

An Exposition on the Algebra and Computation of Persistent Homology

(second version)

An expository paper submitted to
OREGON STATE UNIVERSITY
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE (MS) DEGREE IN MATHEMATICS

Author
Jason Ranoa

Advisor
Dr. Christine Escher

Initially submitted
February 2024
with corrections submitted August 2024.

An Exposition on the Algebra and Computation of Persistent Homology.

ABSTRACT. We discuss the algebra behind the matrix reduction algorithm for persistent homology, as presented in the paper “*Computing Persistent Homology*” by Afra Zomorodian and Gunnar Carlsson, in the lens of the more modern characterization of persistence modules as functors from a poset category to a category of vector spaces over a field adopted by authors such as Peter Bubenik, Frederik Chazal, and Ulrich Bauer.

AUTHOR’S NOTE FOR THE SECOND VERSION.

This copy of the paper is a corrected version of the expository paper submitted in February 2024 before my final oral examination (defense). Because of time constraints and other personal and health issues, a non-negligible number of errors were not fixed or were not caught in time. Listed below are the major differences and changes:

- (i) **Section 2.3** and **Section 2.5** contain significant changes and corrections in the exposition and commentary. The definitions and results presented remain mostly the same, e.g. with minor changes in notation.
- (ii) The entirety of **Section 2.4** is re-written. When I was first writing this section, I only had a tenuous grasp of graded module theory, and this was reflected in the definitions and explanations I provided. I was also missing a number of important definitions (e.g. graded submodule) and results (e.g. quotients of graded modules by graded submodules are graded).
- (iii) Introductions to **Chapter 3** and to **Chapter 4** are added.
- (iv) Errors in the commentary and examples of **Section 3.2** were fixed. Additional commentary and images were included at several points.
- (v) **Section 3.3** had several significant errors, e.g. missing definitions for objects used in the next chapter, incorrect examples. Some crucial results, e.g. Proposition 3.3.18, were also not identified in the previous version.
- (vi) Section 4.3 “*Graded Invariant Factor Decompositions and SNDs*” of the old version was divided into three sections for better exposition: **Section 4.3**, **Section 4.4**, and **Section 4.5**. The content is mostly unchanged, with some additional commentary and minor changes to formatting.
- (vii) **Section 4.6** “*Matrix Calculation of Homology of Graded Chain Complexes*” was missing in the old version because I forgot to uncomment the `import` line in my \LaTeX set-up and somehow missed this in the review. This section contains a brief explanation about how the results for ungraded chain complexes extend nicely for graded chain complexes, and another example calculation of persistent homology.
- (viii) The list of symbols in **Appendix A1** is now grouped by topic. Corrections made in the previous chapters are also reflected here. Several errors and formatting issues were corrected in **Appendix A2**. Some comments made in the previous chapters involving modules were collected and moved to **Appendix A3**.

Please note that a significant part of the edits done in this version have not been reviewed or approved by my advisor (Dr. Christine Escher) as these were made after I had graduated.

— Jason Ranoa

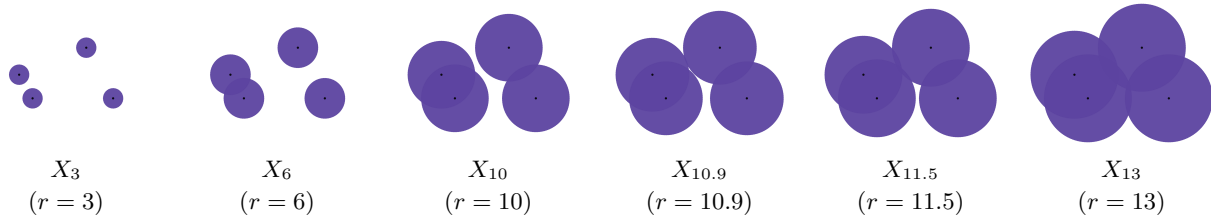
Table of Contents

Introduction	4
1 Simplicial Homology	6
1.1 Simplicial Complexes and Geometric Realizations	7
1.2 Simplicial Homology with Coefficients in a PID	14
1.3 Functorial Constructions in Simplicial Homology	22
2 Introduction to Persistence Theory	26
2.1 Persistence Modules as Functors	28
2.2 The Category of Persistence Modules	33
2.3 Interval Decompositions of Persistence Modules	43
2.4 The Category of Graded Modules over Polynomial Rings	53
2.5 The Equivalence between Persistence Modules and Graded Modules	62
3 Filtrations and Persistent Homology	74
3.1 Filtrations of Simplicial Complexes	75
3.2 The Persistent Homology of Filtrations	79
3.3 Simplicial Persistent Homology	89
4 Calculation by Matrices	99
4.1 The Structure Theorem and Smith Normal Decompositions	101
4.2 Matrix Calculation of Homology of Chain Complexes	113
4.3 The Graded Structure Theorem and SNDs in the Graded Case	121
4.4 Matrix Reduction of Graded Matrices	127
4.5 An Ungraded SND Algorithm in the Graded Case	136
4.6 Matrix Calculation of Homology of Graded Chain Complexes	148
Acknowledgements	152
Appendices	153
A1 List of Symbols	153
A2 Matrices over Euclidean Domains	156
A3 Notes on Ring and Module Theory	161
A4 A Brief Review of Categories and Functors	164
References	168

Introduction

Topological data analysis (TDA) is a relatively new field of study that seeks real-world applications of the theory of algebraic topology. One such application involves the identification of features on finite datasets (as finite metric subspaces of \mathbb{R}^N) by investigating a family of topological spaces constructed using said dataset over some parameter.

For example, given the dataset $\mathbb{X} = \{x_1, x_2, x_3, x_4\} \subseteq \mathbb{R}^2$ with $x_1 = (2, 22)$, $x_2 = (6, 15)$, $x_3 = (22, 30)$, and $x_4 = (30, 15)$, we can construct a family $\{X_r : r \in \mathbb{R}_{\geq 0}\}$ of topological subspaces of \mathbb{R}^2 by taking the union of 2-disks centered at each x_i with radius $r \in [0, \infty)$. This construction produces 6 subspaces distinct up to homeomorphism, as illustrated below:

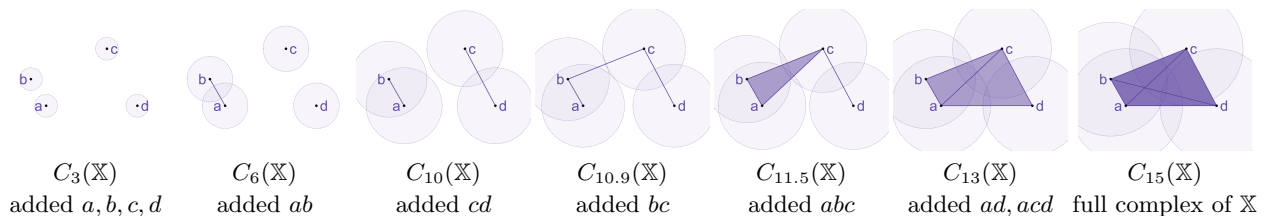


The collection of homology groups $H_n(X_r)$ with $r \in \mathbb{R}_{\geq 0}$ and $n \in \mathbb{N}_0$ and the images of the maps $H_n(X_r) \rightarrow H_n(X_s)$ on homology induced by the inclusions $i^{s,r} : X_r \rightarrow X_s$ is what we call the *persistent homology* of the family $\{X_r\}$. This persistent homology then is used to determine certain characteristics of the dataset \mathbb{X} . In the example above, we can determine that the points of \mathbb{X} are near one another since the space X_r merges into one path component as early as $r = 10.9$.

In practice, the calculation of persistent homology happens at the level of simplicial complexes. Following the example above, an abstract simplicial complex $C_r(\mathbb{X})$ called the *Čech complex* of \mathbb{X} with parameter r is constructed for each $r \in \mathbb{R}_{\geq 0}$ using the following rule:

For each subset S of \mathbb{X} with $n = \text{card}(S) \geq 1$, S is an $(n - 1)$ -simplex of $C_r(\mathbb{X})$ if and only if the collection of the 2-disks centered at each $x_i \in S$ with radius r has non-trivial intersection as subspaces of \mathbb{R}^2 .

There are seven distinct abstract simplicial complexes in the family $\{C_r(\mathbb{X}) : r \in \mathbb{R}_{\geq 0}\}$ of Čech complexes, as illustrated below. For clarity, the elements $\{x_1, x_2, x_3, x_4\}$ of \mathbb{X} are denoted as $\{a, b, c, d\}$ respectively when used as vertices of $C_r(\mathbb{X})$.



The collection $\{H_n(C_r(\mathbb{X})) : r \in \mathbb{R}_{\geq 0}\}$ of simplicial homology groups then determines the persistent homology of \mathbb{X} by the *Nerve Theorem*, which primarily states that $H_n(X_r) \cong H_n(C_r(\mathbb{X}))$ as abelian groups for all $r \in \mathbb{R}_{\geq 0}$ and $n \in \mathbb{N}_0$.

In this expository paper, we consider a more general case and discuss the persistent homology $H_n(K_\bullet; \mathbb{F})$ of \mathbb{N}_0 -indexed filtrations $K_\bullet = \{K_t\}_{t \in \mathbb{N}_0}$ of abstract simplicial complexes (wherein $K_t \subseteq K_s$ whenever $t \leq s$). More specifically, we talk about the algebraic background needed for the *matrix reduction algorithm for persistent homology* discussed in the following paper:

The concepts discussed in this paper were then used as basis for a number of software packages such as Ripser [Bau21], a C++ package written by Ulrich Bauer that calculates the persistent homology of Vietoris-Rips filtrations, taking finite metric subspaces of \mathbb{R}^N as input.

This expository paper talks about the concepts in [ZC05] in the perspective of category theory. We have included an introductory-level discussion of category theory in Appendix A4. For general texts on category theory, we recommend *Category Theory in Context* [Rie16] by Emily Riehl and *Introduction to Homological Algebra* [Rot08] by Joseph J. Rotman.

Much like how the homology groups $H_n(K; R)$ of a simplicial complex K with coefficients in a principal ideal domain (PID) R are encoded as R -modules, we characterize the persistent homology $H_n(K_\bullet; \mathbb{F})$ of a filtration K_\bullet with coefficients in a field \mathbb{F} as a *persistence module* over \mathbb{F} . This paper is organized as follows:

- In **Chapter 1**: We provide necessary background information involving simplicial complexes and simplicial homology with coefficients in a PID R , as well as identify a number of results that will be useful in the characterization of persistent homology.
- In **Chapter 2**: We define persistence modules over a field \mathbb{F} as a *diagram*, i.e. as functors $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$, describe the interval decomposition of a persistence module (which is unique up to persistence isomorphism), and present a category equivalence between the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules and the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ of graded $\mathbb{F}[x]$ -modules.
- In **Chapter 3**: We introduce simplicial filtrations as functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$ and characterize the persistent homology of these filtrations as persistence modules. We also describe how simplicial homology can be extended to the case of persistence modules, which we call *simplicial persistent homology*, and discuss the objects in $\mathbf{GrMod}_{\mathbb{F}[x]}$ brought about by applying the category equivalence between persistence modules and graded modules.
- In **Chapter 4**: We discuss how the invariant factor decomposition of the homology of free chain complexes of R -modules can be calculated using the Smith Normal Decomposition (SND) of matrices over R and how this calculation can be extended to the case of graded $\mathbb{F}[x]$ -modules. We also discuss how this method of calculation, along with *simplicial persistent homology*, can be used to find the interval decomposition of persistence modules corresponding to persistent homology.

Chapter 1. Simplicial Homology

In this chapter, we review a number of definitions and results involving the simplicial homology of simplicial complexes. In the field of persistent homology theory, the term *simplicial complex* is conventionally used to refer to an *abstract simplicial complex*

Note that this chapter is not supposed to be a rigorous or thorough treatment of simplicial homology theory but instead covers constructions that are immediately relevant in our discussion of persistent homology. As such, proofs for most of the propositions and theorems in this chapter are not provided in this paper but can be found in standard algebraic topology textbooks. This chapter is divided into the following sections:

Section 1.1. Simplicial Complexes and Geometric Realizations

We discuss abstract simplicial complexes, which are representations of certain topological spaces, and review some terminology and results.

Note that, after this section, we often drop the modifier *abstract* when referring to an *abstract simplicial complex*, following convention.

Section 1.2. Simplicial Homology with Coefficients in a PID

We provide definitions for the constructions involved in the simplicial homology of (abstract) simplicial complexes and identify homology as an invariant of the homeomorphism type of topological spaces.

Section 1.3. Functorial Constructions in Simplicial Homology

We define the category **A-Simp** of (abstract) simplicial complexes and simplicial maps and discuss why the constructions discussed in the previous section correspond to functors **A-Simp** \rightarrow **Mod**_{*R*} and **A-Simp** \rightarrow **Ch-Mod**_{*R*} where **Mod**_{*R*} refers to the category of *R*-modules and *R*-module homomorphisms and **Ch-Mod**_{*R*} refers to the corresponding chain complex category.

This functorial perspective of simplicial homology will be relevant in Chapter 3 when we discuss persistent homology.

Listed below are the main references used in this chapter (in order of decreasing relevance):

1. *An Introduction to Algebraic Topology* [Rot88] by Joseph J. Rotman.
2. *Algebraic Topology* [Hat02] by Allen Hatcher.
3. *Elements of Algebraic Topology* [Mun93] by James Munkres.

Section 1.1. Simplicial Complexes and Geometric Realizations

In this section, we define a combinatorial representation of topological spaces called a *simplicial complex*. There are generally two types of simplicial complexes used in algebraic topology:

1. Abstract Simplicial Complexes (given in Definition 1.1.1), which are combinatorial representations. We have found that, in the field of persistent homology, the term *simplicial complex* conventionally refers to this type of simplicial complex. We follow this convention in this expository paper.
2. Geometric Simplicial Complexes (given in Definition 1.1.8), which are representations of topological spaces as a collection of subsets of \mathbb{R}^N for sufficiently high N . Note that, in the more general field of algebraic topology, the term *simplicial complex* usually refers to this type of simplicial complex, as seen in [Hat02] and [Mun93].

For clarity, we do not drop the modifier *abstract* and *geometric* when discussing simplicial complexes in this section. However, under certain finiteness conditions, there is a correspondence between these two types of simplicial complexes.

We start with a definition for abstract simplicial complexes, taken from [Rot08, p141].

Definition 1.1.1. An **abstract simplicial complex** K is a collection of nonempty subsets of some set V such that the following properties are satisfied:

- i. For all $v \in V$, $\{v\} \in K$, i.e. all vertices are in K .
- ii. For all sets $\tau \in K$, if σ is non-empty subset of τ , then $\sigma \in K$.

The set V is called the **vertex set** of K , denoted $\text{Vert}(K) := V$, and an element $v \in V$ is called a **vertex** of K . If V is finite, K is called a **finite simplicial complex**.

A **simplex** σ is a set in K . If $\text{card}(\sigma) = n + 1$, then $\sigma \in K$ is called an n -**simplex** of K with **dimension** $n \in \mathbb{N}_0$ denoted by $\dim(\sigma) = n$. A **face** of σ is another simplex τ of K such that τ is a subset of σ , denoted $\tau \subseteq \sigma$. If $\dim(\tau) = \dim(\sigma) - 1$, τ is called a **facet** of σ .

The **dimension** of K , denoted $\dim(K)$, is the maximum dimension of its simplices. In this case, we say K is n -**dimensional** and write $\dim(K) = n$. A simplicial complex L is called a **subcomplex** of K if every simplex $\tau \in L$ is in K .

Remarks. (1) Note that, in the context of calculating simplicial homology (as in Section 1.2) and simplicial persistent homology (later in Section 3.3), we often assume that the simplicial complex K is finite. While we make an effort to identify this finiteness condition when relevant, some papers (e.g. [ZC05]) implicitly assume that simplicial complexes are finite.

- (2) Outside of this section, we often denote n -simplices using string notation as opposed to set notation for brevity, e.g. we write abc to refer to the 2-simplex $\{a, b, c\}$. We justify this convention later in Section 1.2 under Remark 1.2.2 in the context of simplicial homology.

Note that $\text{card}(A)$ refers to the cardinality of some set A . In this paper, we have chosen not to use the conventional notation of $|A|$ for cardinality since the operator $|-|$ is used differently in the context of simplicial complexes. Note that for simplices σ , $\text{card}(\sigma) = \dim(\sigma) + 1$. We state some remarks about the vertex set $V := \text{Vert}(K)$ of a simplicial complex K :

1. The vertex set V is often not explicitly identified in a specification of K since V can be identified by

$$\begin{aligned} V &= \bigcup_{\sigma \in K} \sigma & \text{and} & & V &= \bigcup \{\sigma \in K : \text{card}(\sigma) = 1\} \\ \text{i.e. as the union of all simplices of } K & & & & \text{i.e. as the union of all 0-simplices of } K \end{aligned}$$

2. Observe that K can be seen as a subset of the power set 2^V of V . Therefore, if V is finite, then 2^V is finite and $K \subseteq 2^V$ is also finite as a set.

As a sidenote, the notation 2^V for the power set of V comes from the fact that for each set $P \in 2^V$, the set P can be seen as a collection of choices, one for each $v \in V$, whether either v is included in P or not included in P (i.e. two possible choices for each $v \in V$).

While the vertex set $V := \text{Vert}(K)$ of an abstract simplicial complex can consist of points in \mathbb{R}^N for sufficiently high N , the elements of V are often defined and interpreted to be indeterminates. Consequently, the vertex set V and the simplicial complex K generally do not have inherently topological or geometric characteristics. Hence, the modifier *abstract* in the name *abstract simplicial complexes*. As a sidenote, some references call abstract simplicial complexes as *combinatorial simplicial complexes* for the same reason. We provide an example of an abstract simplicial complex below.

Example 1.1.2. Let $V = \{a, b, c, d\}$ be a set of indeterminates. The collection K , as given below, is an abstract simplicial complex:

$$K = \left\{ \begin{array}{l} \{a\}, \{b\}, \{c\}, \{d\}, \\ \{a, b\}, \{a, c\}, \{b, c\}, \{a, d\}, \{b, d\}, \\ \{a, b, c\}, \{a, b, d\} \end{array} \right\}$$

Observe that the non-empty subsets of $\{a, b, c\}$, i.e. the sets $\{a, b\}$, $\{b, c\}$, $\{a, c\}$ and $\{a\}$, $\{b\}$, $\{c\}$ are all in K . The same applies to those of $\{a, b, d\}$.

Abstract simplicial complexes can be related to each other using simplicial maps. We provide a definition below, taken from [Rot88, p141].

Definition 1.1.3. Let K and L be simplicial complexes.

- i. A function $f : K \rightarrow L$ is called a **simplicial map** if for all n -simplices $\sigma = \{v_0, \dots, v_n\}$ of K , $f(\sigma) = \{f(v_0), \dots, f(v_n)\}$ is a simplex of L (not necessarily of dimension n).
- ii. If there exists a simplicial map $f : K \rightarrow L$ such that σ is an n -simplex of K if and only if $f(\sigma)$ is an n -simplex of L , then we call f a **simplicial isomorphism** and say K is **isomorphic** to L . Note that $f : K \rightarrow L$ is a simplicial isomorphism if and only if $f : K \rightarrow L$ is a bijection (between sets of sets).

Note that a simplicial map $f : K \rightarrow L$ is determined by its restriction $f : \text{Vert}(K) \rightarrow \text{Vert}(L)$ into the vertex sets of K and L .

Observe that a simplicial isomorphism is a correspondence between the vertex sets of two simplicial complexes. As such, we can interpret a simplicial isomorphism to be a *relabeling* or *renaming* of the vertices of a simplicial complex. We provide an example of this *renaming* below:

Example 1.1.4. Let K be as given in Example 1.1.2 and define the simplicial complex L using the simplicial map $f : \text{Vert}(K) \rightarrow \text{Vert}(L)$ given by $a \mapsto v_1$, $b \mapsto v_2$, $c \mapsto v_3$, and $d \mapsto v_4$. Then, L is as follows:

$$K = \left\{ \begin{array}{l} \{a\}, \{b\}, \{c\}, \{d\}, \\ \{a, b\}, \{a, c\}, \{b, c\}, \{a, d\}, \{b, d\}, \\ \{a, b, c\}, \{a, b, d\} \end{array} \right\} \xrightarrow{f} L = \left\{ \begin{array}{l} \{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \\ \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_4\}, \{v_2, v_4\}, \\ \{v_1, v_2, v_3\}, \{v_1, v_2, v_4\} \end{array} \right\}$$

We can also define L first as above and define the simplicial map $f : K \rightarrow L$ by $a \mapsto v_1$, $b \mapsto v_2$, $c \mapsto v_3$, and $d \mapsto v_4$, i.e. writing a correspondence between the vertices of K to those of L . In this case, f defines a simplicial

isomorphism.

To gain topological information from abstract simplicial complexes, we relate them to geometric simplicial complexes. As suggested by the modifier *geometric*, these are composed of geometric spaces like lines, triangles, and their n -dimensional analogs and are, therefore, more suited for topology-related interpretations, e.g. as spaces in \mathbb{R}^N . Before we provide a definition for geometric n -simplices, we list some notation and terminology for affine spaces and convex hull below, taken from [Rot88, Chapter 2].

1. A subset $A \subseteq \mathbb{R}^N$ is an **affine space** if for every pair of distinct points $x, y \in A$, the line passing through x and y is contained in A . Equivalently, $A \subseteq \mathbb{R}^N$ is an affine space if there exists $a \in \mathbb{R}^N$ and a vector subspace V of \mathbb{R}^N such that $A = a + V = \{a + v : v \in V\}$, i.e. an affine space is a translated vector subspace of \mathbb{R}^N .
2. Given a set $X \subseteq \mathbb{R}^N$, the **affine hull** $\text{aff}(X)$ of X is the smallest affine space in \mathbb{R}^N that contains X . If $X = \{x_0, x_1, \dots, x_n\}$, $\text{aff}(X)$ is given by

$$\begin{aligned}\text{aff}(X) &= \left\{ \sum_{i=0}^n t_i x_i \in \mathbb{R}^N : \sum_{i=0}^n t_i = 1 \right\} = \left\{ x_0 + \sum_{i=1}^n t_i (x_i - x_0) : t_1, \dots, t_n \in \mathbb{R} \right\} \\ &= x_0 + \text{span}\{x_1 - x_0, \dots, x_n - x_0\}\end{aligned}$$

We call $\text{aff}(X)$ the affine hull **spanned** by X .

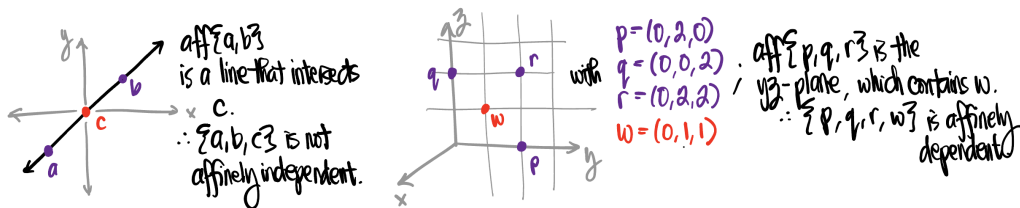
3. A set $X = \{x_0, x_1, \dots, x_n\}$ in \mathbb{R}^N is called **affinely independent** if the only subset Y of X such that $\text{aff}(Y) = \text{aff}(X)$ is X itself. Equivalently, X is affinely independent if the set $\{x_1 - x_0, \dots, x_n - x_0\}$ is linearly independent. X is called **affinely dependent** if X is not affinely independent.
4. Let $A \subseteq \mathbb{R}^N$ be an affine space such that $A = a + V$ for some $a \in \mathbb{R}^N$ and subspace V of \mathbb{R}^N with dimension n (as a vector subspace). Then, the **dimension** of A is n , denoted $\dim(A) = n$, and we call A an **n -plane**. If A is spanned by an affinely independent set $\{x_0, \dots, x_n\}$, then $\dim(A) = n$.
5. The **convex hull** $\text{conv}(X)$ of a set $X \subseteq \mathbb{R}^N$ is the smallest convex set in \mathbb{R}^N containing X . If $X = \{x_0, \dots, x_n\}$, then $\text{conv}(X)$ is given as follows:

$$\text{conv}(X) = \text{conv}(\{x_0, \dots, x_n\}) = \left\{ \sum_{i=0}^n t_i x_i \in \mathbb{R}^N : \sum_{i=0}^n t_i = 1 \text{ with } t_i \geq 0 \right\}$$

Observe that $\text{conv}(X) \subseteq \text{aff}(X)$.

Note that there are various properties listed above that need to be proven, e.g. an affine space is translated vector subspace and $\dim(A)$ is well-defined. We refer to [Rot88, Chapter 2] for such properties. We provide an example below involving affine spaces and the affine independence/dependence of sets.

Example 1.1.5. The sets $A = \{a, b, c\} \subset \mathbb{R}^2$ and $P = \{p, q, r, w\} \subset \mathbb{R}^3$, as illustrated below, are affinely dependent sets. The affine dependence of A can be shown by the non-empty intersection of the affine hull of $\{a, b\}$ (a line) with $\{c\}$ and that of P by the non-empty intersection of the affine hull of $\{p, q, r\}$ (the yz -plane) and $\{w\}$.



Symbolically, $\text{aff}\{A\} = \text{aff}\{a, b\}$ and $\text{aff}\{P\} = \text{aff}\{p, q, r\}$ despite $\{A\} \neq \{a, b\}$ and $\{P\} \neq \{p, q, r\}$. Note that sets $\{a, b\} = A \setminus \{c\}$ and $\{p, q, r\} = P \setminus \{w\}$ are both affinely independent.

Having these definition for affine planes and convex hulls allow a nice characterization of geometric n -simplices, as stated below.

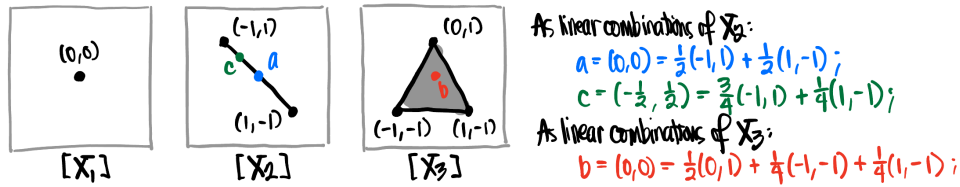
Definition 1.1.6. A **geometric n -simplex** $\sigma = [x_0, \dots, x_n]$ in \mathbb{R}^N is the convex hull $\sigma := \text{conv}(x_0, \dots, x_n)$ of some affinely independent set $X = \{x_0, \dots, x_n\}$ of points in \mathbb{R}^N . The elements x_0, \dots, x_n are called the **vertices** of σ and we say that σ is **spanned** by X . The **dimension** $\dim(\sigma)$ of σ is given by $\dim(\sigma) := n$. When n is arbitrary or unambiguous, we may refer to σ as a **geometric simplex**.

A simplex τ spanned by some $Y \subseteq X$ is called a **face** of σ . When $\dim(\tau) = \dim(\sigma) - 1$, τ is called a **facet** of σ . The **(geometric) boundary** of σ is the collection of all its facets.

Remark. Observe that the terminology for geometric simplices is similar to those of (abstract) simplices. This choice is intentional.

We provide an example below.

Example 1.1.7. Let $X_1 = \{(0, 0)\}$, $X_2 = \{(-1, 1), (1, -1)\}$, and $X_3 = \{(0, 1), (-1, -1), (1, -1)\}$ be collections of points in \mathbb{R}^2 . All X_1, X_2, X_3 are affinely independent. Illustrated below are the simplices spanned by X_1, X_2, X_3 respectively with selected points expressed as linear combinations of the corresponding vertex sets.



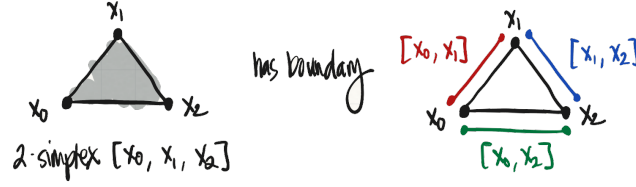
As a consequence of the affine independence requirement of the geometric simplices, it can be proven that each geometric n -simplex is a homeomorphic copy of every other geometric n -simplex, not necessarily living in the same ambient space \mathbb{R}^N . See [Rot88, Theorem 2.10, Exercise 2.2.6]. Intuitively, this means that there is a distinct geometric shape expected of an n -simplex for each $n \in \mathbb{N}_0$. For example:

1. Any geometric 0-simplex is a point.
2. Any geometric 1-simplex is a line, homeomorphic to the 1-disk $D^1 = [-1, 1] \subseteq \mathbb{R}$.
3. Any geometric 2-simplex is a filled-in triangle, homeomorphic to the 2-disk D^2 (a circle with its interior).
4. Any geometric 3-simplex is a solid tetrahedron, homeomorphic to the 3-disk D^3 (a solid sphere).

Thus, there is usually no reference to the ambient space \mathbb{R}^N and we assume that the ambient dimension N is large enough, similarly to how we refer to other topological constructs like S^n (the n -sphere). More generally, an n -simplex is homeomorphic to the n -disk D^n for all $n \in \mathbb{N}_0$ (with D^0 taken to be point $\mathbb{R}^0 = \{0\}$). This relationship also extends to the geometric boundary of n -simplices for $n \geq 1$. For example:

1. The boundary of a geometric 1-simplex $[x_0, x_1]$ is the set $\{[x_1], [x_2]\}$ of 0-simplices, which is homeomorphic to the 1-sphere $S^0 = \{-1, 1\} \subseteq \mathbb{R}$. Note that the topological boundary of $D^1 = [-1, 1]$ as a topological subspace of \mathbb{R} is given by $\partial(D^1) = S^0$.
2. The boundary of a geometric 2-simplex $[x_0, x_1, x_2]$ is the set $\{[x_0, x_1], [x_1, x_2], [x_0, x_2]\}$ of 1-simplices. Observe that the union of these 1-simplices form a loop, which is homeomorphic to the 1-sphere S^1 , the

topological boundary of the 2-disk D^2 as a topological subspace of \mathbb{R}^2 , as illustrated below:



That is, the union of the boundary of a geometric n -simplex is homeomorphic to the topological boundary $\partial D^n = S^{n-1}$ of the n -disk as a subset of \mathbb{R}^n where S^n refers to the n -sphere $S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}$. Once we have established the relationship between geometric simplicial complexes and abstract simplicial complexes, this notion of boundary extends to the simplices of abstract simplicial complexes as well. We see this later in Section 1.2 in the context of simplicial homology.

A geometric simplicial complex is then a collection of these geometric simplices with certain properties. We state this in more detail below.

Definition 1.1.8. A **geometric simplicial complex** K is a collection of simplices in \mathbb{R}^N such that

- i. If $\sigma \in K$, every face of σ also belongs in K .
- ii. If $\sigma, \tau \in K$, then $\sigma \cap \tau$ is either empty or a common face of σ and of τ .

We write $\text{Vert}(K)$ to denote the **vertex set** of K given by the union of all geometric 0-simplices of K . The Euclidean space \mathbb{R}^N is called the **ambient space** of K . A simplicial complex L is called a **subcomplex** of K if $\text{Vert}(L) \subseteq \text{Vert}(K)$. If K is a finite set, then K is called a **finite geometric simplicial complex**.

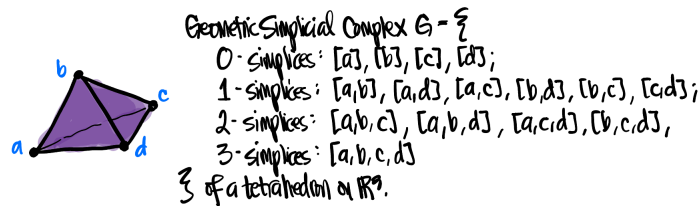
The **underlying space** or **polytope** $|K|$ of K is the topological subspace of \mathbb{R}^N given by $|K| := \bigcup_{\sigma \in K} \sigma$, i.e. the union of all simplices of K . A topological space X is called a **polyhedron** if there exists a geometric simplicial complex K and a homeomorphism $h : |K| \rightarrow X$. In this case, the pair (K, h) is called a **triangulation** of X .

Remark. In this expository paper, we somewhat abuse notation for triangulations of topological spaces by not specifying the homeomorphism $h : |K| \rightarrow X$, i.e. we often say that K is a triangulation of X .

We want to emphasize that geometric simplicial complexes are a specific type of representation of topological spaces and that both conditions, as given in Definition 1.1.8, have to be satisfied. In particular, an arbitrary collection of simplices $\{\sigma_i\}_{i \in I}$ does not make a simplicial complex. We provide some examples below.

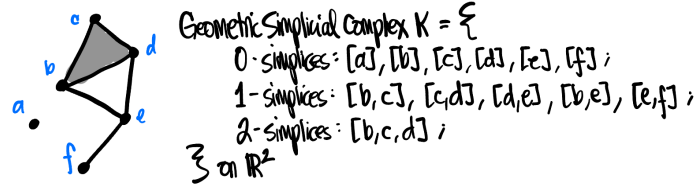
Example 1.1.9. The first two collections of simplices listed below correspond to geometric simplicial complexes but the third does not.

1. The solid tetrahedron in \mathbb{R}^3 as a geometric simplicial complex.



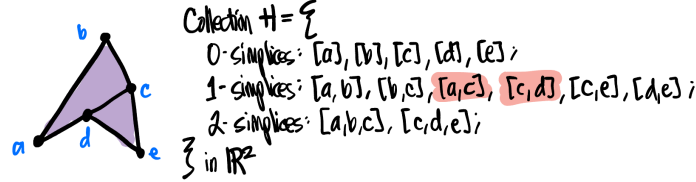
Note that a, b, c, d are points in \mathbb{R}^3 with unidentified coordinates and are not indeterminates.

2. A geometric simplicial complex in \mathbb{R}^2 .



As with the previous item, a, b, c, d, e, f are points in \mathbb{R}^2 .

3. The collection H , as illustrated below, is not a geometric simplicial complex since it violates condition (ii) of Definition 1.1.8. One such pair that violates condition (ii) is the pair $[a, c]$ and $[c, d]$ since their intersection, $[c, d]$, is not a face of $[a, c]$.



Observe that simply adding the simplex $[a, b, c, d]$ to H does not fix the issue. The set $\{a, b, c, d\}$ is not an affinely independent set and, therefore, cannot generate a simplex. That is, the simplex $[a, b, c, d]$ is not defined.

Abstract simplicial complexes and geometric simplicial complexes are related using the following concept:

Definition 1.1.10. The **vertex scheme** K_V of a geometric simplicial complex K_G is the abstract simplicial complex given by $K_V := \{A \subseteq V : A \text{ spans some simplex in } K_G\}$ with $\text{Vert}(K_V) = \text{Vert}(K_G)$ (as sets). Given a simplicial complex K_A , we say that geometric simplicial complex K_G is a **geometric realization** of K_A if the vertex scheme K_V of K_G is isomorphic to K_A .

Observe that, while we can always generate an abstract simplicial complex from a geometric simplicial complex, the converse may not be true. However, if the abstract simplicial complex in question has finite dimension, there always exist a corresponding geometric simplicial complex. We state this as a theorem below.

Theorem 1.1.11. Existence and Uniqueness (up to Homeomorphism) of Geometric Realizations

Every abstract simplicial complex K with dimension $\dim(K) = d$ has a geometric realization in \mathbb{R}^{2d+1} . Furthermore, the polytopes of the geometric realizations of K are unique up to homeomorphism.

Remark. We refer to [Mat03, Theorem 1.6.1] for proof of the existence claim and [Mat03, Proposition 1.5.4] for the claim of uniqueness up to homeomorphism.

The proof of the existence claim relies on constructing a geometric simplicial complex K_G using the *moment curve*. Label the vertices of K by $\text{Vert}(K) = \{v_i\}_{i \in I}$. By assumption, $\dim(\sigma) \leq d$ and $\text{card}(\sigma) \leq d + 1$ for all simplices $\sigma \in K$. We can determine a lower bound on the dimension N of the ambient space \mathbb{R}^N needed for K_G by considering an upper bound on $\text{card}(\sigma \cup \tau)$ with σ and τ distinct simplices of K :

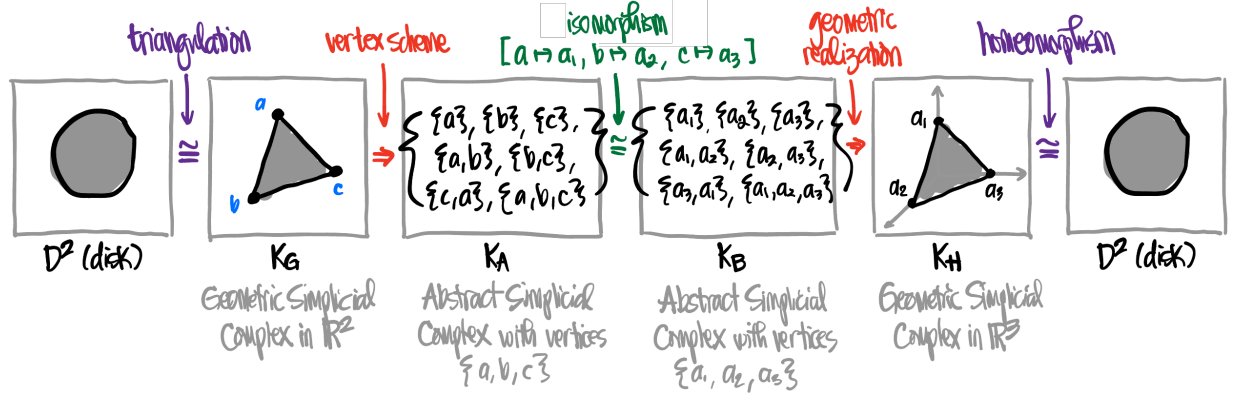
$$\text{card}(\sigma \cup \tau) \leq \text{card}(\sigma) + \text{card}(\tau) \leq (d + 1) + (d + 1) = 2d + 2$$

We define $N = 2d + 1$ and consider the moment curve $\gamma(t) : \mathbb{R} \rightarrow \mathbb{R}^{2d+1}$ given by $\gamma(t) = (t, t^2, \dots, t^{2d+1})$. The moment curve has the property that for any set $G \subseteq \gamma(\mathbb{R})$ with $\text{card}(G) \geq 2d + 2$, any subset of G of cardinality $2d + 2$ is affinely independent (i.e. G is in general position). We construct a vertex

set $X = \{x_i\}_{i \in I}$ of K_G by assigning a unique $t_i \in \mathbb{R}$ for each vertex $v_i \in \text{Vert}(K)$ and defining $x_i := \gamma(t_i)$. Then, we define K_G by the condition: $\{v_{i_0}, \dots, v_{i_n}\}$ is an n -simplex of K if and only if $[x_{i_0}, \dots, x_{i_n}]$ is a geometric n -simplex of K_G .

Observe that this result tells us that every abstract simplicial complex of finite dimension determines a unique topological space by a geometric simplicial complex and that, assuming the dimension is finite, we can go back and forth between representations as convenient. We provide an example below.

Example 1.1.12. The figure below illustrates a number of representations of the 2-disk D^2 .



The correspondence between these representations are identified below:

1. The geometric simplicial complex K_G in \mathbb{R}^2 with $\text{Vert}(K_G) = \{a, b, c\}$ is a triangulation of D^2 with vertex scheme given by the abstract simplicial complex K_A .
2. We relate K_A to another abstract simplicial complex K_B with $\text{Vert}(K_B) = \{a_1, a_2, a_3\}$ using a simplicial isomorphism. We can interpret this to be renaming of the vertices a, b, c of K_A to a_1, a_2, a_3 of K_B respectively.
3. K_B has a geometric realization given by the geometric simplicial complex K_H in \mathbb{R}^3 with $\text{Vert}(K_H) = \{e_1, e_2, e_3\}$ (the standard basis vectors on \mathbb{R}^3). Observe that K_H is a triangulation of D^2 either by inspection of K_H or by application of the uniqueness result of Theorem 1.1.11 on $|K_G| \cong D^2$.

Observe that by Theorem 1.1.11, K_A with $d = 2$ is guaranteed to have a geometric realization in $\mathbb{R}^{2d+1} = \mathbb{R}^5$. Note that the theorem does not state that $2d + 1$ is the minimal dimension and, therefore, this example does not contradict said theorem.

Section 1.2. Simplicial Homology with Coefficients in a PID

In this section, we discuss the homology of simplicial complexes with coefficients in a PID R . Note that the examples presented in this paper mostly consider cases wherein the coefficient ring is $R = \mathbb{Z}$ (in this case, the homology group is an abelian group) or $R = \mathbb{F}$ is a field (in this case, the homology group is a \mathbb{F} -vector space).

The construction of simplicial homology involves the creation of a chain complex of R -modules from an abstract simplicial complex. We use the following definition for chain complexes, taken from [Rot88, p317].

A *chain complex* $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ of R -modules is an \mathbb{Z} -indexed collection of R -modules C_n and R -module homomorphisms $\partial_n : C_n \rightarrow C_{n-1}$ such that for all $n \in \mathbb{Z}$, $\partial_n \circ \partial_{n+1} = 0$. The R -modules C_n are generally called *chain groups* and the homomorphisms ∂_n are called *differentials* or *connecting homomorphisms*.

We follow the convention in [Rot88] and use an asterisk (*) as the “placeholder” for the index of a chain complex. In contrast, other references, e.g. [Rie16; Rot08], use a bullet (•) and write $C_\bullet = (C_n, \partial_n)_{n \in \mathbb{Z}}$. We reserve the use of bullets for persistence modules, introduced later in Chapter 2.

We start with the definition for the chain groups in simplicial homology, adapted from [Rot88, Chapter 7].

Definition 1.2.1. Let K be a simplicial complex and R a PID. For each $n \in \mathbb{Z}$ with $n \geq 0$, define the n^{th} **simplicial chain group** $C_n(K; R)$ of K with coefficients in R to be the R -module with the following presentation:

Generators: All $(n+1)$ -tuples (v_0, v_1, \dots, v_n) such that $v_i \in \text{Vert}(K)$ and $\{v_0, \dots, v_n\}$ is an n -simplex of K .

Relations: For $n = 0$, none. For $n \geq 1$, $(v_0, v_1, \dots, v_n) - (\text{sgn } \pi)(v_{\pi(0)}, v_{\pi(1)}, \dots, v_{\pi(n)})$ for all permutations $\pi : [n] \rightarrow [n]$ with $[n] = \{0, 1, \dots, n\}$ where $\text{sgn } \pi = \pm 1$ refers to the parity of π .

For $n \in \mathbb{Z}$ with $n \leq -1$, define $C_n(K; R)$ to be the trivial R -module. An n -**chain** is an element of $C_n(K; R)$. If $R = \mathbb{Z}$, we often write $C_n(K) := C_n(K; \mathbb{Z})$ and call $C_n(K)$ the n^{th} **simplicial chain group** of K , i.e. without reference to the ring \mathbb{Z} .

- Remarks.* (1) For brevity, we usually use the term “chain group” when referring to a simplicial chain group if the relation to a simplicial complex is clear from context, e.g. we say $C_n(K; R)$ is a chain group of K if it has been stated that K is a simplicial complex.
- (2) The $(n+1)$ -tuple (v_0, v_1, \dots, v_n) described above can be seen as a total order (alternatively, linear order) on the n -simplex $\sigma := \{v_0, v_1, \dots, v_n\}$. In the statement above, it just so happens that the indexing i on the vertices v_i seems to follow the total order on $[n] = \{0, \dots, n\} \subseteq \mathbb{N}_0$. Note that there cannot be repeated vertices on (v_0, \dots, v_n) . Otherwise, $\text{card}(\sigma) < n$ and σ cannot be an n -simplex.

In the context of simplicial homology, the index $n \in \mathbb{Z}$ is usually called the *dimension* since each n -simplex corresponds to an n -dimensional object. We follow this convention throughout this paper to distinguish n from other indices. Note that the dimension being possibly negative for simplicial chain groups is not an issue since $C_n(K; R)$ for $n < 0$ is trivial. In fact, some references like [Hat02; Mun93] do not define such groups since chain complexes in said references are not typically discussed outside a topological perspective.

Observe that Definition 1.2.1 defines simplicial chain groups as quotient modules and that the elements of $C_n(K; R)$ are cosets of formal sums of orderings of n -simplices of K . Relative to the conventional coset notation, elements of $C_n(K; R)$ should be written as sums of terms of the form $[(v_0, \dots, v_n)]$ where (v_0, \dots, v_n) is a specific ordering of the n -simplex $\{v_0, \dots, v_n\}$. To avoid the abundance of grouping symbols and for brevity, we introduce alternative notation below.

Remark 1.2.2. For each n -simplex $\sigma = \{v_0, v_1, \dots, v_n\}$ of some simplicial complex K , define the string $v_0 v_1 \cdots v_n$ and the symbol $[v_0, \dots, v_n]$ to refer to the coset $[(v_0, \dots, v_n)] \in C_n(K; R)$, specifically with coset representative (v_0, \dots, v_n) .

Note that we prefer using the string notation outside this section since we will be defining simplicial homology groups later. That is, we prefer writing $v_0 v_1 \cdots v_n$ for elements of $C_n(K; R)$ so that we write $[v_0 v_1 \cdots v_n]$, as opposed to $[[v_0, v_1, \dots, v_n]]$, for the elements of the simplicial homology group $H_n(K; R)$ (defined later in Definition 1.2.9). The notation $[v_0, \dots, v_n]$ seems to be taken from that of geometric simplices, as stated in Definition 1.1.6.

Next, we state a result about simplicial chain groups being free modules, taken from [Rot88, Lemma 7.10].

Proposition 1.2.3. Let K be a simplicial complex and R a PID. For all $n \in \mathbb{Z}$, $C_n(K; R)$ is a free R -module.

Proof. For $n \leq -1$, $C_n(K; R)$ is trivial and therefore free. Assume $n \geq 0$. Let G_n be the free R -module generated by $(n+1)$ -tuples (v_0, \dots, v_n) such that $\{v_0, \dots, v_n\}$ is an n -simplex of K . Let S_n be the (free) submodule of G_n generated by elements of the form $(v_0, \dots, v_n) - (\text{sgn } \pi)(v_{\pi(0)}, \dots, v_{\pi(n)})$ with $\{v_0, \dots, v_n\}$ an n -simplex of K and $\pi : [n] \rightarrow [n]$ a permutation. By definition, $C_n(K; R) = G_n / S_n$.

For each n -simplex $\sigma = \{v_0, \dots, v_n\}$ of K : define $G_{n,\sigma}$ to be the free R -module generated by all $(n+1)$ -tuples of σ and $S_{n,\sigma}$ to be the (free) submodule of $G_{n,\sigma}$ generated by elements of the form $(v_0, \dots, v_n) - (\text{sgn } \pi)(v_{\pi(0)}, \dots, v_{\pi(n)})$. Then,

$$G_n = \bigoplus_{\substack{\sigma \in K \\ \dim(\sigma)=n}} G_{n,\sigma} \quad , \quad S_n = \bigoplus_{\substack{\sigma \in K \\ \dim(\sigma)=n}} S_{n,\sigma} \quad \text{and} \quad C_n(K; R) = \frac{G_n}{S_n} = \bigoplus_{\substack{\sigma \in K \\ \dim(\sigma)=n}} \left(\frac{G_{n,\sigma}}{S_{n,\sigma}} \right)$$

Therefore, it suffices to check that each $(G_{n,\sigma} / S_{n,\sigma})$ is a free R -module.

Let $\sigma = \{v_0, \dots, v_n\}$ be an n -simplex of K . Note that by labelling the vertices of σ , we are implicitly imposing a total order on σ . Let $(\sigma) := (v_0, \dots, v_n)$. Let $\pi : [n] \rightarrow [n]$ be a permutation and let $(\sigma_\pi) \in G_{n,\sigma}$ refer to the tuple $(\sigma_\pi) := (v_{\pi(0)}, \dots, v_{\pi(n)})$. If π is an identity permutation, then $(\sigma) - (\text{sgn } \pi)(\sigma_\pi) = (\sigma) - (1)(\sigma) = 0$. Otherwise, $(\sigma) - (\text{sgn } \pi)(\sigma_\pi) \neq 0$.

Let $B_{n,\sigma}$ be a set of elements in $G_{n,\sigma}$ such that $(\sigma) - (\text{sgn } \pi)(\sigma_\pi) \in B_{n,\sigma}$ if and only if $\pi : [n] \rightarrow [n]$ is not the identity permutation on $[n]$. Then, $B_{n,\sigma}$ is a basis of $S_{n,\sigma}$ and $\{(\sigma)\} \cup B_{n,\sigma}$ is a basis of $G_{n,\sigma}$. Therefore, the quotient $G_{n,\sigma} / S_{n,\sigma}$ is a free R -module with basis $\{[\sigma]\}$ with $[\sigma] := (\sigma) + S_{n,\sigma}$.

Therefore, $C_n(K; R)$ is a free R -module. A basis B_n of $C_n(K; R)$ can be generated from the union of the bases $\{[\sigma]\}$ for each summand $G_{n,\sigma} / S_{n,\sigma}$. \blacksquare

Remark. In the proof above, we consider the direct sums to correspond to *internal direct sums*, wherein the elements of the direct sum need not be tuples. For example, $R\langle a \rangle \oplus R\langle b \rangle = \{r_1 a + r_2 b : r_1, r_2 \in R\}$ (internal), as opposed to $R\langle a \rangle \oplus R\langle b \rangle = \{(r_1 a, r_2 b) : r_1, r_2 \in R\}$ (external). However, since these two notions uniquely determine an R -module up to R -module isomorphism, these are used somewhat interchangeably.

We want to emphasize that in most texts including [Hat02] and [Rot88], simplicial chain groups are typically defined on *oriented* simplicial complexes. Below, we provide definitions for these notions, relative to our definition of chain groups as stated in Definition 1.2.1.

Definition 1.2.4. Given an n -simplex $\sigma = \{v_0, \dots, v_n\}$ of K , the coset $[v_0, \dots, v_n] := [(v_0, \dots, v_n)] \in C_n(K; R)$ is called an **oriented n -simplex**. An **orientation** on a simplicial complex K is a total order

on the vertex set $\text{Vert}(K)$ of K . An **oriented simplicial complex** is a simplicial complex equipped with an orientation.

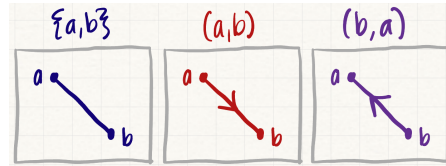
Remark. Recall that a total order \leq on a finite set $A = \{a_0, \dots, a_n\}$ corresponds to a linear ordering of the same set, i.e. an ordered sequence (a_0, a_1, \dots, a_n) of elements of A defines a unique total order $a_0 < a_1 < \dots < a_n$ of A . Following that convention, we usually equip a orientation to a simplicial complex K by using a tuple (v_0, \dots, v_N) corresponding to a total order on $\text{Vert}(K) = \{v_0, \dots, v_N\}$. Note that if the vertices of K are indexed with a totally ordered set (e.g. some subset of \mathbb{Z}), the total order on $\text{Vert}(K)$ is often defined to correspond to the total order on said indexing set.

Observe that Definition 1.2.1 does not require that the simplicial complex be equipped with an orientation beforehand. Here, we interpret orientation on the simplices of K (not of the simplicial complex K) to be more of a consequence of the definition of chain groups. In particular, the relation $(v_0, \dots, v_n) - (\text{sgn } \pi)(v_{\pi(0)}, \dots, v_{\pi(n)}) \in C_n(K; R)$ corresponding to an n -simplex $\sigma = \{v_0, \dots, v_n\}$ of K with $n \geq 1$ in Definition 1.2.1 corresponds to the following equivalence relation \sim on the set of all total orders (written as tuples) on $\sigma = \{v_0, \dots, v_n\}$:

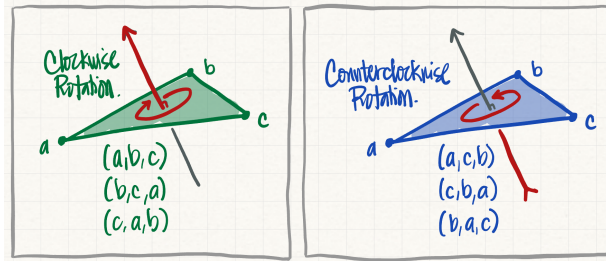
$$\begin{aligned} (v_0, \dots, v_n) &\sim (v_{\pi(0)}, \dots, v_{\pi(n)}) && \text{if } \pi : [n] \rightarrow [n] \text{ is an even permutation} \\ (v_0, \dots, v_n) &\sim -(v_{\pi(0)}, \dots, v_{\pi(n)}) && \text{if } \pi : [n] \rightarrow [n] \text{ is an odd permutation} \end{aligned}$$

Note that parity (i.e. $\text{sgn } \pi$) is not defined for permutations $\pi : \{0\} \rightarrow \{0\}$ on one element. The relation \sim produces two equivalence classes for each n -simplex σ of K with $n \geq 1$, which are often interpreted to correspond to *orientations* of the n -simplex σ as a *geometric* object (hence the name). We list some of these interpretations below:

For $n = 1$: The geometric simplex corresponding to a 1-simplex $\{a, b\}$ is a line segment. Then, an orientation on $\{a, b\}$ is interpreted to be a choice in direction of said line segment, as illustrated below:



For $n = 2$: The geometric simplex corresponding to a 2-simplex $\{a, b, c\}$ is a triangle (more specifically, the convex hull of a triangle). An orientation on this simplex is often interpreted to be a choice in the direction of rotation about an axis normal to the affine 2-plane spanned by $\{a, b, c\}$. The $6 = 3!$ possible total orders on $\{a, b, c\}$ partition into two orientations, as illustrated below:



Note that which rotation is considered “counterclockwise” is determined by the right hand rule on the *chosen* direction of the normal vector. This rotation can also be interpreted as a choice of direction of the vector normal to the plane, denoted in the illustration above by the red part of the axis.

Then, multiplication of $[v_0, \dots, v_n]$ by $-1 \in R$ is interpreted to be a reversal in orientation. Unfortunately, this interpretation loses effectiveness in higher dimensions ($n \geq 3$) as it becomes harder to describe or visualize the

orientation of n -dimensional objects. It does, however, provide a justification for the definition of boundary maps.

Instead, we interpret an orientation on K to correspond to a choice of basis on each of the chain groups $C_n(K; R)$ of K . Since each n -simplex $\sigma = \{v_0, \dots, v_n\}$ of a simplicial complex K is a subset of the vertex set $\text{Vert}(K)$ of K , an orientation on K imposes a total order on σ by restriction. Consequently, an orientation on a simplicial complex K uniquely identifies an $(n+1)$ -tuple for each n -simplex of K . This motivates the following definition:

Definition 1.2.5. Let K be a simplicial complex and R a PID. Let \leq' be a total order on $\text{Vert}(K)$, i.e. an orientation on K . For each $n \in \mathbb{Z}$ with $n \geq 0$, let the **standard ordered basis** \mathcal{K}_n of $C_n(K; R)$ **induced by the orientation on K** be the following set, ordered lexicographically with respect to \leq' :

$$\mathcal{K}_n := \left\{ [v_0, \dots, v_n] : \{v_0, \dots, v_n\} \text{ is an } n\text{-simplex of } K \text{ and } v_0 \leq' \dots \leq' v_n \right\}$$

with $[v_0, \dots, v_n]$ as described in Remark 1.2.2, i.e. the symbol $[v_0, \dots, v_n]$ and the string $v_0 v_1 \dots v_n$ both refer to the coset in $C_n(K; R)$ specifically with coset representative (v_0, \dots, v_n) .

Remark. Observe that there is an immediate one-to-one correspondence between the elements of \mathcal{K}_n and the n -simplices of K . Earlier in the proof of Proposition 1.2.3, $C_n(K; R)$ is shown to be a direct sum of free R -modules $G_{n,\sigma}/S_{n,\sigma}$, one for each n -simplex σ of K with basis determined by some arbitrarily chosen total order (expressed as a tuple) on σ . \mathcal{K}_n simply specifies which total order of σ is chosen for each summand $G_{n,\sigma}/S_{n,\sigma}$. That is, the claim of \mathcal{K}_n being a basis of $C_n(K; R)$ is given by the same proof as Proposition 1.2.3.

Note that the lexicographic order on \mathcal{K}_n is not entirely critical for the construction of chain groups. For example, [Rot88] defines an orientation on K more generally to be a partial order on K that restricts to a total order on σ for each simplex $\sigma \in K$. In this case, we are not guaranteed that every vertex v_k and w_k of K are comparable and lexicographic order may not be well-defined. Then, \mathcal{K}_n can only be considered a *standard basis*, as opposed to a standard *ordered* basis. However, since we will be calculating simplicial homology using matrices in Chapter 4, having a “default” order on \mathcal{K}_n becomes convenient for exposition.

Since we have defined an orientation on a simplicial complex K to be a total order \leq' on $\text{Vert}(K)$, a lexicographic order on \mathcal{K}_n is well-defined. More specifically, we refer to a lexicographic order on the tuples (v_0, \dots, v_n) of the coset representative of $C_n(K; R)$. Given two n -simplices $\{v_0, \dots, v_n\}$ and $\{w_0, \dots, w_n\}$ of K :

$$[v_0, \dots, v_n] \leq [w_0, \dots, w_n] \quad \text{if and only if} \quad \left\{ \begin{array}{l} \text{there exists } k \in [n] \text{ such that} \\ v_i = w_i \text{ for all } i = 0, \dots, k-1 \\ \text{and } v_k \leq' w_k \end{array} \right\}$$

In practice, this means that we compare the k^{th} elements of (v_0, \dots, v_n) and of (w_0, \dots, w_n) in increasing index $k \in \{0, \dots, n\}$ until the vertices v_k and w_k are different. Then, whichever coset comes first in the order of \mathcal{K}_n is determined by whether v_k or w_k comes first in the order of $\text{Vert}(K)$ (i.e. the orientation on K).

Below, we provide an example wherein we identify the chain groups of a simplicial complex K along with the standard bases on said chain groups.

Example 1.2.6. Let K be a simplicial complex on the full simplex of $V = \{a, b, c, d\}$, i.e. the n -simplices of K are exactly the subsets of V of cardinality $n+1$ and K has the geometric realization of a solid tetrahedron.

Equip K with the orientation by $V = (a, b, c, d)$. Observe that the total order on $\text{Vert}(K) = V$ is denoted using a tuple. Given below are the standard ordered bases \mathcal{K}_n on the chain groups $C_n(K) = C_n(K; \mathbb{Z})$ of the

oriented simplicial complex K :

$$\begin{aligned}\mathcal{K}_0 &= ([a], [b], [c], [d]) & \mathcal{K}_2 &= ([a, b, c], [a, b, d], [a, c, d], [b, c, d]) \\ \mathcal{K}_1 &= ([a, b], [a, c], [a, d], [b, c], [b, d], [c, d]) & \mathcal{K}_3 &= ([a, b, c, d])\end{aligned}$$

with $\mathcal{K} = \emptyset$ for $n \in \mathbb{Z}$ with $n \neq 0, 1, 2, 3$. Observe that for $n = 2$, the oriented n -simplex $[a, b, c]$ comes before $[a, b, d]$ since the first two entries match and we have $c \leq d$ as elements of V on the third entry. Similarly, $[a, c, d]$ is listed before $[b, c, d]$ since on the first entry, we have $a \leq b$ on the first entry.

Following Remark 1.2.2, we may write $[v_0, \dots, v_n]$ as a string $v_0 \cdots v_n$ of vertices for brevity. Then, we can describe the n^{th} chain groups $C_n(K)$ of K as follows:

$$C_n(K) = \begin{cases} \mathbb{Z}\langle a, b, c, d \rangle & \text{if } n = 0 \\ \mathbb{Z}\langle ab, ac, ad, bc, bd, cd \rangle & \text{if } n = 1 \\ \mathbb{Z}\langle abc, abd, acd, bcd \rangle & \text{if } n = 2 \\ \mathbb{Z}\langle abcd \rangle & \text{if } n = 3 \\ 0 & \text{otherwise} \end{cases}$$

Examples of elements of $C_2(K)$ include $\sigma_1 = 2abc - abd$ and $\sigma_2 = 3acd + 7abc - 10bcd$.

Next, we provide a definition for the boundary map related to the simplicial chain groups.

Definition 1.2.7. Let K be a simplicial complex and R a PID. For each $n \in \mathbb{Z}$ with $n \geq 0$, define the n^{th} **boundary map** or the n^{th} **boundary homomorphism** $\partial_n : C_n(K; R) \rightarrow C_{n-1}(K; R)$ to be the R -module homomorphism given by

$$\partial_n([v_0, \dots, v_n]) := \sum_{i=0}^n (-1)^i [v_0, \dots, \hat{v}_i, \dots, v_n] = \sum_{i=0}^n (-1)^i [v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_n] \quad (\text{E1})$$

where \hat{v}_i indicates the removal of the vertex v_i in the ordering (v_0, \dots, v_n) . For $n < 0$, $\partial_n : C_n(K; R) \rightarrow C_{n-1}(K; R)$ can only be the trivial homomorphism since $C_n(K; R) = 0$.

An n -**cycle** is an element of $\ker(\partial)_n \subseteq C_n(K; R)$ and an n -**boundary** is an element of $\text{im}(\partial_{n+1}) \subseteq C_{n+1}(K; R)$.

Remark. It can be verified that Equation (E1) is well-defined for any choice of coset representative of $\sigma \in C_n(K; R)$ (relative to Definition 1.2.1) and produces a well-defined homomorphism. In particular, for any oriented n -simplex $[\sigma] = [v_0, \dots, v_n]$ with an arbitrary coset representative (v_0, \dots, v_n) and for any permutation $\pi : [n] \rightarrow [n]$, it can be verified that

$$\begin{aligned}\partial_n([v_0, \dots, v_n]) &:= \sum_{i=0}^n (-1)^i [v_0, \dots, \hat{v}_i, \dots, v_n] = (\text{sgn } \pi) \sum_{i=0}^n (-1)^i [v_{\pi(0)}, \dots, \hat{v}_{\pi(i)}, \dots, v_{\pi(n)}] \\ &=: (\text{sgn } \pi) \partial_n([v_{\pi(0)}, \dots, v_{\pi(n)}]).\end{aligned}$$

with the first and second equality being given by Equation (1).

Next, we state the property that determines that the collection of chain groups and boundary maps form a chain complex.

Proposition 1.2.8. For each $n \in \mathbb{Z}$, $\partial_n \partial_{n+1} = 0$. That is, $\text{im}(\partial_{n+1}) \subseteq \ker(\partial_n)$.

Remark. This can be proven by direct calculation on a coset $[v_0, \dots, v_{n+1}] \in C_{n+1}(K; R)$ with representative (v_0, \dots, v_{n+1}) . A similar argument is presented under the proof of [Rot88, Theorem 4.6].

Observe that this property on $\{\partial_n\}$ allows us to take the quotient of the R -module $\ker(\partial_n)$ by its submodule $\text{im}(\partial_{n+1})$. This quotient is exactly the simplicial homology of a simplicial complex. We state this in more detail below.

Definition 1.2.9. The **simplicial chain complex** of a simplicial complex K with coefficients in a PID R is the chain complex $C_*(K; R) := (C_n(K; R), \partial_n)_{n \in \mathbb{Z}}$ of simplicial chain groups $C_n(K; R)$ and boundary maps $\partial_n : C_n(K; R) \rightarrow C_{n-1}(K; R)$, illustrated as the following sequence:

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}(K; R) \xrightarrow{\partial_{n+1}} C_n(K; R) \xrightarrow{\partial_n} C_{n-1}(K; R) \xrightarrow{\partial_{n-1}} \cdots$$

For each $n \in \mathbb{Z}$, define the n^{th} **simplicial homology group** $H_n(K; R)$ and the n^{th} **Betti number** $\beta_n(K; R)$ of K with coefficients in R as follows:

$$H_n(K; R) := \frac{\ker(\partial_n)}{\text{im}(\partial_{n+1})} \quad \text{and} \quad \beta_n(K; R) := \text{rank}(H_n(K; R))$$

If $R = \mathbb{Z}$, we write $C_*(K) := C_*(K; \mathbb{Z})$ and $H_n(K) := H_n(K; \mathbb{Z})$.

Remark. We follow the convention by [Rot88] and use $(*)$ as the “placeholder” of the index $n \in \mathbb{Z}$ of the chain complex $C_* = (C_n, \partial_n)$ with chain groups C_n and differentials $\partial_n : C_n \rightarrow C_{n-1}$.

Observe that if the simplicial complex K is a finite simplicial complex, then the n^{th} chain group $C_n(K; R)$ must also be finitely generated for all $n \in \mathbb{Z}$. In this case, $H_n(K; R)$ is a finitely-generated module over a PID R and the Structure Theorem (Theorem 4.1.1) for finitely-generated modules over R applies. We discuss this and a method of calculating the homology of free chain complexes using matrices over R later in Chapter 4.

One of the key characteristics of simplicial homology is that it is an invariant of the homeomorphism type of topological spaces. We state this in a theorem below.

Theorem 1.2.10. Let K and L be simplicial complexes and let R be a PID. If the geometric realizations of K and L are homeomorphic as topological spaces, then $H_n(K; R) \cong H_n(L; R)$ as R -modules.

Remark. We refer to [Rot88, Theorem 7.13] for a proof. This theorem actually generalizes to the case of *homotopy equivalence*, which is an equivalence weaker than a homeomorphism.

Note that there are other ways to define the homology of a topological space. With simplicial homology, we can define the homology group $H_n(X; R)$ of a *topological space* X to be the homology group $H_n(K; R)$ of a triangulation K of X . Observe that this is well-defined by Theorem 1.2.10, i.e. $H_n(K; R)$ is determined up to isomorphism regardless of the triangulation K . Note that this definition requires that a triangulation of X exist in the first place. Homology can also be defined using other representations of spaces, e.g. singular homology of topological spaces, simplicial homology of Δ -complexes, and cellular homology of CW-complexes. The key point here is that all these homology theories must produce isomorphic homology groups as with simplicial homology. We refer to the discussion in [Hat02, Theorem 2.27] about singular homology and [Hat02, Theorem 2.35] about cellular homology.

We can take advantage of this by comparing results from other homology theories against results from

simplicial homology. This becomes particularly useful when the triangulations of a relatively simple space are cumbersome or unwieldy. We provide an example of this below.

Example 1.2.11. Let K be the abstract simplicial complex given as follows, with simplices written as strings (following Remark 1.2.2) for brevity and with vertex indices always taken modulo 3.

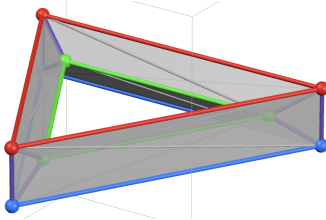
$$\begin{aligned}
 \text{nine 0-simplices of } K & : \quad \text{Vert}(K) = \{a_0, a_1, a_2, b_0, b_1, b_2, c_0, c_1, c_2\} \\
 \text{twenty-seven 1-simplices of } K & : \quad \left\{ \begin{array}{ccc} a_i a_{i+1} & b_i b_{i+1} & c_i c_{i+1} \\ a_i b_i & b_i c_i & c_i a_i \\ a_i b_{i+1} & b_i c_{i+1} & c_i a_{i+1} \end{array} \text{ for } i = 0, 1, 2 \right\} \\
 \text{eighteen 2-simplices of } K & : \quad \left\{ \begin{array}{ccc} a_i a_{i+1} b_{i+1}, & b_i b_{i+1} c_{i+1}, & c_i c_{i+1} a_{i+1}, \\ a_i b_i b_{i+1}, & b_i c_i c_{i+1}, & c_i a_i a_{i+1} \end{array} \text{ for } i = 0, 1, 2 \right\}
 \end{aligned}$$

We can interpret K to be a geometric simplicial complex in \mathbb{R}^3 by defining the vertices of K as follows:

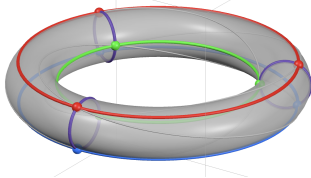
$$\begin{aligned}
 \text{For each } i \in \{0, 1, 2\} : \quad a_i &:= T\left(\frac{2\pi}{3}i, \frac{\pi}{3}\right), \quad b_i := T\left(\frac{2\pi}{3}i, -\frac{\pi}{3}\right), \quad c_i := T\left(\frac{2\pi}{3}i, -\pi\right), \\
 \text{with } T(u, v) &= \begin{pmatrix} \cos(u)(R + r \cos(v)) \\ \sin(u)(R + r \cos(v)) \\ r \sin(v) \end{pmatrix} \text{ with } r = 1 \text{ and } R = 5.
 \end{aligned}$$

We illustrate K as a geometric simplicial complex below and claim that there exists a triangulation $|K| \rightarrow T^2$ to the torus T^2 parametrized by $T(u, v)$ with $(u, v) \in \mathbb{R} \times \mathbb{R}$.

The Geometric
Simplicial Complex K :



The Torus T^2
by $T(u, v)$, $(u, v) \in \mathbb{R} \times \mathbb{R}$:



We use the following color scheme for the simplices of K and their corresponding images in T^2 under the triangulation $|K| \rightarrow T^2$.

- red : In K , the vertices a_i and the three 1-simplices $a_i a_{i+1}$ for $i \in \{0, 1, 2\}$. In T^2 , the major circle given by $T(u, \frac{\pi}{3})$ with $u \in \mathbb{R}$.
 - blue : In K , the vertices b_i and the three 1-simplices $b_i b_{i+1}$ for $i \in \{0, 1, 2\}$. In T^2 , the major circle given by $T(u, -\frac{\pi}{3})$ with $u \in \mathbb{R}$.
 - green : In K , the vertices c_i and the three 1-simplices $c_i c_{i+1}$ for $i \in \{0, 1, 2\}$. In T^2 , the major circle given by $T(u, -\pi)$ with $u \in \mathbb{R}$.
 - purple : In K , the nine 1-simplices $a_i b_i$, $b_i c_i$, and $c_i a_i$ for $i = 0, 1, 2$. In T^2 , the three minor circles connecting the vertices a_i, b_i, c_i given by $T(\frac{2\pi}{3}i, v)$ for $v \in \mathbb{R}$ for each $i = 0, 1, 2$.
 - gray : In K , all eighteen 2-simplices (triangles) and the nine 1-simplices $a_i b_{i+1}$, $b_i c_{i+1}$, and $c_i a_{i+1}$ for $i \in \{0, 1, 2\}$.
- For each $i \in \{0, 1, 2\}$, the thin gray loop in T^2 that connects the vertices a_i, b_{i+1}, c_{i+2} are given by $T(t, \frac{\pi}{3} + \frac{2\pi}{3}i - t)$ with $t \in \mathbb{R}$.

Observe that K has $9 + 27 + 18 = 54$ simplices and calculating its homology groups by hand can be very cumbersome to do by hand, e.g. $C_1(K; \mathbb{Z})$ has 27 basis elements. Since K is a triangulation of T^2 , we can use Theorem 1.2.10 and the known homology groups of T^2 to determine $H_n(K; \mathbb{Z})$ up to \mathbb{Z} -module isomorphism as

follows:

$$H_n(K_T; \mathbb{Z}) \cong H_n(T^2; \mathbb{Z}) = H_n(T^2) \cong \begin{cases} \mathbb{Z} & \text{if } n = 0 \\ \mathbb{Z} \oplus \mathbb{Z} & \text{if } n = 1 \\ \mathbb{Z} & \text{if } n = 2 \\ 0 & \text{if } n \geq 3 \end{cases}$$

For contrast, [Hat02, Example 2.3, p106] uses delta-complexes to calculate $H_n(T^2)$ and the chain groups corresponding to $C_0(T^2), C_1(T^2), C_2(T^2)$ have 1, 3, and 2 basis elements respectively. As a sidenote, it has been proven that the minimal triangulation of T^2 consists of fourteen 2-simplices (triangles).

Section 1.3. Functorial Constructions in Simplicial Homology

In this section, we discuss the functorial nature of the construction of the simplicial chain complex $C_*(K; R) = (C_n(K; R), \partial_n)_{n \in \mathbb{Z}}$ for a simplicial complex K , as discussed in Section 1.2. For reference, we briefly discuss categories and functors in Appendix A4. We consider the following categories in this section:

1. The category **Mod** $_R$ of modules over a PID R and R -module homomorphisms.
2. The category **Ch-Mod** $_R$ of chain complexes of R -modules and of chain maps.

Note that the category **Ch-Mod** $_R$ has an accompanying *chain homology functor* $H_n(-) : \mathbf{Ch-Mod}_R \rightarrow \mathbf{Mod}_R$ with the following assignments: Let $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ and $A_* = (A_n, \alpha_n)_{n \in \mathbb{Z}}$ be chain modules with R -modules C_n and A_n and differentials $\partial_n : C_n \rightarrow C_{n-1}$ and $\alpha_n : A_n \rightarrow A_{n-1}$.

1. $H_n(-)$ maps C_* to its n^{th} homology group $H_n(C_*)$ defined by $H_n(C_*) = \ker(\partial_n) / \text{im}(\partial_{n+1})$. Note that this is well-defined since $\text{im}(\partial_{n+1}) \subseteq \ker(\partial_n)$ by the definition of chain complexes.
2. Given a chain map $f_* : C_* \rightarrow A_*$ with $f_* = \{f_n : C_n \rightarrow A_n\}_{n \in \mathbb{Z}}$, $H_n(-)$ maps f_* to the R -module homomorphism $H_n(C_*) \rightarrow H_n(A_*)$ induced by the map $f_n : C_n \rightarrow A_n$ and the quotient (or cokernel) construction. Note that the map on $H_n(C_*) \rightarrow H_n(A_*)$ is well-defined since $f_{n-1}\partial_n = \alpha_n \circ f_n$ by definition of a chain map.

We claim that simplicial homology, as discussed in Section 1.2, corresponds to a composition of functors: one that creates the simplicial chain complex $C_*(K; R)$ of a simplicial complex K and the n^{th} homology functor on **Ch-Mod** $_R$. Recall that a functor $F : \mathbf{C} \rightarrow \mathbf{D}$ between categories **C** and **D** is an assignment of objects and morphisms of **C** to those of **D** such that the following properties are satisfied:

- (F1) For each morphism $f : X \rightarrow Y$ in **C**, the functor maps f to a morphism with domain $F(X)$ and codomain $F(Y)$, i.e. $f \mapsto F(f) : F(X) \rightarrow F(Y)$.
- (F2) The functor respects *identity* of objects and of morphisms: For all objects X in **C** with identity morphism $\text{id}_X : X \rightarrow X$, the morphism $F(\text{id}_X) : F(X) \rightarrow F(X)$ is exactly the identity of $F(X)$ in **D**, i.e. $F(\text{id}_X) = \text{id}_{F(X)}$.
- (F3) The functor respects *composition* of morphisms: For all composable pairs of morphisms $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, $F(g \circ f) = F(g) \circ F(f)$.

In this section, we argue that the construction of the n^{th} chain group $C_n(K; R)$ of a simplicial complex K and of the simplicial chain complex $C_*(K; R)$ correspond to the functors by defining the morphism assignments for both and showing that the properties above are satisfied. To start, we state a result involving the formation of the category of simplicial complexes.

Theorem 1.3.1. Abstract simplicial complexes and simplicial maps form a well-defined category.

Remark. This theorem is taken from [Rot88, Theorem 7.7], which states “a routine check” as proof. We believe this refers to checking that the axioms listed in the definition of category, i.e. as listed Definition A4.1, are satisfied. Since simplicial maps can be seen as functions between sets of sets, the composition law and the identity morphism designed to each abstract simplicial complex follow those on the category **Set** of sets and functions. Then, it suffices to check that composition is unital with identity morphisms and is associative, i.e. Definition A4.1(iv,v).

Definition 1.3.2. Denote the category given in Theorem 1.3.1 by **A-Simp** and call it the **category of simplicial complexes and simplicial maps**, or the **category of simplicial complexes** for convenience.

We want to show that the construction of the chain groups in Definition 1.2.1 correspond to a functor $\mathbf{A-Simp} \rightarrow \mathbf{Mod}_R$ with \mathbf{Mod}_R being the category of R -modules and R -module homomorphisms. Definition 1.2.1 determines the object assignment. Below, we provide a definition for the corresponding morphism assignment, taken from [Rot88].

Definition 1.3.3. Let $f : K \rightarrow L$ be a simplicial map between simplicial complexes K and L . For each $n \in \mathbb{N}_0$, define the **homomorphism** $f_{n,\#} : C_n(K; R) \rightarrow C_n(L; R)$ **on the n^{th} chain groups induced by f** as follows:

$$f_{n,\#}([\sigma]) = f_{n,\#}([v_0, v_1, \dots, v_n]) := \begin{cases} [f(v_0), f(v_1), \dots, f(v_n)] & \text{if } f(\sigma) \text{ is an } n\text{-simplex in } L \\ 0 & \text{otherwise} \end{cases}$$

where (v_0, \dots, v_n) is the coset representative of $[\sigma] \in C_n(K; R)$ with $\sigma = \{v_0, \dots, v_n\}$ and $(f(v_0), \dots, f(v_n))$ that of $f_{n,\#}([\sigma]) \in C_n(L; R)$ of the n -simplex $f(\sigma) = \{f(v_0), \dots, f(v_n)\}$.

Remark. For brevity, we often suppress the dimension n in $f_{n,\#}$, i.e. we write $f_{\#}$ for $f_{n,\#}$.

Note that this construction is defined on the coset representatives of $C_n(K; R)$. Consequently, if L is equipped with an orientation and the standard basis $L[n]$ by Definition 1.2.5 is used, appropriate sign changes might be needed for $[f(v_0), \dots, f(v_n)] \in C_n(L; R)$, i.e.

$$[f(v_0), \dots, f(v_n)] = (\text{sgn } \pi) [f(v_{\pi(0)}), \dots, f(v_{\pi(n)})]$$

where $\pi : [n] \rightarrow [n]$ permutes the vertices of $f(\sigma)$ such that $(f(v_{\pi(0)}), \dots, f(v_{\pi(n)}))$ is ordered with respect to the orientation on L .

Observe that the definition $f_{n,\#} : C_n(K; R) \rightarrow C_n(L; R)$ above already satisfies property F1. Below, we state that this also satisfies the other two properties required for functors.

Proposition 1.3.4. Let R be a PID and let $n \in \mathbb{Z}$.

- i. Let K be a simplicial complex with identity simplicial map $\text{id}_K : K \rightarrow K$. Then, $(\text{id}_K)_{\#} = \text{id}_{C_n(K; R)}$.
- ii. Let $f : K_1 \rightarrow K_2$ and $g : K_2 \rightarrow K_3$ be simplicial maps on simplicial complexes K_1 , K_2 , and K_3 . Then, $(g \circ f)_{\#} = g_{\#} \circ f_{\#}$.

Remark. This can be proven by calculation on some arbitrary element $[\sigma] = [v_0, \dots, v_n]$ of $C_n(K; R)$ or $C_n(K_1; R)$ with $\sigma = \{v_0, \dots, v_n\}$ an n -simplex of K .

Since Definition 1.3.3 satisfies properties F1, F2, and F3 for functors, we can define the construction of the n^{th} chain group $C_n(K; R)$ as a functor, as given below.

Definition 1.3.5. For each $n \in \mathbb{Z}$, define n^{th} **simplicial chain group functor** $C_n(-; R) : \mathbf{A-Simp} \rightarrow \mathbf{Mod}_R$ with coefficients in a PID R as follows:

- i. The object assignment maps a simplicial complex K to the n^{th} chain group $C_n(K; R)$ as given by Definition 1.2.1.
- ii. The morphism assignment maps a simplicial map $f : K \rightarrow L$ between simplicial complexes K and L to map $f_{\#} : C_n(K; R) \rightarrow C_n(L; R)$ given by Definition 1.3.3.

If $R = \mathbb{Z}$, we write $C_n(-) := C_n(-; \mathbb{Z})$.

To construct the functor $\mathbf{A-Simp} \rightarrow \mathbf{Ch-Mod}_R$, we claim that we can simply collect all maps $f_{n,\#} :$

$C_n(K; R) \rightarrow C_n(L; R)$ on the chain groups induced by a simplicial map $f : K \rightarrow L$ to define a chain map between the simplicial chain complexes $C_*(K; R)$ and $C_*(L; R)$. The following result allows us to do this:

Proposition 1.3.6. Let $f : K \rightarrow L$ be a simplicial map between simplicial complexes K and L . For all $n \in \mathbb{Z}$,

$$f_{n-1, \#} \circ \partial_n^K = \partial_n^L \circ f_{n, \#}$$

where $\partial_n^K : C_n(K; R) \rightarrow C_{n-1}(K; R)$ and $\partial_n^L : C_n(L; R) \rightarrow C_{n-1}(L; R)$ refer to the n^{th} simplicial boundary map of K and L respectively, i.e. the following diagram commutes:

$$\begin{array}{ccc} C_n(K; R) & \xrightarrow{\partial_n^K} & C_{n-1}(K; R) \\ f_{n, \#} \downarrow & & \downarrow f_{n-1, \#} \\ C_n(L; R) & \xrightarrow{\partial_n^L} & C_{n-1}(L; R) \end{array}$$

That is, the collection $\{f_{n, \#}\}_{n \in \mathbb{Z}}$ determines a chain map $\{f_{n, \#}\} : C_*(K; R) \rightarrow C_*(L; R)$.

Remark. This can be proven by direct calculation on an arbitrary element $[v_0, \dots, v_n] \in C_n(K; R)$ with coset representative (v_0, \dots, v_n) on an n -simplex $\{v_0, \dots, v_n\}$ of K . Note that boundary maps, as given in Definition 1.2.7, are well-defined on any choice of coset representative.

Observe that the chain map given by $\{f_{n, \#}\}_{n \in \mathbb{Z}}$, as denoted above, already satisfies the domain and codomain property of functors as stated in Property F1. The fact that $\{f_{n, \#}\}_{n \in \mathbb{Z}}$ respects identity as in Property F2 and respects composition as in Property F3 follows from those properties being satisfied by $f_{n, \#}$ for each $n \in \mathbb{Z}$ separately. Therefore, we can define a functor on $\mathbf{A-Simp} \rightarrow \mathbf{Ch-Mod}_R$ as follows:

Definition 1.3.7. Define the n^{th} **simplicial chain complex functor** $C_*(-; R) : \mathbf{A-Simp} \rightarrow \mathbf{Ch-Mod}_R$ with coefficients in a PID R as follows:

- i. $C_*(-; R)$ maps a simplicial complex K to the simplicial chain complex $C_*(K; R) = (C_n(K; R), \partial_n)_{n \in \mathbb{Z}}$ as given by Definition 1.2.9.
- ii. $C_*(-; R)$ maps a simplicial map $f : K \rightarrow L$ between simplicial complexes K and L to the collection $\{f_{n, \#}\}_{n \in \mathbb{Z}}$ of maps $f_{n, \#} : C_n(K; R) \rightarrow C_n(L; R)$ as given by Definition 1.3.3.

If $R = \mathbb{Z}$, we write $C_*(-) := C_*(-; \mathbb{Z})$ for brevity.

Bringing this all together, we have the following proposition:

Proposition 1.3.8. Let K be a simplicial complex and R a PID. For all $n \in \mathbb{Z}$:

$$H_n(K; R) = \left(H_n \circ C_*(-; R) \right) (K) = \left(H_n(C_*(K; R)) \right)$$

where $H_n(K; R)$ is the simplicial chain group of K by Definition 1.2.9, $H_n : \mathbf{Ch-Mod}_R \rightarrow \mathbf{Mod}_R$ is the chain homology functor, and $C_*(-; R)$ is the simplicial chain complex functor by Definition 1.3.7.

Proof. The object assignment by $H_n(-) \circ C_*(-; R)$ is exactly as described in Definition 1.2.9 for $H_n(K; R)$. Note that this proposition disregards the morphism assignment of the relevant functors. ■

For convenience, we define the composition of the simplicial chain complex functor and the homology functor to be the simplicial chain group functor.

Definition 1.3.9. For each $n \in \mathbb{Z}$, define the n^{th} **simplicial homology functor** $H_n(-; R) : \mathbf{A-Simp} \rightarrow \mathbf{Mod}_R$ with coefficients in a PID R as the following composition of functors:

$$H_n(-; R) := H_n \circ C_n(-; R)$$

where $H_n : \mathbf{Ch-Mod}_R \rightarrow \mathbf{Mod}_R$ refers to the chain homology functor. Given a simplicial map $f : K \rightarrow L$, let $f_* : H_n(K; R) \rightarrow H_n(L; R)$ be the map on homology induced by application of the functor $H_n(-; R)$.

Below, we identify a result that will be useful later in Section 3.3 in the context of persistent homology.

Lemma 1.3.10. Let L be a subcomplex of some simplicial complex K and let $i : L \rightarrow K$ be the corresponding inclusion map. For each $n \in \mathbb{Z}$:

- i. The n^{th} chain group $C_n(L; R)$ of L is a submodule of $C_n(K; R)$ and for all $[\sigma] \in C_n(L; R)$, $i_{n,\#}([\sigma]) = \text{id}_{C_n(K; R)}([\sigma]) = [\sigma]$ where $\text{id}_{C_n(K; R)}$ refers to the identity map on $C_n(K; R)$.
- ii. For all $[\sigma] \in C_n(L; R)$, $\partial_n^L([\sigma]) = \partial_n^K([\sigma])$ where $\partial_n^L : C_n(L; R) \rightarrow C_{n-1}(L; R)$ refers to the boundary map on L and $\partial_n^K : C_n(K; R) \rightarrow C_{n-1}(K; R)$ that on K .

Proof. Let $n \in \mathbb{Z}$. If $n < 0$, then $C_n(L; R) = 0$, $C_n(K; R) = 0$, and the proposition is trivially satisfied. Assume $n \geq 0$. Since $i : L \rightarrow K$ is an inclusion map, $i(\sigma) = \text{id}_K(\sigma)$ for all n -simplices $\sigma \in L$. By Proposition 1.3.4, $i_{n,\#} = \text{id}(C_n(K; R)) = (\text{id}_K)_{\#}$. By Proposition 1.3.6, for $[\sigma] \in C_n(L; R)$:

$$(i_{n,\#} \circ \partial_n^L)([\sigma]) = (\partial_n^K \circ i_{n,\#})([\sigma]) = \partial_n^K([\sigma])$$

as desired. ■

Chapter 2. Introduction to Persistence Theory

The study of persistence modules or *persistence theory* first came about in the study of persistent homology. Herbert Edelsbrunner, David Letscher, and Afra Zomorodian in the paper *Topological Persistence and Simplification* [ELZ02] had first defined persistence in terms of homology classes of some collection of spaces. In particular, persistence theory is interested in the following quantity:

$$\text{rank}\left(H_n(X_t; R) \xrightarrow{i_*} H_n(X_s; R)\right)$$

where X_t and X_s are two topological spaces such that $X_t \subseteq X_s$ and R is some ring. Here, the term *persistence* refers to how homology classes in $H_n(X_t; R)$ map to $H_n(X_s; R)$ under the homomorphism i_* induced by the inclusion $i : X_t \rightarrow X_s$, where we say a homology class $[\sigma] \in H_n(X_t; R)$ *persists* if $i_*([\sigma]) \neq 0$.

Persistence theory was later generalized. For example, Afra Zomorodian and Gunnar Carlsson in [ZC05] (published in February 2005) then introduced a definition for persistence modules in terms of R -modules and considered persistent homology to be a specific example of a persistence module. In particular, a persistence module was defined to be a collection $\{M_t\}_{t \in \mathbb{N}_0}$ of R -modules together with a collection of homomorphisms $\{\varphi_t : M_t \rightarrow M_{t+1}\}_{t \in \mathbb{N}_0}$. Relative to this characterization, persistence theory studies the following quantity:

$$\text{rank}\left(M_t \xrightarrow{\varphi_t} M_{t+1}\right).$$

However, definitions for certain constructions involving persistence modules may seem arbitrary or unmotivated given this collection definition. For example, the direct sum between persistence modules $\{M_t\}_{t \in \mathbb{N}_0}$ and $\{N_t\}_{t \in \mathbb{N}_0}$ is defined to be done pointwise as follows:

$$\{M_t\}_{t \in \mathbb{N}_0} \oplus_{\text{Pers}} \{N_t\}_{t \in \mathbb{N}_0} := \left\{M_t \oplus_{\text{Mod}} N_t\right\}_{t \in \mathbb{N}_0}$$

where \oplus_{Pers} refers to the direct sum of persistence modules and \oplus_{Mod} to that of R -modules. This notion of direct sum between persistence modules is then related to that between graded $R[x]$ -modules, without much justification as to how the definition for \oplus_{Pers} is consistent with the categorical definition (so as to be consistent with that of \oplus_{Mod}).

Later, Bubenik and Scott in the paper *Categorification of Persistence Modules* [BS14] (presented in May 2014) introduced a new definition for persistence modules in terms of category theory: one that covers the ideas of Zomorodian and Carlsson, provides a more robust foundation to the theory, and is generalizable to concepts such as zigzag persistence and multiparameter persistence. This new definition in terms of functors is often used in the current literature for persistence theory.

In this chapter, we discuss persistence theory relative to this characterization of persistence modules as functors, and present results involving persistence modules relevant to concepts discussed in [ZC05]. This chapter is organized as follows:

Section 2.1. Persistence Modules as Functors.

We define persistence modules over a field \mathbb{F} as functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$ and introduce a type of persistence module called a *finite-type* persistence module.

Section 2.2. The Category of Persistence Modules.

We discuss the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules over \mathbb{F} , defined to be the category of functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$. We also provide definitions for several algebraic constructions involving persistence modules, e.g. isomorphisms, direct sums, and chain complexes, arising from this characterization as a functor category and briefly discuss how these definitions are consistent with their corresponding categorical definitions.

Section 2.3. Interval Decompositions of Persistence Modules.

We introduce a specific type of direct sum decomposition of a persistence module called an *interval decomposition*, which can be proven to be unique up to persistence isomorphism, and discuss the notion of a *persistence barcode*.

Section 2.4. The Category of Graded Modules over Polynomial Rings

We discuss the category $\mathbf{GrMod}_{R[x]}$ of graded modules over the polynomial ring $R[x]$ with R a PID and review relevant definitions, terminology, and results involving these graded modules. We also introduce *graded invariant factor decompositions* for graded $\mathbb{F}[x]$ -modules, which correspond to invariant factor decompositions that respect the graded structure, and present the *Graded Structure Theorem*.

Section 2.5. The Equivalence between Persistence Modules and Graded Modules

We present an isomorphism of categories between $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ and discuss how interval decompositions of persistence modules over \mathbb{F} correspond to graded invariant factor decompositions of graded $\mathbb{F}[x]$ -modules. We also talk about the correspondence between the algebraic constructions in $\mathbf{Pers}_{\mathbb{F}}$ and those of $\mathbf{GrMod}_{\mathbb{F}[x]}$ resulting from this isomorphism of categories.

The majority of the definitions and results we present below involving persistence modules are taken or adapted from the following papers, listed in increasing order of their initial publication dates:

1. *The Categorification of Persistence Modules* [BS14] by Bubenik and Scott.
2. *The Structure and Stability of Persistence Modules* [CSGO13] by Chazal, de Silva, Glisse, and Oudot.
3. *The Observable Structure of Persistence Modules* [CCS14] by Chazal, Crawley-Boevey, and de Silva.
4. *Homological Algebra for Persistence Modules* [BM21] by Bubenik and Milićević.

Note that most of the concepts in this chapter are discussed in terms of category theory. Some introductory category theory definitions and results are presented in Appendix A4. For a more detailed treatment of category theory, we recommend reading *Category Theory in Context* by Emily Riehl [Rie16] and *Introduction to Homological Algebra* by Joseph Rotman [Rot08].

Section 2.1. Persistence Modules as Functors

We start this section by providing a functor definition for persistence modules, adapted from [CCS14, Section 1.3]. For reference, $\mathbf{Poset}(\mathbb{N}_0, \leq)$ refers to the category induced by the partially ordered set (poset) of nonnegative integers $\mathbb{N}_0 = \{n \in \mathbb{Z} : n \geq 0\}$ under the \leq relation (see Definition A4.3) and $\mathbf{Vect}_{\mathbb{F}}$ refers to the category of vector spaces over a scalar field \mathbb{F} and linear maps.

Definition 2.1.1. A **persistence module** $(V_{\bullet}, \alpha_{\bullet})$ over a field \mathbb{F} is a functor of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$. For convenience, we may write V_{\bullet} to refer to $(V_{\bullet}, \alpha_{\bullet})$. When the field \mathbb{F} is arbitrary or unambiguous, we may refer to V_{\bullet} as a **persistence module**.

We identify the following terminology for certain features of persistence modules:

- i. For each $t \in \mathbb{N}_0$, define $V_t := (V_{\bullet}, \alpha_{\bullet})(t)$, i.e. V_t is the vector space over \mathbb{F} obtained by evaluating the functor $(V_{\bullet}, \alpha_{\bullet})$ on the object t of $\mathbf{Poset}(\mathbb{N}_0, \leq)$. The vector spaces of $(V_{\bullet}, \alpha_{\bullet})$ generally refer to the collection $\{V_t : t \in \mathbb{N}_0\}$ of vector spaces.
- ii. A **structure map** of $(V_{\bullet}, \alpha_{\bullet})$ refers to a linear map $\alpha_{s,t} : V_t \rightarrow V_s$ by $\alpha_{s,t} := (V_{\bullet}, \alpha_{\bullet})(t \rightarrow s)$ for $t, s \in \mathbb{N}_0$ with $t \leq s$, i.e. $\alpha_{s,t}$ is obtained by evaluating the functor $(V_{\bullet}, \alpha_{\bullet})$ on the morphism $t \rightarrow s$ of $\mathbf{Poset}(\mathbb{N}_0, \leq)$. For brevity, we may write $\alpha_t : V_t \rightarrow V_{t+1}$ to refer to $\alpha_t := \alpha_{t+1,t} = V_{\bullet}(t \rightarrow t+1)$.
- iii. We call $\mathbf{Poset}(\mathbb{N}_0, \leq)$ the **indexing category** of V_{\bullet} and \mathbb{N}_0 the **indexing set** of V_{\bullet} . We may also refer to the index $t \in \mathbb{N}_0$ as the **parameter** or **scale** of the vector space V_t in the persistence module V_{\bullet} .

Remarks. (1) The location of the bullet (\bullet) in $(V_{\bullet}, \alpha_{\bullet})$ determine the location of the indices of the vector spaces and structure maps in the notation, e.g. if $(W^{\bullet}, \gamma^{\bullet})$ were a persistence module, we denote its vector spaces as W^t and its structure maps by $\gamma^{s,t} : W^t \rightarrow W^s$. Unlike in the case of chain complexes and cochain complexes, we use the same definition for persistence modules regardless of the location of the bullet (\bullet) (i.e. as a subscript or superscript).

(2) In this expository paper, we use an asterisk ($*$) as the “placeholder” for the index $n \in \mathbb{Z}$ for a chain complex $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$, following [Rot88], and we reserve the use of bullets (\bullet) for persistence modules. This distinction will be important after Definition 2.2.10, where we introduce chain complexes of persistence modules.

(3) For the case of structure maps, the indices $t, s \in \mathbb{N}_0$ of $\alpha_{s,t}$ are written in right-to-left order, following the notation for function composition, so that compositions of structure maps are written like $\alpha_{s,r} \circ \alpha_{r,t} = \alpha_{s,t}$ (with equality given later by Lemma 2.1.3). Note that, in this paper, when we say $\alpha_{s,t}$ is a structure map, then it is implied that $t, s \in \mathbb{N}_0$ and $t \leq s$.

We would like to point out that the definition given in [CCS14, Section 1.3] applies to a more general family of persistence modules since it allows for different indexing categories for the domain category $\mathbf{Poset}(I, \leq)$ and different categories such as \mathbf{Mod}_R for the codomain category. That is, a persistence module V_{\bullet} is defined as a functor $V_{\bullet} : \mathbf{Poset}(I, \leq) \rightarrow \mathbf{Mod}_R$ where (I, \leq) is some partially ordered set and R is some (commutative) ring. Observe that if R is a field, then an R -module is an R -vector space by definition, i.e. $\mathbf{Mod}_{\mathbb{F}}$ and $\mathbf{Vect}_{\mathbb{F}}$ are the same category. We bring attention to this since homology is typically introduced using \mathbb{Z} coefficients. Therefore, it may be natural to consider functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Mod}_{\mathbb{Z}}$, so that we can consider a persistence module where each index in \mathbb{N}_0 corresponds to the n^{th} homology group with coefficients in \mathbb{Z} of some space for some fixed $n \in \mathbb{N}_0$. In this paper, we restrict persistence modules to codomain category $\mathbf{Vect}_{\mathbb{F}}$ since $\mathbb{Z}[x]$ is not a principal ideal domain (PID) but $\mathbb{F}[x]$ for any field \mathbb{F} is, which allows us to use the Graded Structure Theorem (Theorem 4.3.1) later in Section 2.5.

We would also like to emphasize that since we introduce persistence theory to discuss persistent homology more concretely, we will primary use $\mathbb{F} = \mathbb{Q}$ and $\mathbb{F} = \mathbb{Z}_p$ for our examples. We discuss this choice in more detail

later in Remark 3.2.4 under Section 3.2. Below, we provide an example of a persistence module over \mathbb{Q} .

Example 2.1.2. Let $(A_\bullet, \alpha_\bullet)$ be a persistence module over \mathbb{Q} with the vector spaces A_t given as follows:

$$A_t = \begin{cases} \mathbb{Q}\langle a_1, a_2 \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle b \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle c_1, c_2, c_3 \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle d \rangle & \text{if } t \geq 3 \end{cases}$$

and the structure maps $\alpha_{s,t} : A_t \rightarrow A_s$ of A_\bullet are defined as follows:

$$\begin{array}{lll} \alpha_{1,0} : A_0 \rightarrow A_1 & \alpha_{2,1} : A_1 \rightarrow A_2 & \alpha_{t,2} : A_2 \rightarrow A_3 \text{ with } t \geq 3 \\ a_1 \mapsto 2b & b \mapsto 3c_1 + 2c_2 + c_3 & c_1 \mapsto d \\ a_2 \mapsto -3b & & c_2 \mapsto 0 \\ & & c_3 \mapsto 2d \\ \\ \alpha_{2,0} : A_0 \rightarrow A_2 & \alpha_{t,1} : A_1 \rightarrow A_t \text{ with } t \geq 3 & \alpha_{t,0} : A_0 \rightarrow A_t \text{ with } t \geq 3 \\ a_1 \mapsto 6c_1 + 4c_2 + 2c_3 & b \mapsto 5d & a_1 \mapsto 10d \\ a_2 \mapsto -9c_1 - 6c_2 - 3c_3 & & a_2 \mapsto -15d \end{array}$$

For all $t \in \mathbb{N}_0$, the structure map $\alpha_{t,t} : A_t \rightarrow A_t$ is the identity map on A_t . For all $t, s \geq 3$, the structure map $\alpha_{s,t} : A_t \rightarrow A_s$ is the identity map on $A_t = \mathbb{Q}\langle d \rangle$, i.e. $d \mapsto d$.

Observe that the specification for the structure maps above satisfy the functorial property of composition. By definition of $\mathbf{Poset}(\mathbb{N}_0, \leq)$ (as given in Definition A4.3), there is exactly one morphism $t \rightarrow s$ for any $t, s \in \mathbb{N}_0$ with $t \leq s$. Therefore, the composition $(t \rightarrow r) \circ (r \rightarrow s)$ of morphisms in $\mathbf{Poset}(\mathbb{N}_0, \leq)$ with $t \leq r \leq s$ must be the morphism $t \rightarrow s$. We talk about this in more detail below in Lemma 2.1.3.

For example, any composition of structure maps resulting in a domain of $A_0 = \mathbb{Q}\langle a_1, a_2 \rangle$ and a codomain of $A_3 = \mathbb{Q}\langle d \rangle$ must exactly be the structure map $\alpha_{3,0} : A_0 \rightarrow A_3$ (i.e. domain, codomain, and assignments/evaluation must be the same). We list some of these compositions below, evaluated on $2a_1 \in A_0$.

$$\begin{array}{ll} \alpha_{3,0}(2a_1) = 2(10d) & = 20d \\ (\alpha_{3,2} \circ \alpha_{2,0})(2a_1) = \alpha_{3,2}(12c_1 + 8c_2 + 4c_3) = 12d + 8d & = 20d \\ (\alpha_{3,1} \circ \alpha_{1,0})(2a_1) = \alpha_{3,1}(4b) & = 20d \\ (\alpha_{3,2} \circ \alpha_{2,1} \circ \alpha_{1,0})(2a_1) = (\alpha_{3,2} \circ \alpha_{2,1})(2b) = \alpha_{3,2}(6c_1 + 4c_2 + c_3) = 12d + 8d & = 20d \\ (\alpha_{3,0} \circ \alpha_{0,0})(2a_1) = \alpha_{3,0}(2a_1) & = 20d \\ (\alpha_{3,2} \circ \alpha_{2,2} \circ \alpha_{2,0})(2a_1) = (\alpha_{3,2} \circ \alpha_{2,2})(12c_1 + 8c_2 + 4c_3) = \alpha_{3,2}(12c_1 + 8c_2 + 4c_3) = 12d + 8d & = 20d \end{array}$$

Defining persistence modules as functors is a succinct way to impose additional algebraic structure to the construction. As mentioned in the introduction, [ZC05, Definition 3.2] defines a persistence module as a collection of R -modules $\{V^t\}$ and homomorphisms $\{\varphi^t : V_t \rightarrow V_{t+1}\}$, both indexed by $t \in \mathbb{N}_0$. Observe that the collection definition does not explicitly state, for example, that there exists a linear map $V_2 \rightarrow V_4$. While it may be somewhat natural to assume the linear map $\alpha_{4,2} : V_2 \rightarrow V_4$ is given by the composition $\alpha_{4,2} = \alpha_3 \circ \alpha_2$, having to state it as a separate condition can be somewhat confusing or cumbersome and can also feel arbitrary. In comparison, by assuming that said collection corresponds to a functor, we get a number of properties as a consequence of the functor definition. We list some below.

Lemma 2.1.3. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module.

- i. For all $t \in \mathbb{N}_0$, the structure map $\alpha_{t,t} : V_t \rightarrow V_t$ is the identity map on V_t .
- ii. For all $t, r, s \in \mathbb{N}_0$ with $t \leq r \leq s$, $\alpha_{s,t} = \alpha_{s,r} \circ \alpha_{r,t}$. That is, the following diagram commutes:

$$\begin{array}{ccc} V_t & \xrightarrow{\alpha_{s,t}} & V_s \\ & \searrow \alpha_{r,t} \quad \# \quad \nearrow \alpha_{s,r} & \\ & V_r & \end{array}$$

- iii. For all $t, s \in \mathbb{N}_0$ with $t \leq s$, there is exactly one linear map $V_t \rightarrow V_s$.

Proof. Let $n \in \mathbb{N}_0$. By definition of $\mathbf{Poset}(\mathbb{N}_0, \leq)$, the unique morphism $n \rightarrow n$ must be the identity morphism on n . By definition, a functor must pass $V_\bullet(n \rightarrow n)$ to the identity map on V_n . Therefore, the structure map $V_\bullet(n \rightarrow n) = \alpha_{n,n} = \text{id}_{V_n}$.

Let $t, s, r \in \mathbb{N}_0$ such that $t \leq r \leq s$. By definition of $\mathbf{Poset}(\mathbb{N}_0, \leq)$, we have the equality $(s \leftarrow r) \circ (r \leftarrow t) = (s \leftarrow t)$ as morphisms of $\mathbf{Poset}(\mathbb{N}_0, \leq)$. Since functors must respect composition (indicated below by \star), we have

$$\alpha_{s,t} = V_\bullet(s \leftarrow t) = V_\bullet((s \leftarrow r) \circ (r \leftarrow t)) \stackrel{\star}{=} V_\bullet(s \leftarrow r) \circ V_\bullet(r \leftarrow t) = \alpha_{s,r} \circ \alpha_{r,t}$$

Claim (iii) is a re-statement of Claim (ii), wherein all linear maps with domain V_t and domain V_s must be equal. \blacksquare

Note that Lemma 2.1.3 allows us to represent persistence modules as sequences and use sequences to refer to a persistence module unambiguously. In particular, we can describe a persistence module V_\bullet as the following sequence:

$$V_0 \xrightarrow{\alpha_0} V_1 \xrightarrow{\alpha_1} V_2 \xrightarrow{\alpha_2} V_3 \xrightarrow{\alpha_3} V_4 \xrightarrow{\alpha_4} \dots$$

with α_t being the linear map $\alpha_t : V_t \rightarrow V_{t+1}$ for all $t \in \mathbb{N}_0$. As a sidenote, this sequence representation applies to any diagram with $\mathbf{Poset}(\mathbb{N}_0, \leq)$ as the domain category. Note that we use the term *diagram* to refer to functors of the form $\mathbf{Poset}(I, \leq) \rightarrow \mathbf{C}$, as described in Definition A4.6.

Conversely, Lemma 2.1.3 also implies that there is enough information in the sequence (i.e. the collection of vector spaces $\{V_n\}$ and morphisms $\{\alpha_t : V_t \rightarrow V_{t+1}\}$) to determine a persistence module V_\bullet . In particular, for a given persistence module V_\bullet , the structure maps of the form $\alpha_{s,t} : V_t \rightarrow V_s$ with $s \neq t + 1$ are uniquely determined by the set of $\{\alpha_t\}_{t \in \mathbb{N}_0}$ of linear maps. Therefore, the collection definition given in [ZC05, Definition 3.2] is compatible with the functor definition given in Definition 2.1.1. We state this as a lemma below.

Proposition 2.1.4. Let $\{W_t\}_{t \in \mathbb{N}_0}$ and $\{\varphi : W_t \rightarrow W_{t+1}\}_{t \in \mathbb{N}_0}$ be \mathbb{N}_0 -indexed collections of \mathbb{F} -vector spaces and linear maps respectively. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module over \mathbb{F} such that $V_t = W_t$ for all $t \in \mathbb{N}_0$ and the structure maps of $(V_\bullet, \alpha_\bullet)$ of the form $\alpha_t : V_t \rightarrow V_{t+1}$ are given by $\alpha_t = \varphi_t$. Then, there is exactly one choice for the remaining structure maps of $(V_\bullet, \alpha_\bullet)$, as listed below:

- i. For all $t \in \mathbb{N}_0$, the structure map $\alpha_{t,t} : V_t \rightarrow V_t$ must be the identity map $\text{id}_{V_t} : V_t \rightarrow V_t$ of V_t .
- ii. For all $t \in \mathbb{N}_0$ and $k \geq 2$, the structure map $\alpha_{t+k,t} : V_t \rightarrow V_{t+k}$ is given by the following composition:

$$\alpha_{s,t} = \alpha_{s-1} \circ \alpha_{s-2} \circ \dots \circ \alpha_{t+1} \circ \alpha_t$$

That is, the pair of collections $\{W_t\}$ and $\{\varphi_t\}$ unambiguously determines a persistence module $(W_\bullet, \varphi_\bullet)$.

Proof. This follows immediately from the definition of $\mathbf{Poset}(\mathbb{N}_0, \leq)$ and that of functors. The proof is

similar to that of Lemma 2.1.3. ■

Note that by Lemma 2.1.3 and Proposition 2.1.4, we get a correspondence between the set of persistence modules $(V_\bullet, \alpha_\bullet)$ and the set of pairs of collections $(\{V_t\}, \{\alpha_t : V_t \rightarrow V_{t+1}\})$. For convenience, we often define persistence modules using Proposition 2.1.4. As an example, we describe the persistence module $(A_\bullet, \alpha_\bullet)$ given in Example 2.1.2 using a collection of vector spaces and of linear maps. We also describe the sequence corresponding to $(A_\bullet, \alpha_\bullet)$.

Example 2.1.5. Let $(B_\bullet, \beta_\bullet)$ be a persistence module over \mathbb{Q} with vector spaces B_t given as follows:

$$B_t = \begin{cases} \mathbb{Q}\langle a_1, a_2 \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle b \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle c_1, c_2, c_3 \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle d \rangle & \text{if } t \geq 3 \end{cases}$$

Let the structure maps $\beta_t : B_t \rightarrow B_{t+1}$ of B_\bullet be as given below:

$$\begin{array}{llll} \beta_0 : B_0 \rightarrow B_1 & \beta_1 : B_1 \rightarrow B_2 & \beta_2 : B_2 \rightarrow B_3 & \beta_t : B_t \rightarrow B_{t+1} \text{ for } t \geq 3 \\ a_1 \mapsto 2b & b \mapsto 3c_1 + 2c_2 + c_3 & c_1 \mapsto d & d \mapsto d \\ a_2 \mapsto -3b & & c_2 \mapsto 0 & \\ & & c_3 \mapsto 2d & \end{array}$$

Use Proposition 2.1.4 to define the remaining structure maps of B_\bullet . Then, $B_\bullet = A_\bullet$ (as functors) where A_\bullet is the persistence module defined in Example 2.1.2. Observe that the specification for B_\bullet is much shorter than that for A_\bullet . Also, A_\bullet corresponds to the following sequence:

$$\begin{array}{ccccccc} A_0 & & A_1 & & A_2 & & A_3 & & A_4 \\ \parallel & & \parallel & & \parallel & & \parallel & & \parallel \\ \mathbb{Q}\langle a_1, a_2 \rangle & \xrightarrow{\alpha_0} & \mathbb{Q}\langle b \rangle & \xrightarrow{\alpha_1} & \mathbb{Q}\langle c_1, c_2, c_3 \rangle & \xrightarrow{\alpha_2} & \mathbb{Q}\langle d \rangle & \xrightarrow{\alpha_3} & \mathbb{Q}\langle d \rangle \xrightarrow{\alpha_4} \dots \end{array}$$

Observe that Definition 2.1.1 allows for persistence modules that have non-finite characteristics. For example, some of the vector spaces V_t of a persistence module V_\bullet may be infinite-dimensional. We can also have a persistence module $(W_\bullet, \gamma_\bullet)$ wherein none of the structure maps $\gamma_t : W_t \rightarrow W_{t+1}$ are isomorphisms or can be made into isomorphisms by a change of basis on W_t . In this case, we are considering an infinite number of vector spaces. Persistence modules of this nature do present some problems. Indeed, the results presented in [ZC05] only consider a specific type of persistence module, one that is finite in specific ways. We use a definition of such, adapted from [ZC05, Definition 3.3].

Definition 2.1.6. A persistence module $(V_\bullet, \alpha_\bullet)$ is **constant on an interval** $I \subseteq \mathbb{N}_0$ if for all $t, s \in I$ with $t \leq s$, the structure map $\alpha_{s,t} : V_t \rightarrow V_s$ is a vector space isomorphism. A persistence module V_\bullet is called **finite-type** if all of its vector spaces V_t are finite-dimensional and V_\bullet is constant on $[N, \infty)$ for some $N \in \mathbb{N}_0$.

Remark. Since persistence theory is a relatively new field, there exist terms that are commonly used in the literature but are defined differently depending on the author(s). The term *finite-type* is one of these.

Later in Proposition 2.5.14, using the definition above for *finite-type*, we state that interval decompositions (defined in Definition 2.3.4) exist for finite-type persistence modules. In contrast, [BS14, Definition 4.1] and [CSGO13] define the term *finite-type* to refer to persistence modules

for which an interval decomposition exists. The term *tame* in [BS14] serves the same function as *finite-type* in [ZC05], i.e. it serves as a characterization of persistence modules that have interval decompositions. However, the term *tame* also has its problems, as discussed in [CSGO13, p5].

All of the examples we will present in this expository involve finite-type persistence modules (not including the one example given below for comparison). As we will discuss in Section 3.2, under the assumption that the simplicial complex in question is finite, the persistence modules generated by said simplicial complex will necessarily be of finite-type. This assumption is justified since most practical applications of persistent homology, particularly those involving calculation, have simplicial complexes generated from finite datasets. The usual construction involves letting the dataset V be the vertex set of the to-be constructed simplicial complex K . Since K is a subset of the power set 2^V and power sets of finite sets are finite, K must also be finite. This explains why some of the introductory literature on persistent homology describe persistence modules as finite sequences, i.e. persistence modules are described to be finite sequences as given below:

$$V_0 \xrightarrow{\alpha_0} V_1 \xrightarrow{\alpha_1} V_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_{N-1}} V_N$$

Using our definition, we would interpret the sequence above to correspond to a persistence module V_\bullet that is constant on $[N, \infty)$ and assume that $V_n = V_N$ for all $n \in [N+1, \infty)$. As an example, we look at the persistence module given earlier in Example 2.1.2.

Example 2.1.7. Let $(A_\bullet, \alpha_\bullet)$ be as defined in Example 2.1.2. Then, A_\bullet is a finite-type persistence module and is constant on $[3, \infty)$. Also, A_\bullet corresponds to the following sequence:

$$\begin{array}{ccccccc} & A_0 & & A_1 & & A_2 & & A_3 \\ & \parallel & & \parallel & & \parallel & & \parallel \\ \mathbb{Q}\langle a_1, a_2 \rangle & \xrightarrow{\alpha_0} & \mathbb{Q}\langle b \rangle & \xrightarrow{\alpha_1} & \mathbb{Q}\langle c_1, c_2, c_3 \rangle & \xrightarrow{\alpha_2} & \mathbb{Q}\langle d \rangle \end{array}$$

For comparison, we provide an example below of a persistence module that is not finite-type.

Example 2.1.8. Let $A = \{a_n\}_{n \in \mathbb{N}_0}$ be a set of indeterminates. Define the persistence module $(F_\bullet, \eta_\bullet)$ over \mathbb{Q} using Proposition 2.1.4 as follows:

1. For each $t \in \mathbb{N}_0$, $F_t := \mathbb{Q}\langle a_0, a_1, \dots, a_t \rangle$.
2. For each $t \in \mathbb{N}_0$, define the structure map $\eta_t : F_t \rightarrow F_{t+1}$ by $\eta_t(a_i) = a_i$ for all $i \in \{0, \dots, t\}$.

Then, the following sequence represents the persistence module $(F_\bullet, \eta_\bullet)$:

$$\begin{array}{ccccccc} F_0 & & F_1 & & F_2 & & F_3 \\ \parallel & & \parallel & & \parallel & & \parallel \\ \mathbb{Q}\langle a_0 \rangle & \xrightarrow[\substack{a_0 \mapsto a_0}}{\eta_0} & \mathbb{Q}\langle a_0, a_1 \rangle & \xrightarrow[\substack{a_i \mapsto a_i, i=0,1}}{\eta_1} & \mathbb{Q}\langle a_0, a_1, a_2 \rangle & \xrightarrow[\substack{a_i \mapsto a_i, i=0,1,2}}{\eta_2} & \mathbb{Q}\langle a_0, a_1, a_2, a_3 \rangle \xrightarrow{\eta_3} \cdots \end{array}$$

While all vector spaces F_t are finite-dimensional, V_\bullet fails to be constant on the interval $[N, \infty)$ for any $N \in \mathbb{N}_0$. We identify two ways we can see this:

1. For all $t \in \mathbb{N}_0$, $\eta_t : F_t \rightarrow F_{t+1}$ cannot be an isomorphism since $\eta_t^{-1}(\{a_{t+1}\}) = \emptyset$, i.e. η_t is not surjective.
2. Any two vector spaces F_t and F_s of V_\bullet with $t \neq s$ cannot be isomorphic since $\dim(F_t) = t+1 \neq s+1 = \dim(F_s)$.

Therefore, F_\bullet is **not** a finite-type persistence module.

Section 2.2. The Category of Persistence Modules

One of the key goals of persistence theory is to characterize persistence modules that admit a special decomposition called an interval decomposition. Note that we use the term *(direct sum) decomposition* to refer to an isomorphism to a direct sum, similar to the case of vector spaces, modules, and of chain complexes. However, to understand decompositions, we first need to define the notions of isomorphisms and of direct sums between persistence modules.

In persistence theory, these algebraic constructions are defined by forming the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules over \mathbb{F} as a *functor category* with *natural transformations* as morphisms. Below, we provide definitions specific for the category of persistence modules, adapted from [CCS14, Definition 1.3].

Definition 2.2.1. The category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules over a field \mathbb{F} is the category of functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$ and natural transformations, i.e. $\mathbf{Pers}_{\mathbb{F}}$ consists of the following:

- i. The objects in $\mathbf{Pers}_{\mathbb{F}}$ are persistence modules over \mathbb{F} , as given in Definition 2.1.1.
- ii. The morphisms in $\mathbf{Pers}_{\mathbb{F}}$ are persistence morphisms, defined as follows:

A **persistence morphism** $\varphi_{\bullet} : (V_{\bullet}, \alpha_{\bullet}) \rightarrow (W_{\bullet}, \gamma_{\bullet})$ between persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ over \mathbb{F} is a collection of linear maps $\varphi_{\bullet} = (\varphi_t : V_t \rightarrow W_t)_{t \in \mathbb{N}_0}$ such that for all $t, s \in \mathbb{N}_0$ with $t \leq s$, the composition relation $\gamma_{s,t} \circ \varphi_t = \varphi_s \circ \alpha_{s,t}$ is satisfied, i.e. the following diagram commutes:

$$\begin{array}{ccc} V_t & \xrightarrow{\alpha_{s,t}} & V_s \\ \varphi_t \downarrow & & \downarrow \varphi_s \\ W_t & \xrightarrow{\gamma_{s,t}} & W_s \end{array}$$

Remarks. (1) In this paper, writing the symbol φ_{\bullet} to represent a persistence morphism determines that φ_t refers to the corresponding linear map $\varphi_t : V_t \rightarrow W_t$ for each $t \in \mathbb{N}_0$, i.e. the bullet (\bullet) in φ_{\bullet} is replaced by the index $t \in \mathbb{N}_0$, much like in the case of persistence modules as remarked under Definition 2.1.1.

- (2) By [Rie16, Section 1.7, p44] and [Rot08, Example 1.19(i), p27], for any two categories \mathbf{C} and \mathbf{D} , there exists a category denoted $\mathbf{D}^{\mathbf{C}}$ consisting of functors of the form $\mathbf{C} \rightarrow \mathbf{D}$ as objects and natural transformations between functors. This category is called a *functor category*. We use this result for the claim of $\mathbf{Pers}_{\mathbb{F}}$ being a well-defined category.

As a sidenote, [BS14] denotes $\mathbf{Pers}_{\mathbb{F}}$ as $\mathbf{Vect}_{\mathbb{F}}^{(\mathbb{N}_0, \leq)}$ where (\mathbb{N}_0, \leq) refers to the poset category $\mathbf{Poset}(\mathbb{N}_0, \leq)$ (see remarks under Definition A4.3 for the notation of poset categories).

We provide an example of a persistence morphism below. Note that not all collections of linear maps $\{\varphi_t : V_t \rightarrow W_t\}_{t \in \mathbb{N}_0}$ produce a persistence morphism $\varphi_{\bullet} : V_{\bullet} \rightarrow W_{\bullet}$. We need to check that the composition relation, i.e. the commuting squares and diagrams, is satisfied.

Example 2.2.2. Let $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ be persistence modules over \mathbb{Q} with the following \mathbb{Q} -vector spaces defined over indeterminates A, B, C for V_t and a, b, c for W_t :

$$V_t = \begin{cases} 0 & \text{if } t = 0 \\ \mathbb{Q}\langle A \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle A, B \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle A, B, C \rangle & \text{if } t \geq 3 \end{cases} \quad \text{and} \quad W_t = \begin{cases} \mathbb{Q}\langle a, b \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle a, b, c \rangle & \text{if } t \geq 1 \end{cases}$$

The structure maps $\alpha_t : V_t \rightarrow V_{t+1}$ of V_\bullet and $\gamma_t : W_t \rightarrow W_{t+1}$ of W_\bullet are determined by the identity maps on $\mathbb{Q}\langle A, B, C \rangle$ and on $\mathbb{Q}\langle a, b, c \rangle$ respectively, i.e. $\alpha_t(X) = \text{id}_{\mathbb{Q}\langle A, B, C \rangle}(X)$ for all $X \in V_t$ and $\gamma_t(x) = \text{id}_{\mathbb{Q}\langle a, b, c \rangle}(x)$ for all $x \in W_t$.

Define a family $\{\varphi_t\}_{t \in \mathbb{N}_0}$ of linear maps $\varphi_t : V_t \rightarrow W_t$ as follows: Let $\Phi : \mathbb{Q}\langle A, B, C \rangle \rightarrow \mathbb{Q}\langle a, b, c \rangle$ be given by $A \mapsto b - a$, $B \mapsto a - c$, and $C \mapsto c - b$. Define $\varphi_t : V_t \rightarrow W_t$ by $\varphi_t(X) := \Phi(X)$ for all $X \in V_t$, i.e. we have the following assignments for φ_t :

$$\begin{aligned}\varphi_t(A) &= b - a & \text{for } t \geq 1 \\ \varphi_t(B) &= a - c & \text{for } t \geq 2 \\ \varphi_t(C) &= c - b & \text{for } t \geq 3\end{aligned}$$

Note that $\varphi_0 : V_0 \rightarrow W_0$ is necessarily the trivial map since the domain V_0 is the trivial vector space. We can confirm that the family $\{\varphi_t\}_{t \in \mathbb{N}_0}$ satisfies the commutativity requirement as follows: for all $X \in V_t$ with $t \geq 1$:

$$(\varphi_s \circ \alpha_{s,t})(X) = (\Phi \circ \text{id}_{\mathbb{Q}\langle A, B, C \rangle})(X) = \Phi(X) = (\text{id}_{\mathbb{Q}\langle a, b, c \rangle} \circ \Phi)(X) = (\gamma_{s,t} \circ \varphi_t)(X)$$

This is illustrated in the diagram below:

$$\begin{array}{ccccccc} (V_\bullet, \alpha_\bullet) : & \begin{array}{c} (t=0) \\ 0 \end{array} & \xrightarrow{\alpha_0} & \begin{array}{c} (t=1) \\ \mathbb{Q}\langle A \rangle \end{array} & \xrightarrow{\alpha_1} & \begin{array}{c} (t=2) \\ \mathbb{Q}\langle A, B \rangle \end{array} & \xrightarrow{\alpha_2} & \begin{array}{c} (t=3) \\ \mathbb{Q}\langle A, B, C \rangle \end{array} & \xrightarrow{\alpha_3} & \dots \\ & \downarrow \varphi_0 & & \downarrow \varphi_1 & & \downarrow \varphi_2 & & \downarrow \varphi_3 & & \\ (W_\bullet, \gamma_\bullet) : & \mathbb{Q}\langle a, b \rangle & \xrightarrow{\gamma_0} & \mathbb{Q}\langle a, b, c \rangle & \xrightarrow{\gamma_1} & \mathbb{Q}\langle a, b, c \rangle & \xrightarrow{\gamma_2} & \mathbb{Q}\langle a, b, c \rangle & \xrightarrow{\gamma_3} & \dots \end{array}$$

$\varphi_1 : A \mapsto b - a$
 $\varphi_2 : \begin{array}{l} A \mapsto b - a \\ B \mapsto a - c \end{array}$
 $\varphi_3 : \begin{array}{l} A \mapsto b - a \\ B \mapsto a - c \\ C \mapsto c - b \end{array}$

Therefore, the set $\{\varphi_t\}_{t \in \mathbb{N}_0}$ determines a persistence morphism $\varphi_\bullet : (V_\bullet, \alpha_\bullet) \rightarrow (W_\bullet, \gamma_\bullet)$.

Observe that, given two persistence morphisms $\varphi_\bullet : (A_\bullet, \alpha_\bullet) \rightarrow (B_\bullet, \beta_\bullet)$ and $\psi_\bullet : (B_\bullet, \beta_\bullet) \rightarrow (C_\bullet, \gamma_\bullet)$ between persistence modules $(A_\bullet, \alpha_\bullet)$, $(B_\bullet, \beta_\bullet)$, and $(C_\bullet, \gamma_\bullet)$, the collection of compositions $\{\psi_t \circ \varphi_t\}_{t \in \mathbb{N}_0}$ also determines a persistence morphism $\psi_\bullet \circ \varphi_\bullet : A_\bullet \rightarrow C_\bullet$. Relative to the commuting squares, the squares of φ_\bullet and ψ_\bullet together produce the squares for the composition, as illustrated below:

$$\begin{array}{ccc} \left(\begin{array}{c} \text{commuting square of} \\ \varphi_\bullet : (A_\bullet, \alpha_\bullet) \rightarrow (B_\bullet, \beta_\bullet) \end{array} \right) & & \left(\begin{array}{c} \text{commuting square of} \\ \psi_\bullet : (B_\bullet, \beta_\bullet) \rightarrow (C_\bullet, \gamma_\bullet) \end{array} \right) \\ \begin{array}{ccc} A_t & \xrightarrow{\alpha_{s,t}} & A_s \\ \varphi_t \downarrow & \# & \downarrow \varphi_s \\ B_t & \xrightarrow{\beta_{s,t}} & B_s \end{array} & \text{with} & \begin{array}{ccc} B_t & \xrightarrow{\beta_{s,t}} & B_s \\ \psi_t \downarrow & \# & \downarrow \psi_s \\ C_t & \xrightarrow{\gamma_{s,t}} & C_s \end{array} \end{array} \quad \text{implies} \quad \begin{array}{ccc} \left(\begin{array}{c} \text{commuting square for} \\ \text{the composition } \psi_\bullet \circ \varphi_\bullet \end{array} \right) \\ \begin{array}{ccc} A_t & \xrightarrow{\alpha_{s,t}} & A_s \\ \psi_t \circ \varphi_t \downarrow & \# & \downarrow \psi_s \circ \varphi_s \\ C_t & \xrightarrow{\gamma_{s,t}} & C_s \end{array} \end{array}$$

Therefore, the composition $\psi_\bullet \circ \varphi_\bullet : (A_\bullet, \alpha_\bullet) \rightarrow (C_\bullet, \gamma_\bullet)$ of persistence morphisms is also a persistence morphism. Note that this is part of the proof of **Pers** _{\mathbb{F}} being a well-defined category.

Since the category **Pers** _{\mathbb{F}} of persistence modules is defined to be a functor category of functors of the form **Poset**(\mathbb{N}_0, \leq) \rightarrow **Vect** _{\mathbb{F}} , there is a natural extension of the algebraic constructions in **Vect** _{\mathbb{F}} to the case of **Pers** _{\mathbb{F}} . In particular, we can define constructions in **Pers** _{\mathbb{F}} by pointwise evaluation for each $t \in \mathbb{N}_0$ using the corresponding constructions in **Vect** _{\mathbb{F}} . Note that we need to be careful here since the constructions in **Pers** _{\mathbb{F}} have to respect the structure of the morphisms. We discuss this in more detail for the following constructions in **Pers** _{\mathbb{F}} : isomorphism relations, direct sum operations, subobject relations, kernel, image, and cokernel of morphisms, and chain complexes.

We want to point out that this extension process (for lack of a better term) is not unique to persistence

modules. More generally, this seems to be a common property of functor categories with functors of the form $\mathbf{Poset}(I, \leq) \rightarrow \mathbf{A}$ where \mathbf{A} is some *abelian* category. Note that \mathbf{Mod}_R over a PID R and $\mathbf{Vect}_{\mathbb{F}}$ over a field \mathbb{F} are examples of abelian categories. Abelian categories are outside the scope of this paper but for those interested, this is discussed in more detail in [Rot08, Section 5.5: Proposition 5.93 and Corollary 5.94] and in [Wei95, Appendix A.4]. This might explain why most of the literature we have read on persistence theory do not explicitly describe these constructions, e.g. [CCS14; BM21] wherein a reference to $\mathbf{Pers}_{\mathbb{F}}$ being an *abelian* or *Grothendieck* category seems to suffice.

Note that an equality between persistence modules in $\mathbf{Pers}_{\mathbb{F}}$ corresponds to an equality between functors. That is, two persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ are equal if $V_t = W_t$ as \mathbb{F} -vector spaces for all $t \in \mathbb{N}_0$ and $\alpha_{s,t} = \gamma_{s,t}$ as linear maps for all $t, s \in \mathbb{N}_0$ with $t \leq s$. We start with a definition for isomorphisms between persistence modules.

Definition 2.2.3. A **persistence isomorphism** $\varphi_{\bullet} : V_{\bullet} \rightarrow W_{\bullet}$ between persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ over \mathbb{F} is a persistence morphism $\varphi_{\bullet} : V_{\bullet} \rightarrow W_{\bullet}$ with $\varphi_{\bullet} = (\varphi_t : V_t \rightarrow W_t)_{t \in \mathbb{N}_0}$ such that for all $t \in \mathbb{N}_0$, the linear map $\varphi_t : V_t \rightarrow W_t$ is an \mathbb{F} -vector space isomorphism. If such a persistence isomorphism exists, we say that V_{\bullet} and W_{\bullet} are **isomorphic** (as persistence modules) and write $V_{\bullet} \cong W_{\bullet}$.

Remark. If needed, we may write $\overset{\text{Pers}}{\cong}$ instead of \cong to emphasize that the isomorphism relation is of the category $\mathbf{Pers}_{\mathbb{F}}$. We use this notation sometimes in Chapter 4 where we talk about isomorphisms between R -modules, between graded $\mathbb{F}[x]$ -modules, and between persistence modules in the same context.

We want to emphasize that a persistence isomorphism has to be given by a persistence morphism and that the structure maps of the persistence modules cannot be ignored. That is, given two persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$, a collection $\{\varphi_t\}_{t \in \mathbb{N}_0}$ of \mathbb{F} -vector space isomorphisms $\varphi_t : V_t \rightarrow W_t$ generally does not make a persistence isomorphism. Below, we provide an example of a collection $\{\varphi_t\}_{t \in \mathbb{N}_0}$ of vector spaces isomorphisms that is not a persistence morphism, and another collection $\{\psi_t\}_{t \in \mathbb{N}_0}$ that is a persistence isomorphism.

Example 2.2.4. Define the persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ over \mathbb{Q} as follows: For all $t \in \mathbb{N}_0$, define $V_t := \mathbb{Q}\langle a, b \rangle$ and $W_t := \mathbb{Q}\langle x, y \rangle$ with indeterminates a, b, x , and y . For $t = 0$, define the structure map $\alpha_0 : V_0 \rightarrow V_1$ by $\alpha_0(a) := a + b$ and $\alpha_0(b) := b$. For $t \geq 1$, let $\alpha_t : V_t \rightarrow V_{t+1}$ be the identity map on $\mathbb{Q}\langle a, b \rangle$. For all $t \in \mathbb{N}_0$, let $\gamma_t : W_t \rightarrow W_{t+1}$ be the identity map on $\mathbb{Q}\langle x, y \rangle$.

Part (a). For each $t \in \mathbb{N}_0$, let $\varphi_t : V_t \rightarrow W_t$ be given by $a \mapsto x$ and $b \mapsto y$. Since $V_t = \mathbb{Q}\langle a, b \rangle$ and $W_t = \mathbb{Q}\langle x, y \rangle$ for all $t \in \mathbb{N}_0$, each φ_t is a vector space isomorphism with obvious inverse. However, the collection $\{\varphi_t\}_{t \in \mathbb{N}_0}$ is **not** a persistence morphism since

$$x + y = (\varphi_1 \circ \alpha_0)(a) = (\varphi_1 \circ \alpha_{1,0})(a) \neq (\gamma_{1,0} \circ \varphi_0)(a) = (\gamma_0 \circ \varphi_0)(a) = x$$

That is, the following diagram does not commute:

$$\begin{array}{ccc} V_0 = \mathbb{Q}\langle a, b \rangle & \xrightarrow[\substack{a \mapsto a+b, b \mapsto b}]{\alpha_{1,0}} & \mathbb{Q}\langle a, b \rangle = V_1 \\ \substack{a \mapsto x \\ b \mapsto y} \downarrow \varphi_0 & & \varphi_1 \downarrow \substack{a \mapsto x \\ b \mapsto y} \\ W_0 = \mathbb{Q}\langle x, y \rangle & \xrightarrow[\substack{x \mapsto x, y \mapsto y}]{\gamma_{1,0}} & \mathbb{Q}\langle x, y \rangle = W_1 \end{array}$$

Therefore, $\{\varphi_t\}_{t \in \mathbb{N}_0}$ is **not** a persistence isomorphism.

Part (b). It turns out that there exists a persistence isomorphism between $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$. This becomes clearer if we perform a change of basis on V_t and W_t for $t \in \mathbb{N}_0$. Observe that $\mathbb{Q}\langle a, b \rangle = \mathbb{Q}\langle c, b \rangle$ with $c := a + b$, i.e. $\{c, b\}$ is a basis of $\mathbb{Q}\langle a, b \rangle$. Then, we can express the persistence module

$(V_\bullet, \alpha_\bullet)$ as follows. Note that $\alpha_t = \text{id}_{\mathbb{Q}\langle a, b \rangle}$ for $t \geq 1$.

$$(V_\bullet, \alpha_\bullet) : \quad \begin{array}{ccccccc} (t=0) & & (t=1) & & (t=2) & & (t=3) \\ \mathbb{Q}\langle a, b \rangle & \xrightarrow[\substack{a \mapsto a+b=:c \\ b \mapsto b}}{\alpha_0} & \mathbb{Q}\langle c, b \rangle & \xrightarrow[\substack{c \mapsto c \\ b \mapsto b}}{\alpha_1} & \mathbb{Q}\langle c, b \rangle & \xrightarrow[\substack{c \mapsto c \\ b \mapsto b}}{\alpha_2} & \mathbb{Q}\langle c, b \rangle \longrightarrow \dots \end{array}$$

Similarly, observe that $\mathbb{Q}\langle x, y \rangle = \mathbb{Q}\langle z, y \rangle$ with $z := x + y$, i.e. $\{z, y\}$ is a basis for $\mathbb{Q}\langle x, y \rangle$. We apply this change of basis on W_t for all $t \in \mathbb{N}_0$. Since $\gamma_t = \text{id}_{\mathbb{Q}\langle x, y \rangle}$ for all $t \in \mathbb{N}_0$, the persistence module $(W_\bullet, \gamma_\bullet)$ can be illustrated as follows:

$$(W_\bullet, \gamma_\bullet) : \quad \begin{array}{ccccccc} (t=0) & & (t=1) & & (t=2) & & (t=3) \\ \mathbb{Q}\langle z, y \rangle & \xrightarrow[\substack{z \mapsto z \\ y \mapsto y}}{\gamma_0} & \mathbb{Q}\langle z, y \rangle & \xrightarrow[\substack{z \mapsto z \\ y \mapsto y}}{\gamma_1} & \mathbb{Q}\langle z, y \rangle & \xrightarrow[\substack{z \mapsto z \\ y \mapsto y}}{\gamma_2} & \mathbb{Q}\langle z, y \rangle \longrightarrow \dots \end{array}$$

For $t = 0$, define $\psi_0 : V_0 \rightarrow W_0$ by $a \mapsto z$ and $b \mapsto y$. For $t \geq 1$, define $\psi_t : V_t \rightarrow W_t$ by $c \mapsto z$ and $b \mapsto y$. We claim that $\psi_\bullet = \{\psi_t\}_{t \in \mathbb{N}_0}$ determines a persistence morphism $\psi_\bullet : (V_\bullet, \alpha_\bullet) \rightarrow (W_\bullet, \gamma_\bullet)$ since these two diagrams commute (with $t \geq 1$ for the diagram on the right hand side):

$$\begin{array}{ccc} V_0 = \mathbb{Q}\langle a, b \rangle & \xrightarrow[\substack{a \mapsto c, b \mapsto b}}{\alpha_{1,0}} & \mathbb{Q}\langle c, b \rangle = V_1 \\ \downarrow \substack{a \mapsto z \\ b \mapsto y} \psi_0 & & \downarrow \substack{c \mapsto z \\ b \mapsto y} \psi_1 \\ W_0 = \mathbb{Q}\langle z, y \rangle & \xrightarrow[\substack{z \mapsto z, y \mapsto y}}{\gamma_{1,0}} & \mathbb{Q}\langle z, y \rangle = W_1 \end{array} \quad \text{and} \quad \begin{array}{ccc} V_t = \mathbb{Q}\langle c, b \rangle & \xrightarrow[\substack{\text{(identity)}}]{\alpha_t} & \mathbb{Q}\langle c, b \rangle = V_{t+1} \\ \downarrow \substack{\text{(identity)}} \psi_t & & \downarrow \substack{\text{(identity)}} \psi_{t+1} \\ W_t = \mathbb{Q}\langle z, y \rangle & \xrightarrow[\substack{\text{(identity)}}]{\gamma_t} & \mathbb{Q}\langle z, y \rangle = W_{t+1} \end{array}$$

Observe that, unlike the case for $\{\varphi_t\}_{t \in \mathbb{N}_0}$, $\psi_1 \circ \alpha_0$ and $\gamma_0 \circ \psi_0$ agree at $a \in V_0$ for $\{\psi_t\}_{t \in \mathbb{N}_0}$:

$$(\psi_1 \circ \alpha_0)(a) = \psi_1(c) = z = \gamma_0(z) = (\gamma_0 \circ \psi_0)(a)$$

Therefore, $\psi_\bullet = \{\psi_t\}_{t \in \mathbb{N}_0}$ is a persistence morphism $\psi_\bullet : V_\bullet \rightarrow W_\bullet$. Since each linear map $\psi_t : V_t \rightarrow W_t$ is a linear isomorphism, ψ_\bullet is a persistence isomorphism and $(V_\bullet, \alpha_\bullet) \cong (W_\bullet, \gamma_\bullet)$.

Next, we provide a characterization of the direct sum of persistence modules.

Definition 2.2.5. The **direct sum** $(V_\bullet, \alpha_\bullet) \oplus (W_\bullet, \beta_\bullet) =: (U_\bullet, \gamma_\bullet)$ of two persistence modules $(V_\bullet, \alpha_\bullet)$ and $(W_\bullet, \beta_\bullet)$ over a field \mathbb{F} is the persistence module over \mathbb{F} with vector spaces given by $U_t := V_t \oplus W_t$ (i.e. a direct sum of vector spaces) for all $t \in \mathbb{N}_0$ and structure maps $\gamma_{s,t} = \alpha_{s,t} \oplus \beta_{s,t}$ (i.e. the unique linear map induced by the direct sum) for all $t, s \in \mathbb{N}_0$ with $t \leq s$, i.e.

$$\begin{array}{ccc} \gamma_{s,t} : U_t := V_t \oplus W_t & \rightarrow & V_s \oplus W_s =: U_s \\ (v_t, w_t) & \mapsto & (\alpha_{s,t}(v_t), \beta_{s,t}(w_t)) \end{array}$$

Observe that this definition extends to *finite* direct sums of persistence modules. Since direct sums of persistence modules are given by those of vector spaces, properties of the direct sum of vector spaces extend to the case of persistence modules. We identify some of these below:

1. Direct sums of persistence modules are *commutative*, i.e. $(V_\bullet, \alpha_\bullet) \oplus (W_\bullet, \beta_\bullet) \cong (W_\bullet, \beta_\bullet) \oplus (V_\bullet, \alpha_\bullet)$.

2. Finite direct sums of persistence modules are *associative*, i.e.

$$\left((V_{\bullet}, \alpha_{\bullet}) \oplus (W_{\bullet}, \beta_{\bullet}) \right) \oplus (U_{\bullet}, \gamma_{\bullet}) \cong (V_{\bullet}, \alpha_{\bullet}) \oplus \left((W_{\bullet}, \beta_{\bullet}) \oplus (U_{\bullet}, \gamma_{\bullet}) \right)$$

Therefore, a statement such as $V_{\bullet} \oplus W_{\bullet} \oplus U_{\bullet}$ is unambiguous (up to persistence isomorphism).

3. The zero persistence module 0_{\bullet} with trivial vector spaces and trivial structure maps is the identity of the direct sum operation in $\mathbf{Pers}_{\mathbb{F}}$, i.e. for any persistence module $(V_{\bullet}, \alpha_{\bullet})$, $(V_{\bullet}, \alpha_{\bullet}) \oplus 0_{\bullet} \cong (V_{\bullet}, \alpha_{\bullet})$.
4. The distinction between *internal* direct sums and *external* direct sums of vector spaces also extend to the case of persistence modules, That is, there is no difference between the two constructions up to vector space isomorphism when there are a finite number of non-trivial summands:
 - (a) An internal direct sum $A = A_1 \oplus A_2$ on a vector space A is defined if A_1 and A_2 are both subspaces of A and that $A_1 \cap A_2 = \{0\}$. Here, the elements of A are generally not of the form (a_1, a_2) with $a_1 \in A_1$ and $a_2 \in A_2$. For example, $A = \mathbb{R}^2$, $A_1 = \text{span}\{(1, 1)\}$, and $A_2 = \text{span}\{(1, -1)\}$.
 - (b) An external direct sum $A := A_1 \oplus A_2$ of vector spaces A_1 and A_2 refers to direct/Cartesian product $A_1 \times A_2$ with a vector space structure induced by those of A_1 and of A_2 . Here, the spaces A_1 and A_2 are interpreted to be distinct vector spaces and the element of A are exactly of the form (a_1, a_2) with $a_1 \in A_1$ and $a_2 \in A_2$. Note that the direct sum and the direct product of a finite collection of vector spaces produce isomorphic vector spaces.

For the case of persistence modules, the collection of vector space isomorphisms between these two constructions for each index $t \in \mathbb{N}_0$ forms a persistence isomorphism.

In this paper, we prefer to interpret direct sums as internal direct sums if possible and avoid denoting the elements of the direct sum as tuples.

We provide an example of a direct sum of two persistence modules below.

Example 2.2.6. Let the persistence modules $(A_{\bullet}, \alpha_{\bullet})$ and $(B_{\bullet}, \beta_{\bullet})$ over \mathbb{Z}_2 be given as follows, with indeterminates a_i for each $i \in \{1, 2, 3\}$ and b_j for each $j \in \{1, 2, 3\}$:

$$A_t = \begin{cases} \mathbb{Z}_2 \langle a_1, a_2 \rangle & \text{if } t = 0 \\ \mathbb{Z}_2 \langle a_1, a_2, a_3 \rangle & \text{if } t = 1 \\ \mathbb{Z}_2 \langle a_2, a_3 \rangle & \text{if } t \geq 2 \end{cases} \quad \text{with} \quad \begin{cases} \alpha_0 : A_0 \rightarrow A_1 & \text{by } a_1 \mapsto a_1, \quad a_2 \mapsto a_2 \\ \alpha_1 : A_1 \rightarrow A_2 & \text{by } a_1 \mapsto 0, \quad a_2 \mapsto a_2, \quad a_3 \mapsto a_3 \\ \alpha_t : A_t \rightarrow A_{t+1} & \text{by } a_2 \mapsto a_2, \quad a_3 \mapsto a_3 \quad \text{for } t \geq 2 \end{cases}$$

$$B_t = \begin{cases} \mathbb{Z}_2 \langle b_1 \rangle & \text{if } t = 0 \\ \mathbb{Z}_2 \langle b_1, b_2 \rangle & \text{if } t = 1 \\ \mathbb{Z}_2 \langle b_1, b_2, b_3 \rangle & \text{if } t \geq 2 \end{cases} \quad \text{with} \quad \begin{cases} \beta_0 : B_0 \rightarrow B_1 & \text{by } b_1 \mapsto b_1 \\ \beta_1 : B_1 \rightarrow B_2 & \text{by } b_1 \mapsto b_1, \quad b_2 \mapsto b_2 \\ \beta_t : B_t \rightarrow B_{t+1} & \text{by } b_1 \mapsto b_1, \quad b_2 \mapsto b_2, \quad b_3 \mapsto b_3 \quad \text{for } t \geq 2 \end{cases}$$

These are illustrated below:

$$\begin{array}{c}
A_{\bullet}: \quad \begin{array}{c} A_0 \\ \parallel \\ \mathbb{Z}_2 \langle a_1, a_2 \rangle \end{array} \xrightarrow[\substack{a_1 \mapsto a_1 \\ a_2 \mapsto a_2}]{\alpha_0} \begin{array}{c} A_1 \\ \parallel \\ \mathbb{Z}_2 \langle a_1, a_2, a_3 \rangle \end{array} \xrightarrow[\substack{a_1 \mapsto 0, \\ a_2 \mapsto a_2, \\ a_3 \mapsto a_3}]{\alpha_1} \begin{array}{c} A_2 \\ \parallel \\ \mathbb{Z}_2 \langle a_2, a_3 \rangle \end{array} \xrightarrow[\text{id}]{\alpha_2} \dots \\
B_{\bullet}: \quad \begin{array}{c} B_0 \\ \parallel \\ \mathbb{Z}_2 \langle b_1 \rangle \end{array} \xrightarrow[\substack{b_1 \mapsto b_1}]{\beta_0} \begin{array}{c} B_1 \\ \parallel \\ \mathbb{Z}_2 \langle b_1, b_2 \rangle \end{array} \xrightarrow[\substack{b_1 \mapsto b_1, \quad b_2 \mapsto b_2}]{\beta_1} \begin{array}{c} B_2 \\ \parallel \\ \mathbb{Z}_2 \langle b_1, b_2, b_3 \rangle \end{array} \xrightarrow[\text{id}]{\beta_2} \dots
\end{array}$$

Let $(Y_\bullet, \gamma_\bullet) := (A_\bullet, \alpha_\bullet) \oplus (B_\bullet, \beta_\bullet)$. The vector spaces Y_t of $(Y_\bullet, \gamma_\bullet)$ are given as follows:

$$Y_t = \begin{cases} \mathbb{Z}_2\langle a_1, a_2 \rangle \oplus \mathbb{Z}_2\langle b_1 \rangle & \cong \mathbb{Z}_2\langle a_1, a_2, b_1 \rangle & \text{if } t = 0 \\ \mathbb{Z}_2\langle a_1, a_2, a_3 \rangle \oplus \mathbb{Z}_2\langle b_1, b_2 \rangle & \cong \mathbb{Z}_2\langle a_1, a_2, a_3, b_1, b_2 \rangle & \text{if } t = 1 \\ \mathbb{Z}_2\langle a_2, a_3 \rangle \oplus \mathbb{Z}_2\langle b_1, b_2, b_3 \rangle & \cong \mathbb{Z}_2\langle a_2, a_3, b_1, b_2, b_3 \rangle & \text{if } t \geq 2 \end{cases}$$

Observe that the isomorphism relations above are valid since a_i 's and b_j 's are defined to be indeterminates and that $\mathbb{Z}_2\langle a_1, a_2, a_3 \rangle \cap \mathbb{Z}_2\langle b_1, b_2, b_3 \rangle = \{0\}$.

The structure maps $\gamma_t : Y_t \rightarrow Y_{t+1}$ can be described by collecting the assignments of α_t on the basis elements a_1, a_2, a_3 and those of β_t on b_1, b_2, b_3 (whichever is included in the direct sum). For example, $\gamma_1 : Y_1 \rightarrow Y_2$ has the following assignments:

$$\gamma_1 : Y_1 \rightarrow Y_2 \quad \text{is given by} \quad \begin{cases} a_1 \mapsto 0, & a_2 \mapsto a_2, & a_3 \mapsto a_3, \\ b_1 \mapsto b_1, & b_2 \mapsto b_2, \end{cases}$$

Since $\alpha_t = \text{id}_{A_2}$ and $\beta_t = \text{id}_{B_2}$ for all $t \geq 2$, $(A_\bullet, \alpha_\bullet)$ and $(B_\bullet, \beta_\bullet)$ are both constant on $[2, \infty)$ and the same applies to the direct sum $(Y_\bullet, \gamma_\bullet)$.

For an arbitrary category \mathbf{C} , the terms *subobject*, *kernel*, *cokernel*, and *image* are generally defined to be morphisms of \mathbf{C} satisfying certain properties (as opposed to objects). For an arbitrary category \mathbf{C} , a subobject of an object x in \mathbf{C} refers to an injective morphism $y \hookrightarrow x$ with codomain x [Rie16, Definition 4.6.8]. For some of the categories that are more accessible to the introductory learner, subobjects $y \hookrightarrow x$ of an object x (assuming the morphism agrees with the identity $\text{id}_x : x \rightarrow x$, i.e. $y \hookrightarrow x$) are characterized by the domain y and are labeled using a term specific to the category. We list some examples below.

1. In the category $\mathbf{Vect}_{\mathbb{F}}$ of vector spaces over a field \mathbb{F} , subobjects correspond to (*vector*) *subspaces* and each vector subspace W of a vector space V has a corresponding inclusion map $W \hookrightarrow V$. Since W has to be a vector space (i.e. an object of $\mathbf{Vect}_{\mathbb{F}}$), the inclusion $W \hookrightarrow V$ must be a linear map, i.e. a morphism in $\mathbf{Vect}_{\mathbb{F}}$.
2. In the category $\mathbf{A-Simp}$ of (abstract) simplicial complexes and simplicial maps (as given in Definition 1.3.2), subobjects coincide with (simplicial) subcomplexes. Given a subcomplex L of a simplicial complex K , the set-wise inclusion map $L \hookrightarrow K$ is also a simplicial map (as described in Definition 1.1.3).
3. In the category \mathbf{Top} of topological spaces, subobjects are (topological) subspaces. Given any subset Y of a topological space X , Y can be equipped with the subspace topology of X . The set-wise inclusion map $Y \hookrightarrow X$ is then made a continuous map relative to this topology on Y .

Below, we provide a corresponding definition for subobjects on the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules.

Definition 2.2.7. Fix a field \mathbb{F} . Let $(V_\bullet, \alpha_\bullet)$ be a persistence module over a field \mathbb{F} . A persistence module $(W_\bullet, \gamma_\bullet)$ over \mathbb{F} is called a **persistence submodule** or **submodule** of $(V_\bullet, \alpha_\bullet)$ if the following are true:

- i. W_t is a vector subspace of V_t for all $t \in \mathbb{N}_0$.
- ii. The structure maps $\gamma_{s,t} : W_t \rightarrow W_s$ of $(W_\bullet, \gamma_\bullet)$ are exactly the structure maps of $\alpha_{s,t} : V_t \rightarrow V_s$ of $(V_\bullet, \alpha_\bullet)$ with domain and codomain restricted to W_t and W_s respectively.

That is, the collection $\{i_t\}_{t \in \mathbb{N}_0}$ of inclusion maps $i_t : W_t \hookrightarrow V_t$ forms a persistence morphism $(W_\bullet, \gamma_\bullet) \hookrightarrow (V_\bullet, \alpha_\bullet)$. In this case, we write $(W_\bullet, \gamma_\bullet) \subseteq (V_\bullet, \alpha_\bullet)$ (or $W_\bullet \subseteq V_\bullet$ for convenience).

This definition of subobjects for persistence modules allows characterizations of kernels, cokernels, and

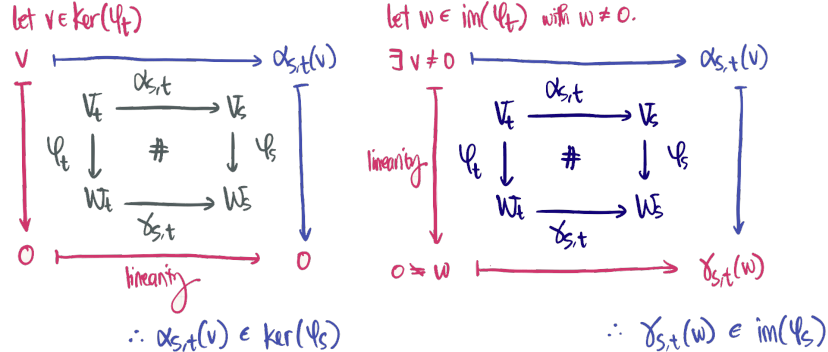
images of persistence morphisms like those of linear maps in $\mathbf{Vect}_{\mathbb{F}}$, i.e. as objects in the category as opposed to morphisms. In particular, we extend the set-wise definition of kernels, images, and cokernels in $\mathbf{Vect}_{\mathbb{F}}$ to the case of $\mathbf{Pers}_{\mathbb{F}}$. The composition relation on the structure maps of persistence modules (i.e. the commuting squares condition) guarantee that the resulting collection of vector subspaces form a persistence submodule. We state this in more detail below.

Definition 2.2.8. Let $\varphi_{\bullet} : (V_{\bullet}, \alpha_{\bullet}) \rightarrow (W_{\bullet}, \gamma_{\bullet})$ be a persistence morphism between two persistence modules $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ over a field \mathbb{F} with $\varphi_{\bullet} = \{\varphi_t : V_t \rightarrow W_t\}_{t \in \mathbb{N}_0}$.

- i. The **kernel** $\ker(\varphi_{\bullet}) =: (K_{\bullet}, \kappa_{\bullet})$ of φ_{\bullet} is the submodule of $(V_{\bullet}, \alpha_{\bullet})$ with $K_t = \ker(\varphi_t) \subseteq V_t$ for all $t \in \mathbb{N}_0$.
- ii. The **image** $\text{im}(\varphi_{\bullet}) =: (B_{\bullet}, \beta_{\bullet})$ of φ_{\bullet} is the submodule of $(W_{\bullet}, \gamma_{\bullet})$ with $B_t = \text{im}(\varphi_t) \subseteq W_t$ for all $t \in \mathbb{N}_0$.
- iii. The **cokernel** $\text{coker}(\varphi_{\bullet}) =: (C_{\bullet}, \lambda_{\bullet})$ of φ_{\bullet} is the persistence module with vector spaces given by $C_t := \text{coker}(\varphi_t) = W_t / \text{im}(\varphi_t)$ for all $t \in \mathbb{N}_0$ and structure maps $\lambda_{s,t} : C_t \rightarrow C_s$ being the linear maps induced by $\gamma_{s,t} : W_t \rightarrow W_s$ and the cokernel/quotient construction on $\mathbf{Vect}_{\mathbb{F}}$, i.e. $\eta_{s,t}$ maps $[w] \rightarrow [\gamma_{s,t}(w)]$ with $w \in W_t$, for all $t, s \in \mathbb{N}_0$ with $t \leq s$.

Remark. The symbols K , B , and C for the kernel, image, and cokernel respectively are used here for convenience, i.e. we do not typically use these letters for the kernel, image, and cokernel respectively.

Observe that the vector spaces for the kernel, image, and cokernel of persistence morphisms are all well-defined since the kernel, image, and cokernel of linear maps are well-defined. For the kernel and image, the question is whether the resulting structure maps by Definition 2.2.7 as persistence submodules are well-defined. We can show this using a diagram chase using a diagram chase on the commuting squares of $(V_{\bullet}, \alpha_{\bullet})$ and $(W_{\bullet}, \gamma_{\bullet})$ on the indices t and s , as illustrated below.



The diagram chase on the left implies that $\alpha_{s,t}(\ker(\varphi_t)) \subseteq \ker(\varphi_s) \subseteq V_s$ and that the restriction of the domain V_t to $\ker(\varphi_t)$ and the codomain V_s to $\ker(\varphi_s)$ of the structure map $\alpha_{s,t} : V_t \rightarrow V_s$ is well-defined.

The diagram chase on the right implies that $\gamma_{s,t}(\text{im}(\varphi_t)) \subseteq \text{im}(\varphi_s) \subseteq W_s$. This tells us that the restriction of the domain and codomain of the structure map $\gamma_{s,t} : W_t \rightarrow W_s$ to $\text{im}(\varphi_t)$ and $\text{im}(\varphi_s)$ respectively is well-defined. This also tells us that the structure maps of $\text{coker}(\varphi_{\bullet})$ are well-defined since

$$\gamma_{s,t}(w + \text{im}(\varphi_t)) = \gamma_{s,t}(w) + \text{im}(\varphi_s) \in W_s$$

and that, for any coset representative $w \in \text{im}(\varphi_t)$ of the trivial element $[w] = 0 \in W_t / \text{im}(\varphi_t)$, $\lambda_{s,t}([w]) = [\gamma_{s,t}(w)] = 0 \in W_s / \text{im}(\varphi_s)$.

We provide an example below where the kernel, image, and cokernel of a persistence morphism are identified.

Example 2.2.9. Let $(A_\bullet, \alpha_\bullet)$ and $(X_\bullet, \chi_\bullet)$ be persistence modules over \mathbb{Q} with vector spaces A_n and X_n defined as follows:

$$A_t = \begin{cases} \mathbb{Q}\langle a \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle a, b \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle a, b, c, d \rangle & \text{if } t \geq 2 \end{cases} \quad \text{and} \quad X_t = \begin{cases} \mathbb{Q}\langle x, y \rangle & \text{if } t = 0, 1 \\ \mathbb{Q}\langle x, y, z \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle x, y, z, w \rangle & \text{if } t \geq 3 \end{cases}$$

where the structure maps $\alpha_{s,t} : A_t \rightarrow A_s$ of A_\bullet and $\chi_{s,t} : X_t \rightarrow X_s$ of X_\bullet are restrictions of the identity maps on $\mathbb{Q}\langle a, b, c, d \rangle$ and $\mathbb{Q}\langle x, y, z, w \rangle$ respectively. Define the map $\Phi : \mathbb{Q}\langle a, b, c, d \rangle \rightarrow \mathbb{Q}\langle x, y, z, w \rangle$ by

$$[\Phi] = \begin{matrix} & a & b & c & d \\ \begin{matrix} x \\ y \\ z \\ w \end{matrix} & \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad \text{or equivalently} \quad \begin{cases} a \mapsto x \\ b \mapsto -x \\ c \mapsto x + y + z \\ d \mapsto y + z \end{cases}$$

For each $t \in \mathbb{N}_0$, define $\varphi_t : A_t \rightarrow X_t$ by $\varphi_t(f) = \Phi(f)$ for all $f \in A_t$. Observe that the codomains of φ_t with $t = 0, 1, 2, 3$ are well-defined. We claim that $\varphi_\bullet = \{\varphi_t\}_{t \in \mathbb{N}_0}$ is a persistence morphism $\varphi_\bullet : (A_\bullet, \alpha_\bullet) \rightarrow (X_\bullet, \chi_\bullet)$.

The vector spaces of the persistence module $\ker(\varphi_\bullet)$ by $\ker(\varphi_\bullet)(t) := \ker(\varphi_t)$ and of the persistence module $\text{im}(\varphi_\bullet)$ by $\text{im}(\varphi_\bullet)(t) := \text{im}(\varphi_t)$ are as follows:

$$\ker(\varphi_t) = \begin{cases} 0 & \text{if } t = 0 \\ \mathbb{Q}\langle a + b \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle a + b, a - c + d \rangle & \text{if } t \geq 2 \end{cases} \quad \text{and} \quad \text{im}(\varphi_t) = \begin{cases} \mathbb{Q}\langle x \rangle & \text{if } t = 0, 1 \\ \mathbb{Q}\langle x, y + z \rangle & \text{if } t \geq 2 \end{cases}$$

Since the vector subspace relations $A_t \subseteq A_s$ and $X_t \subseteq X_s$ for $t \leq s$ are preserved by φ_t and φ_s , the structure maps $\ker(\varphi_\bullet)(t \rightarrow s)$ and $\text{im}(\varphi_\bullet)(t \rightarrow s)$, which are inclusion maps as restrictions of identity maps, are both well-defined. Observe that $\ker(\varphi_\bullet)$ is a persistence submodule of $(A_\bullet, \alpha_\bullet)$ and $\text{im}(\varphi_\bullet)$ is a persistence submodule of $(X_\bullet, \chi_\bullet)$.

The vector spaces of $\text{coker}(\varphi_\bullet)$ by $\text{coker}(\varphi_\bullet)(t) = \text{coker}(\varphi_t)$ can be described as follows:

$$\text{coker}(\varphi_t) = \frac{X_t}{\text{im}(\varphi_t)} = \begin{cases} \mathbb{Q}\langle x, y \rangle / \mathbb{Q}\langle x \rangle & \cong \mathbb{Q}\langle [y] \rangle & \cong \mathbb{Q} & \text{if } t = 0 \\ \mathbb{Q}\langle x, y \rangle / \mathbb{Q}\langle x \rangle & \cong \mathbb{Q}\langle [y] \rangle & \cong \mathbb{Q} & \text{if } t = 1 \\ \mathbb{Q}\langle x, y, z \rangle / \mathbb{Q}\langle x, y + z \rangle & \cong \mathbb{Q}\langle [z] \rangle & \cong \mathbb{Q} & \text{if } t = 2 \\ \mathbb{Q}\langle x, y, z, w \rangle / \mathbb{Q}\langle x, y + z \rangle & \cong \mathbb{Q}\langle [z], [w] \rangle & \cong \mathbb{Q}^2 & \text{if } t \geq 3 \end{cases}$$

Since vector spaces have no torsion, the ranks of the $\text{coker}(\varphi_t)$ for all $t \in \mathbb{N}_0$ is determined by the rank of X_t minus the rank of $\text{im}(\varphi_t)$. The structure maps of $\text{coker}(\varphi_\bullet)$ are illustrated by the following sequence:

$$\begin{array}{ccccccc} \text{coker}(\varphi_0) & & \text{coker}(\varphi_1) & & \text{coker}(\varphi_2) & & \text{coker}(\varphi_3) \\ \parallel & & \parallel & & \parallel & & \parallel \\ \mathbb{Q}\langle [y] \rangle & \xrightarrow{[y] \mapsto [y]} & \mathbb{Q}\langle [y] \rangle & \xrightarrow{\substack{[y] \mapsto 0 \\ [y] \text{ becomes trivial} \\ \text{on addition of } y + z \\ \text{to } \text{im}(\varphi_1)}} & \mathbb{Q}\langle [z] \rangle & \xrightarrow{\substack{[z] \mapsto [z] \\ [z] \text{ is unaffected} \\ \text{by addition of } w \\ \text{to } X_3}} & \mathbb{Q}\langle [z], [w] \rangle & \xrightarrow{\substack{\text{structure maps} \\ \text{are identity maps} \\ \text{for } t \geq 3}} & \cdots \end{array}$$

Note that defining the cokernel of morphisms in an abelian category defines the quotient operation on said category. For example, in the category of \mathbf{Mod}_R of modules over a PID R , the quotient module X/Y of an

R -module X by its submodule Y is exactly the cokernel of the inclusion map $Y \hookrightarrow X$, i.e.

$$X / Y = X / \text{im}(Y \hookrightarrow X) = \text{coker}(Y \hookrightarrow X)$$

Consequently, Definition 2.2.8 also tells us that the quotient operation on a persistence module by its submodule is done pointwise for each $t \in \mathbb{N}_0$. Since we have definitions for persistence submodule relations and quotients of persistence modules, we can form chain complexes of persistence modules.

More generally, given any abelian category \mathbf{A} , there is a corresponding category $\mathbf{Ch-A}$ of chain complexes in \mathbf{A} , defined very similarly as in the case of chain complexes of R -modules. For those interested, we refer to [Rot08, Section 5.5] and [Wei95, Chapter 1] for a more general and detailed discussion. We state what this means specifically for $\mathbf{Pers}_{\mathbb{F}}$ below.

Definition 2.2.10. The **category of persistence complexes over a field \mathbb{F}** , denoted $\mathbf{Ch-Pers}_{\mathbb{F}}$, refers to the category of chain complexes on $\mathbf{Pers}_{\mathbb{F}}$, with objects and morphism described below:

- i. The objects of $\mathbf{Ch-Pers}_{\mathbb{F}}$ are persistence complexes, defined as follows:

A **persistence complex** $(V_{\bullet}, \alpha_{\bullet}, \partial_{\bullet})$ over \mathbb{F} is a \mathbb{Z} -indexed collection of persistence modules $(V_n^{\bullet}, \alpha_n^{\bullet})$ over \mathbb{F} and persistence morphisms $\partial_n^{\bullet} : (V_n^{\bullet}, \alpha_n^{\bullet}) \rightarrow (V_{n-1}^{\bullet}, \alpha_{n-1}^{\bullet})$ such that $\partial_n^{\bullet} \circ \partial_{n-1}^{\bullet} = 0^{\bullet}$ for all $n \in \mathbb{Z}$, where 0^{\bullet} denotes the zero persistence morphism. Illustrated below is $(V_{\bullet}, \alpha_{\bullet}, \partial_{\bullet})$ as a filtered sequence of persistence modules and persistence morphisms:

$$\cdots \xrightarrow{\partial_{n+2}^{\bullet}} V_{n+1}^{\bullet} \xrightarrow{\partial_{n+1}^{\bullet}} V_n^{\bullet} \xrightarrow{\partial_n^{\bullet}} V_{n-1}^{\bullet} \xrightarrow{\partial_{n-1}^{\bullet}} \cdots$$

For convenience, we may suppress the structure maps of V^{\bullet} and write $(V_{\bullet}, \partial_{\bullet}) := (V_{\bullet}, \alpha_{\bullet}, \partial_{\bullet})$.

- ii. The morphisms of $\mathbf{Ch-Pers}_{\mathbb{F}}$ are persistence chain morphisms, defined as follows:

A **persistence chain morphism** $f_{\bullet}^{\bullet} : (V_{\bullet}, \partial_{\bullet}) \rightarrow (W_{\bullet}, \delta_{\bullet})$ between two persistence complexes $(V_{\bullet}, \partial_{\bullet})$ and $(W_{\bullet}, \delta_{\bullet})$ over \mathbb{F} is a \mathbb{Z} -indexed collection of persistence morphisms $f_n^{\bullet} : V_n^{\bullet} \rightarrow W_n^{\bullet}$ such that for all $n \in \mathbb{Z}$, $f_{n-1}^{\bullet} \circ \partial_n^{\bullet} = \delta_n^{\bullet} \circ f_n^{\bullet}$, i.e. we have the following commutative squares:

$$\begin{array}{ccc} V_n^{\bullet} & \xrightarrow{\partial_n^{\bullet}} & V_{n-1}^{\bullet} \\ f_n^{\bullet} \downarrow & \# & \downarrow f_{n-1}^{\bullet} \\ W_n^{\bullet} & \xrightarrow{\delta_n^{\bullet}} & W_{n-1}^{\bullet} \end{array}$$

Observe that the notation for the persistence complex $(V_{\bullet}, \partial_{\bullet})$ has two “placeholder” indices: an index $t \in \mathbb{N}_0$ denoted by the bullet (\bullet) and an index $n \in \mathbb{Z}$ denoted by the asterisk ($*$). Since each $n \in \mathbb{Z}$ identifies a persistence module $(V_n^{\bullet}, \alpha_n^{\bullet})$ and each $t \in \mathbb{N}_0$ identifies an \mathbb{F} -vector space V_n^t , a persistence complex $(V_{\bullet}, \partial_{\bullet})$ corresponds to the following commutative grid of vector spaces:

$$\begin{array}{ccccccc}
& & V_{n+1}^\bullet & & V_n^\bullet & & V_{n-1}^\bullet \\
& & \vdots & & \vdots & & \vdots \\
& & \uparrow \alpha_{n+1}^{t+1} & & \uparrow \alpha_n^{t+1} & & \uparrow \alpha_{n-1}^{t+1} \\
V_*^{t+1} : & \cdots & \xrightarrow{\partial_{n+2}^{t+1}} & V_{n+1}^{t+1} & \xrightarrow{\partial_{n+1}^{t+1}} & V_n^{t+1} & \xrightarrow{\partial_n^{t+1}} & V_{n-1}^{t+1} & \xrightarrow{\partial_{n-1}^{t+1}} & \cdots \\
& & \uparrow \alpha_{n+1}^t & & \uparrow \alpha_n^t & & \uparrow \alpha_{n-1}^t \\
V_*^t : & \cdots & \xrightarrow{\partial_{n+2}^t} & V_{n+1}^t & \xrightarrow{\partial_{n+1}^t} & V_n^t & \xrightarrow{\partial_n^t} & V_{n-1}^t & \xrightarrow{\partial_{n-1}^t} & \cdots \\
& & \uparrow \alpha_{n+1}^{t-1} & & \uparrow \alpha_n^{t-1} & & \uparrow \alpha_{n-1}^{t-1} \\
V_*^{t-1} : & \cdots & \xrightarrow{\partial_{n+2}^{t-1}} & V_{n+1}^{t-1} & \xrightarrow{\partial_{n+1}^{t-1}} & V_n^{t-1} & \xrightarrow{\partial_n^{t-1}} & V_{n-1}^{t-1} & \xrightarrow{\partial_{n-1}^{t-1}} & \cdots \\
& & \uparrow \alpha_{n+1}^{t-2} & & \uparrow \alpha_n^{t-2} & & \uparrow \alpha_{n-1}^{t-2} \\
& & \vdots & & \vdots & & \vdots
\end{array}$$

Note that each column represents a single persistence module and each row represents a chain complex of vector spaces and linear maps. Persistence complexes will be significant later in Section 3.3 of Chapter 3 where we extend the construction of simplicial homology of simplicial complexes to the case of persistence modules.

The category **Ch-A** of chain complexes on an abelian category **A** also brings with it a family of chain homology functors $H_n(-) : \mathbf{Ch-A} \rightarrow \mathbf{A}$, one for each $n \in \mathbb{Z}$. We state a corresponding definition specific to the category **Pers_F** below.

Definition 2.2.11. For each $n \in \mathbb{Z}$, the n^{th} **chain homology functor** $H_n : \mathbf{Ch-Pers_F} \rightarrow \mathbf{Pers_F}$ on **Ch-Pers_F** sends a persistence complex $(V_\bullet, \partial_\bullet)$ to its n^{th} **chain homology** $H_n(V_\bullet, \partial_\bullet)$, which is the persistence module given by

$$H_n(V_\bullet, \partial_\bullet) = \frac{\ker(\partial_n^\bullet)}{\text{im}(\partial_{n+1}^\bullet)} = \text{coker}\left(\text{im}(\partial_{n+1}^\bullet) \hookrightarrow \ker(\partial_n^\bullet)\right)$$

Section 2.3. Interval Decompositions of Persistence Modules

Now that we have definitions for isomorphisms and direct sums between persistence modules, it is possible for us to talk about decompositions of persistence modules. Note that we use the term *decomposition*, or more specifically *direct sum decomposition*, in the same way we would for vector spaces and modules. That is, a decomposition of a persistence module consists of a (finite) direct sum of other persistence modules such that the direct sum is isomorphic to the original persistence module.

In persistence theory, we are interested in a unique form of decomposition called *interval decomposition*. In this section, we define the notions of interval modules and interval decompositions and provide examples for each. To start, we provide a definition for interval modules, adapted from [CSGO13, Section 1.4].

Definition 2.3.1. Let \mathbb{F} be a field. The J -**interval module** \mathbb{I}_{\bullet}^J over \mathbb{F} for some interval $J \subseteq \mathbb{N}_0$ is the persistence module $\mathbb{I}_{\bullet}^J = (\mathbb{I}_{\bullet}^J, i_{\bullet}^J) : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$ with vector spaces \mathbb{I}_t^J and structure maps $i_{s,t}^J : \mathbb{I}_t^J \rightarrow \mathbb{I}_s^J$ given as follows:

$$\mathbb{I}_t^J = \begin{cases} \mathbb{F} & \text{if } t \in J \\ 0 & \text{if } t \notin J \end{cases} \quad \text{for all } t \in \mathbb{N}_0 \quad \text{and} \quad i_{s,t}^J = \begin{cases} \text{id}_{\mathbb{F}} & \text{if } t, s \in J \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } t, s \in \mathbb{N}_0 \text{ with } t \leq s.$$

If the interval J and the field \mathbb{F} are arbitrary, we may drop references to both and refer to \mathbb{I}_{\bullet}^J as an **interval module**. If $J = [a, b)$, we may write $\mathbb{I}_{\bullet}^{[a,b)}$ to refer to \mathbb{I}_{\bullet}^J . Similarly, we may write $\mathbb{I}_{\bullet}^{[a,\infty)}$ if $J = [a, \infty)$.

Remark. The condition that $i_{s,t}^J = \text{id}_{\mathbb{F}}$ if $t, s \in J$ is stated as $i_{s,t} = 1$ in [CSGO13, Section 1.4]. In this case, $i_{s,t}^J = 1$ refers to the linear map $\mathbb{F} \rightarrow \mathbb{F}$ by $k \mapsto 1 \cdot k = k$ for all $k \in \mathbb{F}$.

Since we are only using \mathbb{N}_0 as the indexing set for persistence modules, we can characterize all intervals in \mathbb{N}_0 as pairs of values. In particular, any interval in \mathbb{N}_0 can be represented using exactly one of the two representations below:

$$[a, b) = \{n \in \mathbb{N}_0 : a \leq n < b\} \quad \text{or} \quad [a, \infty) = \{n \in \mathbb{N}_0 : a \leq n\}$$

Observe that, assuming we allow b in $[a, b)$ to have ∞ as a value (i.e. $b \in \mathbb{N}_0 \cup \{\infty\}$), every interval in \mathbb{N}_0 can be unambiguously represented by the two endpoints a and b . Thus, in some papers (e.g. [ZC05; Bau21]), the interval $[a, b)$ is represented using an ordered pair (a, b) .

Since we will be working at the level of persistence isomorphisms, it will be helpful to have a characterization of persistence modules that are isomorphic to interval modules. We provide such a characterization below in terms of ranks.

Lemma 2.3.2. A persistence module $(V_{\bullet}, \alpha_{\bullet})$ over \mathbb{F} is isomorphic to the interval module \mathbb{I}_{\bullet}^J for some interval $J \subseteq \mathbb{N}_0$ if and only if the vector spaces V_t and structure maps $\alpha_{s,t} : V_t \rightarrow V_s$ satisfy the following:

$$\text{rank}(V_t) = \begin{cases} 1 & \text{if } t \in J \\ 0 & \text{if } t \notin J \end{cases} \quad \text{and} \quad \text{rank}(\alpha_{s,t}) = \begin{cases} 1 & \text{if } t, s \in J \\ 0 & \text{otherwise} \end{cases} \quad (\text{E1})$$

where $\text{rank}(-)$ refers to the rank of \mathbb{F} -vector spaces and of \mathbb{F} -linear maps.

Proof. Let $(V_{\bullet}, \alpha_{\bullet})$ be a persistence module over \mathbb{F} .

For the forward direction, assume that $V_{\bullet} \cong \mathbb{I}_{\bullet}^J$ for some interval $J \subseteq \mathbb{N}_0$ with $\mathbb{I}_{\bullet}^J = (\mathbb{I}_{\bullet}^J, i_{\bullet}^J)$. Then, there exists a persistence isomorphism $\varphi_{\bullet} : V_{\bullet} \rightarrow \mathbb{I}_{\bullet}^J$ with $\varphi_{\bullet} = \{\varphi_t : V_t \rightarrow \mathbb{I}_t^J\}_{t \in \mathbb{N}_0}$. By Definition 2.2.3, $\varphi_t : V_t \rightarrow \mathbb{I}_t^J$ is an \mathbb{F} -vector space isomorphism for each $t \in \mathbb{N}_0$. Since linear

isomorphisms preserve the ranks of vector spaces, we have the following for each $t \in \mathbb{N}_0$:

$$\text{rank}(V_t) = \text{rank}(\varphi_t(V_t)) = \text{rank}(\mathbb{I}_t^J) = \begin{cases} \text{rank}(\mathbb{F}) = 1 & \text{if } t \in J \\ \text{rank}(0) = 0 & \text{otherwise} \end{cases}$$

Similarly, composition with linear isomorphisms preserve the ranks of linear maps. Then, for all $t, s \in \mathbb{N}_0$ with $t \leq s$:

$$\text{rank}(\alpha_{s,t}) = \text{rank}(\varphi_s \circ \alpha_{s,t} \circ \varphi_t) = \text{rank}(i_{s,t}^J) = \begin{cases} \text{rank}(\text{id}_{\mathbb{F}}) = 1 & \text{if } t, s \in J \\ \text{rank}(0) = 0 & \text{otherwise} \end{cases}$$

Therefore, V_\bullet satisfies (E1) for the interval $J \subseteq \mathbb{N}_0$.

For the backwards direction, assume that $(V_\bullet, \alpha_\bullet)$ satisfies (E1) for some interval $J \subseteq \mathbb{N}_0$. To construct the persistence isomorphism $\varphi_\bullet : V_\bullet \rightarrow \mathbb{I}_\bullet^J$ with $\mathbb{I}_\bullet^J = (\mathbb{I}_\bullet^J, i_\bullet^J)$, we build the linear maps $\varphi_t : V_t \rightarrow \mathbb{I}_t^J$ inductively. For clarity, let $1_t \in \mathbb{F}$ refer to the multiplicative identity of \mathbb{F} , specifically as the vector space of \mathbb{I}_\bullet^J at index $t \in J$, i.e. $i_{s,t}^J(1_t) = \text{id}_{\mathbb{F}}(1_t) = 1_s$ for all $t, s \in J$ with $t \leq s$.

1. Let $a = \min(J)$, which exists by the well-ordering principle on \mathbb{N}_0 . That is, $J = [a, b)$ for some $b \in \mathbb{N}_0$ or $J = [a, \infty)$. Choose a nonzero $\sigma_a \in V_a$, which exists by assumption of $\text{rank}(V_a) = 1$. Observe that $\{\sigma_a\}$ is a basis of V_a . Define the linear map $\varphi_a : V_a \rightarrow \mathbb{I}_a^J = \mathbb{F}$ by $\sigma_a \mapsto 1_a$.
2. Let $t \in J$ with $t \neq a$. Let $\sigma_t \in V_t$ be such that $\alpha_{t,a}(\sigma_a) = \sigma_t$. By assumption, $\text{rank}(V_t) = 1$ and $\text{rank}(\alpha_{a,t}) = 1$. Then, $\sigma_t \neq 0$ and $\{\sigma_t\}$ is a basis for V_t . Define the linear map $\varphi_t : V_t \rightarrow \mathbb{I}_t^J = \mathbb{F}$ by $\sigma_t \mapsto 1_t$.
3. Let $t \in \mathbb{N}_0$ with $t \notin J$. By assumption, $\text{rank}(V_t) = 0$ and V_t is the trivial vector space. Then, $\varphi_t : V_t \rightarrow \mathbb{I}_t^J = 0$ can only be the trivial map.

Let $t, s \in \mathbb{N}_0$ with $t \leq s$. By Definition 2.2.1(ii), we need to show that the following composition relation is satisfied: $\varphi_s \circ \alpha_{s,t} = i_{s,t}^J \circ \varphi_t$ where $\alpha_{s,t} : V_t \rightarrow V_s$ is a structure map of V_\bullet and $i_{s,t}^J : \mathbb{I}_t^J \rightarrow \mathbb{I}_s^J$ is that of \mathbb{I}_\bullet^J . We examine three cases:

1. Assume $t, s \in J$. Then, V_t and V_s are non-trivial \mathbb{F} -vector spaces with bases $\{\sigma_t\}$ and $\{\sigma_s\}$ respectively. By Lemma 2.1.3(ii), $\sigma_s = \alpha_{s,a}(\sigma_a) = (\alpha_{s,t} \circ \alpha_{t,a})(\sigma_a) = \alpha_{s,t}(\sigma_t)$. Then,

$$(\varphi_s \circ \alpha_{s,t})(\sigma_t) = \varphi_s(\sigma_s) = 1_s = i_{s,t}^J(1_t) = (i_{s,t}^J \circ \varphi_t)(\sigma_t)$$

Therefore, the relation $\varphi_s \circ \alpha_{s,t} = i_{s,t}^J \circ \varphi_t$ is satisfied.

2. Assume $t \notin J$. Then, V_t is the trivial vector space by assumption of $\text{rank}(V_t) = 0$ and the relation $\varphi_s \circ \alpha_{s,t} = i_{s,t}^J \circ \varphi_t$ is trivially satisfied.
3. Assume $s \notin J$. Then, $\mathbb{I}_s^J = 0$ and the maps $\varphi_s : V_s \rightarrow \mathbb{I}_s^J$ and $i_{s,t}^J : \mathbb{I}_t^J \rightarrow \mathbb{I}_s^J$ are necessarily zero. Then, the composition relation $\varphi_s \circ \alpha_{s,t} = i_{s,t}^J \circ \varphi_t = 0$ is trivially satisfied.

Therefore, $\varphi_\bullet : V_\bullet \rightarrow \mathbb{I}_\bullet^J$ is a persistence isomorphism and $V_\bullet \cong \mathbb{I}_\bullet^J$. ■

As discussed in Section 2.1, the structure maps of a persistence module $(V_\bullet, \alpha_\bullet)$ are uniquely determined by the collection $\{\alpha_t : V_t \rightarrow V_{t+1}\}_{t \in \mathbb{N}_0}$ of linear maps. Let the interval $J \subseteq \mathbb{N}_0$ be given by $J = [a_k, b_k)$ for some $a_k \in \mathbb{N}_0$ and $b_k \in [a_k, \infty) \cup \{\infty\}$. The condition (E1) in Lemma 2.3.2 can equivalently be stated as follows:

$$\text{rank}(V_t) = \begin{cases} 1 & \text{if } t \in [a, b) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \text{rank}(\alpha_t) = \begin{cases} 1 & \text{if } t \in [a, b-1) \\ 0 & \text{otherwise} \end{cases} \quad (\text{E2})$$

where $[a, b-1)$ is interpreted to be $[a, \infty)$ if $b = \infty$. Since the structure map $\alpha_{b-1} : V_{b-1} \rightarrow V_b$ has trivial

codomain $V_b = 0$, it must also be trivial. Lemma 2.3.2 also implies that if $V_\bullet \cong \mathbb{I}_\bullet^{[a,b]}$, then any nonzero element $\sigma_a \in V_a$ will satisfy $\alpha_{t,a}(\sigma) \neq 0$ for all $t \in [a, b-1]$. We illustrate this below relative to the sequence representation of V_\bullet , with the additional assumption that $b \in \mathbb{N}_0$:

$$\begin{array}{ccccccccccc} 0 = V_{a-1} & \xrightarrow{\alpha_{a-1}} & V_a & \xrightarrow{\alpha_a} & V_{a+1} & \xrightarrow{\alpha_{a+1}} & V_{a+2} & \longrightarrow & \cdots & \longrightarrow & V_{b-1} & \xrightarrow{\alpha_{b-1}} & V_b = 0 \\ & & \sigma_a & \longmapsto & v_{a+1} & \longmapsto & v_{a+2} & \longmapsto & \cdots & \longmapsto & v_{b-1} & \longmapsto & 0 \\ & & & & \Downarrow & & \Downarrow & & & & \Downarrow & & \\ & & & & 0 & & 0 & & & & 0 & & \end{array}$$

where $v_t := \alpha_{t,a}(\sigma_a)$ for $t \geq a+1$. Observe that for all $t \in [a, b]$, $\{v_t\}$ is a basis for V_t since $v_t \neq 0$ and $\text{rank}(V_t) = 1$ by assumption. For $t \geq b$, $v_t = 0$ and V_t is the trivial vector space. We give an example of a persistence module that is isomorphic to an interval module below.

Example 2.3.3. Let $(Q_\bullet, \gamma_\bullet)$ be a persistence module over \mathbb{Q} with vector spaces Q_t and structure maps $\gamma_t : Q_t \rightarrow Q_{t+1}$ given as follows:

$$Q_t = \begin{cases} \mathbb{Q}\langle a \rangle & \text{if } t = 3 \\ \mathbb{Q}\langle b \rangle & \text{if } t = 4 \\ \mathbb{Q}\langle c \rangle & \text{if } t = 5 \\ 0 & \text{if } t \notin \{3, 4, 5\} \end{cases} \quad \text{and} \quad \begin{cases} \gamma_3 : Q_3 \rightarrow Q_4 & \text{by } a \mapsto 2b \\ \gamma_4 : Q_4 \rightarrow Q_5 & \text{by } b \mapsto \frac{1}{3}c \\ \gamma_5 : Q_5 \rightarrow Q_6 & \text{by } c \mapsto 0 \end{cases}$$

Note that for all $t \notin \{3, 4, 5\}$, the structure map $\gamma_t : Q_t \rightarrow Q_{t+1}$ is necessarily the zero map since $Q_t = 0$.

By Lemma 2.3.2, $Q_\bullet \cong \mathbb{I}_\bullet^J$ with $J = \{3, 4, 5\} = [3, 6]$. We can also determine a corresponding persistence isomorphism $\varphi_\bullet : Q_\bullet \rightarrow \mathbb{I}_\bullet^{[3,6]}$ using the same arguments presented in the proof of Lemma 2.3.2. Choose $a \in Q_3 = \mathbb{Q}\langle a \rangle$. Then, the images of a under the structure maps of Q_\bullet is illustrated as follows:

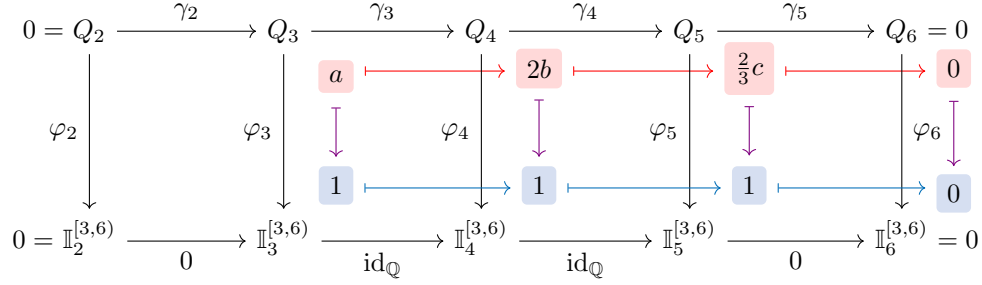
$$\begin{array}{ccccccc} 0 = Q_2 & \xrightarrow{\gamma_2} & Q_3 & \xrightarrow{\gamma_3} & Q_4 & \xrightarrow{\gamma_3} & Q_5 & \xrightarrow{\gamma_3} & Q_6 = 0 \\ & & a & \longmapsto & 2b & \longmapsto & \frac{2}{3}c & \longmapsto & 0 \end{array}$$

Observe that $\{2b\}$ and $\{\frac{2}{3}c\}$ are bases of $Q_4 = \mathbb{Q}\langle b \rangle$ and $Q_5 = \mathbb{Q}\langle c \rangle$ respectively. The linear maps $\varphi_t : Q_t \rightarrow \mathbb{I}_t^{[3,6]}$ of the persistence isomorphism φ_\bullet are then given as follows:

$$\begin{array}{ccc} \varphi_3 : Q_3 = \mathbb{Q}\langle a \rangle \rightarrow \mathbb{I}_3^{[3,6]} = \mathbb{Q} & , & \varphi_4 : Q_4 = \mathbb{Q}\langle b \rangle \rightarrow \mathbb{I}_4^{[3,6]} = \mathbb{Q} & , & \varphi_5 : Q_5 = \mathbb{Q}\langle c \rangle \rightarrow \mathbb{I}_5^{[3,6]} = \mathbb{Q} \\ a \mapsto 1 & & 2b \mapsto 1 & & \frac{2}{3}c \mapsto 1 \end{array}$$

Note that for all $t \notin [3, 6]$, the linear map $\varphi_t : Q_t \rightarrow \mathbb{I}_t^{[3,6]}$ is the trivial map since $Q_t = 0$ and $\mathbb{I}_t^{[3,6]} = 0$.

Let $\mathbb{I}_\bullet^{[3,6]} = (\mathbb{I}_\bullet^{[3,6]}, i_\bullet)$, i.e. let $i_{s,t} : \mathbb{I}_t^{[3,6]} \rightarrow \mathbb{I}_s^{[3,6]}$ denote the structure maps of $\mathbb{I}_\bullet^{[3,6]}$. Since φ_\bullet is a persistence morphism, the commutativity relation $\varphi_s \circ \gamma_{s,t} = i_{s,t} \circ \varphi_t$ is satisfied for all $t, s \in \mathbb{N}_0$ with $t \leq s$. This can be visualized using the following diagram, where highlighted in red are elements of red Q_t and in blue are those of blue $\mathbb{I}_t^{[3,6]}$.



Next, we provide a definition for interval decompositions, adapted from [CSGO13, Section 1.5].

Definition 2.3.4. An **interval decomposition** of a persistence module $(V_\bullet, \alpha_\bullet)$ over a field \mathbb{F} is a finite direct sum $\bigoplus_k \mathbb{I}_\bullet^{J_k}$ of interval modules where $\{J_k\}_{k=1}^m$ is some multiset of intervals in \mathbb{N}_0 such that

$$V_\bullet \cong \bigoplus_{k=1}^m \mathbb{I}_\bullet^{J_k} \cong \left(\mathbb{I}_\bullet^{J_1} \right) \oplus \left(\mathbb{I}_\bullet^{J_2} \right) \oplus \cdots \oplus \left(\mathbb{I}_\bullet^{J_m} \right)$$

We say that V_\bullet **admits an interval decomposition** if there exists an interval decomposition of V_\bullet .

Remark. Some authors refer to $\{J_k\}$ as a collection instead of a multiset. While we may conventionally use the term *collection* to refer to a set or some set-like object, we will use the term *collection* to refer to multiset whenever we are talking about interval decompositions. For example, the collection $\{[1, 4), [1, 4)\}$ corresponds to the interval decomposition $\mathbb{I}_\bullet^{[1, 4)} \oplus \mathbb{I}_\bullet^{[1, 4)}$.

Observe that the definition of the term “interval decomposition” does not require that a persistence isomorphism $\varphi_\bullet : V_\bullet \rightarrow \bigoplus_k \mathbb{I}_\bullet^{J_k}$ be explicitly given, only that at least one exists. This convention is justified by the following uniqueness theorem.

Theorem 2.3.5. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module over \mathbb{F} . Assume that there exists two interval decompositions for V_\bullet given as follows:

$$V_\bullet \cong \bigoplus_{k \in K} \mathbb{I}_\bullet^{J_k} \quad \text{and} \quad V_\bullet \cong \bigoplus_{m \in M} \mathbb{I}_\bullet^{L_m}$$

where $\{J_k : k \in K\}$ and $\{L_m : m \in M\}$ are two multisets of intervals in \mathbb{N}_0 . Then, there exists a bijection $\pi : K \rightarrow M$ between the indexing sets K and M such that for all $k \in K$, $J_k = L_{\pi(k)}$ as intervals in \mathbb{N}_0 .

Remark. A proof is available under [CSGO13, Theorem 1.3].

The existence of the bijection $\pi : K \rightarrow M$, as denoted in the theorem above, implies that the intervals of $\{J_k\}$ are exactly the intervals of $\{L_m\}$. If the multisets $\{J_k\}$ and $\{L_m\}$ are linearly ordered, then the bijection $\pi : K \rightarrow M$ corresponds to a permutation or re-ordering of the intervals in $\{J_k\}$. This explains why some authors describe interval decompositions as being unique up to permutation.

Furthermore, this implies that the multiset $\{J_k\}$ of intervals that determine the interval decomposition $\bigoplus_k \mathbb{I}_\bullet^{J_k}$ is unique up to persistence isomorphism. This uniqueness result explains why we can use the article *the* when talking about interval decompositions, e.g. we talk of *the* interval decomposition of a persistence module. Note that the persistence isomorphism $V_\bullet \rightarrow \bigoplus_{k \in K} \mathbb{I}_\bullet^{J_k}$ behind the isomorphism relation $V_\bullet \cong \bigoplus_{k \in K} \mathbb{I}_\bullet^{J_k}$ is not

generally unique. This motivates the following terminology.

Definition 2.3.6. The **persistence barcode** $\text{Bar}(V_\bullet)$ of a persistence module $(V_\bullet, \alpha_\bullet)$ is the multiset $\{J_k\}_{k=1}^m$ of intervals in \mathbb{N}_0 such that $\bigoplus_{k=1}^m \mathbb{I}_\bullet^{J_k}$ is an interval decomposition of V_\bullet , i.e. $V_\bullet \cong \bigoplus_{k=1}^m \mathbb{I}_\bullet^{J_k}$. Note that $\text{Bar}(V_\bullet)$ does not exist if V_\bullet does not have an interval decomposition.

The Structure Theorem (Theorem 4.1.1) for finitely-generated modules over a PID R states that the invariant factors of a finitely-generated R -module are uniquely determined up to R -module isomorphism. An equivalent way of saying this is that the invariant factors of an R -module (if an invariant factor decomposition exists) is an invariant of the isomorphism type of R -modules.

In the same vein, the persistence barcode (if it exists) is an invariant of the isomorphism type of persistence modules. Furthermore, the persistence barcode is a concise characterization of the ranks of the vector spaces and structure maps of a persistence module. We state this in more detail below.

Proposition 2.3.7. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module over \mathbb{F} . Assume V_\bullet admits an interval decomposition and let $\text{Bar}(V_\bullet)$ be given by the multiset $\text{Bar}(V_\bullet) = \{J_k\}_{k=1}^m$ of intervals. Then,

$$\text{rank}(V_t) = \text{card}\left\{J \in \text{Bar}(V_\bullet) : t \in J\right\} \quad (\text{E3})$$

$$\text{rank}(\alpha_{s,t}) = \text{card}\left\{J \in \text{Bar}(V_\bullet) : [t, s] = [t, s+1] \subseteq J\right\} \quad (\text{E4})$$

where $\text{rank}(-)$ refers to the rank of \mathbb{F} -vector spaces and \mathbb{F} -linear maps and $\text{card}(-)$ refers to the number of elements in a set, i.e. set cardinality.

Proof. Assume that $V_\bullet \cong \bigoplus_{k=1}^m \mathbb{I}_\bullet^{J_k}$ for some multiset $\{J_k\}_{k=1}^m$ of intervals in \mathbb{N}_0 . Then, there exists a persistence isomorphism $\varphi_\bullet : V_\bullet \rightarrow \bigoplus_k \mathbb{I}_\bullet^{J_k}$ with linear isomorphisms $\varphi_t : V_t \rightarrow \bigoplus_{k=1}^m \mathbb{I}_t^{J_k}$ for $t \in \mathbb{N}_0$.

For (E3): Let $t \in \mathbb{N}_0$. By definition of interval module, we have the following for each $k \in \{1, \dots, m\}$,

$$\text{rank}\left(\mathbb{I}_t^{J_k}\right) = \begin{cases} \text{rank}(\mathbb{F}) = 1 & \text{if } t \in J_k \\ \text{rank}(0) = 0 & \text{otherwise} \end{cases}$$

Since $\varphi_t : V_t \rightarrow W_t$ is a linear isomorphism and by properties of the direct sum of vector spaces:

$$\text{rank}(V_t) = \text{rank}(\varphi_t(V_t)) = \text{rank}\left(\bigoplus_{k=1}^m \mathbb{I}_t^{J_k}\right) = \sum_{k=1}^m \text{rank}\left(\mathbb{I}_t^{J_k}\right) = \text{card}\left\{J \in \text{Bar}(V_\bullet) : t \in J\right\}$$

Therefore, Equation (E3) is satisfied.

For (E4): Let $t, s \in \mathbb{N}_0$ with $t \leq s$. For each $k \in \{1, \dots, m\}$, let $\gamma_{s,t}^k$ refer to the structure map $\gamma_{s,t}^k : \mathbb{I}_t^{J_k} \rightarrow \mathbb{I}_s^{J_k}$ of the interval module $\mathbb{I}_\bullet^{J_k}$. Note that $t, s \in J_k$ if and only if $[t, s] \subseteq J_k$. Therefore,

$$\text{rank}(\gamma_{s,t}^k) = \begin{cases} \text{rank}(\text{id}_{\mathbb{F}}) = 1 & \text{if } [t, s] \subseteq J_k \\ \text{rank}(0) = 0 & \text{otherwise} \end{cases}$$

Let $\gamma_{s,t} : \bigoplus_{k=1}^m \mathbb{I}_t^{J_k} \rightarrow \bigoplus_{k=1}^m \mathbb{I}_s^{J_k}$ be the structure map of $\bigoplus_{k=1}^m \mathbb{I}_\bullet^{J_k}$ on $t \rightarrow s$. By definition of the direct sum of persistence modules, $\gamma_{s,t} = \bigoplus_{k=1}^m \gamma_{s,t}^k$ where $\bigoplus_{k=1}^m \gamma_{s,t}^k$ is the linear map induced by the direct sum operation between \mathbb{F} -vector spaces. Since φ_\bullet must be a persistence morphism, the

composition relation $\gamma_{s,t} \circ \varphi_t = \varphi_s \circ \alpha_{s,t}$ is satisfied. Therefore,

$$\begin{aligned} \text{rank}(\alpha_{s,t}) &= \text{rank}\left((\varphi_s)^{-1} \circ \gamma_{s,t} \circ \varphi_t\right) = \text{rank}(\gamma_{s,t}) = \text{rank}\left(\bigoplus_{k=1}^m \gamma_{s,t}^k\right) \\ &= \sum_{k=1}^m \text{rank}(\gamma_{s,t}^k) = \text{card}\left\{J \in \text{Bar}(V_\bullet) : [t, s] = [t, s+1] \subseteq J\right\} \end{aligned}$$

Therefore, Equation (E4) is also satisfied. \blacksquare

It can be proven that finite-type persistence modules admit interval decompositions. We state this later in Proposition 2.5.14 under Section 2.5 and prove it using an equivalence of categories.

The calculation of the interval decomposition of a given finite-type persistence module $(V_\bullet, \alpha_\bullet)$ requires that both the vector spaces V_t and the structure maps $\alpha_{s,t} : V_t \rightarrow V_s$ for all $t \in \mathbb{N}_0$ be considered, for lack of a better term, *simultaneously* for all indices $t \in \mathbb{N}_0$. In practice, this means that the decomposition of the vector spaces V_t must be *compatible* with the structure maps, in that the collection of vector space isomorphisms involving each V_t must form a persistence morphism. We give an example of this below.

Example 2.3.8. Let $(V_\bullet, \gamma_\bullet)$ be a persistence module over \mathbb{Z}_5 with vector spaces given as follows:

$$V_t = \begin{cases} \mathbb{Z}_5\langle x_1 \rangle & \text{if } t \in [0, 13) \\ \mathbb{Z}_5\langle y_1, y_2 \rangle & \text{if } t \in [13, 21) \\ \mathbb{Z}_5\langle z_1, z_2, z_3 \rangle & \text{if } t \in [21, \infty) \end{cases}$$

For $t \neq 12$ and $t \neq 20$, let $\gamma_t : V_t \rightarrow V_{t+1}$ be the identity map on V_t , which is well-defined since $V_t = V_{t+1}$. Define the structure map $\gamma_t : V_t \rightarrow V_{t+1}$ of V_\bullet for $t = 12$ and $t = 20$ as follows:

$$\begin{array}{ll} \gamma_{12} : V_{12} \rightarrow V_{13} & \gamma_{20} : V_{20} \rightarrow V_{21} \\ x_1 \mapsto y_2 & y_1 \mapsto z_1 \\ & y_2 \mapsto z_3 \end{array}$$

To determine the interval decomposition of C_\bullet , we need to find a decomposition of C_\bullet into persistence submodules such each submodule is isomorphic to an interval module. For each $t \in \mathbb{N}_0$, A_t , B_t , and C_t be vector subspaces of V_t given as follows:

$$A_t = \begin{cases} \mathbb{Z}_5\langle x_1 \rangle & \text{if } t \in [0, 13) \\ \mathbb{Z}_5\langle y_2 \rangle & \text{if } t \in [13, 21) \\ \mathbb{Z}_5\langle z_3 \rangle & \text{if } t \in [21, \infty) \end{cases} \quad B_t = \begin{cases} 0 & \text{if } t \in [0, 13) \\ \mathbb{Z}_5\langle y_1 \rangle & \text{if } t \in [13, 21) \\ \mathbb{Z}_5\langle z_1 \rangle & \text{if } t \in [21, \infty) \end{cases} \quad C_t = \begin{cases} 0 & \text{if } t \in [0, 13) \\ 0 & \text{if } t \in [13, 21) \\ \mathbb{Z}_5\langle z_2 \rangle & \text{if } t \in [21, \infty) \end{cases}$$

Observe that, for all $t \in \mathbb{N}_0$, A_t is a \mathbb{Z}_5 -vector subspace of V_t and that the structure map $\gamma_t : V_t \rightarrow V_{t+1}$ of V_\bullet satisfies $\gamma_t(A_t) \subseteq A_{t+1} \subseteq V_{t+1}$. Therefore, the collection $\{A_t\}_{t \in \mathbb{N}_0}$ of vector spaces determines a persistence submodule A_\bullet of V_\bullet with $A_\bullet(t) = A_t$ and the structure map $A_t \rightarrow A_{t+1}$ defined by restricting the domain and codomain of the structure map $\gamma_t : V_t \rightarrow V_{t+1}$ of V_\bullet for all $t \in \mathbb{N}_0$. Similarly, the collections $\{B_t\}_{t \in \mathbb{N}_0}$ and $\{C_t\}_{t \in \mathbb{N}_0}$ of \mathbb{Z}_5 -vector spaces determine persistence submodules B_\bullet and C_\bullet of V_\bullet respectively.

Since for each $t \in \mathbb{N}_0$, $V_t \cong A_t \oplus B_t \oplus C_t$ as \mathbb{Z}_5 -vector spaces (with the direct sum interpreted as an internal direct sum of vector spaces), we have the following decomposition of V_\bullet as a persistence module:

$$V_\bullet \cong A_\bullet \oplus B_\bullet \oplus C_\bullet$$

with the persistence isomorphism $i_\bullet : A_\bullet \oplus B_\bullet \oplus C_\bullet \rightarrow V_\bullet$ given by the inclusion maps $i_t : A_t \oplus B_t \oplus C_t \hookrightarrow V_t$ for all $t \in \mathbb{N}_0$. This decomposition of V_\bullet is illustrated in the following diagram:

$$\begin{array}{ccccccccccc}
& (t=0) & \cdots & (t=12) & (t=13) & \cdots & (t=20) & (t=21) & \cdots \\
V_\bullet : & \mathbb{Z}_5 \langle x_1 \rangle & \longrightarrow \cdots \longrightarrow & \mathbb{Z}_5 \langle x_1 \rangle & \longrightarrow \mathbb{Z}_5 \langle y_1, y_2 \rangle & \longrightarrow \cdots \longrightarrow & \mathbb{Z}_5 \langle z_1, z_2, z_3 \rangle & \longrightarrow \mathbb{Z}_5 \langle z_1, z_2, z_3 \rangle & \longrightarrow \cdots \\
\parallel & \parallel & & \parallel & \parallel & & \parallel & \parallel & \\
A_\bullet : & \mathbb{Z}_5 \langle x_1 \rangle & \longrightarrow \cdots \longrightarrow & \mathbb{Z}_5 \langle x_1 \rangle & \longrightarrow \mathbb{Z}_5 \langle y_2 \rangle & \longrightarrow \cdots \longrightarrow & \mathbb{Z}_5 \langle y_2 \rangle & \longrightarrow \mathbb{Z}_5 \langle z_3 \rangle & \longrightarrow \cdots \\
\oplus & & & \oplus & & & \oplus & \oplus & \\
B_\bullet : & & & & \mathbb{Z}_5 \langle y_1 \rangle & \longrightarrow \cdots \longrightarrow & \mathbb{Z}_5 \langle y_1 \rangle & \longrightarrow \mathbb{Z}_5 \langle z_1 \rangle & \longrightarrow \cdots \\
\oplus & & & & & & & \oplus & \\
C_\bullet : & & & & & & & \mathbb{Z}_5 \langle z_2 \rangle & \longrightarrow \cdots
\end{array}$$

From the diagram above, we can conclude that the persistence submodules A_\bullet , B_\bullet , and C_\bullet of V_\bullet are isomorphic to interval modules as follows:

For $A_\bullet \cong \mathbb{I}_\bullet^{[0, \infty)}$: Let $A_\bullet = (A_\bullet, \alpha_\bullet)$, i.e. denote the structure maps of A_\bullet by $\alpha_t : A_t \rightarrow A_{t+1}$. Recall that $\alpha_t(a) = \gamma_t(a)$ for all $a \in A_t$ and for all $t \in \mathbb{N}_0$. Observe that for all $t \in \mathbb{N}_0 = [0, \infty)$, $\text{rank}(A_t) = 1$ and $\text{rank}(\alpha_t) = 1$. Then, $A_\bullet \cong \mathbb{I}_\bullet^{[0, \infty)}$ by Lemma 2.3.2, relative to characterization by Equation (E2).

Alternatively, we can construct the persistence isomorphism $\varphi_\bullet^A : A_\bullet \rightarrow \mathbb{I}_\bullet^{[0, \infty)}$ explicitly as follows: Denote the multiplicative identity of \mathbb{Z}_5 as 1. Then, for each $t \in \mathbb{N}_0$, define the linear map $\varphi_t^A : A_t \rightarrow \mathbb{I}_t^{[0, \infty)}$ as follows:

If $t \in [0, 13)$: Let $\varphi_t^A : A_t = \mathbb{Z}_5 \langle x_1 \rangle \rightarrow \mathbb{I}_t^{[0, \infty)} = \mathbb{Z}_5$ be given by $x_1 \mapsto 1$.

If $t \in [13, 21)$: Let $\varphi_t^A : A_t = \mathbb{Z}_5 \langle y_2 \rangle \rightarrow \mathbb{I}_t^{[0, \infty)} = \mathbb{Z}_5$ be given by $y_2 \mapsto 1$.

If $t \in [21, \infty)$: Let $\varphi_t^A : A_t = \mathbb{Z}_5 \langle z_3 \rangle \rightarrow \mathbb{I}_t^{[0, \infty)}$ be given by $z_3 \mapsto 1$.

It should be straightforward to check that the linear maps $\varphi_t^A : A_t \rightarrow \mathbb{I}_t^{[0, \infty)}$ commute with the structure maps $\alpha_t : A_t \rightarrow A_{t+1}$ of A_\bullet and those of $\mathbb{I}_\bullet^{[0, \infty)}$. Therefore, φ_\bullet^A by $\varphi_\bullet^A = \{\varphi_t^A\}_{t \in \mathbb{N}_0}$ is a well-defined persistence morphism. Since each φ_t^A is a vector space isomorphism, φ_\bullet^A is a persistence isomorphism.

For $B_\bullet \cong \mathbb{I}_\bullet^{[13, \infty)}$: Let $B_\bullet = (B_\bullet, \beta_\bullet)$, i.e. denote the structure maps of B_\bullet by $\beta_t : B_t \rightarrow B_{t+1}$. Observe that if $t \in [13, \infty)$, $\text{rank}(B_t) = 1$ and $\text{rank}(\beta_t) = 1$. If $t \in [0, 12)$, B_t is trivial and $\text{rank}(B_t) = 0$ and $\text{rank}(\beta_t) = 0$. Then, $B_\bullet \cong \mathbb{I}_\bullet^{[13, \infty)}$ by Lemma 2.3.2.

The persistence isomorphism $\varphi_\bullet^B : B_\bullet \rightarrow \mathbb{I}_\bullet^{[13, \infty)}$ can be constructed similarly as with the case of φ_\bullet^A , i.e.

If $t \in [0, 13)$: B_t is trivial and $\varphi_t^B : B_t \rightarrow \mathbb{I}_t^{[13, \infty)} = 0$ can only be the trivial map.

If $t \in [13, 21)$: The map $\varphi_t^B : B_t = \mathbb{Z}_5 \langle y_1 \rangle \rightarrow \mathbb{I}_t^{[13, \infty)} = \mathbb{Z}_5$ is given by $y_1 \mapsto 1$.

If $t \in [21, \infty)$: The map $\varphi_t^B : B_t = \mathbb{Z}_5 \langle z_1 \rangle \rightarrow \mathbb{I}_t^{[13, \infty)} = \mathbb{Z}_5$ is given by $z_1 \mapsto 1$.

For $C_\bullet \cong \mathbb{I}_\bullet^{[21, \infty)}$: Let $C_\bullet = (C_\bullet, \psi_\bullet)$. Note that if $t \in [21, \infty)$, $\text{rank}(C_t) = 1$ and $\text{rank}(\psi_t : C_t \rightarrow C_{t+1}) = 1$. If $t \in [0, 21)$, $\text{rank}(C_t) = 0$ and $\text{rank}(\psi_t : C_t \rightarrow C_{t+1}) = 0$. By Lemma 2.3.2, $C_\bullet \cong \mathbb{I}_\bullet^{[21, \infty)}$.

The persistence isomorphism $\varphi_\bullet^C : C_\bullet \rightarrow \mathbb{I}_\bullet^{[21, \infty)}$ is given as follows: If $t \in [21, \infty)$, $\varphi_t^C : C_t = \mathbb{Z}_5 \langle z_2 \rangle \rightarrow \mathbb{I}_t^{[21, \infty)} = \mathbb{Z}_5$ is given by $z_2 \mapsto 1$. If $t \in [0, 21)$, $C_t = 0$ and φ_t^C can only be the trivial map.

Therefore, we have the following decomposition for V_\bullet :

$$V_\bullet \cong A_\bullet \oplus B_\bullet \oplus C_\bullet \cong \mathbb{I}_\bullet^{[0,\infty)} \oplus \mathbb{I}_\bullet^{[13,\infty)} \oplus \mathbb{I}_\bullet^{[21,\infty)}$$

Consequently, the persistence barcode $\text{Bar}(V_\bullet)$ of V_\bullet is given by $\text{Bar}(V_\bullet) = \{[0, \infty), [13, \infty), [21, \infty)\}$.

Observe that, in the above example, the persistence barcode of V_\bullet corresponds to a collection of bases for each V_t that is compatible with the structure maps. In particular, Example 2.3.8 is written such that elements of the bases $\{x_1, x_2, x_3\}$, $\{y_1, y_2\}$, and $\{z_1, z_2, z_3\}$ for V_t with $t \in [0, 13)$, $t \in [13, 21)$, and $t \in [21, \infty)$ respectively are mapped either to $1 \in \mathbb{Z}_5$ or $0 \in \mathbb{Z}_5$ in the given interval decomposition for V_\bullet .

Unfortunately, this nice correspondence does not generally apply to arbitrary finite-type persistence modules. In the example below, we define the vector spaces of a persistence module using some set of indeterminates but show an interval decomposition of said persistence modules that requires a change of basis for the vector spaces involved.

Example 2.3.9. Let $(W_\bullet, \gamma_\bullet)$ be a persistence module over \mathbb{Z}_3 with vector spaces given as follows:

$$W_t = \begin{cases} \mathbb{Z}_3\langle a_1 \rangle & \text{if } t = 0 \\ \mathbb{Z}_3\langle b_1, b_2 \rangle & \text{if } t = 1 \\ \mathbb{Z}_3\langle c_1, c_2 \rangle & \text{if } t = 2 \\ \mathbb{Z}_3\langle d_1 \rangle & \text{if } t \geq 3 \end{cases}$$

Let the structure maps of W_\bullet of the form $\gamma_t : W_t \rightarrow W_{t+1}$ for $t \in \mathbb{N}_0$ be given by

$$\begin{array}{llll} \gamma_0 : W_0 = \mathbb{Z}_3\langle a_1 \rangle \rightarrow W_1 & \gamma_1 : W_1 \rightarrow W_2 & \gamma_2 : W_2 \rightarrow W_3 & \gamma_t : W_t \rightarrow W_{t+1} \text{ for all } t \geq 3 \\ a_1 \mapsto 2b_1 + b_2 & b_1 \mapsto c_1 + c_2 & c_1 \mapsto 0 & d_1 \mapsto d_1 \\ & b_2 \mapsto c_1 + 2c_2 & c_2 \mapsto d_1 & \end{array}$$

To determine the interval decomposition of W_\bullet , we must first find a decomposition of W_\bullet into persistence submodules such that each submodule is isomorphic to an interval module. For each $t \in \mathbb{N}_0$, define the \mathbb{Z}_3 -vector spaces U_t and V_t as follows:

$$A_t = \begin{cases} \mathbb{Z}_3\langle a_1 \rangle & \text{if } t = 0 \\ \mathbb{Z}_3\langle 2b_1 + b_2 \rangle & \text{if } t = 1 \\ \mathbb{Z}_3\langle c_2 \rangle & \text{if } t = 2 \\ \mathbb{Z}_3\langle d_1 \rangle & \text{if } t \geq 3 \end{cases} \quad \text{and} \quad B_t = \begin{cases} 0 & \text{if } t = 0 \\ \mathbb{Z}_3\langle b_1 + b_2 \rangle & \text{if } t = 1 \\ \mathbb{Z}_3\langle 2c_1 \rangle & \text{if } t = 2 \\ 0 & \text{if } t \geq 3 \end{cases}$$

For the collections $\{A_t\}_{t \in \mathbb{N}_0}$ and $\{B_t\}_{t \in \mathbb{N}_0}$ to determine persistence submodules $A_\bullet = (A_\bullet, \alpha_\bullet)$ and $B_\bullet = (B_\bullet, \beta_\bullet)$ of W_\bullet respectively, we must show that linear maps $\alpha_t : A_t \rightarrow A_{t+1}$ and $\beta_t : B_t \rightarrow B_{t+1}$ obtained by appropriately restricting the domain and codomain of $\gamma_t : W_t \rightarrow W_{t+1}$ are well-defined. This is shown by the following calculations:

For $\{A_t\}_{t \in \mathbb{N}_0}$: It suffices to consider the images of the basis elements a_1 , $2b_1 + b_2$, and c_2 of A_0 , A_1 , and A_2 respectively under the appropriate structure map γ_t of W_\bullet :

$$\begin{array}{llll} \gamma_0(a_1) & = & \cdots & = 2b_1 + b_2 \\ \gamma_1(2b_1 + b_2) & = 2(c_1 + c_2) + (c_1 + 2c_2) = 2c_1 + 2c_2 + c_1 + 2c_2 = & c_2 & \\ \gamma_2(c_2) & = & \cdots & = d_1 \end{array}$$

Note that for all $t \geq 3$, γ_3 is the identity map on $V_t = \mathbb{Z}_3\langle d_1 \rangle$. Then, for all $t \in \mathbb{N}_0$, $\gamma_t(A_t) \subseteq A_{t+1}$ and the linear map $\alpha_t : A_t \rightarrow A_{t+1}$ obtained by restricting $\gamma_t : W_t \rightarrow W_{t+1}$ is well-defined. Therefore, $A_\bullet = (A_\bullet, \alpha_\bullet)$ by $A_\bullet(t) = A_t$ and $A_\bullet(t \rightarrow t+1) = \alpha_t$ is a persistence submodule of W_\bullet .

For $\{B_t\}_{t \in \mathbb{N}_0}$: Note that for $t \neq 1, 2$, B_t is trivial. It suffices to consider the image of $b_1 + b_2 \in B_1$ and $2c_1 \in B_2$ under the appropriate structure map γ_t of W_\bullet :

$$\begin{aligned}\gamma_1(b_1 + b_2) &= (c_1 + c_2) + (c_1 + 2c_2) = 2c_1 + 3c_2 = 2c_1 \\ \gamma_2(2c_1) &= 2(0) = 0\end{aligned}$$

Then, $\gamma_1(B_1) \subseteq \mathbb{Z}_3\langle 2c_1 \rangle = B_2$ and $\gamma_2(B_2) \subseteq \mathbb{Z}_3\langle 0 \rangle = B_3$, and the linear maps $\beta_1 : B_1 \rightarrow B_2$ and $\beta_2 : B_1 \rightarrow B_2$ obtained by restricting $\gamma_1 : W_1 \rightarrow W_2$ and $\gamma_2 : W_2 \rightarrow W_3$ are well-defined. For $t \neq 1, 2$, define $\beta_t : B_t \rightarrow B_{t+1}$ to be the trivial map. Therefore, $B_\bullet = (B_\bullet, \beta_\bullet)$ by $B_\bullet(t) = B_t$ and $B_\bullet(t \rightarrow t+1) = \beta_t$ is a persistence submodule of W_\bullet .

Observe that for $t = 0$ and for $t \geq 3$, $A_t \oplus B_t \cong A_t = V_t$ since B_t is trivial. For the case of $t = 2$, we have that $\{2b_1 + b_2, b_1 + b_2\}$ is a basis for $W_2 = \mathbb{Z}_3\langle b_1, b_2 \rangle$ by the following calculation:

$$\begin{aligned}(1)(2b_1 + b_2) + (2)(b_1 + b_2) &= 2b_1 + b_2 + 2b_1 + 2b_2 = 4b_1 + 3b_2 = b_1 \\ (2)(2b_1 + b_2) + (2)(b_1 + b_2) &= 4b_1 + 2b_2 + 2b_1 + 2b_2 = 6b_1 + 4b_2 = b_2\end{aligned}$$

Similarly, $\{2c_1, c_2\}$ is a basis for $W_3 = \mathbb{Z}_3\langle c_1, c_2 \rangle$ since $2(2c_1) = 4c_1 = c_1$. Therefore, $W_t = A_t \oplus B_t$ for all $t \in \mathbb{N}_0$ (taken as an internal direct sum) and we have the following persistence isomorphism relation:

$$W_\bullet \cong A_\bullet \oplus B_\bullet$$

This is illustrated in the diagram below:

$$\begin{array}{ccccccccc} & (t=0) & & (t=1) & & (t=2) & & (t=3) & & \cdots \\ W_\bullet : & \mathbb{Z}_5\langle a_1 \rangle & \longrightarrow & \mathbb{Z}_5\langle b_1, b_2 \rangle & \longrightarrow & \mathbb{Z}_5\langle c_1, c_2 \rangle & \longrightarrow & \mathbb{Z}_5\langle d_1 \rangle & \longrightarrow & \cdots \\ \parallel & \parallel & & \parallel & & \parallel & & \parallel & & \\ A_\bullet : & \mathbb{Z}_5\langle a_1 \rangle & \longrightarrow & \mathbb{Z}_5\langle 2b_1 + b_2 \rangle & \longrightarrow & \mathbb{Z}_5\langle c_2 \rangle & \longrightarrow & \mathbb{Z}_5\langle d_1 \rangle & \longrightarrow & \cdots \\ \oplus & \oplus & & \oplus & & \oplus & & \oplus & & \\ B_\bullet : & 0 & \longrightarrow & \mathbb{Z}_3\langle b_1 + b_2 \rangle & \longrightarrow & \mathbb{Z}_3\langle 2c_1 \rangle & \longrightarrow & 0 & \longrightarrow & \cdots \end{array}$$

The submodules A_\bullet and B_\bullet of W_\bullet are also isomorphic to interval modules, as described below:

For $A_\bullet \cong \mathbb{I}_\bullet^{[0, \infty)}$: The persistence isomorphism $\varphi_\bullet^A : A_\bullet \rightarrow \mathbb{I}_\bullet^{[0, \infty)}$ is given as follows:

For $t = 0$: The map $\varphi_0^A : A_0 \rightarrow \mathbb{I}_0^{[0, \infty)} = \mathbb{Z}_3$ is given by $a_1 \mapsto 1 \in \mathbb{Z}_3$.

For $t = 1$: The map $\varphi_1^A : A_1 \rightarrow \mathbb{I}_1^{[0, \infty)} = \mathbb{Z}_3$ is given by $2b_1 + b_2 \mapsto 1 \in \mathbb{Z}_3$.

For $t = 2$: The map $\varphi_2^A : A_2 \rightarrow \mathbb{I}_2^{[0, \infty)} = \mathbb{Z}_3$ is given by $c_2 \mapsto 1 \in \mathbb{Z}_3$.

For $t \geq 2$: The map $\varphi_t^A : A_t \rightarrow \mathbb{I}_t^{[0, \infty)} = \mathbb{Z}_3$ is given by $d_1 \mapsto 1 \in \mathbb{Z}_3$.

We can confirm that the maps $\varphi_t^A : A_t \rightarrow \mathbb{I}_t^{[0, \infty)}$ commutes with the structure maps $\alpha_t : A_t \rightarrow A_{t+1}$ of A_\bullet and those of $\mathbb{I}_\bullet^{[0, \infty)}$. We can also conclude that $A_\bullet \cong \mathbb{I}_\bullet^{[0, \infty)}$ using Lemma 2.3.2 since $\text{rank}(A_t) = 1$ and $\text{rank}(\alpha_t : A_t \rightarrow A_{t+1}) = 1$ for all $t \in \mathbb{N}_0$.

For $B_\bullet \cong \mathbb{I}_\bullet^{[1, 3)}$: The persistence isomorphism $\varphi_\bullet^B : B_\bullet \rightarrow \mathbb{I}_\bullet^{[1, 3)}$ is given as follows:

For $t = 0$: B_0 is trivial and $\varphi_0^B : B_0 \rightarrow \mathbb{I}_\bullet^{[1,3)} = 0$ is necessarily the trivial map.

For $t = 1$: The map $\varphi_1^B : B_1 \rightarrow \mathbb{I}_1^{[1,3)} = \mathbb{Z}_3$ is given by $b_1 + b_2 \mapsto 1 \in \mathbb{Z}_3$.

For $t = 2$: The map $\varphi_2^B : B_2 \rightarrow \mathbb{I}_2^{[1,3)} = \mathbb{Z}_3$ is given by $2c_1 \mapsto 1 \in \mathbb{Z}_3$.

For $t \geq 2$: B_t and $\mathbb{I}_t^{[1,3)}$ are both trivial and φ_t^B is necessarily the trivial map.

Observe that the maps $\{\varphi_t^B\}_{t \in \mathbb{N}_0}$ commute with the structure maps of B_\bullet and those of $\mathbb{I}_\bullet^{[1,3)}$. In particular, we have the following for the case of $t = 1$ and $t = 2$:

$$\begin{aligned} (\varphi_2^B \circ \beta_1)(b_1 + b_2) &= \varphi_2^B(2c_1) = 1 = \text{id}_{\mathbb{Z}_3}(1) = (\text{id}_{\mathbb{Z}_3} \circ \varphi_1^B)(b_1 + b_2) \\ (\varphi_3^B \circ \beta_2)(2c_1) &= \varphi_3^B(0) = 0 = \text{id}_{\mathbb{Z}_3}(0) = (\text{id}_{\mathbb{Z}_3} \circ \varphi_2^B)(2c_1) \end{aligned}$$

Alternatively, we can use Lemma 2.3.2 (relative to its characterization by Equation (E2)) to determine that $B_\bullet \cong \mathbb{I}_\bullet^{[1,3)}$ since we have the following:

$$\text{rank}(B_t) = \begin{cases} 1 & \text{if } t \in [1, 3) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \text{rank}(\beta_t) = \begin{cases} 1 & \text{if } t \in [1, 3 - 1) = 1 \\ 0 & \text{otherwise} \end{cases}$$

Therefore, we have the following persistence isomorphism relation:

$$W_\bullet \cong A_\bullet \oplus B_\bullet \cong \mathbb{I}_\bullet^{[0,\infty)} \oplus \mathbb{I}_\bullet^{[1,3)}$$

and the persistence barcode of W_\bullet is given by $\text{Bar}(W_\bullet) = \{[0, \infty), [1, 3)\}$.

In Chapter 4, we discuss a method of calculating the interval decomposition of finite-type persistence modules that are the n^{th} chain homology of a chain complex of finite-type persistence modules.

Section 2.4. The Category of Graded Modules over Polynomial Rings

Let R be a commutative ring with identity $1_R \in R$. One of the key results in persistence theory is the equivalence between the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules and the category $\mathbf{GrMod}_{R[x]}$ of graded $R[x]$ -modules, discussed in Section 2.5.

However, we have found that definitions and descriptions for graded modules, particularly those involving the *category* of graded modules, seem to be sparse in introductory-level abstract algebra texts, and that rigorous definitions often only appear in more specialized fields such as commutative algebra and homological algebra. To avoid confusion, we identify a number of definitions, notation, and results involving graded $R[x]$ -modules in this section that are used throughout the paper. We use the following texts as our primary references.

1. *Graded Syzygies* [Pee11] by Irena Peeva.
2. *Methods of Graded Rings* [NO04] by Constantin Năstăsescu and Freddy Van Oystaeyen.

To start, most of the literature regarding graded modules and persistence modules use the term *action* to refer to the scalar multiplication operation on a module. We state a definition of *action* below.

Definition 2.4.1. The **action** of R on an R -module M is a biadditive group action $\cdot : R \times M \rightarrow M$ that defines the scalar multiplication operation on M .

Note that the requirements that the *action* $\cdot : R \times M \rightarrow M$ be a group action and be biadditive satisfy the usual conditions for scalar multiplication on an R -module. We state this in more detail below:

1. The condition that $\cdot : R \times M \rightarrow M$ be a *group action* requires that \cdot must satisfy the identity and compatibility axioms with respect to the multiplication operation $\cdot_R : R \times R \rightarrow R$ on R (as a ring), i.e. for all $m \in M$ and $r, s \in R$,

$$1_R \cdot m = m \quad \text{and} \quad r \cdot (s \cdot m) = (r \cdot_R s) \cdot m$$

2. The condition that $\cdot : R \times M \rightarrow M$ be *biadditive* requires that \cdot be linear on the first and second arguments with respect to addition $+_R : R \times R \rightarrow R$ on R (as a ring) and to addition $+_M : M \times M \rightarrow M$ on M (as an abelian group), i.e. for all $r, s \in R$ and $m, n \in M$,

$$\begin{aligned} (r +_R s) \cdot m &= (r \cdot m) +_M (s \cdot m) \\ r \cdot (m +_M n) &= (r \cdot m) +_M (r \cdot n) \end{aligned}$$

The two equalities above are sometimes called the distributivity properties of scalar multiplication (of modules).

Next, we provide definitions for the standard grading on $R[x]$ taken from [Pee11, Section 1].

Definition 2.4.2. The **standard grading** on the polynomial ring $R[x]$ is given by polynomial degree.

- i. A nonzero element $f \in R[x]$ is **homogeneous** if $f = ax^t$ for some nonzero $a \in R$ and $t \in \mathbb{N}_0$, i.e. f is a monomial. In this case, we say that $f = ax^t$ is **homogeneous of degree t** and write $\deg_h(f) = t$.
- ii. For all $t \in \mathbb{N}_0$ with $t \geq 1$, define Rx^t as the subring of $R[x]$ consisting of all monomials of degree t and the zero polynomial, i.e. $Rx^t = \{rx^t : r \in R\}$. Define $Rx^0 := R$.
- iii. Under the standard grading, the **homogeneous component of $R[x]$ of degree $t \in \mathbb{N}_0$** is the subring Rx^t of $R[x]$. Observe that $R[x] = \bigoplus_{t \in \mathbb{N}_0} Rx^t$ as abelian groups with \oplus taken as an internal direct sum.

In this paper: when $R[x]$ is viewed as a graded ring, we always equip $R[x]$ with the standard grading.

Remark. For a general definition for graded rings (i.e. not specific to $R[x]$ under the standard grading), see [Pee11; NO04]. Note that the grading on a ring R can be defined relative to any abelian group G . Here, the degrees of homogeneous elements take values in G and say that R is G -graded.

In our case, we consider the standard grading on $R[x]$ to be an \mathbb{N}_0 -grading on $R[x]$ since we only expect the degrees of homogeneous elements of $R[x]$ to take values in \mathbb{N}_0 . We point this out since, in some references, e.g. [Web85; Mar; NO04], $R[x]$ is described to be \mathbb{Z} -graded with the homogeneous components of $R[x]$ of degree $t \in \mathbb{Z}$ with $t < 0$ being trivial abelian groups. Here, $R[x]$ is often seen as a subring of the Laurent polynomial ring $R[x, x^{-1}]$.

With $R[x]$ equipped with the standard grading, observe that we denote the degree of $f \in R[x]$ by $\deg_h(f)$, as opposed to $\deg(f)$. Note that, conventionally, $\deg(f)$ denotes the degree of a polynomial $f \in R[x]$, disregarding the equipped grading on $R[x]$. More specifically, for all nonzero $f \in R[x]$, we have that:

$$\deg_h(f) = \begin{cases} \deg(f) & \text{if } f \text{ is homogeneous, i.e. } f \text{ is a monomial} \\ \text{undefined} & \text{otherwise} \end{cases}$$

We identify two advantages to this approach:

1. Writing $\deg_h(f)$ emphasizes that $f \in R[x]$ is to be considered as an element of a graded ring. This additional notation suggests that $\deg_h(f)$ and $\deg(f)$, while related or similar, are two distinct properties.

We have found this observation to be important when first learning graded module theory especially since, relative to a more general definition of graded ring, $R[x]$ may equipped with a grading other than the standard grading. For example, we may equip $R[x]$ with the trivial grading wherein all elements of $R[x]$ are defined to be homogeneous of degree $0 \in \mathbb{N}_0$. In this case, we have that $\deg_h(3x^2) = 0$ despite $\deg(3x^2) = 2$. While we do not consider cases like these in this paper, it is helpful to know this distinction for other applications.

2. Writing a relation in terms of $\deg_h(-)$ emphasizes that the relation is generally only valid when the arguments are homogeneous elements of a graded ring. This will be relevant later in Section 4.3 where we require and expect nonzero elements to be homogeneous.

The notion of graded rings gives rise to the notion of graded modules. In this paper, since we only consider the graded structure of modules over $R[x]$ equipped with the standard grading, we use a definition that is specific to this type of modules. We state this below.

Definition 2.4.3. An $R[x]$ -module M is a **graded $R[x]$ -module** or **graded module** if there exists an \mathbb{N}_0 -indexed family $\{M_t\}_{t \in \mathbb{N}_0}$ of additive subgroups M_t of M such that two conditions are satisfied:

- i. The underlying abelian group M decomposes into $M = \bigoplus_{t \in \mathbb{N}_0} M_t$ with equality and (internal) direct sum taken to be of abelian groups (seen as additive groups).
- ii. For all $s, t \in \mathbb{N}_0$, $Rx^s \cdot M_t \subseteq M_{s+t}$, i.e. the action of $R[x]$ on M respects the grading on $R[x]$ and on M .

For all $t \in \mathbb{N}_0$, we call the additive subgroup $M_t \subseteq M$ the **homogeneous component of M of degree t** . We call the direct sum $\bigoplus_{t \in \mathbb{N}_0} M_t$ the **homogeneous decomposition of M** . A nonzero element $f \in M$ is **homogeneous** if there exists $t \in \mathbb{N}_0$ such that $f \in M_t$. In this case, we say that f is **homogeneous of degree t** and write $\deg_h(f) = t$. A subset $A \subseteq M$ is **homogeneous** if each element $a \in A$ is homogeneous (not necessarily of the same degree). A $R[x]$ -submodule N of M is called a **graded submodule** if N is also a graded $R[x]$ -module.

Remarks. (1) The statement $Rx^s \cdot M_t$ is specific to the ring being $R[x]$ equipped with the standard grading.

This condition can be written more generally as $R_s M_t \subseteq M_{s+t}$ where R_s is the homogeneous component of degree $s \in \mathbb{N}_0$ of the graded ring $R = \bigoplus_{t \in \mathbb{N}_0} R_t$.

- (2) The remarks about G -grading of rings below Definition 2.4.2 also apply to graded modules. More specifically, if the graded ring is considered to be G -graded, then the graded modules over said ring is also G -graded. Since we consider $R[x]$ to be \mathbb{N}_0 -graded in this paper, this means that our graded $R[x]$ -modules are also \mathbb{N}_0 -graded.

By the definition above, the homogeneous decomposition of a graded $R[x]$ -module M refers to a direct sum of abelian groups. In practice, we usually define graded $R[x]$ -modules at the level of R -modules. That is, we define a graded $R[x]$ -module using two collections of objects:

1. A family $\{M_t\}_{t \in \mathbb{N}_0}$ of R -modules such that $M_t \cap M_s = \{0\}$ whenever $t \neq s$, i.e. each R -module M_t is seen as having nonzero elements distinct from the other R -modules in said family. Note that we usually force the condition $M_t \cap M_s = \{0\}$ by using the family $\{M_t x^t\}_{t \in \mathbb{N}_0}$ constructed from $\{M_t\}$.

We then define M to be the (internal) direct sum $M = \bigoplus_{t \in \mathbb{N}_0} M_t$ of R -modules. This induces an R -module structure on M . Observe that the nonzero elements in one summand M_t cannot be generated using sums of R -multiples of elements from other summands, i.e. M_s with $t \neq s$. If R is a field \mathbb{F} , then M is a (graded) \mathbb{F} -vector space. As a sidenote, a *graded \mathbb{F} -vector space* is an \mathbb{F} -vector space (i.e. \mathbb{F} -module) that satisfies Condition (i) of Definition 2.4.3. Note that all graded modules over $\mathbb{F}[x]$ are graded \mathbb{F} -vector spaces. However, not all graded \mathbb{F} -vector spaces form graded modules, e.g. the action of $\mathbb{F}[x]$ may not be defined.

2. A collection of assignments of the form $x \cdot m_t := m_{t+1}$, defining the product between $x \in R[x]$ of degree 1 and each element $m_t \in M_t$ of degree $t \in \mathbb{N}_0$ to some element $m_{t+1} \in M_{t+1}$ of degree $t+1$.

These assignments are then extended linearly over the action of R on each M_t , the ring operation on $R[x]$, and the addition operation on M as an R -module. If the action of $R[x]$ is well-defined, then M becomes an $R[x]$ -module.

The condition that $M_t \cap M_s = \{0\}$ whenever $t \neq s$ satisfies the condition that M is a direct sum $\bigoplus_{t \in \mathbb{N}_0} M_t$, i.e. each element $m \in M$ can be represented *uniquely* as a sum of homogeneous elements. The set of assignments in the form $x \cdot m_t := m_{t+1}$ forces the action of $R[x]$ on M to satisfy $Rx \cdot M_t \subseteq M_{t+1}$. If the assignments result in a well-defined action of $R[x]$ on M , then the condition that $Rx^s \cdot M_t \subseteq M_{s+t}$ is satisfied. Therefore, the resulting $R[x]$ -module M is also a graded $R[x]$ -module. We provide an example of this construction below.

Example 2.4.4. For each $t \in \mathbb{N}_0$, define the \mathbb{Z} -module M_t by $M_t := \mathbb{Z}\langle a, b \rangle x^t = \{k_1 a x^t + k_2 b x^t : k_1, k_2 \in \mathbb{Z}\}$ where $\mathbb{Z}\langle a, b \rangle$ is the free \mathbb{Z} -module with basis $\{a, b\}$. Then, the direct sum below results in an \mathbb{Z} -module:

$$M = \bigoplus_{t \in \mathbb{N}_0} M_t = \bigoplus_{t \in \mathbb{N}_0} \mathbb{Z}\langle a, b \rangle x^t = \left\{ \sum_{t=0}^{\infty} (k_t^a a x^t + k_t^b b x^t) : \text{only finitely many } k_t^a, k_t^b \in \mathbb{Z} \text{ are nonzero} \right\}$$

Define the action of $\mathbb{Z}[x]$ on M , which we denote as $\star : \mathbb{Z}[x] \times M \rightarrow M$ in this example for clarity, by linearly extending the following set of assignments on M :

$$x \star a x^t := 2a x^{t+1} \quad \text{and} \quad x \star b x^t := 3b x^{t+1} \quad \text{for all } t \in \mathbb{N}_0$$

Here, linearly extending means that we use the compatibility and biadditivity axioms, as discussed under Definition 2.4.1, to calculate assignments of $\star : \mathbb{Z}[x] \times M \rightarrow M$ not of the form $x \star a x^t$ and $x \star b x^t$. Below, we provide examples of this linear extension.

In the equality $\boxed{=}$, we extend the assignments $x \star ax^t$ and $x \star bx^t$ over ...

$$\begin{aligned}
x \cdot (ax^3 + bx^5) &\boxed{=} (x \star ax^3) + (x \star bx^5) = 2ax^4 + 3bx^6 && \dots \text{ addition } + \text{ on } M \text{ as a } \mathbb{Z}\text{-module.} \\
x \cdot 3ax^{11} &\boxed{=} 3 \cdot (x \star ax^{11}) = 3 \cdot (2ax^{12}) = 6ax^{12} && \dots \text{ the action } \cdot \text{ of } \mathbb{Z} \text{ on } M_{11} = \mathbb{Q}\langle a, b \rangle x^{11}. \\
x^2 \star ax^3 &= (x \cdot x) \star ax^3 \boxed{=} x \star (x \star ax^3) = x \star (2ax^4) = 4ax^5 && \dots \text{ ring operation } \cdot \text{ on } \mathbb{Z}[x]. \\
(x + x^2) \star b &\boxed{=} (x \star bx^0) + (x^2 \star bx^0) = 3bx + 9bx^2 && \dots \text{ addition } + \text{ on } \mathbb{Z}[x] \text{ as a ring.}
\end{aligned}$$

We claim that this describes a well-defined action of $\mathbb{Z}[x]$ on M , which makes M a $\mathbb{Z}[x]$ -module. We can verify that the decomposition $M = \bigoplus_{t \in \mathbb{N}_0} \mathbb{Z}\langle a, b \rangle x^t$, along with the action \star on M , satisfies both conditions in Definition 2.4.3. This determines M to be a graded $\mathbb{Z}[x]$ -module with homogeneous component of degree $t \in \mathbb{N}_0$ given by $\mathbb{Z}\langle a, b \rangle x^t$. We list some examples of elements of M and determine their homogeneity or degree:

1. $m_1 = 2ax + 3bx^2$ is not a homogeneous element of M , i.e. $\deg_h(2ax + 3bx^2)$ is undefined. Note that it may feel natural to extend $\deg(-) : R[x] \rightarrow \mathbb{N}_0$ using powers of x , e.g. we may consider that $\deg(2ax + 3bx^2) = 2$. However, the distinction between $\deg_h(-)$ and $\deg(-)$ in the case of $R[x]$ also holds in the case of M .
2. $m_2 = 4ax^5 - 3bx^5 = (4a - 3b)x^5$ is a homogeneous element of M with $\deg_h(m_2) = 5$.
3. $m_3 = -7a$ is a homogeneous element with $\deg_h(m_3) = 0$.

Note that, by construction, the homogeneous elements of M are exactly monomials of the form fx^t with nonzero $f \in \mathbb{Z}\langle a, b \rangle$ and $t \in \mathbb{N}_0$.

We want to point out that the elements of an $R[x]$ -module M are generally not polynomials in x (unlike the example above). Consequently, if the $R[x]$ -module M is graded, the degrees of the homogeneous elements of M are not generally immediately apparent from the notation for said elements.

It would be very helpful (calculation-wise) if we can, without loss of generality, assume that the notation of the elements of a graded $R[x]$ -module also carry information about the degree. To formalize this, we need a notion of similarity or equivalence between graded $R[x]$ -modules. First, we provide a definition for the family of homomorphisms between graded modules, which in turn induces a definition for isomorphisms between graded modules.

Definition 2.4.5. Let $\varphi : M \rightarrow N$ be an $R[x]$ -module homomorphism between graded $R[x]$ -modules M and N . Let M_t and N_t refer to the homogeneous component of M and N of degree $t \in \mathbb{N}_0$ respectively.

- i. We call $\varphi : M \rightarrow N$ a **graded $R[x]$ -module homomorphism** or **graded homomorphism** if for all $t \in \mathbb{N}_0$, $\varphi(M_t) \subseteq N_t$. That is, φ sends homogeneous elements of M to either zero or homogeneous elements of N of the same degree.
- ii. If $\varphi : M \rightarrow N$ is both an $R[x]$ -module isomorphism and a graded homomorphism, then $\varphi : M \rightarrow N$ is called a **graded $R[x]$ -module isomorphism** or **graded isomorphism** and we say that M and N are **graded isomorphic**, denoted $M \cong N$.

Graded modules and graded homomorphisms give rise to a *category* of graded modules. We state this as a theorem below.

Theorem 2.4.6. Let R be a PID. Graded $R[x]$ -modules and graded $R[x]$ -module homomorphisms form a well-defined category, denoted $\mathbf{GrMod}_{R[x]}$ and called the **category of graded $R[x]$ -modules**.

Remark. We refer to [NO04, Section 2.2] for the claim of $\mathbf{GrMod}_{R[x]}$ being a well-defined category. Note that the category of graded modules can be more generally defined using grading over any abelian group G and over some G -graded ring R .

For our case, the symbol $\mathbf{GrMod}_{R[x]}$ refers specifically to the category of \mathbb{N}_0 -graded modules over $R[x]$ equipped with the standard grading. Note that while it is possible to define $\mathbf{GrMod}_{R[x]}$ to be of \mathbb{Z} -graded modules (following remarks under Definition 2.4.2), the category equivalence presented in Section 2.5 between $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ is not valid for said definition.

Observe that all graded $R[x]$ -modules and all graded $R[x]$ -module homomorphisms are $R[x]$ -modules and $R[x]$ -module homomorphisms respectively. Equivalently, we can also say that $\mathbf{GrMod}_{R[x]}$ is a *subcategory* of $\mathbf{Mod}_{R[x]}$ or that there exists a forgetful functor $\mathbf{GrMod}_{R[x]} \rightarrow \mathbf{Mod}_{R[x]}$ (see [NO04, Section 2.5]).

Later in Section 4.3, we use this subcategory relationship to investigate an algorithm involving $\mathbb{F}[x]$ -modules and see if it can be extended to the case of graded $\mathbb{F}[x]$ -modules. The main hurdle here is that the graded structure of the graded modules involved may not be preserved across multiple steps, which may result in an isomorphism that is only true at the level of $R[x]$ -modules and not at the level of graded $R[x]$ -modules. The distinction between isomorphisms in $\mathbf{Mod}_{R[x]}$ and graded isomorphisms in $\mathbf{GrMod}_{R[x][x]}$ becomes relevant when we use the category equivalence discussed in Section 2.5.

In Chapter 4, we will talk about different levels of isomorphisms involving graded $R[x]$ -modules. To avoid ambiguity, we identify some notation below.

Remark 2.4.7. When given an equality ($=$) or isomorphism relation (\cong) involving $R[x]$ -modules, the following symbols denote the category in which the relation applies.

- i. The symbol \mathbf{Ab} in $\stackrel{\mathbf{Ab}}{=}$ or $\stackrel{\mathbf{Ab}}{\cong}$ denotes a relation either in the category \mathbf{AbGrp} of abelian groups, \mathbf{Mod}_R of R -modules, or $\mathbf{Vect}_{\mathbb{F}}$ of \mathbb{F} -vector spaces. In these cases, only the additive structure and (if defined) the action of R on the modules involved are respected by the relations.

In this paper, we often use this shorthand to describe the homogeneous decomposition of graded $\mathbb{F}[x]$ -modules, wherein an accompanying collection of assignments in the form $x \cdot mx^t$ is expected.

- ii. The symbol \mathbf{Mod} in $\stackrel{\mathbf{Mod}}{=}$ or $\stackrel{\mathbf{Mod}}{\cong}$ denotes a relation in the category $\mathbf{Mod}_{R[x]}$ of $R[x]$ -modules, i.e. the relation respects the action of $R[x]$ on the $R[x]$ -modules involved (and not just R) but it may not respect the graded structure (if such exists).
- iii. The symbol \mathbf{GrMod} in $\stackrel{\mathbf{GrMod}}{=}$ or $\stackrel{\mathbf{GrMod}}{\cong}$ denotes a relation in the category $\mathbf{GrMod}_{R[x]}$ of graded $R[x]$ -modules, i.e. the relation involves graded $R[x]$ -modules and is given by a graded isomorphism.

Note that the symbols \mathbf{Ab} , \mathbf{Mod} , and \mathbf{GrMod} can be seen as increasing restrictions on the relation. For example, the relation $M \stackrel{\mathbf{GrMod}}{=} N$ implies $M \stackrel{\mathbf{Mod}}{=} N$ which implies $M \stackrel{\mathbf{Ab}}{=} N$. The converse is not generally true.

Next, we state the result that allows us to assume without loss of generality that the notation for the elements of a graded $R[x]$ -module carry information about their homogeneity or degree.

Proposition 2.4.8. Let M be a graded $R[x]$ -module and denote its homogeneous component of degree $t \in \mathbb{N}_0$ by $M_t \subseteq M$. Then, there exists a graded $R[x]$ -module M' such that the homogeneous component of M' of degree $t \in \mathbb{N}_0$ is given by $M_t x^t$ and M' is graded isomorphic to M .

Proof. Assume M is a graded $R[x]$ -module. By definition, there exists a family $\{M_t\}_{t \in \mathbb{N}_0}$ of additive sub-

groups M_t of M such that

$$M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} M_t \quad \text{and} \quad Rx^s \cdot M_t \subseteq M_{t+s}$$

where \cdot denotes the action of $R[x]$ on M . The homogeneous component of M of degree $t \in \mathbb{N}_0$ is given by M_t . By restriction of the action of $R[x]$ on M , each additive subgroup M_t is also an R -module.

For each $t \in \mathbb{N}_0$, define the R -module $M_t x^t$ by the canonical isomorphism $\eta_t : M_t \rightarrow M_t x^t$ by $m_t \mapsto m_t x^t$ for all $m_t \in M_t$. Since $M_t x^t \cap M_s x^s = \{0\}$ whenever $t \neq s$, $\bigoplus_{t \in \mathbb{N}_0} M_t x^t$ is a well-defined (internal) direct sum of R -modules. Define the R -module M' by

$$M' := \bigoplus_{t \in \mathbb{N}_0} M_t x^t$$

Then, for all $m \in M'$, m decomposes uniquely into $\sum_{t \in \mathbb{N}_0} m_t x^t$ with $m_t \in M_t$ for all $t \in \mathbb{N}_0$ such that only finitely many m_t 's are nonzero.

Let $\eta := \bigoplus_{t \in \mathbb{N}_0} \eta_t$ be the R -module homomorphism $\eta : M \rightarrow M'$ induced by the direct sum operation between R -modules. Observe that η must be an R -module isomorphism. Let $\star : R[x] \times M \rightarrow M$ be given by $x^s \star m := \eta(x^s \cdot \eta^{-1}(m))$ for all $m \in M'$ and $s \in \mathbb{N}_0$. More specifically, if $m \in M'$ decomposes into $m = \sum_{t \in \mathbb{N}_0} m_t x^t$ with $m_t \in M_t$ for all $t \in \mathbb{N}_0$, then

$$x^s \star m = x^s \star \sum_{t \in \mathbb{N}_0} m_t x^t = \sum_{t \in \mathbb{N}_0} \eta(x^s \cdot \eta^{-1}(m_t x^t)) = \sum_{t \in \mathbb{N}_0} (x^s \cdot m_t) x^{t+s}$$

Since $Rx^s \cdot M_t \subseteq M_{t+s}$, $(x^s \cdot m_t) \in M_{t+s} x^{t+1}$. Since η is an R -module isomorphism, \star forms a well-defined action of $R[x]$ on M' and makes M' an $R[x]$ -module.

Let $t, s \in \mathbb{N}_0$. Since $Rx^s \star M_t x^t = (Rx^s \cdot M_t) x^{t+s} \subseteq (M_{t+s}) x^{t+1} = M_{t+s} x^{t+s}$, M' is a graded $R[x]$ -module with homogeneous component of degree $t \in \mathbb{N}_0$ given by $M_t x^t$. Since $\eta(M_t) = M_t x^t$, η is a graded $R[x]$ -module homomorphism. Since η is an $R[x]$ -module isomorphism, η is a graded isomorphism. Therefore, M and M' are graded isomorphic. \blacksquare

Observe that the fact that $M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} M_t$ is a direct sum of R -modules is critical here since it implies that each nonzero homogeneous element of M has a unique degree. Then, using the proposition above, we can assume the following:

Remark 2.4.9. Let M be a graded $R[x]$ -module. In this paper, assume without loss of generality that there exists a family $\{M_t\}_{t \in \mathbb{N}_0}$ of R -modules such that

$$M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} M_t x^t \quad \text{and} \quad Rx^s \cdot M_t \subseteq M_{s+t} \text{ for all } t, s \in \mathbb{N}_0$$

and the homogeneous component of M of degree $t \in \mathbb{N}_0$ is exactly $M_t x^t$. Therefore, a nonzero element $m \in M$ is homogeneous of degree $t \in \mathbb{N}_0$ if and only if $m = m_t x^t$ for some $m_t \in M_t$.

Note that some $R[x]$ -module isomorphisms, if between two graded $R[x]$ -modules, can be made into graded $R[x]$ -module isomorphisms by applying appropriate shifting in grading. Below, we provide a definition for this grading shift.

Definition 2.4.10. Let M be a graded $R[x]$ -module and assume, following Remark 2.4.9, that the homogeneous

component of M of degree t is given by $M_t x^t$ for some R -module M_t for all $t \in \mathbb{N}_0$. For each $k \in \mathbb{N}_0$, define the **k -upwards shift** $\Sigma^k M$ of a graded $R[x]$ -module M to be the graded $R[x]$ -module given as follows.

- i. The underlying abelian group of $\Sigma^k M$ is given by $\Sigma^k M := \bigoplus_{t \in \mathbb{N}_0} M_t x^{t+k} = \bigoplus_{t=k}^{\infty} M_{t-k} x^t$.
- ii. The action of $R[x]$ on $\Sigma^k M$ satisfies $x \cdot ax^{t+k} = bx^{t+k+1}$ whenever $x \cdot_M ax^t = bx^{t+1}$ with $ax^t \in M_t x^t \subseteq M$, $bx^{t+1} \in M_{t+1} x^{t+1}$ where \cdot_M refers to the action of $R[x]$ on M .

The homogeneous component of $\Sigma^k M$ of degree $t \in \mathbb{N}_0$ is given by $M_{t-k} x^t$ if $t \geq k$ and trivial if $t < k$. The index $k \in \mathbb{N}_0$ is called the **grading shift** or **shift** of $\Sigma^k M$.

Remark. In [Pee11; NO04], the k -upwards shift $\Sigma^k M$ of M is denoted as $M(-k)$ since the homogeneous component of $\Sigma^k M = M(-k)$ of degree $t \geq k$ is exactly the homogeneous component of M of degree $t - k$. The notation $\Sigma^k M$ follows that of [ZC05].

One key application of upwards shifts involves the isomorphism between $R[x]$ and $\Sigma^k R[x]$. Let $\varphi : R[x] \rightarrow \Sigma^k R[x]$ be given by $rx^t \mapsto rx^{t+k}$. We claim that φ is an $R[x]$ -module homomorphism with inverse $rx^t \mapsto rx^{t-k}$. Note that $\Sigma^k R[x]$ has no elements of the form rx^t with $t < k$. Observe that if $k \geq 1$, φ sends the homogeneous elements of $R[x]$ of degree $t \in \mathbb{N}_0$ to the homogeneous component of $\Sigma^k R[x]$ of degree $t + k \neq t$. Therefore, φ is not a graded isomorphism. In other words, we have that

$$\text{If } k \geq 1: \quad R[x] \stackrel{\mathbf{Mod}}{\cong} \Sigma^k R[x] \quad \text{but} \quad R[x] \stackrel{\mathbf{GrMod}}{\not\cong} \Sigma^k R[x]$$

Note that this affects how the Structure Theorem on finitely-generated $\mathbb{F}[x]$ -modules in the category $\mathbf{Mod}_{\mathbb{F}[x]}$ translates to the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ of graded $\mathbb{F}[x]$ -modules.

Next, we identify a number of useful results involving algebraic constructions in $\mathbf{GrMod}_{R[x]}$, particularly those that extend, in a specific sense, those of the ungraded category $\mathbf{Mod}_{R[x]}$ of $R[x]$ -modules.

Proposition 2.4.11. An $R[x]$ -submodule N of a graded $R[x]$ -module M is a graded $R[x]$ -module if and only if there exists a homogeneous system of generators for N , i.e. there exists a homogeneous set of elements in M that generates N .

Remark. The first steps of the proof can be found under [Mar, Proposition 2.1]. This proposition is also listed as [Pee11, Exercise 2.8].

Proposition 2.4.12. Kernels and images of graded homomorphisms are graded, i.e.

If $\varphi : M \rightarrow N$ is a graded $R[x]$ -module homomorphism between graded $R[x]$ -modules M and N , then $\ker(M)$ is graded submodule of M and $\text{im}(M)$ is a graded submodule of N .

Remark. A proof is available under [Pee11, Proposition 2.9, p8-9].

Proposition 2.4.13. Quotients of graded modules over graded submodules are graded, i.e.

Let M be a graded $R[x]$ -module and N a graded submodule of M . Following Remark 2.4.9, let the homogeneous component of M and N of degree $t \in \mathbb{N}_0$ be given by $M_t x^t$ and $N_t x^t$ with R -modules M_t and N_t respectively. Then, M/N is a graded $R[x]$ -module with homogeneous component of degree $t \in \mathbb{N}_0$ given by $(M_t / N_t) x^t$.

Remark. Note that $M/N = \text{coker}(N \hookrightarrow M)$ and the inclusion map $N \hookrightarrow M$ is a graded homomorphism. [Pee11, Proposition 2.9, p8-9] states that cokernels of graded homomorphisms are also graded along with a proof.

Proposition 2.4.14. Direct sums of graded modules are graded, i.e.

Let M and N be graded $R[x]$ -modules. Following Remark 2.4.9, let the homogeneous component of M and N of degree $t \in \mathbb{N}_0$ be given by $M_t x^t$ and $N_t x^t$ with R -modules M_t and N_t respectively. Then, $M \oplus N$ is a graded $R[x]$ -module with homogeneous component of degree $t \in \mathbb{N}_0$ given by $(M_t \oplus N_t)x^t$.

Remark. This is discussed in [NO04, p20] and listed as [Mar, Exercise 1.4]. [Pee11] does not seem to explicitly state this but assumes this in other results such as [Pee11, Exercise 2.9, Theorem 2.10].

Next, we state a result involving upward shifts in grading and quotients.

Proposition 2.4.15. Let M be a graded $R[x]$ -module and let N be a graded submodule of M . For all $s \in \mathbb{N}_0$,

$$\Sigma^s(M/N) \stackrel{\text{GrMod}}{\cong} (\Sigma^s M) / (\Sigma^s N)$$

That is, the upwards shifts distribute over quotients.

Proof. Since N is a graded submodule of M , then $\Sigma^s N$ must also be a graded submodule of $\Sigma^s M$ and the quotient $\Sigma^s M / \Sigma^s N$ is well-defined. Denote the homogeneous component of degree $t \in \mathbb{N}_0$ of a graded module by addition of the subscript t . For all $t < s$, $(\Sigma^s M)_t = 0$, $(\Sigma^s N)_t = 0$, and $(\Sigma^s(M/N))_t = 0$. For all $t \geq s$, we have that

$$(\Sigma^s(M/N))_t = (M/N)_{t-s} = M_{t-s} / N_{t-s} = (\Sigma^s M)_t / (\Sigma^s N)_t = ((\Sigma^s M) / (\Sigma^s N))_t$$

by Definition 2.4.10 (for $\Sigma^k M$) and by Proposition 2.4.15 (involving quotients). ■

Recall that $\mathbb{F}[x]$ is a PID for any field \mathbb{F} . The Structure Theorem (Theorem 4.1.1) on $\mathbf{Mod}_{\mathbb{F}[x]}$ guarantees the existence and uniqueness of *invariant factor decompositions* for finitely generated $\mathbb{F}[x]$ -modules up to $\mathbb{F}[x]$ -module isomorphism. In Section 4.3, we state a corresponding result in $\mathbf{GrMod}_{\mathbb{F}[x]}$ which we call the *The Graded Structure Theorem* (Theorem 4.3.1) for *graded invariant factor decompositions* (given in Definition 4.3.2) for finitely generated graded modules. For convenience, statements of Theorem 4.3.1 and Definition 4.3.2 are included below.

Theorem 4.3.1. The Graded Structure Theorem.

Let M be a finitely generated graded $\mathbb{F}[x]$ -module over $\mathbb{F}[x]$ for some field \mathbb{F} . Then, there exists a finite direct sum of shifted cyclic graded submodules of $\mathbb{F}[x]$ that is graded isomorphic to M as follows:

$$M \stackrel{\text{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]$$

with indices $s_1, \dots, s_r, \dots, s_m \in \mathbb{N}_0$ and non-zero, non-unit $x^{t_1}, x^{t_2}, \dots, x^{t_r} \in \mathbb{F}[x]$ such that the divisibility relation $x^{t_1} \mid x^{t_2} \mid \cdots \mid x^{t_r}$ is satisfied. Furthermore, the collection $\{(x^{t_1}, s_1), \dots, (x^{t_r}, s_r), (0, s_{r+1}), \dots, (0, s_m)\}$ is uniquely determined by M up to graded isomorphism.

Remarks. (1) We use the term *cyclic* similarly in the case of $R[x]$ -modules, i.e. a cyclic graded module is a graded module that can be generated by a single element. It can be verified that cyclic graded modules can be generated by a single *homogeneous* element as well.

(2) Note that the ideal $(x^t) = \mathbb{F}[x] \cdot x^t$ is a graded submodule of $\mathbb{F}[x]$ since it is generated by a

homogeneous element x^t . Therefore, the quotient $\mathbb{F}[x]/(x^t)$ results in a graded module. In this case, we call (x^t) a *graded ideal* of $\mathbb{F}[x]$.

Definition 4.3.2. Let M be a finitely generated $\mathbb{F}[x]$ -module and let the following direct sum decomposition of M be as denoted in the Graded Structure Theorem (Theorem 4.3.1):

$$M \stackrel{\text{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]$$

This decomposition is called the **graded invariant factor decomposition** of M . The **invariant factors** of M are given by $x^{t_1}, \dots, x^{t_r} \in \mathbb{F}[x]$ (which are non-zero and non-unit) and the **grading shifts** of M by $s_1, \dots, s_m \in \mathbb{N}_0$.

The notion of grading on $R[x]$ -modules and $R[x]$ -module homomorphisms also extend to the case of chain complexes. We provide relevant definitions below.

Definition 2.4.16. A chain complex $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ of $R[x]$ -modules is a **graded chain complex** if for all $n \in \mathbb{Z}$, C_n is a graded $R[x]$ -module and $\partial_n : C_n \rightarrow C_{n-1}$ is a graded homomorphism. A chain map $\varphi_* : C_* \rightarrow A_*$ with $\varphi_* = (\varphi_n : C_n \rightarrow A_n)_{n \in \mathbb{Z}}$ between graded chain complexes C_* and A_* is a **graded chain map** if each homomorphism φ_n is a graded homomorphism.

The collection of graded chain complexes and graded chain maps also form a chain complex category and a corresponding chain homology functor. We identify some notation below.

Definition 2.4.17. Let $\mathbf{Ch-GrMod}_{R[x]}$ denote the **category of graded chain complexes** of graded $R[x]$ -modules and graded chain maps. To each $n \in \mathbb{Z}$, let $H_n(-)$ refer to the n^{th} **chain homology functor** $H_n(-) : \mathbf{Ch-GrMod}_{R[x]} \rightarrow \mathbf{GrMod}_{R[x]}$.

Remark. For the claim of $\mathbf{Ch-GrMod}_{R[x]}$ being well-defined, we refer to [NO04, Section 2.2] which claims that $\mathbf{GrMod}_{R[x]}$ is an *abelian category*, and to [Rot08, Section 5.5] which states that each abelian category has a corresponding category of chain complexes and family of homology functors. Note that Proposition 2.4.13 states that the n^{th} chain homology of a graded chain complex, as the cokernel of a graded homomorphism, must be a graded module.

Note that, much like how $\mathbf{GrMod}_{R[x]}$ can be seen as a subcategory of $\mathbf{Mod}_{R[x]}$, the category $\mathbf{Ch-GrMod}_{R[x]}$ of graded chain complexes can be seen as a subcategory of the category $\mathbf{Ch-Mod}_{R[x]}$ of chain complexes of $R[x]$ -modules.

Lastly, the additional structure required for the grading for graded $R[x]$ -modules also brings about graded versions of certain properties of ungraded $R[x]$ -modules. For example, [NO04, p21] defines a graded $\mathbb{F}[x]$ -module to be *graded-free* or *gr-free* if it has a homogeneous basis, and provides a counterexample in which a graded module over some graded ring (other than $R[x]$) with a non-homogeneous basis is not graded-free. For clarity, we explicitly make the following remark.

Remark 2.4.18. When we state that a graded $R[x]$ -module M is a **free graded** module, we mean that M is both a graded module and a free $R[x]$ -module, i.e. M has a basis that is not necessarily homogeneous. In this paper, if M has a homogeneous basis, we do not use the term *graded-free* for M to avoid ambiguity.

Section 2.5. The Equivalence between Persistence Modules and Graded Modules

In this section, we discuss the isomorphism of categories between the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules over \mathbb{F} (see Definition 2.2.1) and the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ of \mathbb{N}_0 -graded $\mathbb{F}[x]$ -modules (see Theorem 2.4.6). We also discuss how this isomorphism allows us to correspond the algebraic constructions in $\mathbf{Pers}_{\mathbb{F}}$ (e.g. direct sums, persistence isomorphisms, chain complexes) to those in $\mathbf{GrMod}_{\mathbb{F}[x]}$.

Before we begin, we want to emphasize that the results presented in this section rely heavily on concepts involving category theory and homological algebra, and proofs for said results are not provided in the following discussion. Since we mainly need to *use* the results in this section for the calculation of persistent homology, it should suffice to know the definitions involving categories and functors presented in Appendix A4 and the following definition for an *isomorphism of categories*:

Given two categories \mathbf{C} and \mathbf{D} with identity functors $\text{id}_{\mathbf{C}} : \mathbf{C} \rightarrow \mathbf{C}$ and $\text{id}_{\mathbf{D}} : \mathbf{D} \rightarrow \mathbf{D}$ respectively, an **isomorphism of categories** or **category isomorphism** between \mathbf{C} and \mathbf{D} refers to a pair of functors $F : \mathbf{C} \rightarrow \mathbf{D}$ and $G : \mathbf{D} \rightarrow \mathbf{C}$ such that $G \circ F = \text{id}_{\mathbf{C}}$ and $F \circ G = \text{id}_{\mathbf{D}}$ [Rie16, p20].

Roughly speaking, the functors F and G describe a bijective correspondence between the objects and morphisms of \mathbf{C} and those of \mathbf{D} respectively.

Remark. An *equivalence of categories* or a *category equivalence* between categories \mathbf{C} and \mathbf{D} refers to a pair of functors $F : \mathbf{C} \rightarrow \mathbf{D}$ and $G : \mathbf{D} \rightarrow \mathbf{C}$ such that there exists a *natural isomorphism* between the functors $G \circ F$ and $\text{id}_{\mathbf{C}}$ and another between $F \circ G$ and $\text{id}_{\mathbf{D}}$. The existence of these natural isomorphisms is usually denoted by $G \circ F \cong \text{id}_{\mathbf{C}}$ and $F \circ G \cong \text{id}_{\mathbf{D}}$.

We point this out since [ZC05, Theorem 3.1] claims the existence of an *equivalence of categories* between $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$. For our discussion, it should suffice to know that the existence of an isomorphism of categories is a stronger requirement than that of an equivalence of categories, i.e. a category isomorphism implies a category equivalence but the converse is not generally true, and that the results we present below also apply for the weaker case of category equivalence.

In this section, the isomorphism of categories between $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ is given by the functors $\Gamma_{\mathbf{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ (described in Definition 2.5.1) and $\Gamma_{\mathbf{Pers}} : \mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ (described in Definition 2.5.6). Then, we present the claim that $\Gamma_{\mathbf{GrMod}}$ and $\Gamma_{\mathbf{Pers}}$ indeed determine an isomorphism of categories in Theorem 2.5.9. Finally, the propositions following Theorem 2.5.9 describe the correspondence between algebraic constructions.

We begin with a description of $\Gamma_{\mathbf{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ below, adapted from [BM21, p8].

Definition 2.5.1. Fix a field \mathbb{F} . Define $\Gamma_{\mathbf{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ to be the following assignment of the objects and morphisms of $\mathbf{Pers}_{\mathbb{F}}$ to those of $\mathbf{GrMod}_{\mathbb{F}[x]}$ respectively.

- i. *Object Assignment:* Let $(V_{\bullet}, \alpha_{\bullet})$ be a persistence module over \mathbb{F} . Define $\Gamma_{\mathbf{GrMod}}(V_{\bullet}, \alpha_{\bullet}) = \Gamma_{\mathbf{GrMod}}(V_{\bullet})$ to be the graded $\mathbb{F}[x]$ -module given as follows:

$$\Gamma_{\mathbf{GrMod}}(V_{\bullet}) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} V_t x^t = \left\{ \sum_{t=0}^{\infty} v_t x^t : v_t \in V_t \text{ for all } t \in \mathbb{N}_0 \right\}$$

where only finitely many of the v_t 's are nonzero for each element of $\Gamma_{\mathbf{GrMod}}(V_{\bullet})$. Let the action of $\mathbb{F}[x]$ on $\Gamma_{\mathbf{GrMod}}(V_{\bullet})$ be given by $x^s \cdot v_t x^t := \alpha_{t+s,t}(v_t) x^{s+t}$. Then, for each $t \in \mathbb{N}_0$, the homogeneous component of $\Gamma_{\mathbf{GrMod}}(V_{\bullet}, \alpha_{\bullet})$ of degree t is given exactly by $V_t x^t$.

- ii. *Morphism Assignment:* Let $\varphi_{\bullet} : (V_{\bullet}, \alpha_{\bullet}) \rightarrow (W_{\bullet}, \gamma_{\bullet})$ be a persistence morphism between persistence modules over \mathbb{F} with $\varphi_{\bullet} = (\varphi_t : V_t \rightarrow W_t)_{t \in \mathbb{N}_0}$. Define $\Gamma_{\mathbf{GrMod}}(\varphi_{\bullet})$ to be the graded $\mathbb{F}[x]$ -homomorphism

$\mathbf{\Gamma}_{\text{GrMod}}(\varphi_\bullet) : \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \rightarrow \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ given by

$$\sum_{t=0}^{\infty} v_t x^t \mapsto \sum_{t=0}^{\infty} \varphi_t(v_t) x^t$$

Remarks. (1) The notation $\mathbf{\Gamma}_{\text{GrMod}}$ is not standard or convention in persistence literature. We introduced this notation since we had found that explicitly identifying the “conversion” of persistence modules to graded modules to be helpful in understanding the theory.

On a less serious note, we chose uppercase gamma ($\mathbf{\Gamma}$) since it looks like the letter **T**. In this case, the notation $\mathbf{\Gamma}_{\text{GrMod}}$ references the conversion of a persistence module **to** a graded module in $\mathbf{GrMod}_{\mathbb{F}[x]}$. For contrast, [ZC05] denotes the functor corresponding to $\mathbf{\Gamma}_{\text{GrMod}}$ simply as α (lowercase alpha), i.e. they would write $\alpha(V_\bullet)$ instead of $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$.

(2) The statement of $\mathbf{\Gamma}_{\text{GrMod}}$ above follows Remark 2.4.9, where we explicitly include the power x^t on the notation for the elements of graded modules to identify the degree $t \in \mathbb{N}_0$. For contrast, [ZC05] and [BM21] defines $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ as $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) := \bigoplus_{t \in \mathbb{N}_0} V_t$, i.e. the summands are V_t instead of $V_t x^t$. In this case, the direct sum of \mathbb{F} -vector spaces is to be interpreted as an *external direct sum* and the elements of $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ are \mathbb{N}_0 -indexed tuples (v_0, v_1, \dots) with the entry at index $t \in \mathbb{N}_0$ being an element of V_t .

The assignment $\mathbf{\Gamma}_{\text{GrMod}}$ as denoted above determines a functor, as stated below.

Proposition 2.5.2. Fix a field \mathbb{F} . The object and morphism assignment $\mathbf{\Gamma}_{\text{GrMod}}$ by Definition 2.5.1 determines a functor $\mathbf{\Gamma}_{\text{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$.

Remark. A brief discussion of the proof is available in [BM21, p8], which relies on the commutativity requirement on the structure maps of persistence modules (see Lemma 2.1.3) and on the linear maps of persistence morphisms (see Definition 2.2.1(ii)).

One way to prove the proposition above involves verifying that the assignments are well-defined (in that, they indeed produce graded modules and graded homomorphisms) and that the functorial axioms are satisfied (see Definition A4.4(i-iv)). In the following discussion, we roughly explain why the object and morphism assignment by $\mathbf{\Gamma}_{\text{GrMod}}$ produces graded $\mathbb{F}[x]$ -modules and graded $\mathbb{F}[x]$ -module homomorphisms respectively.

About the Object Assignment of $\mathbf{\Gamma}_{\text{GrMod}}(-)$: The construction of the graded $\mathbb{F}[x]$ -module by $\mathbf{\Gamma}_{\text{GrMod}}(-)$, as stated in Definition 2.5.1(i), can be understood to happen in three stages: Let $V_\bullet = (V_\bullet, \alpha_\bullet)$ be a persistence module over some field \mathbb{F} .

1. The statement $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} V_t x^t$ first defines $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ as a (graded) \mathbb{F} -vector space.

For each $t \in \mathbb{N}_0$, V_t is an \mathbb{F} -vector space by assumption of $(V_\bullet, \alpha_\bullet)$ being a persistence module and $V_t x^t$ is a \mathbb{F} -vector space given by the canonical linear isomorphism $V_t \rightarrow V_t x^t$ by $v \mapsto v x^t$ for all $v \in V_t$. Then, the direct sum $\bigoplus_{t=0}^{\infty} V_t x^t$ of \mathbb{F} -vector spaces induces an action of \mathbb{F} on the set $\bigoplus_{t=0}^{\infty} V_t x^t$ (as discussed in Remark 2.4.9), making $\bigoplus_{t=0}^{\infty} V_t x^t$ an \mathbb{F} -module, i.e. a \mathbb{F} -vector space.

Note that the characterization of the elements of $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ as $\sum_{t=0}^{\infty} v_t x^t$ refers to how we prefer to interpret the elements of graded modules as formal sums of powers of x^t with coefficients in V_t , per Remark 2.4.9.

2. The assignments $x^s \cdot v_t x^t = \alpha_{t+s,t}(v_t) x^{t+s}$ determine the action of $\mathbb{F}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ and makes $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ a $\mathbb{F}[x]$ -module, not necessarily graded at this point.

For clarity, let $\star : \mathbb{F}[x] \times M \rightarrow M$ with $M := \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ be given by $x^s \star v_t x^t := \alpha_{t+s,t}(v_t) x^{t+s}$, i.e. we denote the supposed action of $\mathbb{F}[x]$ on M by \star . Since $\alpha_{t+s,t}(V_t) \subseteq V_{t+s}$ by definition of V_\bullet as a persistence module, $\alpha_{t+s,t}(v_t) x^{t+s} \in M$.

Recall that, by Definition 2.4.1, an action of $\mathbb{F}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ must be a biadditive group action. Since the restriction of \star onto $\mathbb{F} \times M \rightarrow M$ describes the action of \mathbb{F} on M and M has been shown to be a \mathbb{F} -vector space, it should suffice to show that for all $t, s, r \in \mathbb{N}_0$ and for all $v_t \in V_t$,

$$x^r \star (x^s \star v_t x^t) = (x^r \cdot x^s) \star v_t x^t$$

where \cdot refers to the multiplication on $\mathbb{F}[x]$. By Lemma 2.1.3, we have the commutativity relation $\alpha_{t+s+r,t} = \alpha_{t+s+r,t+s} \circ \alpha_{t+s,t}$ (the structure maps of V_\bullet commute with each other). Then,

$$\begin{aligned} x^r \star (x^s \star v_t x^t) &= x^r \star \alpha_{t+s,t}(v_t) x^{t+s} = (\alpha_{t+s+r,t+s} \circ \alpha_{t+s,t})(v_t) x^{t+s+r} \\ &= \alpha_{t+s+r,t}(v_t) x^{t+s+r} = x^{s+r} \star v_t x^t = (x^r \cdot x^s) \star v_t x^t \end{aligned}$$

Therefore, \star is a well-defined action of $\mathbb{F}[x]$ on M and $M = \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ is a $\mathbb{F}[x]$ -module.

3. Finally, the specified domain and codomain of the structure maps $\alpha_{s,t} : V_t \rightarrow V_s$ of V_\bullet for each $t \in \mathbb{N}_0$ makes $M = \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ a *graded* $\mathbb{F}[x]$ -module. More specifically, the relation $\alpha_{t+s,t}(V_t) \subseteq V_{t+s}$ implies that the action \star of $\mathbb{F}[x]$ on M satisfies the following relation for all $t, s \in \mathbb{N}_0$:

$$\mathbb{F} x^s \star V_t x^t = \alpha_{t+s,t}(V_t) x^{t+s} \subseteq V_{t+s} x^{t+s}$$

This, together with the decomposition $M = \bigoplus_{t \in \mathbb{N}_0} V_t x^t$, satisfies the conditions of Definition 2.4.3 and determines that $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ is a *graded* $\mathbb{F}[x]$ -module with homogeneous component of degree $t \in \mathbb{N}_0$ given exactly by $V_t x^t$.

Recall that, by Proposition 2.1.4, the structure maps $\alpha_{s,t} : V_t \rightarrow V_s$ of V_\bullet are determined uniquely by the collection $\{\alpha_t : V_t \rightarrow V_{t+1}\}_{t \in \mathbb{N}_0}$. Similarly, we can unambiguously describe the action \star of $\mathbb{F}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ using assignments of the following form:

$$x \star v_t x^t = \alpha_t(v_t) x^{t+1} \quad \text{for all} \quad v_t x^t \in V_t x^t \subseteq \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \quad \text{with} \quad \alpha_t : V_t \rightarrow V_{t+1}$$

As a sidenote, [ZC05] uses this simpler characterization to define the object assignment of the functor $\mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$. We give an example of the object assignment of $\mathbf{\Gamma}_{\text{GrMod}}$ in action below.

Example 2.5.3. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module over \mathbb{Q} with vector spaces given by $V_t = \mathbb{Q}\langle a \rangle$ and linear maps $\alpha_t : V_t \rightarrow V_{t+1}$ given by $a \mapsto a$ for all $t \in \mathbb{N}_0$. Then, the graded $\mathbb{F}[x]$ -module $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ is given by

$$\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} V_t x^t = \bigoplus_{t \in \mathbb{N}_0} \mathbb{Q}\langle a \rangle x^t = \mathbb{Q}[x]\langle a \rangle$$

with the action of $\mathbb{Q}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ given by $x \cdot a x^t = \alpha_t(a) x^{t+1} = a x^{t+1}$. Note that $\mathbb{Q}[x]\langle a \rangle$ is the free $\mathbb{F}[x]$ -module generated by the indeterminate a with a interpreted to have degree 0. Since $x \cdot a x^t = a x^{t+1}$, it can be verified that the map $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \rightarrow \mathbb{Q}[x]$ by $a x^t \mapsto x^t \in \mathbb{F}[x]$ is a graded isomorphism. That is, $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ is graded isomorphic to $\mathbb{Q}[x]$.

The graded isomorphism in the example above is relatively straightforward to determine since the chosen bases for the vector spaces $\{V_t\}_{t \in \mathbb{N}_0}$ are “compatible” with the structure maps and the ring operation in $\mathbb{Q}[x]$. Here, we interpret a “compatible” choice to mean the following: given a persistence module $(V_\bullet, \alpha_\bullet)$, the structure map $\alpha_t : V_t \rightarrow V_{t+1}$ sends a basis element a_t of V_t to another basis element a_{t+1} of V_{t+1} so that the

action of $\mathbb{F}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ is given by $x \cdot a_t x^t \mapsto b_{t+1} x^{t+1}$. Loosely speaking, the bases are chosen such that multiplication by x on an element vx^t of $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ changes the basis element and adds 1 to the power of x^t . We provide an example below where the immediate choice for bases is not “compatible” in this same sense.

Example 2.5.4. Let $(W_\bullet, \gamma_\bullet)$ be a persistence module over \mathbb{Q} with vector spaces given by $W_t = \mathbb{Q}\langle a \rangle$ and linear maps $\gamma_t : W_t \rightarrow W_{t+1}$ by $a \mapsto 2a$ for all $t \in \mathbb{N}_0$. Define M to be the graded $\mathbb{Q}[x]$ -module given by $M := \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$. Then, M has the following underlying abelian group:

$$\mathbf{\Gamma}_{\text{GrMod}}(W_\bullet) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} W_t x^t = \bigoplus_{t \in \mathbb{N}_0} \mathbb{Q}\langle a \rangle x^t = \mathbb{Q}[x]\langle a \rangle$$

For clarity, let $\star : \mathbb{Q}[x] \times M \rightarrow M$ denote the action of $\mathbb{Q}[x]$ on M and let \cdot refer to the usual scalar multiplication on $\mathbb{Q}[x]$ given by $x \cdot kx^t = kx^{t+1}$ for all $k \in \mathbb{F}$ and $t \in \mathbb{N}_0$. Then, \star is given by $x \star \gamma_t(a)x^t = 2ax^{t+1}$ and more generally by $x^s \star ax^t = 2^s ax^{t+1}$ for all $t, s \in \mathbb{N}_0$.

Define a map $f : M \rightarrow \mathbb{Q}[x]$ by $ax^t \mapsto x^t$ for all $t \in \mathbb{N}_0$. Note that f is only well-defined as a linear map with M and $\mathbb{Q}[x]$ being seen as \mathbb{Q} -vector spaces. In particular, the assignment $ax^t \mapsto x^t$ fails to commute with scalar multiplication over $\mathbb{Q}[x]$, i.e. $x \cdot f(ax^t) \neq f(x \star ax^t)$ for all $t \in \mathbb{N}_0$ since $x \cdot f(ax^t) = x \cdot x^t = x^{t+1}$ but $f(x \star ax^t) = f(2ax^{t+1}) = 2x^{t+1}$.

In contrast, define the map $g : M \rightarrow \mathbb{Q}[x]$ by $ax^t \mapsto 2^{-t}x^t$. Since $\{ax^t\}$ and $\{2^{-t}x^t\}$ both serve as bases for $\mathbb{Q}\langle a \rangle x^t$ and $\mathbb{Q}x^t$ respectively, g is well-defined as a linear map between \mathbb{Q} -vector spaces. Observe that for all $t \in \mathbb{N}_0$, $g(x \star ax^t) = x \cdot g(ax^t)$ since

$$\begin{aligned} g(x \star ax^t) &= g(2ax^{t+1}) = 2^{-(t+1)} \cdot 2x^{t+1} = 2^{-t}x^{t+1} \\ x \cdot g(ax^t) &= x \cdot 2^{-t}x^t = 2^{-t}x^{t+1} \end{aligned}$$

Since the action \star of $\mathbb{F}[x]$ commutes with g , g is also a $\mathbb{Q}[x]$ -module homomorphism. Since for all $t \in \mathbb{N}_0$, g satisfies $\mathbb{Q}x \star \mathbb{Q}\langle a \rangle x^t \subseteq \mathbb{Q}\langle a \rangle x^{t+1}$, g is also a graded $\mathbb{F}[x]$ -module homomorphism. Since g has an obvious inverse by $x^t \mapsto 2^t ax^t$, $M = \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ and $\mathbb{Q}[x]$ are graded isomorphic with graded isomorphism given by $g : M \rightarrow \mathbb{Q}[x]$. Note that the graded isomorphism by $g : M \rightarrow \mathbb{Q}[x]$ may be easier to see if we choose the basis element $b_t := 2^t a$ for the vector space W_t for each $t \in \mathbb{N}_0$.

About the Morphism Assignment of $\mathbf{\Gamma}_{\text{GrMod}}(-)$: The construction of the graded $\mathbb{F}[x]$ -module homomorphism by $\mathbf{\Gamma}_{\text{GrMod}}$ as stated in Definition 2.5.1(ii) can also be understood to happen in three stages. Let $(V_\bullet, \alpha_\bullet)$ and $(W_\bullet, \gamma_\bullet)$ be persistence modules over \mathbb{F} . Let $\varphi_\bullet : V_\bullet \rightarrow W_\bullet$ be a persistence morphism with $\varphi_\bullet = (\varphi_t : V_t \rightarrow W_t)_{t \in \mathbb{N}_0}$.

1. At the level of \mathbb{F} -modules, i.e. \mathbb{F} -vector spaces:

For each $t \in \mathbb{N}_0$, $\varphi_t : V_t \rightarrow W_t$ must be a well-defined linear map between \mathbb{F} -vector spaces V_t and W_t by definition of persistence morphism. The canonical isomorphisms $V_t \rightarrow V_t x^t$ and $W_t \rightarrow W_t x^t$ given by $v \mapsto vx^t$ for $v \in V_t$ and $w \mapsto wx^t$ for $w \in W_t$ respectively define the linear map $\varphi_t^{\text{Gr}} : V_t x^t \rightarrow W_t x^t$ given by $v_t x^t \mapsto \varphi_t(v_t)x^t$. Then, the (internal) direct sum operation on \mathbb{F} -vector spaces induces $\varphi^{\text{Gr}} : \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \rightarrow \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ to be a linear map $\varphi^{\text{Gr}} : \bigoplus_{t \in \mathbb{N}_0} V_t x^t \rightarrow \bigoplus_{t \in \mathbb{N}_0} W_t x^t$ between (graded) \mathbb{F} -vector spaces given by $\varphi^{\text{Gr}} = \bigoplus_{t \in \mathbb{N}_0} \varphi_t^{\text{Gr}}$.

2. At the level of $\mathbb{F}[x]$ -modules, not necessarily graded at this point:

For φ^{Gr} to be a well-defined $\mathbb{F}[x]$ -module homomorphism, it must be shown that the action of $\mathbb{F}[x]$ on both $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ and $\mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ commute with φ^{Gr} . Since φ^{Gr} is a well-defined linear map, it should

suffice to check that the following is true:

$$\varphi^{\text{Gr}}(x^s \star_V v) = x^s \star_W \varphi^{\text{Gr}}(v) \quad \text{for all } t, s \in \mathbb{N}_0 \text{ and } v \in \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$$

where \star_V and \star_W refer to the action of $\mathbb{F}[x]$ on $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ and $\mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ respectively. This condition can be equivalently stated as follows:

$$(\varphi_{t+s} \circ \alpha_{t+s,t})(v_t) = (\gamma_{t+s,t} \circ \varphi_t)(v_t) \quad \text{for all } t, s \in \mathbb{N}_0 \text{ and for all } v_t \in V_t$$

Observe that, by Definition 2.2.1(ii), a persistence morphism φ_\bullet must form commuting squares with the structure maps of V_\bullet and of W_\bullet , i.e. $\varphi_{t+s} \circ \alpha_{t+s,t} = \gamma_{t+s,t} \circ \varphi_t$ for all $t, s \in \mathbb{N}_0$. Therefore, $\varphi^{\text{Gr}} : \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet) \rightarrow \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ is a well-defined $\mathbb{F}[x]$ -module homomorphism.

3. At the level of graded $\mathbb{F}[x]$ -modules:

Since $\varphi^{\text{Gr}}(V_t x^t) \subseteq W_t x^t$ for all $t \in \mathbb{N}_0$ and $V_t x^t$ and $W_t x^t$ are exactly the homogeneous components of $\mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ and $\mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$ of degree $t \in \mathbb{N}_0$ respectively, φ^{Gr} is also a well-defined *graded* $\mathbb{F}[x]$ -module homomorphism.

We provide an example of the morphism assignment of $\mathbf{\Gamma}_{\text{GrMod}}(-)$ below.

Example 2.5.5. Let $(V_\bullet, \alpha_\bullet)$ and $(W_\bullet, \alpha_\bullet)$ be persistence modules over \mathbb{Q} with vector spaces given as follows with distinct indeterminates a, b , and ab :

$$V_t = \begin{cases} 0 & \text{if } t < 5 \\ \mathbb{Q}\langle ab \rangle & \text{if } t \geq 5 \end{cases} \quad \text{and} \quad W_t = \begin{cases} 0 & \text{if } t < 2 \\ \mathbb{Q}\langle a, b \rangle & \text{if } t \geq 2 \end{cases}$$

Define $\alpha_t : V_t \rightarrow V_{t+1}$ by $ab \mapsto ab$ for all $t \geq 5$ and $\gamma_t : W_t \rightarrow W_{t+1}$ by $a \mapsto a$ and $b \mapsto b$ for all $t \geq 3$. Let V^{Gr} and W^{Gr} be graded $\mathbb{Q}[x]$ -modules given by $V^{\text{Gr}} := \mathbf{\Gamma}_{\text{GrMod}}(V_\bullet)$ and $W^{\text{Gr}} := \mathbf{\Gamma}_{\text{GrMod}}(W_\bullet)$. We claim that V^{Gr} and W^{Gr} are graded isomorphic to the following:

$$V^{\text{Gr}} \cong^{\text{GrMod}} \Sigma^5 \mathbb{Q}[x]\langle ab \rangle = \mathbb{Q}[x]\langle abx^5 \rangle \quad \text{and} \quad W^{\text{Gr}} \cong^{\text{GrMod}} \Sigma^3 \mathbb{Q}[x]\langle a, b \rangle = \mathbb{Q}[x]\langle ax^3, bx^3 \rangle$$

where Σ^k refers to the upward shift in grading (see Definition 2.4.10). That is, the relation is valid at the level of graded $\mathbb{Q}[x]$ -module homomorphisms and that the action of $\mathbb{Q}[x]$ on V^{Gr} and W^{Gr} is given by $x \cdot abx^t = abx^{t+1}$ for all $t \geq 5$ and $x \cdot ax^t = ax^{t+1}, x \cdot bx^t = bx^{t+1}$ for all $t \geq 3$ respectively.

For each $t \in \mathbb{N}_0$, define $\varphi_t : V_t \rightarrow W_t$ to be the restriction of the map $\Phi : \mathbb{Q}\langle ab \rangle \rightarrow \mathbb{Q}\langle a, b \rangle$ given by $ab \mapsto b - a$. We claim that $\varphi_\bullet = (\varphi_t)_{t \in \mathbb{N}_0}$ defines a persistence morphism $\varphi_\bullet : V_\bullet \rightarrow W_\bullet$. Let $\varphi^{\text{Gr}} : V^{\text{Gr}} \rightarrow W^{\text{Gr}}$ be the graded $\mathbb{Q}[x]$ -module homomorphism given by $\varphi^{\text{Gr}} := \mathbf{\Gamma}_{\text{GrMod}}(\varphi_\bullet)$. Then, for all $t \in \mathbb{N}_0$ with $t \geq 5$:

$$\varphi^{\text{Gr}}(abx^t) = \varphi_t(ab)x^t = (b - a)x^t$$

Observe that φ^{Gr} sends homogeneous elements of V^{Gr} of degree $t \geq 5$ to homogeneous elements of W^{Gr} of the same degree. Note that V^{Gr} has no homogeneous elements of degree $t < 5$.

Next, we provide a description of $\mathbf{\Gamma}_{\text{Pers}} : \mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ below.

Definition 2.5.6. Fix a field \mathbb{F} . Define $\mathbf{\Gamma}_{\text{Pers}} : \mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ to be the following assignment of the objects and morphisms of $\mathbf{GrMod}_{\mathbb{F}[x]}$ to those of $\mathbf{Pers}_{\mathbb{F}}$ respectively.

i. *Object Assignment:* Let M be a graded $\mathbb{F}[x]$ -module and let its homogeneous component be given by

$M_t x^t$ for some \mathbb{F} -vector space M_t for all $t \in \mathbb{N}_0$. Then, $M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} M_t x^t$.

Define $\mathbf{\Gamma}_{\text{Pers}}(M) =: (V_\bullet, \alpha_\bullet)$ to be the persistence module over \mathbb{F} such that for all $t \in \mathbb{N}_0$, the vector space V_t is given by $V_t := M_t$, and for all $t, s \in \mathbb{N}_0$ with $t \leq s$, the structure map $\alpha_{s,t} : V_t \rightarrow V_s$ is given by $\alpha_{s,t}(v_t) = v_s$ whenever $x^{s-t} \cdot v_t x^t = v_s x^s$ with $v_s \in M_s$.

- ii. *Morphism Assignment:* Let $f : M \rightarrow N$ be a graded $\mathbb{F}[x]$ -module homomorphism between graded $\mathbb{F}[x]$ -modules M and N . Let the homogeneous component of M and N be given by $M_t x^t$ and $N_t x^t$ for some \mathbb{F} -vector spaces M_t and N_t for all $t \in \mathbb{N}_0$ respectively.

Define $\mathbf{\Gamma}_{\text{Pers}}(f) =: \varphi_\bullet$ to be the persistence morphism $\varphi_\bullet : V_\bullet \rightarrow W_\bullet$ between persistence morphisms $(V_\bullet, \alpha_\bullet) := \mathbf{\Gamma}_{\text{Pers}}(M)$ and $(W_\bullet, \gamma_\bullet) := \mathbf{\Gamma}_{\text{Pers}}(N)$ with $\varphi_t : V_t \rightarrow W_t$ given by $\varphi(v_t) := w_t$ whenever $f(v_t x^t) = w_t x^t$ for all $v_t \in V_t$ and for all $t \in \mathbb{N}_0$.

Remarks. (1) As is the case of $\mathbf{\Gamma}_{\text{GrMod}}$, the notation $\mathbf{\Gamma}_{\text{Pers}}$ is not standard or convention in persistence literature. In fact, notation for the functor $\mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ in the category isomorphism was not explicitly identified in [ZC05].

- (2) The graded modules and graded homomorphisms follow Remark 2.4.9, where x^t is added to the notation of the elements of graded modules to help identify the degree of homogeneous elements.

Proposition 2.5.7. Fix a field \mathbb{F} . The object and morphism assignment $\mathbf{\Gamma}_{\text{Pers}}$ by Definition 2.5.6 determines a functor $\mathbf{\Gamma}_{\text{Pers}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$.

Remark. A brief discussion of the proof is available in [BM21, p8].

Note that the arguments given for $\mathbf{\Gamma}_{\text{GrMod}}$ producing well-defined graded modules and graded homomorphisms, when presented in reverse order, also tells us that application of $\mathbf{\Gamma}_{\text{Pers}}$ results in well-defined persistence modules and persistence morphisms. We provide an example of $\mathbf{\Gamma}_{\text{Pers}}$ in action below.

Example 2.5.8. Let M be a graded $\mathbb{Q}[x]$ -module given as follows:

$$M \stackrel{\text{GrMod}}{=} \Sigma^4 \left(\mathbb{Q}[x] / \mathbb{Q}[x] \langle x^5 \rangle \right)$$

Since $(x^5) = \mathbb{Q}[x] \langle x^5 \rangle$ is generated by a homogeneous element of $\mathbb{Q}[x]$, M must be a graded $\mathbb{Q}[x]$ -module by Proposition 2.4.11. Let $\{M_t\}_{t \in \mathbb{N}_0}$ be the family of \mathbb{Q} -vector spaces such that $M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} M_t x^t$ and $M_t x^t$ is the homogeneous component of M of degree $t \in \mathbb{N}_0$ for all $t \in \mathbb{N}_0$. Then, M_t is given as follows:

$$M_t x^t = \begin{cases} 0 & \text{if } t < 4 \\ \mathbb{Q} x^t & \text{if } t \in \{4, 5, \dots, 8\} = [4, 9) \\ 0 & \text{if } t \geq 9 \end{cases}$$

That is, M consists of \mathbb{Q} -linear combinations of x^4, x^5, x^6, x^7 and x^8 . The action of $\mathbb{Q}[x]$ on M is also given by $x \cdot x^t = x^{t+1}$ for all $t \in \{4, \dots, 8\}$ and by $x \cdot x^8 = 0$.

Let $(V_\bullet, \alpha_\bullet)$ be the persistence module over \mathbb{Q} given by $(V_\bullet, \alpha_\bullet) := \mathbf{\Gamma}_{\text{Pers}}(M)$. Then, the vector spaces of V_\bullet are given as follows:

$$V_t = \begin{cases} 0 & \text{if } t < 4 \\ \mathbb{Q} & \text{if } t \in [4, 9) \\ 0 & \text{if } t \geq 9 \end{cases}$$

For $t \in \{4, 5, \dots, 7\}$, the structure map $\alpha_t : V_t \rightarrow V_{t+1}$ is given by $V_t \ni 1 \mapsto 1 \in V_{t+1}$, i.e. $\alpha_t = \text{id}_{\mathbb{Q}}$. For $t = 8$, the structure map $\alpha_8 : V_8 \rightarrow V_9$ sends every vector in V_8 to the zero vector in $V_9 = 0$. For $t \notin [4, 9)$, $\alpha_t : V_t \rightarrow V_{t+1}$ is trivial since $V_t = 0$. Equivalently, we have that for all $t, s \in \mathbb{N}_0$ with $t \leq s$:

$$\alpha_{s,t} = \begin{cases} \text{id}_{\mathbb{Q}} & \text{if } t \in [4, 9) \\ 0 & \text{otherwise} \end{cases}$$

We claim that V_{\bullet} is isomorphic to the interval module $\mathbb{I}_{\bullet}^{[4,9)}$ over \mathbb{Q} as persistence modules.

As a sidenote, we can consider the elements of M to be \mathbb{Q} -linear combinations of x^t with $t \in [4, 9)$ since (x^5) being a graded submodule of $\mathbb{Q}[x]$ means that we can consider the homogeneous component of degree $t \in \mathbb{N}_0$ separately. More specifically: Let $N = (x^5)$. The homogeneous component $N_t x^t$ of N is given by $N_t x^t = \mathbb{Q} x^t$ if $t \geq 5$ and $N_t x^t = 0$ otherwise. Then, following Proposition 2.4.14,

$$\frac{\mathbb{Q}[x]}{(x^5)} \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} \left(\frac{\mathbb{Q} x^t}{N_t x^t} \right) = \bigoplus_{t \in \mathbb{N}_0} \left(\frac{\mathbb{Q}}{N_t} \right) x^t = \left(\bigoplus_{t=0}^4 (\mathbb{Q}/0) x^t \right) \oplus \left(\bigoplus_{t=5}^{\infty} (\mathbb{Q}/\mathbb{Q}) x^t \right) = \bigoplus_{t=0}^4 \mathbb{Q} x^t$$

Note that we write $\stackrel{\text{Ab}}{=}$ since the direct sums only consider the action of \mathbb{F} on the quotient module, not the action of $\mathbb{F}[x]$. Alternatively, since (x^5) is a graded submodule, the quotient $\mathbb{Q}[x] / \mathbb{Q}\langle x^5 \rangle$ makes all \mathbb{Q} -multiples of powers x^t with $t \geq 5$ trivial and each coset of $\mathbb{Q}[x] / \mathbb{Q}\langle x^5 \rangle$ is represented uniquely by a \mathbb{Q} -linear combination of x^t with $t \in \{1, \dots, 4\}$.

Finally, we state our theorem involving the isomorphism of categories between that of \mathbb{N}_0 -indexed persistence modules over \mathbb{F} and \mathbb{N}_0 -graded $\mathbb{F}[x]$ -modules.

Theorem 2.5.9. Fix a field \mathbb{F} . The pair Γ_{GrMod} and Γ_{Pers} determine an isomorphism of categories between the category $\mathbf{Pers}_{\mathbb{F}}$ of persistence modules over \mathbb{F} and the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ on \mathbb{N}_0 -graded $\mathbb{F}[x]$ -modules, i.e.

$$\Gamma_{\text{Pers}} \circ \Gamma_{\text{GrMod}} = \text{id}_{\mathbf{Pers}_{\mathbb{F}}} \quad \text{and} \quad \Gamma_{\text{GrMod}} \circ \Gamma_{\text{Pers}} = \text{id}_{\mathbf{GrMod}_{\mathbb{F}[x]}}$$

where $\text{id}_{\mathbf{Pers}_{\mathbb{F}}}$ and $\text{id}_{\mathbf{GrMod}_{\mathbb{F}[x]}}$ denote the identity functors on $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ respectively. Note that this also implies that Γ_{GrMod} and Γ_{Pers} form an equivalence of categories.

Remarks. (1) This theorem is a special case of [BM21, Theorem 2.21] where $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ are denoted as $\mathbf{Mod}_R^{\mathbf{P}}$ and $\text{Gr}^{\mathbf{P}}\text{-Mod}_{R[U_0]}$ respectively with $\mathbf{P} = \mathbf{Poset}(\mathbb{N}_0, \leq)$, $R = \mathbb{F}$, and $\mathbb{F}[x] = R[U_0]$.

(2) A weaker version of this theorem is presented in [ZC05, Theorem 3.1], which only claims a category equivalence between the subcategory of $\mathbf{Pers}_{\mathbb{F}}$ limited to *finite-type* persistence modules over \mathbb{F} and the subcategory of $\mathbf{GrMod}_{\mathbb{F}[x]}$ limited to *finitely-generated* graded $\mathbb{F}[x]$ -modules. [ZC05] claims that the Artin-Rees theorem in commutative algebra is sufficient for the proof.

Alternatively, [CK18] provides a detailed proof of [ZC05, Theorem 3.1] (which it calls the ZC Representation Theorem) without using the Artin-Rees theorem. A sketch of an alternative proof that uses the Artin-Rees theorem is provided in [CK18, Appendix C].

About the Correspondence between Algebraic Constructions.

Theorem 2.5.9 is particularly significant in this paper since it allows us to correspond algebraic constructions in $\mathbf{Pers}_{\mathbb{F}}$, i.e. involving persistence modules, to those in $\mathbf{GrMod}_{\mathbb{F}[x]}$, i.e. involving graded modules. In the second half of this section, we state the propositions that tell us which properties of the original persistence complex

are preserved as we go from the category of persistence modules, to that of graded modules, then back to that of persistence modules.

We begin with a statement involving isomorphisms in $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$.

Proposition 2.5.10. The functors $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ and $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$ preserve and reflect isomorphisms, i.e.

- i. A persistence morphism $\varphi_{\bullet} : V_{\bullet} \rightarrow W_{\bullet}$ between persistence modules V_{\bullet} and W_{\bullet} is a persistence isomorphism if and only if $\mathbf{\Gamma}_{\mathbf{GrMod}}(\varphi_{\bullet}) : \mathbf{\Gamma}_{\mathbf{GrMod}}(V_{\bullet}) \rightarrow \mathbf{\Gamma}_{\mathbf{GrMod}}(W_{\bullet})$ is a graded $\mathbb{F}[x]$ -module isomorphism.
- ii. A graded $\mathbb{F}[x]$ -module homomorphism $f : M \rightarrow N$ between graded $\mathbb{F}[x]$ -modules M and N is a graded isomorphism if and only if $\mathbf{\Gamma}_{\mathbf{Pers}}(f) : \mathbf{\Gamma}_{\mathbf{Pers}}(M) \rightarrow \mathbf{\Gamma}_{\mathbf{Pers}}(N)$ is a persistence isomorphism.

Proof. For part (i): For the forward direction, assume $\varphi_{\bullet} : V_{\bullet} \rightarrow W_{\bullet}$ is a persistence isomorphism. By [Rie16, Lemma 1.3.8], which states that all functors preserve isomorphisms, $\mathbf{\Gamma}_{\mathbf{GrMod}}(\varphi_{\bullet}) : \mathbf{\Gamma}_{\mathbf{GrMod}}(V_{\bullet}) \rightarrow \mathbf{\Gamma}_{\mathbf{GrMod}}(W_{\bullet})$ must be a graded $\mathbb{F}[x]$ -module isomorphism. For the converse, assume that $\mathbf{\Gamma}_{\mathbf{GrMod}}(\varphi_{\bullet})$ is a graded isomorphism. By Theorem 2.5.9, $\mathbf{\Gamma}_{\mathbf{Pers}} \circ \mathbf{\Gamma}_{\mathbf{GrMod}} = \text{id}_{\mathbf{Pers}_{\mathbb{F}}}$ and

$$(\mathbf{\Gamma}_{\mathbf{Pers}} \circ \mathbf{\Gamma}_{\mathbf{GrMod}})(V_{\bullet}) = V_{\bullet}, \quad (\mathbf{\Gamma}_{\mathbf{Pers}} \circ \mathbf{\Gamma}_{\mathbf{GrMod}})(W_{\bullet}) = W_{\bullet}, \quad \text{and} \quad (\mathbf{\Gamma}_{\mathbf{Pers}} \circ \mathbf{\Gamma}_{\mathbf{GrMod}})(\varphi_{\bullet}) = \varphi_{\bullet}.$$

By [Rie16, Lemma 1.3.8] on $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$, φ_{\bullet} must be a persistence isomorphism. A similar argument applies for part (ii). \blacksquare

Remark. This results holds even if $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ and $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$ only form an equivalence of categories since [Rie16, Theorem 1.5.9] states that any functor defining an equivalence of categories is full and faithful and [Rie16, Exercise 1.5.iv] states that full and faithful functors reflect isomorphisms.

Since a graded invariant factor decomposition of a graded module is given by a graded isomorphism, Proposition 2.5.10 tells us that the application of $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$ on such a decomposition will produce a persistence isomorphism. Moreover, we can show that this graded decomposition corresponds to an interval decomposition. Below, we describe the relationship between interval modules and the cyclic summands of graded invariant factor decompositions.

Lemma 2.5.11. Fix a field \mathbb{F} . The interval persistence modules correspond to shifted cyclic graded $\mathbb{F}[x]$ -modules, i.e. for all $t, s \in \mathbb{N}_0$,

$$\mathbb{I}_{\bullet}^{[s, \infty)} = \mathbf{\Gamma}_{\mathbf{Pers}}\left(\Sigma^s \mathbb{F}[x]\right) \quad \text{and} \quad \mathbb{I}_{\bullet}^{[s, s+t)} = \mathbf{\Gamma}_{\mathbf{Pers}}\left(\Sigma^s \left(\mathbb{F}[x] / (x^t)\right)\right)$$

where \mathbb{I}_{\bullet}^J denotes the interval modules over the interval $J \subseteq \mathbb{N}_0$ (see Definition 2.3.1), $\Sigma^s(-)$ denotes an upwards s -shift in grading (see Definition 2.4.10) and $(x^t) = \mathbb{F}[x]\langle x^t \rangle$ is the graded $\mathbb{F}[x]$ -module generated by x^t .

Proof. Let $t, s \in \mathbb{N}_0$. First, we want to show that $\mathbb{I}_{\bullet}^{[s, \infty)} \cong \mathbf{\Gamma}_{\mathbf{Pers}}(\Sigma^s \mathbb{F}[x])$. Note that $\Sigma^s \mathbb{F}[x] = \mathbb{F}[x]\langle x^s \rangle$ is a graded $\mathbb{F}[x]$ -module with homogeneous component given by $\mathbb{F}x^r$ if $r \geq s$ and trivial otherwise. The action of $\mathbb{F}[x]$ on $\Sigma^s \mathbb{F}[x]$ is given by $x \cdot x^r = x^{r+1}$ for all $r \geq s$. Equivalently, $x^{q-r} \cdot x^r = x^q$ for all $r, q \in [s, \infty)$ with $r \leq q$. Let $(V_{\bullet}, \alpha_{\bullet}) := \mathbf{\Gamma}_{\mathbf{Pers}}(\Sigma^s \mathbb{F}[x])$. For all $r, q \in \mathbb{N}_0$ with $r \leq q$, the vector space V_r and the structure map $\alpha_{q,r} : V_r \rightarrow V_q$ are given as follows:

$$V_r = \begin{cases} \mathbb{F} & \text{if } r \in [s, \infty) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \alpha_{q,r} = \begin{cases} \text{id}_{\mathbb{F}} & \text{if } r, q \in [s, \infty) \\ 0 & \text{otherwise} \end{cases}$$

By Definition 2.3.1, $V_{\bullet} = \mathbb{I}_{\bullet}^{[s, \infty)}$.

Next, we want to show that $\mathbb{I}_{\bullet}^{[s, s+t]} = \mathbf{\Gamma}_{\text{Pers}}(M)$ with $M := \Sigma^s(\mathbb{F}[x] / (x^t))$. By Proposition 2.4.13, the homogeneous component of $\mathbb{F}[x] / (x^t)$ of degree $r \in \mathbb{N}_0$ is given by the quotient of the homogeneous component of $\mathbb{F}[x]$ of degree r by that of $(x^t) = \mathbb{F}[x]\langle x^t \rangle$. That is, the homogeneous component $M_r x^r$ of M of degree $r \in \mathbb{N}_0$ is as follows:

$$M_r x^r = \begin{cases} 0 & \cong 0 & \text{if } r \in [0, s) \\ \mathbb{F}x^r / 0 & \cong \mathbb{F}x^r & \text{if } r \in [s, s+t) \\ \mathbb{F}x^r / \mathbb{F}x^r & \cong 0 & \text{if } r \in [s+t, \infty) \end{cases}$$

The action of $\mathbb{F}[x]$ on M is given by $x \cdot x^r = x^{r+1}$ for $r \in [s, s+t-1)$ and by $x \cdot x^{s+t-1} = 0$. Equivalently, for all $r, q \in \mathbb{N}_0$ with $r \leq q$, $x^{q-r} \cdot x^r = x^q$ if $r, q \in [s, s+t)$ and $x^{q-r} \cdot x^r = 0$ otherwise. Let $(V_{\bullet}, \alpha_{\bullet}) := \mathbf{\Gamma}_{\text{Pers}}(M)$. Then, the vector spaces V_r and structure maps $\alpha_{q,r} : V_r \rightarrow V_q$ of V_{\bullet} are given as follows:

$$V_r = \begin{cases} \mathbb{F} & \text{if } r \in [s, s+t) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \alpha_{q,r} = \begin{cases} \text{id}_{\mathbb{F}} & \text{if } r, q \in [s, s+t) \text{ with } r \leq q \\ 0 & \text{otherwise} \end{cases}$$

By Definition 2.3.1, $V_{\bullet} = \mathbb{I}_{\bullet}^{[s, s+t]}$. ■

Note that we can also determine the homogeneous component of degree $r \in \mathbb{N}_0$ of $M := \Sigma^s(\mathbb{F}[x] / (x^t))$ by using Proposition 2.4.15 and distributing the upwards shift $\Sigma^s(-)$ across the quotient, i.e.

$$\Sigma^s \left(\frac{\mathbb{F}[x]}{(x^t)} \right) \stackrel{\text{GrMod}}{\cong} \frac{\Sigma^s \mathbb{F}[x]}{\Sigma^s (x^t)} = \frac{\mathbb{F}[x]\langle x^s \rangle}{\mathbb{F}[x]\langle x^{t+s} \rangle} = \left\{ kx^r \in \mathbb{F}[x] : k \in \mathbb{F} \text{ and } r \in [s, s+t) \right\}$$

More rigorously, any coset of M as a quotient module can be represented uniquely by \mathbb{F} -linear combinations of powers x^r with $r \in [s, s+t)$. Following Remark 2.4.9, the only non-trivial homogeneous components of M are those of degree $r \in [s, s+t)$ given exactly by $\mathbb{F}x^r$.

To state the correspondence between graded invariant factor decompositions and interval decompositions of persistence modules, we need the following result involving finite direct sums.

Proposition 2.5.12. The functors $\mathbf{\Gamma}_{\text{GrMod}}(-)$ and $\mathbf{\Gamma}_{\text{Pers}}(-)$ distribute over finite direct sums, i.e. given persistence modules V_{\bullet} and W_{\bullet} and graded $\mathbb{F}[x]$ -modules M and N ,

$$\begin{aligned} \mathbf{\Gamma}_{\text{GrMod}}(V_{\bullet} \oplus_{\mathbf{P}} W_{\bullet}) &\stackrel{\text{GrMod}}{\cong} \mathbf{\Gamma}_{\text{GrMod}}(V_{\bullet}) \oplus_{\mathbf{G}} \mathbf{\Gamma}_{\text{GrMod}}(W_{\bullet}) \\ \mathbf{\Gamma}_{\text{Pers}}(M \oplus_{\mathbf{G}} N) &\stackrel{\text{Pers}}{\cong} \mathbf{\Gamma}_{\text{Pers}}(M) \oplus_{\mathbf{P}} \mathbf{\Gamma}_{\text{Pers}}(N) \end{aligned}$$

where $\oplus_{\mathbf{G}}$ refers to a direct sum of graded $\mathbb{F}[x]$ -modules and $\oplus_{\mathbf{P}}$ refers to that of persistence modules.

Remark. This can be seen as a consequence $\mathbf{\Gamma}_{\text{GrMod}}(-)$ and $\mathbf{\Gamma}_{\text{Pers}}(-)$ forming an equivalence of categories by Theorem 2.5.9. We are unable to provide a proof of this (or a reference for such) at this moment but we have outlined two possible arguments below.

1. We can use [Rie16, Lemma 3.3.6], which roughly states that an equivalence of categories preserves all limits and colimits, as described in [Rie16, Chapter 3]. Then, as argued in the proof of [Rie16, Proposition 4.5.10], direct sums and kernels in abelian categories are finite limits. Since $\mathbf{\Gamma}_{\text{GrMod}}(-)$ and $\mathbf{\Gamma}_{\text{Pers}}(-)$ form an equivalence of categories by Theorem 2.5.9, these must both preserve finite direct sums.
2. Alternatively, we can use [Rie16, Proposition 4.4.5], which roughly states that any equivalence of categories determine an adjoint equivalence (also see [Rie16, Footnote 37, p30]), and [Rie16,

Corollary 4.5.11], which then implies that these functors are exact and additive.

Note that [Rie16] defines an additive functor to be a functor that preserves direct sums. In contrast, [Rot88, Corollary 5.88] states this to be a property of additive functors. A more detailed discussion on properties preserved by exact and additive functors can also be found in [Bai17] with the above result given as part of [Bai17, Theorem 27].

We are unsure if the above arguments apply for arbitrary direct sums. However, this is not an issue for us since arbitrary direct sums of persistence modules are not needed for the matrix reduction algorithm for persistent homology.

Below, we state the correspondence between interval decompositions of persistence modules and graded invariant factor decompositions of graded modules.

Corollary 2.5.13. Let $(V_\bullet, \alpha_\bullet)$ be a persistence module. Assume that $M := \Gamma_{\text{GrMod}}(V_\bullet)$ admits the following graded invariant factor decomposition, as described in the Graded Structure Theorem (Theorem 4.3.1):

$$\Gamma_{\text{GrMod}}(V_\bullet) = M \stackrel{\text{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \dots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \dots \oplus \Sigma^{s_m} \mathbb{F}[x]$$

with invariant factors $\{x^{t_1}, \dots, x^{t_r}\}$ and grading shifts $\{s_1, \dots, s_r, \dots, s_m\}$. Then, V_\bullet admits the following interval decomposition, as defined in Definition 2.3.4,

$$V_\bullet = \Gamma_{\text{Pers}}(M) \stackrel{\text{Pers}}{\cong} \mathbb{I}_\bullet^{[s_1, s_1+t_1)} \oplus \dots \oplus \mathbb{I}_\bullet^{[s_r, s_r+t_r)} \oplus \mathbb{I}_\bullet^{[s_{r+1}, \infty)} \oplus \dots \oplus \mathbb{I}_\bullet^{[s_m, \infty)}$$

with persistence barcode given by $\text{Bar}(V_\bullet) = \{[s_1, s_1 + t_1), \dots, [s_r, s_r + t_r)\} \cup \{[s_{r+1}, \infty), \dots, [s_m, \infty)\}$.

Proof. Proposition 2.5.10 implies that the graded isomorphism of M to its graded invariant factor decomposition determines a persistence isomorphism by application of $\Gamma_{\text{Pers}}(-)$. Then, we distribute $\Gamma_{\text{Pers}}(-)$ over the graded decomposition using Proposition 2.5.12 and apply Lemma 2.5.11 to each cyclic summand. ■

The Graded Structure Theorem (Theorem 4.3.1) guarantees the existence of graded invariant factor decompositions for finitely generated graded $\mathbb{F}[x]$ -modules. This, along with Corollary 2.5.13, gives us a benchmark for the existence of interval decompositions. We state this in more detail below.

Proposition 2.5.14. Let (V_\bullet, α) be a persistence module over \mathbb{F} . If V_\bullet is a finite-type persistence module, then $\Gamma_{\text{GrMod}}(V_\bullet)$ is a finitely generated graded $\mathbb{F}[x]$ -module and V_\bullet admits an interval decomposition.

Proof. Assume $(V_\bullet, \alpha_\bullet)$ is a finite-type persistence module. By Definition 2.1.6, the vector space V_t is finite dimensional for all $t \in \mathbb{N}_0$ and V_\bullet is constant on $[N, \infty)$ for some $N \in \mathbb{N}_0$. For each $t \in [0, N]$, let \mathcal{B}_t be a basis of V_t . Let $\mathcal{B} := \bigcup_{t=0}^N \mathcal{B}_t$. Note that \mathcal{B} is a finite set.

Let $M := \Gamma_{\text{GrMod}}(V_\bullet)$. By definition of $\Gamma_{\text{GrMod}}(-)$, the homogeneous component of M of degree $t \in \mathbb{N}_0$ is given exactly by $V_t x^t$ and the action of $\mathbb{F}[x]$ on M is given by $x^s \cdot v_t x^t = \alpha_{s+t,t}(v_t) x^{t+s}$ with $\alpha_{s+t,t} : V_t \rightarrow V_{s+t}$ for all $v_t \in V_t$ and $t \in \mathbb{N}_0$. Let $t \in [0, N]$. Note that for all $b \in \mathcal{B}_t$, $b x^t \in V_t x^t \subseteq M$ by construction. Define $\mathcal{B}_t x^t := \{b x^t \in M : b \in \mathcal{B}_t\}$ and $\mathcal{B}^{\text{Gr}} := \bigcup_{t=0}^N \mathcal{B}_t x^t$. We want to show that \mathcal{B}^{Gr} generates M , i.e. $M = \{f(x) \cdot b x^t : f(x) \in \mathbb{F}[x], b x^t \in \mathcal{B}^{\text{Gr}}\}$.

Since $M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} V_t x^t$, $m = \sum_{t \in \mathbb{N}_0} v_t x^t$ for a unique set $\{v_t\}_{t \in \mathbb{N}_0}$ of elements $v_t \in V_t$, only finitely many of which are nonzero. Then, it suffices to check if each homogeneous component $v_t x^t$ is generated by \mathcal{B}^{Gr} . Let $m = v_t x^t \in V_t x^t \subseteq M$ for some $t \in \mathbb{N}_0$.

1. Assume $t \in [0, N]$. Since \mathcal{B}_t is a basis of V_t by assumption, $\mathcal{B}_t x^t$ must also be a basis of $V_t x^t$, viewing $V_t x^t$ as an \mathbb{F} -vector space. Therefore, there exists an \mathbb{F} -linear combination in $\mathcal{B}_t x^t \subseteq \mathcal{B}^{\text{Gr}}$ that equals $v_t x^t$.
2. Assume $t \in [N + 1, \infty)$. By assumption of V_\bullet being constant on $[N, \infty)$, the structure map $\alpha_{t,N} : V_N \rightarrow V_t$ must be an \mathbb{F} -vector space isomorphism. Let $v_N \in V_N$ such that $\alpha_{t,N}(v_N) = v_t$. Then, $v_N x^N \in M$ and $x^{t-N} \cdot v_N x^N = \alpha_{t,N}(v_N) x^t = v_t x^t$.

Therefore, any element $m \in M$ is an $\mathbb{F}[x]$ -linear combination on \mathcal{B}^{Gr} . Since $M \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} V_t x^t$ by definition, all $\mathbb{F}[x]$ -linear combinations on \mathcal{B}^{Gr} must be on M . Then,

$$M = \mathbb{F}[x] \mathcal{B}^{\text{Gr}} = \left\{ \sum_{t=0}^{\infty} f_t(x) \cdot b_t x^t : f_t(x) \in \mathbb{F}[x] \text{ finitely many of which are nonzero and } b_t x^t \in \mathcal{B}^{\text{Gr}} \right\}$$

and \mathcal{B} is a finite system of generators for M . Therefore, M is a finitely generated graded $\mathbb{F}[x]$ -module and admits a graded invariant factor decomposition by the Graded Structure Theorem (Theorem 4.3.1). By Corollary 2.5.13, there exists an interval decomposition for V_\bullet . \blacksquare

Remark. If we consider functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Mod}_{\mathbb{Z}}$ to be persistence modules and let $\mathbf{Pers}_{\mathbb{Z}}$ be the corresponding functor category, we can modify Theorem 2.5.9 to state an isomorphism of categories between $\mathbf{Pers}_{\mathbb{Z}}$ and the category $\mathbf{GrMod}_{\mathbb{Z}[x]}$ of \mathbb{N}_0 -graded $\mathbb{Z}[x]$ -modules. However, while the graded $\mathbb{Z}[x]$ -module constructed by applying $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ to some finite-type persistence module may be finitely generated, the Graded Structure Theorem still would not apply to said graded $\mathbb{Z}[x]$ -module since $\mathbb{Z}[x]$ is not a PID. Therefore, we cannot use Proposition 2.5.14 for the existence of interval decompositions for persistence modules of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Mod}_{\mathbb{Z}}$.

Finally, we state how the chain homology of persistence modules and that of graded modules interact with the equivalence of categories given by $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ and $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$ below.

Proposition 2.5.15. The functors $\mathbf{\Gamma}_{\mathbf{GrMod}}$ and $\mathbf{\Gamma}_{\mathbf{Pers}}$ preserve chain complexes and commute with the homology functor, i.e. let $H_n^{\text{Pers}}(-) : \mathbf{Ch-Pers}_{\mathbb{F}} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ denote the n^{th} homology functor on persistence complexes and let $H_n^{\text{Gr}}(-) : \mathbf{Ch-GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ denote that on graded chain complexes. Then,

- i. Let $(V_\bullet, \partial_\bullet) = (V_n^\bullet, \partial_n^\bullet)_{n \in \mathbb{Z}}$ be a persistence complex with persistence modules V_n^\bullet over \mathbb{F} and persistence morphisms $\partial_n^\bullet : V_n^\bullet \rightarrow V_{n-1}^\bullet$. For all $n \in \mathbb{Z}$, there exists a graded isomorphism such that

$$\mathbf{\Gamma}_{\mathbf{GrMod}}\left(H_n^{\text{Pers}}(V_\bullet, \partial_\bullet)\right) \stackrel{\mathbf{GrMod}}{\cong} H_n^{\text{Gr}}(M_\bullet, d_\bullet)$$

where $(M_\bullet, d_\bullet) = (M_n, d_n)_{n \in \mathbb{Z}}$ is the graded chain complex generated by component-wise application of $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ with $M_n := \mathbf{\Gamma}_{\mathbf{GrMod}}(V_n^\bullet)$ and $d_n : M_n \rightarrow M_{n-1}$ by $d_n := \mathbf{\Gamma}_{\mathbf{GrMod}}(\partial_n^\bullet)$ for all $n \in \mathbb{Z}$.

- ii. Let $(M_\bullet, d_\bullet) = (M_n, d_n)_{n \in \mathbb{Z}}$ be a graded chain complex with graded $\mathbb{F}[x]$ -modules M_n and graded homomorphisms $d_n : M_n \rightarrow M_{n-1}$. For all $n \in \mathbb{Z}$, there exists a persistence isomorphism such that

$$\mathbf{\Gamma}_{\mathbf{Pers}}\left(H_n^{\text{Gr}}(M_\bullet, d_\bullet)\right) \stackrel{\mathbf{Pers}}{\cong} H_n^{\text{Pers}}(V_\bullet, \partial_\bullet)$$

where $(V_\bullet, \partial_\bullet) = (V_n^\bullet, \partial_n^\bullet)_{n \in \mathbb{Z}}$ is the persistence complex generated by component-wise application of $\mathbf{\Gamma}_{\mathbf{Pers}}$, with $V_n^\bullet := \mathbf{\Gamma}_{\mathbf{Pers}}(M_n)$ and $\partial_n^\bullet := \mathbf{\Gamma}_{\mathbf{Pers}}(d_n)$ for all $n \in \mathbb{Z}$.

Remark. We are unable to provide a rigorous proof (or a reference for such) at this moment. However, if we can claim that $\mathbf{\Gamma}_{\mathbf{GrMod}}(-)$ and $\mathbf{\Gamma}_{\mathbf{Pers}}(-)$ are exact functors, following the remarks under Proposition

2.5.12, then the above result is stated as part of [Bai17, Theorem 27]. A similar result is listed as [Rot08, Exercise 6.8, p339], which claims that exact additive functors between categories of modules over different rings commute with homology. We believe this applies more generally to exact additive functors between abelian categories, as discussed on the introduction of [Rot88, Chapter 6].

We use the proposition above later in Section 3.3 in the context of simplicial homology.

Chapter 3. Filtrations and Persistent Homology

Persistent homology theory is interested in the characterization, calculation, and representation of the persistent homology of filtrations of topological spaces.

Generally speaking, a *filtration* refers to any collection $\{C_t : t \in \Lambda\}$ of objects, indexed over some set Λ , along with a set of subobject relations. The indexing set Λ is usually equipped with a partial order \leq that determines the expected set of subobject relations on $\{C_t : t \in \Lambda\}$, i.e. the subobject relation $C_t \subseteq C_s$ is present if and only if $t \leq s$ in the poset (Λ, \leq) . Then, the persistent homology of a filtration $\{X_t : t \in \Lambda\}$ of topological spaces X_t refers to some characterization of the following collections over all dimensions $n \in \mathbb{N}_0$:

$$\left\{ \begin{array}{c} \text{homology groups} \\ H_n(X_t) : t \in \Lambda \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{c} \text{maps on homology induced by inclusions} \\ H_n(X_t) \rightarrow H_n(X_s) : t, s \in \Lambda \text{ with } t \leq s \end{array} \right\}$$

In practice, the homology groups are taken with coefficients in a field \mathbb{F} (usually $\mathbb{F} = \mathbb{Z}_p$ for prime p), and the calculation of persistent homology is done at the level of ranks, i.e. we want to find the following quantities:

$$\left\{ \text{rank}(H_n(X_t; \mathbb{F})) : t \in \Lambda \right\} \quad \text{and} \quad \left\{ \text{rank}\left(H_n(X_t; \mathbb{F}) \xrightarrow{i_*} H_n(X_s; \mathbb{F})\right) : t, s \in \Lambda \text{ with } t \leq s \right\}$$

These are then represented succinctly using a multiset of intervals in Λ (as a poset) called the *persistence barcode* of the filtration, which is different albeit very similar to the persistence barcode of a persistence module.

In this chapter, we discuss the key ideas and constructions presented in the paper *Computing Persistent Homology* [ZC05] by Afra Zomorodian and Gunnar Carlsson. In particular, we restrict our attention to \mathbb{N}_0 -indexed filtrations $K_\bullet := \{K_t\}_{t \in \mathbb{N}_0}$ of (abstract) simplicial complexes K_t such that $K_t \subseteq K_{t+1}$ for all $t \in \mathbb{N}_0$ and study the following sequence of homology groups and induced maps:

$$H_n(K_0; \mathbb{F}) \xrightarrow{i_0^*} H_n(K_1; \mathbb{F}) \xrightarrow{i_1^*} H_n(K_2; \mathbb{F}) \xrightarrow{i_2^*} H_n(K_3; \mathbb{F}) \xrightarrow{i_3^*} \dots$$

where $i_t : K_t \rightarrow K_{t+1}$ denotes the inclusion map for all $t \in \mathbb{N}_0$. We discuss these notions relative to the characterization of persistence modules presented in Chapter 2, i.e. as functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$. This chapter is structured as follows:

Section 3.1. We characterize simplicial filtrations, i.e. filtrations of simplicial complexes, as functors of the form $K_\bullet : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$ where $\mathbf{A-Simp}$ denotes the category of (abstract) simplicial complexes and simplicial maps. We also discuss the notion of *finite-type filtrations*.

Section 3.2. We formalize the notion of the persistent homology of a simplicial filtration K_\bullet with coefficients in a field \mathbb{F} by defining a persistence module over \mathbb{F} called the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F}) : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$ for each dimension $n \in \mathbb{Z}$.

We also identify a number of relevant terminology and interpretations involving $H_n(K_\bullet; \mathbb{F})$, and show that the interval decomposition of $H_n(K_\bullet; \mathbb{F})$ exists assuming K_\bullet is a finite-type filtration.

Section 3.3. We extend the simplicial chain complex of (abstract) simplicial complexes, as discussed in Section 1.2, to the case of persistence modules and construct a chain complex of persistence modules called the simplicial persistence complex $C_*(K_\bullet; \mathbb{F}) = (C_n(K_\bullet; \mathbb{F}), \partial_n^\bullet)_{n \in \mathbb{Z}}$. The n^{th} chain homology of this persistence complex is then shown to be isomorphic to $H_n(K_\bullet; \mathbb{F})$.

We also discuss how the isomorphism of categories between $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$ discussed in Section 2.5 allows us to calculate the n^{th} chain homology of $C_*(K_\bullet; \mathbb{F})$ at the level of graded $\mathbb{F}[x]$ -modules. This will serve as the basis for the *matrix reduction algorithm for persistent homology*, later discussed in Chapter 4.

Section 3.1. Filtrations of Simplicial Complexes

In this section, we discuss specific \mathbb{N}_0 -indexed collections of simplicial complexes $\{K_t\}_{t \in \mathbb{N}_0}$ called *filtrations* and provide a characterization of these as functors of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$.

Definition 3.1.1. A **simplicial filtration** K_\bullet of a simplicial complex K is a functor $K_\bullet : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$ with the following properties:

- i. For all $t \in \mathbb{N}_0$, $K_\bullet(t)$ is a subcomplex of K . For brevity, we often write $K_t := K_\bullet(t)$, i.e. the bullet is replaced with $t \in \mathbb{N}_0$. The **index** $t \in \mathbb{N}_0$ of the simplicial complex K_t in K_\bullet is sometimes called the **scale** or **parameter** of K_t in K_\bullet .
- ii. For all $t, s \in \mathbb{N}_0$ with $t \leq s$, K_t is a subcomplex of K_s and $K_\bullet(t \rightarrow s) : K_t \rightarrow K_s$ is exactly the inclusion map $K_t \hookrightarrow K_s$. Let $i^{s,t} : K_t \rightarrow K_s$ and $i^t : K_t \rightarrow K_{t+1}$ denote the inclusion maps $i^{s,t} := K_\bullet(t \rightarrow s)$ and $i^t := K_\bullet(t \rightarrow t+1)$ respectively.
- iii. For each simplex $\sigma \in K$, there must exist $t \in \mathbb{N}_0$ such that σ is a simplex of K_t .

When the relation to a simplicial complex K is clear from context, we may say K_\bullet is a filtration of K for brevity.

Remark. Condition (iii) above determines that each simplicial filtration K_\bullet corresponds to a unique simplicial complex, i.e. $K = \bigcup_{t \in \mathbb{N}_0} K_\bullet(t) = \bigcup_{t \in \mathbb{N}_0} K_t$ if K_\bullet is a filtration of K . It serves a similar purpose as condition Definition 1.1.1(i) for simplicial complexes.

Let K_\bullet be a filtration of a simplicial complex K . Observe that the subcomplex relation $K_t \subseteq K_s$ for all $t, s \in \mathbb{N}_0$ with $t \leq s$ stated in Definition 3.1.1(ii) implies that for all $t \in \mathbb{N}_0$, K_t is a subcomplex of K_{t+1} . Therefore, the simplicial filtration K_\bullet determines the following nested sequence of simplicial complexes:

$$K_0 \subseteq K_1 \subseteq K_2 \subseteq K_3 \subseteq K_4 \subseteq \cdots$$

We can also show that increasing nested sequences of simplicial complexes determine simplicial filtrations, as we expect it should. We state this in a proposition below.

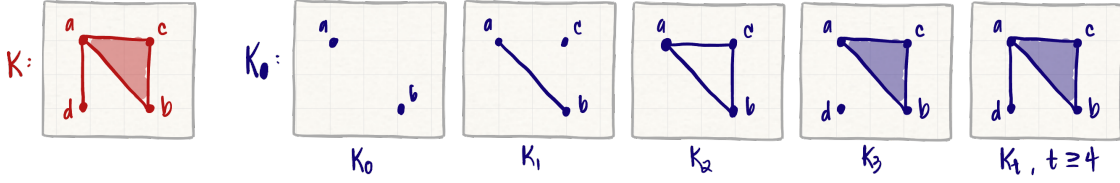
Proposition 3.1.2. Let $\{K_t\}_{t \in \mathbb{N}_0}$ be a collection of simplicial complexes such that $K_t \subseteq K_{t+1}$ for all $t \in \mathbb{N}_0$. Then, $\{K_t\}_{t \in \mathbb{N}_0}$ determines a filtration K_\bullet of the simplicial complex $K := \bigcup_{t \in \mathbb{N}_0} K_t$ by $K_\bullet(t) := K_t$.

Proof. Note that the union of simplicial complexes is a simplicial complex. Therefore, K is well-defined. For each $t \in \mathbb{N}_0$, let $K_\bullet(t) := K_t$. Observe that K_t is a subcomplex of K . For all $t, s \in \mathbb{N}_0$ with $t \leq s$, define $K_\bullet(t \rightarrow s)$ to be the inclusion map $K_t \hookrightarrow K_s$, which exists since $K_t \subseteq K_s$ by assumption.

Let $t, r, s \in \mathbb{N}_0$ such that $t \leq r \leq s$. Since the composition $(K_t \hookrightarrow K_s) \circ (K_s \hookrightarrow K_r)$ of inclusion maps is exactly the inclusion $K_t \hookrightarrow K_r$, the composition axiom as stated in Definition A4.4(iii) is satisfied. Since the inclusion map $K_t \hookrightarrow K_t$ is exactly the identity map, the identity axiom as stated in Definition A4.4(iv) is satisfied. Therefore, K_\bullet is a well-defined functor. The conditions of Definition 3.1.1 are also all satisfied. ■

The subcomplex relations also allow us to characterize a simplicial filtration K_\bullet by identifying the simplicial complex K_0 at index $0 \in \mathbb{N}_0$ and specifying the collection $\{\sigma_{t,i}\}$ of simplices that is appended to K_{t-1} to form $K_t = K_{t-1} \cup \{\sigma_{t,i}\}$ for each $t \geq 1$. Note that each K_t is required to be a simplicial complex by definition and sequentially appending an arbitrary set of simplices to some simplicial complex L generally does not result in a simplicial filtration. We provide an example of a simplicial filtration below.

Example 3.1.3. Let the simplicial complex K with vertex set $\text{Vert}(K) = \{a, b, c, d\}$ and let the filtration K_\bullet on K be given by the following illustrations:

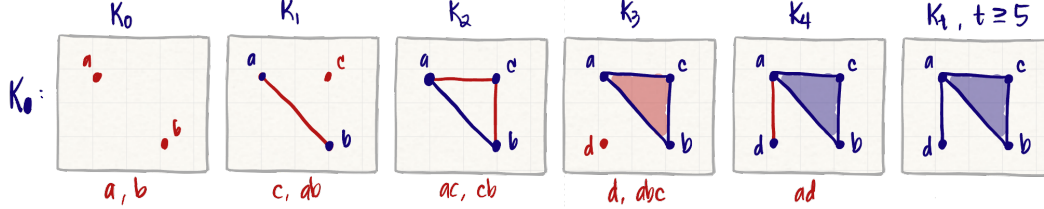


The simplicial complexes K_t of K_\bullet can be described abstractly as follows, with the simplices of K_t written as strings of vertices, following the remarks under Definition 1.1.1.

$$K_t = \begin{cases} \{a, b\} & \text{if } t = 0 \\ \{a, b, c, ab\} & \text{if } t = 1 \\ \{a, b, c, ab, ac, bc\} & \text{if } t = 2 \\ \{a, b, c, d, ab, ac, bc, abc\} & \text{if } t = 3 \\ \{a, b, c, d, ab, ac, bc, ad, abc\} & \text{if } t \geq 4 \end{cases}$$

Note that the abstract description of K_t can be cumbersome to work with by hand, e.g. confirming that each element corresponds to a simplicial complex and that the collection corresponds to a filtration can become tedious even for a relatively small number of simplices.

In the illustration below, we describe the filtration K_\bullet by specifying which simplices are added as we increase the index $t \in \mathbb{N}_0$. In particular, for each $t \in \mathbb{N}_0$, the simplices colored in red are the simplices in K_t that are not present in K_{t-1} , with K_{-1} interpreted to be \emptyset . These simplices are also listed in red below the illustration of each K_t .



Observe that there are no simplices added to K_t for $t \geq 5$.

We would like to point out that the subcomplex relations on the simplicial complexes of a filtration K_\bullet can be seen as a total order on the set $\{K_t\}_{t \in \mathbb{N}_0}$ induced by the total order on \mathbb{N}_0 . That is, we can define a total order \leq_c on the set of simplicial complexes by defining $K \leq_c L$ if and only if K is a subcomplex of L . Since functors with $\mathbf{Poset}(\mathbb{N}_0, \leq)$ as the domain category can generally be represented using a sequence with arrows (see discussion on poset categories in Appendix A4), we can interpret an arrow $K_t \rightarrow K_s$ to refer to the subcomplex relation $K_t \subseteq K_s$, i.e.

$$K_0 \xrightarrow{i^0} K_1 \xrightarrow{i^1} K_2 \xrightarrow{i^2} K_3 \xrightarrow{i^3} \dots \quad \text{corresponds to} \quad K_0 \subseteq K_1 \subseteq K_2 \subseteq K_3 \subseteq \dots$$

One consequence of this is the use of colloquial language when talking about filtrations. For example, *going up* a filtration usually implies going from a simplicial complex K_t from some $t \in \mathbb{N}_0$ and then considering a simplicial complex K_s at a higher index $t < s$.

Note that in most introductory literature for persistent homology, simplicial filtrations are generally represented using *finite* nested sequences. That is, filtrations are sometimes described to be finite collections

$\{K_0, K_1, \dots, K_T\}$ of simplicial complexes such that

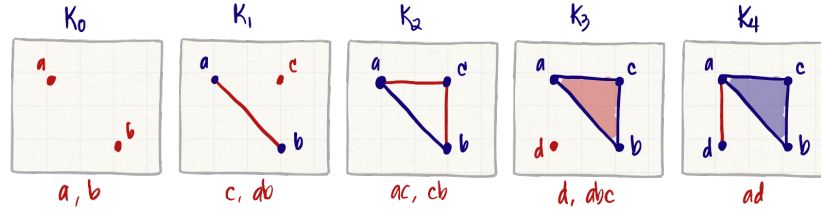
$$K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_T$$

This motivates the following terminology.

Definition 3.1.4. Let K_\bullet be a simplicial filtration. We say that K_\bullet is **constant on an interval** $I \subseteq \mathbb{N}_0$ if for all $t, s \in I$ with $t \leq s$, $K_t = K_s$. We say that K_\bullet is **finite-type** if K_t is a finite simplicial complex for all $t \in \mathbb{N}_0$, and there exists $T \in \mathbb{N}_0$ such that K_\bullet is constant on $[T, \infty)$.

Note that these definitions mimic those of Definition 2.1.6, which defines the same terms for the case of persistence modules. Observe that if K_\bullet is a finite-type filtration, then there can only be finitely many distinct simplices in K_\bullet . As we will see in Section 3.3, this finite-type condition on filtrations implies that the persistence modules constructed for simplicial persistent homology are also of finite-type. We provide an example of a finite-type filtration below.

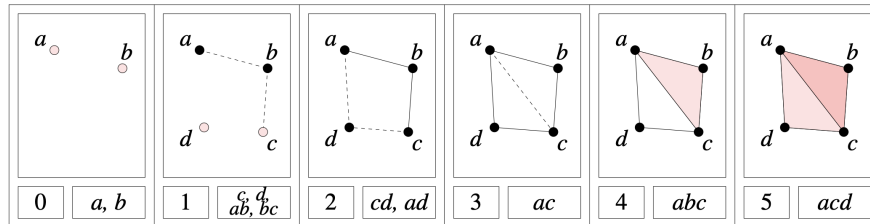
Example 3.1.5. The simplicial filtration K_\bullet defined in Example 3.1.3 is a finite-type filtration that is constant on $[4, \infty)$. In other words, knowing that K_\bullet is constant on $[4, \infty)$, we can determine K_\bullet using only the following illustrations of simplicial complexes:



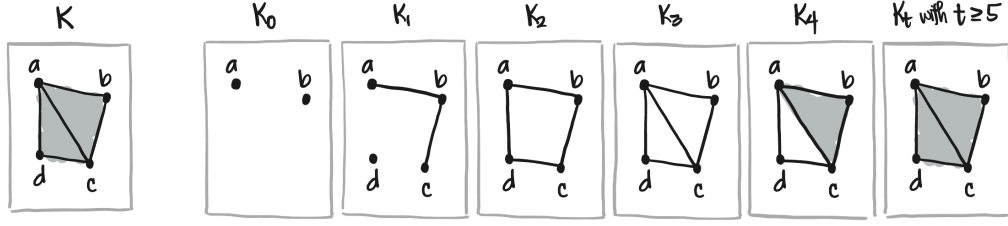
The assumption that K_\bullet is constant on $[4, \infty)$ determines that for all $t \geq 5$, $K_t = K_4$.

In practice, simplicial filtrations are usually constructed to be filtrations on *finite* simplicial complexes. Relative to Definition 3.1.1, we assume that a filtration described this way is constant on $[T, \infty)$. We provide an example of this below.

Example 3.1.6. Given below is a copy of [ZC05, Figure 1], which describes a nested sequence of simplicial complexes $K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_5$. Under each illustration of a simplicial complex, the index $t \in \{0, 1, \dots, 5\}$ in K_t is denoted on the bottom-left, and the list of simplices present in K_t but not on K_{t-1} is on the bottom-right (with $K_{-1} := \emptyset$). The 0-simplices in said list for each K_t are drawn with **light red** shaded circles, the 1-simplices with *dashed lines*, and the 2-simplices with **light red** shaded triangles.



Relative to Definition 3.1.1, this determines a filtration K_\bullet on the simplicial complex $K := K_5$ with $K_\bullet(t) = K_t$ illustrated below for all $t \in \mathbb{N}_0$:



Observe that K_\bullet is constant on $[5, \infty)$.

Filtrations on finite simplicial complexes are necessarily of finite-type. We state this in more detail below.

Lemma 3.1.7. Let K_\bullet be a filtration on a simplicial complex K . If K is a finite simplicial complex, then K_\bullet is a finite-type filtration and there exists some $T \in \mathbb{N}_0$ such that K_\bullet is constant on $[T, \infty)$.

Proof. Assume that K is a finite simplicial complex and let K_\bullet be some filtration on K . For all $t \in \mathbb{N}_0$, $K_\bullet(t) = K_t$ must be a finite simplicial complex since $K_t \subseteq K$ by Definition 3.1.1(i) and K is finite by assumption.

Assume, for the sake of contradiction, that there does not exist $T \in \mathbb{N}_0$ such that K_\bullet is constant on $[T, \infty)$. Let $r_0 = 0$. For each $t \geq 1$, let $r_t \in \mathbb{N}_0$ be such that $r_{t-1} \leq r_t$ and $K_\bullet(r_{t-1}) \neq K_\bullet(r_t)$. Note that r_t exists since K_\bullet cannot be constant on $[r_{t-1}, \infty)$ by assumption. Since $K_\bullet(r_{t-1}) \subseteq K_\bullet(r_t)$ by Definition 3.1.1(ii), there exists a simplex $\sigma_t \in K$ such that $\sigma_t \in K_\bullet(r_t)$ and $\sigma_t \notin K_\bullet(r_{t-1})$. This inductive process constructs a collection $\{\sigma_0, \sigma_1, \sigma_2, \dots\}$ of simplices of K such that each σ_t is distinct. Since this process can continue indefinitely by assumption of the non-existence of $T \in \mathbb{N}_0$, we can construct an infinite collection $\{\sigma_t\}_{t=0}^\infty$ of distinct simplices of K . Since $\bigcup_{t=0}^\infty \sigma_t \subseteq K$, K must be infinite. This contradicts the assumption that K is finite. ■

As a final remark, we want to point out that, like with persistence modules, the definition of a filtration can be generalized to allow for any poset category to be the domain category. That is, we can define filtrations to be functors of the form $\mathbf{Poset}(P, \leq) \rightarrow \mathbf{A-Simp}$ for any poset (P, \leq) .

Section 3.2. The Persistent Homology of Filtrations

In this section, we define the persistent homology of filtrations in terms of persistence modules and discuss the interval decompositions of persistence modules in the context of persistent homology. We start with a definition of persistent homology of filtrations, adapted from [Bau21].

Definition 3.2.1. Let $n \in \mathbb{Z}$. The n^{th} **persistent homology module** $H_n(K_\bullet; \mathbb{F})$ of a filtration K_\bullet **with coefficients in** \mathbb{F} is the persistence module $H_n(K_\bullet; \mathbb{F}) : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{F}}$ given by the following functor composition:

$$H_n(K_\bullet; \mathbb{F}) := H_n(-; \mathbb{F}) \circ K_\bullet$$

where $H_n(-; \mathbb{F}) : \mathbf{A-Simp} \rightarrow \mathbf{Vect}_{\mathbb{F}}$ denotes the n^{th} simplicial homology functor with \mathbb{F} coefficients. The n^{th} **persistence barcode** $\text{Bar}_n(K_\bullet; \mathbb{F})$ of K_\bullet **with coefficients in** \mathbb{F} is the persistence barcode of $H_n(K_\bullet; \mathbb{F})$ as a persistence module, i.e. $\text{Bar}_n(K_\bullet; \mathbb{F}) := \text{Bar}(H_n(K_\bullet; \mathbb{F}))$.

We have some remarks regarding the notation and terminology of the vector spaces and structure maps of $H_n(K_\bullet; \mathbb{F})$, taken as a persistence module, relative to our characterization of persistence modules by Definition 2.1.1.

1. Observe that $H_n(K_\bullet; \mathbb{F}) = H_n(K_t; \mathbb{F})$ for all $t \in \mathbb{N}_0$. This is consistent with the convention of replacing the bullet (\bullet) of a persistence module $V_\bullet = (V_\bullet, \alpha_\bullet)$ to denote the vector space V_t at index $t \in \mathbb{N}_0$. We may also call the index $t \in \mathbb{N}_0$, in the context of $H_n(K_\bullet; \mathbb{F})$ as a persistent homology module, as the *scale* or *parameter* of $H_n(K_t; \mathbb{F})$.
2. The structure maps of a filtration K_\bullet , as stated in Definition 3.1.1, are denoted as $i^{s,t} : K_t \rightarrow K_s$. Following the notation for the induced maps on homology, the structure maps of $H_n(K_\bullet; \mathbb{F})$ are denoted as $i_*^{s,t} : H_n(K_t; \mathbb{F}) \rightarrow H_n(K_s; \mathbb{F})$. Relative to the notation $V_\bullet = (V_\bullet, \alpha_\bullet)$ for an arbitrary persistence module, $H_n(K_\bullet; \mathbb{F}) = (H_n(K_\bullet; \mathbb{F}), i_*)$.

Since we may be dealing with homology classes in $H_n(K_t; \mathbb{F})$ of $H_n(K_\bullet; \mathbb{F})$ across different indices $t \in \mathbb{N}_0$, we identify some alternative notation below.

Remark 3.2.2. We modify the coset notation for the homology classes to include the index $t \in \mathbb{N}_0$ as a subscript when discussed in the context of $H_n(K_\bullet; \mathbb{F})$, i.e. we write $[\sigma]_t$ to refer to the homology class $[\sigma] \in H_n(K_t; \mathbb{F})$ as an element of the vector space of $H_n(K_\bullet; \mathbb{F})$ specifically at index $t \in \mathbb{N}_0$.

Note that σ in $[\sigma]_t \in H_n(K_t; \mathbb{F})$ denotes a cycle representative $\sigma \in \ker(\partial_n^t)$, with $\partial_n^t : C_n(K_t; \mathbb{F}) \rightarrow C_{n-1}(K_t; \mathbb{F})$ referring to the n^{th} simplicial boundary map of K_t .

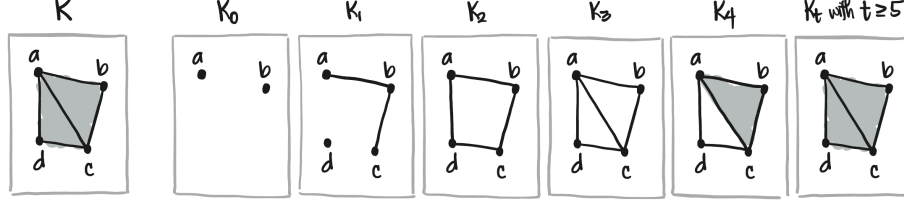
Since we defined persistent homology modules using functor composition, we can interpret the persistent homology module $H_n(K_\bullet; \mathbb{F})$ to be calculated using simplicial homology first at each index $t \in \mathbb{N}_0$, and second by assembling the resulting homology groups into a persistence module, i.e. we have two separate operations, illustrated as follows:

$$\begin{array}{ccccc} \mathbf{Poset}(\mathbb{N}_0, \leq) & \xrightarrow{\text{filtration } K_\bullet} & \mathbf{A-Simp} & \xrightarrow{\text{simplicial homology } H_n(-; \mathbb{F})} & \mathbf{Vect}_{\mathbb{F}} \\ t & \longmapsto & K_t & \longmapsto & H_n(K_t; \mathbb{F}) \\ \text{(scale or index)} & & \text{(simplicial complex)} & & \text{(homology group)} \end{array}$$

As such, the terminology for simplicial complexes and simplicial homology also extends to the homology groups of a persistent homology module at some specific index $t \in \mathbb{N}_0$.

Below, we provide an example of a persistent homology module of a filtration.

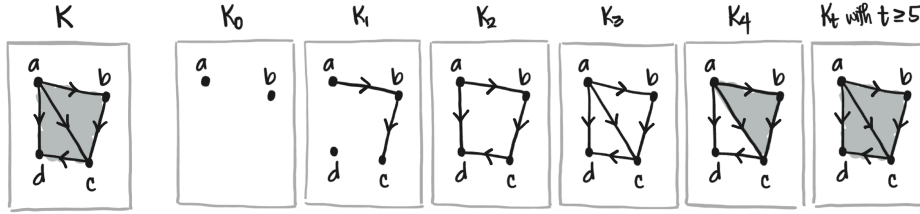
Example 3.2.3. Let the simplicial complex K and the filtration K_\bullet on K be as illustrated below:



We consider the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{Q})$ of K_\bullet with \mathbb{Q} coefficients in all dimensions $n \in \mathbb{N}_0$.

Part A. Pointwise Calculation of Homology Groups with Rational Coefficients

We consider the chain groups $C_n(K_t; \mathbb{Q})$ and boundary maps $\partial_n^t : C_n(K_t; \mathbb{Q}) \rightarrow C_{n-1}(K_t; \mathbb{Q})$ of K_t at each scale $t \in \mathbb{N}_0$ separately. Orient each simplicial complex K_t with the vertex order $\text{Vert}(K) = (a, b, c, d)$ restricted on $\text{Vert}(K_t)$. Illustrated below is the induced orientation on the 1-simplices of K and those of K_t for all $t \in \mathbb{N}_0$:



For all $n \in \mathbb{N}_0$, the n^{th} homology group $H_n(K_t; \mathbb{Q})$ of K_t with \mathbb{Q} coefficients is then calculated as usual with

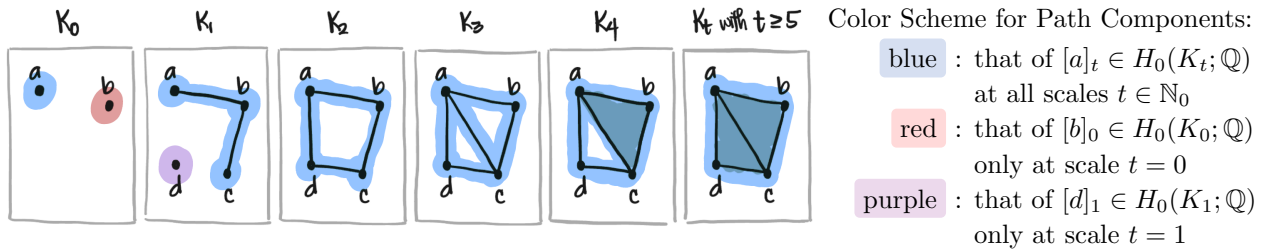
$$H_n(K_t; \mathbb{Q}) = \ker(\partial_n^t) / \text{im}(\partial_{n+1}^t)$$

Observe that $H_n(K_t; \mathbb{Q})$ is a \mathbb{Q} -vector space for all $n \in \mathbb{N}_0$. Since $C_n(K_t; \mathbb{Q}) = 0$ for all $n \geq 3$, $\ker(\partial_n^t) = 0$ and $H_n(K_t; \mathbb{Q}) = 0$ for all $n \geq 2$. Then, $H_0(K_t; \mathbb{Q})$ and $H_1(K_t; \mathbb{Q})$ for all $t \in \mathbb{N}_0$ are as follows:

$$H_0(K_t; \mathbb{Q}) \cong \begin{cases} \mathbb{Q}\langle [a]_0, [b]_0 \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle [a]_1, [d]_1 \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle [a]_t \rangle & \text{if } t \geq 2 \end{cases} \quad \text{and} \quad H_1(K_t; \mathbb{Q}) = \begin{cases} 0 & \text{if } t = 0, 1 \\ \mathbb{Q}\langle [ab + bc + cd - ad]_2 \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle [ab + bc - ac]_3, [ac + cd - ad]_3 \rangle & \text{if } t = 3 \\ \mathbb{Q}\langle [ac + cd - ad]_4 \rangle & \text{if } t = 4 \\ 0 & \text{if } t \geq 5 \end{cases}$$

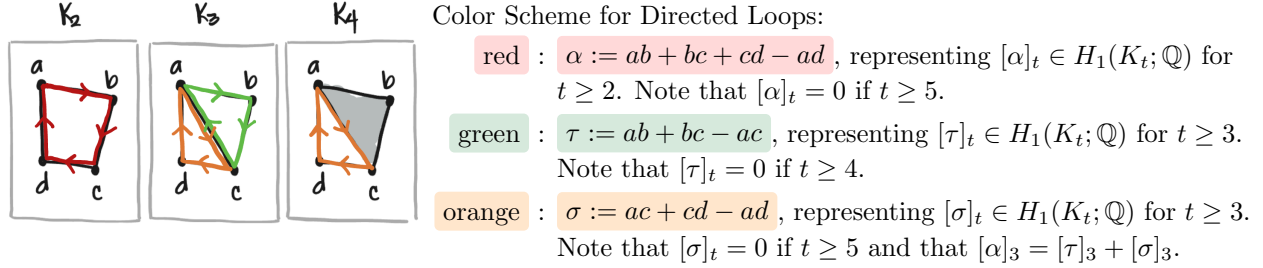
Part B. Interpretation of Homology Groups with Rational Coefficients

The interpretation of the homology classes of $H_0(K_t; \mathbb{Z})$ as corresponding to path components, see [Hat02, Proposition 2.6, p109], also applies to homology in \mathbb{Q} coefficients. Illustrated below are the path components of K_t relating to basis elements of $H_0(K_t; \mathbb{Q})$ for all $t \in \mathbb{N}_0$.



The cycle representatives of $H_1(K_t; \mathbb{Z})$ correspond to directed loops in K_t , consisting of \mathbb{Z} -multiples of oriented 1-simplices. The same interpretation also applies to $H_1(K_t; \mathbb{Q})$ since we can scale any cycle representative

such that it consists of \mathbb{Z} -multiples of oriented 1-simplices. Illustrated below are the loops in K_t with $t = 2, 3, 4$ corresponding to homology classes in $H_1(K_t; \mathbb{Q})$:



Negative multiples of oriented 1-simplices denote a reversal in direction. For example, $ad \in C_1(K_2; \mathbb{Z})$ is a line from vertex a to vertex d but $-ad \in C_1(K_2; \mathbb{Z})$ is a line from d to a . The addition of an oriented 1-simplex to an oppositely oriented copy of itself can be interpreted to be an annihilation of said simplex, e.g. $ad + (-ad) = 0$.

Part C. Formation of Persistent Homology Modules

Let $n \in \mathbb{N}_0$. The n^{th} persistent homology module $H_n(K_\bullet; \mathbb{Q})$ has vector spaces given by $H_n(K_\bullet; \mathbb{Q})(t) = H_n(K_t; \mathbb{Q})$ for all $t \in \mathbb{N}_0$. The structure maps of $H_n(K_\bullet; \mathbb{Q})$ are the maps on homology induced by inclusions $i^{s,t} : K_t \rightarrow K_s$ for all $t, s \in \mathbb{N}_0$ with $t \leq s$. For now, it suffices to know that these structure maps exist.

Observe that $H_n(K_\bullet; \mathbb{Q})$ is the trivial persistence module over \mathbb{Q} for $n \geq 2$ since $H_n(K_t; \mathbb{Q}) = 0$ for all $t \in \mathbb{N}_0$. We talk more about $H_0(K_\bullet; \mathbb{Q})$ and $H_1(K_\bullet; \mathbb{Q})$ in the following examples in this section.

Remark 3.2.4. The preferred coefficient field for persistent homology is $\mathbb{F} = \mathbb{Z}_p$ with prime p . The majority of the examples in this paper use $\mathbb{F} = \mathbb{Q}$ for the coefficient field for convenience.

The reason for \mathbb{Z}_p being the preferred coefficient field is that numbers in \mathbb{Z}_p can be represented exactly using a finite number of bits. For example, the Ripser package described in [Bau21] stores numbers in \mathbb{Z}_p as unsigned 16-bit integers and allows for calculation of persistent homology in \mathbb{Z}_p coefficients with $p \leq 2^{16} - 1$. Furthermore, a number in \mathbb{Z}_2 can be represented using one bit and arithmetic operations in \mathbb{Z}_2 can be done as bit operations, which are extremely fast computationally. This explains why Cohen-Steiner, et al. in [CEM06] and other authors (particularly those in the field of computer science) often define persistent homology with coefficients in $\mathbb{F} = \mathbb{Z}_2$, i.e. their persistence modules are of the form $\mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{Vect}_{\mathbb{Z}_2}$.

However, operations in \mathbb{Z}_p can be somewhat cumbersome to do by hand, which became an issue when we were creating examples for this expository paper. As a compromise, we use the field $\mathbb{F} = \mathbb{Q}$ for most of our examples since the relation between homology groups over \mathbb{Q} and those of \mathbb{Z}_p is relatively straightforward. For reference, see [Hat02, Corollary 3A.6].

The persistence barcode of a persistence module, as discussed in Chapter 2, characterizes the ranks of the vector spaces and the ranks of the structure maps of the persistence module. The same applies for persistent homology modules, assuming interval decompositions and persistence barcodes exist. We state an existence result below.

Proposition 3.2.5. Let K_\bullet be a filtration on a finite simplicial complex K and fix a field \mathbb{F} . For each $n \in \mathbb{N}_0$, the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F})$ is of finite-type and therefore admits an interval decomposition.

Proof. By Lemma 3.1.7, K_t is a finite simplicial complex for all $t \in \mathbb{N}_0$ and there exists $T \in \mathbb{N}_0$ such that K_\bullet is constant on $[T, \infty)$. For $t \in [0, T]$, $H_n(K_t; \mathbb{F})$ must be finitely generated since K_t is a

finite simplicial complex. For all $t \in [T, \infty)$, $H_n(K_T; \mathbb{F}) = H_n(K_t; \mathbb{F})$ since $K_T = K_t$. Therefore, $H_n(K_\bullet; \mathbb{F})$ is a finite-type persistence module that is constant on $[T, \infty)$. By Proposition 2.5.14, $H_n(K_\bullet; \mathbb{F})$ admits an interval decomposition. ■

The calculation of this interval decomposition is the goal of the matrix reduction algorithm for persistent homology. Below, we identify terminology used in [Bau21] and other persistent homology literature involving these interval decompositions.

Definition 3.2.6. Let K_\bullet be a filtration on a finite simplicial complex K . Let the persistence barcode $\text{Bar}_n(K_\bullet; \mathbb{F})$ of the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F})$ of K_\bullet with coefficients in a field \mathbb{F} be given as follows:

$$\text{Bar}_n(K_\bullet; \mathbb{F}) = \text{Bar}(H_n(K_\bullet; \mathbb{F})) = \left\{ [b_j, d_j] \right\}_{j=1}^r \cup \left\{ [a_i, \infty) \right\}_{i=1}^m$$

where $\{[b_j, d_j]\}_{j=1}^r$ and $\{[a_i, \infty)\}_{i=1}^m$ both consist of intervals in \mathbb{N}_0 with $b_j, d_j, a_i \in \mathbb{N}_0$.

- i. The pair (b_j, d_j) of indices from the interval $[b_j, d_j]$ is called an **index persistence pair** of K_\bullet . In this case, b_j is called a **birth index** and d_j is called its corresponding **death index**.
- ii. The index a_i in the interval $[a_i, \infty)$ is called an **essential birth index** or **essential index** of K_\bullet .

By Theorem 2.3.5, persistence barcodes of persistence modules are unique up to persistence isomorphism. Therefore, the collection of index persistence pairs and essential indices from $H_n(K_\bullet; \mathbb{F})$ can be considered an invariant of the isomorphism type of persistence modules.

As briefly discussed in Section 3.1, it seems to be common practice to assume the index $t \in \mathbb{N}_0$ represents some time value. Roughly speaking, the terms *birth index* and *death index* correlate to how homology classes are created (alternatively, born) or become trivial (alternatively, die or get destroyed) as we increase the time parameter $t \in \mathbb{N}_0$. The term *persistence* then corresponds to how long a homology class that is born at some index $t_B \in \mathbb{N}_0$ remains non-trivial (alternatively, lives) as we increase the index $t \geq t_B$.

Then, the phrase “as we increase the time parameter” refers to the application of the structure map $i_*^{s,t} : H_n(K_t; \mathbb{F}) \rightarrow H_n(K_s; \mathbb{F})$. We can also colloquially say that we are going up the filtration K_\bullet from time t to time $s \geq t$. To avoid ambiguity, we provide definitions for the terms such as *lives*, *dies*, and *persists* below.

Definition 3.2.7. Let K_\bullet be a simplicial filtration. Let $[\sigma]_t \in H_n(K_t; \mathbb{F})$ be a non-trivial homology class with cycle representative $\sigma \in \ker(\partial_n^t) \subseteq C_n(K_t; \mathbb{F})$. Let $s \in \mathbb{N}_0$ such that $t \leq s$.

- i. We say that $[\sigma]_t$ **lives** or **persists in** K_s if $i_*^{s,t}([\sigma]_t) \neq 0$.
- ii. If $[\tau]_t \in H_n(K_t; \mathbb{F})$ is a non-trivial homology class such that $[\tau]_t \neq [\sigma]_t$ and $i_*^{s,t}([\sigma]_t) = i_*^{s,t}([\tau]_t) \neq 0$, we say that $[\tau]_t$ and $[\sigma]_t$ **merge in** K_s .
- iii. We say that $[\sigma]_t$ is **born in** K_B or **at index** $B \in \mathbb{N}_0$ if

$$B = \max \left\{ b \in \mathbb{N}_0 : b \leq t \text{ and } [\sigma]_t \notin \text{im}(i_*^{t,b-1}) \right\}$$

where $\text{im}(i_*^{t,-1})$ is taken to be the trivial vector space.

- iv. We say that $[\sigma]_t$ **dies** or **is destroyed in** K_D or **at index** $D \in \mathbb{N}_0$ if

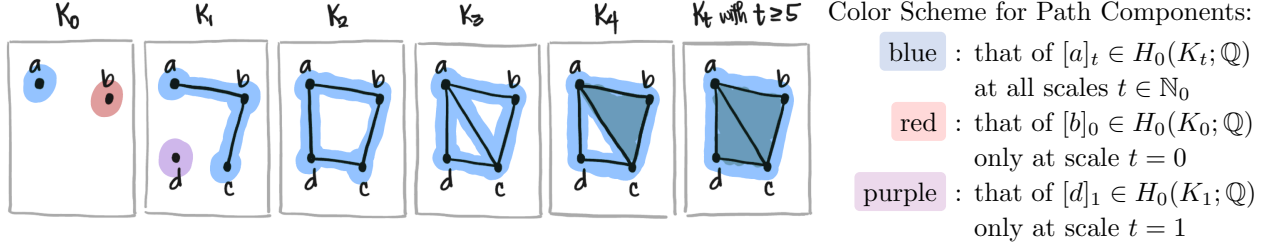
$$D = \min \left\{ d \in \mathbb{N}_0 : t \leq d \text{ and } i_*^{d,t}([\sigma]_t) = 0 \right\}$$

If such a minimum does not exist, then we say that $[\sigma]_t$ **does not die in the filtration**. Note that if $[\sigma]_t$ is destroyed in K_D , then $i^{s,t}([\sigma]_t) = (i^{s,D} \circ i^{D,t})([\sigma]_t) = i^{s,D}(0) = 0$ for any $s \geq D$.

- v. We say that $[\sigma]_t$ has **persistence** or **lifespan** $D - B \in \mathbb{N}_0$ if $[\sigma]_t$ is born at index $B \in \mathbb{N}_0$ and dies at index $D \in \mathbb{N}_0$. If $[\sigma]_t$ does not die in K_\bullet , then persistence is interpreted to be ∞ .

We provide an example of these terms in use below, relative to the 0th persistent homology module.

Example 3.2.8. Let K_\bullet be as given in Example 3.2.3. For convenience, we copied the illustration involving $H_0(K_t; \mathbb{Q})$ and the path components of K_t at each $t \in \mathbb{N}_0$ below:



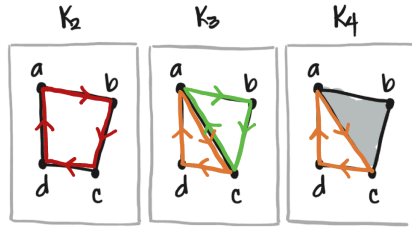
We consider some 0th homology classes in $H_0(K_t; \mathbb{Q})$ for varying $t \in \mathbb{N}_0$ below.

1. The homology classes $[a]_0, [b]_0 \in H_0(K_0; \mathbb{Q})$ persist to K_2 since $i^{2,0}([a]_0) = [a]_2 \neq 0$ and $i^{2,0}([b]_0) = [b]_2 \neq 0$. Moreover, $[a]_0$ and $[b]_0$ live or persist to K_t for all $t \geq 1$. At index $t = 0$, $[a]_0$ is highlighted in blue and $[b]_0$ in red in the illustration above.
2. The homology classes $[a]_0, [b]_0 \in H_0(K_0; \mathbb{Q})$ merge in K_1 since $i^{1,0}([a]_0) = [a]_1 = [b]_1 = i^{1,0}([b]_0)$. That is, the 0-cycles a and b represent distinct homology classes in K_0 but represent the same homology class in K_1 . Furthermore, for all $t \geq 1$, $[a]_t$ and $[b]_t$ represent the same path component, highlighted in blue in the illustration above.
3. The homology class $[a]_3 \in H_0(K_3; \mathbb{Q})$ is born in K_0 since $\text{im}(i_*^{3,0}) = \mathbb{Q}\langle[a]_3\rangle$. Similarly, the homology class $[d]_4 \in H_0(K_3; \mathbb{Q})$ is born in K_0 since $[d]_4 = [a]_4 = i^{4,0}([a]_0)$. These homology classes are highlighted in blue in the illustration above.
4. The homology class $[d]_1 \in H_0(K_1; \mathbb{Q})$, specifically at index $t = 1$ and highlighted in purple above, is born at index $t = 1$ since $\text{im}(i_*^{1,0}) = \mathbb{Q}\langle[a]_1\rangle$ and $[d]_1 \neq [a]_1$. Note that for all $t \geq 2$, $[d]_t$ corresponds to the path component highlighted in blue.

Observe that none of the homology classes listed above die in the filtration.

Let K_\bullet be a filtration on a simplicial complex K . Observe that, in dimension $n = 0$, a 0th homology class in $H_0(K_t; \mathbb{F})$ for any $t \in \mathbb{N}_0$ cannot die in a filtration since path components do not disappear as we increase the time parameter. If K is path connected, then all of these 0th homology classes merge into one homology class at sufficiently high $T \in \mathbb{N}_0$. Persistence becomes more interesting in dimensions $n \geq 1$ since it becomes possible for homology classes to become trivial as we increase the time parameter. We provide an example below on dimension $n = 1$.

Example 3.2.9. Let K_\bullet be as given in Example 3.2.3. For convenience, we copied the illustration involving $H_0(K_t; \mathbb{Q})$ and directed loops in K_t below. Note that if $t \notin \{2, 3, 4\}$, then $H_1(K_t; \mathbb{F}) = 0$.

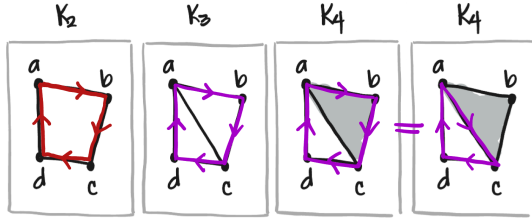


Color Scheme for Directed Loops:

- red : $\alpha := ab + bc + cd - ad$, representing $[\alpha]_t \in H_1(K_t; \mathbb{Q})$ for $t \geq 2$. Note that $[\alpha]_t = 0$ if $t \geq 5$.
- green : $\tau := ab + bc - ac$, representing $[\tau]_t \in H_1(K_t; \mathbb{Q})$ for $t \geq 3$. Note that $[\tau]_t = 0$ if $t \geq 4$.
- orange : $\sigma := ac + cd - ad$, representing $[\sigma]_t \in H_1(K_t; \mathbb{Q})$ for $t \geq 3$. Note that $[\sigma]_t = 0$ if $t \geq 5$ and that $[\alpha]_3 = [\tau]_3 + [\sigma]_3$.

We consider some 1st homology classes in $H_1(K_t; \mathbb{Q})$ with $t \in \{2, 3, 4\}$ below.

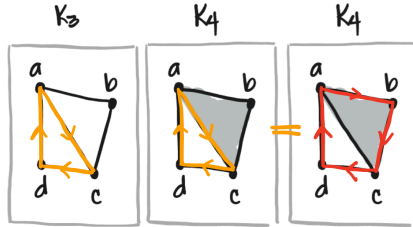
- $[\alpha]_2 = [ab + bc + cd - ad]_2 \in H_1(K_2; \mathbb{Q})$ at index $t = 2$ is born at index $t = 2$, dies at index $t = 5$, and has a lifespan of 3. Illustrated below is the image of $[\alpha]_2$ on $H_n(K_t; \mathbb{Q})$ for $t = 3, 4$.



Color Scheme:

- red : α in K_2 , representing $[\alpha]_2$.
- purple : α in K_3 and K_4 , representing $[\alpha]_3$ and $[\alpha]_4$ respectively.

- $[\alpha]_3 = [ab + bc + cd - ad]_3 \in H_1(K_3; \mathbb{Q})$ at index $t = 3$ is born at index $t = 2$ since $i_*^{3,2}([\alpha]_2) = [\alpha]_3$.
- $[\alpha]_4 = [ab + bc + cd - ad]_4 \in H_1(K_4; \mathbb{Q})$ is born at index $t = 3$ since $i_*^{4,3}([\alpha]_3) = [\sigma]_4 = [\alpha]_4$ with $\sigma = ac + cd - ad$, as illustrated below:



Color Scheme:

- red : α in K_4 , representing $[\alpha]_4$.
- orange : (middle) $\sigma \in K_4$, representing $[\sigma]_4 = [\alpha]_4$
(left) σ in K_3 , representing $[\sigma]_3$.

- $[\tau]_3 = [ab + bc - ac]_3 \in H_1(K_3; \mathbb{Q})$ at index $t = 3$ is born at the same index $t = 3$. Since the 2-simplex abc is added to K_4 , $\partial_2^4(abc) = bc - ac + ab = \tau$ and τ becomes a 1-boundary in K_4 . Therefore, $i_*^{4,3}([\tau]_3) = [\tau]_4 = 0$ and $[\tau]_3$ dies in K_4 .
- $[\sigma]_3 \in H_1(K_3; \mathbb{Q})$ at index $t = 3$ is born at index $t = 3$, dies at index $t = 5$, and has persistence $5 - 3 = 2$.

Let K_\bullet be a filtration on a finite simplicial complex K and assume that K_\bullet is constant on $[T, \infty)$. Note that $T \in \mathbb{N}_0$ exists by Lemma 3.1.7 and that $K = K_T$. The persistence barcode $\text{Bar}_n(K_\bullet; \mathbb{F})$ of a filtration K_\bullet with \mathbb{F} coefficients determines the interval decomposition $H_n(K_\bullet; \mathbb{F})$ as a persistence module. Let $\text{Bar}_n(K_\bullet; \mathbb{F})$ be given as follows

$$\text{Bar}_n(K_\bullet; \mathbb{F}) = \text{Bar}(H_n(K_\bullet; \mathbb{F})) = \left\{ [s_1, s_1 + t_1), \dots, [s_r, s_r + t_r), [s_{r+1}, \infty), \dots, [s_m, \infty) \right\}$$

for some $s_1, \dots, s_m \in \mathbb{N}_0$ and $t_1, \dots, t_r \in \mathbb{N}_0$. The intervals $[s_j, s_j + t_j)$ and $[s_j, \infty)$ in $\text{Bar}_n(K_\bullet; \mathbb{F})$ are denoted following Corollary 2.5.13, where we state the correspondence between graded invariant factor decompositions in $\mathbf{GrMod}_{\mathbb{F}[x]}$ and interval decompositions in $\mathbf{Pers}_{\mathbb{F}}$. Therefore, there exists a persistence isomorphism as follows:

$$\varphi_\bullet : H_n(K_\bullet; \mathbb{F}) \rightarrow \left(\mathbb{I}_{\bullet}^{[s_1, s_1 + t_1)} \oplus \dots \oplus \mathbb{I}_{\bullet}^{[s_r, s_r + t_r)} \oplus \mathbb{I}_{\bullet}^{[s_{r+1}, \infty)} \oplus \dots \oplus \mathbb{I}_{\bullet}^{[s_m, \infty)} \right)$$

This persistence isomorphism identifies a collection of homology classes on varying indices $t \in \mathbb{N}_0$ by:

$$[\sigma_k]_{s_k} \in (\varphi_{s_k})^{-1}(\mathbb{I}_{\bullet}^{J_k}(s_k)) \subseteq H_n(K_{s_k}; \mathbb{F}) \quad \text{with} \quad J_k := \begin{cases} [s_k, s_k + t_k) & \text{if } k \in \{1, \dots, r\} \\ [s_k, \infty) & \text{if } k \in \{r+1, \dots, m\} \end{cases}$$

Note that $(\varphi_{s_k})^{-1}(\mathbb{I}_{\bullet}^{J_k}(s_k))$ must be a one-dimensional \mathbb{F} -vector space and there exists a non-trivial homology class $[\sigma_k]_{s_k}$ for each $k \in \{1, \dots, m\}$. The birth, death, and persistence of $[\sigma_k]_{s_k}$ as a homology class in $H_0(K_{\bullet}; \mathbb{F})$ correspond to the indices in the intervals $[s_k, s_k + t_k)$ or $[s_k, \infty)$. In particular:

1. For $k \in \{1, \dots, r\}$, the homology class $[\sigma_k]_{s_k}$ from $\mathbb{I}_{\bullet}^{J_k} = \mathbb{I}_{\bullet}^{[s_k, s_k + t_k)}$ is born at index s_k , dies at index $s_k + t_k$, and has persistence t_k . Furthermore, the behavior of the structure maps match the interval $[s_k, s_k + t_k)$, i.e.

$$i_*^{t, s_k}([\sigma_k]_{s_k}) = \begin{cases} [\sigma_k]_t \neq 0 & \text{if } t \in [s_k, s_k + t_k) \\ 0 & \text{if } t \geq s_k + t_k \end{cases}$$

Observe that the bounded interval $[s_k, s_k + t_k)$ corresponds to homology classes that become trivial in $K = K_T$, i.e. $[\sigma_k]_T = i_*^{T, s_k}([\sigma_k]_{s_k}) = 0$ as an element of $H_n(K_T; \mathbb{F}) = H_n(K_T; \mathbb{F})$.

2. For $k \in \{r+1, \dots, m\}$, the homology class $[\sigma_k]_{s_k}$ from $\mathbb{I}_{\bullet}^{J_k} = \mathbb{I}_{\bullet}^{[s_k, \infty)}$ is born at index s_k and does not die in the filtration, i.e. for all $t \geq s_k$, $i_*^{t, s_k}([\sigma_k]_{s_k}) = [\sigma_k]_t \neq 0$.

Note that the unbounded intervals $[s_k, \infty)$ correspond to homology classes that do not die in $K = K_T$, i.e. $[\sigma_k]_T = i_*^{T, s_k}([\sigma_k]_{s_k}) \neq 0$ in $H_n(K; \mathbb{F}) = H_n(K_T; \mathbb{F})$ for all $k \in \{r+1, \dots, m\}$. This might explain why s_k in the interval $[s_k, \infty)$ is called an *essential* birth index of K_{\bullet} .

The problem here is that such a collection $\{[\sigma_1]_{s_1}, \dots, [\sigma_m]_{s_m}\}$ of homology classes is not generally immediately obvious by examination of the homology groups $H_n(K_t; \mathbb{F})$ at every $t \in \mathbb{N}_0$ since the structure maps, i.e. the maps on homology induced by inclusion, cannot be ignored.

The matrix reduction algorithm for persistent homology, discussed in Chapter 4, not only determines the persistence barcode of $H_n(K_{\bullet}; \mathbb{F})$ but also determines possible cycle representatives corresponding to each interval in said barcode. In the two examples below, we interpret results that will be calculated in Chapter 4.

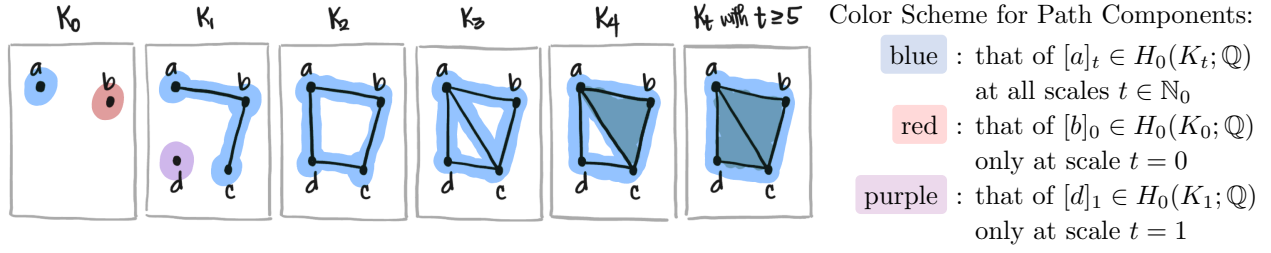
Example 3.2.10. Let K_{\bullet} be given as in Example 3.2.3. Given below is a summary of the results of the calculation involving $H_0(K_{\bullet}; \mathbb{Q})$ started in Example 4.5.3 and finished in Example 4.5.8:

Interval Module	Interval	Birth Index	Cycle Representative	Persistence
$\mathbb{I}_{\bullet}^{[0,1)}$	$J_1 = [0, 1)$	$s_1 = 0$	$\sigma_1 = b - a$	$t_1 = 1 - 0 = 1$
$\mathbb{I}_{\bullet}^{[1,2)}$	$J_2 = [1, 2)$	$s_2 = 1$	$\sigma_2 = d - a$	$t_2 = 2 - 1 = 1$
$\mathbb{I}_{\bullet}^{[0,\infty)}$	$J_3 = [0, \infty)$	$s_3 = 0$	$\sigma_3 = a$	$t_3 = \infty$

The interval decomposition of $H_0(K_{\bullet}; \mathbb{Q})$ is then calculated to be

$$H_0(K_{\bullet}; \mathbb{Q}) \stackrel{\text{Pers}}{\cong} \mathbb{I}_{\bullet}^{[0,\infty)} \oplus \mathbb{I}_{\bullet}^{[0,1)} \oplus \mathbb{I}_{\bullet}^{[1,2)}$$

Observe that each interval module $\mathbb{I}_{\bullet}^{J_k}$ corresponds to a homology class that is born at index s_k and dies at index $s_k + t_k$. However, these homology classes do not correspond directly to the path components of each K_t . For comparison, we copied the illustration in Example 3.2.3 involving $H_0(K_{\bullet}; \mathbb{Q})$:

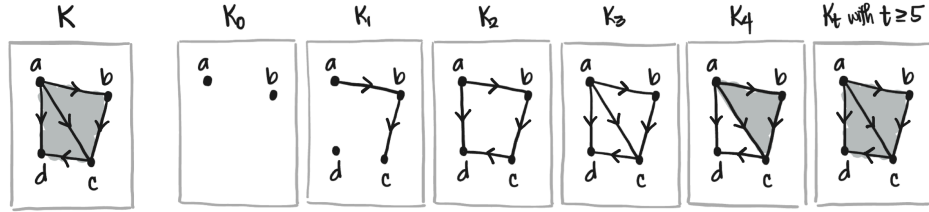


Observe that $\mathbb{I}_{\bullet}^{[0,\infty)}$ corresponds to the homology classes $[a]_t$ with $t \in \mathbb{N}_0$ and has a direct relation to the path component of K_t containing the 0-cycle a . However, the same does not apply for $\mathbb{I}_{\bullet}^{[0,1)}$ and $\mathbb{I}_{\bullet}^{[1,2)}$, which correspond to the homology classes $[b - a]_0$ and $[d - a]_1$ respectively.

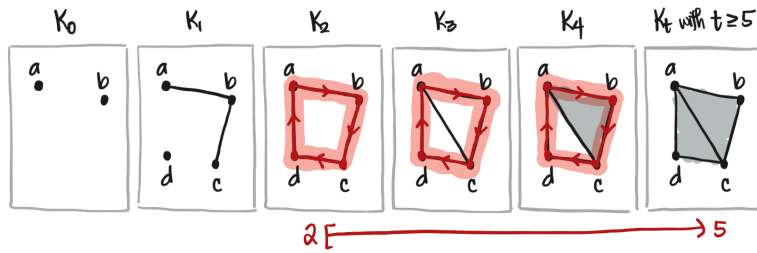
Example 3.2.11. Let K_{\bullet} be given as in Example 3.2.3. A summary of the results calculated in Example 4.6.1 involving the interval decomposition of $H_1(K_{\bullet}; \mathbb{Q})$ is given below:

Interval Module	Interval	Birth Index	Cycle Representative	Persistence
$\mathbb{I}_{\bullet}^{[2,5)}$	$J_1 = [2, 5)$	$s_1 = 2$	$\beta_1 = ab + bc - ad + cd$	$t_1 = 5 - 2 = 3$
$\mathbb{I}_{\bullet}^{[3,4)}$	$J_2 = [3, 4)$	$s_2 = 3$	$\beta_2 = -ab - bc + ac$	$t_2 = 4 - 3 = 1$

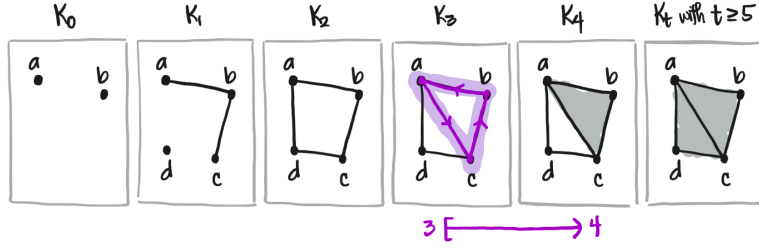
The interval decomposition of $H_1(K_{\bullet}; \mathbb{Q})$ is given by $H_1(K_{\bullet}; \mathbb{Q}) = \mathbb{I}_{\bullet}^{[2,5)} \oplus \mathbb{I}_{\bullet}^{[3,4)}$. For convenience, we copied the illustration in Example 3.2.3 describing the orientation on the 1-simplices of K_t at each $t \in \mathbb{N}_0$ induced by the vertex order $\text{Vert}(K) = (a, b, c, d)$ below.



The first interval module $\mathbb{I}_{\bullet}^{[2,5)}$ corresponds to the homology classes $[\beta_1]_t$ with $t \in [2, 5)$. Observe that the 1-cycle $\beta_1 = ab + bc - ad + cd$, illustrated in red below for $t \in [2, 5)$, first appears at index $s_1 = 2$. The 2-cycle $abc + acd$, that makes β_1 a 1-boundary and the homology class $[\beta_1]_t$ trivial, first appears at index $s_1 + t_1 = 2 + 3 = 5$.



The second interval module $\mathbb{I}_{\bullet}^{[3,4)}$ corresponds to the homology classes $[\beta_2]_t$ with $t \in [3, 4)$. Observe that the 1-cycle $\beta_2 = -ab - bc + ac$, illustrated below in purple for $t \in [3, 4)$, first appears at $s_2 = 3$. The 2-cycle $-abc$, which makes β_2 a 1-boundary and the homology class $[\beta_2]_t$ trivial, first appears at index $s_2 + t_2 = 4$.



Note that the 1-cycle $ac + cd - ad$ is not represented in either interval module for $H_1(K_\bullet; \mathbb{Q})$, despite being a valid cycle representative for homology classes in $H_1(K_t; \mathbb{Q})$ for $t \geq 3$.

For the sake of completion, we include an alternate characterization for persistent homology below, one that is prevalent in persistent homology literature, e.g. in [ZC05; CEM06]. While this perspective is not as useful as the functor definition of persistent homology (Definition 3.2.1) for rigorously discussing the origins of the matrix reduction algorithm (as presented in Chapter 4), it does offer some more insight to the significance of persistence barcodes. The following definitions are taken from [ZC05, Section 2.6].

Definition 3.2.12. Let K_\bullet be a simplicial filtration and let \mathbb{F} be a field. The p -persistent n^{th} homology group $H_n(K_t; p; \mathbb{F})$ with coefficients in \mathbb{F} of the simplicial complex K_t in K_\bullet is the \mathbb{F} -vector space given by

$$H_n(K_t; p; \mathbb{F}) = i_*^{t+p, t} \left(H_n(K_t; \mathbb{F}) \right) \subseteq H_n(K_{t+p}; \mathbb{F})$$

and the rank of $H_n(K_t; p; \mathbb{F})$ is called the p -persistent n^{th} Betti number $\beta_n(K_t; p; \mathbb{F})$ of K_t with coefficients in \mathbb{F} .

Observe that the characterization of persistent homology as a persistence module $H_n(K_\bullet; \mathbb{F})$ by Definition 3.2.1 accounts for the p -persistent homology groups $H_n(K_t; p; \mathbb{F})$ for all $p, t \in \mathbb{N}_0$ and is a more concise way to describe how all of these homology groups are related. Furthermore, the persistence barcode $\text{Bar}_n(K_\bullet; \mathbb{F})$ of a filtration K_\bullet is a concise representation of the ranks of all p -persistent homology groups. In other words, the persistence barcode encodes the evolution of the Betti numbers of K_t across the filtration K_\bullet . We state this in more detail below.

Corollary 3.2.13. Let K_\bullet be a filtration on a finite simplicial complex K . For all $n \in \mathbb{N}_0$, the n^{th} Betti numbers $\beta_n(K_t; \mathbb{F})$ of the simplicial complex K_t and the n^{th} p -persistent Betti numbers $\beta_n(K_t; p; \mathbb{F})$ of K_t in K_\bullet are determined by the persistence barcode $\text{Bar}_n(K_\bullet; \mathbb{F})$ as follows:

$$\begin{aligned} \beta_n(K_t; \mathbb{F}) &= \text{card} \left\{ J \in \text{Bar}_n(K_\bullet; \mathbb{F}) : t \in J \right\} \\ \beta_n(K_t; p; \mathbb{F}) &= \text{card} \left\{ J \in \text{Bar}_n(K_\bullet; \mathbb{F}) : [t, t+p] = [t, t+p+1) \subseteq J \right\} \end{aligned}$$

Proof. By Proposition 3.2.5, an interval decomposition for $H_n(K_\bullet; \mathbb{F})$ exists. Applying Proposition 2.3.7 on the \mathbb{F} -vector space $H_n(K_t; \mathbb{F})$ for all $t \in \mathbb{N}_0$, we have that

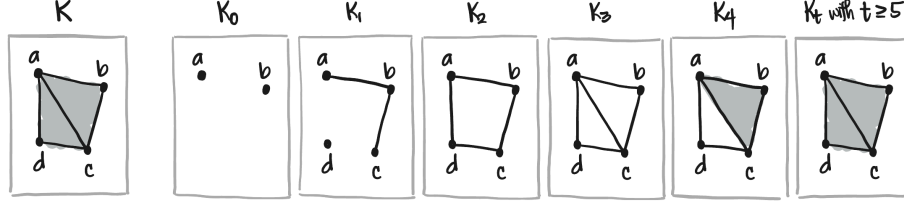
$$\text{rank}(H_n(K_t; \mathbb{F})) = \beta_n(K_t; \mathbb{F}) = \text{card} \left\{ J \in \text{Bar}_n(K_\bullet; \mathbb{F}) : t \in J \right\}.$$

Applying Proposition 2.3.7 on structure maps $i_*^{t+p, t}$ of $H_n(K_t; \mathbb{F})$ for all $t, p \in \mathbb{N}_0$, we have that

$$\text{rank}(i_*^{t+p, t}) = \text{rank} \left(i_*^{t+p, t} (H_n(K_t; \mathbb{F})) \right) = \text{rank} (H_n(K_t; p; \mathbb{F})) = \beta_n(K_t; p; \mathbb{F})$$

$$= \text{card} \left\{ J \in \text{Bar}_n(K_\bullet; \mathbb{F}) : [t, t+p] = [t, t+p+1] \subseteq J \right\}. \quad \blacksquare$$

Example 3.2.14. Let K_\bullet be given as in Example 3.2.3. For convenience, the illustration of the simplicial complex K and the filtration K_\bullet is copied below:



Following the discussion on Example 3.2.10, the interval decomposition $H_0(K_\bullet; \mathbb{Q}) \cong \mathbb{I}_\bullet^{[0,1)} \oplus \mathbb{I}_\bullet^{[1,2)} \oplus \mathbb{I}_\bullet^{[0,\infty)}$ implies that

$$\text{Bar}_0(K_\bullet; \mathbb{Q}) = \left\{ [0, \infty), [0, 1), [1, 2) \right\}$$

Below, we compare the p -persistent 0^{th} homology group $H_0(K_t; p; \mathbb{Q})$ and the p -persistent 0^{th} Betti number $\beta_0(K_t; p; \mathbb{Q})$ for selected $p, t \in \mathbb{N}_0$ and compare these against the results expected by Corollary 3.2.13.

1. Let $t = 1$. Since $K_t = K_1$ has two path components, we expect $\beta_0(K_1; \mathbb{Q}) = 2$. There are 2 intervals in $\text{Bar}_0(K_\bullet; \mathbb{Q})$ that contain $t = 1$: $[0, \infty), [1, 2) \in \text{Bar}_0(K_\bullet; \mathbb{Q})$.
2. Let $t = 0$ and $p = 1$. Since $K_t = K_0$ has three path components and $K_{t+p} = K_1$ has two path components, we expect that $H_0(K_0; 1; \mathbb{Q}) \cong \mathbb{Q}^2$ and $\beta_0(K_0; 1; \mathbb{Q}) = 2$. Observe that there are 2 intervals in $\text{Bar}_0(K_\bullet; \mathbb{Q})$ containing the interval $[t, t+p] = [0, 1]$: $[0, \infty), [1, 2) \in \text{Bar}_0(K_\bullet; \mathbb{Q})$.
3. Let $t = 1$ and $p = 3$. Since $K_t = K_1$ has two path components and $K_{t+p} = K_4$ has one path component, we expect that $H_0(K_1; 3; \mathbb{Q}) \cong \mathbb{Q}$ and $\beta_0(K_1; 3; \mathbb{Q}) = 3$. Observe that there is only 1 interval in $\text{Bar}_0(K_\bullet; \mathbb{Q})$ containing $[t, t+p] = [1, 4]$: $[0, \infty) \in \text{Bar}_0(K_\bullet; \mathbb{Q})$.

Section 3.3. Simplicial Persistent Homology

In the previous section, we interpreted the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F})$ of a simplicial filtration K_\bullet as done pointwise, with each index $t \in \mathbb{N}_0$ considered separately. In this section, we consider an approach wherein we consider all indices $t \in \mathbb{N}_0$ of K_\bullet simultaneously. In particular, we extend the notions of simplicial chain groups, boundary maps, and chain complexes to the case of *persistence modules*. For brevity, we will refer to this extension as *simplicial persistent homology*.

We want to emphasize that the persistence modules and persistence complexes we introduce in this section are not exactly new. Rather, they come as a natural consequence of the functorial nature of simplicial homology. Earlier in Section 1.3, we discussed how the calculation of the simplicial homology can be seen as a composition of functors, as illustrated below for the simplicial complex K_t of K_\bullet :

$$\begin{array}{ccccc}
 \mathbf{A-Simp} & \xrightarrow{C_*(-; R)} & \mathbf{Ch-Vect}_{\mathbb{F}} & \xrightarrow{H_n(-)} & \mathbf{Vect}_{\mathbb{F}} \\
 K_t & \longmapsto & C_*(K_t; \mathbb{F}) = (C_n(K_t; R), \partial_n)_{n \in \mathbb{Z}} & \longmapsto & H_n(K_t; \mathbb{F}) \\
 (\text{simplicial complex}) & & (\text{simplicial chain complex}) & & (\text{simplicial homology})
 \end{array}$$

In simplicial persistent homology, we leave the parameter $t \in \mathbb{N}_0$ in K_\bullet arbitrary. We present a loose visualization of this below, with the parts highlighted in red to be discussed in this section.

$$\begin{array}{ccccc}
 [\mathbf{Poset}(\mathbb{N}_0, \leq), \mathbf{A-Simp}] & \longrightarrow & [\mathbf{Poset}(\mathbb{N}_0, \leq), \mathbf{Ch-Vect}_{\mathbb{F}}] & \longrightarrow & [\mathbf{Poset}(\mathbb{N}_0, \leq), \mathbf{Vect}_{\mathbb{F}}] = \mathbf{Pers}_{\mathbb{F}} \\
 & & = \mathbf{Ch-Pers}_{\mathbb{F}} & & \\
 K_\bullet & \longmapsto & C_*(K_\bullet; \mathbb{F}) & \longmapsto & H_n(K_\bullet; \mathbb{F}) \\
 (\text{filtration}) & & (\text{simplicial persistence complex?}) & & (\text{persistent homology module})
 \end{array}$$

As we construct new objects in $\mathbf{Pers}_{\mathbb{F}}$, we also investigate the corresponding graded structures in $\mathbf{GrMod}_{\mathbb{F}[x]}$ resulting from the application of $\Gamma_{\mathbf{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ to these objects and discuss which properties are preserved under this category equivalence.

We start with a definition of *filtered chain modules* of simplicial filtrations, which are simplicial chain groups extended to the case of persistence modules. Note that this definition is adapted from [Bau21].

Definition 3.3.1. Let K_\bullet be a filtration on a simplicial complex K and fix a field \mathbb{F} . For each $n \in \mathbb{Z}$, define the n^{th} **filtered chain module** $C_n(K_\bullet; \mathbb{F})$ of K_\bullet with coefficients in \mathbb{F} to be the persistence module over \mathbb{F} given by the following functor composition:

$$C_n(K_\bullet; \mathbb{F}) := C_n(-; \mathbb{F}) \circ K_\bullet$$

where $C_n(-; \mathbb{F}) : \mathbf{A-Simp} \rightarrow \mathbf{Vect}_{\mathbb{F}}$ refers to the n^{th} chain group functor with coefficients in \mathbb{F} .

Remarks. (1) $C_n(K_\bullet; \mathbb{F})$ is well-defined since the codomain category of $K_\bullet : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$ and the domain category of $C_n(-; \mathbb{F}) : \mathbf{A-Simp} \rightarrow \mathbf{Vect}_{\mathbb{F}}$ match. Note that K_\bullet is a functor by decomposition and $C_n(-; \mathbb{F})$ is a well-defined functor as discussed above Definition 1.3.5.

(2) The modifier “*filtered*” implies that the chain groups are those from a simplicial *filtration*. We decided to use the term “*module*”, instead of calling $C_n(K_\bullet; \mathbb{F})$ a filtered chain *group*, to emphasize that $C_n(K_\bullet; \mathbb{F})$ is to be interpreted mainly as a persistence *module*.

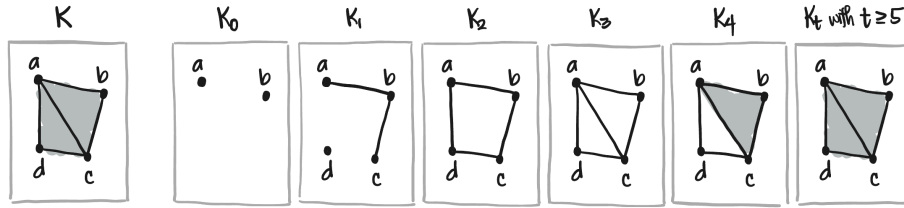
Observe that $C_n(K_\bullet; \mathbb{F})(t) = C_n(K_t; \mathbb{F})$ for all $t \in \mathbb{N}_0$ and the notation involving the vector spaces of $C_n(K_\bullet; \mathbb{F})$ should work as expected. For example, passing the parameter $t \in \mathbb{N}_0$ to $C_n(K_\bullet; \mathbb{F})$ returns the simplicial chain group of K_t of K_\bullet . This is also consistent with the convention of writing $V_t := V_\bullet(t)$ for an arbitrary persistence module V_\bullet , as introduced in Definition 2.1.1.

As given in Definition 3.1.1, the inclusion maps of K_\bullet are denoted as $i^{s,t} : K_t \rightarrow K_s$, with the shorthand of $i^t = i^{t+1,t} : K_t \rightarrow K_{t+1}$. Since the structure maps of $C_n(K_\bullet; \mathbb{F})$ must be the maps on the simplicial chain groups induced by inclusions, we can extend the hash (#) notation identified in Definition 1.3.3 and denote the structure maps of $C_n(K_\bullet; \mathbb{F})$ as follows:

$$\begin{aligned} i_{\#}^{s,t} &= C_n(K_\bullet; \mathbb{F})(t \rightarrow s) & \text{with} & & i_{\#}^{s,t} : C_n(K_t; \mathbb{F}) \rightarrow C_n(K_s; \mathbb{F}) \\ \text{and} \quad i_{\#}^t &= C_n(K_\bullet; \mathbb{F})(t \rightarrow t+1) & \text{with} & & i_{\#}^t : C_n(K_t; \mathbb{F}) \rightarrow C_n(K_{t+1}; \mathbb{F}) \end{aligned}$$

We provide an example of a filtered chain module below.

Example 3.3.2. Let K_\bullet and K be given as in Example 3.2.3 and orient K with the vertex order $\text{Vert}(K) = (a, b, c, d)$. For convenience, an illustration of K and K_\bullet (without orientation) is copied below:



The 0th and 1st filtered chain modules $C_0(K_\bullet; \mathbb{Q})$ and $C_1(K_\bullet; \mathbb{Q})$ of K_\bullet with rational coefficients have the following vector spaces, described relative to their respective standard bases:

$$C_0(K_t; \mathbb{Q}) = \begin{cases} \mathbb{Q}\langle a, b \rangle & \text{if } t = 0 \\ \mathbb{Q}\langle a, b, c, d \rangle & \text{if } t \geq 1 \end{cases} \quad \text{and} \quad C_1(K_t; \mathbb{Q}) = \begin{cases} 0 & \text{if } t = 0 \\ \mathbb{Q}\langle ab, bc \rangle & \text{if } t = 1 \\ \mathbb{Q}\langle ab, bc, cd, ad \rangle & \text{if } t = 2 \\ \mathbb{Q}\langle ab, bc, cd, ad, ac \rangle & \text{if } t \geq 3 \end{cases}$$

Below, we list some examples of 0-chains and 1-chains at selected values of $t \in \mathbb{N}_0$:

1. The 0-chain $2a + b$ is an element of $C_0(K_t; \mathbb{Q})$ for all $t \in \mathbb{N}_0$. In contrast, the chain $a + 2b + c$ is not an element of $C_0(K_0; \mathbb{Q})$ since the simplex c is not in K_0 .
2. The 1-chain $ab + bc - ac$ first appears in $C_1(K_3; \mathbb{Q})$, in that $t = 3$ is the minimal index for which $ab + bc - ac \in C_1(K_t; \mathbb{Q})$. Observe that $ab + bc - ac \in C_1(K_s; \mathbb{Q})$ for any $s \geq 3$.
3. The 1-chain $2ab - bc$ as an element of $C_1(K_1; \mathbb{Q})$ is mapped to $i_{\#}^{s,1}(2ab - bc) = 2ab - bc \in C_1(K_s; \mathbb{Q})$ for any $s \geq 1$.

Given a filtration K_\bullet of a simplicial complex K , the vector spaces $C_n(K_t; \mathbb{F})$ of the filtered chain module $C_n(K_\bullet; \mathbb{F})$ can all be considered subspaces of the simplicial chain group $C_n(K; \mathbb{F})$ of K . Additionally, the structure maps of $C_n(K_\bullet; \mathbb{F})$ are all inclusions and agree with the identity map on $C_n(K; \mathbb{F})$. We state this in more detail below.

Corollary 3.3.3. Let K_\bullet be a filtration of some simplicial complex K . For all $t, s \in \mathbb{N}_0$ with $t \leq s$, $C_n(K_t; \mathbb{F})$ is a vector subspace of $C_n(K; \mathbb{F})$ and the structure map $i_{\#}^{s,t} : C_n(K_t; \mathbb{F}) \rightarrow C_n(K_s; \mathbb{F})$ satisfies $i_{\#}^{s,t}(\sigma) = \text{id}_{C_n(K; \mathbb{F})}(\sigma) = \sigma$ for all $\sigma \in C_n(K_t; \mathbb{F})$.

Proof. Let $t, s \in \mathbb{N}_0$ with $t \leq s$. By definition of K_\bullet , $K_t \subseteq K_s \subseteq K$. By Lemma 1.3.10(i), $C_n(K_t; \mathbb{F}) \subseteq C_n(K_s; \mathbb{F}) \subseteq C_n(K; \mathbb{F})$ and for all n -chains $\sigma \in C_n(K_t; \mathbb{F})$, $i_{\#}^{s,t}(\sigma) = \text{id}_{C_n(K_s; \mathbb{F})}(\sigma) = i_{\#}^{[s]}(\sigma) = \text{id}_{C_n(K; \mathbb{F})}(\sigma) = \sigma$ where $i^{[s]}$ refers to the inclusion map $i^{[s]} : K_s \rightarrow K$. ■

The corollary above does not take any imposed orientation on either K or on any K_t of K_\bullet into account. However, for convenient calculation, we usually set an orientation on K and let all simplicial complexes K_t of K_\bullet inherit said orientation by restriction. In this case, the standard basis for any $C_n(K_t; \mathbb{F})$ at index $t \in \mathbb{N}_0$ (relative to the inherited orientation) must be a subset of that of $C_n(K; \mathbb{F})$. This property, along with the corollary above, makes the graded module obtained by applying $\mathbf{\Gamma}_{\text{GrMod}}$ to the filtered chain modules have a relatively uncomplicated structure. Below, we identify notation for said graded module.

Definition 3.3.4. For each $n \in \mathbb{Z}$, the n^{th} **graded chain module** of a simplicial filtration K_\bullet with coefficients in a field \mathbb{F} is the graded $\mathbb{F}[x]$ -module given by $C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) := \mathbf{\Gamma}_{\text{GrMod}}(C_n^{\text{Gr}}(K_\bullet; \mathbb{F}))$. We call an element of $C_n(K_\bullet; \mathbb{F})$ a **filtered n -chain**.

Remark. The term “graded chain module” is not used in most of the literature for persistent homology theory. We introduced this term in this paper to emphasize the difference between $C_n(K_\bullet; \mathbb{F})$ as a persistence module and $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ as a graded module. The symbol $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is added for brevity.

Following the description of $\mathbf{\Gamma}_{\text{GrMod}}$ given in Definition 2.5.1, we have the following set description for the graded chain module $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$:

$$C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} C_n(K_t; \mathbb{F})x^t$$

Then, a filtered n -chain $\sigma(x) \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is an $\mathbb{F}[x]$ -formal sum $\sigma(x) = \sum_{t \in \mathbb{N}_0} \sigma_t x^t$ of n -chains such that $\sigma_t \in C_n(K_t; \mathbb{F})$, i.e. an n -chain at index $t \in \mathbb{N}_0$, for all $t \in \mathbb{N}_0$. Note that the direct sum characterization also implies that only finitely many of σ_t can be non-trivial.

Observe that this formal sum notation for filtered n -chains makes identifying the homogeneity and degree of elements in $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ straightforward. This becomes useful since we prefer dealing with homogeneous elements of graded chain modules, as we will see later in Chapter 4. To avoid losing this benefit, we usually avoid using the indeterminate x of $\mathbb{F}[x]$ as a vertex of the simplicial complex K .

Before we proceed with an example, we have identified two useful properties of $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ resulting from Corollary 3.3.3 below.

Lemma 3.3.5. Let K_\bullet be a filtration of some simplicial complex K . Let $n \in \mathbb{Z}$.

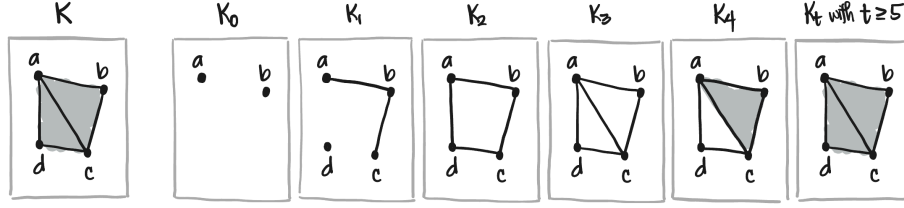
- i. For all $t \in \mathbb{N}_0$, $C_n(K_t; \mathbb{F})x^t$ is a vector subspace of $C_n(K; \mathbb{F})x^t$. Moreover, if a filtered n -chain is of the form $\sigma x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ with $\sigma \in C_n(K; \mathbb{F})$ and $t \in \mathbb{N}_0$, then $\sigma \in C_n(K_t; \mathbb{F})$.
- ii. The action of $\mathbb{F}[x]$ on $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ satisfies $x^s \cdot \sigma x^t = \sigma x^{t+s}$ for all $\sigma x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ and for all $t, s \in \mathbb{N}_0$.

Proof. For (i): $C_n(K_t; \mathbb{F})x^t \subseteq C_n(K; \mathbb{F})x^t$ since $C_n(K_t; \mathbb{F})$ is a vector subspace of $C_n(K; \mathbb{F})$ by Corollary 3.3.3. Assume there exists $\sigma x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ with $\sigma \in C_n(K; \mathbb{F})$ and $t \in \mathbb{N}_0$. We can assume, without loss of generality, $x^t \notin C_n(K_s; \mathbb{F})$ for all $s \in \mathbb{N}_0$. Therefore, σx^t can only be present in $C_n(K_t; \mathbb{F})x^t$, the homogeneous component of $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ of degree t . Therefore, $\sigma \in C_n(K_t; \mathbb{F})$.

For (ii): Let $t, s \in \mathbb{N}_0$ and let $\sigma x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$. Then, $\sigma \in C_n(K_t; \mathbb{F})x^t$. By definition of $\mathbf{\Gamma}_{\text{GrMod}}$, $x^s \cdot \sigma x^t = i_{\#}^{t+s,s}(\sigma)x^{t+s}$. By Corollary 3.3.3, $i_{\#}^{t+s,s}(\sigma)x^{t+s} = \sigma x^{t+s}$. ■

We provide an example of a filtered chain module viewed as a graded $\mathbb{Q}[x]$ -module below.

Example 3.3.6. Let K and K_\bullet be as given in Example 3.2.3, illustrations of which are copied below for convenience. Orient K and each K_t by (a, b, c, d) , with the vertex set restricted when appropriate.



A description of the 0th graded chain module $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ of K_\bullet is given below. Note that the direct sums below are to be interpreted as internal direct sums of \mathbb{Q} -vector spaces.

$$\begin{aligned} C_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) &= \mathbf{I}_{\text{GrMod}}(C_0(K_\bullet; \mathbb{Q})) \stackrel{\text{Ab}}{=} \bigoplus_{t \in \mathbb{N}_0} C_0(K_t; \mathbb{Q})x^t = \mathbb{Q}\langle a, b \rangle \oplus \left(\bigoplus_{t=1}^{\infty} \mathbb{Q}\langle a, b, c, d \rangle x^t \right) \\ &= \left\{ a \cdot f_1(x) + b \cdot f_2(x) + c \cdot x f_3(x) + d \cdot x f_4(x) : f_i(x) \in \mathbb{Q}[x] \text{ for } i = 1, 2, 3, 4 \right\} \end{aligned}$$

Using Lemma 3.3.5, we know that the action of $\mathbb{Q}[x]$ on $C_0^{\text{Gr}}(K_\bullet; \mathbb{F})$ satisfies the following:

$$\begin{aligned} x \cdot ax^t &= ax^{t+1} & \text{and} & & x \cdot bx^t &= bx^{t+1} & \text{for all } t \in \mathbb{N}_0 \\ x \cdot cx^t &= cx^{t+1} & \text{and} & & x \cdot dx^t &= dx^{t+1} & \text{for all } t \geq 1 \end{aligned}$$

Listed below are some filtered 0-chains in $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$, along with some comments.

1. The filtered 0-chain $\sigma_1(x) := 2ax + 3bx^3$ corresponds to the pair of the 0-chain $2a \in C_0(K_1; \mathbb{Q})$ at index 1 and the 0-chain $3b \in C_0(K_3; \mathbb{Q})$ at index 3. Note that $\sigma_1(x)$ is not a homogeneous element since $\deg_h(2ax) = 1 \neq 3 = \deg_h(3bx^3)$. In this case, $\deg_h(\sigma_1(x))$ is undefined.
2. The element c is not a filtered 0-chain of K_\bullet since $c \notin C_0(K_0; \mathbb{Q}) = \mathbb{Q}\langle a, b \rangle$. In contrast, for all $t \geq 1$, $cx^t \in C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ since c is a vertex in K_t .
3. The filtered 0-chain $\sigma_2(x) := (a + 2b - c)x^4$ corresponds to the 1-chain $a + 2b - c$ as an element of $C_0(K_4; \mathbb{Q})$ at index 4 and $\deg_h(\sigma_2(x)) = 4$, i.e. $\sigma_2(x)$ is homogeneous of degree 4.
4. The filtered 0-chain $a \in C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ generates the $\mathbb{Q}[x]$ -submodule $\mathbb{Q}[x]\langle a \rangle$ consisting of elements of the form $ax^t \in C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ with $t \in \mathbb{N}_0$. Observe that $\mathbb{Q}[x]\langle a \rangle$ is a graded submodule of $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$.

Recall that simplicial chain groups $C_n(K; R)$ with coefficients in a PID R are free R -modules. We have a similar result for graded chain modules.

Proposition 3.3.7. Let K_\bullet be a simplicial filtration. For all $n \in \mathbb{N}_0$, $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is free.

Proof. Let $n \in \mathbb{Z}$. If $n \geq -1$, then $C_n(K_\bullet; \mathbb{F})$ is trivial and therefore free. Assume $n \geq 0$. To determine that $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is free, it suffices to show that it is torsion-free as a $\mathbb{F}[x]$ -module. Let $\sigma x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$. Then, $\sigma \in C_n(K_t; \mathbb{F})$. Note that $C_n^{\text{Gr}}(K_\bullet; \mathbb{Q})$ is a \mathbb{F} -vector space and is torsion-free. If $x^s \cdot \sigma x^t = 0$ for some $s \in \mathbb{N}_0$, then $i_{\#}^{t+s, s}(\sigma) = 0$, contradicting Lemma 3.3.5. Therefore, $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is torsion-free and, therefore, free. ■

Earlier in Definition 1.2.5, we identified a natural choice for basis on the chain groups of a simplicial complex based on its equipped orientation. Similarly, if K_\bullet is a filtration on a simplicial complex K , an orientation on K induces a natural choice of basis on its graded chain modules. We state this as a proposition below.

Proposition 3.3.8. Let K_\bullet be a filtration of an oriented finite simplicial complex K . For each $n \in \mathbb{N}_0$, $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is free with the following homogeneous basis:

$$\mathcal{K}_n^{\text{Gr}} := \left\{ \sigma_1 x^{t_1}, \sigma_2 x^{t_2}, \dots, \sigma_m x^{t_m} \right\} \quad \text{with} \quad t_i := \min \left\{ t \in \mathbb{N}_0 : \sigma_i \in C_n(K_t; \mathbb{F}) \right\} \quad \text{for all } i \in \{1, \dots, m\} \quad (\text{E1})$$

where $K[n] = \{\sigma_1, \dots, \sigma_m\}$ is the standard basis of $C_n(K_t; \mathbb{F})$ induced by the orientation on K (see Defn. 1.2.5).

Proof. Fix $n \in \mathbb{N}_0$. Since K is a finite simplicial complex, $K[n]$ is finite with $m := \text{card}(K[n]) < \infty$ and we can label the oriented n -simplices in $K[n]$ by $\{\sigma_1, \dots, \sigma_m\}$. Let $i \in \{1, \dots, m\}$. By Definition 3.1.1(iii), there must exist some $t \in \mathbb{N}_0$ such that σ_i corresponds to an n -simplex in K and $\sigma_i \in C_n(K_t; \mathbb{F})$. Since \mathbb{N}_0 is bounded below, a minimal $t_i \in \mathbb{N}_0$ for σ_i such that $\sigma_i \in C_n(K_{t_i}; \mathbb{F})$ must exist. Then, $\sigma_i x^{t_i} \in C_n(K_{t_i}; \mathbb{F}) x^{t_i} \subseteq C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ is homogeneous with degree t_i . Therefore, $\mathcal{K}_n^{\text{Gr}}$, as given above, is well-defined as a set.

Assume, without loss of generality, that $t_1 \leq t_2 \leq \dots \leq t_m$, i.e. $\mathcal{K}_n^{\text{Gr}}$ is indexed in order of increasing degree. We need to show that $\mathcal{K}_n^{\text{Gr}}$ generates $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ and that $\mathcal{K}_n^{\text{Gr}}$ is $\mathbb{F}[x]$ -linearly independent.

Fix $t \in \mathbb{N}_0$ and consider the vector subspace $C_n(K_t; \mathbb{F}) x^t$ of $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$. Let $r \in \{1, \dots, m\}$ be maximum such that $t_r \leq t$. By minimality of the t_i 's, σ_i corresponds to an n -simplex in K_t if and only if $i \in \{1, \dots, r\}$. Let K_t inherit the orientation on K by restriction and let $K_t[n]$ be the standard basis of $C_n(K_t; \mathbb{F})$ induced by this new orientation. Then, $K_t[n] = \{\sigma_1, \dots, \sigma_r\}$ and the following set is a basis for $C_n(K_t; \mathbb{F})$:

$$\mathcal{B}(t) := \left\{ \sigma_1 x^t, \sigma_2 x^t, \dots, \sigma_r x^t \right\} = \left\{ x^{t-t_1} \cdot \sigma_1 x^{t_1}, x^{t-t_2} \cdot \sigma_2 x^{t_2}, \dots, x^{t-t_r} \cdot \sigma_r x^{t_r} \right\} \quad (\text{E2})$$

Since $t - t_i \geq 0$ for each $i \in \{1, \dots, r\}$, each element in $C_n(K_t; \mathbb{F})$ is equal to a $\mathbb{F}[x]$ -linear combination in $\mathcal{K}_n^{\text{Gr}}$. Since $t \in \mathbb{N}_0$ is arbitrary, $\mathcal{K}_n^{\text{Gr}}$ generates $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$.

Consider the $\mathbb{F}[x]$ -linear dependence relation $f(x) := \sum_{i=1}^m f_i \cdot \sigma_i x^{t_i} = 0$ for $\mathcal{K}_n^{\text{Gr}}$ with $f_1, \dots, f_m \in \mathbb{F}[x]$. Recall that $C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) = \bigoplus_{t \in \mathbb{N}_0} C_n(K_t; \mathbb{F}) x^t$. For each $i \in \{1, \dots, m\}$, let $\{k_{i,t}\}_{t \in \mathbb{N}_0}$ be such that $f_i(x) \cdot \sigma_i x^{t_i} = \sum_{t=0}^\infty k_{i,t} \cdot \sigma_i x^t$ with $k_{i,t} \in \mathbb{F}$ for all $t \in \mathbb{N}_0$. Observe that, for all $i \in \{1, \dots, m\}$, only finitely many $\{k_{i,t}\}_{t \in \mathbb{N}_0}$ are nonzero and $k_{i,t} = 0$ if $t < t_i$. Then, $f(x)$ decomposes into

$$f(x) = \sum_{i=1}^m f_i \cdot \sigma_i x^{t_i} = \sum_{i=1}^m \left(\sum_{t=0}^\infty k_{i,t} \cdot \sigma_i x^t \right) = \sum_{t=0}^\infty \left(\sum_{i=1}^m k_{i,t} \cdot \sigma_i x^t \right)$$

and the homogeneous component of $f(x)$ of degree $t \in \mathbb{N}_0$ is given by $\sum_{i=1}^m k_{i,t} \cdot \sigma_i x^t$. To solve linear dependence relation $f(x) = 0$, it suffices to consider $\sum_{i=1}^m k_{i,t} \cdot \sigma_i x^t = 0$ separately for each $t \in \mathbb{N}_0$. Let $t \in \mathbb{N}_0$. Observe that $\sum_{i=1}^m k_{i,t} \cdot \sigma_i x^t \in C_n(K_t; \mathbb{F}) x^t$ is a \mathbb{F} -linear combination in $\mathcal{B}(t)$, with $\mathcal{B}(t)$ as defined in Equation (E2). Since $\mathcal{B}(t)$ is a basis for $C_n(K_t; \mathbb{F}) x^t$, $k_{1,t} = k_{2,t} = \dots = k_{m,t} = 0$. Since t is arbitrary, $k_{i,t} = 0$ for all $t \in \mathbb{N}_0$. Then, for each $i \in \{1, \dots, m\}$,

$$f_i = \sum_{t=0}^\infty k_{i,t} x^{t-t_i} = \sum_{t=t_i}^\infty k_{i,t} x^{t-t_i} = \sum_{t=t_i}^\infty 0 \cdot x^{t-t_i} = 0$$

Therefore, $f_1 = f_2 = \dots = f_m$ and $\mathcal{K}_n^{\text{Gr}}$ is $\mathbb{F}[x]$ -linearly independent.

Since $\mathcal{K}_n^{\text{Gr}}$ is $\mathbb{F}[x]$ -linearly independent, generates $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$, and consists of homogeneous elements, $\mathcal{K}_n^{\text{Gr}}$ is a homogeneous $\mathbb{F}[x]$ -basis for $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$. ■

In the example calculations presented later in Chapter 4, we often use the basis $\mathcal{K}_n^{\text{Gr}}$ as a starting point.

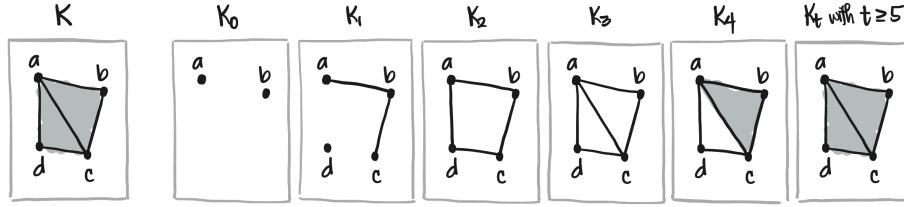
For convenience, we provide a name for this basis below.

Definition 3.3.9. Let K_\bullet be a filtration of an oriented finite simplicial complex K . For each $n \in \mathbb{N}_0$, let $\mathcal{K}_n^{\text{Gr}}$, as denoted in Proposition 3.3.8 and ordered first in increasing degree then by lexicographic order of the vertices of K , be the **standard ordered basis** of $C_n^{\text{Gr}}(K_\bullet; \mathbb{Q})$ **induced by the orientation on K** .

Remark. Other authors may use a different ordering e.g. [ZC05] orders $\mathcal{K}_n^{\text{Gr}}$ in *decreasing* degree instead. We introduced the notion for standard *ordered* basis here for convenience, as we will deal with coordinate matrices later in Chapter 4.

In the example below, we identify the basis $\mathcal{K}_n^{\text{Gr}}$ corresponding to the given graded chain modules.

Example 3.3.10. Let K and K_\bullet be as given in Example 3.2.3, illustrations of which are copied below for convenience, and orient K by the vertex order (a, b, c, d) .



Provided below are the ordered bases \mathcal{K}_n for $C_n^{\text{Gr}}(K_\bullet; \mathbb{Q})$ for $n = 0, 1, 2$, as described in Proposition 3.3.8.

$$\mathcal{K}_0^{\text{Gr}} = (a, b, cx, dx), \quad \mathcal{K}_1^{\text{Gr}} = (abx, bcx, adx^2, cdx^2, acx^3), \quad \mathcal{K}_2^{\text{Gr}} = (abcx^4, acdx^5)$$

Therefore, we can describe $C_n^{\text{Gr}}(K_\bullet; \mathbb{Q})$ for $n = 0, 1, 2$ as follows:

$$C_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) = \mathbb{Q}[x] \langle a, b, cx, dx \rangle, \quad C_1 = \mathbb{Q}[x] \langle abx, bcx, adx^2, cdx^2, acx^3 \rangle, \quad C_2 = \mathbb{Q}[x] \langle abcx^4, acdx^5 \rangle$$

Standard ordered bases allow us to represent filtered n -chains using coordinate matrices and vectors. Following the ordering of $\mathcal{K}_0^{\text{Gr}}$ and $\mathcal{K}_1^{\text{Gr}}$ denoted above, we listed some examples of these coordinate vectors below.

$$\begin{aligned} \sigma_1 &= 2ax + 3bx^3 = (a)(2x) + (b)(2x^3) &\Rightarrow [\sigma_1] &= \begin{pmatrix} a \\ b \\ cx \\ dx \end{pmatrix} \begin{pmatrix} 2x \\ 3x^3 \\ 0 \\ 0 \end{pmatrix} \\ \sigma_2 &= (ab + bc - ac)x^4 \\ &= (abx)(x^3) + (bcx)(x) + (acx^3)(-x) &\Rightarrow [\sigma_2] &= \begin{pmatrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{pmatrix} \begin{pmatrix} 2x \\ 3x^3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \sigma_3 &= (cd - ad + ac)x^5 \\ &= (adx^2)(-x^3) + (cdx^2)(x^3) + (acx^3)(x^2) &\Rightarrow [\sigma_3] &= \begin{pmatrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -x^3 \\ x^3 \\ x^2 \end{pmatrix} \end{aligned}$$

The functorial nature of the simplicial chain complex construction also allows us to extend the simplicial boundary maps to the case of persistence modules. More specifically, the collection of simplicial boundary maps form a persistence morphism between the filtered chain groups. We state this in more detail below.

Corollary 3.3.11. Let K_\bullet be a simplicial filtration and let $n \in \mathbb{Z}$. For each $t \in \mathbb{N}_0$, let $\partial_n^t : C_n(K_t; \mathbb{F}) \rightarrow C_{n-1}(K_t; \mathbb{F})$ be the n^{th} simplicial boundary map of K_t . Then, the collection $\{\partial_n^t\}_{t \in \mathbb{N}_0}$ determines a persistence morphism $\partial_n^\bullet : C_n(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}(K_\bullet; \mathbb{F})$.

Proof. Fix $n \in \mathbb{N}_0$ and let $t, s \in \mathbb{N}_0$ with $t \leq s$. By Proposition 1.3.6, $i_{n-1, \#}^{s,t} \circ \partial_n^t = \partial_n^s \circ i_{n, \#}^{s,t}$ for all $t, s \in \mathbb{N}_0$ with $t \leq s$ with $i_{n, \#}^{s,t} : C_n(K_t; \mathbb{F}) \rightarrow C_n(K_s; \mathbb{F})$ and $i_{n-1, \#}^{s,t} : C_{n-1}(K_t; \mathbb{F}) \rightarrow C_{n-1}(K_s; \mathbb{F})$ referring to structure maps of $C_n(K_\bullet; \mathbb{F})$ and $C_{n-1}(K_\bullet; \mathbb{F})$ respectively. That is, the following diagram commutes:

$$\begin{array}{ccc} C_n(K_t; R) & \xrightarrow{\partial_n^t} & C_{n-1}(K_t; R) \\ i_{n, \#}^{s,t} \downarrow & & \downarrow i_{n-1, \#}^{s,t} \\ C_n(K_s; R) & \xrightarrow{\partial_n^s} & C_{n-1}(K_s; R) \end{array}$$

Therefore, $\partial_n^\bullet = (\partial_n^t)_{t \in \mathbb{N}_0}$ is a persistence morphism by Definition 2.2.1. ■

We name these persistence morphisms, along with the corresponding graded homomorphism given by application of $\Gamma_{\text{GrMod}}(-)$, below.

Definition 3.3.12. For each $n \in \mathbb{Z}$, define the n^{th} **filtered boundary morphism** ∂_n^\bullet of a simplicial filtration K_\bullet to be the persistence morphism $\partial_n^\bullet : C_n(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}(K_\bullet; \mathbb{F})$ given by $\partial_n^\bullet = (\partial_n^t)_{t \in \mathbb{N}_0}$ where $\partial_n^t : C_n(K_t; \mathbb{F}) \rightarrow C_{n-1}(K_t; \mathbb{F})$ is the n^{th} simplicial boundary map of K_t .

Definition 3.3.13. Let K_\bullet be a filtration of a simplicial complex. For each $n \in \mathbb{Z}$, define the n^{th} **graded boundary morphism** ∂_n^{Gr} of K_\bullet to be the graded $\mathbb{F}[x]$ -module homomorphism $\partial_n^{\text{Gr}} : C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}^{\text{Gr}}(K_\bullet; \mathbb{F})$ given by $\partial_n^{\text{Gr}} := \Gamma_{\text{GrMod}}(\partial_n^\bullet)$. A **filtered n -cycle** is an element of $\ker(\partial_n^{\text{Gr}})$ and a **filtered n -boundary** that of $\text{im}(\partial_{n+1}^{\text{Gr}})$. If K is oriented and finite, define the n^{th} **graded boundary matrix** $[\partial_n^{\text{Gr}}]$ to be the matrix of ∂_n^{Gr} relative to the standard ordered bases \mathcal{K}_n and \mathcal{K}_{n-1} .

Below, we identify a useful characterization of filtered and graded boundary morphisms of a filtration K_\bullet relative to the simplicial boundary morphism of a simplicial complex K .

Corollary 3.3.14. Let K_\bullet be a filtration of a simplicial complex K and let $n \in \mathbb{N}_0$.

- i. For all $t \in \mathbb{N}_0$, $\partial_n^t(\sigma) = \partial_n(\sigma)$ for all $\sigma \in C_n(K_t; \mathbb{F})$ and $t \in \mathbb{N}_0$.
- ii. The n^{th} graded boundary map satisfies $\partial_n^{\text{Gr}}(\sigma) = \sum_{t=0}^{\infty} \partial_n(\sigma_t)x^t$ for all $\sigma = \sum_{t=0}^{\infty} \sigma_t x^t \in C_n(K_\bullet; \mathbb{F})$.

Proof. Fix $n \in \mathbb{N}_0$. For (i): Since $\partial_n^\bullet : C_n(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}(K_\bullet; \mathbb{F})$ is a persistence morphism, we have that for all $t \in \mathbb{N}_0$ and $\sigma \in C_n(K_t; \mathbb{F})$,

$$\partial_n^t(\sigma) = \left(\partial_n \circ i_{n, \#}^{[t]} \right)(\sigma) = \left(i_{n-1, \#}^{[t]} \circ \partial_n \right)(\sigma) = \left(\text{id}_{C_{n-1}(K_t; \mathbb{F})} \circ \partial_n \right)(\sigma) = \partial_n(\sigma)$$

where $i^{[t]} : K_t \rightarrow K$ refers to the inclusion map. For (ii): For all $\sigma = \sum_{t=0}^{\infty} \sigma_t x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$,

$$\partial_n^{\text{Gr}}(\sigma) = \partial_n \left(\sum_{t=0}^{\infty} \sigma_t x^t \right) \stackrel{(*)}{=} \sum_{t=0}^{\infty} \partial_n^t(\sigma_t) x^t = \sum_{t=0}^{\infty} \partial_n(\sigma_t) x^t$$

with $(*)$ given by definition of the morphism assignment of Γ_{GrMod} . ■

The corollary above tells us that, when evaluating filtered n -chains against graded boundary morphisms, we can basically ignore the added x^t in the notation and determine the boundary of an oriented n -simplex as usual. We have an example below.

Example 3.3.15. Let K and K_\bullet be as given in Example 3.2.3 and equip K with the orientation by (a, b, c, d) . Let $\sigma_1 := (ab + bc - ac)x^4 \in C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$. Then, σ_1 is a filtered 1-cycle by the following calculation:

$$\partial_1^{\text{Gr}}(\sigma_1) = \partial_1^{\text{Gr}}((ab + bc - ac)x^4) = \partial_1(ab + bc - ac)x^4 = ((b - a) + (c - b) - (c - a))x^4 = (0)x^4 = 0$$

Additionally, σ_1 is a filtered 1-boundary since $\sigma_1 \in \text{im}(\partial_2^{\text{Gr}})$ by the following calculation:

$$\partial_2^{\text{Gr}}(abcx^4) = \partial_2(abc)x^4 = (bc - ac + ab)x^4 = \sigma_1$$

Let $\sigma_2 := (ab + bc - ac)x^3 \in C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$. Then, $\partial_1^{\text{Gr}}(\sigma_2) = \partial_1(ab + bc - ac)x^3 = 0$ and σ_2 is a filtered 1-cycle. However, σ_2 is not a filtered 1-boundary since abc is not in K_3 and $abcx^3 \notin C_2^{\text{Gr}}(K_\bullet; \mathbb{Q})$.

We need to verify that the \mathbb{Z} -indexed collection of filtered chain modules $C_n(K_\bullet; \mathbb{Q})$ and filtered boundary morphisms $\partial_n^\bullet : C_n(K_\bullet; \mathbb{Q}) \rightarrow C_{n-1}(K_\bullet; \mathbb{Q})$ does indeed determine a persistence complex. We also need to do the same for the graded chain modules and graded boundary morphisms.

Proposition 3.3.16. Let K_\bullet be a simplicial filtration. Then, $\partial_{n-1}^\bullet \circ \partial_n^\bullet = 0_\bullet$ and $\partial_{n-1}^{\text{Gr}} \circ \partial_n^{\text{Gr}} = 0$ for all $n \in \mathbb{Z}$.

Proof. Fix $n \in \mathbb{N}_0$. For the persistence modules case: For all $t \in \mathbb{N}_0$, $(\partial_{n-1}^\bullet \circ \partial_n^\bullet)_t = \partial_{n-1}^\bullet \circ \partial_n^\bullet = 0$ as \mathbb{F} -linear maps. Therefore, $\partial_n^\bullet \circ \partial_{n-1}^\bullet = 0_\bullet$, with 0_\bullet denoting the zero persistence morphism. For the graded module case: For all $\sigma = \sum_{t=0}^\infty \sigma_t x^t \in C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$, we have the following:

$$(\partial_{n-1}^{\text{Gr}} \circ \partial_n^{\text{Gr}})(\sigma) = \partial_{n-1}^{\text{Gr}}\left(\sum_{t=0}^\infty \partial_n(\sigma_t)x^t\right) = \sum_{t=0}^\infty (\partial_{n-1} \circ \partial_n)x^t = \sum_{t=0}^\infty 0 \cdot x^t = 0$$

Therefore, $\partial_{n-1}^{\text{Gr}} \circ \partial_n^{\text{Gr}} = 0$. ■

Now, we define the *simplicial persistence complex* of a simplicial filtration, along with its corresponding chain complex of graded modules.

Definition 3.3.17. Define the **simplicial persistence complex** $C_*(K_\bullet; \mathbb{F})$ and the **simplicial graded chain complex** $C_*^{\text{Gr}}(K_\bullet; \mathbb{F})$ of a simplicial filtration K_\bullet with coefficients in \mathbb{F} as follows:

$$C_*(K_\bullet; \mathbb{F}) := \left(C_n(K_\bullet; \mathbb{F}), \partial_n^\bullet \right)_{n \in \mathbb{Z}} \quad \text{and} \quad C_*^{\text{Gr}}(K_\bullet; \mathbb{F}) := \left(C_n^{\text{Gr}}(K_\bullet; \mathbb{F}), \partial_n^{\text{Gr}} \right)_{n \in \mathbb{Z}}$$

The simplicial persistence complex $C_*(K_\bullet; \mathbb{F})$ can be visualized as the following sequence of persistence modules and persistence morphisms:

$$\cdots \xrightarrow{\partial_{n+2}^\bullet} \boxed{C_{n+1}(K_\bullet; \mathbb{F})} \xrightarrow{\partial_{n+1}^\bullet} \boxed{C_n(K_\bullet; \mathbb{F})} \xrightarrow{\partial_n^\bullet} \boxed{C_{n-1}(K_\bullet; \mathbb{F})} \xrightarrow{\partial_{n-1}^\bullet} \cdots$$

The change in perspective from having the index $t \in \mathbb{N}_0$ of a filtration K_\bullet take precedence to that for the index $n \in \mathbb{N}_0$ of the chain complex C_* can be visualized using the following commutative diagram. In the diagram

below, each row fixes the index $t \in \mathbb{N}_0$ of the filtration and represents the simplicial chain complex $C_*(K_t; \mathbb{F})$ of K_t . Similarly, each column fixes the dimension $n \in \mathbb{Z}$ and corresponds to a filtered chain module $C_n(K_\bullet; \mathbb{F})$. For clarity, we suppressed the index t in $i_\#^{s,t}$ and highlighted the vector spaces of $C_{n+1}(K_\bullet; \mathbb{F})$ in red, those of $C_n(K_\bullet; \mathbb{F})$ in blue, and $C_{n-1}(K_\bullet; \mathbb{F})$ in green.

$$\begin{array}{ccccccc}
 & \vdots & & \vdots & & \vdots & \\
 & i_\# \uparrow & & i_\# \uparrow & & i_\# \uparrow & \\
 \cdots & \xrightarrow{\partial_{n+2}} & C_{n+1}(K_{t+1}; \mathbb{F}) & \xrightarrow{\partial_{n+1}} & C_n(K_{t+1}; \mathbb{F}) & \xrightarrow{\partial_n} & C_{n-1}(K_{t+1}; \mathbb{F}) \xrightarrow{\partial_{n-1}} \cdots \\
 & i_\# \uparrow & & i_\# \uparrow & & i_\# \uparrow & \\
 \cdots & \xrightarrow{\partial_{n+2}} & C_{n+1}(K_t; \mathbb{F}) & \xrightarrow{\partial_{n+1}} & C_n(K_t; \mathbb{F}) & \xrightarrow{\partial_n} & C_{n-1}(K_t; \mathbb{F}) \xrightarrow{\partial_{n-1}} \cdots \\
 & i_\# \uparrow & & i_\# \uparrow & & i_\# \uparrow & \\
 \cdots & \xrightarrow{\partial_{n+2}} & C_{n+1}(K_{t-1}; \mathbb{F}) & \xrightarrow{\partial_{n+1}} & C_n(K_{t-1}; \mathbb{F}) & \xrightarrow{\partial_n} & C_{n-1}(K_{t-1}; \mathbb{F}) \xrightarrow{\partial_{n-1}} \cdots \\
 & i_\# \uparrow & & i_\# \uparrow & & i_\# \uparrow & \\
 & \vdots & & \vdots & & \vdots &
 \end{array}$$

Finally, we present the result that allows us to use both the simplicial persistence complex and the simplicial graded complex in our calculations.

Proposition 3.3.18. Let K_\bullet be a simplicial filtration. For all $n \in \mathbb{Z}$, we have that

$$H_n(K_\bullet; \mathbb{F}) \cong H_n^{\text{Pers}}(C_*(K_\bullet; \mathbb{F})) \cong (\Gamma_{\text{Pers}} \circ H_n^{\text{Gr}})(C_*^{\text{Gr}}(K_\bullet; \mathbb{F})) \quad \text{as persistence modules over } \mathbb{F}$$

where $H_n^{\text{Pers}} : \mathbf{Ch-Pers}_{\mathbb{F}} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ and $H_n^{\text{Gr}} : \mathbf{Ch-GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ refer to the chain homology functors on persistence complexes and graded chain complexes respectively.

Proof. Recall that $H_n(-; \mathbb{F}) = H_n^{\text{Pers}} \circ C_*(-; \mathbb{F})$ as functors $\mathbf{A-Simp} \rightarrow \mathbf{Vect}_{\mathbb{F}}$, where $H_n(-; \mathbb{F})$ and $C_*(-; \mathbb{F})$ are the simplicial homology and simplicial chain complex functors respectively. Since functor composition is associative when defined, we have the following:

$$H_n(K_\bullet; \mathbb{F}) = H_n(-; \mathbb{F})(K_\bullet) = (H_n^{\text{Pers}} \circ C_*(-; \mathbb{F}))(K_\bullet) = H_n^{\text{Pers}}(C_*(K_\bullet; \mathbb{F}))$$

As stated in Proposition 2.5.15, Γ_{GrMod} and Γ_{Pers} preserve chain complexes and chain homology. In particular, both functors commute with the chain homology functors. Then,

$$\begin{aligned}
 H_n^{\text{Pers}} \circ C_*(K_\bullet; \mathbb{F}) &= \text{id}_{\mathbf{Pers}_{\mathbb{F}}} \circ H_n^{\text{Pers}} \circ C_*(K_\bullet; \mathbb{F}) = \Gamma_{\text{Pers}} \circ \Gamma_{\text{GrMod}} \circ H_n^{\text{Pers}} \circ C_*(K_\bullet; \mathbb{F}) \\
 &= \Gamma_{\text{Pers}}(\Gamma_{\text{GrMod}} \circ H_n^{\text{Pers}} \circ C_*(K_\bullet; \mathbb{F})) = \Gamma_{\text{Pers}}(H_n^{\text{Gr}} \circ \Gamma_{\text{GrMod}} \circ C_*(K_\bullet; \mathbb{F})) \\
 &= (\Gamma_{\text{Pers}} \circ H_n^{\text{Gr}})(C_*^{\text{Gr}}(K_\bullet; \mathbb{F})) \quad \blacksquare
 \end{aligned}$$

The proposition above is a fundamental result behind the derivation of the *matrix reduction algorithm for persistent homology*, as presented in [ZC05], since it allows us to calculate at the level of graded modules and return back to that of persistence modules post-calculation. For convenience, we name the chain homology of the graded persistence complex.

Definition 3.3.19. For each $n \in \mathbb{Z}$, define the n^{th} **graded homology module** $H_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ with coefficients in a field \mathbb{F} of a simplicial filtration K_\bullet as $H_n^{\text{Gr}}(K_\bullet; \mathbb{F}) := H_n \circ C_*^{\text{Gr}}(K_\bullet; \mathbb{F})$, i.e. the n^{th} chain homology of the simplicial graded chain complex.

Chapter 4. Calculation by Matrices

The paper *Calculating Persistent Homology* [ZC05] by Afra Zomorodian and Gunnar Carlsson describes how the *matrix reduction algorithm for persistent homology* is based on a method of calculating invariant factor decompositions of finitely-generated $\mathbb{F}[x]$ -modules using matrices over $\mathbb{F}[x]$. Certain assumptions on the $\mathbb{F}[x]$ -modules in question allow this calculation to be simulated using matrices over \mathbb{F} . In this chapter, we frame the discussion in [ZC05] relative to the theoretical foundation established in the previous chapters.

Fix a field \mathbb{F} and let K_\bullet be a simplicial filtration of a finite simplicial complex K . As established in Section 3.3, the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F})$ of K_\bullet with coefficients in \mathbb{F} can be calculated using the following persistence isomorphism relation:

$$H_n(K_\bullet; \mathbb{F}) \cong \left(\Gamma_{\text{Pers}} \circ H_n \circ \Gamma_{\text{GrMod}} \right) \left(C_*(K_\bullet; \mathbb{F}) \right)$$

where $C_*(K_\bullet; \mathbb{F}) = (C_m(K_\bullet; \mathbb{F}); \partial_m^\bullet)_{m \in \mathbb{Z}}$ is the simplicial persistence complex of K_\bullet , $H_n : \mathbf{Ch-GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ is the n^{th} chain homology functor on graded chain complexes, and $\Gamma_{\text{GrMod}} : \mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ and $\Gamma_{\text{Pers}} : \mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ refer to the category equivalence discussed in Section 2.5.

Following the arguments in [ZC05], the *matrix reduction algorithm for persistent homology* calculates persistent homology at the level of graded modules. In particular, the algorithm determines the *graded invariant factor decomposition* of the following graded $\mathbb{F}[x]$ -module by matrix reduction:

$$H_n^{\text{Gr}}(K_\bullet; \mathbb{F}) := \left(H_n \circ \Gamma_{\text{GrMod}} \right) \left(C_*(K_\bullet; \mathbb{F}) \right)$$

The resulting decomposition then determines the *interval decomposition* of the persistence module $H_n(K_\bullet; \mathbb{F})$, similarly as in Lemma 2.5.11 and Corollary 2.5.13.

In this chapter, we take a more general view and explore how matrix reduction can be used to find graded invariant factor decompositions of the n^{th} chain homology of graded chain complexes. Let R be a PID and \mathbb{F} be a field. This chapter is structured as follows:

In Section 4.1. The Structure Theorem and Smith Normal Decompositions

We consider the Structure Theorem for Finitely Generated Modules over a PID R in the category \mathbf{Mod}_R , i.e. disregarding grading (if it exists), and discuss how *invariant factor decompositions* of said modules can be calculated using presentations and a matrix factorization called Smith Normal Decomposition (SND). Note that our examples in this section use $R = \mathbb{Z}$ for comparison to the graded case in the later sections.

In Section 4.2. Matrix Calculation of Homology of Ungraded Chain Complexes

We consider chain complexes $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ in $\mathbf{Ch-Mod}_R$ such that for all $n \in \mathbb{Z}$, C_n is a free R -module of finite-rank. For each $n \in \mathbb{Z}$, we present an existence result involving the existence of a decomposition of C_n into three free direct summands:

$$C_n \cong K_n^{\text{free}} \oplus K_n^{\text{tor}} \oplus \frac{C_n}{\ker(\partial_n)}$$

such that the free component and torsion component of $H_n(C_*)$ are given by $F(H_n(C_*)) \cong K_n^{\text{free}}$ and $T(H_n(C_*)) \cong K_n^{\text{tor}} / \text{im}(\partial_{n+1})$ respectively. We also discuss how these components can be determined from specific SNDs of the matrices of the differentials $\partial_{n+1} : C_{n+1} \rightarrow C_n$ and $\partial_n : C_n \rightarrow C_{n-1}$.

In Section 4.3. The Graded Structure Theorem and SNDs in the Graded Case

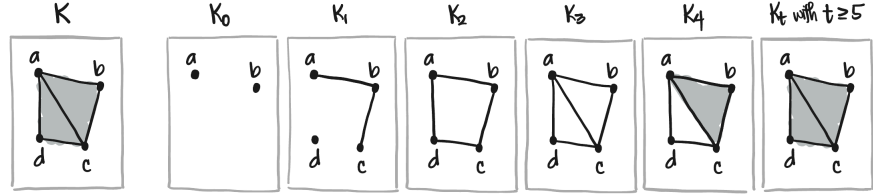
We present the Graded Structure Theorem for Finitely-Generated $\mathbb{F}[x]$ -modules in the category $\mathbf{GrMod}_{\mathbb{F}[x]}$. We also discuss how this theorem can be considered a special case of the

Structure Theorem in $\mathbf{Mod}_{\mathbb{F}[x]}$ and how the method of calculating invariant factor decompositions by SNDs, limited to *graded presentations*, can be used to find *graded* invariant factor decompositions of graded $\mathbb{F}[x]$ -modules. For convenience, we use $\mathbb{F} = \mathbb{Q}$ for our examples.

In Section 4.4. Matrix Reduction of Graded Matrices

We consider matrices of homomorphisms in graded presentations, which we call *graded matrices* for brevity. We discuss how specific matrix operations, i.e. elementary permutations, elementary dilations, and non-trivial elimination operations in matrices over $\mathbb{F}[x]$, preserve the homogeneity of graded matrices.

We also consider the simplicial filtration K_\bullet presented in Example 3.2.3, i.e. that in [ZC05, Figure 1]. An illustration of K_\bullet is copied below for convenience.



In particular, we perform matrix reduction on $[\partial_1^{\text{Gr}}]$ and $[\partial_2^{\text{Gr}}]$, the matrices of the graded boundary maps $\partial_1^{\text{Gr}} : C_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) \rightarrow C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ and $\partial_2^{\text{Gr}} : C_2^{\text{Gr}}(K_\bullet; \mathbb{Q}) \rightarrow C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ respectively, and discuss why the homogeneity of said matrices are preserved after the listed matrix operations.

In Section 4.5. An Ungraded SND Algorithm in the Graded Case

We present a general algorithm for finding SNDs of graded matrices, adapted from an algorithm for finding SNDs of matrices over a PID R , and discuss why the SNDs resulting from this algorithm can be used to determine *graded* invariant factor decompositions.

We also use this algorithm to determine the graded invariant factor decomposition of the graded $\mathbb{Q}[x]$ -module $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ of K_\bullet and use said result, along with Lemma 2.5.11 and Corollary 2.5.13, to determine the interval decomposition of the persistent homology module $C_0(K_\bullet; \mathbb{Q})$ of K_\bullet in dimension $n = 0$.

In Section 4.6. Matrix Calculation of Homology of Graded Chain Complexes

We consider graded chain complexes $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ in $\mathbf{Ch-GrMod}_{\mathbb{F}[x]}$ such that C_n is a free graded $\mathbb{F}[x]$ -module of finite rank. We briefly discuss why the decomposition of C_n into the free summands K_n^{free} , K_n^{tor} , and $C_n / \ker \partial_n$ and how the method of finding ungraded invariant factor decompositions of chain homology in $\mathbf{Ch-Mod}_{\mathbb{F}[x]}$ extends to the graded case.

Some of the notation used involving matrices and matrix reduction are identified Appendix A2. In particular, we bring emphasis to the elementary matrices $E_{\text{swap}}^{[n]}(k_1, k_2)$, $E_{\text{dilate}}^{[n]}(k, \mu)$, and $E_{\text{add}}^{[n]}(k_j, k_i; \alpha) \in \text{GL}(n, R)$ described in Definition A2.5.

Section 4.1. The Structure Theorem and Smith Normal Decompositions

The Structure Theorem for Finitely Generated Modules over a PID, which we call the **Structure Theorem** in this paper for convenience, proves the existence and uniqueness of invariant factor decompositions of certain modules. In this section, we discuss one of the proofs of the Structure Theorem, as presented in [DF03], and how the method of calculating invariant factor decompositions can re-stated as a matrix calculation using a matrix factorization called a *Smith Normal Decomposition*, defined later in this section in Definition 4.1.8. To start, we provide a statement of the Structure Theorem below.

Theorem 4.1.1. The Structure Theorem for Finitely Generated Modules over a PID.

Let M be a finitely generated module over some PID R . There exists a set of **invariant factors** $\{d_i\}_{i=1}^n$ of non-invertible elements $d_i \in R$ with divisibility relation $d_1 \mid d_2 \mid \cdots \mid d_n$ such that M is isomorphic to a direct sum of cyclic modules as follows:

$$M \cong \bigoplus_{i=1}^n R / (d_i) = R / (d_1) \oplus R / (d_2) \oplus \cdots \oplus R / (d_n)$$

This direct sum is called the **invariant factor decomposition** of M and is unique up to isomorphism. The invariant factors d_i are unique up to multiplication by units.

Remark. We refer to [DF03, Theorem 12.5] and [DF03, Theorem 12.9] for the proofs of the existence and uniqueness claims respectively. Note that a key characteristic of a PID is that all of its ideals can be generated by a single element. Given $d \in R$ with R a PID, we write (d) to refer to the ideal generated by d by $(d) := Rd = \{rd : r \in R\}$.

The divisibility relation $d_1 \mid d_2 \mid \cdots \mid d_n$ on the invariant factors $\{d_i\}$ in the Structure Theorem is sometimes stated in terms of proper ideals, wherein each (d_i) must be a proper ideal of R , i.e. $(d_i) \neq R$, and we have the following decreasing sequence of ideals:

$$(d_1) \supseteq (d_2) \supseteq \cdots \supseteq (d_n)$$

Here, we use the term *decreasing* relative to the subset relation. Given $a, b \in R$ such that a divides b , i.e. $a \mid b$, there must exist $q \in R$ such that $aq = b$. Then, any element $rb \in (b)$ with $r \in R$ must also be in (ra) since $rb = r(aq) = (rq)a$ and $rq \in R$. Therefore, $(b) \subseteq (a)$ or equivalently, $(a) \supseteq (b)$.

Since the ideals (d_i) must be proper ideals, this means that the invariant factors d_i cannot be invertible. More specifically, if $a \in R$ is invertible, then $(a) = (a^{-1}a) = (1) = R$ where $a^{-1} \in R$ refers to the multiplicative inverse of a in R . If $d_i \in R$ were invertible, then the summand $R / (d_i) = R / R$ would be the trivial module and, therefore, can be removed from the direct sum without invalidating the isomorphism.

We also want to emphasize that the Structure Theorem allows the invariant factors d_i to be zero. Note that, relative to the decreasing sequence of ideals, this means that $d_i = 0$ would occur at the end of the sequence since $(0) \subseteq (a)$ for any $a \in R$. Consequently, some references prefer to only consider nonzero d_i 's for the invariant factors and state Theorem 4.1.1 as follows:

$$M \cong R^f \oplus R / (d_1) \oplus R / (d_2) \oplus \cdots \oplus R / (d_r)$$

with $f + r = n$ and $d_i \neq 0$ for all $i \in \{1, \dots, r\}$. In this case, we call R^f the *free component* of M and $f \in \mathbb{N}_0$ the rank of M . We call the remaining part of the decomposition $\bigoplus_{i=1}^r R / (d_i)$ the *torsion component* of M . We prefer the statement in Theorem 4.1.1 since it is more suitable with the matrix calculation we present in this section.

The proof for the existence of invariant factor decompositions presented in [DF03, Theorem 12.5] relies on a system of generators and relations of a finitely generated module over a PID. We provide an alternate characterization of these systems below.

Definition 4.1.2. A **presentation** of a module M over a PID R is an exact sequence

$$F_S \xrightarrow{\varphi} F_G \xrightarrow{\pi} M \rightarrow 0$$

of free R -modules F_S and F_G with homomorphisms $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$. We call F_G and F_S the **module of generators** and **module of relations** respectively. Given a basis A of F_G and S of F_S , we call $\pi(A) \subseteq M$ and $(\varphi \circ \pi)(S) \subseteq M$ a **system of generators and relations for M** respectively. A **finite presentation** of M is a presentation wherein both F_S and F_G have finite rank.

When we say that a presentation of M is given by $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$, we refer to the exact sequence as given above.

Remark. We usually indicate that a presentation $\varphi : F_S \rightarrow F_G$ is finite by listing a finite basis $S = (\sigma_1, \dots, \sigma_n)$ of F_S and $A = (\alpha_1, \dots, \alpha_m)$ of F_G . Also, this notion of presentation is not generally compatible with the notion of presentation of groups, particularly in the case of non-abelian groups.

In the proof presented in [DF03, Theorem 12.5], a presentation of M is constructed using a set of generators $\{a_1, \dots, a_m\}$ of M , which exists by assumption of M being finitely generated. Let $\{\alpha_1, \dots, \alpha_m\}$ be a set of indeterminates. Define the module of generators to be $F_G = R\langle\alpha_1, \dots, \alpha_m\rangle$ and the homomorphism $\pi : F_G \rightarrow M$ by $\alpha_j \mapsto a_j$ for $j \in \{1, \dots, m\}$, i.e. α_j is essentially a relabeling of $a_j \in M$. Note that we use different labels for a_j and α_j since $a_j \in M$ may be a torsion element of M , i.e. there may exist $r \in R$ such that $r \cdot a_j = 0$, but $\alpha_j \in F_G$ cannot be since it is an element of the free R -module F_G . Then, the following exact sequence is a finite presentation for M :

$$\begin{array}{ccccccc} F_S & & & F_G & & & \\ \parallel & & & \parallel & & & \\ \ker(\pi) & \xrightarrow{\varphi} & R\langle\alpha_1, \dots, \alpha_m\rangle & \xrightarrow{\pi} & M & \longrightarrow & 0 \end{array}$$

where $\varphi : \ker(\pi) \rightarrow R\langle\alpha_1, \dots, \alpha_m\rangle$ is taken to be the inclusion map. Since $\ker(\pi)$ is a submodule of a finitely generated module over a PID, $\ker(\pi)$ must also be finitely generated. Note that this is not true in general if R is not a PID. Then, $\ker(\pi)$ is a free module with finite basis, i.e. of finite rank. Then, the presentation above determines M by the following isomorphism:

$$M \stackrel{(1)}{=} \operatorname{im}(\pi) \stackrel{(\star)}{\cong} F_G / \ker(\pi) \stackrel{(2)}{=} F_G / \operatorname{im}(\varphi)$$

where (\star) is given by the first isomorphism theorem on $\pi : F_G \rightarrow M$, and (1) and (2) are both given by the exactness of the sequence $F_S \rightarrow F_G \rightarrow M \rightarrow 0$, i.e. $\operatorname{im}(\pi) = \ker(M \rightarrow 0) = M$ and $\ker(\pi) = \operatorname{im}(\varphi)$ respectively. Note that the homomorphism $\varphi : F_S \rightarrow F_G$ determines M up to isomorphism since $M \cong F_G / \operatorname{im}(\varphi)$.

Observe that this isomorphism holds true even if F_S is not exactly $\ker(\pi)$, i.e. F_S only needs to contain $\ker(\pi)$ as a submodule, or if $\varphi : F_S \rightarrow F_G$ is not an inclusion map. That is, the isomorphism holds for arbitrary presentations given by Definition 4.1.2. We discuss state this in more detail later in Proposition 4.1.5.

The next step in the proof of [DF03, Theorem 12.5] involves finding a basis on $\ker(\pi)$ and F_G , as denoted in Definition 4.1.2, of a presentation such that certain properties are fulfilled. The existence of such a basis is guaranteed by the following theorem, taken from [DF03, Theorem 12.5].

Theorem 4.1.3. Invariant Factor Theorem for Submodules.

Let M be a free module over a PID R with $\text{rank}(M) = m$ and let L be a submodule of M . Then, L is a free submodule with $\text{rank}(L) = r \leq m$ and there exists a basis $B = \{\beta_1, \dots, \beta_m\}$ of M and nonzero elements $d_1, \dots, d_r \in R$ with divisibility relation $d_1 \mid d_2 \mid \dots \mid d_r$ such that $\{d_1\beta_1, \dots, d_r\beta_r\}$ is a basis of L . Furthermore, the elements d_1, \dots, d_r are unique up to multiplication by units.

Remark. For a proof of the existence and uniqueness claims, see [DF03, Theorem 12.4] and [DF03, Theorem 12.9] respectively.

With M and L as denoted in the theorem above, we have the following decompositions for M and L using the bases $\{\beta_1, \dots, \beta_m\}$ and $\{d_1\beta_1, \dots, d_r\beta_r\}$:

$$\begin{aligned} M &\cong R\langle\beta_1\rangle \oplus \dots \oplus R\langle\beta_r\rangle \oplus R\langle\beta_{r+1}\rangle \oplus \dots \oplus R\langle\beta_m\rangle \\ L &\cong R\langle d_1\beta_1\rangle \oplus \dots \oplus R\langle d_r\beta_r\rangle \end{aligned}$$

Note that the elements $d_1, \dots, d_r \in R$ given by this theorem may still be invertible as elements of R , e.g. $d_j = 1$ for some $j \in \{1, \dots, r\}$. Observe that each basis element $d_j\beta_j$ of L is then associated with a unique basis element β_j of M . Since for each $j \in \{1, \dots, r\}$, $R\langle d_j\beta_j\rangle$ is a submodule of $R\langle\beta_j\rangle$, we can characterize the quotient module M/L by considering each pair of $R\langle\beta_j\rangle$ and $R\langle d_j\beta_j\rangle$ as a torsion summand of M/L . We state the result that allows us to do this below.

Lemma 4.1.4. Let M and N be modules over a ring R . Let A be a submodule of M and B that of N . Then,

$$\frac{M \oplus N}{A \oplus B} \cong \left(\frac{M}{A} \right) \oplus \left(\frac{N}{B} \right)$$

Proof. Let $\pi_1 : M \rightarrow M/A$ and $\pi_2 : N \rightarrow N/B$ be canonical quotient homomorphisms. Note that $\ker(\pi_1) = A$, $\ker(\pi_2) = B$ and that both π_1 and π_2 are surjective, i.e. $\text{im}(\pi_1) = M/A$, $\text{im}(\pi_2) = N/B$. The direct sum of modules induces a homomorphism $\pi = \pi_1 \oplus \pi_2$ as follows:

$$\begin{aligned} \pi : M \oplus N &\rightarrow (M/A) \oplus (N/B) \\ (m, n) &\mapsto (\pi_1(m), \pi_2(n)) = (m + A, n + B) \end{aligned}$$

Then, $\ker(\pi) = \ker(\pi_1) \oplus \ker(\pi_2) = A \oplus B$ and $\text{im}(\pi) = \text{im}(\pi_1) \oplus \text{im}(\pi_2) = (M/A) \oplus (N/B)$. By the first isomorphism theorem on π ,

$$\text{im}(\pi) = \left(\frac{M}{A} \right) \oplus \left(\frac{N}{B} \right) \cong \frac{M \oplus N}{\ker(\pi)} = \frac{M \oplus N}{A \oplus B}.$$

■

We then apply Theorem 4.1.3 on the image of a finite presentation of a finitely generated module over a PID. Note that the basis from this theorem allows us to apply Lemma 4.1.4. We state this in more detail below.

Proposition 4.1.5. Let M be a finitely generated module over a PID R . Let $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ correspond to a presentation of M with $\text{rank}(F_S) = n$ and $\text{rank}(F_G) = m$. Then, there exists a basis $\{\beta_1, \dots, \beta_m\}$ of F_G and nonzero values $d_1, \dots, d_r \in R$ with divisibility relation $d_1 \mid d_2 \mid \dots \mid d_r$ such that $\{d_1\beta_1, \dots, d_r\beta_r\}$ is a basis for $\text{im}(\varphi)$ and

$$M \cong \frac{F_G}{\text{im}(\varphi)} \cong \left(\frac{R\langle\beta_1\rangle}{R\langle d_1\beta_1\rangle} \right) \oplus \dots \oplus \left(\frac{R\langle\beta_r\rangle}{R\langle d_r\beta_r\rangle} \right) \oplus R\langle\beta_{r+1}\rangle \oplus \dots \oplus R\langle\beta_m\rangle$$

Proof. Since images of module homomorphisms are submodules of the codomain, $\text{im}(\varphi)$ is a submodule of F_G . Let the basis $B = \{\beta_1, \dots, \beta_m\}$ of F_G and nonzero elements $d_1, \dots, d_r \in R$ be given by Theorem 4.1.3 on $\text{im}(\varphi) =: L$. Then, $\text{im}(\varphi) \cong R\langle d_1\beta_1 \rangle \oplus \dots \oplus R\langle d_r\beta_r \rangle$ and $F_G = R\langle \beta_1 \rangle \oplus \dots \oplus R\langle \beta_m \rangle$. Observe that for each $j \in \{1, \dots, r\}$, $R\langle d_1\beta_j \rangle$ is a submodule of $R\langle \beta_j \rangle$. We apply Lemma 4.1.4 on $M \cong F_G / \text{im}(\varphi)$ as follows:

$$\begin{aligned} M &\cong \frac{F_G}{\text{im}(\varphi)} \cong \frac{R\langle \beta_1 \rangle \oplus \dots \oplus R\langle \beta_m \rangle}{R\langle d_1\beta_1 \rangle \oplus \dots \oplus R\langle \beta_r \rangle} \\ &\cong \frac{R\langle \beta_1 \rangle \oplus \dots \oplus R\langle \beta_m \rangle}{R\langle d_1\beta_1 \rangle \oplus \dots \oplus R\langle d_r\beta_r \rangle \oplus \underbrace{0 \oplus \dots \oplus 0}_{m-r \text{ times}}} \\ &\cong \left(\frac{R\langle \beta_1 \rangle}{R\langle d_1\beta_1 \rangle} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_r \rangle}{R\langle d_r\beta_r \rangle} \right) \oplus \left(\frac{R\langle \beta_{r+1} \rangle}{0} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_m \rangle}{0} \right) \\ &\cong \left(\frac{R\langle \beta_1 \rangle}{R\langle d_1\beta_1 \rangle} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_r \rangle}{R\langle d_r\beta_r \rangle} \right) \oplus R\langle \beta_{r+1} \rangle \oplus \dots \oplus R\langle \beta_m \rangle \end{aligned}$$

where the trivial R -module is denoted by 0. ■

Note that, in the statement of the Structure Theorem (Theorem 4.1.1), the invariant factors d_1, \dots, d_n may include zero elements at the end of the sequence while the elements d_1, \dots, d_r , as denoted in Proposition 4.1.5, are defined to be nonzero. The notation does suggest that d_1, \dots, d_r are related to the invariant factors. By Proposition 4.1.5, the module of generators F_G has $\text{rank}(F_G) = m$. The zero invariant factors correspond to the free summands of M . Since $R\langle 0\beta_j \rangle = \{r(0\beta_j) : r \in R\} = \{0\}$ becomes trivial, we can define additional elements d_{r+1}, \dots, d_m to be zero and present the decomposition from Proposition 4.1.5 as follows:

$$M \cong \bigoplus_{j=1}^m \left(\frac{R\langle \beta_j \rangle}{d_j \beta_j} \right) = \underbrace{\left(\frac{R\langle \beta_1 \rangle}{R\langle d_1\beta_1 \rangle} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_r \rangle}{R\langle d_r\beta_r \rangle} \right)}_{\text{either torsion } R\text{-modules or trivial}} \oplus \underbrace{\left(\frac{R\langle \beta_{r+1} \rangle}{R\langle d_{r+1}\beta_{r+1} \rangle} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_m \rangle}{R\langle d_m\beta_m \rangle} \right)}_{\text{these are free } R\text{-modules since } d_j = 0}$$

This direct sum is then transformed into an invariant factor decomposition by replacing each summand into either a copy of R or a cyclic ideal of R . We state the required isomorphisms for these below.

Lemma 4.1.6. Let $R\langle a \rangle$ be a free module over a PID R with basis $\{a\}$ and let $d \in R$ be nonzero. Then, $R\langle a \rangle \cong R$. If d is invertible, then $R\langle a \rangle / R\langle da \rangle \cong 0$, i.e. the trivial module. Otherwise, $R\langle a \rangle / R\langle da \rangle \cong R / (d)$.

Proof. Let $f : R\langle a \rangle \rightarrow R$ be given by $a \mapsto 1$ where $1 \in R$ refers to the identity element of R . Observe that f is a homomorphism with inverse $r \mapsto ra \in R\langle a \rangle$ for all $r \in R$. Then, f is an isomorphism and $R\langle a \rangle \cong R$. We then examine two cases below.

1. Assume that $d \in R$ is invertible, i.e. there exists $d^{-1} \in R$ such that $d^{-1}d = 1$. Since $R\langle da \rangle \subseteq R\langle a \rangle$, it suffices to show $R\langle da \rangle \subseteq R\langle a \rangle$ to claim that $R\langle da \rangle = R\langle a \rangle$. Each element $ra \in R\langle a \rangle$ with $r \in R$ is generated by $s = d^{-1}r$ in $R\langle da \rangle$ as follows:

$$(s)(da) = (d^{-1}r)(da) = d^{-1}d(ra) = ra.$$

Note that PIDs are commutative by definition. Therefore, $R\langle da \rangle = R\langle a \rangle$ and $R\langle a \rangle / R\langle da \rangle = 0$.

2. Let $\pi : R \rightarrow R / (d)$ be the canonical quotient map. By the 1st isomorphism theorem on $(\pi \circ f)$:

$$R / (d) = \text{im}(\pi \circ f) \cong \frac{R\langle a \rangle}{\ker(\pi \circ f)} = \frac{R\langle a \rangle}{\{ra : r \in (d)\}} = \frac{R\langle a \rangle}{R\langle da \rangle}.$$

■

Then, by applying Lemma 4.1.6 to the direct sum resulting from Proposition 4.1.5 and removing the trivial summands, we get an invariant factor decomposition of a module over a PID. Since the nonzero values d_1, \dots, d_r , as denoted in Proposition 4.1.5, must satisfy the divisibility condition $d_1 \mid \dots \mid d_r$, any invertible values must occur at the beginning of the list. That is, we have some k such that $1 \leq k \leq r$ and the elements d_1, \dots, d_k are invertible (contributing to trivial summands) and the elements d_{k+1}, \dots, d_r are not invertible. Then, we get an invariant factor decomposition as follows:

$$\begin{aligned}
M &\cong \overbrace{\left(\frac{R\langle\beta_1\rangle}{R\langle d_1\beta_1\rangle} \right) \oplus \dots \oplus \left(\frac{R\langle\beta_k\rangle}{R\langle d_k\beta_k\rangle} \right)}^{\text{trivial summands}} \oplus \overbrace{\left(\frac{R\langle\beta_{k+1}\rangle}{R\langle d_{k+1}\beta_{k+1}\rangle} \right) \oplus \dots \oplus \left(\frac{R\langle\beta_r\rangle}{R\langle d_r\beta_r\rangle} \right)}^{\text{torsion summands}} \oplus \overbrace{R\langle\beta_{r+1}\rangle \oplus \dots \oplus R\langle\beta_m\rangle}^{\text{free summands}} \\
&\cong \underbrace{(\text{nothing in here})}_{\text{trivial summands}} \oplus \underbrace{\frac{R}{(d_{k+1})} \oplus \dots \oplus \frac{R}{(d_r)}}_{\text{torsion summands}} \oplus \underbrace{R \oplus \dots \oplus R}_{f := m - r \text{ times}}
\end{aligned}$$

The matrix calculation for invariant factor decompositions essentially comes from re-stating the Invariant Factor Theorem for Submodules (Theorem 4.1.3) in terms of matrices. We state this in more detail below.

Proposition 4.1.7. Let R be a PID. Let $\varphi : N \rightarrow M$ be an R -module homomorphism between free R -modules N and M with $\text{rank}(N) = n$ and $\text{rank}(M) = m$. Then, there exists a basis $T = (\tau_1, \dots, \tau_n)$ of N , a basis $B = (\beta_1, \dots, \beta_m)$ of M , and nonzero elements $d_1, \dots, d_r \in R$ with $r = \text{rank}(\text{im}(\varphi)) \leq n$ and divisibility relation $d_1 \mid d_2 \mid \dots \mid d_r$ such that

$$\varphi(\tau_i) = \begin{cases} d_i \beta_i & \text{if } i \in \{1, \dots, r\} \\ 0 & \text{if } i \in \{r+1, \dots, n\} \end{cases}$$

That is, the matrix $[\varphi]_{B,T}$ of φ relative to T and B is given by the following block matrix

$$[\varphi]_{B,T} = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(R) \quad \text{with} \quad D_r = \text{diag}(d_1, \dots, d_r) = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & d_r \end{pmatrix}$$

Furthermore, the elements d_1, \dots, d_r are unique by multiplication of units.

Proof. We provide an outline of the proof given for [AW92, Proposition 4.3.20] below.

Since $\text{im}(\varphi)$ is a submodule of M , the Invariant Factor Theorem for Submodules (Theorem 4.1.3) applies. Let $B = (\beta_1, \dots, \beta_m)$ be the basis of M and $\{d_1, \dots, d_r\}$ with $r \leq m$ be the set of nonzero elements given by Theorem 4.1.3. Then, $\{d_1\beta_1, \dots, d_r\beta_r\}$ is a basis for $\text{im}(\varphi)$ and the divisibility relation $d_1 \mid \dots \mid d_r$ is satisfied. Note that d_1, \dots, d_r are also unique up to multiplication by units.

For each $i \in \{1, \dots, r\}$, choose $\tau_i \in N$ such that $\varphi(\tau_i) = d_i\beta_i$. Note that $N \cong \ker(\varphi) \oplus (N / \ker(\varphi))$ and that $\{\tau_1, \dots, \tau_r\}$ is a basis of $N / \ker(\varphi)$. Since $\ker(\varphi)$ is a free submodule of N with $n = \text{rank}(\ker \varphi) + r$, there exists a basis $\{\tau_{r+1}, \dots, \tau_n\}$ of $\ker(\varphi)$. One then proves that $\{\tau_1, \dots, \tau_r, \tau_{r+1}, \dots, \tau_n\}$ is a basis of N by showing linear independence. Then, by construction, $\varphi(\tau_i) = d_i\beta_i$ if $i \in \{1, \dots, r\}$ and $\varphi(\tau_i) = 0$ otherwise.

The matrix representation follows from the definition of matrices of homomorphisms. For reference, see Definition A3.8 in Appendix A3. Let $i \in \{1, \dots, n\}$ be a column index. By definition,

$\text{col}_i([\varphi]_{B,T}) = [\varphi(\tau_i)]_B$, i.e. the coordinate vector of $\varphi(\tau_i)$ relative to B . If $i \in \{1, \dots, r\}$:

$$[\varphi]_{B,T}(j, i) = [\varphi(\tau_i)]_B(j) = \begin{cases} d_i & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

If $i \in \{r+1, \dots, n\}$, then $\text{col}_i([\varphi]_{B,T}) = [\varphi(\tau_i)]_B$ would be the zero column since $\varphi(\tau_i) = 0$. \blacksquare

The proposition above is then used to prove the existence of Smith Normal Decompositions of matrices, defined below. Note that this definition is taken from from [AW92, Section 5.3].

Definition 4.1.8. The **Smith Normal Decomposition (SND)** of a matrix $A \in M_{m,n}(R)$ is a triple (U, D, V) of matrices $U \in \text{GL}(m, R)$, $V \in \text{GL}(n, R)$, $D \in M_{m,n}(R)$ such that

$$U^{-1}AV = D = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad D_r = \text{diag}(d_1, \dots, d_r)$$

with $r = \text{rank}(A)$ and nonzero elements $d_1, d_2, \dots, d_r \in R$ that satisfy divisibility relation $d_1 \mid d_2 \mid \dots \mid d_r$. We call $D \in M_{m,n}(R)$ the **Smith Normal Form (SNF)** of $A \in M_{m,n}(R)$.

We do warn readers that, in most references, the matrix $U \in \text{GL}(m, R)$ in the Smith Normal Decomposition (U, D, V) of $A \in M_{m,n}(R)$ is defined such that $UAV = D$, as opposed to our definition where U corresponds to the factorization $U^{-1}AV = D$. We have decided to change the definition for $U \in \text{GL}(m, R)$ so that the factorization $U^{-1}AV = D$ is more compatible with our application, i.e. we should interpret $U^{-1}AV = D$ as corresponding to some R -module homomorphism equipped with change of bases on the domain and codomain, as we will see later in this section in Proposition 4.1.11.

Next, we state the result for the existence and uniqueness of Smith Normal Decompositions. The main idea here involves creating an R -module homomorphism from the given matrix and using Proposition 4.1.7. Note that definitions for elementary matrices over R , including elementary dilations, are given in Definition A2.5.

Theorem 4.1.9. Let $A \in M_{m,n}(R)$ be a matrix over a PID R . Then, a Smith Normal Decomposition (U, D, V) of A exists and the Smith Normal Form D of A is unique up to elementary dilations over R , i.e. the diagonal elements of D are unique up to multiplication by units in R .

Proof. We refer to the proof in [AW92, Theorem 5.3.1] and provide an outline below.

Let $\mathcal{S} = (\sigma_1, \dots, \sigma_n)$ denote the standard ordered basis on R^n and let $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$ be that on R^m . Define the R -module homomorphism $\varphi : R^n \rightarrow R^m$ such that $[\varphi]_{\mathcal{A}, \mathcal{S}} = A$, i.e. A is the matrix of φ relative to \mathcal{A} and \mathcal{S} . By Proposition 4.1.7, there exists a basis $\mathcal{T} = (\tau_1, \dots, \tau_n)$ of N , a basis $\mathcal{B} = (\beta_1, \dots, \beta_m)$ of M , and nonzero elements $d_1, \dots, d_r \in R$ with $r = \text{rank}(A) \leq n$ such that

$$[\varphi]_{\mathcal{B}, \mathcal{T}} = \begin{pmatrix} D_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(R) \quad \text{with} \quad D_r = \text{diag}(d_1, \dots, d_r)$$

Let $[\text{id}_M]_{\mathcal{B}, \mathcal{A}} \in \text{GL}(m, R)$ be the matrix of the identity map $\text{id}_M : M \rightarrow M$ relative to \mathcal{A} and \mathcal{B} . Note that $[\text{id}_M]_{\mathcal{B}, \mathcal{A}}$ is invertible with inverse $[\text{id}_M]_{\mathcal{A}, \mathcal{B}}$. Similarly, let $[\text{id}_N]_{\mathcal{T}, \mathcal{S}} \in \text{GL}(n, R)$ be the matrix of the identity map $\text{id}_N : N \rightarrow N$ relative to \mathcal{T} and \mathcal{S} . Then, we can relate $[\varphi]_{\mathcal{A}, \mathcal{S}}$ and $[\varphi]_{\mathcal{B}, \mathcal{T}}$ by the following matrix equation:

$$[\text{id}_M]_{\mathcal{B}, \mathcal{A}} [\varphi]_{\mathcal{A}, \mathcal{S}} [\text{id}_N]_{\mathcal{S}, \mathcal{T}} = [\varphi]_{\mathcal{B}, \mathcal{T}}$$

Let $U \in \text{GL}(m, R)$, $D \in \text{M}_{m,n}(R)$, and $V \in \text{GL}(n, R)$ be given by

$$U = [\text{id}_M]_{\mathcal{A}, \mathcal{B}} \quad D = [\varphi]_{\mathcal{B}, \mathcal{T}} \quad V = [\text{id}_N]_{\mathcal{S}, \mathcal{T}}$$

Observe that the divisibility relation on the entries d_1, \dots, d_r is satisfied by Proposition 4.1.7. We can confirm that the matrix factorization is correct by the following calculation:

$$U^{-1}AV = \left([\text{id}_M]_{\mathcal{A}, \mathcal{B}}\right)^{-1} \left([\varphi]_{\mathcal{A}, \mathcal{S}}\right) \left([\text{id}_N]_{\mathcal{S}, \mathcal{T}}\right) = [\text{id}_M]_{\mathcal{B}, \mathcal{A}} [\varphi]_{\mathcal{A}, \mathcal{S}} [\text{id}_N]_{\mathcal{S}, \mathcal{T}} = [\varphi]_{\mathcal{B}, \mathcal{T}} = D$$

Since the elements d_1, \dots, d_r satisfy the divisibility relation $d_1 \mid \dots \mid d_r$ and $U^{-1}AV = D$, (U, D, V) is an SND of A .

Note that the uniqueness of the Smith Normal Form D of A (up to multiplication by elementary dilations over R) is implied by the uniqueness of the invariant factors up to multiplication by units, as stated in the Invariant Factor Theorem for Submodules (Theorem 4.1.3). ■

Below, we provide an example of two SNDs of a matrix over \mathbb{Z} . Note that for both SNDs, the Smith Normal Forms on both SNDs only differ by an elementary dilation, e.g. multiplication by (-1) since (-1) is a unit in \mathbb{Z} , but the matrices U and V on an SND (U, D, V) are not generally unique.

Example 4.1.10. Let $A \in \text{M}_{4,3}(\mathbb{Z})$ be given as follows:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 4 \end{pmatrix}$$

By Mathematica (a software system), A admits the following Smith Normal Decomposition (U_1, D_1, V_1) where $(U_1)^{-1}AV_1 = D_1$ and the matrices $U_1 \in \text{GL}(4, \mathbb{Z})$, $D_1 = \text{M}_{4,3}(\mathbb{Z})$, $V_1 \in \text{GL}(3, \mathbb{Z})$ are given by

$$U_1 = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 7 & 3 & 2 & 1 \\ 1 & 4 & 0 & 0 \end{pmatrix} \quad D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad V_1 = \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

We can confirm this by doing the following calculation:

$$(U_1)^{-1}AV_1 = \begin{pmatrix} 4 & 0 & 0 & -3 \\ -1 & 0 & 0 & 1 \\ -13 & 1 & 0 & 9 \\ 1 & -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \dots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} = D_1$$

Another Smith Normal Decomposition of A is (U_2, D_2, V_2) with $U_2 \in \text{GL}(4, \mathbb{Z})$, $D_2 = \text{M}_{4,3}(\mathbb{Z})$, and $V_2 \in \text{GL}(3, \mathbb{Z})$ given as follows:

$$U_2 = \begin{pmatrix} 6 & 4 & -1 & 0 \\ 15 & 13 & -3 & 0 \\ 24 & 22 & -5 & 1 \\ 7 & 4 & -1 & 0 \end{pmatrix} \quad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix} \quad V_2 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

As with (U_1, D_1, V_1) , we can confirm that (U_2, D_2, V_2) is a valid factorization by doing the following calculation:

$$(U_2)^{-1}A(V_2) = \begin{pmatrix} -1 & 0 & 0 & 1 \\ -6 & 1 & 0 & 3 \\ -31 & 4 & 0 & 18 \\ 1 & -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{pmatrix} = D_2$$

Observe that the diagonal elements $D_2(1, 1) = 1, D_2(2, 2) = 1, D_2(3, 3) = -3$ of D_2 obey the divisibility rule where d_1 divides d_2 and d_2 divides d_3 . Also, the first two diagonal elements of D_1 and D_2 match and for the 3rd diagonal element, we have $D_2(3, 3) = -3 = (-1)D_1(3, 3) = (-1)(3)$.

As implied in the proof of Theorem 4.1.9, given a homomorphism $\varphi : N \rightarrow M$, the nonzero diagonal elements d_1, \dots, d_r of the Smith Normal Form D of the matrix $[\varphi]_{A,S}$ are exactly the nonzero elements guaranteed by the Invariant Factor Theorem on Submodules (Theorem 4.1.3). We believe this is partly why the nonzero diagonal elements of the Smith Normal Form of a matrix are sometimes called the *invariant factors* of the matrix $[\varphi]_{A,S}$. To avoid confusion between the invariant factors of modules (which cannot be invertible elements) and that of matrices (which can be invertible elements), we will only use invariant factors in the context of modules, i.e. the nonzero elements of the SNF of matrices are not called invariant factors in this paper.

This means that the problem of finding invariant factor decompositions can now be expressed as a matrix factorization problem, i.e. that of finding an SND. Below, we describe how we should interpret an SND of $[\varphi]_{A,S}$, as denoted above, in order to apply Proposition 4.1.5 to calculate invariant factor decompositions.

Proposition 4.1.11. Let $\varphi : N \rightarrow M$ be a module homomorphism between free R -modules N and M with ordered bases $S = (\sigma_1, \dots, \sigma_n)$ and $A = (\alpha_1, \dots, \alpha_m)$ respectively. Let (U, D, V) be an SND of $[\varphi]_{A,S} \in M_{m,n}(R)$ and let $d_i \in R$ be given by $d_i = D(i, i) \neq 0$ for $i \in \{1, \dots, r\}$ with $r = \text{rank}([\varphi]_{A,S})$.

Then, $V \in \text{GL}(n, R)$ determines a basis $T = (\tau_1, \dots, \tau_n)$ of N given by $[\tau_i]_S = \text{col}_i(V)$ for $i \in \{1, \dots, n\}$. Similarly, $U \in \text{GL}(m, R)$ determines a basis $B = (\beta_1, \dots, \beta_m)$ of M given by $[\beta_j]_A = \text{col}_j(U)$ for $j \in \{1, \dots, m\}$. Furthermore, the divisibility relation $d_1 \mid d_2 \mid \cdots \mid d_r$ is satisfied and $D = [\varphi]_{B,T}$, i.e.

$$\varphi(\tau_i) = \begin{cases} d_i \beta_i & \text{if } i \in \{1, \dots, r\} \\ 0 & \text{if } i \in \{r+1, \dots, n\} \end{cases}$$

Note that $\{\varphi(\tau_1), \dots, \varphi(\tau_r)\} = \{d_1 \beta_1, \dots, d_r \beta_r\}$ forms a basis of $\text{im}(\varphi) \subseteq M$ and $\{\tau_{r+1}, \dots, \tau_n\}$ determines a basis for $\ker(\varphi) \subseteq N$.

Proof. Let $e_i^{[n]} \in M_{n,1}(R)$ denote the i^{th} standard basis (column) vector of R^n and let $e_j^{[m]} \in M_{m,1}(R)$ be the j^{th} standard basis (column) vector of R^m .

Since $V \in \text{GL}(n, R)$ and $U \in \text{GL}(m, R)$ are invertible matrices, they can be used as change of basis matrices. Define a basis $T = (\tau_1, \dots, \tau_n)$ of N by $V = [\text{id}_N]_{S,T}$. Note that $[\text{id}_N]_{T,S} = V^{-1}$. Similarly, define a basis $B = (\beta_1, \dots, \beta_m)$ of M by $[\text{id}_M]_{A,B} = U$. Note that $[\text{id}_M]_{B,A} = U^{-1}$. Then, the coordinate vectors of each τ_i and β_j relative to S and T respectively are given below:

$$\begin{aligned} [\tau_i]_S &= [\text{id}_N]_{S,T} [\tau_i]_T = V e_i^{[n]} = \text{col}_i(V) \in M_{n,1}(R) \quad \text{for } i \in \{1, \dots, n\} \\ [\beta_j]_A &= [\text{id}_M]_{A,B} [\beta_j]_B = ([\text{id}_M]_{B,A})^{-1} e_j^{[m]} = U e_j^{[m]} = \text{col}_j(U) \in M_{m,1}(R) \quad \text{for } j \in \{1, \dots, m\} \end{aligned}$$

Since (U, D, V) is an SND of $[\varphi]_{A,S}$, $U^{-1}[\varphi]_{A,S}V = D$. Then, $[\varphi]_{B,T} = D$ by the following calculation:

$$D = U^{-1}[\varphi]_{A,S}V = ([\text{id}_M]_{A,B})^{-1}([\varphi]_{A,S})([\text{id}_N]_{S,T}) = [\text{id}_M]_{B,A}[\varphi]_{A,S}[\text{id}_N]_{S,T} = [\varphi]_{B,T}$$

Then, for each $i \in \{1, \dots, n\}$, the coordinate vector of $\varphi(\tau_i)$ relative to B is as follows:

$$[\varphi(\tau_i)]_B = [\varphi]_{B,T}[\tau_i]_T = De_i^{[n]} = \text{col}_i(D) = \begin{cases} d_i e_i^{[m]} & \text{if } i \in \{1, \dots, r\} \\ 0 & \text{if } i \in \{r+1, \dots, n\} \end{cases}$$

Therefore, $\varphi(\tau_i) = d_i \beta_i$ if $i \in \{1, \dots, r\}$ and $\varphi(\tau_i) = 0$ otherwise. Observe that $\{\varphi(\tau_1), \dots, \varphi(\tau_r)\} = \{d_1 \beta_1, \dots, d_r \beta_r\}$ is a basis of $\text{im}(\varphi)$ and $\{\tau_{r+1}, \dots, \tau_n\}$ is a basis for $\ker(\varphi)$. The divisibility relation $d_1 \mid d_2 \mid \dots \mid d_r$ is satisfied by definition of SND. ■

Since an SND (U, D, V) of the matrix $[\varphi]_{A,S}$ of a homomorphism $\varphi : N \rightarrow M$, as denoted above, gives us the results of the Invariant Factor Theorem for Submodules (Theorem 4.1.3), i.e. we can identify the basis $B = (\beta_1, \dots, \beta_m)$ of M and the basis $\{d_1 \beta_1, \dots, d_r \beta_r\}$ of $\text{im}(\varphi)$, we can use Proposition 4.1.5 to calculate invariant factor decompositions.

Observe that given a finite presentation $\varphi : F_S \rightarrow F_G$, $\pi : F_G \rightarrow M$ of an R -module M , the basis $B = (\beta_1, \dots, \beta_m)$ and the nonzero elements $d_1, \dots, d_r \in R$ given by the above proposition (Proposition 4.1.11) are exactly the information required in the hypothesis of Proposition 4.1.5. In particular, if given ordered bases $S = (\sigma_1, \dots, \sigma_n)$ of F_S and $A = (\alpha_1, \dots, \alpha_m)$, an SND (U, D, V) of $[\varphi]_{A,S}$ determines the basis $B = (\beta_1, \dots, \beta_m)$ by $[\beta_j]_A = \text{col}_j(U)$ and the nonzero elements $d_j = D(j, j)$ and yields the following isomorphism by application of Proposition 4.1.5:

$$\begin{aligned} M &\cong \frac{F_G}{\text{im}(F_S)} \cong \frac{R\langle \alpha_1, \dots, \alpha_m \rangle}{\text{im}(F_S)} \\ &\cong \frac{R\langle \beta_1, \dots, \beta_m \rangle}{R\langle d_1 \beta_1, \dots, d_r \beta_r \rangle} && \text{by Proposition 4.1.11} \\ &\cong \left(\frac{R\langle \beta_1 \rangle}{R\langle d_1 \beta_1 \rangle} \right) \oplus \dots \oplus \left(\frac{R\langle \beta_r \rangle}{R\langle d_r \beta_r \rangle} \right) \oplus R\langle \beta_{r+1} \rangle \oplus \dots \oplus R\langle \beta_m \rangle && \text{by Proposition 4.1.5} \\ &\cong \frac{R}{(d_1)} \oplus \dots \oplus \frac{R}{(d_k)} \oplus R \oplus \dots \oplus R && \text{by Lemma 4.1.6} \end{aligned}$$

where $k \in \{1, \dots, r\}$ is chosen such that d_{k+1}, \dots, d_r are all the invertible elements of the set $\{d_1, \dots, d_r\}$. We provide an example of this calculation below, using a homomorphism between \mathbb{Z} -modules.

Example 4.1.12. Let $N = \mathbb{Z}\langle a, b, c, d \rangle$ and $M = \langle x, y, z \rangle$. Define the module homomorphism $\varphi : N \rightarrow M$ by

$$a \mapsto x - y, \quad b \mapsto y + z, \quad c \mapsto 2x, \quad d \mapsto z.$$

Then, the matrix of φ relative to the bases $A = (a, b, c, d)$ of N and $X = (x, y, z)$ of M is as follows:

$$[\varphi]_{X,A} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} \\ \begin{matrix} x \\ y \\ z \end{matrix} & \begin{pmatrix} 1 & 0 & 2 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix}$$

Using Mathematica, the triple (U, D, V) of matrices below corresponds to an SND of $[\varphi]_{X,A}$:

$$U = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad V = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

The validity of this factorization can be confirmed by performing the following calculation:

$$U^{-1}[\varphi]_{X,A}V = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} = \dots = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = D$$

From $V \in \text{GL}(4, \mathbb{Z})$, we determine the basis $\mathcal{A} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ of N by $[\alpha_i]_A = \text{col}_i(V)$ as follows:

$$\begin{array}{llll} [\alpha_1]_A = \text{col}_1(V) & [\alpha_2]_A = \text{col}_2(V) & [\alpha_3]_A = \text{col}_3(V) & [\alpha_4]_A = \text{col}_4(V) \\ \alpha_1 = a & \alpha_2 = b & \alpha_3 = -a - b + c + d & \alpha_4 = -2a - 2b + c + 2d \end{array}$$

From $U \in \text{GL}(3, \mathbb{Z})$, we get the basis $\mathcal{Y} = (\gamma_1, \gamma_2, \gamma_3)$ of M by $[\gamma_i] = \text{col}_i(U)$ as follows:

$$\begin{array}{lll} [\gamma_1]_X = \text{col}_1(U) & [\gamma_2]_X = \text{col}_2(U) & [\gamma_3]_X = \text{col}_3(U) \\ \gamma_1 = x - y & \gamma_2 = y + z & \gamma_3 = x \end{array}$$

We can confirm that $D = [\varphi]_{\mathcal{Y}, \mathcal{A}}$ by the following calculations:

$$\begin{array}{ll} \varphi(\alpha_1) = \varphi(a) = x - y & = \gamma_1 \\ \varphi(\alpha_2) = \varphi(b) = y + z & = \gamma_2 \\ \varphi(\alpha_3) = \varphi(-a - b + c + d) = -(x - y) - (y + z) + 2x + z = x & = \gamma_3 \\ \varphi(\alpha_4) = \varphi(-2a - 2b + c + 2d) = -2(x - y) - 2(y + z) + 2x + 2(z) & = 0 \end{array}$$

Then, we can express the kernel and image of φ as follows:

$$\begin{array}{ll} \ker(\varphi) = \mathbb{Z}\langle \alpha_4 \rangle = \mathbb{Z}\langle -2a - 2b + c + 2d \rangle & \text{with } \text{rank}(\ker \varphi) = 1 \\ \text{im}(\varphi) = \mathbb{Z}\langle D(1, 1)\gamma_1, D(2, 2)\gamma_2, D(3, 3)\gamma_3 \rangle = \mathbb{Z}\langle x - y, y + z, x \rangle & \text{with } \text{rank}(\text{im } \varphi) = 3 \end{array}$$

While we have established that Smith Normal Decompositions always exist for arbitrary PIDs, we have not yet discussed an algorithm for their calculation. However, if we restrict the PID R to be a Euclidean domain, we can use the following result.

Proposition 4.1.13. Any invertible matrix over a Euclidean domain R can be expressed as a finite product of elementary matrices. That is, $\text{GL}(n, R)$ is generated by the elementary matrices over R of degree n .

Remarks. (1) For a proof, see [AW92, Theorem 5.2.10]. Note that definitions for elementary matrices (i.e. elementary dilations, elementary permutations, and elementary transvections) on arbitrary Euclidean domains are given in Definition A2.5.

(2) The paper *Products of Elementary Matrices and Non-Euclidean Principal Ideal Domains* [CZZ18] conjectures that this is not generally true if the ring R is a PID but not a Euclidean domain.

Since elementary matrices correspond to row and column operations (as discussed in Proposition A2.7 for row operations and in Proposition A2.8 for column operations), it is possible to calculate SNDs by doing matrix reduction assuming the matrices are over Euclidean domains. We will not present a general algorithm for calculating SNDs of matrices over Euclidean domains in this paper and we refer to [AW92, Remark 5.3.4] for those interested. Below, we provide an example of this reduction process in a matrix over \mathbb{Z} .

Example 4.1.14. Let $A \in M_{3,4}(\mathbb{Z})$ be the (3×4) -matrix given by

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

We can calculate an SND of A by the following row and column reduction operation, represented by the sequence (A_n) of matrices. Let $A_0 = A$. The elements of A_n highlighted in **red** are elements that are to be eliminated using an elementary transvection, in **blue** is the pivot used for said transvection, and in **orange** the pivot multiplier.

$$\begin{aligned} A_1 &:= A_0 E_{\text{add}}^{[4]}(1, 2; -2) = \begin{pmatrix} \text{blue } 1 & \text{red } 2 & 0 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & \text{orange } -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ A_2 &:= A_1 E_{\text{add}}^{[4]}(1, 4; -1) = \begin{pmatrix} \text{blue } 1 & 0 & 0 & \text{red } 1 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \text{orange } -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ A_3 &:= A_2 E_{\text{add}}^{[4]}(2, 4; -1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \text{blue } 3 & 0 & \text{red } 3 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \text{orange } -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ A_4 &:= A_3 E_{\text{add}}^{[4]}(3, 4; -1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & \text{blue } 1 & \text{red } 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \text{orange } -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

The matrix A_4 is almost is Smith Normal Form. We need to apply row and column permutations. Highlighted in **green** are the elements that are switched in each A_n .

$$\begin{aligned} A_5 &:= E_{\text{swap}}^{[3]}(2, 3) A_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \text{green } 3 & 0 & 0 \\ 0 & 0 & \text{green } 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix} \\ A_6 &:= A_5 E_{\text{swap}}^{[4]}(2, 3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \text{green } 3 & 0 \\ 0 & \text{green } 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} =: D \end{aligned}$$

The matrices $U \in GL(3, \mathbb{Z})$ and $V \in GL(4, \mathbb{Z})$ are given by:

$$\begin{aligned} V &:= E_{\text{add}}^{[4]}(1, 2; -2) E_{\text{add}}^{[4]}(1, 4; -1) E_{\text{add}}^{[4]}(2, 4; -1) E_{\text{add}}^{[4]}(2, 4; -1) E_{\text{swap}}^{[4]}(2, 3) = \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ U &:= \left(E_{\text{swap}}^{[3]}(2, 3) \right)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

Observe that U and V are invertible since they are both a finite product of elementary matrices, which are also invertible. Therefore, the triple (U, D, V) represent an SND of A .

The method described in [AW92, Remark 5.3.4] mostly involves the elimination of specific entries in a matrix, like in the case of matrices over \mathbb{R} . However, since nonzero elements of an arbitrary Euclidean domain are not generally invertible, the entries used to eliminate other entries along the same row or column (usually called a *pivot*) have to be chosen with care, i.e. we cannot use any nonzero element as in the case of matrices over \mathbb{R} . The main difference lies in the family of elementary dilations over R which, as described in Definition A2.5, are defined on the group of units R^\times of R . In the case of $R = \mathbb{R}$, any nonzero element of \mathbb{R} can be used to construct an elementary dilation. So, any nonzero element $a \in \mathbb{R}$ of a matrix over \mathbb{R} can be used to eliminate any entry in the same row or column since a can be reduced by 1 using an elementary dilation by $a^{-1} = \frac{1}{a}$. In contrast, the only units of \mathbb{Z} are $\{-1, 1\}$ and an element such as $2 \in \mathbb{Z}$ cannot be reduced to 1.

Consequently, there is no direct analog of row or column echelon form for matrices over \mathbb{Z} and over $\mathbb{F}[x]$. In particular, even if a matrix over \mathbb{Z} or $\mathbb{F}[x]$ may appear to be in row or column echelon form (as we conventionally define those), the pivots of that matrix generally do not correspond to the diagonal elements of its Smith Normal Form. Note that this applies more generally for Euclidean domains that are not fields. We provide an example of this problem below.

Example 4.1.15. Let $A \in M_{3,4}(\mathbb{Z})$ be the matrix given by

$$A = \begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 7 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Note that the matrix A is in row echelon form with pivots 2, 7, 3 (in the row echelon sense) above highlighted in red. An SND (U, D, V) of A , calculated by Mathematica, is given below:

$$U = I_3 \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} \quad V = \begin{pmatrix} -10 & -6 & 0 & -21 \\ -2 & -1 & 0 & -4 \\ 7 & 4 & 0 & 14 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Observe that the nonzero diagonal elements of D are 1, 1, 3, not 2, 7, 3. Note that this is because we cannot do dilation operations on A as a matrix over \mathbb{Z} like the one below:

$$A' = E_{\text{dilate}}^{[3]} \left(1, \frac{1}{2} \right) A = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 7 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{3}{2} & 0 \\ 0 & 7 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

Note that $\frac{1}{2}$ is not an element of \mathbb{Z} and therefore, $E_{\text{dilate}}^{[3]} \left(1, \frac{1}{2} \right)$ is not an elementary matrix over \mathbb{Z} . However, if we consider A as a matrix over \mathbb{R} , we can calculate an SND $(U_{\mathbb{R}}, D_{\mathbb{R}}, V_{\mathbb{R}})$ of A as follows:

$$U_{\mathbb{R}} = I_3 \quad D_{\mathbb{R}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad V_{\mathbb{R}} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{3}{2} \\ 0 & \frac{1}{7} & 0 & -\frac{2}{7} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

Observe that $V_{\mathbb{R}}$ is not a matrix over \mathbb{Z} . As a sidenote, since all nonzero field elements are invertible, any nonzero diagonal element of the Smith Normal Form of any matrix over a field \mathbb{F} can always be made into $1 \in \mathbb{F}$ by multiplication of an appropriate elementary dilation over \mathbb{F} .

Section 4.2. Matrix Calculation of Homology of Chain Complexes

Let $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ be a chain complex of free modules C_n over a PID R with differentials $\partial_n : C_n \rightarrow C_{n+1}$ and assume that each C_n is of finite rank. By the assumption on $\text{rank}(C_n)$, the n^{th} homology $H_n(C_*)$ of C_* is finitely generated and therefore admits an invariant factor decomposition by the Structure Theorem (Theorem 4.1.1). In this section, we discuss how SNDs of the matrices $[\partial_{n+1}]$ and $[\partial_n]$ of $\partial_{n+1} : C_{n+1} \rightarrow C_n$ and $\partial_n : C_n \rightarrow C_{n-1}$ respectively (relative to some chosen ordered bases) can be used to calculate these decompositions for $H_n(C_*)$.

The method we discussed for calculating invariant factor decompositions in Section 4.1 starts with a finite presentation for the R -module in question. The following proposition identifies a finite presentation for $H_n(C_*)$ that arises naturally from the definition $H_n(C_*) := \ker(\partial_n) / \text{im}(\partial_{n+1})$.

Proposition 4.2.1. Let $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ be a chain complex of free modules C_n over a PID R and differentials $\partial_n : C_n \rightarrow C_{n-1}$. Then, the n^{th} homology group $H_n(C_*)$ of C_* admits the following presentation

$$C_{n+1} \xrightarrow{\partial_{n+1}} \ker(\partial_n) \xrightarrow{\pi} H_n(C_*) \longrightarrow 0$$

with π being the canonical projection onto $H_n(C_*)$.

Proof. Note that since C_* is a chain complex, i.e. $\text{im}(\partial_{n+1}) \subseteq \ker(\partial_n)$, the restriction of the codomain of $\partial_{n+1} : C_{n+1} \rightarrow C_n$ onto $\ker(\partial_n)$ is well-defined. Since $\ker(\partial_n)$ is a submodule of a free module C_n over a PID R , $\ker(\partial_n)$ is a free R -module for all $n \in \mathbb{Z}$ [DF03, Theorem 12.1]. By definition of homology of chain complexes, $H_n(C_*) = \ker(\partial_n) / \text{im}(\partial_{n+1})$. Then, $\text{im}(\partial_{n+1}) = \ker(\pi)$, and $\text{im}(\pi) = \ker(H_n(C_*) \rightarrow 0) = H_n(C_*)$. Therefore, Equation (1) is an exact sequence of free R -modules and is a presentation of $H_n(C_*)$. ■

We want to emphasize that, in the proposition above, the codomain of the $(n+1)^{\text{th}}$ differential ∂_{n+1} is restricted to $\ker(\partial_n)$. For clarity, we write $\partial_{n+1}^{\ker} : C_{n+1} \rightarrow \ker(\partial_n)$ to refer to $\partial_{n+1} : C_{n+1} \rightarrow C_n$ with this codomain restriction. Since it is rarely the case that $\ker(\partial_n) = C_n$, the matrices $[\partial_{n+1}^{\ker}]$ (i.e. with the codomain restriction) and $[\partial_{n+1}]$ (i.e. without the codomain restriction) are generally not the same. For example, since $\text{rank}(\ker \partial_n) \leq \text{rank}(C_n)$, the number of rows of $[\partial_{n+1}^{\ker}]$ is less than or equal to that of $[\partial_n]$. It may also be the case that the basis of $\ker(\partial_n)$ used for $[\partial_{n+1}^{\ker}]$ is not a subset of the basis of C_n used for $[\partial_{n+1}]$.

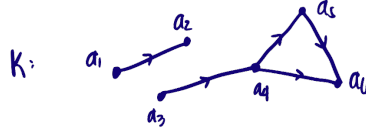
This means that the results of Section 4.1 only immediately apply to the matrix $[\partial_{n+1}^{\ker}]$. More specifically, if we were to use an SND of $[\partial_{n+1}]$, then said calculation would correspond to the following sequence:

$$C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\pi} H_n(C_*) \longrightarrow 0$$

Note that this sequence is generally not exact and, therefore, is not a presentation of $H_n(C_*)$. Consequently, in order for $H_n(C_*)$ to be computable from $[\partial_{n+1}]$, we need to determine how SNDs of $[\partial_{n+1}]$ correspond to those of $[\partial_{n+1}^{\ker}]$.

Before we discuss this further, we provide an example calculation of the 0^{th} homology group $H_0(K) = H_0(K; \mathbb{Z})$ of a simplicial complex K below. Note that, by definition, the 0^{th} boundary map $\partial_0 : C_0(K) \rightarrow 0$ has trivial codomain. Therefore, $\ker(\partial_0) = C_0(K) = C_0(K; \mathbb{Z})$ and the method of calculation by Section 4.1 is directly applicable to the matrix $[\partial_1]$. Note that an orientation on K induces a standard ordered basis on the chain groups of K .

Example 4.2.2. Define the simplicial complex K as illustrated below and equip K with the orientation given by $\text{Vert}(K) = (a_1, a_2, a_3, a_4, a_5, a_6)$.



The following sequence is a finite presentation for the 0th homology group $H_0(K) = H_0(K; \mathbb{Z})$ of K with coefficients in \mathbb{Z} :

$$C_1 \xrightarrow{\partial_1} C_0 \xrightarrow[\ker(\partial_0)]{\pi} H_0(K) \longrightarrow 0$$

Therefore, we can calculate $H_0(K)$ using an SND (U_1, D_1, V_1) of the matrix $[\partial_1]$ of $\partial_1 : C_1(K) \rightarrow C_0(K)$ relative to some ordered bases. The orientation on K by $\text{Vert}(K) = (a_1, a_2, a_3, a_4, a_5, a_6)$ determines the following ordered bases of $C_0(K)$ and $C_1(K)$:

0-simplices: $(a_1, a_2, a_3, a_4, a_5, a_6)$ and 1-simplices: $(a_1a_2, a_3a_4, a_4a_5, a_4a_6, a_5a_6)$

Then, the matrix $[\partial_1] \in M_{6,5}(\mathbb{Z})$ relative to the standard ordered bases can be calculated as follows:

$$[\partial_1] = \begin{matrix} & a_1a_2 & a_3a_4 & a_4a_5 & a_4a_6 & a_5a_6 \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{matrix} & \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$

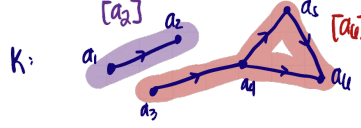
Given below is an SND (U_1, D_1, V_1) of $[\partial_1]$ calculated using Mathematica. Since $C_0 = \ker(\partial_0)$, all six columns of U_1 correspond to a basis of $\ker(\partial_0)$. Highlighted in green are the columns of $U_1 \in \text{GL}(6, \mathbb{Z})$ corresponding to torsion summands of $H_0(K)$ and their corresponding diagonal elements in D_1 and in blue are the columns of U_1 corresponding to free summands of $H_0(K)$.

$$U_1 = \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{matrix} \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \quad D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad V_1 = \begin{matrix} a_1a_2 \\ a_3a_4 \\ a_4a_5 \\ a_4a_6 \\ a_5a_6 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then, by Proposition 4.1.5, we can calculate $H_0(K)$ as follows:

$$\begin{aligned} H_0(K) &= \frac{C_0(K)}{\text{im}(\partial_1)} \cong \frac{\mathbb{Z}\langle a_2 - a_1, a_4 - a_5, a_5 - a_4, a_6 - a_4, a_2, a_6 \rangle}{\mathbb{Z}\langle (1)(a_2 - a_1), (1)(a_4 - a_3), (1)(a_5 - a_4), (1)(a_6 - a_4) \rangle} \\ &= \frac{\mathbb{Z}\langle a_2 - a_1 \rangle}{\mathbb{Z}\langle a_2 - a_1 \rangle} \oplus \frac{\mathbb{Z}\langle a_4 - a_3 \rangle}{\mathbb{Z}\langle a_4 - a_3 \rangle} \oplus \frac{\mathbb{Z}\langle a_5 - a_4 \rangle}{\mathbb{Z}\langle a_5 - a_4 \rangle} \oplus \frac{\mathbb{Z}\langle a_6 - a_4 \rangle}{\mathbb{Z}\langle a_6 - a_4 \rangle} \oplus \mathbb{Z}\langle a_2 \rangle \oplus \mathbb{Z}\langle a_6 \rangle \\ &= \mathbb{Z}\langle a_2 \rangle \oplus \mathbb{Z}\langle a_6 \rangle \cong \mathbb{Z}^2 \end{aligned}$$

Observe that this result matches with the expected interpretation of the 0th homology group of K , wherein K has two path components represented by the homology classes $[a_2]$ and $[a_6]$ as illustrated below:



We now consider $H_n(C_*)$ and the matrix $[\partial_{n+1}^{\ker}]$ corresponding to the presentation of $H_n(C_*)$ by Proposition 4.2.1, i.e. ∂_{n+1}^{\ker} refers to ∂_{n+1} with codomain restricted to $\ker(\partial_n)$. Following the calculation presented in Section 4.1, an SND $(U_{n+1}^{\ker}, D_{n+1}^{\ker}, V_{n+1}^{\ker})$ of $[\partial_{n+1}^{\ker}]$ determines a basis $B = (\beta_1, \dots, \beta_{r+f})$ of $\ker(\partial_n)$ with $\text{rank}(\ker \partial_n) = r + f$ by $[\beta_j] = \text{col}_j(U_{n+1}^{\ker})$ and nonzero elements $d_1, \dots, d_r \in R$ by $d_j = D_{n+1}^{\ker}(j, j)$ such that $\{d_1\beta_1, \dots, d_r\beta_r\}$ is a basis for $\text{im}(\partial_{n+1})$.

Observe that we can express the partition of the basis B into $(\beta_1, \dots, \beta_r)$ and $(\beta_{r+1}, \dots, \beta_m)$ as a decomposition $\ker(\partial_n) = K_n^{\text{tor}} \oplus K_n^{\text{free}}$ of $\ker(\partial_n)$ where K_n^{tor} and K_n^{free} are free R -submodules given by:

$$K_n^{\text{tor}} := R\langle \beta_1, \dots, \beta_r \rangle \quad \text{and} \quad K_n^{\text{free}} := R\langle \beta_{r+1}, \dots, \beta_m \rangle$$

As the notation suggests, the submodule K_n^{tor} accounts for all elements in $\ker(\partial_n)$ that map to torsion elements or trivial in $H_n(C_*)$ and the submodule K_n^{free} accounts for the elements of $\ker(\partial_n)$ that remain free in $H_n(C_*)$. We illustrate this below with the **torsion** component of $H_n(C_*)$ highlighted in **green** and **free** component of $H_n(C_*)$ in **blue**.

$$\begin{aligned} H_n(C_*) &\cong \frac{\ker(\partial_n)}{\text{im}(\partial_{n+1})} \cong \frac{K_n^{\text{tor}} \oplus K_n^{\text{free}}}{\text{im}(\partial_{n+1})} \cong \frac{K_n^{\text{tor}}}{\text{im}(\partial_{n+1})} \oplus K_n^{\text{free}} \\ &\cong \frac{R\langle \beta_1 \rangle}{R\langle d_1\beta_1 \rangle} \oplus \dots \oplus \frac{R\langle \beta_r \rangle}{R\langle d_r\beta_r \rangle} \oplus R\langle \beta_{r+1} \rangle \oplus \dots \oplus R\langle \beta_m \rangle \end{aligned}$$

In order for $H_n(C_*)$ to be computable from $[\partial_{n+1}]$, we need to show that a basis $B = (\beta_1, \dots, \beta_m)$ of $\ker(\partial_n)$ and elements $d_1, \dots, d_r \in R$ as described above can be calculated from some SND $(U_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$ with U_{n+1} inducing a basis on C_n , not on $\ker(\partial_n)$. The existence of such an SND of $[\partial_{n+1}]$ is given by the following result, taken from [Mun93, Theorem 11.4].

Theorem 4.2.3. Let $C_* = (C_n, \partial_n)$ be a chain complex of free modules C_n over a PID R with differentials $\partial_n : C_n \rightarrow C_{n-1}$ and assume each C_n is of finite rank. For each $n \in \mathbb{N}_0$, C_n decomposes into the following direct sum of free R -modules:

$$C_n = K_n^{\text{tor}} \oplus K_n^{\text{free}} \oplus \frac{C_n}{\ker(\partial_n)} \quad \text{i}$$

where $\ker(\partial_n) = K_n^{\text{free}} \oplus K_n^{\text{tor}}$, $\text{im}(\partial_{n+1}) \subseteq K_n^{\text{tor}}$, and $K_n^{\text{tor}} / \text{im}(\partial_{n+1})$ is the torsion component of $H_n(C_*)$. Furthermore, there exists an SND $(U_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$ such that the basis $B = (\beta_1, \dots, \beta_q)$ of C_n by $[\beta_j] = \text{col}_j(U_{n+1})$ partitions into three sets of bases as follows:

$$K_n^{\text{tor}} = R\langle \beta_1, \dots, \beta_r \rangle \quad , \quad K_n^{\text{free}} = R\langle \beta_{r+1}, \dots, \beta_{r+f} \rangle \quad \text{and} \quad \frac{C_n}{\ker(\partial_n)} = R\langle \beta_{r+f+1}, \dots, \beta_q \rangle$$

Note that $\text{im}(\partial_{n+1}) = R\langle d_1\beta_1, \dots, d_r\beta_r \rangle$ with $d_1, \dots, d_r \in R$ corresponding to the nonzero entries of D_n .

Remark. For a proof, see [Mun93, Theorem 11.4] where they write W_n , V_n , and U_n to refer to K_n^{tor} , K_n^{free} , and $C_n / \ker(\partial_n)$ respectively. Note that since $K_n^{\text{tor}} / \text{im}(\partial_{n+1})$ has to be the torsion component of

$H_n(C_*)$, we have the following by definition of torsion component:

$$H_n(C_*) = \frac{\ker(\partial_n)}{\operatorname{im}(\partial_{n+1})} \cong K_n^{\text{free}} \oplus \frac{K_n^{\text{tor}}}{\operatorname{im}(\partial_{n+1})}.$$

with $K_n^{\text{free}} = R\langle\beta_{r+1}, \dots, \beta_{r+f}\rangle \cong R^f$ consists of the torsion-free elements of $H_n(C_*)$.

The existence of this decomposition of C_n into direct summands guarantees that the information we get about $\operatorname{im}(\partial_{n+1})$ from an SND of $[\partial_{n+1}]$ is compatible to that about $\ker(\partial_n)$ from an SND of $[\partial_n]$. Let $\operatorname{rank}(C_n) = q$ and let the ranks of the direct summands of C_n by Theorem 4.2.3 be given as follows:

$$\operatorname{rank}(K_n^{\text{tor}}) = r \quad ; \quad \operatorname{rank}(K_n^{\text{free}}) = f \quad ; \quad \operatorname{rank}(C_n / \ker(\partial_n)) = q$$

Let $(W_{n+1}, D_{n+1}, V_{n+1})$ be an arbitrary SND of $[\partial_{n+1}]$. The first r columns of the matrix $W_{n+1}D_{n+1}$ determines a basis $(d_1\alpha_1, \dots, d_r\alpha_r)$ of $\operatorname{im}(\partial_{n+1})$ (or equivalently, the first r columns of W_{n+1} and the first r rows of D_{n+1}). Let $A = (\alpha_1, \dots, \alpha_m)$ be the basis of C_n given by $[\alpha_j] = \operatorname{col}_j(W_{n+1})$. Part of Theorem 4.2.3 tells us that the basis A of C_n partitions into two sets: the subset $(\alpha_1, \dots, \alpha_r)$ is a basis of the direct summand K_n^{tor} of C_n and the other subset $(\alpha_{r+1}, \dots, \alpha_m)$ form a basis of the direct sum $K_n^{\text{free}} \oplus (C_n / \ker \partial_n)$, as illustrated below:

$$\begin{array}{c} \overbrace{\left(\begin{array}{c|c|c} \begin{array}{c} \vdots \\ [\alpha_1] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\alpha_r] \\ \vdots \end{array} & \begin{array}{c} \vdots \\ [\alpha_{r+1}] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\alpha_m] \\ \vdots \end{array} \end{array} \right)^{-1}}^{W_{n+1} \in \operatorname{GL}(m, R)} \quad [\partial_{n+1}]V_{n+1} = \overbrace{\begin{pmatrix} \begin{array}{c} d_1 \\ \vdots \\ d_r \end{array} & & & & \\ & \ddots & & & \\ & & d_r & & \\ & & & 0 & \cdots & 0 \\ & & & \vdots & \ddots & \vdots \\ & & & 0 & \cdots & 0 \end{pmatrix}}^{D_{n+1} \in M_{q,(-)}(R)} \end{array}$$

$\underbrace{\hspace{10em}}_{K_n^{\text{tor}}} \quad \underbrace{\hspace{10em}}_{K_n^{\text{free}} \oplus (C_n / \ker \partial_n)}$

Note that, in general, the basis $(\alpha_{r+1}, \dots, \alpha_m)$ of $K_n^{\text{free}} \oplus (C_n / \ker(\partial_n))$ generally does not partition into bases for each summand.

In contrast, let (U_n, D_n, V_n) be an arbitrary SND of $[\partial_n]$. The matrix $V_n \in \operatorname{GL}(m, R)$ determines a basis $K = (\kappa_1, \dots, \kappa_m)$ of C_n by $[\kappa_i] = \operatorname{col}_i(V_n)$ such that $(\kappa_1, \dots, \kappa_q)$ is a basis of $C_n / \ker \partial_n$ and $(\kappa_{q+1}, \dots, \kappa_m)$ is a basis of $\ker(\partial_n) = K_n^{\text{tor}} \oplus K_n^{\text{free}}$, as illustrated below.

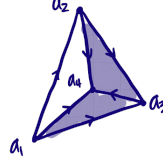
$$\begin{array}{c} \overbrace{\left(\begin{array}{c|c|c} \begin{array}{c} \vdots \\ [\kappa_1] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\kappa_q] \\ \vdots \end{array} & \begin{array}{c} \vdots \\ [\kappa_{q+1}] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\kappa_m] \\ \vdots \end{array} \end{array} \right)}^{V_n \in \operatorname{GL}(m, R)} = \overbrace{\begin{pmatrix} (*)_1 & & & & \\ & \ddots & & & \\ & & (*)_q & & \\ & & & 0 & \cdots & 0 \\ & & & \vdots & \ddots & \vdots \\ & & & 0 & \cdots & 0 \end{pmatrix}}^{D_n \in M_{(-),q}(R)} \end{array}$$

$\underbrace{\hspace{10em}}_{C_n / \ker(\partial_n)} \quad \underbrace{\hspace{10em}}_{\ker(\partial_n) = K_n^{\text{tor}} \oplus K_n^{\text{free}}}$

The problem here is that the basis $(\kappa_{q+1}, \dots, \kappa_m)$ generally does not partition into a basis of K_n^{tor} and a basis of K_n^{free} .

What Theorem 4.2.3 guarantees is the existence of an SND $(U_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$ wherein a basis of K_n^{free} can be identified, possibly in comparison with another SND of $[\partial_n]$. Then, rows of the SNF D_{n+1} of $[\partial_{n+1}]$ corresponding to basis elements of $C_n / \ker \partial_n$ can be removed to get the SNF D_{n+1}^{\ker} of $[\partial_{n+1}^{\ker}]$, the matrix of ∂_{n+1} with the codomain restriction into $\ker(\partial_n)$. We provide an example of such an SND below, in a calculation of the 1st homology group of a simplicial complex.

Example 4.2.4. Let K be the simplicial complex illustrated below with orientation $\text{Vert}(K) = (a_1, a_2, a_3, a_4)$.



The 1st homology group $H_1(K) = H_1(K; \mathbb{Z})$ of K can be calculated from a Smith Normal Decomposition of $[\partial_2]$ with $\partial_2 : C_2(K) \rightarrow C_1(K)$. The orientation on K induces a standard ordered bases on the chain groups of K . We identify these bases for $C_0(K)$, $C_1(K)$, and $C_2(K)$ below.

$$\begin{aligned} \text{four 0-simplices} & : (a_1, a_2, a_3, a_4) \\ \text{six 1-simplices} & : (a_1a_2, a_1a_3, a_1a_4, a_2a_3, a_2a_4, a_3a_4) \\ \text{two 2-simplices} & : (a_1a_3a_4, a_2a_3a_4) \end{aligned}$$

The matrices $[\partial_1]$ and $[\partial_2]$ of the boundary maps $\partial_1 : C_1(K) \rightarrow C_0(K)$ and $\partial_2 : C_2(K) \rightarrow C_1(K)$ respectively relative to these bases are given below.

$$[\partial_1] = \begin{matrix} & a_1a_2 & a_1a_3 & a_1a_4 & a_2a_3 & a_2a_4 & a_3a_4 \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad \text{and} \quad [\partial_2] = \begin{matrix} & a_1a_3a_4 & a_2a_3a_4 \\ \begin{matrix} a_1a_2 \\ a_1a_3 \\ a_1a_4 \\ a_2a_3 \\ a_2a_4 \\ a_3a_4 \end{matrix} & \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \end{matrix}$$

Given below is an SND (U_2, D_2, V_2) of $[\partial_2]$, with the columns of U_2 corresponding to K_1^{tor} (following the notation in Theorem 4.2.3) highlighted in green.

$$U_2 = \begin{matrix} & \begin{matrix} 0 & 0 & 1 & 0 & 1 & 0 \end{matrix} \\ \begin{matrix} a_1a_2 \\ a_1a_3 \\ a_1a_4 \\ a_2a_3 \\ a_2a_4 \\ a_3a_4 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad D_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad V_2 = I_2$$

Columns 3 to 6 of U_2 determine a basis of $K_1^{\text{free}} \oplus (C_1 / \ker \partial_1)$. We can confirm that the basis $B = (\beta_1, \dots, \beta_6)$ of C_1 by $[\beta_i] = \text{col}_i(U_2)$ partitions into bases for K_1^{tor} , K_1^{free} , and $C_1 / \ker \partial_1$ as given in Theorem 4.2.3 by considering the SND (U_1, D_1, V_1) of $[\partial_1]$ given as follows:

$$U_1 = \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} \begin{pmatrix} -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad D_1 = \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad V_1 = \begin{matrix} a_1 a_2 \\ a_1 a_3 \\ a_1 a_4 \\ a_2 a_3 \\ a_2 a_4 \\ a_3 a_4 \end{matrix} \begin{pmatrix} \begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{matrix} & \overbrace{\begin{matrix} 0 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 1 \end{matrix}}^{K_1^{\text{tor}}} & \underbrace{\begin{matrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{matrix}}_{K_1^{\text{free}}} \end{pmatrix}$$

$C_1 / \ker \partial_1$
 K_1^{free}

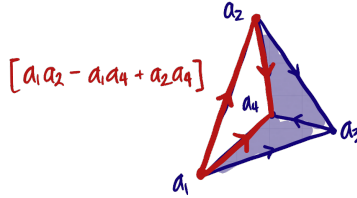
Observe that the matrix U_2 is the matrix V_1 up to column permutation. The basis $B = (\beta_1, \dots, \beta_6)$ of C_1 given by $[\beta_i] = \text{col}_i(U_2)$ is described below, grouped relative to the decomposition of C_1 .

$$\begin{aligned}
K_1^{\text{tor}} : \quad & \sigma_1 = a_1 a_3 - a_1 a_4 + a_3 a_4 \text{ with } d_1 = D_2(1, 1) = 1 \\
& \sigma_2 = a_2 a_3 - a_2 a_4 + a_3 a_4 \text{ with } d_2 = D_2(2, 2) = 1 \\
K_1^{\text{free}} : \quad & \sigma_3 = a_1 a_2 - a_1 a_4 + a_2 a_4 \\
& \sigma_4 = a_1 a_4 \\
C_1 / \ker \partial_1 : \quad & \sigma_5 = a_1 a_2 \\
& \sigma_6 = a_3 a_4
\end{aligned}$$

We can calculate $H_1(K)$ by disregarding the rows of D_2 corresponding to the basis elements of $C_1 / \ker \partial_1$, i.e. by only considering the basis elements of K_1^{free} and K_1^{tor} as follows:

$$\begin{aligned}
H_1(K) = \frac{\ker \partial_1}{\text{im } \partial_2} &\cong \frac{K_1^{\text{tor}}}{\text{im}(\partial_1)} \oplus K_1^{\text{free}} \cong \frac{\mathbb{Z}\langle \sigma_1 \rangle}{\mathbb{Z}\langle d_1 \sigma_1 \rangle} \oplus \frac{\mathbb{Z}\langle \sigma_2 \rangle}{\mathbb{Z}\langle d_2 \sigma_2 \rangle} \oplus \mathbb{Z}\langle \sigma_3 \rangle \\
&\cong \mathbb{Z}\langle a_1 a_2 - a_1 a_4 + a_2 a_4 \rangle \cong \mathbb{Z}
\end{aligned}$$

Note that both summands of K_1^{tor} map to trivial elements in $H_1(K)$ since both $\sigma_1, \sigma_2 \in \text{im } \partial_2$. The cycle representative σ_3 of the homology class $[\sigma_3] \in H_1(K)$ of K is illustrated below:



Once the decomposition of C_n into the three direct summands is established, an SND $(U_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$ with the required properties can be generated from any SND $(W_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$. The first r columns of W_{n+1} , as discussed earlier, determines a basis $(\alpha_1, \dots, \alpha_r)$ of K_n^{tor} by $[\alpha_i] = \text{col}_i(W_{n+1})$. Since K_n^{free} and $C_n / \ker(\partial_n)$ are direct summands of C_n , there must exist bases $(\gamma_1, \dots, \gamma_f)$ of K_n^{free} and $(\omega_1, \dots, \omega_q)$ of $C_n / \ker(\partial_n)$ that is a subset of some basis of C_n . Observe that B , as given below, is a basis of C_n .

$$B := \left\{ \underbrace{\alpha_1, \dots, \alpha_r}_{\text{basis of } K_n^{\text{tor}}}, \underbrace{\gamma_1, \dots, \gamma_f}_{\text{basis of } K_n^{\text{free}}}, \underbrace{\omega_1, \dots, \omega_q}_{\text{basis of } C_n / \ker \partial_n} \right\}$$

A matrix $U_{n+1} \in \text{GL}(m, R)$ such that $(U_{n+1}, D_{n+1}, V_{n+1})$ is an SND of $[\partial_{n+1}]$ and is as given by Theorem 4.2.3

can be defined as follows:

$$U_{n+1} = \left(\begin{array}{c|c|c} \begin{array}{c} \vdots \\ [\alpha_1] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\alpha_r] \\ \vdots \end{array} & \begin{array}{c} \vdots \\ [\gamma_1] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\gamma_f] \\ \vdots \end{array} & \begin{array}{c} \vdots \\ [\omega_1] \\ \vdots \end{array} & \cdots & \begin{array}{c} \vdots \\ [\omega_q] \\ \vdots \end{array} \end{array} \right)$$

K_n^{tor}

K_n^{free}

$C_n / \ker \partial_n$

Observe that since $\text{im}(\partial_{n+1}) \subseteq K_n^{\text{tor}}$, there should be no non-trivial elements of C_{n+1} mapping into K_n^{free} and $C_n / \ker \partial_n$. Then, the expression of $[\partial_{n+1}]$ relative to the basis B of C_n (with the same initial basis of C_{n+1}) should be given exactly by D_{n+1} . That is, replacing the basis elements $\alpha_{r+1}, \dots, \alpha_m$ of $K_n^{\text{free}} \oplus (C_n / \ker \partial_n)$ with any basis element leaves the matrix D_{n+1} undisturbed. We can confirm this by showing that the equation $(W_{n+1})^{-1}[\partial_{n+1}]V_{n+1} = D_{n+1}$ from the SND $(W_{n+1}, D_{n+1}, V_{n+1})$ implies that $(U_{n+1})^{-1}[\partial_{n+1}]V_{n+1} = D_{n+1}$.

Observe that if cycle representatives for the homology groups are not required, we can calculate the invariant factor decomposition of homology groups of free chain complexes using any SNDs of $[\partial_{n+1}]$ (for the nonzero invariant factors) and of $[\partial_n]$ (for $\text{rank}(\partial_n)$). We state this as a result below.

Corollary 4.2.5. Let $C_* = (C_n, \partial_n)$ be a chain complex of finitely-generated free modules C_n over a PID R with differentials $\partial_n : C_n \rightarrow C_{n-1}$. For each $n \in \mathbb{N}_0$, the n^{th} homology group $H_n(C_*)$ can be calculated using the SNF D_{n+1} of $[\partial_{n+1}]$ as follows:

$$H_n(C_*) \cong R^f \oplus \frac{R}{R\langle d_1 \rangle} \oplus \cdots \oplus \frac{R}{R\langle d_r \rangle}$$

where $\{d_1, \dots, d_r\}$ with $r = \text{rank}(\partial_{n+1})$ are exactly the nonzero diagonal elements of D_{n+1} and $f + r = \text{rank}(\ker \partial_n)$. Note that $\text{rank}(\ker \partial_n)$ can be determined from the SNF of $[\partial_n]$.

Proof. Let $(U_{n+1}, D_{n+1}, V_{n+1})$ be the SND of $[\partial_{n+1}]$ as denoted in Theorem 4.2.3, i.e. the basis $B = (\beta_1, \dots, \beta_m)$ of C_n induced by $[\beta_j] = \text{col}_j(U_{n+1})$ with $U_{n+1} \in \text{GL}(m, R)$ partitions into three bases: $(\beta_1, \dots, \beta_r)$ of K_n^{tor} , $(\beta_{r+1}, \dots, \beta_{r+f})$ of K_n^{free} , and $(\beta_{r+f+1}, \dots, \beta_m)$ of $C_n / \ker \partial_n$. The nonzero entries of D_{n+1} are given by d_1, \dots, d_r with $r = \text{rank}(\partial_{n+1})$ and $d_j = D_{n+1}(j, j)$. Since $\ker(\partial_n) = K_n^{\text{tor}} \oplus K_n^{\text{free}}$, $\text{rank}(\partial_n) = r + f$. Then,

$$\begin{aligned} H_n(C_*) &= \frac{\ker \partial_n}{\text{im } \partial_{n+1}} = \frac{K_n^{\text{tor}}}{\text{im}(\partial_{n+1})} \oplus K_n^{\text{free}} = \frac{R\langle \beta_1, \dots, \beta_r \rangle}{R\langle d_1 \beta_1, \dots, d_r \beta_r \rangle} \oplus R\langle \beta_{r+1}, \dots, \beta_{r+f} \rangle \\ &= \frac{R}{R\langle d_1 \rangle} \oplus \frac{R}{R\langle d_1 \rangle} \oplus R^f \end{aligned}$$

Since the SNF D_{n+1} of $[\partial_{n+1}]$ is unique up to multiplication by elementary dilations over R , any SND of $[\partial_{n+1}]$ will share the same SNF. ■

This result is particularly powerful since the SNDs of $[\partial_{n+1}]$ as guaranteed by Theorem 4.2.3 can be cumbersome to find. In addition, most software packages only provide the SNF of matrices over \mathbb{Z} as output, not the full SND.

We re-calculate $H_1(K)$ of Example 4.2.4 again using this corollary below. Note that the SNDs we provide in the example below are the SNDs returned by the `SmithDecomposition[-]` method with the matrices $[\partial_1]$ and $[\partial_2]$ as input. That is, the SNDs presented in Example 4.2.4 are specifically calculated to match Theorem 4.2.3.

Example 4.2.6. Let the simplicial complex K be as given in Example 4.2.4. An SND (U'_2, D'_2, V'_2) of $[\partial_2]$ is given below with the columns of U'_2 corresponding to K_1^{tor} highlighted in green and the columns corresponding to the direct sum $K_1^{\text{free}} \oplus (C_1 / \ker \partial_1)$ are highlighted in purple.

$$U_2 = \begin{matrix} & \overbrace{\begin{matrix} 0 & 0 & 0 & 1 & 0 & 0 \end{matrix}}^{K_1^{\text{tor}}} \\ \begin{matrix} a_1 a_2 \\ a_1 a_3 \\ a_1 a_4 \\ a_2 a_3 \\ a_2 a_4 \\ a_3 a_4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad D'_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad V'_2 = I_2$$

$\underbrace{\hspace{10em}}_{K_1^{\text{free}} \oplus (C_1 / \ker \partial_1)}$

An SND (U'_1, D'_1, V'_1) of $[\partial_1]$ is also given below, with the columns of V'_1 corresponding to $C_1 / \ker \partial_1$ highlighted in red and the columns corresponding to $\ker \partial_1$ in orange.

$$U'_1 = \begin{matrix} \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} -1 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad D'_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad V'_1 = \begin{matrix} & \overbrace{\begin{matrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}^{K_1^{\text{tor}} \oplus K_1^{\text{free}}} \\ \begin{matrix} a_1 a_2 \\ a_1 a_3 \\ a_1 a_4 \\ a_2 a_3 \\ a_2 a_4 \\ a_3 a_4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$\underbrace{\hspace{10em}}_{C_1 / \ker \partial_1}$

We can determine the following facts from these two SNDs:

1. From D'_1 , $\text{rank}(\ker \partial_1) = 6 - 3 = 3$ and from D'_2 , $r = \text{rank}(\partial_2) = 2$. Therefore, $f = 1$.
2. Let $A = (\alpha_1, \dots, \alpha_6)$ be a basis of C_1 by $[\alpha_j] = \text{col}_j(U'_2)$. We can confirm that the basis $(\alpha_3, \dots, \alpha_6)$ of $K_1^{\text{free}} \oplus (C_1 / \ker \partial_1)$ does not partition into bases for each summand since $\partial_1(\alpha_j) \neq 0$ for $j \in \{3, \dots, 6\}$ but $\text{rank}(K_1^{\text{free}}) = f = 1$.
3. Let $K = (\kappa_1, \dots, \kappa_6)$ be a basis on C_1 by $[\kappa_i] = \text{col}_i(V'_1)$. Since the basis elements $\kappa_4, \kappa_5, \kappa_6$ do not appear as basis elements in A , i.e. the columns of V'_1 are not columns of U'_2 . Note that it is not immediately clear if one of $\kappa_4, \kappa_5, \kappa_6$ generates K_1^{free} .

However, if we do not need to calculate cycle representatives, we have enough information to calculate $H_1(K)$ up to isomorphism using Corollary 4.2.5: From D'_2 , we have $r = 2$, $d_1 = 1$, $d_2 = 1$. From D'_1 , we have $\text{rank}(\ker \partial_1) = 3$ and $f = 3 - 2 = 1$. Then,

$$H_1(K) \cong \mathbb{Z}^f \oplus \left(\frac{\mathbb{Z}}{d_1 \mathbb{Z}} \right) \oplus \left(\frac{\mathbb{Z}}{d_2 \mathbb{Z}} \right) = \mathbb{Z}^{3-2} \oplus \left(\frac{\mathbb{Z}}{\mathbb{Z}} \right) \oplus \left(\frac{\mathbb{Z}}{\mathbb{Z}} \right) = \mathbb{Z}.$$

Observe that this matches with our calculation for $H_1(K)$ in Example 4.2.4.

Section 4.3. The Graded Structure Theorem and SNDs in the Graded Case

In Section 4.1, we discussed a method of calculating the invariant factor decompositions of finitely generated modules over a PID R , i.e. finitely generated modules in the category $\mathbf{Mod}_{\mathbb{F}[x]}$. In this section, we extend this method to the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ of *graded* $\mathbb{F}[x]$ -modules, wherein invariant factor decompositions need to be given by graded isomorphisms. For reference, we discuss the category $\mathbf{GrMod}_{\mathbb{F}[x]}$ and the notation we use involving graded $\mathbb{F}[x]$ -modules earlier in Section 2.4.

To start, we state **Graded Structure Theorem** in $\mathbf{GrMod}_{\mathbb{F}[x]}$ for graded $\mathbb{F}[x]$ -modules, which is the theorem corresponding to the Structure Theorem (Theorem 4.1.1) in the ungraded category $\mathbf{Mod}_{\mathbb{F}[x]}$ of $\mathbb{F}[x]$ -modules.

Theorem 4.3.1. The Graded Structure Theorem.

Let M be a finitely generated graded $\mathbb{F}[x]$ -module over $\mathbb{F}[x]$ for some field \mathbb{F} . Then, there exists a finite direct sum of shifted cyclic graded submodules of $\mathbb{F}[x]$ that is graded isomorphic to M as follows:

$$M \stackrel{\mathbf{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]$$

with indices $s_1, \dots, s_r, \dots, s_m \in \mathbb{N}_0$ and non-zero, non-unit $x^{t_1}, x^{t_2}, \dots, x^{t_r} \in \mathbb{F}[x]$ such that the divisibility relation $x^{t_1} \mid x^{t_2} \mid \cdots \mid x^{t_r}$ is satisfied. Furthermore, the collection $\{(x^{t_1}, s_1), \dots, (x^{t_r}, s_r), (0, s_{r+1}), \dots, (0, s_m)\}$ is uniquely determined by M up to graded isomorphism.

Remark. This theorem is stated in [ZC05, Theorem 2.1] without proof. The paper [Loe23] by Clara Löh provides a detailed proof, with this theorem listed as [Loe23, Theorem 5.1].

Much like the case of the Structure Theorem in $\mathbf{Mod}_{\mathbb{F}[x]}$, we give the decomposition guaranteed by the Graded Structure Theorem a special name.

Definition 4.3.2. Let M be a finitely generated $\mathbb{F}[x]$ -module and let the following direct sum decomposition of M be as denoted in the Graded Structure Theorem (Theorem 4.3.1):

$$M \stackrel{\mathbf{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]$$

This decomposition is called the **graded invariant factor decomposition** of M . The **invariant factors** of M are given by $x^{t_1}, \dots, x^{t_r} \in \mathbb{F}[x]$ (which are non-zero and non-unit) and the **grading shifts** of M by $s_1, \dots, s_m \in \mathbb{N}_0$.

Remark. The term *invariant factor* for graded $\mathbb{F}[x]$ -modules is deliberately chosen since it corresponds to that of $\mathbb{F}[x]$ -modules, i.e. the invariant factors of a graded $\mathbb{F}[x]$ -module M are exactly the invariant factors of M viewed as a $\mathbb{F}[x]$ -module (disregarding grading). Note that some references, e.g. [Loe23], define invariant factors to include zero elements. To avoid confusion, we sometimes use the term *nonzero invariant factor* for clarity.

In this section, we discuss how this graded invariant factor decomposition can be calculated, under certain restrictions, using a method similar to that for non-graded decompositions as presented in Section 4.1. Let M be a finitely generated graded $\mathbb{F}[x]$ -module. Since $\mathbb{F}[x]$ is a PID, the non-graded Structure Theorem (Theorem

4.1.1) guarantees the existence of an invariant factor decomposition for M as a $\mathbb{F}[x]$ -module as follows:

$$M \cong \left(\frac{\mathbb{F}[x]}{(f_1)} \right) \oplus \cdots \oplus \left(\frac{\mathbb{F}[x]}{(f_r)} \right) \oplus \mathbb{F}[x] \oplus \cdots \oplus \mathbb{F}[x]$$

with nonzero invariant factors $f_1, \dots, f_r \in \mathbb{F}[x]$. However, this isomorphism may not correspond to that between graded $\mathbb{F}[x]$ -modules. For example, if $f_j \in \mathbb{F}[x]$ is not a homogeneous element for some $j \in \{1, \dots, r\}$, then the ideal (f_j) is not a graded ideal of $\mathbb{F}[x]$ and the quotient $\mathbb{F}[x]/(f_j)$ is not necessarily a graded ring. If we assume that each $f_j \in \mathbb{F}[x]$ is homogeneous of the form $f_j = x^{t_j}$ and each homogeneous element $q \in M$ is mapped to a homogeneous element in one of the summands, then we can shift each summand such that the degree of $q \in M$ matches that of its image in the decomposition as follows:

$$M \stackrel{\text{GrMod}}{\cong} \sum^{\deg_h(q_1)} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \sum^{\deg_h(q_r)} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \sum^{\deg_h(q_{r+1})} \mathbb{F}[x] \oplus \cdots \oplus \sum^{\deg_h(q_{r+1})} \mathbb{F}[x] \quad (\text{E1})$$

where $q_j \in M$ is chosen such that q_j corresponds to the generator 1_j of each summand. We claim that, assuming we start with a presentation that respects the graded structure, then a non-graded invariant factor decomposition can be transformed (for lack of a better word) to a graded invariant factor decomposition. Below, we provide a definition for these kinds of presentations.

Definition 4.3.3. A **graded presentation** of a graded $\mathbb{F}[x]$ -module M is a presentation

$$F_S \xrightarrow{\varphi} F_G \xrightarrow{\pi} M \rightarrow 0$$

on M such that F_S and F_G are free graded $\mathbb{F}[x]$ -modules and the homomorphisms $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ are graded homomorphisms.

A graded presentation by $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ of a graded $\mathbb{F}[x]$ -module M determines M up to non-graded isomorphism by $M \cong F_G / \text{im}(\varphi)$, as discussed under Definition 4.1.2. Since $\text{im}(\varphi)$ is a graded submodule of F_G , the quotient $F_G / \text{im}(\varphi)$ is a graded $\mathbb{F}[x]$ -module that inherits the grading of F_G and the graded presentation also determines M up to graded isomorphism. As with the non-graded case, the calculation of graded invariant factor decompositions start with graded presentations. Below, we provide an existence claim for these graded presentations.

Lemma 4.3.4. There exists a graded presentation for any finitely generated graded $\mathbb{F}[x]$ -module.

Proof. Let M be a finitely generated graded $\mathbb{F}[x]$ -module. By assumption of M being graded, there must exist a homogeneous system of generators $\{a_1 x^{s_1}, \dots, a_m x^{s_m}\}$ of M with $\deg_h(a_j x^{s_j}) = s_j \in \mathbb{N}_0$ for all $j \in \{1, \dots, m\}$. Note that we write $a_j x^{s_j} \in M$ such that a_j is an element of the \mathbb{F} -vector space M_{s_j} with $M_{s_j} x^{s_j}$ being the homogeneous component of M of degree s_j . We can construct a graded presentation for M as follows:

Define the module of generators F_G to be the free graded $\mathbb{F}[x]$ -module with homogeneous basis given by $A = (\alpha_1 x^{s_1}, \dots, \alpha_m x^{s_m})$ with $\deg_h(\alpha_j x^{s_j}) = s_j$ and define $\pi : F_G \rightarrow M$ by $\alpha_j x^{s_j} \mapsto a_j x^{s_j}$ for $j \in \{1, \dots, m\}$. Then, π is a graded homomorphism. Observe that if the set of generators of M were not homogeneous, then the requirement that $\deg_h(\alpha_j x^{s_j}) = \deg_h(a_j x^{s_j})$ may not be fulfilled since $\deg_h(a_j x^{s_j})$ may be undefined. For reference, F_G can be expressed as a direct sum of shifted graded free $\mathbb{F}[x]$ -modules, each with one basis element, as follows:

$$F_G = \mathbb{F}[x]\langle \alpha_1 x^{s_1} \rangle \oplus \cdots \oplus \mathbb{F}[x]\langle \alpha_m x^{s_m} \rangle = \sum^{s_1} \mathbb{F}\langle \alpha_1 \rangle \oplus \cdots \oplus \sum^{s_m} \mathbb{F}\langle \alpha_m \rangle \quad (\text{E2})$$

Define the module of relations F_S by $F_S := \ker(\pi)$ and define $\varphi : F_S \rightarrow F_G$ to be the inclusion of $\ker(\pi)$ into F_G . Note that $\ker(\pi)$ is a graded submodule of F_S as it is the kernel of the graded homomorphism $\pi : F_G \rightarrow M$, and φ is graded as the identity map. As with the non-graded case, $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ correspond to a non-graded presentation of M . Since F_S, F_G are graded $\mathbb{F}[x]$ -modules and φ, π are graded homomorphisms, this presentation of M is also a graded presentation of M . ■

Let M be a finitely generated graded $\mathbb{F}[x]$ -module. Let a graded presentation of M be given by $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ with bases $S = (\sigma_1(x), \dots, \sigma_n(x))$ of F_S and $A = (\alpha_1(x), \dots, \alpha_m(x))$ of F_G . In the non-graded case, an SND (U, D, V) of $[\varphi]_{A,S}$, which exists since $\mathbb{F}[x]$ is a PID, gives us the following by Proposition 4.1.11:

1. A basis $T = (\tau_1(x), \dots, \tau_n(x))$ of F_G by $[\tau_i(x)]_S = \text{col}_i(V)$.
2. A basis $B = (\beta_1(x), \dots, \beta_m(x))$ of F_S by $[\beta_j(x)]_A = \text{col}_j(U)$.
3. The set of nonzero invariant factors $f_1, \dots, f_r \in \mathbb{F}[x]$ by $f_j = D(j, j)$.

Note that we write $\sigma_i(x)$ and $\alpha_j(x)$ here to emphasize that these elements are polynomials in x , i.e. the degree of $\sigma_i(x)$ and of $\alpha_j(x)$ are generally not zero.

For the process described for Equation (E1) to make sense, i.e. shifting the summands of the (non-graded) invariant factor decomposition, the basis $B = (\beta_j(x))$ must be a homogeneous basis. Otherwise, $\deg_h(\beta_j(x))$ and $\Sigma^{\deg_h(\beta_j(x))} \mathbb{F}[x]$ would be undefined. If we have that B is homogeneous and that the nonzero invariant factors $\{f_j\}$ are also homogeneous, then we can calculate a graded decomposition for M that would later correspond to a graded invariant factor decomposition. We state this in more detail below.

Proposition 4.3.5. Let M be a finitely generated graded $\mathbb{F}[x]$ -module. Let $\varphi : F_S \rightarrow F_G, \pi : F_G \rightarrow M$ correspond to a finite graded presentation on M with $\text{rank}(F_S) = n$ and $\text{rank}(F_G) = m$.

If there exists a homogeneous basis $\{\beta_1 x^{s_1}, \dots, \beta_m x^{s_m}\}$ of F_G and nonzero elements $x^{t_1}, \dots, x^{t_r} \in \mathbb{F}[x]$ with divisibility relation $x^{t_1} \mid \dots \mid x^{t_r}$ such that $\{\beta_1 x^{s_1+t_1}, \dots, \beta_r x^{s_r+t_r}\}$ is a basis for $\text{im}(\varphi)$, then we have the following graded isomorphism on M :

$$M \cong^{\text{GrMod}} \left(\frac{\mathbb{F}[x] \langle \beta_1 x^{s_1} \rangle}{\mathbb{F}[x] \langle \beta_1 x^{s_1+t_1} \rangle} \right) \oplus \dots \oplus \left(\frac{\mathbb{F}[x] \langle \beta_r x^{s_r} \rangle}{\mathbb{F}[x] \langle \beta_r x^{s_r+t_r} \rangle} \right) \oplus \mathbb{F}[x] \langle \beta_{r+1} x^{s_{r+1}} \rangle \oplus \dots \oplus \mathbb{F}[x] \langle \beta_m x^{s_m} \rangle \quad (\text{E3})$$

Note that we write the basis elements $\beta_j x^{s_j} \in F_G$ such that $\deg_h(\beta_j x^{s_j}) = s_j$.

Proof. Assume there does exist a homogeneous basis $\{\beta_j x^{s_j}\}_{j=1}^m$ of F_G and nonzero elements $\{x^{t_j}\}_{j=1}^r \in \mathbb{F}[x]$. For $j \in \{1, \dots, r\}$: $\mathbb{F}[x] \langle \beta_j x^{s_j+t_j} \rangle$ is a graded submodule of $\mathbb{F}[x] \langle \beta_j x^{s_j} \rangle$ since it has a homogeneous basis. Therefore, the quotient module $\mathbb{F}[x] \langle \beta_j x^{s_j} \rangle / \mathbb{F}[x] \langle \beta_j x^{s_j+t_j} \rangle$ is graded with homogeneous components of degree $q \in \{s_j, \dots, t_j - 1\}$ given by $[k \cdot \beta_j x^q]$ for $k \in \mathbb{F}$ nonzero.

By Proposition 4.1.5, (E3) is a $\mathbb{F}[x]$ -module isomorphism. For $j \in \{1, \dots, r\}$: the isomorphism by (1) maps $[\beta_j x^q] \in \mathbb{F}[x] \langle \beta_j x^{s_j} \rangle / \mathbb{F}[x] \langle \beta_j x^{s_j+t_j} \rangle$ with $q \in \{s_j, \dots, t_j - 1\}$ to $\pi(\beta_j x^q) \in M$. Since π is a graded homomorphism, we have that

$$\deg_h([\beta_j x^q]) = q = \deg_h(\beta_j x^q) = \deg_h(\pi(\beta_j x^q))$$

Similarly, for $j \in \{r+1, \dots, m\}$, the isomorphism by (1) maps $\beta_j x^q \in \mathbb{F}[x] \langle \beta_j x^{s_j} \rangle$ with $q \geq s_j$ to $\pi(\beta_j x^q)$ and $\deg_h(\beta_j x^q) = q = \deg_h(\pi(\beta_j x^q))$. Therefore, (E3) is a graded isomorphism. ■

We provide an example of this calculation below. Note that we have not yet proven that a homogeneous basis $\{\beta_j x^{s_j}\}$ and a set of nonzero invariant factors $\{x^{t_j}\}$ with the desired properties generally exists.

Example 4.3.6. Let M be a graded $\mathbb{Q}[x]$ -module and define a graded presentation on M as follows: Let $S = \{\sigma_1 x, \sigma_2 x^6, \sigma_3 x^8, \sigma_4 x^{10}\}$ be a basis on the $\mathbb{Q}[x]$ -module F_S of relations and let $A = \{\alpha_1 x, \alpha_2 x^2, \alpha_3 x^3, \alpha_4 x^4\}$ be a basis on the $\mathbb{Q}[x]$ -module F_G of generators. Observe that since S and A are homogeneous bases, F_S and F_G are both graded $\mathbb{Q}[x]$ -modules. Let $\varphi : F_S \rightarrow F_G$ be given by the following matrix:

$$[\varphi]_{A,S} = \begin{pmatrix} 1 & 4x^5 & 0 & 3x^9 \\ 0 & x^4 & -x^6 & 0 \\ 0 & 0 & x^5 & -2x^7 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We can confirm by inspection that φ is a graded homomorphism. Assume that $M = F_G / \text{im}(\varphi)$, i.e. π is the canonical quotient homomorphism. An SND (U_1, D_1, V_1) of $[\varphi]_{A,S}$ is given below:

$$U_1 = \begin{pmatrix} 1 & 2x & -x^2 & 0 \\ 0 & 1 & 0 & -x^2 \\ 0 & 0 & 1 & 2x \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^4 & 0 & 0 \\ 0 & 0 & x^5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad V_1 = \begin{pmatrix} 1 & -2x^5 & -5x^7 & -11x^9 \\ 0 & 1 & x^2 & 2x^4 \\ 0 & 0 & 1 & 2x^2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We can confirm this matrix factorization is correct by the following calculation:

$$(U_1)^{-1}[\varphi]_{A,S}V_1 = \begin{pmatrix} 1 & -2x & x^2 & -4x^3 \\ 0 & 1 & 0 & x^2 \\ 0 & 0 & 1 & -2x \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4x^5 & 0 & 3x^9 \\ 0 & x^4 & -x^6 & 0 \\ 0 & 0 & x^5 & -2x^7 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2x^5 & -5x^7 & -11x^9 \\ 0 & 1 & x^2 & 2x^4 \\ 0 & 0 & 1 & 2x^2 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \cdots = D_1$$

The matrix $U_1 \in \text{GL}(4, \mathbb{Q}[x])$ induces a new basis $B = (\beta_i x^{t_i})$ on F_G by $[\beta_i x^{t_i}]_A = \text{col}_i(U)$, given below, along with its corresponding diagonal element $d_i = D_1(i, i)$. Note that the basis elements of B can be written as $\beta_i x^{t_i}$ with $\deg_h(\beta_i x^{t_i}) = t_i$ since the basis B consists of homogeneous elements (confirmed post-calculation).

$$\begin{aligned} \beta_1 x^{t_1} &= (1)(\alpha_1)x &= (\alpha_1)x &\text{ with } d_1 = 1 \\ \beta_2 x^{t_2} &= (2x)(\alpha_1 x) + (1)(\alpha_2 x^2) &= (2\alpha_1 + \alpha_2)x^2 &\text{ with } d_2 = x^4 \\ \beta_3 x^{t_3} &= (-x^2)(\alpha_1 x) + (1)(\alpha_3 x^3) &= (-\alpha_1 + \alpha_3)x^3 &\text{ with } d_3 = x^5 \\ \beta_4 x^{t_4} &= (-x^2)(\alpha_2 x^2) + (2x)(\alpha_3 x^3) + (1)(\alpha_4 x^4) &= (-\alpha_2 + 2\alpha_3 + \alpha_4)x^4 &\text{ with } d_4 = 0 \end{aligned}$$

By applying Proposition 4.1.5, we have the following (non-graded) isomorphism on M . Note that since B is a homogeneous basis and the nonzero elements $d_i \in \mathbb{F}[x]$ are also homogeneous, the following decomposition is also a graded isomorphism by Proposition 4.3.5:

$$M \stackrel{\text{GrMod}}{\cong} \left(\frac{\mathbb{Q}[x]\langle \beta_1 x \rangle}{\mathbb{Q}[x]\langle \beta_1 x \rangle} \right) \oplus \left(\frac{\mathbb{Q}[x]\langle \beta_2 x^2 \rangle}{\mathbb{Q}[x]\langle x^4 \cdot \beta_2 x^2 \rangle} \right) \oplus \left(\frac{\mathbb{Q}[x]\langle \beta_3 x^3 \rangle}{\mathbb{Q}[x]\langle x^5 \cdot \beta_3 x^3 \rangle} \right) \oplus \mathbb{Q}[x]\langle \beta_4 x^4 \rangle$$

Since the only units of $\mathbb{Q}[x]$ are nonzero field elements, the nonzero entries of the SNF D_1 of $[\varphi]_{A,S}$ have to be homogeneous elements by uniqueness of SNFs. However, we can do row operations on the SNF $(U_1)^{-1}[\varphi]_{A,S}V_1 = D_1$ without disturbing the nonzero elements of D , i.e. 1, x^4 and x^5 . In particular, we can add any $\mathbb{Q}[x]$ -multiple of the 4th row of D_1 to rows 1, 2, and 3 of D_1 without changing D_1 . However, this will affect the matrix that determines the basis for F_G . For example, we can do the following:

$$E_{\text{add}}^{[4]}(1, 4; x^{10})D_1 = \begin{pmatrix} 1 & 0 & 0 & x^{10} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^4 & 0 & 0 \\ 0 & 0 & x^5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^4 & 0 & 0 \\ 0 & 0 & x^5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = D_1$$

Then, we can get another SND (U_2, D_1, V_1) of $[\varphi]_{A,S}$ by considering the following factorization:

$$D_1 = E_{\text{add}}^{[4]}(1, 4; x^{10}) D_1 = E_{\text{add}}^{[4]}(1, 4; x^{10}) (U_1)^{-1} [\varphi]_{A,S} V_1 = \underbrace{\left(U_1 E_{\text{add}}^{[4]}(1, 4; -x^{10}) \right)^{-1}}_{\text{Set this as } (U_2)^{-1}} [\varphi]_{A,S} V_1$$

Note that the inverse of the elementary transvection $E_{\text{add}}^{[4]}(1, 4; x^{10})$ is given by $E_{\text{add}}^{[4]}(1, 4; -x^{10})$. Then, the matrix $U_2 \in \text{GL}(4, \mathbb{Q}[x])$ determines a new basis $P = (p_1(x), \dots, p_4(x))$ of F_G as follows:

$$U_2 = \begin{pmatrix} 1 & 2x & -x^2 & -x^{10} \\ 0 & 1 & 0 & -x^2 \\ 0 & 0 & 1 & 2x \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{aligned} p_1(x) &= \beta_1 x \\ p_2(x) &= \beta_2 x^2 \\ p_3(x) &= \beta_3 x^3 \\ p_4(x) &= (-x^{10})(\alpha_1 x) + \beta_4 x^4 = -\alpha_1 x^{11} + \beta_4 x^4 \end{aligned}$$

Then, the isomorphism by Proposition 4.1.5 on the SND (U_2, D_1, V_1) of $[\varphi]_{A,S}$, given below, is **not** a graded isomorphism:

$$M \stackrel{\text{Mod}}{\cong} \left(\frac{\mathbb{F}[x]\langle \beta_1 x \rangle}{\mathbb{F}[x]\langle \beta_1 x \rangle} \right) \oplus \left(\frac{\mathbb{F}[x]\langle \beta_2 x^2 \rangle}{\mathbb{F}[x]\langle x^4 \cdot \beta_2 x^4 \rangle} \right) \oplus \left(\frac{\mathbb{F}[x]\langle \beta_3 x^3 \rangle}{\mathbb{F}[x]\langle x^5 \cdot \beta_3 x^5 \rangle} \right) \oplus \mathbb{F}[x]\langle -\alpha_1 x^{11} + \beta_4 x^4 \rangle$$

In particular, the summand $\mathbb{F}[x]\langle -\alpha_1 x^{11} + \beta_4 x^4 \rangle$ fails to be a graded $\mathbb{Q}[x]$ -module since it cannot be generated by a homogeneous element.

The decomposition given by Proposition 4.3.5 is then transformed into a graded invariant factor decomposition by applying specific graded isomorphisms to each summand in the direct sum. We state these isomorphisms in the following lemma.

Lemma 4.3.7. Let $\mathbb{F}[x]\langle \alpha x^s \rangle$ be the free graded $\mathbb{F}[x]$ -module generated by the element αx^s with $\deg(\alpha x^s) = s \in \mathbb{N}_0$. Let $x^t \in \mathbb{F}[x]$ for some $t \in \mathbb{N}_0$ with $t \neq 0$. Then, we have the following graded isomorphisms:

$$\mathbb{F}[x]\langle \alpha x^s \rangle \stackrel{\text{GrMod}}{\cong} \Sigma^s \mathbb{F}[x] \quad \text{and} \quad \frac{\mathbb{F}[x]\langle \alpha x^s \rangle}{\mathbb{F}[x]\langle x^t \cdot \alpha x^s \rangle} \stackrel{\text{GrMod}}{\cong} \Sigma^s \left(\mathbb{F}[x] / (x^t) \right)$$

Note that if $t = 0$, then $x^t = 1$ is a unit of $\mathbb{F}[x]$ and $\mathbb{F}[x]\langle \alpha x^s \rangle / \mathbb{F}[x]\langle x^t \cdot \alpha x^s \rangle = 0$.

Proof. Let $\varphi : \mathbb{F}[x]\langle \alpha x^s \rangle \rightarrow \Sigma^s \mathbb{F}[x]$ be given by $\alpha x^s \mapsto (x^s)(1)$. This is an isomorphism with inverse $(x^s)(1) \mapsto \alpha x^s$. Let $f \in \mathbb{F}[x]\langle \alpha x^s \rangle$ be a homogeneous element. Then, $f = k \cdot \alpha x^{s+r}$ for some $k \in \mathbb{F}$ nonzero and $r \in \mathbb{N}_0$ and

$$\deg_{\text{h}}(f) = \deg_{\text{h}}(k \cdot \alpha x^{s+r}) = s + r = \deg_{\text{h}}((x^{s+r})(1)) = \deg_{\text{h}}(\varphi(f))$$

The quotients $\mathbb{F}[x] / (x^t)$ and $\mathbb{F}[x]\langle \alpha x^s \rangle / \mathbb{F}[x]\langle \alpha x^{s+t} \rangle$ are graded $\mathbb{F}[x]$ -modules with grading determined by their coset representatives, well-defined since (x^t) and $\mathbb{F}[x]\langle \alpha x^{s+t} \rangle$ are graded submodules respectively. Then, the isomorphism $\mathbb{F}[x]\langle \alpha x^s \rangle / \mathbb{F}[x]\langle \alpha x^{s+t} \rangle \rightarrow \Sigma^s (\mathbb{F}[x] / (x^t))$ given by $[\alpha x^s] \rightarrow (x^s)[1]$ is graded similarly as in the case of φ . ■

By this lemma, the graded decomposition given by Proposition 4.3.5 can be transformed into a graded invariant factor decomposition as follows, with summands that have a unit for x^{t_j} , i.e. $t_j = 0$, removed from

the decomposition.

$$\begin{aligned}
M &\stackrel{\text{GrMod}}{\cong} \left(\frac{\mathbb{F}[x]\langle\beta_1 x^{s_1}\rangle}{\mathbb{F}[x]\langle\beta_1 x^{s_1+t_1}\rangle} \right) \oplus \cdots \oplus \left(\frac{\mathbb{F}[x]\langle\beta_r x^{s_r}\rangle}{\mathbb{F}[x]\langle\beta_r x^{s_r+t_r}\rangle} \right) \oplus \mathbb{F}[x]\langle\beta_{r+1} x^{s_{r+1}}\rangle \oplus \cdots \oplus \mathbb{F}[x]\langle\beta_m x^{s_m}\rangle \\
&\stackrel{\text{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]\langle\beta_1\rangle}{\mathbb{F}[x]\langle\beta_1 x^{t_1}\rangle} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]\langle\beta_r\rangle}{\mathbb{F}[x]\langle\beta_r x^{t_r}\rangle} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x]\langle\beta_{r+1}\rangle \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]\langle\beta_m\rangle \\
&\stackrel{\text{GrMod}}{\cong} \Sigma^{s_1} \left(\frac{\mathbb{F}[x]}{(x^{t_1})} \right) \oplus \cdots \oplus \Sigma^{s_r} \left(\frac{\mathbb{F}[x]}{(x^{t_r})} \right) \oplus \Sigma^{s_{r+1}} \mathbb{F}[x] \oplus \cdots \oplus \Sigma^{s_m} \mathbb{F}[x]
\end{aligned}$$

Observe that the set of grading shifts (s_1, \dots, s_m) are identified from the degree of the homogeneous basis element $\deg(\beta_j x^{s_j})$. We provide an example of calculating a graded invariant factor decomposition of a module below.

Example 4.3.8. Let M be the graded $\mathbb{Q}[x]$ -module as given in Example 4.3.6. We have calculated the following graded decomposition on M by Proposition 4.3.5:

$$M \stackrel{\text{GrMod}}{\cong} \left(\frac{\mathbb{Q}[x]\langle\beta_1 x\rangle}{\mathbb{Q}[x]\langle\beta_1 x\rangle} \right) \oplus \left(\frac{\mathbb{Q}[x]\langle\beta_2 x^2\rangle}{\mathbb{Q}[x]\langle x^4 \cdot \beta_2 x^2\rangle} \right) \oplus \left(\frac{\mathbb{Q}[x]\langle\beta_3 x^3\rangle}{\mathbb{Q}[x]\langle x^5 \cdot \beta_3 x^3\rangle} \right) \oplus \mathbb{Q}[x]\langle\beta_4 x^4\rangle$$

Then, by application of Lemma 4.3.7, the graded invariant factor decomposition on M is given as follows:

$$M \stackrel{\text{GrMod}}{\cong} \Sigma^2 \left(\mathbb{Q}[x] / (x^4) \right) \oplus \Sigma^3 \left(\mathbb{Q}[x] / (x^5) \right) \oplus \Sigma^4 \mathbb{Q}[x]$$

Note that the first summand with the basis element $\beta_1 x$ becomes a trivial module. Recall that in Chapter 2, we discussed how graded $\mathbb{Q}[x]$ -modules correspond to persistence modules over \mathbb{Q} . Then, by Corollary 2.5.13, we have the following isomorphism:

$$\Gamma_{\text{Pers}}(M) \stackrel{\text{Pers}}{\cong} \mathbb{I}_{\bullet}^{[2,2+4)} \oplus \mathbb{I}_{\bullet}^{[3,3+5)} \oplus \mathbb{I}_{\bullet}^{[4,\infty)} = \mathbb{I}_{\bullet}^{[2,6)} \oplus \mathbb{I}_{\bullet}^{[3,8)} \oplus \mathbb{I}_{\bullet}^{[4,\infty)}$$

where $\stackrel{\text{Pers}}{\cong}$ refers to an isomorphism in the category $\mathbf{Pers}_{\mathbb{Q}}$ of persistence modules over \mathbb{Q} .

Section 4.4. Matrix Reduction of Graded Matrices

Let M be a graded $\mathbb{F}[x]$ -module with a graded presentation given by $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$. In Proposition 4.3.5, we stated that if there exists a homogeneous basis $\{\beta_j x^{s_j}\}_{j=1}^m$ of F_G and nonzero invariant factors $x^{t_1}, \dots, x^{t_r} \in \mathbb{F}[x]$ such that $\{\beta_1 x^{s_1+t_1}, \dots, \beta_r x^{s_r+t_r}\}$ is a basis for $\text{im}(\varphi)$, then the graded invariant factor decomposition of M can be calculated from said basis. In this section and in Section 4.5, we argue that we can always calculate such a basis, assuming we know of homogeneous bases A and S of F_S and F_G respectively.

Our argument makes use of the general algorithm for the calculation of Smith Normal Decompositions described in [AW92, Remark 5.3.4] for matrices over arbitrary Euclidean domains, as discussed in Section 4.1 after Example 4.1.14. This algorithm involves strategically performing matrix reduction on $[\varphi]_{A,S}$, with V being the product of elementary matrices corresponding to column reduction and U^{-1} being that for row reduction. Since $[\varphi]_{A,S}$ is a matrix of a graded homomorphism, there are a number of restrictions on the entries of $[\varphi]_{A,S}$ that makes matrix reduction more straightforward and allows a more simplified version of the algorithm by [AW92, Remark 5.3.4].

Assume A and S are homogeneous bases. In this section, we consider four types of elementary matrices used in matrix reduction and discuss how matrix reduction preserves the homogeneity of these bases. Then, in Section 4.5, we present *Algorithm 4.5.1. Matrix Reduction Algorithm for Graded SNDs*, which calculates the desired SND (U, D, V) of $[\varphi]_{A,S}$ using only said types of elementary matrices.

Our first result is about the entries of matrices over $\mathbb{F}[x]$ corresponding to graded presentations.

Lemma 4.4.1. Let $\varphi : N \rightarrow M$ be a graded homomorphism between two free graded $\mathbb{F}[x]$ -modules N and M with ordered homogeneous bases $S = (\sigma_1(x), \dots, \sigma_n(x))$ and $A = (\alpha_1(x), \dots, \alpha_m(x))$ respectively. Let $[\varphi]$ be the matrix of φ relative to S and A , i.e. $[\varphi] := [\varphi]_{A,S} \in M_{m,n}(\mathbb{F}[x])$.

For all row indices $j \in \{1, \dots, m\}$ and column indices $i \in \{1, \dots, n\}$ such that $[\varphi](j, i) \neq 0$, $[\varphi](j, i)$ is a homogeneous element of $\mathbb{F}[x]$ and

$$\deg_h(\alpha_j(x)) + \deg_h([\varphi](j, i)) = \deg_h(\sigma_i(x))$$

where $\deg_h(\alpha_j(x))$ and $\deg_h(\sigma_i(x))$ are defined by homogeneity of the bases S and A .

Proof. For brevity, write $\sigma_i := \sigma_i(x)$ for $i \in \{1, \dots, n\}$ and $\alpha_j := \alpha_j(x)$ for $j \in \{1, \dots, m\}$.

Let $i \in \{1, \dots, n\}$ be a column index such that $\varphi(\sigma_i) \neq 0$. Since $\varphi : N \rightarrow M$ is a graded homomorphism by assumption, $\text{col}_i([\varphi])$ is not a zero column and $\deg_h(\sigma_i) = \deg_h(\varphi(\sigma_i))$. By definition of matrix of module homomorphisms, we have that $\text{col}_i([\varphi]) = [\varphi(\sigma_i)]_A$ and

$$\varphi(\sigma_i) = \sum_{j=1}^m (\text{col}_i([\varphi])(j) \alpha_j) = \sum_{j=1}^m [\varphi](j, i) \alpha_j$$

Let $j \in \{1, \dots, m\}$ be a row index such that $[\varphi](j, i) \neq 0$. Since $\varphi(\sigma_i)$ is homogeneous of degree $\deg_h(\varphi(\sigma_i))$, the summand $[\varphi](j, i) \alpha_j$ must be homogeneous of the same degree. Moreover, $[\varphi](j, i)$ must also be homogeneous. Otherwise, the product $[\varphi](j, i) \alpha_j$ would not be homogeneous. Then, we have that:

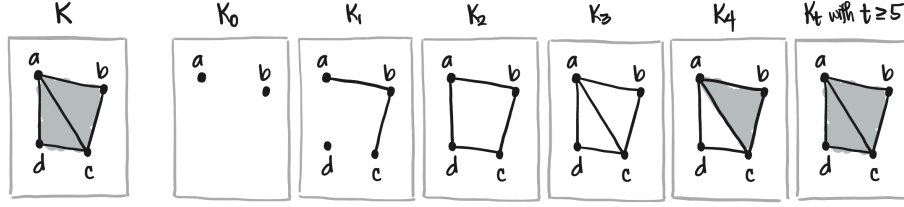
$$\deg_h(\sigma_i) = \deg_h(\varphi(\sigma_i)) = \deg_h([\varphi](j, i) \alpha_j) = \deg_h([\varphi](j, i)) + \deg_h(\alpha_j).$$

■

Below, we provide an example of a graded homomorphism involving the persistent homology of some

filtration, as discussed in Section 3.3.

Example 4.4.2. Let K_\bullet and K be given as in Example 3.2.3 and orient K with the vertex order $\text{Vert}(K) = (a, b, c, d)$. For convenience, an illustration of K and K_\bullet (without orientation) is copied below:



The 1st and 2nd graded chain modules of K_\bullet with coefficients in \mathbb{Q} are given as follows, relative to the ordered bases induced by the orientation $\text{Vert}(K) = (a, b, c, d)$ on the simplicial complex K :

$$C_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) = \mathbb{Q}[x] \langle abx, bcx, adx^2, cdx^2, acx^3 \rangle \quad \text{and} \quad C_2^{\text{Gr}}(K_\bullet; \mathbb{Q}) = \mathbb{Q}[x] \langle abcx^4, acdx^5 \rangle$$

The matrix $[\partial_2^{\text{Gr}}]$ of the graded boundary map $\partial_2^{\text{Gr}} : C_2^{\text{Gr}}(K_\bullet; \mathbb{Q}) \rightarrow C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ relative to the standard ordered bases is given as follows:

$$[\partial_2^{\text{Gr}}] = \begin{matrix} & \begin{matrix} abcx^4 & acdx^5 \end{matrix} \\ \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} & \begin{pmatrix} x^3 & 0 \\ x^3 & 0 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & x^2 \end{pmatrix} \end{matrix}$$

Observe that the nonzero entries of $[\partial_2^{\text{Gr}}]$ satisfy the degree relation stated in Lemma 4.4.1. We list these below.

$$\begin{aligned} [\partial_2^{\text{Gr}}](1, 1) &= x^3 & : \quad \deg_h(x^3) + \deg_h(abx) &= \deg_h(abcx^4) \\ [\partial_2^{\text{Gr}}](2, 1) &= x^3 & : \quad \deg_h(x^3) + \deg_h(bcx) &= \deg_h(abcx^4) \\ [\partial_2^{\text{Gr}}](3, 2) &= -x^3 & : \quad \deg_h(-x^3) + \deg_h(adx^2) &= \deg_h(acdx^5) \\ [\partial_2^{\text{Gr}}](4, 2) &= x^3 & : \quad \deg_h(x^3) + \deg_h(cdx^2) &= \deg_h(acdx^5) \\ [\partial_2^{\text{Gr}}](5, 1) &= -x & : \quad \deg_h(-x) + \deg_h(acx^3) &= \deg_h(abcx^4) \\ [\partial_2^{\text{Gr}}](5, 2) &= x^2 & : \quad \deg_h(x^2) + \deg_h(acx^3) &= \deg_h(acdx^5) \end{aligned}$$

Note that Lemma 4.4.1 does not apply for zero entries since $\deg_h(0)$ is undefined.

Since Lemma 4.4.1 implies that the nonzero elements of the matrices over $\mathbb{F}[x]$ are homogeneous and that their degrees are fixed by the degrees of the initial homogeneous bases (with the added assumptions stated on the same lemma), we can show that elimination operations done on the matrix reduction algorithm for graded SNDs must also preserve the homogeneity of the initial bases, i.e. the bases that they induce are also homogeneous assuming we start with homogeneous bases.

Column operations on a matrix correspond to multiplication of said matrix on the *right* by an elementary matrix (see Proposition A2.8 in Appendix A2). For clarity, we identify certain terminology involving column reduction of matrices.

Definition 4.4.3. A **column reduction operation** on a matrix $T \in M_{m,n}(\mathbb{F}[x])$ consists of the following:

- i. a fixed row index $r \in \{1, \dots, m\}$,

- ii. a nonzero **target entry** $T(r, k)$ with **target column index** $k \in \{1, \dots, n\}$ to be eliminated in T ,
- iii. a nonzero **pivot entry** $T(r, p)$ with **pivot column index** $p \in \{1, \dots, n\}$,
- iv. a **pivot multiplier** $f \in \mathbb{F}[x]$ such that $T(r, k) + f \cdot T(r, p) = 0$, and
- v. an elementary transvection $V := E_{\text{add}}^{[n]}(p, k; f)$, multiplied to T on the right. Note that the pivot column index p is the first argument and the target column index k is the second.

Then, the entry $T(r, k)$ is eliminated in the product $TV = TE_{\text{add}}^{[n]}(p, k; f)$, i.e. $(TV)(r, k) = 0$.

Following the notation above, we can show that $(TV)(r, k) = 0$ by looking at the columns of the product TV and using a column-wise description of $V = E_{\text{add}}^{[n]}(p, k; f)$: for all column indices $i \in \{1, \dots, n\}$,

$$\text{col}_i(TV) = T \text{col}_i(V) = T \text{col}_i\left(E_{\text{add}}^{[n]}(p, k; f)\right) = \begin{cases} \text{col}_k(T) + f \cdot \text{col}_p(T) & \text{if } i = k \\ \text{col}_i(T) & \text{otherwise} \end{cases}$$

Note that T and TV can differ only on the k^{th} column. Then, the $(r, k)^{\text{th}}$ entry of TV would be zero, i.e. the target entry $T(r, k)$ is eliminated as follows:

$$\text{col}_k(T)(r) + f \cdot \text{col}_p(T)(r) = T(r, k) + f \cdot T(r, p) = 0$$

For arbitrary matrices over $\mathbb{F}[x]$, elimination operations as above generally cannot be done since f exists if and only if the pivot entry $T(r, p)$ divides the target entry $T(r, k)$. However, assuming $T = [\varphi]$ is as given in Lemma 4.4.1, the elements $[\varphi](r, p)$ and $[\varphi](r, k)$ are homogeneous and we have the following equivalence:

$$\left(\begin{array}{c} [\varphi](r, p) \mid [\varphi](r, k) \\ \text{divisibility relation} \end{array} \right) \quad \text{if and only if} \quad \left(\begin{array}{c} \deg_h([\varphi](r, p)) \leq \deg_h([\varphi](r, k)) \\ \text{degree relation} \end{array} \right)$$

Then, $f \in \mathbb{F}[x]$ is given by $f = -[\varphi](r, k) / [\varphi](r, p)$ assuming that $\deg_h([\varphi](r, p)) \leq \deg_h([\varphi](r, k))$. The question now is whether the basis induced by $V \in \text{GL}(n, \mathbb{F}[x])$ and the remaining entries of $\text{col}_k([\varphi]V)$ correspond to homogeneous elements. The following proposition addresses this.

Proposition 4.4.4. Column Reduction Preserves Homogeneity.

Let $\varphi : N \rightarrow M$ be a graded homomorphism between two free graded $\mathbb{F}[x]$ -modules N and M with ordered homogeneous bases $S = (\sigma_1(x), \dots, \sigma_n(x))$ and $A = (\alpha_1(x), \dots, \alpha_m(x))$ respectively. Let $[\varphi]$ be the matrix of φ relative to S and A , i.e. $[\varphi] := [\varphi]_{A,S} \in M_{m,n}(\mathbb{F}[x])$.

Let $p, k \in \{1, \dots, n\}$ be distinct column indices and $r \in \{1, \dots, m\}$ be a row index such that the target $[\varphi](r, k)$ and the pivot $[\varphi](r, p)$ are both nonzero and $\deg_h([\varphi](r, p)) \leq \deg_h([\varphi](r, k))$. If $f \in \mathbb{F}[x]$ is nonzero and homogeneous such that

$$\deg_h([\varphi](r, k)) = \deg_h(f) + \deg_h([\varphi](r, p))$$

then the basis $T = (\tau_1(x), \dots, \tau_n(x))$ of N induced by the elementary transvection $V := E_{\text{add}}^{[n]}(p, k; f)$ by $[\tau_i]_S = \text{col}_i(V)$ is homogeneous.

Proof. Let $f \in \mathbb{F}[x]$ be nonzero and homogeneous such that $\deg_h([\varphi](r, k)) = \deg_h(f) + \deg_h([\varphi](r, p))$. For brevity, let $\sigma_i := \sigma_i(x)$, $\tau_i := \tau_i(x)$ and $\alpha_j := \alpha_j(x)$.

Since $V := E_{\text{add}}^{[n]}(p, k; f)$ differs from the identity matrix I_n only in the $(p, k)^{\text{th}}$ entry, we have that $\tau_i = \sigma_i$ and τ_i is homogeneous for all $i \in \{1, \dots, n\}$ with $i \neq a$. We need to check that $\tau_k \in N$ is also homogeneous. Since $[\tau_k]_S = \text{col}_k(V) = \text{col}_k(I_n) + f \cdot \text{col}_p(I_n)$, we have that $\tau_k = \sigma_k + f \cdot \sigma_p$.

For τ_k to be homogeneous, we need to show that $\deg_h(\sigma_k) = \deg_h(f \cdot \sigma_p)$:

$$\begin{aligned}
\deg_h(f \cdot \sigma_p) &= \deg_h(f) + \deg_h(\sigma_p) && \text{by homogeneity of } f \text{ and } \sigma_p \\
&= \deg_h([\varphi](r, k)) - \deg_h([\varphi](r, p)) && \text{by assumption on } f \in \mathbb{F}[x] \\
&\quad + \deg_h(\sigma_p) \\
&= \deg_h([\varphi](r, k)) - \deg_h([\varphi](r, p)) \\
&\quad + \deg_h([\varphi](r, p)) + \deg_h(\alpha_r) && \text{by Lemma 4.4.1 on the } [\varphi](r, p) \neq 0 \\
&= \deg_h([\varphi](r, k)) + \deg_h(\alpha_r) \\
&= \deg_h(\sigma_k) && \text{by Lemma 4.4.1 on } [\varphi](r, k) \neq 0
\end{aligned}$$

Since $\deg_h(\sigma_k) = \deg_h(f \cdot \sigma_p)$, $\tau_k \in N$ is homogeneous with $\deg_h(\tau_k) = \deg_h(\sigma_k)$. Therefore, $T = (\tau_i)_{i=1}^n$ is a homogeneous basis of N . ■

Note that the statement of Proposition 4.4.4 applies even when $f \in \mathbb{F}[x]$, as denoted above, is not used to eliminate entries, i.e. it is not required that $[\varphi](r, k) + f \cdot [\varphi](r, p) = 0$, only that $[\varphi](r, k) + f \cdot [\varphi](r, p)$ results in either zero or a homogeneous element of the same degree. While this is not critical for the reduction algorithm, it does allow some steps that may be more preferable when working by hand.

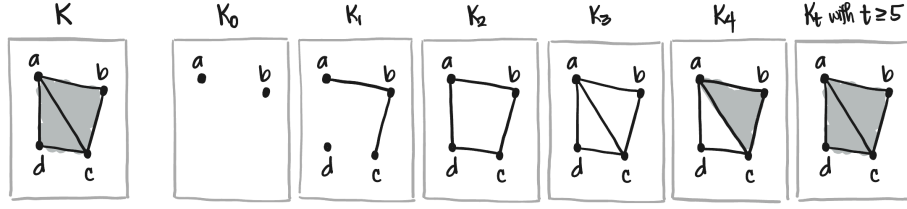
Observe that the key part of the proof of Proposition 4.4.4 is the existence of a row index $r \in \{1, \dots, m\}$ such that both $[\varphi](r, k)$ and $[\varphi](r, p)$ are nonzero, in that they determine the degree of the element $f \in \mathbb{F}[x]$. Note that we use the degree relation from Lemma 4.4.1 on both $[\varphi](r, k)$ and $[\varphi](r, p)$, which require them to both be nonzero. If such a row index r does not exist, then it is possible for the nonzero entries of $[\varphi]V$ to be homogeneous but the basis $T = (\tau_i)$ induced by V to not be a homogeneous basis. A trivial example would involve a zero column on $[\varphi]$, wherein if the nonzero entries of $[\varphi]$ are homogeneous and $\text{col}_p([\varphi])$ is the zero column, then $[\varphi]E_{\text{add}}^{[n]}(p, k f)$ for any column index k and any element $f \in \mathbb{F}[x]$ would also have entries that are either zero or homogeneous. We saw a similar issue happen with row reduction earlier in Example 4.3.6 wherein the matrix $U_2 \in \text{GL}(4, \mathbb{F}[x])$ produced a non-homogeneous basis since the 4th row is the zero row. Fortunately, column operations of this type are generally not required for matrix reduction since these tend to create more nonzero entries, i.e. the opposite goal of matrix reduction.

As denoted above, $[\varphi]V = [\varphi]_{A,S}V = [\varphi]_{A,T}$, i.e. $[\varphi]V$ corresponds to the matrix of φ relative to the bases $T = (\tau_i(x))_{i=1}^n$ of N and $A = (\alpha_j(x))_{j=1}^m$ of M . Since we have shown that T is a homogeneous basis of N , Lemma 4.4.1 also applies for $[\varphi]V = [\varphi]_{A,T}$, i.e. for all nonzero entries $[\varphi]_{A,T}(j, i)$,

$$\deg_h(\alpha_j(x)) + \deg_h([\varphi]_{A,T}(j, i)) = \deg_h(\tau_i(x)).$$

An immediate consequence of this is that we can do elimination by column reduction finitely many times and the resulting matrices from said operation will also preserve the homogeneity of the initial bases. Below, we provide an example of this column reduction process in action. Note that the following example does not use the algorithm by [AW92, Remark 5.3.4] (Algorithm 4.5.1) since we have not established that row reduction preserves homogeneity. Instead, this example more closely resembles the column reduction algorithm discussed in the next section.

Example 4.4.5. Let K_\bullet and K be given as in Example 3.2.3 and as illustrated below.



The 0th and 1st graded chain modules of K_\bullet with coefficients in \mathbb{Q} are given as follows, relative to ordered bases induced by the orientation $\text{Vert}(K) = (a, b, c, d)$ on the simplicial complex K :

$$C_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) = \mathbb{Q}[x]\langle a, b, cx, dx \rangle \quad \text{and} \quad C_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) = \mathbb{Q}[x]\langle abx, bcx, adx^2, cdx^2, acx^3 \rangle$$

Below, we identify the matrix $[\partial_1^{\text{Gr}}]$ of the 1st graded boundary morphism $\partial_1^{\text{Gr}} : C_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) \rightarrow C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ relative to the same ordered bases:

$$[\partial_1^{\text{Gr}}] = \begin{matrix} & abx & bcx & adx^2 & cdx^2 & acx^3 \\ \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} & \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & 0 & x & 0 \end{pmatrix} \end{matrix}$$

Given below is a sequence Q_1, Q_2, \dots of matrices $Q_n \in M_{4,5}(\mathbb{F}[x])$ resulting from successive column reduction operations on $Q_0 := [\partial_1^{\text{Gr}}]$, with the following color scheme:

chosen pivot entry, target entry for elimination, pivot multiplier, affected column

$$\begin{aligned} Q_1 &:= Q_0 E_{\text{add}}^{[5]}(3, 4; 1) = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -x & 0 & -x^2 & -x^2 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \\ Q_2 &:= Q_1 E_{\text{add}}^{[5]}(2, 4; -x) = \begin{pmatrix} -x & 0 & -x^2 & -x^2 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -x & 0 & -x^2 & -x^2 & -x^3 \\ x & -x & 0 & x^2 & 0 \\ 0 & 1 & 0 & 0 & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \\ Q_3 &:= Q_2 E_{\text{add}}^{[5]}(1, 4; -x) = \begin{pmatrix} -x & 0 & -x^2 & -x^2 & -x^3 \\ x & -x & 0 & x^2 & 0 \\ 0 & 1 & 0 & 0 & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -x & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \\ Q_4 &:= Q_3 E_{\text{add}}^{[5]}(2, 5; -x^2) = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x^2 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & x^3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \end{aligned}$$

$$Q_5 := Q_4 E_{\text{add}}^{[5]}(1, 5; -x^2) = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & x^3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & -x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -x & 0 & -x^2 & 0 & 0 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix}$$

Then, we have $Q_5 = [\partial_1^{\text{Gr}}]V$ with $V \in \text{GL}(5, \mathbb{F}[x])$ given as follows, with each successive elementary transvection matrix multiplied on the right of Q_i :

$$V = \overbrace{E_{\text{add}}^{[5]}(3, 4; 1)}^{\text{first column elimination}} \underbrace{E_{\text{add}}^{[5]}(2, 4; -x)}_{\text{second column elimination}} \cdots \overbrace{E_{\text{add}}^{[5]}(1, 5; -x^2)}^{\text{fifth and last column elim.}} = \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & -x & -x^2 \\ 0 & 1 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then, $V \in \text{GL}(5, \mathbb{F}[x])$ induces a new basis $T = (\tau_1(x), \dots, \tau_5(x))$ by $[\tau_i(x)] = \text{col}_i(V)$, i.e.

$$\begin{aligned} \tau_1(x) &= (1)(abx) = (ab)x \\ \tau_2(x) &= (1)(bcx) = (bc)x \\ \tau_3(x) &= (1)(adx^2) = (ad)x^2 \\ \tau_4(x) &= (-x)(abx) + (-x)(bcx) + (1)(adx^2) = (-ab - bc + ad)x^2 \\ \tau_5(x) &= (-x^2)(abx) + (-x^2)(bcx) + (1)(acx^3) = (-ab - bc + ac)x^3 \end{aligned}$$

Observe that T is a homogeneous basis for $C_2^{\text{Gr}}(K_{\bullet}; \mathbb{Q})$ and that the nonzero entries of Q_5 are as expected by Lemma 4.4.1. For example, at the $(1, 3)^{\text{th}}$ entry of Q_5 , we have that

$$\deg_h(a) + \deg_h(Q_5(1, 3)) = 0 + \deg_h(-x^2) = 3 = \deg_h(\tau_3(x)) = \deg_h(adx^2)$$

We have a similar result for elimination by row reduction. Row operations on a matrix correspond to multiplication of said matrix on the *left* by an elementary matrix (Proposition A2.7 in Appendix A2). For clarity, we identify certain terminology involving elimination by row reduction below.

Definition 4.4.6. A **row reduction operation** on a matrix $T \in M_{m,n}(\mathbb{F}[x])$ consists of the following:

- i. a fixed column index $c \in \{1, \dots, n\}$,
- ii. a nonzero **target entry** $T(k, c)$ with **target row index** $k \in \{1, \dots, m\}$ to be eliminated in T ,
- iii. a nonzero **pivot entry** $T(p, c)$ with **pivot row index** $p \in \{1, \dots, m\}$,
- iv. a **pivot multiplier** $f \in \mathbb{F}[x]$ such that $T(k, c) + f \cdot T(p, c) = 0$, and
- v. an elementary transvection $U^{-1} := E_{\text{add}}^{[m]}(k, p; f)$, multiplied to T on the left. Note that the pivot row index p is the *second* argument and the target row index k is the *first*, and that $U = E_{\text{add}}^{[m]}(k, p; -f)$.

Then, the entry $T(k, c)$ is eliminated in the product $U^{-1}T = E_{\text{add}}^{[m]}(k, p; f)T$, i.e. $(U^{-1}T)(k, c) = 0$.

Remark. We define $U^{-1} := E_{\text{add}}^{[m]}(k, p; f)$, as opposed to setting $U := E_{\text{add}}^{[m]}(k, p; f)$, since we want to interpret the resulting matrix as a change of basis on the codomain of homomorphisms.

Following the notation above, we can show that $(U^{-1}T)(k, c) = 0$ by examining the the rows of the product

$U^{-1}T \in M_{m,n}(\mathbb{F}[x])$ and using a row-wise description of $U^{-1} = E_{\text{add}}^{[m]}(k, p; f)$: for all row indices $j \in \{1, \dots, m\}$,

$$\text{row}_j(U^{-1}T) = \text{row}_j(U^{-1})T = \text{row}_j\left(E_{\text{add}}^{[m]}(k, p; f)\right)T = \begin{cases} \text{row}_k(T) + f \cdot \text{row}_p(T) & \text{if } j = k \\ \text{row}_j(T) & \text{otherwise} \end{cases}$$

As with the case of column reduction, when $T = [\varphi]$ is as given in Lemma 4.4.1, the elimination process becomes more straightforward. We have the following result involving row reduction of matrices given by Lemma 4.4.1.

Proposition 4.4.7. Row Reduction Preserves Homogeneity.

Let $\varphi : N \rightarrow M$ be a graded homomorphism between two free graded $\mathbb{F}[x]$ -modules N and M with ordered homogeneous bases $S = (\sigma_1, \dots, \sigma_n)$ and $A = (\alpha_1, \dots, \alpha_m)$ respectively. Let $[\varphi]$ be the matrix of φ relative to S and A , i.e. $[\varphi] := [\varphi]_{A,S} \in M_{m,n}(\mathbb{F}[x])$.

Let $p, k \in \{1, \dots, m\}$ be distinct row indices and $c \in \{1, \dots, n\}$ be a column index such that the pivot $[\varphi](p, c)$ and the target $[\varphi](k, c)$ are both nonzero and $\deg_h([\varphi](p, c)) \leq \deg_h([\varphi](k, c))$. If $f \in \mathbb{F}[x]$ is nonzero and homogeneous such that

$$\deg_h([\varphi](k, c)) = \deg_h(f) + \deg_h([\varphi](p, c))$$

then the basis $B = (\beta_1(x), \dots, \beta_m(x))$ of M induced by the elementary transvection $U := E_{\text{add}}^{[n]}(k, p; -f)$ by $[\beta_j]_A = \text{col}_j(U)$ is homogeneous.

Proof. Let $f \in \mathbb{F}[x]$ be nonzero and homogeneous such that $\deg_h([\varphi](k, c)) = \deg_h(f) + \deg_h([\varphi](p, c))$. For brevity, let $\sigma_i := \sigma_i(x)$, $\alpha_j := \alpha_j(x)$, and $\beta_j := \beta_j(x)$. Note that while $U := E_{\text{add}}^{[n]}(k, p; -f) \in \text{GL}(m, \mathbb{F}[x])$ corresponds to a row operation, the basis $B = (\beta_1, \dots, \beta_m)$ is determined by the columns of U .

For $j \in \{1, \dots, m\}$ with $j \neq p$, $\text{col}_j(U) = \text{col}_j(I_m)$ and $\beta_j = \alpha_j$ is homogeneous. Consider the p^{th} column of U . Since $[\beta_p]_A = \text{col}_p(U) = \text{col}_p(I_m) - f \cdot \text{col}_k(I_m)$, $\beta_p = \alpha_p - f \cdot \alpha_k$. For $\beta_p \in M$ to be homogeneous, we need to show that $\deg_h(\alpha_p) = \deg_h(-f \cdot \alpha_k)$. Note that $\deg_h(-f) = \deg_h(f)$.

$$\begin{aligned} \deg_h(f \cdot \alpha_k) &= \deg_h(f) + \deg_h(\alpha_k) && \text{by homogeneity of } f \text{ and } \alpha_k \\ &= \deg_h([\varphi](k, c)) - ([\varphi](p, c)) && \text{by assumption on } f \in \mathbb{F}[x] \\ &\quad + \deg_h(\alpha_k) \\ &= \deg_h([\varphi](k, c)) - \deg_h([\varphi](p, c)) \\ &\quad + \deg_h(\sigma_c) - \deg_h([\varphi](k, c)) && \text{by Lemma 4.4.1 on } [\varphi](k, c) \neq 0 \\ &= -\deg_h([\varphi](p, c)) + \deg_h(\sigma_c) \\ &= \deg_h(\alpha_p) && \text{by Lemma 4.4.1 on } [\varphi](p, c) \neq 0 \end{aligned}$$

Since $\deg_h(\alpha_p) = \deg_h(-f \cdot \alpha_k)$, $\beta_p \in M$ is homogeneous with $\deg_h(\beta_p) = \deg_h(\alpha_p)$. Therefore, $B = (\beta_j)_{j=1}^m$ is a homogeneous basis. \blacksquare

Observe that, as denoted above, $U^{-1}[\varphi] = U^{-1}[\varphi]_{A,S} = [\varphi]_{B,S}$, i.e. $U^{-1}[\varphi]$ is the matrix of φ relative to the bases $S = (\sigma_i(x))_{i=1}^n$ of N and $B = (\beta_j(x))_{j=1}^m$ of M . Since we have shown that B is a homogeneous basis, then Lemma 4.4.1 applies to $U^{-1}[\varphi] = [\varphi]_{B,S}$, i.e. for all nonzero entries $[\varphi]_{B,T}(j, i)$:

$$\deg_h(\beta_j(x)) + \deg_h([\varphi]_{B,S}(j, i)) = \deg_h(\sigma_i(x)).$$

As with the case for column reduction, we can do elimination by row reduction finitely many times and the resulting matrices will also preserve the homogeneity of the initial bases and that of its entries. Below, we

continue Example 4.4.5 and perform row reduction.

Example 4.4.8. We continue from Example 4.4.5. We currently have that $Q_5 = [\partial_1^{\text{Gr}}]V$ with

$$Q_5 = \begin{pmatrix} -x & 0 & -x^2 & 0 & 0 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \quad [\partial_1^{\text{Gr}}] = \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \quad V = \begin{pmatrix} 1 & 0 & 0 & -x & -x^2 \\ 0 & 1 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We perform the following row operations on Q_5 , as written below with the following color scheme:

chosen pivot entry , target entry for elimination , pivot multiplier , affected row

$$\begin{aligned} Q_6 &:= E_{\text{add}}^{[4]}(1, 4; x)Q_5 = \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x & 0 & -x^2 & 0 & 0 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} = \begin{pmatrix} -x & 0 & 0 & 0 & 0 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \\ Q_7 &:= E_{\text{add}}^{[4]}(2, 3; x)Q_6 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x & 0 & 0 & 0 & 0 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} = \begin{pmatrix} -x & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \\ Q_8 &:= E_{\text{add}}^{[4]}(1, 2; 1)Q_7 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \end{aligned}$$

Note that the elementary transvection for each row operation is multiplied on the left of Q_i . Let $U \in \text{GL}(4, \mathbb{F}[x])$ such that $U^{-1}Q_5 = U^{-1}[\partial_1^{\text{Gr}}]V$. Then, U^{-1} is given as follows:

$$U^{-1} = \overbrace{E_{\text{add}}^{[4]}(1, 2; 1)}^{\text{third row elimination}} \cdot \underbrace{E_{\text{add}}^{[4]}(2, 3; x)}_{\text{second row elimination}} \cdot \overbrace{E_{\text{add}}^{[4]}(1, 4; x)}^{\text{first row elimination}} = \begin{pmatrix} 1 & 1 & x & x \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Since the inverse of $E_{\text{add}}^{[m]}(a, b; f)$ is given by $E_{\text{add}}^{[m]}(a, b; -f)$, we can calculate U as follows:

$$\begin{aligned} U &= \left(E_{\text{add}}^{[4]}(1, 2; 1) E_{\text{add}}^{[4]}(2, 3; x) E_{\text{add}}^{[4]}(1, 4; x) \right)^{-1} = \overbrace{E_{\text{add}}^{[4]}(1, 4; -x)}^{\text{inverse corr. to first row elimination}} \underbrace{E_{\text{add}}^{[4]}(2, 3; -x)}_{\text{inverse corr. to second row elimination}} \overbrace{E_{\text{add}}^{[4]}(1, 2; -1)}^{\text{inverse corr. to third row elimination}} \\ &= \begin{pmatrix} a & -1 & 0 & -x \\ b & 0 & 1 & 0 \\ cx & 0 & 0 & 1 \\ dx & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Then, the basis $B = (\beta_1(x), \dots, \beta_4(x))$ of $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ given by $[\beta_j(x)] = \text{col}_j(U)$ is as follows:

$$\begin{aligned}\beta_1(x) &= (1)(a) = a \\ \beta_2(x) &= (-1)(a) + (1)(b) = -a + b \\ \beta_3(x) &= (-x)(b) + (1)(cx) = (-b + c)x \\ \beta_4(x) &= (-x)(a) + (1)(dx) = (-a + d)x\end{aligned}$$

Observe that B is a homogeneous basis and that the nonzero entries of Q_8 are as expected by Lemma 4.4.1.

The last two types of elementary matrices involved in matrix reduction are elementary permutations and elementary dilations. In the proposition below, we claim that these two kinds of matrices also preserve homogeneity of bases. Note that, unlike the case for elementary transvections in column and row reduction, the elementary permutations and dilations are not restricted by the entries of the matrix being reduced.

Proposition 4.4.9. Let N be a free graded $\mathbb{F}[x]$ -module with ordered homogeneous basis $S = (\sigma_1, \dots, \sigma_n)$. If $V \in \text{GL}(n, \mathbb{F}[x])$ is an elementary permutation or elementary dilation over $\mathbb{F}[x]$, then the basis $T = (\tau_1, \dots, \tau_n)$ of N by $[\tau_i]_S = \text{col}_i(V)$ is homogeneous.

Proof. We examine two cases. Assume that $V = E_{\text{swap}}^{[n]}(k_1, k_2)$ is an elementary permutation with distinct indices $k_1, k_2 \in \{1, \dots, n\}$. Note that if $k_1 = k_2$, then $V = I_n$ and T is trivially a homogeneous basis. For $i \in \{1, \dots, n\}$ with $i \neq k_1$ and $i \neq k_2$, $\tau_i = \sigma_i$. Since $\text{col}_{k_1}(V) = \text{col}_{k_2}(I_n)$, $\tau_{k_1} = \sigma_{k_2}$ is homogeneous. Similarly, $\text{col}_{k_2}(V) = \text{col}_{k_1}(I_n)$ and $\tau_{k_2} = \sigma_{k_1}$ is homogeneous. Therefore, T is a homogeneous basis.

Assume that $V = E_{\text{dilate}}^{[n]}(k, \mu)$ is an elementary dilation over $\mathbb{F}[x]$ with index $k \in \{1, \dots, n\}$ and $\mu \in (\mathbb{F}[x])^\times = \mathbb{F}[x] \setminus \{0\}$. For $i \in \{1, \dots, n\}$ with $i \neq k$, $\tau_i = \sigma_i$. Since $\text{col}_k(V) = \mu \cdot \text{col}_k(I_n)$ and μ is a nonzero homogeneous element of degree 0, $\tau_k = \mu \sigma_k$ and $\deg_h(\tau_k) = \deg_h(\mu \sigma_k) = \deg_h(\sigma_k)$. Therefore, T is a homogeneous basis. ■

Then, multiplication on the left (for row reduction) and on the right (for column reduction) by an elementary dilation or elementary permutation over $\mathbb{F}[x]$ will preserve the homogeneity of the initial bases and Lemma 4.4.1 applies to the resulting matrix.

Putting it all together, any reduction algorithm on a matrix over $\mathbb{F}[x]$ that satisfy Lemma 4.4.1 that involves a finite product of elementary permutations, elementary transvections, and elementary transvections that satisfy Proposition 4.4.4 or Proposition 4.4.7 over $\mathbb{F}[x]$ will produce a matrix that has zero or homogeneous elements and the bases they produce will also be homogeneous (assuming we start with homogeneous bases).

Section 4.5. An Ungraded SND Algorithm in the Graded Case

In this section, we present and discuss the algorithm for finding SNDs given [AW92, Remark 5.3.4] optimized for the case of matrices of graded homomorphisms relative to homogeneous bases.

Algorithm 4.5.1. Matrix Reduction Algorithm for Graded SNDs.

Input : $[\varphi] \in M_{m,n}(\mathbb{F}[x])$ as given in the hypothesis of Lemma 4.4.1, i.e. φ is a graded $\mathbb{F}[x]$ -module homomorphism relative to homogeneous bases.
Output : A Smith Normal Decomposition (U, D, V) of $[\varphi]$ with $U \in GL(m, \mathbb{F}[x])$, $V \in GL(n, \mathbb{F}[x])$ and $D \in M_{m,n}(\mathbb{F}[x])$.

```

1. set  $D_0 := [\varphi]$ .
2. for increasing indices  $k := 1, \dots, \min(m, n)$  do
3.   if  $D_{k-1}(j, i) = 0$  for all  $j \in \{k, \dots, m\}$  and  $i \in \{k, \dots, n\}$  then
4.     set  $D_{\min(m,n)} := D_{k-1}$ .
5.     for indices  $i := k, \dots, \min(m, n)$  do set  $U_i := I_m$  and  $V_i := I_n$ .
6.     break. ▷ stop the for-loop,  $D_{\min(m,n)}$  is in SNF.
7.   find a row index  $r_k \in \{k, \dots, m\}$  and a column index  $c_k \in \{k, \dots, m\}$  such that
      
$$\deg_h(D_{k-1}(r_k, c_k)) = \min \left\{ \deg_h(D_{k-1}(j, i)) : j \in \{k, \dots, m\}, i \in \{k, \dots, n\}, D_{k-1}(j, i) \neq 0 \right\}.$$

8.   set  $U_{k,k} := E_{\text{swap}}^{[m]}(k, r_k)$ ,  $V_{k,k} := E_{\text{swap}}^{[n]}(k, c_k)$  and  $W_k := U_{k,k} D_{k-1} V_{k,k}$ .
9.   for increasing column index  $i := k+1, \dots, n$  do
10.    set  $f_{k,i} := -W_k(k, i) / W_k(k, k)$  and  $V_{k,i} := E_{\text{add}}^{[n]}(k, i; f_{k,i})$ .
11.   set  $V_k := V_{k,k} V_{k,k+1} \cdots V_{k,n}$ .
12.   for increasing row index  $j := k+1, \dots, m$  do
13.    set  $g_{k,j} := -W_k(j, k) / W_k(k, k)$  and  $U_{k,j} := E_{\text{add}}^{[m]}(j, k; -g_{k,j})$ . ▷  $(U_{k,j})^{-1} = E_{\text{add}}^{[m]}(j, k; f_{k,k})$ 
14.   set  $U_k := U_{k,k} U_{k,k+1} \cdots U_{k,m}$ . ▷  $(U_k)^{-1} = (U_{k,m})^{-1} (U_{k,m-1})^{-1} \cdots (U_{k,k})^{-1}$ 
15.   set  $D_k := (U_k)^{-1} D_{k-1} V_k$ .
16. set  $U := U_1 U_2 \cdots U_{\min(m,n)}$ ,  $V := V_1 V_2 \cdots V_{\min(m,n)}$ .
17. set  $D := D_{\min(m,n)}$ .
18. return  $(U, D, V)$ 

```

Remarks. (1) This algorithm is written such that no variable is re-defined, e.g. for $[\varphi] \in M_{3,4}(\mathbb{F}[x])$, the matrix D_2 is defined once in Line 15 at the $k = 2$ loop and not re-defined in later loops. This is done to help with exposition.

(2) We refer to [AW92, Remark 5.3.4] for the correctness of Algorithm 4.5.1, i.e. that the result (U, D, V) is indeed an SND of $[\varphi]$ (as denoted above).

For the rest of this section, we discuss how the assumption of $[\varphi]$ being given as in Lemma 4.4.1 allows certain optimizations in the algorithm and how the matrix operations in the algorithm preserve the homogeneity of the initial matrices. In particular, we examine 4 major steps that are done for each k^{th} iteration of the for-loop starting in Line 2 of Algorithm 4.5.1:

Step A. Check if the matrix D_k is in Smith Normal Form. If not, continue.

Step B. Choose an appropriate pivot and perform the appropriate row and column permutations.

Step C. Eliminate the entries to the right of the pivot by column reduction.

Step D. Eliminate the entries below the pivot by row reduction.

We expand on what happens for each of these steps below.

Step A. Checking the Result of the Previous Iteration: Lines 3 to 7 of Algorithm 4.5.1.

The algorithm in [AW92, Remark 5.3.4] examines the submatrix $D_{k-1}[k:m, k:n]$ of D_{k-1} , where D_{k-1} is the result of the previous iteration, and $D_{k-1}[k:m, k:n]$ is obtained by removing columns $1, \dots, k-1$ and rows $1, \dots, k-1$ of D_{k-1} , i.e. D_{k-1} can be expressed as the following block matrix:

$$D_{k-1} = \begin{pmatrix} \text{diag}(d_1, \dots, d_{k-1}) & 0 \\ 0 & D_{k-1}[k:m, k:n] \end{pmatrix}$$

where d_1, \dots, d_{k-1} are to be the nonzero diagonal elements of the SNF D of $[\varphi]$. There are two cases:

Case 1: If the submatrix $D_{k-1}[k:m, k:n]$ is the zero matrix, then D_{k-1} is already in Smith Normal Form and we do not need to do any more matrix operations. That is, D_{k-1} is the Smith Normal Form of $[\varphi]$. Note that we set $D_{\min(m,n)} := D_{k-1}$ in Line 4, as opposed to defining $D := D_{k-1}$ right there, for consistency. The **break** keyword in Line 6 tells us to stop the for-loop and proceed directly to Line 16. Then, at Line 17, we define $D := D_{\min(m,n)}$.

Case 2: If the submatrix $D_{k-1}[k:m, k:n]$ is not the zero matrix, then D_{k-1} is not in Smith Normal Form. The goal of the k^{th} iteration is then to find matrices $U_k \in \text{GL}(m, \mathbb{F}[x])$ and $V_k \in \text{GL}(n, \mathbb{F}[x])$ such that

$$(U_k)^{-1} D_{k-1} V_k = D_k = \begin{pmatrix} \text{diag}(d_1, \dots, d_{k-1}, d_k) & 0 \\ 0 & D_k[k+1:m, k+1:n] \end{pmatrix}$$

where D_k is the end result of this iteration and passed to the next, i.e. the matrix D_k determines the k^{th} diagonal element of the Smith Normal Form of D . We apply the appropriate elementary matrices on D_{k-1} to achieve this.

Below, we provide an example of a matrix on which Algorithm 4.5.1 stops at some k^{th} iteration and another where it does not and continues.

Example 4.5.2. Let $A \in \text{M}_{4,4}(\mathbb{F}[x])$ and $B \in \text{M}_{4,4}(\mathbb{F}[x])$ be given below:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5x & 2 & 0 \\ 0 & x^2 & 3x & 3 \\ 0 & 4 & -2x & 2x^2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We can view the matrix A to be the matrix D_1 relative to Algorithm 4.5.1, i.e. this is the matrix that is processed in the iteration with $k = 2$. Then, A is not in Smith Normal Form and we consider the submatrix $A[2:4, 2:4]$, highlighted in red below:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5x & 2 & 0 \\ 0 & x^2 & 3x & 3 \\ 0 & 4 & -2x & 2x^2 \end{pmatrix} \quad \text{with} \quad A[2:4, 2:4] = \begin{pmatrix} 5x & 2 & 0 \\ x^2 & 3x & 3 \\ 4 & -2x & 2x^2 \end{pmatrix}$$

Since $A[2:4, 2:4]$ is not the zero matrix, we proceed with the calculation.

In contrast, we can assume that the matrix B corresponds to the matrix D_2 relative to Algorithm 4.5.1, i.e. the matrix processed in the iteration with $k = 3$. Then, Algorithm 4.5.1 concludes that B is in Smith Normal

Form since the submatrix $B[3:4, 3:4]$, highlighted in red below, is the zero matrix.

$$B = D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then, the for-loop stops here and sets $D_{\min(m,n)} = D_2 = B$ in Line 4 and $D = B$ in Line 17 of Algorithm 4.5.1.

Step B. On Selection of Pivot Elements: Lines 7-8 of Algorithm 4.5.1

The next step involves finding an element in the submatrix $D_{k-1}[k:m, k:n]$ that can be used to eliminate any nonzero entry in $D_{k-1}[k:m, k:n]$ as necessary. For convenience, we call this element the *pivot*, consistent with the terminology for our reduction operations by Definition 4.4.3 and Definition 4.4.6.

This pivot, given in Line 7 of Algorithm 4.5.1 by $D_{k-1}(r_k, c_k)$, is an element that has minimal degree across the nonzero entries of D_{k-1} , where the term *degree* refers to the Euclidean function $\deg(-)$ on a Euclidean domain. Since the pivot is chosen to be of minimal degree, it can be shown that any nonzero element of $D_{k-1}[k:m, k:n]$ can be eliminated by a finite number of elementary row/column operations. Note that if the matrix $[\varphi]$ is over a field \mathbb{F} , then any nonzero element can serve as the pivot since $\deg(\mu) = 1$ for all nonzero $\mu \in \mathbb{F}$.

Observe that, by Lemma 4.4.1, the nonzero entries of $[\varphi]$ are homogeneous and are of the form kx^t for some $k \in \mathbb{F}$ and $t \in \mathbb{N}_0$. Then, the degree $\deg(-)$ on $\mathbb{F}[x]$ as a Euclidean domain and the degree $\deg_h(-)$ on $\mathbb{F}[x]$ as a graded ring agree on the entries of the matrices involved. Moreover, if the chosen pivot is $kx^s \in \mathbb{F}[x]$ for some $k \in \mathbb{F}$ and $s \in \mathbb{N}_0$, the pivot multiplier required to eliminate a nonzero entry $k'x^t$ can immediately be determined by $k'x^t / kx^s = (k'/k)x^{t-s}$ since $t \leq s$ by minimality of the degree of the chosen pivot. This is reflected in the Step C involving column reduction and Step D involving row reduction.

We then apply appropriate permutations on D_{k-1} such that the pivot $D_{k-1}(r_k, c_k)$ is found on the $(k, k)^{\text{th}}$ entry. Relative to the submatrix $D_{k-1}[k:m, k:n]$, we want the pivot to be the $(1, 1)^{\text{th}}$ entry post-permutation. The row permutation is given by $U_{k,k} := E_{\text{swap}}^{[m]}(k, r_k)$, which swaps rows k and r_k of D_{k-1} in the product $(U_{k,k})^{-1}D_{k-1}$. Note that $(E_{\text{swap}}^{[m]}(k, r_k))^{-1} = E_{\text{swap}}^{[m]}(k, r_k)$. The column permutation is given by $V_{k,k} := E_{\text{swap}}^{[n]}(k, c_k)$, which swaps columns k and c_k of $U_{k,k}D_{k-1}$ in the product $U_{k,k}D_{k-1}V_{k,k}$. We save this permuted matrix as W_k and perform the matrix operations on W_k .

Note that, by Proposition 4.4.9 (permutations preserve homogeneity), the homogeneity of the entries of W_k and of the bases induced by the multiplication of the permutation matrices $U_{k,k} = E_{\text{swap}}^{[m]}(k, r_k)$ and $V_{k,k} = E_{\text{swap}}^{[n]}(k, c_k)$ are preserved. We provide an example of this calculation below.

Example 4.5.3. Let $[\partial_1^{\text{Gr}}] \in M_{4,5}(\mathbb{F}[x])$ be as given in Example 4.4.5. For convenience, the description of $[\partial_1^{\text{Gr}}]$ is provided below. We apply Algorithm 4.5.1 on $[\partial_1^{\text{Gr}}]$. Following the notation in the algorithm, we have that $D_0 := [\partial_1^{\text{Gr}}]$. At the $k = 1$ iteration, the entry $D_0(3, 2)$, highlighted in blue below, can serve as the pivot.

$$D_0 := [\partial_1^{\text{Gr}}] = \begin{matrix} & \begin{matrix} abx & bcx & adx^2 & cdx^2 & acx^3 \end{matrix} \\ \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} & \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \end{matrix}$$

Then, we permute the rows of D_0 by $U_{1,1} = E_{\text{swap}}^{[4]}(1, r_1)$ with $r_1 := 3$ and columns of D_0 by $V_{1,1} = E_{\text{swap}}^{[5]}(1, c_1)$

with $c_1 := 2$ as follows. Highlighted in **green** are the rows and columns of the **pivot** $D_0(r_1, c_1) = 1$ and in **orange** are those of the **entry** $D_0(1, 1) = -x$.

$$\begin{aligned}
 U_{1,1}D_0V_{1,1} &= E_{\text{swap}}^{[4]}(1, 3)D_0E_{\text{swap}}^{[5]}(1, 2) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 & x & x^2 \\ -x & x & 0 & 0 & 0 \\ 0 & -x & -x^2 & 0 & -x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} =: W_1
 \end{aligned}$$

with $W_1 \in M_{4,5}(\mathbb{F}[x])$ as denoted in Algorithm 4.5.1.

Step C. On Elimination by Column Reduction: Lines 9-11 of Algorithm 4.5.1

The matrices $V_{k,k+1}, V_{k,k+2}, \dots, V_{k,n} \in \text{GL}(n, \mathbb{F}[x])$ correspond to column operations that eliminate the entries to the left of the pivot $W_k(k, k)$. Recall that any nonzero homogeneous element of $\mathbb{F}[x]$ divides any nonzero homogeneous element of equal or greater degree. Since the pivot $W_k(k, k)$ is of minimal degree by construction, we can eliminate any nonzero entry $W_k(k, i)$ to the left of $W_k(k, k)$ with $i \in \{k+1, \dots, n\}$ by addition of $f_{k,i} \cdot W_k(k, k)$ with $f_{k,i} \in \mathbb{F}[x]$ homogeneous.

Note that since none of the transvections $V_{k,i} := E_{\text{add}}^{[n]}(k, i; f_{k,i})$ act on the k^{th} columns of W_k and all of them add a $f_{k,i}$ -multiple of the k^{th} column of W_k , the order in which $V_{k,i}$ is multiplied to W_k does not matter for the definition of $V_{k,i}$. In particular, $\text{col}_k(W_k) = \text{col}_k(W_k V_{k,i})$ for any $i \in \{k+1, \dots, n\}$. Therefore, we can define $f_{k,i}$ for each transvection $V_{k,i}$ relative to W_k , as done in Line 10, as opposed to a matrix product like $W_k V_{k,k+1} \cdots V_{k,i-1}$ that accounts for matrix operations that are already done.

Observe that if $W_k(k, i) = 0$, then there is nothing to eliminate. In that case, $f_{k,i} = 0$, $E_{\text{add}}^{[n]}(k, i; 0) = I_n$, and homogeneity is trivially preserved. However, if $W_k(k, i) \neq 0$, then $f_{k,i}$ must be homogeneous with

$$\deg_{\text{h}}(f_{k,i}) + \deg_{\text{h}}(W_k(k, k)) = \deg_{\text{h}}(W_k(k, i)).$$

Then, Proposition 4.4.4 (column reduction preserves homogeneity) applies and these column reduction operations preserve the homogeneity of the bases and of the entries. We provide an example below.

Example 4.5.4. Continue Example 4.5.3. For reference, a description of $W_1 \in M_{4,5}(\mathbb{Q}[x])$ is given below, with the **pivot entry** $W_1(1, 1)$ highlighted in **blue** and the **target elements**, i.e. the elements to the left of the pivot, highlighted in **red**.

$$W_1 = \begin{pmatrix} 1 & 0 & 0 & x & x^2 \\ -x & x & 0 & 0 & 0 \\ 0 & -x & -x^2 & 0 & -x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix}$$

Below, we define the elementary transvections $E_{\text{add}}^{[5]}(1, i; f_{1,i})$ for $i \in \{2, \dots, 5\}$ used to eliminate the entry $W_1(1, i)$ with the **pivot multiplier** $f_{1,i}$ highlighted in **orange**.

$$\begin{aligned}
i = 2 : \quad f_{1,2} &= -\frac{W_1(1,2)}{W_1(1,1)} = 0 & \text{and} \quad V_{1,2} &:= E_{\text{add}}^{[5]}(1, 2; 0) = I_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
i = 3 : \quad f_{1,3} &= -\frac{W_1(1,3)}{W_1(1,1)} = 0 & \text{and} \quad V_{1,3} &:= E_{\text{add}}^{[5]}(1, 3; 0) = I_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
i = 4 : \quad f_{1,4} &= -\frac{W_1(1,4)}{W_1(1,1)} = -x & \text{and} \quad V_{1,4} &:= E_{\text{add}}^{[5]}(1, 3; -x) = \begin{pmatrix} 1 & 0 & 0 & -x & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
i = 5 : \quad f_{1,5} &= -\frac{W_1(1,5)}{W_1(1,1)} = -x^2 & \text{and} \quad V_{1,4} &:= E_{\text{add}}^{[5]}(1, 3; -x^2) = \begin{pmatrix} 1 & 0 & 0 & 0 & -x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

As a sidenote, we can calculate the matrix product $(V_{1,2})(V_{1,3})(V_{1,4})(V_{1,5})$ or any re-ordering of the product without explicitly identifying each $E_{\text{add}}^{[5]}(1, i; f_{1,i})$ as follows:

$$(V_{1,2})(V_{1,3})(V_{1,4})(V_{1,5}) = \begin{pmatrix} 1 & f_{1,2} & f_{1,3} & f_{1,4} & f_{1,5} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -x & -x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then, the column reduction step in the $k = 1$ loop produces the following matrix, with the columns highlighted in orange being the columns affected by the column reduction.

$$(W_1)(V_{1,2}) \cdots (V_{1,5}) = \begin{pmatrix} 1 & 0 & 0 & x & x^2 \\ -x & x & 0 & 0 & 0 \\ 0 & -x & -x^2 & 0 & -x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -x & -x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -x & x & 0 & x^2 & x^3 \\ 0 & -x & -x^2 & 0 & -x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix}$$

Note that the definition of V_1 by Line 11 of Algorithm 4.5.1 includes the column permutation in its definition. Note that since $V_{1,1} = E_{\text{swap}}^{[5]}(1, 2)$ is multiplied on the left of $(V_{1,2}) \cdots (V_{1,5})$, it permutes the rows of $(V_{1,2}) \cdots (V_{1,5})$ instead. Highlighted below are Row 1 and Row 2 of $(V_{1,2}) \cdots (V_{1,5})$ and where they are mapped to after the row permutation.

$$V_1 = (V_{1,1})(V_{1,2})(V_{1,3})(V_{1,4})(V_{1,5}) = E_{\text{swap}}^{[5]}(1, 2) \begin{pmatrix} 1 & 0 & 0 & -x & -x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The matrix $V_1 \in \text{GL}(5, \mathbb{Q}[x])$ induces a basis $T = (\tau_1(x), \dots, \tau_5(x))$ of $C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ given as follows:

$$\begin{aligned}\tau_1(x) &= (1)(bcx) &= (bc)x \\ \tau_2(x) &= (1)(abx) &= (ab)x \\ \tau_3(x) &= (1)(adx^2) &= (ad)x^2 \\ \tau_4(x) &= (-x)(bcx) + (1)(cdx^2) &= (-bc + cd)x^2 \\ \tau_5(x) &= (-x^2)(bcx) + (acx^3) &= (-bc + ac)x^3\end{aligned}$$

Observe that the basis $T = (\tau_i(x))$ is a homogeneous basis.

Step D. On Elimination by Row Reduction: Lines 12-14 of Algorithm 4.5.1

The matrices $U_{k,k+1}, U_{k,k+2}, \dots, U_{k,m} \in \text{GL}(m, \mathbb{F}[x])$ correspond to row operations that eliminate the entries under the pivot entry $W_k(k, k)$. As with column reduction, since the elements of W_k are homogeneous by Lemma 4.4.1 and the pivot $W_k(k, k)$ is chosen to be of minimal degree, each entry $W_k(j, k)$ under the pivot can be eliminated by the addition of $g_{k,j}W_k(k, k)$ with $g_{k,j} \in \mathbb{F}[x]$ homogeneous and $j \in \{k+1, \dots, m\}$.

As mentioned on Item 4, the column operations in Lines 9-11 of Algorithm 4.5.1 do not affect the k^{th} column of W_k . Then, the entries below the pivot $W_k(k, k)$, i.e. entries indexed by (j, k) with $j \in \{k+1, \dots, m\}$, are unaffected by the column reduction. Therefore, we can define the transvections $U_{k,j}$ that eliminate the $(j, k)^{\text{th}}$ entry of W_k relative to W_k , as opposed to the matrix $W_k V_k$ post-column reduction. Furthermore, all of the row operations by the transvection $(U_{k,j})^{-1} := E_{\text{add}}^{[m]}(j, k; g_{k,j})$ also do not affect the k^{th} row of W_k . Therefore, we can define the elements $g_{k,j}$ such that $W_k(j, k) + g_{k,j} \cdot W_k(k, k) = 0$ relative to W_k , as done in Line 13 of Algorithm 4.5.1. As a sidenote, the order in which the row operations $E_{\text{add}}^{[m]}(j, k; g_{k,j})$ with $j \in \{k+1, \dots, m\}$ and the column operations $E_{\text{add}}^{[n]}(k, i; f_{k,i})$ with $i \in \{k+1, \dots, n\}$ do not affect the definition of $g_{k,j}$ and $f_{k,i}$.

Since we have labeled the matrix $U_k \in \text{GL}(m, \mathbb{F})$ in Item (1) such that $(U_k)^{-1}D_{k-1}V_k = D_k$, i.e. $(U_k)^{-1}$ more closely describes the row operations done on $W_k = U_{k,k}D_{k-1}V_{k,k}$, we will have to take the inverse of the matrix product $(U_{k,m})^{-1}(U_{k,m-1})^{-1} \dots (U_{k,k+1})^{-1}$ resulting from the row operations done on D_{k-1} if we want an expression for U_k . To simplify this calculation, we can apply the property that $(AB)^{-1} = B^{-1}A^{-1}$ for any pair of invertible matrices $A, B \in \text{GL}(m, \mathbb{F}[x])$. Then, the matrix reduction on D_{k-1} is as follows:

$$\begin{aligned}D_k &= \overbrace{(U_{k,m})^{-1}(U_{k,m-1})^{-1} \dots (U_{k,k+1})^{-1}}^{\text{row reduction on } W_k} \overbrace{(U_{k,k})^{-1} (D_{k-1}) (V_{k,k})}^{W_k} \overbrace{(V_{k,k+1})(V_{k,k+2}) \dots (V_{k,n})}^{\text{column reduction on } W_k} \\ &= \left(\underbrace{(U_{k,k})(U_{k,k+1}) \dots (U_{k,m-1})(U_{k,m})}_{U_k} \right)^{-1} (D_{k-1}) \underbrace{(V_{k,k})(V_{k,k+1})(V_{k,k+2}) \dots (V_{k,n})}_{V_k}\end{aligned}$$

Observe that if $W_k(j, k) = 0$, then there is nothing to eliminate and $E_{\text{add}}^{[m]}(j, k; 0) = I_m$ trivially preserves homogeneity. If $W_k(j, k) \neq 0$, $g_{k,j} \in \mathbb{F}[x]$ must be homogeneous and satisfy the following equation:

$$\deg_h(g_{k,j}) + \deg_h(W_k(k, k)) = \deg_h(W_k(j, k))$$

Then, Proposition 4.4.7 (row reduction preserves homogeneity) applies and these row operations preserve the homogeneity of the bases and of the entries. We continue the calculation of the SNF of $[\partial_1^{\text{Gr}}]$ below.

Example 4.5.5. Continue Example 4.5.4. Listed below is the matrix W_1 , with the pivot entry $W_1(1, 1)$

highlighted in blue and the target entries, i.e. the entries below the chosen pivot, in red.

$$W_1 = \begin{pmatrix} \textcolor{blue}{1} & 0 & 0 & x & x^2 \\ \textcolor{red}{-x} & x & 0 & 0 & 0 \\ \textcolor{red}{0} & -x & -x^2 & 0 & -x^3 \\ \textcolor{red}{0} & 0 & x & -x & 0 \end{pmatrix}$$

Below, we define the elementary transvections $E_{\text{add}}^{[4]}(j, i; f_{1,i})$ for $j \in \{2, \dots, 4\}$ used to eliminate the entry $W_1(j, 1)$, with the pivot multiplier $g_{1,j}$ highlighted in orange.

$$\begin{aligned} j=2 : \quad g_{1,2} &= -\frac{W_1(2,1)}{W_1(1,1)} = x \quad \text{and} \quad (U_{1,2})^{-1} := E_{\text{add}}^{[4]}(2, 1; \textcolor{orange}{x}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \textcolor{orange}{x} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ j=3 : \quad g_{1,3} &= -\frac{W_1(3,1)}{W_1(1,1)} = 0 \quad \text{and} \quad (U_{1,3})^{-1} := E_{\text{add}}^{[4]}(3, 1; \textcolor{orange}{0}) = I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \textcolor{orange}{0} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ j=4 : \quad g_{1,4} &= -\frac{W_1(4,1)}{W_1(1,1)} = 0 \quad \text{and} \quad (U_{1,4})^{-1} := E_{\text{add}}^{[4]}(4, 1; \textcolor{orange}{0}) = I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \textcolor{orange}{0} & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Like in the case of $(V_{1,2}) \cdots (V_{1,5})$, the matrix product $(U_{1,4})^{-1}(U_{1,3})^{-1}(U_{1,2})^{-1} = (U_{1,2}U_{1,3}U_{1,4})^{-1}$ can be identified explicitly without identifying each matrix $U_{1,j}$ as follows:

$$(U_{1,4})^{-1}(U_{1,3})^{-1}(U_{1,2})^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \textcolor{orange}{g_{1,2}} & 1 & 0 & 0 \\ \textcolor{orange}{g_{1,3}} & 0 & 1 & 0 \\ \textcolor{orange}{g_{1,4}} & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \textcolor{orange}{x} & 1 & 0 & 0 \\ \textcolor{orange}{0} & 0 & 1 & 0 \\ \textcolor{orange}{0} & 0 & 0 & 1 \end{pmatrix}$$

Since the inverses of elementary transvections are known, we can also calculate the product $(U_{1,2})(U_{1,3})(U_{1,4})$, i.e. the inverse of the matrix calculated above, directly as follows:

$$(U_{1,2})(U_{1,3})(U_{1,4}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \textcolor{orange}{-g_{1,2}} & 1 & 0 & 0 \\ \textcolor{orange}{-g_{1,3}} & 0 & 1 & 0 \\ \textcolor{orange}{-g_{1,4}} & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \textcolor{orange}{-x} & 1 & 0 & 0 \\ \textcolor{orange}{0} & 0 & 1 & 0 \\ \textcolor{orange}{0} & 0 & 0 & 1 \end{pmatrix}$$

To calculate $U_1 := (U_{1,1})(U_{1,2})(U_{1,3})(U_{1,4})$ with $U_{1,1} = E_{\text{swap}}^{[4]}(1, 3)$ from Example 4.5.5, we can permute the rows of the product $(U_{1,2})(U_{1,3})(U_{1,4})$ as follows:

$$U_1 = (U_{1,1})(U_{1,2})(U_{1,3})(U_{1,4}) = E_{\text{swap}}^{[4]}(1, 3) \begin{pmatrix} \textcolor{lightgreen}{1} & 0 & 0 & 0 \\ \textcolor{orange}{-x} & 1 & 0 & 0 \\ \textcolor{orange}{0} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} \begin{pmatrix} \textcolor{orange}{0} & 0 & 1 & 0 \\ \textcolor{orange}{-x} & 1 & 0 & 0 \\ \textcolor{lightgreen}{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The matrix $U_1 \in \text{GL}(4, \mathbb{Q}[x])$ determines a basis $B = (\beta_1(x), \dots, \beta_4(x))$ of $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q}[x])$ by $[\beta_j(x)] = \text{col}_j(U_1)$ as given below:

$$\begin{aligned} \beta_1(x) &= (-x)(b) + (1)(cx) = (-b + c)x \\ \beta_2(x) &= (1)(b) = b \\ \beta_3(x) &= (1)(a) = a \\ \beta_4(x) &= (1)(dx) = (d)x \end{aligned}$$

We can confirm by direct calculation that the entries to the right and below the pivot $W_1(1,1)$ in the product $(U_1)^{-1}D_0V_1 =: D_1$ are zero, highlighted in red below:

$$\begin{aligned}
 (U_1)^{-1}D_0V_1 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ x & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & x & x^2 \\ -x & x & 0 & 0 & 0 \\ 0 & -x & -x^2 & 0 & -x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & x^2 & x^3 \\ 0 & -x & -x^2 & 0 & x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} =: D_1
 \end{aligned}$$

Final Step. Aggregating the Results: Lines 16-17 of Algorithm 4.5.1

Observe that in the k^{th} iteration, the zero entries found below and to the right of the entries d_1, \dots, d_{k-1} of D_{k-1} with $d_i = D_{k-1}(i, i)$ remain zero on D_k , i.e. the permutation, column reduction, and row reduction operations done on D_{k-1} leave columns $1, \dots, k-1$ and rows $1, \dots, k-1$ of D_{k-1} undisturbed. Then, at each iteration with $k \in \{1, \dots, \min(m, n)\}$, the matrix $[\varphi]$ gets cleared row by row and column by column like so:

$$\begin{aligned}
 D_0 := [\varphi] &\rightsquigarrow D_1 := (U_1)^{-1}D_0V_1 && D_2 := (U_2)^{-1}D_1V_2 && \rightsquigarrow \dots \\
 \begin{pmatrix} * & * & \dots & * & * & \dots & * \\ * & * & \dots & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * & * & \dots & * \\ * & * & \dots & * & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * & * & \dots & * \end{pmatrix} && \rightsquigarrow \begin{pmatrix} d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & * & \dots & * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * & * & \dots & * \\ 0 & * & \dots & * & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * & * & \dots & * \end{pmatrix} && \rightsquigarrow \begin{pmatrix} d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & * & * & \dots & * \\ 0 & 0 & \dots & * & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & * & * & \dots & * \end{pmatrix} && \rightsquigarrow \dots \\
 \dots &\rightsquigarrow D_{k-1} := (U_{k-1})^{-1}D_{k-2}V_{k-1} && D_k := (U_k)^{-1}D_{k-1}V_k && \rightsquigarrow \left(\text{Repeat until SNF} \right. \\
 \begin{pmatrix} d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{k-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & * & \dots & * \end{pmatrix} && \rightsquigarrow \begin{pmatrix} d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_{k-1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & d_k & \dots & * \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & * & \dots & * \end{pmatrix} && \rightsquigarrow \left(\text{Repeat until SNF} \right. \\
 && && && \text{is achieved} && \left. \right)
 \end{aligned}$$

We also get the matrices $U \in \text{GL}(m, \mathbb{F}[x])$ and $V \in \text{GL}(n, \mathbb{F})$ by unfolding the matrix reduction as follows:

$$\begin{aligned}
 D &:= D_{\min(m,n)} \\
 &= (U_{\min(m,n)})^{-1} (D_{\min(m,n)-1}) (V_{\min(m,n)}) \\
 &\quad \vdots \\
 &= (U_{\min(m,n)})^{-1} \dots (U_{k+1})^{-1} (U_k)^{-1} (D_{k-1}) (V_k)(V_{k+1}) \dots (V_{\min(m,n)}) \\
 &\quad \vdots \\
 &= (U_{\min(m,n)})^{-1} \dots (U_k)^{-1} \dots (U_2)^{-1} (U_1)^{-1} D_0 (V_1)(V_2) \dots (V_k) \dots (V_{\min(m,n)})
 \end{aligned}$$

$$= \underbrace{\left((U_1)(U_2) \cdots (U_k) \cdots (U_{\min(m,n)}) \right)^{-1}}_U [\varphi] \underbrace{(V_1)(V_2) \cdots (V_k) \cdots (V_{\min(m,n)})}_V$$

Below, we finish the calculation on $[\partial_1^{\text{Gr}}]$ started on Example 4.5.3.

Example 4.5.6. Continue from Example 4.5.5. In this example, we finish the calculation of an SND (U, D, V) of $[\partial_1^{\text{Gr}}]$ by Algorithm 4.5.1. For reference, $[\varphi] =: D_0 \in M_{4,5}(\mathbb{Q}[x])$ is given below.

$$D_0 := [\partial_1^{\text{Gr}}] = \begin{matrix} & abx & bcx & adx^2 & cdx^2 & acx^3 \\ \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} & \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \end{matrix}$$

From Example 4.5.5, we have calculated the matrices $U_1 \in \text{GL}(4, \mathbb{Q}[x])$, $V_1 \in \text{GL}(5, \mathbb{Q}[x])$, and $D_1 \in M_{4,5}(\mathbb{Q}[x])$ as labeled in Algorithm 4.5.1. For convenience, we copied the results below.

$$U_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -x & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & x^2 & x^3 \\ 0 & -x & -x^2 & 0 & x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \quad V_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

For the $k = 2$ loop of Algorithm 4.5.1:

We perform matrix operations on D_1 . For brevity, we provide the matrices $U_2 \in \text{GL}(4, \mathbb{Q}[x])$ and $V_2 \in \text{GL}(5, \mathbb{Q}[x])$ that account for all row and column operations on D_1 respectively. Note that since $D_1(2, 2) = x$ is already of minimal degree across the nonzero entries $D_1(j, i)$ with $j \in \{2, \dots, 4\}$ and $i \in \{2, \dots, 5\}$, $D_1(2, 2)$ can serve as the pivot and we do not need to permute D_1 . Highlighted in blue is the pivot $D_1(2, 2) = x$, in red are the entries to be eliminated by row reduction, and in orange are those to be eliminated by column operations.

$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & x^2 & x^3 \\ 0 & -x & -x^2 & 0 & x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix}$$

We set $U_2 \in \text{GL}(4, \mathbb{Q}[x])$ for the row operations and $V_2 \in \text{GL}(5, \mathbb{Q})$ for the column operations as follows, with their respective pivot multipliers highlighted in orange and red respectively.

$$(U_2)^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{with} \quad U_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad V_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

For the $k = 3$ loop of Algorithm 4.5.1:

Given below is $D_2 := (U_2)^{-1} D_1 V_2$, with the chosen pivot entry $D_1(4, 3) = x$ highlighted in blue. Since

the pivot is not given by the $(3, 3)^{\text{th}}$ entry of D_2 , we need to permute the rows and columns of D_2 .

$$D_2 := (U_2)^{-1} D_1 V_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & x^2 & x^3 \\ 0 & -x & -x^2 & 0 & x^3 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & -x^2 & x^2 & 0 \\ 0 & 0 & x & -x & 0 \\ 0 & 0 & -x^2 & x^2 & 0 \end{pmatrix}$$

We have the following calculations for $U_3 \in \text{GL}(4, \mathbb{Q}[x])$ and $V_3 \in \text{GL}(5, \mathbb{Q}[x])$ below. Highlighted in green are the entries and rows related to row permutations, those in red for column reduction, and those in orange for row reduction.

$$\begin{aligned} W_3 &:= E_{\text{swap}}^{[4]}(3, 4) D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & -x^2 & x^2 & 0 \\ 0 & 0 & x & -x & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x & -x & 0 \\ 0 & 0 & -x^2 & x^2 & 0 \end{pmatrix} \\ D_3 &:= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & x & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x & -x & 0 \\ 0 & 0 & -x^2 & x^2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ (U_3)^{-1} &:= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & x & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & x \end{pmatrix} \\ U_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & x & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & x & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & x & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -x & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -x & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ V_3 &:= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

For the $k = 4$ loop of Algorithm 4.5.1:

Observe that D_3 is already in Smith Normal Form. Then, Algorithm 4.5.1 stops and returns the matrices U and V as denoted below and the matrix $D := D_3$ as the Smith Normal Form of $D_0 = [\partial_1^{\text{Gr}}]$.

$$\begin{aligned} U &= (U_1)(U_2)(U_3) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -x & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -x & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -1 & -x & 1 \\ b & -x & 1 & 0 \\ cx & 1 & 0 & 0 \\ dx & 0 & 0 & 1 \end{pmatrix} \\ V &= (V_1)(V_2)(V_3) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} abx & 0 & 1 & 0 & -x \\ bcx & 1 & 0 & 0 & -x \\ adx^2 & 0 & 0 & 1 & 1 \\ cdx^2 & 0 & 0 & 0 & 1 \\ acx^3 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

We can confirm that $U^{-1}[\partial_1^{\text{Gr}}]V = D$ by the following calculation:

$$U^{-1}[\partial_1^{\text{Gr}}]V = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & x & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & x & 0 \end{pmatrix} \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & x & x^2 \\ 0 & 0 & x & -x & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & -x & -x^2 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = D$$

Then, the matrices $U \in \text{GL}(4, \mathbb{Q}[x])$ and $V \in \text{GL}(5, \mathbb{Q}[x])$ induce bases $B = (\beta_1(x), \dots, \beta_4(x))$ of $C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ by $[\beta_j] = \text{col}_j(U)$ and $T = (\tau_1(x), \dots, \tau_5(x))$ of $C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ by $[\tau_i] = \text{col}_i(V)$ as follows:

$$\begin{aligned} \beta_1(x) &= (-x)(b) + (1)(cx) = (-b + c)x \\ \beta_2(x) &= (-1)(a) + (1)(b) = (-a + b) \\ \beta_3(x) &= (-x)(a) + (1)(dx) = (-a + d)x \\ \beta_4(x) &= (1)(a) = (a) \end{aligned}$$

$$\begin{aligned} \tau_1(x) &= (1)(bcx) = (bc)x \\ \tau_2(x) &= (1)(abx) = (ab)x \\ \tau_3(x) &= (1)(adx^2) = (ad)x^2 \\ \tau_4(x) &= (-x)(abx) + (-x)(bcx) + (1)(adx^2) + (1)(cdx^2) = (-ab + bc + ad + cd)x^2 \\ \tau_5(x) &= (-x^2)(abx) + (-x^2)(bcx) + (1)(acx^3) = (-ab + bc + ac)x^3 \end{aligned}$$

Observe that both B and T are homogeneous bases and that the nonzero elements $1, x, x \in \mathbb{Q}[x]$ of the SNF D are also homogeneous.

With Algorithm 4.5.1 established, we can now state the existence claim for the Graded Structure Theorem (Theorem 4.3.1) in relation to the SNDs compatible with Proposition 4.3.5.

Proposition 4.5.7. Let M be a finitely generated graded $\mathbb{F}[x]$ -module with graded presentation given by $\varphi : F_S \rightarrow F_G$ and $\pi : F_G \rightarrow M$ with homogeneous bases $S = (\sigma_1(x), \dots, \sigma_n(x))$ of F_S and $A = (\alpha_1(x), \dots, \alpha_m(x))$ of F_G .

Then, there exists an SND (U, D, V) of $[\varphi]_{A,S}$ such that the basis $B = (\beta_1(x), \dots, \beta_m(x))$ of F_G by $[\beta_j(x)] = \text{col}_j(U)$ is a homogeneous basis and the nonzero elements $d_1(x), \dots, d_r(x)$ of the SNF D by $d_j(x) = D(j, j)$ are also homogeneous. Furthermore, the basis $B = \{\beta_j(x)\}_{j=1}^m$ of F_G and the nonzero elements $\{d_j(x)\}_{j=1}^r$ satisfy the hypothesis of Proposition 4.3.5 and can be used to calculate the graded invariant factor decomposition of M .

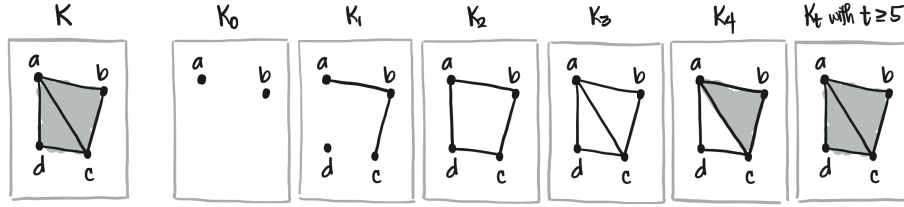
Proof. Since $[\varphi]_{A,S}$ satisfies Lemma 4.4.1, Algorithm 4.5.1 will yield an SND (U, D, V) with the required properties. Note that Proposition 4.4.4 states that the properties given by Lemma 4.4.1 are preserved in the column reduction operations done by Algorithm 4.5.1, Proposition 4.4.7 states this for row reduction, and Proposition 4.4.9 states this for the permutations and dilations.

The divisibility relation $d_1(x) \mid \cdots \mid d_r(x)$ is guaranteed by how the pivots in each iteration in Algorithm 4.5.1 are chosen, i.e. $\deg_h(d_1(x)) \leq \deg_h(d_2(x)) \leq \cdots \leq \deg_h(d_r(x))$ since each $d_r(x)$ is chosen to be of minimal degree. Note that while $d_j(x)$ is generally not in the form x^{t_j} with $\deg_h(d_j(x)) = t_j$ as required by Proposition 4.3.5, there exists a nonzero $\mu_j \in \mathbb{F}$ such that $\mu \cdot d_j(x) = x^{t_j}$. Then, the SNF D of $[\varphi]$ can be multiplied by appropriate elementary dilations over $\mathbb{F}[x]$ to have $d_j(x) = x^{t_j}$ for all $j \in \{1, \dots, r\}$. ■

While we have not strictly established this for arbitrary graded chain complexes of free graded $\mathbb{F}[x]$ -modules, we can continue the calculation on $[\partial_1^{\text{Gr}}]$ on Example 4.5.6 and determine the 0^{th} persistent homology module

of the filtration K_\bullet .

Example 4.5.8. Let K_\bullet and K be given as in Example 3.2.3 and orient K with the vertex order $\text{Vert}(K) = (a, b, c, d)$. For convenience, an illustration of K and K_\bullet (without orientation) is copied below:



We claim that the following sequence is a graded presentation of the 0^{th} graded persistent homology module $H_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ of K_\bullet with coefficients in \mathbb{Q} :

$$C_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) \xrightarrow{\partial_1^{\text{Gr}}} C_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) \xrightarrow{\pi} H_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) \longrightarrow 0$$

where π refers to the canonical quotient map. The justification for this is similar to the non-graded case by Proposition 4.2.1 with $\ker(\partial_0^{\text{Gr}}) = C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$.

An SND (U, D, V) of $[\partial_1^{\text{Gr}}]$ in the graded presentation of $H_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$, as calculated in Example 4.5.6, is given below. The columns of U and the diagonal entries of D highlighted in blue correspond to a basis of $\text{im}(\varphi)$, and the columns of U highlighted in red correspond to the basis elements that map to the free component of $H_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$.

$$U = \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} \begin{pmatrix} 0 & -1 & -x & 1 \\ -x & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad V = \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} \begin{pmatrix} 0 & 1 & 0 & -x & -x^2 \\ 1 & 0 & 0 & -x & -x^2 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then, $H_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$ is calculated using Proposition 4.3.5 as follows, with the trivial and torsion summands highlighted in blue and the free summands in red.

$$\begin{array}{l} H_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) \xrightarrow{\text{GrMod}} \frac{\mathbb{Q}[x]\langle(-b+c)x\rangle}{\mathbb{Q}[x]\langle(1)(-b+c)x\rangle} \oplus \frac{\mathbb{Q}[x]\langle-a+b\rangle}{\mathbb{Q}[x]\langle(x)(-a+b)\rangle} \oplus \frac{\mathbb{Q}[x]\langle(-a+d)x\rangle}{\mathbb{Q}[x]\langle(x)(-a+d)x\rangle} \oplus \mathbb{Q}[x]\langle a \rangle \\ \xrightarrow{\text{GrMod}} (0) \oplus \left(\frac{\mathbb{Q}[x]\langle-a+b\rangle}{\mathbb{Q}[x]\langle(x)(-a+b)\rangle} \right) \oplus \Sigma^1 \left(\frac{\mathbb{Q}[x]\langle-a+d\rangle}{\mathbb{Q}[x]\langle(x)(-a+d)\rangle} \right) \oplus \mathbb{Q}[x]\langle a \rangle \\ \xrightarrow{\text{GrMod}} \frac{\mathbb{Q}[x]}{(x)} \oplus \Sigma^1 \left(\frac{\mathbb{Q}[x]}{(x)} \right) \oplus \mathbb{Q}[x] \end{array}$$

We can also determine the 0^{th} persistence barcode $\text{Bar}(K_\bullet; \mathbb{Q})$ of the filtration K_\bullet by calculating the interval decomposition of the 0^{th} persistent homology module $H_0(K_\bullet; \mathbb{Q})$ by Corollary 2.5.13 as follows:

$$\begin{array}{llll} H_0^{\text{Gr}}(K_\bullet; \mathbb{Q}) & \xrightarrow{\text{GrMod}} & \frac{\mathbb{Q}[x]}{(x)} \oplus \Sigma^1 \left(\frac{\mathbb{Q}[x]}{(x)} \right) \oplus \mathbb{Q}[x] & \text{as graded } \mathbb{Q}[x]\text{-modules} \\ H_0(K_\bullet; \mathbb{Q}) & \xrightarrow{\text{Pers}} & \mathbb{I}_\bullet^{[0,1)} \oplus \mathbb{I}_\bullet^{[1,2)} \oplus \mathbb{I}_\bullet^{[0,\infty)} & \text{as persistence modules over } \mathbb{Q} \\ \text{Bar}(K_\bullet; \mathbb{Q}) & \ni & [0, 1) \quad , \quad [1, 2) \quad , \quad [0, \infty) & \text{as intervals in a persistence barcode} \end{array}$$

Section 4.6. Matrix Calculation of Homology of Graded Chain Complexes

Let $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ be a chain complex of free graded $\mathbb{F}[x]$ -modules C_n of finite rank with graded differentials $\partial_n : C_{n-1} \rightarrow C_n$. As with the case of chain complexes of free R -modules in Section 4.2, we can calculate the n^{th} homology of C_* using SNDs of $[\partial_{n+1}]$. Recall that this calculation relies on the existence of a decomposition of C_n by Theorem 4.2.3 into three direct summands as denoted below:

$$C_n \stackrel{\text{Mod}}{\cong} K_n^{\text{tor}} \oplus K_n^{\text{free}} \oplus (C_n / \ker \partial_n)$$

We discussed that this decomposition can be represented by a specific SND $(U_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$ and we have also established that this SND can be calculated using an arbitrary SND $(W_{n+1}, D_{n+1}, V_{n+1})$ of $[\partial_{n+1}]$. Then, using the same arguments for Smith Normal Decompositions of matrices over $\mathbb{F}[x]$ given by Lemma 4.4.1 in Section 4.3, we claim that the results of Section 4.2 also extend to the graded case. That is, we have the following decomposition on C_n as a *graded* $\mathbb{F}[x]$ -module:

$$C_n \stackrel{\text{GrMod}}{\cong} K_n^{\text{tor}} \oplus K_n^{\text{free}} \oplus (C_n / \ker \partial_n)$$

with graded $\mathbb{F}[x]$ -modules K_n^{tor} , K_n^{free} , and $C_n / \ker \partial_n$ defined similarly as in Theorem 4.2.3.

In Section 3.3, we discussed how the persistent homology of filtrations can be expressed as the homology of graded chain complexes. That is, given a filtration K_\bullet , we can determine the n^{th} persistent homology module $H_n(K_\bullet; \mathbb{F})$ of K_\bullet with coefficients in \mathbb{F} by the calculating the n^{th} simplicial persistent homology module:

$$H_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \cong_{\text{GrMod}} \frac{\ker(\partial_n^{\text{Gr}})}{\text{im}(\partial_{n+1}^{\text{Gr}})}$$

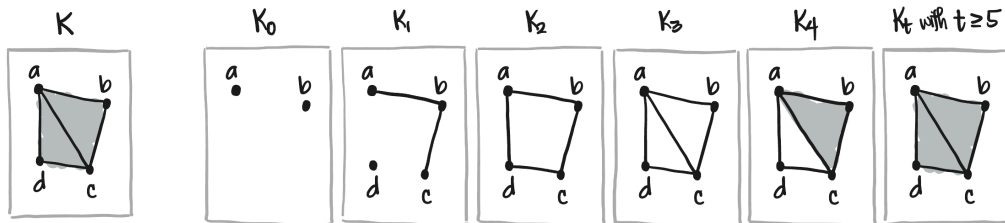
where, for all $n \in \mathbb{Z}$, the filtered n^{th} chain module $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ of K_\bullet is a graded $\mathbb{F}[x]$ -module and graded n^{th} boundary map $\partial_n^{\text{Gr}} : C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}^{\text{Gr}}(K_\bullet; \mathbb{F})$ is a graded homomorphism. Then, $H_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ admits the following graded presentation:

$$C_{n+1}^{\text{Gr}}(K_\bullet; \mathbb{F}) \xrightarrow{\partial_{n+1}^{\text{Gr}}} \ker(\partial_n^{\text{Gr}}) \xrightarrow{\pi} H_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \rightarrow 0$$

Since the arguments for the calculation of graded chain complexes follow exactly as that of Section 4.2 (with the addition of the modifier *graded* for $R = \mathbb{F}[x]$), we only present examples and some comments about calculation in this section.

Earlier in Example 4.5.8, we calculated the 0^{th} persistent homology module $H_0(K_\bullet; \mathbb{Q})$ of some filtration K_\bullet where $\ker(\partial_0^{\text{Gr}}) = C_0^{\text{Gr}}(K_\bullet; \mathbb{Q})$. In the example below, we calculate the 1^{st} persistent homology module $H_1(K_\bullet; \mathbb{Q})$ of the same filtration.

Example 4.6.1. Let K be the simplicial complex given below with orientation $\text{Vert}(K) = (a, b, c, d)$. Let K_\bullet be a filtration on K with K_t given as follows:



To calculate $H_1(K_\bullet; \mathbb{Q})$, we consider the matrices of the graded boundary maps $\partial_2^{\text{Gr}} : C_2^{\text{Gr}}(K_\bullet; \mathbb{Q})$ and $\partial_1^{\text{Gr}} : C_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$, which determine the torsion and free component of $H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ respectively. Provided below are

$[\partial_1^{\text{Gr}}]$ and $[\partial_2^{\text{Gr}}]$ relative to the standard bases.

$$[\partial_1^{\text{Gr}}] = \begin{matrix} & abx & bcx & adx^2 & cdx^2 & acx^3 \\ \begin{matrix} a \\ b \\ cx \\ dx \end{matrix} & \begin{pmatrix} -x & 0 & -x^2 & 0 & -x^3 \\ x & -x & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & x^2 \\ 0 & 0 & x & x & 0 \end{pmatrix} \end{matrix} \quad \text{and} \quad [\partial_2^{\text{Gr}}] = \begin{matrix} & abcx^4 & acdx^5 \\ \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} & \begin{pmatrix} x^3 & 0 \\ x^3 & 0 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & x^2 \end{pmatrix} \end{matrix}$$

We want to find an SNDs (U_1, D_1, V_1) and (U_2, D_2, V_2) of $[\partial_1^{\text{Gr}}]$ and $[\partial_2^{\text{Gr}}]$ respectively such that a basis compatible with the following decomposition of $C_1^{\text{Gr}}(K_{\bullet}; \mathbb{Q})$ can be identified from said SNDs:

$$C_1^{\text{Gr}}(K_{\bullet}; \mathbb{Q}) \cong^{\text{GrMod}} K_1^{\text{tor}} \oplus K_1^{\text{free}} \oplus \frac{C_1^{\text{Gr}}(K_{\bullet}; \mathbb{Q})}{\ker(\partial_1^{\text{Gr}})}$$

with K_1^{tor} and K_1^{free} as denoted in Theorem 4.2.3. Recall that

$$\ker(\partial_1^{\text{Gr}}) \cong K_1^{\text{tor}} \oplus K_1^{\text{free}} \quad \text{and} \quad T(H_1^{\text{Gr}}(K_{\bullet}; \mathbb{Q})) \cong K_1^{\text{tor}} / \text{im}(\partial_2^{\text{Gr}})$$

where $T(-)$ denotes the torsion submodule of a graded $\mathbb{F}[x]$ -module. Note that if we want to identify cycle representatives, the basis we get for K_1^{tor} from the SND of $[\partial_2^{\text{Gr}}]$ must be identifiable in the basis for $\ker(\partial_1^{\text{Gr}})$ we get from the SND of $[\partial_1^{\text{Gr}}]$. To get SNDs with this property, we do matrix reduction differently than described by Algorithm 4.5.1.

For the SND (U_2, D_2, V_2) of $[\partial_2^{\text{Gr}}]$: Define $T_0 := [\partial_2^{\text{Gr}}]$ and let T_k denote matrix $[\partial_2^{\text{Gr}}]$ after k column or row elimination operations. We use the following color scheme in the calculation below:

chosen pivot , entry to be eliminated , pivot multiplier , affected row or column

$$\begin{aligned} T_1 &:= T_0 \cdot E_{\text{add}}^{[2]}(1, 2; x) = \begin{pmatrix} x^3 & 0 \\ x^3 & 0 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & x^2 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x^3 & x^4 \\ x^3 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} \\ T_2 &:= E_{\text{add}}^{[5]}(2, 5; x^2) \cdot T_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x^2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^3 & x^4 \\ x^3 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} = \begin{pmatrix} x^3 & x^4 \\ 0 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} \\ T_3 &:= E_{\text{add}}^{[5]}(1, 5; x^2) \cdot T_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & x^2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^3 & x^4 \\ 0 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} = \begin{pmatrix} 0 & x^4 \\ 0 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} \\ T_4 &:= E_{\text{add}}^{[5]}(3, 4; 1) \cdot T_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x^4 \\ 0 & x^4 \\ 0 & -x^3 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} = \begin{pmatrix} 0 & x^4 \\ 0 & x^4 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
T_5 &:= E_{\text{add}}^{[5]}(2, 4; -x) \cdot T_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -x & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x^4 \\ 0 & x^4 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} = \begin{pmatrix} 0 & x^4 \\ 0 & 0 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} \\
T_6 &:= E_{\text{add}}^{[5]}(1, 4; -x) \cdot T_5 = \begin{pmatrix} 1 & 0 & 0 & -x & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x^4 \\ 0 & 0 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix}
\end{aligned}$$

Observe that the matrix T_6 is only a few elementary permutations away from being in Smith Normal Form. Instead of doing row and column permutations on T_6 , we stop here and work with the SND (U_2, D_2, V_2) of $[\partial_2^{\text{Gr}}]$ as given below, with the columns of U_2 corresponding to a basis of K_1^{tor} and the entries in D_2 corresponding to the invariant factors of $H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ highlighted in red and purple.

$$\begin{aligned}
V_2 &= E_{\text{add}}^{[2]}(1, 2; x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad D_2 = T_6 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & x^3 \\ -x & 0 \end{pmatrix}, \\
U_2 &= \left(E_{\text{add}}^{[5]}(1, 4; -x) \cdot E_{\text{add}}^{[5]}(2, 4; -x) \cdot E_{\text{add}}^{[5]}(3, 4; 1) \cdot E_{\text{add}}^{[5]}(1, 5; x^2) \cdot E_{\text{add}}^{[5]}(2, 5; x^2) \right)^{-1} \\
&= E_{\text{add}}^{[5]}(2, 5; -x^2) \cdot E_{\text{add}}^{[5]}(1, 5; -x^2) \cdot E_{\text{add}}^{[5]}(3, 4; -1) \cdot E_{\text{add}}^{[5]}(2, 4; x) \cdot E_{\text{add}}^{[5]}(1, 4; x) \\
&= abx \begin{pmatrix} 1 & 0 & 0 & x & -x^2 \\ bcx & 0 & 1 & 0 & -x^2 \\ adx^2 & 0 & 0 & -1 & 0 \\ cdx^2 & 0 & 0 & 1 & 0 \\ acx^3 & 0 & 0 & 0 & 1 \end{pmatrix} \\
&\quad \underbrace{\begin{pmatrix} x & -x^2 \\ -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}}_{K_1^{\text{tor}}}
\end{aligned}$$

Let $\beta_1 x^{s_1} \in C_n^{\text{Gr}}(K_\bullet; \mathbb{Q})$ be given by $[\beta_1 x^{s_1}] = \text{col}_4(U_2)$ and $[\beta_2 x^{s_2}] = \text{col}_5(U_2)$, following the color scheme above. Then, we have the following information for the graded invariant factor decomposition of $K_1^{\text{tor}} / \text{im}(\partial_2^{\text{Gr}})$:

Filtered Cycle	Cycle in $C_1(K; \mathbb{Q})$	Degree of Invariant Factor	Grading Shift
$\beta_1 x^{s_1} = (ab + bc - ad + cd)x^2$	$\beta_1 = ab + bc - ad + cd$	$t_1 = \deg_h(x^3) = 3$	$s_1 = 2$
$\beta_2 x^{s_2} = (-ab - bc + ac)x^3$	$\beta_2 = -ab - bc + ac$	$t_1 = \deg_h(-x) = 1$	$s_2 = 3$

Then, the torsion component $T(H_1^{\text{Gr}}(K_\bullet; \mathbb{Q}))$ of $H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$ is given by:

$$T\left(H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})\right) \stackrel{\text{GrMod}}{\cong} \frac{K_1^{\text{tor}}}{\text{im}(\partial_2^{\text{Gr}})} \cong \frac{\mathbb{Q}[x]\langle \beta_1 x^2 \rangle}{\mathbb{Q}[x]\langle x^3 \cdot \beta_1 x^2 \rangle} \oplus \frac{\mathbb{Q}[x]\langle \beta_2 x^3 \rangle}{\mathbb{Q}[x]\langle x \cdot \beta_2 x^3 \rangle} \cong \Sigma^2 \left(\frac{\mathbb{Q}[x]}{(x^3)} \right) \oplus \Sigma^3 \left(\frac{\mathbb{Q}[x]}{(x)} \right)$$

To determine the free component $F(H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})) \cong K_1^{\text{free}}$ of $H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$, we need information about $\ker(\partial_1^{\text{Gr}}) = K_1^{\text{free}} \oplus K_1^{\text{tor}}$, which we can get from an SND of $[\partial_1^{\text{Gr}}]$. Given below is the SND (U_1, D_1, V_1) of

$[\partial_1^{\text{Gr}}]$ calculated in Example 4.4.5 (column reduction) and Example 4.4.8 (row reduction):

$$U_1 = \begin{pmatrix} 1 & 1 & x & x \\ 0 & 1 & x & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad D_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 \end{pmatrix} \quad V_1 = \begin{matrix} abx \\ bcx \\ adx^2 \\ cdx^2 \\ acx^3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & x & -x^2 \\ 0 & 1 & 0 & x & -x^2 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$\underbrace{\begin{pmatrix} x & -x^2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\ker(\partial_1^{\text{Gr}})}$

Since the 4th and 5th columns of the SNF D_1 of $[\partial_1^{\text{Gr}}]$ are **zero columns**, the 4th and 5th columns of V_1 determine a basis of $\ker(\partial_1^{\text{Gr}})$. Observe that these two columns are exactly the columns corresponding to $\beta_1 x^{s_1}$ and $\beta_2 x^{s_2}$. Then, $K_1^{\text{free}} = 0$ and the free component of $H_1^{\text{Gr}}(K_\bullet, \mathbb{Q})$ is also trivial. Therefore,

$$H_1^{\text{Gr}}(K_\bullet; \mathbb{Q}) \cong^{\text{GrMod}} F\left(H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})\right) \oplus T\left(H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})\right) \cong 0 \oplus \Sigma^2\left(\frac{\mathbb{Q}[x]}{(x^3)}\right) \oplus \Sigma^3\left(\frac{\mathbb{Q}[x]}{(x)}\right)$$

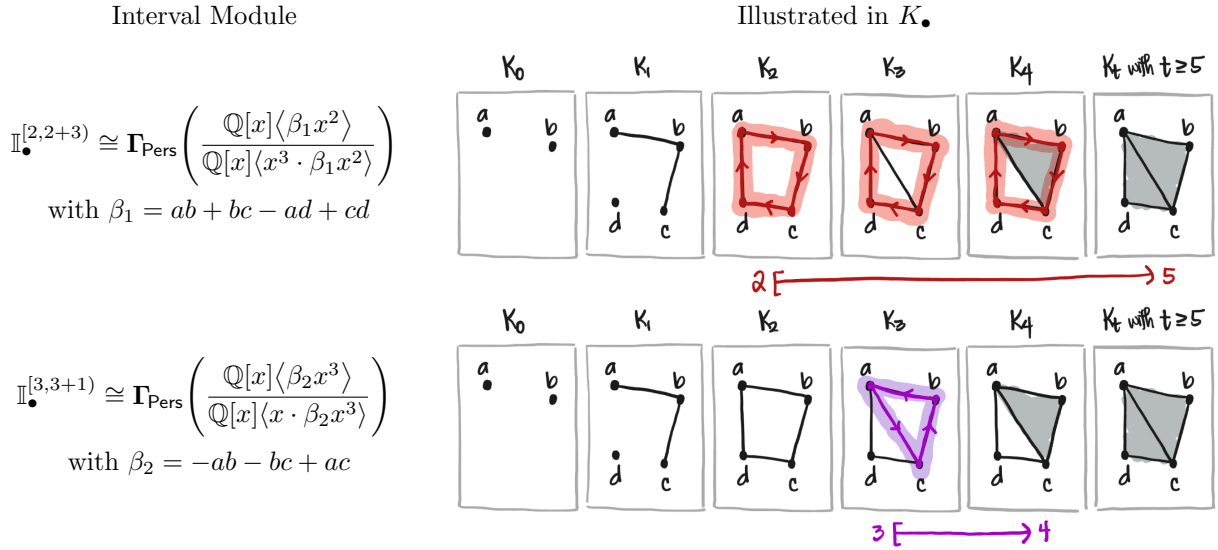
By application of $\mathbf{\Gamma}_{\text{Pers}}(-)$ on the graded invariant factor decomposition of $H_1^{\text{Gr}}(K_\bullet; \mathbb{Q})$, we get the following for the interval decomposition and persistence barcode $\text{Bar}_1(K_\bullet; \mathbb{Q})$ of $H_1(K_\bullet; \mathbb{Q})$:

$$H_1(K_\bullet; \mathbb{Q}) \cong^{\text{Pers}} \mathbf{\Gamma}_{\text{Pers}}\left(\Sigma^2\left(\frac{\mathbb{Q}[x]}{(x^3)}\right) \oplus \Sigma^3\left(\frac{\mathbb{Q}[x]}{(x)}\right)\right) \cong \mathbb{I}_\bullet^{[2,2+3]} \oplus \mathbb{I}_\bullet^{[3,3+1]}$$

and

$$\text{Bar}_1(K_\bullet; \mathbb{Q}) = \{[2, 5], [3, 4]\}$$

The intervals in the 1st persistent barcode $\text{Bar}_1(K_\bullet; \mathbb{Q})$ of K_\bullet over \mathbb{Q} correspond to the 1st filtered homology classes illustrated below:



Acknowledgements

This expository paper is the culmination of my two and a half year's stay at Oregon State University. I will be the first to admit that this paper is far from perfect and there is so much more that can be done to improve it. The weaknesses of this paper are reflective of my journey as a graduate student and a learner of mathematics. Before I started my program at OSU, I had very little to no knowledge of topics like category theory, graded module theory, algebraic topology, and homological algebra – all of which I had to learn in sufficient depth to write this paper. Mistakes were a huge part of my journey but I like to see my recognition of these mistakes as part of growth and as a testament to how much I have improved from where I started.

Despite this paper's weaknesses, I'd like to think that this paper has some meaningful contribution to mathematics. Persistent homology theory is a new and young field of mathematics, one that emerged only a few decades ago. A huge part of the writing process involved gathering all these seemingly disparate papers and results, and trying to make sense of the differences in definitions, convention, and notation. This paper is my attempt at reconciling these differences and making the connections between all these algebraic concepts clearer and more concrete. I'd like to think that, in the future, what I had written would help another student progress further in their learning journey.

The strengths of this paper are also reflective of my support system, without whom I could not have completed this paper and my Master's degree. I'd like to express my sincere gratitude to the following people:

1. My advisor, Dr. Christine Escher, for her unwavering support and guidance throughout my career at OSU, undeterred by my failings and weaknesses.
2. Dr. Chad Giusti and Dr. Ren Guo, for their encouragement, patience, and understanding as part of my graduate committee, especially during the times I had faced hardships and roadblocks.
3. The mathematics community at OSU, including my fellow graduate students, the graduate and undergraduate math faculty, and the math support staff, for making me feel welcome at OSU.
4. My family and friends, for their support throughout the challenges I've faced, whether they be academic, financial, emotional, or medical in nature.

Appendix A1. List of Symbols

Symbols Involving Sets, Modules, and Graded Modules.

\mathbb{N}_0	the nonnegative integers, i.e. $\mathbb{N}_0 = \{n \in \mathbb{Z} : n \geq 0\} = \{0, 1, 2, \dots\}$.	$R[x]$	the polynomial ring with indeterminate x , equipped with the standard grading when grading is considered, see Definition 2.4.2.
R	unless otherwise specified, typically used to refer to an arbitrary principal ideal domain (PID).	Mx^t	the R -module isomorphic to M by $mx^t \mapsto m \in M$; may be used to identify the degree of a homogeneous component of a graded $R[x]$ -module, see Remark 2.4.9.
R^\times	the group of units (invertible elements) of some ring R .	$\deg(-)$	the degree of a polynomial with indeterminate x , not necessarily as an element of a graded $\mathbb{F}[x]$ -module.
\mathbb{F}	typically used to refer to an arbitrary field.	$\deg_h(-)$	the degree of a homogeneous element of a graded $\mathbb{F}[x]$ -module, usually used when a relation is only valid for homogeneous elements, agrees with $\deg(-)$ for nonzero homogeneous elements and is undefined for non-homogeneous elements and zero.
$F(M)$	the free component of an R -module M .	id_A	usu. the identity function on some set A .
$T(M)$	the torsion component of an R -module M .		
$\frac{M}{N}$	usu. the quotient module with some R -module M and submodule N of M ; equiv. to M/N .		
$R\langle - \rangle$	refers to the free R -module formed by R -formal sums of elements in the argument, see Definition A3.7.		

Symbols for Relevant Categories.

C	usu. used to refer to an arbitrary category.	Mod $_{R[x]}$	the category of ungraded $R[x]$ -modules, i.e. disregarding any graded structure (if such exists).
Top	the category of topological spaces and continuous maps.	GrMod $_{R[x]}$	the category of graded modules over $R[x]$ and graded $R[x]$ -module homomorphisms, see Theorem 2.4.6.
Poset (R, \leq)	the category induced by a poset (R, \leq) , see Definition A4.3, conventionally denoted as (R, \leq) in category theory; not to be confused with the category Poset of posets and order-preserving maps (not used in this paper).	Pers $_{\mathbb{F}}$	the category of persistence modules over \mathbb{F} and persistence morphisms, see Definition 2.2.1.
A-Simp	the category of (abstract) simplicial complexes and simplicial maps, see Definition 1.3.2.	Ch-A	the category of chain complexes of an abelian category A and chain maps; in this paper, $*$ is used for the index $n \in \mathbb{Z}$, e.g. $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ is a chain complex.
A	usu. for an arbitrary abelian category.	Ch-GrMod $_{\mathbb{F}[x]}$	the category of graded chain complexes, see Definition 2.4.17.
AbGrp	the category of abelian groups and group homomorphisms.	Ch-Pers $_{\mathbb{F}}$	the category of persistence complexes, see Definition 2.2.10.
Mod $_R$	the category of right R -modules and R -module homomorphisms for some ring R , the modifier “right” is usu. dropped when R is commutative.	$H_n(-)$	the n^{th} chain homology functor Ch-A \rightarrow A in some abelian category A . For A = GrMod $_{\mathbb{F}[x]}$, see Definition 2.4.17. For A = Pers $_{\mathbb{F}}$, see Definition 2.2.10(iii).
Mod $_{\mathbb{Z}}$	category of \mathbb{Z} -modules, equiv. to AbGrp .	$\text{id}_{\mathbf{C}}$	the identity functor on a category C .

Symbols and Shorthands for Specific Binary Relations.

\oplus	refers to the direct sum operation on an abelian category \mathbf{A} ; when the category \mathbf{A} is ambiguous or emphasized, an accompanying relation $\cong_{\mathbf{A}}$ is identified in the relevant line. For $\mathbf{A} = \mathbf{Mod}_{\mathbb{F}[x]}$ or $\mathbf{A} = \mathbf{GrMod}_{\mathbb{F}[x]}$, see Remark 2.4.7. For $\mathbf{A} = \mathbf{Pers}_{\mathbb{F}}$, see Definition 2.2.5.	$\stackrel{\text{Ab}}{=}$	denotes an equality at the level of abelian groups or $\mathbb{F}[x]$ -vector spaces, see Remark 2.4.7.
$\cong_{\mathbf{C}}$	denotes an isomorphism relation on a category \mathbf{C} , usu. used when the category in which the relation is considered is ambiguous or emphasized.	$\stackrel{\text{Mod}}{\cong}$	shorthand for an isomorphism relation in $\mathbf{Mod}_{R[x]}$, disregarding any graded structure (if such exists), see Remark 2.4.7.
		$\stackrel{\text{GrMod}}{\cong}$	shorthand for an isomorphism relation in $\mathbf{GrMod}_{\mathbb{F}[x]}$, see Remark 2.4.7.
		$\stackrel{\text{Pers}}{\cong}$	shorthand for an isomorphism relation in $\mathbf{Pers}_{\mathbb{F}}$, see remark under Definition 2.2.3.

Notation involving Simplicial Complexes and Simplicial Homology. See Chapter 1.

K	unless otherwise specified, usually used to refer to an arbitrary simplicial complex.	∂_n	usu. refers to a simplicial boundary map $\partial_n : C_n(K; R) \rightarrow C_{n-1}(K; R)$, see Definition 1.2.7; also used to denote differentials of an arbitrary chain complex $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$.
$\text{Vert}(K)$	the vertex set of a simplicial complex K , see Definition 1.1.1.	$C_*(K; R)$	the simplicial chain complex $C_*(K; R) = (C_n(K; R), \partial_n)_{n \in \mathbb{Z}}$ with boundary maps $\partial_n : C_n(K; R) \rightarrow C_{n-1}(K; R)$, see Definition 1.2.9.
\mathcal{K}_n	the standard ordered basis of $C_n(K; R)$ of a simplicial complex K relative to some given orientation on $\text{Vert}(K)$, see Definition 1.2.5.	$C_*(K)$	shorthand for $C_*(K; \mathbb{Z})$.
$C_n(K; R)$	the n^{th} simplicial chain group of a simplicial complex K , see Definition 1.2.1.	$C_*(-; R)$	the simplicial chain complex functor $\mathbf{A-Simp} \rightarrow \mathbf{Ch-Mod}_R$, see Definition 1.3.7.
$C_n(K)$	shorthand for $C_n(K; \mathbb{Z})$.	$H_n(K; R)$	the n^{th} simplicial homology group of a simplicial complex K with coeff. in R , see Definition 1.2.9.
$C_n(-; R)$	the n^{th} simplicial chain group functor $\mathbf{A-Simp} \rightarrow \mathbf{Mod}_R$ with coefficients in R , see Definition 1.3.5.	$H_n(K)$	shorthand for $H_n(K; \mathbb{Z})$.
$f_{n,\#}$	usu. the map $C_n(K; R) \rightarrow C_n(L; R)$ on the n^{th} simplicial chain groups induced by the simplicial map $f : K \rightarrow L$, see Definition 1.3.3.	$H_n(-; R)$	usu. refers to the n^{th} simplicial homology functor $\mathbf{A-Simp} \rightarrow \mathbf{Mod}_R$ with coeff. in R , see Definition 1.3.9.
$f_{\#}$	shorthand for $f_{n,\#}$, used when the dimension n is arbitrary or unambiguous.	$\beta_n(K; R)$	the n^{th} Betti number of a simplicial complex K with coefficients in a PID R , i.e. $\beta_n(K; R) = \text{rank}(H_n(K; R))$.
$[v_0, \dots, v_n]$	an oriented n -simplex, i.e. a basis element of an n^{th} chain group $C_n(K; R)$ with ordering (v_0, \dots, v_n) , see Definition 1.2.5.	$\beta_n(K)$	shorthand for $\beta_n(K; \mathbb{Z})$.
$v_0 \cdots v_n$	string shorthand for the oriented n -simplex $[v_0, \dots, v_n]$, see Remark 1.2.2.	$[\sigma]$	usu. refers to a homology class in $H_n(K; R)$ with the n -chain σ an R -formal sum of oriented n -simplices, typically written in string notation.

Notation involving Persistence Modules. See Chapter 2.

$(V_\bullet, \alpha_\bullet)$	a persistence module with vector spaces V_t and structure maps $\alpha_{s,t} : V_t \rightarrow V_s$, see Definition 2.1.1.	\mathbb{I}_\bullet^J	a J -interval (persistence) module with $J \subseteq \mathbb{N}_0$ an interval, see Definition 2.3.1.
V_\bullet	shorthand for $(V_\bullet, \alpha_\bullet)$, see Definition 2.1.1.	$\text{Bar}(V_\bullet)$	the persistence barcode of a persistence module V_\bullet , see Definition 2.3.6.
φ_\bullet	usu. a persistence morphism $\varphi_\bullet : V_\bullet \rightarrow W_\bullet$ with $\varphi_\bullet = (\varphi_t)_{t \in \mathbb{N}_0}$, see Definition 2.2.1(ii).	Γ_{GrMod}	the functor $\mathbf{Pers}_{\mathbb{F}} \rightarrow \mathbf{GrMod}_{\mathbb{F}[x]}$ in the isomorphism of categories bet. $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$, see Definition 2.5.1.
$(V_\bullet^\bullet, \alpha_\bullet^\bullet, \partial_\bullet^\bullet)$	a persistence complex, i.e. a chain complex of persistence modules $(V_n^\bullet, \alpha_n^\bullet)$ with differentials $\partial_n^\bullet : V_n^\bullet \rightarrow V_{n-1}^\bullet$ over $n \in \mathbb{Z}$; see Definition 2.2.10.	Γ_{Pers}	the functor $\mathbf{GrMod}_{\mathbb{F}[x]} \rightarrow \mathbf{Pers}_{\mathbb{F}}$ in the isomorphism of categories bet. $\mathbf{Pers}_{\mathbb{F}}$ and $\mathbf{GrMod}_{\mathbb{F}[x]}$, see Definition 2.5.6.
$(V_\bullet^\bullet, \partial_\bullet^\bullet)$	shorthand for $(V_\bullet^\bullet, \alpha_\bullet^\bullet, \partial_\bullet^\bullet)$.		

Notation involving Filtrations and Persistent Homology. See Chapter 3.

K_\bullet	a filtration $K_\bullet : \mathbf{Poset}(\mathbb{N}_0, \leq) \rightarrow \mathbf{A-Simp}$ of a simplicial complex K , see Definition 3.1.1.	$C_*(K_\bullet; \mathbb{F})$	the simplicial persistence complex of K_\bullet with $C_*(K_\bullet; \mathbb{F}) = (C_n(K_\bullet; \mathbb{F}), \partial_n^\bullet)_{n \in \mathbb{Z}}$, see Definition 3.3.17.
$C_n(K_\bullet; \mathbb{F})$	the n^{th} filtered chain module of a filtration K_\bullet with coefficients in \mathbb{F} , see Definition 3.3.1.	$C_*^{\text{Gr}}(K_\bullet; \mathbb{F})$	the simp. graded chain complex of K_\bullet with $C_*^{\text{Gr}}(K_\bullet; \mathbb{F}) = (C_n^{\text{Gr}}(K_\bullet; \mathbb{F}), \partial_n^{\text{Gr}})_{n \in \mathbb{Z}}$, see Definition 3.3.17.
$C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$	the n^{th} graded chain module of a filtration K_\bullet with coefficients in \mathbb{F} , see Definition 3.3.4.	$H_n(K_\bullet; \mathbb{F})$	the n^{th} persistent homology module of K_\bullet with coeff. in \mathbb{F} , see Definition 3.2.1.
$\mathcal{K}_n^{\text{Gr}}$	the standard ordered basis of $C_n^{\text{Gr}}(K_\bullet; \mathbb{F})$ induced by the orientation on K , see Definition 3.3.9.	$H_n^{\text{Gr}}(K_\bullet; \mathbb{F})$	the n^{th} graded homology module of a filtration K_\bullet , see Definition 3.3.19.
∂_n^\bullet	the n^{th} filtered boundary morphism $\partial_n^\bullet : C_n(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}(K_\bullet)$, see Definition 3.3.12.	$[-]_t$	a homology class in a persistent homology module at scale $t \in \mathbb{N}_0$, see Defn. 3.2.1.
∂_n^{Gr}	the n^{th} graded boundary map $\partial_n^{\text{Gr}} : C_n^{\text{Gr}}(K_\bullet; \mathbb{F}) \rightarrow C_{n-1}^{\text{Gr}}(K_\bullet; \mathbb{F})$, see Definition 3.3.13.	$\text{Bar}_n(K_\bullet; \mathbb{F})$	the n^{th} persistence barcode of a filtration K_\bullet with coefficients in \mathbb{F} , see Defn. 3.2.1.
		$H_n(K_t; p; \mathbb{F})$	the p -persistent n^{th} homology group with coefficients in \mathbb{F} , see Defn. 3.2.12.
		$\beta_n(K_t; p; \mathbb{F})$	the p -persistent n^{th} Betti number with coefficients in \mathbb{F} , see Defn. 3.2.12.

Notation involving Matrices and Smith Normal Decompositions (SNDs). See Appendix A2.

$[v]_S$	the coordinate matrix of $v \in M$ relative to a basis S of some free R -module M , see Definition A3.8.	I_n	the identity matrix in $\text{GL}(n, R)$.
$[\varphi]_{A,S}$	the matrix of a module homomorphism $\varphi : M \rightarrow N$ relative to bases S and A of R -modules M and N respectively.	$E_{\text{swap}}^{[n]}(-)$	an elementary permutation of degree n over some ring R , see Definition A2.5.
(U, D, V)	usu. refers to an SND of some matrix A such that $U^{-1}AV = D$, see Defn. 4.1.8.	$E_{\text{dilate}}^{[n]}(-)$	an elementary dilation of degree n over some ring R , see Definition A2.5.
		$E_{\text{add}}^{[n]}(-)$	an elementary transvection of degree n over some ring R , see Definition A2.5.

Appendix A2. Matrices over Euclidean Domains

In this paper, we discuss algorithms over matrices over some Euclidean domain R , i.e. matrices whose entries are elements of R . Note that Euclidean domains have multiplicative identities, typically denoted by 1 or 1_R , which are unique. Euclidean domains are also commutative rings, i.e. the multiplication operation is commutative.

Note that most definitions involving matrices over R are generally the same as that over \mathbb{R} . For example, since R has identity 1, identity matrices and invertible matrices are defined in the same way as that of $\text{GL}(n, \mathbb{R})$. Below, we identify notation involving families of matrices used in this paper.

Definition A2.1. Let R be a Euclidean domain. Let $M_{m,n}(R)$ refer to the **collection of all $(m \times n)$ -matrices**, i.e. having m rows and n columns, **over** R . If $m = n$, we write $M_n(R)$ to refer to $M_{n,n}(R)$. Let $\text{GL}(n, R)$ refer to the **general linear group of degree n over R** , i.e. the collection of all invertible $(n \times n)$ -matrices over R .

Next, we identify notation for certain components of some given matrix.

Definition A2.2. Let $A \in M_{m,n}(R)$. Let $A(j, i)$ denote the **entry** of A at the j^{th} row and i^{th} column.

If $n = 1$, we call $A \in M_{m,1}(R)$ a **column vector** and write $A(j) := A(j, 1)$ for the j^{th} entry of A . If $m = 1$, we call $A \in M_{1,n}(R)$ a **row vector** and write $A(i) := A(1, i)$ for the i^{th} entry of A .

Let $\text{col}_i(A) \in M_{m,1}(R)$ refer to the column vector corresponding to the i^{th} **column of** $A \in M_{m,n}(R)$, i.e. for all $j \in \{1, \dots, m\}$, $\text{col}_i(A)(j) = A(j, i)$ for fixed i . Let $\text{row}_j(A) \in M_{1,n}(R)$ refer to the row vector corresponding to the j^{th} **row of** $A \in M_{m,n}(R)$, i.e. for all $i \in \{1, \dots, n\}$, $\text{row}_j(A)(i) = A(j, i)$ for fixed j .

Note that indices of the rows and columns of matrices always start at 1. We emphasize this since, for some objects referenced in this paper, e.g. persistence modules, the indexing starts at 0. In this paper, we make an effort to use j and i to refer to an index of some row and column respectively but this is sometimes not possible, e.g. in cases where i or j have been defined beforehand.

We also talk about diagonal matrices in this paper. For brevity, we identify notation for describing diagonal matrices by the elements on their diagonal.

Definition A2.3. Let $D = \text{diag}(d_1, d_2, \dots, d_n)$ with $d_i \in R$ for all $i \in \{1, \dots, n\}$ refer to matrix $D \in M_n(R)$ given by $D(i, j) = d_i$ if $i = j$ and $D(i, j) = 0$ otherwise.

We also talk about block matrices in this paper. We identify the notation used for these below.

Definition A2.4. Let $A \in M_{m_1, n_1}(R)$, $B \in M_{m_1, n_2}(R)$, $C \in M_{m_2, n_1}(R)$, $D \in M_{m_2, n_2}(R)$. We write

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

to describe the matrix $X \in M_{m,n}(R)$ with $m = m_1 + m_2$ and $n = n_1 + n_2$. We call X a **block matrix** and the matrices A, B, C, D the **blocks** of X . When a block is given by a zero matrix, we usually do not write the dimensions of the zero matrix and assume its dimensions are appropriately defined.

Listed below are three families of matrices that are generalizations of the *elementary matrices* over \mathbb{R} , i.e. those used in matrix reduction and are usually called Type (I), Type (II), and Type (III) matrices. Below, we state a definition for these families of matrices, taken from [AW92, Definition 4.1.8]. Note that interpretation

of the left or right multiplication of these matrices are given later in Proposition A2.7 (row operations) and Proposition A2.8 (column operations).

Definition A2.5. The **elementary matrices over R of degree n** consists of elementary permutations, elementary dilations, and elementary transvections, defined below. Note that all indices are elements of $\{1, \dots, n\}$.

- i. An **elementary permutation** $E_{\text{swap}}^{[n]}(k_1, k_2) \in \text{GL}(n, R)$ on two indices $k_1, k_2 \in \{1, \dots, n\}$ is the matrix obtained by interchanging rows k_1 and row k_2 (or equivalently, columns k_1 and k_2) of the identity matrix I_n , i.e.

$$\text{row}_j \left(E_{\text{swap}}^{[n]}(k_1, k_2) \right) = \begin{cases} \text{row}_{k_2}(I_n) & \text{if } j = k_1 \\ \text{row}_{k_1}(I_n) & \text{if } j = k_2 \\ \text{row}_j(I_n) & \text{otherwise} \end{cases} \quad \text{or} \quad \text{col}_i \left(E_{\text{swap}}^{[n]}(k_1, k_2) \right) = \begin{cases} \text{col}_{k_2}(I_n) & \text{if } i = k_1 \\ \text{col}_{k_1}(I_n) & \text{if } i = k_2 \\ \text{col}_i(I_n) & \text{otherwise} \end{cases}$$

Elementary permutations are involutions, i.e. the inverse of $E_{\text{swap}}^{[n]}(k_1, k_2)$ is $E_{\text{swap}}^{[n]}(k_1, k_2)$.

- ii. An **elementary transvection** $E_{\text{add}}^{[n]}(k_j, k_i; \alpha) \in \text{GL}(n, R)$ by $\alpha \in R$ on the $(k_j, k_i)^{\text{th}}$ entry with $k_j, k_i \in \{1, \dots, n\}$ and $k_j \neq k_i$ is the matrix obtained by taking the identity matrix and replacing the $(k_j, k_i)^{\text{th}}$ entry with α , i.e.

$$E_{\text{add}}^{[n]}(k_j, k_i; \alpha)(j, i) = \begin{cases} \alpha & \text{if } j = k_j \text{ and } i = k_i \\ I_n(j, i) & \text{otherwise} \end{cases}$$

We may call α the **transvection multiplier** of $E_{\text{add}}^{[n]}(k_j, k_i; \alpha)$. The inverse of an elementary transvection $E_{\text{add}}^{[n]}(k_j, k_i; \alpha)$ is the transvection $E_{\text{add}}^{[n]}(k_j, k_i; -\alpha)$ where $-\alpha$ is the additive inverse of α in R .

- iii. An **elementary dilation** $E_{\text{dilate}}^{[n]}(k, \mu) \in \text{GL}(n, R)$ on index $k \in \{1, \dots, n\}$ by a unit $\mu \in R^\times$ is the matrix obtained by replacing the k^{th} diagonal element of I_n with μ , i.e.

$$E_{\text{dilate}}^{[n]}(k, \mu)(j, i) = \begin{cases} \mu & \text{if } j = k \text{ and } i = k \\ I_n(j, i) & \text{otherwise} \end{cases}$$

We sometimes call μ the **dilation multiplier** of $E_{\text{dilate}}^{[n]}(k, \mu)$. The inverse of an elementary dilation $E_{\text{dilate}}^{[n]}(k, \mu)$ is the dilation $E_{\text{dilate}}^{[n]}(k, \mu^{-1})$ with μ^{-1} the multiplicative inverse of μ in R .

One reason why elementary matrices are considered significant is due to the following result for matrices over Euclidean domains, stated below.

Proposition A2.6. Any invertible matrix over a Euclidean domain R can be expressed as a finite product of elementary matrices. That is, $\text{GL}(n, R)$ is generated by the elementary matrices over R of degree n .

Remark. For a proof, see under [AW92, Theorem 5.2.10]. As a sidenote,

Observe that, given two Euclidean domains R_1 and R_2 , the family of elementary permutations on R_1 and R_2 are defined very similarly since only the multiplicative identities are used to generate them and these identities are unique for each ring. For example, the elementary permutations of matrices over \mathbb{Z} are exactly the elementary permutations of matrices of \mathbb{R} since they share identity elements.

We provide a visual description of these elementary permutations below. Observe that the order of the arguments k_1 and k_2 in $E_{\text{swap}}^{[n]}(k_1, k_2)$ does not matter, i.e. $E_{\text{swap}}^{[n]}(k_1, k_2) = E_{\text{swap}}^{[n]}(k_2, k_1)$.

$$\begin{array}{c}
\begin{array}{c} i=1 \quad \dots \quad k_1 \quad \dots \quad k_2 \quad \dots \quad n \triangleleft \text{column indices} \\
j=1 \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ k_1 & 0 & 0 & 1 & 0 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ k_2 & 0 & 1 & 0 & 0 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots \\ n & 0 & \dots & 0 & \dots & 0 & 1 \end{pmatrix} \\
\triangle \\
\text{row indices} \end{array}
\end{array}$$

Color Scheme:

- indices k_1 and k_2
- entries in I_n set to 1
- entries in I_n set to 0
- with uncolored entries agreeing with I_n

Elementary transvections are relatively straightforward to work with since they allow any ring element to be the transvection multiplier. Note that if $\alpha = 0$ is the additive identity of R , then $E_{\text{add}}^{[n]}(k_j, k_i; \alpha) = I_n \in \text{GL}(n, R)$. Also, note that the order of arguments in the notation $E_{\text{add}}^{[n]}(k_j, k_i; \alpha)$ is important, unlike for elementary permutations, i.e. $E_{\text{add}}^{[n]}(k_j, k_i; \alpha) \neq E_{\text{add}}^{[n]}(k_i, k_j; \alpha)$.

Provided below is a visual description of $E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha)$. As the labels suggest, the first two arguments $k_{\text{row}}, k_{\text{col}}$ determine that the $(k_{\text{row}}, k_{\text{col}})^{\text{th}}$ entry of $E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha)$ is multiplier α .

$$\begin{array}{c}
\begin{array}{c} i=1 \quad \dots \quad k_{\text{row}} \quad \dots \quad k_{\text{col}} \quad \dots \quad n \triangleleft \text{column indices} \\
j=1 \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ k_{\text{row}} & 0 & 1 & \alpha & 0 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ k_{\text{col}} & 0 & 0 & 1 & 0 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots \\ n & 0 & \dots & 0 & \dots & 0 & 1 \end{pmatrix} \\
\triangle \\
\text{row indices} \end{array}
\end{array}$$

Color Scheme:

- row index k_{row}
- column index k_{column}
- transvection multiplier α
- with uncolored entries agreeing with I_n

An elementary transvection $E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha)$ can also be described row-wise as follows. Note that only row k_{row} of $E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha)$ differs from the identity matrix I_n .

$$\text{row}_j \left(E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha) \right) = \begin{cases} \text{row}_j(I_n) + \alpha \cdot \text{row}_{k_{\text{col}}}(I_n) & \text{if } j = k_{\text{row}} \\ \text{row}_j(I_n) & \text{otherwise} \end{cases}$$

It also has the following column-wise description. Note that only column k_{col} of $E := E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha)$ differs from I_n .

$$\text{col}_i \left(E_{\text{add}}^{[n]}(k_{\text{row}}, k_{\text{col}}; \alpha) \right) = \begin{cases} \text{col}_i(I_n) + \alpha \cdot \text{col}_{k_{\text{row}}}(I_n) & \text{if } i = k_{\text{col}} \\ \text{col}_i(I_n) & \text{otherwise} \end{cases}$$

Let R be a commutative ring with identity and S be a subring of R . Since S and R must share an identity element, the elementary dilations over R are exactly those over S . Given an elementary dilation $E_{\text{add}}^{[n]}(k_j, k_i; r)$ over R with $r \in R$, it should be clear that if $r \in S$, then $E_{\text{add}}^{[n]}(k_j, k_i; r)$ is also an elementary dilation over S . More specifically, $E_{\text{add}}^{[n]}(k_j, k_i; r) \in \text{GL}(n, S)$.

However, we do not have this nice relationship for elementary dilations. That is, an elementary dilation

$E_{\text{dilate}}^{[n]}(k, r)$ over R with $r \in R^\times$ is not generally an elementary dilation over S even if $r \in S$. In particular, $E_{\text{dilate}}^{[n]}(k, r)$ is that over S only if r is invertible in S . This distinction becomes extremely significant when considering which elementary matrices we can use for matrix reduction. We consider some relevant cases below.

1. If $R = \mathbb{F}$ is a field, then any nonzero element of \mathbb{F} can be used as the dilation multiplier μ in $E_{\text{dilate}}^{[n]}(k, \mu)$. For example, row reduction of a matrix with \mathbb{R} entries allows the multiplication of any entry by a nonzero constant. This is because $E_{\text{dilate}}^{[n]}(k, a)$ is a valid elementary dilation over \mathbb{R} for any nonzero $a \in \mathbb{R}$.
2. In the case of $R = \mathbb{Z}$, the family of elementary dilations over \mathbb{Z} is much smaller than that of \mathbb{R} , despite $\mathbb{Z} \subseteq \mathbb{R}$, since the only units of \mathbb{Z} are 1 and -1 . So, when performing row reduction on matrices over \mathbb{Z} , we can only multiply each row by either 1 (which does nothing) or -1 .

As another example, let $M = E_{\text{dilate}}^{[3]}(1, 2)$ with multiplier $\mu = 2$. Observe that M is a valid elementary dilation over \mathbb{R} since $\mu = 2$ has an inverse of $\frac{1}{2} \in \mathbb{R}$. Given below is a more visual description of M and $M^{-1} = E_{\text{dilate}}^{[3]}(1, \frac{1}{2})$.

$$M = E_{\text{dilate}}^{[3]}(1, 2) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad M^{-1} = E_{\text{dilate}}^{[3]}(1, \frac{1}{2}) = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Since $\frac{1}{2} \notin \mathbb{Z}$, $M^{-1} \notin M_{3,3}(\mathbb{Z})$ and M^{-1} is not a valid elementary dilation over \mathbb{Z} .

3. In the case of $R = \mathbb{Q}[x]$, the elementary dilations over \mathbb{Q} and those of $\mathbb{Q}[x]$ are exactly the same since they share the same group of units, i.e. $\mathbb{Q}^\times = (\mathbb{Q}[x])^\times = \mathbb{Q} \setminus \{0\}$.

So, when performing row reduction on a matrix over $\mathbb{Q}[x]$, we cannot use an elementary dilation to reduce the powers since $x^t \notin \mathbb{Q}[x]$ for $t \in \mathbb{Z}$ with $t < 0$. For example, let $M \in M_{3,2}(\mathbb{Q}[x])$ be as given below.

$$M = \begin{pmatrix} x^3 & 0 & 0 \\ 0 & x & 1 \end{pmatrix}$$

The following row operation on M is not valid as a matrix operation over $\mathbb{Q}[x]$:

$$E_{\text{dilate}}^{[n]}(1, x^{-3})M = \begin{pmatrix} x^{-3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^3 & 0 & 0 \\ 0 & x & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 1 \end{pmatrix}$$

In particular, $E_{\text{dilate}}^{[n]}(1, x^{-3}) \notin M_{2,2}(\mathbb{Q}[x])$ and we cannot multiply the first row of M by x^{-3} .

We provide a visual description of elementary dilations below.

$$E_{\text{dilate}}^{[n]}(k, \mu) = \begin{matrix} & \begin{matrix} i=1 & \dots & k-1 & k & k+1 & \dots & n \end{matrix} \triangleleft \text{column indices} \\ \begin{matrix} j=1 \\ \vdots \\ k-1 \\ k \\ k+1 \\ \vdots \\ n \end{matrix} \triangleup \text{row indices} & \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & & 1 & 0 & 0 & & 0 \\ k & 0 & 0 & \mu & 0 & & 0 \\ 0 & & 0 & 0 & 1 & & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \end{matrix}$$

Color Scheme:
row or column index
dilation multiplier μ
 with uncolored entries
 agreeing with I_n

Below, we provide a characterization of row and column operations on matrices relative to these elementary matrices.

Proposition A2.7. Elementary Row Operations on a matrix $A \in M_{m,n}(R)$ correspond to multiplying A by some elementary matrix in $GL(m, R)$ on the left. In particular:

- i. Row Permutation or Row Swapping.** The product $E_{\text{swap}}^{[m]}(k_1, k_2)A$ corresponds to swapping the k_1^{th} and k_2^{th} rows of A , i.e.

$$\text{row}_j \left(E_{\text{swap}}^{[n]}(k_1, k_2)A \right) = \begin{cases} \text{row}_{k_2}(A) & \text{if } j = k_1 \\ \text{row}_{k_1}(A) & \text{if } j = k_2 \\ \text{row}_j(A) & \text{otherwise} \end{cases}$$

- ii. Row Dilation or Row Multiplication.** The product $E_{\text{dilate}}^{[m]}(k, \mu)A$ corresponds to multiplying the k^{th} row by a unit $\mu \in R^\times$, i.e.

$$\text{row}_j \left(E_{\text{dilate}}^{[n]}(k, \mu)A \right) = \begin{cases} \mu \text{row}_j(A) & \text{if } j = k \\ \text{row}_j(A) & \text{otherwise} \end{cases}$$

- iii. Row Addition.** With $\alpha \in R$ and column indices $k \neq p$, the product $E_{\text{add}}^{[m]}(k, p, \alpha)A$ corresponds to adding an α -multiple of the p^{th} row of A to the k^{th} row of A , i.e.

$$\text{row}_j \left(E_{\text{add}}^{[n]}(k, p; \alpha)A \right) = \begin{cases} \text{row}_k(A) + \alpha \text{row}_p(A) & \text{if } j = k \\ \text{row}_j(A) & \text{otherwise} \end{cases}$$

Note that the first argument k in $E_{\text{add}}^{[m]}(k, p, \alpha)A$ is the target row.

Proposition A2.8. Elementary Column Operations on a matrix $A \in M_{m,n}(R)$ correspond multiplying A by some elementary matrix in $GL(n, R)$ on the *left*. In particular:

- i. Column Permutation or Column Swapping.** The product $AE_{\text{swap}}^{[n]}(k_1, k_2)$ corresponds to swapping the k_1^{th} and k_2^{th} column of A , i.e.

$$\text{col}_i \left(AE_{\text{swap}}^{[n]}(k_1, k_2) \right) = \begin{cases} \text{col}_{k_2}(A) & \text{if } i = k_1 \\ \text{col}_{k_1}(A) & \text{if } i = k_2 \\ \text{col}_i(A) & \text{otherwise} \end{cases}$$

- ii. Column Dilation or Column Multiplication.** The product $AE_{\text{dilate}}^{[n]}(k, \mu)$ corresponds to multiplying the k^{th} column of A by a unit $\mu \in R^\times$, i.e.

$$\text{col}_i \left(AE_{\text{dilate}}^{[n]}(k, \mu) \right) = \begin{cases} \mu \text{col}_k(A) & \text{if } i = k \\ \text{col}_i(A) & \text{otherwise} \end{cases}$$

- iii. Column Addition.** With $\alpha \in R$ and column indices $k \neq p$, the product $E_{\text{add}}^{[n]}(p, k; \alpha)$ corresponds to adding an α -multiple of the p^{th} column of A to the k^{th} column of A , i.e.

$$\text{col}_i \left(AE_{\text{add}}^{[n]}(p, k; \alpha) \right) = \begin{cases} \text{col}_k(A) + \alpha \text{col}_p(A) & \text{if } i = k \\ \text{col}_i(A) & \text{otherwise} \end{cases}$$

Note that the second argument k in $E_{\text{add}}^{[n]}(p, k; \alpha)$ is the target column.

Appendix A3. Notes on Ring and Module Theory

For reference, we have listed some relevant basic definitions and notation involving rings and modules. These are mostly taken from the texts *Algebra: An Approach to Module Theory* [AW92] by William Adkins and Steven Weintraub, and *Abstract Algebra* [DF03] by David Dummit and Richard Foote.

Definition A3.1. Let R be a commutative ring. An **ideal** S of a ring R is a subring of R such that $RS \subseteq S$ and $SR \subseteq S$, i.e. the subring S is closed under left and right multiplication of elements of R . An ideal $S \subseteq R$ is called **principal** if S can be generated by a single element, i.e. there exists $a \in R$ such that $S = Ra = \{ra : r \in R\}$. In this case, we write $S = (a)$ and say S is **generated** by $a \in R$.

Remark. In some references, $S = (a)$ is written as $S = \langle a \rangle$, i.e. with angle brackets. We avoid using this notation in this paper for clarity, e.g. in cases where torsion is present. Instead, we use angle brackets as described in Definition A3.7, i.e. $\langle a \rangle$ is usually interpreted as $\mathbb{Z}\langle a \rangle$.

Below, we provide a definition for PIDs. Note that, in this paper, we almost always require that the ring R be at least a PID.

Definition A3.2. An **integral domain** is a nonzero commutative ring R with (multiplicative) identity $1_R \in R$ such that product of any two nonzero elements of R is nonzero. A **principal ideal domain** R is an integral domain such that every ideal of R is principal.

Remark. Note that some references do not necessarily require integral domains to have identities. For contrast, an example of a commutative ring without identity is $2\mathbb{Z}$.

We also have references to *Euclidean domains* in this paper but we have determined that an exact definition for such is not relevant. It should suffice to know that all Euclidean domains are PIDs and are equipped with some sort of division algorithm, in which quotients and remainders are well-defined and unique. Examples of Euclidean domains (and therefore, PIDs) include the integers \mathbb{Z} , all fields \mathbb{F} , and polynomial rings $\mathbb{F}[x]$ for any field \mathbb{F} . Note that the polynomial ring $\mathbb{Z}[x]$ is not a PID. For example, the ideal $(2, x) = \{2f + xg : f, g \in \mathbb{Z}[x]\} \subseteq \mathbb{Z}[x]$ is not principal.

Next, we identify terminology involving properties of modules.

Definition A3.3. An R -module M is called **cyclic** if it can be generated by a single element $m \in M$, i.e. $Rm = \{rm : r \in R\} = M$.

Note that R can be viewed as an cyclic R -module generated by its identity element $1 \in R$. Similarly, given $d \in R$, the quotient $R/(d)$ is a cyclic R -module with underlying abelian group $R/(d)$ and action $s \cdot [r] = [rs]$ generated by the coset $[1]$.

We also talk about notions of torsion and free involving modules over a PID. We provide definitions involving these below.

Definition A3.4. Let M be a module over a commutative ring R . A subset $B \subseteq M$ is called a **basis** of M if each element of M can be written as a unique R -linear combination of elements of B . An R -module M is called **free** if it has a basis.

Definition A3.5. Let M be a module over a commutative ring R . We say that $m \in M$ is a **torsion element** of M if there exists some nonzero $r \in R$ such that $rm = 0$. If M has no torsion elements, then M is called **torsion-free**. If every element of M is torsion, then M is called a **torsion module**.

Modules over PIDs are considered “well-behaved” in that torsion and free can be considered distinct notions in such modules. A rigorous discussion of how this works is outside the scope of this paper and we refer to [DF03, Chapter 12] and [AW92, Chapter 3] for more details. Listed below are a number of relevant results for reference.

1. Let R be an integral domain and let M be an R -module. Then, the subset $T(M) := \{m \in M : rm = 0 \text{ for some nonzero } r \in R\} \subseteq M$ is a submodule of M and the quotient $M/T(M)$ is torsion-free [AW92, Proposition 3.2.18].
2. Let R be an integral domain and let M be an R -module. If M is free, then M is torsion-free [AW92, Proposition 3.4.8]. Conversely: If M is a finitely generated torsion-free module over a PID R , then M is free [AW92, Theorem 3.6.6]. Therefore, a module over a PID is free if and only if it is torsion-free.
3. Let M be a module over a PID R . If M is finitely generated, then $M \cong T(M) \oplus M/T(M)$ [AW92, Corollary 3.4.17]. Since $M/T(M)$ is torsion-free and therefore free, each element $m \in M$ decomposes uniquely into a “free” component and “torsion” component.
4. Let R be a commutative ring with identity. If M is a free R -module with a finite basis, then every basis of M has the same number of elements [AW92, Corollary 3.6.18]. This allows the notion of “rank” to be well-defined for free R -modules and implies that all free R -modules of the same rank are isomorphic.
5. If M and N are finitely generated modules over a PID R , then $M \cong N$ if and only if $T(M) \cong T(N)$ and $\text{rank}(M/T(M)) = \text{rank}(N/T(N))$ [AW92, Corollary 3.6.20]. This determines that the “free part” of M , i.e. the summand $M/T(M)$ is unique up to isomorphism.

This motivates the following terminology for finitely generated modules over a PID.

Definition A3.6. Let M be a finitely generated module over a PID R . Define the **torsion component** $T(M)$ (also **torsion submodule**) and the **free component** $F(M)$ of M to be submodules of M given as follows:

$$T(M) := \left\{ m \in M : rm = 0 \text{ for some nonzero } r \in R \right\} \quad \text{and} \quad F(M) := M/T(M)$$

Let the **rank** of M , denoted $\text{rank}(M)$, be the cardinality of any basis of $F(M)$.

In Section 4.1 and Section 4.2, we examine a proof of the Structure Theorem (Theorem 4.1.1) in \mathbf{Mod}_R and discuss how we can calculate invariant factor decompositions for finitely generated modules over a PID using presentations and matrix reduction.

In this expository paper, we usually define R -modules (particularly in examples) using formal sums of some set of indeterminates for convenience. We include relevant definitions below.

Definition A3.7. Let R be a PID. Let $A = \{a_1, a_2, \dots, a_n\}$ be some set of indeterminates.

- i. An **R -formal sum in A** is an expression in the form $a = \sum_{i=1}^n r_i a_i = r_1 a_1 + r_2 a_2 + \dots + r_n a_n$ for some $r_1, \dots, r_n \in R$. When R is unambiguous, we may refer to $\sum_{i=1}^n r_i a_i$ as a formal sum in A . We may also say that a is a **formal sum of elements of A** .
- ii. The **R -module generated by A** is the free R -module with basis in correspondence with A . We usually

write elements of $R\langle A \rangle$ as formal sums in A with coefficients in R , i.e.

$$R\langle A \rangle := \left\{ \sum_{i=1}^n r_i a_i : r_i \in R, a_i \in A \text{ for all } i = 1, \dots, n \right\}$$

When an R -module M is generated this way, we may write $M = R\langle A \rangle$ or $M = R\langle a_1, \dots, a_n \rangle$.

Remark. We somewhat abuse notation and also use $R\langle - \rangle$ to generate free submodules of $R\langle A \rangle$. For example, let $M = R\langle a, b \rangle$. Then, we may write $R\langle 2b \rangle$ to refer to the submodule $R\langle 2b \rangle = \{r \cdot 2b : r \in R\} = (2b)$, treating $2b$ as an element of $R\langle b \rangle$ rather than an indeterminate distinct from b . In this case, we prefer to write $R\langle 2b \rangle$ to emphasize that $R\langle 2b \rangle$ is free.

Below, we identify notation for matrices related to free modules and homomorphisms between free modules, adapted from [AW92, Section 4.3].

Definition A3.8. Let M and N be modules over a commutative ring R with ordered bases $\mathcal{A} = (\alpha_1, \dots, \alpha_m)$ and $\mathcal{S} = (\sigma_1, \dots, \sigma_n)$ respectively. Let $\varphi : N \rightarrow M$ be a module homomorphism.

- i. The **coordinate vector** $[\alpha]_{\mathcal{A}} \in M_{m,1}(R)$ of $\alpha = \sum_{j=1}^m r_j \alpha_j \in M$ **relative to** \mathcal{A} is the column vector given by $[\alpha]_{\mathcal{A}}(j) = r_j$ for $j \in \{1, \dots, m\}$, i.e. $[\alpha]_{\mathcal{A}} = (r_1, r_2, \dots, r_m)^\top$.
- ii. The **matrix** $[\varphi]_{\mathcal{A},\mathcal{S}} \in M_{m,n}(R)$ of φ **relative to** \mathcal{A} and \mathcal{S} is the matrix given by

$$\text{col}_i[\varphi]_{\mathcal{A},\mathcal{S}} = [\varphi(\sigma_i)]_{\mathcal{A}} \quad \text{for all } i \in \{1, \dots, n\}.$$

If \mathcal{A} and \mathcal{S} are the standard bases for M and N respectively (if such are defined), then we may suppress the subscripts and write $[\alpha]$ and $[\varphi]$ for $[\alpha]_{\mathcal{A}}$ and $[\varphi]_{\mathcal{S},\mathcal{A}}$ respectively.

Appendix A4. A Brief Review of Categories and Functors

Below, we list a number of basic definitions and results relevant in this expository paper, mostly taken from the text *Category Theory in Context* [Rie16, Chapter 1] by Emily Riehl.

Note that there are some topics in category theory that are relevant to what is discussed in this paper (albeit some tangentially) that are not included in this appendix, e.g. abelian categories and homological algebra. For said topics, we recommend *Introduction to Homological Algebra* [Rot08, Chapters 1 and 5] by Joseph Rotman and *Introduction to Homological Algebra* [Wei95, Chapter 1 and Appendix A] for introductory reading.

Definition A4.1. [Rie16, Definition 1.1.1] A **category \mathbf{C}** consists of a class $\text{Ob}(\mathbf{C})$ of **objects** and a class $\text{Hom}(\mathbf{C})$ of **morphisms** such that all the following conditions are satisfied:

- i. Each morphism has specified **domain** and **codomain** objects. This relationship is typically denoted by $f : X \rightarrow Y$, where $f \in \text{Hom}(\mathbf{C})$ with X and Y as its domain and codomain objects respectively.
- ii. Each object $X \in \text{Ob}(\mathbf{C})$ has a designated **identity morphism** $1_X : X \rightarrow X$.
- iii. For any pairs for morphisms $f, g \in \text{Hom}(\mathbf{C})$ such that $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ (i.e. f and g are **composable**), there exists a specified **composite morphism** $g \circ f \in \text{Hom}(\mathbf{C})$ such that $gf : X \rightarrow Z$. The collection of these assignments is usually referred as the **composition law** of \mathbf{C} .
- iv. Composition is unital with identity morphisms, i.e. for any morphism $f \in \text{Hom}(\mathbf{C})$ with $f : X \rightarrow Y$, $f \circ 1_X = f$ and $f = 1_Y \circ f$.
- v. Composition is associative, i.e. for any $f, g, h \in \text{Hom}(\mathbf{C})$ with $f : X \rightarrow Y$, $g : Y \rightarrow Z$, and $h : Z \rightarrow W$, the composites $h \circ (g \circ f)$ and $(h \circ g) \circ f$ must be equal.

A **subcategory \mathbf{D}** of some category \mathbf{C} is another category whose classes of objects and of morphisms are subclasses of those of \mathbf{C} .

Remark. For our purposes, it suffices to know that classes act similarly to sets, in that they are collections of objects. They are distinct from sets to avoid paradoxes such as Russell's paradox.

The class of morphisms and the composition law of a category determines the notion of similarity, i.e. isomorphism relations, in said category. Below, we include a category-level definition of isomorphisms.

Definition A4.2. A morphism $f : X \rightarrow Y$ in a category \mathbf{C} is called an **isomorphism** if there exists another morphism $g : Y \rightarrow X$ such that $g \circ f = \text{id}_X$ and $f \circ g = \text{id}_Y$. The objects X and Y are **isomorphic** in \mathbf{C} if there exists a isomorphism $f : X \rightarrow Y$. In this case, we write $X \cong_{\mathbf{C}} Y$. If the category \mathbf{C} is clear from context, we usually write $X \cong Y$ for brevity.

Note that, in general, equality between objects and morphisms in a category \mathbf{C} are considered in the abstract sense. This is because not all categories are concrete categories. Loosely speaking, a category \mathbf{C} is a *concrete category* if each object $X \in \text{Ob}(\mathbf{C})$ are sets (usually with some additional structure). If a category is not concrete, then notions such as *injective maps* and *surjective maps*, as defined in set theory, are ambiguous. For morphisms that act similarly, new category-level terms are introduced. For example, injective maps are *monomorphisms* and surjective maps are *epimorphisms*.

With the exception of $\mathbf{Poset}(I, \leq)$ (see Definition A4.3 below), all relevant categories discussed in this expository paper are concrete categories. For clarity, we prefer the set-level terminology when applicable. Listed below are some of these categories, taken from [Rie16, Example 1.1.3] and [BS14]. Note that the composition law on the categories below are given by the usual function composition.

1. **Set** denotes the category of sets (as objects) and functions (as morphisms). Isomorphisms in this category are *bijections*, as is usually defined in set theory.
2. **Top** denotes the category of topological spaces (as objects) and continuous maps (as morphisms). Isomorphisms in this category are homeomorphisms. Recall that composition of continuous maps are continuous.
3. Let \mathbb{F} be a field. **Vect** $_{\mathbb{F}}$ denotes the category of \mathbb{F} -vector spaces (as objects) and \mathbb{F} -linear maps (as morphisms). Isomorphisms in **Vect** $_{\mathbb{F}}$ are usually called isomorphisms, although we may call these vector space isomorphisms for clarity, **Vect** $_{\mathbb{F}}$ has a subcategory denoted **FinVect** $_{\mathbb{F}}$ composed of finite-dimensional \mathbb{F} -vector spaces, with the class of morphisms appropriately restricted.
4. Let R be a ring. **Mod** $_R$ (or ${}_R\mathbf{Mod}$) refers to the category of right (or left) R -modules as objects and R -module homomorphisms as morphisms. Note that if R is a commutative ring, then **Mod** $_R = {}_R\mathbf{Mod}$.

Since we generally require R to be a PID in this paper, R is commutative and we refer to **Mod** $_R$ as simply the category of R -modules. We also prefer using the symbol **Mod** $_R$ for formatting purposes. As with **Vect** $_{\mathbb{F}}$, we may refer to isomorphisms in **Mod** $_R$ R -module isomorphisms for clarity.

5. Let R be a PID. **Ch-Mod** $_R$ denotes the category of chain complexes of R -modules as objects and chain maps as morphisms. In this paper, we use the following definition for chain complexes: A chain complex C_* of R -modules is a \mathbb{Z} -indexed collection of R -modules C_n and R -module homomorphisms $\partial_n : C_n \rightarrow C_{n-1}$. In this case, we write $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ and call ∂_n the *differentials* of C_* .

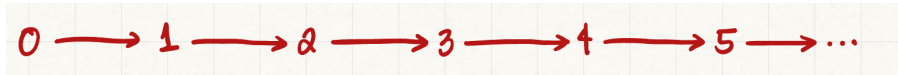
In this paper, for brevity, we usually call categories only by their objects, e.g. we say **Mod** $_R$ is the category of R -modules, without reference to R -module homomorphisms.

Poset categories play a huge role in this expository paper. We provide a definition taken from [Rie16, Example 1.1.4], but with notation slightly changed for clarity.

Definition A4.3. The **poset category** **Poset** (I, \leq) , also called **indexing category**, induced by a poset (I, \leq) is the category constructed as follows:

- i. The elements of I are exactly the objects of **Poset** (I, \leq) .
- ii. For all $a, b \in I$, there exists a unique morphism $a \rightarrow b$ if and only if $a \leq b$.
- iii. Given all $a \in I$, the morphism $a \rightarrow a$ is the identity morphism of the object a .
- iv. For all $a, b, c \in I$ with $a \leq b \leq c$, the composition law on the morphism is given by $(c \leftarrow a) = (c \leftarrow b) \circ (b \leftarrow a)$.

Posets can usually be described and defined using simple directed graphs. In particular, a poset (I, \leq) corresponds to a simple directed graph where the vertices represent the elements of I and the arrows represent the relations in \leq . For example, the poset (\mathbb{N}_0, \leq) has the corresponding directed graph:



Note that \mathbb{N}_0 is a totally ordered set and therefore, is also a poset. This interpretation also extends to the induced category **Poset** (I, \leq) , wherein the vertices and arrows of the graph represent the objects and morphisms of **Poset** (I, \leq) respectively.

Since posets have to be antisymmetric, the only isomorphisms on **Poset** (I, \leq) are the identity morphisms. Given any $a, b \in I$ with $a \leq b$, there cannot be a relation $b \leq a$ unless $a = b$. In that case, there does not exist a morphism $b \rightarrow a$. In other words, the notions of equality and isomorphisms are the same in **Poset** (I, \leq) .

Functors are also present in a significant amount in this expository paper. Roughly speaking, functors are functions between categories, but instead of just having assignments on the objects, there are also assignments on the morphisms. We provide a definition of functors below taken from [Rie16, Definition 1.3.1]

Definition A4.4. A functor $F : \mathbf{C} \rightarrow \mathbf{D}$ between categories \mathbf{C} and \mathbf{D} consists of the following:

- i. An (object) assignment of each object X in \mathbf{C} to some object A in \mathbf{D} , denoted $F(X) = A$.
- ii. A (morphism) assignment of each morphism $f : X \rightarrow Y$ in \mathbf{C} to a morphism $h : F(X) \rightarrow F(Y)$ in \mathbf{D} , denoted $F(f) = h$. Note that the domain and codomain of the morphism assignment are determined by the object assignment.
- iii. The functor must respect composition, i.e. for any composable pair f and g in \mathbf{C} , $F(f)$ and $F(g)$ must also be composable in \mathbf{D} , i.e. $F(f \circ g) = F(f) \circ F(g)$.
- iv. The functor must respect identity maps, i.e. the functor assigns the identity morphism id_X of any object X in \mathbf{C} to the identity morphism $\text{id}_{F(X)}$ of $F(X)$ in \mathbf{D} , i.e. $F(\text{id}_X) = \text{id}_{F(X)}$.

We call \mathbf{C} the **domain category** and \mathbf{D} as the **codomain category**. The last two conditions listed above are often called **functorial properties**.

Remark. More generally, the statement above defines a *covariant functor*. There is another type of functor called a *contravariant functor* but it is not relevant to this paper.

A number of constructions in algebraic topology can be considered functors, which is appropriate given that category theory is said to be motivated by observations on algebraic topology theory.

Example A4.5. Listed below are some functors related to homology and simplicial homology, taken from [Rie16, Example 1.3.2] and [Hat02, Section 2.3].

1. The n^{th} **singular homology functor** $H_n(-; R) : \mathbf{Top} \rightarrow \mathbf{Mod}_R$ sends a topological space X to its n^{th} homology group $H_n(X; R)$ with coefficients in a PID R and $n \in \mathbb{N}_0$. When $R = \mathbb{Z}$, the n^{th} homology functor is usually written as $H_n(-)$, suppressing the reference to coefficient ring \mathbb{Z} . The morphism assignment sends an inclusion map $i : X \rightarrow Y$ to an induced homomorphism $i_* : H_n(X; R) \rightarrow H_n(Y; R)$ as denoted in [Hat02].
2. Abstractly, homology can be calculated from any chain complex. For each $n \in \mathbb{Z}$, we have the n^{th} **chain homology functor** $H_n(-) : \mathbf{Ch-Mod}_R \rightarrow \mathbf{Mod}_R$ that sends a chain complex $C_* = (C_n, \partial_n)_{n \in \mathbb{Z}}$ of R -modules to its n^{th} homology group by the following construction:

$$H_n(C_*) = \ker(\partial_n) / \text{im}(\partial_{n+1})$$

The morphism assignment sends the chain map $\varphi_* : C_* \rightarrow A_*$ with $\varphi_* = (\varphi_n : C_n \rightarrow A_n)_{n \in \mathbb{Z}}$ to the homomorphism $H_n(C_*) \rightarrow H_n(A_*)$ induced by the cokernel operation on $\varphi_n : C_n \rightarrow A_n$ and $\varphi_{n+1} : C_{n+1} \rightarrow A_{n+1}$.

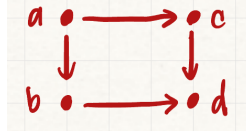
3. The “free” functor $F : \mathbf{Set} \rightarrow \mathbf{Mod}_R$ sends a set X to the free R -module $R\langle X \rangle$, with X treated as a set of indeterminates (see Definition A3.7) and sends a function $f : X \rightarrow Y$ to the unique homomorphism $R\langle X \rangle \rightarrow R\langle Y \rangle$ defined by mapping basis elements to basis elements.

As mentioned earlier, poset categories $\mathbf{Poset}(I, \leq)$ can be described using graphs. In the same vein, we can illustrate commutative diagrams as functors with some poset category as the domain category. We provide a definition below taken from [Rie16, Definition 1.6.4].

Definition A4.6. A **diagram** is a functor $F : \mathbf{Poset}(I, \leq) \rightarrow \mathbf{C}$ from some poset category $\mathbf{Poset}(I, \leq)$ to some other category \mathbf{C} .

Given a diagram $F : \mathbf{Poset}(I, \leq) \rightarrow \mathbf{C}$, the vertices of the directed graph corresponding to $\mathbf{Poset}(I, \leq)$ are replaced by objects and the arrows by morphisms. The transitivity of (I, \leq) and the composition axiom for functors require the resulting diagram (in the non-category theory-sense) be commutative. We provide a simple example involving squares:

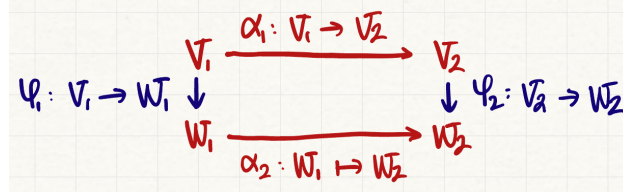
Example A4.7. Let the poset (S, \leq) be given by the following directed graph:



Let V_1, V_2, W_1, W_2 be \mathbb{R} -vector spaces. Let $\alpha_1 : V_1 \rightarrow V_2, \alpha_2 : W_1 \rightarrow W_2$ and $\varphi_1 : V_1 \rightarrow W_1, \varphi_2 : V_2 \rightarrow W_2$ be linear maps. Then, the requirement that $\varphi_2 \circ \alpha_1 = \alpha_2 \circ \varphi_1$ can be restated as follows: Let $F : \mathbf{Poset}(S, \leq) \rightarrow \mathbf{Vect}_{\mathbb{R}}$ be an assignment on the objects and morphisms of $\mathbf{Poset}(S, \leq)$ as follows:

$$\begin{array}{llll} F(a) = V_1 & F(c) = V_2 & F(b) = W_1 & F(d) = W_2 \\ F(a \rightarrow b) = \varphi_1 & F(a \rightarrow c) = \alpha_1 & F(b \rightarrow d) = \alpha_2 & F(c \rightarrow d) = \varphi_2 \end{array}$$

We illustrate the assignment F as a diagram below:



Then, $\varphi_2 \circ \alpha_1 = \alpha_2 \circ \varphi_1$ if and only if F represents a functor. By definition of functor, the morphisms in the form $x \rightarrow x$ are mapped to identity linear maps. In terms of compositions, we have the following chain of equalities:

$$\varphi_2 \circ \alpha_1 = F(d \leftarrow c) \circ F(c \leftarrow a) \stackrel{*}{=} F(d \leftarrow a) \stackrel{*}{=} F(d \leftarrow b) \circ F(b \leftarrow a) = \alpha_2 \circ \varphi_1$$

The equalities labeled by \star are due to the uniqueness of morphisms in $\mathbf{Poset}(S, \leq)$.

References

- [AW92] William Adkins and Steven Weintraub. *Algebra. An Approach via Module Theory*. Springer, 1992.
- [Bai17] James Bailie. “Abelian Categories and Mitchell’s Embedding Theorem”. In: (2017). This is unpublished. A copy is posted on Bailie’s GitHub page, linked here: <https://jameshbailie.github.io/Papers/>.
- [Bau21] Ulrich Bauer. “Ripser. Efficient Computation of Vietoris-Rips Persistence Barcodes”. In: *Journal of Applied and Computational Topology* 5 (June 2021). DOI: [10.1007/s41468-021-00071-5](https://doi.org/10.1007/s41468-021-00071-5). URL: <https://arxiv.org/abs/1908.02518>.
- [BM21] Peter Bubenik and Nikola Milićević. “Homological Algebra for Persistence Modules”. In: *Foundations of Computational Mathematics* 21.5 (Jan. 2021), pp. 1233–1278. DOI: [10.1007/s10208-020-09482-9](https://doi.org/10.1007/s10208-020-09482-9). URL: <https://arxiv.org/abs/1905.05744>.
- [BS14] Peter Bubenik and Jonathan A. Scott. “Categorification of Persistent Homology”. In: *Discrete & Computational Geometry* 51.3 (Jan. 2014), pp. 600–627. DOI: [10.1007/s00454-014-9573-x](https://doi.org/10.1007/s00454-014-9573-x). URL: <https://arxiv.org/abs/1205.3669>. The first version of this paper was submitted to ArXiv on May 2012. A revised copy was added on January 2014.
- [CCS14] Frederic Chazal, William Crawley-Boevey, and Vin de Silva. *The Observable Structure of Persistence Modules*. 2014. arXiv: [1405.5644](https://arxiv.org/abs/1405.5644) [math.RT]. A free copy is posted on Chazal’s website, linked here: <https://geometrica.saclay.inria.fr/team/Fred.Chazal/>.
- [CEM06] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. “Vines and Vineyards by Updating Persistence in Linear Time”. In: vol. 2006. June 2006, pp. 119–126. DOI: [10.1145/1137856.1137877](https://doi.org/10.1145/1137856.1137877). URL: <https://www.mrzv.org/publications/vineyards/socg06/>.
- [CK18] René Corbet and Michael Kerber. “The Representation Theorem of Persistence, Revisited and Generalized”. In: *Journal of Applied and Computational Topology* 2.1-2 (July 2018), pp. 1–31. DOI: [10.1007/s41468-018-0015-3](https://doi.org/10.1007/s41468-018-0015-3). URL: <https://doi.org/10.1007/s41468-018-0015-3>.
- [CSGO13] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Oudot. *The Structure and Stability of Persistence Modules*. 2013. arXiv: [1207.3674](https://arxiv.org/abs/1207.3674) [math.AT]. URL: <https://arxiv.org/abs/1207.3674>. A textbook of the same title has been published by Springer. The link on Chazal’s website, <https://geometrica.saclay.inria.fr/team/Fred.Chazal/>, redirects to the ArXiv page.
- [CZZ18] L. Cossu, P. Zanardo, and U. Zannier. “Products of elementary matrices and non-Euclidean principal ideal domains”. In: *Journal of Algebra* 501 (2018), pp. 182–205. ISSN: 0021-8693. DOI: <https://doi.org/10.1016/j.jalgebra.2017.11.051>. URL: <https://www.sciencedirect.com/science/article/pii/S002186931730666X>.
- [DF03] D.S. Dummit and R.M. Foote. *Abstract Algebra*. Wiley, 2003. ISBN: 9780471433347.
- [ELZ02] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. “Topological Persistence and Simplification”. In: *Discrete & Computational Geometry* 2002 (Jan. 2002), pp. 511–533. DOI: [10.1007/s00454-002-2885-2](https://doi.org/10.1007/s00454-002-2885-2). A free copy is posted by Shmuel Weinberger, linked here: <https://math.uchicago.edu/~shmuel/AAT-readings/Data%20Analysis%20Edelsbrunner-Letscher-Zomorodian.pdf>.
- [Hat02] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002. ISBN: 0-521-79540-0. A free copy is posted by Hatcher, linked here: <https://pi.math.cornell.edu/~hatcher/>.
- [Loe23] Clara Loeh. *A comment on the structure of graded modules over graded principal ideal domains in the context of persistent homology*. 2023. arXiv: [2301.11756](https://arxiv.org/abs/2301.11756) [math.AC].

- [Mar] Tom Marley. “Graded Rings and Modules”. In: (). There are lecture notes posted in <https://math.unl.edu/tom-marley> for a course in the University of Nebraska.
- [Mat03] Jiří Matoušek. *Using the Borsuk-Ulam Theorem. Lectures on Topological Methods in Combinatorics and Geometry*. Springer Book Archive, 2003.
- [Mun93] James Munkres. *Elements of Algebraic Topology*. CRC Press, 1993.
- [NO04] C. Nastasescu and F. van Oystaeyen. *Methods of Graded Rings*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2004. ISBN: 9783540409984.
- [Pee11] Irena Peeva. *Graded Syzygies*. Algebra and Applications. Springer London, 2011. DOI: [10.1007/978-0-85729-177-6](https://doi.org/10.1007/978-0-85729-177-6).
- [Rie16] Emily Riehl. *Category Theory in Context*. Dover Publications, 2016. ISBN: 978-0-486-82080-4. A free copy is posted on Emily Riehl’s Github page, linked here: emilyriehl.github.io/files/context.pdf,
- [Rot08] Joseph J. Rotman. *An Introduction to Homological Algebra*. Universitext. Springer New York, 2008. ISBN: 9780387683249.
- [Rot88] Joseph J. Rotman. *An Introduction to Algebraic Topology*. 1st ed. Graduate Texts in Mathematics 119. Springer-Verlag New York, 1988. ISBN: 9780387966786,0387966781.
- [Web85] Cary Webb. “Decomposition of Graded Modules”. In: *Proceedings of the American Mathematical Society* 94.4 (Aug. 1985). A free copy from the AMS is posted here: <https://www.ams.org/journals/proc/1985-094-04/S0002-9939-1985-0792261-6/>.
- [Wei95] Charles A. Weibel. *An Introduction to Homological Algebra*. Cambridge Studies in Advanced Mathematics 38. Cambridge University Press, 1995. ISBN: 0521559871.
- [ZC05] Afra Zomorodian and Gunnar Carlsson. “Computing Persistent Homology”. In: *Discrete & Computational Geometry* 33.2 (Feb. 2005), pp. 249–274. URL: <https://doi.org/10.1007/s00454-004-1146-y>. A free copy from the Geometric Computation Group, part of the CS department at Stanford University, is available here: <https://geometry.stanford.edu/papers/zc-cph-05/zc-cph-05.pdf>.