Cover page

Exam information
NMAA60002E - Mathematics Thesis 30 ECTS,
Department of Mathematical Sciences - Kontrakt:108830
(Brendan Michael Collins)

Handed in by
Brendan Michael Collins
dbz917@alumni.ku.dk

Exam administrators
Eksamensteam, SCIENCE Uddannelse
eksamen@science.ku.dk

Assessors
Jesper Michael Møller
Examiner
moller@math.ku.dk
+4535320691

Iver Mølgaard Ottosen
Co-examiner
ottosen@math.aau.dk

Hand-in information

Titel: Persistent Homology and Noise
Må besvarelsen gøres til genstand for udlån: Yes

FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Persistent Homology and Noise

Author
Brendan Collins

Supervisor
Prof. Jesper Michael Møller

Thesis for the Master degree in Mathematics.
Department of Mathematical Sciences,
University of Copenhagen

September 2, 2019

Abstract

This thesis aims to provide an overview of the theory of persistent homology, and an insight into how it can be applied. Persistent homology allows one to analyse a data set by using techniques of algebraic topology to reveal the shape of the data. It has found applications in many areas from neuroscience to machine learning.

We cover the theory of one dimensional persistence, which was developed in the mid 2000's and still underpins the current literature. We then move on to the case of multidimensional persistence, which has proven to be a more complicated topic and so has garnished more attention and has been approached in many different ways. Here we primarily consider noise systems, introduced in [17]. These are used to develop pseudometrics, and in turn stable invariants on the space of tame multidimensional persistence modules. We introduce the idea of Serre noise systems, a particular type of noise system which are of theoretical interest. We then demonstrate how noise systems are of practical interest by using them to arrive at a stable invariant and applying this invariant to analyse a dataset.

Acknowledgements

I would like to thank my supervisor Jesper Michael Møller for introducing me to persistent homology and for all of his time, effort and help in preparing this thesis. I am also thankful to Wojciech Chacholski for all his help in understanding his work and for making software available to me which I used in this thesis, as well as making me feel welcome during my visit to KTH. Finally I would like to thank all my family and friends for their support and welcome distractions.

Contents

Chapter 1

Introduction

Topological data analysis (TDA), and in particular persistent homology, is a rapidly growing field that has garnished a lot of recent attention. The goal is to use tools from algebraic topology to be able to describe the shape of a point cloud, which for practical purposes often comes from some data set. Traditional invariants of a space, such as the homology and homotopy groups, are not useful for distinguishing point clouds since these groups are trivial on any such space. This is where persistent homology comes in. The primary idea is to look at the changing homology of a simplicial complex (such as the Vietoris-Rips complex of a point cloud) as some parameter ε changes. This changing homology tells us something about the underlying shape of the data. In particular, generators of the homology group of the simplicial complex which persist for a large range of ε values are generaly considered to correspond to the features of the data. On the other hand, those generators which die off quickly reflect noise in the data. While there have been challenges to this viewpoint, and indeed particualr examples where this is not the correct interpretation [26], it provides a good overview of the motivation behind the methods of persistent homology. There have been many successful attempts to apply persistent homology to real world data [27],[28],[29]. The most well known example is probably the discovery of a specific form of breast cancer, for which the survival rates are very good [25].

The organisation of this thesis is as follows: In section 3 we discuss one-dimensional persistence, largely folowing [1]. Note that the dimension here does not refer to the dimension of the space over which we work, but rather the number of parameters in the aforementioned simplicial complex. We use the one-dimensional case to introduce ideas such as persistence modules, which are functors $F : \mathbb{R} \to \text{Vect}_k$, and a pseudometric on these modules, called the interleaving distance. We demonstrate that decomposable persistence modules admit a complete invariant called the barcode or persistence diagram. We then provide a proof the isometry theorem, which shows that the barcode is stable, in the sense that a pair of persistence modules which are close to one another in the interleaving distance have similar barcodes. In section 4 we investigate generalized persistence modules. These are functors $F : P \to C$, where $(P, \leq)$ is a poset, and C is an arbitrary category. We work with a generalization of the interleaving distance. Much of this chapter serve as a primer for chapter 5, which covers multidimensional persistence. Here we use the interleaving distance and superlinear families defined in chapter 4 extensively. We will see that the multidimensional case is more complicated than the one dimesional case. Largely, this is because we have no multidimensional equivalent of the barcode; in fact it is impossible to define a complete invariant on multidimensional persistence modules [15]. As a result, much of the work is focused on coming up with good incomplete invariants which reflect the properties of a data set that we are interested in. In pursuit of this goal, our attention turns to developing pseudometrics on the space of multidimensional persistence modules. We investigate noise systems, introduced in [17], as a systematic means of concocting more pseudometrics. Finally, we investigate the connection between these noise systems and Serre categories and Serre localizations. Chapter 6 contains two applications of the tools discussed in the thesis, primarily using ideas from chapter 5. The first is a "textbook" example, based on an exercise in a course taught by Wojciech Chacholski at KTH, Stockholm on TDA. We are able to create a signature for 3 shapes in the plane - a square, triangle and circle. These signatures take the form of positive valued functions, and plotting these functions shows them to be visibly distinct. We do the same with another circle with some points added to the center to serve as noise. This noisy circle has a signature which is indistinguishable from that of the original circle. The second application analyses a data set of heart disease patients, and uses a similar signature to attampt to distinguish healthy patients from those with heart disease. The thesis contains some novel contributions, namely the proof of Theorem 5.6.5, as well as the discussion in section 5.8.

Chapter 2

Preliminaries and Background

Given a poset P, denote by $P \cup \infty$ the same poset with a maximum element ∞ attached. We denote by $\mathbb{R}^r_{\geq 0}$ the r-tuples of real numbers in which each entry is greater than or equal to 0. Similarly for $\mathbb{Q}^r_{\geq 0}, \mathbb{N}^r_{\geq 0}$.
A **multiset** is a set X in which elements can occur with multiplicity greater than 1. A **multisubset** of X is a function $f : X \to \mathbb{N}$. The value $f(x)$ is called the multiplicity of $x \in T$. The function f is called finite if it has finite support, in which case we can define the rank of f to be the sum $\sum_{x\in X} f(x)$. Let $\mathrm{Mult}(X)$ denote all the multisubsets of X.
Given a small category I and a field k, for an object $i \in I$, define the functor

$$k(i, -) : I \to \mathrm{Vect}_k$$

to be the one which maps $j \in I$ to the k-vector space spanned by $\mathrm{Hom}_I(i, j)$. In particular, if I is a poset then $k(i, j) \in \{0, k\}$.
In the case that I is a poset, we can define an interval $[a, b) \subset I$ by

$$[a, b) = \{i \in I | a \leq i \leq b \text{ and } i \neq b\}.$$

Let $[a, b) \subset I$ be a half-open interval. Then we define an **interval functor** $[a, b) : I \to R - mod$, for some ring R, by

$$[a, b)_t = \begin{cases} R & \text{if } t \in [a, b) \\ 0 & \text{otherwise} \end{cases}$$

and

$$[a, b)(s \leq t) = \begin{cases} 1 & \text{if } s, t \in J \\ 0 & \text{otherwise} \end{cases}$$

It will always be clear from context whether $[a, b)$ refers to an interval or an interval functor.
We denote by $U_i : I \to R$-mod, the unique functor given by

$$U_i(j) = \begin{cases} R & i = j \\ 0 & \text{otherwise} \end{cases}$$

and $U_i(i \leq i) = \mathrm{id}$, and U_i is 0 on all other maps.
Let $\{V_i\}_{i\in I}$ be a family of k-vector spaces for some field k. Then functors of the form $F = \bigoplus_{i\in I} k(i, -) \otimes V_i$ are called **free**. We call the vector spaces V_i the **components** of F. A free functor F is said to be of **finite type** if each V_i is finite dimensional. The **support** of $F = \bigoplus_{i\in I} k(i, -) \otimes V_i$ is the set

$$\mathrm{supp}(F) = \{i \in I | V_i \neq 0\}.$$

and F is said to be of finite support if this set is finite. If F is of finite type and it has finite support, then we can define the rank of F as $\mathrm{rank}(F) = \sum_{i\in I} \dim V_i$. The 0-Betti diagram of F is defined to be

the multiset of objects in I with multiplicities given by $\beta_0 F(i) = \dim_k V_i$.
A morphism $\phi : X \to Y$ in a category C is called **minimal** if any morphism $f : X \to X$ satisfying $\phi = \phi \circ f$ is an isomorphism. A natural transformation $\phi : \tilde{F} \to F$ of functors in $\mathrm{Fun}(I, C)$ is called a **minimal cover** of F if $\tilde{F}$ is free and ϕ is both minimal and an epimorphism. All minimal covers of a given functor are isomorphic [17]. In the case that $F \in \mathrm{Fun}(I, \mathrm{Vect}_k)$, then if $\tilde{F}$ is a minimal cover of F and $\tilde{F}$ is of finite type and finite rank, then so is F. We define the rank of F to be the rank of $\tilde{F}$. Similarly the support and 0-Betti diagram of F are defined to be those of $\tilde{F}$. Note that none of the above depends on the choice of minimal cover for F.
Given a functor $F : I \to \mathrm{Vect}_k$, there is a bijection between the set of natural transformations $k(i, -) \to F$ and the vector space $F(i)$ given by

$$\mathrm{Nat}(k(i,-), F) \to F(i), \quad f \mapsto f_i(\mathrm{id}_k)$$

Therefore any $f \in F(i)$ yields a unique natural transformation $f : k(i, -) \to F$ (which is also denoted by f) such that $f_i(\mathrm{id}_k) = f$. A **minimal set of generators** for F is a set $\{f_1 \in F(i_1), \cdots f_n \in F(i_n)\}$ such that the induced natural transformation

$$\bigoplus_{m=1}^{n} f_m : \bigoplus_{m=1}^{n} k(i_m, -) \to F$$

is a minimal cover of F. A functor has a minimal set of generators if and only if it is of finite rank, in which case the number of generators is given by $\mathrm{rank}\, F$. The 0-th Betti diagram of F is given by the map $\beta_0 F : I \to \mathbb{N}$ which maps i to the number of generators in $F(i)$.

Example 2.0.1. Let $F : \mathbb{N}^2_{\geq 0} \to \mathrm{Vect}_k$ be a functor with the following representation

$$F = \begin{array}{ccccccc} k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & k & \longrightarrow & k \end{array}$$

When we say that this diagram represents the functor F, we mean that this shows the values of F in $[0,3] \times [0,3] \subset \mathbb{N}^2_{\geq 0}$, and that all the maps which are not necessarily zero are the identity. Also all the maps which are not shown are the identity. To represent a functor $\mathbb{R}^2_{\geq 0} \to \mathrm{Vect}_k$, we show the values of the functor at the critical points. It is clear that the free functor $\tilde{F} = k((0,2), -) \oplus k((1,1), -) \oplus k((2,0), -)$ covers F, in the sense that there is a surjective map $\tilde{F} \to F$.

$$\tilde{F} = \begin{array}{ccccccc} k & \longrightarrow & k^2 & \longrightarrow & k^3 & \longrightarrow & k^3 \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k^2 & \longrightarrow & k^3 & \longrightarrow & k^3 \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k^2 & \longrightarrow & k^2 \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & k & \longrightarrow & k \end{array}$$

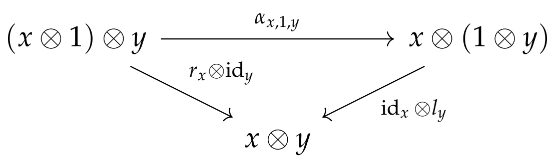
Visually, it is also clear that this is a minimal cover of F. It follows that rank $F = 3$. ◊

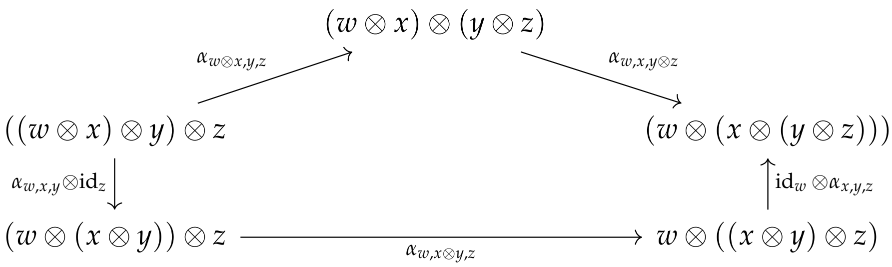
Throughout this thesis we will use the language of category theory. The prerequisite knowledge of the theory does not go far beyond the definition of a functor and a category, but there are some definitions that will be used later on which we will state here.

Definition 2.0.2. A monoidal category is a category C equipped with the following data

- A functor
$$\otimes : C \times C \to C$$
called the tensor product
- An object $1_C \in C$ which is the unit for the tensor functor.
- A natural isomorphism
$$\alpha : ((-) \otimes (-)) \otimes (-) \to (-) \otimes ((-) \otimes (-))$$
with components given by
$$\alpha_{x,y,z} : (x \otimes y) \otimes z \to x \otimes (y \otimes z)$$
- Natural isomorphisms
$$l : 1 \otimes (-) \to (-)$$
$$r : (-) \otimes 1 \to (-)$$
with components given by
$$l_x : 1 \otimes x \to x, \quad r_x : x \otimes 1 \to x$$

And these natural isomorphisms must be such that the following diagrams commute for all $x, y, w, z \in C$





Definition 2.0.3. A lax monoidal functor is a functor $F : C \to D$ between monoidal categories $(C, \otimes_C, 1_C), (D, \otimes_D, 1_D)$, together with a natural transformation with components

$$\phi_{x,y} : Fx \otimes_D Fy \to F(x \otimes_x y)$$

and a morphsim $\psi : 1_D \to F(1_C)$, such that the following diagrams commute:
Associativity:

$$\begin{array}{ccc}
(F(x) \otimes_D F(y)) \otimes_D F(z) & \xrightarrow{\alpha^D_{F(x),F(y),F(z)}} & F(x) \otimes_D (F(y) \otimes_D F(z)) \\
\downarrow{\scriptstyle \phi_{x,y} \otimes \mathrm{id}} & & \downarrow{\scriptstyle \mathrm{id} \otimes \phi_{y,z}} \\
F(x \otimes_C y) \otimes_C F(z) & & F(x) \otimes_D F(y \otimes_C z) \\
\downarrow{\scriptstyle \phi_{x \otimes_C y, z}} & & \downarrow{\scriptstyle \phi_{x, y \otimes_C z}} \\
F((x \otimes_C y) \otimes_C z) & \xrightarrow[F(\alpha^C_{x,y,x})]{} & F(x \otimes_C (y \otimes_C z))
\end{array}$$

Unitality:

$$\begin{array}{ccc}
1_D \otimes_D F(x) & \xrightarrow{\psi \otimes \mathrm{id}} & F(1_C) \otimes_D F(x) \\
\downarrow{\scriptstyle l_{F(x)D}} & & \downarrow{\scriptstyle \phi_{1_C,x}} \\
F(x) & \xleftarrow[F(l^C_x)]{} & F(1_C \otimes_C x)
\end{array}
\qquad
\begin{array}{ccc}
F(x) \otimes_d 1_d & \xrightarrow{\mathrm{id} \otimes \psi} & F(x) \otimes_D F(1_C) \\
\downarrow{\scriptstyle r_{F(x)D}} & & \downarrow{\scriptstyle \phi_{x,1_C}} \\
F(x) & \xleftarrow[F(r^C_x)]{} & F(x \otimes_C 1_C)
\end{array}$$

The natural transformation ϕ and the morphism ψ are called the monoidal coherence maps. An oplax monoidal functor is a lax monoidal functor $C^{\mathrm{op}} \to D^{\mathrm{op}}$.

Definition 2.0.4. Let C be an essentially small abelian category. The Grothendieck group of C, denoted $G(C)$, is the abelian group with one generator for each isomorphism class $[a]$ of objects $a \in C$, and one relation $[b] \sim [a] + [c]$ for every short exact sequence

$$0 \to a \to b \to c \to 0.$$

$G(C)$ has a natural translation invariant ordering given by $[a] \leq [b]$ whenever there is a morphism $[a] \to [b]$ which forms the second morphism of a short exact sequence as above.

Definition 2.0.5. Given a functors $F : C \to D, G : C \to C'$, the right Kan extension of F along G is a pair $(\tilde{F}, \phi)$ where $\tilde{F}$ is a functor $\tilde{F} : C' \to D$, and ϕ is a natural transformation $\phi : \tilde{F} \circ G \to F$ which is universal in the folowing sense: Given a pair (H, ψ) which satisfies the same conditions as (F, ϕ), there exists a unique natural transformation $\rho : H \to \tilde{F}$ such that $\psi = \phi \circ \rho$.

$$\begin{array}{ccc}
C' & \overset{H}{\searrow} & \\
{\scriptstyle G}\uparrow & \overset{\tilde{F}}{\searrow} & \\
C & \xrightarrow[F]{} & D
\end{array}$$

We will use a number of ideas from topology throughout the thesis. The reader is assumed to be broadly familiar with simplicial complexes and homology, otherwise [22] provides a comprehensive introduction.

Definition 2.0.6. A multifiltered space is a topological space X along with a family of subspaces $\{X_v\}_{v \in \mathbb{N}^r}$ such that $X_v \subset X_w$ whenever $v \leq w$

Definition 2.0.7. A pseudometric on a metric space X is a map $d : X \times X \to \mathbb{R}_{\geq 0}$ such that

- $d(x,x) = 0 \forall x \in X.$
- $d(x,y) = d(y,x) \forall x, y \in X.$
- $d(x,y) \leq d(x,z) + d(z,y) \forall x, y, x \in X.$

In particular, it is a metric without the requirement that $d(x, y) = 0 \implies x = y$. Its is called an extended pseudometric if its codomain is $\mathbb{R}_{\geq 0} \cup \infty$.

We will also be using some ideas from algebra, particularly multigraded rings and modules.

Definition 2.0.8. An r-graded ring is a ring R with a decomposition of abelian groups

$$R \cong \bigoplus_{v \in \mathbb{N}^r} R_v$$

such that $R_u \cdot R_v \subset R_{u+v}$

A 1-graded ring is just called a graded ring.

Definition 2.0.9. A multigraded module over R is a module M with a vector space decomposition

$$M = \bigoplus_{v \in \mathbb{N}^r} M_v$$

such that $R_w \cdot M_v \subset M_{w+v}$. A module homorphism that preserves the multigrading is a homorphism of multigraded modules. Elements of the summands M_v are called homogeneous elements.

Chapter 3

One-dimensional Persistence

3.1 Motivation

The theory of one-dimensional persistence is well studied. Here we give an introduction to this theory and build up to the proof of the isometry theorem, one of the most important theorems in one-dimensional persistence. Fix a ring R with unit throughout this section.

Definition 3.1.1. Let X be a simplicial complex with a filtration

$$0 = X_0 \subset X_1 \subset X_2 \subset \cdots \subset X_n = X.$$

Then the (p,q)-persistent kth homology group of X with coefficients in R, denoted $H_k^{p,q}(X,R)$, is defined as the image of the map

$$i : H_k(X_p, R) \to H_k(X_q, R)$$

where i is the map induced from the inclusion $X_p \hookrightarrow X_q$.

There is another definition of persistent homology which is perhaps a bit more illustrative. First of all, recall that given a simplicial complex X, we can associate to it a chain complex

$$\cdots \xrightarrow{d_{i+2}} C_{i+1} \xrightarrow{d_{i+1}} C_i \xrightarrow{d_i} C_{i-1} \longrightarrow \cdots$$

where C_i is the free group on the i-simplices of X. Then we have the groups $Z_k(X) = Z_k(C_\bullet) = \ker d_k$ and $B_k(X) = B_k(C_\bullet) = \operatorname{im} d_{k+1}$. Then of course $H_k(X,R) = Z_k(X)/B_k(X)$. We can also define the (p,q)-persistent kth homology group of a filtered complex X as

$$H_k^{p,q}(X) \cong Z_k(X_p, R)/(B_k(X_q, R) \cap Z_k(X_p, R)).$$

From this definition we see that $H_k^{p,q}$ has a generator for every k-cycle in X_p which is not mapped to a boundary in X_q by the inclusion $X_p \hookrightarrow X_q$. In particular, this definition is the same as the one above. Note that $H_k^{p,q}(X_i)$ is an R module since it is defined as the image of an R-module homomorphism. The idea behind persistent homology is that if we fix p, and increase q we kill short-lived generators, which are the "topological noise" of the complex.

Example 3.1.2 (Vietoris-Rips complex)**.** Given a collection of points $V \subset \mathbb{R}^n$ (or indeed any metric space), the Vietoris-Rips complex of these points for a value $\varepsilon \in \mathbb{R}$, is the simplicial complex whose k-simplices correspond to $(k+1)$-tuples of points in V which are pairwise within a distance ε. An example of how persistent homology works is to look at a filtered complex where for a given set of points, each step in the filtration is the Vietoris-Rips complex of those points for a different ε value. Then you can examine which homology generators are present for a large range of ε values. These are called features of the set of points (typically a data set). The homology generators which only exist for a small range of ε values are called noise.

$\lozenge$

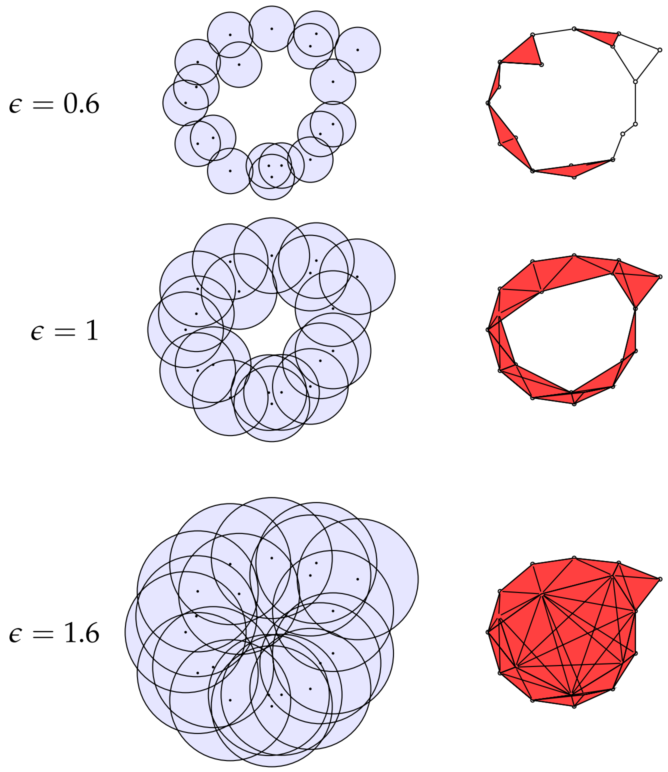


FIGURE 3.1: The Vietoris-Rips complex for varying values of ϵ for some points. These points roughly form an annulus concentrated around the unit circle. We see that the hole in the center persists for a large range of ϵ, while smaller holes such as the one present for $\epsilon = 0.6$ vanish quickly.

Example 3.1.3 (Sub-level sets)**.** Let X be a topological space, and let $f : X \to \mathbb{R}$ be a function. For each $t \in \mathbb{R}$, we have the sublevel set

$$(X, f)_t = f^{-1}((-\infty, t]) = \{x \in X | f(x) \leq t\}$$

Note that if $a \leq b$, then $f^{-1}((-\infty, a]) \subset f^{-1}((-\infty, b])$, and so this defines a real-indexed filtration of X. $\Diamond$

Example 3.1.4 (Čech complex)**.** Given a collection of points in $V \subset \mathbb{R}^n$ (or indeed any metric space), the Čech complex of these points for a value $\varepsilon \in \mathbb{R}$, is the simplicial complex whose k-simplices correspond to $(k+1)$-tuples of points in V whose closed $\varepsilon/2$-neighhbourhoods have a common point of intersection. $\Diamond$

Definition 3.1.5. Let $(P, \leq)$ be a partially ordered set. A P-persistence module is a family of P-indexed R-modules $(M_i)_{i \in P}$ along with module homomorphisms

$$v_s^t : M_s \to M_t$$

whenever $s \leq t$ such that $v_s^t \circ v_r^s = v_r^t$ and v_t^t is the identity on M_t.
In particular, a persistence module is a functor from $(P, \leq)$ to the category of R-modules. We will sometimes use both the index notation and functor notation i.e. M_t and $M(t)$ mean the same thing. The $R-$modules M_i are called the **components** of M, and the maps $v_s^t = M(s \leq t)$ are called the **structure maps**.

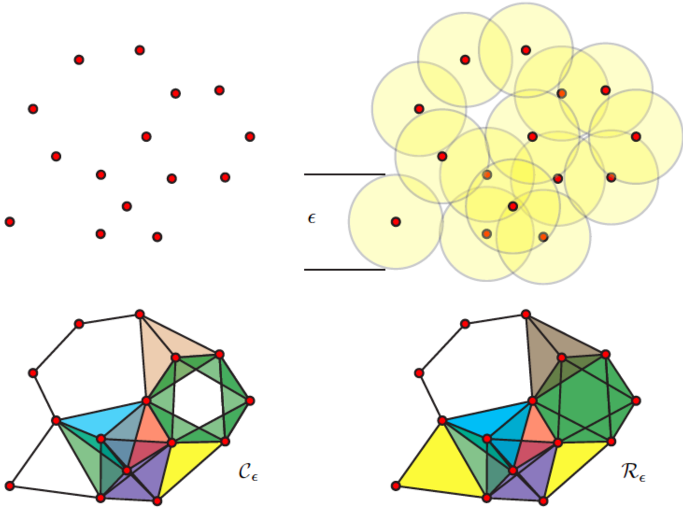


FIGURE 3.2: [8] Comparison of the Čech complex (left) and Vietoris-Rips complex (right) of a set of points in the plane for a particular value of ε.

We will generally be concerned with persistence modules over $\mathbb{R}$ or some subset of $\mathbb{R}$. If a persistence module is indexed over $\mathbb{N}$ it is called a discrete persistence module.
If M is a P-persistence module and $Q \subset P$, then we get a Q-persistence module by considering only those R-modules and maps in M with indices in Q. This is called the **restriction of M to Q**, denoted M_Q.

Given a persistence module M, we denote by $M[a]$ the **a-shift of** M, defined as

$$M[a]_i = M_{i+a}, \quad M[a](s \leq t) = M(s+a \leq t+a)$$

Note that a discrete persistence module M is entirely defined by the diagram

$$\cdots \xrightarrow{M(t-2 \leq t-1)} M_{t-1} \xrightarrow{M(t-1 \leq t)} M_t \xrightarrow{M(t \leq t+1)} M_{t+1} \xrightarrow{M(t+1 \leq t+2)} \cdots$$

Persistence modules are the central algebraic objects of study in the field of persistent homology.

Definition 3.1.6. A morphism of persistent modules M, N is a natural transformation $\phi : M \to N$.

We can now consider the collection of R-persistence modules as a category, with morphisms as per the definition above. As with any category, we define an isomorphism of persistence modules to be a morphism with an inverse. This category is an abelian category, and it inherits this structure from the category of R modules. For example, given two persistence modules M and N, their product is the persistence module with modules given by $M_i \Pi N_i$, and structure maps given by $M(s \leq t) \Pi N(s \leq t)$.

Example 3.1.7. Given a filtered complex

$$0 \subset X_1 \subset \cdots \subset X_n = X$$

we can define a persistence module $M : (\mathbb{N}, \leq) \to R - mod$ by

$$M_i = H_k(X_i, R), \quad M(i \leq j) = \iota_i^j.$$

Where ι_i^j is induced from the inclusion $X_i \hookrightarrow X_j$. This is the canonical example to keep in mind as it is in practice how persistent homology is applied to a data set. ♢

Example 3.1.8. Interval Modules Let $[a, b) \subset \mathbb{R}$ be a half-open interval. Then the interval functor $[a, b) : \mathbb{R} \to R$-mod is a persistence module. This will turn out to be an important example. ◊

We will often want to impose some finiteness conditions on persistence modules, the following are some such conditions.

Definition 3.1.9. A persistence module M is called q-tame if

$$r_s^t := \operatorname{rank} M(s \leq t) < \infty \text{ whenever } s < t$$

Definition 3.1.10. A persistence module M is locally finite if it is a direct sum of interval modules, and any bounded subset of $\mathbb{R}$ intersects only finitely many of the intervals.

Definition 3.1.11. An $\mathbb{R}$-persistence module M is of finite type if there exists some $m \in \mathbb{R}$ such that $M(s \leq t)$ is an isomorphism whenever $s, t \geq m$.

Definition 3.1.12. An $\mathbb{R}$-persistence module M is of finitely presented (generated) type if it is of finite type and each M_i is finitely presented (generated).

We recall here the definition of a finitely presented module

Definition 3.1.13. An R module A is finitely generated if there is a surjective module morphism $\phi : R^n \to A$ for some $n \in \mathbb{N}$. A is called finitely presented if $\ker \phi$ is also finitely generated, in which case we have an exact sequence

$$R^m \longrightarrow R^n \xrightarrow{\phi} A \longrightarrow 0$$

3.2 The representation theorem

Given a discrete persistence module M, we can associate with it a graded $R[t]$-module, where $R[t]$ is given the usual grading ,via the map

$$\alpha(M) := \bigoplus_{i \in \mathbb{N}} M_i$$

where the R-module structure is the sum of the structures of each M_i, and where multiplication by t is defined by $t \cdot m_i := M(i \leq i+1)(m_i)$ for $m_i \in M_i$. Similarly, given an $R[t]$-module $\bigoplus_{i \in \mathbb{N}} M_i$, we get a discrete persistence module by defining a functor β from $R[t]$-mod to the category of discrete persistence modules, where

$$\beta\left(\bigoplus_{i \in \mathbb{N}} M_i\right) = M$$

where $M(i \leq i+1)(m_i) := t \cdot m_i$.
In [6], Zomorodian and Carlsson state the following theorem

Theorem 3.2.1. *The functors α and β define an equivalence of categories between the category of discrete persistence modules of finitely generated type over R and the category of finitely generated non-negatively graded modules over $R[t]$*

This statement is not entirely correct, as can be demonstrated via the following counterexample, due to [7]. Let $R = \mathbb{Z}[x_1, x_2, \ldots]$ and consider the graded $R[t]$-module $M = \bigoplus_{i \in \mathbb{N}} M_i$, where $M_i = R/(x_1, \ldots, x_i)$, and where multiplication by t is the map $M_i \to M_{i+1}$ given by $t \cdot f := f \mod x_{i+1}$. Now the module M is generated by $\{1\}$, but the persistence module $\beta(M)$ is not of finite type, since the inclusion $M_i \to M_{i+1}$ is never an isomorphism.

It turns out that R needs to be commutative and Noetherian for theorem 3.2.1 to hold. As can be seen from the counterexample, the issue is that $\beta(M)$ is not generally of finitely generated type whenever M is finitely generated. In [7], Corbet and Kerber prove the following version of the representation theorem

Theorem 3.2.2. *The category of finitely presented graded $R[t]$-modules is isomorphic to the category of discrete persistence modules over R of finitely presented type.*

We note that this theorem, along with the following lemma, gives us theorem 3.2.1 in the case that R is commutative and Noetherian.

Lemma 3.2.3. *If R is a Noetherian ring, then every finitely generated R-module is finitely presented. If R is also commutative, then $R[t]$ is Noetherian.*

The representation theorem now guides us in choosing suitable rings over which to take persistence modules. If R is not a field, we have no simple description for the modules over $R[t]$. When $R = k$, a field, then $k[t]$ is a PID, so every every graded module over $k[t]$ is of the form

$$\left(\bigoplus_{j=1}^{m} t^{\beta_j} \cdot (k[t]/(t^{n_j}))\right) \oplus \left(\bigoplus_{i=1}^{n} t^{\alpha_i} \cdot k[t]\right). \tag{3.1}$$

In the case where a persistence module M has its components given by homology in degree k, as in example 3.1.9, then if we work over a field k, $\alpha(M)$ will have the above form. The interpretation of 3.1 is that the free parts correspond to homology generators which are born at the parameter α_i and persist (i.e. do not die). The torsion elements correspond to homology generators which appear at a parameter β_i and die at $\beta_i + n_j$. We note also that in this case the persistent homology groups $H_n^{p,q}$ will be vector spaces over k, and are therefore fully determined by their rank. The persistence of M is then the data of the intervals $[\beta_j, \beta_j + n_j)$, and we can visualize this via what we call barcodes, which we will meet later. In particular, the persistence module corresponding to 3.1 is

$$\left(\bigoplus_{j=1}^{m} [\beta_j, \beta_j + n_j)\right) \oplus \left(\bigoplus_{i=1}^{n} k(\alpha_i, -)\right).$$

Due to this nice interpretation, we will work over a fixed field k from now on, unless otherwise specified.

3.3 The role of interval modules

Definition 3.3.1. The direct sum of two persistence modules M, N is defined by

$$(M \oplus N)_i = M_i \oplus N_i, \quad (M \oplus N)(v \leq w) = M(v \leq w) \oplus N(v \leq w)$$

Definition 3.3.2. A persistence module M is said to be indecomposable if the only direct sum decompositions of M are of the form $M = M_1 \oplus M_2$ are where either $M_1 = 0$ or $M_2 = 0$.

We will see that interval modules are the "building blocks" for persistence modules. Here we state some facts about interval modules.

Lemma 3.3.3. *Let $[a, b)$ be an interval $\mathbb{R}$-persistence module. Then* $\mathrm{End}([a, b)) = k$.

Lemma 3.3.4. *Interval modules are indecomposable.*

Theorem 3.3.5. *If a persistence module M can be decomposed into interval modules in two separate ways*

$$M \cong \bigoplus_{l \in L} [a_l, b_l) \cong \bigoplus_{n \in N} [a_n, b_n)$$

Then there is a bijection $\sigma : L \to N$ such that $J_l = K_{\sigma(n)}$ for all l.

We wish to know when a persistence module can be written as a direct sum of interval modules, and also with what multiplicity each interval module occurs. This will give us all the information we need to construct the barcode of the module.
We now introduce some definitions due to [4], which will only be used in the statement of theorems 3.3.9 and 3.3.11.

Definition 3.3.6. A T-persistence module M is said to have the descending chain condition if for all $t, s_1, s_2, \ldots \in T \subset \mathbb{R}$ with $t \geq s_1 > s_2 > \cdots$, the chain

$$M_t \supset \operatorname{im} M(s_1 \leq t) \supset \operatorname{im} M(s_2 \leq t) \supset \cdots$$

stabilizes, and for all $t, r_1, r_2, \ldots \in T$ with $t \leq \cdots < r_2 < r_1$, the chain

$$M_t \supset \ker M(t \leq r_1) \supset \ker M(t \leq r_2) \supset \cdots$$

stabilizes.

Definition 3.3.7. A cut for $T \subset \mathbb{R}$ is a pair $c = (c^-, c^+)$ of subsets of T such that $T = c^- \cup c^+$ and $s < t$ for all $s \in c^-$ and $t \in c^+$. Let M be a T-persistence module.

If c is a cut with $t \in c^+$, we define the following subspaces of M_t:

$$\operatorname{im}^-_{ct} = \bigcup_{s \in c^-} \operatorname{im} M(s \leq t), \quad \operatorname{im}^+_{ct} = \bigcap_{s \in c^+, s \leq t} \operatorname{im} M(s \leq t),$$

and if $t \in c^-$, we define the following subspaces of M_t:

$$\ker^-_{ct} = \bigcup_{r \in c^-, t \leq r} \ker M(t \leq r), \quad \ker^+_{ct} = \bigcap_{r \in c^+} \ker M(t \leq r).$$

If c^- is empty, we set $\operatorname{im}^-_{ct} = 0$, and if c^+ is empty we set $\ker^+_{ct} = M_t$.
If $I \subset T$ is an interval, there are uniquely determined cuts l and u with $I = l^+ \cap u^-$. In particular

$$l^- = \{t : t < s \forall s \in I\}, \quad l^+ = \{t : t \geq s \text{ for some } s \in I\}$$
$$u^+ = \{t : t > s \forall s \in I\}, \quad u^- = \{t : t \leq s \text{ for some } s \in I\}.$$

Definition 3.3.8. Let M and T be as before. For $I \subset T$ an interval, and for $l = (l^-, l^+)$ and $u = (u^-, u^+)$ the unique cuts determined by I, we define the following subspaces of M_t:

$$V^-_{It} = (\operatorname{im}^-_{lt} \cap \ker^+_{ut}) + (\operatorname{im}^+_{lt} \cap \ker^-_{ut}) \quad V^+_{It} = \operatorname{im}^+_{lt} \cap \ker^+_{ut}$$

The maps $M(s \leq t)$ induce maps on $V^\pm_{Is} \to V^\pm_{It}$, and using these maps we can define the inverse limit $V^\pm_I = \varprojlim_{t \in I} V^\pm_{It}$. We are now ready to state our first theorem about the decomposition of modules as a sum of interval modules.

Theorem 3.3.9. *Any persistence module M with the descending chain condition is a direct sum of interval modules. Also, the multiplicity of $[a, b)$ as a summand of M is equal to the dimension of $V^+_{[a,b)} / V^-_{[a,b)}$.*

Remark. The proof of this statement in [4] assumes that M is a T-persistence module where T has a countable subset which is dense in the order topology on T i.e. where T is separable. Since we are working with $T \subset \mathbb{R}$, this will always be the case. This is because $\mathbb{R}$ is a separable metric space, and the order topology on $\mathbb{R}$ is the same as the standard topology. It is a general fact that any subspace of a separable metric space is separable.

We note in particular that if M is such that each M_i is finite dimensional, or if M is a T-persistence module where T is finite, then it satisfies the descending chain condition. We say that M is of length n if T has cardinality equal to n. In such cases we may always assume that $T = \{1, 2, \ldots n\}$. In this case we have an easier way of determining the interval decomposition of M.

Definition 3.3.10. Let M be a peristence module of length n. The right filtration of M, denoted $R(M)$, is the n-tuple given by

$$R(M) = (\operatorname{im} M(1 \leq n), \operatorname{im} M(2 \leq n), \ldots, \operatorname{im} M(n-1, n), M_n)$$

The easiest case is when all the structure maps are injective. In this case, M is called **streamlined**. As we have seen from previous examples, this is the most common case in applications.

Theorem 3.3.11. *[3] Let M be a streamlined persistence module of length n, and define the n-tuple $c = (c_1, c_n, \ldots c_n)$ such that c_i is the dimension of the ith component of $R(M)$. Then we have an isomorphism of persistence modules*

$$M \cong \bigoplus_{1 \leq i \leq n} c_i[i, n).$$

We can make a similar statement when M is not necessarily streamlined.

Definition 3.3.12. If $M = M_1 \to M_2 \to \cdots \to M_n$, then for $m \leq n$ let $M/m = M_1 \to M_2 \to \cdots M_m$ be the truncation of M to length m.

Theorem 3.3.13. *[3] Let M be a persistence module of length n. For $m < n$, define the m-tuple $(c_1^m, c_2^m, \ldots, c_m^m)$ such that c_i^m is the dimension of $R(M/m)_i \cap \ker M(m \leq m+1)$, i.e. the dimension of $\operatorname{im} M(i, m) \cap \ker M(m \leq m+1)$. When $m = n$, then $c_i^n = c_i$ where c_i is as in the previous definition. Then we have an isomorphism*

$$M \cong \bigoplus_{1 \leq i \leq m \leq n} c_i^m[i, m)$$

Example 3.3.14. Consider the persistence module $M : \{1, 2, \cdots, n\} \to Vect_k$ given by

$$M_t = \begin{cases} k^2 & t = 2, 3 \\ k & \text{otherwise} \end{cases}$$

where $M(1 \leq 2) = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $M(3 \leq 4) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and all other structure maps are the identity. By inspection, we can see that

$$M \cong [1, n] \oplus [2, 3].$$

We can verify this using the methods described above. Setting the interval $I = [2, 3]$, to find its multiplicity we need to compute V_I^+, V_I^-. First we see that the cuts u, l for the interval I are given by

$$\begin{aligned} l^- = \{1\} \quad & l^+ = \{2, 3, \ldots n\} \\ u^- = \{1, 2, 3\} \quad & l^+ = \{4, 5, \ldots n\}. \end{aligned}$$

These cuts determine the following values

$$\begin{aligned} \mathrm{im}^+_{l2} &= k^2 & \mathrm{im}^+_{l3} &= k^2 \\ \mathrm{im}^-_{l2} &= k \oplus 0 & \mathrm{im}^-_{l3} &= k \oplus 0 \\ \ker^+_{u2} &= 0 \oplus k & \ker^+_{u3} &= 0 \oplus k \\ \ker^-_{u2} &= 0 & \ker^-_{u3} &= 0. \end{aligned}$$

And now we can compute

$$\begin{aligned} V^+_{I2} = k, V^+_{I3} = k &\implies V^+_I = k \\ V^-_{I2} = 0, V^-_{I3} = k &\implies V^-_I = 0. \end{aligned}$$

This gives us the result that we were expecting. Similarly we can show that the multiplicity of $[1, n]$ is 1, and all other intervals have multiplicity 0. Using the second method described above, we first note that the right filtration of M is $R(M) = (k, k, \ldots, k)$. Now $M \cong \bigoplus_{1 \leq i \leq m \leq n} c_i^m [i, m]$ as per Theroem 3.3.13, where

$$\begin{aligned} &c_2^3 = R(M/3)_2 \cap \ker M(3 \leq 4) = k^2 \cap k = k, \\ &c_1^n = c_1 = k. \end{aligned}$$

Since $\ker M(m \leq m+1) = 0$ whenever $m \neq 3$, the only other interval with non-zero mulitplicity must be of the form $[i, 3]$. Now $c_1^3 = (k \oplus 0) \cap (0 \oplus k) = 0$, and so $M \cong [1, n] \oplus [2, 3]$. ♢

So to recap, we now have an explicit decomposition of a persistence module M whenever M is of length n for some $n \in \mathbb{N}$, and whenever M satisfies the descending chain condition. In particular, this includes the case where each M_i is finite dimensional. Thus we are at the very least justified in using persistent homology in practical applications, where every thing is necessarily finite.

Example 3.3.15. We now consider an example of a persistence module with no interval decomposition, due to [23]. Define a $\mathbb{Z}_{\leq 0}$-persistence module M by

$$\begin{aligned} M_0 &= \{\text{sequences of real numbers } (x_1, x_2, \ldots)\} \\ M_{-n} &= \{\text{sequences with } x_1 = \cdots = x_n = 0\} \end{aligned}$$

The maps $M(-m \leq -n)$ are the inclusion maps $M_{-m} \hookrightarrow M_{-n}$. Assume M has an interval decomposition. Since M is streamlined, each of the intervals must be of the form $[-n, 0]$ or $(-\infty, 0]$. Since $\dim(M_{-n}/M_{-n-1}) = 1$, each interval module occurs with multiplicity 1. Since $\bigcap_{n \in \mathbb{Z}_{\leq 0}} M_{-n} = \{0\}$, the interval $(-\inf, 0]$ does not occur at all. This would imply that $M \cong \bigoplus_{n \geq 0} [-n, 0]$, but this contradicts the fact that $\dim(M_0)$ is uncountable. Note that M clearly does not satisfy the descending chain condition. ♢

Definition 3.3.16. If a persistence module M can be decomposed as a direct sum of interval modules

$$M \cong \bigoplus_{n \in N} [a_n, b_n)$$

Then we define the persistence barcode of M to be the multiset

$$\mathrm{bar}(M) = \{(a_n, b_n) \in \mathbb{R}^2 | n \in N\}.$$

Theorem 3.3.5 tells us that this is invariant under isomorphism, and indeed every possible multiset of intervals in $T \subset \mathbb{R}$ gives an isomorphism class of T-persistence modules. There are two common ways to represent the data of $\mathrm{bar}(M)$: a barcode diagram, and by plotting the points of $\mathrm{bar}(M)$ in the plane, often referred to as a persistence diagram.

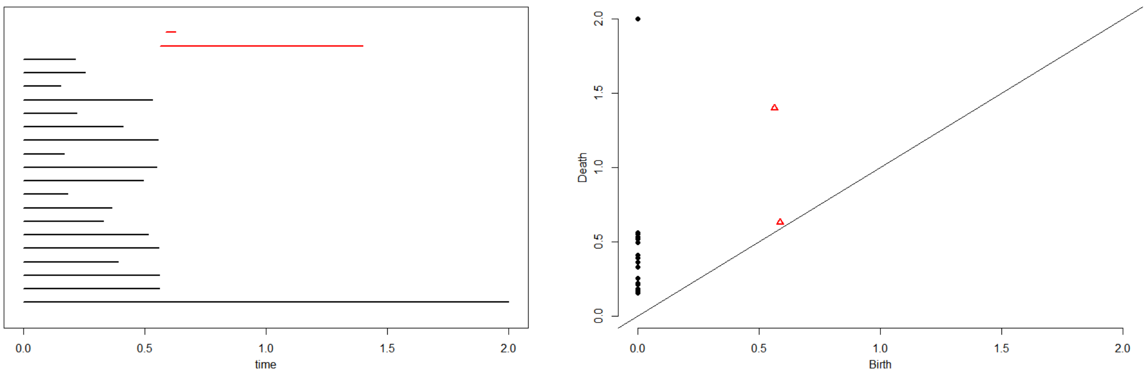
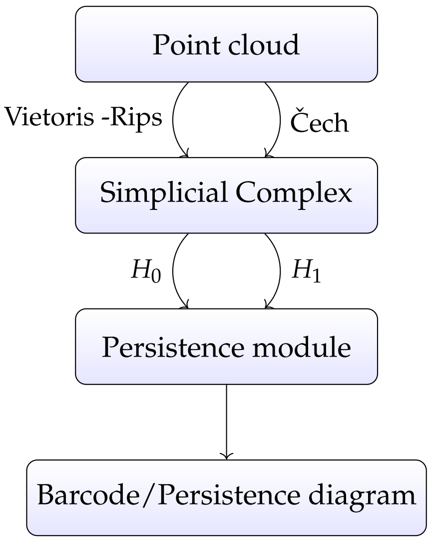


FIGURE 3.3: The barcode (left) and persistence diagram (right) of the points in figure 3.1

Example 3.3.17. Consider again the points in figure 3.1. Their persistence diagram and barcodes are given by the two diagrams in figure 3.3 above. In both diagrams black represents H_0 and red represents H_1. We can see that there is clearly one generator of H_1 which persists for a long range of ε, representing the hole in the ring of points. The short lived generator of H_1 is noise, as it does not contribute to our image of these points as a circle. Note that there will always be a H_0 generator which does not die. $\Diamond$

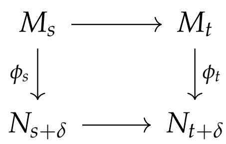
So now we have seen the full process of going from a data set to a visualisation of the shape of the data, sometimes referred to as the "persistence pipeline". Given a data set in the form of a point cloud, we build the Vietoris-Rips complex for varying values of our parameter ε. We apply the functors $H_i(-)$ to the complex at each value of ε, which gives us a persistence module. We then compute the barcode of this module and plot it.



3.4 Interleaving

The idea of interleaving was first introduced in [5].

Definition 3.4.1. Let M, N be R-persistence modules, and let $\delta \in \mathbb{R}$. A persistence module homomorphism of degree δ is a family of maps $\phi_t : M_t \to N_{t+\delta}$ such that the following diagram commutes for all $s < t$



where the unlabelled maps are the structure maps.

So in other words it is a persistence module homomorphism $M \to N[\delta]$. We denote the degree δ morphisms $M \to N$ by $\mathrm{Hom}^{\delta}(M, N)$, and we define $\mathrm{End}^{\delta}(M)$ similarly. Composition is then given by a map

$$\mathrm{Hom}^{\delta_1}(M, N) \times \mathrm{Hom}^{\delta_2}(L, M) \to \mathrm{Hom}^{\delta_1+\delta_2}(L, N).$$

Example 3.4.2. The shift map $1_M^{\delta} \in \mathrm{End}^{\delta}(M)$ is the collection of the structure maps of M of the form $M(s \leq s + \delta)$. ♢

Note that if $\phi \in \mathrm{Hom}^{\varepsilon}(M, N)$ for any $\varepsilon \in \mathbb{R}$, then $\phi \circ 1_M^{\delta} = 1_N^{\delta} \circ \phi$ for all $\delta \geq 0$. This is just a reformulation of the commutativity condition in definition 3.4.1.

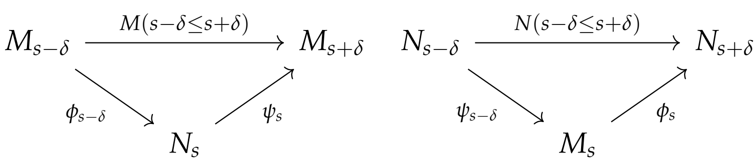
Definition 3.4.3. Two persistence modules are δ-interleaved if there are homomorphisms $\phi \in \mathrm{Hom}^{\delta}(M, N)$, $\psi \in \mathrm{Hom}^{\delta}(N, M)$ such that

$$\psi \circ \phi = 1_M^{2\delta}, \quad \phi \circ \psi = 1_N^{2\delta}$$

Explicitly, we have maps

$$\phi_t : M_t \to N_{t+\delta}, \quad \psi_t : N_t \to M_{t+\delta}$$

which satisfy the commutativity condition in definition 3.4.1, and the following diagrams commute.



To write things more succinctly, a δ-interleaving between two persistence modules M and N is the following data

$$\phi \circ 1_M^{\varepsilon} = 1_N^{\varepsilon} \circ \phi, \quad \psi \circ 1_N^{\varepsilon} = 1_M^{\varepsilon} \circ \psi, \quad \psi \circ \phi = 1_M^{2\delta}, \quad \phi \circ \psi = 1_N^{2\delta} \tag{3.2}$$

where $\phi \in \mathrm{Hom}^{\delta}(M, N), \psi \in \mathrm{Hom}^{\delta}(N, M)$ and the first two relations hold for any $\varepsilon \in \mathbb{R}$.

Example 3.4.4. Let X be a topological space, and $f, g : X \to \mathbb{R}$ be functions. Consider the persistence modules we get by applying $H_n(-)$ to the filtrations $((X, f)_t)_{t\in\mathbb{R}}$ and $((X, g)_t)_{t\in\mathbb{R}}$ of X. We denote these by $H_n((X, f))$ and $H_n((X, g))$ respectively, where $H_n((X, f))_t = H_n((X, f)_t)$, where here we are viewing the persistence module as a functor from $(\mathbb{R}, \leq)$. Note that we leave out the coefficients, since we have already stated that we are working over a fixed field k. If $||f - g||_\infty < \delta$, then $H_n((X, f))$ and $H_n(X, g))$ are δ-interleaved. Note that we have an inclusion $(X, f)_t \subset (X, g)_{t+\delta}$. To see this, let $x \in (X, f)_t$, so $f(x) \leq t$. Now if $g(x) \leq f(x)$, then $g(x) \leq t \implies g(x) \leq t + \delta$, and so $x \in (X, g)_{t+\delta}$. If $f(x) < g(x)$, then $g(x) - f(x) < \delta \implies g(x) < f(x) + \delta \implies g(x) < t + \delta$, and so once again we have $x \in (X, g)_{t+\delta}$. Similarly, we have an inclusion $(X, g)_t \subset (X, f)_{t+\delta}$. These inclusions induce morphisms of degree δ

$$\phi : H_n((X, f)) \to H_n((X, g)), \quad \psi : H_n((X, g)) \to H_n((X, f))$$

Since these maps are induced by inclusion maps, the interleaving relations 3.2 are satisfied. ♢

An interleaving of two persistence modules can itself be associated to a persistence module over a particular poset. Consider the partial order on $\mathbb{R}^2$

$$(a_1, b_1) \leq (a_2, b_2) \Leftrightarrow a_1 \leq a_2 \text{ and } b_1 \leq b_2.$$

For any $x \in \mathbb{R}$, define the x-shifted diagonal as

$$\Delta_x = \{(a, b) \in \mathbb{R}^2 | \, b - a = 2x\}$$

The category $(\Delta_x, \leq)$ is equivalent to $(\mathbb{R}, \leq)$ via the map $t \to (t - x, t + x) \in \Delta_x$.

Proposition 3.4.5 ([1]). *Let $x, y \in \mathbb{R}$. Persistence modules M, N are $|y - x|$-interleaved if and only if there is a $\Delta_x \cup \Delta_y$-persistence module L such that $L|_{\Delta_x} = M$ and $L|_{\Delta_y} = N$.*

3.5 Interpolation and the isometry theorem

Theorem 3.5.1. *Any persistence module L over $\Delta_{x_0} \cup \Delta_{x_1}$ extends to a persistence module $\tilde{L}$ over the diagonal strip*

$$\Delta_{[x_0, x_1]} = \{(p, q) \in \mathbb{R}^2 | x_0 \leq q - p \leq 2x_1\}$$

Proof. We can interpret this theorem as a functor extension theorem by regarding the posets $\Delta_{x_0} \cup \Delta_{x_1}$ and $\Delta_{[x_0, x_1]}$ as categories. Persistence modules over these posets are functors to the category of vector spaces, so the theorem asserts the existence of an extension $\tilde{L}$ of any functor L in the following diagram

$$\begin{array}{ccc} \Delta_{[x_0,x_1]} & & \\ \uparrow & \overset{\tilde{L}}{\searrow} & \\ \Delta_{x_0} \cup \Delta_{x_1} & \xrightarrow{L} & \mathrm{Vect}_k \end{array}$$

With this view, we can now apply Kan extensions to prove the theorem. It is a well known fact about Kan extensions that given a functors $F : C \to D, G : C \to C'$, then if C is small and D is complete, then F has a right Kan extension along G. Similarly, if D is cocomplete then F has a left Kan extension along G [24]. Upon noting that Vect_k is both complete and co-complete, we immediately get the theorem. Note that this also proves that $\tilde{L}$ may in general not be unique. □

We now get the interpolation lemma of [5] as a corollary of theorem 3.5.1.

Lemma 3.5.2 (Interpolation lemma). *Suppose M, N are a δ-interleaved pair of persistence modules. Then there exists a family of persistence modules $(M^x | x \in [0, \delta])$ such that $M^0 = M$ and $M^\delta = N$, and M^x, M^y are $|x - y|$-interleaved for all $x, y \in [0, \delta]$.*

Proof. If M, N are δ-interleaved, then there is a persistence module L over $\Delta_0 \cup \Delta_\delta$ such that $L|_{\Delta_0} = M$ and $L|_{\Delta_\delta} = N$. By theorem 3.5.1, this extends to a persistence module $\tilde{L}$ over the strip $\Delta_{[0,\delta]}$. We then can define a family of persistence modules by $M^x = \tilde{L}|_{\Delta_x}$, and M^x, M^y are $|y - x|$- interleaved for all $x, y \in [0, \delta]$ by proposition 3.4.5. □

3.5.1 Interleaving distance

If M and N are a pair of δ-interleaved persistence modules, then they are $(\delta+\varepsilon)$-interleaved for all $\varepsilon > 0$. Indeed, if ϕ, ψ are the δ-interleaving maps, then

$$\phi' = \phi \circ 1_M^{\varepsilon} = 1_N^{\varepsilon} \circ \phi$$
$$\psi' = \psi \circ 1_N^{\varepsilon} = 1_M^{\varepsilon} \circ \psi$$

give a $(\delta+\varepsilon)$-interleaving. We wish to find the minimum δ for which two persistent modules are δ-interleaved. However as we will see, this minimum is not always attained. This leads us to the following definition.

Definition 3.5.3. Two persistence modules M, N are called δ^+-interleaved if they are $(\delta+\varepsilon)$-interleaved for all $\varepsilon > 0$.

It is clear from the definition that δ-interleaved implies δ^+-interleaved. In general, the converse does not hold.

Example 3.5.4. Two persistence modules are 0-interleaved iff they are isomorphic. ♢

Example 3.5.5. Let M, N be two non isomorphic persistence modules which have all structure maps equal to the zero map. Then M and N are 0^+-interleaved, but not 0-interleaved. The zero maps provide an ε-interleaving for all $\varepsilon > 0$ since $1_M^{2\varepsilon} = 1_N^{2\varepsilon} = 0$. ♢

Definition 3.5.6. The interleaving distance between two persistence modules is defined as

$$\begin{aligned} d_I(M,N) &= \inf\{\delta \mid M, N \text{ are } \delta\text{- interleaved}\} \\ &= \min\{\delta \mid M, N \text{ are } \delta^+\text{- interleaved}\} \end{aligned}$$

Proposition 3.5.7 ([1]). *The interleaving distance satisfies the triangle inequality. In particular, given three persistence modules M, N and L, we have*

$$d_I(M,L) = d_I(M,N) + d_I(N,L)$$

So the interleaving distance is a pseudometric. Indeed, it is an extended pseudometric if we define $d_I(F,G) = \infty$ whenever F, G are not δ-interleaved for any δ. But as the next example will show, it is not a metric since $d_I(M,N) = 0$ does not imply that $M \cong N$.

Example 3.5.8. The four interval modules $[p,q], [p,q), (p,q], (p,q)$ are 0^+ interleaved, but not isomorphic.

Proof. It is clear that none of the four interval modules are isomorphic. Consider the persistence modules $[p,q], [p,q)$. We wish to show that these modules are ε-interleaved for every $\varepsilon > 0$. This amounts to showing that the following diagrams commute for every $\varepsilon > 0$

$$\begin{array}{ccccc} [p,q]_t & \longrightarrow & [p,q]_{t+2\varepsilon} & \quad & [p,q)_t & \longrightarrow & [p,q)_{t+2\varepsilon} \\ & \searrow \quad \nearrow & & & & \searrow \quad \nearrow & \\ & [p,q)_{t+\varepsilon} & & & & [p,q]_{t+\varepsilon} & \end{array}$$

where all the maps are the identity map where possible, or otherwise 0. Consider the left hand diagram. We have four cases:

- Whenever $t \notin [p,q]$ two of the map are zero, so the diagram commutes.

- If $t = q$, then $t + \varepsilon \notin [p, q)$, and $t + 2\varepsilon \notin [p, q]$, so the diagram commutes because all the maps are 0.
- If $t + 2\varepsilon \notin [p, q]$, then again two of the map are zero, so the diagram commutes.
- If $t \in [p, q)$ and $t + 2\varepsilon \in [p, q]$, then $t + \varepsilon \in [p, q)$, therefore all the maps are the identity, so the diagram commutes.

The right hand diagram commutes for similar reasons. It can be shown that the other interval modules are pairwise 0^+-interleaved by the same argument. □

◊

Proposition 3.5.9. *Let $(M^i | i \in I), (N^i | i \in I)$ be two families of persistence modules for some indexing set I. Let*

$$M = \bigoplus_{i \in I} M^i \qquad N = \bigoplus_{i \in I} N^i.$$

Then

$$d_I(M, N) \leq \sup_{i \in I} d_I(M^i, N^i)$$

Proof. Given interleaving maps ϕ_i, ψ_i for each pair M_i, N_i, then $\phi = \oplus_{i \in I} \phi_i, \psi = \oplus_{i \in I} \psi_i$ are interleaving maps for M, N. So it is clear that any upper bound on the values $d_I(M_i, N_i)$ is also an upper bound for $d_I(M, N)$. □

3.5.2 Bottleneck distance

Recall that given a persistence module M, the multiset $\text{bar}(M)$ is composed of points in the plane corresponing to intervals $[p, q)$. In particular, it is a subset of the extended open half plane

$$\mathbb{H}^\circ = \{(p, q) | -\infty \leq p < q \leq \infty\}$$

We wish to describe two barcodes as being close if there is a bijection between them which does not move any point too far. For this purpose we use the l^∞ norm in the plane:

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

Points at infinity are compared in the following way:

$$\begin{aligned} d^\infty((-\infty, q), (-\infty, s)) &= |q - s| \\ d^\infty((p, \infty), (r, \infty)) &= |p - r| \\ d^\infty((-\infty, \infty), (-\infty, \infty)) &= 0 \end{aligned}$$

The extended open half plane is divided into 4 strata, which are $\mathbb{H}^\circ$ intersected with the following spaces

$$\mathbb{R} \times \mathbb{R}, \{-\infty\} \times \mathbb{R}, \mathbb{R} \times \{\infty\}, \{-\infty\} \times \{\infty\}.$$

Distances under d^∞ between points in different strata are defined to be infinite. We relate points to the diagonal as follows:

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

Proposition 3.5.10. *Let $M = [p, q), N = [r, s)$. Then*

$$d_I(M, N) \leq d^\infty((p, q), (r, s)).$$

Proof. Let p, q, r, s be finite. We must show that if $\delta > \max(|p-r|, |q-s|)$ then M, N are δ-interleaved. Define families of linear maps

$$\begin{aligned} \phi &= (\phi_t : M_t \to N_{t+\delta}) \\ \psi &= (\psi_t : N_t \to M_{t+d}) \end{aligned}$$

by

$$\phi_t = \begin{cases} \text{id} & M_t = N_{t+\delta} = k \\ 0 & \text{otherwise} \end{cases}$$

and ψ_t is defined similarly. We have to show that these satisfy the interleaving relations (3.2). We first need to show that ϕ and ψ are in fact persistent module homomorphisms. Showing ϕ is a homomorphism amounts to showing that the following diagram commutes

$$\begin{array}{ccc} M_t & \longrightarrow & M_{t+\eta} \\ {\scriptstyle \phi_t}\downarrow & & \downarrow{\scriptstyle \phi_{t+\delta}} \\ N_{t+\delta} & \longrightarrow & N_{t+\eta+\delta} \end{array}$$

This clearly commutes as long as the two following criteria are met

$$\begin{aligned} N_{t+\delta} = 0 &\Leftrightarrow \text{ one of the other modules } = 0 \\ M_{t+\eta} = 0 &\Leftrightarrow \text{ one of the other modules } = 0 \end{aligned}$$

Consider the case $N_{t+\delta} = 0$. If none of the others are 0, then $p \leq t$. Also we have $t + \delta < r$ or $s \leq t + \delta$. The latter cannot be true since by assumption we have $r \leq t + \eta + \delta < s$, therefore the former must be the case. This implies that $t - p + \delta < r - p \implies \delta < r - p$. But $\delta > r - p$ by assumption, so this is impossible. Similarly, if $M_{t+\eta} = 0$ and none of the other modules are zero, then $q \leq t + \eta, t + \eta + \delta \leq s$. But $\delta > s - q$, so this is impossible. Therefore ϕ is a module homomorphism, and by a similar argument so is ψ.
We also require the following relations to be satisfied:

$$\psi \circ \phi = 1_M^{2\delta}, \quad \phi \circ \psi = 1_N^{2\delta}$$

The first is equivalent to the following diagram commuting

$$\begin{array}{ccccc} M_t & & \longrightarrow & & M_{t+2\delta} \\ & {\scriptstyle \phi_t}\searrow & & \nearrow{\scriptstyle \psi_{t+\delta}} & \\ & & N_{t+\delta} & & \end{array}$$

This diagram commutes as long as we don't have $N_{t+\delta} = 0$ and $M_t = M_{t+2\delta} = k$. If that were the case, the top row would imply $p \leq t$ and $t + 2\delta < q$. Now note that if $r \leq p$, then since $N_{t+\delta} = 0$ and $M_{t+2\delta} = k$, then $s \leq t + \delta < t + 2\delta < q$. However this is impossible since $\delta > |q - s|$. So $r > p$. By a similar argument, $q > s$. So we have $\delta > r - p$ and $\delta > q - s$. Therefore must be true that

$$r \leq p + \delta \leq t + \delta \leq q - \delta < s$$

and so $N_{t+\delta} = k$, a contradiction. It follows that $\psi \circ \phi = 1_M^{2\delta}$, and similarly $\phi \circ \psi = 1_N^{2\delta}$. The case where p, q, r, s, t are allowed to be infinite is similar. $\square$

Note that this proposition provides another proof of the statement in example 3.5.8, since the necessary $<$ terms can be changed to $\leq$, and nothing in the proof changes.

Proposition 3.5.11. *Let $M = [p, q)$, and let 0 denote the zero module. Then*

$$d_I(M, 0) = \frac{1}{2}(q - p)$$

Proof. Let $\delta \geq 0$. If there is a δ-interleaving, then the interleaving maps must be zero maps. So $\psi \circ \phi = 1_M^{2\delta} = 0$. This is true whenever $\delta > \frac{1}{2}(q - p)$. □

Definition 3.5.12. Given two sets A, B, a partial matching between them is a collection of pairs $\Gamma \subset A \times B$ such that

- For every $\alpha \in A$ there is at most one $\beta \in B$ s.t. $(\alpha, \beta) \in \Gamma$.
- For every $\beta \in B$ there is at most one $\alpha \in A$ s.t. $(\alpha, \beta) \in \Gamma$.

Such a partial matching is called a δ-matching if in addition the following is true

- $(\alpha, \beta) \in \Gamma \implies d^\infty(a, b) \leq \delta$
- If $\alpha \in A$ is unmatched (so there is no $\beta \in B$ s.t $(\alpha, \beta) \in \Gamma$), then $d^\infty(\alpha, \Delta) \leq \delta$.
- If $\beta \in B$ is unmatched, then $d^\infty(\beta, \Delta) \leq \delta$.

In practice we are interested in δ-matchings between diagrams of persistence modules, which are multisets rather than sets. We can go from multisets to sets by distinguishing each occurence of an element in the multiset. For example, if $x \in X$ occurs with multiplicity n, this gives rise to n distinct elements $x_1, \ldots, x_n$ in X considered as a set.

Definition 3.5.13. The bottleneck distance between two multisets A, B in $\mathbb{H}^\circ$ is defined as

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

Proposition 3.5.14 ([1])**.** *The bottleneck distance satisfies the triangle inequality, i.e.*

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

for three multisets A, B, C.

Theorem 3.5.15. *Let M, N be decomposable persistence modules. Then*

$$d_I(M, N) \leq d_b(\text{dgm}(M), \text{dgm}(N))$$

Proof. Let Γ be a δ-matching between $\text{bar}(M), \text{bar}(N)$. Now each point in the barcodes correspond to an interval summand in the decomposition of the persistence modules. Write

$$M = \bigoplus_{j \in J} M^j, \quad N = \bigoplus_{j \in J} N^j$$

so that each pair (M^j, N^j) is either:

- A pair of matched intervals under the δ-matching M.
- M^j is unmatched and $N^j = 0$.
- N^j is unmatched and $M^j = 0$.

Then $d_I(M^j, N^j) \leq \delta$ for all $j \in J$ by Proposition 3.5.10. Then by Proposition 3.5.9, $d_I(M, N) \leq \delta$, as required. □

3.5.3 Persistence measure

Definition 3.5.16. Let M be an $\mathbb{R}$-persistence module. Then the persistence measure of M, denoted μ_M, is the function

$$\mu_M(R) = \text{multiplicity of } [b,c) \text{ in } M_{\{a,b,c,d\}}$$

where $R = [a,b] \times [c,d] \subset \mathbb{R}^2$ with $a < b < c < d$.

Proposition 3.5.17. *Let $M = [p,q)$ and $R = [a,b] \times [c,d]$ where $a < b < c < d$. Then*

$$\mu_M(R) = \begin{cases} 1 & [b,c] \subset [p,q) \subset (a,d) \\ 0 & \textit{otherwise} \end{cases}$$

Proof. It is clear that $M_{\{a,b,c,d\}}$ is either an interval or 0. Thus $\mu_M(R) \leq 1$. Moreover, $\mu_M(R) = 1$ precisely when $M_{\{a,b,c,d\}}$ is 0 on a,d and k on b,c. This happens iff $[b,c] \subset [p,q] \subset (a,d)$. Note that we stated the theorem for a half open interval M, but the proof clearly also works when M is open or closed. □

Corollary 3.5.18. *Let M be a decomposable $\mathbb{R}$-persistence module. Then $\mu_M(R) = |\operatorname{bar}(M) \cap R|$*

Proposition 3.5.19 ([1])**.** *Let M be a persistence module, $a < b < c < d$. If $r_b^c < \infty$, and $R = [a,b] \times [c,d]$. Then*

$$\mu_M(R) = r_b^c - r_a^c - r_b^d + r_a^d$$

The following proposition is the reason why we call μ_M a measure.

Proposition 3.5.20 ([1])**.** *μ_M is additive under horizontal and vertical splitting of rectangles, i.e.*

$$\mu_M([a,b] \times [c,d]) = \mu_M([a,p] \times [c,d]) + \mu_M([p,b] \times [c,d])$$
$$\mu_M([a,b] \times [c,d]) = \mu_M([a,b] \times [c,q]) + \mu_M([a,b] \times [q,d])$$

where $a < p < b < c < q < d$.

Since persistence diagrams sit in $\mathbb{H}^\circ$, we also want to be able to talk about intervals with $a = -\infty$ and $d = \infty$. We do so by setting

$$M_{-\infty} = 0, M_\infty = 0 \text{ for any } \mathbb{R}\text{-module } M.$$

Then using Proposition 3.5.19 gives

$$\mu_M([-\infty,b] \times [c,\infty]) = r_b^c$$
$$\mu_M([a,b] \times [c,\infty]) = r_b^c - r_a^c$$
$$\mu_M([-\infty,b] \times [c,d]) = r_b^c - r_b^d$$

3.6 The isometry theorem

Theorem 3.6.1 (The Isometry Theorem)**.** *Let M, N be q-tame persistence modules. Then*

$$d_I(M,N) = d_b(\operatorname{bar}(M), \operatorname{bar}(N))$$

The isometry theorem is in fact a combination of two theorems; the stability theorem and the converse stability theorem.

3.6.1 The stability theorem

Theorem 3.6.2 (The Stability Theorem)**.** *Let M, N be q-tame persistence modules. Then*

$$d_I(M, N) \geq d_b(\mathrm{bar}(M), \mathrm{bar}(N))$$

This theorem, along with the converse stability theorem will give us the Isometry Theorem. The proof will follow closely the strategy of [1], [5]. The original proof [9] considered only the case where the persistence modules arose from the sublevel sets of functions $f, g : X \to \mathbb{R}$ where X is a triangulable topological space and f, g were continuous and had only finitely many critical values. This was generalised by [5] to the case which we are interested in, namely for q-tame persistence modules. It was further generalised in [1] to a statement about rectangular measures, a notion introduced in the same paper. Here we will provide a proof in the same generality as [5] while avoiding some technical tools introduced in [5], [1]. The primary tools we need are the Interpolation lemma, which we have already proven, and the box lemma (Lemma 3.6.5).

Definition 3.6.3. Let $R = [a, b] \times [c, d]$ be a rectangle in $\mathbb{H}^\circ$. The δ-thickening of R is the rectangle

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

Given the above situation, we write

$$A = a - \delta, \quad B = b + \delta, \quad C = c - \delta, \quad D = d + \delta$$

Lemma 3.6.4 (Box lemma [1], [9])**.** *Let M, N be a pair of δ-interleaved persistence modules. Let R be a rectangle whose δ-thickening R^δ lies above the diagonal. Then $\mu_M(R) \leq \mu_N(R^\delta)$, and similarly $\mu_N(R) \leq \mu_M(R^\delta)$.*

This result extends to rectangles whose δ-thickening doesn't lie above the diagonal by setting $\mu_M(R^\delta) = \infty$ in such cases. We are now ready to prove the stability theorem.

proof of The Stability Theorem. By the interpolation lemma, we know that there exists an interpolating family of persistence modules $(M^x | x \in [0, \delta])$ such that $M^0 = M, M^\delta = N$ and M^x, M^y are $|y - x|$-interleaved.
We must split the proof into 4 cases; a case for each strata of $\mathbb{H}^\circ$. Assume for the moment that $\mathrm{bar}(M), \mathrm{bar}(N) \subset \mathbb{H}^\circ \cap \mathbb{R}^2$.

Claim 3.6.5. *The following two properties hold:*

- *If $\alpha \in \mathrm{bar}(M^x)$ and $d^\infty(\alpha, \Delta) > |x - y|$, then there exists a $\beta \in \mathrm{bar}(M^y)$ with $d^\infty(\alpha, \beta) \leq |y - x|$.*
- *If $\beta \in \mathrm{bar}(M^y)$ and $d^\infty(\beta, \Delta) > |x - y|$, then there exists an $\alpha \in \mathrm{bar}(M^x)$ with $d^\infty(\alpha, \beta) \leq |y - x|$.*

Proof. We prove only the first statement, since the second follows by symmetry. First, some notation: Let $S_\delta(\alpha)$ denote the square centred at α with side length 2δ. In other words, if $\alpha = (p, q)$ then:

$$S_\delta(\alpha) = [p - \delta, p + \delta] \times [q - \delta, q + \delta].$$

Also let $\eta = |x - y|$. Now, given an α as in the claim, let $\varepsilon > 0$ be small enough such that $\eta + \varepsilon < d^\infty(\alpha, \Delta)$. Then the box lemma gives

$$1 \leq \mu_{M^x}(S_\varepsilon(\alpha)) \leq \mu_{M^y}(S_{\eta+\varepsilon}(\alpha)).$$

Since this is true for all sufficiently small $\varepsilon > 0$, there is at least one point $\beta \in \mathrm{bar}(M^y)$ which is contained in $S_\eta(\alpha)$. □

Claim 3.6.6. *The theorem holds if* $\mathrm{bar}(M^x)$ *has finite cardinality for each* $x \in [0, \delta]$.

Proof. We will prove this using the following fact:

- We claim that for every $x \in [0, \delta]$, there is a real number $\delta_x > 0$ such that $\mathrm{bar}(M^x), \mathrm{bar}(M^y)$ are $|y - x|$-matched whenever $|y - x| < \delta_x$.

To show this, we first enumerate the distinct elements of $\mathrm{bar}(M^x)$ by $\alpha_1, \alpha_2, \ldots \alpha_k$, and we denote their respective multiplicities by $n_1, n_2, \ldots n_k$. Then we can define

$$\delta_x = 1/2 \min(\{d^\infty(\alpha_i, \Delta) | \alpha_i \in \mathrm{bar}(M^x)\} \cup \{d^\infty(\alpha_i, \alpha_j) | \alpha_i, \alpha_j \in \mathrm{bar}(M^x)\})$$

Now assume $|y - x| < \delta_x$. We want to show that $\mathrm{bar}(M^x), \mathrm{bar}(M^y)$ are $|y - x|$-matched. Let $\eta = |y - x|$, and define

$$\Delta^\eta := \{\alpha \in \mathbb{H}^\circ | d^\infty(\alpha, \Delta) \leq \eta\}.$$

By the claim 3.6.6, we know that $\mathrm{bar}(M^y)$ is contained in

$$\Delta^\eta \cup S_\eta(\alpha_1) \cup \cdots \cup S_\eta(\alpha_k).$$

By the definition of δ_x, the terms in the above union are all disjoint. Now let $\varepsilon > 0$ be small enough such that $2\eta + \varepsilon < 2\delta_x$. Then by the box lemma we have

$$n_i = \mu_{M^x}(S_\varepsilon(\alpha_i)) \leq \mu_{M^y}(S_{\eta+\varepsilon}(\alpha_i)) \leq \mu_{M^x}(S_{2\eta+\varepsilon}(\alpha_i)) = n_i.$$

Since the two outside terms are equal, they must also be equal to the middle term. Thus $S_\eta(\alpha_i)$ contains the same number of points (with multiplicity) of $\mathrm{bar}(M^X)$ as $\mathrm{bar}(M^y)$. By mapping each of the n_i copies of α_i to a point of $\mathrm{bar}(M^y)$ that is contained in $S_\eta(\alpha_i)$, we can define a bijection, and in fact an η-matching between the two multisets.
Equipped with this statement, we are ready to prove that the theorem holds in the finite case. Let

$$m = \sup\{x \in [0, \delta] | \,\mathrm{bar}(M^0), \mathrm{bar}(M^x) \text{ are } x - \text{matched}\}$$

We wish to show that $m = \delta$. First of all, if $\delta_0 \geq \delta$, then $m = \delta$. So assume $\delta_0 < \delta$. It is clear that m is positive since $m \geq \delta_0$. Let $0 < m' < m$ be such that $\mathrm{bar}(M^0), \mathrm{bar}(M^{m'})$ are m'-matched and $m - m' < \delta_m$. Then applying the triangle inequality, we see that $\mathrm{bar}(M^0), \mathrm{bar}(M^m)$ are m-matched. Now suppose that $m < \delta$. We again apply fact 1 to $0 < m < m''$ where $m'' - m < \delta_m$ to deduce that $\mathrm{bar}(M^0)\,\mathrm{bar}(M^{m''})$ are m''-matched. However this contradicts the maximality of m, thus we must have $m = \delta$, and thus $\mathrm{bar}(M^0), \mathrm{bar}(M^\delta)$ are δ-matched. □

Claim 3.6.7. *The theorem holds in the general case, i.e. the persistence diagrams need not have finite cardinality.*

Proof. Let $\mathcal{U} = (U_n)_{n \in \mathbb{N}}$ be an open cover of $\mathbb{H}^\circ \cap \mathbb{R}^2$ such that $U_n \subset U_{n+1}$ and each U_n has compact closure. Since M^0, M^δ are q-tame, we have

$$r_b^c = \mu_M([-\infty, b] \times [c, \infty]) < \infty$$

and it follows that $|\,\mathrm{bar}(M^0) \cap U_n|, |\,\mathrm{bar}(M^\delta) \cap U_n|$ are finite for all $n \in \mathbb{N}$. Denote by $M^{0,n}, M^{\delta,n}$ the persistence modules which have $\mathrm{bar}(M^0) \cap U_n$, $\mathrm{bar}(M^\delta) \cap U_n$ as their persistence diagrams, respectively. Then we can choose the open cover $\mathcal{U}$ such that the interleaving maps between M^0, M^δ restrict to δ-interleaving maps between $M^{0,n}, M^{\delta,n}$. So by the previous claim, we have a δ-matching between

$\text{bar}(M^0) \cap U_n, \text{bar}(M^\delta) \cap U_n$ for every $n \in N$. Call this matching Γ_n. We can now take a limit of these matchings to get a δ-matching Γ between $\text{bar}(M^0), \text{bar}(M^\delta)$ as follows: We have indicator functions

$$\chi : \text{bar}(M^0) \times \text{bar}(M^\delta) \to \{0, 1\}$$
$$\chi_n : \text{bar}(M^{0,n}) \times \text{bar}(M^{\delta,n}) \to \{0, 1\}$$

for Γ and Γ_n respectively. Take a fixed enumeration $((\alpha_j, \beta_j))_{j \in J}$ of the countable set $\text{bar}(M^0) \times \text{bar}(M^\delta)$. We will construct a descending chain

$$\mathbb{N} = N_0 \supset N_1 \supset N_2 \supset \cdots$$

of infinite subsets of $\mathbb{N}$ such that $\chi_n(\alpha_j, \beta_j)$ takes the same value for all $n \in N_j$. We then define $\chi(\alpha_j, \beta_j)$ to be this common value. We will then define Γ via this indicator function.
We define N_j inductively: Once N_{j-1} is defined at least one of the two sets

$$\{n \in N_{j-1} | \chi_n(\alpha_j, \beta_j) = 0\} \text{ and } \{n \in N_{j-1} | \chi_n(\alpha_j, \beta_j) = 1\}$$

has infinite cardinality, and that will be our N_j.

Claim 3.6.8. *If $F \subset \text{bar}(M^0) \times \text{bar}(M^\delta)$ is finite, then there is a $j \geq 1$ s.t.*

$$\chi(\alpha, \beta) = \chi_n(\alpha, \beta) \forall (\alpha, \beta) \in F, \forall n \in N_j$$

Proof. Choose j *s.t.* $\{(\alpha_1, \beta_1), \ldots, (\alpha_j, \beta_j)\}$ contains F as a subset □

To see that this Γ is in fact a δ-matching, we first note that $(\alpha, \beta) \in \Gamma \implies d^\infty(\alpha, \beta) < \delta$ since this is true in every Γ_n. To see that each α is matched with at most one β, suppose $\chi(\alpha, \beta) = \chi(\alpha, \beta') = 1$ for $\beta \neq \beta' \in \text{bar}(M^\delta)$. By the Claim 3.6.9, there is some n such that $\chi_n(\alpha, \beta) = \chi_n(\alpha, \beta') = 1$, which contradicts the assumption that Γ_n is a partial matching. Similarly each β is matched with at most one α.
Now suppose that $\alpha \in M^0$ with $d^\infty(\alpha, \Delta) > \delta$. The square $S_\delta(\alpha)$ is compact in $\mathbb{H}^\circ \cap \mathbb{R}^2$, and is therefore contained in U_n for large enough n and so α is matched in Γ_n for sufficiently large n. Now α has only finitely many neighbours within a distance δ, call them $\beta_1, \ldots, \beta_k$, in the locally finite set $\text{bar}(M^\delta)$, so by the Claim 3.6.9 there are infinitely many $n \in N$ such that $\chi(\alpha, \beta_i) = \chi_n(\alpha, \beta_i)$ for every $1 \leq i \leq k$. This implies that for large enough n

$$\chi(\alpha, \beta_i) = \chi_n(\alpha, \beta_i) = 1$$

for some i. Thus α is matched. By symmetry, the necessary statements for β are also true. Thus Γ is the required δ-matching □

This completes the proof when $\text{bar}(M), \text{bar}(N) \subset \mathbb{H}^\circ \cap \mathbb{R}^2$. At the point $(-\infty, \infty)$, we define

$$\mu_M((-\infty, \infty)) = \lim_{e \to \infty} \mu_M([-\infty, -e] \times [e, \infty]) = \min_e \mu_M([-\infty, -e] \times [e, \infty])$$

We then have a box lemma at the point at infinity.

Lemma 3.6.9. *Let M, N be persistence modules which satisfy the box lemma. Then*

$$\mu_M((-\infty, \infty)) = \mu_N((-\infty, \infty)).$$

It follows that the points of $\text{bar}(M^0), \text{bar}(M^\delta)$ which are at the point at infinity get matched with each other, and apart from that the matching Γ is unchanged.
For the cases of $\mathbb{H}^\circ \cap \{-\infty\} \times \mathbb{R}$ and $\mathbb{H}^\circ \cap \mathbb{R} \times \{\infty\}$, the same basic proof works once we introduce some modifications. Similarly to above, we define

$$\mu_{M^x}(-\infty, [c,d]) = \lim_{b \to -\infty} \mu_{M^x}([-\infty, b] \times [c,d]) = \min_b \mu_{M^x}([-\infty, b] \times [c,d])$$
$$\mu_{M^x}([a,b], \infty) = \lim_{c \to \infty} \mu_{M^x}([a,b] \times [c,\infty]) = \min_c \mu_{M^x}([a,b] \times [c,\infty]).$$

Note that since $R \subset S \implies \mu_M(R) \leq \mu_M(S)$, the above limits and the limit in the definition of $\mu_{M^x}(-\infty, \infty)$ always exist. We also have a box lemma for the lines at infinity.

Lemma 3.6.10. *Let M, N be persisence modules which satisfy the box lemma. Let a, b, c, d, A, B, C, D be as before. Then we have the following "box inequalities" at the lines at infinity:*

$$\mu_M(-\infty, [c,d]) \leq \mu_N(-\infty, [C,D])$$
$$\mu_M([a,b], \infty) \leq \mu_N([A,B], \infty).$$

We also have the same inequalities when you swap the measures, i.e.

$$\mu_N(-\infty, [c,d]) \leq \mu_M(-\infty, [C,D])$$
$$\mu_N([a,b], \infty) \leq \mu_M([A,B], \infty).$$

Now the only changes that need to be made to the above proof is to replace the square $S_\eta(\alpha)$ with intervals $[\alpha - \eta, \alpha + \eta]$. □

The Stability Theorem shows that persistence modules are stable in the sense that persistence modules which are close in the interleaving distance have similar barcodes.

3.6.2 The converse stability theorem

We now prove the converse stability theorem, following [1]. There will be some statements given are required for the proof of the converse stability theorem, none of which we will need again.
We begin by approximating persistent modules by better behaved ones via a procedure called smoothing.

Definition 3.6.11. Let M be a persistence module and $\varepsilon > 0$ a real number. Then the ε-smoothing of M, denoted M^ε is the image of the persistence module homomorphism

$$1_M^{2\varepsilon} : M[-\varepsilon] \to M[\varepsilon]$$

Proposition 3.6.12. *Let M be a persistence module and $\varepsilon > 0$. Then*

$$d_I(M, M^\varepsilon) \leq \varepsilon$$

Proof. We have the following factorisation of $1_M^{2\varepsilon}$:

$$M[-\varepsilon] \to M^\varepsilon \to M[\varepsilon]$$

Where the first map is surjective and the second map is injective at each index t. These maps define an ε-interleaving between M, M^ε. □

Example 3.6.13. Let $M = [p, q)$. Then

$$M^{\varepsilon} = \begin{cases} [p+\varepsilon, q-\varepsilon) & p+\varepsilon < q-\varepsilon \\ 0 & \text{otherwise} \end{cases}$$

This is clear from the definition of M^{ε}. ♢

Proposition 3.6.14. *The barcode of M^{ε} is obtained from* bar(M) *by applying the translation*

$$T_{\varepsilon} : (p, q) \to (p+\varepsilon, q-\varepsilon)$$

to the part of $\mathbb{H}^{\circ}$ which lies above the diagonal $\{(t-\varepsilon, t+\varepsilon)\}$, and is constant everywhere else.

Corollary 3.6.15. *Let M be a q-tame persistence module. Then*

$$d_b(\text{bar}(M), \text{bar}(M^{\varepsilon})) \leq \varepsilon$$

Proof. An ε-matching between the persistence diagrams is given by

$$\text{bar}(M) \to \text{bar}(M^{\varepsilon}), \quad (p, q) \mapsto (p+\varepsilon, q-\varepsilon)$$

This is a bijection apart from the unmatched points of bar(M). All such points lie on or below the line $\{(t-\varepsilon, t+\varepsilon) | t \in \mathbb{R}\}$. Such points lie within a distance ε of the diagonal, and therefore vanish under the above map. □

Proposition 3.6.16 ([1])**.** *If a persistence module M is q-tame then M^{ε} is locally finite. In particular, M^{ε} has an interval decomposition.*

Theorem 3.6.17 (Converse Stability Theorem)**.** *Let M, N be q-tame persistence modules. Then*

$$d_I(M, N) \leq d_b(\text{bar}(M), \text{bar}(N))$$

Proof. M, N are q-tame $\implies$ $M^{\varepsilon}, N^{\varepsilon}$ are decomposable. So by theorem 3.5.15, we know that $d_I(M^{\varepsilon}, N^{\varepsilon}) \leq d_b(\text{bar}(M^{\varepsilon}), \text{bar}(N^{\varepsilon}))$. Then

$$\begin{aligned} d_I(M, N) &\leq d_I(M, M^{\varepsilon}) + d_I(M^{\varepsilon}, N) \\ &\leq d_I(M, M^{\varepsilon}) + d_I(M^{\varepsilon}, N^{\varepsilon}) + d_I(N^{\varepsilon}, N) \\ &\leq d_I(M^{\varepsilon}, N^{\varepsilon}) + 2\varepsilon \\ &\leq d_b(\text{bar}(M^{\varepsilon}), \text{bar}(N^{\varepsilon})) + 2\varepsilon \\ &\leq d_b(\text{bar}(M^{\varepsilon}), \text{bar}(M)) + d_b(\text{bar}(M), \text{bar}(N^{\varepsilon})) + 2\varepsilon \\ &\leq d_b(\text{bar}(M^{\varepsilon}), \text{bar}(M)) + d_b(\text{bar}(M), \text{bar}(N)) + d_b(\text{bar}(N), \text{bar}(N^{\varepsilon})) + 2\varepsilon \\ &\leq d_b(\text{bar}(M), \text{bar}(N)) + 4\varepsilon \end{aligned}$$

Where above we have used the fact that $M^{\varepsilon}, N^{\varepsilon}$ satisfy the converse stability theorem, the triangle inequality for both the bottleneck distance and the interleaving distance, and corollary 3.6.16. Since this is true for all $\varepsilon > 0$, we have

$$d_I(M, N) \leq d_b(\text{bar}(M), \text{bar}(N))$$

as required. □

Finally, there is a characterisation of q-tame persistence modules in terms of locally finite modules.

Definition 3.6.18. A persistence module M is said to be approximated in the interleaving distance by a family of persistence modules $(M^i)_{i\in I}$ if for all $\varepsilon > 0$, there is an $i \in I$ such that

$$d_I(M, M^i) \leq \varepsilon$$

Theorem 3.6.19. *A persistence module M is q-tame if and only if it can be approximated in the interleaving distance by a family of locally finite modules.*

Proof. If M is q-tame then it is approximated by modules of the form M^ε, which are all locally finite. For the converse, suppose that M is approximated by locally finite modules. Fix $b < c$ and let L be a locally finite module such that $d_I(M, L) \leq \varepsilon$ for some $\varepsilon < (c - b)/2$. Then

$$\begin{aligned} r_b^c &= \operatorname{rank}(M_b \to M_c) \\ &= \operatorname{rank}(M_b \to L_{b+\varepsilon} \to L_{c-\varepsilon} \to M_c) \\ &\leq \dim(L_{b+\varepsilon}) \\ &< \infty \end{aligned}$$

□

Chapter 4

Generalized Persistence

Definition 4.0.1. Let $(P, \leq)$ be a poset, and C an arbitrary category. Then a generalized persistence module (referred to simply as a persistence module in this chapter) is a functor

$$M : P \to C.$$

Such a persistence module is called a persistence module in C over P. The collection of persistence modules in C over P is itself a category, denoted C^P.

In this section we will discuss generalized persistence modules. We first introduce a generalization of the interleaving and then of the persistence diagram that applies in a limited case. The former will allow us to define the interleaving for multidimensional persistence modules in chapter 5, and the latter will show that we can sometimes consider persistence modules over rings that are not fields.

Definition 4.0.2. A translation on a poset P is a function $T : P \to P$ which is monotone and satisfies $x \leq T(x)$ for all $x \in P$. In particular, a translation on P is a functor $T : P \to P$ for which there is a natural transformation $\mathrm{id} \implies T$. Since all morphisms in P are unique, this natural transformation is unique. We denote the set of all translations on P by Trans_P.

Example 4.0.3. The map

$$T_\delta : \mathbb{R} \to \mathbb{R} \quad t \mapsto t + \delta$$

is a translation on $(\mathbb{R}, \leq)$. ◊

Example 4.0.4 (Sublevel set)**.** Given a function $f : X \to \mathbb{R}$ on a topological space X, we have a generalized persistence module given by

$$(X, f) : \mathbb{R} \to X, \quad t \to f^{-1}((-\infty, t]).$$

Then the sublevel set persistent homology that we have encoutered before is the composite of two generalized persistence modules

$$\mathbb{R} \xrightarrow{(X,f)_{(-)}} Top \xrightarrow{H_k(-)} \mathrm{Vect}_k\,.$$

◊

Lemma 4.0.5. *Let $F, G : C \to P$ be two functors into a poset P viewd as a category. Then there is a natural transformation $\phi : F \to G \Leftrightarrow F(x) \leq G(x)\ \forall x \in C$. This natural transformation is unique when it exists.*

Proof. There exists a family of maps $(\phi_x) : F(x) \to G(x)$ if and only if $F(x) \leq G(x)$ for all $x \in C$. If such a family exists, then the following diagram automatically commutes since all maps in P are unique:

$$\begin{array}{ccc} F(x) & \xrightarrow{F(\alpha)} & F(y) \\ {\scriptstyle \phi_x}\downarrow & & \downarrow{\scriptstyle \phi_y} \\ G(x) & \xrightarrow[G(\alpha)]{} & G(y). \end{array}$$

And so $(\phi_x)_{x\in P}$ is a natural transformation. Since all the maps are unique, this natural transformation is also unique. □

Given a poset $(P, \leq)$, the set Trans_P has the following structure:

- It is a monoid with respect to composition, where the identity functor is the identity.
- It is a poset with respect to the relation

$$T \leq T' \Leftrightarrow T(x) \leq T'(x)\ \forall x \in P \Leftrightarrow \exists \text{ a natural transformation } T \to T'.$$

Definition 4.0.6. Let $(P, \leq)$ be a poset, $T, T' \in \mathrm{Trans}_P, C$ an arbitrary category, and $M, N \in C^P$ two persistence modules. Then a $(T, T')-$interleaving between M, N is a pair of natural transformations

$$\phi : M \to N \circ T, \quad \psi : N \to M \circ T'$$

such that

$$\psi \circ \phi = M \circ \eta_{T' \circ T}, \quad \phi \circ \psi = N\eta_{T \circ T'}. \tag{4.1}$$

Where we let η_F denote the unique natural transformation $\mathrm{id} \to F$. Explicitly, we have the following commuting diagrams for all $x \leq y \in P$

$$\begin{array}{ccc} M(x) & \longrightarrow & M(y) \\ {\scriptstyle \phi_x}\searrow & & \searrow{\scriptstyle \phi_y} \\ N(Tx) & \longrightarrow & N(Ty) \end{array} \qquad \begin{array}{ccc} & M(T'x) & \longrightarrow & M(T'y) \\ {\scriptstyle \psi_x}\nearrow & & {\scriptstyle \psi_y}\nearrow & \\ N(x) & \longrightarrow & N(y) & \end{array}$$

$$\begin{array}{ccc} M(x) & \longrightarrow & M(T' \circ Tx) \\ {\scriptstyle \phi_x}\searrow & & \nearrow{\scriptstyle \psi_{Tx}} \\ & N(Tx) & \end{array} \qquad \begin{array}{ccc} & M(T'x) & \\ {\scriptstyle \psi_x}\nearrow & & \searrow{\scriptstyle \phi_{T'x}} \\ N(x) & \longrightarrow & N(T \circ T'x) \end{array}$$

A (T, T) interleaving is simply called a T-interleaving.

Example 4.0.7. Let M, N be $\mathbb{R}$-persistence modules in the normal (non-generalized) sense. A morphism $M \to N \circ T_\delta$ is a morphism in $\mathrm{Hom}^\delta(M, N)$, and a T_δ-interleaving is a δ-interleaving. ◊

Lemma 4.0.8. *Let P, Q be posets, $M, N : P \to Q$ persistence modules in Q^P and $T, T' \in \mathrm{Trans}_P$. Then*

$$M, N \textit{ are } (T, T') - \textit{interleaved} \Leftrightarrow M(x) \leq N \circ T(x), N(x) \leq M \circ T'(x)\ \forall x \in P.$$

Proof. The existence of the maps $\phi_x : M(x) \to N \circ T(x), \psi_x : N(x) \to M \circ T'(x)$ is equivalent to the inequalities. All the relevant diagrams commute since all the maps are unique. □

The interleaving relation is functorial, monotone and satisfies a triangle inequality.

Proposition 4.0.9 (Functoriality)**.** *Let P be a poset, $T, T' \in Trans_P$, C, D be arbitrary categories, $F : C \to D$ a functor and $M, N \in C^P$. Then M, N are (T, T')-interleaved $\implies$ FM, FN are (T, T')-interleaved.*

Proposition 4.0.10 (Monotonicity)**.** *Let P be a poset, $T_1, T_2, T_1', T_2' \in \text{Trans}_P$, C a category, and $M, N \in C^P$. Suppose $T_1 \leq T_2, T_1' \leq T_2'$. Then M, N are (T_1, T_1')-interleaved $\implies$ M, N are (T_2, T_2')-interleaved.*

Proposition 4.0.11 (Triangle inequality)**.** *Let P be a poset, $M, N, L \in C^P$ for a category C, and $T_1, T_2, T_1', T_2' \in \text{Trans}_P$. Then M, N are (T_1, T_1')-interleaved, and N, L are (T_2, T_2')-interleaved $\implies$ M, L are $(T_2 \circ T_1, T_2' \circ T_1')$-interleaved.*

4.1 Sublinear projections and superlinear families

In this section we investigate when it makes sense to say that two generalized persistence modules are δ-interleaved for some $\delta \in \mathbb{R}_{\geq 0}$. Fix a poset $(P, \leq)$ and a category C throughout this section.

4.1.1 Sublinear projections

Definition 4.1.1. A sublinear projection is a map $\omega : \text{Trans}_P \to [0, \infty]$ such that

- $\omega_{\text{id}} = 0$.
- $\omega_{T_1 \circ T_2} \leq \omega_{T_1} + \omega_{T_2} \; \forall T_1, T_2 \in \text{Trans}_P$.

A sublinear projection is called monotone if $\omega_T \leq \omega_{T'}$ whenever $T \leq T'$. Given any sublinear projection ω, its monotone hull is the sublinear projection given by

$$\hat{\omega}_T = \inf\{\omega_{T'} | T' \geq T\}.$$

It is clear from this definition that $\hat{\omega}$ is monotone.

Definition 4.1.2. A translation $T \in \text{Trans}_P$ is called a δ-translation with respect to ω if $\omega_T \leq \delta$. Persistence modules $M, N \in C^P$ are δ-interleaved w.r.t. ω if they are (T, T')-interleaved for a pair of δ-translations T, T'.

Definition 4.1.3. The interleaving distance w.r.t. a sublinear projection ω is defined as

$$d^{\omega}(M, N) = \inf\{\delta \in [0, \infty) | M, N \text{ are } \delta\text{-interleaved w.r.t. } \omega\}.$$

If the set on the right hand side is empty, we set $d^{\omega}(M, N) = \infty$.

From now on we fix a sublinear projection $\omega : \text{Trans}_P \to [0, \infty]$. Note that if M, N are δ-interleaved w.r.t ω, then they are $\delta + \epsilon$-interleaved for all $\epsilon \in [0, \infty]$. To see this, just observe that if a translation T is a δ-translation, then it is a $\delta + \epsilon$-translation for all $\epsilon \in [0, \infty]$.

Proposition 4.1.4 ([10])**.** *Let $\hat{\omega}$ be the monotone hull of ω. Then $d^{\omega} = d^{\hat{\omega}}$.*

This proposition guarantees that we can always assume that a sublinear projection is monotone without affecting the resulting interleaving distance. Of course, for any of this discussion to make sense, we need the following theorem.

Theorem 4.1.5. *d^{ω} is a pseudometric.*

Proof. It is clear that d^{ω} is symmetric, and $d^{\omega}(M, M) = 0$ since any persistence module is 0-interleaved with itself. So we just need to check the triangle inequality. Let M, N be δ_1-interleaved, and N, L be δ_2-interleaved. Let T_i, T_i' be δ_i-translations such that M, N are (T_1, T_1')-interleaved and N, L are (T_2, T_2')-interleaved. Then by the triangle inequality for the interleaving relation, M, L are $(T_2 \circ T_1, T_2' \circ T_1')$-interleaved, and $\omega_{T_2 \circ T_1}, \omega_{T_2' \circ T_2'} \leq \delta_1 + \delta_2$, and so the triangle inequality is satisfied. □

4.1.2 Lawvere metrics

Definition 4.1.6. A Lawvere metric space is a set X together with a function $d : X \times X \to [0, \infty]$ such that $d(x, x) = 0 \; \forall x \in X$, and d satisfies the triangle inequality.

Proposition 4.1.7. *Let d be a Lawvere metic on the poset P. Then*

$$\omega_T^d = \sup\{d(x, T(x)) | x \in P\}$$

defines a sublinear projection.

Proof. Clearly $\omega_{\mathrm{id}}^d = 0$ and $\omega_T^d \in [0, \infty] \; \forall T \in \mathrm{Trans}_P$. For sublinearity, we see that

$$d(x, T_1 \circ T_2(x)) \leq d(x, T_1 x) + d(T_1 x, T_2 \circ T_1 x) \leq \omega_{T_1}^d + \omega_{T_2}^d$$

so $\omega_{T_1 \circ T_2}^d \leq \omega_{T_1}^d + \omega_{T_2}^d$. □

So considering theorem 4.1.5, we see that any Lawvere metric on P gives rise to a pseudometric on C^P.

Example 4.1.8. We can derive the regular interleaving distance of $\mathbb{R}$-persistence modules as follows: Let ω be the sublinear projection associated to the standard metric $d(x, y) = |x - y|$ on $\mathbb{R}$. If T is a δ-translation w.r.t. ω, then

$$x \leq T(x) \leq x + \delta$$

for every $x \in \mathbb{R}$, and so $T \leq T_\delta$. Then two $\mathbb{R}$-persistence modules M, N are δ-interleaved w.r.t. ω $\Leftrightarrow$ M, N are (T, T')-interleaved for a pair of δ-translations T, T' $\Leftrightarrow$ M, N are T_δ-interleaved. It follows that d^ω is the usual interleaving distance for $\mathbb{R}$-persistence modules. ♢

4.1.3 Superlinear families

Now we look at the dual approach to sublinear projections.

Definition 4.1.9. A superlinear family over P is a function $\Omega : [0, \infty) \to \mathrm{Trans}_P$ which satisfies $\Omega_{\delta_1} \circ \Omega_{\delta_2} \leq \Omega_{\delta_1 + \delta_2}$

The definition implies that Ω_0 is the identity element when we view a superlinear family as a monoid with the operation of composition. To see this, observe that

$$\Omega_\delta \leq \Omega_0 \circ \Omega_\delta \leq \Omega_\delta$$

and

$$\Omega_\delta \leq \Omega_\delta \circ \Omega_0 \leq \Omega_\delta.$$

We also have that any superlinear family is monotone. Indeed

$$\Omega_{\delta_1} \leq \Omega_{\delta_2 - \delta_1} \circ \Omega_{\delta_1} \leq \Omega_{\delta_2}$$

whenever $\delta_1 \leq \delta_2$. We now fix a superlinear family over the poset P.

Definition 4.1.10. The interleaving distance between $M, N \in C^P$ w.r.t. Ω is defined as

$$d^\Omega(M, N) = \inf\{\delta \in [0, \infty) | M, N \text{ are } \Omega_\delta\text{-interleaved}\}.$$

We let $d^\Omega(M, N) = \infty$ if the set on the right hand side is empty.

Once again we have the key fact that a superlinear family corresponds to a psuedometric on C^P.

Theorem 4.1.11. *d^Ω is a pseudometric on C^P.*

Proof. Again, it is clear that d^Ω is symmetric, and $d^\Omega(M, M) = 0$ since any persistence module is Ω_0-interleaved with itself. For the triangle inequality, let M, N be Ω_{δ_1}-interleaved, and N, L be Ω_{δ_2}-interleaved. Then M, L are $\Omega_{\delta_2} \circ \Omega_{\delta_1}$-interleaved by the triangle inequality for the interleaving relation. This implies that M, L are $\Omega_{\delta_1+\delta_2}$-interleaved since $\Omega_{\delta_2} \circ \Omega_{\delta_1} \leq \Omega_{\delta_1+\delta_2}$ by superlinearity. Then once we apply the monotonicity of the interleaving relation, we get the desired interleaving. □

Theorem 4.1.12. *Let $M, N \in C^P$, D be an arbitrary category, and $F : C \to D$ be a functor. If Ω is a superlinear family over P, then*

$$d^\Omega(FM, FN) \leq d^\Omega(M, N).$$

Proof. If M, N are Ω_δ-interleaved then FM, FN are Ω_δ-interleaved by functoriality. □

We can get a superlinear family from a sublinear projection, as the next theorem shows.

Theorem 4.1.13. *Let ω be a sublinear projection on P. Suppose for every $\delta \geq 0$, there is a translation Ω_δ with the property that $\omega_{\Omega_\delta} \leq \delta$ which is largest in the sense that $\omega_T \leq \delta \implies T \leq \Omega_\delta$. Then $\Omega = (\Omega_\delta)_\delta$ is a superlinear family, and $d^\omega = d^\Omega$.*

Proof. To show superlinearlity, we note that

$$\omega_{\Omega_{\delta_1} \circ \Omega_{\delta_2}} \leq \omega_{\Omega_{\delta_1}} + \omega_{\Omega_{\delta_2}} \leq \delta_1 + \delta_2$$

and this implies $\Omega_{\delta_1} \circ \Omega_{\delta_2} \leq \Omega_{\delta_1+\delta_2}$ by the 'largest' assumption. To see that the two interleaving distances are the same, we have to show that the following statements are eqiuvalent

- M, N are (T, T')-interleaved for δ-tranlations T, T'.
- M, N are Ω_δ-interleaved.

The direction $\Leftarrow$ is trivial. For the other direction, because of the 'largest' assumption, T, T' are δ-translation $\implies T, T' \leq \Omega_\delta$. Then monotonicity of the interleaving relation gives us the result. □

We will see that the connection between superlinear families and sublinear projections goes deeper.

4.1.4 Monoidal structures

If we view the poset P as a category in the usual way, then it is clear that Trans_P is a monoidal category where the tensor product is given by composition. So $T \otimes T' = T \circ T'$ and $(T_1 \leq T_2) \otimes (T_1' \leq T_2') = T_1 \circ T_1' \leq T_2 \circ T_2'$. The tensor unit is id_P.
Similarly, both $[0, \infty]$ and $[0, \infty)$ are monoidal categories with the tensor product given by

$$x \otimes y = x + y, \quad (x \leq y) \otimes (w \leq z) = x + w \leq y + z$$

and the tensor unit is 0.

Lemma 4.1.14. *A super linear family corresponds to a lax monoidal functor*

$$\Omega : ([0, \infty), +, 0) \to (\mathrm{Trans}_P, \circ, \mathrm{id})$$

Proof. The fact that Ω is monotone guarantees that it is a functor. The maps $\Omega_a \circ \Omega_b \leq \Omega_{a+b}$ and the map $\mathrm{id} \leq \Omega_0$, which exists by definition, are the monoidal coherence maps. Since all the maps involved are unique, all the relevant diagrams commute. □

Lemma 4.1.15. *A monotone sublinear projection $\omega : \mathrm{Trans}_P \to [0, \infty]$ corresponds to an oplax monidal functor.*

Proof. The sublinearity relation $\omega_{T \circ T'} \leq \omega_T + \omega_{T'}$ and that $\omega_{\mathrm{id}} = 0$ give the monoidal coherence maps. Choosing ω to be monotone ensures that it is a functor. □

As we have mentioned, requiring that ω be monotone is not a very limiting condition, since we can replace any sublinear projection by its monotone hull and we get the same interleaving distance. We are now ready to define an adjunction between sublinear projections and superlinear families.

Definition 4.1.16. Let $\omega : \mathrm{Trans}_P \to [0, \infty], \Omega : [0, \infty) \to \mathrm{Trans}_P$ be arbitrary functions. We say that $\omega \dashv \Omega$ if

$$\omega_T \leq \delta \Leftrightarrow T \leq \Omega_\delta \; \forall \delta \in [0, \infty), T \in \mathrm{Trans}_P .$$

If we append a terminal object to Trans_P and define Ω_∞ to be this terminal object, then the above relation corresponds with the definition of an adjunction.

Proposition 4.1.17 ([10])**.** *Let $\omega : \mathrm{Trans}_P \to [0, \infty], \Omega : [0, \infty) \to \mathrm{Trans}_P$ be arbitrary functions such that $\omega \dashv \Omega$. Then:*

1. $T \leq \Omega_{\omega_T} \; \forall T \in \mathrm{Trans}_P$.
2. $\omega_{\Omega_\delta} \leq \delta$.
3. *ω is monotone.*
4. *Ω is monotone.*

We also have the following general fact about monoidal functors:

Proposition 4.1.18 ([11])**.** *A functor with a right adjoint is oplax monoidal $\Leftrightarrow$ the right adjoint is a lax monoidal functor.*

Applying the proposition in this context, if we have arbitrary functions $\omega : \mathrm{Trans}_P \to [0, \infty], \Omega : [0, \infty) \to \mathrm{Trans}_P$, such that $\omega \dashv \Omega$, then ω is a sublinear projection $\Leftrightarrow$ Ω is a superlinear family.

Theorem 4.1.19. *If ω is a sublinear projection, and Ω is a superlinear family such that $\omega \dashv \Omega$, then $d^\omega = d^\Omega$.*

Proof. First, recall

$$\begin{aligned} d^\Omega(M, N) &= \inf\{\delta \in [0, \infty) | M, N \text{ are } \Omega_\delta\text{-interleaved}\}. \\ d^\omega(M, N) &= \inf\{\delta | M, N \text{ are } \delta\text{-interleaved w.r.t. } \omega\} \\ &= \inf\{\delta | M, N \text{ are } (T, T')\text{-interleaved where } T, T' \text{ are } \delta\text{-translations w.r.t. } \omega\} \end{aligned}$$

If M, N are (T, T')-interleaved where $\omega_T, \omega_{T'} \leq \delta$, then $T, T' \leq \Omega_\delta$, and so M, N are Ω_δ-interleaved by monotonicity. Thus $d^\Omega(M, N) \leq d^\omega(M, N)$. For the converse inequality, note that if M, N are Ω_δ-interleaved, then $d^\omega(M, N) \leq \delta$ since $\omega_{\Omega_\delta} \leq \delta$. □

So whenever we can define a sublinear projection $\mathrm{Trans}_P \to [0, \infty]$, or a superlinear family $[0, \infty) \to \mathrm{Trans}_P$, we have a way to define the notion of a δ-interleaving.

Chapter 5

Multidimensional Persistence

In chapter 3, we looked at persistence diagrams over subsets of $\mathbb{R}$. We now consider what happens when we index over subsets of $\mathbb{R}^n$.

Definition 5.0.1. A multidimensional persistence module is a functor

$$M : P^n \to \text{Vect}_k$$

where P is a subset of $\mathbb{R}$. In particular it is simply a P^n-persistence module.

While we have stated this definition in terms of any arbitrary subset of $\mathbb{R}$, in practice we will be restricting ourselves to $P = \mathbb{R}, \mathbb{Q}, \mathbb{N}$. We will see that multidimensional persistence modules do not lend themselves to the general treatment that their 1-dimensional cousins do.

5.1 Complications with multidimensional persistence

The ideal situation would be to find a complete invariant for multidimensional persistence, such as the barcode for 1-dimensional persistence. In particular, we want a parameterization which assigns non-isomorphic persistence modules to different points in the parameter space, and isomorphic ones to the same point. We also want this parameter to always map into the same space, i.e. we want it to always return an integer, a vector space, etc. (in the case of the barcode, the parameter space is always a multiset in $\mathbb{H}^\circ$). As it turns out, such an invariant does not exist [15]. Just as in the one-dimensional case, we first wish to use algebraic structures to classify multidimensional persistence modules. Throughout this section, let R denote the polynomial ring $k[x_1, \ldots, x_r]$. Given a vector $v = (v_1, \ldots, v_r) \in \mathbb{N}^r$, we denote by x^v the monomial $x_1^{v_1} \cdots x_r^{v_r}$. R is an r-graded ring with the grading $R_v = kx^v$.

Definition 5.1.1. Given an $\mathbb{N}^r$-persistence module M, we can define an r-graded module over R by

$$\alpha(M) = \bigoplus_{v \in \mathbb{N}^r} M_v$$

where the action $x^{v-u} : M_u \to M_v$ is the map $M(u \leq v)$.

Theorem 5.1.2 ([15])**.** *The correspondence α defines an equivalence of categories between the category of finite $\mathbb{N}^r$-persistence modules over k and the category of finitely generated n-graded modules over R.*

So far our treatment of multidimensional persistence has been a simple extension of the one dimensional case. However this more or less ends here. In the 1-dimensional case, the algebraic object at hand - $k[x]$-modules - had a very simple classification, given by the structure theorem. Unfortunately in the case of r-graded R-modules, the corresponding classification is more complicated. A full discussion of this classification can be seen in [15].

5.2 Multidimensional persistent homology

Let X be a multifiltered simplicial complex, and let $\{X_v\}_{v\in\mathbb{N}^r}$ be a multifiltration of X where each X_v is a simplicial complex. Consider the functor C_n from the category of simplicial complexes to Vect_k which sends a simplicial complex Y to the k-vector space generated by the n-simplices in Y. Applying C_n to the multifiltration $\{X_v\}_v$ gives us a family of vector spaces $\{C_n(X_v)\}_v$ and linear inclusions $\{C_n(X_v \to X_w)\}_{v\leq w}$. These determine an r-graded R-module

$$C_n(X) = \bigoplus_v C_n(X_v)$$

with module action given by

$$x^w \cdot C_n(X_v) = \operatorname{im} C_n(X_v \to X_{v+w}).$$

Definition 5.2.1. The n-chain module of a multifiltered space X is the r-graded R-module $C_n(X)$.

Now consider the functor $H_n(-)$ from simplicial complexes to Vect_k which sends a complex to its nth simplicial homology group with coefficients in k. Applying this functor to the multifiltration $\{X_v\}$, we get a family of vector spaces $\{H_n(X_v)\}$ and linear maps $\{H_n(X_v \to X_w)\}_{v\leq w}$. This determines an r-graded R-module

$$H_n(X) = \bigoplus_v H_n(X_v)$$

with module action given by

$$x^w \cdot H_n(X_v) = \operatorname{im} H_n(X_v \to X_{v+w}).$$

Definition 5.2.2. The n-multipersistent homology module of a multifiltered space X is the module $H_n(X)$.

Let $d = \dim X$ where X is a multifiltered simplicial complex. The modules $C_n(X)$ fit into a chain complex of r-graded R-modules

$$C_\bullet(X): \quad 0 \longrightarrow C_d(X) \xrightarrow{\delta_d} \cdots \longrightarrow C_1(X) \xrightarrow{\delta_1} C_0(X) \longrightarrow 0.$$

This is the direct sum of the corresponding chain complexes $\{C_\bullet(X_v)\}_{v\in\mathbb{N}^r}$. Note that the maps δ_i here are the direct sums of the normal boundary maps $C_n(X_v) \to C_{n-1}(X_v)$. Now let $\sigma \in X_v$ be an n-face, and let $a \in C_n(X_v)$ be the corresponding basis element. Then $x^{e_j} \cdot a = C_n(v \leq v + e_j)(a)$, where e_j is the jth standard basis vector, corresponds to the same face σ since $C_n(v \leq v + e_j)$ is the inclusion map. Thus $x^{e_j} \cdot \delta_n = \delta_n \cdot x^{e_j}$. We know that $\delta_n(C_n(X_v)) \subset C_{n-1}(X_v)$ and $x^{e_j}(C_n(X_v)) \subset C_n(X_{v+e_j})$. So $C_\bullet(X)$ is double graded:

$$C_\bullet(X) = \bigoplus_n C_n(X) = \bigoplus_n \bigoplus_{v\in\mathbb{N}^r} C_n(X_v) = \bigoplus_{v\in\mathbb{N}^r} \bigoplus_n C_n(X_v) = \bigoplus_{v\in\mathbb{N}^r} C(X_v).$$

Proposition 5.2.3 ([15]). *Multidimensional homology modules are the homology modules of the chain complexes $C_\bullet$.*

Proof.

$$\bigoplus_{v\in\mathbb{N}^r} H_n(X_v) = H_n\left(\bigoplus_{v\in\mathbb{N}^r} X_v\right) = H_n(C_\bullet).$$

□

5.3 Searching for invariants

Since the search for a complete invariant on the space of multidimensionanal persistence modules is a fruitless one, we turn our attention to finding invariants that are still useful in applications. Additionally, we would like to produce continuous invariants so that some statistical analysis can be done on the invariants. One of the disadvantages of the barcode as an invariant is that it does not lend itself to such analysis. For example, it does not even make sense to take the average of two barcodes. The key to finding continuous invariants will turn out to be generating pseudometrics on an appropriate subset of persistence modules. From now on, we will follow the literature and focus on persistence modules over $\mathbb{Q}^r_{\geq 0}$ for $r \geq 1$ [17], [18],[19],[20].
Given a set T, in particular where T is a collection of multidimensional persistence modules, we want to produce useful and applicable invariants on this set. A multiset $f \in \text{Mult}(T)$ is an invariant. We wish to "stabilize" f to get a continuous invariant. Given an extended pseudometric on T, $d : T \times T \to \mathbb{R} \cup \infty$, we construct a function $\hat{f} : T \to \text{Mult}(\mathbb{Q}_{\geq 0})$, called the hierarchical stabilization of f [17]. For $x \in T, \tau \in \mathbb{Q}_{\geq 0}$, define

$$\hat{f}(x)(\tau) := \min\{f(y) | d(x,y) \leq \tau\}.$$

Note that if $\tau \leq \varepsilon$, then $\hat{f}(\tau) \geq \hat{f}(\varepsilon)$. Thus $\hat{f}(x)$ can be viewed as a functor $T \to \mathbb{N}^{op}$. We say that $\hat{f}(x), \hat{f}(y)$ are δ-interleaved if $\hat{f}(x)(\varepsilon) \geq \hat{f}(y)(\varepsilon + \delta)$, and $\hat{f}(y)(\varepsilon) \geq \hat{f}(x)(\varepsilon + \delta)$ for all $\varepsilon \in \mathbb{Q}_{\geq 0}$.

Proposition 5.3.1. *Given a function $f : T \to \mathbb{N}$ for some set T, and a pseudometric d on T, the function $\hat{f}$ is 1-Lipschitz, i.e.*

$$d_I(\hat{f}(x), \hat{f}(y)) \leq d(x,y)$$

where d_I is the interleaving distance.

Proof. Assume that $d(x,y) < \infty$, or else there is nothing to prove. Let $\varepsilon, \tau \in \mathbb{Q}_{\geq 0}$ *s.t.* $d(x,y) \leq \varepsilon$. By the triangle inequality, $B(y,\tau) \subset B(x, \tau + \varepsilon)$ and $B(x,\tau) \subset B(y, \tau + \varepsilon)$. Therefore $\hat{f}(y)(\tau) \geq \hat{f}(x)(\tau + \varepsilon)$ and $\hat{f}(x)(\tau) \geq \hat{f}(y)(\tau + \varepsilon)$. In particular, $\hat{f}(x), \hat{f}(y)$ are ε-interleaved. The result follows. □

5.4 Tame functors

We want to define tame functors as functors which have only a discrete number of critical points. To do this, we first take a look at functors in $\text{Fun}(\mathbb{N}_{\geq 0}, \text{Vect}_k)$.

Definition 5.4.1. The radical of a functor $F : \mathbb{N}^r_{\geq 0} \to \text{Vect}_k$, denoted $\text{rad}\, F$, is the functor whose value at v is the subspace of $F(v)$ given by the sum of all the images of the maps $F(u \leq v)$ for all $u \leq v$.

The functor $F/\text{rad}\, F$ is isomorphic to a functor of the form $\bigoplus(U_v \otimes V_v)$, where $\{V_v\}$ is a sequence of vector spaces indexed over $\mathbb{N}^r_{\geq 0}$, and U_v is as in section 2. Functors of such a form are called **semisimple**.

Example 5.4.2. Consider the free functor $F = k((1,0), -) \oplus k((0,1), -) : \mathbb{N}^2 \to \text{Vect}_k$.

$$\begin{array}{ccccc} k & \longrightarrow & k^2 & \longrightarrow & k^2 \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k^2 & \longrightarrow & k^2 \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k \end{array}$$

Its radical is given by the following functor

$$\begin{array}{ccccc} k & \longrightarrow & k^2 & \longrightarrow & k^2 \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k^2 & \longrightarrow & k^2 \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & k \end{array}$$

And the quotient $F/\operatorname{rad} F$ is given by

$$\begin{array}{ccccc} 0 & \longrightarrow & 0 & \longrightarrow & 0 \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & 0 & \longrightarrow & 0 \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & 0 \end{array}$$

$\diamondsuit$

Proposition 5.4.3 ([17])**.** *Any functor* $F : \mathbb{N}^r_{\geq 0} \to \mathrm{Vect}_k$ *has a minimal cover*

This proposition means that it makes sense to consider the rank, support and 0-th Betti diagram of persistence modules over $\mathbb{N}^r$.

Let $\alpha \in \mathbb{Q}$. We have a functor $\alpha : \mathbb{N}^r \to \mathbb{Q}^r$ which is coordinate-wise multiplication by α. We also have a functor

$$\lfloor \alpha^{-1} \rfloor : \mathbb{Q}^r \to \mathbb{N}^r, \quad (v_1, \dots, v_r) \mapsto (\lfloor \frac{v_1}{\alpha} \rfloor, \dots, \lfloor \frac{v_r}{\alpha} \rfloor).$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. $\lfloor \alpha^{-1} \rfloor \alpha : \mathbb{N}^r \to \mathbb{N}^r$ is the identity.

Definition 5.4.4. A functor $G : \mathbb{Q}^r \to \mathrm{Vect}_k$ is called α-tame if it is isomorphic to $F\lfloor \alpha^{-1} \rfloor$ for some functor $F : \mathbb{N}^r \to \mathrm{Vect}_k$, i.e. $F \cong G\alpha : \mathbb{N}^r \to \mathrm{Vect}_k$. A functor is called tame if it is α-tame for some $\alpha \in \mathbb{Q}$. We denote the set of all tame functors $\mathbb{Q}^r \to \mathrm{Vect}_k$ by $\mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$. This forms a category when we take the morphisms to be natural transformations.

Proposition 5.4.5 ([17])**.** *Let* $F \in \mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$ *be* α*-tame such that* $F\alpha / \operatorname{rad} F\alpha \cong \bigoplus_w (U_w \otimes V_w)$ *for a sequence of* k*-vector spaces* $\{V_w\}$*. Then*

a) $supp(F) = \{\alpha w | w \in \mathrm{supp}(F\alpha)\}$

b) $\operatorname{rank} F = \operatorname{rank} F\alpha = \sum \dim_k V_w$

c) $\beta_0 F(v) = \begin{cases} \beta_0(F\alpha)(w) = \dim_k V_w & v = \alpha w \\ 0 & \textit{otherwise} \end{cases}$

Thus we can talk about the rank, support, and 0-th Betti diagram of tame persistence modules. The function $\operatorname{rank} : \mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k) \to \mathbb{N}$ is an invariant of tame persistence modules, and it will be part of the input of the hierarchichal stabilization process. We will stablize these functions with respect to some pseudometric on $\mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$. Of course one pseudometric we can work with is the interleaving metric. We have the superlinear family

$$\Omega_\varepsilon(v) = v + \vec{\varepsilon}$$

where $\vec{\varepsilon}$ is the vector in $\mathbb{Q}^r_{\geq 0}$ with ε in every entry. We can then define two persistence modules to be ε-interleaved if they are Ω_ε-interleaved. This definition of interleaving corresponds to the definition in [20]. We wish to have another way of generating pseudometrics on tame functors. To define these pseudometrics, we need to look at noise systems.

5.5 Noise systems

We now discuss noise systems, first introduced in [17]. The idea is to define some notion of how far away a tame persistence modules is from the zero functor. Conceptually, we think of those persistence modules which are close to the zero module as noise.

Definition 5.5.1. A noise system in $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$ is a collection $\mathcal{C} = \{C_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ of sets of tame functors such that

- The zero functor is in C_ε for all $\varepsilon \in \mathbb{Q}_{\geq 0}$.
- $\tau \leq \varepsilon \implies C_\tau \subset C_\varepsilon$.
- Given an exact sequence of tame functors

$$0 \to F_0 \to F_1 \to F_2 \to 0$$

  then

  1. $F_1 \in C_\varepsilon \implies F_0, F_2 \in C_\varepsilon$.
  2. If $F_0 \in C_\varepsilon, F_2 \in C_\tau \implies F_1 \in C_{\tau+\varepsilon}$.

Given two noise systems $\mathcal{C} = \{C_\varepsilon\}_\varepsilon, \mathcal{D} = \{D_\varepsilon\}_\varepsilon$, we say that $\mathcal{C} \subset \mathcal{D}$ if $C_\varepsilon \subset D_\varepsilon$ for all ε. With this relation, the collection of all noise systems in $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$ forms a poset.

Example 5.5.2 (Standard noise in the direction of a cone)**.** A cone is a subset of $\mathbb{Q}^r$ made up of all linear combinations of some finite collection of elements in $\mathbb{Q}^r$, where these linear combinations have coefficients in $\mathbb{Q}_{\geq 0}$. Let $V \subset \mathbb{Q}^r_{\geq 0}$ be a cone. Define

$$V_\varepsilon = \{F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) | \text{ For any } u \in \mathbb{Q}^r_{\geq 0}, x \in F(u)\, \exists\, w \in V s.t. \,||w|| = \varepsilon \text{ and } x \in \ker(u \leq u + w)\}.$$

To see that this is a noise system, first note that the 0 functor is clearly in V_ε for all ε. Let $\tau \leq \varepsilon$. If $x \in \ker F(u \leq u+w)$ then $x \in \ker F(u \leq u + \frac{\varepsilon}{\tau} w)$, since $w \leq \frac{\varepsilon}{\tau} w$. As $||\frac{\varepsilon}{\tau} w|| = \frac{\varepsilon}{\tau}||w||$, if we have $||w|| = \tau$, then the inclusion $V_\tau \subset V_\varepsilon$ is clear.
Consider now the short exact sequence

$$0 \to F \to G \to H \to 0.$$

If $G \in V_\varepsilon$, then by naturality $F, H \in V_\epsilon$. Now assume $F \in V_\varepsilon, H \in V_\tau$. Consider $x \in G(u)$. Then its image $x_1 \in H(u)$ is in $\ker H(u \leq u+w)$ for some $w \in V$ with $||w|| = \tau$. Therefore $G(u \leq u+w)$ maps x to an element $x_2 \in \ker(G(u+w) \to H(u+w)) = \text{im}(F(u+w) \to G(u+w))$. Let $\hat{x}_2$ denote the element in $F(u+w)$ which is mapped to x_2.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & F(u) & \longrightarrow & G(u) & \longrightarrow & H(u) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & F(u+w) & \longrightarrow & G(u+w) & \longrightarrow & H(u+w) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & F(u+w+w') & \longrightarrow & G(u+w+w') & \longrightarrow & H(u+w+w') & \longrightarrow & 0
\end{array}$$

We can thus find a $w' \in V$ with $||w'|| = \varepsilon$ such that $\hat{x}_2 \in \ker F(u+w \leq u+w+w')$. It follows that $x \in \ker G(u \leq u+w+w')$. Since $||w+w'|| \leq ||w|| + ||w'|| = \tau + \varepsilon$, it follows that $x \in \ker G(u \leq u + \frac{\tau+\varepsilon}{||w+w'||}(w+w'))$. Therefore $G \in V_{\tau+\varepsilon}$. Note that $\frac{\tau+\varepsilon}{||w+w'||}(w+w') \in V$ since V is a cone. $\Diamond$

Example 5.5.3 (Domain noise). Let $\mathcal{X} = \{X_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ be a sequence of non-decreasing subsets of $\mathbb{Q}^r_{\geq 0}$, where by non-decreasing we mean that $\tau \leq \varepsilon \implies X_\tau \subset X_\varepsilon$. For a persistence module $F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k)$, we define its domain to be

$$\text{Domain}(F) = \{v \in \mathbb{Q}^r_{\geq 0} | F(v) \neq 0\}.$$

Then the domain noise associated to $\mathcal{X}$ is defined by

$$\mathcal{X}_\varepsilon = \{F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) | \, \text{Domain}(F) \subset X_\varepsilon\}.$$

The fact that the sequence $\{X_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ is non-decreasing ensures that this is in fact a noise system. It is also clear that the domain noise is closed under direct sums. $\diamondsuit$

We will use noise systems to define pseudometrics on $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$. For the rest of this section, we fix a noise system $\mathcal{C} = \{\mathcal{C}_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ in $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$.

Definition 5.5.4. An ε-equivalence betweeen two tame functors F, G is a natural transformation $\phi : F \to G$ such that $\ker \phi, \operatorname{coker} \phi \in \mathcal{C}_\varepsilon$. [1]

Definition 5.5.5. Let $\varepsilon \in \mathbb{Q}_{\geq 0}$, and $F, G \in \text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$. F and G are called ε-close if there is a tame functor H and natural transformations $\phi : H \to F$, $\psi : H \to G$ such that ϕ is a τ-equivalence and ψ is a μ-equivalence where $\tau + \mu \leq \varepsilon$.

Note that if there is an ε-equivalence between F, G, then they are ε-close. Being ε-close is a reflexive and symmetric relation, but it is not transitive. We need some properties about ε-equivalences in order to show that being ε-close defines a pseudometric on $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$.

Proposition 5.5.6. *Consider the following commuting square in* $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$*:*

$$\begin{array}{ccc} H & \xrightarrow{\phi} & F \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\psi'} \\ G & \xrightarrow[\phi']{} & E \end{array} \tag{5.1}$$

Then (5.1) *is a pushout square and ϕ is an ε-equivalence $\implies$ ϕ' is an ε-equivalence. Also* (5.1) *is a pullback square and ϕ' is an ε-equivalence $\implies$ ϕ is an ε equivalence.*

Proof. We will prove only the first statement since the proof of the second is similar. Our approach will be to show that $\operatorname{coker} \phi \cong \operatorname{coker} \phi'$ and that $\ker \phi'$ is a quotient of $\ker \phi$. Then the result follows since each component of a noise system is closed under taking quotients. Since $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$ is a functor category, taking pushouts is performed object-wise, so we consider the following pushout diagram in Vect_k:

$$\begin{array}{ccc} H(x) & \xrightarrow{\phi_x} & F(x) \\ {\scriptstyle\psi_x}\downarrow & & \downarrow{\scriptstyle\psi'_x} \\ G(x) & \xrightarrow[\phi'_x]{} & E(x) \end{array} \tag{5.2}$$

In particular, $E(x) \cong F(x) \oplus G(x)/\sim$, where $\sim$ is the equivalence relation generated by $\phi_x(v) \sim \psi_x(v)$ for all $v \in H(x)$, and ϕ_x, ψ_x are the projection maps. Now $\operatorname{coker} \phi'_x \cong (F(x) \oplus G(x)/\sim) \,/\, \operatorname{im} \phi'_x$. In

[1] In [17], an ε-equivalence between F, G is defined to be a natural transformation $\phi : F \to G$ such that $\ker \phi \in \mathcal{C}_\tau, \operatorname{coker} \phi \in \mathcal{C}_\mu$ where $\tau + \mu \leq \varepsilon$. We use a slightly different definition here because this formulation is more natural in the context of connecting noise systems to Serre categories, as we will do later.

particular, $\operatorname{coker} \phi'_x \cong F(x) \oplus G(x)/\sim$, where now additionally $(0, w) \sim (0, 0)$ for any $w \in G(x)$. Therefore $\phi_x(v) \sim \psi_x(v) \sim (0, 0)$. So we have $\operatorname{coker} \phi'_x \cong F(x)/\operatorname{im} \phi_x \cong \operatorname{coker} \phi_x$. For the kernels, note that if $w \in \ker(\phi'_x)$, then $w = \psi_x(v) \sim \phi_x(v) \sim 0$ for some $v \in H(x)$. Therefore $\ker(\phi'_x) \cong \psi(\ker(\phi_x))$, which by the first isomorphism theorem is a quotient of $\ker \phi_x$, as required. □

Proposition 5.5.7. *Let $\phi : F \to G$, $\psi : G \to H$ be natural transformations in* $\mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$*. If ϕ is an ε_1-equivalence and ψ is an ε_2-equivalence, then $\psi \circ \phi$ is an $(\varepsilon_1 + \varepsilon_2)$-equivalence.*

Proof. First note that we have the following exact sequences

$$0 \to \ker \phi \to \ker \psi \circ \phi \to \ker \psi$$

$$\operatorname{coker} \phi \to \operatorname{coker} \psi \circ \phi \to \operatorname{coker} \psi \to 0$$

Let $\ker \phi, \operatorname{coker} \phi \in C_{\varepsilon_1}$, and $\ker \psi, \operatorname{coker} \psi \in C_{\varepsilon_2}$. Now the image of $\ker \psi \circ \phi \to \ker \psi$ from the first exact sequence is contained in C_{ε_2}, since it is a subfunctor of $\ker \psi$. Now if we replace $\ker \psi$ by this image, the first sequence becomes a full short exact sequence, therefore $\ker \psi \circ \phi \in C_{\varepsilon_1+\varepsilon_2}$. Similarly it is true that $\operatorname{coker} \psi \circ \phi \in C_{\varepsilon_1+\varepsilon_2}$. Therefore $\psi \circ \phi$ is an $(\varepsilon_1 + \varepsilon_2)$-equivalence. □

Corollary 5.5.8. *Let F, G be tame functors and $\tau, \mu \in \mathbb{Q}_{\geq 0}$. Then TFAE*

a) There is a tame functor H and natural transformations $\phi : H \to F, \psi : H \to G$ such that ϕ is a τ-equivalence and ψ is a μ-equivalence.

b) There is a tame functor E and natural transformations $\phi' : G \to E, \psi' : F \to E$ such that ϕ' is a τ-equivalence and ψ' is a μ-equivalence.

Proof. Consider the diagram (5.1) as a pushout, and assume a) is true. Then by Proposition 5.5.6, ϕ' is a τ-equivalence and ψ' is a μ-equivalence. The other direction is similar. □

Proposition 5.5.9. *Given tame functors F, G, L such that F, G are ε_1-close and G, L are ε_2-close, then F, L are $(\varepsilon_1 + \varepsilon_2)$-close.*

Proof. From the statement, we have natural transformations

$$\phi_1 : H_1 \to F, \quad \psi_1 : H_1 \to G$$

Such that ϕ_1 is a τ_1-equivalence and ψ_1 is a μ_1-equivalence, with $\tau_1 + \mu_1 \leq \varepsilon_2$. And similarly we have

$$\phi_2 : H_2 \to G, \quad \psi_2 : H_2 \to L$$

Where ϕ_2 is a τ_2-equivalence and ψ_2 is a μ_2-equivalence, with $\tau_2 + \mu_2 \leq \varepsilon_2$. Now consider the following pullback square

$$\begin{array}{ccc} H_3 & \xrightarrow{\phi_3} & H_1 \\ {\scriptstyle \psi_3}\downarrow & & \downarrow{\scriptstyle \psi_1} \\ H_2 & \xrightarrow[\phi_2]{} & G \end{array}$$

By Proposition 5.5.6, ϕ_3 is a τ_2-equivalence and ψ_3 is a μ_1-equivalence. So, $\phi_1 \circ \phi_3 : H_3 \to F$ is a $(\tau_1 + \tau_2)$-equivalence, and $\psi_2 \circ \psi_3 : H_3 \to L$ is a $(\mu_1 + \mu_2)$-equivalence. It follows that F, L are $(\varepsilon_1 + \varepsilon_2)$-close. □

Definition 5.5.10. Let F, G be tame functors. Then we define a pseudometric on $\mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$ by

$$d(F, G) = \inf\{\varepsilon \in \mathbb{Q}_{\geq 0} | F, G \text{ are } \varepsilon - \text{close }\}$$

From the discussion in this section, it is clear that this defines a pseudometric. We are now ready to apply the hierarchical stabilization process. We have equipped the set $\mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$ with an invariant given by the rank function, and a pseudometric d dependent on the noise system $\mathcal{C}$. Let $B(F, \varepsilon)$ denote the closed ball of radius ε with respect to the pseudometric d. Then we get the stabilized rank invariant

$$\widehat{\mathrm{rank}} : \mathit{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k) \to \mathrm{Mult}(\mathbb{Q}), F \mapsto (\varepsilon \mapsto \min\{\mathrm{rank}(G) | G \in B(F, \varepsilon)\}).$$

5.6 Persistence contours

We have a method of generating noise systems, called persistence contours.

Definition 5.6.1. A persistence contour is a functor $(\mathbb{Q}^r_{\geq 0} \cup \infty) \times \mathbb{Q}_{\geq 0} \to \mathbb{Q}^r_{\geq 0} \cup \infty$ such that for any $v \in \mathbb{Q}^r_{\geq 0} \cup \infty, \varepsilon, \tau \in \mathbb{Q}_{\geq 0}$, we have

- $v \leq C(v, \varepsilon)$.
- $C(C(v, \varepsilon), \tau) \leq C(v, \tau + \varepsilon)$.
- $C(\infty, \varepsilon) = \infty$ for all ε

We have in fact already met persistence contours before. Recall the definition of a superlinear family over a poset P.

Definition 5.6.2. A superlinear family over P is a function $\Omega : [0, \infty) \to \mathrm{Trans}_P$ which satisfies $\Omega_{\delta_1} \circ \Omega_{\delta_2} \leq \Omega_{\delta_1 + \delta_2}$

Given a persistence contour C, we can define a superlinear family over $P = \mathbb{Q}^r_{\geq 0} \cup \infty$ by setting $\Omega_\varepsilon(v) = C(v, \varepsilon)$, and vice versa. The first condition of C being a persistence contour corresponds to Ω_ε being a translation, and the second condition is superlinearity. In this way persistence contours can be viewed as a superlinear families. From now on we will only use this viewpoint, and we use the term persistence contour, or simply contour, to refer to a superlinear family over $P = \mathbb{Q}^r_{\geq 0} \cup \infty$ for some r.

Example 5.6.3. The **standard contour** is given by the map

$$\Omega_\varepsilon(v) = v + \vec{\varepsilon}.$$

Similarly, we can define a persistence contour by

$$\Omega_\varepsilon(v) = v + f(\varepsilon)$$

where $f : \mathbb{Q}_{\geq 0} \to \mathbb{Q}^r_{\geq 0}$ is non-decreasing and superlinear, in the sense that $f(\tau + \varepsilon) \leq f(\tau) + f(\varepsilon)$. ♢

Definition 5.6.4. Given a persistence contour Ω, define $C_\varepsilon \subset \mathrm{Tame}(\mathbb{Q}^r, \mathrm{Vect}_k)$ to be the collection of finitely generated tame functors for which $F(v \leq \Omega_\varepsilon(v))$ is the zero homorphism whenever $\Omega_\varepsilon(v) < \infty$. The collection $\{C_\varepsilon\}_\varepsilon$ is the noise system associated to the contour Ω.

Proposition 5.6.5 ([17])**.** *The noise system associated to a persistence contour Ω is in fact a noise system.*

Proof. Clearly $0 \in C_\varepsilon$ for all ε. Let $\tau \leq \varepsilon$, and $G \in C_\tau$. Then for any $v \in \mathbb{Q}^r_{\geq 0} \cup \infty$, $\Omega_\tau(v) \leq \Omega_\varepsilon(v)$. Then since $G(v \leq \Omega_\tau(v))$ is trivial, so is the composition $G(v \leq \Omega_\varepsilon(v)) = G(\Omega_\tau(v) \leq \Omega_\varepsilon(v)) \circ G(v \leq$

$\Omega_\tau(v))$. Therefore $G \in C_\varepsilon$, and so $C_\tau \subset C_\varepsilon$.
Now consider an exact sequence of tame functors

$$0 \to G_0 \to G_1 \to G_2 \to 0$$

Let $G_1 \in C_\varepsilon$, so $G_1(v \leq \Omega_\varepsilon(v)) = 0$. Now $G_1(v) \to G_2(v)$ is an epimorphism, and $G_0(\Omega_\varepsilon) \to G_1(\Omega_\varepsilon)$ is a monomorphism. Then by naturality, we see that $G_0(v \leq \Omega_\varepsilon(v))$ and $G_2(v \leq \Omega_\varepsilon(v))$ are trivial. In particular, this becomes clear when we consider the following commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & G_0(v) & \longrightarrow & G_1(v) & \longrightarrow & G_2(v) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle 0} & & \downarrow & & \\
0 & \longrightarrow & G_0(\Omega_\varepsilon(v)) & \longrightarrow & G_1(\Omega_\varepsilon(v)) & \longrightarrow & G_2(\Omega_\varepsilon(v)) & \longrightarrow & 0
\end{array}$$

Now let $G_0 \in C_\tau, G_2 \in C_\varepsilon$. Now consider the following commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & G_0(v) & \longrightarrow & G_1(v) & \longrightarrow & G_2(v) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle 0} & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & G_0(\Omega_\tau(v)) & \longrightarrow & G_1(\Omega_\tau(v)) & \longrightarrow & G_2(\Omega_\tau(v)) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow{\scriptstyle 0} & & \\
0 & \longrightarrow & G_0(\Omega_\varepsilon \circ \Omega_\tau(v)) & \longrightarrow & G_1(\Omega_\varepsilon \circ \Omega_\tau(v)) & \longrightarrow & G_2(\Omega_\varepsilon \circ \Omega_\tau(v)) & \longrightarrow & 0
\end{array}$$

Then commutativity implies that the composition of the two central vertical maps is trivial. Then in particular, since $\Omega_\varepsilon \circ \Omega_\tau \leq \Omega_{\tau+\varepsilon}$, we have $G(v \leq \Omega_{\varepsilon+\tau}(v)) = 0$. □

We now prove a claim which was stated without proof in [17]. This result demonstrates how the pseudometric generated by the noise system associated to a persistence contour is related to the interleaving distance to the interleaving distance.

Theorem 5.6.6. *Let $F, G : \mathbb{Q}^r_{\geq 0} \to \mathrm{Vect}_k$ be tame persistence modules, and Ω a persistence contour. Then $d_I(F,G) \leq \delta \implies F, G$ are 3δ-close w.r.t the noise system generated by Ω. Also if F, G are δ-close, then they are 2δ-interleaved. In particular*

$$\frac{1}{2} d_I(F,G) \leq d(F,G) \leq 3 d_I(F,G)$$

Proof. First assume that F, G are δ-interleaved, i.e. Ω_δ-interleaved. So we have natural transformations $\phi : F \to G \circ \Omega_\delta$, $\psi : G \to F \circ \Omega_\delta$ such that $\phi \circ \psi = G \circ \eta_{\Omega_{2\delta}}$, and $\psi \circ \phi = F \circ \eta_{\Omega_{2\delta}}$, where η_T is the unique natural transformation between the identity functor and the translation T. Our goal will be to show that ϕ is a 2δ-equivalence and that $\lambda : G \to G \circ \Omega_\delta$ is a δ-equivalence, where λ is precomposition by η_{Ω_δ}. Then using the triangle inequality, we will be able to conclude that F, G are 3δ-close.
To see that $\lambda : G \to G \circ \Omega_\delta$ is a δ-equivalence, consider the following diagram

$$\begin{array}{ccccccc}
\ker \lambda_x & \longrightarrow & G(x) & \xrightarrow{\lambda_x} & G(\Omega_\delta(x)) & \longrightarrow & \operatorname{coker} \lambda_x \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\ker \lambda_{\Omega_\delta(x)} & \longrightarrow & G(\Omega_\delta(x)) & \xrightarrow[\lambda_{\Omega_\delta(x)}]{} & G(\Omega_{2\delta}(x)) & \longrightarrow & \operatorname{coker} \lambda_{\Omega_\delta(x)}
\end{array}$$

Notice that in the central square, both maps coming from $G(x)$ are λ_x, so by commutativity, the vertical map between the kernels must be trivial. Similarly, the map between the cokernels must be trivial. Therefore $\ker \lambda, \operatorname{coker} \lambda \in C_\delta$, and it follows that $G, G \circ \Omega_\delta$ are δ-close.
To see that ϕ is a 2δ-equivalence, recall that we have the following exact sequences:

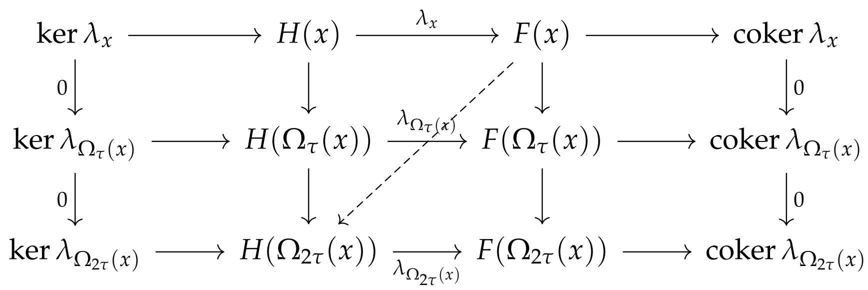
$$0 \to \ker \phi \to \ker \psi \circ \phi \to \ker \psi$$

$$\operatorname{coker} \psi \to \operatorname{coker} \phi \circ \psi \to \operatorname{coker} \phi \to 0$$

Now recall that $(\psi \circ \phi)_x = F \circ \eta_{\Omega_{2\delta}(x)} = F(x \leq \Omega_{2\delta}(x))$. Therefore $\ker \psi \circ \phi \in C_{2\delta}$. In the first exact sequence above, if we replace $\ker \psi$ by the image of the last map, we get a short exact sequence. It follows that $\ker \phi \in C_{2\delta}$, and similarly $\operatorname{coker} \phi \in C_{2\delta}$. In particular, $F, G \circ \Omega_\delta$ are 2δ-close.
Now assume that F, G are δ-close. To show that they are 2δ-interleaved, it suffices to show that any δ-equivalence gives rise to a 2δ-interleaving. To see this, note that since F, G are δ-close, we have the following diagram:

$$\begin{array}{ccc} & L & \\ {}^{\theta}\swarrow & & \searrow^{\rho} \\ F & & G \end{array}$$

Where L is a tame functor and θ, ρ are τ- and μ-equivalences respectively, where $\tau + \mu \leq \delta$. Then if we can say that L, F are τ-interleaved, and L, G are μ-interleaved, it will follow by the triangle inequality that F, G are δ-interleaved.
Let $H \in \operatorname{Tame}(\mathbb{Q}^r, \operatorname{Vect}_k)$, and $\lambda : H \to F$ be a τ-equivalence. So $\ker \lambda, \operatorname{coker} \lambda \in C_\tau$. In particular, we have the following diagram



We have a natural transformation $H \to F \circ \Omega_{2\tau}$ given by the composition $\lambda \circ H(x \leq \Omega_{2\tau}(x))$. So the existence of a 2τ interleaving between H, F is equivalent to the existence of the dashed arrow in the above diagram such that the resulting diagram commutes. Fix an element $z \in F(x)$. Then by commutativity of the upper right square, $z' := F(x \leq \Omega_\tau(x))(z) \in \operatorname{im} \lambda_{\Omega_\tau(x)}$. Now consider $w_1, w_2 \in H(\Omega_\tau(x))$ such that $\lambda_{\Omega_\tau(x)}(w_i) = z'$ for $i = 1, 2$. Then in particular $w_1 - w_2 \in \ker \lambda_{\Omega_\tau(x)}$. Hence, by commutativity of the bottom left square, we have $H(\Omega_\tau(x) \leq \Omega_{2\tau}(x))(w_1) = H(\Omega_\tau(x) \leq \Omega_{2\tau}(x))(w_2) =: w$. Then the map given by $z \mapsto w$ is a well defined linear function, which by the naturality of λ, defines a natural map $F \to H \circ \Omega_{2\tau}$. In particular, it defines a 2τ-interleaving between F, H. □

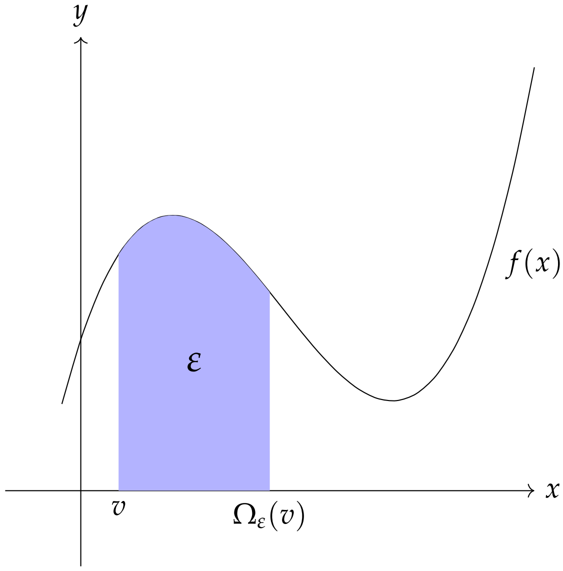
Proposition 5.6.7 ([17]). *Let $V = \operatorname{Cone}(v_1, \ldots, v_n)$ be a cone in $\mathbb{Q}^r_{\geq 0}$. If $v_1, \ldots, v_n$ are linearly independent, then then the standard noise in the direction of V is the same as the standard noise in the direction of $v_1 + \cdots + v_n$.*

This proposition implies that theorem 5.6.5 also holds when we replace the noise system generated by a persistence contour by the standard noise in the direction of the full cone $\mathbb{Q}^r_{\geq 0}$, which is the claim made in [17].[2]

In one dimension, we can generate contours from positive, Lebesgue measurable functions. We will call such functions **density functions**.

Example 5.6.8. Given a density function f and some $\varepsilon \in \mathbb{Q}_{\geq 0}$, we can uniquely define a persistence contour Ω by

$$\varepsilon = \int_{v}^{\Omega_\varepsilon(v)} f(x)dx.$$



To see that this defines a contour, consider $\varepsilon, \tau \in \mathbb{Q}_{\geq 0}$. Then

$$\begin{aligned} \varepsilon + \tau &= \int_{v}^{\Omega_\varepsilon(v)} f(x)dx + \int_{\Omega_\varepsilon(v)}^{\Omega_\tau(\Omega_\varepsilon(v))} f(x)dx \\ &= \int_{v}^{\Omega_\tau(\Omega_\varepsilon(v))} f(x)dx \end{aligned}$$

However,

$$\varepsilon + \tau = \int_{v}^{\Omega_{\varepsilon+\tau}(v)} f(x)dx$$

by definition. Therefore $\Omega_{\varepsilon+\tau}(v) = \Omega_\tau(\Omega_\varepsilon(v))$, and so Ω is a persistence contour. A contour defined in this way is said to be of **distance type**. Note that if f is the constant function $f(x) = 1$, then the distance type contour we get is the standard contour. ♢

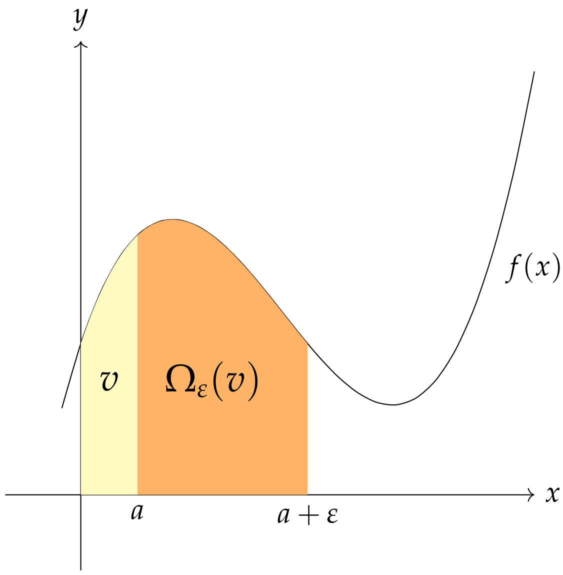
Example 5.6.9. Given a density function f, and a value $v \in \mathbb{Q}_{\geq 0}$, there is a unique $a \in \mathbb{Q}_{\geq 0}$ such that

$$v = \int_{0}^{a} f(x)dx.$$

Then we can define a contour Ω by

$$\Omega_\varepsilon(v) = \int_{a}^{a+\varepsilon} f(x)dx.$$

[2] The statement in [17] is that if d is the pseudometric associated to the standard noise in the direction of the full cone $\mathbb{Q}^r_{\geq 0}$, then $d_I(F,G) \leq d(F,G) \leq 6d_I(F,G)$. The discrepancy is due to the different choice of definition for an ε-equivalence that we have used.



Monotonicity of integrals guarantees that Ω satisfies the first condition of being a contour. Now, by definition

$$\Omega_\tau(\Omega_\varepsilon(v)) = \int_a^{a+\varepsilon+\tau} f(x)dx = \Omega_{\varepsilon+\tau}(v)$$

and so Ω also satisfies the second condition of being a contour. Contours generated in this way are said to be of **shift type**. Just as in the previous example, if $f(x) = 1$, then Ω is the standard contour. $\Diamond$

We will now see that persistence contours give rise to a particularly nice set of noise systems, which we call simple noise systems.

5.7 Denoising and simple noise systems

Definition 5.7.1. Let $\{C_\varepsilon\}_\varepsilon$ be a noise system in $\text{Tame}(\mathbb{Q}^r, \text{Vect}_k)$, and F a tame and finitely generated functor. A denoising of F is a sequence of functors $\{\text{denoise}(F)_\varepsilon\}_\varepsilon$ such that for any $\varepsilon \in \mathbb{Q}_{\geq 0}$:

- $\text{denoise}(F)_\varepsilon \in B(F, \varepsilon)$.
- $\text{rank}(\text{denoise}(F)_\varepsilon) = \widehat{\text{rank}}F(\varepsilon)$.

So a denoising of F at scale ε is a functor in $B(F, \varepsilon)$ that realises the minimum rank. Thus a denoising of F with respect to a particular noise system is a functor which realises the stable rank of F. Therefore any efficient way of producing a denoising of a functor would also result in an efficient computation of the stable rank.
Given a tame persistence module F, we define its τ-neighbourhood w.r.t to some noise system $\{C_\varepsilon\}_\varepsilon$ as

$$B_\subset(F, \tau) = \{F' \subset F | F' \text{ tame and } F/F' \in C_\tau\}.$$

If the noise system is closed under direct sums, in the sense that each C_ϵ is closed under direct sums, then if the above set has a minimal element w.r.t. inclusion, this minimal element is unique [19]. If this minimal element exists, we call it F^τ. Note that whenever F^τ exists, the condition that $F/F' \in C_\tau$ is equivalent to saying that $F^\tau \subset F'$. To see this, note that if $F^\tau \subset F'$, then there is the short exact sequence

$$0 \to F'/F^\tau \to F/F^\tau \to F/F' \to 0.$$

Therefore $B_{\subset}(F,\tau)$ can be identified with the set

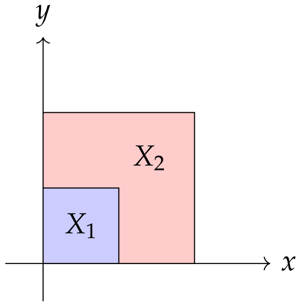
$$\{F' \subset F | F' \text{ tame and } F^{\tau} \subset F' \subset F\}$$

Definition 5.7.2. A noise system $\{S_\varepsilon\}_\varepsilon$ is called simple if

- It is closed under direct sums.
- For any finitely generated functor $G \in \text{Tame}(\mathbb{Q}^r_{\geq 0})$, the set $B_{\subset}(G,\tau)$ has a minimal element for any $\tau \in \mathbb{Q}_{\geq 0}$, which we denote G^τ.
- $\text{rank}\, G^\tau \leq \text{rank}\, G$.

Note that the third condition is always satisfied in the one dimensional case, but not in higher dimensions.

Example 5.7.3. Let $\mathcal{X} = \{X_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$, where $X_\varepsilon = [0,\varepsilon) \times [0,\varepsilon) \subset \mathbb{Q}^2_{\geq 0}$, and let $\{\mathcal{X}_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ be the associated domain noise system. Let $F = k((0,0),-)$. Then F^τ is the minimal subfunctor with respect to inclusion such that $\text{Domain}(F/F^\tau) \subset X_\tau$. In particular, we must have $F = F^\tau$ on $\mathbb{Q}^2_{\geq 0} \backslash X_\tau$. So for $\tau = 0$,it is clear that $F = F^\tau$. For all other values, F^τ is given by the functor $k((0,\varepsilon),-) \oplus k((\varepsilon,0),-)$. Therefore, $\text{rank}\, F^\tau \geq \text{rank}\, F$. It follows that the domain noise is not simple.



$$F = \begin{array}{ccccc} k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k & \longrightarrow & k \end{array} \qquad F^1 = \begin{array}{ccccc} k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k \end{array}$$

◊

It turns out that simple noise systems admit a simplified way of computing the stable rank, particularly in the one-dimensional case. Given a noise system $\{C_\epsilon\}_\epsilon$, the approach will be to compute $\widehat{\text{rank}} F(\tau)$ by considering the tame functors in $B_{\subset}(F,\tau)$.

Theorem 5.7.4 ([17],[21])**.** *Let $\{S_\varepsilon\}_\varepsilon$ be a simple noise system, and $G \in \text{Tame}(\mathbb{Q}^r_{\geq 0})$ be finitely generated. Then*

$$\widehat{\text{rank}} G(\tau) = \min\{\text{rank}\, F | F \in B_{\subset}(G,\tau)\}$$

Therefore in order to calculate the stable rank of F, it is enough to look at the functors in $B_{\subset}(F,\tau)$. So if we can produce a family of simple noise systems, and construct explicitly the functors F^τ, we will be able to compute the stable rank invariant. It turns out that in the one dimensional case, we can do just this.

Let G be a finitely generated tame persistence module (of any dimension), and let Ω be a persistence contour, and $\mathcal{C} = \{C_\varepsilon\}_\varepsilon$ the associated noise system. Choose a minimal set of generators $\{g_s \in G(v_s)\}_{s=1}^n$. Let $h_s = G(v_s \leq \Omega_\tau(v_s))(g_s)$. Now define $G^\tau \subset G$ to be the functor generated by the elements h_s. This reuse of notation is justified by the fact that the two definitions of G^τ coincide when our noise system is simple. To see this, irst we will show that $G/G^\tau \in C_\tau$. For $v \in \mathbb{Q}_{\geq 0}$, any $x \in G(v)$ can be written as $x = \sum_{v_s \leq v} \lambda_s G(v_s \leq v)(g_s)$ for some coefficents λ_s. So

$$\begin{aligned} G(v \leq \Omega_\tau(v))(x) &= \sum_{v_s \leq v} \lambda_s G(v \leq \Omega_\tau(v)) \circ G(v_s \leq v)(g_s) \\ &= \sum_{v_s \leq v} \lambda_s G(\Omega_\tau(v_s) \leq \Omega_\tau(v) \circ G(v_s \leq \Omega_\tau(v_s))(g_s) \\ &= \sum_{v_s \leq v} \lambda_s G(\Omega_\tau(v_s) \leq \Omega_\tau(v))(h_s). \end{aligned}$$

And in particular, x is mapped into $G^\tau(v)$. To check minimality, let $G' \in B_\subset(G, \tau)$. In particular, $G/G' \in C_\tau$, so $G(v_s \leq \Omega_\tau(v_s))$ maps the generator g_s to an element in $G'(v)$, and so in particular h_s belongs to $G'(v)$. Therefore $G^\tau \subset G'$.

Example 5.7.5. Let Ω be the standard contour, and $F : \mathbb{N}^2_{\geq 0} \to \text{Vect}_k$ be the following persistence module:

$$F = \begin{array}{ccccc} k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ k & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k \end{array}$$

Then F has generators which lie in $F(0,1)$ and $F(1,0)$, and any set of such generators is minimal. Similarly, if we compute F^1, we see that a minimal set of generators is made up of elements of $F(1,2)$ and $F(2,1)$, so F^1 looks like the following diagram in the range $[0,2] \times [0,2]$:

$$F^1 = \begin{array}{ccccc} 0 & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & 0 \end{array}$$

Now by Theorem 5.7.4, $\widehat{\text{rank}} F(1)$ can be computed by looking at the ranks of the functors G such that $F^1 \subset G \subset F$. We can see that the free functor $G = k((1,1), -)$ satisfies this criteria, and so $\widehat{\text{rank}} F(1) = \text{rank}\, G = 1$.

$$G = \begin{array}{ccccc} 0 & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & k & \longrightarrow & k \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \longrightarrow & 0 & \longrightarrow & 0 \end{array}$$

$\Diamond$

Proposition 5.7.6 ([19])**.** *If $F \in \mathrm{Tame}(\mathbb{Q}_{\geq 0}, \mathrm{Vect}_k)$ is finitely generated, then*

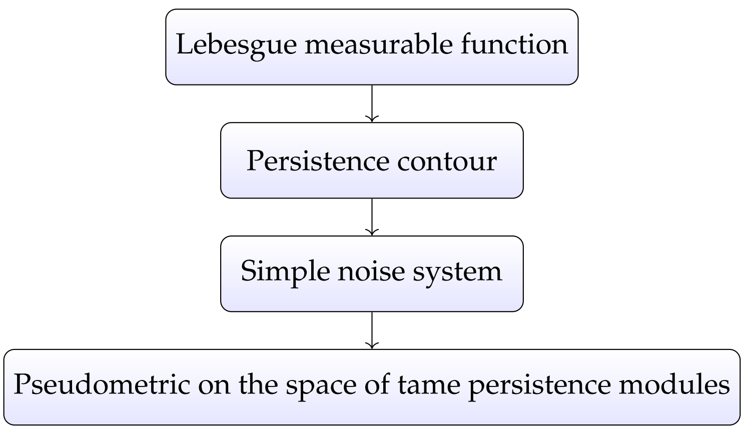
$$\widehat{\mathrm{rank}} F(\tau) = F^{\tau}$$

where the stabililisation of the rank is w.r.t the noise system generated by a contour.

In particular, $\mathrm{denoise}(F)_{\tau} = F^{\tau}$ in the one dimensional case. We recall that finitely generated tame one-dimensional persistence modules admit an interval decomposition, and so a corresponding barcode. In the multi-dimensional case the situation is more complicated and there is in general no canonical choice of a denoising, so the focus is on directly computing the stable rank.

Example 5.7.7. Let $\{S_{\varepsilon}\}$ be the simple noise system associated to the standard contour in one dimension. Consider the free functor in one generator $F = k(i, -) \in \mathrm{Tame}(\mathbb{Q}_{\geq 0}, \mathrm{Vect}_k)$ for some $i \in \mathbb{Q}_{\geq 0}$. Then $F^{\tau} = k(i + \tau, -)$. For $G = [a, b) \in \mathrm{Tame}(\mathbb{Q}_{\geq 0}, \mathrm{Vect}_k)$, an interval module in one dimension, we have $G^{\tau} = [a + \tau, b)$. It follows that for a decomposable one dimesional persistence module H, the subfunctor H^{τ} has rank given by the number of bars in the barcode decomposition of H of length strictly greater than τ. This explicit description of a denoising for one dimensional persistence modules allows the stable rank to be computed very efficiently, since we can compute the barcode efficiently. $\Diamond$

It turns out that simple noise systems are in bijective correspondence with persistence contours. We have already seen that contours give rise to noise systems, and the discussion in this section section shows that these noise systems are simple. The map which sends a persistence contour to its associated simple noise system turns out to be a bijection between the set of contours and simple noise systems [19]. This bijection, along with the methods for generating contours described in example 6.7 and 6.8 give us a pipeline for generating a wide variety of metrics between persistence modules. By taking any choice of positive valued, Lebesgue measurable function, we can generate a contour. This contour is associated to a unique simple noise system, from which we derive a new pseudometric. This pipeline allows us to very easily vary the nature of the pseudometric we wish to use, so it is easier to find an appropriate metric for any given application. The question of finding a good pseudometric for a particular data set is still open, with trial and error being the primary method available at the moment.



5.8 Noise systems as Serre categories

We will now give a brief introduction to the theory of Serre categories.

Definition 5.8.1. Let A be an abelian category. A Serre subcategory of A is a non-empty full subcategory C of A such that given any SES in A

$$0 \to F \to G \to H \to 0$$

$F, H \in \mathrm{Ob}(\mathcal{C}) \Leftrightarrow G \in \mathrm{Ob}(\mathcal{C})$.

Lemma 5.8.2. *Let A be an abelian category, and C be a Serre subcategory. Then there exists an abelian category A/C and an exact functor*

$$\pi : A \to A/C$$

which is essentially surjective, and is such that $\pi(F)$ is the zero object of A/C iff $F \in \mathrm{Ob}(C)$, i.e. C is the kernel of π.

So Serre subcategories are used to define an analogy to quotients on categories. The pair $(A/C, \pi)$ has the following universal property: Given an exact functor $\pi' : A \to B$ such that C is in the kernel of π', then there is a unique exact functor $\phi : A/C \to B$ such that $\phi \circ \pi = \pi'$.

$$\begin{array}{ccc} A & \xrightarrow{\quad \pi' \quad} & B \\ & {}_{\pi}\searrow \quad \nearrow^{\phi} & \\ & A/C & \end{array}$$

We in fact have an explicit construction of the category A/C. Define a C-equivalence in A to be a morphism $\psi : F \to G$ in A such that $\ker \psi, \operatorname{coker} \psi \in \mathcal{C}$. Then A/C is obtained by inverting all C-equivalences [30].

The definition of a noise system looks very similar to that of a Serre category. And indeed we use noise systems to define the functors which we consider noise at a scale ε. It therefore makes sense to want to take the quotient $\mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)/S_\tau$ for a noise system $\{S_\varepsilon\}_\varepsilon$. However this is not always possible, since in general S_ε is not a Serre subcategory; while it is true that given an SES in $\mathrm{Tame}(\mathbb{Q}^r_{\geq 0})$

$$0 \to F \to G \to H \to 0$$

then $G \in S_\varepsilon \implies F, H \in S_\varepsilon$, the converse is not in general true. What we can say is that the components S_0 and $\bigcup_\varepsilon S_\varepsilon$ are always Serre subcategories of $\mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)$.

Example 5.8.3. Let $\{S_\varepsilon\}_\varepsilon$ be the noise system associated to the standard contour. Then recall that an ε-equivalence is a morphism $\phi : F \to G$ such that the following diagram commutes for all $x \in \mathbb{Q}^r_{\geq 0}$

$$\begin{array}{ccccccc} \ker \phi_x & \longrightarrow & F(x) & \xrightarrow{\phi_x} & G(x) & \longrightarrow & \operatorname{coker} \phi_x \\ {\scriptstyle 0}\downarrow & & \downarrow & & \downarrow & & \downarrow{\scriptstyle 0} \\ \ker \phi_{x+\vec{\varepsilon}} & \longrightarrow & F(x+\vec{\varepsilon}) & \xrightarrow{\phi_{x+\vec{\varepsilon}}} & G(x+\vec{\varepsilon}) & \longrightarrow & \operatorname{coker} \phi_{x+\vec{\varepsilon}} \end{array}$$

where the vertical maps are the structure maps of F and G. Then following the construction above, $\pi_0 : S_0 \to \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)/S_0$ inverts all 0-equivalences. However 0-equivalences in this case are already isomorphisms, so the categories $\mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)$ and $\mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)/S_0$ are equivalent. The other case to look at is $\pi_\infty : \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k) \to \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)/\bigcup_\varepsilon S_\varepsilon$. Here, any ε-equivalence is mapped to an isomorphism, so any two tame persistence modules become isomorphic whenever they are ε-interleaved for any ε. ♢

From the above example it is clear that we would like more control over the size of the noise we can quotient out. For this we want to introduce the notion of a Serre noise system.

Definition 5.8.4. A Serre noise system, or simple a Serre system is a noise system $\{S_\varepsilon\}_\varepsilon$ where every component S_ε is a Serre subcategory of the category of tame functors.

Example 5.8.5. As was mentioned previously, the domain noise is closed under direct sums. However it is also true that it is a Serre system. Consider the support noise defined by $\mathcal{X} = \{X_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}} \subset \mathbb{Q}^r_{\geq 0}$. Then given an exact sequence of tame persistence modules,

$$0 \to F \to G \to H \to 0$$

if $F, H \in \mathcal{X}_\varepsilon$, then in particular $F, H = 0$ on $\mathbb{Q}^r_{\geq 0} \backslash X_\varepsilon$. Then by exactness, G is also trivial there, and so $G \in X_\varepsilon$. If we consider the functor $\pi_\varepsilon : \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) \to \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k)/\mathcal{X}_\varepsilon$, this functor sends ε-equivalences to isomorphisms. In particular, two tame persistence modules which are pointwise isomorphic on $\mathbb{Q}^r_{\geq 0} \backslash X_\varepsilon$ are mapped to isomorphic objects under π_ε. ◊

Given an exact sequence of tame functors

$$0 \to F \to G \to H \to 0$$

it is always the case that $G(v) \cong F(v) \oplus H(v)$ since all short exact sequences in Vect_k split. In a Serre system, we can additionally say that any functors belonging to the same short exact sequence belong to the same component of the system. In other words, the fact that $G(v) \cong F(v) \oplus H(v)$ holds implies that F, G, H are contained in the same component of the Serre system. Therefore, in a Serre system $\mathcal{S} = \{S_\varepsilon\}_\varepsilon$ it is sufficient to look at persistence modules pointwise when we want to compute $\text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k)/S_\varepsilon$ *i.e.* we can ignore the structure maps. Thus the definition of the components S_ε do not depend on the structure maps. Conversely, any such components which satisfy the conditions of being a noise system must define a Serre system.

Example 5.8.6 (Dimension noise). Let $\{n_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$ be a sequence of natural numbers where $n_0 = 0$ and $n_\tau + n_\varepsilon \leq n_{\tau+\varepsilon}$. We can define a noise system, called the **dimension noise** by

$$S_\varepsilon = \{F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) \,|\, \dim F \leq n_\varepsilon\}.$$

The fact that this defines a noise system depends only on the conditions we have placed on the sequence $\{n_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}}$. It is also clear that this defines a Serre system. ◊

A natural question to ask is if there are any Serre systems which are simple. The answer is yes, and in fact we have a way of generating such systems. Let $\{C_\varepsilon\}_\varepsilon$ be a noise system. Then for $\varepsilon \in \mathbb{Q}_{\geq 0}$, define

$$X_\varepsilon = \{v \in \mathbb{Q}^r_{\geq 0} | k(v, -) \in C_\varepsilon\} \subset \mathbb{Q}^r_{\geq 0}.$$

Note that $\tau \leq \varepsilon \implies C_\tau \subset C_\varepsilon \implies X_\tau \subset X_\varepsilon$. Therefore, we can consider the domain noise $\mathcal{X}$ with respect to $\{X_\varepsilon\}_\varepsilon$, in particular

$$\begin{aligned} \mathcal{X}_\varepsilon &= \{F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) | F(v) = 0 \text{ if } v \notin X_\varepsilon\} \\ &= \{F \in \text{Tame}(\mathbb{Q}^r_{\geq 0}, \text{Vect}_k) | F(v) = 0 \text{ if } k(v, -) \notin C_\varepsilon\}. \end{aligned}$$

Consider $w \in X_\varepsilon$. Then $k(w, -) \in C_\varepsilon \implies k(v, -) \in C_\varepsilon$ for all $w \leq v$ since $k(v, -)$ is a subfunctor of $k(w, -)$. Therefore $v \in X_\varepsilon$ for all $w \leq v$. Now $k(w, v) \neq 0$ iff $w \leq v$. But as we have seen, any v such that $k(v, -) \notin C_\varepsilon$ must be less that w. It follows that $k(w, -) \in \mathcal{X}_\varepsilon$. Similarly, if F has a minimal set of generators which all belong to X_ε, then $F \in \mathcal{X}_\varepsilon$. It was shown in [19] that this system is simple, and considering the above characterization of Serre systems it is clearly Serre. The associated persistence contour is given by

$$\Omega_\tau(v) = \begin{cases} v & k(v, -) \notin C_\tau \\ \infty & \text{otherwise} \end{cases}$$

So given any noise system, we can generate a system which is both simple and Serre. Explicitly, we see that

$$\begin{aligned}\mathcal{X}_\varepsilon &= \{F \in \mathrm{Tame}(\mathbb{Q}^r_{\geq 0} | F(v) = 0 \text{ if } k(v,-) \notin C_\varepsilon\} \\ &= \{F \in \mathrm{Tame}(\mathbb{Q}^r_{\geq 0} | F(v \leq v) = 0 \text{ if } k(v,-) \notin C_\varepsilon\} \\ &= \{F \in \mathrm{Tame}(\mathbb{Q}^r_{\geq 0} | F(v \leq \Omega_\varepsilon(v)) = 0\}\end{aligned}$$

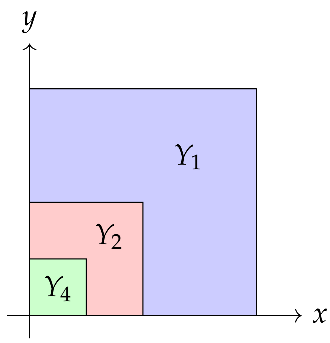
where we use the fact that $F(v \leq v)$ is always the identity, and that the definition of the noise system associated to a contour ignores those $v \in \mathbb{Q}^r$ where $\Omega_\varepsilon(v) = \infty$. From this example, it is clear that that any contour of the form

$$\Omega_\varepsilon(v) = \begin{cases} v & P_\varepsilon(v) \\ \infty & \neg P_\varepsilon(v) \end{cases}$$

corresponds to a Serre system, where here we use notation from first order logic, and $P_\varepsilon(v)$ means that v satisfies some property P_ε of the elements of $\mathbb{Q}^r_{\geq 0}$. In order for Ω to be contour, the sequence of properties $\{P_\varepsilon\}_\varepsilon$ must satisfy the condition that $P_{\varepsilon+\tau}(v) \implies P_\tau(v)$ and $P_\varepsilon(v)$ $\forall v \in \mathbb{Q}^r_{\geq 0}$. In particular, any such sequence of properties of vectors in $\mathbb{Q}^r_{\geq 0}$ yields a simple Serre system.

Example 5.8.7. Consider the decreasing sequence of sets $\{Y_\varepsilon\}_{\varepsilon \in \mathbb{Q}_{\geq 0}} \subset \mathbb{Q}^2_{\geq 0}$ given by

$$Y_\varepsilon = \begin{cases} [0, \frac{1}{\varepsilon}) \times [0, \frac{1}{\varepsilon}) & \varepsilon \geq 1 \\ [0,1) \times [0,1) & \varepsilon \leq 1. \end{cases}$$



Then we can define a simple Serre system $\mathcal{S}$ by

$$S_\varepsilon = \{F \in \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k) | F(v) = 0 \text{ if } v \in Y_\varepsilon\}.$$

The associated contour is then given by

$$\Omega_\varepsilon(v) = \begin{cases} v & v \in Y_\varepsilon \\ \infty & \text{otherwise.} \end{cases}$$

Consider the tame functor $F = k((0,1/4),-) \oplus k((1/2,0),-)$. Then F^τ has generators in $F(\Omega_\tau((0,1/4)))$ and $F(\Omega_\tau((1/2,0)))$. In particular,

$$F^\tau = \begin{cases} F & \tau \leq 2 \\ k((0,1/4),-) & 2 < \tau \leq 4 \\ 0 & \tau > 4. \end{cases}$$

The exact functor $\pi_\varepsilon : \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k) \to \mathrm{Tame}(\mathbb{Q}^r_{\geq 0}, \mathrm{Vect}_k)/S_\varepsilon$ identifies persistece modules which are pointwise isomorphic on Y_ε. ♢

Chapter 6

Applications

Both of the analyses in this chapter we performed using python, largely using software packages [33] [32].

6.1 Textbook example

We will now use the techniques described in chapter 5 to differentiate between different shapes in the plane. Consider a circle, triangle and square in the plane, as in figure 6.1. Note that all of these shapes are homeomorphic to a circle, and so are indistinguishable via traditional techniques of algebraic topology.

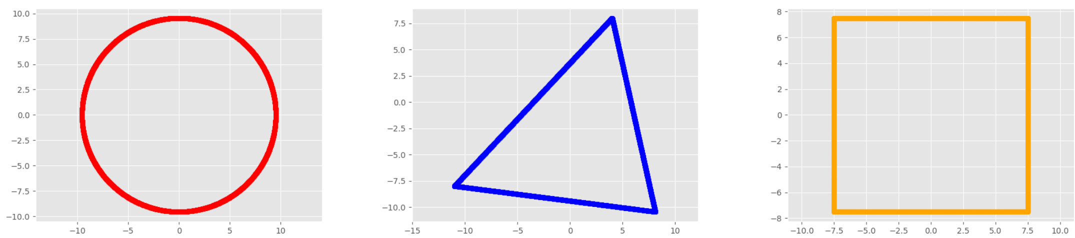


FIGURE 6.1

In order to build a Vietoris-Rips complex from which we can build $H_0(-)$ and $H_1(-)$ persistence modules, we first randomly generate 5000 points centered around each shape, which gives us the point clouds in 6.2.

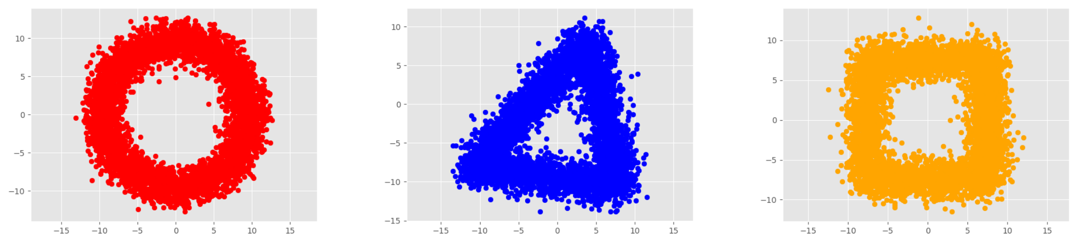


FIGURE 6.2

The strategy will be to take a sample of 100 points from each shape, form the Vietoris-Rips complex of the resulting point cloud, and from this compute $\widehat{\text{rank}}H_0(-)$ and $\widehat{\text{rank}}H_1(-)$. We will then repeat this 2000 times, and take the average of the stable ranks and plot them. This process will provide a signature for each shape which clearly distinguishes them visually. We will also do the same

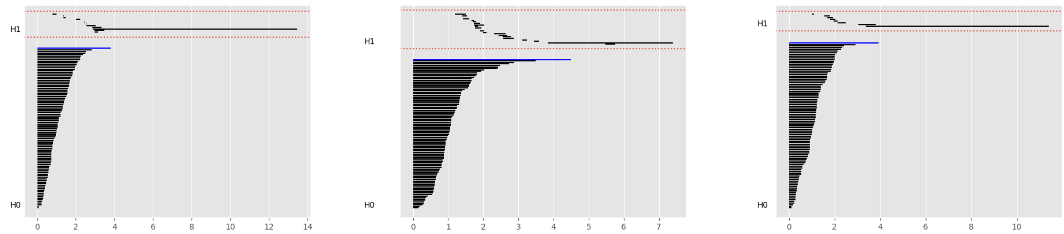


FIGURE 6.3: The barcode for the circle (left), triangle (middle) and square (right)

for a fourth shape, which is just the circle with some noise added, as shown in figure 6.4. The process that we have described will produce a signature for the "noisy" circle which is indistinguishable from that of the circle. This shows that the stable rank is an effective tool at ignoring noise in the data, which as we mentioned in the introduction has been one of the motivations for persuing persistent homology from the beginning. The barcodes for a sample of 100 points from the square,triangle and circle are shown in figure . Note that the barcode of the noisy circle would be very different from the circle. In fact, adding just a single point to the center of the circle would drastically change its barcode.

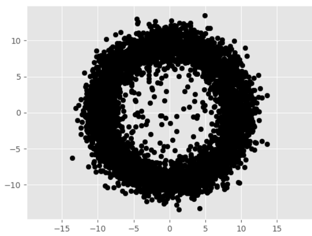


FIGURE 6.4: The noisy circle

For each sample of 100 points, we compute $\widehat{\text{rank}}H_0(-)$, $\widehat{\text{rank}}H_1(-)$ and $\widehat{\text{rank}}H_0(-)/\widehat{\text{rank}}H_1(-)$. This third quantity is what will comprise the aforementioned signature of each shape. For example, figure 6.5 shows the graph of $\widehat{\text{rank}}H_0(-)$, $\widehat{\text{rank}}H_0(-)$ and $\widehat{\text{rank}}H_0(-)/\widehat{\text{rank}}H_1(-)$ for one sample of 100 points taken from the triangle.

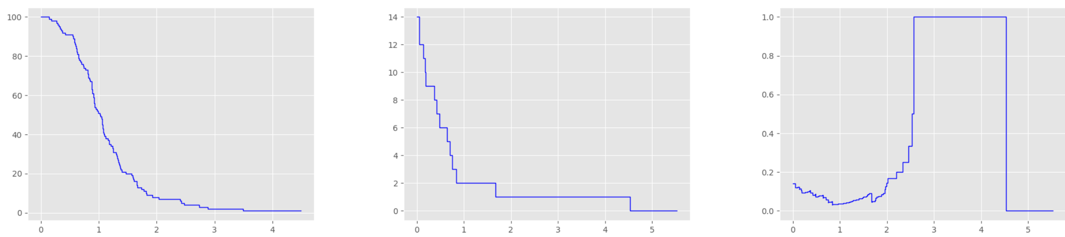


FIGURE 6.5: The stable rank for H_0 (left), H_1 (middle) and H_1/H_0 (right) of a sample of 100 points from the triangle

After doing this 2000 times for each shape, we will have 2000 different stable rank functions for reach shape. We take the average of these and plot them. It is clear from figure 6.6 that each shape has a distinct signature, and that the circle and noisy circle are indistinguishable using this method.

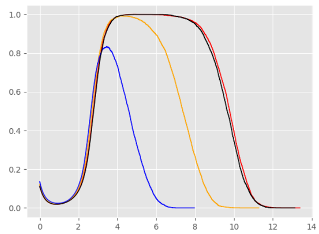


FIGURE 6.6: Plots of the averaged $\widehat{\text{rank}}H_1/\widehat{\text{rank}}H_0$ functions for the different shapes. blue = triangle, orange = square, red = circle, black = noisy circle.

6.2 Heart disease data

We now look at some real life data concerning heart disease patients, available at [31]. The data contains information about 303 people, and each data point is comprised of 14 variables including age, sex and resting blood pressure. Of the patients in question, 165 had heart disease and 138 did not. Our goal will be to find a suitable invariant and metric to distinguish between a sample of healthy people and a sample of people with heart disease. To do so, we first split the data into training and test data. The former contained 202 of the data points, and the latter 101 points. We then further split these sets by whether or not the patient had heart disease. So we had four data sets, denoted TestPos, TestNeg, TrainPos and TrainNeg, where "Pos" indicated potients with heart disease, and "Neg" indicates those without. Then using the same method as in the textbook example, we computed the average stable rank of the H_0 and H_1 persistence modules with respect to the standard contour for TestPos, TestNeg, TrainPos and TrainNeg. This time we took 400 samples of 20 points to compute the average. The plot of these averages is shown in figure 6.7. Visually, $\widehat{\text{rank}}H_1$ appears to do a good job

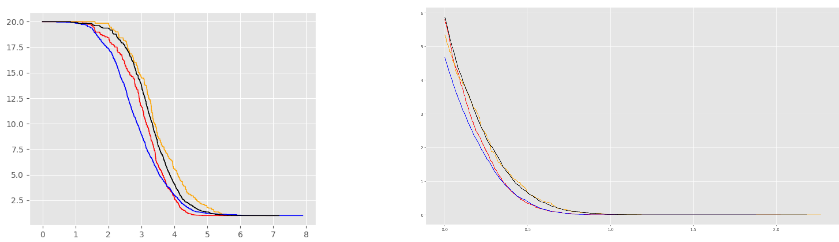


FIGURE 6.7: Plots of the average of 400 samples for $\widehat{\text{rank}}H_0$ and $\widehat{\text{rank}}H_1$. red = TestPos, orange = TestNeg, blue = TrainPos, black = TestNeg.

of identifying and distinguishing between TestPos and TestNeg. We wish to measure to what degree this is actually the case. Let TestPos_{H_i} denote the function $\widehat{\text{rank}}H_i$ for TestPos, and similarly for TestNeg, TrainPos and TrainNeg, and let $d_1(-,-)$ denote the metric associated with the L^1 norm. Then we want to compute $d_1(\text{TestPos}_{H_i}, \text{TrainPos}_{H_i}), d_1(\text{TestPos}_{H_i}, \text{TrainNeg}_{H_i}), d_1(\text{TestNeg}_{H_i}, \text{TrainPos}_{H_i})$, and $d_1(\text{TestNeg}_{H_i}, \text{TrainNeg}_{H_i})$. If the stable rank invariant (w.r.t the standard contour) for H_i is a good invariant for distinguishing healthy and unhealthy samples, then we should see

$$d_1(\text{TestPos}_{H_i}, \text{TrainPos}_{H_i}) < d_1(\text{TestPos}_{H_i}, \text{TrainNeg}_{H_i})$$

and

$$d_1(\text{TestNeg}_{H_i}, \text{TrainNeg}_{H_i}) < d_1(\text{TestNeg}_{H_i}, \text{TrainPos}_{H_i}).$$

The confusion matrices for these distances are shown in figure 6.8. We can see from these that the stable rank of H_0 is unable to definitively identify a sample of people with heart disease, but does a good job of recognising a healthy sample. Confirming the visual inspection, the stable rank of H_1 provides a better overall assessment, correctly categorizing both TestPos and TestNeg.

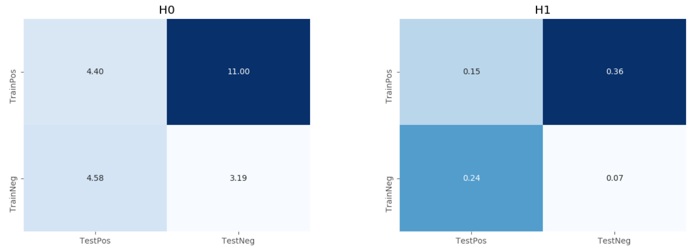


FIGURE 6.8: Confusion matrices for d_1

We can repeat this process, but this time stabilizing the rank w.r.t the noise system generated by the contour of distance type associated with the function shown in figure 6.9.

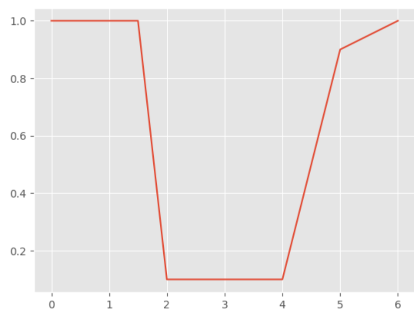


FIGURE 6.9

This produces new functions $\widehat{\text{rank}}H_0$, $\widehat{\text{rank}}H_1$. The graphs of these functions are shown in figure 6.10. The confusion matrices for the distances between the $\widehat{\text{rank}}H_0, \widehat{\text{rank}}H_1$ functions are given in figure 6.11. This contour puts greater emphasis on the homology generators that are present for mid-range values of the parameter ϵ in the Vietoris-Rips complex. We see that in this case, $\widehat{\text{rank}}H_0$ acts as a better signature for identifying samples of people with heart disease than when we stabilize the rank with respect to the standard contour.

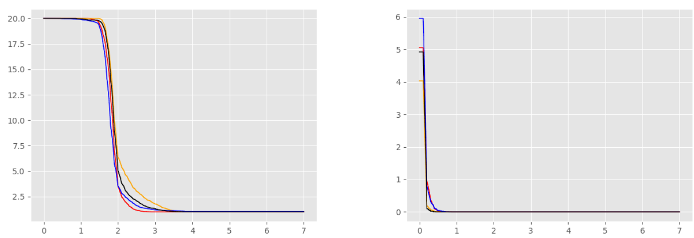


FIGURE 6.10: $\widehat{\text{rank}}H_0$ (left) and $\widehat{\text{rank}}H_1$ (right) w.r.t the contour defined by the function in figure 6.9.

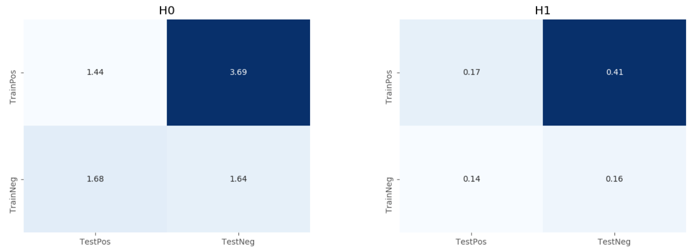


FIGURE 6.11: Confusion matrices for d_1 w.r.t. the new contour.

There is another way of measuring the effectiveness of the signatures $\widehat{\text{rank}}H_0$ and $\widehat{\text{rank}}H_1$. The procedure is to take the average of the stable rank across 400 samples of 20 points for TrainPos and TrainNeg, as before. Then, given a sample of 20 points from TestPos, we compute the function rank H_i of the resulting Vietoris-Rips complex and measure the distance in the L^1 norm to $\text{TrainPos}_{H_i}, \text{TrainNeg}_{H_i}$. If the distance to TrainPos_{H_i} is less than the distance to TrainNeg_{H_i}, we consider this a success. Then repeating this process for 400 samples of TestPos, we obtain a success rate for the signature $\widehat{\text{rank}}H_i$. We then do the same thing for TestNeg. For the standard contour, we get the success rates shown in the confusion matrices in figure 6.12. The values given in these confusion matrices are the succes rates as percentages.

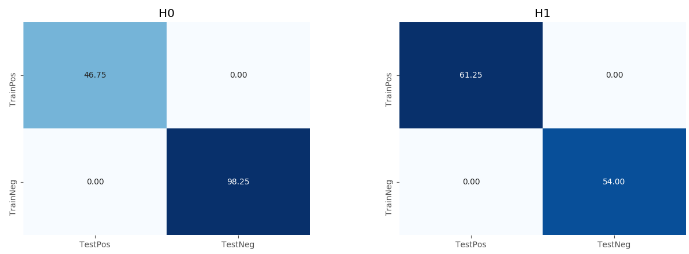


FIGURE 6.12: Success rate confusion matrices w.r.t the standard contour

Using this process makes it easier to compare the relative effectiveness of the different signatures. From figure 6.12, we can see that $\widehat{\text{rank}}H_0$ successfully recognises healthy patients very effectively, whereas $\widehat{\text{rank}}H_1$ is better at identifying those with heart disease. Figure 6.13 gives the corresponding confusion matrix when we replace the standard contour by the contour of distance type associated with the function shown in figure 6.9. This agrees with the previous observation that stabilizing w.r.t. this contour improves $\widehat{\text{rank}}H_0$ as a signature for detecting patients with heart disease.

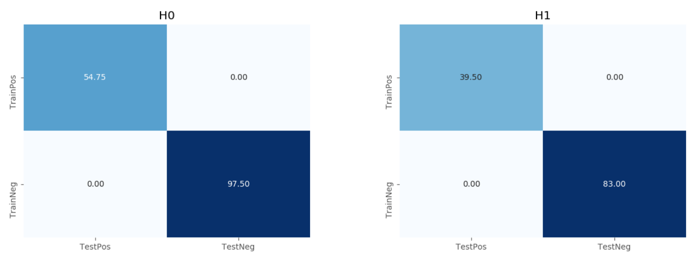


FIGURE 6.13: Success rate confusion matrices w.r.t the new contour.

The general idea of the approach taken so far is to look at a sample of the population and be able to discern if those people are at risk of heart disease. But of course, we would also like to know what the underlying cause of heart disease is, or at the very least which characteristics correlate strongly with the presence of heart disease. Ideally, the methods discussed so far should reflect when a particular variable or group of variables are key indicators of heart disease. Based on previous analyses of this data set [31], some of the most prominent variables in this regard have been found to be the level of chest pain experienced by a patient, the number of major blood vessels they have, their maximum heart rate, whether or not they experience exercise induced angina, and two other variables which are related to a patient's electrocardiogram (ECG). If we run the same analysis as above w.r.t the standard contour for the dataset restricted to only the above variables, we get the success rates shown in figure 6.14.

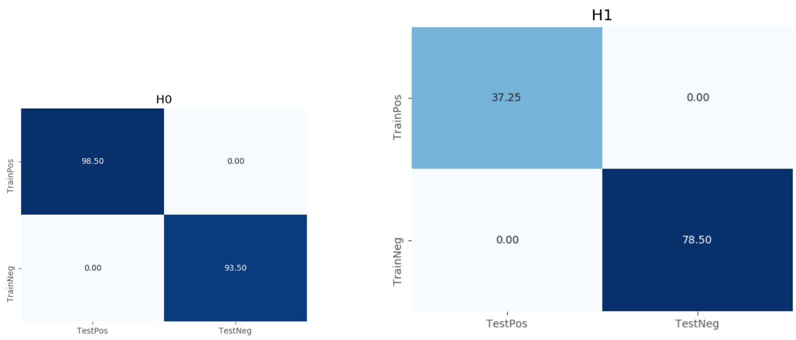


FIGURE 6.14: Success rate confusion matrices w.r.t. the standard contour when we restrict the dataset to the most important variables.

We can clearly see that in this context $\widehat{\text{rank}}H_0$ is an extremely effective signature.

Bibliography

[1] Chazal, F., de Silva, V.,Glisse, M. and Oudot, S. *The Structure and Stability of Persistence Modules*. Springer, 2016.

[2] Azumaya, G. *Corrections and supplementaries to my paper concerning Krull-Remak-Schmidt's theorem*. orem. Nagoya Mathematical Journal, 1:117-124, 1950.

[3] Carlsson, G. and de Silva, V. *Zigzag persistence*. Foundations of Computational Mathematics, 10(4):367-405, 2010.

[4] Crawley-Boevey, W. *Decomposition of pointwise finite-dimensional persistence modules*. Journal of Algebra and Its Applications, Vol. 14, No. 05, 1550066, 2015.

[5] Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L. and Oudot, S. *Proximity of Persistence Modules and their Diagrams*. Proceeding of the 25th Annual ACM Symposium on Computational Geometry (SoCG), pages 237-246, 2009.

[6] Zomorodian, A. and Carlsson, G.: *Computing persistent homology*. Discrete and Computational Geometry, 33(2):249-274, 2005.

[7] Corbet, R. and Kerber, M.: *The Representation Theorem of Persistent Homology Revisited and Generalized*. Preprint `arXiv:1707.08864v3 [math.AT]`, 2018.

[8] Ghrist, R.: *Barcodes: The persistent topology of data*. Bull. Amer. Math. Soc. 45 61-75, 2008.

[9] Cohen-Steiner, D., Edelsbrunner, H. Harer, J.:*Stability of Persistent Diagrams*. Discrete and Computational Geometry, 37:103–120 , 2007.

[10] Bubenik, P., De Silva, V., Scott, J.: *Metrics for generalized persistence modules*. Found. Comput. Math. 15(6), 1501–1531 (2015)

[11] Schwede, S. and Shipley, B.: *Equivalences of monoidal model categories*. Algebr. Geom. Topol. 3 (2003), 287–334

[12] Etingof, P., Gelaki, S., Nikshych, D. and Ostrik, V.:*Tensor categories*. Mathematical Surveys and Monographs, Volume 205, American Mathematical Society, 2015

[13] Patel, A.: *Generalized persistence diagrams*. J. Appl. Comput. Topol. 1(3), 397–419, 2018

[14] McCleary, A. and Patel, A.: *Botleneck Stability for Generalized Persistence Diagrams*. Preprint `arXiv:1806.00170 [math.AT]`, 2018

[15] Gunnar, C. and Zomorodian, A.: *The Theory of Multidimensional Persistence*. Discrete & Computational Geometry. 42:71-93, 2009

[16] Chacholski, W. and Riihimaki, H.: *Metrics and stabilization in one parameter persistence*. Preprint `arXiv:1904.02905 [math.AT]`, 2019

[17] Scolamiero, M., Chacholski, W., Lundman, A. et al.: *Multidimensional persistence and noise*. Found Comput Math 17: 1367, 2017

[18] Riihimaki, H. and Chacholski, W.: *Generalized persistence analysis based on stable rank invariant*. Preprint `arXiv:1807.01217 [cs.CG]`, 2018

[19] Gäfvert, O. and Chacholski, W.:*Stable invariants for multidimensional persistence*. Preprint `arXiv:1703.03632 [math.AT]`, 2017

[20] Lesnick, M.: *The theory of the interleaving distance on multidimensional persistence modules*. Found Comput Math 15: 613, 2015

[21] Gäfvert, O.: *Algorithms for multidimensional persistence*. Master's thesis, 2016

[22] Hatcher, A.: *Algebraic Topology*. Cambridge, Cambridge University Press, 2002

[23] Webb, C.: *Decomposition of graded modules*. Proc. Am. Math. Soc. 94: 565-571, 1986

[24] Mac Lane, S.: *Categories for the working mathematician*. New York, Springer-Verlag New York, 1978

[25] Nicolau, S., Levine, A. amd Carlsson, G.: *Topology based data analysis identifies a subgroup of breast with a unique mutational profile and excellent survival*. Proceedings of the National Academy of Sciences 108(17):7265-70, 2011

[26] Bendich, P. et al.: *Persistent homology analysis of brain artery trees*. Ann Appl Stat. 10(1): 198–218, 2016

[27] Keller, B., Lesnick, M. and Willke, T.: *Persistent homology for virtual screening*. ChemRxiv preprint, 2018

[28] Topaz, C., Ziegelmeier, L. and Halverson, T.: *Toplogical data analysis of biological aggregation models*. PloS One, 10(5):e0126383, 2015

[29] Singh, G., et al.: *Topological analysis of population activity in visual cortex*. Journal of Vision 8(8):11, 1–18, 2008

[30] *The stacks project*. `https://stacks.math.columbia.edu/tag/02MN`

[31] *Heart disease UCI*. `https://www.kaggle.com/ronitf/heart-disease-uci`

[32] Chacholski, W.: *stableRANK*. Software, not yet distributed.

[33] Gäfvert, O.:*Topcat*. `https://github.com/olivergafvert/topcat`