Stability of Zigzag Persistence with Respect to a Reflection-type Distance

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Mathematical Sciences in the Graduate School of The Ohio State University

By

Alex Elchesen, B.S.

Graduate Program in Mathematical Sciences

The Ohio State University

2017

Master’s Examination Committee:

Facundo Mémoli, Advisor

Matthew Kahle
Abstract

We use the reflection functors introduced by Bernstein, Gelfand, and Ponomarev in [4] to define a metric on the space of all zigzag modules of a given length, which we call the reflection distance. We show that the reflection distance between two given zigzag modules of the same length is an upper bound for the bottleneck distance between the persistence diagrams of the given modules. We also extend our distance to weighted zigzag modules and prove an analogous result in this setting.
This thesis is dedicated to everyone ever.
Acknowledgments

First and foremost, I want to thank my advisor Facundo Mémoli for introducing me to applied algebraic topology and topological data analysis, for the huge amounts of support and guidance he has given me, and for the endless patience he seems to be able to muster up for me. I also want to thank him for his part in creating a community both with his involvement in the TGDA seminars and within his TGDA group.

Within the TGDA group, I would like to thank Woojin Kim for many beneficial conversations about zigzag persistence. I would also like to thank Samir Chowdhury and Zane Smith for their comments when I was presenting on algebraic stability. I would also like to thank Matthew Kahle for sitting on my committee and for giving me a lot of helpful career advice.

Outside of the TGDA community, there are too many people with whom helpful discussions were had; special acknowledgments go out to Jake Blomquist and Gabe Bainbridge for useful discussions on category theory, to Jared O’Neal for discussions about the practicality of what we do, and to Jacob Miller for discussions about all the things.

I would like to acknowledge support from the National Science Foundation Research Training Grant on Algebraic Topology and its Applications grant number
NSF-RTG #DMS-1547357 and the departmental Rhodus Graduate Fellowship which both supported me during the bulk of the time that this thesis was written.

Thanks go out to my closest friends Jacob Miller, Steven All, Komla Gnona, Marien Simeni Simeni, and Alex Credico who kept me sane over the last two years. Also, I want to shout out to Josiah Hounyo, Kossi Houngbedji, and Xiaokai Sun who have all gone on to do great things but have still remained great friends.

Finally, I would like to thank my parents and my grandma for their support during this program and, of course, throughout my life.
Vita

March 8, 1991 .............................................. Born - Painseville, Ohio

2015 .................................................. B.S. Mathematics, The Ohio State University
2015-present ................................. Graduate Teaching and Research Associate, The Ohio State University

Fields of Study

Major Field: Mathematical Sciences
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Persistence</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Stability</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Zigzag Persistence and Stability</td>
<td>5</td>
</tr>
<tr>
<td>2. Mathematical Preliminaries</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Relations</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Decorated Endpoints</td>
<td>9</td>
</tr>
<tr>
<td>2.3 Topics from Linear Algebra</td>
<td>10</td>
</tr>
<tr>
<td>2.3.1 Quotient Spaces</td>
<td>10</td>
</tr>
<tr>
<td>2.3.2 Induced Linear Maps</td>
<td>11</td>
</tr>
<tr>
<td>2.3.3 Kernels and Cokernels</td>
<td>12</td>
</tr>
<tr>
<td>2.3.4 Free Vector Spaces</td>
<td>12</td>
</tr>
<tr>
<td>2.4 Basic Category Theory</td>
<td>13</td>
</tr>
<tr>
<td>2.4.1 Categories</td>
<td>13</td>
</tr>
<tr>
<td>2.4.2 Functors</td>
<td>14</td>
</tr>
<tr>
<td>2.4.3 Natural Transformations</td>
<td>15</td>
</tr>
<tr>
<td>2.4.4 Some Categories and Constructions</td>
<td>15</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Types of Morphisms</td>
</tr>
<tr>
<td>2.4.6</td>
<td>Limits and Colimits</td>
</tr>
<tr>
<td>2.4.7</td>
<td>Products and Coproducts</td>
</tr>
<tr>
<td>2.4.8</td>
<td>Pullbacks and Pushouts</td>
</tr>
<tr>
<td>3.1</td>
<td>Zigzag Modules</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Morphisms Between Zigzag Modules</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Interval Modules and the Zero Module</td>
</tr>
<tr>
<td>3.2</td>
<td>Decompositions of Zigzag Modules</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Multisets</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Persistence Barcodes and Persistence Diagrams</td>
</tr>
<tr>
<td>3.3</td>
<td>Reflection Functors</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Sinks and Sources</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Type Reflections and Type Reversals</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Sink and Source Reflection Functors</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Sink Reflection Functors</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Source Reflection Functors</td>
</tr>
<tr>
<td>3.3.6</td>
<td>Properties of Reflection Functors</td>
</tr>
<tr>
<td>3.4</td>
<td>Extroversion and Introversion Functors</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Forward and Backward Flows</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Type Extroversion and Introversion</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Extroversion and Introversion Functors</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Properties of Extroversion and Introversion Functors</td>
</tr>
<tr>
<td>4.1</td>
<td>Transformations of Zigzag Modules</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Arrow Reversals</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Equivalence of Zigzag Modules which Differ by an Arrow Reversal</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Sequences of Compatible Functors</td>
</tr>
<tr>
<td>4.2</td>
<td>The Reflection Distance</td>
</tr>
<tr>
<td>4.3</td>
<td>The Bottleneck Distance and Reflective Stability</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Matchings</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The Bottleneck Distance</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Stability of Persistence Diagrams with Respect to the $p = 1$ Reflection Distance</td>
</tr>
<tr>
<td>4.4</td>
<td>Weighted Zigzag Modules</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Decorated Interval Partitions</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Weighted Zigzag Modules</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.5 The Weighted Reflection Distance</td>
<td>77</td>
</tr>
<tr>
<td>4.5.1 The Weighted Cost of a Composition of Reflections</td>
<td>77</td>
</tr>
<tr>
<td>4.5.2 The Weighted Reflection Distance between Zigzag Modules Weighted by the Same Interval Partition</td>
<td>78</td>
</tr>
<tr>
<td>4.5.3 The General Weighted Reflection Distance</td>
<td>79</td>
</tr>
<tr>
<td>4.6 Diagrams of Weighted Zigzag Modules Stability</td>
<td>87</td>
</tr>
<tr>
<td>Bibliography</td>
<td>91</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Left: Visualization of the persistence diagram of the zigzag module $V = \mathbb{I}([1, 4]) \oplus \mathbb{I}([1, 2]) \oplus \mathbb{I}([2, 3]) \oplus \mathbb{I}([2, 3]) \oplus \mathbb{I}([3, 3])$. Right: Persistence barcode visualization of the same zigzag module.</td>
<td>37</td>
</tr>
<tr>
<td>3.2 Left: Points in the diagram of a zigzag module $V$ move according to the arrows when the reflection functors $R^+_k$ or $R^-_k$ are applied. The point $(k, k)$ corresponding to the summand $\mathbb{I}([k, k])$ is killed. Right: The corresponding picture for barcodes.</td>
<td>46</td>
</tr>
<tr>
<td>3.3 Points in the persistence diagram of a zigzag module $V$ move according to the arrows when the extroversion functor $R^<em>_k$ (left) or introversion functor $R^</em>_k$ (right) are applied. The point $(k, k)$ corresponding to the summand $\mathbb{I}([k, k])$ is annihilated by both of these functors.</td>
<td>57</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Persistence

Persistent homology is a tool for studying data sets in a scale-independent way. One begins with a data set and a method for producing a nested family of simplicial complexes, parametrized by the positive real numbers - such a family of simplicial complexes is referred to as a filtration. Applying a homology functor with coefficients coming from some field to this filtration, one ends up with a collection of vector spaces and maps between these vector spaces. Such objects are referred to as persistence modules and are the main object of study in the theory of persistence. More precisely, a persistence module is:

- a collection of vector spaces \( \{ V_t \}_{t \in \mathbb{R}} \), indexed by \( \mathbb{R} \), together with
- a collection of linear transformations \( \phi_V(s, t) = \phi(s, t) : V_s \to V_t \) defined for all \( s \leq t \) such that \( \phi(t, t) = \text{Id}_{V_t} \) and \( \phi(s, t) \circ \phi(r, s) = \phi(r, t) \) for all \( r \leq s \leq t \).

A persistence module is said to be pointwise finite dimensional if \( V_t \) is a finite dimension vector space for all \( t \in \mathbb{R} \). Pointwise finite dimensional persistence modules decompose into direct sums of “interval modules” [10]. These are especially simple persistence modules which are naturally parametrized by intervals of the real line (see...
or \[10\] for the definitions of interval modules in the case of persistence or see \[3\] below for the definition of interval zigzag modules). Such a decomposition uniquely determines a multiset of real intervals associated to a given persistence module. This multiset is referred to as the barcode of the persistence module and provides an isomorphism invariant for persistence modules.

1.2 Stability

Since one would like to use such invariants to analyze data, stability is a highly desirable property of any such method. Informally, a process which takes data as input and provides some invariant as output is stable if whenever the input data is perturbed slightly, the resulting invariant changes only slightly. A standard metric for measuring the closeness of two persistence barcodes is called the bottleneck distance \(d_b\) (see Chapter \[4\] for the definition of the bottleneck distance). Stability results then take different forms depending on how the underlying data is acquired. Below we provide a survey of several stability results of this nature: we do not give all of the formal definitions but instead refer the reader to the original sources of these results.

Let \(X\) be a topological space and let \(f : X \to \mathbb{R}\) be a continuous function. For each \(a \leq b \in \mathbb{R}\), we have an inclusion of the sub-level sets

\[f^{-1}((\infty, a]) \subseteq f^{-1}((\infty, b]).\]

This gives us a growing sequence of topological spaces indexed by \(\mathbb{R}\), which we call the sub-level set filtration. We obtain a persistence module by applying a homology functor with field coefficients to the sub-level sets and inclusion maps. We then have the following:
Theorem 1.2.1 (Classical Stability Theorem,[9]). Let $X$ be a topological space and let $f$ and $g$ be “tame” functions on $X$. Then

$$d_b^\infty(D(f), D(g)) \leq \| f - g \|_\infty,$$

where $D(f), D(g)$ are the persistence barcodes of the persistence modules obtained by applying homology to the sub-levelset filtrations of $f, g$, respectively.

We refer the reader to [9] for the precise definitions of tame functions. Theorem 1.2.1 says that whenever the functions $f$ and $g$ are uniformly $\epsilon$-close then their persistence barcodes are $\epsilon$-close in the bottleneck distance. In other words, small changes of the data result in only small changes of the persistence barcodes produced.

As another example, suppose that we are given a finite metric space $(X, d_X)$. One common method of obtaining a filtration from such a metric space is to define for each $\alpha > 0$ a simplicial complex $\mathcal{R}_\alpha(X, d_X)$ where $\sigma = \{x_1, \ldots, x_N\}$ is a simplex of $\mathcal{R}_\alpha(X, d_X)$ if and only if $\text{diam}(\sigma) < \alpha$. The resulting filtration is called the Rips filtration and is denoted $\mathcal{R}(X, d_X)$. For a fixed field $F$ and $k \in \mathbb{N}$ we denote the persistence barcode obtained by applying a $k$-dimensional homology functor with coefficients from $F$ by $D_k \mathcal{R}(X)$. Then the following can be said for the stability of such barcodes:

Theorem 1.2.2 ([8]). For any finite metric spaces $(X, d_X)$ and $(Y, d_Y)$, and for any $k \in \mathbb{N}$,

$$d_b^\infty(D_k \mathcal{R}(X), D_k \mathcal{R}(Y)) \leq 2 \cdot d_{GH}((X, d_X), (Y, d_Y)),$$

where $d_{GH}$ is the Gromov-Hausdorff distance.

The metric $d_{GH}$ appearing in Theorem 1.2.2 is called the Gromov-Hausdorff distance and is a common metric on the space of all compact metric spaces; we refer the
reader to [8] for the precise definition. If $X, Y \subset \mathbb{R}^n$ are finite subsets of $\mathbb{R}^n$, viewed as metric subspaces of $\mathbb{R}^n$, then an immediate consequence of the definition of the Gromov-Hausdorff distance is that

$$d_{\text{GH}}((X, d_X), (Y, d_Y)) \leq d_H(X, Y),$$

where $d_H$ is the Hausdorff distance in $\mathbb{R}^n$. Thus, for point clouds $X, Y \subset \mathbb{R}^n$, a consequence of Theorem 1.2.2 is

$$d_{\text{b}}^\infty(D_k\mathcal{R}(X), D_k\mathcal{R}(Y)) \leq 2 \cdot d_H(X, Y).$$

This can be used to infer the following: given a point cloud $X \subset \mathbb{R}^n$ and an $\epsilon$-perturbation $\Delta X$ of $X$ (measured using the Hausdorff distance), the persistence barcodes of $\mathcal{R}(X)$ and $\mathcal{R}(\Delta X)$ are $2\epsilon$-close in the bottleneck distance. Once again, small changes in the input data (the point cloud) result in small changes in the persistence barcode. This property of persistence is extremely desirable from the point of view of data analysis since data is almost always present with some inherent error.

Another notion of stability is known as algebraic stability. The authors of [7] introduced the notion of an $\epsilon$-interleaving between two persistence modules, which is intuitively an “approximate isomorphism”. A notion closely related to the bottleneck distance is the notion of an $\epsilon$-matching between persistence barcodes. We then have the following:

**Theorem 1.2.3** (Algebraic Stability Theorem: [7], [3]). *If two pointwise finite dimensional persistence modules $V$ and $W$ are $\epsilon$-interleaved then there exists an $\epsilon$-matching between their respective barcodes $D(V)$ and $D(W)$.***

The interleaving distance $d_I$ between persistence modules $V$ and $W$ is defined to be the infimal $\epsilon$ for which an $\epsilon$-interleaving exists between $V$ and $W$ exists. Similarly,
it is a matter of fact that the bottleneck distance between the persistence barcodes $D(V)$ and $D(W)$ is equal to the infimal $\epsilon$ for which an $\epsilon$-matching exists between $D(V)$ and $D(W)$. Then an immediate corollary of Theorem 1.2.3 is:

**Corollary 1.2.1.** For any pointwise finite dimensional persistence modules $V$ and $W$ we have

$$d_b(D(V), D(W)) \leq d_I(V, W).$$

In fact, it turns out that $d_I = d_b$ (see [3] for the precise statements). The Algebraic Stability Theorem says that when persistence modules themselves are close (as measured with the interleaving distance) then their corresponding persistence barcodes are also close. Such results are useful because it is often possible to say that similar data produces persistence modules that are close in the interleaving distance.

1.3 **Zigzag Persistence and Stability**

Informally, zigzag modules are persistence modules in which we allow the linear maps to point either left or right; the precise definition is given in Chapter 3. Some work has been done to extend the notion of interleavings to the setting of zigzag modules. Recently, the authors of [2] were able to provide an algebraic stability result for a special class of multidimensional persistence modules called *block decomposable modules*. Using an analogous notion of $\epsilon$-interleaving for multidimensional persistence, they we able to show that if two block decomposable modules are $\frac{5}{2}\epsilon$-interleaved then their exists and $\epsilon$-matching between their “block barcodes”. Bjerkevik was able to reduce the constant factor $\frac{5}{2}$ to 1 even more recently in [1].

The authors were able to define a functor mapping zigzag modules to block decomposable modules. Using this functor to induce notions of $\epsilon$-interleavings between
zigzag modules and $\epsilon$-matchings between their barcodes, they obtained an algebraic stability result for zigzag persistence.

The approach we take to be able to state our stability result is, at least superficially, quite different from the approach which makes use of the notion of interleaving. Our approach is to transform zigzag modules “into” each other, assigning costs to such transformations. The cost of the most efficient transformations defines what we call the reflection distance between zigzag modules. We then show that the reflection distance between two given zigzag modules is an upper bound for the bottleneck distance between their respective persistence barcodes, which for zigzag modules are defined analogously to the persistence barcodes of ordinary persistence modules.
Chapter 2: Mathematical Preliminaries

2.1 Relations

Fix a set $X$. A binary relation on $X$ (often shortened to simply relation) is any nonempty subset $R \subseteq X \times X$. We give names to several important properties that a relation might have:

- **Reflexitivity.** $R$ is said to be reflexive if $(x, x) \in R$ for all $x \in X$.

- **Irreflexivity.** $R$ is said to be irreflexive if $(x, x) \notin X$ for all $x \in X$. Note that, in general, the property of being irreflexive is not the same as the property of not being reflexive.

- **Symmetry.** $R$ is said to be symmetric if $(x, y) \in R$ implies $(y, x) \in R$.

- **Antisymmetry.** $R$ is said to be antisymmetric if $(x, y), (y, x) \in R$ implies $x = y$.

- **Transitivity.** $R$ is said to be transitive if $(x, y), (y, z) \in R$ implies $(x, z) \in R$.

Other types of relations can defined by specifying that any number of properties hold (so long as these properties are consistent with each other). The following types of relations will be used extensively throughout:
• An *equivalence relation* is a relation which is reflexive, symmetric, and transitive.

• A *partial order* is a relation which is reflexive, antisymmetric, and transitive. A pair \((X, P)\) where \(X\) is a set and \(P\) is a partial order on \(X\) is called a *partially ordered set* or *poset* for short.

Other types of relations will be defined when necessary.

If an arbitrary collection of relations \(\{R_\alpha\}_\alpha\) on a fixed set \(X\) all satisfy the same property \(P\) from the list above then so does their intersection \(\bigcap_{\alpha} R_\alpha\). This observation provides a convenient way of describing the “smallest” relation satisfying a given property and containing a given relation. More formally and more generally, suppose \(P\) is any property of relations. Suppose also that for any arbitrary collection \(\{R_\alpha\}_\alpha\) of relations with each \(R_\alpha\) having property \(P\), the intersection \(\bigcap_{\alpha} R_\alpha\) also has property \(P\). Then for any relation \(R\) over a set \(X\) we define the *\(P\)-closure* of \(R\) to be the relation

\[
R^P := \bigcap_{R \subseteq S \subseteq X \times X} S.
\]

By definition \(R \subseteq R^P\) and by assumption \(R^P\) has property \(P\). Moreover, \(R^P\) is minimal in the sense that if \(S\) is any relation on \(X\) containing \(R\) and having property \(P\) then \(R^P \subseteq S\). Note also that \(R = R^P\) if and only if \(R\) has property \(P\). We explicitly describe some of the most common closures in which cases it is possible to find more convenient descriptions:

• The **reflexive closure** of \(R\) is the relation

\[
R^\Delta := \bigcap_{R \subseteq S \subseteq X \times X} S.
\]
It can be shown that $R^\Delta = R \cup \Delta$ where $\Delta := \Delta_X = \{(x, x) \mid x \in X\}$, hence the notation.

- The *symmetric closure* of $R$ is the relation

$$R^\sigma := \bigcap_{R \subseteq S \subseteq X \times X \text{ symmetric}} S.$$  

It can be shown that $R^\sigma = R \cup \{(y, x) \mid (x, y) \in X\}$.

- The *transitive closure* of $R$ is the relation

$$R^+ := \bigcap_{R \subseteq S \subseteq X \times X \text{ transitive}} S.$$  

For a finite set $X$, the transitive closure is the set of all $(x, y) \in X \times X$ for which there exists an $N \in \mathbb{N}$ and a sequence $(x_0, x_1), \ldots, (x_{N-1}, x_N) \in R$ with $x_0 = x$ and $x_N = y$.

- The *equivalence relation generated by* $R$ (or *equivalence closure*) is the equivalence relation

$$R^\equiv := \bigcap_{R \subseteq S \subseteq X \times X \text{ an equivalence}} S.$$  

It can be shown that $R^\equiv = ((R^\Delta)^\sigma)^+$.

### 2.2 Decorated Endpoints

Here we introduce a convenient notation that will allow us to denote all types of intervals of the real line in a unified way.

Let $D = \{-, +\}$ and define the set of *decorated endpoints* $E$ by

$$E = (\mathbb{R} \times D) \cup \{-\infty, \infty\}. $$
For each $t \in \mathbb{R}$, we denote the decorated endpoints $(t, -)$ and $(t, +)$ by $t^-$ and $t^+$, respectively. We define a total order on $E$ by declaring that $- < +$, taking the lexicographic ordering on $\mathbb{R} \times D$, and declaring $-\infty$ and $\infty$ to be the minimal and maximal elements, respectively, of $E$.

Let $\mathcal{I}_\mathbb{R}$ denote the collection of (non-empty) intervals of the real line. Then there is a bijection between the set $\{(b, d) \in E \times E \mid b < d\}$ and $\mathcal{I}_\mathbb{R}$ which is specified by the following table:

<table>
<thead>
<tr>
<th></th>
<th>$t^-$</th>
<th>$t^+$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-\infty$</td>
<td>$(-\infty, t)$</td>
<td>$(-\infty, t)$</td>
<td>$(-\infty, \infty)$</td>
</tr>
<tr>
<td>$s^-$</td>
<td>$[s, t)$</td>
<td>$[s, t)$</td>
<td>$[s, \infty)$</td>
</tr>
<tr>
<td>$s^+$</td>
<td>$(s, t)$</td>
<td>$(s, t]$</td>
<td>$(s, \infty)$</td>
</tr>
</tbody>
</table>

For $b, d \in E$ with $b < d$, we write $\langle b, d \rangle$ to denote the corresponding interval.

### 2.3 Topics from Linear Algebra

#### 2.3.1 Quotient Spaces

Let $V$ be a vector space over a field $F$ and let $W \subseteq V$ be a subspace. We define a relation $R_W$ on $V$ by declaring that $(v, v') \in R_W$ if and only if $v - v' \in W$. Using linearity of the subspace $W$, it is easy to verify that $R_W$ defines an equivalence relation on $V$. The equivalence class of $v \in V$ under this equivalence relation will be denoted by $[v]_W$. The quotient set $V/W$ of equivalence classes can be given a vector space structure as follows: we define $[v_1]_W + [v_2]_W := [v_1 + v_2]_W$ for all $v_1, v_2 \in V$ and $\alpha[v]_W = [\alpha v]_W$ for all $v \in V$ and $\alpha \in F$. Note that if $[v_1]_W = [v'_1]_W$ and $[v_2]_W = [v'_2]_W$ then $v_1 - v'_1, v_2 - v'_2 \in W$ so that $(v_1 + v_2) - (v'_1 + v'_2) \in W$ and hence $[v_1 + v_2]_W = [v'_1 + v'_2]_W$. Similarly, if $[v]_W = [v']_W$ then $v - v' \in W$ so that by linearity
of the subspace $W$, $\alpha(v - v') = \alpha v - \alpha v' \in W$ and hence $\alpha[v]_W = \alpha[v']_W$. Thus our definitions of addition of equivalence classes and of multiplication of equivalence classes by scalars are well-defined. The operations make $V/W$ into a vector space which we call the \textit{quotient (vector) space} of $V$ by $W$.

### 2.3.2 Induced Linear Maps

Suppose we have a linear map $f : V_1 \to V_2$ between vector spaces $V_1$ and $V_2$. Suppose also that $W_1 \leq V_1$ and $W_2 \leq V_2$ are subspaces satisfying $f(W_1) \subseteq W_2$. Then $f$ induces a map

$$
\overline{f} : V_1/W_1 \to V_2/W_2, \quad [v]_{W_1} \mapsto [f(v)]_{W_2}.
$$

We verify that this map is a well-defined linear map: suppose that $[v]_{W_1} = [v']_{W_1}$ and recall that this is true if and only if $v - v' \in W_1$. Then $f(v - v') = f(v) - f(v') \in f(W_1) \subseteq W_2$ so that $\overline{f}([v]_{W_1}) = [f(v)]_{W_2} = [f(v')]_{W_2} = \overline{f}([v']_{W_1})$ and thus $\overline{f}$ is well-defined. Linearity of $\overline{f}$ follows in a straightforward way from the linearity of $f$. The map $\overline{f}$ defined in this way is called the \textit{linear map induced by $f$} between the quotient spaces $V_1/W_1$ and $V_2/W_2$.

**Proposition 2.3.1.** Let $V_1, V_2, V_3$ be vector spaces with subspaces $W_1 \leq V_1$, $W_2 \leq V_2$, $W_3 \leq V_3$ and let $f : V_1 \to V_2$ and $g : V_2 \to V_3$ be linear maps such that $f(W_1) \subseteq W_2$ and $g(W_2) \subseteq W_3$. Then $(g \circ f)(W_1) \subseteq W_3$ and the induced linear maps $\overline{f} : V_1/W_1 \to V_2/W_2$, $\overline{g} : V_2/W_2 \to V_3/W_3$, and $\overline{g} \circ \overline{f} : V_1/W_1 \to V_3/W_3$ satisfy $\overline{g} \circ \overline{f} = \overline{g \circ f}$.

**Proof.** Since $f(W_1) \subseteq W_2$ and $g(W_2) \subseteq W_3$ we have $(g \circ f)(W_3) = g(f(W_1)) \subseteq g(W_2) \subseteq W_3$. Now for any $[v]_{W_1} \in V_1/W_1$ we have

$$
\overline{g} \circ \overline{f}([v]_{W_1}) = [g(f(v))]_{W_3} = \overline{g}([f(v)]_{W_2}) = \overline{g}(\overline{f}([v]_{W_1})) = (\overline{g} \circ \overline{f})([v]_{W_1}),
$$

11
so that $g \circ f = \overline{g} \circ \overline{f}$.

2.3.3 Kernels and Cokernels

Let $f : V \to W$ be a linear map between vectors spaces $V$ and $W$. The kernel of $f$, denoted ker($f$), is the subspace of $V$ defined by

$$\text{ker}(f) := \{v \in V \mid f(v) = 0\}.$$  

The cokernel of $f$, denoted coker($f$), is defined to be the quotient space

$$\text{coker}(f) := W/\text{im}(f).$$

2.3.4 Free Vector Spaces

Fix a field $F$ and a finite set $X$. Define

$$F(X) := \{f : X \to F \mid |f^{-1}(F\backslash\{0\})| < \infty\},$$

the set of all functions from $X$ to $F$ which take nonzero values on only finitely many elements of $X$ (such functions are said to have finite support). If $f, g \in F(X)$ then it is clear that $f + g \in F(X)$. For each $\gamma \in F$ and $f \in F(X)$ we define the scalar multiple of $f$ by $\gamma$ to be the function $\gamma f : X \to F$ given by $(\gamma f)(x) = \gamma(f(x))$ for all $x \in X$. It is easy to see that we also have $\gamma f \in F(X)$ whenever $f \in F(X)$. With these operations of addition and scalar multiplication, $F(X)$ becomes a vector space called the free vector space of $X$ over $F$. The zero element of $F(X)$ is simply the zero function.

For each $x_0 \in X$ define $\delta_{x_0} : X \to F$ by setting

$$\delta_{x_0}(x) = \begin{cases} 1 & x = x_0 \\ 0 & x \neq x_0. \end{cases}$$
Clearly $\delta_{x_0} \in F(X)$ for all $x_0 \in X$. Moreover, if $f \in F(X)$ then we can write

$$f = \sum_{x \in f^{-1}(F\backslash\{0\})} f(x)\delta_x.$$  

That is, every element of $F$ can be written as a finite linear combination of functions from the set $\Delta(X) = \{\delta_x \in F(X) \mid x \in X\}$. Moreover, the elements of $\Delta$ are linearly independent in the sense that for any finite linear combination

$$f = \sum_k \alpha_k \delta_{x_k},$$

$f = 0$ implies $\alpha_k = 0$ for all $k$. Thus $\Delta(X)$ forms a basis for $F(X)$.

There is an obvious bijection between $X$ and $\Delta(X)$ given by $x \mapsto \delta_x$. It follows that $\dim(F(X)) = |X|$. Moreover, we will identify each $\delta_x$ with $x$. Thus elements of $F(X)$ will be written as finite formal sums

$$\sum_k \alpha_k x_k$$

of elements in $X$ with coefficients in $F$.

### 2.4 Basic Category Theory

#### 2.4.1 Categories

A category $C$ consists of a collection of objects, denoted $\text{Ob}(C)$, a collection of morphisms, denoted $\text{Hom}(C)$, two functions $\text{dom}_C = \text{dom} : \text{Hom}(C) \to \text{Ob}(C)$ and $\text{cod}_C = \text{cod} : \text{Hom}(C) \to \text{Ob}(C)$ associating to each morphism a domain and codomain, respectively, and a function $\text{id} : \text{Ob}(C) \to \text{Hom}(C)$ associating to each object an identity morphism, satisfying the following axioms:

1. If $f, g \in \text{Hom}(C)$ are such that $\text{cod}(f) = \text{dom}(g)$, then there exists a morphism $g \circ f \in \text{Hom}(C)$ with $\text{dom}(g \circ f) = \text{dom}(f)$ and $\text{cod}(g \circ f) = \text{cod}(g)$,
2. If $f, g, h \in \text{Hom}(\mathcal{C})$ with $\text{cod}(f) = \text{dom}(g)$ and $\text{cod}(g) = \text{dom}(h)$ then $(h \circ g) \circ f = h \circ (g \circ f)$,

3. For each $X \in \mathcal{C}$, $\text{dom}(\text{id}(X)) = X = \text{cod}(\text{id}(X))$,

4. For any $f \in \text{Hom}(\mathcal{C})$ with $\text{dom}(f) = X$ and $\text{cod}(f) = Y$, we have $f \circ \text{id}(X) = f = \text{id}(Y) \circ f$.

We use the notation $f : X \to Y$ to denote a morphism $f \in \text{Hom}(\mathcal{C})$ with $\text{dom}(f) = X$ and $\text{cod}(f) = Y$. We visualize a morphism $f : X \to Y$ pictorially as an arrow pointing from $X$ to $Y$:

$$X \xrightarrow{f} Y$$

We denote the collection of all morphisms with domain $X$ and codomain $Y$ by $\text{Hom}_\mathcal{C}(X, Y) := \text{Hom}(X, Y)$. For each $X \in \text{Ob}(\mathcal{C})$ we denote the identity morphism of $X$ by $\text{id}_X := \text{id}(X)$.

### 2.4.2 Functors

Let $\mathcal{C}, \mathcal{D}$ be categories. A functor $F$ from $\mathcal{C}$ to $\mathcal{D}$, denoted $F : \mathcal{C} \to \mathcal{D}$, is a pair of mappings $F : \text{Ob}(\mathcal{C}) \to \text{Ob}(\mathcal{D})$ and $F : \text{Hom}(\mathcal{C}) \to \text{Hom}(\mathcal{D})$ (note that both mappings are denoted by $F$) such that

1. For any $X, Y \in \mathcal{C}$ and for any morphism $f : X \to Y$, we have $F(f) : F(X) \to F(Y)$,

2. For any $X \in \mathcal{C}$, $F(\text{id}_X) = \text{id}_{F(X)}$,

3. For any $X, Y, Z \in \mathcal{C}$ and for any morphisms $f : X \to Y$ and $g : Y \to Z$, we have $F(g \circ f) = F(g) \circ F(f)$.
For any pair of objects \( X, Y \in \text{Ob}(\mathcal{C}) \), the functor \( F \) induces a map \( F_{X,Y} : \text{Hom}_\mathcal{C}(X, Y) \to \text{Hom}_\mathcal{D}(F(X), F(Y)) \) given by \( f \mapsto F(f) \). \( F \) is said to be faithful when \( F_{X,Y} \) is injective for all \( X, Y \), full when \( F_{X,Y} \) is surjective for all \( X, Y \), and fully faithful when \( F_{X,Y} \) is bijective for all \( X, Y \) in \( \mathcal{C} \).

### 2.4.3 Natural Transformations

Let \( F \) and \( G \) be functors between the categories \( \mathcal{C} \) and \( \mathcal{D} \). A natural transformation from \( F \) to \( G \) is a family of morphisms \( \eta := \{ \eta_X : F(X) \to G(X) \}_{X \in \text{Ob}(\mathcal{C})} \subset \text{Hom}(\mathcal{D}) \), indexed by the objects of \( \mathcal{C} \), such that the diagram

\[
\begin{array}{ccc}
F(X) & \xrightarrow{F(f)} & F(Y) \\
\downarrow \eta_X & & \downarrow \eta_Y \\
G(X) & \xrightarrow{G(f)} & G(Y)
\end{array}
\]

commutes for all objects \( X, Y \in \text{Ob}(\mathcal{C}) \) and for any morphism \( f : X \to Y \). The morphisms \( \eta_X \) are called the components of the natural transformation.

### 2.4.4 Some Categories and Constructions

We list here the categories and constructions of categories that will be used throughout this thesis:

1. **The Category of Vector Spaces.** Fix a field \( F \). The collection of all vector spaces over \( F \) forms a category whose morphisms are linear transformations. We denote this category by \( \text{Vec}(F) = \textbf{Vec} \). We denote by \( \text{vec}(F) = \textbf{vec} \) the full subcategory of \( \textbf{Vec} \) whose objects are finite-dimensional vector spaces.

2. **Poset Categories.** We may view a partially ordered set \((X, P)\) as a category whose objects are elements of \( X \) and whose morphisms are pairs \((x, y) \in P\).
The morphism \((x, y)\) is assigned domain \(x\) and codomain \(y\). The composition of \((y, z) \circ (x, y)\) is defined to be \((x, z)\) and exists by transitivity of the partial order. The identity morphism of \(x\) is \((x, x)\) and exists by reflexivity of the partial order.

3. **Functor Categories.** Let \(C\) and \(D\) be categories. We define a new category \(D^C\) whose objects are functors from \(C\) to \(D\) and whose morphisms are natural transformations. The composition of natural transformations is defined by component-wise by composition of components.

4. **The Opposite Category.** Let \(C\) be any category. We define the *opposite category* of \(C\), denoted \(C^{\text{op}}\), to be the category whose objects and morphisms are the same as those of \(C\) but whose domain and codomain functions are interchanged. That is, \(\text{Ob}(C^{\text{op}}) := \text{Ob}(C)\) and \(\text{Hom}(C^{\text{op}}) := \text{Hom}(C)\) but \(\text{dom}_{C^{\text{op}}} := \text{cod}_C\) and \(\text{cod}_{C^{\text{op}}} := \text{dom}_C\).

### 2.4.5 Types of Morphisms

Let \(f : X \to Y\) be a morphism between \(X, Y \in \text{Ob}(C)\). We call \(f\) a *monomorphism* if for any object \(W \in \text{Ob}(C)\) and for any pair of morphisms \(g_1, g_2 : W \to X\) with \(f \circ g_1 = f \circ g_2\), we have \(g_1 = g_2\).

Dually, the morphism \(f\) is said to be an *epimorphism* if for any object \(Z \in \text{Ob}(C)\) and for any pair of morphisms \(h_1, h_2 : Y \to Z\) with \(h_1 \circ f = h_2 \circ f\), we have \(h_1 = h_2\).

Monomorphisms and epimorphisms are dual in the sense that \(f \in \text{Hom}(C)\) is a monomorphisms if and only if \(f\) is an epimorphism in the opposite category \(C^{\text{op}}\).

The morphism \(f\) is called an *isomorphism* if there exists a morphism \(g : Y \to X\) such that \(f \circ g = \text{id}_Y\) and \(g \circ f = \text{id}_X\). Note that an isomorphism is always both a
monomorphism and an epimorphism, but the converse is not true in general. However, in many of the categories we will work with such as \textbf{Set} and \textbf{Vec}, morphisms which are both monomorphisms and epimorphisms are indeed isomorphisms.

2.4.6 Limits and Colimits

Fix categories $\mathcal{J}$ and $\mathcal{C}$. A \textit{diagram of shape} $\mathcal{J}$ is a functor $D : \mathcal{J} \to \mathcal{C}$. This definition is merely a change of terminology to emphasize a change of perspective; when we refer to a functor $D : \mathcal{J} \to \mathcal{C}$ as a diagram, we are viewing $\mathcal{J}$ as an \textit{index category} indexing some subsets of the objects and arrows of $\mathcal{C}$.

A \textit{cone} over a diagram $D : \mathcal{J} \to \mathcal{C}$ is an object $N \in \text{Ob}(\mathcal{C})$ together with a collection of morphisms $\lambda := \{\lambda_J : N \to D(J) \mid J \in \text{Ob}(\mathcal{J})\}$, indexed by the objects of $\mathcal{J}$, such that for any morphism $f : A \to B$ in $\text{Hom}(\mathcal{J})$ we have $\lambda_B = D(f) \circ \lambda_A$. We denote a cone by $(N, \lambda)$.

The \textit{limit} of the diagram $D$ is a cone $(L, \phi)$ over $D$ such that for any other cone $(N, \lambda)$ over $D$, there exists a unique morphism $\psi : N \to L$ in $\text{Hom}(\mathcal{C})$ with $\lambda_A = \phi_A \circ \psi$ for all $A \in \text{Ob}(\mathcal{J})$.

In general, the limit of a diagram may not exist. A category $\mathcal{C}$ is said to \textit{have all $\mathcal{J}$-limits} if every diagram $D : \mathcal{J} \to \mathcal{C}$ has a limit.

A \textit{cocone} over $D$ is an object $N \in \text{Ob}(\mathcal{C})$ together with a collection of morphisms $\lambda := \{\lambda_J : D(J) \to N \mid J \in \text{Ob}(\mathcal{J})\}$, indexed by the objects of $\mathcal{J}$, such that for any
morphism $f : A \to B$ in $\text{Hom}(\mathcal{J})$ we have $\lambda_B = D(f) \circ \lambda_A$. We denote a cocone by $(N, \lambda)$.

The **colimit** of the diagram $D$ is a cocone $(C, \phi)$ over $D$ such that for any cocone $(N, \lambda)$ over $D$, there exists a unique morphism $\psi : C \to N$ in $\text{Hom}(\mathcal{C})$ with $\lambda_A = \psi \circ \phi_A$ for all $A \in \text{Ob}(\mathcal{J})$.

![Diagram of colimit](image)

As is the case with limits, the colimit of a diagram may not exist. A category $\mathcal{C}$ is said to **have all $\mathcal{J}$-colimits** if every diagram $D : \mathcal{J} \to \mathcal{C}$ has a colimit.

**Theorem 2.4.1.** Let $\mathcal{J}$ and $\mathcal{C}$ be categories where $\mathcal{C}$ contains all $\mathcal{J}$-limits, let $D^1, D^2 \in \mathcal{C}^\mathcal{J}$, and let $L^1, L^2 \in \mathcal{C}$ denote the limits of $D^1$, $D^2$, respectively. If there exists a natural transformation $\eta : D^1 \to D^2$ all of whose components are monomorphisms, then $(L^1, \eta_J \circ \lambda^1_J)$ is a cone for $D^2$ and the unique morphism $\psi : L^1 \to L^2$ satisfying $\eta_J \circ \lambda^1_J = \lambda^2_J \circ \psi$ for all $J \in \mathcal{J}$ is a monomorphism.

**Proof.**
Let $W \in C$ and let $g_1, g_2 : W \to L^1$ be parallel morphisms such that $\psi \circ g_1 = \psi \circ g_2$. We will show that we necessarily have $g_1 = g_2$ which in turn will show that $\psi$ is a monomorphism.

Note that $(W, \lambda_I \circ g_1)$ and $(W, \lambda_I \circ g_2)$ are both cones over $D^1$, and evidently $g_1$ and $g_2$ are the respective unique morphisms for which these cones factor through $(L, \lambda_I)$. By commutativity of the diagram above, we have $\eta_I \circ \lambda_I \circ g_1 = \lambda_I^2 \circ (\psi \circ g_1)$ and $\eta_I \circ \lambda_I \circ g_2 = \lambda_I^2 \circ (\psi \circ g_2)$ for all $I \in \mathcal{J}$. Since $\psi \circ g_1 = \psi \circ g_2$, we in fact have $\eta_I \circ \lambda_I \circ g_1 = \eta_I \circ \lambda_I \circ g_2$ for all $I \in \mathcal{J}$. Since each $\eta_I$ is assumed to be a monomorphism, we have $\lambda_I \circ g_1 = \lambda_I \circ g_2$ for all $I \in \mathcal{J}$. Thus $(W, \lambda_I \circ g_1)$ and $(W, \lambda_I \circ g_2)$ define the same cone over $D^1$, and so we must in fact have $g_1 = g_2$. □

2.4.7 Products and Coproducts

Let $C$ be a category and let $X, Y \in \text{Ob}(C)$. The product of $X$ and $Y$, denoted $X \times Y$ if it exists, is an object in $C$ together with morphisms $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y$ satisfying the following universal property: for any object $N \in \text{Ob}(C)$ for which there exists morphisms $f : N \to X$ and $g : N \to Y$, there exists a unique morphism $\psi$ such that the diagram

$$
\begin{array}{ccc}
N & \xrightarrow{\exists \psi} & Y \\
\downarrow{g} & & \downarrow{\pi_2} \\
X \times Y & \xleftarrow{\pi_1} & X
\end{array}
$$

commutes.

The coproduct of $X$ and $Y$, denoted $X \coprod Y$ if it exists, is an object in $C$ together with morphisms $p_1 : X \to X \coprod Y$ and $p_2 : Y \to X \coprod Y$ such that for any object $N \in \text{Ob}(C)$ for which there exists morphisms $f : X \to X \coprod Y$ and $g : Y \to X \coprod Y$, there exists
a unique morphisms $\psi : X \coprod Y \to N$ such that the diagram

\[
\begin{array}{c}
\xymatrix{
X \ar[r]^{p_1} & X \coprod Y \ar[r]^{p_2} & Y \\
N \ar[ru]^{\exists \psi} \ar[ru]^{g} \ar[ru]_{f}
}
\end{array}
\]

commutes. The coproduct is dual to the product, i.e. the coproduct of $X$ in $Y$ taken
in $\mathcal{C}$ is equal to the product of $X$ and $Y$ taken in $\mathcal{C}^{op}$.

Now suppose that $f : X_1 \to Y_1$ and $g : X_2 \to Y_2$ are morphisms in $\text{Hom}(\mathcal{C})$ and
suppose that the products $X_1 \times X_2$ and $Y_1 \times Y_2$ exist in $\mathcal{C}$. Then we have the following
diagram:

\[
\begin{array}{c}
\xymatrix{
X_1 \ar[r]^{\pi_1^X} & X_1 \times X_2 \ar[r]^{\pi_2^X} & X_2 \\
Y_1 \ar[r]_{\pi_1^Y} \ar[ru]_{f \circ \pi_1^X} & Y_1 \times Y_2 \ar[r]_{\pi_2^Y} \ar[ru]_{g \circ \pi_2^X} & Y_2
}
\end{array}
\]

where $\pi_1^X, \pi_2^X$ and $\pi_1^Y, \pi_2^Y$ are the canonical projection morphisms for $X_1 \times X_2$ and
$Y_1 \times Y_2$, respectively. We then define the product $f \times g$ to be the unique morphism
$\psi : X_1 \times X_2 \to Y_1 \times Y_2$ such that $f \circ \pi_1^X = \pi_1^Y \circ \psi$ and $g \circ \pi_2^X = \pi_2^Y \circ \psi$.

The next proposition shows that the product of two functors in a functor category exists and is obtained by taking products componentwise, provided the target category contains all products:

**Proposition 2.4.1.** Let $F_1, F_2 \in \mathcal{D}^\mathcal{C}$ be functors and suppose that the product exists
between any two objects in $\mathcal{D}$. Then the product $F_1 \times F_2$ exists in $\mathcal{D}^\mathcal{C}$ and is defined
on objects by $(F_1 \times F_2)(X) = F_1(X) \times F_2(X)$ for all $X \in \text{Ob}(\mathcal{C})$ and on morphisms
$(F_1 \times F_2)(f) = F_1(f) \times F_2(f)$ for all $f \in \text{Hom}(\mathcal{C})$. The projection morphisms $\rho^1 : F_1 \times F_2 \to F_1$ and $\rho^1 : F_1 \times F_2 \to F_2$ are the natural transformations defined componentwise
as the projection morphisms for $(F_1 \times F_2)(X)$.

20
Proof. First we check that $F_1 \times F_2$ as defined in the statement of the proposition is a well-defined functor. It is straightforward to verify that $\text{id}_{F_1(X)} \times \text{id}_{F_2(X)} = \text{id}_{F_1(X) \times F_2(X)}$ for all $X \in \text{Ob}(\mathcal{C})$ so that

$$(F_1 \times F_2)(\text{id}_X) = F_1(\text{id}_X) \times F_2(\text{id}_X) = \text{id}_{F_1(X)} \times \text{id}_{F_2(X)}$$

$$= \text{id}_{F_1(X) \times F_2(X)} = \text{id}_{(F_1 \times F_2)(X)}.$$

Next, consider morphisms $f : X \to Y$ and $g : Y \to Z$ in $\text{Hom}(\mathcal{C})$. The commutativity of the diagram

$$
\begin{array}{cccc}
F_1 & \xleftarrow{\pi_1^X} & F_1(X) \times F_2(X) & \xrightarrow{\pi_2^X} & F_2(X) \\
\downarrow{F_1(f)} & & \downarrow{F_1(f) \times F_2(f)} & & \downarrow{F_2(f)} \\
F_1(Y) & \xleftarrow{\pi_1^Y} & F_1(Y) \times F_2(Y) & \xrightarrow{\pi_2^Y} & F_2(Y) \\
\downarrow{F_1(g)} & & \downarrow{F_1(g) \times F_2(g)} & & \downarrow{F_2(g)} \\
F_1(Z) & \xleftarrow{\pi_1^Z} & F_1(Z) \times F_2(Z) & \xrightarrow{\pi_2^Z} & F_2(Z)
\end{array}
$$

implies that $(F_1(g) \circ F_1(f)) \times (F_2(g) \circ F_2(f)) = (F_2(g) \times F_2(g)) \circ (F_1(f) \times F_2(f))$. Hence

$$(F_1 \times F_2)(g \circ f) = F_1(g \circ f) \times F_2(g \circ f)$$

$$= (F_1(g) \circ F_1(f)) \times (F_2(g) \circ F_2(f))$$

$$= (F_1(g) \times F_2(g)) \circ (F_1(f) \times F_2(f)) = (F_1 \times F_2)(g) \circ (F_1 \times F_2)(f).$$

Thus $F_1 \times F_2$ is indeed a functor.

Next, we check that the collections of morphisms $\rho^1 = \{\pi_1^X : (F_1 \times F_2)(X) \to F_1(X) \mid X \in \text{Ob}(\mathcal{C})\}$ and $\rho^2 = \{\pi_2^X : (F_1 \times F_2)(X) \to F_2(X) \mid X \in \text{Ob}(\mathcal{C})\}$ define natural transformations from $F_1 \times F_2$ to $F_1$ and $F_2$, respectively. But this follows immediately from the action of $F_1 \times F_2$ on morphisms: if $f : X \to Y$ is a morphisms
in \text{Hom}(\mathcal{C}) then by definition \((F_1 \times F_2)(f) = F_1(f) \times F_2(f)\) is the unique morphism such that

\[
F_1(f) \circ \rho^1_X = F_1(f) \circ \pi^X_1 \\
= \pi^Y_1 \circ (F_1 \times F_2)(f) = \rho^1_Y \circ (F_1 \times F_2)(f)
\]

so that \(\rho^1\) is a natural transformation from \(F_1 \times F_2\) to \(F_1\). Symmetrically, we have \(F_2(f) \circ \rho^2_X = \rho^2_Y \circ (F_1 \times F_2)(f)\) so that \(\rho^2\) is a natural transformation from \(F_1 \times F_2\) to \(F_2\).

Finally, we check that \(F_1 \times F_2\) together with the natural transformations \(\rho^1\) and \(\rho^2\) indeed satisfy the universal property of products. Let \(F \in \mathcal{D}\) and let \(\eta^1 : F \to F_1\) and \(\eta^2 : F \to F_2\) be natural transformations. For each \(X \in \text{Ob}(\mathcal{C})\), let \(\psi_X : F(X) \to (F_1 \times F_2)(X) = F_1(X) \times F_2(X)\) be the unique morphisms such that \(\eta^1_X = \rho^1_X \circ \psi_X\) and \(\eta^2_X = \rho^2_X \circ \psi_X\). We will show that \(\psi = \{\psi_X \mid X \in \text{Ob}(\mathcal{C})\}\) is a natural transformation from \(F\) to \(F_1 \times F_2\) satisfying \(\eta^1 = \rho^1 \circ \psi\) and \(\eta^2 = \rho^2 \circ \psi\), and that \(\psi\) is the unique natural transformation with this property.

Indeed, we want to show that for any objects \(X\) and \(Y\) from \(\mathcal{C}\), the red square in the diagram below commutes, given that all other squares in this diagram commute. Consider the maps \(F_1(f) \circ \eta^1_X : F(X) \to F_1(Y)\) and \(F_2(f) \circ \eta^2_X : F(X) \to F_2(Y)\). We have

\[
F_1(f) \circ \eta^1_X = F_1(f) \circ (\rho^1_X \circ \psi_X) \\
= (F_1(f) \circ \rho^1_X) \circ \psi_X \\
= [\rho^1_Y \circ (F_1 \times F_2)(f)] \circ \psi_X = \rho^1_Y \circ [(F_1 \times F_2)(f) \circ \psi_X]
\]
and, symmetrically, \( F_2(f) \circ \eta^2_X = \rho^2_Y \circ ((F_1 \times F_2)(f) \circ \psi_X) \). On the other hand, we have

\[
F_1(f) \circ \eta^1_X = \eta^1_Y \circ F(f) = (\rho^1_Y \circ \psi_Y) \circ F(f) = \rho^1_Y \circ (\psi_Y \circ F(f))
\]

and, symmetrically, \( F_1(f) \circ \eta^2_X = \rho^2_Y \circ (\psi_Y \circ F(f)) \). By the universal property of products, there is exactly one map \( \phi : F(X) \to (F_1 \times F_2)(Y) \) satisfying both \( F_1(f) \circ \eta^1_X = \rho^1_Y \circ \phi \) and \( F_2(f) \circ \eta^2_X = \rho^2_Y \circ \phi \), and so we must have

\[
\phi = (F_1 \times F_2)(f) \circ \psi_X = \psi_Y \circ F(f).
\]

Thus \( \psi \) does in fact define a natural transformation from \( F \) to \( F_1 \times F_2 \). Since \( \rho^1_X \circ \psi_X = \eta^1_X \) and \( \rho^2_X \circ \psi_X = \eta^2_X \) for all \( X \in \text{Ob}(C) \), we have \( \rho^1 \circ \psi = \eta^1 \) and \( \rho^2 \circ \psi = \eta^2 \). Moreover, if \( \xi = \{ \xi_X \mid X \in \text{Ob}(C) \} \) is any other natural transformation satisfying \( \rho^1 \circ \xi = \eta^1 \) and \( \rho^2 \circ \xi = \eta^2 \) then \( \rho^1_X \circ \xi_X = \eta^1_X \) and \( \rho^2_X \circ \xi_X = \eta^2_X \) for all \( X \in \text{Ob}(C) \). By the uniqueness of each \( \psi_X \), we have \( \xi_X = \psi_X \) for all \( X \in \text{Ob}(C) \) and thus \( \xi = \psi \). This completes the proof.

Next we show that the (co)limit of the (co)product of two diagrams is isomorphic to the (co)product of the (co)limit:
Theorem 2.4.2. Let $\mathcal{J}$ and $\mathcal{C}$ be categories where $\mathcal{C}$ contains all products and $\mathcal{J}$-limits, and let $D^1 : \mathcal{J} \to \mathcal{C}$ and $D^2 : \mathcal{J} \to \mathcal{C}$ be functors. Let $D^\times : \mathcal{J} \to \mathcal{C}$ denote the product $D^1 \times D^2$ in the functor category $\mathcal{C}^\mathcal{J}$. If $L^1$, $L^2$, and $L^\times$ denote the limits of $D^1$, $D^2$, and $D^\times$, respectively, then $L^\times \cong L^1 \times L^2$. Dually, if $C^1$, $C^2$, and $C^\times$ denote the colimits of $D^1$, $D^2$, and $D^\times$, respectively, then $C^\times \cong C^1 \coprod C^2$.

Proof.

For each $J \in \mathcal{J}$, let $p^1_J : D^\times(J) \to D^1(J)$ and $p^2_J : D^\times(J) \to D^2(J)$ denote the canonical projection morphisms. Note that $(L^\times, p^1_J \circ \lambda^\times_J)$ is a cone for $D^1$ so that there exists a unique morphism $p^1_L : L^\times \to L^1$ satisfying $\lambda^1_J \circ p^1_L = p^1_J \circ \lambda^\times_J$ for all $J \in \mathcal{J}$. Similarly, $(L^\times, p^2_J \circ \lambda^\times_J)$ is a cone for $D^2$ so that there exists a unique morphism $p^2_L : L^\times \to L^2$ satisfying $\lambda^2_J \circ p^2_L = p^2_J \circ \lambda^\times_J$ for all $J \in \mathcal{J}$.

Now if $W \in \mathcal{C}$ and $\omega^1 : W \to L^1$ and $\omega^2 : W \to L^2$ then $\lambda^1_J \circ \omega^1 : W \to D^1(J)$ and $\lambda^2_J \circ \omega^2 : W \to D^2(J)$ for all $J \in \mathcal{J}$. By the universality of $D^\times(J)$, for each $J \in \mathcal{J}$ there exists a unique morphism $\alpha_J : W \to D^\times(J)$ such that $\lambda^1_J \circ \omega^1 = p^1_J \circ \alpha_J$ and $\lambda^2_J \circ \omega^2 = p^2_J \circ \alpha_J$.

Next, we note that $(W, \alpha_J)$ is a cone for $D^\times$ so there exists a unique morphism $\psi : W \to L^\times$ such that $\lambda^\times_J \circ \psi = \alpha_J$ for all $J \in \mathcal{J}$.
It can be checked that \((W, p^1_J \circ \alpha_J)\) is a cone for \(D^1\) which implies that \(\omega^1\) is the unique morphism satisfying \(\lambda^1_J \circ \omega^1 = p^1_J \circ \alpha_J\). Similarly, \((W, p^2_J \circ \alpha_J)\) is a cone for \(D^2\) which implies that \(\omega^2\) is the unique morphism satisfying \(\lambda^2_J \circ \omega^2 = p^2_J \circ \alpha_J\).

Then \(p^1_J \circ \alpha_J = p^1_J \circ (\lambda^x_J \circ \psi) = (p^1_J \circ \lambda^x_J) \circ \psi = (\lambda^1_J \circ p^1_L) \circ \psi = \lambda^1_J \circ (p^1_L \circ \psi)\). By uniqueness of \(\omega^1\), we have \(p^1_J \circ \psi = \omega^1\). Similarly we have \(p^2_J \circ \psi = \omega^2\).

Now suppose that \(\psi' : W \to L^x\) satisfies \(p^1_L \circ \psi' = \omega^1\) and \(p^2_L \circ \psi' = \omega^2\). Then \(\lambda^1_J \circ \omega^1 = \lambda^1_J \circ (p^1_L \circ \psi') = (\lambda^1_J \circ p^1_L) \circ \psi' = (p^1_J \circ \lambda^x_J) \circ \psi' = p^1_J \circ (\lambda^x_J \circ \psi')\) for all \(J \in \mathcal{J}\). Similarly, we have \(\lambda^2_J \circ \omega^2 = \lambda^2_J \circ (\lambda^x_J \circ \psi')\) for all \(J \in \mathcal{J}\). By the uniqueness of \(\alpha_J\), we have \(\lambda^x_J \circ \psi' = \alpha_J\) for all \(J \in \mathcal{J}\). By the uniqueness of \(\psi\) as the morphism by which \((W, \alpha_J)\) factors through \((L^x, \lambda^x_J)\), we have \(\psi' = \psi\).

Thus we have shown that the object \(L^x\) and morphisms \(p^1_L : L^x \to L^1\) and \(p^2_L : L^x \to L^2\) have the property that for any \(W \in \mathcal{C}\) and for any morphisms \(\omega^1 : W \to L^1\) and \(\omega^2 : W \to L^2\), there exists a unique morphism \(\psi : W \to L^x\) such that \(\omega_1 = p^1_L \circ \psi\) and \(\omega_2 = p^2_L \circ \psi\). This is exactly the universal property for products, and so we have \(L^x \cong L^1 \times L^2\).

To prove the dual statement, consider the opposite functors \((D^1)^{\text{op}} : \mathcal{J}^{\text{op}} \to \mathcal{C}^{\text{op}}\), \((D^2)^{\text{op}} : \mathcal{J}^{\text{op}} \to \mathcal{C}^{\text{op}}\), and \((D^x)^{\text{op}} : \mathcal{J}^{\text{op}} \to \mathcal{C}^{\text{op}}\). Then \(C^1\), \(C^2\), and \(C^x\) are the limits of \((D^1)^{\text{op}}\), \((D^2)^{\text{op}}\), and \((D^x)^{\text{op}}\), respectively. By the proceeding discussion, \(C^x \cong C^1 \times C^2\), where this product is taken in \(\mathcal{C}^{\text{op}}\). It follows that \(C^x \cong C^1 \coprod C^2\), this coproduct being taken in \(\mathcal{C}\), as claimed.

\(\Box\)

Remark 2.4.1. In the category \textbf{Vec} of vector spaces over a given field, the product and coproduct coincide and are just the direct sum of vector spaces. Thus, the limit or colimit of the direct sum of two diagrams is isomorphic to the direct sum of the limits or colimits, respectively, of the two diagrams.
2.4.8 Pullbacks and Pushouts

Let \( \mathcal{C} \) be any category and consider a diagram of the form

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y & \leftarrow \xleftarrow{g} Z
\end{array}
\] (2.1)

The pullback of such a diagram is an object \( P \in \text{Ob}(\mathcal{C}) \) together with two morphisms \( p_1 : P \to X \) and \( p_2 : P \to Z \) such that \( f \circ p_1 = g \circ p_2 \), and such that for any object \( Q \in \text{Ob}(\mathcal{C}) \) and morphisms \( q_1 : Q \to X \) and \( q_2 : Q \to Z \) with \( f \circ q_1 = f \circ q_2 \), there exists a unique morphism \( \psi : Q \to P \) such that \( q_1 = p_1 \circ \psi \) and \( q_2 = p_2 \circ \psi \). Pictorially, the following diagram commutes:

\[
\begin{array}{ccc}
Q & \xrightarrow{q_2} & Z \\
\phantom{Q} & \xrightarrow{\exists \psi} & \\
\phantom{Q} & \xleftarrow{p_1} & P \\
\phantom{Q} & \downarrow & \phantom{p_2} \downarrow & \phantom{g} \downarrow \phantom{Z} \\
X & \xrightarrow{f} & Y & \xleftarrow{g} Z
\end{array}
\]

Note that the pullback is essentially the same as the limit of diagram 2.1 except that we forget the morphism from \( P \) to \( Y \). However, there is no loss since commutativity completely determines this map. In other words, the pullback contains the same information as the limit of diagram 2.1.

Dually, the pushout of a diagram of the form

\[
\begin{array}{ccc}
X & \xleftarrow{f} & Y & \xrightarrow{g} Z
\end{array}
\] (2.2)

is an object \( P \in \text{Ob}(\mathcal{C}) \) together with two morphisms \( p_1 : X \to P \) and \( p_2 : Z \to P \) such that \( p_1 \circ f = p_2 \circ g \), and such that for any object \( Q \in \text{Ob}(\mathcal{D}) \) and morphisms \( q_1 : X \to Q \) and \( q_2 : Q \to Z \) with \( q_1 \circ f = q_2 \circ g \), there exists a unique morphism
\( \psi : P \to Q \) such that \( q_1 = \psi \circ p_1 \) and \( q_2 = \psi \circ p_1 \). Pictorially, the following diagram commutes:

The pushout is essentially the same as the colimit of diagram 2.2 in the sense that the two notions provide the same information.
Chapter 3: Zigzag Persistence

We fix a field $F$. All vector spaces throughout will be finite dimensional over $F$.

3.1 Zigzag Modules

A zigzag module $V = (V_i, p_i)$ is a finite sequence

$$V_1 \leftarrow p_1 \rightarrow V_2 \leftarrow p_2 \rightarrow \cdots \leftarrow p_{n-2} \rightarrow V_{n-1} \leftarrow p_{n-1} \rightarrow V_n$$

(3.1)

of vector spaces and linear transformations between them. An arrow $V_i \xleftarrow{p_i} V_{i+1}$ represents either a forward linear map $V_i \xrightarrow{p_i} V_{i+1}$ or a backward linear map $V_i \xleftarrow{p_i} V_{i+1}$, but never both. The $p_i$ are referred to as structure maps. Note that the $i$-th structure map $p_i$ either has domain $V_i$ or $V_{i+1}$ and codomain $V_{i+1}$ or $V_i$ respectively.

We will use the notation $p_i : V_{i_1} \to V_{i_2}$ when the direction of $p_i$ has not been specified. In other words, $i_1, i_2 \in \{i, i + 1\}$ with $i_1 \neq i_2$, $\text{dom}(p_i) = V_{i_1}$, and $\text{cod}(p_i) = V_{i_2}$.

The length of a zigzag module is the length of the sequence (3.1) above. We will denote the collection of all zigzag modules of length $n$ by $n$-Mod. A finite sequence of the symbols $\rightarrow$ and $\leftarrow$, indicating the directions of the linear maps in (3.1) as read from left to right, is called the type $\tau$ of the zigzag module. Formally, the type of a zigzag module of length $n$ is a sequence $\tau \in \{\rightarrow, \leftarrow\}^{n-1}$. We will use the notation $\mathcal{T}_n := \{\rightarrow, \leftarrow\}^{n-1}$. 

28
Remark 3.1.1. Sequences in $T_n$ have length $n - 1$, not length $n$. This is so that zigzag modules in $n$-Mod have types in $T_n$.

We define a map

$$\text{type} := \text{type}_n : n\text{-Mod} \to T_n$$

by setting $\text{type}(V)$ to be the type of $V$. We denote the collection of all zigzag modules of type $\tau$ by $\text{Mod}_\tau$, that is

$$\text{Mod}_\tau := \{V \in n\text{-Mod} \mid \text{type}(V) = \tau\}.$$ 

Note that $\text{Mod}_\tau = \text{type}^{-1}(\tau)$ and $n\text{-Mod} = \bigcup_{\tau \in T_n} \text{Mod}_\tau$. A zigzag module of type $\tau$ is also called a $\tau$-module.

Example 3.1.1. Let $n = 3$. Then

$$T_n = \{(\rightarrow, \rightarrow), (\rightarrow, \leftarrow), (\leftarrow, \rightarrow), (\leftarrow, \leftarrow)\}.$$ 

Consider the zigzag modules

$$V_1 = F \xrightarrow{id} F \xrightarrow{id} F \quad V_2 = F \xrightarrow{id} F \xleftarrow{0} 0$$

$$V_3 = F \xleftarrow{0} 0 \xrightarrow{0} F \quad V_4 = 0 \xleftarrow{0} F \xrightarrow{id} F$$

where $F$ is viewed as a 1-dimensional vector space over itself and $0$ denotes the trivial vector space. Each of these zigzag modules is an element of $n$-Mod and we have

$\text{type}(V_j) = \tau_j$ for $j = 1, 2, 3, 4$.

3.1.1 Morphisms Between Zigzag Modules

Fix zigzag modules $V = (V_i, p_i)$ and $W = (W_i, q_i)$ in $\text{Mod}_\tau$. A morphism from $V$ to $W$ is a collection $\phi = \{\phi_i : V_i \to W_i\}_{i=1}^n$ of linear transformations such that the
commutes. We denote a morphism from $V$ to $W$ by $\phi : V \to W$. The linear maps $\phi_i$ comprising the morphism $\phi$ are called the *components* of $\phi$. Composition of morphisms is defined component-wise and the identity morphism $\text{id}_V : V \to V$ is the morphism all of whose components are identity maps. With these definitions in place, for every $n \in \mathbb{N}$ and for each $\tau \in \mathcal{T}_n$, the collection $\text{Mod}_\tau$ of $\tau$-modules together with the collection of all morphisms between them forms a category denoted $\text{Mod}_\tau$. We call a morphism $\phi$ an *isomorphism*, *monomorphism*, or *epimorphism* if all of the $\phi_i$ are either bijective, injective, or surjective, respectively. If $\phi : V \to W$ is an isomorphism we say that $V$ and $W$ are *isomorphic* and write $V \cong W$.

Proposition 3.1.1. Let $V \in \text{Mod}_\tau$. A $\tau$-module $W$ is isomorphic to a submodule of $V$ if and only if there exists a monomorphism $f : W \to V$.

Proof. Suppose that $W$ is isomorphic to a submodule of $U$ of $V$ and let $\phi : W \to U$ be any isomorphism. It is not hard to see that the inclusion $\iota : U \hookrightarrow V$, defined component-wise by the inclusions $\iota_i : U_i \hookrightarrow V_i$, is a monomorphism so that the composition $f := \iota \circ \phi : W \hookrightarrow V$ is the desired monomorphism.
Conversely, suppose that \( f : W \hookrightarrow V \) is a monomorphism. Let \( U_k = f_k(W_k) \subseteq V_k \) for all \( k \). Let \( p_i : V_{i_1} \to V_{i_2} \) be a structure map and suppose that \( u \in U_{i_1} \). Then there exists \( w \in W_{i_1} \) such that \( f_{i_1}(w) = u \) and hence \( p_i(u) = p_i(f_{i_1}(w)) = f_{i_2}(q_i(w)) \in f_{i_2}(W_{i_2}) = U_{i_2} \). Thus \( U = (U_i, r_i) \) is a submodule of \( V \), where \( r_i = p_i|_{U_{i_1}} \). Moreover, for any \( w \in W_{i_1} \) we have \( r_i(f_{i_1}(w)) = p_i(f_{i_1}(w)) = f_{i_2}(q_i(w)) \). Thus \( f : W \to U \) is a morphism of \( \tau \)-modules. Since each \( f_i : W_i \to U_i = f_i(W_i) \) is bijective, we see that \( W \cong U \).

The previous proposition justifies us working with monomorphisms between zigzag modules instead of working with submodules directly. We write \( W \leq V \) whenever there exists a monomorphism \( f : W \to V \). Note that \( W \cong V \) if and only if \( W \leq V \) and \( V \leq W \).

### 3.1.2 Interval Modules and the Zero Module

Fix \( n \in \mathbb{N} \) and \( \tau \in \mathcal{T}_n \). For each pair \( b, d \in \{1, \ldots, n\} \) with \( b \leq d \) we define a zigzag module \( \mathbb{I}_{\tau}([b, d]) = (I_i, p_i) \in \text{Mod}_\tau \), called the interval \( \tau \)-module on \([b, d]\), by setting

\[
I_i := \begin{cases} 
F & b \leq i \leq d \\
0 & \text{otherwise}
\end{cases}
\quad \text{and} \quad
p_i := \begin{cases} 
\text{id}_F & b \leq i < d \\
0 & \text{otherwise}
\end{cases}
\]

When the type is clear we will drop the subscript \( \tau \) and just write \( \mathbb{I}([b, d]) \). Interval \( \tau \)-modules of the form \( \mathbb{I}_{\tau}([k, k]) \) are called simple interval \( \tau \)-modules. We also define the zero module \( \mathcal{O}_\tau = (Z_i, z_i) \) of type \( \tau \) to be the \( \tau \)-module with \( Z_i = 0 \) and \( z_i = 0 \) for all \( i \).

**Example 3.1.2.** Let \( n = 3 \) and let \( \tau = (\to, \leftarrow) \in \mathcal{T}_n \). There are \( 6 = \binom{n+1}{2} \) interval \( \tau \)-modules given by
$$I_\tau([1,1]) = F \xrightarrow{0} 0 \xleftarrow{0} 0 \quad I_\tau([1,2]) = F \xrightarrow{id} F \xleftarrow{0} 0$$

$$I_\tau([1,3]) = F \xrightarrow{id} F \xleftarrow{id} F \quad I_\tau([2,2]) = 0 \xrightarrow{0} F \xleftarrow{0} 0$$

$$I_\tau([2,3]) = 0 \xrightarrow{0} F \xleftarrow{id} F \quad I_\tau([3,3]) = 0 \xrightarrow{0} 0 \xleftarrow{0} F$$

### 3.2 Decompositions of Zigzag Modules

The direct sum of two $\tau$-modules $X = (X_i, \alpha_i)$ and $Y = (Y_i, \beta_i)$ is a $\tau$-module $X \oplus Y = (Z, \gamma_i)$ where $Z_i = X_i \oplus Y_i$ and where $\gamma_i = \alpha_i \oplus \beta_i$ for all $i$. We say that $W$ is a summand of $V$ whenever there exists a $\tau$-module $U$ such that $V \cong W \oplus U$ and we write $W \preceq V$. The relation $\preceq$ defines a partial order on $\text{Mod}_\tau$. The $\tau$-module $V$ is said to be decomposable if there exists nonzero $\tau$-modules $W$ and $U$ such that $V \cong W \oplus U$, and is said to be indecomposable otherwise.

**Proposition 3.2.1.** If $(W_i, q_i) = W \preceq V = (V_i, p_i)$ and if the $k$-th structure map $p_k$ of $V$ is bijective, injective, or surjective, then the $k$-th structure map $q_k$ of $W$ is also bijective, injective, or surjective, respectively.

**Proof.** Let $U = (U_i, r_i)$ be such that $V \cong W \oplus U$. Suppose $p_k = q_k \oplus r_k : V_{k_1} \to V_{k_2}$ is injective. If $q_k(v) = 0$ then $p_k(v, 0) = (q_k(v), r_k(0)) = (0, 0)$ so that, by injectivity of $p_k$, $(v, 0) = (0, 0)$ and hence $v = 0$. Thus $\ker(q_k) = 0$ so that $q_k$ is injective.

Now suppose that $p_k$ is surjective. Then in particular, for any $x \in W_{k_2}$ there exists some $(w, u) \in W_{k_1} \oplus U_{k_1}$ such that $p_k(w, u) = (q_k(w), r_k(u)) = (x, 0)$, and hence $w \in W_{k_1}$ satisfies $q_k(w) = x$. Since $x \in W_{k_2}$ was arbitrary, we see that $q_k$ is surjective.

If $p_k$ is bijective then the fact the $q_k$ is bijective follows from the previous two claims. \qed
Theorem 3.2.1 (Krull-Remak-Schmidt, Gabriel [6, 12, 17]). For each \( n \in \mathbb{N} \) and for every \( \tau \in \mathcal{T}_n \), the indecomposable \( \tau \)-modules are precisely the interval \( \tau \)-modules. Moreover, every \( V \in \text{Mod}_\tau \) decomposes as a direct sum of interval \( \tau \)-modules. This decomposition is unique up to the order in which the summands appear.

Carlsson and de Silva first stated this theorem in the context of zigzag modules in [6]. As those authors point out, the existence and uniqueness of a decomposition of zigzag modules into indecomposables is proven in almost the exact same way as for ordinary modules; this is the classical Krull-Remak-Schmidt theorem for modules, a proof of which can be found in [17]. Gabriel solved the problem of describing the indecomposable zigzag modules, albeit in a more general setting, in [12]. Bernstein, Gelfand, and Ponomarev soon after provided an alternate proof of Gabriel’s Theorem in [4], using their so-called “reflection functors” which, perhaps unsurprisingly, play a pivotal role later in this paper in the definition of our reflection distance.

3.2.1 Multisets

A multiset is a pair \( \mathcal{M} = (S, m) \) where \( S \) is any set and \( m : S \to \mathbb{N} \) is called the multiplicity function which assigns to each \( s \in S \) its multiplicity \( m(s) \). The representation of a multiset \( \mathcal{M} = (S, m) \) is defined to be the set

\[
\{(s, k) \in S \times \mathbb{N} \mid 1 \leq k \leq m(s)\}.
\]

Informally, we think of a multiset as a set in which an element \( s \) is allowed to appear \( m(s) \) times. Given that \( M \) is the representation of some multiset \( \mathcal{M} = (S, m) \), we can retrieve the underlying set \( S \) via the projection function \( \pi_1 : M \to S \) defined as \( (s, k) \mapsto s \) by noticing that \( \pi_1(M) = S \). We will always drop the indexing component, referring to \( s \) instead of \( (s, k) \). To then distinguish a multiset representation from
an ordinary set, we will replace the curly brackets “\{−\}” by double curly brackets “\{\{−\}\}”. For instance, the multiset representation \{(a, 1), (a, 2), (b, 1), (c, 1)\} is written \{a, a, b, c\}.

Given two multisets \(M_1 = (S_1, m_1)\) and \(M_2 = (S_2, m_2)\), we define their union, denoted \(M_1 \sqcup M_2\), to be the multiset \((T, \ell)\) where \(T = S_1 \cup S_2\) and \(\ell : T \to \mathbb{N}\) is defined as

\[
\ell = \begin{cases} 
  m_1 + m_2 & \text{on } S_1 \cap S_2 \\
  m_1 & \text{on } S_1 \setminus S_2 \\
  m_2 & \text{on } S_2 \setminus S_1.
\end{cases}
\]

If \(M_1\) and \(M_2\) are the representations of \(\mathcal{M}_1\) and \(\mathcal{M}_2\), respectively, then \(M_1 \sqcup M_2\) is defined to be the representation of \(\mathcal{M}_1 \sqcup \mathcal{M}_2\).

We say that \(\mathcal{M}_2 = (S_2, m_2)\) is a multiset of the \(\mathcal{M}_1 = (S_1, m_1)\), and write \(\mathcal{M}_2 \subseteq \mathcal{M}_1\), if \(S_2 \subseteq S_1\) and \(m_2(s) \leq m_1(s)\) for all \(s \in S_2\). Note that \(\mathcal{M}_1 \subseteq \mathcal{M}_2\) as multisets if and only if \(M_2 \subseteq M_1\) in the usual set-theoretic sense. Finally, note that we have \(M_1, M_2 \subseteq M_1 \sqcup M_2\) and \(M_1 \subseteq M_1 \sqcup M_2, M_2 \subseteq M_1 \sqcup M_2\).

**Example 3.2.1.** Let \(S_1 = \{a, b, c\}\), \(S_2 = \{a, c\}\) and define \(m_1 : S_1 \to \mathbb{N}, m_2 : S_2 \to \mathbb{N}\) by

\[
m_1(x) = \begin{cases} 
  2, & \text{if } x = a \\
  1, & \text{if } x = b \\
  3, & \text{if } x = c
\end{cases}
\]
\[
m_2(x) = \begin{cases} 
  1, & \text{if } x = a \\
  2, & \text{if } x = c
\end{cases}
\]

Then \(\mathcal{M}_1 = (S_1, m_1)\) and \(\mathcal{M}_2 = (S_2, m_2)\) are multisets with respective multiset representations \(M_1 = \{\{a, a, b, c, c\}\}\) and \(M_2 = \{\{a, c\}\}\). Moreover, \(\mathcal{M}_2 \subseteq \mathcal{M}_1\) and \(\mathcal{M}_2 \subseteq M_1\). The union of \(\mathcal{M}_1\) and \(\mathcal{M}_2\) is the multiset \(\mathcal{M}_1 \sqcup \mathcal{M}_2 = (T, \ell)\) where \(T = S_1 \cup S_2 = \{a, b, c\}\) and

\[
\ell = \begin{cases} 
  m_1 + m_2 & \text{on } \{a, c\} \\
  m_1 & \text{on } \{c\}
\end{cases}
\]

so that \(\ell(x) = \begin{cases} 
  3, & \text{if } x = a \\
  1, & \text{if } x = b \\
  5, & \text{if } x = c.
\end{cases}\)

The multiset representation of \(\mathcal{M}_1 \sqcup \mathcal{M}_2\) is \(\mathcal{M}_1 \sqcup \mathcal{M}_2 = \{\{a, a, a, b, c, c, c, c, c\}\}\).
3.2.2 Persistence Barcodes and Persistence Diagrams

Fix an \( n \in \mathbb{N} \) and a type \( \tau \in \mathcal{T}_n \). By Theorem 3.2.1 every \( \tau \)-module \( V \in \text{Mod}_\tau \) has a decomposition of the form

\[
V \cong I_\tau([b_1, d_1]) \oplus \cdots \oplus I_\tau([b_N, d_N]). \tag{3.2}
\]

We define the persistence diagram of \( V \) to be the multiset representation

\[
\text{Dgm}(V) := \{ (b_i, d_i) \in \mathbb{N} \times \mathbb{N} \mid 1 \leq i \leq N \},
\]

whose elements are ordered pairs of endpoints defining the interval modules in the decomposition 3.2. In particular, we always have the decomposition

\[
V \cong \bigoplus_{(b, d) \in \text{Dgm}(V)} I([b, d]). \tag{3.3}
\]

Persistence diagrams are thus isomorphism invariants of zigzag modules. That is, for fixed type \( \tau \), a \( \tau \)-module determines and is determined up to isomorphism by its persistence diagram.

There is a simple relationship between the persistence diagram of a zigzag module and the persistence diagram of any of its summands:

**Theorem 3.2.2.** Fix \( n \in \mathbb{N} \) and \( \tau \in \mathcal{T}_n \). If \( W, V \in \text{Mod}_\tau \) with \( W \preceq V \) then \( \text{Dgm}(W) \subseteq \text{Dgm}(V) \).

**Proof.** Since \( W \preceq V \) there exists \( U \in \text{Mod}_\tau \) such that \( V \cong W \oplus U \). Using the decomposition 3.3. we have

\[
V \cong \left[ \bigoplus_{(b, d) \in \text{Dgm}(W)} I([b, d]) \right] \oplus \left[ \bigoplus_{(b, d) \in \text{Dgm}(U)} I([b, d]) \right].
\]

By the uniqueness statement of Theorem 3.2.1 \( \text{Dgm}(V) = \text{Dgm}(W) \sqcup \text{Dgm}(U) \) so that \( \text{Dgm}(W) \subseteq \text{Dgm}(V) \). \qed
It is also sometimes convenient to consider the persistence barcode of \( V \), defined to be the multiset representation

\[
\text{Pers}(V) := \{[b_1, d_1], \ldots, [b_N, d_N]\},
\]

whose elements are the integer intervals defining the interval modules in the decomposition \( \text{3.2} \). The interval decomposition of a zigzag module \( V \) is then expressed as

\[
V \cong \bigoplus_{J \in \text{Pers}(V)} I(J).
\]  

(3.4)

Persistence diagrams and persistence barcodes encode the same information about an interval decomposition; both \( \text{Dgm}(V) \) and \( \text{Pers}(V) \) are determined by \( V \) and determine \( V \) up to isomorphism. The distinction is made only to indicate a change in perspective: we visualize the persistence diagram of \( V \) by plotting its points in the \( \mathbb{N} \times \mathbb{N} \) plane, indicating multiplicities when necessary; the persistence barcode of \( V \) is visualized by drawing the intervals in \( \text{Pers}(V) \) over the \( \mathbb{N} \)-axis in an arbitrary order.

Both of these visualizations are illustrated in the following:

**Example 3.2.2.** Let \( n = 4 \) and let \( \tau = (\rightarrow, \leftarrow, \rightarrow) \in \mathcal{T}_n \). Consider the \( \tau \)-module

\[
V = I_\tau([1, 4]) \oplus I_\tau([1, 2]) \oplus I_\tau([2, 3]) \oplus I_\tau([2, 3]) \oplus I_\tau([3, 3]).
\]

We have

\[
\text{Dgm}(V) = \{(1, 4), (1, 2), (2, 3), (2, 3), (3, 3)\}
\]

and

\[
\text{Pers}(V) = \{[1, 4], [1, 2], [2, 3], [2, 3], [3, 3]\}.
\]

The persistence diagram and persistence barcode of \( V \) are plotted in Figure \( \text{3.7} \).
3.3 Reflection Functors

3.3.1 Sinks and Sources

Fix \( n \in \mathbb{N} \) and let \( \tau \in \mathcal{T}_n \). A zigzag module \( V = (V_i, p_i) \in \text{Mod}_\tau \) has a sink at index \( k \in \{2, \ldots, n - 1\} \) if it has the form

\[
V = \cdots \xleftarrow{p_{k-2}} V_{k-1} \xrightarrow{p_{k-1}} V_k \xleftarrow{p_k} V_{k+1} \xrightarrow{p_{k+1}} \cdots.
\]

In addition, we say that \( V \) has a sink at index 1 or index \( n \) if the maps \( p_1 \) or \( p_{n-1} \) are of the form \( V_1 \xleftarrow{p_1} V_2 \) or \( V_{n-1} \xrightarrow{p_{n-1}} V_n \), respectively.

Similarly, \( V = (V_i, p_i) \in \text{Mod}_\tau \) has a source at index \( k \in \{2, \ldots, n - 1\} \) if it has the form

\[
V = \cdots \xrightarrow{p_{k-2}} V_{k-1} \xleftarrow{p_{k-1}} V_k \xrightarrow{p_k} V_{k+1} \xleftarrow{p_{k+1}} \cdots,
\]

and has a source at index 1 or \( n \) if the maps \( p_1 \) or \( p_n \) are of the form \( V_1 \xrightarrow{p_1} V_2 \) or \( V_{n-1} \xleftarrow{p_{n-1}} V_n \), respectively. In other words, a \( \tau \)-module \( V \) has a sink at index
$k \in \{1, \ldots, n\}$ if none of the linear maps $p_i$ have domain $V_k$, and $V$ has a source at index $k$ if none of the linear maps $p_i$ have codomain $V_k$.

Note that the property of having a sink or source at a given index depends only on the type of the zigzag module in question; that is, if a $\tau$-module $V$ has a sink or source at index $k \in \{1, \ldots, n\}$ then any other $\tau$-module will also have, respectively, a sink or source at index $k$. Thus, we say that a type $\tau \in \mathcal{T}_n$ has a sink or source at index $k$ if any $\tau$-module has, respectively, a sink or source at index $k$.

**Remark 3.3.1.** When we refer to a type $\tau \in \mathcal{T}_n$ as having a sink or source at index $k \in \{1, \ldots, n\}$, we mean that any $\tau$-module has a sink or source, respectively, at index $k$. The word “index” here does not refer to the $k$-th component of the sequence of arrows defining $\tau$.

**Example 3.3.1.** Let $n = 4$ and let $\tau = (\rightarrow, \leftarrow, \rightarrow) \in \mathcal{T}_n$. Then $\tau$ has sinks at indices 2 and 4 and sources at indices 1 and 3.

### 3.3.2 Type Reflections and Type Reversals

For each $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$, define

$$\mathcal{T}_n^{k, +} := \{\tau \in \mathcal{T}_n \mid \tau \text{ has a sink at index } k\},$$

and

$$\mathcal{T}_n^{k, -} := \{\tau \in \mathcal{T}_n \mid \tau \text{ has a source at index } k\}.$$ 

We let $\mathcal{T}_n^{k, \pm} = \mathcal{T}_n^{k, +} \cup \mathcal{T}_n^{k, -}$ denote the set of all types $\tau \in \mathcal{T}_n$ with either a sink or a source at index $k$. We define the $k$-th reflection map $\sigma_k : \mathcal{T}_n^{k, \pm} \to \mathcal{T}_n^{k, \pm}$ by setting $\sigma_k \tau := \sigma_k(\tau)$ to be the type obtained as follows:

38
1. If $1 < k < n$ then $\sigma_k \tau$ is obtained from $\tau$ by interchanging the $(k - 1)$-st and $k$-th entries of the sequence defining $\tau$.

2. If $k = 1$ then $\sigma_k \tau$ is obtained from $\tau$ by replacing $\to$ by $\leftarrow$ or vice versa in the 1st entry of the sequence defining $\tau$.

3. If $k = n$ then $\sigma_k \tau$ is obtained from $\tau$ by replacing $\to$ by $\leftarrow$ or vice versa in the $(n - 1)$-st entry of the sequence defining $\tau$.

Note that if $\tau \in T_{n}^{k,+}$ then $\sigma_k \tau \in T_{n}^{k,-}$ and vice versa.

**Example 3.3.2.** Let $n = 4$ and let $\tau = (\to, \leftarrow, \to) \in T_n$ as in Example 3.3.1. We have

$$\sigma_1 \tau = (\leftarrow, \leftarrow, \to) \quad \sigma_2 \tau = (\leftarrow, \to, \to)$$

$$\sigma_3 \tau = (\to, \to, \leftarrow) \quad \sigma_4 \tau = (\to, \leftarrow, \leftarrow).$$

For each $k \in \{1, \ldots, n - 1\}$, we define the $k$-th reversal map $r_k : T_n \to T_n$ by setting $r_k \tau = r_k(\tau)$ to be the type obtained from $\tau$ by replacing $\to$ by $\leftarrow$ or vice versa in the $k$-th entry of the sequence defining $\tau$.

**Example 3.3.3.** Again let $n = 4$ and $\tau = (\to, \leftarrow, \to) \in T_n$ as in Examples 3.3.1 and 3.3.2. We have

$$r_1 \tau = (\leftarrow, \leftarrow, \to), \quad r_2 \tau = (\to, \to, \to), \quad r_3 \tau = (\to, \leftarrow, \leftarrow).$$

### 3.3.3 Sink and Source Reflection Functors

In this section, we define the so-called “reflection functors” introduce by Bernstein, Gelfand, and Ponomarev in [4] as a tool for proving Gabriel’s Theorem. For us, reflection functors will provide a means of transforming a zigzag module into a new zigzag module of a different type.

Throughout this section we fix $n \in \mathbb{N}$. 

39
3.3.4 Sink Reflection Functors

Fix \( k \in \{2, \ldots, n - 1\} \) and let \( \tau \in \mathcal{T}_n^{k, +} \). We define a map

\[
R^+_k : \text{Mod}_\tau \to \text{Mod}_{\sigma_k \tau}
\]

as follows: given the zigzag module

\[
V = \cdots \leftarrow V_{k-1} \overset{p_{k-1}}{\longrightarrow} V_k \overset{p_k}{\leftarrow} V_{k+1} \overset{p_{k+1}}{\longrightarrow} \cdots
\]

in \( \text{Mod}_\tau \), we define a new zigzag module

\[
R^+_k(V) = \cdots \leftarrow V_{k-1} \overset{p_{k-2}}{\leftarrow} V_k \overset{\beta_{k-1}}{\leftarrow} V_k^+ \overset{\alpha_k}{\longrightarrow} V_{k+1} \overset{p_{k+1}}{\leftarrow} \cdots
\]

in \( \text{Mod}_{\sigma_k \tau} \) by setting \( V_k^+ := \ker(h) \), where \( h : V_{k-1} \oplus V_{k+1} \to V_k \) is given by \( v \oplus v' \mapsto p_{k-1}(v) + p_k(v') \), and by defining the linear transformations \( \beta_{k-1} \) and \( \alpha_{k+1} \) to be, respectively, the canonical projections \( \pi_{k-1}, \pi_{k+1} : V_{k-1} \oplus V_{k+1} \to V_k \) restricted to \( \ker(h) \).

If \( \phi : V \to W \) is a morphism of \( \tau \)-modules, we define

\[
\psi = R^+_k(\phi) : R^+_k(V) \to R^+_k(W)
\]

by setting \( \psi_j = \phi_j \) for all \( j \neq k \) and by defining \( \psi_k \) to be the restriction of \( \phi_{k-1} \oplus \phi_{k+1} : V_{k-1} \oplus V_{k+1} \to W_{k-1} \oplus W_{k+1} \) to \( \ker(h) \).

When \( k = 1 \) or \( k = n \), we construct \( R^+_k(V) \) by extending \( V \) to the left or right, respectively, by appending the zero vector space connected by the zero map so that the resulting zigzag module has a sink at the node corresponding to \( V_k \), applying the reflection map defined above at this node, and then removing the appended zero space. Explicitly, if \( V \in n\text{-Mod} \) with \( \tau = \text{type}(V) \in \mathcal{T}_n^{1, +} \) so that \( V \) is of the form

\[
V = V_1 \leftarrow V_2 \overset{p_1}{\longrightarrow} \cdots
\]
then we define $R_1^+(V) \in \text{Mod}_{\sigma_1\tau}$ to be the zigzag module

$$R_1 V = V_1^+ \xrightarrow{\alpha_1} V_2 \xleftarrow{p_2} \cdots$$

where $V_1^+ := \ker(h)$ and where $h : 0 \oplus V_2 \to V_1$ is the map $0 \oplus v \mapsto 0 + p_1(v)$. Thus we just have $V_1^+ = \ker(p_1)$. The map $\alpha_1 : V_1^+ \to V_2$ is then defined to be the inclusion $\ker(p_1) \hookrightarrow V_2$.

Similarly, if $V \in n\text{-Mod}$ with $\tau = \text{type}(V) \in T_{n,+}^n$ so that $V$ is of the form

$$V = \cdots \xrightarrow{p_{n-2}} V_{n-1} \xleftarrow{p_{n-1}} V_n$$

then we define $R_n^+(V) \in \text{Mod}_{\sigma_n\tau}$ to be the zigzag module

$$R_n^+(V) = \cdots \xrightarrow{p_{n-2}} V_{n-1} \xleftarrow{\beta_{n-1}} V_n^+$$

where $V_n^+ := \ker(h)$ and where $h : V_{n-1} \oplus 0 \to V_n$ is the map $0 \oplus v \mapsto p_{n-1}(v) + 0$. Thus we have $V_n^+ = \ker(p_{n-1})$. The map $\beta_{n-1} : V_n^+ \to V_{n-1}$ is then defined to be the inclusion $\ker(p_{n-1}) \hookrightarrow V_{n-1}$.

**Remark 3.3.2.** For each $k \in \{1, \ldots, n\}$, the above definitions make $R_k^+$ into a functor from $\text{Mod}_\tau$ to $\text{Mod}_{\sigma_k\tau}$, called the sink reflection functor at index $k$.

### 3.3.5 Source Reflection Functors

The source reflection functor is defined dually to the sink reflection functor. Fix $k \in \{2, \ldots, n-1\}$ and let $\tau \in T_{n,k-}^n$. We define a map

$$R_k^- : \text{Mod}_\tau \to \text{Mod}_{\sigma_k\tau}$$

as follows: given the zigzag module

$$V = \cdots \xleftarrow{p_{k-2}} V_{k-1} \xrightarrow{p_{k-1}} V_k \xrightarrow{p_k} V_{k+1} \xleftarrow{p_{k+1}} \cdots$$
in \text{Mod}_\tau, we define a new zigzag module
\[ R_k^-(V) = \cdots \xleftarrow{p_{k-2}} V_{k-1} \xrightarrow{\alpha_{k-1}} V_k^- \xleftarrow{\beta_k} V_{k+1} \xleftarrow{p_{k+1}} \cdots \]
in \text{Mod}_{\sigma_k \tau} as follows: \( V_k^- = \text{coker}(h) \) where \( h : V_k \to V_{k-1} \oplus V_{k+1} \) is given by \( v \mapsto (p_{k-1}(v), p_k(v)) \). The linear maps \( \alpha_{k-1} \) and \( \beta_k \) are given by the canonical quotient map \( q : V_{k-1} \oplus V_{k+1} \to \text{coker}(h) \) restricted to the images of the canonical inclusions \( \iota_{k-1} : V_{k-1} \hookrightarrow V_{k-1} \oplus V_{k+1} \) and \( \iota_{k+1} : V_{k+1} \hookrightarrow V_{k-1} \oplus V_{k+1} \), respectively.

If \( \phi : V \to W \) is a morphism of zigzag modules, we define
\[ \psi = R_k^-(\phi) : R_k^-(V) \to R_k^-(W) \]
by setting \( \psi_j = \phi_j \) for all \( j \neq k \) and by defining \( \psi_k \) to be the map induced from \( \phi_{k-1} \oplus \phi_{k+1} : V_{k-1} \oplus V_{k+1} \to W_{k-1} \oplus W_{k+1} \) by passing to the quotient.

In the cases \( k = 1 \) and \( k = n \), the map \( R_k^-(V) \) is defined in the same manner as \( R_k^+ \) was defined above. In full detail, if \( V \in n\text{-Mod} \) with \( \tau = \text{type}(V) \in T_{1}^{n} \) so that \( V \) has the form
\[ V = V_1 \xrightarrow{p_1} V_2 \xleftarrow{p_2} \cdots \]
then we define \( R_1^-(V) \in \text{Mod}_{\sigma_1 \tau} \) to be the zigzag module
\[ R_1^-(V) = V_1^- \xleftarrow{\beta_1} V_2 \xleftarrow{p_2} \cdots \]
where \( V_1^- := \text{coker}(h) \) and where \( h : V_1 \to 0 \oplus V_2 \) is the map \( v \mapsto (0, p_1(v)) \). The map \( \beta_1 : V_2 \to V_1^- \) is defined to be the restriction of the canonical quotient map \( p : 0 \oplus V_2 \to \text{coker}(h) \) restricted to the image of the canonical inclusion \( \iota_2 : V_2 \hookrightarrow 0 \oplus V_2 \).

It is not hard to check that the zigzag module obtained is isomorphic to the one whose first vector space is defined to be \( \text{coker}(p_1) \) and whose first structure map \( q_1 : V_2 \to \text{coker}(p_1) = V_2/\text{im}(p_1) \) is the canonical quotient map.
For $V \in n\text{-Mod}$ with $\tau = \text{type}(V) \in \mathcal{T}_n^{m,-}$, the zigzag module $R_n^-(V) \in \text{Mod}_{\sigma_n,\tau}$ is defined in a completely analogous way.

**Remark 3.3.3.** For each $k \in \{1, \cdots, n\}$, the above definitions make $R_k^-$ into a functor from $\text{Mod}_\tau$ to $\text{Mod}_{\sigma_k,\tau}$, called the source reflection functor at index $k$.

### 3.3.6 Properties of Reflection Functors

The next proposition shows that reflections are universal in a precise sense and formalizes the sense in which the sink and source reflection functors are dual to each other:

**Proposition 3.3.1.** Fix $n \in \mathbb{N}$ and $k \in \{2, \ldots, n-1\}$, and let $\text{FinVec}$ denote the category of finite dimensional vector spaces.

1.) $V_k^+, \beta_{k-1}$, and $\alpha_{k+1}$, as defined in section 3.3.4, form the pullback of the diagram

\[
\begin{array}{ccc}
V_{k-1} & \xrightarrow{p_{k-1}} & V_k & \xleftarrow{-p_k} & V_{k+1}
\end{array}
\]

in the category $\text{FinVec}$.

2.) $V_k^-$, $\alpha_{k-1}$, and $\beta_{k+1}$, as defined in section 3.3.5, form the pushout of the diagram

\[
\begin{array}{ccc}
V_{k-1} & \xleftarrow{p_{k-1}} & V_k & \xrightarrow{-p_k} & V_{k+1}
\end{array}
\]

in the category $\text{FinVec}$.

**Proof.** 1.) It is an exercise in both [11] and [13] to show that the pullback of the diagram defined by the linear maps $p_{k-1} : V_{k-1} \rightarrow V_k$ and $-p_k : V_{k+1} \rightarrow V_k$ is the subspace of $V_{k-1} \oplus V_{k+1}$

\[
U = \{v \oplus v' \in V_{k-1} \oplus V_{k+1} \mid p_{k-1}(v) = -p_k(v')\},
\]

43
together with the canonical projections. Evidently, \( U = \ker(h) = V_k^+ \) where \( h : V_{k-1} \oplus V_{k+1} \to V_k \) is as defined in section 3.3.4.

2.) It is also an exercise in [11] and [13] to show that the pushout of the linear maps \( p_{k-1} : V_k \to V_{k-1} \) and \(-p_k : V_k \to V_{k+1}\) is given by the quotient \( (V_{k-1} \oplus V_{k+1})/W\), where \( W \) is the subspace

\[
W = \{(g_{k-1}(v), f_{k+1}(v)) \in V_{k-1} \oplus V_{k+1} \mid v \in V_k\},
\]
together with the linear maps defined as the canonical quotient projection map restricted to the canonical inclusions of \( V_{k-1} \) and \( V_{k+1} \) into \( V_{k-1} \oplus V_{k+1} \). Evidently, this is just the definition of \( V_k^- \), \( \alpha_{k-1} \), and \( \beta_{k+1} \).

Proposition 3.3.2. Fix \( k \in \{1, \ldots, n\} \), let \( \tau \in \mathcal{T}_n^{\pm} \), and let \( V, W \in \text{Mod}_\tau \). If \( W \leq V \) then \( \mathcal{R}_k^\pm(W) \leq \mathcal{R}_k^\pm(V) \). It follows that if \( V \cong W \) then \( \mathcal{R}_k^\pm(V) \cong \mathcal{R}_k^\pm(W) \).

Proof. If \( W \leq V \) then by Proposition 3.1.1 there is a monomorphism \( j : W \hookrightarrow V \). Since pullbacks and pushouts are limits and colimits, respectively, of the appropriate diagram, we may apply Theorem 2.4.1 to obtain a monomorphism \( \mathcal{R}_k^\pm(j) : \mathcal{R}_k^\pm(W) \hookrightarrow \mathcal{R}_k^\pm(V) \). The second statement follows from the fact that \( V \cong W \) if and only if \( W \leq V \) and \( V \leq W \).

Proposition 3.3.3. Fix \( k \in \{1, \ldots, n\} \), let \( \tau \in \mathcal{T}_n^{\pm} \), and let \( V_1, V_2 \in \text{Mod}_\tau \). Then \( \mathcal{R}_k^\pm(V_1 \oplus V_2) \cong \mathcal{R}_k^\pm(V_1) \oplus \mathcal{R}_k^\pm(V_2) \). This statement generalizes to the sum of any finite number of \( \tau \)-modules.

Proof. This follows from Proposition 3.3.1 together with Theorem 2.4.2. Induction is used to extend the statement to arbitrary finite sums.
Corollary 3.3.1. Fix $k \in \{1, \ldots, n\}$, let $\tau \in \mathcal{T}_n^{k,\pm}$, and let $V, W \in \text{Mod}_\tau$. Suppose that $W \preceq V$. Then $R_k^\pm(W) \preceq R_k^\pm(V)$.

Proof. If $W \preceq V$ then $V \cong W \oplus U$ for some $U \in \text{Mod}_\tau$ so that by Proposition 3.3.3, $R_k^\pm(V) \cong R_k^\pm(W \oplus U) \cong R_k^\pm(W) \oplus R_k^\pm(U)$ and hence $R_k^\pm(W) \preceq R_k^\pm(V)$. \hfill \qed

Corollary 3.3.2. Fix $k \in \{1, \ldots, n\}$, let $\tau \in \mathcal{T}_n^{k,\pm}$, and let $V \in \text{Mod}_\tau$. Then we have $R_k^\pm(V) \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} R_k^\pm(I_{\tau}([b,d]))$.

Proof. Write $V \cong \bigoplus_{(b,d) \in \text{Pers}(V)} I_{\tau}([b,d])$. Using Propositions 3.3.2 and 3.3.3 we have $R_k^\pm(V) \cong R_k^\pm \left( \bigoplus_{(b,d) \in \text{Dgm}(V)} I_{\tau}([b,d]) \right) \cong \bigoplus_{(b,d) \in \text{Pers}(V)} R_k^\pm(I_{\tau}([b,d]))$. \hfill \qed

Corollary 3.3.2 together with the following theorem describe exactly how reflections of zigzag modules effect their persistence barcodes/diagrams:

Theorem 3.3.1 ([4], [14]). Let $k \in \{2, \ldots, n-1\}$ and let $\tau \in \mathcal{T}_n^{k,\pm}$ so that $\sigma_k \tau \in \mathcal{T}_n^{k,-}$. The reflection functors $R_k^+$ and $R_k^-$ induce mutually inverse bijections between the isomorphism classes of interval $\tau$-modules and the isomorphism classes of interval $\sigma_k \tau$-modules, with the exception of the simple interval modules $I_{\tau}([k,k])$ and $I_{\sigma_k \tau}([k,k])$ which are annihilated by these functors. These functors act on the interval modules as follows:

\[
I_{\tau}([k,k]), I_{\sigma_k \tau}([k,k]) \rightarrow 0
\]

\[
I_{\tau}([b,k-1]) \leftrightarrow I_{\sigma_k \tau}([b,k]) \quad \text{for } b \leq k - 1
\]

\[
I_{\tau}([k+1,d]) \leftrightarrow I_{\sigma_k \tau}([k,d]) \quad \text{for } d \geq k + 1
\]

\[
I_{\tau}([b,d]) \leftrightarrow I_{\sigma_k \tau}([b,d]) \quad \text{otherwise.}
\]
The exceptional cases $k = 1$ and $k = n$ behave similarly:

\[ I_\tau([1,1]), I_{\sigma_1\tau}([1,1]) \rightarrow 0 \]

\[ I_\tau([1,d]) \leftrightarrow I_{\sigma_1\tau}([2,d]) \quad \text{for } d \geq 2 \]

and

\[ I_\tau([n,n]), I_{\sigma_n\tau}([n,n]) \rightarrow 0 \]

\[ I_\tau([b,n]) \leftrightarrow I_{\sigma_n\tau}([b,n-1]) \quad \text{for } b \leq n-1. \]

A statement completely analogous to Theorem 3.3.1 holds in the case that $\tau \in T_n^{k,-}$.

---

**Figure 3.2**: Left: Points in the diagram of a zigzag module $V$ move according to the arrows when the reflection functors $R^+_k$ or $R^-_k$ are applied. The point $(k,k)$ corresponding to the summand $I([k,k])$ is killed. Right: The corresponding picture for barcodes.

---

**Example 3.3.4.** Let $n = 4$, let $\tau = (\rightarrow, \leftarrow, \rightarrow) \in T_n$, and let $V$ be the $\tau$-module

\[ V = I_\tau([1,4]) \oplus I_\tau([1,2]) \oplus I_\tau([2,3]) \oplus I_\tau([2,3]) \oplus I_\tau([3,3]) \]

which was consider in Example 3.2.2. Evidently, $\tau \in T_n^{2,+}$. Applying the sink reflection functor at index 2 and using Corollary 3.3.2 and Theorem 3.3.1 we have

\[ R^+_2(V) \cong I_{\sigma_2\tau}([1,4]) \oplus I_{\sigma_2\tau}([1,1]) \oplus 3I_{\sigma_2\tau}([3,3]). \]
Similarly, we have \( \tau \in T_n^{3-} \) and

\[
\mathcal{R}_3^- (V) \cong I_{\sigma_3 \tau}([1, 4]) \oplus I_{\sigma_3 \tau}([1, 2]) \oplus 2I_{\sigma_3 \tau}([2, 2]).
\]

3.4 Extroversion and Introversion Functors

3.4.1 Forward and Backward Flows

Fix an \( n \in \mathbb{N} \) and let \( \tau \in T_n \). A zigzag module \( V = (V_i, p_i) \in \text{Mod}_\tau \) is said to flow forwards at index \( k \in \{2, \ldots, n-1\} \) if it has the form

\[
V = \cdots \xleftarrow{p_{k-2}} V_{k-1} \xrightarrow{p_{k-1}} V_k \xrightarrow{p_k} V_{k+1} \xleftarrow{p_{k+1}} \cdots.
\]

In addition, we say that \( V \) flows forwards at index 1 or index \( n \) if the maps \( p_1 \) or \( p_{n-1} \) are of the form \( V_1 \xrightarrow{p_1} V_2 \) or \( V_{n-1} \xleftarrow{p_{n-1}} V_n \), respectively.

Similarly, \( V = (V_i, p_i) \in \text{Mod}_\tau \) is said to flow backwards at index \( k \in \{2, \ldots, n-1\} \) if it has the form

\[
V = \cdots \xrightarrow{p_{k-2}} V_{k-1} \xleftarrow{p_{k-1}} V_k \xleftarrow{p_k} V_{k+1} \xrightarrow{p_{k+1}} \cdots.
\]

and flows backwards at index 1 or \( n \) if the maps \( p_1 \) or \( p_n \) are of the form \( V_1 \xleftarrow{p_1} V_2 \) or \( V_{n-1} \xrightarrow{p_{n-1}} V_n \), respectively.

Just as with sinks and sources, the property of having a forward or backward flow at a given index depends only on the type of the zigzag module in question. Thus, we say that a type \( \tau \in T_n \) flows forwards or flows backwards at index \( k \) if any \( \tau \)-module flows forwards or flows backwards, respectively, at index \( k \).

Example 3.4.1. Let \( n = 5 \) and let \( \tau = (\rightarrow, \rightarrow, \leftarrow, \leftarrow) \in T_n \). Then \( \tau \) flows forwards at indices 1 and 2, and flows backwards at indices 4 and 5.
3.4.2 Type Extroversion and Introversion

For each \( n \in \mathbb{N} \) and \( k \in \{1, \ldots, k\} \), define

\[
\mathcal{T}^{k,\to}_n := \{ \tau \in \mathcal{T}_n \mid \tau \text{ flows forwards at index } k \},
\]

and

\[
\mathcal{T}^{k,\leftarrow}_n := \{ \tau \in \mathcal{T}_n \mid \tau \text{ flows backwards index } k \}.
\]

For notational convenience later, we let

\[
\mathcal{T}^{k,*}_n := \mathcal{T}^{k,\bullet}_n := \mathcal{T}^{k,\to}_n \cup \mathcal{T}^{k,\leftarrow}_n
\]
denote the set of all types \( \tau \in \mathcal{T}_n \) which flow forwards or backwards at index \( k \). The reason for fixing two notations for the same set will become clear later on.

For each \( k \), we define two maps which take as input a type with a flow at index \( k \) and give as output a type with either a source or sink at index \( k \), leaving the other arrows unchanged. The formal definitions follow:

We define the \( k \)-th extroversion map \( \psi_k : \mathcal{T}^{k,*}_n = \mathcal{T}^{k,\bullet}_n \to \mathcal{T}^{k,\leftarrow}_n \) by setting \( \psi_k \tau := \psi_k(\tau) \) to be the type obtained as follows:

1. For \( 1 < k < n \),

   (a) if \( \tau \in \mathcal{T}^{k,\to}_n \) then \( \psi_k \tau \) is obtained from \( \tau \) by replacing \( \to \) by \( \leftarrow \) in the \((k-1)\)-st entry of the sequence defining \( \tau \),

   (b) if \( \tau \in \mathcal{T}^{k,\leftarrow}_n \) then \( \psi_k \tau \) is obtained from \( \tau \) by replacing \( \leftarrow \) by \( \to \) in the \( k \)-th entry of the sequence defining \( \tau \).

2. For \( k = 1 \),

   (a) if \( \tau \in \mathcal{T}^{1,\to}_n \) then we set \( \psi_k \tau = \tau \),

   (b) if \( \tau \in \mathcal{T}^{1,\leftarrow}_n \) then we set \( \psi_k \tau = \tau \).

48
(b) if $\tau \in \mathcal{T}^{k,\leftarrow}_n$ then $\psi_k \tau$ is obtained from $\tau$ by replacing $\leftarrow$ by $\rightarrow$ in the 1st entry of the sequence defining $\tau$.

3. For $k = n$,

(a) if $\tau \in \mathcal{T}^{k,\rightarrow}_n$ then $\psi_k \tau$ is obtained from $\tau$ by replacing $\rightarrow$ by $\leftarrow$ in the $(n - 1)$-st entry of the sequence defining $\tau$,

(b) if $\tau \in \mathcal{T}^{k,\leftarrow}_n$ then we set $\psi_k \tau = \tau$.

Example 3.4.2. Let $n = 5$ and let $\tau = (\rightarrow, \rightarrow, \leftarrow, \leftarrow) \in \mathcal{T}_n$ as in Example 3.4.1. Then we have

$$
\psi_1 \tau = \tau = (\rightarrow, \rightarrow, \leftarrow, \leftarrow) \quad \psi_2 \tau = (\leftarrow, \rightarrow, \leftarrow, \leftarrow)
$$

$$
\psi_3 \tau = (\rightarrow, \rightarrow, \leftarrow, \rightarrow) \quad \psi_5 \tau = (\rightarrow, \rightarrow, \leftarrow, \leftarrow).
$$

We define the $k$-th introversion map $\xi_k : \mathcal{T}^{k,*}_n = \mathcal{T}^{k,*}_n \rightarrow \mathcal{T}^{k,+}_n$ similarly: we set $\xi_k \tau := \xi_k(\tau)$ to be the type obtained as follows:

1. For $1 < k < n$,

(a) if $\tau \in \mathcal{T}^{k,\rightarrow}_n$ then $\xi_k \tau$ is obtained from $\tau$ by replacing $\rightarrow$ by $\leftarrow$ in the $k$-th entry of the sequence defining $\tau$,

(b) if $\tau \in \mathcal{T}^{k,\leftarrow}_n$ then $\xi_k \tau$ is obtained from $\tau$ by replacing $\leftarrow$ by $\rightarrow$ in the $(k - 1)$-st entry of the sequence defining $\tau$.

2. For $k = 1$,

(a) if $\tau \in \mathcal{T}^{k,\rightarrow}_n$ then $\xi_k \tau$ is obtained from $\tau$ by replacing $\rightarrow$ by $\leftarrow$ in the 1st entry of the sequence defining $\tau$,

(b) if $\tau \in \mathcal{T}^{k,\leftarrow}_n$ then we set $\xi_k \tau = \tau$. 

49
3. For \( k = n \),

(a) if \( \tau \in T_k^{n,\rightarrow} \) then we set \( \xi_k \tau = \tau \),

(b) if \( \tau \in T_k^{n,\leftarrow} \) then \( \xi_k \tau \) is obtained from \( \tau \) by replacing \( \leftarrow \) by \( \rightarrow \) in the

\((n-1)\)-st entry of the sequence defining \( \tau \),

**Example 3.4.3.** Let \( n = 5 \) and let \( \tau = (\rightarrow, \rightarrow, \leftarrow, \leftarrow) \in T_n \) as in Examples 3.4.1 and 3.4.2. Then we have

\[
\begin{align*}
\psi_1 \tau &= (\leftarrow, \rightarrow, \leftarrow, \leftarrow) \\
\psi_2 \tau &= (\leftarrow, \rightarrow, \leftarrow, \leftarrow) \\
\psi_3 \tau &= (\rightarrow, \leftarrow, \rightarrow) \\
\psi_5 \tau &= (\rightarrow, \rightarrow, \leftarrow, \leftarrow) \\
\psi_4 \tau &= (\rightarrow, \rightarrow, \leftarrow, \leftarrow)
\end{align*}
\]

3.4.3 Extroversion and Introversion Functors

We introduce now two more classes of functors which act on zigzag modules with a flow at a given index. The motivation for the following definitions is Proposition 3.3.1 which showed that the reflection functors are obtained by considering limits or colimits of diagrams of the form \( X \rightarrow Y \leftarrow Z \) or \( X \leftarrow Y \rightarrow Z \), respectively. The definitions that follow are related analogously to limits and colimits of diagrams of the form \( X \rightarrow Y \rightarrow Z \) and \( X \leftarrow Y \leftarrow Z \).

Suppose that \( V = (V_i, p_i) \in \text{Mod}_\tau \) for some \( \tau \in T_n^{k,*} = T_n^{k,*} \). Then we can extract either the subdiagram

\[
\begin{array}{ccc}
V_{k-1} & \xrightarrow{p_{k-1}} & V_k & \xrightarrow{p_k} & V_{k+1} \\
\end{array}
\]

or the subdiagram

\[
\begin{array}{ccc}
V_{k-1} & \xleftarrow{p_{k-1}} & V_k & \xleftarrow{p_k} & V_{k+1} \\
\end{array}
\]
depending on whether $\tau \in T_{n}^{k,\rightarrow}$ or $\tau \in T_{n}^{k,\leftarrow}$, respectively. In either case, we may compute the limit $(L, \lambda_j)$ of the extracted diagram and then consider the new diagram

\[ V_{k-1} \xleftarrow{\lambda_1} L \xrightarrow{\lambda_3} V_{k+1}. \]  

(3.7)

We define $\mathcal{R}_k^\star(V) \in \text{Mod}_{\psi_k} \tau$ to be the zigzag module obtained by replacing the appearance of the subdiagram (3.5) or (3.6) by diagram (3.7). To ensure that $\mathcal{R}_k^\star$ is well-defined for $k = 1$ and $k = n$, we use the convention $V_0 = V_{n+1} = 0$, extending our zigzag module accordingly, computing the appropriate transformation, and then removing the auxiliary nodes $V_0$ or $V_n$.

Similarly, if $(C, \gamma_j)$ is the colimit of either diagram (3.5) or (3.6), then we consider the new diagram

\[ V_{k-1} \xrightarrow{\gamma_1} L \xleftarrow{\gamma_3} V_{k+1}. \]  

(3.8)

and we define $\mathcal{R}_k^\star(W) \in \text{Mod}_{\xi_k} \tau$ to be the zigzag module obtained by replacing the appearance of the subdiagram (3.5) or (3.6) by diagram (3.8).

Now let $V, W \in \text{Mod}_{\tau}$ and let $\phi : V \rightarrow W$ by a morphism of $\tau$-modules. Denote the limits of the diagrams $V_i \xrightarrow{f_{vi}} V_k \xrightarrow{g_{vi}} V_{i'}$ and $W_i \xrightarrow{f_{wi}} W_k \xrightarrow{g_{wi}} W_{i'}$ by $(L_V, \{v_i, v_k, v_{i'}\})$ and $(L_W, \{w_i, w_k, w_{i'}\})$, respectively. Then we have the following commutative diagram:

```
   L_V
   |  \exists \mu
   |    \phi_v
   |      \phi_i
   V_i  \xrightarrow{f_v}  V_k  \xrightarrow{g_v}  V_{i'}
   |      \phi_{i'}
    \phi
   W_i  \xrightarrow{f_w}  W_k  \xrightarrow{g_w}  W_{i'}
```

51
From this diagram, we see that \((L_V, \{v_i, v_k, v_{i'}\})\) is a cone for \(W_i \xrightarrow{f_w} W_k \xrightarrow{g_w} W_{i'}\) so that by the universality of \(L_W\), there exists a unique morphism \(\mu\) making the diagram commute. We then define \(\psi = R^*_k : R^*_k(V) \to R^*_k(W)\) by setting \(\psi_j = \phi_j\) for all \(j \neq k\) and \(\phi_k := \mu\). The fact that \(\psi\) is a well-defined morphism follows from the commutativity of the boldened portion of the diagram above.

Similarly, if \((C_V, \{v_i, v_k, v_{i'}\})\) and \((C_W, \{w_i, w_k, w_{i'}\})\) denote the colimits of the diagrams \(V_i \xrightarrow{f_v} V_k \xrightarrow{g_v} V_{i'}\) and \(W_i \xrightarrow{f_w} W_k \xrightarrow{g_w} W_{i'}\), respectively, then we have the following commutative diagram:

From this diagram, we see that \((C_W, \{w_i, w_k, w_{i'}\})\) is a co-cone for \(V_i \xrightarrow{f_v} V_k \xrightarrow{g_v} V_{i'}\) so that by the universality of \(C_V\), there exists a unique morphism \(\eta\) making the diagram commute. We then define \(\psi = R^\bullet_k : R^\bullet_k(V) \to R^\bullet_k(W)\) by setting \(\psi_j = \phi_j\) for all \(j \neq k\) and \(\phi_k := \eta\). Again, \(\psi\) is well-defined morphism since the boldened portion of the diagram commutes.

For each \(k \in \{1, \ldots, n\}\) and for each \(\tau \in \mathcal{T}_n^{k,*} = \mathcal{T}_n^{k,*}\), the above definition makes \(R^*_k\) into a functor from \(\text{Mod}_\tau\) to \(\text{Mod}_{\psi_k\tau}\), which we call the extroversion functor at index \(k\). Similarly, \(R^\bullet_k\) is a functor from \(\text{Mod}_\tau\) to \(\text{Mod}_{\xi_k\tau}\), which we call the introversion functor at index \(k\).

The limit and colimit of diagram (3.5) are particularly simple to describe:
Fact 3.4.1. The limit \((L, \lambda_j)\) of diagram (3.3) is the vector space \(L = V_{k-1}\) together with the linear maps \(\lambda_1 = \text{id}_{V_{k-1}}, \lambda_2 = p_{k-1},\) and \(\lambda_3 = p_k \circ p_{k-1}\). The colimit \((C, \gamma_j)\) of this diagram is the vector space \(C = V_{k+1}\) together with the linear maps \(\lambda_1 = p_k \circ p_{k-1}, \lambda_2 = p_k,\) and \(\lambda_3 = \text{id}_{V_{k+1}}\).

Proof. It can be directly verified that the given vector spaces and linear maps form universal cones or co-cones. \(\square\)

3.4.4 Properties of Extroversion and Introversion Functors

Recall that \(\mathcal{T}_n^{k, *} = \mathcal{T}_n^{k, \cdot}\) both denote the set of types \(\tau \in \mathcal{T}_n\) which flow either forward or backward at index \(k\). Regardless of the direction in which \(\tau \in \mathcal{T}_n^{k, *}, \mathcal{T}_n^{k, \cdot}\) flows at index \(k\), we can apply either the extroversion or introversion functor to a zigzag module \(V \in \text{Mod}_\tau\). We use the notation \(\mathcal{R}_k^{\ast / \cdot}\) to denote an unspecified choice of either \(\mathcal{R}_k^*\) or \(\mathcal{R}_k^\cdot\). In the next chapter, our choice of the use of \(\mathcal{T}_n^{k, *}, \mathcal{T}_n^{k, \cdot}\) will indicate our intention of applying either \(\mathcal{R}_k^*\) or \(\mathcal{R}_k^\cdot\), respectively.

Proposition 3.4.1. Fix \(k \in \{1, \ldots, n\}\), let \(\tau \in \mathcal{T}_n^{k, *}, \mathcal{T}_n^{k, \cdot}\), and let \(V_1, V_2 \in \text{Mod}_\tau\).
Then \(\mathcal{R}_k^*(V_1 \oplus V_2) \cong \mathcal{R}_k^*(V_1) \oplus \mathcal{R}_k^*(V_2)\) and \(\mathcal{R}_k^\cdot(V_1 \oplus V_2) \cong \mathcal{R}_k^\cdot(V_1) \oplus \mathcal{R}_k^\cdot(V_2)\). These statements generalizes to the sum of any finite number of \(\tau\)-modules.

Proof. We may assume without loss of generality that \(\tau \in \mathcal{T}_n^{k, \rightarrow}\). Then \(V_1, V_2 \in \text{Mod}_\tau\) have the forms

\[
V_1 = \cdots \xleftarrow{p_{k-2}} V_{k-1} \xrightarrow{p_{k-1}} V_k \xrightarrow{p_k} V_{k+1} \xleftarrow{p_{k+1}} \cdots
\]

and

\[
V_2 = \cdots \xrightarrow{q_{k-2}} V_{k-1} \xrightarrow{q_{k-1}} V'_k \xrightarrow{q_k} V'_{k+1} \xleftarrow{q_{k+1}} \cdots
\]
By Fact 3.4.1, we have
\[ R^\star_k(V_1) = \cdots \leftarrow V_{k-1} \leftarrow \text{id} \rightarrow V_{k-1} \rightarrow V_{k+1} \leftarrow \text{id} \rightarrow V_{k+1} \leftarrow \cdots \]
and
\[ R^\star_k(V_2) = \cdots \leftarrow V_{k-1} \leftarrow \text{id} \rightarrow V_{k-1} \rightarrow V_{k+1} \leftarrow \text{id} \rightarrow V_{k+1} \leftarrow \cdots . \]

On the other hand, we have
\[ V_1 \oplus V_2 = \cdots \leftarrow V_{k-1} \leftarrow V_{k-1}' \leftarrow \text{id} \rightarrow V_{k-1} \rightarrow V_{k+1} \leftarrow V_{k+1}' \leftarrow \cdots , \]
so that
\[ R^\star_k(V_1 \oplus V_2) = \cdots \leftarrow V_{k-1} \leftarrow V_{k-1}' \leftarrow \text{id} \rightarrow V_{k-1} \rightarrow V_{k+1} \leftarrow V_{k+1}' \leftarrow \text{id} \rightarrow V_{k+1} \leftarrow \cdots . \]

Using the fact that \((p_k \oplus q_k) \circ (p_{k-1} \oplus q_{k-1}) = (p_k \circ p_{k-1}) \oplus (q_k \circ q_{k-1})\), we evidently have
\[ R^\star_k(V_1 \oplus V_2) \cong R^\star_k(V_1) \oplus R^\star_k(V_2). \] The analogous result for \( R^\star_k \) is proved similarly, and both statements extend to any finite sum of \( \tau \)-modules by induction.

**Corollary 3.4.1.** Fix \( k \in \{1, \ldots, n\} \), let \( \tau \in T_n^{k,\star} = T_n^{k,\bullet} \), and let \( V, W \in \text{Mod}_r \). Suppose that \( W \preceq V \). Then \( R^\star_k(W) \preceq R^\star_k(V) \).

**Proof.** If \( W \preceq V \) then \( V \cong W \oplus U \) for some \( U \in \text{Mod}_r \) so that by Proposition 3.4.1
\[ R^\star_k(W) \cong R^\star_k(W \oplus U) \cong R^\star_k(W) \oplus R^\star_k(U) \] and hence \( R^\star_k(W) \preceq R^\star_k(V) \). □

**Corollary 3.4.2.** Fix \( k \in \{1, \ldots, n\} \), let \( \tau \in T_n^{k,\star} = T_n^{k,\bullet} \), and let \( V \in \text{Mod}_r \). Then we have
\[ R^\star_k(V) \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} R^\star_k(L_r([b, d])). \]
Proof. Write $V \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} I_{\tau}([b,d])$. Using Propositions 3.4.1 we have

$\mathcal{R}_k^{*/} (V) \cong \mathcal{R}_k^{*/} \left( \bigoplus_{(b,d) \in \text{Dgm}(V)} I_{\tau}([b,d]) \right) \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} \mathcal{R}_k^{*/} (I_{\tau}([b,d])).$

\[\square\]

**Theorem 3.4.1.** Let $\tau \in \mathcal{T}_n^{k,*} = \mathcal{T}_n^{k,*}$:

(a) Let $k \in \{2, \ldots, n-1\}$ and let $\tau' = \psi_k \tau$. The extroversion functor $R_k^*$ acts as follows on the interval $\tau$-modules:

<table>
<thead>
<tr>
<th>$I_{\tau}([k,k])$</th>
<th>$\xrightarrow{R_k^*} O_{\tau'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{\tau}([b,k-1])$</td>
<td>$\rightarrow I_{\tau'}([b,k])$ for $b \leq k - 1$</td>
</tr>
<tr>
<td>$I_{\tau}([k,d])$</td>
<td>$\rightarrow I_{\tau'}([k+1,d])$ for $d \geq k + 1$</td>
</tr>
<tr>
<td>$I_{\tau}([b,d])$</td>
<td>$\rightarrow I_{\tau'}([b,d])$ otherwise</td>
</tr>
</tbody>
</table>

(b) In the exceptional cases $k = 1$ and $k = n$ we have

<table>
<thead>
<tr>
<th>$I_{\tau}([1,1])$</th>
<th>$\xrightarrow{R_1^*} O_{\tau'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{\tau}([1,d])$</td>
<td>$\rightarrow I_{\psi_1 \tau}([2,d])$ for $d \geq 2$</td>
</tr>
<tr>
<td>$I_{\tau}([b,d])$</td>
<td>$\rightarrow I_{\psi_1 \tau}([b,d])$ otherwise</td>
</tr>
</tbody>
</table>

and

<table>
<thead>
<tr>
<th>$I_{\tau}([n,n])$</th>
<th>$\xrightarrow{R_n^*} O_{\tau'}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{\tau}([b,n-1])$</td>
<td>$\rightarrow I_{\psi_n \tau}([b,n])$ for $b \leq n - 1$.</td>
</tr>
<tr>
<td>$I_{\tau}([b,d])$</td>
<td>$\rightarrow I_{\psi_n \tau}([b,d])$ otherwise.</td>
</tr>
</tbody>
</table>
(2.) (a) Let \( k \in \{2, \ldots, n-1\} \) and let \( \tau' = \xi_k \tau \). The introversion functor \( R^*_k \) acts as follows on the interval \( \tau \)-modules:

\[
I_\tau([k,k]) \xrightarrow{R^*_k} O_{\tau'} \\
I_\tau([b,k]) \rightarrow I_{\tau'}([b,k-1]) \quad \text{for } b \leq k-1 \\
I_\tau([k+1,d]) \rightarrow I_{\tau'}([k,d]) \quad \text{for } d \geq k+1 \\
I_\tau([b,d]) \rightarrow I_{\tau'}([b,d]) \quad \text{otherwise.}
\]

(b) In the exceptional cases \( k = 1 \) and \( k = n \) we have

\[
I_\tau([1,1]) \xrightarrow{R^*_1} O_{\tau'} \\
I_\tau([2,d]) \rightarrow I_{\xi_1 \tau}([1,d]) \quad \text{for } d \geq 2 \\
I_\tau([b,d]) \rightarrow I_{\xi_1 \tau}([b,d]) \quad \text{otherwise.}
\]

and

\[
I_\tau([n,n]) \xrightarrow{R^*_n} O_{\tau'} \\
I_\tau([b,n]) \rightarrow I_{\xi_n \tau}([b,n-1]) \quad \text{for } b \leq n-1. \\
I_\tau([b,d]) \rightarrow I_{\xi_n \tau}([b,d]) \quad \text{otherwise.}
\]

Proof. The proof is just a straightforward application of Fact 3.4.1 to each interval \( \tau \)-module. \(\square\)
Figure 3.3: Points in the persistence diagram of a zigzag module $\mathcal{V}$ move according to the arrows when the extroversion functor $\mathcal{R}_k^\ast$ (left) or introversion functor $\mathcal{R}_k^\ast$ (right) are applied. The point $(k, k)$ corresponding to the summand $\mathcal{I}([k, k])$ is annihilated by both of these functors.
Chapter 4: The Reflection Distance

4.1 Transformations of Zigzag Modules

Recall that for $n \in \mathbb{N}$, $n\text{-Mod} = \bigcup_{\tau \in \mathcal{T}_n} \text{Mod}_\tau$ denotes the collection of all zigzag modules of length $n$. We will first identify zigzag modules $V, W \in n\text{-Mod}$ which differ only in the direction of arrows representing isomorphisms. For example, we want to regard the following length 4 zigzag modules

\[
\begin{array}{c}
0 & \xrightarrow{0} & F & \xrightarrow{id} & F & \xrightarrow{0} & 0 \\
0 & \xrightarrow{0} & F & \xleftarrow{id} & F & \xrightarrow{0} & 0
\end{array}
\]

as being the same. The goal of this section is to establish notation for dealing with zigzag modules which are to be regarded as equivalent.

4.1.1 Arrow Reversals

Fix $\tau \in \mathcal{T}_n$. For each $k \in \{1, \ldots, n-1\}$ we define $\text{Mod}^{\text{iso},k}_\tau \subset \text{Mod}_\tau$ by setting

\[
\text{Mod}^{\text{iso},k}_\tau := \{ V = (V_i, p_i) \mid p_k \text{ is an isomorphism.} \}.
\]

Recall the type reversal map $r_k : \mathcal{T}_n \to \mathcal{T}_n$ which reverses the $k$-th arrow of a given $\tau \in \mathcal{T}_n$. We now define a map

\[
\mathcal{A}_k : \text{Mod}^{\text{iso},k}_\tau \to \text{Mod}^{\text{iso},k}_{r_k \tau}
\]

by setting $\mathcal{A}_k(V) = (V_i, q_i) \in \text{Mod}_{r_k \tau}$, where $q_i = p_i$ for $i \neq k$ and $q_k = p_k^{-1}$.

58
Remark 4.1.1. If $V_i \xrightarrow{p} V_i'$ is an isomorphism appearing in $V$ then $V$ is isomorphic to the zigzag module in which the arrow $V_i \xrightarrow{p} V_i'$ is replaced by the arrow $V_i \xrightarrow{id} V_i$. Thus we may assume that all isomorphisms are the identity map, in which case $A_k(\mathbb{V})$ is obtained by changing the type of $\mathbb{V}$ from $\tau$ to $r_k \tau$ but leaving all of the vector spaces $V_i$ and linear maps $p_i$ unchanged.

Remark 4.1.2. Note that $A_k \circ A_k = \text{id}_{\text{Mod}_{\text{iso,k}}} \text{Mod}$ since reversing the direction of an isomorphism twice leaves the zigzag module unchanged.

Remark 4.1.3. Arrow reversals operate on “arrows”, i.e., they operate on the linear maps defining a zigzag module. In comparison, the reflection functors operate on “vertices” and adjacent arrows.

Fact 4.1.1. Let $\tau \in \mathcal{T}_n$ and suppose that $V, W \in \text{Mod}_{\text{iso,i}}$ for some $i \in \{1, \ldots, n-1\}$ with $V \cong W$. Then

$$A_i(V) \cong A_i(W).$$

Proof. It can be checked that any isomorphism $\phi : V \to W$ also serves as an isomorphism between $A_i(V)$ and $A_i(W)$. \hfill \Box

Proposition 4.1.1. Let $\tau \in \mathcal{T}_n$ and suppose that $V, W \in \text{Mod}_{\text{iso,i}}$ for some $i \in \{1, \ldots, n-1\}$. Then

$$A_i(V \oplus W) = A_i(V) \oplus A_i(W).$$

Proof. The result follows immediately after noticing that $(p_i \oplus q_i)^{-1} = p_i^{-1} \oplus q_i^{-1}$, where $p_i$ and $q_i$ denote the $i$-th structure maps of $V$ and $W$, respectively. \hfill \Box

Proposition 4.1.2. For $\tau \in \mathcal{T}_n$, $i, j \in \{1, \ldots, n-1\}$, and $V = (V_i, p_i) \in \text{Mod}_{\text{iso,i}} \cap \text{Mod}_{\text{iso,j}}$, we have

$$A_i A_j(V) = A_j A_i(V).$$
Proof. If \( i = j \) then the result follows immediately. If \( i \neq j \), then the result follows by noting that \( A_i \) and \( A_j \) operate on different linear maps \( p_i \) and \( p_j \), so that the order in which they are applied does not matter.

\[ \]  

Proposition 4.1.3. Let \( \tau \in T_n \) and suppose that \( V, W \in \text{Mod}_\tau \) with \( W \preceq V \). If \( V \in \text{Mod}^{\text{iso},i}_\tau \) for some \( i \in \{1, \ldots, n-1\} \) then \( W \in \text{Mod}^{\text{iso},i}_\tau \) and  
\[ A_i(W) \preceq A_i(V). \]

Proof. If \( W \preceq V \in \text{Mod}^{\text{iso},i}_\tau \) then the \( i \)-th structure map of \( V \) is an isomorphism so that, by Proposition 3.2.1, the \( i \)-th structure map of \( W \) is an isomorphism as well and hence \( W \in V \in \text{Mod}^{\text{iso},i}_\tau \). Now if \( U \in \text{Mod}_\tau \) is such that \( V \cong W \oplus U \) then by Proposition 4.1.1 we have  
\[ A_i(V) \cong A_i(W \oplus U) = A_i(W) \oplus A_i(U), \]
and thus \( A_i(W) \preceq A_i(V). \)

4.1.2 Equivalence of Zigzag Modules which Differ by an Arrow Reversal

We now define an equivalence relation on \( n \)-Mod, formalizing our discussion at the beginning of this section. For \( V, W \in n \text{-Mod} \) we write \( V \sim W \) if and only if either \( V \cong W \) or there is a finite sequence \( k_1, \ldots, k_j \) of indices in \( \{1, \ldots, n-1\} \) such that  
\[ W \cong A_{k_j} A_{k_{j-1}} \cdots A_{k_1}(V). \]

In words, \( V \sim W \) if \( W \) can be obtained, up to isomorphism, from \( V \) by reversing some (possibly empty) set of arrows representing isomorphisms. Using Fact 4.1.1 and
Proposition 4.1.2, it is straightforward to check that this is in fact defines an equivalence relation on the isomorphism classes of \( n \)-Mod. We will denote the equivalence class of \( V \) under the equivalence relation \( \sim \) by \( \tilde{V} \).

**Remark 4.1.4.** Note that for any \( \tau, \tau' \in T_n \), \( O_\tau \sim O_{\tau'} \). Hence, in what follows we drop the subscript indicating type and denote the zero module of any type by \( O \).

The proof of the next proposition is sketch out by Oudot in [18]; we give the full details here:

**Proposition 4.1.4 ([18]).** Let \( V_1, V_2 \in n\text{-Mod} \). If \( V_1 \sim V_2 \) then

\[
Dgm(V_1) = Dgm(V_2).
\]

**Proof.** Let \( \tau = \text{type}(V_1) \) and write \( V_1 \cong \bigoplus_{(b,d) \in Dgm(V)} I_\tau([b,d]) \). Since \( V_1 \sim V_2 \), there is a sequence of indices \( k_1, k_2, \ldots, k_j \in \{1, \ldots, n - 1\} \) such that

\[
V_2 \cong A_{k_j}A_{k_j-1}\cdots A_{k_1}(V_1).
\]

Let \( \tau' \in T_n \) be defined by \( r_{k_j}r_{k_j-1}\cdots r_{k_1} \tau \) and consider the zigzag module \( W \in \text{Mod}_{\tau'} \) defined by

\[
W := \bigoplus_{(b,d) \in Dgm(V)} I_{\tau'}([b,d]).
\]

By definition of \( W \), we have \( Dgm(W) = Dgm(V_1) \). We claim that \( W \cong V_2 \). To see this, notice that if \( V_i \xrightarrow{p} V_j \) is a structure map of \( V_1 \), with \( p \) being an isomorphism and \( i, j \) being consecutive integers in \( \{1, \ldots, n\} \), and if \( d := \dim(V_i) = \dim(V_j) \), then there is an isomorphism \( \psi : V_i \rightarrow \bigoplus_{m=1}^d \mathbb{F} \) such that the diagram

\[
\begin{array}{ccc}
V_i & \xrightarrow{p} & V_j \\
\downarrow{\psi} & & \downarrow{\psi op^{-1}} \\
\bigoplus_{m=1}^d \mathbb{F} & \xrightarrow{id} & \bigoplus_{m=1}^d \mathbb{F}
\end{array}
\]

61
commutes. This is true if and only if the diagram
\[
\begin{array}{c}
V_i \xleftarrow{p_i^{-1}} V_j \\
\downarrow \psi \quad \quad \quad \downarrow \psi \circ p_i^{-1} \\
\bigoplus_{m=1}^d \ker F \xleftarrow{id} \bigoplus_{m=1}^d \ker F
\end{array}
\]
commutes. Applying this principal to every square at which an arrow reversal is
applied, we see that \( W \cong V_2 \). Hence \( \text{Dgm}(V_1) = \text{Dgm}(W) = \text{Dgm}(V_2) \) by Theorem
3.2.1.

**Proposition 4.1.5.** Let \( k \in \{1, \ldots, n\} \), let \( \circ \in \{+,-,\ast,\bullet\} \), and let \( \tau \in \mathcal{T}_n^{k,\circ} \). Suppose that \( V = (V_i, p_i) \in \text{Mod}_{\tau}^{\text{iso},j} \) for some \( j \in \{1, \ldots, n-1\} \) with \( j \not\in \{k-1, k\} \). Then
\[
\mathcal{A}_j \mathcal{R}_k^\circ(V) = \mathcal{R}_k^\circ \mathcal{A}_j(V).
\]

**Proof.** The assumptions made on the indices \( k \) and \( j \) assure that the maps \( \mathcal{A}_j \) and \( \mathcal{R}_k^\circ \) do not alter the same structure maps \( p_i \), and so the order in which they are applied to \( V \) is inconsequential.

**Corollary 4.1.1.** Let \( k \in \{1, \ldots, n\} \), \( \circ \in \{+,-,\ast,\bullet\} \), \( \tau \in \mathcal{T}_n^{k,\circ} \), and \( V_1, V_2 \in \text{Mod}_{\tau} \). If \( V_1 \sim V_2 \) then
\[
\mathcal{R}_k^\circ(V_1) \sim \mathcal{R}_k^\circ(V_2).
\]

**Proof.** Since \( V_1 \sim V_2 \), there is a sequence \( k_1, \ldots, k_j \in \{1, \ldots, n-1\} \) such that \( V_2 \cong \mathcal{A}_{k_1} \mathcal{A}_{k_{\ell-1}} \cdots \mathcal{A}_{k_1}(V_1) \). Since \( V_1 \) and \( V_2 \) both have either a sink, source, or flow in the same direction at index \( k \), we may assume without loss of generality that \( k_i \not\in \{k-1, k\} \) for all \( i = 1, \ldots, \ell \). Then, applying Proposition 4.1.5 \( \ell \) times, we have
\[
\mathcal{R}_k^\circ(V_2) \cong \mathcal{R}_k^\circ(\mathcal{A}_{k_1} \mathcal{A}_{k_{\ell-1}} \cdots \mathcal{A}_{k_1}(V_1))
\]
\[
\cong \mathcal{A}_{k_1} \mathcal{A}_{k_{\ell-1}} \cdots \mathcal{A}_{k_1}(\mathcal{R}_k^\circ(V_1)),
\]
62
so that $\mathcal{R}_k^o(V_1) \sim \mathcal{R}_k^o(V_2)$. \hfill \square

**Definition 4.1.1.** We define a relation $\preceq$ on $n$-Mod by declaring $W \preceq V$ if and only if there exists a zigzag module $W' \in n$-Mod with $W \sim W'$ and $W' \preceq V$.

In words, $W \preceq V$ if we can obtain a summand of $V$ by reversing any number of the arrows of $W$ representing isomorphisms.

**Fact 4.1.2.** $\preceq$ is a preorder on $n$-Mod. Moreover, $V \sim W$ if and only if $W \preceq V$ and $V \preceq W$ so that $\preceq$ induces a partial order on $n$-Mod/ $\sim$.

**Proof.** Since $V \sim V$ and $V \leq V$, we have $V \preceq V$. If $V_1 \preceq V_2$ and $V_2 \preceq V_3$ then there are zigzag modules $W_1$ and $W_2$ such that $V_1 \sim W_1 \preceq V_2$ and $V_2 \sim W_2 \preceq V_3$. It follows from Proposition 4.1.1 that $V_1 \sim W_2'$ for some $W_2' \preceq W_2 \preceq V_3$. By transitivity of the relation $\preceq$, we have $V_1 \sim W_2' \preceq V_3$ so that $V_1 \preceq V_3$. This shows that $\preceq$ is a preorder.

Now if $V \sim W$ then $V \sim W \preceq W$ and $W \sim V \preceq V$ so that $W \preceq V$ and $V \preceq W$. Conversely, if $W \preceq V$ and $V \preceq W$ then there exists zigzag modules $W' \preceq W$ and $V' \preceq V$ such that $W \sim V'$ and $V \sim W'$. Let $X_1, X_2 \in n$-Mod be such that

$$V \cong V' \oplus X_1 \quad \text{and} \quad W \cong W' \oplus X_2.$$ 

Since $W \sim V'$, there is a composition of arrow reversals $A$ such that $A(W) \cong V'$, and similarly there is a composition of arrow reversals $B$ such that $B(V) \cong W'$. Hence
\( W \cong W' \oplus X_2 \cong B(V) \oplus X_2 \). Then we have

\[
V' \cong A(W) \\
\cong A(B(V) \oplus X_2) \\
\cong A(B(V)) \oplus A(X_2) \\
\cong A(B(V' \oplus X_1)) \oplus A(X_2) \\
\cong A(B(V')) \oplus A(B(X_1)) \oplus A(X_2).
\]

Using the above isomorphisms together with Theorem 3.2.2 and Proposition 4.1.4, we have

\[
Dgm(V') = Dgm(A(B(V')) \oplus A(B(X_1)) \oplus A(X_2)) \\
= Dgm(A(B(V'))) \sqcup Dgm(A(B(X_1))) \sqcup Dgm(A(X_2)) \\
= Dgm(V') \sqcup Dgm(X_1) \sqcup Dgm(X_2)
\]

so that \( Dgm(X_1) = \emptyset = Dgm(X_2) \). Hence \( X_1 \sim O \sim X_2 \) so that in fact have

\( V \cong V' \) and \( W \cong W' \).

Thus \( V \sim W \), completing the proof.

4.1.3 Sequences of Compatible Functors

Consider now a finite sequence \( \mathcal{R}^{s_1}_{k_1}, \mathcal{R}^{s_2}_{k_2}, \ldots, \mathcal{R}^{s_\ell}_{k_\ell} \) of functors, where \( s_j \in \{+, -, \star, \bullet\} \) and \( k_j \in \{1, \ldots, n\} \) for each \( j \). We say that the composition \( \mathcal{R} = \mathcal{R}^{s_\ell}_{k_\ell} \circ \mathcal{R}^{s_{\ell-1}}_{k_{\ell-1}} \circ \cdots \circ \mathcal{R}^{s_1}_{k_1} \) and a zigzag module \( V \in n\text{-Mod} \) are compatible with each other if we can apply the reflections \( \mathcal{R}^{s_j}_{k_j} \) inductively as follows:

1. Suppose there exists some \( \tau_1 \in \mathcal{T}^{k_1,s_1}_n \) and \( V_1 \in \text{Mod}_{\tau_1} \) with \( V_1 \sim V \). In this case we define \( \mathcal{R}^{s_1}_{k_1}(V) := \mathcal{R}^{s_1}_{k_1}(V_1) \).
2. Assume that \( R_{s_{j-1}^{j-1}} \cdots R_{s_{1}^{1}}(V) \) has been defined for \( j \in \{2, \ldots, \ell\} \) and suppose there exists some \( \tau_j \in \mathcal{T}^{k_j,s_j}_n \) and \( V_j \in \text{Mod}_{r_j} \) with \( V_j \sim R_{s_{j-1}^{j-1}} \cdots R_{s_{1}^{1}}(V) \).

In this case we define \( R_{s_j^{jk_j}} \cdots R_{s_1^{k_1}}(V) := R_{s_j^{jk_j}}(V_j) \).

By Corollary 4.1.1 modulo the relation \( \sim \), the result of applying a compatible composition of functors to \( V \) does not depend on the choice of \( V_1 \ldots, V_\ell \). As a result, if \( W \) and \( W' \) are both obtained from \( V \) by applying a composition of compatible functors then \( W \sim W' \).

We define

\[
\text{comp}(\mathcal{R}) := \{ V \in n\text{-Mod} \mid \mathcal{R} \text{ and } V \text{ are compatible} \}.
\]

so that \( \mathcal{R} : \text{comp}(\mathcal{R}) \to n\text{-Mod} \). If \( \mathcal{R}_1 \) and \( \mathcal{R}_2 \) are arbitrary compositions of functors then we have \( \text{comp}(\mathcal{R}_2 \circ \mathcal{R}_1) = \text{comp}(\mathcal{R}_2) \cap \text{comp}(\mathcal{R}_1) \).

The second part of Proposition 3.3.3 has an analogue for the relation \( \succeq \):

**Proposition 4.1.6.** Let \( k \in \{1, \ldots, n\} \), \( \circ \in \{+, -, \ast, \bullet\} \), \( \tau \in \mathcal{T}_n^{k,\circ} \), and \( W \in \text{Mod}_r \).

If \( V \succeq W \) then \( \mathcal{R}_k^\circ(V) \) is well-defined and \( \mathcal{R}_k^\circ(V) \succeq \mathcal{R}_k^\circ(W) \). More generally, if \( \mathcal{R} \) is any composition of functors with \( W \in \text{comp}(\mathcal{R}) \) then \( V \in \text{comp}(\mathcal{R}) \) and \( \mathcal{R}(V) \succeq \mathcal{R}(W) \).

**Proof.** Since \( V \succeq W \), there exists a zigzag module \( V_1 \in \text{Mod}_r \) such that \( V_1 \sim V \) and \( V_1 \preceq W \). In particular, \( \mathcal{R}_k^\circ(V) = \mathcal{R}_k^\circ(V_1) \) is defined by the construction given above. By Corollaries 3.3.1 and 3.4.1 we have \( \mathcal{R}_k^\circ(V_1) \preceq \mathcal{R}_k^\circ(W) \) so that \( \mathcal{R}_k^\circ(V) \succeq \mathcal{R}_k^\circ(W) \).

Iterating this gives the second statement. \( \square \)
4.2 The Reflection Distance

Let \( V \in n\text{-Mod} \) and let \( R = R_{k_t}^{\ell_t} \circ R_{k_{t-1}}^{\ell_{t-1}} \circ \cdots \circ R_{k_1}^{\ell_1} : \text{comp}(R) \to n\text{-Mod} \) be a non-trivial composition of \( \ell \) functors. For each \( p \in [1, \infty] \) define the \( p \)-cost of \( R \) by

\[
C_p(R) := \begin{cases} 
\ell^{1/p}, & \text{for } p \in [1, \infty) \\
1, & \text{for } p = \infty
\end{cases}
\]

and define \( C_p(id_V) = 0 \) for each \( V \in n\text{-Mod} \).

For each \( p \in [1, \infty] \), we define the function \( d^p_R : n\text{-Mod} \times n\text{-Mod} \to \mathbb{R} \) by setting

\[
d^p_R(V, W) := \min \{ \max \{ C_p(R_1), C_p(R_2) \} \mid R_1(V) \preceq W \text{ and } R_2(W) \preceq V \},
\]

where the minimum is taken over pairs \( (R_1, R_2) \) with \( R_1 \) and \( R_2 \) being compositions of functors with \( V \in \text{comp}(R_1) \) and \( W \in \text{comp}(R_2) \).

**Fact 4.2.1.** For any \( p \in [1, \infty] \) we have \( C_p(R_2 \circ R_1) \leq C_p(R_1) + C_p(R_2) \).

**Proof.** This is trivial when \( p = \infty \). For \( p \in [1, \infty) \), this follows from the inequality \((a + b)^{1/p} \leq a^{1/p} + b^{1/p}\) for all nonnegative integers \( a, b \).

**Proposition 4.2.1.** For any \( V \in n\text{-Mod} \) there exists a composition of functors \( R \) with \( V \in \text{comp}(R) \) such that \( R(V) = O \).

**Proof.** Let \( \tau = \text{type}(V) \). It suffices to show that for any interval module \( I_\tau([b, d]) \), there is a composition of functors \( R \) which is compatible with \( V \) and such that \( R(I_\tau([b, d])) = O \). For if this is the case then we can iteratively annihilate interval summands of \( V \) until we arrive at \( O \).

We demonstrate that this can be done as follows: the interval \( \tau \)-module \( I_\tau([b, d]) \) has either a sink, source, forward flow, or backward flow at index \( d \). We can thus apply either a sink reflection functor, source reflection functor, introversion functor,
or extroversion functor, respectively, to obtain an interval module supported over \([b, d - 1]\). In this way, we obtain a composition \(R_1 = R_b^{s_{d-b+1}} \circ R_b^{s_{d-b}} \circ \cdots \circ R_d^{s_2} \circ R_d^{s_1}\) with \(V \in \text{comp}(R_1)\) such that \(I_r([b, d]) = 0\). Apply this procedure again to an interval summand of \(R_1(V)\) to obtain a composition of functors \(R_2\) annihilating this summand. In this way, we obtain a composition \(R = R_m \circ \cdots \circ R_1\), where \(m \leq |\text{Pers}(V)|\). By construction, we have \(R(V) = 0\).

\[\text{Remark 4.2.1.} \quad \text{In light of Theorems 3.3.1 and 3.4.1, it is clear that it takes at least } d - b + 1 \text{ moves to annihilate an interval module supported over } [b, d].\]

\[\text{Theorem 4.2.1.} \quad \text{For each } n \in \mathbb{N} \text{ and for all } p \in [1, \infty], \text{ the function } d^p_R \text{ is a pseudo-metric on } n\text{-Mod}. \text{ Moreover, } d^p_R(V, W) = 0 \text{ if and only if } V \sim W \text{ so that } d^p_R \text{ induces a metric on } n\text{-Mod}/\sim.\]

\[\text{Proof.} \quad \text{The facts that } d^p_R \text{ is a non-negative and symmetric function follows immediately. If } V = W \text{ then certainly } V \preceq W \text{ and } W \preceq V \text{ so that } d^p_R(V, W) = 0 \text{ and hence } d^p_R(V, V) = 0 \text{ for all } V \in n\text{-Mod}.\]

Next we verify the triangle inequality. Let \(V_1, V_2, V_3 \in n\text{-Mod}. \) Consider first the case \(p \in [1, \infty)\) and suppose that \(d^p_R(V_1, V_3) = m_3^{1/p}, d^p_R(V_1, V_2) = m_2^{1/p},\) and \(d^p_R(V_2, V_3) = m_1^{1/p}\) for some non-negative integers \(m_1, m_2, \) and \(m_3.\) Then there exists compositions of functors \(R_1\) and \(R_2,\) with \(V_1 \in \text{comp}(R_1), V_2 \in \text{comp}(R_2)\) such that

\[R_1(V_1) \preceq V_2, \quad R_2(V_2) \preceq V_1 \quad \text{with} \quad C_p(R_1) \leq m_2^{1/p}, \quad C_p(R_2) \leq m_2^{1/p}.\]

Similarly, there exists compositions of functors \(R'_1\) and \(R'_2,\) with \(V_2 \in \text{comp}(R'_1), V_3 \in \text{comp}(R'_2)\) such that

\[R'_1(V_2) \preceq V_3, \quad R'_2(V_3) \preceq V_2 \quad \text{with} \quad C_p(R'_1) \leq m_3^{1/p}, \quad C_p(R'_2) \leq m_3^{1/p}.\]
Then by Proposition 4.1.6 $\mathcal{R}_1' \circ \mathcal{R}_1(V_1) \not\preceq \mathcal{R}_1'(V_2) \not\preceq V_3$. Similarly, $\mathcal{R}_2 \circ \mathcal{R}_2'(V_3) \not\preceq \mathcal{R}_2(V_2) \not\preceq V_1$. Hence, by transitivity of the preorder $\not\preceq$, we have

$$\mathcal{R}_1' \circ \mathcal{R}_1(V_1) \not\preceq V_3 \quad \text{and} \quad \mathcal{R}_2 \circ \mathcal{R}_2'(V_3) \not\preceq V_1. \quad (4.1)$$

By Fact 4.2.1 $C_p(\mathcal{R}_1' \circ \mathcal{R}_1) \leq C_p(\mathcal{R}_1') + C_p(\mathcal{R}_1)$ and $C_p(\mathcal{R}_2 \circ \mathcal{R}_2') \leq C_p(\mathcal{R}_2) + C_p(\mathcal{R}_2')$ so that

$$C_p(\mathcal{R}_1' \circ \mathcal{R}_1) \leq m_2^{1/p} + m_3^{1/p} \quad \text{and} \quad C_p(\mathcal{R}_2 \circ \mathcal{R}_2') \leq m_2^{1/p} + m_3^{1/p}. \quad (4.2)$$

Lines (4.1) and (4.2) together imply

$$d_p^p(V_1, V_3) \leq m_2^{1/p} + m_3^{1/p} = d_p^p(V_1, V_2) + d_p^p(V_2, V_3).$$

The case $p = \infty$ follows easily by using the fact that $d_p^\infty$ takes values in $\{0, 1\}$ and checking that no choice of $V_1, V_2, V_3$ can break the triangle inequality.

Now if $V \sim W$ then, by Fact 4.1.2, $V \not\preceq W$ and $W \not\preceq V$ so that $d_p^p(V, W) = 0$. Conversely, if $V \not\sim W$ then either $V \not\preceq W$ or $W \not\preceq V$. In any case, we must apply some non-trivial reflection to either $V$ or $W$, incurring a non-zero cost so that $d_p^p(V, W) > 0$. □

4.3 The Bottleneck Distance and Reflective Stability

4.3.1 Matchings

Let $S$ and $T$ be sets. A matching between $S$ and $T$ is a relation $M \subseteq S \times T$ such that

1. For any $s \in S$, there is at most one $t \in T$ such that $(s, t) \in M$,

2. For any $t \in T$, there is a most one $s \in S$ such that $(s, t) \in M$. 

68
We denote a matching by $M : S \rightarrow T$. Equivalently, a matching is a bijection $M : S' \rightarrow T'$ for some subsets $S' \subseteq S$ and $T' \subseteq T$. From this point of view, $S'$ is called the coimage of $M$, denoted $S' = \text{coim}(M)$, and $T'$ is called the image of image of $M$, denoted $T' = \text{im}(M)$. A matching is said to finite if $|\text{coim}(M)| = |\text{im}(M)|$ is finite. If $(s, t) \in M$ then $s$ and $t$ are said to be matched. Points in $S \sqcup T$ which are not matched are said to be unmatched.

**Lemma 4.3.1** (Matching Lemma). Let $S$ and $T$ be sets and let $f : S \rightarrow T$ and $g : T \rightarrow S$ be finite matchings. Then there exists a matching $M : S \rightarrow T$ such that

1. $\text{coim}(f) \subseteq \text{coim}(M)$,
2. $\text{coim}(g) \subseteq \text{im}(M)$,
3. If $M(s) = t$ then either $f(s) = t$ or $g(t) = s$.

**Proof.** The proof is inspired by a proof of the Cantor–Schröder–Bernstein theorem given in [19]. Let $s \in \text{coim}(f)$, $t \in \text{coim}(g)$, and consider sequences of the form

$$\cdots \xrightarrow{g} f^{-1}(g^{-1}(s)) \xrightarrow{f} g^{-1}(s) \xrightarrow{g} s \xrightarrow{f} f(s) \xrightarrow{g} g(f(s)) \xrightarrow{f} \cdots$$

and

$$\cdots \xrightarrow{f} g^{-1}(f^{-1}(t)) \xrightarrow{g} f^{-1}(t) \xrightarrow{f} t \xrightarrow{g} g(t) \xrightarrow{f} f(g(t)) \xrightarrow{g} \cdots,$$

where we allow these sequences to terminate to the right or left when undefined. We refer to these sequences as the orbits of $s$ or $t$. Since $f$ and $g$ are finite matchings, such sequences either terminate on both the left and right, or are infinite but periodic. By injectivity of $f$ and $g$, every element of $\text{coim}(f) \sqcup \text{coim}(g)$ appears in exactly one orbit. Moreover, every orbit falls into one of the following five classes:
1. \( s \rightarrow t \rightarrow \cdots \rightarrow s \rightarrow y, \)

2. \( s \rightarrow t \rightarrow \cdots \rightarrow s \rightarrow t \rightarrow x, \)

3. \( t \rightarrow s \rightarrow \cdots \rightarrow t \rightarrow x, \)

4. \( t \rightarrow s \rightarrow \cdots \rightarrow t \rightarrow s \rightarrow y, \)

5. \( s \rightarrow t \rightarrow \cdots \rightarrow s \rightarrow t \rightarrow t \rightarrow x, \)

where the \( s \)'s and \( t \)'s represent elements of \( \text{coim}(f) \) and \( \text{coim}(g) \), respectively, arrows represent either \( f \) or \( g \), and \( y \)'s and \( x \)'s represent elements of \( \text{im}(f) \setminus \text{coim}(g) \) and \( \text{im}(g) \setminus \text{coim}(f) \), respectively. We define a matching \( M : S \rightarrow T \) as follows: for each \( i = 1, \ldots, 5 \) let

\[
S_i^{\text{co}} := \{ s \in \text{coim}(f) \mid s \text{ appears in an orbit of type } i \}
\]

and

\[
T_i^{\text{co}} := \{ t \in \text{coim}(g) \mid t \text{ appears in an orbit of type } i \}.\]

Then \( \text{coim}(f) = \bigcup S_i^{\text{co}} \) and \( \text{coim}(g) = \bigcup T_i^{\text{co}} \). We partition \( S_1^{\text{co}} \) and \( T_3^{\text{co}} \) further by defining

\[
S_1^{\text{co}\setminus p} := \{ s \in S_1^{\text{co}} \mid f(s) \in \text{coim}(g) \}, \quad S_1^p := \{ s \in S_1^{\text{co}} \mid f(s) \notin \text{coim}(g) \},
\]

\[
T_3^{\text{co}\setminus p} := \{ t \in T_3^{\text{co}} \mid g(t) \in \text{coim}(f) \}, \quad \text{and} \quad T_3^p := \{ t \in T_3^{\text{co}} \mid g(t) \notin \text{coim}(f) \},
\]

so that \( S_1 = S_1^{\text{co}\setminus p} \cup S_1^p \) and \( T_3 = T_3^{\text{co}\setminus p} \cup T_3^p \) (here, the letter “\( p \)” stands for “penultimate”). The image, coimage, and mapping of \( M \) is specified by the diagram

\[1\]

The idea of partitioning the image and coimage of our matching and specifying the matching on the parts diagrammatically is inspired by [?], where the authors define a matching in the same way.
coim(M) := \begin{align*}
S_1^\text{co} &\cup S_2^\text{co} \cup S_3^\text{co} \cup S_4^\text{co} \cup S_5^\text{co} \cup S_1^\text{p} \cup g(T_3^p)
\end{align*}
im(M) := \begin{align*}
T_1^\text{co} &\cup T_2^\text{co} \cup T_3^\text{co} \cup T_4^\text{co} \cup T_5^\text{co} \cup T_3^p \cup f(S_1^p)
\end{align*}

Note that \( f(S_1^\text{co}) = T_1^\text{co} \), \( f(S_2^\text{co}) = T_2^\text{co} \), \( g(T_3^p) = S_3^\text{co} \), \( g(T_4^\text{co}) = S_4^\text{co} \), and \( f(S_5^\text{co}) = T_5^\text{co} \) so that \( M \) is surjective. Since \( f \) and \( g \) are injective, we see that \( M \) is a bijection between each of the parts specified and thus defines a matching. Moreover, \( \text{coim}(M) = \text{coim}(f) \cup \text{g}(T_3^p) \supseteq \text{coim}(f) \) and \( \text{im}(M) = \text{coim}(g) \cup f(S_1^p) \supseteq \text{coim}(g) \).

Property (3) evidently holds by the definition of \( M \).

\[ \square \]

### 4.3.2 The Bottleneck Distance

Let \( S \) and \( T \) be multiset representations with underlying set \( \mathbb{R}^2 \) and let \( M : S \rightarrow T \) be a matching. For each \( p \in [1, \infty] \) define

\[
c_p(M) := \max \left\{ \max_{(s,t) \in M} \| s - t \|_p, \max_{r \in S \cup T \text{ unmatched}} \frac{|r_y - r_x|}{2^{1-1/p}} \right\},
\]

where we use the convention \( 1/\infty = 0 \), where \( r_x \) and \( r_y \) denote the \( x \) and \( y \) coordinates, respectively, of the point \( r \in \mathbb{R}^2 \), and \( \| \cdot \|_p : \mathbb{R}^2 \rightarrow \mathbb{R} \) denotes the usual \( \ell^p \)-norm on \( \mathbb{R}^2 \). We then define the \textit{generalized bottleneck distance} between \( S \) and \( T \) by

\[
d^p_b(S, T) := \inf_{M : S \rightarrow T} c_p(M),
\]

where the infimum is taken over all matchings between \( S \) and \( T \).

**Remark 4.3.1.** The ordinary bottleneck distance (as defined in [3]) is just a special case of the generalized bottleneck distance when \( p = \infty \). Since \( \|x\|_\infty \leq \|x\|_1 \) for all \( x \in \mathbb{R}^2 \), we see that \( d^\infty_b(S, T) \leq d^1_b(S, T) \) for all multiset representations \( S \) and \( T \).
Given a multisubset representation $S$ with underlying set $\mathbb{R}^2$, and given $p \in [1, \infty]$ and $\eta > 0$, we define

$$S^\eta_p := \left\{ s \in S \mid \frac{|s_y - s_x|}{2^{1-1/p}} > \eta \right\}.$$

**Lemma 4.3.2.** Fix $p \in [1, \infty)$, let $S$ and $T$ be multisubsets with underlying set $\mathbb{R}^2$, and let $M : S \rightarrow T$ be a matching such that

1. $S^\eta_p \subseteq \text{coim}(M)$,
2. $T^\eta_p \subseteq \text{im}(M)$,
3. if $M(s) = t$ then $\|s - t\|_p \leq \eta$.

Then $c_p(M) \leq \eta$.

**Proof.** Property (3) guarantees that $\max_{(s,t) \in M} \|s - t\|_p \leq \eta$. If $r \in S \sqcup T$ is unmatched then $r \notin S^\eta_p \sqcup T^\eta_p$ so that $\frac{|r_y - r_x|}{2^{1-1/p}} \leq \eta$. That $c_p(M) \leq \eta$ now follows immediately from equation (4.13).

### 4.3.3 Stability of Persistence Diagrams with Respect to the $p = 1$ Reflection Distance

In this section, we show that the 1-bottleneck distance between the persistence diagrams of two given zigzag modules is bounded above by the 1-reflection distance between the zigzag modules themselves.

**Lemma 4.3.3.** Let $\tau \in \mathcal{T}_n$ and $\mathcal{V} \in \text{Mod}_\tau$. If $\mathcal{R}(\mathcal{V}) = O$ then

$$C_1(\mathcal{R}) \geq \max_{(b,d) \in \text{Dgm}(\mathcal{V})} |d - b|.$$

Moreover, if $\mathcal{R}(\mathcal{I}([b,d])) = \mathcal{I}([b',d'])$ then

$$C_1(\mathcal{R}) \geq \| (b,d) - (b',d') \|_1 = |d' - d| + |b' - b|.$$
Proof. We have
\[ O = \mathcal{R}(V) = \mathcal{R} \left( \bigoplus_{(b,d) \in \text{Dgm}(V)} \mathcal{I}_{\tau}([b,d]) \right) \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} \mathcal{R}(\mathcal{I}_\tau([b,d])) \]
so that \( \mathcal{R}(\mathcal{I}_\tau([b,d])) = O \) for all \((b,d) \in \text{Dgm}(V)\). Hence \( C_1(\mathcal{R}) \geq d - b + 1 > |d - b| \) for all \((b,d) \in \text{Dgm}(V)\), proving the first claim.

The second claim should be evident from Figures 3.3.6 and 3.4.4: the point \((b,d) \in \text{Dgm}(\mathcal{I}([b,d]))\) moves along the integer lattice to the point \((b',d')\) as each functor comprising \(\mathcal{R}\) is sequentially applied, each additional application of a functor contributing 1 to \(C_1(\mathcal{R})\) and moving the corresponding point in the diagram to an adjacent node on the integer lattice. \(\square\)

**Theorem 4.3.1.** Let \(V, W \in \tau\text{-Mod}\). Then

\[ d^1_b(Dgm(V), Dgm(W)) \leq d^1_K(V, W). \]

Proof. We show that for any pair of reflections \((\mathcal{R}_1, \mathcal{R}_2)\) with \(V \in \text{comp}(\mathcal{R}_1)\) and \(W \in \text{comp}(\mathcal{R}_2)\) such that \(\mathcal{R}_1(V) \preceq W\) and \(\mathcal{R}_2(W) \preceq V\), there is a matching \(M : \text{Dgm}(W) \not\rightarrow \text{Dgm}(V)\) with \(c_1(M) \leq \max\{C_1(\mathcal{R}_1), C_1(\mathcal{R}_2)\}\).

Let \(\eta = \max\{C_1(\mathcal{R}_1), C_1(\mathcal{R}_2)\}\). Consider the pairs of multisubsets

\[ V_{\eta}^1 = \{ (b,d) \in \text{Dgm}(V) \mid |d - b| > \eta \}, \]
\[ W_{\eta}^1 = \{ (b,d) \in \text{Dgm}(W) \mid |d - b| > \eta \} \]
and

\[ \mathcal{I} = \{ (b,d) \in \text{Dgm}(V) \mid \mathcal{R}_1(\mathcal{I}([b,d])) \neq O \}, \]
\[ \mathcal{J} = \{ (b',d') \in \text{Dgm}(W) \mid \mathcal{R}_2(\mathcal{I}([b',d'])) \neq O \}. \]

Note that by Lemma 4.3.3, \(V_{\eta}^1 \subseteq \mathcal{I}\) and \(W_{\eta}^1 \subseteq \mathcal{J}\).
Let $\alpha_1 : I \to \text{Dgm}(R_1(V))$ and $\alpha_2 : J \to \text{Dgm}(R_2(W))$ be the injections given by $\alpha_i(b,d) = (b',d')$ if and only if $R_i(I[b,d]) = I[b',d']$ for each $i \in \{1, 2\}$. By Theorem 3.2.2 and Proposition 4.1.4, we have $\text{Dgm}(R_1(V)) \subseteq \text{Dgm}(W)$ and $\text{Dgm}(R_2(W)) \subseteq \text{Dgm}(V)$. Let $j_1$ and $j_2$ denote the respective inclusion maps. Then $M_1 := j_1 \circ \alpha_1 : I \to \text{Dgm}(W)$ and $M_2 := j_2 \circ \alpha_2 : J \to \text{Dgm}(V)$ are injective and hence can be viewed as matchings $M_1 : \text{Dgm}(V) \nrightarrow \text{Dgm}(W)$ and $M_2 : \text{Dgm}(W) \nrightarrow \text{Dgm}(V)$. Note that $V_{\eta_1} \subseteq \text{coim}(M_1)$ and $W_{\eta_1} \subseteq \text{coim}(M_2)$.

Let $M : \text{Dgm}(V) \nrightarrow \text{Dgm}(W)$ be the matching constructed from $M_1$ and $M_2$ as in the proof of the Matching Lemma 4.3.1. This matching has the following properties:

1. $\text{coim}(M_1) \subseteq \text{coim}(M)$,
2. $\text{coim}(M_2) \subseteq \text{im}(M)$,
3. if $M(b,d) = (b',d')$ then either $M_1(b,d) = (b',d')$ or $M_2(b,d) = (b',d')$.

In particular, we have $V_{\eta_1} \subseteq \text{coim}(M)$ and $W_{\eta_1} \subseteq \text{coim}(M)$. Also, if $M(b,d) = (b',d')$ then either $R_1(I([b,d])) = I([b',d'])$ or $R_2(I([b,d])) = I([b',d'])$. By Lemma 4.3.3, $\|(b,d) - (b',d')\|_1 \leq \max\{C_1(R_1), C_2(R_2)\} = \eta$. Hence by Lemma 4.3.2, $c_1(M) \leq \eta$, completing the proof.

**Remark 4.3.2.** It is not hard to see that $d^1_b(\text{Dgm}(I([b,d])), \text{Dgm}(O)) = d - b$ while $d^1_b(I([b,d]), O) \geq d - b + 1$ so that even for interval zigzag modules, the bound provided by Theorem 4.3.1 is not tight.

**Corollary 4.3.1.** Under the conditions of the Theorem 4.3.1,

$$d^\infty_b(\text{Pers}(V), \text{Pers}(W)) \leq d^1_R(V, W).$$

**Proof.** This follows from the fact that $d^\infty_b \leq d^1_b$. \qed
4.4 Weighted Zigzag Modules

4.4.1 Decorated Interval Partitions

Fix a bounded interval \( \langle a, b \rangle \subset \mathbb{R} \), where \( a, b \) are decorated endpoints. By an 
*decorated interval partition* \( \mathcal{I} \) of \( \langle a, b \rangle \) we mean a finite, strictly increasing sequence 
\[
a = s_0 < s_1 < \ldots < s_{n-1} < s_n = b
\]
of decorated endpoints starting at \( a \) and ending at \( b \). Such a sequence gives rise to a 
collection 
\[
\mathcal{I} = \{\langle s_{i-1}, s_i \rangle\}_{i=1}^n
\]
of disjoint intervals whose union is \( \langle a, b \rangle \). Conversely, if \( \mathcal{I} = \{\langle b_i, d_i \rangle\}_{i=1}^n \) is any 
collection of mutually disjoint intervals whose union is \( \langle a, b \rangle \), ordered in the obvious 
way, then we must have \( b_1 = a, d_n = b \), and \( b_{i+1} = d_i \) for all \( i = 1, \ldots, n - 1 \). Thus 
\( \mathcal{I} \) determines the sequence \( a = b_1 < b_2 < \ldots < b_{n-1} < b_n < d_n = b \) of decorated 
endpoints. In this way, we establish a one-to-one correspondence between finite, 
strictly increasing sequences of decorated endpoints from \( a \) to \( b \) and collections of 
mutually disjoint intervals whose union is \( \langle a, b \rangle \). Passing from an decorated interval 
partition \( \mathcal{I} \) to the corresponding collection \( \mathcal{I} \) of intervals will always be indicated by 
the bar, and conversely.

A *refinement* of an decorated interval partition \( \mathcal{I} \) of \( \langle a, b \rangle \) is an decorated interval 
partition \( \mathcal{J} \) of \( \langle a, b \rangle \) with \( \mathcal{I} \subseteq \mathcal{J} \). It is not hard to see that if \( \mathcal{I} \) and \( \mathcal{J} \) are the 
corresponding collections of intervals partitioning \( \langle a, b \rangle \) then for each \( J \in \mathcal{J} \) there 
exists \( I \in \mathcal{I} \) with \( J \subseteq I \). That is, \( \mathcal{I} \) is a refinement of \( \mathcal{I} \) in the set-theoretic sense of 
a refinement of a partition.
Example 4.4.1. Let $a = 0^-$ and $b = 2^+$ so that $(a, b) = [0, 2] \subset \mathbb{R}$. One decorated interval partition of $[0, 2]$ is given by

$$\mathcal{I} = \{0^-, .5^-, 1^+, 2^+\}.$$

The corresponding collection $\mathcal{I}$ of subintervals partitioning $[0, 2]$ is

$$\mathcal{I} = \{[0, .5), [.5, 1], (1, 2]\}.$$

Now let $\mathcal{J} \supset \mathcal{I}$ be the decorated interval partition of $[0, 2]$ given by

$$\mathcal{J} = \{0^-, .5^-, .5^+, 1^+, 1.5^+, 2^+\}.$$

The corresponding collection $\mathcal{J}$ of subintervals partitioning $[0, 2]$ is then

$$\mathcal{J} = \{[0, .5), [.5, .5], (.5, 1], (1, 1.5], (1.5, 2]\}.$$

Evidently, $\mathcal{J}$ refines $\mathcal{I}$ in the ordinary sense of one partition refining another.

### 4.4.2 Weighted Zigzag Modules

Fix an interval $(a, b)$ and let $\mathcal{I}$ be a decorated interval partition of $(a, b)$ into $n$ parts. Let

$$\mathcal{I}-\text{Mod} := \{(V, \mathcal{I}) \mid V \in n-\text{Mod}\}.$$

Elements of $\mathcal{I}$-Mod are called \textit{weighted zigzag module} weighted over $(a, b)$ by $\mathcal{I}$. The pair $(V, \mathcal{I}) \in \mathcal{I}$-Mod will be denoted by $V_\mathcal{I}$ or $(V_\mathcal{i}, p_\mathcal{i}, \mathcal{I})$. There are then obvious maps

$$(\mathcal{I}) : n-\text{Mod} \to \mathcal{I}$-Mod$$

given by $V \mapsto V_\mathcal{I} = (V, \mathcal{I})$ and

$$\Pi : \mathcal{I}$-Mod \to n-\text{Mod}$$
given by $V_\mathcal{I} = (V, \mathcal{I}) \mapsto V$. Evidently, the maps $(-)_\mathcal{I}$ and $\Pi$ are inverses of each other. The zigzag module $V$ is called the unweighted zigzag module underlying $V_\mathcal{I}$.

We also define

$$\langle a, b \rangle\text{-Mod} := \bigcup_{\mathcal{I}} \mathcal{I}\text{-Mod},$$

with this union being taken over all decorated interval partitions $\mathcal{I}$ of $\langle a, b \rangle$.

Given the weighted zigzag module $V_\mathcal{I} = (V, \mathcal{I})$ with $V = (V_i, p_i)$, we associate to each $V_i$ the interval $i$-th interval $I_i = \langle s_{i-1}, s_i \rangle$. We let $\omega_i := \ell(I_i) = \pi_1(s_i) - \pi_1(s_{i-1})$ denote the length of the $i$-th interval in $\mathcal{I}$, and we refer to the $\omega_i$ as weights. Hence, we interpret the weighted zigzag module $V_\mathcal{I}$ as representing the zigzag module $V$ being parameterized over $\langle a, b \rangle$, with this parametrization taking constant vector space values over each interval $I_i$.

### 4.5 The Weighted Reflection Distance

#### 4.5.1 The Weighted Cost of a Composition of Reflections

Let $V_\mathcal{I} = (V, \mathcal{I}) \in \mathcal{I}\text{-Mod}$ for some interval partition $\mathcal{I}$ of $\langle a, b \rangle$ into $n$ parts. Suppose that $\mathcal{R} = R_{k_t}^{s_t} \circ R_{k_{t-1}}^{s_{t-1}} \circ \cdots \circ R_{k_1}^{s_1}$ is a composition of functors with $V \in \text{comp}(\mathcal{R})$, as defined in section 4.1.

For each $p \in [1, \infty]$ define the $\mathcal{I}$-weighted $p$-cost of $\mathcal{R}$ by

$$C^p_\mathcal{I}(\mathcal{R}) := \begin{cases} \left( \sum_{i=1}^\ell \omega^p_{k_i} \right)^{1/p}, & \text{for } p \in [1, \infty) \\ \max_i \omega_i, & \text{for } p = \infty \end{cases}$$

and define $C^p_\mathcal{I}(\text{id}_V) := 0$, where $\text{id} = \text{id}_V$ is the identity morphisms for $V$.

**Fact 4.5.1.** For any compositions of reflections $\mathcal{R}_1, \mathcal{R}_2$ and for any $p \in [1, \infty]$ we have $C^p_\mathcal{I}(\mathcal{R}_2 \circ \mathcal{R}_1) \leq C^p_\mathcal{I}(\mathcal{R}_1) + C^p_\mathcal{I}(\mathcal{R}_2)$. 

77
Proof. Just as in the proof of Fact 4.2.1, this observation follows from the inequality 
\((a + b)^{1/p} \leq a^{1/p} + b^{1/p}\) for all \(a, b \geq 0\). \qed

4.5.2 The Weighted Reflection Distance between Zigzag Modules Weighted by the Same Interval Partition

Consider two weighted zigzag modules \(V_I = (V_i, p_i, I), W_I = (W_i, q_i, I) \in \mathcal{I}\text{-Mod}\).

Let \(V = (V_i, p_i)\) and \(W = (W_i, q_i)\) in \(n\text{-Mod}\) be the respective unweighted zigzag modules underlying \(V_I\) and \(W_I\).

For each \(p \in [1, \infty]\), we define a function \(d^p_{R,I}: \mathcal{I}\text{-Mod} \times \mathcal{I}\text{-Mod} \to \mathbb{R}\) by setting
\[
d^p_{R,I}(V_I, W_I) := \min_{(R_1, R_2)} \left\{ \max\{C^p_{I}(R_1), C^p_{I}(R_2)\} \mid R_1(V) \preceq W\text{ and } R_2(W) \preceq V \right\},
\]
where the minimum is taken over pairs \((R_1, R_2)\) with \(R_1\) and \(R_2\) being compatible compositions of reflections for \(V\) and \(W\), respectively. Note that this function is almost identical to \(d^p_R\) defined in section 4.2 except that the costs of the compositions of reflections are computed using the weighted \(p\)-cost.

Theorem 4.5.1. For any bounded interval \((a, b) \subset \mathbb{R}\), interval partition \(I\) of \((a, b)\), and \(p \in [1, \infty]\), \(d^p_{R,I}\) is a pseudometric on \(\mathcal{I}\text{-Mod}\). Moreover, \(d^p_{R,I}(V_I, W_I) = 0\) if and only if \(V_I \sim W_I\) so that \(d^p_{R,I}\) induces a metric on \(\mathcal{I}\text{-Mod}/\sim\).

Proof. The proof is almost identical to that of the unweighted case. As was the case there, the facts that \(d^p_{R,I}\) is a non-negative and symmetric function follows immediately. If \(V_I = W_I\) then \(V = W\) so that \(V \preceq W\) and \(W \preceq V\). Thus \(d^p_{R,I}(V_I, V_I) = 0\) for all \(V_I \in \mathcal{I}\text{-Mod}\).

Now let \(V^1_I, V^2_I, V^3_I \in \mathcal{I}\text{-Mod}\) have underlying zigzag modules \(V^1, V^2, V^3 \in n\text{-Mod}\), respectively. Let \(p \in [1, \infty]\) and suppose that
\[
d^p_R(V^1_I, V^2_I) = \epsilon_1 \quad \text{and} \quad d^p_R(V^2_I, V^3_I) = \epsilon_2.
\]
for some $\epsilon_1, \epsilon_2 \in \mathbb{R}^+$. Then there exists compositions of functors $\mathcal{R}_1$ and $\mathcal{R}_2$, with $V^1 \in \text{comp}(\mathcal{R}_1)$, $V^2 \in \text{comp}(\mathcal{R}_2)$ such that

$$\mathcal{R}_1(V^1) \preceq V^2, \quad \mathcal{R}_2(V^2) \preceq V^1 \quad \text{and} \quad C^w_\mathcal{P}(\mathcal{R}_1) \leq \epsilon_1, \quad C^w_\mathcal{P}(\mathcal{R}_2) \leq \epsilon_1.$$ 

Similarly, there exists compositions of functors $\mathcal{R}'_1$ and $\mathcal{R}'_2$, with $V^2 \in \text{comp}(\mathcal{R}'_1)$, $V^3 \in \text{comp}(\mathcal{R}'_2)$ such that

$$\mathcal{R}'_1(V^2) \preceq V^3, \quad \mathcal{R}'_2(V^3) \preceq V^2 \quad \text{and} \quad C^w_\mathcal{P}(\mathcal{R}'_1) \leq \epsilon_2, \quad C^w_\mathcal{P}(\mathcal{R}'_2) \leq \epsilon_2.$$ 

By Proposition 4.1.6, $\mathcal{R}'_1 \circ \mathcal{R}_1(V_1) \preceq \mathcal{R}'_1(V_2) \preceq V_3$ and $\mathcal{R}_2 \circ \mathcal{R}'_2(V_3) \preceq \mathcal{R}_2(V_2) \preceq V_1$. By transitivity of the preorder $\preceq$, we have

$$\mathcal{R}'_1 \circ \mathcal{R}_1(V_1) \preceq V_3 \quad \text{and} \quad \mathcal{R}_2 \circ \mathcal{R}'_2(V_3) \preceq V_1. \quad (4.4)$$

By Fact 4.5.1, $C^p_\mathcal{I}(\mathcal{R}'_1 \circ \mathcal{R}_1) \leq C^p_\mathcal{I}(\mathcal{R}'_1) + C^p_\mathcal{I}(\mathcal{R}_1)$ and $C^p_\mathcal{I}(\mathcal{R}_2 \circ \mathcal{R}'_2) \leq C^p_\mathcal{I}(\mathcal{R}_2) + C^p_\mathcal{I}(\mathcal{R}'_2)$ so that

$$C^p_\mathcal{I}(\mathcal{R}'_1 \circ \mathcal{R}_1) \leq \epsilon_1 + \epsilon_2 \quad \text{and} \quad C^p_\mathcal{I}(\mathcal{R}_2 \circ \mathcal{R}'_2) \leq \epsilon_1 + \epsilon_2. \quad (4.5)$$

Equations (4.4) and (4.5) together imply

$$d^p_{\mathcal{R},\mathcal{I}}(V^1_{\mathcal{I}}, V^2_{\mathcal{I}}) \leq \epsilon_1 + \epsilon_2 = d^p_{\mathcal{R},\mathcal{I}}(V^1_{\mathcal{I}}, V^2_{\mathcal{I}}) + d^p_{\mathcal{R},\mathcal{I}}(V^2_{\mathcal{I}}, V^3_{\mathcal{I}}).$$

Now if $V \sim W$ then, by Fact 4.1.2, $V \preceq W$ and $W \preceq V$ so that $d^p_{\mathcal{R},\mathcal{I}}(V, W) = 0$. Conversely, if $V \not\sim W$ then either $V \preceq W$ or $W \preceq V$. In any case, we must apply some non-trivial reflection to either $V$ or $W$, incurring a non-zero cost so that $d^p_{\mathcal{R},\mathcal{I}}(V, W) > 0$. \[ \square \]

### 4.5.3 The General Weighted Reflection Distance

We wish to compare weighted zigzag modules weighted by different interval partitions of an interval $\langle a, b \rangle$. Consider a weighted zigzag module $V_{\mathcal{I}} = (V_i, p_i, \mathcal{I}) \in \mathcal{M}$.
I-Mod, where $\mathcal{I}$ partitions $\langle a, b \rangle$ into $n$ parts, and let $\mathcal{J}$ be a refinement of $\mathcal{I}$ partitioning $\langle a, b \rangle$ into $m \geq n$ parts. We define a new weighted zigzag module $V_{\mathcal{J}(\mathcal{I})} := (V', \mathcal{J}) = V'_\mathcal{J}$, where $V' = (V'_i, p'_i) \in m$-Mod is defined as follows: let $J_i$ be the $i$-th interval of $\mathcal{J}$ and let $I_k \in \mathcal{I}$ be such that $J_i \subseteq I_k$. Then we define $V'_i := V_k$. If the $i$-th and $(i + 1)$-st intervals of $\mathcal{J}$ are contained in a single $I \in \mathcal{I}$ then we define $p'_i = \text{id}_{V_i}$. Otherwise, the $i$-th and $(i + 1)$-st intervals of $\mathcal{J}$ are contained in consecutive intervals of $\mathcal{I}$, say the $j$-th and $(j + 1)$-st intervals of $\mathcal{I}$, in which case we define $p_i := p_j$. The definition is best illustrated with an example:

**Example 4.5.1.** Consider the interval $[0, 4]$ and decorated interval partition

$$\mathcal{I} = \{0^-, 2^+, 4^+\}$$

with corresponding set of subintervals

$$\mathcal{I} = \{[0, 2], (2, 4]\},$$

as in Example 4.4.1. Consider also the zigzag module $V = (V_i, p_i)$ of length 2 given by

$$V = F \leftarrow^{\pi_1} F^2,$$

where $\pi_1 : F^2 \to F$ is the projection onto the first coordinate, given explicitly by $(x, y) \mapsto x$. Then the tuple $V_\mathcal{I} = (V_i, p_i, \mathcal{I})$ is a weighted zigzag module, weighted over $[0, 4]$ by $\mathcal{I}$. We can represent the situation visually with the following picture:

$$V_\mathcal{I} : \begin{array}{c}
\xrightarrow{\pi_1} \end{array} [0 \quad x \quad 2 \quad 4 \quad F^2]$$
Consider the refinement
\[ \mathcal{J} = \{0^-, 1^-, 2^+, 3^+, 4^+\} \]
of \( \mathcal{I} \), which has corresponding set of subintervals
\[ \overline{\mathcal{J}} = \{[0, 1), [1, 2], (2, 3], (3, 4]\}. \]
The weighted zigzag module \( \mathbb{V}_\mathcal{J}(\mathcal{I}) = (\mathbb{V}'_{\mathcal{I}}, p'_I, \mathcal{J}) \) is that whose underlying unweighted zigzag module \( \mathbb{V}' = (\mathbb{V}'_{\mathcal{I}}, p'_I) \) is given by
\[ \mathbb{V}' = \mathbb{F} \xleftarrow{\text{id}_\mathbb{F}} \mathbb{F} \xleftarrow{\pi_1} \mathbb{F}^2 \xleftarrow{\text{id}_\mathbb{F}^2} \mathbb{F}^2. \]
Then we have the following pictoral representation of \( \mathbb{V}_\mathcal{J}(\mathcal{I}) \):

Note that this picture for \( \mathbb{V}_\mathcal{J}(\mathcal{I}) \) is obtained from the picture for \( \mathbb{V}_\mathcal{I} \) by refining the subintervals corresponding to \( \mathcal{I} \) and assigning to each new subinterval the vector space assigned to its parent subinterval in \( \mathcal{I} \) by \( \mathbb{V}_\mathcal{I} \). When two adjacent subintervals of \( \mathcal{J} \) lie in adjacent subintervals of \( \mathcal{I} \), the linear map between the corresponding vector spaces on these subintervals is that inherited from \( \mathbb{V}_\mathcal{I} \). Otherwise, two adjacent subintervals of \( \mathcal{J} \) lie in the same subinterval of \( \mathcal{I} \), so that these intervals inherit the same vector space from \( \mathbb{V}_\mathcal{I} \). In this case, the linear map between the corresponding vector spaces on these subintervals is taken to be the appropriate identity map with the direction being assigned arbitrarily.
Lemma 4.5.1. Let $I$ be an interval partition of $\langle a, b \rangle$. If $J$ is an interval partition of $\langle a, b \rangle$ which refines $I$ then

$$d_{p,\mathcal{J}}^p (V_{\mathcal{J}(I)}, W_{\mathcal{J}(I)}) \leq 2 \cdot d_{p,\mathcal{I}}^p (V_{\mathcal{I}}, W_{\mathcal{I}})$$

for any $V_{\mathcal{I}}, W_{\mathcal{I}} \in I$-Mod and for all $p \in [1, \infty]$.

Proof. Let $V_{\mathcal{I}} \in I$-Mod be any zigzag module weighted by $I$, let $V$ denote the unweighted zigzag module underlying $V_{\mathcal{I}}$, and let $V'$ denote the unweighted zigzag module underlying $V_{\mathcal{J}(I)}$. We will show that for any composition of functors $\mathcal{R}$ with $V \in \text{comp}(\mathcal{R})$, there exists a composition of functors $\mathcal{X}$ with $V \in \text{comp}(\mathcal{X})$ for which $\mathcal{X}(V')$ is the zigzag module underlying $\mathcal{R}(V)_{\mathcal{J}(I)}$ and which satisfies

$$C_{\mathcal{J}}^p (\mathcal{X}) \leq 2 \cdot C_{\mathcal{I}}^p (\mathcal{R}).$$

That is, we will find $\mathcal{X}$ which fits the following diagram:

$$
\begin{array}{ccc}
V & \xrightarrow{\mathcal{R}} & \mathcal{R}(V) \\
\downarrow (-)_{\mathcal{I}} & & \downarrow (-)_{\mathcal{I}} \\
V_{\mathcal{I}} & \xrightarrow{\mathcal{R}} & \mathcal{R}(V_{\mathcal{I}}) = \mathcal{R}(V)_{\mathcal{I}} \\
\downarrow \mathcal{J}(I) & & \downarrow \mathcal{J}(I) \\
V_{\mathcal{J}(I)} = V'_{\mathcal{J}} & \xrightarrow{\mathcal{X}} & \mathcal{R}(V)_{\mathcal{J}(I)} = \mathcal{X}(V')_{\mathcal{J}} \\
\downarrow (-)_{\mathcal{J}} & & \downarrow (-)_{\mathcal{J}} \\
V' & \xrightarrow{\mathcal{X}} & \mathcal{X}(V')
\end{array}
$$

We argue that we can always find $\mathcal{X}$ as follows: suppose we wish to apply $\mathcal{R}_k^+$ to $V_{\mathcal{I}}$. We do so by applying $\mathcal{R}_k^+$ to $V$ and then setting $\mathcal{R}_k^+(V_{\mathcal{I}}) = \mathcal{R}_k^+(V)_{\mathcal{I}}$. To apply $\mathcal{R}_k^+$ to $V$, we replace the subdiagram

$$V_{k-1} \xrightarrow{p_{k-1}} V_k \xleftarrow{p_k} V_{k+1} \quad (4.6)$$
of $V$ by a new diagram of the form

$$V_{k-1} \xleftarrow{\alpha_{k-1}} V_k^+ \xrightarrow{p_k} V_{k+1}. \quad (4.7)$$

Now corresponding to diagram 4.6 is a subdiagram of $V'$ which has the form

$$V_{k-1} \xrightarrow{p_{k-1}} V_k \xleftarrow{id} V_k \xrightarrow{id} \cdots \xleftarrow{id} V_k \xrightarrow{id} V_k \xleftarrow{p_k} V_{k+1},$$

where the interior nodes are duplicated as a result of the refinement operation. Note that the direction of the identity maps is irrelevant modulo the relation $\sim$. We can thus apply an extroversion functor at the leftmost $V_k$ appearing in this diagram, obtaining the diagram

$$V_{k-1} \xleftarrow{id} V_{k-1} \xrightarrow{p_{k-1}} V_k \xleftarrow{id} \cdots \xleftarrow{id} V_{k-1} \xrightarrow{p_{k-1}} V_k \xleftarrow{p_k} V_{k+1}.$$

We repeat this process from left to right until we arrive at the diagram

$$V_{k-1} \xleftarrow{id} V_{k-1} \xleftarrow{id} V_{k-1} \xrightarrow{id} \cdots \xrightarrow{id} V_{k-1} \xrightarrow{p_{k-1}} V_k \xrightarrow{id} V_k \xrightarrow{p_k} V_{k+1}.$$  

Note the appearance of 4.6 as the rightmost subdiagram. We now apply the sink reflection functor at the second-to-rightmost node, obtaining

$$V_{k-1} \xleftarrow{id} V_{k-1} \xleftarrow{id} V_{k-1} \xrightarrow{id} \cdots \xrightarrow{id} V_{k-1} \xleftarrow{\alpha_{k-1}} V_k^+ \xrightarrow{\beta_k} V_{k+1}.$$

Next, we apply an extroversion functor at the rightmost occurrence of $V_{k-1}$, obtaining

$$V_{k-1} \xleftarrow{id} V_{k-1} \xleftarrow{id} V_{k-1} \xrightarrow{id} \cdots \xrightarrow{id} V_{k-1} \xleftarrow{\alpha_{k-1}} V_k^+ \xleftarrow{id} V_k^+ \xrightarrow{p_k} V_{k+1}.$$  

Lastly, we repeat the last step from right to left until we arrive at the diagram

$$V_{k-1} \xrightarrow{\alpha_{k-1}} V_k^+ \xleftarrow{id} V_k^+ \xleftarrow{id} \cdots \xleftarrow{id} V_k^+ \xleftarrow{id} V_k^+ \xrightarrow{\beta_k} V_{k+1}. \quad (4.8)$$
But now notice that diagram 4.8 is exactly the subdiagram corresponding to diagram 4.7 after refinement. Let us denote the composition of functors corresponding to the above sequence of operations on $V'$ by $\mathcal{X}^+_{k'}$. If $\omega_k$ denotes the weight of $V_k$ in $V$ and the corresponding weights in $V'$ after refinement are denoted by $\omega_{ki}$ for some $i \in \{1, \ldots, m\}$ so that $\sum_i \omega_{ki} = \omega_k$, then we have

$$C^p_J(\mathcal{X}^+_{k'}) = \left(\sum_\ell \omega^p_{k\ell}\right)^{1/p} \leq \sum_\ell \omega_{k\ell} \leq 2\omega_k = 2 \cdot C^p_I(\mathcal{R}^+_k).$$

A completely analogous procedure can be carried out if we wish to apply a source reflection functor $\mathcal{R}^-_k$ to $V_I$. In this case, one mimics the above procedure, but instead moves from right to left using introversion functors, applies a source reflection functor, and then moves from left to right applying introversion functors again. We will denote this composition of functors by $\mathcal{X}^-_k$, and we have the analogous inequality for the source reflection functor:

$$C^p_J(\mathcal{X}^-_k) \leq 2 \cdot C^p_I(\mathcal{R}^-_k).$$

In the case that we want to apply either an extroversion functor $\mathcal{R}^*_k$ or an introversion functor $\mathcal{R}^*_k$ to $V$, it is not hard to see that the corresponding operations can be carried out on $V'$ at the same cost. In the case of an extroversion functor we may apply extroversions from left to right in the corresponding subdiagram of $V'$, and for introversion functors, we apply introversions from right to left on the appropriate subdiagram of $V'$. In this way, we obtain compositions of functors, either $\mathcal{X}^*_k$ or $\mathcal{X}^*_k$, with

$$C^p_J(\mathcal{X}^*_k) = C^p_I(\mathcal{R}^*_k) \quad \text{or} \quad C^p_J(\mathcal{X}^*_k) = C^p_I(\mathcal{R}^*_k).$$

84
Putting this all together, suppose that we wish to apply the composition of functors
\[ R = R_{k_\ell}^{s_\ell} \circ R_{k_{\ell-1}}^{s_{\ell-1}} \circ \cdots \circ R_{k_1}^{s_1} \]
to \( V \). We have
\[ C_T^p(R) = \left( \sum_{i=1}^\ell \omega_{k_i}^p \right)^{1/p}. \]
By the above procedures, we can find a corresponding sequence
\[ \mathcal{X} = \mathcal{X}_{k_\ell}^{s_\ell} \circ \mathcal{X}_{k_{\ell-1}}^{s_{\ell-1}} \circ \cdots \circ \mathcal{X}_{k_1}^{s_1} \]
of compositions of functors so that \( \mathcal{X}(V') \) is the unweighted zigzag module underlying the \( J \)-refinement of \( R(V_I) \). The \( J \)-weighted \( p \)-cost of \( \mathcal{X} \) is
\[ C_J^p(\mathcal{X}) = \left( \sum_i \sum_j \omega_{k_i,m_j}^p \right)^{1/p}, \]
where \( \sum_j \omega_{k_i,m_j} \leq 2\omega_{k_i} \). Then we have
\[ C_J^p(\mathcal{X}) = \left( \sum_i \sum_j \omega_{k_i,m_j}^p \right)^{1/p} \leq \left( \sum_i \left( \sum_j \omega_{k_i,m_j} \right)^p \right)^{1/p} \]
\[ \leq \left( \sum_i (2\omega_{k_i})^p \right)^{1/p} \leq 2 \left( \sum_i \omega_{k_i}^p \right)^{1/p} = 2 \cdot C_T^p(R), \]
which completes the proof.

**Definition 4.5.1.** For each \( p \in [1, \infty] \), we define a function \( d_{R}^{p,\omega} : (a,b)\text{-Mod} \times (a,b)\text{-Mod} \to \mathbb{R} \) by setting
\[ d_{R}^{p,\omega}(V_I, W_J) := \inf_{K} d_{R,K}^{p,\omega}(V_{K(I)}, W_{K(J)}), \]
where the infimum is taken over all interval partitions \( K \) of \( (a,b) \) which refine both \( I \) and \( J \).
Theorem 4.5.2. Fix a bounded interval \( \langle a, b \rangle \subset \mathbb{R} \). For any \( p \in [1, \infty] \) and for any \( V_I, W_J, U_K \in \langle a, b \rangle \text{-Mod} \), the function \( d_{R}^{p, \omega} \) satisfies

1. \( d_{R}^{p, \omega}(V_I, W_J) \geq 0 \),

2. \( d_{R}^{p, \omega}(V_I, W_J) = d_{R}^{p, \omega}(W_J, V_I) \),

3. \( d_{R}^{p, \omega}(V_I, W_J) \leq 2 \cdot (d_{R}^{p, \omega}(V_I, U_K) + d_{R}^{p, \omega}(U_K, W_J)) \).

Proof. Non-negativity and symmetry of \( d_{R}^{p, \omega} \) follow immediately from the non-negativity and symmetry of each \( d_{R}^{p, \omega} \). We are left only to show that the relaxed triangle inequality holds. Let \( V_I, W_J, U_K \in \langle a, b \rangle \text{-Mod} \) and suppose that \( \eta, \eta' \in \mathbb{R} \) are such that

\[
d_{R}^{p, \omega}(V_I, U_K) < \eta \quad \text{and} \quad d_{R}^{p, \omega}(U_K, W_J) < \eta'.
\]

Then there exist decorated interval partitions \( \mathcal{A} \) and \( \mathcal{B} \) of \( \langle a, b \rangle \) such that \( \mathcal{A} \) refines both \( I \) and \( K \), \( \mathcal{B} \) refines both \( K \) and \( J \), and

\[
d_{R, \mathcal{A}}^{p, \omega}(V_{A(I)}, U_{A(K)}) < \eta \quad \text{and} \quad d_{R, \mathcal{B}}^{p, \omega}(U_{B(K)}, W_{B(J)}) < \eta'.
\]

Let \( \mathcal{C} = \mathcal{A} \cup \mathcal{B} \) be the common refinement of \( \mathcal{A} \) and \( \mathcal{B} \) so that \( \mathcal{C} \) refines \( I \), \( J \), and \( K \). Then using Lemma 4.5.1, we have

\[
d_{R}^{p, \omega}(V_I, W_J) \leq d_{R, \mathcal{C}}^{p, \omega}(V_{C(I)}, W_{C(J)})
\]

\[
\leq d_{R, \mathcal{C}}^{p, \omega}(V_{C(I)}, U_{C(K)}) + d_{R, \mathcal{C}}^{p, \omega}(U_{C(K)}, W_{C(J)})
\]

\[
\leq 2 \cdot d_{R, \mathcal{A}}^{p, \omega}(V_{A(I)}, U_{A(K)}) + 2 \cdot d_{R, \mathcal{B}}^{p, \omega}(U_{B(K)}, W_{B(J)}) < 2(\eta + \eta').
\]

Since \( \eta > d_{R, \mathcal{A}}^{p, \omega}(V_A, U_A) \) and \( \eta' > d_{R, \mathcal{A}}^{p, \omega}(V_A, U_A) \) were arbitrary, we have

\[
d_{R}^{p, \omega}(V_I, W_J) \leq 2 \cdot (d_{R}^{p, \omega}(V_I, U_{L}) + d_{R}^{p, \omega}(U_{L}, W_J)).
\]

\( \square \)
Remark 4.5.1. The function \( d_{R}^{p,\omega} \) satisfies a weaker form of the triangle inequality, sometimes called the \( q \)-relaxed triangle inequality; in this case \( q = 2 \). Functions which satisfy the axioms of a metric but with the triangle inequality replaced with a \( q \)-relaxed triangle inequality are sometimes called quasimetrics. In our case, different weighted zigzag modules can have zero distance with respect to \( d_{R}^{p,\omega} \), so an appropriate name for such a function might be “quasi-pseudometric”.

4.6 Diagrams of Weighted Zigzag Modules Stability

In this section we will consider the interval \( \langle a, b \rangle = [0, 1] \). Consider a weighted zigzag module \( V_{I} \) for some decorated interval partition \( I = \{s_0 < s_1 < \ldots < s_{n+1}\} \) of \([0, 1]\) and let \( V \) be the underlying unweighted zigzag module. We define a map

\[
\phi_{I} : \text{Dgm}(V) \to \mathbb{R}^2, \quad (b, d) \mapsto (s_{b-1}, s_d),
\]

and we define the \textit{weighted persistence diagram} of \( V_{I} \) to be the multiset representation

\[
\text{Dgm}(V_{I}) := \phi_{I}(\text{Dgm}(V)) = \{ \phi_{I}(b, d) \mid (b, d) \in \text{Dgm}(V) \}.
\]

Proposition 4.6.1. Suppose \( I \) and \( J \) are interval partitions of \([0, 1]\) with \( J \) refining \( I \). Then \( \text{Dgm}(V_{I}) = \text{Dgm}(V_{J(I)}) \).

Proof. This follows from the definition of refinement. \( \Box \)

Lemma 4.6.1. Let \( I = \{0 = s_0 < s_1 < \cdots < s_{n-1} < s_n = 1\} \) be a decorated interval partition of \([0, 1]\) and let \( V_{I} \in [0, 1]\text{-Mod} \) with underlying zigzag module \( V \). If \( R(V) = 0 \) then

\[
C_{I}^{1}(R) \geq \max_{(x, y) \in \text{Dgm}(V_{I})} |y - x|.
\]
Moreover, if \( \mathcal{R}(I([b, d])) = I([b', d']) \) then
\[
C^1_K(\mathcal{R}) \geq \|s_{b-1} - s_d\|_1 = |s_d - s_{d'}| + |s_{b-1} - s_{b'-1}|.
\]

Proof. We have
\[
O = \mathcal{R}(V) = \mathcal{R}\left( \bigoplus_{(b,d) \in \text{Dgm}(V)} I_\tau([b, d]) \right) \cong \bigoplus_{(b,d) \in \text{Dgm}(V)} \mathcal{R}(I_\tau([b, d]))
\]
so that \( \mathcal{R}(I_\tau([b, d])) = O \) for all \((b, d) \in \text{Dgm}(V)\). Then \( C^1_K(\mathcal{R}) \geq s_d - s_{b-1} \) for all \((b, d) \in \text{Dgm}(V)\). Since \((b, d) \in \text{Dgm}(V)\) if and only if \((s_{b-1}, s_d) \in \text{Dgm}(V_I)\), we see that \( C^1_K(\mathcal{R}) \geq |y - x| \) for all \((x, y) \in \text{Dgm}(V_I)\), proving the first claim.

The second claim corresponds to the fact that points in the weighted persistence diagram move vertically or horizontally in the plane when a transforming functor is applied to \( V \), and the distance that a point moves is exactly equal to the weighted 1-cost of the transformation applied.

Theorem 4.6.1. Let \( V_I, W_J \in [0, 1]\text{-Mod} \). Then
\[
d^1_b(\text{Dgm}(V_I), \text{Dgm}(W_J)) \leq d^1_{\mathcal{R}}(V_I, W_J).
\]

Proof. Suppose that \( d^1_{\mathcal{R}}(V_I, W_J) < \eta \). As in the proof of 4.3.1, we will construct a matching \( N : \text{Dgm}(V_I) \rightarrow \text{Dgm}(W_J) \) with \( c_1(N) < \eta \). Since \( d^1_{\mathcal{R}}(V_I, W_J) < \eta \), there exists an interval partition \( K = \{s_0 < s_1 < \ldots < s_{m-1} < s_m\} \) of \([0, 1]\) which refines both \( I \) and \( J \) and such that
\[
d^1_{\mathcal{R},K}(V_{K(I)}, W_{K(J)}) < \eta.
\]
Let \( V' \) and \( W' \) be the zigzag modules underlying \( V_{K(I)} \) and \( W_{K(J)} \), respectively. Then there exists compositions of functors \( \mathcal{R}_1 \in \text{comp}(V') \) and \( \mathcal{R}_2 \in \text{comp}(W') \) such that
\[
C^1_K(\mathcal{R}_1) < \eta \quad \text{and} \quad C^1_K(\mathcal{R}_2) < \eta
\]
\[
\mathcal{R}_1(\mathcal{V}') \precsim \mathcal{W}' \quad \text{and} \quad \mathcal{R}_2(\mathcal{W}') \precsim \mathcal{V}'.
\]

Let
\[
\mathcal{X}^1_{\eta} := \{(b, d) \in \text{Dgm}(\mathcal{V}') \mid |s_d - s_{b-1}| > \eta\}, \quad \mathcal{Y}^1_{\eta} := \{(b, d) \in \text{Dgm}(\mathcal{W}') \mid |s_d - s_{b-1}| > \eta\},
\]

let
\[
\mathcal{V}^1_{\eta} := \phi_{\mathcal{K}}(\mathcal{X}^1_{\eta}) \quad \text{and} \quad \mathcal{W}^1_{\eta} := \phi_{\mathcal{K}}(\mathcal{Y}^1_{\eta}),
\]

and note that
\[
\mathcal{V}^1_{\eta} = \{(x, y) \in \text{Dgm}(\mathcal{V}_x) \mid |y - x| > \eta\}, \quad \mathcal{W}^1_{\eta} = \{(x, y) \in \text{Dgm}(\mathcal{W}_x) \mid |y - x| > \eta\}.
\]

Also let
\[
A = \{(b, d) \in \text{Dgm}(\mathcal{V}') \mid \mathcal{R}_1(\mathcal{I}(b, d)) \neq \emptyset\}, \quad B = \{(b, d) \in \text{Dgm}(\mathcal{W}') \mid \mathcal{R}_2(\mathcal{I}(b, d)) \neq \emptyset\}.
\]

Note that by the first part of Lemma 4.6.1, \(\mathcal{X}^1_{\eta} \subseteq A\) and \(\mathcal{Y}^1_{\eta} \subseteq B\).

As in the proof of Theorem 4.3.1, we can define matchings
\[
M_1 : \text{Dgm}(\mathcal{V}') \to \text{Dgm}(\mathcal{W}') \quad \text{and} \quad M_2 : \text{Dgm}(\mathcal{W}') \to \text{Dgm}(\mathcal{V}')
\]
given by \(M_i(b, d) = (b', d')\) if and only if \(\mathcal{R}_i(\mathcal{I}(b, d)) = \mathcal{I}(b', d')\) for \(i \in \{1, 2\}\).

Note that \(\mathcal{X}^1_{\eta} \subseteq \text{coim}(M_1)\) and \(\mathcal{Y}^1_{\eta} \subseteq \text{coim}(M_2)\). The matchings \(M_1\) and \(M_2\) induce matchings
\[
N_1 : \text{Dgm}(\mathcal{V}_K) \to \text{Dgm}(\mathcal{W}_K) \quad \text{and} \quad N_2 : \text{Dgm}(\mathcal{W}_K) \to \text{Dgm}(\mathcal{V}_K)
\]
given by \(N_i(s_{b-1}, s_d) = (s_{b'-1}, s_{d'})\) if and only if \(M_i(b, d) = (b', d')\) for \(i = 1, 2\). By the Matching Lemma 4.3 there exists a matching \(N : \text{Dgm}(\mathcal{V}_K) \to \text{Dgm}(\mathcal{W}_K)\) with the following properties:
1. \( \text{coim}(N_1) \subseteq \text{coim}(N) \),

2. \( \text{coim}(N_2) \subseteq \text{im}(N) \),

3. if \( N(s_{b-1}, s_d) = (s_{b'-1}, s_{d'}) \) then either \( N_1(s_{b-1}, s_d) = (s_{b'-1}, s_{d'}) \) or \( N_2(s_{b-1}, s_d) = (s_{b'-1}, s_{d'}) \).

In particular, we have \( V^n \subseteq \text{coim}(N) \) and \( W^n \subseteq \text{im}(N) \). Also, if \( N(s_{b-1}, s_d) = (s_{b'-1}, s_{d'}) \) then either \( R_1(I([b, d])) = I([b', d']) \) or \( R_2(I([b, d])) = I([b', d']) \). By the second part of Lemma 4.3.3,

\[
\| (s_{b-1}, s_d) - (s_{b'-1}, s_{d'}) \|_1 \leq \max \{ C^1_K(R_1), C^1_K(R_2) \} < \eta.
\]

Hence by Lemma 4.3.2, \( c_1(N) < \eta \), completing the proof. \( \square \)

**Remark 4.6.1.** As pointed out above, we have only shown that the function \( d^{1,\omega}_R \) satisfies a 2-relaxed version of the triangle inequality. It is, however, possible to induce from \( d^{1,\omega}_R \) a legitimate pseudometric \( \rho^{1,\omega}_R \) using a construction sometimes called the maximal sub-dominant pseudometric (see [20] for the definition and [5] for the definition in the ultrametric case). The maximal subdominant pseudometric \( \rho^{1,\omega}_R \) of \( d^{1,\omega}_R \) is the unique pseudometric on \([0, 1]\)-Mod with the property that \( \rho^{1,\omega}_R \leq d^{1,\omega}_R \) and \( d \leq \rho^{1,\omega}_R \) for any other pseudometric \( d \) on \([0, 1]\)-Mod. In particular, \( \rho^{1,\omega}_R \) satisfies

\[
d^1_b(\text{Dgm}(V_I), \text{Dgm}(W_J)) \leq \rho^{1,\omega}_R(V_I, W_J)
\]

for all \( V_J, W_J \in [0, 1]\)-Mod.
Bibliography


