DISSERTATION

GENERALIZATIONS OF PERSISTENT HOMOLOGY

Submitted by

Alexander J. McCleary

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2021

Doctoral Committee:

Advisor: Amit Patel

Henry Adams
Asa Ben-Hur
Chis Peterson

ABSTRACT

GENERALIZATIONS OF PERSISTENT HOMOLOGY

Persistent homology typically starts with a filtered chain complex and produces an invariant called the persistence diagram. This invariant summarizes where holes are born and die in the filtration. In the traditional setting the filtered chain complex is a chain complex of vector spaces filtered over a totally ordered set. There are two natural directions to generalize the persistence diagram: we can consider filtrations of more general chain complexes and filtrations over more general partially ordered sets. In this dissertation we develop both of these generalizations by defining persistence diagrams for chain complexes in an essentially small abelian category filtered over any finite lattice.

ACKNOWLEDGEMENTS

I would like to thank first and foremost my advisor, Amit Patel. Amit's guidance has been a real blessing to me. His creativity and intuition in finding new directions is a big inspiration for me. I would also like to thank all my friends and family for making this process much more enjoyable.

DEDICATION

This dissertation is dedicated to Huxley.

TABLE OF CONTENTS

Chapter 1

Introduction

Persistent homology is a tool for studying filtered chain complexes. Roughly speaking, a filtered chain complex is a nested family of chain complexes. Filtered chain complexes arise naturally in data analysis from Rips complexes [1], Čech complexes [2], and alpha complexes [3]. These are all different ways to convert a finite metric space into a filtered chain complex that captures information about the metric at different scale parameters. As the chain complexes in a filtration get bigger, holes open up and holes close up. Persistent homology captures this behavior with the persistence diagram which encodes where each hole first opens up and where it first closes (if it closes). The persistence diagram condenses a lot of the meaningful topological structure of a filtered simplicial complex into a much more usable form, giving valuable insights into the shape of data [4, 5]. Persistent homology also has deep applications to pure math in areas such as fractal geometry [6–8] and symplectic geometry [9, 10]. One of the most important results in persistent homology is bottleneck stability which says that the construction of the persistence diagram is 1-Lipschitz. In practice, data is inherently noisy so this result is crucial for applications.

Homology is used to convert a filtered simplicial complex into a persistence diagram. Traditionally, for persistence, the homology is defined with field coefficients. In [11], a method for defining persistence diagrams with coefficients in a ring was introduced. This allows for the study of torsion in data, capturing more information from a filtered simplicial complex. In [12], we showed that these generalized persistence diagrams satisfy bottleneck stability as well.

One of the unfortunate weaknesses of the filtered simplicial complexes used in topological data analysis is that they are sensitive to outliers. Because of this, it is often times desirable to incorporate a density parameter into the filtration. This leads to multiparameter persistent homology, an area rich with potential applications but lacking a robust theory [13]. In [14], we showed that multiparameter persistence diagrams satisfy a form of stability and a notion of functoriality. In fact, the results from this paper extend beyond multiparameter persistence to persistent homology over any finite lattice. This structure has numerous potential applications to areas such as dynamic metric spaces [15] and the persistent homology transform [16].

1.1 Outline

We start this dissertation with the relevant background information. In Chapter 2 we introduce lattices and metric lattices and prove a few lemmas that are useful for proving some of our bigger results later on. We then start a brief introduction to category theory in Section 2.2. Beginning with categories and functors, we cover all the way through Kan extensions, abelian categories, and Grothendieck groups. Lastly, this chapter concludes with some background on chain complexes and homology in Section 2.3. We allow our chain complexes to lie in any essentially small abelian category $\mathcal{A}$, allowing for more general homology theories to be used in practice.

Chapter 3 focuses on filtered chain complexes. For a fixed chain complex X_* we define a category Fil of filtrations of X_*. We define the edit distance between filtrations in Fil. Our filtrations differ from the filtrations traditionally used in persistent homology in that we allow X_* to be filtered over any finite lattice P and not simply a totally ordered finite lattice. In the setting where P is totally ordered, we take advantage of this extra structure to define the interleaving distance between filtrations.

In Chapter 4, we define our main invariant: the persistence diagram. We first define birth-death functions of filtrations. These intermediary invariants capture information about where holes are born and die in a filtration. Birth death functions lie in a category $\mathsf{Mon}(\mathcal{G})$ with functors $\mathsf{ZB}_i : \mathsf{Fil} \to \mathsf{Mon}(\mathcal{G})$ taking any filtration to its i-th birth-death function. We then define persistence diagrams as the Möbius inversion of birth-death functions. Persistence diagrams lie in a category $\mathsf{Fnc}(\mathcal{G})$ and the Möbius inversion turns out to be a functor $\mathsf{MI} : \mathsf{Mon}(\mathcal{G}) \to \mathsf{Fnc}(\mathcal{G})$. We define the edit distance between functions in $\mathsf{Fnc}(\mathcal{G})$ and, when the underlying lattice is totally ordered, we also define the bottleneck distance between functions in $\mathsf{Fnc}(\mathcal{G})$.

Chapter 5 focuses on stability results. These results are crucial for applications. In Section 5.1 we show that the construction of the persistence diagram of a filtration is stable with respect to the edit distance. Section 5.2 focuses on the special case of filtrations over totally ordered lattices. Here we prove that the construction of persistence diagrams is stable with respect to the bottleneck distance. Lastly, in Section 5.3 we show that the edit distance and bottleneck distance are strongly equivalent for functions in $\mathsf{Fnc}(\mathcal{G})$ where the underlying lattice is totally ordered.

Chapter 2

Preliminaries

Here we introduce the necessary background information. We start with a short introduction to lattices followed by an introduction to category theory.

2.1 Lattices

A *poset* is a set P with a relation $\leqslant$ that is

- reflexive: $a \leqslant a$ for any $a \in P$
- antisymmetric: if $a \leqslant b$ and $b \leqslant a$ then $a = b$
- transitive: if $a \leqslant b$ and $b \leqslant c$ then $a \leqslant c$.

A poset P is *totally ordered* if for any $a, b \in P$, either $a \leqslant b$ or $b \leqslant a$. For two elements $a, b \in P$ in a poset, we write $a < b$ to mean $a \leqslant b$ and $a \neq b$. For any $a \leqslant c$, the *interval* $[a, c] \subseteq P$ is the subposet consisting of all $b \in P$ such that $a \leqslant b \leqslant c$. The poset P has a *bottom* if there is an element $\bot \in P$ such that $\bot \leqslant a$ for all $a \in P$. The poset P has a *top* if there is an element $\top \in P$ such that $a \leqslant \top$ for all $a \in P$. A function $\alpha : P \to Q$ between two posets is *monotone* if for all $a \leqslant b$, $\alpha(a) \leqslant \alpha(b)$.

The *meet* of two elements $a, b \in P$ in a poset, written $a \wedge b$, is the greatest lower bound of a and b. The *join* of two elements $a, b \in P$ in a poset, written $a \vee b$, is the least upper bound of a and b. The poset P is a *lattice* if both joins and meets exist for all pairs of elements in P. A lattice is *bounded* if it contains both a top and a bottom. If P is a finite lattice, then the existence of meets and joins implies that P has a bottom and a top, respectively. Therefore all finite lattices are bounded. A function $\alpha : P \to Q$ between two bounded lattices is a *bounded lattice function* if $\alpha(\top) = \top$, $\alpha(\bot) = \bot$, and for all $a, b \in P$, $\alpha(a \vee b) = \alpha(a) \vee \alpha(b)$ and $\alpha(a \wedge b) = \alpha(a) \wedge \alpha(b)$. Note that bounded lattice functions are monotone. This is because $a \leqslant b$ if and only if $a \wedge b = a$,

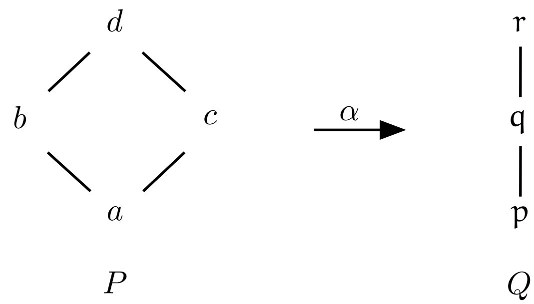


Figure 2.1: Above are the Hasse diagrams of two finite metric lattices P and Q. The metrics d_P and d_Q assigns to every pair of elements the length (i.e. the number of edges) of the shortest path between them. For example, $d_P(a, d) = 2$ and $d_Q(p, q) = 1$. The function $\alpha : P \to Q$ defined as $\alpha(a) = \alpha(b) = p$ and $\alpha(c) = \alpha(d) = r$ is a bounded lattice function. The distortion of α is $||\alpha|| = 1$.

and therefore

$$\alpha(a) = \alpha(a \wedge b) = \alpha(a) \wedge \alpha(b) \Longrightarrow \alpha(a) \leqslant \alpha(b).$$

Example 2.1.1: See Figure 2.1 for two examples of finite metric lattices P and Q and a morphism of finite metric lattices $\alpha : P \to Q$. The distortion of α is $||\alpha|| = 1$. Forthcoming examples will build on this one example.

Proposition 2.1.2: Let P and Q be finite lattices and $\alpha : P \to Q$ a bounded lattice function. Then for all $a \in Q$, the pre-image $\alpha^{-1}[\bot, a]$ has a maximal element.

Proof. The pre-image is non-empty because $\alpha(\bot) = \bot$. The pre-image is finite because P is finite. For any two elements b and c in the pre-image, both $b \vee c$ and $b \wedge c$ are also in the pre-image because

$$\alpha(b \vee c) = \alpha(b) \vee \alpha(c) \leqslant a \vee a = a \qquad \alpha(b \wedge c) = \alpha(b) \wedge \alpha(c) \leqslant a \wedge a = a.$$

Thus $\alpha^{-1}[\bot, a]$ is a finite lattice and all finite lattices have a unique maximal element. □

For a finite lattice P, let $\bar{P} := \{[a, b] \subseteq P : a \leqslant b\}$ be its set of intervals. The product order on $P \times P$ restricts to a partial order $\preceq$ on $\bar{P}$ as follows: $[a, b] \preceq [c, d]$ if $a \leqslant c$ and $b \leqslant d$. The join of two intervals is $[a, b] \vee [c, d] = [a \vee c, b \vee d]$, and the meet of two intervals is

$[a,b] \wedge [c,d] = [a \wedge c, b \wedge d]$. All this makes $\bar{P}$ a finite lattice. Its bottom element is $[\bot, \bot]$ and its top element is $[\top, \top]$.

A bounded lattice function $\alpha : P \to Q$ between two finite lattices induces a bounded lattice function $\bar{\alpha} : \bar{P} \to \bar{Q}$ as follows. For an interval $[a,b] \in \bar{P}$, let $\bar{\alpha}\big([a,b]\big) := \big[\alpha(a), \alpha(b)\big]$. We have

$$\begin{aligned}
\bar{\alpha}\big([a,b] \wedge [c,d]\big) &= \bar{\alpha}\big([a \wedge c, b \wedge d]\big) = \big[\alpha(a \wedge c), \alpha(b \wedge d)\big] \\
&= \big[\alpha(a) \wedge \alpha(c), \alpha(b) \wedge \alpha(d)\big] = \bar{\alpha}\big([a,b]\big) \wedge \bar{\alpha}\big([c,d]\big). \\
\bar{\alpha}\big([a,b] \vee [c,d]\big) &= \bar{\alpha}\big([a \vee c, b \vee d]\big) = \big[\alpha(a \vee c), \alpha(b \vee d)\big] \\
&= \big[\alpha(a) \vee \alpha(c), \alpha(b) \vee \alpha(d)\big] = \bar{\alpha}\big([a,b]\big) \vee \bar{\alpha}\big([c,d]\big) \\
\bar{\alpha}\big([\bot,\bot]\big) &= \big[\alpha(\bot), \alpha(\bot)\big] = [\bot,\bot] \\
\bar{\alpha}\big([\top,\top]\big) &= \big[\alpha(\top), \alpha(\top)\big] = [\top,\top].
\end{aligned}$$

Thus $\bar{\alpha}$ is a bounded lattice function. Further, for any pair of bounded lattice functions $\alpha : P \to Q$ and $\beta : Q \to R$, $\overline{\beta \circ \alpha} = \bar{\beta} \circ \bar{\alpha}$.

Example 2.1.3: The morphism of finite metric lattices $\alpha : P \to Q$ in Example 2.1.1 induces the morphism of finite metric lattices $\bar{\alpha} : \bar{P} \to \bar{Q}$ in Figure 2.2. The distortion of $\bar{\alpha}$ is $||\bar{\alpha}|| = 1$.

2.1.1 Metric Lattices

A *finite metric lattice* is a tuple (P, d_P) where P is a finite lattice and $d_P : P \times P \to \mathbb{R}^{\geqslant 0}$ a metric. A *morphism of finite metric lattices* $\alpha : (P, d_P) \to (Q, d_Q)$ is a bounded lattice function $\alpha : P \to Q$. The *distortion* (see [17, Definition 7.1.4]) of a morphism $\alpha : (P, d_P) \to (Q, d_Q)$, denoted $||\alpha||$, is

$$||\alpha|| := \max_{a,b \in P} \big| d_P(a,b) - d_Q\big(\alpha(a), \alpha(b)\big) \big|.$$

To minimize notation, we will write finite metric lattices (P, d_P) simply as P with the implied metric d_P.

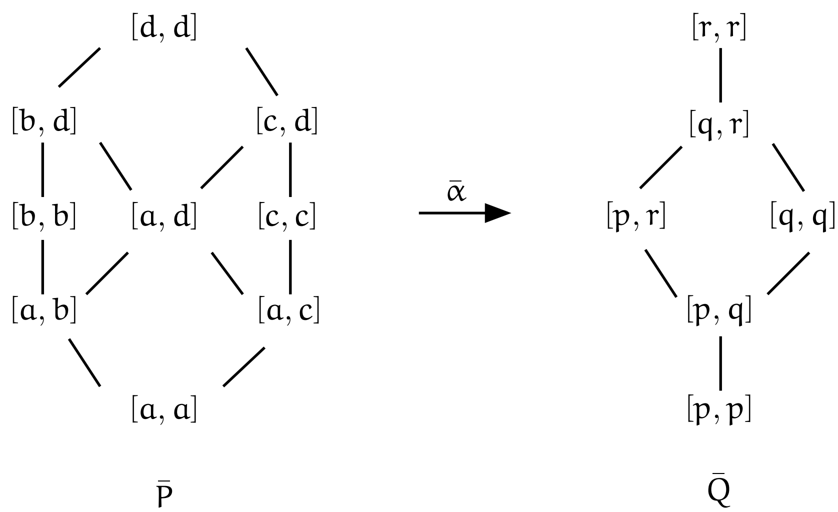


Figure 2.2: Above are the Hasse diagrams of the lattices $\bar{\mathsf{P}}$ and $\bar{\mathsf{Q}}$ where P and Q are from Example 2.1.1. The morphism $\alpha : \mathsf{P} \to \mathsf{Q}$ from the same example extends to the morphism $\bar{\alpha} : \bar{\mathsf{P}} \to \bar{\mathsf{Q}}$ as follows. The function $\bar{\alpha}$ sends $\{[a,a],[a,b],[b,b]\}$ to $\{[p,p]\}$, $\{[a,c],[a,d],[b,d]\}$ to $\{[p,r]\}$, and $\{[c,c],[c,d],[d,d]\}$ to $\{[r,r]\}$. The distortion of $\bar{\alpha}$ is $||\bar{\alpha}|| = ||\alpha|| = 1$.

For every finite metric lattice P, we have the finite metric lattice of intervals $\bar{\mathsf{P}}$ where

$$d_{\bar{\mathsf{P}}}\big([a,b],[c,d]\big) := \max\big\{d_{\mathsf{P}}(a,c), d_{\mathsf{P}}(b,d)\big\}.$$

A morphism $\alpha : \mathsf{P} \to \mathsf{Q}$ of finite metric lattices induces a morphism of finite metric lattices $\bar{\alpha} : \bar{\mathsf{P}} \to \bar{\mathsf{Q}}$. Define the *distortion* of $\bar{\alpha}$ as

$$||\bar{\alpha}|| := \max_{[a,b],[c,d] \in \bar{\mathsf{P}}} \Big| \max\big\{d_{\mathsf{P}}(a,c), d_{\mathsf{P}}(b,d)\big\} - \max\big\{d_{\mathsf{Q}}(\alpha(a),\alpha(c)), d_{\mathsf{Q}}(\alpha(b),\alpha(d))\big\}\Big|.$$

Proposition 2.1.5 says that the two distortions $||\alpha||$ and $||\bar{\alpha}||$ are equal. Its proof requires the following lemma.

Lemma 2.1.4: [18, Lemma 3 page 31] For all non-negative real numbers $w, x, y, z \in \mathbb{R}^{\geqslant 0}$,

$$\big|\max(w,x) - \max(y,z)\big| \leqslant \max\big(|w-y|, |x-z|\big).$$

Proposition 2.1.5: Let $\alpha : P \to Q$ be a bounded lattice function between two finite metric lattices and let $\bar{\alpha} : \bar{P} \to \bar{Q}$ be the induced bounded lattice function on intervals. Then $||\bar{\alpha}|| = ||\alpha||$.

Proof. First we show $||\bar{\alpha}|| \geqslant ||\alpha||$. If $||\alpha|| = \varepsilon$, then there are elements $a, b \in P$ such that $\varepsilon = \left| d_P(a,b) - d_Q(\alpha(a), \alpha(b)) \right|$. For the intervals $[a, a]$ and $[b, b]$, we have

$$\left| \max\left\{ d_P(a,b), d_P(a,b) \right\} - \max\left\{ d_Q(\alpha(a),\alpha(b)), d_Q(\alpha(a),\alpha(b)) \right\} \right| = \varepsilon$$

proving the claim. Now we show $||\bar{\alpha}|| \leqslant ||\alpha||$ using Lemma 2.1.4:

$$\begin{aligned} ||\bar{\alpha}|| &:= \max_{[a,b],[c,d]\in\bar{P}} \left| \max\left\{ d_P(a,c), d_P(b,d) \right\} - \max\left\{ d_Q(\alpha(a),\alpha(c)), d_Q(\alpha(b),\alpha(d)) \right\} \right| \\ &\leqslant \max_{[a,b],[c,d]\in\bar{P}} \left\{ \left| d_P(a,c) - d_Q\big(\alpha(a),\alpha(c)\big) \right|, \left| d_P(b,d) - d_Q\big(\alpha(b),\alpha(d)\big) \right| \right. \\ &= ||\alpha||. \end{aligned}$$

□

2.2 Categories

In this section we provide a brief introduction to category theory. First we review categories and functors followed by an overview of abelian categories and then a little background on chain complexes and homology.

Categories were introduced in 1942 by Samuel Eilenberg and Saunders Mac Lane to study mathematical structures in a very broad sense. They have helped unify concepts from fields as diverse as topology and algebra to geometry and combinatorics. Category theory provides an elegant setting to study many different concepts, and persistent homology is no exception. Much of the theory presented here is written in the language of category theory.

A category consists of objects and morphisms. The objects often represent mathematical structures such as: sets, vector spaces, and topological spaces while the morphisms represent the appropriate structure preserving maps between them such as: functions, linear transformations and continuous maps respectively.

Definition 2.2.1: A (locally small) **category** C consists of:

- A class of objects $\mathrm{ob}(\mathsf{C})$.
- For any objects $a, b \in \mathrm{ob}(\mathsf{C})$ a set $\hom(a, b)$ of morphisms from a to b. We will often use the notation $f : a \to b$ for a morphism $f \in \hom(a, b)$.
- A composition rule for morphisms. That is, for any objects $a, b, c \in \mathrm{ob}(\mathsf{C})$ a map $\circ$: $\hom(b, c) \times \hom(a, b) \to \hom(a, c)$ satisfying:
  - For each object x, there is an identity morphism 1_x with the property that for each morphism $f : a \to b$, $1_b \circ f = f = f \circ 1_a$.
  - Composition is associative; explicitly, for any $f : c \to d$, $g : b \to c$, and $h : a \to b$, $f \circ (g \circ h) = (f \circ g) \circ h$.

Example 2.2.2: The category of sets Set has sets as objects and set maps as morphisms. The category of abelian groups Ab has abelian groups as objects and group homomorphisms as morphisms.

Example 2.2.3: A lattice P can be viewed as a category with elements of P as the objects and a unique morphism $f : a \to b$ if and only if $a \leqslant b$. Composition of morphisms follows from the transitivity of $\leqslant$.

Definition 2.2.4: A morphism $f : a \to b$ in a category C is an **isomorphism** if there exists a morphism $g : b \to a$ such that $f \circ g = 1_b$ and $g \circ f = 1_a$. An **isomorphism class** $[a]$ is the class of all objects that are isomorphic to a in C.

Definition 2.2.5: A category C is **small** if the objects of C form a set and **essentially small** if the isomorphism classes of objects form a set.

Definition 2.2.6: A morphism $f : a \to b$ in a category C is a **monomorphism** if for any object $c \in \mathrm{ob}(\mathsf{C})$ and any morphisms $g_1, g_2 : c \to a$, the equation $f \circ g_1 = f \circ g_2$ holds only if $g_1 = g_2$. Similarly, f is an **epimorphism** if for any object $c \in \mathrm{ob}(\mathsf{C})$ and any morphisms $g_1, g_2 : b \to c$ the equation $g_1 \circ f = g_2 \circ f$ holds only if $g_1 = g_2$.

Example 2.2.7: In Vec a monomorphism is simply an injective linear transformation and an epimorphism is a surjective linear transformation. In Top a monomorphism is an injective continuous map and an epimorphism is a surjective continuous map.

Definition 2.2.8: A **functor** $\mathsf{F} : \mathsf{C} \to \mathsf{D}$ from a category C to a category D consists of:

- For each $a \in \mathrm{ob}(\mathsf{C})$, an object $\mathsf{F}(a) \in \mathrm{ob}(\mathsf{D})$.
- For each $f \in \hom(a, b)$, a morphism $\mathsf{F}(f) \in \hom\big(\mathsf{F}(a), \mathsf{F}(b)\big)$ satisfying:
  - $\mathsf{F}(1_a) = 1_{\mathsf{F}(a)}$ for each $a \in \mathrm{ob}(\mathsf{C})$.
  - $\mathsf{F}(f \circ g) = \mathsf{F}(f) \circ \mathsf{F}(g)$ for any $a, b, c \in \mathrm{ob}(\mathsf{C})$ and any $f : b \to c$ and $g : a \to b$.

Example 2.2.9: There is a functor $\mathsf{F} : \mathsf{Vec} \to \mathsf{Set}$ that takes a vector space to its underlying set and a linear transformation to its underlying set map. There is another functor $\mathsf{G} : \mathsf{FinSet} \to \mathsf{Vec}$ taking any finite set to the vector space with that set as a basis and any set map to the induced linear transformation.

We now introduce natural transformations. Just as functors act as morphisms between categories, natural transformations act as morphisms between functors. Natural transformations are necessary for defining certain concepts in category theory such as Kan extensions.

Definition 2.2.10: Let C and D be categories and $\mathsf{F}, \mathsf{G} : \mathsf{C} \to \mathsf{D}$ functors from C to D. A **natural transformation** $\eta : \mathsf{F} \to \mathsf{G}$ consists of morphisms $\eta_a : \mathsf{F}(a) \to \mathsf{G}(a)$ for each $a \in \mathrm{ob}(\mathsf{C})$ such that for any $f : a \to b$ in C, the following diagram commutes in D

$$\begin{array}{ccc} \mathsf{F}(a) & \xrightarrow{\eta_a} & \mathsf{G}(a) \\ \big\downarrow{\scriptstyle \mathsf{F}(f)} & & \big\downarrow{\scriptstyle \mathsf{G}(f)} \\ \mathsf{F}(b) & \xrightarrow{\eta_b} & \mathsf{G}(b). \end{array}$$

We are now ready to define limits and colimits. These are incredibly important constructions in category theory. Examples of limits and colimits include direct sums of vector spaces and kernels of linear transformations.

Definition 2.2.11: A **diagram** in a category C is a functor $\mathsf{F} : \mathsf{J} \to \mathsf{C}$ where J is any small category. A **cone** over F consists of an object $N \in \text{ob}(\mathsf{C})$ and morphisms $\psi_a : N \to F(a)$ for each $a \in \text{ob}(\mathsf{J})$ such that for any $f : a \to b$ in J, $F(f) \circ \psi_a = \psi_b$. A **limit** of F is a cone (L, ϕ) over F such that for any cone (N, ψ) over F there is a unique morphism $u : N \to L$ with $\phi_a \circ u = \psi_a$ for any $a \in \mathsf{J}$.

Example 2.2.12: Let J be a small category with only the identity morphisms and $F : \mathsf{J} \to \mathsf{Set}$ a diagram. The limit of F is the cartesian product of the sets $F(a)$ over all $a \in \text{ob}(\mathsf{J})$ and the morphisms ϕ_a are given by projecting the cartesian product onto $F(a)$.

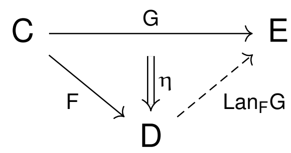
The dual notion of a limit is that of a colimit.

Definition 2.2.13: Let $F : \mathsf{J} \to \mathsf{C}$ be a diagram. A **co-cone** under F consists of an object $Q \in \text{ob}(\mathsf{C})$ and morphisms $\psi_a : F(a) \to Q$ for any $a \in \text{ob}(\mathsf{J})$ such that for any $f : a \to b$ in J, $\psi_b \circ F(f) = \psi_a$. A **colimit** of F is a co-cone (C, ϕ) under F such that for any other co-cone (Q, ψ) under F there is a unique morphism $u : C \to Q$ satisfying $u \circ \phi_a = \psi_a$ for any $a \in \text{ob}(\mathsf{J})$.

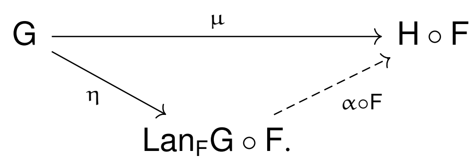
Example 2.2.14: Let $F : \mathsf{J} \to \mathsf{Set}$ be defined as in Example 2.2.12. The coproduct of F is the disjoint union of the sets $F(a)$ over all $a \in \text{ob}(\mathsf{J})$. The morphisms ϕ_a are the inclusions of $F(a)$ into the disjoint union.

One of the most useful constructions in category theory is the Kan extension. This tool will allow us to define constructible filtrations, interpolate between constructible filtrations and express morphisms between filtrations in a natural way.

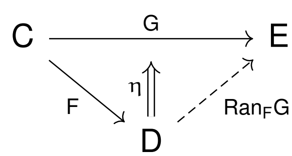
Definition 2.2.15: Let C, D, and E be categories with functors $F : \mathsf{C} \to \mathsf{D}$ and $G : \mathsf{C} \to \mathsf{E}$. The **left Kan extension** of G along F consists of a functor $\mathrm{Lan}_F G : \mathsf{D} \to \mathsf{E}$ and a natural transformations $\eta : G \to \mathrm{Lan}_F G \circ F$



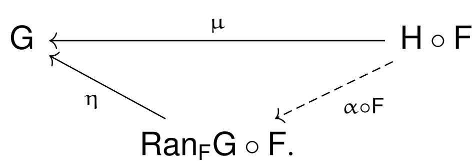
such that for any other functor $\mathsf{H} : \mathsf{D} \to \mathsf{E}$ with a natural transformation $\mu : \mathsf{G} \to \mathsf{H} \circ \mathsf{F}$ there is a unique natural transformation $\alpha : \mathsf{Lan_F G} \to \mathsf{H}$ that makes the following diagram commute



Dually, the **right Kan extension** of G along F consists of a functor $\mathsf{Ran_F G} : \mathsf{D} \to \mathsf{E}$ and a natural transformations $\eta : \mathsf{Ran_F G} \circ \mathsf{F} \to \mathsf{G}$



such that for any other functor $\mathsf{H} : \mathsf{D} \to \mathsf{E}$ with a natural transformation $\mu : \mathsf{H} \circ \mathsf{F} \to \mathsf{G}$ there is a unique natural transformation $\alpha : \mathsf{H} \to \mathsf{Ran_F G}$ that makes the following diagram commute



2.2.1 Abelian Categories

Many categories that arise in practice have more structure than a basic category. In the category of vector spaces hom-sets are more than just sets, they form vector spaces. Linear transformations can be added and scalar multiplied. Additionally, finite products and coproducts always exist in Vec. Similarly, hom-sets in the category of abelian groups also form abelian groups and finite products and coproducts always exist. One way to classify this extra structure is with abelian categories. In this section we will define abelian categories and look at a few examples.

Definition 2.2.16: An object $I \in \mathrm{ob}(\mathsf{C})$ is **initial** if for any $a \in \mathrm{ob}(\mathsf{C})$ there is a unique morphism from I to a. An object $T \in \mathrm{ob}(\mathsf{C})$ is **terminal** if for any $a \in \mathrm{ob}(\mathsf{C})$ there exists a unique morphism from a to T. An object $0 \in \text{ob}(\mathsf{C})$ is a **zero object** if it is both an initial and a terminal object.

Example 2.2.17: A finite lattice P viewed as a category has an initial object $\perp$ and a terminal object $\top$. There is no 0 object in P unless $\perp = \top$.

Example 2.2.18: In Vec a 0 object is a 0-dimensional vector space. There is exactly one morphism from any vector space to a 0-dimensional vector space and one map from a 0-dimensional space to any other vector space.

Remark 2.2.19: If J is the category with no objects and $\mathsf{F} : \mathsf{J} \to \mathsf{C}$ is a diagram then the colimit of F is an initial object in C and the limit of F is a terminal object.

Definition 2.2.20: A category C is **additive** if it has all finite products and each hom-set forms an abelian group with composition being a bilinear map $\circ : \text{hom}(b, c) \times \text{hom}(a, b) \to \text{hom}(a, c)$. Composition being bilinear means that for any $a, b, c \in \text{ob}(\mathsf{C})$ and any $f, g \in \text{hom}(a, b)$ and $h, i \in \text{hom}(b, c)$, we have that $(h + i) \circ f = (h \circ f) + (i \circ f)$ and $h \circ (f + g) = (h \circ f) + (h \circ g)$.

Definition 2.2.21: A **kernel** of a morphism $f : a \to b$ in C is a limit of the diagram on below on the left. A **cokernel** of f is a colimit of the diagram on the right.

$$\begin{array}{ccc} \ker(f) & \dashrightarrow & a \\ \downarrow & & \downarrow f \\ 0 & \longrightarrow & b \end{array} \qquad\qquad \begin{array}{ccc} a & \xrightarrow{f} & b \\ \downarrow & & \downarrow \\ 0 & \dashrightarrow & \mathsf{coker}(f) \end{array}$$

Definition 2.2.22: An **abelian category** is a category $\mathcal{A}$ such that:

- $\mathcal{A}$ is additive.
- Every morphism f in $\mathcal{A}$ has a kernel and a cokernel.
- Every monomorphism is a kernel of a morphism and every epimorphism is a cokernel of a morphism.

Example 2.2.23: Many familiar categories from algebra are abelian including the category of vector spaces, the category of abelian groups, and the category of left R-modules over a fixed ring R.

Remark 2.2.24: Mitchell's embedding theorem shows that every essentially small abelian category is equivalent to a full subcategory of $R\text{-}\mathsf{Mod}$ for some ring R [19].

Definition 2.2.25: An image of a morphism $f : a \to b$ in an abelian category is defined to be $\mathsf{im}(f) = \ker(\mathsf{coker}(f))$.

There is a way to define the image via a universal property as well however it is less intuitive and, in the setting of abelian categories, equivalent.

Definition 2.2.26: A **short exact sequence** in an abelian category is a sequence of objects and maps

$$0 \longrightarrow a \xrightarrow{f} b \xrightarrow{g} c \longrightarrow 0$$

such that $\ker(f) = 0$, $\mathsf{coker}(g) = 0$, and $\mathsf{im}(f) = \ker(g)$.

Definition 2.2.27: Let $\mathcal{A}$ be an essentially small abelian category. The **Grothendieck group** of $\mathcal{A}$ is the abelian group $\mathcal{G}(\mathcal{A})$ generated by isomorphism classes of objects in $\mathcal{A}$ and by the relations $[a] + [c] = [b]$ for any short exact sequence $0 \to a \to b \to c \to 0$. The Grothendieck group has a natural translation invariant partial ordering where $[a] \leqslant [b]$ if there is an object $c \in \mathcal{C}$ such that $[b] - [a] = [c]$. For each $a \hookrightarrow b$, we have $a \oplus c \hookrightarrow b \oplus c$ for any object c in $\mathcal{C}$. This makes $\leqslant$ a translation invariant partial ordering.

Example 2.2.28: Here are three examples of $\mathcal{A}$ with their Grothendieck groups.

- Let Vec be the category of finite dimensional k-vector spaces for some fixed field k. Every finite dimensional k-vector space is isomorphic to k^n for some natural number $n \geqslant 0$. This means that the free abelian group generated by the set of isomorphism classes in Vec is $\bigoplus_n \mathbb{Z}$ over all $n \geqslant 0$. Since every short exact sequence in Vec splits, the only relations are of the form $[A] + [B] = [C]$ whenever $A \oplus B \cong C$. Therefore $\mathcal{G}(\mathsf{Vec}) \cong \mathbb{Z}$ where the translation invariant partial ordering $\preceq$ is the usual total ordering on the integers.

- Let FinAb be the category of finite abelian groups. A finite abelian group is isomorphic to

$$\frac{\mathbb{Z}}{p_1^{n_1}\mathbb{Z}} \oplus \cdots \oplus \frac{\mathbb{Z}}{p_k^{n_k}\mathbb{Z}}$$

  where each p_i is prime. The free abelian group generated by the set of isomorphism classes in FinAb is $\bigoplus_{(p,n)} \mathbb{Z}$ over all pairs (p, n) where p is prime and $n \geqslant 0$ a natural number. Every primary cyclic group $\dfrac{\mathbb{Z}}{p^n\mathbb{Z}}$ fits into a short exact sequence

$$0 \longrightarrow \frac{\mathbb{Z}}{p^{n-1}\mathbb{Z}} \xrightarrow{\times p} \frac{\mathbb{Z}}{p^n\mathbb{Z}} \xrightarrow{/} \frac{\mathbb{Z}}{p\mathbb{Z}} \longrightarrow 0$$

  giving rise to a relation $\left[\dfrac{\mathbb{Z}}{p^n\mathbb{Z}}\right] = \left[\dfrac{\mathbb{Z}}{p^{n-1}\mathbb{Z}}\right] + \left[\dfrac{\mathbb{Z}}{p\mathbb{Z}}\right]$. By induction, $\left[\dfrac{\mathbb{Z}}{p^n\mathbb{Z}}\right] = n\left[\dfrac{\mathbb{Z}}{p\mathbb{Z}}\right]$. Therefore $\mathcal{G}(\mathsf{FinAb}) \cong \bigoplus_p \mathbb{Z}$ where p is prime. For two elements $[a], [b] \in \mathcal{G}(\mathsf{FinAb})$, $[a] \preceq [b]$ if the multiplicity of each prime factor of $[a]$ is at most the multiplicity of each prime factor of $[b]$.

- Let Ab be the category of finitely generated abelian groups. A finitely generated abelian group is isomorphic to

$$\mathbb{Z}^m \oplus \frac{\mathbb{Z}}{p_1^{n_1}\mathbb{Z}} \oplus \cdots \oplus \frac{\mathbb{Z}}{p_k^{n_k}\mathbb{Z}}$$

  where each p_i is prime. The free abelian group generated by the set of isomorphism classes in Ab is $\bigoplus_m \mathbb{Z} \oplus \bigoplus_{(p,n)} \mathbb{Z}$ over all natural numbers $m \geqslant 0$ and over all pairs (p, n) where p is prime and $n \geqslant 0$ a natural number. In addition to the short exact sequences in FinAb, we have

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\times p} \mathbb{Z} \xrightarrow{/} \frac{\mathbb{Z}}{p\mathbb{Z}} \longrightarrow 0$$

$$0 \longrightarrow \mathbb{Z}^m \longrightarrow \mathbb{Z}^{m+n} \longrightarrow \mathbb{Z}^n \longrightarrow 0$$

giving rise to the relations $\left[\dfrac{\mathbb{Z}}{p\mathbb{Z}}\right] = [0]$ and $[\mathbb{Z}^m] + [\mathbb{Z}^n] = [\mathbb{Z}^{m+n}]$. Therefore $\mathcal{G}(\mathsf{Ab}) \cong \mathbb{Z}$ where $\preceq$ is the usual total ordering on the integers. Unfortunately all torsion is lost.

Remark 2.2.29: Although the Grothendieck group can be applied to categories with infinitely generated objects, it is often times trivial due to the Eilenberg Swindle. Consider, for example, the category of (possibly infinite dimensional) vector spaces over a fixed field. For any vector space V, we have that $V \oplus \bigoplus_{i=1}^{\infty} V \cong \bigoplus_{i=1}^{\infty} V$ and so $[V] + [\bigoplus_{i=1}^{\infty} V] = [\bigoplus_{i=1}^{\infty} V]$. Therefore $[V] = 0$ and the Grothendieck group of this category is the trivial group.

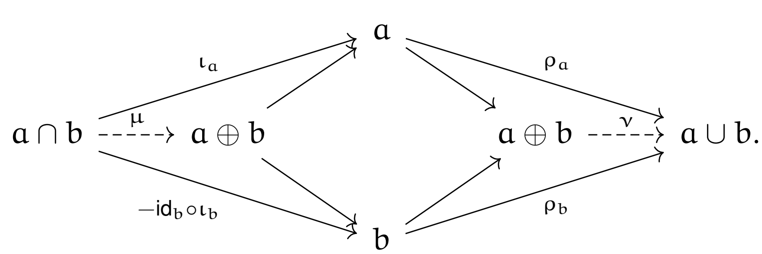
When two objects a and b in an abelian category have monomorphisms $f : a \hookrightarrow c$ and $g : b \hookrightarrow c$ into a third object c, we can take the union and intersection of a and b. The intersection is given by the limit of the diagram

$$\begin{array}{ccc} a \cap b & \dashrightarrow & a \\ \downarrow & & \downarrow f \\ b & \xhookrightarrow{g} & c. \end{array}$$

The union of a and b is given by the colimit of the diagram

$$\begin{array}{ccc} a \cap b & \longrightarrow & a \\ \downarrow & & \downarrow \\ b & \dashrightarrow & a \cup b. \end{array}$$

Any two subobjects $a \hookrightarrow c$ and $b \hookrightarrow c$ fit into a short exact sequence as follows. Consider the following diagram where the dashed arrows come from the universal properties of a biproduct:



The morphism μ is a monomorphism and ν is an epimorphism with kernel $a \cap b$ giving us the short exact sequence

$$0 \longrightarrow a \cap b \xrightarrow{\ \mu\ } a \oplus b \xrightarrow{\ \nu\ } a \cup b \longrightarrow 0.$$

For example in Vec, $\mu : x \mapsto (x, -x)$ and $\nu : (x, y) \mapsto x + y$. The map μ is injective, ν is surjective, and $\ker \mu = \mathrm{im}\, \nu = a \cap b$ making the above sequence exact. The corresponding relation $[a \cap b] + [a \cup b] = [a \oplus b]$ in $\mathcal{G}$ can be rewritten as an inclusion-exclusion formula

$$[a \cup b] = [a] + [b] - [a \cap b]. \tag{2.1}$$

2.3 Chain Complexes and Homology

Homology was developed to study topological spaces from an algebraic point of view. There are many different homology theories. Of particular interest for this dissertation is simplicial homology. Given a simplicial complex K and a fixed ring R, simplicial homology assigns to it an R-module $H_i(K, R)$ for each $i \in \mathbb{N}$ that describes the independent i-dimensional holes in K. Simplicial homology has the advantage of being easily computable with linear algebra techniques. To define homology, we first need to define chain complexes.

Definition 2.3.1: Let $\mathcal{A}$ be an abelian category. A **chain complex** X_* in $\mathcal{A}$ consists of objects X_i for all $i \in \mathbb{Z}$ and morphisms $d_i : X_i \to X_{i-1}$ satisfying $d_i \circ d_{i+1} = 0$ for all $i \in \mathbb{Z}$. A morphism $f : X_* \to Y_*$ between chain complexes consists of morphisms $f_i : X_i \to Y_i$ in $\mathcal{A}$ such that $d_i \circ f_i = f_{i-1} \circ d_i$. The **category of chain complexes in** $\mathcal{A}$ is the category $\mathsf{Ch}(\mathcal{A})$ consisting of chain complexes and chain complex morphisms.

We can associate to any chain complex X_* two types of objects: the cycle objects and the boundary objects. Intuitively, the i-dimensional cycle object of a chain complex is generated by each (possibly filled in) hole while the i-dimensional boundary object is generated by the filled in holes.

Definition 2.3.2: For any chain complex X_* its i-dimensional **cycle object** is $Z_i(X_*) := \ker d_i$ and its i-dimensional **boundary object** is $B_i(X_*) := \mathrm{im}\, d_{i+1}$.

Since $d_i \circ d_{i+1} = 0$ it must be that $B_i(X_*) \subseteq Z_i(X_*)$. This allows us to define the i-dimensional homology of X_* as the quotient of the two objects.

Definition 2.3.3: The i^{th}-**homology** of a chain complex X_* is $H_i(X_*) := \frac{Z_i(X_*)}{B_i(X_*)}$.

Example 2.3.4: Consider the 2-simplex Δ_2. It's simplicial chain complex with coefficients in $\mathbb{R}$ is

$$\cdots \longleftarrow X_{-1} = 0 \longleftarrow X_0 = \mathbb{R}^3 \longleftarrow X_1 = \mathbb{R}^3 \longleftarrow X_2 = \mathbb{R} \longleftarrow X_3 = 0 \longleftarrow \cdots$$

The homology groups of X_* are $H_0(X_*) = \mathbb{R}$ and $H_i(X_*) = 0$ for $i \neq 0$.

Chapter 3

Filtrations

Filtrations are the main objects of study in persistent homology. They arise via data from constructions such as the Rips complexes [1], Čech complexes [2], and alpha complexes [3] and in pure math from sublevel sets of functions among other sources. In this dissertation, we'll study filtrations of a fixed chain complex X_*.

Definition 3.0.1: The **category of sub-complexes of** X_* is the category $\mathsf{Sub}(X_*)$ with objects chain complexes Y_* such that there is an inclusion of chain complexes $Y_* \hookrightarrow X_*$. The morphisms in $\mathsf{Sub}(X_*)$ are inclusions of chain complexes $Y_* \hookrightarrow Z_*$.

Definition 3.0.2: Let P be a finite metric lattice and X_* a chain complex. A filtration of X_* indexed by P, or simply a P**-filtration of** X_*, is a functor $\mathsf{F} : P \to \mathsf{Sub}(X_*)$. That is, for all $a \in P$, $\mathsf{F}(a)$ is a subcomplex of X_* and for all $a \leqslant b$, $\mathsf{F}(a \leqslant b)$ is an inclusion of $\mathsf{F}(a)$ into $\mathsf{F}(b)$. Further, we require that $\mathsf{F}(\top) = X_*$.

Definition 3.0.3: A **filtration-preserving morphism** is a triple $(\mathsf{F}, \mathsf{G}, \alpha)$ where $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ and $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$ are P and Q-filtrations of X_*, respectively, and $\alpha : P \to Q$ is a bounded lattice function satisfying the following axiom: for all $a \in Q$, $\mathsf{G}(a) = \mathsf{F}(a^\star)$ where $a^\star := \max \alpha^{-1}[\bot, a]$:

$$\begin{array}{ccc} P & \xrightarrow{\alpha} & Q \\ & {\scriptstyle \mathsf{F}} \searrow \quad \swarrow {\scriptstyle \mathsf{G}} & \\ & \mathsf{Sub}(X_*). & \end{array}$$

Remark 3.0.4: A more sophisticated but an equivalent definition of a filtration-preserving morphism is the following. A filtration-preserving morphism is a triple $(\mathsf{F}, \mathsf{G}, \alpha)$ where $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ and $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$ are P and Q-filtrations of X_*, respectively, and $\alpha : P \to Q$ is a

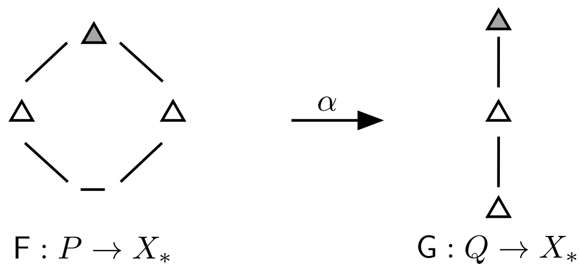
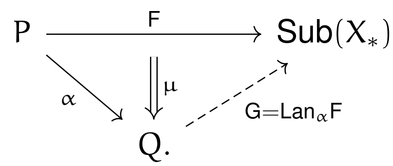


Figure 3.1: Filtrations F and G of the chain complex X_* defined in Example 2.3.4 along with a filtration-preserving morphism α as described in Example 2.1.1.

bounded lattice function such that G is the left Kan extension of F along α, written $\mathsf{G} = \mathsf{Lan}_\alpha \mathsf{F}$:



By construction of the left Kan extension,

$$\mathsf{Lan}_\alpha \mathsf{F}(a) := \mathsf{colim}_{\mathsf{Sub}(X_*)} \mathsf{F}|_{\alpha^{-1}[\bot, a]}$$

for all $a \in Q$ [20]. By Proposition 2.1.2, $\alpha^{-1}[\bot, a]$ has a maximal element $a^\star$ and therefore $\mathsf{Lan}_\alpha \mathsf{F}(a)$ is equal to $\mathsf{F}(a^\star)$. For all $a \leqslant b$ in Q, $a^\star \leqslant b^\star$ inducing the inclusion $\mathsf{Lan}_\alpha \mathsf{F}(a \leqslant b)$. The natural transformation $\mu : \mathsf{F} \Rightarrow \mathsf{G} \circ \alpha$ is gotten as follows. For $c \in P$, let $a := \alpha(c)$. Since $c \leqslant a^\star$ and $\mathsf{Lan}_\alpha \mathsf{F}(a)$ is equal to $\mathsf{F}(a^\star)$, we get the inclusion $\mu(c) : \mathsf{F}(c) \hookrightarrow \mathsf{G} \circ \alpha(c) = \mathsf{G}(a)$.

Example 3.0.5: Let $\alpha : P \to Q$ be the bounded lattice function described in Example 2.1.1. Consider the two filtrations $\mathsf{F} : P \to \Delta K$ and $\mathsf{G} : Q \to \Delta K$ of the 2-simplex K in Figure 3.1. The triple $(\mathsf{F}, \mathsf{G}, \alpha)$ is a filtration-preserving morphism $\alpha : \mathsf{F} \to \mathsf{G}$.

Proposition 3.0.6: If $(\mathsf{F}, \mathsf{G}, \alpha)$ and $(\mathsf{G}, \mathsf{H}, \beta)$ are filtration-preserving morphisms, then $(\mathsf{F}, \mathsf{H}, \beta \circ \alpha)$ is a filtration-preserving morphism.

Proof. Suppose $\mathsf{F} : P \to \mathsf{Sub}(X_*)$, $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$, and $\mathsf{H} : R \to \mathsf{Sub}(X_*)$. For all $a \in R$, $\mathsf{H}(a) = \mathsf{G}(a^\star)$ where $a^\star := \max \beta^{-1}[\bot, a]$. Furthermore, $\mathsf{G}(a^\star) = \mathsf{F}(a^{\star\star})$ where

$a^{\star\star} := \max \alpha^{-1}[\bot, a^{\star}]$. Since $a^{\star\star} = \max(\beta \circ \alpha)^{-1}[\bot, a]$, we have that $\mathsf{H}(a) = \mathsf{F}(a^{\star\star})$. Thus $(\mathsf{F}, \mathsf{H}, \beta \circ \alpha)$ is a filtration-preserving morphism. □

Definition 3.0.7: For a fixed chain complex X_* let $\mathsf{Fil}(X_*)$ be the category whose objects are P-filtrations of X_*, over all finite metric lattices P, and whose morphisms are filtration-preserving morphisms. We call $\mathsf{Fil}(X_*)$ the **category of filtrations of** X_*.

3.0.1 Edit Distance Between Filtrations

A *path* between two filtrations F and H in Fil is a finite sequence

$$\mathsf{F} \xleftrightarrow{\alpha_1} \mathsf{G}_1 \xleftrightarrow{\alpha_2} \cdots \xleftrightarrow{\alpha_3} \mathsf{G}_{n-1} \xleftrightarrow{\alpha_n} \mathsf{H}$$

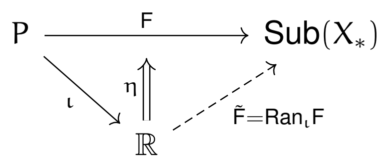
where $\leftrightarrow$ denotes a filtration-preserving morphism in either direction. The *length* of a path is the sum $\sum_{i=1}^{n} ||\alpha_i||$ of the distortions of all the bounded lattice functions. Note that the filtration $\Omega : \star \to \mathsf{Sub}(X_*)$ is terminal in Fil. This implies that any two filtrations in Fil are connected by a path.

Definition 3.0.8: The **edit distance** $d_{\mathsf{Fil}}(\mathsf{F}, \mathsf{H})$ between any two filtrations in Fil is the length of the shortest path between F and H.

3.1 Totally Ordered Filtrations

In the special case where the finite metric lattice P is totally ordered, the additional structure allows for additional constructions. Here, P can be embedded in $\mathbb{R}$ allowing us to define the *interleaving distance* and interpolate between filtrations.

Given a totally-ordered, finite metric lattice P, we embed P into $\mathbb{R}$ by setting $\bot_P := 0$ and the remaining elements are forced by the metric (i.e. if $\bot_P < a$ and $d_P(\bot_P, a) = r$ then $a := r$). In this way we view P as a subset of $\mathbb{R}$. We can then extend a filtration $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ to a functor $\tilde{\mathsf{F}} : \mathbb{R} \to \mathsf{Sub}(X_*)$ by defining $\tilde{\mathsf{F}}(x) := \max_{\substack{a \in P \\ a \leqslant x}} \mathsf{F}(a)$. This extension turns out to be the right Kan extension of F along the embedding $\iota : P \hookrightarrow \mathbb{R}$.



While, in general, functors from $\mathbb{R} \to \mathsf{Sub}(X_*)$ can be quite unwieldy, the functors that arise in this way satisfy a finiteness condition: they only change isomorphism type in finitely many places. We call these functors constructible.

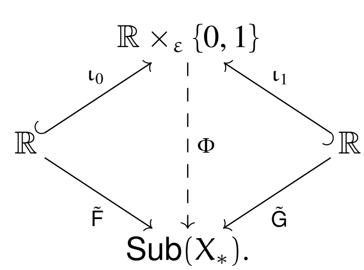
Definition 3.1.1: A **constructible filtration** is a functor $\tilde{\mathsf{F}} : \mathbb{R} \to \mathsf{Sub}(X_*)$ such that there exists a finite, totally-ordered metric lattice P and a filtration $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ with F extended to $\mathbb{R}$ equal to $\tilde{\mathsf{F}}$.

There is a natural distance between constructible filtrations called the *interleaving distance* [21]. For any $\varepsilon \geqslant 0$, let $\mathbb{R} \times_\varepsilon \{0,1\}$ be the poset $\big(\mathbb{R} \times \{0\}\big) \cup \big(\mathbb{R} \times \{1\}\big)$ where $(p,t) \leqslant (q,s)$ if

- $t = s$ and $p \leqslant q$, or
- $t \neq s$ and $p + \varepsilon \leqslant q$.

Let $\iota_0, \iota_1 : \mathbb{R} \hookrightarrow \mathbb{R} \times_\varepsilon \{0,1\}$ be the poset maps $\iota_0 : p \mapsto (p, 0)$ and $\iota_1 : p \mapsto (p, 1)$.

Definition 3.1.2: An ε**-interleaving** between two constructible filtrations $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ is a functor Φ that makes the following diagram commute up to a natural isomorphism:



(3.1)

Two constructible filtrations $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ are ε**-interleaved** if there is an ε-interleaving between them. The **interleaving distance** $d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$ between $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ is the infimum over all $\varepsilon \geqslant 0$ such that $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ are ε-interleaved. This infimum is attained since both $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ are constructible. If $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ are not interleaved, then we let $d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}}) = \infty$.

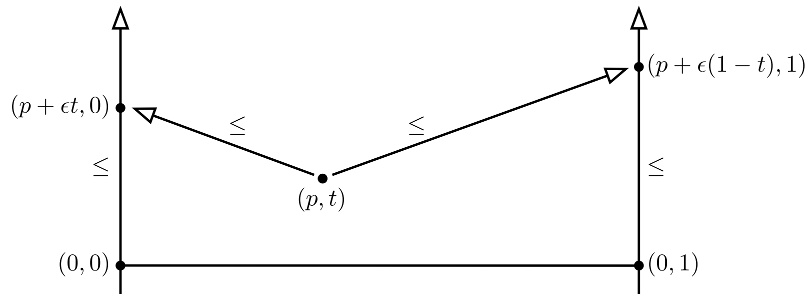


Figure 3.2: An illustration of the poset relation on $\mathbb{R} \times_\varepsilon [0, 1]$.

Proposition 3.1.3 (Interpolation [22])**:** Let $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ be two ε-interleaved constructible filtrations. Then there exists a one-parameter family of constructible filtrations $\{\tilde{\mathsf{K}}_t\}_{t\in[0,1]}$ such that $\tilde{\mathsf{K}}_0 \cong \tilde{\mathsf{F}}$, $\tilde{\mathsf{K}}_1 \cong \tilde{\mathsf{G}}$, and $d_I(\tilde{\mathsf{K}}_t, \tilde{\mathsf{K}}_s) \leqslant \varepsilon|t - s|$.

Proof. Let $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ be ε-interleaved by Φ as in Definition 3.1.2. Define $\mathbb{R} \times_\varepsilon [0, 1]$ as the poset with the underlying set $\mathbb{R} \times [0, 1]$ and $(p, t) \leqslant (q, s)$ whenever $p + \varepsilon|t - s| \leqslant q$. Note that $\mathbb{R} \times_\varepsilon \{0, 1\}$ naturally embeds into $\mathbb{R} \times_\varepsilon [0, 1]$ via $\iota : (p, t) \mapsto (p, t)$. See Figure 3.2. Finding $\{\tilde{\mathsf{K}}_t\}_{t\in[0,1]}$ is equivalent to finding a functor Ψ that makes the following diagram commute up to a natural isomorphism:

$$\begin{array}{ccc} \mathbb{R} \times_\varepsilon \{0, 1\} & \xrightarrow{\ \Phi\ } & \mathsf{C} \\ {\scriptstyle\iota}\big\downarrow & \nearrow_{\Psi} & \\ \mathbb{R} \times_\varepsilon [0, 1]. & & \end{array}$$

This functor Ψ is the right Kan extension of Φ along ι for which we now give an explicit construction. For convenience, let $S := \mathbb{R} \times_\varepsilon \{0, 1\}$ and $\mathsf{T} := \mathbb{R} \times_\varepsilon [0, 1]$. For $(p, t) \in \mathsf{T}$, let $S \uparrow (p, t)$ be the subposet of S consisting of all elements $(p', t') \in S$ such that $(p, t) \leqslant (p', t')$. The poset $S \uparrow (p, t)$, for any $p \in \mathbb{R}$ and $t \notin \{0, 1\}$, has two minimal elements: $(p + \varepsilon t, 0)$ and $\big(p + \varepsilon(1 - t), 1\big)$. For $t \in \{0, 1\}$, the poset $S \uparrow (p, t)$ has one minimal element, namely (p, t). Let $\Psi\big((p, t)\big) := \lim \Phi|_{S\uparrow(p,t)}$. For $(p, t) \leqslant (q, s)$, the poset $S \uparrow (q, s)$ is a subposet of $S \uparrow (p, t)$. This subposet relation allows us to define the morphism $\Psi\big((p, t) \leqslant (q, s)\big)$ as the universal morphism between the two limits. Note that $\Psi\big((p, 0)\big)$ is isomorphic to $\tilde{\mathsf{F}}(p)$ and $\Psi\big((p, 1)\big)$ is isomorphic to $\tilde{\mathsf{G}}(p)$.

We now argue that each functor $\tilde{\mathsf{K}}_t := \Psi(\cdot, t)$ is constructible. Suppose F is a P-filtration and G is a Q-filtration. As we increase p while keeping t fixed, the limit $\tilde{\mathsf{K}}_t(p)$ changes only when one of the two minimal objects of $\mathsf{P} \uparrow (p, t)$ changes isomorphism type. This occurs when p is in $P - \varepsilon t := \{x - \varepsilon t \,|\, x \in P\}$ or $Q - \varepsilon(1 - t)$. Therefore $\tilde{\mathsf{K}}_t$ is a constructible filtration. □

Chapter 4

Persistence Diagrams

The main invariant in persistent homology is the persistence diagram. Given a P-filtration $\mathsf{F} : P \to \mathsf{Sub}(X_*)$, its i-dimensional persistence diagram is a map $\partial\mathsf{ZB}_i\mathsf{F} : \bar{P} \to \mathbb{Z}$ that, loosely speaking, assigns to any interval $I = [i_1, i_2]$ in $\bar{P}$ the number of independent i-cycles that were born at i_1 and died at i_2. In order to define persistence diagrams, we first need to define an intermediary invariant: the birth-death function. This function summarizes where cycles were born and die in a filtration. Both birth-death functions and persistence diagrams naturally lie within categories. The former lies in the category of monotone $\mathcal{G}$-functions while the latter lies in the category of $\mathcal{G}$-functions. The categories of filtrations and monotone $\mathcal{G}$-functions are linked by the birth-death functor, taking a filtration to its birth-death function. Similarly, the Möbius inversion is a functor from the category of monotone $\mathcal{G}$-functions to the category of $\mathcal{G}$-functions. Taken together, these functors form a functorial pipeline with input a filtration and output a persistence diagram.

4.1 Monotone $\mathcal{G}$-Functions

We now define the category of monotone $\mathcal{G}$-functions over finite metric lattices $\mathsf{Mon}(\mathcal{G})$ and construct the i^{th} birth-death functor $\mathsf{ZB}_i : \mathsf{Fil}(X_*) \to \mathsf{Mon}(\mathcal{G})$. Recall that all of our filtrations are of a fixed chain complex X_* in an abelian category $\mathcal{A}$ with Grothendieck group $\mathcal{G}$.

Definition 4.1.1: Let P and Q be two finite metric lattices and let $f : P \to \mathcal{G}$ and $g : Q \to \mathcal{G}$ be two monotone $\mathcal{G}$-functions. A **monotone-preserving morphism** from f to g is a triple $(f, g, \bar{\alpha})$ where $\bar{\alpha} : \bar{P} \to \bar{Q}$ is a bounded lattice function satisfying the following axiom. For all $I \in \bar{Q}$ and $I^\star := \max \bar{\alpha}^{-1}[\bot, I]$, $g(I) = f(I^\star)$:

$$\begin{array}{ccc} \bar{P} & \xrightarrow{\quad\bar{\alpha}\quad} & \bar{Q} \\ & {\scriptstyle f}\searrow \qquad \swarrow{\scriptstyle g} & \\ & \mathcal{G}. & \end{array}$$

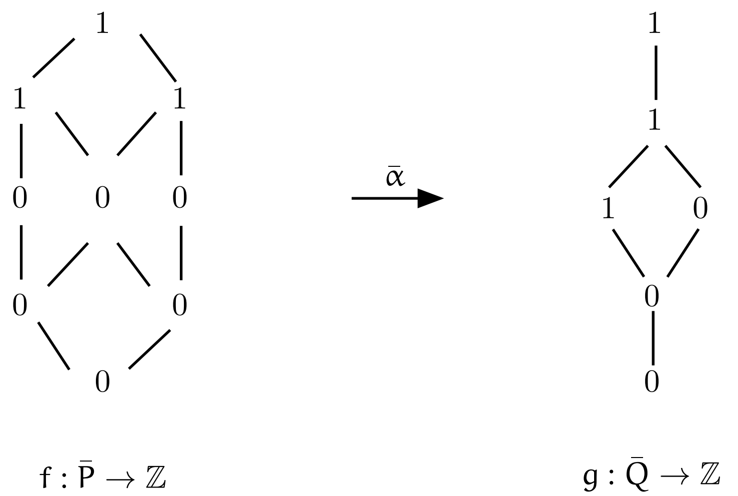
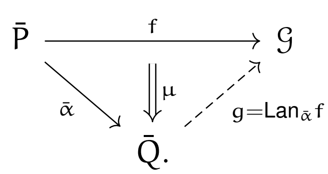


Figure 4.1: Shown are two monotone $\mathbb{Z}$ functions f and g on the metric lattices $\bar{P}$ and $\bar{Q}$ from Example 2.1.3. The triple $(f, g, \bar{\alpha})$, where $\bar{\alpha} : \bar{P} \to \bar{Q}$ is from the same example, is a monotone-preserving morphism from f to g.

Here we require that $\bar{\alpha} : \bar{P} \to \bar{Q}$ be induced by some $\alpha : P \to Q$.

Note that if $(f, g, \bar{\alpha})$ is a monotone-preserving morphism, then $f[\top, \top] = g[\top, \top]$.

Remark 4.1.2: A more sophisticated but an equivalent definition of a monotone-preserving morphism is the following. A **monotone-preserving morphism** is a triple $(f, g, \bar{\alpha})$ where $f : \bar{P} \to \mathcal{G}$ and $g : \bar{Q} \to \mathcal{G}$ are monotone functions and $\bar{\alpha} : \bar{P} \to \bar{Q}$ is a bounded lattice function such that g is the left Kan extension of f along α, written $g = \mathsf{Lan}_{\alpha} f$:



Example 4.1.3: See Figure 4.1 for an example of monotone integral functions f and g on the lattices $\bar{P}$ and $\bar{Q}$, respectively, from Example 2.1.3. The triple $(f, g, \bar{\alpha})$, where $\bar{\alpha} : \bar{P} \to \bar{Q}$ is from the same example, is a monotone-preserving morphism.

Proposition 4.1.4: If $(f, g, \bar{\alpha})$ and $(g, h, \bar{\beta})$ are monotone-preserving morphisms, then $(f, h, \bar{\beta} \circ \bar{\alpha})$ is a monotone-preserving morphism.

Proof. Suppose $f : \bar{P} \to \mathcal{G}$, $g : \bar{Q} \to \mathcal{G}$, and $h : \bar{R} \to \mathcal{G}$. For all $I \in \bar{R}$, $h(I) = g(I^\star)$ where $I^\star := \max \bar{\beta}^{-1}[\bot, I]$. Furthermore, $g(I^\star) = f(I^{\star\star})$ where $I^{\star\star} := \max \bar{\alpha}^{-1}[\bot, I^\star]$. Since $I^{\star\star} = \max(\bar{\beta} \circ \bar{\alpha})^{-1}[\bot, I]$, we have that $h(I) = f(I^{\star\star})$. Thus the composition $(f, h, \bar{\beta} \circ \bar{\alpha})$ is a monotone-preserving morphism. □

Definition 4.1.5: Let $\mathsf{Mon}(\mathcal{G})$ be the category consisting of monotone $\mathcal{G}$-functions $f : \bar{P} \to \mathcal{G}$, over all finite metric lattices P, and monotone-preserving morphisms. We call $\mathsf{Mon}(\mathcal{G})$ the **category of monotone functions**.

4.1.1 Birth-Death Functor

We now construct the birth-death functors. Given a filtration $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ its i-th birth-death function captures information about where i-dimensional cycles are born and die in F.

Definition 4.1.6: Let $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ be an object of $\mathsf{Fil}(X_*)$. For every interval $[a, b] \in \bar{P}$, where $b \neq \top$, let

$$\mathsf{ZB}_i\mathsf{F}[a, b] := \big[\mathsf{Z}_i\mathsf{F}(a) \cap \mathsf{B}_i\mathsf{F}(b)\big]$$

where the intersection is taken inside $\mathsf{Z}_i\mathsf{F}(\top)$. For all other intervals $[a, \top]$, let

$$\mathsf{ZB}_i\mathsf{F}[a, \top] := \big[\mathsf{Z}_i\mathsf{F}(a)\big].$$

The **i-th birth-death function** of F is the function $f_i : \bar{P} \to \mathcal{G}$ that assigns to every interval $[a, b]$ the value $\mathsf{ZB}_i\mathsf{F}[a, b]$.

The reason we force $\mathsf{ZB}_i\mathsf{F}[a, \top]$ to $\big[\mathsf{Z}_i\mathsf{F}(a)\big]$ instead of $\big[\mathsf{Z}_i\mathsf{F}(a) \cap \mathsf{B}_i\mathsf{F}(\top)\big]$ is because we want all cycles to be dead by $\top$. Otherwise, the persistence diagram for F (see Definition **??**) would not see cycles that are born and never die.

Proposition 4.1.7: Let $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ be an object in $\mathsf{Fil}(X_*)$ and $f_i : \bar{P} \to \mathcal{G}$ its i-th birth-death function. Then f_i is monotone.

Proof. For any two intervals $I \preceq J$ in $\bar{P}$, we must show that $f_i(I) \leqslant f_i(J)$. Suppose $I = [a, b]$ and $J = [c, d]$ and $d \neq \top$. Then $\mathsf{Z}_i\mathsf{F}(a) \subseteq \mathsf{Z}_i\mathsf{F}(c)$ and $\mathsf{B}_i\mathsf{F}(b) \subseteq \mathsf{B}_i\mathsf{F}(d)$. Thus $\mathsf{Z}_i\mathsf{F}(a) \cap \mathsf{B}_i\mathsf{F}(b)$ is a subspace of $\mathsf{Z}_i\mathsf{F}(c) \cap \mathsf{B}_i\mathsf{F}(d)$, and therefore $\mathsf{ZB}_i\mathsf{F}[a, b] \leqslant \mathsf{ZB}_i\mathsf{F}[c, d]$. For $J = [c, \top]$, $\mathsf{ZB}_i\mathsf{F}[a, b] \subseteq \mathsf{Z}_i\mathsf{F}(c)$, and therefore $\mathsf{ZB}_i\mathsf{F}[a, b] \leqslant \mathsf{ZB}_i\mathsf{F}[c, \top]$. □

Proposition 4.1.8: Let $(\mathsf{F}, \mathsf{G}, \alpha)$ be a morphism in $\mathsf{Fil}(X_*)$ and f_i and g_i the i-th birth-death functions of F and G, respecively. Then $(f_i, g_i, \bar{\alpha})$ is a morphism in $\mathsf{Mon}(\mathcal{G})$.

Proof. Suppose $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ and $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$. By definition of morphism in $\mathsf{Fil}(X_*)$, $\mathsf{G}(a) \cong \mathsf{F}(a^\star)$, for all $a \in Q$, where $a^\star = \max \bar{\alpha}^{-1}[\perp, a]$. For all intervals $I \in \bar{Q}$, let $I^\star := \max \bar{\alpha}^{-1}[\perp, I]$. If $I = [a, b]$, then $I^\star = [a^\star, b^\star]$ where $b^\star = \max \bar{\alpha}^{-1}[\perp, b]$. The definition of a filtration-preserving morphism implies the following canonical isomorphisms of chain complexes:

$$\mathsf{G}(b) \cong \mathsf{F}(b^\star) \qquad \mathsf{G}(\top) \cong \mathsf{F}(\top) \qquad \mathsf{G}(a) \cong \mathsf{F}(a^\star),$$

which, in turn, implies canonical isomorphisms $\mathsf{Z}_i\mathsf{G}(a) \cong \mathsf{Z}_i\mathsf{F}(a^\star)$ and $\mathsf{B}_i\mathsf{G}(b) \cong \mathsf{B}_i\mathsf{F}(b^\star)$. We have $\mathsf{ZB}_i\mathsf{G}(I) = \mathsf{ZB}_i\mathsf{F}(I^\star)$ and therefore $g_i(I) = f_i(I^\star)$. □

By Propositions 4.1.4, 4.1.7 and 4.1.8, the assignment to each object in $\mathsf{Fil}(X_*)$ its birth-death monotone function is functorial.

Definition 4.1.9: Let $\mathsf{ZB}_i : \mathsf{Fil}(X_*) \to \mathsf{Mon}(\mathcal{G})$ be the functor that assigns to every filtration its i-th birth-death function and to every filtration-preserving morphism the induced monotone-preserving morphism. We call ZB_i the **i-th birth-death functor**.

Example 4.1.10: The functor ZB_1 applied to the filtration-preserving morphism $(\mathsf{F}, \mathsf{G}, \alpha)$ in Example 2.1.3 is the monotone-preserving morphism $(f, g, \bar{\alpha})$ in Example 4.1.3.

4.2 $\mathcal{G}$-Functions

We now define the category of $\mathcal{G}$-functions over finite metric lattices $\mathsf{Fnc}(\mathcal{G})$ and construct the Möbius inversion functor $\mathsf{MI} : \mathsf{Mon}(\mathcal{G}) \to \mathsf{Fnc}(\mathcal{G})$.

Definition 4.2.1: Let P and Q be finite metric lattices and let $\sigma : \bar{P} \to \mathcal{G}$ and $\tau : \bar{Q} \to \mathcal{G}$ be two functions not necessarily monotone. A **charge-preserving morphism** is a triple $(\sigma, \tau, \bar{\alpha})$ where $\sigma : \bar{P} \to \mathcal{G}$ and $\tau : \bar{Q} \to \mathcal{G}$ are functions and $\bar{\alpha} : \bar{P} \to \bar{Q}$ is a bounded lattice function satisfying the following axiom. For all $I \in \bar{Q}$ with $I \neq [q, q]$,

$$\tau(I) = \sum_{J \in \bar{\alpha}^{-1}(I)} \sigma(J). \tag{4.1}$$

If $\bar{\alpha}^{-1}(I)$ is empty, then we interpret the sum as 0. Here we require that $\bar{\alpha}$ be induced by some $\alpha : P \to Q$.

Remark 4.2.2: Our definition of a charge-preserving morphism is related to the definition of a morphism between signed measures. Let (X, Σ_X) and (Y, Σ_Y) be measurable spaces, $\phi : (X, \Sigma_X) \to (Y, \Sigma_Y)$ a measurable map, and $\mu : \Sigma_X \to \mathbb{R}$ a signed measure. Then the pushforward of μ along f is the signed measure $\phi_\# \mu : \Sigma_Y \to \mathbb{R}$ defined as $\phi_\# \mu(U) := \mu\big(\phi^{-1}(U)\big)$. In the category of signed measures, a morphism from (X, Σ_X, μ) to (Y, Σ_Y, ν) is a measurable map $\mu : \Sigma_X \to \Sigma_Y$ such that $\phi_\# \mu = \nu$ [23].

Example 4.2.3: See Figure 4.2 for integral functions σ and τ on the lattices of intervals $\bar{P}$ and $\bar{Q}$, respectively, from Example 2.1.3. The triple $(\sigma, \tau, \bar{\alpha})$, where $\bar{\alpha} : \bar{P} \to \bar{Q}$ is from the same example, is a charge-preserving morphism.

Proposition 4.2.4: If $(\sigma, \tau, \bar{\alpha})$ and $(\tau, \upsilon, \bar{\beta})$ are charge-preserving morphisms, then $(\sigma, \upsilon, \bar{\beta} \circ \bar{\alpha})$ is a charge-preserving morphism.

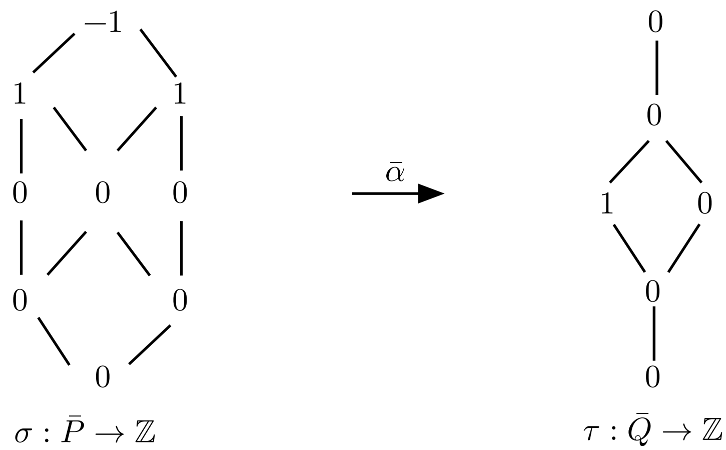


Figure 4.2: Shown are two $\mathbb{Z}$ functions σ and τ on $\bar{P}$ and $\bar{Q}$, respectively, from Example 2.1.3. The bounded lattice function $\bar{\alpha} : \bar{P} \to \bar{Q}$ from the same figure is a charge-preserving morphism from σ to τ.

Proof. Suppose $\sigma : \bar{P} \to \mathcal{G}$, $\tau : \bar{Q} \to \mathcal{G}$, and $\upsilon : \bar{R} \to \mathcal{G}$. For all $I \in \bar{R}$,

$$\upsilon(I) = \sum_{J \in \bar{\beta}^{-1}(I)} \tau(J) = \sum_{J \in \bar{\beta}^{-1}(I)} \sum_{K \in \bar{\beta}^{-1}(J)} \sigma(K) = \sum_{K \in (\bar{\beta} \circ \bar{\alpha})^{-1}(I)} \sigma(K).$$

□

Definition 4.2.5: Let $\mathsf{Fnc}(\mathcal{G})$ be the category of whose objects are functions $\sigma : \bar{P} \to \mathcal{G}$, over all finite metric lattices P, and charge-preserving morphisms. We call $\mathsf{Fnc}(\mathcal{G})$ the **category of $\mathcal{G}$-functions**.

4.2.1 Möbius Inversion Functor

Given any monotone function $f : \bar{P} \to \mathcal{G}$ of $\mathsf{Mon}(\mathcal{G})$, there is a unique function $\sigma : \bar{P} \to \mathcal{G}$ such that

$$f(J) = \sum_{I \in \bar{P} : I \preceq J} \sigma(I) \tag{4.2}$$

for all $J \in \bar{P}$ [24, 25]. The function σ is the called the *Möbius inversion* of f.

Proposition 4.2.6: Let $(f, g, \bar{\alpha})$ be a morphism in $\mathsf{Mon}(\mathcal{G})$, and let σ and τ be the Möbius inversions of f and g, respectively. Then $(\sigma, \tau, \bar{\alpha})$ is a morphism in $\mathsf{Fnc}(\mathcal{G})$.

Proof. Suppose $f : \bar{P} \to \mathcal{G}$ and $g : \bar{Q} \to \mathcal{G}$. We show that

$$\tau(J) = \sum_{K \in \bar{\alpha}^{-1}(J)} \sigma(K)$$

for *all* $J \in \bar{Q}$, and thus $(\sigma, \tau, \bar{\alpha})$ is a charge-preserving morphism. The proof is by induction on the finite metric lattice $\bar{Q}$. By Proposition 2.1.2, the pre-image $\bar{\alpha}^{-1}[\bot, J]$ has a unique maximal element $J^\star$, and $f(J^\star) = g(J)$ by definition of a morphism in $\mathsf{Mon}(\mathcal{G})$.

For the base case, suppose $J = \bot$. Then by Equation (4.2), $g(J) = \tau(J)$. By definition of a morphism in $\mathsf{Mon}(\mathcal{G})$, $g(J) = f(J^\star)$. By Equation (4.2),

$$f(J^\star) = \sum_{K \preceq J^*} \sigma(K) = \sum_{K \in \bar{\alpha}^{-1}(J)} \sigma(K)$$

thus proving the base case.

For the inductive step, suppose $\tau(I) = \sum_{K \in \bar{\alpha}^{-1}(I)} \sigma(K)$, for all $I \prec J$. Then

$$\begin{aligned}
\tau(J) &= \sum_{I \in \bar{Q} : I \preceq J} \tau(I) - \sum_{I \in \bar{Q} : I \prec J} \tau(I) \\
&= g(J) - \sum_{I \in \bar{Q} : I \prec J} \tau(I) && \text{by Equation (4.2)} \\
&= g(J) - \sum_{I \in \bar{Q} : I \prec J} \sum_{K \in \bar{\alpha}^{-1}(I)} \sigma(K) && \text{by Inductive Hypothesis} \\
&= f(J^\star) - \sum_{K \in \bar{P} : \bar{\alpha}(K) \prec J} \sigma(K) \\
&= \sum_{K \in \bar{P} : K \preceq J^\star} \sigma(K) - \sum_{K \in \bar{P} : \bar{\alpha}(K) \prec J} \sigma(K) && \text{by Equation (4.2)} \\
&= \sum_{K \in \bar{P} : \bar{\alpha}(K) \preceq J} \sigma(K) - \sum_{K \in \bar{P} : \bar{\alpha}(K) \prec J} \sigma(K) \\
&= \sum_{K \in \bar{P} : \bar{\alpha}(K) = J} \sigma(K) = \sum_{K \in \bar{\alpha}^{-1}(J)} \sigma(K).
\end{aligned}$$

□

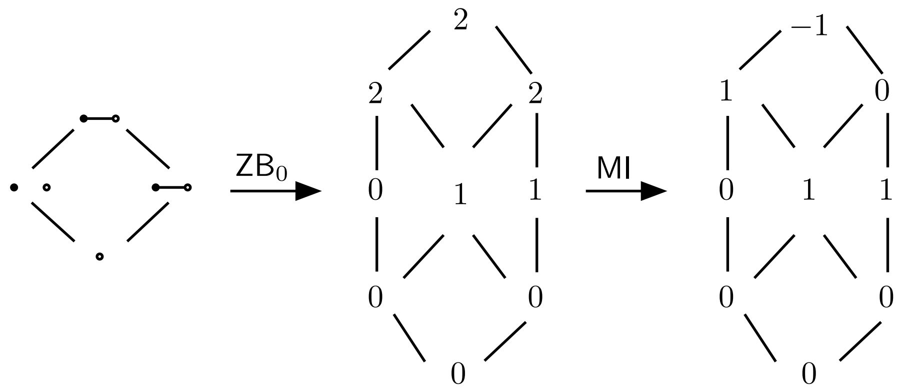


Figure 4.3: Above is a filtration of the 1-simplex, its zeroth birth-death function, and its zeroth persistence diagram.

By Propositions 4.2.4 and 4.2.6, the assignment to every object in $\mathsf{Mon}(\mathcal{G})$ its Möbius inversion is functorial.

Definition 4.2.7: Let $\mathsf{MI} : \mathsf{Mon}(\mathcal{G}) \to \mathsf{Fnc}(\mathcal{G})$ be the functor that assigns to every monotone function its Möbius inversion and to every monotone-preserving morphism the induced charge-preserving morphism. We call MI the **Möbius inversion functor**.

Definition 4.2.8: The i^{th} **persistence diagram** of a filtration $\mathsf{F} : \mathsf{P} \to \mathsf{Sub}(X_*)$ is the Möbius inversion of its i^{th} birth-death function $\mathsf{MI}(\mathsf{ZB}_i\mathsf{F})$.

Example 4.2.9: The functor MI applied to the monotone-preserving morphism $(f, g, \bar{\alpha})$ in Example 4.1.3 is the charge-preserving morphism $(\sigma, \tau, \bar{\alpha})$ in Example 4.2.3.

Example 4.2.10: Consider the example of a filtration $\mathsf{F} : \mathsf{P} \to \Delta K$ in Figure 4.3 where P is the lattice from Example 2.1.1 and K is the 1-simplex. Recall $\bar{\mathsf{P}}$ in Example 2.1.3. Drawn are its zeroth birth-death function $f := \mathsf{ZB}_0 \circ \mathsf{F} : \bar{\mathsf{P}} \to \mathbb{Z}$ and its zeroth persistence diagram $\sigma := \mathsf{MI} \circ \mathsf{ZB}_0 : \bar{\mathsf{P}} \to \mathbb{Z}$. The integer $\sigma[a, d] = 1$ represents the 0-cycle that is born at a, and the integer $\sigma[b, d] = 1$ represents the 0-cycle that is born at b. The integer $\sigma[c, c] = 1$ represents the 0-cycle that is born at c and dies immediately. The integer $\sigma[d, d] = -1$ represents the 0-cycle that was born twice but contributes to just one dimension of the total cycle space.

4.2.2 Edit Distance Between $\mathcal{G}$ Functions

A *path* between two $\mathcal{G}$-functions σ and τ in $\mathsf{Fnc}(\mathcal{G})$ is a finite sequence

$$\sigma \xleftrightarrow{\bar{\alpha}_1} \theta_1 \xleftrightarrow{\bar{\alpha}_2} \cdots \xleftrightarrow{\bar{\alpha}_{n-1}} \theta_{n-1} \xleftrightarrow{\bar{\alpha}_n} \tau$$

where $\leftrightarrow$ denotes a charge-preserving morphism in either direction. The *length* of path is the sum $\sum_{i=1}^{n} ||\bar{\alpha}_i||$ of the distortions of all the bounded lattice functions. Note that any $\mathcal{G}$ function $\omega : \bar{\star} \to \mathbb{Z}$ is terminal in $\mathsf{Fnc}(\mathcal{G})$; see Definition 4.2.1. This means that any two $\mathcal{G}$-functions in $\mathsf{Fnc}(\mathcal{G})$ are connected by a path.

Definition 4.2.11: Define the distance $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau)$ between any two $\mathcal{G}$-functions in $\mathsf{Fnc}(\mathcal{G})$ as the length of the shortest path between σ and τ.

4.3 Persistence Diagrams of Totally Ordered Filtrations

In the case where P is a finite, totally ordered metric lattice embedded in $\mathbb{R}$ as in Section 3.1, we can extend a monotone function $f : \bar{P} \to \mathcal{G}$ to a function $\tilde{f} : \bar{\mathbb{R}} \to \mathcal{G}$ by defining

$$\tilde{f}(I) := \max_{J \in \bar{P} \,:\, J \preceq I} f(J).$$

This allows us to extend birth-death functions from functions defined on $\bar{P}$ to functions defined on $\bar{R}$. Functions obtained in this way are called *constructible*.

Definition 4.3.1: A monotone function $\tilde{f} : \bar{\mathbb{R}} \to \mathcal{G}$ is P**-constructible** if there is a monotone function $f : \bar{P} \to \mathcal{G}$ with f extended to $\bar{\mathbb{R}}$ equal to $\tilde{f}$. We say that $\tilde{f}$ is constructible if there exists a finite, totally ordered metric lattice P such that $\tilde{f}$ is P-constructible.

Proposition 4.3.2: If $\tilde{f} : \bar{\mathbb{R}} \to \mathcal{G}$ is a $P = \{p_0 < \cdots < p_n\}$-constructible monotone function then the function $\tilde{\sigma} : \bar{\mathbb{R}} \to \mathcal{G}$ defined by

$$\tilde{\sigma}(a, b) := \begin{cases} \tilde{f}(p_i, p_j) - \tilde{f}(p_{i-1}, p_j) - \tilde{f}(p_i, p_{j-1}) + \tilde{f}(p_{i-1}, p_{j-1}) & \text{if } [a, b] = [p_i, p_j] \in \bar{P} \\ [0] & \text{otherwise} \end{cases}$$

satisfies the Möbius inversion formula. We interpret terms in this formula with indices out of bounds as 0 $\big($i.e. $\sigma(p_0, p_1)$ would be $\tilde{f}(p_0, p_1) - \tilde{f}(p_0, p_0)\big)$.

Proof. We need to show that for any $J \in \bar{\mathbb{R}}$,

$$\tilde{f}(J) = \sum_{I \in \bar{\mathbb{R}} : I \preceq J} \tilde{\sigma}(I).$$

We can assume without loss of generality that $J \in \bar{P}$ since otherwise we have that for $K = \max_{I \in \bar{P} : I \preceq J} I$, $\tilde{f}(J) = \tilde{f}(K)$ and

$$\sum_{I \in \bar{\mathbb{R}} : I \preceq J} \tilde{\sigma}(I) = \sum_{I \in \bar{\mathbb{R}} : I \preceq K} \tilde{\sigma}(I).$$

Let $J = [p_i, p_j] \in \bar{P}$. Then

$$\begin{aligned} \sum_{I \in \bar{\mathbb{R}} : I \preceq J} \tilde{\sigma}(I) &= \sum_{I \in \bar{P} : I \preceq J} \tilde{\sigma}(I) \\ &= \sum_{t \in \{0,1,\ldots,i\}, s \in \{0,1,\ldots,j\}} \tilde{f}(p_t, p_s) - \tilde{f}(p_{t-1}, p_s) - \tilde{f}(p_t, p_{s-1}) + \tilde{f}(p_{t-1}, p_{s-1}) \\ &= \tilde{f}(p_i, p_j). \end{aligned}$$

□

Remark 4.3.3: Given a filtration $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ over a totally ordered lattice P, its persistence diagram as defined in Definition 4.2.8 is a function $\sigma : \bar{P} \to \mathcal{G}$. On the other hand, embedding

P into $\mathbb{R}$ and computing the persistence diagram of $\tilde{F}$ yields a function from $\bar{\mathbb{R}}$ to $\mathcal{G}$. By Proposition 4.3.2, the support of the latter function lies inside $\bar{P}$ and the two functions agree when restricted to $\bar{P}$. For this reason, we do not distinguish between the two functions.

With the total ordering of P we are able to prove that persistence diagrams are positive. This is not necessarily the case if P is not totally ordered; see Example 4.3. This proposition plays an important role in the proof of bottleneck stability.

Proposition 4.3.4 (Positivity)**:** Let $\tilde{F} : \mathbb{R} \to \mathsf{Sub}(X_*)$ be a P-constructible filtration with i^{th} persistence diagram σ. Then for any $I \in \bar{\mathbb{R}}$ we have that $\sigma(I) \geqslant [0]$.

Proof. If $I \notin \bar{P}$ then $\sigma(I) = [0]$ so assume that $I = [p_t, p_s] \in \bar{P}$. Then

$$\begin{aligned}\sigma(I) &= \mathsf{ZB}_i(p_t, p_s) - \mathsf{ZB}_i(p_{t-1}, p_s) - \mathsf{ZB}_i(p_t, p_{s-1}) + \mathsf{ZB}_i(p_{t-1}, p_{s-1}) \\ &= \big[\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_s)\big] - \big[\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s)\big] \\ &\quad - \big[\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1})\big] + \big[\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_{s-1})\big]\end{aligned}$$

Observe that $\Big(\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s)\Big) \cap \Big(\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1})\Big) = \mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_{s-1})$ so by Equation 2.1

$$\begin{aligned}\big[\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s)\big] + \big[\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1})\big] - \big[\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_{s-1})\big] \\ = \big[\big(\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s)\big) \cup \big(\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1})\big)\big]\end{aligned}$$

and so $\sigma(I) = \big[\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_s)\big] - \big[\big(\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s)\big) \cup \big(\mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1})\big)\big]$. Since

$$\begin{aligned}\mathsf{Z}_i(p_{t-1}) \cap \mathsf{B}_i(p_s) &\hookrightarrow \mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_s) \\ \mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_{s-1}) &\hookrightarrow \mathsf{Z}_i(p_t) \cap \mathsf{B}_i(p_s)\end{aligned}$$

then their union is also a subobject

$$\big(Z_i(p_{t-1}) \cap B_i(p_s)\big) \cup \big(Z_i(p_t) \cap B_i(p_{s-1})\big) \hookrightarrow Z_i(p_t) \cap B_i(p_s).$$

Therefore $\sigma(I) \geqslant [0]$. $\square$

Chapter 5

Stability

In this section we prove our main results: bottleneck stability, edit distance stability, and a proof that the bottleneck distance and edit distance are strongly equivalent in the setting of totally ordered filtrations. Bottleneck stability is a crucial result for applications, implying that persistence diagrams are stable to noise. The bottleneck distance only makes sense for persistence diagrams over totally ordered lattices. For this reason, we introduced the edit distance and prove that persistence diagrams are stable with respect to the edit distance. Finally, we show that the edit distance and bottleneck distance are strongly equivalent for persistence diagrams over totally ordered lattices. This result justifies the edit distance as an extension of the bottleneck distance.

5.1 Edit Distance Stability

The stability of persistence diagrams with respect to the edit distances follows from the functoriality of the pipeline taking a filtration to its persistence diagram.

Theorem 5.1.1: Let $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ and $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$ be filtrations with i-th birth-death functions f and g respectively and i-th persistence diagrams σ and τ. Then $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau) \leqslant d_{\mathsf{Fil}}(\mathsf{F}, \mathsf{G})$.

Proof. Let

$$\mathsf{F} \xleftrightarrow{\alpha_1} \mathsf{H}_1 \xleftrightarrow{\alpha_2} \cdots \xleftrightarrow{\alpha_{n-1}} \mathsf{H}_{n-1} \xleftrightarrow{\alpha_n} \mathsf{G}$$

be a path of length ε between F and G. Let υ_k be the i-th persistence diagram of H_k. Since ZB_i and MI are functors,

$$\sigma \xleftrightarrow{\bar{\alpha}_1} \upsilon_1 \xleftrightarrow{\bar{\alpha}_2} \cdots \xleftrightarrow{\bar{\alpha}_{n-1}} \upsilon_{n-1} \xleftrightarrow{\bar{\alpha}_n} \tau$$

is a path between σ and τ. By Proposition 2.1.5, this path has length at most ε, proving the claim. □

5.2 Bottleneck Stability

Here we return to the setting of Subsection 3.1 with a totally ordered filtration $\mathsf{F} : \mathsf{P} \to \mathsf{Sub}(\mathsf{X}_*)$ and an embedding of P into $\mathbb{R}$.

Definition 5.2.1: For any interval $[a, b] \in \bar{\mathbb{R}}$, let $\left|\left|[a, b]\right|\right|_\infty := \max\left\{|a|, |b|\right\}$. Addition and scalar multiplication of intervals are defined componentwise by $[a, b] + [c, d] = [a + c, b + d]$ and $x[a, b] = [xa, xb]$ for any $x \in \mathbb{R}$.

We now introduce the bottleneck distance between persistence diagrams.

Definition 5.2.2: A **matching** between two non-negative $\mathcal{G}$-functions $\sigma : \bar{\mathsf{P}} \to \mathcal{G}$ and $\tau : \bar{\mathsf{Q}} \to \mathcal{G}$ is a non-negative map $\gamma : \bar{\mathsf{P}} \times \bar{\mathsf{Q}} \to \mathcal{G}$ satisfying

$$\sigma(I) = \sum_{J \in \bar{Q}} \gamma(I, J) \text{ for all } I \neq [p, p] \in \bar{P}$$
$$\tau(J) = \sum_{I \in \bar{P}} \gamma(I, J) \text{ for all } J \neq [q, q] \in \bar{Q}.$$

The **norm** of a matching γ is

$$||\gamma|| := \max_{\left\{I \in \bar{P}, J \in \bar{Q} \,\middle|\, \gamma(I,J) > 0\right\}} ||I - J||_\infty.$$

The **bottleneck distance** between σ and τ is

$$d_B(\sigma, \tau) := \min_{\gamma} ||\gamma||$$

over all matchings γ between σ and τ.

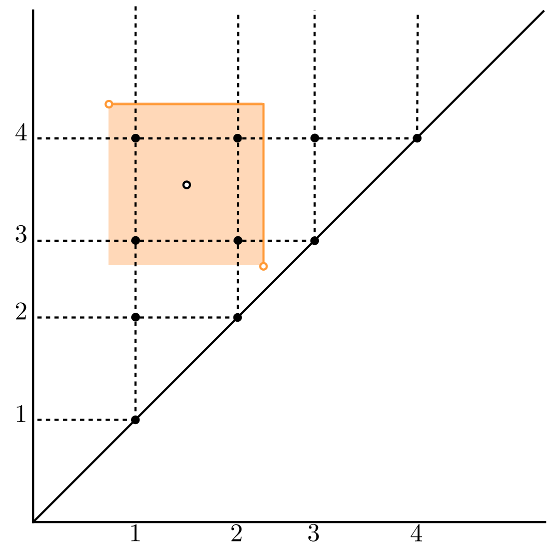


Figure 5.1: The shaded area is the box $\Box_\varepsilon I$ where I is the circle. Note that $\Box_\varepsilon I$ is closed on the top and right, and it is open on the bottom and left.

Definition 5.2.3: For an interval $I = [a, b]$ in $\bar{\mathbb{R}}$ and a value $\varepsilon \geqslant 0$, let

$$\Box_\varepsilon I := \left\{ [x, y] \in \bar{\mathbb{R}} \,\middle|\, a - \varepsilon < x \leqslant a + \varepsilon \text{ and } b - \varepsilon < y \leqslant b + \varepsilon \right\}$$

be the subposet of $\bar{\mathbb{R}}$ consisting of intervals ε-close to I. If I is too close to the diagonal, that is if $b - \varepsilon \leqslant a + \varepsilon$, then we let $\Box_\varepsilon I$ be empty. We call $\Box_\varepsilon I$ the ε-*box* around I. See Figure 5.1.

Lemma 5.2.4: Let $\tilde{f} : \bar{\mathbb{R}} \to \mathcal{G}$ be a constructible, monotone function, $\sigma : \bar{\mathbb{R}} \to \mathcal{G}$ its Möbius inversion, $I = [a, b] \in \bar{\mathbb{R}}$, and $\varepsilon > 0$. If $\Box_\varepsilon I$ is nonempty, then

$$\sum_{J \in \Box_\varepsilon I} \sigma(J) = \tilde{f}(a + \varepsilon, b + \varepsilon) - \tilde{f}(a + \varepsilon, b - \varepsilon) - \tilde{f}(a - \varepsilon, b + \varepsilon) + \tilde{f}(a - \varepsilon, b - \varepsilon).$$

Proof. The equality follow easily from the Möbius inversion formula; see Equation 4.2. We have that

$$\sum_{J \in \Box_\varepsilon I} \sigma(J) = \sum_{\substack{J \in \bar{\mathbb{R}}: \\ J \preceq [a+\varepsilon, b+\varepsilon)}} \sigma(J) - \sum_{\substack{J \in \bar{\mathbb{R}}: \\ J \preceq [a+\varepsilon, b-\varepsilon]}} \sigma(J) - \sum_{\substack{J \in \bar{\mathbb{R}}: \\ J \preceq [a-\varepsilon, b+\varepsilon]}} \sigma(J) + \sum_{\substack{J \in \bar{\mathbb{R}}: \\ J \preceq [a-\varepsilon, b-\varepsilon]}} \sigma(J)$$

$$= \tilde{f}(a+\varepsilon, b+\varepsilon) - \tilde{f}(a+\varepsilon, b-\varepsilon) - \tilde{f}(a-\varepsilon, b+\varepsilon) + \tilde{f}(a-\varepsilon, b-\varepsilon).$$

□

Lemma 5.2.5 (Box Lemma)**:** Let $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ be two ε-interleaved constructible filtrations with i-dimensional birth-death functions $\tilde{f}$ and $\tilde{g}$ and persistence diagrams σ and τ respectively. For any $I \in \bar{\mathbb{R}}$, and $\mu > 0$

$$\sum_{J \in \Box_\mu I} \sigma(J) \leqslant \sum_{J \in \Box_{\mu+\varepsilon} I} \tau(J)$$

whenever $\Box_{\mu+\varepsilon} I$ is nonempty.

Proof. Suppose $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$ are ε-interleaved by Φ in Diagram 3.1. Define $\varphi_r : \tilde{\mathsf{F}}(r) \to \tilde{\mathsf{G}}(r+\varepsilon)$ as $\Phi\big((r,0) \leqslant (r+\varepsilon,1)\big)$ and define $\psi_r : \tilde{\mathsf{G}}(r) \to \tilde{\mathsf{F}}(r+\varepsilon)$ as $\Phi\big((r,1) \leqslant (r+\varepsilon,0)\big)$.

Suppose $I = [a, b]$. By Lemma 5.2.4,

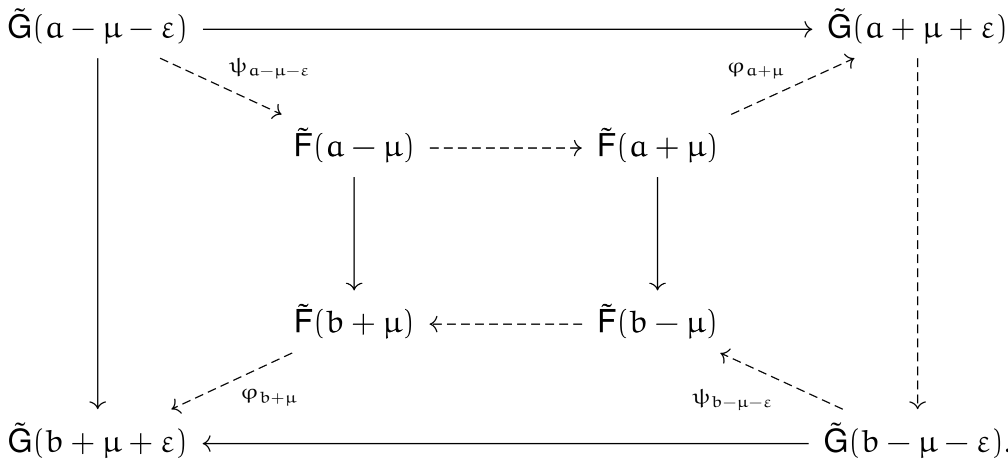
$$\sum_{J \in \Box_\mu I} \sigma(J) = \tilde{f}(a+\mu, b+\mu) - \tilde{f}(a+\mu, b-\mu)$$

$$- \tilde{f}(a-\mu, b+\mu) + \tilde{f}(a-\mu, b-\mu)$$

$$\sum_{J \in \Box_{\mu+\varepsilon} I} \tau(J) = \tilde{g}(a+\mu+\varepsilon, b+\mu+\varepsilon) - \tilde{g}(a+\mu+\varepsilon, b-\mu-\varepsilon)$$

$$- \tilde{g}(a-\mu-\varepsilon, b+\mu+\varepsilon) + \tilde{g}(a-\mu-\varepsilon, b-\mu-\varepsilon)$$

Consider the following commutative diagram where the horizontal and vertical arrows are the appropriate morphisms from $\tilde{\mathsf{F}}$ and $\tilde{\mathsf{G}}$:



Let

$$\mathsf{T} := \{a-\mu-\varepsilon < a-\mu < a+\mu < a+\mu+\varepsilon < b-\mu-\varepsilon < b-\mu < b+\mu < b+\mu+\varepsilon\} \subseteq \mathbb{R}.$$

Let $\mathsf{H} : \mathsf{T} \to \mathsf{Sub}(X_*)$ be the T-filtration determined by the following diagram:

$$\begin{array}{ccc} \mathsf{H}(a-\mu-\varepsilon) = \mathsf{G}(a-\mu-\varepsilon) & \longrightarrow & \mathsf{H}(a-\mu) = \mathsf{F}(a-\mu) \\ & & \downarrow \\ \mathsf{H}(a+\mu) = \mathsf{F}(a+\mu) & \longleftarrow & \mathsf{H}(b-\mu) = \mathsf{F}(b-\mu) \\ \downarrow & & \\ \mathsf{H}(b-\mu-\varepsilon) = \mathsf{G}(b-\mu-\varepsilon) & \longrightarrow & \mathsf{H}(a+\mu+\varepsilon) = \mathsf{G}(a+\mu+\varepsilon) \\ & & \downarrow \\ \mathsf{H}(b+\mu) = \mathsf{F}(b+\mu) & \longleftarrow & \mathsf{H}(b+\mu+\varepsilon) = \mathsf{G}(b+\mu+\varepsilon). \end{array}$$

Here the value of H is given on each value in T and morphisms between adjacent objects are the dashed arrows in the above commutative diagram. Let $\tilde{h}$ be the i-dimensional birth-death function of $\tilde{\mathsf{H}}$ and υ its persistence diagram. We have the following equalities:

$$\tilde{h}(a+\mu, b+\mu) = \tilde{f}(a+\mu, b+\mu)$$
$$\tilde{h}(a+\mu, b-\mu) = \tilde{f}(a+\mu, b-\mu)$$
$$\tilde{h}(a-\mu, b+\mu) = \tilde{f}(a-\mu, b+\mu)$$
$$\tilde{h}(a-\mu, b-\mu) = \tilde{f}(a-\mu, b-\mu)$$
$$\tilde{h}(a+\mu+\varepsilon, b+\mu+\varepsilon) = \tilde{g}(a+\mu+\varepsilon, b+\mu+\varepsilon)$$
$$\tilde{h}(a+\mu+\varepsilon, b-\mu-\varepsilon) = \tilde{g}(a+\mu+\varepsilon, b-\mu-\varepsilon)$$
$$\tilde{h}(a-\mu-\varepsilon, b+\mu+\varepsilon) = \tilde{g}(a-\mu-\varepsilon, b+\mu+\varepsilon)$$
$$\tilde{h}(a-\mu-\varepsilon, b-\mu-\varepsilon) = \tilde{g}(a-\mu-\varepsilon, b-\mu-\varepsilon).$$

By Lemma 5.2.4 along with the above substitutions, we have

$$\sum_{J \in \square_\mu I} \upsilon(J) = \sum_{J \in \square_\mu I} \sigma(J)$$

$$\sum_{J \in \square_{\mu+\varepsilon} I} \upsilon(J) = \sum_{J \in \square_{\mu+\varepsilon} I} \tau(J).$$

By the inclusion $\square_\mu I \subseteq \square_{\mu+\varepsilon} I$ along with Proposition 4.3.4, we have

$$\sum_{J \in \square_\mu I} \upsilon(J) \leqslant \sum_{J \in \square_{\mu+\varepsilon} I} \upsilon(J).$$

This proves the statement.

□

Definition 5.2.6: The **injectivity radius** of a finite, totally ordered metric lattice P is

$$\rho := \min_{a,b \in P : a \neq b} \frac{d_P(a,b)}{2}.$$

Note that if $\mathsf{F} : \mathsf{P} \to \mathsf{Sub}(X_*)$ is a P-filtration with persistence diagram σ, then for any $I \in \bar{\mathsf{P}}$

$$\sigma(I) = \sum_{J \in \Box_\rho I} \sigma(J).$$

Also for any $[a, b] \in \bar{\mathbb{R}}$ if $\sigma(a, b) \neq 0$, then $|a - b| \geqslant 2\rho$.

Lemma 5.2.7 (Easy Bijection)**:** Let $\tilde{\mathsf{F}}$ be a P-constructible filtration with persistence diagram σ and $\rho > 0$ the injectivity radius of P. If $\tilde{\mathsf{G}}$ is a Q-constructible filtration with persistence diagram τ such that $d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}}) < \rho/2$, then $d_B(\sigma, \tau) \leqslant d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$.

Proof. Let $\varepsilon = d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$. Choose a sufficiently small $\mu > 0$ such that $2\mu + 2\varepsilon < \rho$. We construct a matching $\gamma_\mu : \bar{\mathsf{P}} \times \bar{\mathsf{Q}} \to \mathcal{G}$ such that

$$\sigma(I) = \sum_{J \in \bar{\mathsf{Q}}} \gamma_\mu(I, J) \text{ for all } I \neq [p, p] \in \bar{\mathsf{P}} \tag{5.1}$$

$$\tau(J) = \sum_{I \in \bar{\mathsf{P}}} \gamma_\mu(I, J) \text{ for all } J \neq [q, q] \in \bar{\mathsf{Q}}. \tag{5.2}$$

Fix an $I \neq [p, p] \in \bar{\mathsf{P}}$. By Lemma 5.2.5,

$$\sigma(I) = \sum_{J \in \Box_\mu I} \sigma(J) \leqslant \sum_{J \in \Box_{\mu+\varepsilon} I} \tau(J) \leqslant \sum_{J \in \Box_{\mu+2\varepsilon} I} \sigma(J) = \sigma(I).$$

Let $\gamma_\mu(I, J) := \tau(J)$ for all $J \in \Box_{\mu+\varepsilon} I$. Repeat for all $I \in \bar{\mathsf{P}}$. Equation 5.1 is satisfied.

We now check that γ_μ satisfies Equation 5.2. Fix an interval $J = [a, b] \in \bar{\mathsf{Q}}$ with $\tau(J) \neq 0$. If $\frac{b-a}{2} > \mu + \varepsilon$, then by Lemma 5.2.5

$$\tau(J) \leqslant \sum_{I \in \Box_\mu J} \tau(I) \leqslant \sum_{I \in \Box_{\mu+\varepsilon} J} \sigma(I).$$

This means $\gamma_\mu(I, J) = \tau(J)$ for some $I \in \Box_{\mu+\varepsilon} J$. If $\frac{b-a}{2} \leqslant \mu + \varepsilon$, then it must be that $\gamma_\mu(I, J) = 0$ for all $I \neq [p, p] \in \bar{\mathsf{P}}$ for the following reason. Suppose $I = [c, d]$, where $c \neq d$, and $\gamma_\mu\big([c, d], [a, b]\big) \neq 0$. Then $\max\{|a - c|, |b - d|\} \leqslant \mu + \varepsilon$ and therefore $|a - c| \leqslant 3\mu + 3\varepsilon$ which is less than twice the injectivity radius ρ. This means J is unmatched and we may match it to the diagonal. That is, we let $\gamma_\mu\big(\big[\frac{b-a}{2}, \frac{b-a}{2}\big], J\big) := \tau(J)$.

By construction, $\|\gamma_\mu\| \leqslant \mu + \varepsilon$ for all $\mu > 0$ sufficiently small. Therefore $d_B(\sigma, \tau) \leqslant \varepsilon = d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$. □

Theorem 5.2.8 (Bottleneck Stability)**:** Let $\mathsf{F} : P \to \mathsf{Sub}(X_*)$ and $\mathsf{G} : Q \to \mathsf{Sub}(X_*)$ be two filtrations over finite, totally ordered lattices P and Q with i-th persistence diagrams σ and τ respectively. Then $d_B(\sigma, \tau) \leqslant d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$.

Proof. Let $\varepsilon = d_I(\tilde{\mathsf{F}}, \tilde{\mathsf{G}})$. By Proposition 3.1.3, there is a one parameter family of constructible filtrations $\{\tilde{\mathsf{K}}_t\}_{t\in[0,1]}$ such that $d_I(\tilde{\mathsf{K}}_t, \tilde{\mathsf{K}}_s) \leqslant \varepsilon|t-s|$, $\tilde{\mathsf{K}}_0 \cong \tilde{\mathsf{F}}$, and $\tilde{\mathsf{K}}_1 \cong \tilde{\mathsf{G}}$. Let υ_t be the i-dimensional persistence diagram of $\tilde{\mathsf{K}}_t$ for each $t \in [0, 1]$. Each $\tilde{\mathsf{K}}_t$ is constructible with respect to some totally ordered metric lattice P_t, and each P_t has an injectivity radius $\rho_t > 0$. For each $t \in [0, 1]$, consider the open interval

$$U(t) = (t - {}^{\rho_t}/_{4\varepsilon}, t + {}^{\rho_t}/_{4\varepsilon}) \cap [0, 1]$$

By compactness of $[0, 1]$, there is a finite set $Q = \{0 = t_0 < t_1 < \cdots < t_n = 1\}$ such that $\cup_{i=0}^{n} U(t_i) = [0, 1]$. We assume that Q is minimal, that is, there does not exists a pair $t_i, t_j \in Q$ such that $U(t_i) \subseteq U(t_j)$. If this is not the case, simply throw away $U(t_i)$ and we still have a covering of $[0, 1]$. As a consequence, for any consecutive pair $t_i < t_{i+1}$, we have $U(t_i) \cap U(t_{i+1}) \neq \emptyset$. This means

$$t_{i+1} - t_i \leqslant \frac{1}{4\varepsilon}(\rho_{t_{i+1}} + \rho_{t_i}) \leqslant \frac{1}{2\varepsilon} \max\{\rho_{t_{i+1}}, \rho_{t_i}\}$$

and therefore $d_I(\tilde{\mathsf{K}}_{t_i}, \tilde{\mathsf{K}}_{t_{i+1}}) \leqslant \frac{1}{2} \max\{\rho_{t_i}, \rho_{t_{i+1}}\}$. By Lemma 5.2.7,

$$d_B\big(\upsilon_{t_i}, \upsilon_{t_{i+1}}\big) \leqslant d_I\big(\tilde{\mathsf{K}}_{t_i}, \tilde{\mathsf{K}}_{t_{i+1}}\big),$$

for all $0 \leqslant i \leqslant n-1$. Therefore

$$d_B(\sigma,\tau) \leqslant \sum_{i=0}^{n-1} d_B\big(\upsilon_{t_i}, \upsilon_{t_{i+1}}\big) \leqslant \sum_{i=0}^{n-1} d_I\big(\tilde{K}_{t_i}, \tilde{K}_{t_{i+1}}\big) \leqslant \varepsilon.$$

□

5.3 Equivalence of Bottleneck and Edit Distances

We prove that the bottleneck distance defined between persistence diagrams over totally ordered lattices is strongly equivalent to the edit distance.

Theorem 5.3.1: Let P and Q be finite, totally ordered metric lattices embedded in $\mathbb{R}$ with $\bot_P = \bot_Q = 0$ and $\sigma : \bar{P} \to \mathcal{G}$ and $\tau : \bar{Q} \to \mathcal{G}$ two $\mathcal{G}$-functions. Then $d_B(\sigma,\tau) \leqslant d_{\mathsf{Fnc}(\mathcal{G})}(\sigma,\tau) \leqslant 2d_B(\sigma,\tau)$.

Proposition 5.3.2: Let $\sigma : \bar{P} \to \mathcal{G}$ and $\tau : \bar{Q} \to \mathcal{G}$ be non-negative $\mathcal{G}$-functions and γ be an ε-matching between σ and τ. Then γ induces a 1-parameter family of $\mathcal{G}$-functions $\{\upsilon_t\}_{t\in[0,1]}$ with $\upsilon_0 = \sigma$ and $\upsilon_1 = \tau$.

Proof. Let $\bar{S}_t := \big\{(1-t)I + tJ \,\big|\, I \in \bar{P},\ J \in \bar{Q},\ \text{and}\ \gamma(I,J) > 0\big\}$. Define $\upsilon_t : \bar{S}_t \to \mathcal{G}$ to be

$$\upsilon_t(K) := \sum_{\substack{I\in\bar{P},J\in\bar{Q}\\(1-t)I+tJ=K}} \gamma(I,J).$$

At $t = 0$ this reduces to

$$\upsilon_0(K) = \sum_{\substack{I\in\bar{P},J\in\bar{Q}\\I=K}} \gamma(I,J) = \sum_{J\in\bar{Q}} \gamma(K,J) = \sigma(K),$$

for all $K \in \bar{S}_0$, and similarly $\upsilon_1(I) = \tau(I)$. □

As t varies from 0 to 1, there are only finitely many places where the combinatorial structure of υ_t changes. We call these places critical points; see the following definition. These combinatorial changes occur where endpoints of intervals in $\bar{S}_t$ cross or, equivalently, where the cardinality of the set of endpoints changes.

Definition 5.3.3: Let $S_t = \{w \in \mathbb{R} \mid [w, x] \text{ or } [x, w] \in \bar{S}_t\}$ be the set of endpoints of intervals in $\bar{S}_t$. A point $t \in [0, 1]$ is **critical** if for all sufficiently small $\delta > 0$, there exists $s \in (t - \delta, t + \delta)$ with $|S_t| \neq |S_s|$.

Lemma 5.3.4: If $t \in [0, 1]$ is not a critical point, then for any $K \in \bar{S}_t$ there is a unique pair of intervals $I \in \bar{P}$ and $J \in \bar{Q}$ with $\gamma(I, J) > 0$ and $(1 - t)I + tJ = K$.

Proof. Suppose $t \in [0, 1]$ is not critical and there exists $I, I' \in \bar{P}$ and $J, J' \in \bar{Q}$ with $\gamma(I, J) > 0$, $\gamma(I', J') > 0$ and $(1 - t)I + tJ = (1 - t)I' + tJ'$. Then for any t' sufficiently close to t, $(1 - t')I + t'J = (1 - t')I' + t'J'$. Since the interpolation is linear and two lines that intersect in more than one point must be the same line, it follows that $I = I'$ and $J = J'$. □

Lemma 5.3.5: If $\alpha : P \to Q$ is a metric lattice map and $\bar{\alpha} : \bar{P} \to \bar{Q}$ is its induced map on intervals then

$$\max_{I \in \bar{P}} ||I - \bar{\alpha}(I)||_\infty \leqslant ||\bar{\alpha}|| \leqslant 2 \max_{I \in \bar{P}} ||I - \bar{\alpha}(I)||_\infty.$$

Proof. First note that by Proposition 2.1.5, $||\alpha|| = ||\bar{\alpha}||$ and since $\bar{\alpha}$ is induced by α, $\max_{I \in \bar{P}} ||I - \bar{\alpha}(I)||_\infty = \max_{a \in P} |a - \alpha(a)|$ so the inequality reduces to

$$\max_{a \in P} |a - \alpha(a)| \leqslant \max_{a, b \in P} \big| |a - b| - |\alpha(a) - \alpha(b)| \big| \leqslant 2 \max_{a \in P} |a - \alpha(a)|.$$

Note that since $\bot_P = \bot_Q = 0$, each element of P and Q are non-negative. Assume, without loss of generality, that $a \geqslant b$. Then the middle quantity above reduces to

$$\max_{a \geqslant b \in P} \big| a - b - \big(\alpha(a) - \alpha(b)\big) \big| = \max_{a \geqslant b \in P} \big| a - \alpha(a) - \big(b - \alpha(b)\big) \big|$$

Letting $b = 0$ yields the first inequality and the triangle inequality yields the second. □

Lemma 5.3.6: If $t \in [0, 1]$ is not a critical point and $s \in [0, 1]$ is any point with no critical points strictly between t and s, then there is a charge-preserving morphism $(\upsilon_t, \upsilon_s, \bar{\alpha}_{t,s})$ with distortion at most $2\varepsilon|s - t|$.

Proof. We start by defining a map $\alpha_{t,s} : S_t \to S_s$. For any $b \in S_t$ note that since t is not critical, there are unique intervals $I \in \bar{P}$ and $J \in \bar{Q}$ with $\gamma(I, J) > 0$ and either $(1 - t)I + tJ = [a, b]$ or $[b, c]$. If $(1 - t)I + tJ = [a, b]$ then define $\alpha_{t,s}(b)$ to be the right endpoint of the interval $(1 - s)I + sJ$. Similarly, if b is a left endpoint, then we define $\alpha_{t,s}(b)$ to be the left endpoint of $(1 - s)I + sJ$. This map is order preserving since as t varies, endpoints of intervals only cross at critical points and there are no critical points strictly between t and s.

To prove that $\bar{\alpha}_{t,s}$ is charge-preserving, observe that

$$\begin{aligned}
\sum_{K \in \bar{\alpha}_{t,s}^{-1}(L)} \upsilon_t(K) &= \sum_{\substack{I \in \bar{P}, J \in \bar{Q} \\ (1-s)I + sJ = L}} \upsilon_t((1-t)I + tJ) \\
&= \sum_{\substack{I \in \bar{P}, J \in \bar{Q} \\ (1-s)I + sJ = L}} \left(\sum_{\substack{I' \in \bar{P}, J' \in \bar{Q} \\ (1-t)I' + tJ' = (1-t)I + tJ}} \gamma(I', J') \right) \\
&= \sum_{\substack{I \in \bar{P}, J \in \bar{Q} \\ (1-s)I + sJ = L}} \gamma(I, J) = \upsilon_s(L)
\end{aligned}$$

where the third equality follows from Lemma 5.3.4 and the assumption that t is not critical. The distortion of $\bar{\alpha}_{t,s}$ is

$$\begin{aligned}
\|\bar{\alpha}_{t,s}\| &\leqslant 2 \max_{K \in \bar{S}_t} \|K - \bar{\alpha}_{t,s}(K)\|_\infty \\
&= 2 \max_{\substack{I \in \bar{P}, J \in \bar{Q} \\ \gamma(I,J) > 0}} \|(1-t)I + tJ - (1-s)I - sJ\|_\infty \\
&= 2|s - t| \max_{\substack{I \in \bar{P}, J \in \bar{Q} \\ \gamma(I,J) > 0}} \|I - J\|_\infty \leqslant 2\varepsilon|s - t|.
\end{aligned}$$

□

Lemma 5.3.7: For any non-negative $\mathcal{G}$-functions $\sigma : \bar{P} \to \mathbb{Z}$ and $\tau : \bar{Q} \to \mathbb{Z}$ over finite sublattices $P, Q \subseteq \mathbb{R}$, $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau) \leqslant 2d_B(\sigma, \tau)$.

Proof. We show that $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau) \leqslant 2d_B(\sigma, \tau)$ by showing that an ε-matching between σ and τ induces a path between σ and τ of length at most 2ε. For any ε-matching γ between σ and τ, let $\{\upsilon_t\}_{t\in[0,1]}$ be the interpolation induced by γ from Proposition 5.3.2. Let $\{s_0 = 0 < s_1 \cdots < s_n = 1\} \subseteq [0, 1]$ be the set of critical points of the interpolation and choose $\{t_0 < \cdots < t_{n-1}\} \subseteq [0, 1]$ with $0 < t_0 < s_1 < t_1 \cdots < t_{n-1} < 1$. Then the charge-preserving morphisms $\alpha_{t_i, s_{i\pm1}}$ from Lemma 5.3.6 form a path between σ and τ with length at most $\sum_{i=0}^{n-1} 2\varepsilon|s_i - t_i| = 2\varepsilon$. □

Lemma 5.3.8: For any non-negative $\mathcal{G}$-functions $\sigma : \bar{P} \to \mathbb{Z}$ and $\tau : \bar{Q} \to \mathbb{Z}$ over finite sublattices $P, Q \subseteq \mathbb{R}$, $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau) \geqslant d_B(\sigma, \tau)$.

Proof. To show that $d_B(\sigma, \tau) \leqslant d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau)$, we show that a charge-preserving morphism induces a matching. Let (σ, τ, α) be a charge-preserving morphism with distortion ε. Define a matching γ between σ and τ by

$$\gamma(I, J) := \begin{cases} \sigma(I) & \text{if } \alpha(I) = J \\ 0 & \text{otherwise} \end{cases}.$$

Then we have that for any $J \in \bar{Q}$

$$\sum_{I\in\bar{P}} \gamma(I, J) = \sum_{I\in\alpha^{-1}(J)} \sigma(I) = \tau(J)$$

and for any $I \in \bar{P}$

$$\sum_{J\in\bar{Q}} \gamma(I, J) = \alpha(I).$$

Therefore γ is a matching. The norm of γ is

$$||\gamma|| = \max_{I\in\bar{P}, J\in\bar{Q} : \gamma(I,J)>0} ||I - J||_\infty = \max_{I\in\bar{P}} ||I - \alpha(I)||_\infty \leqslant ||\alpha||.$$

□

Theorem 5.3.1 follows immediately from Lemma 5.3.7 and Lemma 5.3.8. The following two examples show that the bounds in Theorem 5.3.1 are tight.

Example 5.3.9: Let $P = \{0 < 1 < 2 < 3\}$ be a totally ordered metric lattice where the distance between two elements is the absolute value of their difference. Let $\sigma, \upsilon : \bar{P} \to \mathbb{Z}$ be two integral functions defined as

$$\sigma[a,b] := \begin{cases} 1 & \text{if } [a,b] = [0,1], [2,3] \\ 0 & \text{otherwise} \end{cases} \qquad \upsilon[a,b] := \begin{cases} 1 & \text{if } [a,b] = [0,2], [1,3] \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 5.2. The bottleneck distance, d_B, between σ and υ is $d_B(\sigma, \upsilon) = 1$. We now compute the edit distance, $d_{\mathsf{Fnc}(\mathcal{G})}$, between σ and υ. Consider a third integral function $\tau : \bar{Q} \to \mathbb{Z}$ where $Q = \{0 < 0.5 < 1 < 1.5 < 2 < 2.5 < 3\}$ is a finite, totally ordered metric lattice where the distance between any two elements is the absolute value of the difference and

$$\upsilon[a,b] := \begin{cases} 1 & \text{if } [a,b] = [0,1.5], [1.5,3] \\ 0 & \text{otherwise.} \end{cases}$$

Let $\alpha : P \to Q$ be the bounded lattice function defined as follows

$$\alpha(0) := 0 \qquad \alpha(1) := 1.5 \qquad \alpha(2) := 1.5 \qquad \alpha(3) := 3.$$

We now have a pair of charge-preserving morphisms $(\sigma, \tau, \bar{\alpha})$ and $(\upsilon, \tau, \bar{\alpha})$. Thus $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \upsilon) \leqslant 2||\bar{\alpha}|| = 2||\alpha|| = 2(0.5) = 1$. Further, this is a shortest path between σ and υ in $\mathsf{Fnc}(\mathcal{G})$. Therefore $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \upsilon) = 1$.

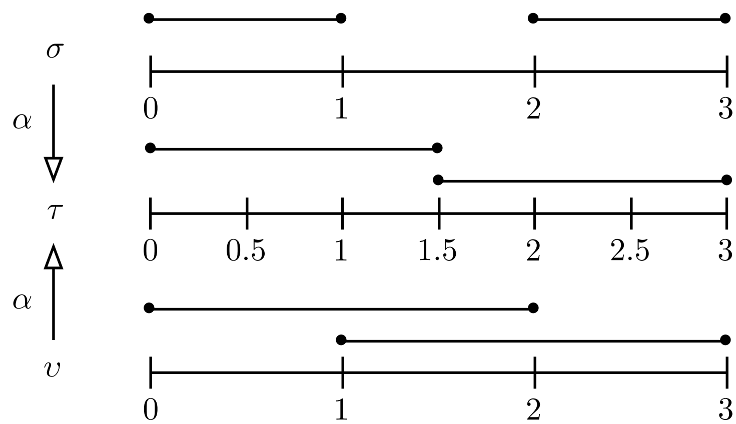


Figure 5.2: Above are three integral functions $\sigma, \upsilon : \bar{P} \to \mathbb{Z}$ and $\tau : \bar{Q} \to \mathbb{Z}$ drawn as barcodes and two charge-preserving morphisms $(\sigma, \tau, \bar{\alpha})$ and $(\upsilon, \tau, \bar{\alpha})$.

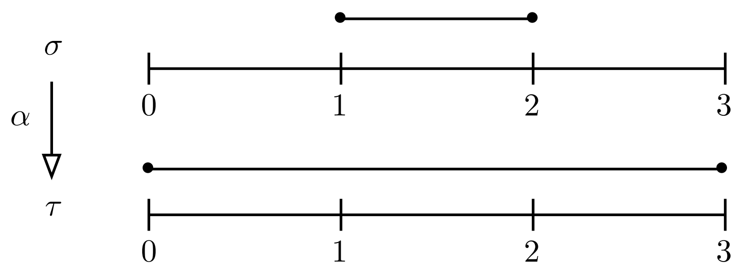


Figure 5.3: Above are two integral functions $\sigma, \tau : \bar{P} \to \mathbb{Z}$ drawn as barcodes and a charge-preserving morphism $(\sigma, \tau, \bar{\alpha})$.

Example 5.3.10: Let P be the metric lattice defined in Example 5.3.9 and $\sigma, \tau : \bar{P} \to \mathbb{Z}$ be defined as

$$\sigma[a,b] := \begin{cases} 1 & \text{if } [a,b] = [1,2] \\ 0 & \text{otherwise} \end{cases} \qquad \upsilon[a,b] := \begin{cases} 1 & \text{if } [a,b] = [0,3] \\ 0 & \text{otherwise.} \end{cases}$$

See Figure 5.3. The bottleneck distance between σ and τ is 1. We now compute the edit distance $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau)$. Let $\alpha : P \to P$ be the bounded lattice function defined by

$$\alpha(0) := 0 \qquad \alpha(1) := 0 \qquad \alpha(2) := 3 \qquad \alpha(3) := 3.$$

The lattice map α induces a charge-preserving morphism $(\sigma, \tau, \bar{\alpha})$ with distortion 2. This is the shortest path between σ and τ in $\mathsf{Fnc}(\mathcal{G})$ so $d_{\mathsf{Fnc}(\mathcal{G})}(\sigma, \tau) = 2$.

Chapter 6

Conclusion

In this dissertation we have significantly broadened the types of filtrations that persistent homology can be applied to. Traditionally persistent homology was only defined for filtrations of chain complexes of vector spaces over finite, totally ordered sets. We extended this to filtrations of chain complexes in an essentially small abelian category valued over finite lattices. Additionally, we developed categories that these filtrations and their persistence diagrams reside in. These advances have tremendous potential for applications.

Expanding the categories that the chain complexes lie in allows for more general homology theories to be used. For example, given a Rips complex, we can now compute its persistence diagram with ring coefficients rather than just field coefficients. This allows for a more nuanced study of torsion that appears in the Rips complex.

Filtrations over more general posets have long been an object of study in persistent homology [26]. In particular, multiparameter persistent homology can be viewed as a special case of this. One of the unfortunate weaknesses of filtrations such as the Rips and Čech complexes is that they are sensitive to outliers. This can be addressed by incorporating a density parameter in the filtration. The resulting filtration is no longer a filtration over a totally ordered set however. For this reason, traditional ways of defining the persistence diagram no longer apply. Our framework has the advantage of being able to handle filtrations such as these.

The categorical structures presented here unlock many new directions and applications of persistent homology. For example, parameterized data sets (such as time varying data) naturally induce morphisms of filtrations between their Rips complexes. Applying the pipeline presented here yields persistence diagrams with morphisms between them. This extra structure allows for new directions in persistent homology, such as cosheaves of persistence diagrams.

Bibliography

[1] Afra Zomorodian. Fast construction of the vietoris-rips complex. *Computers & Graphics*, 34(3):263–271, 2010.

[2] Stefan Dantchev and Ioannis Ivrissimtzis. Efficient construction of the čech complex. *Computers & Graphics*, 36(6):708–713, 2012.

[3] Herbert Edelsbrunner and Ernst P Mücke. Three-dimensional alpha shapes. *ACM Transactions on Graphics (TOG)*, 13(1):43–72, 1994.

[4] Robert Ghrist. Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, 45(1):61–75, 2008.

[5] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[6] Henry Adams, Manuchehr Aminian, Elin Farnell, Michael Kirby, Joshua Mirth, Rachel Neville, Chris Peterson, and Clayton Shonkwiler. A fractal dimension for measures via persistent homology. In *Topological Data Analysis*, pages 1–31. Springer, 2020.

[7] Robert MacPherson and Benjamin Schweinhart. Measuring shape with topology. *Journal of Mathematical Physics*, 53(7):073516, 2012.

[8] Benjamin Schweinhart. Fractal dimension and the persistent homology of random geometric complexes. *Advances in Mathematics*, 372:107291, 2020.

[9] Leonid Polterovich, Daniel Rosen, Karina Samvelyan, and Jun Zhang. *Topological persistence in geometry and analysis*, volume 74. American Mathematical Soc., 2020.

[10] Michael Usher and Jun Zhang. Persistent homology and floer–novikov theory. *Geometry & Topology*, 20(6):3333–3430, 2016.

[11] Amit Patel. Generalized persistence diagrams. *Journal of Applied and Computational Topology*, 1(3-4):397–419, 2018.

[12] Alexander McCleary and Amit Patel. Bottleneck stability for generalized persistence diagrams. *Proceedings of the American Mathematical Society*, 148:3149–3161, Mar 2020.

[13] Gunnar Carlsson and Afra Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, Jul 2009.

[14] Alexander McCleary and Amit Patel. Edit distance and persistence diagrams over lattices. *arXiv preprint arXiv:2010.07337*, 2020.

[15] Woojin Kim. *The Persistent Topology of Dynamic Data*. PhD thesis, The Ohio State University, 2020.

[16] Katharine Turner, Sayan Mukherjee, and Doug M Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[17] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A Course in Metric Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2001.

[18] G. Carlsson, F. Mémoli, A. Ribeiro, and S. Segarra. Axiomatic construction of hierarchical clustering in asymmetric networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 5219–5223, 2013.

[19] Moritz Kerz, Florian Strunk, and Georg Tamme. Algebraic k-theory and descent for blow-ups. *Inventiones mathematicae*, 211(2):523–577, 2018.

[20] Emily Riehl. *Categorical Homotopy Theory*. Cambridge University Press, 2014.

[21] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas Guibas, and Steve Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-fifth*

Annual Symposium on Computational Geometry, SCG '09, pages 237–246, New York, NY, USA, 2009. ACM.

[22] Peter Bubenik, Vin de Silva, and Vidit Nanda. Higher interpolation and extension for persistence modules. *SIAM Journal on Applied Algebra and Geometry*, 1(1):272–284, 2017.

[23] J.L. Doob. *Measure Theory*. Springer-Verlag New York, 1994.

[24] M. Barnabei, A. Brini, and G.-C. Rota. The theory of Möbius functions. *Russian Mathematical Surveys*, 41(3):135–188, 1986.

[25] Gian Carlo Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2(4):340–368, 1964.

[26] Gunnar Carlsson and Afra Zomorodian. The theory of multidimensional persistence. *Discrete & Computational Geometry*, 42(1):71–93, 2009.