_A \text{---} \circ_a \text{---} \bullet_b \text{---} \bullet_B \text{---} \bullet_C \text{---} \bullet_c \text{---} \circ_d \text{---} \circ_D \mid \mathbb{W} \rangle \\ &= \langle \text{-----} \circ_a \text{---} \bullet_b \text{-----} \bullet_c \text{---} \circ_d \text{-----} \mid \mathbb{U} \rangle \\ &= \mu_{\mathbb{U}}(R') \end{aligned}$$

and therefore $m_{\mathbb{W}}(p, q) \leq m_{\mathbb{V}}(p, q)$. The reverse inequality follows by symmetry, so the two multisets must be equal. \square

Therefore, the stability theorem remains true in **Ob**.

Proposition 8.17. *Let \mathbb{V} and \mathbb{W} be q -tame persistence modules. Then \mathbb{V} and \mathbb{W} are ob-isomorphic if and only if their undecorated persistence diagrams are equal.*

Proof. We saw the forwards direction above. For the reverse direction, we know that \mathbb{V} and \mathbb{W} are ob-isomorphic to direct sums of interval modules. We can choose the intervals to be open and nonempty. These intervals are determined by the persistence diagrams of \mathbb{V} and \mathbb{W} . These diagrams are equal, so the two direct sum decompositions are isomorphic. \square

The above proposition shows we have finally come full circle. Just as there is a bijection in standard persistent homology between persistence modules and barcodes, there is a bijection between isomorphism classes of q -tame persistence modules in **Ob** and locally finite persistence diagrams.

References

- [1] Maurice Auslander. “Representation theory of artin algebras II”. In: *Communications in algebra* 1.4 (1974), pp. 269–310.
- [2] Gorô Azumaya et al. “Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt’s theorem”. In: *Nagoya Mathematical Journal* 1 (1950), pp. 117–124.
- [3] M Barot. *Representations of quivers*. URL: http://www.matem.unam.mx/barot/articles/notes_ictp.pdf.
- [4] Peter Bubenik and Jonathan A Scott. “Categorification of persistent homology”. In: *Discrete & Computational Geometry* 51.3 (2014), pp. 600–627.
- [5] Peter Bubenik, Vin de Silva, and Jonathan Scott. “Metrics for generalized persistence modules”. In: *arXiv preprint arXiv:1312.3829* (2013).
- [6] Frédéric Chazal, William Crawley-Boevey, and Vin De Silva. “The observable structure of persistence modules”. In: *arXiv preprint arXiv:1405.5644* (2014).
- [7] Frédéric Chazal, Vin De Silva, and Steve Oudot. “Persistence stability for geometric complexes”. In: *Geometriae Dedicata* (2013), pp. 1–22.
- [8] Frédéric Chazal et al. “Proximity of persistence modules and their diagrams”. In: *Proceedings of the twenty-fifth annual symposium on Computational geometry*. ACM. 2009, pp. 237–246.
- [9] Frédéric Chazal et al. “The structure and stability of persistence modules”. In: *arXiv preprint arXiv:1207.3674* (2012).
- [10] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. “Stability of persistence diagrams”. In: *Discrete & Computational Geometry* 37.1 (2007), pp. 103–120.
- [11] William Crawley-Boevey. “Decomposition of pointwise finite-dimensional persistence modules”. In: *arXiv preprint arXiv:1210.0819* (2012).
- [12] Vin De Silva, Dmitriy Morozov, and Mikael Vejdemo-Johansson. “Dualities in persistent (co)homology”. In: *Inverse Problems* 27.12 (2011), p. 124003.
- [13] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2003. ISBN: 9780471433347.
- [14] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. “Topological persistence and simplification”. In: *Discrete and Computational Geometry* 28.4 (2002), pp. 511–533.
- [15] David Eisenbud et al. *Commutative algebra with a view toward algebraic geometry*. Vol. 27. Springer New York, 1995.
- [16] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.
- [17] Saunders Mac Lane. *Categories for the working mathematician*. Vol. 5. Springer, 1998.

- [18] Nicolae Popescu. *Abelian categories with applications to rings and modules*. Vol. 3. Academic Press London, 1973.
- [19] Mikael Vejdemo-Johansson. “Sketches of a platypus: persistent homology and its algebraic foundations”. In: *arXiv preprint arXiv:1212.5398* (2012).
- [20] Cary Webb. “Decomposition of graded modules”. In: *Proceedings of the American Mathematical Society* 94.4 (1985), pp. 565–571.
- [21] Afra Zomorodian and Gunnar Carlsson. “Computing persistent homology”. In: *Discrete & Computational Geometry* 33.2 (2005), pp. 249–274.

A Implementation in Haskell

The following Haskell code implements the algorithm given by Zomorodian and Carlsson [21] for computing the persistence homology of a filtration of a simplicial complex.

Listing 1: Algebra.hs

```
module Algebra where

import Prelude hiding ((*), (/), (-), (+), recip, negate)

infixl 6 +, -

class AdditiveGroup f where
  zero :: f
  (+) :: f -> f -> f
  (-) :: f -> f -> f
  negate :: f -> f

  a - b = a + negate b
  negate a = zero - a

infixl 7 *, /

class (AdditiveGroup f) => Field f where
  one :: f
  (*) :: f -> f -> f
  (/) :: f -> f -> f
  recip :: f -> f

  a / b = a * recip b
  recip a = one / a

  fromInteger :: Integer -> f
```

Listing 2: Rational.hs

```
{-# LANGUAGE FlexibleInstances , TypeSynonymInstances #-}

module Rational where

import Algebra

instance AdditiveGroup Rational where
  zero = 0
  (+) = (Prelude.+)
  negate = Prelude.negate

instance Field Rational where
  one = 1
  (*) = (Prelude.*)
  recip = Prelude.recip
  fromInteger = fromIntegral
```

Listing 3: Z2.hs

```

module Z2 where

import Prelude hiding ((*), (/), (-), (+), recip, negate)
import Algebra

data Z2 = Zero | One deriving (Show, Eq)

instance AdditiveGroup Z2 where
    zero = Zero
    Zero + Zero = Zero
    Zero + One  = One
    One  + Zero = One
    One  + One  = Zero
    negate = id

instance Field Z2 where
    one = One
    Zero * _ = Zero
    One  * a = a
    recip = id

    fromInteger n | even n    = Zero
                  | otherwise = One

```

Listing 4: Simplex.hs

```

module Simplex where

data Simplex a = Simplex { dimension :: Int, vertices :: [a] }
    deriving (Eq, Ord)

simplex :: [a] -> Simplex a
simplex v = Simplex (length v - 1) v

faces :: Simplex a -> [Simplex a]
faces (Simplex 0 v) = []
faces (Simplex d v) = map (Simplex (d-1)) (deletions v)
    where deletions [] = []
          deletions (x:xs) = xs : map (x:) (deletions xs)

--instance Show (Simplex Char) where
--    show (Simplex d v) = v

instance (Show a) => Show (Simplex a) where
    show (Simplex d v) = concatMap show v

```

Listing 5: Chain.hs

```

{-# LANGUAGE FlexibleInstances , TypeSynonymInstances #-}

module Chain where

import Prelude hiding ((*), (+), (-), (>), fromInteger, negate)
import Simplex

```

```

import Algebra
import qualified Data.Map as M

type Chain f b = M.Map b f

coefficientFor :: (Ord b) => Chain f b -> b -> f
coefficientFor c b = c M.! b

simplexBoundary :: (Field f, Ord a) => Simplex a -> Chain f
  (Simplex a)
simplexBoundary s = M.fromList $ zip (faces s) (map fromInteger
  (cycle [1, -1]))

elements :: Chain f b -> [b]
elements = M.keys

clean :: (Field f, Eq f, Ord b) => Chain f b -> Chain f b
clean = M.filter (/= zero)

empty :: (Field f, Eq f) => Chain f b -> Bool
empty c = all (== zero) (M.elems c)

instance (Field f, Eq f, Ord b) => AdditiveGroup (Chain f b) where
  zero = M.empty
  c + c' = clean $ M.unionWith (+) c c'
  negate = M.map negate

infixl 7 .*
(.*) :: (Field f, Ord b) => f -> Chain f b -> Chain f b
x .* c = M.map (x*) c

```

Listing 6: Barcode.hs

```

module Barcode where

data WithInfinity a = Finite a | Infinity deriving (Eq, Ord)

data Interval a = Interval { start :: a, end :: WithInfinity a }
  deriving (Eq, Ord)

type Barcode a = [Interval a]

instance (Show a) => Show (WithInfinity a) where
  show (Finite a) = show a
  show Infinity = "Inf"

instance (Show a) => Show (Interval a) where
  show (Interval start end) = show start ++ "└─>┘" ++ show end

finiteInterval :: a -> a -> Interval a
finiteInterval start end = Interval start (Finite end)

infiniteInterval :: a -> Interval a
infiniteInterval start = Interval start Infinity

```

Listing 7: PersistentHomology.hs

```
{-# LANGUAGE TemplateHaskell, RankNTypes #-}

module PersistentHomology where

import Prelude hiding ((-), (/), recip)
import Language.Haskell.TH
import Data.List
import Data.Function
import Data.Maybe
import Control.Applicative
import Control.Monad.State
import qualified Data.Map as M
import qualified Data.IntMap as IM
import Control.Lens

import Algebra
import Simplex (Simplex)
import qualified Simplex
import Chain (Chain, (.*))
import qualified Chain
import Barcode

import Debug.Trace

data Slot f a = Slot { _slotSimplex :: Simplex a,
                      _chain :: Maybe (Chain f (Simplex a)),
                      _marked :: Bool } deriving (Show)

makeLenses ''Slot

type Table f a = IM.IntMap (Slot f a)

data PersistentHomology f a = PH { _table :: Table f a,
                                   _indexMap :: M.Map (Simplex a)
                                   Int,
                                   _simplexCount :: Int,
                                   _barcode :: Barcode Int }

makeLenses ''PersistentHomology

type PersistentHomologyS f a = State (PersistentHomology f a)

simplexIndex :: (Ord a) => PersistentHomology f a -> Simplex a ->
Int
simplexIndex ph s = (ph^.indexMap) M.! s

indexSlot :: (Ord a) => Int -> Lens' (PersistentHomology f a) (Slot
f a)
indexSlot i = singular $ table . ix i

slot :: (Ord a) => Simplex a -> Lens' (PersistentHomology f a)
(Slot f a)
slot s f ph = indexSlot (simplexIndex ph s) f ph
```



```

eliminatePivotRows :: (Field f, Eq f, Ord a) => PersistentHomology
  f a -> Chain f (Simplex a) -> (Chain f (Simplex a), Int)
eliminatePivotRows ph c | null markedElements = (zero, 0)
                        | isNothing pivotChain = (c, pivotIndex)
                        | isJust pivotChain = eliminatePivotRows ph
                          reduced
where markedElements = filter (\s -> ph ^. slot s . marked)
  (Chain.elements c)
  pivotIndex = maximum $ map (simplexIndex ph)
    markedElements
  pivotSlot = ph ^. indexSlot pivotIndex
  pivotSimplex = pivotSlot ^. slotSimplex
  pivotChain = pivotSlot ^. chain
  pivotCoefficient = Chain.coefficientFor (fromJust
    pivotChain) pivotSimplex
  chainCoefficient = Chain.coefficientFor c pivotSimplex
  reduced = c - (chainCoefficient / pivotCoefficient) .*
    fromJust pivotChain

markSimplex :: (Ord a) => Simplex a -> PersistentHomologyS f a ()
markSimplex s = slot s . marked .= True

storeSimplexBoundary :: (Ord a) => Int -> Chain f (Simplex a) ->
  PersistentHomologyS f a ()
storeSimplexBoundary i c = indexSlot i . chain .= Just c

blankSlot :: Simplex a -> Slot f a
blankSlot s = Slot s Nothing False

startPH :: PersistentHomology f a
startPH = PH { _table = IM.empty,
               _indexMap = M.empty,
               _simplexCount = 0,
               _barcode = [] }

addSimplex :: (Field f, Eq f, Ord a) => Simplex a ->
  PersistentHomologyS f a ()
addSimplex s = do
  ph <- get
  let index = ph ^. simplexCount
  table.at index ?= blankSlot s
  indexMap.at s ?= index
  simplexCount += 1

  ph <- get
  let dimension = Simplex.dimension s
  boundary = Chain.simplexBoundary s
  (reduced, pivotIndex) = eliminatePivotRows ph boundary

  if Chain.empty reduced then
    markSimplex s
  else do
    storeSimplexBoundary pivotIndex reduced

```

```

barcode %= cons (finiteInterval pivotIndex index)

runPH :: (Field f, Eq f, Ord a) => [Simplex a] ->
    PersistentHomology f a
runPH xs = execState (mapM addSimplex xs) startPH

findInfiniteIntervals :: PersistentHomology f a -> Barcode Int
findInfiniteIntervals ph = map (\(index, sl) -> infiniteInterval
    index) unfilled
    where assocs = IM.assocs (ph^.table)
        unfilled = filter (\(index, sl) -> sl^.marked &&
            isNothing (sl^.chain) ) assocs

finalBarcode :: PersistentHomology f a -> Barcode Int
finalBarcode ph = sort $ ph^.barcode ++ findInfiniteIntervals ph

```

Listing 8: Main.hs

```

module Main where

import PersistentHomology
import Rational
import Z2
import Simplex
import Chain

filtration :: [Simplex Char]
filtration = map simplex ["a", "b", "ab", "c", "ac", "bc"]

main :: IO ()
main = print (finalBarcode (runPH filtration :: PersistentHomology
    Z2 Char))

```