STRUCTURAL FEATURES OF PERSISTENT HOMOLOGY AND THEIR ALGORITHMIC TRANSFORMATIONS

A Dissertation

presented to

the Faculty of the Graduate School

University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

ANDREI PAVLICHENKO

Dr. Jan Segert, Dissertation Supervisor

MAY 2021

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

STRUCTURAL FEATURES OF PERSISTENT HOMOLOGY AND THEIR ALGORITHMIC TRANSFORMATIONS

presented by Andrei Pavlichenko, a candidate for the degree of Doctor of Philosophy

of Mathematics, and hereby certify that in their opinion it is worthy of acceptance.

Professor Jan Segert

Professor Calin Chindris

Professor Stephen Montgomery-Smith

Professor David Retzloff

ACKNOWLEDGEMENTS

I want to express my enormous gratitude to my advisor, Professor Jan Segert for his relentless support throughout my studies and our research projects. His vision, knowledge, and rationalization were not only crucial for settling pivotal questions and various uncertainties in our work, but also left me with a persistent impression of the standard of rigor and utility one should exercise to understand and develop important mathematical principles. This was only fueled further by his great enthusiasm and fun analysis of the topics at hand.

I would also like to thank Professor Calin Chindris, Professor Stephen Montgomery-Smith, and Professor David Retzloff as members of my dissertation committee and the work it took to examine everything in a timely manner.

My occasional conversations with Dr. Chindris and Dr. Montgomery-Smith were insightful and clarifying on some issues both academic and organizational, helping me to ease the planning of important things. I've also enjoyed their classes, and for the same reason I would like to thank other members of the Mathematics faculty with whom I've had the honor to study by watching their well-thought allegories and the zeal with which they worked through proofs and examples during lectures.

I also want to give a big shout out to everyone at the Mathematics department support staff, who have always been very friendly and immensely helpful with even the most puzzling and sudden of the concerns that I've had.

To my friends and colleagues over the years in the Mathematics department and beyond: I am glad for all the pleasant banter and the occasional (moderate) mischief, which will be fondly remembered. “Dot calculus” will not be forgotten.

And of course, I am forever grateful to my family who have observed and encouraged me throughout these years and have remained positive about my plans while helping constructively to keep my goals and dreams in place. I think about our stories and conversations every day, finding joy for near future and course for the far future.

Now, it begins!

Contents

STRUCTURAL FEATURES OF PERSISTENT HOMOLOGY AND THEIR ALGORITHMIC TRANSFORMATIONS

Andrei Pavlichenko

Dr. Jan Segert, Dissertation Supervisor

ABSTRACT

We re-examine the theory and orthodox methods that underlie the study of persistent homology, particularly in its calculation of homological cycle representatives that are associated to persistence diagrams. A common background to the subject covers several aspects: schemes to process input data (embedding it in a low-dimensional manifold), categorical descriptions of persistence objects, and algorithms by which the barcode summarizing the homology is found.

We overview these aspects, focusing on filtered simplicial complexes, traditional computation of persistent homology, and the stability theorem for barcodes. By reformulating these notions in the language of category theory, we can speak more plainly on some recurring notions that are relevant to our discussion. This ultimately sets up for vector space filtrations that prove to be suitable tools for codifying the homology of complexes, including the (co)images and (co)kernels arising from morphisms of complexes.

The main body of work then presents an alternative approach to persistent homology, based on filtrations of vector spaces. We elaborate with an interesting example whose persistent homology is readily computed as a quotient of appropriate filtrations; in the process, we produce a representative basis of homological cycles, a step that is often overlooked in existing literature. The proposed algorithm is also notable

in that it easily handles the calculation of (co)images and (co)kernels for persistent morphisms, supplying us with the same level of detail; while other algorithms do exist for computing the barcodes of these universal objects, such methods are not easily generalizable. Finally, we compute appropriate homological cycles and use a certain algorithmic matching scheme that both implies the usual barcode matching and attempts to better interpret this interesting behavior.

Chapter 1

Introduction

The purpose of this thesis is to reexamine and extend the computational foundation of *persistent homology*, with primary focus on how this is used to figure out structural components of geometric objects constructed from point cloud data as well as those induced by their mappings. Persistent homology is generally considered to be an integral engine in the discipline of *topological data analysis (TDA)*, a peculiar approach to data analysis that has received judicious attention over the last decade from a broad range of practitioners in data science. Some of the more recent studies that have highlighted the use of TDA include [12], [29], [31], [33], [34], [45], [48], and [49]. A cursory of account of this setting is in order.

An inherent part of mastering a given subject is being able to construe its underlying principles and how they interplay with one another. We see this in well-posed scientific theories which propose laws having the power to predict and explain certain phenomena. This helps build better models and instruments for taking on tentative problems, and improves our ability to compare things that adhere to the framework of these laws. Furthermore, this lets us posit new (and possibly conflicting) directions for learning, for we acknowledge that the current body of knowledge may be incomplete.

In analyzing existing data, we typically express these underlying principles through closest-fit continuous models and discrete classification schemes. As real-life data contains stochastic effects (from lack of control on its environment, and possibly experimental error), the language for these principles must either gradually change or become enriched with more degrees of freedom; for instance, linear regression may be expanded to include higher-order effects such as correlation between independent variables, while classification schemes tend to become finer. Still, it is imperative to understand how reliable the predictors of these models presently are prior to determining vectors for their improvement. After all, a signature of a good model is the strength with which the underlying principles are demonstrated to explain observed phenomena, even while the model appears abbreviated for clear presentation and/or computational efficiency.

Various methods of TDA have been shown to excel in extracting meaningful features from noisy data. A notoriously complicated and dispersed task for various machine learning disciplines, feature extraction/recovery aims to produce a model of some system that (1) is supported by sample data (subject to some measure of inaccuracy), while (2) using less resources to specify it (by analogy to data compression). Among all approaches that are used to achieve this, TDA does this on the assumption that continuous (homeomorphic) changes to the coordinate space used to specify our data should not in general entail drastic changes to the totality of observed features. This invokes the topological aspect of persistent homology, which can be summarized as an interplay of its two underlying subjects, topology and homological algebra.

A rigorous treatise of topology can be traced back to work done on Euclidean tilings by Kepler in [35] and Euler's combinatorial treatment of convex polytopes in his famous paper [26]. The subject would steadily develop over the course of the next few centuries, eventually formalizing the notions of "closeness" and "continuity" for very abstract spaces. At the same time, the combinatorial treatment with simplicial polytopes was successfully applied to understanding the structure of real and complex manifolds, driven in the works by B. Riemann and H. Poincare. These early developments culminated in a classification theory for manifolds and various sheaves of functions, which today are widely-taught in courses on complex manifolds and Riemannian geometry.

By early twentieth century, the subject has manifested into broader homology theory. The idea is that geometric features on simplicial polytopes can in certain ways be endowed with algebraic operations, by which they can be combined into complex structures that may represent discernable "features" of a space. Going further, certain classes of ordinary functions on spaces then give rise to operators on these algebras, representable by matrices if the algebras are finitely-generated over the integers $\mathbb{Z}$ (appealing to the combinatorial sense) or over over a field (allowing us to use linear algebra, which is a well-understood subject). The original impetus for this was being able to compute the "boundary" of geometric features, represented by a signed sum of simplices; the process is reminiscent of ordinary matrix reduction.

As the theory of homological algebra developed, some have found it beneficial to reinterpret all these elements in the language of category theory. One reason for this is that categorical constructions, especially those adhering to a universal property (associated to diagram limits and colimits), are ubiquitous in that they do well to handle abstraction and describe various mathematical constructions. This proves useful for our purposes as well, when we transition between the various steps involved in persistent and when we carry out similar calculation over different classes of objects.

The integral homology group contains all the information about a manifold in as much as it can be described by inherent simplicial features; the universal coefficient theorem further allows us to compute homology with values in any other ring (including a field). Nonetheless, the complexity of freely-generated abelian group is often superseded by the simplicity of vector space representations of homology, which it often preferred in experiments. For example, the homological invariants in the latter case are often trivial to capture, such as dimension and operator rank. Doing so with the integers, requires use of factorization such as the Smith Normal Form, with only very limited results about the invariants of abelian maps.

Now, although pure homology and topology are hardly the highlight of this report, it is inevitable for these subjects to find their way into describing persistent homology. It distinctly began to appear at around the start the early 1990s, with separate attempts to recover some information about the homology of a space from either: the sublevel sets of selected Morse functions, or a random point cloud sample; more information can be found in the introduction to [21] and [10]. These attempts founded a paradigm (formulated in [20])) that has been central to the subject ever since: the goal is to reconstruct the homology of a space by enriching it with many "layers" that consolidate into a *filtration*. Significant features of the space – those that point to its connected components, holes, hollow regions (etc.) and indicative of their relative "size" (within the space), are somehow expected to appear in many of these layers.

That is, they “persist”.

A major breakthrough occurred in 2003 with the publication of a seminal algorithm [60] for efficiently computing the summary of persistent homology, called the *barcode* of complex. This was closely followed by the publication of the celebrated Stability Theorem in [15] which ensured that this summary is robust to small changes in data input. Throughout the following years these results have been reviewed and reimplemented, with various versions for both generalized and specialized uses appearing in other prominent papers such as [16], [55], [4], and [9].

One component of homology that is often overlooked is calculating the actual simplicial chain representatives, manifesting as subsets of the underlying simplicial complex. The Carlsson-Zomorodian algorithm, which uses a matrix reduction procedure, does actually compute so-called “creators” and “destroyers” whose pairings correspond to individual indecomposable objects in a barcode. But insofar as forming a coherent basis for the chain complex or some of its important subspaces, these creators and destroyers require additional work. For a large majority of applications, this appears to be acceptable practice: just the barcode of the chain complex matters, while inferential analysis on the cycle representative can be done later (if at all).

This is evident of ordinary homology and its applications as well; for instance, the integral homology of a manifold may be described by “the greatest number” of non-homotopic closed loops, while accounting for possible parity such as that witnessed on a Klein bottle. There are several justifications for doing this; one has already been mentioned, in that homology is often merely intended to classify geometric structures, allowing us to distinguish a 2-torus and a 2-sphere, for example. Other reasons include computational difficulty, such as that involved in solving a differential flow on a manifold to extract usable cycles; another problem is that invariants for persistent homology with integral coefficients is very difficult to find. Finally, there is an epistemic objection that simplicial chains are only *representatives* of a particular homology class up to equivalence; as a consequence, this complicates the problem by needing selecting "optimal" such representatives. All of this is generally true.

However, it shall be counter-challenged that in persistent homology not only is it tractable to find appropriate simplicial chain representatives of barcode invariants, but choosing to omit them from the final output of the calculation obstructs anyone from many insights to be found in the studied data set via the explicit homology generators of the underlying complex. Moreover, this can be done via an algorithm that is accessible to anyone with background material in familiar linear algebra (assuming that a careful methodology was followed). Since the barcode of a filtered space is a homological invariant, this procedure will also produce the expected barcode – though it is interpreted somewhat differently from the usual Jordan-block pairing that is present in the reduced form of the differential matrix.

This carries some other benefits. Not any less than the persistent homology of a space, it is also interesting to compute the persistent homology of other objects that are derived from it – such as the image of a map of chain complexes. Currently, this leads practitioners to use heavy machinery from category theory (which however are of rudimentary importance) in order to begin working the computation. Interpreting this in terms of basic vector spaces operations alleviates this somewhat bulky burden and allows us to use the algorithm directly. This method is also slightly more general than the usual algorithm in the sense that allows us to work with degenerate filtrations. The Carlsson-Zomorodian algorithm requires us to select a basis of the canonical filtration of the chain complex from the start, producing potentially varying results. Again, this is not evident when one is simply looking to compute the and birth/death time of a simplicial feature, which is invariant under filtration-preserving changes of basis.

Hence the primary objective of this thesis is to establish an alternative persistent homology algorithm based on filtration quotients and how it extends to algorithms for the computation of (co)kernels and (co)images of maps. All the necessary background material will be introduced and referenced as needed. Thus, chapter 2 will be committed to discussing the foundations of persistent homology, starting with ordinary homology, the notion of filtrations and chains of vector spaces, and some commonplace theorems in the subject. Chapter 3 will be spent on discussing the category-theoretic background as it relates to the subject, closing off with some pertinent results to the objects we are working with. All of this is to motivate the procedure that is stated and performed on an interesting example in chapter 4.

Chapter 2

Persistent Homology, and its preliminaries

Computing the homology of a topological space requires executing three separate, but equally important tasks:

- writing a finite combinatorial model of the space
- associating an algebraic complex to this model
- doing matrix operations to extract generators for the complex

Persistent homology, which is surveyed in section 2.3, has an additional complication in that the output of this procedure needs to be consistent across multiple different "layers". However, we'll see that choosing our spatial model appropriately (first task above) eventually allows us to perform a single matrix reduction (using field coefficients) at a global level. This is a surprising feature of the subject, ultimately contributing to its versatility and becoming the basis of our calculations in chapter 4.

We begin with core concepts from algebraic topology, for which one is only required to have some prior exposure to Euclidean geometry/topology and linear algebra.

2.1 Cursory intro to ordinary homology

A hands-on example of a topological space is the standard (Euclidean) n-simplex $\Delta^n \subset \mathbb{R}^{n+1}$. Letting $\mathbf{e_1}, \dots, \mathbf{e_{n+1}}$ denote unit vectors on the positive coordinate axis of $\mathbb{R}^{n+1}$, we define Δ^n as the convex hull of the set $\{\mathbf{e_i}\}_{i=1}^{n+1}$; that is

$$\Delta^n = \left\{ t_1\, \mathbf{e_1} + \cdots + t_{n+1}\, \mathbf{e_{n+1}} \ : t_i \in [0,1], \sum_{i=1}^{n+1} t_i = 1 \right\}$$

This is a convex, compact subspace of $\mathbb{R}^{n+1}$ whose elements are uniquely determined by the "coordinates" (t_i). In $\mathbb{R}^{n+1}$, its interior consists of points where $t_i > 0$ for all $1 \leqslant i \leqslant n+1$; the complement of its interior is the topological boundary of Δ^n. Setting $t_i = 0$ for a various subsets of $i \in \{1, \dots, n+1\}$ yields copies of lower-dimensional simplices called *faces* of Δ^n, whose union comprises its boundary. In particular, the points positioned at each $\mathbf{e_i}$ (ie, where $t_i = 1$) are 0-simplices called *vertices*. A thorough treatment of the topological qualities of standard n-simplices is given in [52].

General topological spaces go beyond subsets of $\mathbb{R}^{n+1}$, which are already abundant with examples that aren't easily represented using standard simplices. Nonetheless, classical homology is meant to capture the degree to which the features of a topological space are really embeddings of standard simplices and how well they "connect". This has been achieved through *singular homology*, based on looking at all (finite formal sums of) continuous maps from standard n-simplices to the topological space. Good classical reference texts that examine this approach include [46] and [58], as well as [8] (the last one is recommended if one is interested in differential forms and singular cohomology).

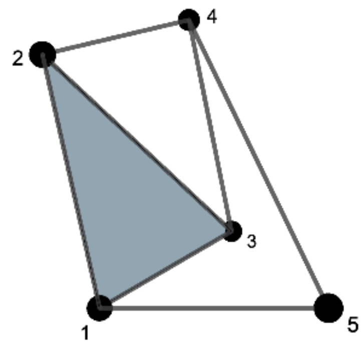
The pressing difficulty with accepting this approach here is computational unfeasibility. The set of all continuous maps from Δ^n to the topological space may be (and for us, will be) uncountably infinite. As such, it is imperative to develop a proper understanding of simplicial complexes, which gives the ability to perform finitely many calculations for completing future examples.

2.1.1 Simplicial Complexes

As shown by the previous discussion, the combinatorics of a standard simplex are described entirely by specifying its set of vertices. In fact, two different embeddings of a standard n-simplex may possibly be distinguished via their image of the vertices. As far as computing homology goes, this distinction will be sufficient.

We start with a finite set, whose elements are understood to be *vertices*. An *(abstract) simplicial complex* K is a collection of nonempty subsets from this set of vertices, with the following property: if $f \in K$ then any subset of f is also in K. Appropriately, the elements of K will be referred to as *faces*. In this setup, the initial set of vertices is sometimes denoted $V(K)$. On occasion, we may take a subcollection $L \subseteq K$ that satisfies being an abstract simplicial complex; we then refer to L as a *subcomplex* of K.

A *geometric realization* $|K|$ of complex K is obtained from a one-to-one mapping of $V(K)$ to the vertices of Δ^N, where $N = |V(K)|$. Every other $f \in K$ is then identified with a face of Δ^N whose vertices are in the image of f under this map. The result is a collection of simplices in bijective correspondence with K; this collection is closed under non-empty intersections of its constituent simplices, as well as under inclusions of every face of a simplex in the collection – these two properties are what merits a collection to be a *simplicial complex*. Finally, taking the union of all the simplices in



$$K = \Big\{\{1\},\{2\},\{3\},\{4\},\{5\},\\ \{1,2\},\{1,3\},\{1,5\},\{2,3\},\{2,4\},\\ \{3,4\},\{4,5\},\{1,2,3\}\Big\}$$

the collection results in the space $|K|$, endowed with the subspace topology from $\mathbb{R}^N$. It is readily verified that different selections of the vertex map result in homeomorphic copies of $|K|$, so the latter is a well-defined space up to homeomorphism.

Example 2.1. The simplicial complex K specified above has the set of vertices $V(K) = \{1, 2, 3, 4, 5\}$. It also has seven 1-simplices and one 2-simplex. Its geometric realization is shown to the left, embedded linearly in $\mathbb{R}^3$.

Now given a topological space X, the first main task (if possible) is to choose a complex K that will act as a "triangulation" (or "tesselation") for X. To be more specific, complex K is required to be such that $|K|$ is *homotopy equivalent* to X, denoted $|K| \simeq X$. Formally, a homotopy equivalence is given by the existence of two continuous functions $f : |K| \to X$ and $g : X \to |K|$ such that $f \circ g$ and $g \circ f$ can be continuously "deformed" to the identity functions $X \to X$ and $|K| \to |K|$, respectively. This *homotopy* of $f \circ g$ and $g \circ f$ to their respective identity functions makes homotopy equivalence of X and $|K|$ a weaker notion than that of a would-be homeomorphism, where instead there is equality of $f \circ g$ and $g \circ f$ to the identity functions; for more details, see section 19 of [46]. While most practical applications proceed with spaces X having some homeomorphic "triangulation" anyway, the assumption that the weaker $|K| \simeq X$ holds instead is necessary in general since X may not possess some desirable properties – particularly that X is Hausdorff and second-countable.

Over more than a century, much research effort has been devoted into obtaining such combinatorial representations of X. The methods can generally be described as inductive processes by which simplicial embeddings are assigned so as to "complete" the space. An illustrative case is with CW-complexes (c.f. [32]) that can be attempted on any Hausdorff space X: fundamentally, the method decomposes X into "n-cells" (whose interior is homeomorphic to that of an n-simplex) satisfying some local finiteness and closure properties. We remark that this way is known to be more general than simplicial complexes, although its derivation for a generic manifold X requires the use of discrete Morse theory (quite an elaborate subject, well surveyed in [41]). Neither of these topics will be pursued henceforth.

In many cases, it is practical to describe a topological space by some choice of an open cover. If $\mathcal{U} = \{U_i\}_{i \in I}$ is an indexed collection of nonempty open sets such that $X \subseteq \bigcup_{i \in I} U_i$ then we may look for "spots" of X where the cover is redundant. Specifically, the sets $I_0 = I$ and $I_k \subseteq \{J \subset I : |J| = k+1\}$ (for all $k \geqslant 1$) describe combinations of indices where $(k+1)$-wise intersections of open sets are nonempty:

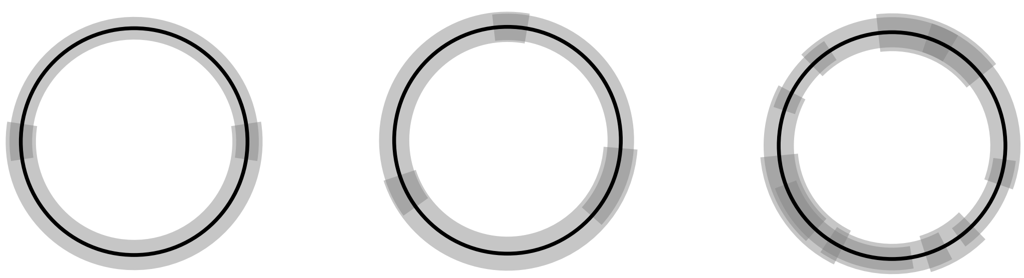
$$\{i_1, \ldots, i_{k+1}\} \in I_k \quad \iff \quad U_{i_1} \cap \ldots \cap U_{i_{k+1}} \neq \varnothing$$

Provided any such selection of $k+1$ indices, any subset of them also implies a nonempty intersection of the respective open sets. In particular, the collection $I_0 \cup I_1 \cup \cdots \cup I_k$ is an abstract simplicial complex for any fixed k; for compact X, we may select a finite cover $\mathcal{U}$ of X to get a maximal collection for some $k \geqslant 1$. The resulting finite complex $K_{\mathcal{U}}$ is known as the *nerve* of the open cover $\mathcal{U}$. The next example demonstrates the well-known criterion of Leray: first recall that a topological space is *contractible* if it is homotopy equivalent to a point. For example, any convex set is contractible, including the standard simplices Δ^n (but not necessarily their unions).

Theorem 2.2. *Let $\mathcal{U}$ be an open cover for a compact set X. If every k-wise intersection of open sets in $\mathcal{U}$ is contractible, then $|K_{\mathcal{U}}| \simeq X$.*

It is customary to refer to a cover $\mathcal{U}$ of X that satisfies the requirement of Theorem 2.2 as a *good cover*. In [22], the authors arrive at a stronger conclusion when X satisfies some additional properties – that $|K_{\mathcal{U}}|$ and X are in fact homeomorphic.

Example 2.3 (1-sphere)**.** Let $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$. Shown below are three open covers of this: (a) by two open sets, (b) by three open sets, and (c) by ten open sets. A greater grayscale intensity indicates nonempty intersections by more open sets.



(a) a “bad” cover (b) a good cover (c) an “uglier” good cover

Figure 2.1: The nerve of the open cover in (a) has two vertices (one for each“hemisphere”) and an edge between them since the open sets have non-empty intersection, so $|K_{\mathcal{U}}| = \Delta^1$. This is inadequate for homology, since Δ^1 is contractible whereas $\mathbb{S}^1$ is not. However, 2-wise intersections of the open sets in (b) are contractible (each is homeomorphic to a 1-simplex) and Theorem 2.2 applies. Shown in (c) is also a good cover, though a more complicated one – some 2-simplices are also present in its nerve, corresponding to regions that are shared by three open sets. This is also a more typical cover, owing to the apparent stochasticity of real data.

The nerve of an open cover will be used in section 2.3 to give a description of the Čech and the Vietoris-Rips filtrations, as used in persistent homology. Alternative methods for generating simplicial complexes do exist however; a special mention is given to the Delaunay triangulation. This method is based on looking at the cell intersections of the Voronoi diagram (for a given vertex set in $\mathbb{R}^n$), and is well-renowned for providing a robust triangulation of the convex hull spanned by the, potentially large, set of vertices. In persistent homology, this has been given rise to the alpha filtrations and its Čech subfiltration; past work on witness complexes [54] has also utilized this construction. Yet another approach is to take an existing simplicial complex and to get a subdivision, thus obtaining another of which the former is a subcomplex. Classically this was mainly associated with barycentric subdivision, by which the barycenters of (selected) existing faces are added as vertices and the faces that contain them are replaced with simplices induced by adjacencies of the new vertices to the old ones. Naturally, this results in a larger simplicial complex that is harder to work with, but is advantageous when two (or more) simplicial structures based on point sets of X are considered.

2.1.2 Simplicial Chains, Cycles, and Boundaries

Having settled on abstract simplicial complexes as a suitable model for geometric data, we can describe their homology. Intuitively, albeit somewhat erroneously, it is to identify subcomplexes on the simplicial complex whose interior is empty (or "hollow"). For instance, the simplicial complex representing the unit circle, as given by the nerve induced from the open cover in (b) of Figure 2.1, is missing a 2-simplex whose boundary is the union of the three edges; adding that to the simplicial collection gives a simplicial realization of the unit disc $\mathbf{D}^2 = \{z \in \mathbb{C} : |z| \leqslant 1\}$.

As shown eventually, this is done entirely through algebra – the trick, originally due to Poincaré, is in how the topological boundary of a simplex is represented. Recalling that the boundary of Δ^n equals the union of its proper faces, we see that its simplicial complex K (a collection that is closed under inclusion of faces) should already encode some of that information. In fact, that union can be taken over faces $\Delta_i^{n-1} = \{(t_1, \dots, t_{n+1}) : t_i = 0\}$ for every $1 \leqslant i \leqslant n+1$, so we can write:

$$\partial_n(\Delta^n) := \Delta_1^{n-1} - \Delta_2^{n-1} + \cdots + (-1)^n \Delta_{n+1}^{n-1} = \sum_{i=1}^{n+1} (-1)^{i-1} \Delta_i^{n-1} \tag{2.1}$$

This is a formal sum that includes all faces of Δ^n having dimension $(n-1)$; it is straightforward to define ∂_j analogously for every j-simplex with $j < n$ (and also setting $\partial_0 \equiv 0$). The coefficients (± 1) have been chosen to highlight the following property: after applying ∂_{j-1} to every summand of $\partial_j(\Delta^j)$ in equation 2.1 the remaining sum should simplify to zero. That is, the topological boundary of a simplicial boundary should be empty (as all points on the boundary of a closed set are in its subspace-induced interior). The coefficients ± 1 are interpreted to represent the spatial orientations of each simplex, in the sense of classical differential calculus.

Abstract simplicial complexes K shall be treated similarly by considering formal sums of faces in the collection. Traditionally, we assume that the vertices $V(K)$ are totally ordered and that this ordering extends to produce an *ordered simplex* for every simplex contained in K. The notation $[v_1, \dots, v_{j+1}]$ refers to an ordered j-simplex whose vertices are $v_1, \dots, v_{n+1}$ in order from least to greatest; for each $1 \leqslant i \leqslant j+1$, the notation $[v_1, \dots, \hat{v}_i, \dots, v_{j+1}]$ refers to the face of $[v_1, \dots, v_{j+1}]$ that omits vertex

v_i. Following the trick above, we define the boundary of any ordered j-simplex:

$$\partial_j\left([v_1, \ldots, v_{j+1}]\right) := \sum_{i=1}^{n+1}(-1)^{i-1}[v_1, \ldots, \hat{v}_i, \ldots, v_{j+1}] \tag{2.2}$$

Now for each $1 \leqslant j \leqslant |V(K)|$, define $C_j = C_j(K)$ to be the set of all finite formal sums of ordered j-simplices contained in K; its elements will simply be called *j-chains* of K. The natural addition operation on formal sums makes each C_j a finitely-generated $\mathbb{Z}$-module, and the (j^{th}) *differential operator* $\partial_j : C_j \to C_{j-1}$ (as given by equation 2.2 for pure j-simplices) extends to other j-chains by linearity. In the next section, we begin seeing how these constructions can work over more general principal ideal domains, particularly fields $\mathbb{F}$. Now, letting C_j be zero unless $1 \leqslant j \leqslant |V(K)|$, this results in the following diagram of $\mathbb{Z}$-modules:

$$\cdots \longleftarrow C_{j-1} \xleftarrow{\partial_j} C_j \xleftarrow{\partial_{j+1}} C_{j+1} \longleftarrow \cdots$$

This diagram is the *chain complex* of K; we sometimes speak of the graded algebra $C_\bullet = \bigoplus_{j\in\mathbb{Z}} C_j$ equipped with the graded differential operator $\partial_\bullet : C_\bullet \to C_\bullet$ having degree -1. The chain complex reproduces that characteristic property where composition of the consecutive differential operators vanishes: $\partial_j \circ \partial_{j+1} = 0$, for all $j \in \mathbb{Z}$. This property has an elementary consequence. Define sets $Z_j = \ker\left(\partial_j : C_j \to C_{j-1}\right)$ and $B_j = \operatorname{im}\left(\partial_{j+1} : C_{j+1} \to C_j\right)$, whose elements are the *j-cycles* and the *j-boundaries*, respectively. Then we have $B_j \subseteq Z_j \subseteq C_j$.

2.1.3 Computing Homology over the Integers and Fields

Following the discussion at the start of the last subsection, we can interpret (and merely at that) the q-chains of a simplicial complex K to generalize some of its subcomplexes; concretely, we can look at the subcomplex of K generated by the summands of any q-chain. In this interpretation, the q-cycles describe subcomplexes possessing a (possibly empty) interior of dimension $q+1$, while the q-boundaries describe such subcomplexes that are contractible. Therefore, it is reasonable to define the (j^{th}) *integral homology* of K as the quotient of modules:

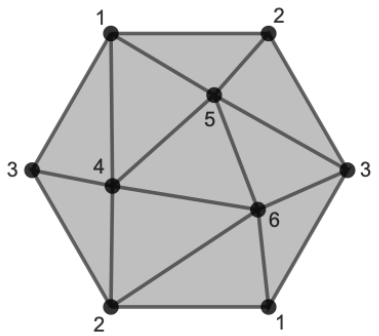
$$H_j(C_\bullet) = Z_j/B_j = \{\xi + B_j : \forall \xi \in Z_j\}$$

Recall: $\xi + B_j = \zeta + B_j$ iff $\xi - \zeta \in B_j$. As finitely-generated $\mathbb{Z}$-modules also, their structure is well-understood by their fundamental theorem (see [37]): for each j there exist natural numbers $\beta_j, \theta_{j2}, \theta_{j3}, \ldots$ a finite number of which are non-zero, such that

$$H_j(C_\bullet) \cong \mathbb{Z}^{\beta_j} \oplus (\mathbb{Z}/2\mathbb{Z})^{\theta_{j2}} \oplus (\mathbb{Z}/3\mathbb{Z})^{\theta_{j3}} \oplus \ldots \tag{2.3}$$

Most applications focus on the *torsion-free* component of homology $\mathbb{Z}^{\beta_j}$, whose rank β_j is the j^{th} *Betti number*. However, torsion is also prominently exhibited in several spaces of interest.

Example 2.4. Revisit $\mathbb{S}^1$ in Example 2.1 and the nerve K of its cover in (b). Here, $C_0 \simeq \mathbb{Z}^3$, $C_1 \simeq \mathbb{Z}^3$, and $C_j = \langle 0 \rangle$ for $j \geqslant 2$; the differential map has its components $\partial_j : C_j \to C_{j-1}$ all equal to zero maps for $j \neq 1$ and $\partial_1 : C_1 \to C_0$ is specified by $(c_{12}, c_{13}, c_{23}) \in \mathbb{Z}^3 \mapsto (-c_{12} - c_{13}, c_{12} - c_{23}, c_{13} + c_{23}) \in \mathbb{Z}^3$. The nullspace of ∂_0 is $Z_0 = C_0$, the nullspace of ∂_1 is $Z_1 \simeq \langle (1, -1, 1) \rangle$, and for all the other maps the nullspace is zero; hence $H_j(\mathbb{S}^1) = 0$ for $j \geqslant 2$. The range B_0 of ∂_1 is generated by the elements specific to $(-1, 1, 0), (-1, 0, 1), (0, -1, 1)$ in $\mathbb{Z}^3$, where the latter is a difference of the previous two; replacing that with $(1, 0, 0)$ gives a generating set of $Z_0 \simeq \mathbb{Z}^3$. Hence, $H_0(\mathbb{S}^1) = Z_0/B_0 = \langle (1, 0, 0) + B_0 \rangle \simeq \mathbb{Z}$ and $H_1(\mathbb{S}^1) = Z_1/B_1 = Z_1/\langle 0 \rangle \simeq \mathbb{Z}$. The 0^{th} homology shows that $\mathbb{S}^1$ is a connected set, while the generator of $H_1(\mathbb{S}^1)$ matches the "hollow loop" shape of the 1-sphere.



$$
\begin{aligned}
C_0 = \quad & \Big\langle [1], [2], [3], [4], [5], [6] \Big\rangle = \Big\{ \sum_{v=1}^{6} c_v \cdot [v] : c_v \in \mathbb{Z} \Big\} \\
& \partial_0([v]) = 0, \;\; \forall v = 1, \ldots, 6 \\
C_1 = \quad & \Big\langle [v, w] : 1 \leqslant i < j \leqslant 6 \Big\rangle = \Big\{ \sum_{v,w} c_{vw} \cdot [v, w] \Big\} \\
& \partial_1([v, w]) = [w] - [v], \;\; \forall v < w \\
C_2 = \quad & \Big\langle [1,2,5], [1,2,6], [1,3,4], [1,3,6], [1,4,5], \\
& \qquad [2,3,5], [2,3,4], [2,4,6], [3,5,6], [4,5,6] \Big\rangle \\
& \partial_2([u, v, w]) = [v, w] - [u, w] + [u, v]
\end{aligned}
$$

Example 2.5. Let $\mathbb{RP}^2 = \mathbf{D}^2/\sim$, where $\mathbf{D}^2 \subseteq \mathbb{C}$ is the closed unit disc with the equivalence relation under which $z \sim w$ if and only if $z = w$ or $|z - w| = 2$. It is suitably represented by the simplicial complex presented above. Note that every 0-boundary can be written as a sum in terms of $\partial_1([1,2]), \partial_1([2,3]), \ldots, \partial_1([5,6])$; these are independent generators of $B_0 \simeq \mathbb{Z}^5$, and since $Z_0 = C_0$ then $H_0(\mathbb{RP}^2) = Z_0/B_0$ is given by $\langle [1] + B_0 \rangle \simeq \mathbb{Z}$. The $j = 1$ homology is more peculiar here; observe that

$$\partial_2([1,2,4])\ ,\ \partial_2([1,2,6])\ ,\ \partial_2([1,3,5])\ ,\ \ldots\ ,\ \partial_2([2,5,6])\ ,\ \partial_2([3,4,6])\ ,\ \partial_2([4,5,6])$$

are all independent generators of $B_1 \simeq \mathbb{Z}^{10}$, the signed sum of which (with $+1$ for the last element or if $[u, v, w]$ has $[2]$ as a vertex, and -1 otherwise) has boundary $2 \cdot ([1,2] - [1,3] + [2,3])$. However, the cycle $\sigma = [1,2] - [1,3] + [2,3]$ itself is not the boundary of any 2-chain, so $(\sigma + B_1) \in H_1(\mathbb{RP}^2)$ is non-trivial with $\big\langle \sigma + B_1 \big\rangle \simeq \mathbb{Z}/2\mathbb{Z}$. Apparently $Z_1 = \langle \sigma \rangle + B_1$, so this completely describes the $j = 1$ homology of $\mathbb{RP}^2$.

Classically, one may calculate the j^{th} homology of a complex by writing the differential $\partial_\bullet$ on $C_\bullet$ (or at least its component $C_{j-1} \leftarrow C_j \leftarrow C_{j+1}$) in matrix form, and computing its *Smith normal form.* The resulting diagonal matrix provides a one-to-one correspondence between generators for the image of $\partial_\bullet$ (the boundary chains) and generators for the coimage of $\partial_\bullet$ (ie, those not in the kernel of $\partial_\bullet$), with the torsion and free group coefficients expressed on the diagonal.

Among the greatest successes of integral homology is the ability to determine the homology groups $H_j(C_\bullet; R)$ with coefficients in any other abelian group R. Of particular importance is when R is some field $\mathbb{F}$, so the underlying complex is endowed with vector space structure. The *universal coefficient theorem* (for homology, see [58]) allows us to compute these:

Theorem 2.6. *For any $\mathbb{Z}$-module R, there is a functorial short exact sequence*

$$0 \to H_j(C_\bullet) \otimes_{\mathbb{Z}} R \to H_j(C_\bullet; R) \to \mathrm{Tor}_1(H_{j-1}(C_\bullet), R) \to 0$$

that splits for every chain complex $C_\bullet$ (albeit not functorially).

In particular, when $R = \mathbb{F}$ is a field, the datum $H_j(C_\bullet; \mathbb{F})$ is an $\mathbb{F}$-vector space whose basis consists of generators for $H_j(C_\bullet; \mathbb{Z})$ in the torsion-free component, as well as in any component $(\mathbb{Z}/k\mathbb{Z})^{\theta_{jk}}$ where k is divisible by the characteristic of $\mathbb{F}$. In other uses, the universal coefficient theorem allows one to find the cohomology of a topological space in any coefficients from a ring R.

Example 2.7. For $\mathbb{S}^1$ in example 2.4 and any field $\mathbb{F}$, we have $H_j(\mathbb{S}^1; \mathbb{F}) \simeq \mathbb{F}$ if $j = 0, 1$ and $H_j(\mathbb{S}^1) \simeq 0$ otherwise. For $\mathbb{RP}^2$ in example 2.5, we have $H_0(\mathbb{RP}^2; \mathbb{F}) \simeq \mathbb{F}$, for any field $\mathbb{F}$; however $H_1(\mathbb{RP}^2; \mathbb{Z}/2\mathbb{Z})$ is a vector space of dimension one, while $H_1(\mathbb{RP}^2; \mathbb{Z}/3\mathbb{Z})$ and $H_1(\mathbb{RP}^2; \mathbb{Q})$ are both zero vector spaces.

As we shall see, field-valued homology groups are generally preferred in practical calculations due to their simplicity, and hence are often calculated directly using methods of linear algebra. An often-taught method for computing homology in field coefficients directly is by writing the differential matrix (or its appropriate component) in Jordan canonical form; the 2×2 Jordan blocks describe matchings between generators of the image and coimage of the differential operator, while the 1×1 Jordan blocks point to generators of the $\mathbb{F}$-valued homology group (larger Jordan blocks do not exist here since $\partial_\bullet \circ \partial_\bullet = 0$). Instead, however, we will be describing different algorithms throughout that are more grounded in matrix reduction.

It is worth mentioning that alternative homology theories (based on different choices of combinatorial models of topological spaces) yield equivalent results. Simplicial homology can be known as the first "successful" homology theory, in the sense that it allowed one to fully, yet rigorously work out homological data for a wide range of known compact spaces. Other formulations can be found to be more general and/or more notationally brief. One is using CW-complexes and cellular homology; for instance, the 1-sphere from example 2.4 can be merely encoded with a 0-cell and a 1-cell, while the space $\mathbb{RP}^2$ only requires a specification of six cells (as opposed to a total of 34 simplices as in example 2.5). Alternative homology theories can be proposed for topological spaces, but as long as they adhere to a set of five Eilenberg-Steenrod axioms (see [24]) then they yield consistent results as far as the theories are applicable. To do this however, we would need to further pursue the notion of *relative homology*, where a pair (X, A) of topological spaces X and $A \subset X$ is considered; for the interested reader, a discussion on this can be found in classical texts cited near the beginning of this section.

2.1.4 Morphisms of Complexes and their Homologies

Sometimes we are less interested in topological spaces themselves than characteristics assigned to them by external functions. With the only relevant constraint being continuity (though not always even that), we'd like to have a function $\Phi : X \to \hat{X}$ between topological spaces be representable as a function between their respective simplicial complexes K and $\hat{K}$; we'll suppose that $X = |K|$ and $\hat{X} = |\hat{K}|$. This function $\phi : K \to \hat{K}$ taking simplices to simplices should do so in a meaningful way. We may attempt this by defining ϕ so that it preserves inclusions of simplices, that is $\phi(f) \subseteq \phi(g)$ in $\hat{K}$ whenever $f \subseteq g$ in K; it is also natural to assume that the function ϕ restricts to a map of vertices $V(K) \to V(\hat{K})$, thus mimicking the behavior of continuous functions. In fact, a map of vertices $V(K) \to V(\hat{K})$ induces a map $\phi : K \to \hat{K}$ by setting $\phi(f)$ to the set-theoretic image of $f \in K$ under the vertex map, somewhat following the idea of embedding $K \to \Delta^{|V(K)|}$ in subsection 2.1.1. This resulting ϕ is known as a *simplicial map*.

It begs the question whether a simplicial map (with its property that faces of K are mapped to faces of $\hat{K}$) is too restrictive, referring to how well it concurs with the actual map Φ. A sensible demand is that $\phi(v) = \Phi(v)$ for all vertices v in K; more broadly, for any $x \in |K|$ and any face $f \in K$ containing x that $\Phi(x)$ is contained in face $\phi(f)$ holds. If such a ϕ can be obtained, then it is a suitable *simplicial approximation* of function Φ, though it is clear that these are very restrictive conditions imposed on K and $\hat{K}$ that already may not be satisfied by the triangulations. Nevertheless, L. Brouwer (who was responsible for introducing the formal notion of a simplicial map [17]) was able to prove the following result.

Theorem 2.8. *Given simplicial complexes $K, \hat{K}$ and some function $\Phi : |K| \to |\hat{K}|$ that is continuous, for some integer $n \geqslant 0$ there exists a simplicial approximation $\phi : \mathrm{Sd}_n K \to \hat{K}$ of Φ, where $\mathrm{Sd}_n K$ is the simplicial complex obtained after n counts of barycentric subdivision performed iteratively on K.*

Evidently, simplicial maps take j-simplices to k-simplices, where $j \geqslant k$ (equality holds when a simplex's vertices are mapped injectively). For each $j \geqslant 0$, define a function ϕ_j on the j-simplices of K mapping to $\hat{K}$ by:

$$\phi_j([v_1, \ldots, v_{j+1}]) = \begin{cases} 0 & \text{, if } \phi(v_s) = \phi(v_t) \text{ for some } s \neq t \\ [\phi(v_1), \ldots, \phi(v_{j+1})] & \text{, otherwise} \end{cases}$$

This definition implies we are only interested in simplices whose dimension is preserved under ϕ; indeed, mapping a simplex to a copy of its proper face is a deformation retract, which induces a homotopy equivalence. Now by linearity, this extends to a homomorphism $\phi_j = C_j(\phi) : C_j(K) \to C_j(\hat{K})$ of j-chains, for all simplicial dimensions j; together, these can be viewed as components of the graded homomorphism $\phi_\bullet = C_\bullet(\phi) : C_\bullet(K) \to C_\bullet(\hat{K})$ that has degree 0, referred to as the *chain map (induced by ϕ)*. The complete picture is displayed on the following diagram.

$$\begin{array}{ccccccccc}
\cdots & \longleftarrow & \hat{C}_{p-1} & \longleftarrow & \hat{C}_p & \longleftarrow & \hat{C}_{p+1} & \longleftarrow & \cdots \\
 & & {\scriptstyle \phi_{p-1}}\Big\uparrow & & {\scriptstyle \phi_p}\Big\uparrow & & {\scriptstyle \phi_{p+1}}\Big\uparrow & & \\
\cdots & \longleftarrow & C_{p-1} & \longleftarrow & C_p & \longleftarrow & C_{p+1} & \longleftarrow & \cdots .
\end{array}$$

Meanwhile, each chain complex is equipped with its own differential structure: $\partial_\bullet$ for $C_\bullet(K)$ and $\hat{\partial}_\bullet$ for $C_\bullet(\hat{K})$, respectively. The chain map displays an important property in that it *commutes* with the differential structures, in the following sense: for every $j \geqslant 1$, $\quad \hat{\partial}_j \circ \phi_j \equiv \phi_{j-1} \circ \partial_j$. Visually, any two homomorphisms obtained by composition of mappings between two fixed modules in the above diagram are equal. An immediate consequence is the following:

- for any j-cycle $c \in Z_j(K)$, $\hat{\partial}_j[\phi_j(c)] = 0$; that is, $\phi_j(Z_j) \subseteq \hat{Z}_j$
- if $b = \partial_j(c)$, then $\phi_{j-1}(b) = \hat{\partial}_j[\phi_j(c)]$; that is, $\phi_j(B_j) \subseteq \hat{B}_j$.

Therefore, the induced quotient map

$$\xi + B_j \in Z_j/B_j \mapsto \phi_j(\xi) + \hat{B}_j \in \hat{Z}_j/\hat{B}_j$$

is well-defined, and represents the j^{th} component of $H_j(\phi_\bullet)$, the induced map of homologies. Note that while the classification of individual spaces by their homology is well-understood from Theorem 2.3, the classification of Z-module homomorphisms is remarkably more difficult. Finally, for any field $\mathbb{F}$, the map of homologies $H_j(\phi_\bullet)$ defines an $\mathbb{F}$-linear map $H_j(C_\bullet(K); \mathbb{F}) \to H_j(C_\bullet(\hat{K}); \mathbb{F})$ by taking elements in a basis of $H_j(C_\bullet(K); \mathbb{F})$ to their images in $H_j(C_\bullet(\hat{K}); \mathbb{F})$.

As we revisit these concepts later, it is helpful to point out their *functorial* character. We began with simplicial complexes K and $\hat{K}$, whose appropriate choice of function is a simplicial map $\phi : K \to \hat{K}$ between them. We proceeded to derive modules $C_\bullet(K)$ and $C_\bullet(\hat{K})$ with a corresponding homomorphism $C_\bullet(\phi_\bullet)$ between them, which appropriately "preserved" their (differential chain) structure. This finally led to producing n^{th} homology modules $H_n(C_\bullet(K))$ and $H_n(C_\bullet(\hat{K}))$, with an induced map $H_n(C_\bullet(\phi))$ acting between them. Going further, we can extract the n^{th} homology in chosen field coefficients:

$$H_n(C_\bullet(\phi); \mathbb{F}) : H_n(C_\bullet(K); \mathbb{F}) \to H_n(C_\bullet(\hat{K}); \mathbb{F})$$

which are vector spaces, with a linear transformations between them. So real details aside, the process can be summarized as transitioning $K \xrightarrow{\phi} \hat{K}$ along each step of the computation. This kind of formalism will be more elaborated on in chapter 3, but is useful to have when doing persistent homology.

Observe that while each "transition" produces compatible structures, some of their familiar properties may not carry over well! For example, the inclusion $\mathrm{Bd}\,\Delta^2 \to \Delta^2$ (where $\mathrm{Bd}\,\Delta^2$ is the topological boundary of Δ^2) produces an injective simplicial map, whereas the induced homology map $H_1(C_\bullet(\mathrm{Bd}\,\Delta^2)) \to H_1(C_\bullet(\Delta^2))$ is surjective.

2.2 Persistence Vector Spaces

As mentioned earlier, persistent homology looks at all the previous constructions on multiple levels, which somehow capture the relative and global distributions of the given geometric data. We embed this new detail into our existing framework.

Let $(T, \leqslant)$ be a partially-ordered set, here assumed to be either the integers $T = \mathbb{Z}$ or the reals $T = \mathbb{R}$.For a fixed field $\mathbb{F}$, a *persistence vector space* ${}_\bullet V$ is specified by a collection of vector spaces $\{{}_tV : t \in T\}$ and also a linear map ${}_sV \to {}_tV$ for any $s \leqslant t$ (which is the identity map if $s = t$); these *structure maps* are required to preserve composition, so if $r \leqslant s \leqslant t$ then $({}_sV \to {}_tV) \circ ({}_rV \to {}_sV) = ({}_rV \to {}_tV)$.

$$\begin{tikzcd} {}_rV \arrow[r] \arrow[rr, bend left] & {}_sV \arrow[r] & {}_tV \end{tikzcd}$$

Figure 2.2: The composition law can be observed on this commutative diagram.

In the case $T = \mathbb{Z}$, we can distinguish *tempered* persistence (vector) spaces [43] for which there exist only finitely-many $s \in \mathbb{Z}$ where the maps ${}_sV \to {}_{s+1}V$ are not vector space isomorphisms. These are well-behaved objects and are commonly encountered in applications. A similar notion exists for the case $T = \mathbb{R}$, whose associated objects are defined as *tame*; see page 25 of [47] for details.

Remark 2.9. It is more common to refer to persistence vector spaces as *persistence modules*. Indeed, a persistence vector space ${}_\bullet V$ can be identified with a left R-module M, whose abelian group structure is given by the direct sum of ${}_t V$ over all $t \in T$. If $T = \mathbb{Z}$, the ring R associated to the module is the polynomial ring $\mathbb{F}[t]$. Left multiplication of t^n by any element of the module corresponds to the pointwise action of the structure map ${}_\bullet V \mapsto {}_{\bullet+n} V$ and extends linearly to the whole polynomial ring. Similarly, when $T = \mathbb{R}$ we associate the ring $R = \mathbb{F}\langle t^{\mathbb{R}} \rangle$, which is $\mathbb{F}$-spanned by all elements t^α with $\alpha \in \mathbb{R}$. Left multiplication is defined similarly, induced by the action of the structure map ${}_\bullet V \mapsto {}_{\bullet+\alpha} V$. For other choices of partially-ordered sets T, corresponding persistence vector spaces can also be identified with appropriate left R-modules – this is evident by the Freyd-Mitchell embedding theorem (see Theorem 3.16). The ring R is to capture the graph structure of the set T, or more precisely, the "transition" between its elements via the partial order relations.

Suppose that ${}_\bullet W$ and ${}_\bullet V$ are persistence vector spaces such that for every $t \in T$ there is an isomorphism ${}_t\varphi : {}_t W \xrightarrow{\sim} {}_t V$. Then, every structure map ${}_s V \to {}_t V$ can be "pulled back" to a map ${}_t\varphi^{-1} \circ ({}_s V \to {}_t V) \circ {}_s\varphi$ from ${}_s W$ to ${}_t W$. If one can find these isomorphisms ${}_\bullet\varphi$ such that this *pullback* of ${}_s V \to {}_t V$ is identical to the structure map ${}_s W \to {}_t W$ (for any $s \leqslant t$), then one has an *isomorphism* of persistent vector spaces and is denoted ${}_\bullet W \simeq {}_\bullet V$. It is clear that any algebraic properties that apply to ${}_\bullet V$ will also apply to persistence vector spaces isomorphic to ${}_\bullet W$, via the pullback (and also its inverse, the *pushforward*).

Alternatively, we may consider a situation where every isomorphism to be replaced by inclusion so that ${}_tW \subseteq {}_tV$ for all $t \in T$. We then say that ${}_\bullet W$ is a *(persistence vector) subspace* of ${}_\bullet V$ if a similar condition on the compatibility of the structure maps holds: specifically that for all $s \leqslant t$, the map ${}_sV \to {}_tV$ is identical to ${}_sW \to {}_tW$ when its domain and range is restricted to ${}_sW$ and ${}_tW$, respectively. Note that subspaces of tempered persistence spaces are also tempered. The *trivial (persistence) vector space* ${}_\bullet\mathbf{0}$, given by ${}_t\mathbf{0} = \mathbf{0}$ with zero maps between them, is clearly a subspace of every ${}_\bullet V$; of course any persistent vector space ${}_\bullet V$ is a subspace of itself.

For any two persistent vector spaces (${}_\bullet U$ and ${}_\bullet W$) parametrized by the same set T, their *direct sum* ${}_\bullet U \oplus {}_\bullet W$ is given by the "pointwise" direct sum: take the collection $\{{}_tU \oplus {}_tW) : t \in T\}$ whose linear map associated to indices $s \leqslant t$ is the direct sum of $({}_sU \to {}_tU)$ and $({}_sW \to {}_tW)$, having block matrix representation

$$({}_sU \to {}_tU) \oplus ({}_sW \to {}_tW) = \begin{pmatrix} {}_sU \to {}_tU & 0 \\ 0 & {}_sW \to {}_tW \end{pmatrix}$$

Any ${}_\bullet V$ that can be written (up to isomorphism) as the direct sum of non-trivial persistence vector spaces ${}_\bullet U$ and ${}_\bullet W$ is called *decomposable*; otherwise, ${}_\bullet V$ is *indecomposable*. When attempting to find a direct sum decomposition ${}_\bullet V$ in terms of its indecomposable summands, one must remember that such a decomposition cannot be unique (even in linear algebra, any two-dimensional vector space can be spanned by any pair of its independent vectors). However, direct sum decompositions are "essentially" unique as a consequence of the famous Krull-Schmidt theorem (see Theorem 7.5 in [37]), a corollary of which for persistence vector spaces is stated below. We shall later restate this as Theorem 3.8 to underline its categorical generality.

Theorem 2.10 (see Theorem 4.2 in [36]). *Take two finite series of indecomposable tempered persistence vector spaces* ${}_\bullet V'_1, \dots, {}_\bullet V'_m$ *and* ${}_\bullet V''_1, \dots, {}_\bullet V''_n$ *such that*

$$ {}_\bullet V'_1 \oplus \cdots \oplus {}_\bullet V'_m \simeq {}_\bullet V''_1 \oplus \cdots \oplus {}_\bullet V''_n $$

Then $m = n$*, and there exists a permutation* π *of all indices such that* ${}_\bullet V'_i \simeq {}_\bullet V''_{\pi(i)}$.

It follows that a persistent vector space can be algebraically understood entirely by its indecomposable summands. In fact, the Krull-Schmidt decomposition also describes indecomposable persistence vector spaces – their set of endomorphisms has the structure of a local ring. While this can be used to fully classify them (up to isomorphism), another approach to do this will be overviewed in subsection 2.2.3.

2.2.1 Filtrations of a Vector Space

A particular example of a persistence vector space is a *filtration* of a fixed vector space V. For each $t \in T$ here, the pointwise elements ${}_tV$ are subspaces of V such that ${}_sV \subseteq {}_tV$ whenever $s \leqslant t$; the map ${}_sV \to {}_tV$ is then taken to be the inclusion map.

Provided that V is of finite dimension, it follows directly that any filtration ${}_tV$ of V is a tempered persistence vector space. Generally, we also speak of *exhaustive* filtrations (if there is an $r \in T$ such that ${}_rV = {}_sV$ for all $r \leqslant s$) and *separated* filtrations (if there is an $x \in T$ such that ${}_sV = \mathbf{0}$ for all $s \leqslant x$) of V.

Subspaces of a filtration (as a persistence vector space) are usually called *subfiltrations*. Now, for any two subfiltrations ${}_\bullet U$ and ${}_\bullet W$ of V we can define their *intersection* and *subspace sum* as filtrations of V in a pointwise manner:

$$ {}_t({}_\bullet U \cap {}_\bullet W) = {}_tU \cap {}_tW \qquad \text{and} \qquad {}_t({}_\bullet U + {}_\bullet W) = {}_tU + {}_tW $$

Elsewhere, these special filtrations are referred to by their categorical terminology as the *pullback* and *pushforward*, respectively.

Remark 2.11. If ${}_\bullet U \cap {}_\bullet W$ is trivial then ${}_\bullet U + {}_\bullet W \simeq {}_\bullet U \oplus {}_\bullet W$ (and also conversely).

Filtrations provide a natural setting for us to work beyond ordinary chain complexes, in situations where they carry extra information about their "levels" (indexed by the set T). On the other hand, the generality of persistence vector spaces allows us to speak of quotients, which provide a powerful and flexible framework for representing the homology groups as we will show in chapter 4.

2.2.2 Morphisms of Persistence Vector Spaces

A *morphism* ${}_\bullet\phi$ of *persistence vector spaces* ${}_\bullet U$ and ${}_\bullet V$ is given pointwise by linear maps ${}_t\phi : {}_tU \to {}_tV$ (for all $t \in T$) that preserve composition with the structure maps of ${}_\bullet U$ and ${}_\bullet V$. That is (with focus on $T = \mathbb{Z}$), the following diagram commutes:

$$\begin{tikzcd}
\cdots \arrow[r] & {}_{t-1}V \arrow[r] & {}_tV \arrow[r] & {}_{t+1}V \arrow[r] & \cdots \\
\cdots \arrow[r] & {}_{t-1}U \arrow[r] \arrow[u, "{}_{t-1}\phi"] & {}_tU \arrow[r] \arrow[u, "{}_t\phi"] & {}_{t+1}U \arrow[r] \arrow[u, "{}_{t+1}\phi"] & \cdots .
\end{tikzcd}$$

We've already looked at some examples, notably subspace inclusions in the case where ${}_tU \subseteq {}_tV$. As another instance, the canonical projection maps from the direct sums ${}_tU \oplus {}_tW$ to components ${}_tU$ and ${}_tW$ (for all $t \in T$) comprise morphisms

$${}_\bullet U \oplus {}_\bullet W \to {}_\bullet U \qquad \text{and} \qquad {}_\bullet U \oplus {}_\bullet W \to {}_\bullet W$$

More generally, a *monomorphism* of persistence vector spaces ${}_\bullet V$ and ${}_\bullet W$ is a morphism ${}_\bullet\phi : {}_\bullet V \to {}_\bullet W$ such that every ${}_t\phi$ is an injective map, while an *epimorphism* of ${}_\bullet V$ and ${}_\bullet W$ is a morphism ${}_\bullet\phi$ such that ${}_t\phi$ is surjective for all $t \in T$. More on this can be found in chapter 3; we leave off for now that these behave like one-to-one and onto linear maps between vector spaces.

At each level $t \in T$, the map ${}_t\phi : {}_tW \to {}_tV$ has an associated kernel subspace $\ker({}_t\phi) \subseteq {}_tW$ and image subspace $\operatorname{im}({}_t\phi) = {}_t\phi\,({}_tW) \subseteq {}_tV$. These form the sequence

$$\cdots \longrightarrow \ker({}_{t-1}\phi) \longrightarrow \ker({}_t\phi) \longrightarrow \ker({}_{t+1}\phi) \longrightarrow \cdots$$

of kernels (to be denoted $\ker{}_\bullet\phi$), and the sequence

$$\cdots \longrightarrow \operatorname{im}({}_{t-1}\phi) \longrightarrow \operatorname{im}({}_t\phi) \longrightarrow \operatorname{im}({}_{t+1}\phi) \longrightarrow \cdots$$

of images (to be denoted $\operatorname{im}{}_\bullet\phi$). Furthermore, the sequence $\operatorname{coker}{}_\bullet\phi$ of cokernels

$$\cdots \longrightarrow \operatorname{coker}({}_{t-1}\phi) \longrightarrow \operatorname{coker}({}_t\phi) \longrightarrow \operatorname{coker}({}_{t+1}\phi) \longrightarrow \cdots$$

with $\operatorname{coker}({}_s\phi) = {}_sV/\operatorname{im}({}_s\phi)$, and the sequence $\operatorname{coim}{}_\bullet\phi$ of coimages

$$\cdots \longrightarrow \operatorname{coim}({}_{t-1}\phi) \longrightarrow \operatorname{coim}({}_t\phi) \longrightarrow \operatorname{coim}({}_{t+1}\phi) \longrightarrow \cdots$$

with $\operatorname{coim}({}_s\phi) = {}_sW/\ker({}_s\phi))$ are used; note that by the rank theorem for vector spaces, $\operatorname{coim}({}_s\phi) \simeq \operatorname{im}({}_s\phi)$. It is not explicitly obvious that there should be (horizontal) structure maps in each one of these four diagrams, but these are fully described by the following result.

Proposition 2.12. *For any morphism ${}_\bullet\phi : {}_\bullet W \to {}_\bullet V$ of persistence vector spaces and any $s \leqslant t$ in T, there are well-defined maps*

$$\ker({}_s\phi) \to \ker({}_t\phi) \qquad\qquad \operatorname{im}({}_s\phi) \to \operatorname{im}({}_t\phi)$$

given by restricting the respective structure maps ${}_sW \to {}_tW$ and ${}_sV \to {}_tV$ to the subspaces $\ker({}_s\phi)$ and $\operatorname{im}({}_s\phi)$. Furthermore, these commute with the inclusions of the kernel and image into ${}_\bullet W$ and ${}_\bullet V$, respectively.

Likewise, there are well-defined quotient maps

$$\mathrm{coim}({}_s\phi) \to \mathrm{coim}({}_t\phi) \qquad\qquad \mathrm{coker}({}_s\phi) \to \mathrm{coker}({}_t\phi)$$

that are induced by the respective structure maps ${}_sW \to {}_tW$ *and* ${}_sV \to {}_tV$*. Furthermore, these commute with the canonical projections of* ${}_\bullet W$ *and* ${}_\bullet V$ *onto the coimage and cokernel, respectively.*

For details, see [53]. We'll revisit this discussion when reestablishing the class of persistence vector spaces as a category in Chapter 3.

2.2.3 The A_n quiver and indecomposable representations

A important factor for the surge of cultural enthusiasm around persistent homology was the ability to predict meaningful constituents of complexes formed by data. This came with the technical observation that indecomposables of persistent vector spaces can be understood through the combinatorics of A_n-type quivers. In brief, these are finite connected graphs with n vertices and $n-1$ directed edges such that each vertex has at most 2 adjacent vertices. If no two edges share the same "head" or "tail" vertex, then we call this a *unidirectional* A_n *quiver*; by contrast, *zigzag* A_n *quivers* only have vertices that act as a "sink" or a "source". Many other quivers of outside interest are associated with Dynkin diagrams, some of which are shown in figure 2.3.

A *quiver representation* is an assignment where a choice of vector space V_i is made for every vertex i of the quiver and a choice of linear map $V_i \to V_j$ is made for every edge from vertex i to vertex j. It should now be apparent that representations of unidirected A_n quivers produce persistence vector spaces with $T = \{1, \dots, n\}$. This will generalize to all tempered persistence vector spaces ${}_\bullet V$ once a value $n > 0$ is

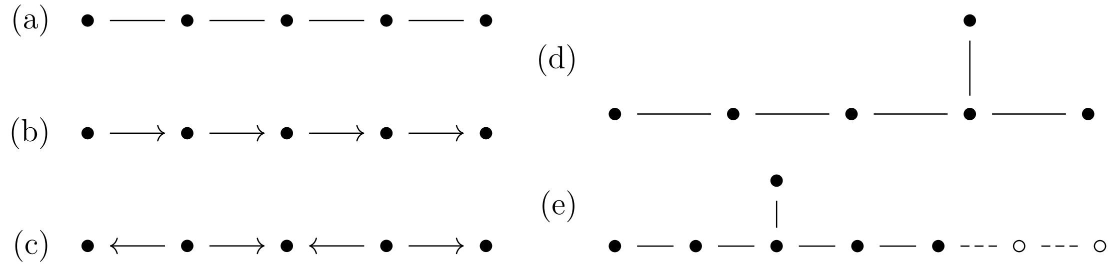


Figure 2.3: (a) shows an A_5 diagram. It is an *underlying graph* for its directed variants, (b) the unidirectional A_5 quiver, and (c) the zigzag A_5 quiver. Other significant quivers include those of D_n-type and E_n-type; for example, (d) shows a D_6 diagram and (e) the E_6, E_7, E_8 diagrams, whose number of vertices should match $n = 6, 7, 8$.

found large enough so that ${}_sV \simeq {}_1V$ ($\forall s \leqslant 1$) and ${}_tV \simeq {}_nV$ ($\forall t \geqslant n$). Furthermore, it has already been established by theorem 2.10 that these can be uniquely written (up to isomorphism) to a finite direct sum of some indecomposables.

We speak of *indecomposable representations* of A_n quivers while referring to indecomposable persistence vector spaces with $T = \{1, \ldots, n\}$. These are simple objects to understand and interpret in applications. Typically, the following result is cited when classifying quivers of "finite representation type".

Theorem 2.13 (Gabriel)**.** *Let Q be a connected quiver.*

1. *Q has finitely-many (isomorphism classes of) indecomposable representations if and only if Q is of type A_n, D_n, E_6, E_7, or E_8.*

2. *If Q is of type A/D/E, then there exists a bijection from the (finite) set of isomorphism classes of indecomposable representations of Q to the set of positive roots of the Tits quadratic form associated with Q.*

For proof and discussion, see pages 203-222 of [53].

Corollary 2.14. *Let ${}_\bullet V$ be a tempered persistence vector space.*

1. *${}_\bullet V$ is indecomposable if and only if there exist $i < j$ in $\mathbb{Z}$ such that*

$$ {}_tV \simeq \begin{cases} \mathbb{F} & \text{, if } i \leqslant t < j \\ \mathbf{0} & \text{, otherwise} \end{cases} \qquad \text{, where } {}_iV \to {}_{j-1}V \text{ is the identity.} $$

2. *Otherwise ${}_\bullet V$ is isomorphic to a finite direct sum of indecomposables, some of which are possibly non-distinct up to isomorphism.*

The indecomposable representations in corollary 2.14 are identified by the *interval* $[i, j)$, and the set of intervals (possibly with repeating elements) that determines the direct sum decomposition of ${}_\bullet V$ is called a *barcode*. Since direct sum decompositions of persistence vector spaces are essentially unique by theorem 2.10, persistent vector spaces have a unique barcode that stays invariant under isomorphism. This makes them a central point for persistent homology.

Another consequence of Theorem 2.13 is for representations of zigzag quivers. Also known as *zig-zag modules*, these sometimes arise when a dataset cannot be cleanly modeled by a filtration. Applications of this arise in probability density estimation, data subsampling, and during qualitative selection of persistent features; a good overview of these applications as well as the theory itself can be found in [11]. More generally, Gabriel's theorem is responsible for identifying the stated quivers as "tame"; the representation theory for these quivers is well-understood, as opposed to "wild" quivers (for which persistent homology is still enigmatic).

2.2.4 The Isometry Theorem, and the Bottleneck Metric

The previous discussion suggests that a persistence vector space ${}_\bullet V$ can be entirely described by a unique barcode $\mathcal{B}({}_\bullet V)$, up to isomorphism (over a fixed field $\mathbb{F}$). Each interval $[i, j)$ in the barcode points to a particular indecomposable summand of ${}_\bullet V$ having the form in 2.14(1) – the only contribution of this summand to the direct sum is a one-dimensional vector space for values of $t \in T$ in the interval. The generator of this one-dimensional vector space is a "feature" that seems to "persist" across over the interval $[i, j)$. It is why some might refer to these generators collectively as *persistent features*.

The interpretation of persistent features in the context of topological data analysis are best left to the next section. Also, the question of "how well" a persistence vector space is described by its barcode will be deferred until section 2.3. For now, we turn to our ability to capture "similarity" of different persistence vector spaces, and how this reflects on their barcodes.

An intrinsic way to perturb a persistent vector space (using $T = \mathbb{Z}$ or $\mathbb{R}$) is by using *ε-shifts*. Given any such persistence vector space ${}_\bullet V$, a (left) ε-shift produces the persistent vector space ${}_{\bullet+\varepsilon}V$, whose data at any $t \in T$ is ${}_{t+\varepsilon}V$ and the structure map corresponding to any $s \leqslant t$ is ${}_{s+\varepsilon}V \to {}_{t+\varepsilon}V$. Observe that there is a natural morphism ${}_\bullet V \to {}_{\bullet+\varepsilon}V$ (whose component at $t \in T$ is just the structure map ${}_tV \to {}_{t+\varepsilon}V$), and that its composition with ${}_{\bullet+\varepsilon}V \to {}_{\bullet+\varepsilon+\delta}V$ produces ${}_\bullet V \to {}_{\bullet+\varepsilon+\delta}V$; these are called *transition morphisms* of ${}_\bullet V$.

The idea is to extend this notion on a pair of persistence vector spaces ${}_\bullet U$ and ${}_\bullet W$, which may or may not be possible. An $\varepsilon-$*interleaving* between them is a choice of two morphisms ${}_\bullet f : {}_\bullet U \to {}_{\bullet+\varepsilon}W$ and ${}_\bullet g : {}_\bullet W \to {}_{\bullet+\varepsilon}U$ with the following property: the compositions $({}_{\bullet+\varepsilon}g) \circ ({}_\bullet f)$ and $({}_{\bullet+\varepsilon}f) \circ ({}_\bullet g)$ produce the transition morphisms ${}_\bullet U \to {}_{\bullet+2\varepsilon}U$ and ${}_\bullet W \to {}_{\bullet+2\varepsilon}W$, respectively (where the definition of the morphisms

${}_{\bullet+\varepsilon}f : {}_{\bullet+\varepsilon}U \to {}_{\bullet+2\varepsilon}W$ and ${}_{\bullet+\varepsilon}g : {}_{\bullet+\varepsilon}W \to {}_{\bullet+2\varepsilon}U$ is naturally implied from ${}_{\bullet}f$ and ${}_{\bullet}g$). This can be viewed in figure 2.4.

$$\begin{tikzcd}
\cdots \arrow[r] & {}_{t}W \arrow[r] \arrow[dr] & {}_{t+\varepsilon}W \arrow[r] \arrow[dr] & {}_{t+2\varepsilon}W \arrow[r] \arrow[dr] & {}_{t+3\varepsilon}W \arrow[r] \arrow[dr] & {}_{t+4\varepsilon}W \arrow[r] \arrow[dr] & {}_{t+5\varepsilon}W \arrow[r] & \cdots \\
\cdots \arrow[r] & {}_{t}U \arrow[r] \arrow[ur] & {}_{t+\varepsilon}U \arrow[r] \arrow[ur] & {}_{t+2\varepsilon}U \arrow[r] \arrow[ur] & {}_{t+3\varepsilon}U \arrow[r] \arrow[ur] & {}_{t+4\varepsilon}U \arrow[r] \arrow[ur] & {}_{t+5\varepsilon}U \arrow[r] & \cdots
\end{tikzcd}$$

Figure 2.4: Here, some components of the persistence vector spaces are shown, with the initial value of parameter t chosen arbitrarily. The up-facing arrows represent components of ${}_{\bullet}f$, while the down-facing arrows represent components of ${}_{\bullet}g$; when the diagram is "shifted" to the left by ε, these arrows also describe ${}_{\bullet+\varepsilon}f$ and ${}_{\bullet+\varepsilon}g$, respectively. The definition of an ε-interleaving implies that this diagram is commutative – in particular, the transition maps ${}_{\bullet}U \to {}_{\bullet+2\varepsilon}U$ and ${}_{\bullet}W \to {}_{\bullet+2\varepsilon}W$ can instead be evaluated by "weaving through" ${}_{\bullet+\varepsilon}W$ and ${}_{\bullet+\varepsilon}U$, respectively.

Needless to say, calculating these morphisms or even showing that they exist (for some fixed $\varepsilon > 0$) stands as a significant challenge. The interleaving property does however obey some rules of regularity. For example, if an ε-interleaving in terms of ${}_{\bullet}f : {}_{\bullet}U \to {}_{\bullet+\varepsilon}W$ and ${}_{\bullet}g : {}_{\bullet}W \to {}_{\bullet+\varepsilon}U$ is given, then we can compose these with the transition morphisms ${}_{\bullet}U \to {}_{\bullet+\delta}U$ and ${}_{\bullet}W \to {}_{\bullet+\delta}W$ to obtain an $(\varepsilon+\delta)$-interleaving. Hence, this suggests the following:

Definition 2.15. The *interleaving distance* $d_I\left({}_{\bullet}U, {}_{\bullet}W\right)$ of two persistence vector spaces ${}_{\bullet}U$ and ${}_{\bullet}W$ is the smallest $\varepsilon > 0$ for which there exists an ε-interleaving between them. If such an $\varepsilon > 0$ doesn't exist then $d_I\left({}_{\bullet}U, {}_{\bullet}W\right) = \infty$.

It is clear that $d_I\left({}_{\bullet}U, {}_{\bullet}W\right) = 0$ if and only if ${}_{\bullet}f$ and ${}_{\bullet}g$ above are isomorphisms of the persistence vector spaces. Thus, a positive interleaving distance indicates some degree of dissimilarity between the chosen persistence vector spaces. In fact, the following result suggests that d_I is a suitable measure of "dissimilarity".

Lemma 2.16. *The function $d_I(\cdot, \cdot)$ is well-defined on equivalence classes of persistence vector spaces, and obeys the following rules:*

1. $d_I({}_\bullet U, {}_\bullet W) \geqslant 0$, *with equality iff* ${}_\bullet U \simeq {}_\bullet W$
2. $d_I({}_\bullet U, {}_\bullet W) = d_I({}_\bullet W, {}_\bullet U)$
3. $d_I({}_\bullet U, {}_\bullet W) \leqslant d_I({}_\bullet U, {}_\bullet V) + d_I({}_\bullet V, {}_\bullet W)$

In particular, d_I is an extended pseudometric on persistence vector spaces.

On the other hand there are barcodes, which are efficient representations of persistence vector spaces. It is good to be equipped with some metric to compare them, be it related or not to the interleaving distance. An instinctual way to compare two sets of intervals is by "matching" most-alike intervals from each barcode into pairs, while keeping some intervals unmatched from both sets – we need to admit this possibility at least because the barcodes may have a different number of intervals.

Let $\delta > 0$ and denote by $\mathcal{B}_U = \mathcal{B}({}_\bullet U)$ and $\mathcal{B}_W = \mathcal{B}({}_\bullet W)$ the respective barcodes of ${}_\bullet U$ and ${}_\bullet W$. A δ-*matching* is a partition of the sets $\mathcal{B}_U$ and $\mathcal{B}_W$ into some subsets $\mathcal{B}_U^{2\delta} \cup (\mathcal{B}_U - \mathcal{B}_U^{2\delta})$ and $\mathcal{B}_W^{2\delta} \cup (\mathcal{B}_W - \mathcal{B}_W^{2\delta})$, respectively, such that:

- for every $[i, j)$ in either $\mathcal{B}_U^{2\delta}$ and $\mathcal{B}_W^{2\delta}$, we have $i + 2\delta > j$;
- there exists a bijection between $(\mathcal{B}_U - \mathcal{B}_U^{2\delta})$ and $(\mathcal{B}_W - \mathcal{B}_W^{2\delta})$, sending interval $[i, j)$ to interval $[i', j')$ with $[i' + \delta, j' - \delta) \subseteq [i, j) \subseteq [i' - \delta, j' + \delta)$.

These two different treatments for interval types are shown in figure 2.5. The principle behind a δ-matching is that it should suffice to "discard" short intervals (from both barcodes) for the remainder to somehow be matched based on the proximity of

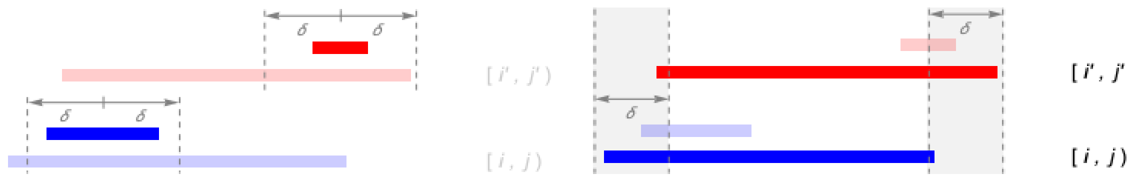


Figure 2.5: Following [4], here shown and detailed are sample barcodes $\mathcal{B}_U$ (blue) and $\mathcal{B}_W$ (red), each with two intervals. On the left are emphasized the "short" intervals of each $\mathcal{B}_{-}^{2\delta}$, having length less than 2δ. The right figure emphasizes the remaining "long" intervals, and depicts a matching of interval $[i, j) \in \mathcal{B}_U - \mathcal{B}_U^{2\delta}$ to an interval $[i', j') \in \mathcal{B}_W - \mathcal{B}_W^{2\delta}$ such that their endpoints are within an error range of δ from each other.

their endpoints. As with interleavings of persistence vector spaces, this is not guaranteed to exist for any $\delta > 0$, but still obeys certain regularity. For one, $\mathcal{B}_U$ and $\mathcal{B}_W$ are δ'-matched if they are δ-matched and $\delta \leqslant \delta'$. Hence:

Definition 2.17. The *bottleneck distance* $d_B(\mathcal{B}_U, \mathcal{B}_W)$ of two barcodes $\mathcal{B}_U$ and $\mathcal{B}_W$ is the smallest $\delta > 0$ for which there exists an δ-matching between them. If such a $\delta > 0$ does not exist then $d_B(\mathcal{B}_U, \mathcal{B}_W) = \infty$.

Similar to the interleaving distance, d_B is an extended metric on the collection of all barcodes (having countably-many intervals); in particular, two barcodes have a bottleneck distance of zero if and only if they are identical.

Remark 2.18. An alternate description of a δ-matching uses a proper bijection γ between certain multisets of $\mathbb{R}^2$ called *persistence diagrams*: each contains the set of points (i, j) corresponding to the intervals $[i, j)$ in one of either barcode (possibly with multiplicity to account for repetition), as well as countably-many points (k, k) on the diagonal corresponding to all "null" intervals $[k, k)$ for $k \in \mathbb{R}$. Then γ is a δ-matching if $\sup\|x - \gamma(x)\|_\infty < \delta$, where the supremum is taken over all points x from the domain of γ and $\|a, b\|_\infty := \max\{|a|, |b|\}$ is the ordinary L^∞ norm on $\mathbb{R}^2$. If

γ is chosen such that it is the identity on a cofinite subset of null intervals, then this notion of a δ-matching agrees with the one above. This was the original viewpoint of persistence diagram matchings, made precise in [15].

Computing the bottleneck distance involves a finite optimization problem. A brute force algorithm would run all permutations of intervals from one barcode to match them with intervals from another, all to select a permutation minimizing the bottleneck distance. A more efficient (polynomial-time) algorithm that works iteratively is outlined on page 241 of [18].

Of course, we are interested in finding the relationship between the two metrics. The derivational similarities between the two functions are hard to miss, and if one is a strong advocate for using category theory in this context then perhaps they would anticipate some bounding inequality between them. The next result surprisingly suggests something much stronger.

Theorem 2.19. *For any two tempered persistence vector spaces ${}_{\bullet}U$ and ${}_{\bullet}W$, having barcodes $\mathcal{B}_U$ and $\mathcal{B}_W$ we have $d_I\left({}_{\bullet}U, {}_{\bullet}W\right) = d_B\left(\mathcal{B}_U, \mathcal{B}_W\right)$.*

This is the so-called "isometry theorem", affirming that the barcode decomposition (or equivalently, the persistence diagrams) of persistence vector spaces is not only a convenient algebraic classification tool for these objects but also maintains enough of the objects' original data for their structural comparison. In particular, this is a useful paradigm in applied persistent homology, where the interleaving distance and its physical meaning for certain sequences of vector spaces can be computed directly with the barcode distance.

In the original paper, the weaker *algebraic stability theorem* was stated where the equality was replaced with inequality bounding the interleaving distance from below. This was a generalization of (and a nod to) the more widely-celebrated stability theorem, to be discussed in the next section. Bubenik and Scott later realized in [5] that this is in fact an isometry, the proof of which relies on using the interval decomposition of each persistence vector space to construct an interleaving as a direct sum of δ-shifts between interval-induced subspaces. The most recent rendition of the proof [4], draws heavily from the categorical structure of the class of persistence vector spaces and how it (nearly functorially) induces matchings between interval modules; these concepts will be explored more later.

It is also worth mentioning that the isometry theorem (and the corresponding stability theorem) can be stated for *generalized* persistence vector spaces – these are objects where the index set T is only assumed to be a partially-ordered set. Here, morphisms between two spaces are defined pointwise in a way that makes the resulting diagram of vector spaces commutative; one may then proceed to define a generalized notion of a shift or a transition morphism. For a background on these, the reader is referred to [9] and [42].

2.3 A Summary of Persistent Homology

Persistent homology extends the theory of classical homology – not only does it compute the homology of a graphical structure (with the goal of describing some data set), but it ultimately allows us to compare representatives of the homology (cycles) in terms of their "prominence". As such, many principles and objects from before remain important pieces in this method of analysis, only being extended to include the notion of "scale". Even then, the application of familiar elements and operations from linear algebra is prevalent when computing field homology.

Of note is the shift from working with vector spaces (or finitely-generated modules) to persistence vector spaces that we have introduced just previously. It is thus suitable to begin with the background of topological data analysis and how that motivates the usage of persistence vector spaces.

2.3.1 Filtrations of Simplicial Complexes

We first lay out the setting of persistent homology. Let K be a simplicial complex, T be a totally-ordered set (same as in section 2.2), and $f : K \to T$ be a function. If f is order-preserving (ie, if $\sigma \subseteq \tau$ are simplices in K then $f(\sigma) \leqslant f(\tau)$), then the resulting preimage ${}_tK := f^{-1}\Big((-\infty, t]\Big)$ is a simplicial complex itself for any $t \in T$. Clearly, there is an inclusion morphism ${}_rK \to {}_sK$ for any $r \leqslant s$ in T. Note also that since K is a finite complex there exist values $t^-, t^+ \in T$ such that ${}_rK = \varnothing$ for all $r \leqslant t^-$ and ${}_sK = K$ for all $s \geqslant t^+$. The resulting sequence is denoted ${}_\bullet K$ and is called a *filtered simplicial complex.*

One way that these complexes are constructed is on top of existing discrete data sets; note that $T = \mathbb{R}$ is predominantly used. Let $X = \{x_1, \ldots, x_n\}$ be a subset of $\mathbb{R}^d$ and let $K = K(X)$ be a *complete* simplicial complex with vertices in X (ie, it is an affine image of the standard simplex Δ^n with vertices identified to elements in X). The function $f : K \to \mathbb{R}$ can be defined in several different ways. In the *Vietoris-Rips filtration*, $f = f_{\text{Rips}}$ assigns 1-simplices $[x_i, x_j]$ to the Euclidean distance from x_i to x_j, while 0-simplices of K are assigned to 0; proceeding recursively on simplex dimension, f assigns an m-simplex to the maximum value of f over its proper faces, for every $m > 1$. Alternatively, the *Čech filtration* assigns to each simplex $\sigma \in K(X)$

the value $f_{\check{\text{C}}\text{ech}}(\sigma) = \max_{p\in|\sigma|} \xi_\sigma(p)$; here, $|\sigma|$ is the convex hull of the vertices in σ and $\xi_\sigma(p)$ is the minimum Euclidean distance of a point $p \in |\sigma|$ to any of the vertices of σ. Note that for any $t \in T$, $f^{-1}_{\check{\text{C}}\text{ech}}\Big((-\infty, t]\Big)$ is a Čech complex of the cover $\mathcal{U}(t)$ whose elements are open d-spheres of radius t centered at points in X. These filtered simplicial complexes are demonstrated in the short example below.

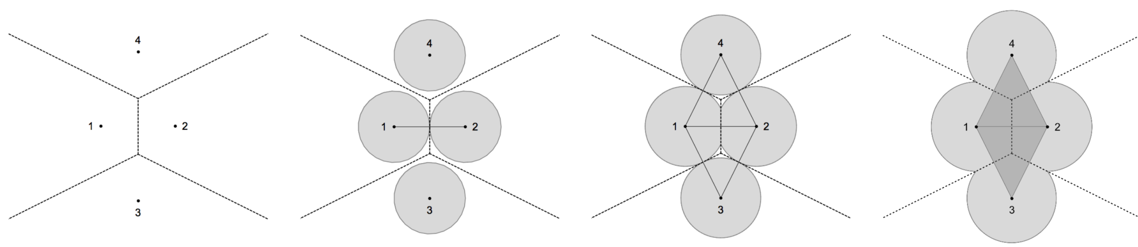


Figure 2.6: This example from [43] shows a four-step real filtration of a Delaunay complex spanned by points 1, 2, 3, and 4 – here, progression from left to right is triggered by increasing the parameter $t \geqslant 0$ (beyond one of several "critical" values). The Voronoi cells that surround each point (whose boundaries are indicated on the left-most image) determine the terminal structure of the complex on the right-most image (including five 1-simplices and two 2-simplices). An allowed 1-simplex $[i, j]$ is then added to the complex when t equals half of the Euclidean distance between vertices i and j; graphically, this is the minimum radius that two closed balls centered at i and j must have to have non-empty intersection. Similarly, an allowed 2-simplex $[i, j, k]$ is added to the complex at the minimum value of t needed such that the intersection of three balls centered at i, j, and k having this radius is non-empty. This specifies the Čech filtration. We only need to note that the Rips filtration would differ by assigning an lower value of the parameter t to the level when the 2-simplices are added to the complex; here, the 2-simplices would be added in the second-to-last image, instead of the last (right-most) image. Furthermore, the terminal complex in the Rips filtration typically typically consists of simplices for every combination of vertices; for example, here we would not omit the 1-simplex $[3, 4]$ from the terminal complex (albeit, it is added to the complex for a relatively large value of the parameter t).

We can use any metric space X in place of subsets of the Euclidean space $\mathbb{R}^d$. More generally, we can begin by taking a function $F : X \to \mathbb{R}$, and consider the sublevel sets ${}_tX := F^{-1}\Big((-\infty, t]\Big)$, with inclusion maps ${}_sX \to {}_tX$ between complexes for any $s \leqslant t$. For how we calculate simplicial homology, we'd also need the assumption that X has a *locally-finite triangulation*, ie. there exists a triangulation of X such that every element of X has an open neighborhood containing it that intersects a finite number of simplices in the triangulation. Even under these assumptions, it still remains to address whether the resulting filtration will give rise to a tractable homology (that is *tame*, as discussed in Chapter 2 of [47]); fortunately, we have a topological criterion [14] to test this:

Theorem 2.20 (also Proposition 2.3 in [47]). *A continuous function $f : X \to \mathbb{R}$ is tame (ie, gives rise to a tame filtration) in any of the following cases:*

1. *the space X has a finite triangulation;*
2. *the space X has a locally-finite triangulation, the preimages of compact intervals under f are compact subsets of X (i.e. f is* proper*), and f has a lower bound on X.*

A good example of a function that satisfies these conditions is a Morse function on a compact X (which can be generically selected among all real-valued maps on X). Then, the resulting filtration ${}_\bullet X$ on X will possess a simple property that there is a finite set of *critical values* $c_1, \ldots, c_m \in \mathbb{R}$ such that ${}_sX \to {}_tX$ is a homotopy equivalence unless $s \leqslant c \leqslant t$ for one of the critical values c. One can see how the theory of tempered persistence vector spaces becomes applicable in this setting.

2.3.2 The Standard Algorithm of Persistent Homology

In 2.1.3, we outlined some conventional methods by which the n^{th} homology of triangulizable space can be computed. Every practical method is based on working with a matrix representation of the differential operator $\partial_\bullet$, and at minimum produces representative modules B_n of all boundary n-chains (ie, the range of ∂_{n+1}) and Z_n of all n-cycles (ie, the nullspace of ∂_n). A complete calculation would also provide a function $B_n \to Z_n$ that identifies boundaries as a subset of the set of cycles, relying on the equation $\partial_\bullet^2 = 0$ (nilpotency). Then, the n^{th} homology with field coefficients is just represented by those n-cycles which are not n-boundaries.

In persistent homology this is more complicated because the sets of cycles and boundaries may vary between different levels of the filtration. Naively, this can be resolved by doing above calculations inductively, where in each repetition the entire complex gets restricted to a fixed level. However, this alone is insufficient to describe how the homology classes at each level "combine" into persistent features. The goal here is to provide for each persistent feature a choice of n-cycle at each level (representing a homology class at that level) such that the natural inclusions which transition from lower to higher levels take one choice of cycle to another.

To do this, a method involving only one round of matrix reduction is used. Let n be the value of the final critical level in the filtration. Take a basis of n-simplices comprising the underlying simplicial complex, and let this basis be ordered consistently with the filtration function applied to each simplex, the order by which the simplices "appear" in the filtration; for simplices which have the same filtration level, there is freedom in choosing their order in the basis (for example, if the simplices are labelled alphabetically then the dictionary order can be used to break ties). This basis spans the vector space C_n, and the same can be done for the spaces C_{n-1} of $(n-1)$-simplices and C_{n+1} of $(n+1)$-simplices. Then, the components $\partial_{n+1} : C_{n+1} \to C_n$ and $\partial_n : C_n \to C_{n-1}$ of the differential operator ∂ can each be put into matrix form with respect to these bases. Observe that restricting these bases to elements having level less than or equal to a fixed level t will produce bases for subspaces of C_{n-1}, C_n, C_{n+1} that represent the $(n-1)$-, the n-, and $(n+1)$- chains of the filtration at level t, respectively; we denote ${}_tC_p$ to be the vector space generated by all p-chains having a level in the filtration that is $\leqslant t$, and ${}_t\partial_p : {}_tC_p \to {}_tC_{p-1}$ to the restriction of ∂_p to this level in the filtration. In fact, these are represented by upper-left submatrices of the given matrix representation of ∂_n and ∂_{n+1}.

This construction allows us to restate the problem that we have posed: while the operators ${}_t\partial_n$ and ${}_t\partial_{n+1}$ can be used to calculate the bases of n-cycles and n-boundaries (resp.) up to a chosen level t, the resulting bases need to restrict to a corresponding basis at every level s if $s \leqslant t$. It is therefore reasonable to work recursively, beginning with reduction at the lower levers and proceeding towards higher levels in the filtration. This can potentially be done in various ways (depending on the complexity, or *degeneracy*, of the filtration), but we will prioritize one algorithm that is straightforward in its execution for any such setup. Because we are working inductively (where the initial step can be assumed to be null, as we'll see shortly) and the filtration is has a finite list of critical points, we begin by supposing that we have have found bases of cycles and boundaries that have the desired consistency across every level $s < t$ up to (but not including) some fixed critical value t.

Write the matrix form of the linear operator ${}_t\partial_n : {}_tC_n \to {}_tC_{n-1}$, with respect to the ordered bases of standard simplices. The column vectors of this matrix represent the image of each basis element of ${}_tC_n$ under the map ∂_n. For each column, the non-zero entries represent constituent $(n-1)$-simplices in the boundary of the selected

n-simplex, and the one closest to the bottom row of the matrix represents the one that appears the latest in the filtration (in other words, the "youngest"); call this non-zero entry the *pivot* of the given column. Perform a similar procedure and analysis on the linear operator ${}_t\partial_{n+1} : {}_tC_{n+1} \to {}_tC_n$.

Now do the following in each matrix: for every column, starting with the left-most one, check to see if the row that its pivot is in is different that the rows the pivots of all columns to the left of it are in. If true (vacuously when working on the left-most column), then repeat this for the column immediately on the right. But if you find two columns whose pivots are on the same row, then subtract a constant multiple of the left column from the right column so that the right column either becomes the zero vector or has its pivot appear in a new row; replace the right column with this new vector in the matrix and keep the left column as it was. Keep track of these operations as you work through each column of either matrix.

This is fundamentally just matrix reduction with permitted column operations – specifically those that adjust columns by linear combinations of columns on their left. The first reduction produces a matrix form of ${}_t\partial_n$ with respect to a new (calculated) basis of ${}_tC_n$ and the old basis of ${}_tC_{n-1}$; the second reduction produces a matrix form of ${}_t\partial_{n+1}$ with respect to a new (calculated) basis of ${}_tC_{n+1}$ and the old basis of ${}_tC_n$. Both matrices have the property that every non-zero column has its pivot in a distinct row from the others. The zero columns in the resulting matrix form of ${}_t\partial_n$ correspond to the new basis elements of ${}_tC_n$ whose $(n-1)$-boundary is null; these form an ordered basis of ${}_tZ_n$. The nonzero columns in the resulting matrix form of ${}_t\partial_{n+1}$ represent a maximal set of elements in the image subspace of ${}_t\partial_{n+1}$ that are linearly independent;

these form an ordered basis of ${}_tB_n$.

It is also easy to see that this process produces consistent results across all critical values of the parameter $t \in \mathbb{R}$: if $s < t$, the image of any standard simplex for the domain space ${}_sC_p$ under the operators ${}_s\partial_p$ and ${}_t\partial_p$ has the same representation in the basis of standard $(p-1)$-simplices; in particular, the non-zero entries of its corresponding column vectors in ${}_s\partial_p$ and ${}_t\partial_p$ are identical (the "longer" vector just has a longer tail of trailing zero entries). Therefore, reduction of the matrix for ${}_t\partial_p$ actually invokes the matrix reduction for ${}_s\partial_p$ (as an upper-left submatrix), with identical column operations on the basis of ${}_sC_p$. This justifies the induction step.

Trivially, $\partial_n \circ \partial_{n+1} = 0$ implies that ${}_t\partial_n \circ {}_t\partial_{n+1} = 0$, and therefore it holds that ${}_tB_n \subseteq {}_tZ_n$ for all values of the parameter $t \in \mathbb{R}$. This does not mean that the selected basis ${}_tB_n$ will be a subset of the selected basis for ${}_tZ_n$; however, certain (allowed) linear combinations with basis elements of ${}_tZ_n$ do form the selected basis of ${}_tB_n$. A crucial observation is that each basis element of ${}_tZ_n$ has a "youngest" n-simplex that is distinct than that of other elements, all because of how this basis was calculated using column operations. It then follows that for every basis element in ${}_tZ_n$ there is at most one basis element of ${}_tB_n$ whose youngest n-simplices match, due to the uniqueness of row-positions for column pivots in the latter. In latter chapters this will be used to derive the matrix representation of the inclusion map ${}_tB_n \hookrightarrow {}_tZ_n$, but it suffices to say for now that this allows to establish the cycle-boundary correspondence that defines the homology of the space.

Then to conclude the previous discussion, the homology (in field coefficients) of the complex at a fixed level $t \in \mathbb{R}$ is given by the quotient ${}_tZ_n / {}_tB_n$, which is isomorphic to a vector space that is generated by basis elements in ${}_tZ_n$ (cycles) with no corresponding basis elements in ${}_tB_n$ (boundaries). Clearly the homology of the complex will vary across different values of t, but we can study its behavior by noting that we have a persistence vector space of field homologies:

$$\cdots \longrightarrow {}_{t_{i-1}}Z_n/{}_{t_{i-1}}B_n \longrightarrow {}_{t_i}Z_n/{}_{t_i}B_n \longrightarrow {}_{t_{i+1}}Z_n/{}_{t_{i+1}}B_n \longrightarrow \cdots$$

Here, the indicated values t_j on the diagram above can be taken to be the critical values of the filtration – only then can new cycle and boundary elements can form in the complex. As discussed before, this persistence vector space has a barcode decomposition – in this context, each interval $[b_k, d_k)$ in the barcode is associated to a distinct homology class that is *born* at $t = b_k$ and *dies* at $t = d_k$ (some homology classes do not formally die, in which case $d_k = \infty$ by convention). But the output of this algorithm above allows to compute these directly – the birth of a new homology class is indicated by the addition of an element to the basis of cycles at $t = b_k$, and its death is indicated by the addition of a corresponding element in the basis of boundaries at $t = d_k$ (deaths at $t = \infty$ occur exactly for cycles that have no corresponding boundary even in the terminal level of the complex). Hence, simply knowing the bases of ${}_tZ_n$ and ${}_tB_n$ provides enough information to deconstruct the persistent homology of the complex. This completes the intent of using the algorithm.

Finally, it should become apparent that the algorithm could have been performed simply on the differentials ∂_n and ∂_{n+1} (of the terminal complex) from the start to produce the sought-after consistent results. This final remark, completely specifies the algorithm and its correctness. We summarize it in pseudocode on the next page; note that "partner" refers to the cycle-boundary correspondence.

Data: an $n \times n$ (differential) matrix $D = (D_1 \dots D_n)$
Result: returns array partner[$\bullet$], upper uni-triangular $n \times n$ matrix U

```
procedure PairCells()
    partner[n]=(0, ..., 0) ; // row positions of each column pivot
    U = I_n ; // change of basis, will satisfy D' = DU is reduced
    foreach 1 ⩽ j ⩽ n do
        EliminateBoundaries(j);
        if DU_j ≠ 0⃗ then
            i=Youngest(j);
            partner[j]=i;
    return ; // see output
procedure EliminateBoundaries(j)
    while DU_j ≠ 0⃗ do
        i=Youngest(j);
        if partner[i]≠ 0 then U_j = U_j + U_partner[i];
        else return; // pivot has been found
    return ; // at this point, U_s is a cycle
function Youngest(c)
    p="LastNonzeroPosition"(DU_c) ;
    return p ;
```

Algorithm 1: Modified ZC algorithm in matrix form (over $\mathbb{Z}_2$), from [60].

This powerful algorithm was originally published by Afra Zomorodian and Gunnar Carlsson in their seminal paper [60]. Its original intent was to explicitly calculate the barcode of the homology sequence for a complex and some of its representatives, but has later been extended/included in other algorithms that require a consistent way to work with bases across multiple levels in some filtration(s). Some of these will be discussed in future chapters.

2.3.3 Stability in Persistence Homology

The purpose of the discussion up until now was to show how a persistence vector space ${}_\bullet V$ can be generated to abbreviate the structure of some point cloud data, followed by how to calculate its unique barcode $\mathcal{B}({}_\bullet V)$ via a series of matrix operations – we already knew this was possible from the discussion in 2.2.3. This, in theory, should allow us to extract distinguishable "patterns", with certain "prominence" that can be quantified using the barcode. In practice however, there are potential complications.

In the course of any experiment, the process of gathering data is naturally susceptible to small perturbations, even in a controlled environment. This can be attributed to various factors including instrumental error, imprecision in the model and hypotheses used, and perhaps partially due to the nature of the experiment itself. As such, different iterations of even the most well-executed experimental setups will produce data sets with some degree of deviation from each other. Naturally, this carries over to every step in the execution of persistent homology for each dataset resulting in two barcodes, each an invariant of either set under vector space isomorphisms. It should not be a surprise that the two barcodes in this setup are very likely to be different – yet on the other hand, the expectation is that small perturbations in data sets shouldn't drastically affect their underlying homological structure, thus raising some questions. Are intervals "stable" under continuous changes? Overall, are barcodes "robust" when perturbed? How "much" perturbation in a data set can we reasonably allow; and even then, how do we actually relate the barcodes of tightly-comparable data sets? Ultimately, we come to realize that being able to produce an invariant like this is a pretty low "benchmark" in practice.

Fortunately, the stability theorem puts these questions (about the validity of using barcodes as meaningful predictors of data features) to rest. An inherent part of its mechanism is the notion of distance – not only for measuring the deviation of data sets (commonly via mean square error, or the chessboard metric), but also that of barcodes themselves. In previous sections we described something like this with the bottleneck distance, in the process of calculating which we also find a "least-energy" matching between intervals of the two barcodes. In particular, a small bottleneck distance implies that any pair of matched intervals have their respective birth and death indices be approximate to each other (while simultaneously discarding short intervals from both sets as "insignificant"). So at the very least, complexes with similar barcodes have relatable "prominent" homological features (associated to intervals of significant length which are matched); as it also turns out, this relatability is made explicit via induced morphisms of such complexes.

The crucial thesis of the stability theorem is that this relatability is tightly controlled by the geometry of the complexes themselves. The original result [15] was actually stated primarily for filtrations of a topological space X induced by the level sets of some continuous real-valued functions f and g; here the distance used was the L^∞ norm on the function space of X, defined by $(f,g) \mapsto ||f-g||_\infty := \sup_{x\in X} |f(x)-g(x)|$.

Theorem 2.21. *Let X be a triangulizable space and $f, g : X \to \mathbb{R}$ be continuous tame functions. Then:*

$$d_B(\mathcal{B}({}_\bullet F), \mathcal{B}({}_\bullet G)) \leqslant ||f-g||_\infty$$

where ${}_\bullet F$ and ${}_\bullet G$ are the sequences of homology spaces induced by the filtration of levels sets of f and g, respectively.

Theorem 2.21 was originally proved by carefully separating box-like sets out of the persistence diagrams of ${}_\bullet F$ and ${}_\bullet G$ (whose bounds were subject to slight perturbations), from which inequalities could be produced that bound their cardinality by each other; controlling the perturbation by the topological proximity of f to g

(which involved convex approximations) then established the result. This idea was further refined in [13] to introduce the notion of ε-interleavings of spaces that could be *interpolated*, exhibiting a certain density with respect to the interleaving metric (a detailed proof of which can be found in [14]); the paper also generalized the original ("strong") interleaving to "weak" interleavings (where index sets other than $\mathbb{R}$ are considered), although utility of strong interleavings still gets greater attention in practice.

A simple proof of Theorem 2.21 uses the Isometry Theorem. Letting $||f-g||_\infty = \varepsilon$, it readily follows that the sublevel sets of f and g are ε-interleaved: for any $t \in T$:

$$f^{-1}\Big((-\infty, t]\Big) \subseteq g^{-1}\Big((-\infty, t+\varepsilon]\Big) \qquad \text{and} \qquad g^{-1}\Big((-\infty, t]\Big) \subseteq f^{-1}\Big((-\infty, t+\varepsilon]\Big)$$

In [5], it is then shown that the homology spaces ${}_\bullet F$ and ${}_\bullet G$ are ε-interleaved and hence $d_I\left({}_\bullet F, {}_\bullet G\right) \leqslant \varepsilon$. The statement of Theorem 2.19 then concludes the proof. It should be noted that until recently [38], proofs of the stability theorem involved a weaker form of the isometry theorem, the algebraic stability theorem: namely, the 1-Lipschitz continuity of the bottleneck metric with respect to interleaving distance. While this can be used instead of Theorem 2.19, the latter result implores a deeper categorical relation between persistence vector spaces and their persistence diagrams by constructing an explicit correspondence between creator/destroyer pairs of interval modules. This structural hallmark deserves to be fully disclosed, later.

It remains to apply Theorem 2.21 in the context of finite point cloud data – specifically, those point clouds that are embedded in Euclidean space $\mathbb{R}^m$ and give rise to either a Rips or a Čech filtration. In order to quantify the discrepancy between two point clouds, one needs to use a metric on sets. The following is customarily used:

recall that the *Hausdorff distance* $d_H(A, B)$ of two subsets A and B in a metric space (with metric ρ) is defined as

$$d_H(A, B) = \max\left\{\sup_{a\in A}\inf_{b\in B}\rho(a, b)\ ,\ \sup_{b\in B}\inf_{a\in A}\rho(a, b)\right\}$$

This becomes zero if and only if $A = B$. Now suppose you have two point clouds X and Y that are subsets of $\mathbb{R}^m$, with $\rho = \rho_2$ given by the Euclidean norm. if K_X and K_Y are their respective m-convex hulls, then the polarization identity on ρ_2 implies that $d_H(K_X, K_Y) = d_H(X, Y)$; in fact, if ρ is given by any other vector space norm then from the basic methods of convex geometry one can show that the d_H-distance between X and Y is an upper bound for the d_H-distance of their convex hulls. Since the Rips and Čech schemes each produce filtrations of some triangulations on the convex hulls of X and Y, this shows it is sufficient to continue using the measurement $d_H(X, Y)$ while working the simplicial complex at hand.

Suppose that X and Y have the same cardinality and that the Hausdorff distance between them is sufficiently small ($\varepsilon > 0$) for there to be an induced pairwise bijection between points of X and of Y. Then a triangulation of K_X induces a corresponding triangulation of K_Y, and vice versa. Using a Rips or Čech filtration scheme produces level functions λ_X and λ_Y on respective simplicial complexes so that the positive difference in the levels of corresponding 1-simplices is at most ε. Higher-dimensional simplices behave similarly; the positive difference in levels of corresponding n-simplices is at most ε in the Rips scheme, and at most $\varepsilon\sqrt{2}$ in the Čech scheme. For the purpose of simplicity, it is sensical to consider that the two filtered complexes arise as filtrations (leveled by functions λ_X and λ_Y) of a single "common" simplicial complex.

The next result then allows us to view this setup in terms of sublevel set filtrations on their "common" simplicial complex.

Lemma 2.22. *Let K be a finite simplicial complex. Suppose that ${}^{1}_{\bullet}K$ and ${}^{2}_{\bullet}K$ are two filtrations of K (over a discrete set $T \subset \mathbb{R}$) with associated level functions $\lambda_1, \lambda_2 : K \to T$, given by $\lambda_i(\sigma) = \min\{t \in T : \sigma \in {}^{i}_{t}K\}$ (for every face $\sigma \in K$). Then there exist tame functions $f_1, f_2 : |K| \to \mathbb{R}$ such that $f_i^{-1}\Big((-\infty, t]\Big)$ is homotopy-equivalent to $|{}^{i}_{t}K|$ for all $t \in T$ and $\max_{c\in|K|} |f_1(c) - f_2(c)| = \max_{\sigma\in K} |\lambda_1(\sigma) - \lambda_2(\sigma)|$.*

Proof. Let K' be a subdivision of K that contains a 0-simplex sampled from the interior of $|\sigma|$ for every face σ of K, such as its barycenter. Then λ_1 and λ_2 each induce (trivially) a unique map $l_1, l_2 : V(K') \to T$, where $V(K')$ is the vertex set of K'. Let $f_1 : |K'| \to \mathbb{R}$ and $f_2 : |K'| \to \mathbb{R}$ be the affine extensions of each respective $V(K') \to T$ to the geometric realization $|K'|$ of K'; that is, $f_i(x) = \sum_{v\in V(\sigma)} \tau_v(x) l_i(v)$ where $x = \sum_{v\in V(\sigma)} \tau_v(x) v$ for some $x \in \sigma \in K'$ and $\sum_{v\in V(\sigma)} \tau_v(x) = 1$ (this assignment is well-defined since K' is a simplicial complex).

It is straightforward to verify that each f_i is continuous on $|K'| \simeq |K|$ and is thus a tame function by Theorem 2.20(1). The homotopy equivalence between $f_i^{-1}\Big((-\infty, t]\Big)$ and $|{}^{i}_{t}K|$ holds since the natural embedding of the latter into $|K'|$ is a *deformation retract* of the former – it is possible to "retract" the points on the topological boundary of $f_i^{-1}\Big((-\infty, t]\Big)$ to "adjacent" points on the boundary of $|{}^{i}_{t}K|$ by a collection of uniformly-parametrized paths in $f_i^{-1}\Big((-\infty, t]\Big)$.

Finally, the max and min values of $f_1(x) - f_2(x) = \sum_{v\in V(\sigma)} \tau_v(x)\Big(l_1(v) - l_2(v)\Big)$ on $|\sigma|$, for any face σ of K', occur at $V(\sigma)$. Hence, $\max_{c\in|K'|} |f_1(c) - f_2(c)|$ equals $\max_{c\in V(K')} |f_1(c) - f_2(c)|$, which is trivially equal to $\max_{\sigma\in K} |\lambda_1(\sigma) - \lambda_2(\sigma)|$. Result follows since $|K'| \simeq |K|$. ∎

This leads us to arrive at two conclusions. Firstly, we can settle the pressing question on the stability of persistent homology when working with some common filtrations in Euclidean space – at least, when some conditions are met. Specifically, for N-point clouds X and Y whose Hausdorff distance $d_H(X, Y) = \varepsilon$ is sufficiently small, Theorem 2.21 guarantees that for a chosen filtration scheme the bottleneck distance between the barcodes associated with K_X and K_Y is a robust metric; recalling the discussion in section 2.2.4, there exist ε-matchings of barcodes under the Rips scheme and $\varepsilon\sqrt{2}$-matchings of barcodes under the Čech scheme.

If this (conceivably stringent) requirement on the point clouds X and Y is not met, one may perform additional preprocessing of the data by available techniques. When X and Y are not of the same size, one may attempt to construct witness complexes of X and Y, whereby the datasets are resampled to reduce clustering of entries; see [54] for the original paper. On the other hand, when working with noisy data X and/or Y it is generally allowable to perform techniques by which the data is "smoothed out", typically by statistical kernels; see [6] for a recent study of this approach. Another issue to be wary of is the sensitivity of the triangulations used on K_X and K_Y to small perturbations overall. In particular, it is known that the Delaunay triangulation (which is often used in practice) may change considerably even for small perturbations of a dataset; some proposals [7] attempt to resolve this currently.

Secondly, Lemma 2.22 shows that the level functions λ associated with (tame) filtered spaces carry enough information to produce meaningful summaries (barcodes) of topological data. So, we shall set the geometrical aspects of these structures aside and refocus on them as tempered persistent vector spaces – with a heavy emphasis on concrete vector space computations.

Chapter 3

Category Theory

So far in our overview, we were able to derive notions from persistent homology via rigorous calculations, going through many constructions before finally obtaining a summary of the data as a barcode. Category theory allows us to then speak of these constructions in general terms, with special focus given to how they compare to each other. Classical texts on the subject include [40] and [1].

An archetypical category is **Set**, a proper collection consisting of all conceivable sets and a collection of all ordinary functions/maps between them. With this minimal description, we will see that many common set operations can be established including products, isomorphisms, and equivalence relations. However, normally when invoking sets, we exercise these operations on its individual elements or subsets; in category theory, there is no natural notion of "elements" that "belong" to a set.

3.1 General Introduction

Formally, a *category* $\mathcal{C}$ consists of a collection of *objects* and a collection of *morphisms* between each ordered pair of objects. By analogy of objects to sets and morphisms to functions, a morphism f going from A to B will be denoted as $f : A \to B$, with $\mathrm{Hom}_{\mathcal{C}}(A, B)$ often used to denote the collection of all such morphisms. Consecutive morphisms may be *composed*: given objects A, B, C and morphisms $f : A \to B$, $g : B \to C$, composition gives a unique morphism $(g \circ f) : A \to C$. We only require that the composition law obeys the associativity property: $(h \circ g) \circ f = h \circ (g \circ f)$, for any morphisms $f : A \to B$, $g : B \to C$, and $h : C \to D$. Finally, we assume that every object admits a unique *identity morphism* $1 = 1_A : A \to A$ such that: $f \circ 1 = f$ and $1 \circ g = g$, for any morphisms $f : A \to B$, $g : C \to A$. This is generally as far as the ongoing analogy extends.

Categories can be constructed from an existing category $\mathcal{C}$. One elementary way involves taking a category $\mathcal{D}$ whose objects form a subcollection of objects in $\mathcal{C}$, and whose every morphism $f : A \to B$ also appears as a morphism in $\mathcal{C}$. Then $\mathcal{D}$ is a *subcategory* of $\mathcal{C}$ (written $\mathcal{D} \subset \mathcal{C}$) if the composition law of $\mathcal{D}$ is also consistent with the composition law of $\mathcal{C}$. Many categories can be formed this way; for instance, one may take the subcategory **set** of **Set** consisting of finite sets as its objects. Another category that is naturally constructed from $\mathcal{C}$ is the *opposite* (or *dual*) category $\mathcal{C}^{op}$, containing all the objects of $\mathcal{C}$ and "reversing" every morphism $f : A \to B$ in $\mathcal{C}$ to get $f^{op} : B \to A$ in $\mathcal{C}^{op}$; however, this will not be used in our discussion.

The intention here is to restate mathematical notions solely using the above language of objects and morphisms; such notions are then "categorical". For instance, with allusion to injective and surjective set functions, we define *monomorphisms* (or simply, *monics*) and *epimorphisms* (or simply *epics*). We say $f : A \to B$ is monic if the equation $f \circ g = f \circ h$ only holds when $g = h$ (given any $g, h : C \to A$). Similarly, we say $f : A \to B$ is epic if the equation $g \circ f = h \circ f$ only holds when $g = h$ (given any $g, h : B \to C$). In a related vein, some morphisms $f : A \to B$ admit some

$g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$; this f is called an *isomorphism*, and g is its *inverse* (it is in fact unique). This allows us to talk about *isomorphic* objects within a category. We remark that any isomorphism is both monic and epic, although (outside of **Set**) the converse may be false.

Another example concerns the existence of certain *initial* and *terminal objects*. Given any object A in category $\mathcal{C}$, an initial object X admits exactly one morphism $X \to A$, and a terminal object Y admits exactly one morphism $A \to Y$. In **Set**, the empty set $\emptyset = \{\}$ is the only initial object with the morphism being set inclusion. On the other hand, any set with just one element in **Set** is a terminal object; given a set A, we define the morphism to be the map sending every element of A to the only value of the singleton set. By contrast, take a look at the category $\mathbf{Vect}(\mathbb{F})$ consisting of all conceivable vector spaces (objects) and all the linear maps between them (morphisms), over some fixed field $\mathbb{F}$. Every vector space has a unique zero-dimensional subspace $\langle 0 \rangle$; this subspace acts as both an initial and a terminal object for $\mathbf{vect}(\mathbb{F})$, since $\langle 0 \rangle \to V$ and $V \to \langle 0 \rangle$ are clearly unique! An object acting both as an initial and a terminal object in a category is rightfully called a *zero object*.

One might be lightly stricken by this contrast between **Set** and $\mathbf{Vect}(\mathbb{F})$ upon the realizing that the latter is a subcategory of the former. Indeed, one might elaborate that **Set** has "more morphisms" than $\mathbf{Vect}(\mathbb{F})$, ensuring that its terminal and initial objects are "distinct enough" (while it's also true that the empty set is not an object of $\mathbf{Vect}(\mathbb{F})$ to begin with). This should highlight the benefit of considering these two as separate categories, in order to emphasize particular facts about each.

Many other advanced notions can be described in more or less the same manner by which some special objects and morphisms can be canonically identified in $\mathcal{C}$, after satisfying some existence and uniqueness conditions. Commonly, *diagrams* are employed to demonstrate these notions in a neat and minimal way, by focusing attention to some subcategory of $\mathcal{C}$ whose objects and morphisms can be indexed by proper sets (what is called a *small category*). If the index sets are finite (or at least countable), then the diagram may be represented by a graph. For example, the following diagram describes a zero object and *zero morphisms* $0_{V,W} : V \to W$ in $\mathbf{Vect}(\mathbb{F})$:

$$\begin{tikzcd} V \arrow[rr, "0_{V,W}"] \arrow[dr, "!"] & & W \\ & Z \arrow[ur, "!"] & \end{tikzcd}$$

Figure 3.1: Here, V and W are any two vector spaces, Z is a representative zero object, and the composition law defines $0_{V,W}$. The exclamation signs are conventionally used to indicate the uniqueness of the given morphism.

A word of caution is due, however – our ability to do prescribe uniqueness of objects has an inherent limitation. Namely, a categorical property that is satisfied by some object will also be satisfied by all other objects isomorphic to it; for this reason, it is common to clarify that the existence and uniqueness of a certain object holds up to isomorphism. For example, note that every vector space (over $\mathbb{F}$) has a zero-dimensional subspace but not all vector spaces are generated by elements from the same universal space; that is to say that a zero object $\mathbf{0}$ does not have some singular presence among all objects of $\mathbf{Vect}(\mathbb{F})$, and can only be represented by a selection of some zero-dimensional vector space (among all its isomorphic copies).

We'd prefer to avoid this technicality with morphisms, which is possible in many cases where the category can be shown to be *locally small* – that is, for any two objects A and B the collection $\mathrm{Hom}(A, B)$ is a proper set. This allows to refer to morphisms as concrete elements, and for any two isomorphic objects A and B we derive natural bijections in $\mathrm{Hom}(A, C) \simeq \mathrm{Hom}(B, C)$ and $\mathrm{Hom}(C, A) \simeq \mathrm{Hom}(C, B)$ (for any object C). Indeed, it is especially evident with zero objects as every vector space V admits unique maps $\langle 0 \rangle \to V$ and $V \to \langle 0 \rangle$.

The ability to declare that some prescribed morphisms in a diagram are unique is a crucial component for stating *universal morphisms* that exhibit a *universal property*. These will be important later for defining several familiar objects such as kernels, cokernels, pullbacks, and direct sums of mathematical objects.

3.1.1 Functors and Equivalences

A *functor* $\mathcal{F}$ between two categories $\mathcal{C}$ and $\mathcal{D}$ (written $\mathcal{F} : \mathcal{C} \to \mathcal{D}$) assigns to every object X from $\mathcal{C}$ an object $\mathcal{F}(X)$ of $\mathcal{D}$. Every functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ is assumed to be one of these two types: *covariant functors* assign to every morphism $A \to B$ a morphism $\mathcal{F}(A) \to \mathcal{F}(B)$, while *contravariant functors* assign the morphism to some $\mathcal{F}(B) \to \mathcal{F}(A)$. Furthermore, the composition laws need to be compatible: if $f : A \to B$ and $g : B \to C$, then $g \circ f$ is assigned to $\mathcal{F}(g) \circ \mathcal{F}(f)$ by covariant functors and to $\mathcal{F}(f) \circ \mathcal{F}(g)$ by contravariant functors. Lastly, a functor $\mathcal{F}$ must preserve identity morphisms: for any object $A \in \mathcal{C}$, if $1_A : A \to A$ and $1_{\mathcal{F}(A)} : \mathcal{F}(A) \to \mathcal{F}(A)$ are identity morphisms then $\mathcal{F}(1_A) = 1_{\mathcal{F}(A)}$.

The presence of a functor allows us to speak of comparable structures within separate categories. In our case, covariant functors shall be predominantly used.

Example 3.1. Consider $\mathbf{Vect}(\mathbb{F}) \to \mathbf{Set}$, assigning any given vector space to its underlying set and every linear map to its underlying function. It is readily verified that this seemingly natural operation is a covariant functor, which we refer to as the *forgetful functor* between $\mathbf{Vect}(\mathbb{F})$ and $\mathbf{Set}$.

More generally, a category $\mathcal{C}$ is called *concrete* if there exists a covariant function $\mathcal{C} \to \mathbf{Set}$ such that such that any two distinct morphisms in $f, g : A \to B$ in $\mathcal{C}$ are assigned to distinct morphisms in $\mathbf{Set}$ (that is, the functor maps the Hom-sets of $\mathcal{C}$ injectively to the Hom-sets of $\mathbf{Set}$); the functor between them can be referred to as a *faithful embedding*. Thus, the objects of a concrete category are endowed with an underlying set structure and its morphisms behave like functions on the set.

Example 3.2. Let $\mathbf{vect}(\mathbb{F})$ be a subcategory of $\mathbf{Vect}(\mathbb{F})$ whose objects are finite-dimensional vector spaces with linear maps as morphisms between them. Then we have an "inclusion" functor $\mathbf{vect}(\mathbb{F}) \to \mathbf{Vect}(\mathbb{F})$, mapping objects and morphisms of the subcategory to themselves as constituents of the larger category. This is clearly a covariant functor that maps Hom-sets in $\mathbf{vect}(\mathbb{F})$ surjectively to their respective Hom-sets in $\mathbf{Vect}(\mathbb{F})$. However, there is no functor $\mathbf{Vect}(\mathbb{F}) \to \mathbf{vect}(\mathbb{F})$ that doesn't assign all objects to $\langle 0 \rangle$.

In the last example, it can be said that $\mathbf{vect}(\mathbb{F})$ is a *full subcategory* of $\mathbf{Vect}(\mathbb{F})$, since the Hom-sets in $\mathbf{vect}(\mathbb{F})$ are identical to their assigned Hom-sets in $\mathbf{Vect}(\mathbb{F})$. The functor between them can be referred to as a *full embedding*.

Every category $\mathcal{C}$ always admits an *identity functor* $\mathbb{1}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$ that simply assigns every object A to A and every morphism $A \to B$ to itself. *Composition* of two functors is also defined in a natural manner. With these notions, it is possible to describe the notion of functor "invertibility" – specifically, two functors $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ are inverses of each other if their composition yields $\mathcal{G} \circ \mathcal{F} = \mathbb{1}_{\mathcal{C}}$ and $\mathcal{F} \circ \mathcal{G} = \mathbb{1}_{\mathcal{D}}$. In that case, the functors $\mathcal{F}$ and $\mathcal{G}$ are both *isomorphisms* of categories $\mathcal{C}$ and $\mathcal{D}$, which themselves are simply referred to as *isomorphic categories*.

An isomorphism of categories implies that both categories are indistinguishable by the behavior of their objects and morphisms. It is a very strong condition that is often unnecessarily restrictive. In particular, one often doesn't need the ubiquity of isomorphisms that any given object has within a category (as hinted at in the introduction), for any one of those representatives is all it takes to demonstrate a property of interest. This leads us to adopt instead the concept of "equivalence" as a more effective concept of categorical similarity.

Example 3.3. Let $\mathcal{E}_1$ be a category with one object C and one morphism $1_C : C \to C$; let $\mathcal{E}_2$ be a category with two objects A and B, whose morphisms include the identity morphisms $1_A, 1_B$ as well as some $f : A \to B$ and $g : B \to A$ – necessarily, it must hold that $g \circ f = 1_A$ and $f \circ g = 1_B$ (that is, A and B are isomorphic in $\mathcal{E}_2$).

$$\mathcal{E}_1 = \left\{ \begin{tikzcd} C \arrow[loop right, "1_C"] \end{tikzcd} \right\} \quad , \quad \mathcal{E}_2 = \left\{ \begin{tikzcd} A \arrow[loop left, "1_A"] \arrow[r, bend left, "f"] & B \arrow[l, bend left, "g"] \arrow[loop right, "1_B"] \end{tikzcd} \right\}$$

There are only two choices of functors $\mathcal{E}_1 \to \mathcal{E}_2$, which assign C to either A or B. On the other hand, there is only one functor $\mathcal{E}_2 \to \mathcal{E}_1$, assigning A and B to C and all four morphisms of $\mathcal{E}_2$ to 1_C. In either case, it does not hold that $(\mathcal{E}_1 \to \mathcal{E}_2) \circ (\mathcal{E}_2 \to \mathcal{E}_1) = \mathbb{1}_{\mathcal{E}_2}$. Hence the two categories are not isomorphic.

First, let us revisit categorical diagrams. Let $\mathcal{J}$ be a *small category*, meaning that its objects are elements of some set J and $\forall x, y \in J$ the collection $J_{x,y} = \text{Hom}(x, y)$ of morphisms also is a set; when J and all the Hom-sets of $\mathcal{J}$ are finite (or at least countable) it is common to represent $\mathcal{J}$ as a directed graph (whose nodes represent objects and vertices represent edges). A *diagram* of *shape* $\mathcal{J}$ in a category $\mathcal{C}$ is then a choice of a covariant functor $\mathcal{J} \to \mathcal{C}$. For example, the diagram in figure 3.1 has (that is, indexed by) a shape with three objects $J = \{1, 2, 3\}$ and three morphisms $\{1 \to 2, 1 \to 3, 2 \to 3\}$ while the function maps $2 \to \langle 0 \rangle$ and maps $1, 3$ to a choice of two vector spaces V, W.

What if there are two such diagrams in a category? That is, when $\mathcal{J}$ is a small category and there are two functors $\mathcal{F}, \mathcal{G} : \mathcal{J} \to \mathcal{C}$, it is of interest to determine whether the induced diagrams in $\mathcal{C}$ are "consistent"; a reasonable way to establish this is by seeking a *natural transformation* between $\mathcal{F}$ and $\mathcal{G}$, a sort of "morphism" $\mathcal{F} \Rightarrow \mathcal{G}$ between objects of one diagram to those of the other. Concretely, the transformation assigns to every object $x \in \mathcal{J}$ a morphism $\eta_x : \mathcal{F}(x) \to \mathcal{G}(x)$ in $\mathcal{C}$ such that: for all morphisms $x \to y$ in $\mathcal{J}$, we have the commutation relation $\eta_y \circ \mathcal{F}(x \to y) = \mathcal{G}(x \to y) \circ \eta_x$. This definition easily extends to the definition of a *natural transformation* η between functors $\mathcal{F}, \mathcal{G} : \mathcal{C} \to \mathcal{D}$ of any two categories, and can be denoted $\eta : \mathcal{F} \Rightarrow \mathcal{G}$.

$$\begin{tikzcd}
\mathcal{F}(x) \arrow[r] \arrow[d, "\eta_x"] & \mathcal{F}(y) \arrow[d, "\eta_y"] \\
\mathcal{G}(x) \arrow[r] & \mathcal{G}(y)
\end{tikzcd}$$

Figure 3.2: For any morphism $x \to y$ in a category $\mathcal{C}$, there is a commutation relation that a natural transformation $\eta : \mathcal{F} \to \mathcal{G}$ must satisfy, which is depicted in the diagram above.

A *natural isomorphism* between functors $\mathcal{F}$ and $\mathcal{G}$ is a natural transformation $\mathcal{F} \Rightarrow \mathcal{G}$ that induces an isomorphism $\mathcal{F}(x) \simeq \mathcal{G}(x)$ for all objects x in the source category; one then typically writes $\mathcal{F} \approx \mathcal{G}$. It is quick to verify that this is symmetric relation, i.e. $\mathcal{F} \approx \mathcal{G}$ if and only if $\mathcal{G} \approx \mathcal{F}$.

We can now talk about equivalence, as opposed to strict isomorphism, between categories. Two functors $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ and $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ are *equivalences of categories* if $\mathcal{G} \circ \mathcal{F} \approx \mathbb{1}_{\mathcal{C}}$ and $\mathcal{F} \circ \mathcal{G} \approx \mathbb{1}_{\mathcal{D}}$. This is a weakening of categorical isomorphism, because we only require that the composition of functors $\mathcal{C} \to \mathcal{D} \to \mathcal{C}$ and $\mathcal{D} \to \mathcal{C} \to \mathcal{D}$ can be (naturally) transformed to the respective identity functors $\mathbb{1}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$ and $\mathbb{1}_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}$, and vice versa.

Indeed, consider again the categories $\mathcal{E}_1$ and $\mathcal{E}_2$ in Example 3.3. For any choice of functor $\mathcal{E}_1 \to \mathcal{E}_2$ we have that $(\mathcal{E}_2 \to \mathcal{E}_1) \circ (\mathcal{E}_1 \to \mathcal{E}_2) = \mathbb{1}_{\mathcal{E}_1}$, while $(\mathcal{E}_1 \to \mathcal{E}_2) \circ (\mathcal{E}_2 \to \mathcal{E}_1)$ assigns both A and B to either A or B; suppose it's the former. Then by assigning $\eta_A = 1_A$ and $\eta_B = f$, we get a transformation η between $(\mathcal{E}_2 \to \mathcal{E}_1) \circ (\mathcal{E}_1 \to \mathcal{E}_2)$ and $\mathbb{1}_{\mathcal{E}_1}$; moreover, these functors are naturally isomorphic since η_A and η_B are isomorphisms. Hence, $\mathcal{E}_1$ and $\mathcal{E}_2$ equivalent categories. This may have already been expected after our previous discussion, since both categories essentially describe a singleton object equipped with the identity map – only $\mathcal{E}_2$ includes two isomorphic copies of it. In fact, it is readily verified that $\mathcal{E}_1$ is not equivalent to any subcategory of $\mathcal{E}_2$ that contains both A and B but excludes any (or both) morphisms f and g.

Example 3.4. Consider $\mathrm{Hom}(\bullet, \mathbb{F}) : \mathbf{vect}(\mathbb{F}) \to \mathbf{vect}(\mathbb{F})$, where a vector space V over field $\mathbb{F}$ is assigned to its dual space $\mathrm{Hom}(V, \mathbb{F}) = V^*$ (the space of linear functionals $V \to \mathbb{F}$) and morphisms $\phi : V \to W$ are taken to their adjoint $\mathrm{Hom}(\phi, \mathbb{F}) = \phi^* : W^* \to V^*$ (defined by setting $\phi^*(g) = g \circ \phi$ for any $g : W \to \mathbb{F}$). This is readily verified to be a contravariant functor, and may be composed with itself to obtain the assignment of every finite-dimensional vector space V with its double dual $V^{**} = (V^*)^*$. It is known from basic linear algebra that $V \simeq V^{**}$, via the map

$x \in V \mapsto \Phi_x \in V^{**}$ where $\Phi_x(f) := f(x)$ for any $f \in V^*$. This induces an equivalence of $\mathbf{vect}(\mathbb{F})$ to its opposite category, a phenomenon known as *dual equivalence*.

Oftentimes, another criterion is used to establish equivalence of categories $\mathcal{C}$ and $\mathcal{D}$. Recall that a functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ is *faithful* if for any two objects $A, B \in \mathcal{C}$ the assignment $\mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(\mathcal{F}(A), \mathcal{F}(B))$ is an injective map, and that $\mathcal{F}$ is *full* if $\mathrm{Hom}_{\mathcal{C}}(A, B) \to \mathrm{Hom}_{\mathcal{D}}(\mathcal{F}(A), \mathcal{F}(B))$ is a surjective map. Furthermore, we define a functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ to be *dense* if every object $D \in \mathcal{D}$ is isomorphic to an object of the form $\mathcal{F}(C)$ for some object $C \in \mathcal{C}$.

Theorem 3.5. *A functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$ induces and equivalence of categories if and only if $\mathcal{F}$ is faithful, full, and dense.*

While this is a convenient criterion when we are solely interested in demonstrating that two categories are equivalent, because it only requires to demonstrate the existence of one functor with the properties in Theorem 3.5. At other times though, it is worthwhile to specify a complementary functor $\mathcal{G} : \mathcal{D} \to \mathcal{C}$ for transferring diagrams in one category to another in explicit form.

3.1.2 Notable Categorical Properties and Elements

With the ability to declare diagrams in any given category, we can easily describe several other properties and constructs in category theory. As before, these will be motivated by familiar constructions from the categories **Set** or $\mathbf{vect}(\mathbb{F})$, while keeping in mind that some of these constructions may or may not exist in a given category.

The notion of a zero object $\langle 0 \rangle$ has already been described. Furthermore, for any objects A, B in a category $\mathcal{C}$ with a zero object we have unique *zero morphisms*

$0_{A,B} : A \to B$ and $0_{B,A} : B \to A$ given by the compositions $A \to \langle 0 \rangle \to B$ and $B \to \langle 0 \rangle \to A$ (see figure 3.1). It should readily follows (from the commutativity of diagrams) that composition with a zero morphism produces another zero morphism; that is, $(B \to C) \circ 0_{A,B} = 0_{A,C}$ and $0_{B,C} \circ (A \to B) = 0_{A,C}$ for any objects A, B, C and some morphisms $A \to B$ and $B \to C$. When no ambiguity occurs, we will simply write $0 = 0_{A,B}$ any two objects A and B. This immediately allows one to speak of kernels and cokernels associated to morphisms in $\mathcal{C}$.

The notion of a kernel is common to categories whose objects possess an algebraic structure, via some binary operation. In $\mathbf{vect}(\mathbb{F})$, the kernel of a linear map $f : V \to W$ is typically defined as the subspace $f^{-1}(0) = \{x \in V : f(x) = 0\}$ of V; equivalently, it is the largest subspace of V on which the map f and the zero map $0_{V,W} : V \to W$ coincide. Drawing from this, the *kernel* of a morphism $f : V \to W$ can be described categorically as an object K outfitted with a morphism $i : K \to V$ such that $f \circ i = 0$ and also: for any other choice of K' and $i' : K' \to V$ such that $f \circ i' = 0$, there is a unique morphism $u : K' \to K$ such that $i' = i \circ u$. See the commutative diagrams below.

$$\begin{tikzcd} K \arrow[r, "i"] \arrow[d, "0"] & V \arrow[dl, "f"] \\ W & \end{tikzcd} \qquad \begin{tikzcd} K' \arrow[dr, dotted, "!u"] \arrow[drr, bend left, "i'"] \arrow[ddr, bend right, "0"'] & & \\ & K \arrow[r, "i"] \arrow[d, "0"] & V \arrow[dl, "f"] \\ & W & \end{tikzcd}$$

Figure 3.3: The term “kernel” refers to the object-morphism pair (K, i), not just K itself. The left diagram demonstrates the basic categorical function of the kernel, while the right diagram portrays the *universal property* that this pair satisfies.

The universal property essentially guarantees that the kernel is uniquely defined up to isomorphism. Indeed, the existence of $u : K \to K'$ and $u' : K' \to K$ would imply that $i = i \circ (u \circ u')$ and $i' = i' \circ (u' \circ u)$, while the uniqueness of such maps implies $u \circ u' = i_K$ and $u' \circ u = i_{K'}$; the two objects K and K' are *canonically isomorphic*, as u and u' are the "preferred" morphisms between them. Furthermore, the universal property can be used to show that the morphism $i : K \to V$ is monic, and may thus be understood as akin to the inclusion $f^{-1}(0) \subseteq V$ of vector spaces.

Similarly, cokernels manifest in a category as object-morphism pairs that get associated to morphisms. The cokernel of a linear map $f : V \to W$ is defined as the quotient vector space $W/f(V)$, where $f(V) \subseteq W$ is the image of f; recall that there is a canonical projection $x \in W \mapsto x + f(V) \in W/f(V)$. So in categorical terms, the *cokernel* of a morphism $f : V \to W$ consists of an object C and a morphism $q : W \to C$ such that $q \circ f = 0$ and also: for any other morphism $q' : W \to C'$ such that $q' \circ f = 0$, there is a unique morphism $v : C \to C'$ such that $q' = v \circ q$. See the commutative diagrams below.

As is the case with the kernel, the universal property guarantees the cokernel to be unique up to isomorphism, and in fact there is a canonical isomorphism between any representative objects C and C' of a cokernel. Furthermore, the universal property

$$\begin{tikzcd}
C & W \arrow[l, "q"'] \\
V \arrow[u, "0"] \arrow[ur, "f"] &
\end{tikzcd}
\qquad
\begin{tikzcd}
C' & & \\
& C \arrow[ul, dotted, "!v"] & W \arrow[l, "q"] \arrow[ull, bend right, "q'"'] \\
& V \arrow[u, "0"] \arrow[ur, "f"'] \arrow[uul, bend left, "0"] &
\end{tikzcd}$$

Figure 3.4: Again, "cokernel" refers to the object-morphism pair (C, q), not just C itself. The left diagram demonstrates the basic categorical function of the cokernel, while the right diagram portrays the *universal property* that it satisfies.

can be used to show that the morphism $q : W \to C$ is epic, and may thus be understood as akin to the canonical projection $W \twoheadrightarrow W/f(V)$ of a vector space onto its quotient.

Some authors define the kernel and cokernel of a morphism by means of *equalizers* and *coequalizers*, respectively. To summarize, given a morphism $f : V \to W$, its kernel is a morphism $i : K \to V$ with $f \circ i = 0_{V,W} \circ i$ satisfying a universal property; the cokernel of f is morphism $q : W \to C$ with $q \circ f = q \circ 0_{V,W}$ satisfying its own universal property. Generally, (co)equalizers may not all exist even if (co)kernels do.

We move on to products and coproducts of objects – these are ubiquitous operations that have recognizable analogs even in the category **Set**.

$$\begin{tikzcd} \prod_{i\in I} X_i \arrow[d, "\pi_k"] \\ X_k \end{tikzcd} \qquad\qquad \begin{tikzcd} & \prod_{i\in I} X_i \arrow[d, "\pi_k"] \\ U \arrow[ur, dotted, "!f"] \arrow[r, "f_k"] & X_k \end{tikzcd}$$

Figure 3.5: The term "product" includes the object $\prod_{i\in I} X_i$ and all the canonical projections π_k. On the left diagram is the basic categorical structure of the product, while on the right diagram is its *universal property.*

For two sets A and B, their product customarily refers to the Cartesian product $A \times B = \{(a, b) : a \in A, b \in B\}$; here, recall that there exist so-called coordinate maps $(a, b) \in A \times B \mapsto a \in A$ and $(a, b) \in A \times B \mapsto b \in B$ and that any two functions $C \to A$ and $C \to B$ can equivalently be specified as a unique function $C \to A \times B$. To generalize, products $\prod_{i\in I} X_i$ can be defined for any arbitrary collection $\{X_i\}_{i\in I}$ of sets (while quietly putting down oppositions to the axiom of choice), which owns the same properties. This can be interpreted into categorical language, so the *product* of a collection of objects $\{X_i\}_{i\in I}$ is defined to be an object $\prod_{i\in I} X_i$ that is equipped with morphisms $\pi_k : \prod_{i \in I} X_i \to X_k$ ($\forall k \in I$) called *canonical projections* such that: for any object U and a collection of morphisms $f_k : U \to X_k$ there exists a unique morphism $f : U \to \prod_{i \in I} X_i$ such that $f_k = \pi_k \circ f$. See Figure 3.5.

$$\begin{tikzcd} \coprod_{i \in I} X_i \\ X_k \arrow[u, "\iota_k"] \end{tikzcd} \qquad \begin{tikzcd} & \coprod_{i \in I} X_i \arrow[dl, dotted, "!g"'] \\ V & X_k \arrow[l, "g_k"] \arrow[u, "\iota_k"'] \end{tikzcd}$$

Figure 3.6: The "coproduct" includes the object $\coprod_{i \in I} X_i$ and all the canonical inclusions ι_k. On the left diagram is the basic categorical structure of the coproduct, while on the right diagram is its *universal property*.

The coproduct of a collection of sets $\{X_i\}_{i \in I}$ nominally refers to their disjoint union $\coprod_{i \in I} X_i = \{(i, x) : x \in X_i\}$; here, there exist maps (sometimes called embeddings) $X_k \to \coprod_{i \in I} X_i$ simply taking $x \mapsto (k, x)$. Categorically then, the *coproduct* of a collection of objects $\{X_i\}_{i \in I}$ is defined to be an object $\coprod_{i \in I} X_i$ that is equipped with morphisms $\iota_k : X_k \to \coprod_{i \in I} X_i$ called *canonical inclusions* ($\forall k \in I$) such that: for any object V and a collection of morphisms $g_k : X_k \to V$ there exists a unique morphism $g : \coprod_{i \in I} X_i \to V$ such that $g_k = g \circ \iota_k$. See Figure 3.6.

Similarly to the case with kernels and cokernels, the universal properties of products and coproducts necessarily guarantee that their underlying objects are uniquely identifiable via a canonical isomorphism. In some categories, the canonical morphisms π_k and ι_k are actually epics and monics (respectively).

There is a parallelism in the way that these constructs have been defined. We begin with a diagram $\mathcal{D}$ in our category $\mathcal{C}$, then introduce an object Z with selected morphisms $\varphi = \{\varphi_X\}_{X \in \mathcal{D}}$ such that the resulting diagram commutes. In two cases of interest, where the morphisms in φ are all either of the form $\varphi_X : Z \to X$ or

$\varphi_X : X \to Z$, we denote the resulting diagrams as $\mathcal{D}^{\triangleleft}(Z, \varphi)$ or $\mathcal{D}^{\triangleright}(Z, \varphi)$ respectively, which are usually called the *cone* or *co-cone* of Z to/from $\mathcal{D}$; these describe the basic structure and function of the selected object Z with respect to $\mathcal{D}$.

The exceptionalism of one of these diagrams $\mathcal{D}^{\triangleleft}(Z, \varphi)$ or $\mathcal{D}^{\triangleright}(Z, \varphi)$ over the others is then warranted in the following way: for any other cone $\mathcal{D}^{\triangleleft}(Z', \varphi')$ (or co-cone $\mathcal{D}^{\triangleright}(Z', \varphi')$) that there exists a unique morphism $u : Z' \to Z$ (or $u : Z \to Z'$) such that $\varphi'_X = \varphi_X \circ u$ (or $\varphi'_X = u \circ \varphi_X$) for all $X \in \mathcal{D}$; equivalently, there is a unique natural transformation $\mathcal{D}^{\triangleleft}(Z', \varphi') \Rightarrow \mathcal{D}^{\triangleleft}(Z, \varphi)$ or $\mathcal{D}^{\triangleright}(Z, \varphi) \Rightarrow \mathcal{D}^{\triangleright}(Z', \varphi)$ that restricts to the identity transformation on subdiagram $\mathcal{D}$ to itself. This property that holds for an exceptional cone $\mathcal{D}^{\triangleleft}(Z, \varphi)$ or co-cone $\mathcal{D}^{\triangleright}(Z, \varphi)$ is what is usually meant by a *universal property*, and the object Z with morphisms φ is called the *limit* or *colimit* of $\mathcal{D}$, respectively.

Thus kernels and products can be defined by limits, while cokernels and coproducts can be defined by colimits. Example 3.6 suggests another example of common constructs from categorical theory that can be defined this way.

Example 3.6. For a pair of morphisms $f : A \to C$ and $g : B \to C$ we define their *pullback* $\hat{P}$ to be the limit of the diagram $\mathcal{D} = \{A \xrightarrow{f} C \xleftarrow{g} B\}$; of the diagrams below, the left portrays $\mathcal{D}^{\triangleleft}(\hat{P}, \cdot)$ and the right portrays the universal property of $\hat{P}$.

$$
\begin{tikzcd}
\hat{P} \arrow[r] \arrow[d] \arrow[dr] & A \arrow[d, "f"] \\
B \arrow[r, "g"] & C
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
\hat{P}' \arrow[dr, dotted, "!u"] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\
& \hat{P} \arrow[r] \arrow[d] \arrow[dr] & A \arrow[d, "f"] \\
& B \arrow[r, "g"] & C
\end{tikzcd}
$$

Similarly, for a pair of morphisms $f : A \to B$ and $g : A \to C$ we define their *pushout* $\check{P}$ to be the colimit of the diagram $\mathcal{D} = \{B \xleftarrow{f} A \xrightarrow{g} C\}$; of the diagrams below, the left portrays $\mathcal{D}^{\triangleright}(\check{P}, \cdot)$ and the right portrays the universal property of $\check{P}$.

$$
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"] \arrow[dr] & B \arrow[d] \\
C \arrow[r] & \check{P}
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"] \arrow[dr] & B \arrow[d] \arrow[ddr, bend left] & \\
C \arrow[r] \arrow[drr, bend right] & \check{P} \arrow[dr, dotted, "!u"] & \\
& & \check{P}'
\end{tikzcd}
$$

The last topic to cover here doesn't involve the existence of certain constructs within a category, but rather constructing novel categories altogether. One may already recall from the beginning of the introduction some methods for constructing categories from existing ones.

Given a (locally small) category $\mathcal{C}$, suppose that for any two objects X and Y there is a congruence relation $\sim$ on the morphism set $\text{Hom}_{\mathcal{C}}(X, Y)$; that is, $\sim$ is reflexive ($f \sim f$, $\forall f : X \to Y$), $\sim$ is symmetric (if $f, g : X \to Y$ satisfy $f \sim g$ then $g \sim f$), and transitive (if $f, g, h : X \to Y$ satisfy $f \sim g$ and $g \sim h$ then $f \sim h$). Each of these relations define equivalence relations on every Hom-set of $\mathcal{C}$, but they won't be accordant to the overall structure of $\mathcal{C}$ unless they satisfy the following: if $f, g : X \to Y$ and $h, k : Y \to Z$ are any morphisms such that $f \sim g$ (in $\text{Hom}_{\mathcal{C}}(X, Y)$) and $g \sim h$ (in $\text{Hom}_{\mathcal{C}}(Y, Z)$), then $f \circ h \sim g \circ k$ in $\text{Hom}_{\mathcal{C}}(X, Z)$. Then the relations $\sim$ define a *(categorical) congruence relation* on $\mathcal{C}$.

The *quotient category* of $\mathcal{C}$ with respect to $\sim$ (denoted $\mathcal{C}/\sim$) is a category whose objects are exactly the objects of $\mathcal{C}$ and whose morphisms are equivalence classes $[f]_{\sim} : X \to Y$ for every $f : X \to Y$ in $\mathcal{C}$. In essence, we are erasing the distinction between some morphisms if they are equivalent with respect to $\sim$. As an immediate effect, one may verify that $X \simeq Y$ in $\mathcal{C}$ only if $X \simeq Y$ in $\mathcal{C}/\sim$ (but not vice versa).

A way to obtain such a category from $\mathcal{C}$ is to apply a functor $\mathcal{F} : \mathcal{C} \to \mathcal{D}$. Then one may declare that $f \sim g$ in $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ if and only if $\mathcal{F}(f) = \mathcal{F}(g)$; one readily checks that this is a congruence relation on $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ and that the equivalence relation preserves the composition law on $\mathcal{C}$. Moreover, every congruence relation $\sim$ on $\mathcal{C}$ may be obtained in such a way: there is a functor $\mathcal{C} \to \mathcal{C}/\sim$ that sends objects $X \mapsto X$ and sends morphisms $f \mapsto [f]_\sim$. This will be utilized in section 3.3.2.

3.1.3 Relevant Examples, as used in literature

In this section, we describe some of the more common categories that occur in discussions on persistent homology.

The category **Set** has already been mentioned as a template for understanding the very notion of a category. Morphisms are given by functions beween sets, with monomoprhisms and epimorphisms described exactly by injective and surjective maps (respectively); going further, isomorphisms are given by bijection between sets. The product of sets A and B is represented by the usual Cartesian product $A \times B$ with canonical projections given by the coordinate functions of $A \times B$ onto A and B; on the other hand, the coproduct of A and B is represented by their disjoint union $A \sqcup B$ whose canonical inclusions are given by $a \in A \mapsto (a, 1) \in A \sqcup B$ and $b \in B \mapsto (b, 2) \in A \sqcup B$. But otherwise, this category lacks many other important notions discussed in section 3.1.2 and so we shall seldom consider it further.

Another category, of which **Set** is a subcategory, is **Rel** whose objects are also the objects of **Set** but with *(binary) relations* between sets as morphisms. Recall that a relation ρ on sets A and B is a subset of $A \times B$, and we may write $a \overset{\rho}{\sim} b$ if and only if $(a,b) \in \rho$; it has a *domain* $\mathrm{dom}(\rho)$ and a *range* $\mathrm{ran}(\rho)$:

$$\mathrm{dom}(\rho) = \{a \in A : a \overset{\rho}{\sim} b,\ \text{for some } b \in B\}$$

$$\mathrm{ran}(\rho) = \{b \in B : a \overset{\rho}{\sim} b,\ \text{for some } a \in A\}$$

Composition of relations $\rho : A \to B$ and $\tau : B \to C$ is given by: $a \overset{\tau\circ\rho}{\sim} c$ if and only if $\exists b \in B$ with $a \overset{\rho}{\sim} b$ and $b \overset{\tau}{\sim} c$. The identity morphism 1_A on any object A is in fact the identity relation $a \overset{1_A}{\sim} a$. An interesting property apparent in **Rel** is that every morphism $f : A \to B$ naturally induces an opposite morphism $f^{op} : B \to A$, given by $f^{op} = \{(b,a) \in B \times A : a \overset{f}{\sim} b\}$.

Category **Rel** has several important features that **Set** does not possess. Notably, the empty set $\varnothing$ in this category behaves both as an initial and a terminal object, so **Rel** comes with a zero object; the induced zero morphisms $0_{A,B} : A \to B$ are then just null relations ($a \sim b$ doesn't hold for any $a \in A$ and $b \in B$). Monomorphisms $\rho : A \to B$ in this category satisfy $\mathrm{dom}(\rho) = A$ and that every $b \in B$ admits at most one $a \in A$ such that $a \overset{\rho}{\sim} b$ holds; epimorphisms $\tau : A \to B$ satisfy $\mathrm{ran}(\tau) = B$ and that every $a \in A$ admits at most one $b \in B$ such that $a \overset{\tau}{\sim} b$ holds. Since isomorphisms are necessarily both monic and epic, these are given by bijections between sets. Otherwise, note that $f : A \to B$ is monic if and only if f^{op} is epic.

The kernel of a morphism $f : A \to B$ is given by $i : K \to A$, where $K = A \backslash \mathrm{dom}(f)$ and i expresses the set inclusion relation $K \subseteq A$; this $i : K \to A$ is actually an injective function. Similarly, the cokernel of $f : A \to B$ is given by $q : B \to C$, where $C = B \setminus \mathrm{ran}(f)$ and q is the relation where $b \overset{q}{\sim} b$ for all $b \notin \mathrm{ran}(f)$.

Since **Set** is a subcategory of **Rel**, it sounds sane at first to declare the product of objects A and B to be their Cartesian product $A \times B$ and their coproduct to be the disjoint union $A \sqcup B$, with the same canonical morphisms as in **Set**. However, this causes a problem: the canonical morphisms $\iota_A : A \to A \sqcup B$ and $\iota_B : B \to A \sqcup B$ from **Set** have opposite morphisms ι_A^{op} and ι_B^{op} in **Rel**, yet there are several morphisms $A \sqcup B \to A \times B$ that make the diagram in Figure 3.5 commute, contradicting the uniqueness part of the universal property. As it turns out, for a finite collection of objects $\{X_i\}_{i \in I}$ in **Rel**, their disjoint union $\coprod_{i \in I} X_i$ appears in both the product and the coproduct of the collection! The canonical inclusions of the coproduct are the morphisms ι_k and canonical projections of the product are the morphisms ι_k^{op}. The property that the product and coproduct are isomorphic will play a large role in section 3.2.

Now, there is a particular subcategory of **Rel** that has been shown to be relevant to persistent homology. The category **Mch** of *partial matchings* between sets originally appeared in [19] and [4] showed that it has a connection to the calculation of the bottleneck distance between barcodes.

Define the objects of **Mch** to be all conceivable sets whose morphisms are relations $\rho : X \to Y$ such that $\mathrm{dom}(\rho) \to \mathrm{ran}(\rho)$ is a bijection. After careful inspection, it is evident that **Mch** does not possess all products and coproducts, as per [5]. However, similar to **Rel**, it does contain a zero object (the empty set), zero morphisms between all objects, and every morphism $f : X \to Y$ has a kernel and a cokernel (which are the same as in **Rel**).

Since our discussion of persistent homology began with geometric considerations, it is reasonable to speak some about the categories that inscribe them. There is the category **SCpx** that contains finite simplicial complexes K as objects which, from section 2.1.1, are given by a finite set $V(K)$ and an inclusion-preserving collection F of subsets in $V(K)$, and with simplicial maps as morphisms between objects. There is the forgetful functor **SCpx** $\to$ **Set** (that assigns a complex K to its vertex set $V(K)$), which by Theorem 3.5 induces an equivalence between: the subcategory of all simplicial complexes K that are isomorphic to standard simplices in **SCpx**, and the full subcategory containing all finite sets in **Set**. Otherwise, **SCpx** doesn't greatly differ from the bigger category in terms of properties discussed thus far. Note that the product of K and L has vertex set $V(K \times L) = V(K) \times V(L)$, whose faces are subsets $\mathfrak{f} \subset \sigma \times \tau$ where σ is any face of K and τ is any face of L; the coproduct of K and L has vertex set $V(K \sqcup L) = V(K) \sqcup V(L)$, whose collection of faces is simply given by the disjoint union of the collection of faces for each complex.

While methodologies differ on the subject, we have seen in the last chapter that it is convenient to have an ordering of the elements in $V(K)$ – the category of simplicial complexes with totally-ordered vertices and whose simplicial morphisms are order-preserving can then be denoted **SSet**. There is a forgetful functor **SSet** $\to$ **SCpx**, but a functor in the reverse direction does not exist – simply choosing an arbitrary order on $V(K)$ of every complex K in **SCpx** will not preserve morphisms between objects for as simple as those isomorphic to Δ^1. The difference between the two categories is also reflected in that the notion of a product in **SSet** is more "refined"; for any K and L in **SSet**, we have $V(K \times L) = V(K) \times V(L)$ and the faces of $K \times L$ are subsets $\mathfrak{f} \subset \sigma \times \tau$ (for simplices σ in K and τ in L) such that there is an induced total order on the elements of $\mathfrak{f}$ via $(a, x) \leqslant (b, y)$ in $\mathfrak{f}$ if and only if $a \leqslant b$ and $x \leqslant y$.

Of course, there is also the prominent category **Top**, containing all conceivable topological spaces as objects with continuous functions between them as morphisms. Just like in **Set**, the initial object is given by the empty set $\varnothing$ and the terminal object given by a singleton space; the product and coproduct of a family of top. spaces is are also defined, represented by the Cartesian product (with the Tychonoff topology) and the disjoint union (the final topology with respect to its canonical inclusions), respectively. In other applications, it is also notable that **Top** also contains pushouts for any diagram $X \xleftarrow{f} A \to Y$ of topological spaces – the resulting object is usually denoted $X \cup_f Y$ when $A \to Y$ is subspace inclusion, representing the outcome of "attaching" Y to X along A.

There are functors **SCpx** $\to$ **Top** and **SSet** $\to$ **Top** that send a complex K to its geometric realization $|K|$, but they are defined differently. For simplicial sets, we can identify the vertices of K with vertices on Δ^N (where N is number of vertices of K); then $|K|$ can be defined as in section 2.1.1, while simplicial maps $\varphi : K \to L$ are identified to the continuous functions $|\varphi| : |K| \to |L|$ that preserve convex combinations when restricted to faces of the complexes. On the other hand, the geometric realization of K in **SCpx** may be defined by taking function spaces $\mathrm{Conv}(\mathfrak{f}) = \{p : \mathfrak{f} \to [0,1] | \sum_{x\in\mathfrak{f}} p(x) = 1\}$ for every face $\mathfrak{f}$, and endowing them with the L^∞-norm topology (or any other equivalent norm); with the obvious inclusions on $\mathrm{Conv}(\mathfrak{f})$ induced by inclusions of faces, we may attach them along common subspaces via the pushout to form the topological space $|K|$. A simplicial map $\varphi : K \to L$ produces a continuous function $|\varphi| : |K| \to |L|$ that assigns $p : \mathfrak{f} \to [0,1]$ to the map $x \in \mathfrak{g} \mapsto \sum_{a\in\varphi^{-1}(x)} p(a)$ where φ is restricted to a map of faces $\mathfrak{f} \to \mathfrak{g}$.

It should be noted that the geometric realization functor behaves oddly with respect to (seemingly harmless) categorical products of simplicial complexes. Indeed, the product $\Delta^1 \times \Delta^1$ of standard 1-simplices has four vertices and its faces are in one-to-one correspondence with faces in Δ^3; however the topological product $[0,1]^2$ of two intervals is only two-dimensional! We have the formula $\Delta^N \times \Delta^M \simeq \Delta^{N+M+NM}$ in **SCpx** so $|\Delta^N| \times |\Delta^M| \simeq |\Delta^N \times \Delta^M|$ only when $N = 0$ or $M = 0$. Nonetheless, for general simplicial complexes K and L the universal property of products allows for a naturally-defined map:

$$|K \times L| \to |K| \times |L|$$

which is generally surjective. On the other hand, the product $K \otimes L$ in **SSet** can be shown to be well-behaved with respect to the geometric realization; that is, $|K \otimes L| \simeq |K| \times |L|$ (via the map shown above). See Proposition 1.25 of [51] for proof and discussion.

Now, some of the most highly-studied categories in this subject are categories of modules. In particular, we have the category $\mathbf{Mod}(R)$ consisting of all conceivable (left) R-modules over a fixed ring R and all R-homomorphisms $f : M \to N$ between any modules M and N. In this category, we re-encounter many of the constructions from section 3.1.2. There is always a zero R-module $\mathbf{0}$ (similar to the zero vector space), which is the zero object of $\mathbf{Mod}(R)$. For an R-homomorphism $f : M \to N$, the kernel $\ker(f)$ is given by the submodule $f^{-1}(\mathbf{0}) \subseteq M$ (equipped with the appropriate inclusion) and the cokernel $\operatorname{coker}(f)$ given by the quotient module $N/f(M)$ (equipped with the appropriate quotient map). Recall that an R-homomorphism f is injective if and only if $\ker(f) \simeq \mathbf{0}$ (as objects) and surjective if and only if $\operatorname{coker}(f) \simeq \mathbf{0}$; these are respectively the monomorphisms and epimorphisms of

$\mathbf{Mod}(R)$. Moreover, every monic $\varphi : K \to M$ represents the kernel of some R-homomorphism (given by $M \to M/\varphi(K)$) and every epic $\psi : N \to C$ represents the cokernel of some R-homomorphism (given by the inclusion $\psi^{-1}(\mathbf{0}) \to N$).

The product and coproduct in $\mathbf{Mod}(R)$ of any two modules M and N exist and are isomorphic – both are given by the direct sum $M \oplus N$ of the modules. The underlying set structure of the direct sum is the Cartesian product $M \times N$, whose canonical projections π_M and π_N are the same ones for the product in $\mathbf{Mod}(R)$; on the other hand, the canonical inclusions of the coproduct as given by $\iota_M : m \mapsto m \oplus 0$ and $\iota_N : n \mapsto 0 \oplus n$. It is also possible to define the span of submodules $N_1, N_2 \subseteq M$ to be the subspace $N_1 + N_2 = \{a + b : a \in N_1, b \in N_2\}$ of M; note that $N_1 \oplus N_2 \simeq N_1 + N_2$ if and only if $N_1 \cap N_2 = \mathbf{0}$. As a special case, consider these constructions when $R = \mathbb{F}$, a field; this is just the category $\mathbf{Vect}(\mathbb{F})$ of vector spaces over $\mathbb{F}$.

It is also understood that all finitely-generated (left) R-modules form a full subcategory $\mathbf{mod}(R)$ of $\mathbf{Mod}(R)$, just like $\mathbf{vect}(\mathbb{F})$ is a full of $\mathbf{Vect}(\mathbb{F})$. This subcategory enjoys all the properties mentioned above, and is a better setting for the examples to be discussed further.

The category $\mathbf{Rep}(Q)$ is closely knit to some categories of left modules (at least as far as persistent homology goes). One may recall from section 2.2.3 that a *quiver* Q possesses the structure of a directed graph, specified by some data (V, E, h, t) where V and E are its sets of vertices and edges and $h, t : E \to V$ assign the head vertex and tail vertex to every edge; a quiver representation X assigns a vector space $X(i)$ for every $i \in V$ and a linear map $X(e) : X(te) \to X(he)$ for every $e \in E$, with the resulting diagram of vector spaces commuting. Morphisms of two representations are natural transformations from one such diagram to another. The subcategory $\mathbf{rep}(Q)$ contains finite-dimensional quiver representations. See [53] for additional details.

Example 3.7. Given a quiver Q, define an abelian group consisting of k-linear combinations of *paths* in Q, which are finite associative strings with characters in E modulo the rules: $0e = 0 = e0$, and $ee' = 0$ unless $te = he'$; including the trivial paths a_i for all $i \in V$ (with $ta_i = i = ha_i$) in the k-linear span and letting associative multiplication of strings be given by their concatenation, a ring $R = kQ$ is obtained called the *path algebra*.

- Functor $\mathbf{Rep}(Q) \to \mathbf{Mod}(kQ)$ assigns quiver representations X of Q to the module M spanned by the vectors of all the $X(i)$, $i \in V$; the ring kQ acts on M by consecutively applying linear transformations $X(e_k)$ taken from non-vanishing stings $e_n \cdots e_1$ when read right-to-left. Assign morphisms accordingly.

- Functor $\mathbf{Mod}(kQ) \to \mathbf{Rep}(Q)$ assigns left kQ-modules M to Q-representations X given by data $X(i) = a_i \cdot M$ for each vertex i and $X(e) : X(te) \to X(he)$ given by $x \mapsto e \cdot x$ for every edge e. Assign morphisms accordingly.

The two functors above yield an equivalence of categories $\mathbf{Mod}(kQ)$ and $\mathbf{Rep}(Q)$; moreover, the two restrict to equivalences of subcategories $\mathbf{rep}(Q)$ and $\mathbf{mod}(kQ)$.

We proceed to formalize some of these notions present in categories of modules (and their equivalent variants) further in the next section.

3.2 Abelian Categories

Quite a few of the examples that were discussed in section 3.1.3 had a common set of desirable properties that prove to be useful in describing and computing persistent homology. The purpose of this section is to give a quick introduction to such categories and some other of their known properties. For general reference, [27] is considered a classical text on the subject, while [59] offers a more modern take.

Historically, a precursor concept for developing this was that of an enriched category. The idea is to take a locally small category $\mathcal{C}$, and reconsider each Hom-set as an object in some other category $\mathcal{S}$. This can be done to further develop the language used to describe morphisms in a category $\mathcal{C}$, and allow one to do more operations with diagrams in $\mathcal{C}$ and establish its inherent properties perhaps more naturally.

While this can be a useful accompaniment, these overt constructions will be avoided in favor of clearer exposition. An interested reader may further consult chapter 3 of [50] for the exact details of categorical enrichment, and also [23] and chapter 2 of [25] for a more focused discussion on monoidal categories.

3.2.1 Basics and important properties

Suppose that $\mathcal{C}$ is a category enriched over $\mathbf{vect}(\mathbb{F})$ in the following sense:

1. For any pair of its objects X and Y, the set $\mathrm{Hom}(X, Y)$ of morphisms $X \to Y$ is a finite-dimensional $\mathbb{F}$-vector space.

2. The composition map $\circ : \mathrm{Hom}(Y, Z) \times \mathrm{Hom}(X, Y) \to \mathrm{Hom}(X, Z)$ is $\mathbb{F}$-bilinear; that is, for any constants $\alpha, \beta \in \mathbb{F}$, we have

$$(\alpha \cdot f_1 + \beta \cdot f_2) \circ g = \alpha \cdot f_1 \circ g + \beta \cdot f_2 \circ g$$

$$f \circ (\alpha \cdot g_1 + \beta \cdot g_2) = \alpha \cdot f \circ g_1 + \beta \cdot f \circ g_2$$

We then call $\mathcal{C}$ an $\mathbb{F}$-*linear (additive) category* if $\mathcal{C}$ contains a zero object $\mathbf{0}$ and for any finite collection $X_1, \ldots, X_n$ of objects their product and coproduct exist and are isomorphic in $\mathcal{C}$: this *biproduct* of $X_1, \ldots, X_n$ will be denoted $X_1 \oplus \cdots \oplus X_n$. Some authors use $\mathbf{Vect}(\mathbb{F})$ to enrich $\mathcal{C}$ in the above definition; we won't use that here.

More generally, $\mathcal{C}$ is called an *additive category* if condition (1) is replaced to say that $\mathrm{Hom}(X, Y)$ is $\mathbb{Z}$-module and (2) only requires the equations to hold for $\alpha, \beta \in \mathbb{Z}$.

By definition, the biproduct (or *direct sum*) $\bigoplus_i X_i = X_1 \oplus \cdots \oplus X_n$ of a collection of objects $X_1, \ldots, X_n$ has both canonical inclusions $\iota_k : X_k \to \bigoplus_i X_i$ and canonical projections $\pi_k : \bigoplus_i X_i \to X_k$. In a category that has all zero morphisms (such as when they are given by a zero object), it is easy to show from the universal properties of product and coproduct that the following relations hold:

$$\pi_j \circ \iota_i = \begin{cases} 1 & , \text{if } i = j \\ 0 & , \text{otherwise} \end{cases} \tag{3.1}$$

Moreover, the above relations can be used to show that: $\iota_1 \circ \pi_1 + \cdots + \iota_n \circ \pi_n = 1$, the identity map on the biproduct.

Morphisms $f : \bigoplus_{i=1}^n X_i \to \bigoplus_{j=1}^m Y_j$ induce morphisms $f_{ij} = \pi_j^Y \circ f \circ \iota_i^X$ of $X_i \to Y_j$; this allows us to write f as an $m \times n$ block matrix whose block component in position (i, j) is given by the morphism f_{ij}. Of course, if we have a collection of morphisms $g_1 : X_1 \to Y_1, \ldots, g_n : X_n \to Y_n$ then this induces a morphism $g : \bigoplus_{i=1}^n X_i \to \bigoplus_{j=1}^n Y_j$ that can be represented by an $n \times n$ block diagonal matrix whose component in position (k, k) is simply g_k; we write $g = g_1 \oplus \cdots \oplus g_n$. Also, note that:

$$(f_1 + f_2) \oplus (g_1 + g_2) = f_1 \oplus g_1 + f_2 \oplus g_1 + f_1 \oplus g_2 + f_2 \oplus g_2$$

where $X_1 \overset{f_1, f_2}{\rightrightarrows} X_2$ and $Y_1 \overset{g_1, g_2}{\rightrightarrows} Y_2$. In an $\mathbb{F}$-linear category, it also holds

$$(\lambda \cdot f) \oplus g = f \oplus (\lambda \cdot g) = \lambda \cdot f \oplus g \quad , \forall \lambda \in \mathbb{F}$$

The existence of such as a well-behaved construct for objects X and Y (particularly that $X \prod Y \simeq X \coprod Y$) is not too idiosyncratic. If $\mathcal{C}$ has all zero morphisms, the universal property of the coproduct (if it exists) would produce morphisms $\pi_X : X \coprod Y \to X$ and $\pi_Y : X \coprod Y \to Y$, which uniquely correspond to morphisms $i : X \to X$, $i : Y \to Y$, $0 : X \to Y$, $0 : Y \to X$. Then immediately, the universal property of the product (if it exists) will produce a unique morphism

$$X \coprod Y \to X \prod Y$$

such that each ι_k is uniquely determined by π_k. That the biproduct exists just means that this is an isomorphism.

That $\mathcal{C}$ is enriched over abelian groups (or vector spaces) can also be used to show the existence of biproducts. Assuming that objects X and Y have a product $X \prod Y$ in $\mathcal{C}$, we can obtain the coproduct by taking morphisms $(1 \oplus 0) : X \oplus \mathbf{0} \simeq X \to X \oplus Y$ and $(0 \oplus 1) : \mathbf{0} \oplus Y \simeq Y \to X \oplus Y$ to be the canonical inclusions, where 1 is the identity morphism; likewise, the product can be obtained from the coproduct $X \coprod Y$ in a similar fashion. Necessarily, these morphisms would satisfy the equations 3.1, and the allow us to take a morphism $f \circ (X \prod Y \to X) + g \circ (X \prod Y \to Y)$ to represent $X \prod Y \to Z$ for any diagram $X \xrightarrow{f} Z \xleftarrow{g} Y$. That is, the existence of a product or coproduct in an abelian-enriched category guarantees the corresponding biproduct.

Conversely, given a category $\mathcal{C}$ where biproducts $X \oplus Y$ exist for all objects X and Y, we can actually specify a binary operation on $\mathrm{Hom}_{\mathcal{C}}(X, Y)$ so that $\mathcal{C}$ is an enriched category! Note that the universal properties of the product and coproduct (respectively) produce unique morphisms $\Delta_X : X \to X \oplus X$ and $\check{\Delta}_Y : Y \oplus Y \to Y$ such that $(X \oplus X \to X) \circ \Delta_X = 1_X$ and $\check{\Delta}_Y \circ (Y \to Y \oplus Y) = 1_Y$; Δ_Z and $\check{\Delta}_Z$ are called the *diagonal* and *codiagonal* morphisms of Z. Then for any $f, g : X \to Y$ we can define $f+g := \check{\Delta}_Y \circ (f \oplus g) \circ \Delta_X$. This operation is associative, commutative, and with respect to which composition of morphisms is bilinear. However, it doesn't quite make each Hom-set into a $\mathbb{Z}$-module (or an $\mathbb{F}$-vector space), since it isn't certain that the binary operation $+$ is invertible (or that composition commutes with scalar multiplication). Hence, categorical enrichment is not necessarily a redundant assumption.

The significance of having biproducts in a category was also underscored in section 2.2.3, where objects could be written (in some sense, uniquely) as direct sums of indecomposable objects which are classifiable. We shall borrow this terminology for the categorical setting. In an additive (or $\mathbb{F}$-linear) category, an object X shall be called *decomposable* if it is isomorphic to a direct sum $Y \oplus Z$ of two nonzero objects Y and Z; an *indecomposable* (object) is then a nonzero object that is not decomposable. We would like to be able to express any object X as a direct sum of some collection of indecomposables – provided that such a decomposition does exist. In addition, we also need to be able to classify all indecomposables as well as know the extent to which such direct sum decompositions are unique.

The "uniqueness" of the decomposition in Theorem 2.10 from section 2.2 is a special case of the following general result for Krull-Schmidt categories:

Theorem 3.8 (see Theorem 4.2 in [36]). *Let X be an object in an additive category, and suppose there are two finite decompositions*

$$X_1 \oplus \cdots \oplus X_m = X = Y_1 \oplus \cdots \oplus Y_n$$

into (nonzero) objects each having local endomorphism rings. Then $m = n$, and there exists a permutation π of all indices such that $X_i' \simeq X_{\pi(i)}''$.

Recall from algebra that a *local ring* is a ring R that has a unique maximal ideal. For an object X in an additive category $\mathcal{C}$, the *endomorphism* ring is the module $\mathrm{Hom}_{\mathcal{C}}(X,X)$ with multiplication given by function composition. It can be shown that $\mathrm{Hom}_{\mathcal{C}}(X,X)$ is local if and only if: for every $f : X \to X$, either f is an iso or $1_X - f$ is an iso. Because of this theorem, we will refer to an additive (or $\mathbb{F}$-linear) category $\mathcal{C}$ as a *Krull-Schmidt* category if every object X in $\mathcal{C}$ satisfies the hypotheses of Theorem 3.8.

An additive (or $\mathbb{F}$-linear) category $\mathcal{C}$ is called a *pre-abelian category* if every morphism $f : X \to Y$ admits a kernel $K \to X$ and a cokernel $Y \to C$ in $\mathcal{C}$. We have an immediate result:

Proposition 3.9. *Let $f : X \to Y$ be a morphism in a pre-abelian category.*

1. *f is a monomorphism if and only if* $\ker(f) = (\mathbf{0} \to X)$*.*

2. *f is an epimorphism if and only if* $\mathrm{coker}(f) = (Y \to \mathbf{0})$*.*

Note that in a pre-abelian category, all equalizers and coequalizers exist: the equalizer of two morphisms $f, g : X \to Y$ equals $\ker(f - g)$ and their coequalizer equals $\mathrm{coker}(f - g)$. This is important, due to the following profound theorem:

Theorem 3.10 (see Exercise C, Chapter 3 in [27])**.** *A category $\mathcal{C}$ has a limit of every finite diagram $\mathcal{D}$ if and only if for any objects A and B it possesses their product $A \prod B$ and all $f, g \in \mathrm{Hom}_{\mathcal{C}}(A, B)$ have an equalizer. Dually, $\mathcal{C}$ has a colimit of every finite diagram $\mathcal{D}$ if and only if it for any objects A and B it possesses their coproduct $A \coprod B$ and all $f, g \in \mathrm{Hom}_{\mathcal{C}}(A, B)$ have a coequalizer.*

This should serve the argument that pre-abelian categories $\mathcal{C}$ admit a lot of structure; any conceivable finite diagram in $\mathcal{C}$ will have a limit and colimit (which by definition are unique, up to isomorphism). Before we get into discussing their relationship, first note that the operations $\ker(\bullet)$ and $\operatorname{coker}(\bullet)$ behave "orderly". Clearly, for any $f : X \to Y$, the kernel of $\ker(f)$ and the cokernel of $\operatorname{coker}(f)$ are just zero moprhisms from/to $\mathbf{0}$. The next result extends this further – note that equality $(=)$ may mean "equality up to isomorphism of objects".

Proposition 3.11. *Let $f : X \to Y$ be a morphism in a pre-abelian category $\mathcal{C}$.*

1. *The kernel $p : K \to X$ of f satisfies $p = \ker(f) = \ker(\operatorname{coker}(p))$.*

2. *The cokernel $q : Y \to C$ of f satisfies $q = \operatorname{coker}(f) = \operatorname{coker}(\ker(q))$.*

Generally, we have that $f : X \to Y$ is a kernel of some morphism in $\mathcal{C}$ if and only if $f = \ker(\operatorname{coker}(f))$; dually, $f : X \to Y$ is a cokernel of some morphism in $\mathcal{C}$ if and only if $f = \operatorname{coker}(\ker(f))$.

For a morphism $f : X \to Y$, define its *image* $\operatorname{im}(f) = \ker(\operatorname{coker}(f))$ and its *coimage* $\operatorname{coim}(f) = \operatorname{coker}(\ker(f))$. These constructs are aptly named; for any morphism $f : M \to N$ in $\mathbf{Mod}(R)$, the image $f(M) \subseteq N$ is clearly the kernel of the quotient map $N \to N/f(M)$, while the coimage manifests as $M \to M/f^{-1}(\mathbf{0})$.

We can then use the universal property of the cokernel (as in $\operatorname{coim}(f)$) and the kernel (as in $\operatorname{im}(f)$) to obtain a unique *natural map* $\tilde{f} : \operatorname{coim}(f) \to \operatorname{im}(f)$, such that the morphism $f : X \to Y$ factors as:

$$X \longrightarrow \operatorname{coim}(f) \xrightarrow{\tilde{f}} \operatorname{im}(f) \longrightarrow Y \tag{3.2}$$

In fact, this $\tilde{f}$ is a natural isomorphism in some categories, making the underlying (universal) objects of the coimage and image indistinguishable. Think to $\mathbf{vect}(\mathbb{F})$, where for any linear map $f : V \to W$ the range $f(V) \subseteq W$ is isomorphic to the quotient $V/f^{-1}(\mathbf{0})$ by the rank-nullity theorem.

Finally, we call $\mathcal{C}$ an *abelian category* if $\mathcal{C}$ is pre-abelian and for every morphism $f : X \to Y$ the natural map $\tilde{f} : \mathrm{coim}(f) \to \mathrm{im}(f)$ in the factorization described by diagram 3.2 is an isomorphism. From this it immediately follows that every morphism $f : X \to Y$ in an abelian category has a factorization $f = m \circ e$, where $m : X \to I$ is a monic map and $e : I \to Y$ is an epic map; this factorization is unique up to isomorphism of object I.

The fact that the image and coimage are thusly related in a category has a number of other useful properties.

Lemma 3.12. *Let $f : X \to Y$ be a morphism in an abelian category $\mathcal{C}$.*

1. *f is a monomorphism if and only if it is the kernel of some $Y \to Z$.*

2. *f is an epimorphism if and only if it is the cokernel of some $Z \to X$.*

3. *f is an isomorphism if and only if it is both monic and epic.*

Moreover, a pre-abelian category is abelian if both criteria 1 and 2 hold in it (also known as normality conditions).

Our primary interest will be in $\mathbb{F}$-linear abelian categories, for which there is another result essentially guaranteeing all objects to have a unique direct sum decomposition. Generally, the well-known Atiyah's criterion (see Theorem 5.5 in [36]) provides a very useful criterion for verifying the Krull-Schmidt property of an abelian category. Separately, the Hom-sets in this type of category satisfy a certain *bi-chain condition*, which when combined Atiyah's criterion produces the following result:

Theorem 3.13 (Atiyah)**.** *A linear abelian category $\mathcal{C}$ is Krull-Schmidt; that is:*

1. *Every object in $\mathcal{C}$ admits a finite decomposition as a sum of indecomposables.*

2. *Every indecomposable has a local endomorphism ring.*

The proof (see [2]) of Atiyah's criterion however is nonconstructive. It neither provides an algorithm to decompose a given object as a direct sum of indecomposables, nor a classification of indecomposables.

3.2.2 Short Exact Sequences in Abelian Categories

In abelian categories, many notions are drawn by means of special important diagrams $\mathcal{D}$ called *chain complexes*; this construction requires that if $f : A \to B$ and $g : B \to C$ are morphisms in $\mathcal{D}$ then $g \circ f = 0$. Such diagrams arise in many practical considerations.

For example, recall that (differential) chain complexes from section 2.1.2 are represented by diagrams of modules $C_\bullet$ which have the form:

$$\cdots \longleftarrow C_{j-1} \xleftarrow{\partial_j} C_j \xleftarrow{\partial_{j+1}} C_{j+1} \longleftarrow \cdots$$

The fact that $\partial_j \circ \partial_{j+1} = 0$ for all indices j implies that $\operatorname{im}(\partial_{j+1}) \subseteq \ker(\partial_j)$ on their underlying objects, from which we can calculate the homology of the complex at position j by taking the quotient of modules. In categorical terms, a pair of morphisms $A \xrightarrow{f} B \xrightarrow{g} C$ in an abelian category can be factored as:

$$\begin{tikzcd}
 & \mathrm{Coim}(f) \arrow[r, "\simeq"'] & \mathrm{Im}(f) \arrow[dr, "\mathrm{im}(f)"'] & & \\
A \arrow[ur, "\mathrm{coim}(f)"'] \arrow[rrr, "f"'] & & & B \arrow[r, "g"'] & C \\
 & & & \mathrm{Ker}(g) \arrow[u, "\mathrm{ker}(g)"'] &
\end{tikzcd}$$

If $A \xrightarrow{f} B \xrightarrow{g} C$ is a chain complex, then $g \circ \mathrm{im}(f) = 0$ (since $\mathrm{coim}(f)$ is an epic and $\mathrm{Coim}(f) \to \mathrm{Im}(f)$ is an iso) and by the universal property of the kernel $\ker(g)$ there exists a morphism $x : \mathrm{Im}(f) \to \mathrm{Ker}(g)$ such that $\mathrm{im}(f) = \ker(g) \circ x$; this morphism is necessarily monic since $\mathrm{im}(f)$ is. Hence, we may treat $\mathrm{Im}(f)$ as a *subobject* of $\mathrm{Ker}(g)$ in any abelian category under this scenario. In a sense, taking "homology" in position of object B will result in $\mathrm{coker}(x : \mathrm{Im}(f) \to \mathrm{Ker}(g))$.

Another example is obtained just by considering a single morphism $f : A \to B$. Find $\ker(f)$ and $\mathrm{coker}(f)$:

$$\mathbf{0} \longrightarrow \mathrm{Ker}(f) \xrightarrow[\ker(f)]{} A \xrightarrow{f} B \xrightarrow[\mathrm{coker}(f)]{} \mathrm{Coker}(f) \longrightarrow \mathbf{0}$$

The zero objects were added in to emphasize that this is a chain complex. In fact:

- The kernel of $\ker(f)$ is zero, i.e. the morphism $0 \to \mathrm{Ker}(f)$;

- The cokernel of $\mathrm{coker}(f)$ is also zero, given by $\mathrm{Coker}(f) \to 0$.

Furthermore, we can find $\mathrm{im}(f) = \ker(\mathrm{coker}(f))$ and $\mathrm{coim}(f) = \mathrm{coker}(\ker(f))$:

$$\begin{tikzcd}
 & & \mathbf{0} \arrow[r] & \mathrm{Im}(f) \arrow[dr, "\mathrm{im}(f)"] & & & \\
\mathbf{0} \arrow[r] & \mathrm{Ker}(f) \arrow[r] & A \arrow[rr, "f"] \arrow[dr, "\mathrm{coim}(f)"'] & & B \arrow[r] & \mathrm{Coker}(f) \arrow[r] & \mathbf{0} \\
 & & & \mathrm{Coim}(f) \arrow[r] & \mathbf{0} & &
\end{tikzcd}$$

Again, it readily follows that:

- The kernel of $\mathrm{im}(f)$ is the morphism $0 \to \mathrm{Im}(f)$;
- The cokernel of $\mathrm{coim}(f)$ is the morphism $\mathrm{Coim}(f) \to 0$.

Lastly, it follows from Proposition 3.11 that the cokernel of $\ker(f)$ is isomorphic to $\mathrm{coim}(f)$ and the cokernel $\mathrm{im}(f)$ is isomorphic to $\mathrm{coker}(f)$. So to interpret this in more familiar terms, at the position of every object in each of the two sequences (after removing f itself), the image of the incoming morphism is equal to the kernel of the outgoing morphism!

Example 3.14. Consider the morphism $x \in \mathbb{Z} \mapsto 2 \cdot x \in \mathbb{Z}$ of $\mathbb{Z}$-modules. Since this is monic, its kernel is zero and hence the image is $\mathbb{Z}/\mathbf{0} = \mathbb{Z}$.

$$\mathbf{0} \to \mathrm{Ker}(f) \to A \to \mathrm{Coim}(f) \to \mathbf{0} \quad = \quad \mathbf{0} \to \mathbf{0} \to \mathbb{Z} \to \mathbb{Z} \to \mathbf{0}$$

On the other hand, the cokernel of the morphism is given by the quotient $\mathbb{Z}/2\mathbb{Z}$ and the image is $2\mathbb{Z} \subseteq \mathbb{Z}$. The sequence $\mathbf{0} \to \mathrm{Im}(f) \to B \to \mathrm{Coker}(f) \to \mathbf{0}$ becomes:

$$\mathbf{0} \to \mathbb{Z} \simeq 2\mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \to \mathbf{0}$$

Sequences $\mathbf{0} \to \mathrm{Ker}(f) \to A \to \mathrm{Coim}(f) \to \mathbf{0}$ and $\mathbf{0} \to \mathrm{Im}(f) \to B \to \mathrm{Coker}(f) \to \mathbf{0}$ will be called the *canonical short exact sequences* associated with $f : A \to B$. In general, a *short exact sequence* is a diagram:

$$\mathbf{0} \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow \mathbf{0}$$

such that the image of every morphism is isomorphic to the kernel of the immediate morphism after. We may equivalently describe it as: f is monic, g is epic, and $\mathrm{im}(f) = \ker(g)$.

Many other short exact sequences are famously used in literature. Another example that exists in $\mathbf{Mod}(R)$ arises from the direct sum of two submodules $P, N \subset M$. Recalling from section 3.1.3 the construction of submodules $P+N$ and $P\cap N$, the universal property of $P\oplus N$ ensures maps $P\cap N \to P\oplus N$ and $P\oplus N \to P+N$ corresponding to: the inclusion $x \in P\cap N \mapsto x \in P$, the negative inclusion $x \in P\cap N \mapsto -x \in N$, and inclusions $P, N \subseteq P + N$. It is not hard to show that $P \cap N \to P \oplus N \to P + N$ is a complex (a diagram-chasing argument that uses equation 3.1), which explicit computations show the following to be a short-exact sequence:

$$\mathbf{0} \longrightarrow P \cap N \xrightarrow{(\iota_P, -\iota_N)} P \oplus N \xrightarrow{\pi_P + \pi_N} P + N \longrightarrow \mathbf{0}$$

This sequence is the foundation for the Mayer-Vietoris formula, a celebrated result in homology theory. We will also briefly revisit this sequence in section 4.1.3.

Two short exact sequences are said to be isomorphic if there is an isomorphism from one diagram to another. Given fixed objects A and C, the short exact sequences of the form $0 \to A \to B \to C \to 0$ can all be classified, up to isomorphism, in certain abelian categories – in fact, all such sequences form the set $\mathrm{Ext}^1(C, A)$ that is an abelian group with respect to the Baer sum. The group is tedious to compute, requiring the existence of projective resolutions of objects in an abelian category; see chapter 3 of [30] for more details.

For every two objects A and C in an abelian category, there is a distinguished short exact sequence $\mathbf{0} \to A \to B \to C \to \mathbf{0}$ that *splits*; here, the sequence is isomorphic to one where $B = A \oplus C$ with $A \to B$ being a canonical inclusion and $B \to C$ being a canonical projection. If $\mathrm{Ext}^1(C, A)$ can be defined, a split short exact sequence corresponds to the identity element of this group.

The second sequence in example 3.14 splits, while the first one doesn't. Generally, the following lemma proves useful for characterizing split short exact sequences.

Lemma 3.15 (Splitting Lemma). *Suppose you are given a short exact sequence below.*

$$\mathbf{0} \longrightarrow A \xrightarrow{\quad f \quad} B \xrightarrow{\quad g \quad} C \longrightarrow \mathbf{0}$$

Then the following are equivalent:

1. *the sequence splits, so $B \simeq A \oplus C$;*
2. *there is a morphism $h : B \to A$ such that $h \circ f = 1_A$ is the identity morphism;*
3. *there is a morphism $k : C \to B$ such that $g \circ k = 1_C$ is the identity morphism.*

3.2.3 Relevant Examples, revisited

The category **Set** possesses product and coproducts, which are usually different, for any pair of objects A and B. The canonical projections $A \overset{\pi_A}{\leftarrow} A \times B \overset{\pi_B}{\rightarrow} B$ are surjective maps (epics), while the canonical inclusions $A \overset{\iota_A}{\rightarrow} A \sqcup B \overset{\iota_B}{\leftarrow} B$ are injections (monics). There is an initial object ($\varnothing$) and terminal objects (sets of cardinality 1), which are different; there is no notion of zero morphisms in **Set**, although one could try to define *"right" zero morphisms* $\zeta_A : \varnothing \to A$ that satisfy certain commutative relations, namely that $(A \to B) \circ \zeta_A = \zeta_B$ for all objects A and B. Nonetheless, kernels and cokernels cannot be defined for general $f : A \to B$. So **Set** does not satisfy most axioms of an additive category.

Category **Rel** (of sets as objects and set relations as morphisms) was shown to possess a zero object ($\varnothing$), and for any two objects A and B their disjoint union $A \sqcup B$ with canonical maps $A \overset{\pi_A}{\leftarrow} A \sqcup B \overset{\pi_B}{\rightarrow} B$ and $A \overset{\iota_A}{\rightarrow} A \sqcup B \overset{\iota_B}{\leftarrow} B$ (where the $\pi_\bullet$ are epic and the $\iota_\bullet$ are monic) forms a biproduct of A and B. Every morphism $f : A \to B$ also has a kernel $p : K \to A$ and a cokernel $q : B \to C$, where p and q induce bijective functions after restriction to their domain and range; since monomorphisms and epimorphisms are generally not bijections (after restriction), the normality conditions stated in Lemma 3.12 do not hold. Furthermore, the additive operation (induced by the product and coproduct) on each set $\mathrm{Hom}(A, B)$ in **Rel** is not invertible, given by $+ : (f, h) \mapsto f \cup g \in \mathrm{Hom}(A, B)$ for any $f, g \in \mathrm{Hom}(A, B)$. In particular, **Rel** is not an additive (or linear additive) category.

The subcategory **Mch** (in **Rel**) of partial matchings inherits a zero object, all zero morphisms, and all kernels and cokernels from **Rel**. In addition, the normality conditions stated in Lemma 3.12 are satisfied: if $f : A \to B$ is a monomorphism then it is a kernel of $B \to (B \setminus \mathrm{ran}(f))$, and if $f : A \to B$ is a epimorphism, then it is a cokernel of $(A \setminus \mathrm{dom}(f)) \to A$. We also have that the underlying objects behind the image and coimage of any morphism $f : A \to B$ are given by $\mathrm{ran}(f)$ and $\mathrm{dom}(f)$, respectively, and f restricts to an isomorphism between them. However as per [5], products and coproducts may not be defined in this category. Hence, **Mch** is not Abelian (or even additive) but rather a *Puppe-exact* category.

By the same reasoning as we did for **Set**, categories **SCpx** and **SSet** are not additive. An interested reader is encouraged to study [28] for an interesting survey on the categorical properties of simplicial sets, including a simplicial set structure on their Hom-sets. Category **Top** is also not additive, although it should be said that some full subcategories of **Top** are pre-abelian; prominent examples include the category of Banach spaces and the category of compact Hausdorff abelian topological groups.

Perhaps the most well-known example of an abelian category (and its conceptual origin) is $\mathbf{Mod}(R)$ of all left R-modules. It "internalizes" its Hom-sets, meaning that every $\mathrm{Hom}(M, N)$ is an abelian group over R itself. The biproducts in $\mathbf{Mod}(R)$ are given by the direct sum, equipped with the canonical projections and inclusions that satisfy equation 3.1. Every morphism $f : M \to N$ admits a kernel and a cokernel; moreover, every monic $K \to M$ induces an isomorphism between K and a submodule of M, and every epic $M \to C$ induces an isomorphism between C and a quotient of M by some subspace (following the first fundamental theorem on homomorphisms). Therefore, monics $K \to M$ are the kernel of $M \to M/K$ and epics $M \to C$ are the cokernel of some inclusion of a subspace into M; by lemma 3.12, this is sufficient to show that $\mathbf{Mod}(R)$ to be abelian.

A special case of abelian categories, the category $\mathbf{vect}(\mathbb{F}) = \mathbf{Mod}(\mathbb{F})$ of vector spaces over field $\mathbb{F}$ has an additional property in that every short exact sequence splits. Indeed, given a short exact sequence $\mathbf{0} \to U \xrightarrow{f} W \xrightarrow{g} V \to \mathbf{0}$ of vector spaces, we have that the injectivity of f implies the range $f(U)$ is generated by the set $\beta = \{f(x) : x \in \beta_U\}$, given any basis β_U of U. Basis β can then be extended to a basis β_W of all of W (something that isn't generally possible with R-modules), allowing us to define a linear map $W \to U$ that takes $f(x) \in \beta \mapsto x \in \beta_U$ and otherwise $y \in \beta_W$ is assigned to 0. By lemma 3.15, this sequence splits; furthermore, it is basic to show that the subspace of W spanned by the basis $\beta_W \setminus \beta$ is the coimage of the surjective map $W \to V$.

The reasoning is similar that $\mathbf{mod}(R)$ is also abelian. It is notable that for any objects M and N in this category, any $f : M \to N$ may be represented by a matrix with coefficients in R, provided that an ordered set of generators is picked for each object. Likewise, $\mathbf{vect}(\mathbb{F}) = \mathbf{mod}(\mathbb{F})$ is an exact category, with every linear map $f : V \to W$ having a matrix representation with coefficients in $\mathbb{F}$ – provided that some bases of V and W have been pre-selected.

Peculiarly, the Freyd-Mitchell embedding theorem (first mentioned in section 2.2) suggests that all abelian categories can be thought of as categories of modules (over some ring). For proof and discussion, see Theorem 7.34 in [27].

Theorem 3.16 (Freyd-Mitchell)**.** *If $\mathcal{A}$ is a small abelian category, then there is a ring R and an exact, full, faithful functor $\mathcal{F} : \mathcal{A} \to \mathbf{Mod}(R)$. In particular, $\mathcal{A}$ is equivalent to a full subcategory of $\mathbf{Mod}(R)$.*

Our main focus now will be to discuss the category $\mathbf{vect}^T(\mathbb{F})$ of persistent vector spaces indexed by T, and some of its subcategories. This will be done in section 3.3.1.

3.3 Pertinent Results

We focus here on the category of persistence vector spaces, which were introduced in section 2.2, as this will be the setting of explicit homological calculations in the next chapter. The terminology and properties of the objects discussed have been introduced in [43], which the reader is implored to review for some of the more explicitly constructive proofs.

Recall that a *persistence vector space* ${}_{\bullet}V$ over a totally-ordered set T is a collection $\{{}_tV : t \in T\}$ of finite-dimensional vector spaces together with linear (transition) maps ${}_sV \to {}_tV$ such that $s \leqslant \forall r \leqslant t$ we have ${}_sV \to {}_tV = ({}_rV \to {}_tV) \circ ({}_sV \to {}_rV)$. One can equivalently restate this in categorical terms: treating T as a category whose objects are elements of T and whose relation $\leqslant$ determines what morphisms exist between them, a persistence vector space ${}_{\bullet}V$ is simply a functor $V : T \to \mathbf{vect}(\mathbb{F})$. A morphism ${}_{\bullet}f : {}_{\bullet}V \to {}_{\bullet}W$ is a collection of linear maps $\{{}_tf : {}_tV \to {}_tW : t \in T\}$ that commute with the transition maps of ${}_tV$ and ${}_tW$, as seen in the diagram below.

$$\begin{tikzcd}
\cdots \arrow[r] & {}_{t-1}W \arrow[r] & {}_tW \arrow[r] & {}_{t+1}W \arrow[r] & \cdots \\
\cdots \arrow[r] & {}_{t-1}V \arrow[r] \arrow[u, "{}_{t-1}f"] & {}_tV \arrow[r] \arrow[u, "{}_tf"] & {}_{t+1}V \arrow[r] \arrow[u, "{}_{t+1}f"] & \cdots .
\end{tikzcd}$$

Equivalently, a morphism ${}_{\bullet}f : {}_{\bullet}V \to {}_{\bullet}W$ is a natural transformation from functor $V : T \to \mathbf{vect}(\mathbb{F})$ to functor $W : T \to \mathbf{vect}(\mathbb{F})$. The objects ${}_{\bullet}V$ together with morphisms ${}_{\bullet}f$ constitute the category $\mathbf{vect}^T(\mathbb{F})$.

Note 3.17. We shall restrict to the case where $T = (\mathbb{Z}, \leqslant)$ or $T = \{1 \leqslant \cdots \leqslant n\}$.

The zero persistence vector space is the sequence ${}_{\bullet}\mathbf{0}$, with ${}_t\mathbf{0} = \mathbf{0}$ for all $t \in T$. For any two persistence vector spaces ${}_{\bullet}V$ and ${}_{\bullet}W$, their direct sum is given by ${}_{\bullet}V \oplus {}_{\bullet}W$, where the object at level t is just ${}_tV \oplus {}_tW$ and for $s \leqslant t$ the transition map is simply $({}_sV \to {}_tV) \oplus ({}_sW \to {}_tW)$. Recall that there are morphisms ${}_\bullet V \oplus {}_\bullet W \to {}_\bullet V$ and ${}_\bullet V \to {}_\bullet V \oplus {}_\bullet W$, given by the level-wise projection and injection to the first coordinate respectively; similar maps exist for ${}_\bullet W$.

Following Proposition 2.12, to every morphism $f : {}_\bullet V \to {}_\bullet W$ there are associated persistence vector spaces $\ker({}_\bullet f)$ and $\operatorname{coker}({}_\bullet f)$; similarly, there are persistence vector spaces $\operatorname{im}({}_\bullet f)$ and $\operatorname{im}({}_\bullet f)$. Note that $\ker({}_\bullet f)$ and $\operatorname{im}({}_\bullet f)$ are subobjects of respective ${}_\bullet V$ and ${}_\bullet W$, while $\operatorname{coker}({}_\bullet f)$ and $\operatorname{coim}({}_\bullet f)$ are quotients of respective ${}_\bullet W$ and ${}_\bullet V$.

We distinguish [43] *tempered persistence vector spaces* ${}_\bullet V$, where all but finitely-many of the structure maps ${}_sV \to {}_tV$ are linear isomorphisms. These objects form a full subcategory $\mathbf{vect}^T_\star(\mathbb{F})$ in $\mathbf{vect}^T(\mathbb{F})$. One reason that this subcategory is important is because of Corollary 2.14, which guarantees that every object can be written as a direct sum of a finite number of indecomposable objects, each of which has a simple form. Every indecomposable object can be described by an *interval* I, which is a subset of T such that $t \in I$ whenever $i \leqslant t < j$ for some $i, j \in I$. To be concrete, an interval $I \subseteq T$ specifies an *interval persistence vector space* (or *interval complex*) ${}_\bullet I$, defined by

$$
{}_tI = \begin{cases} \mathbb{F} & \text{, if } t \in I \\ \mathbf{0} & \text{, otherwise} \end{cases}
$$

where every arrow $\mathbb{F} \to \mathbb{F}$ is the identity morphism $1_{\mathbb{F}}$. Ordinary interval notation is used to denote intervals I; for example, the interval persistence vector space ${}_\bullet[1, 4)$ is the diagram of vector spaces

$$
\begin{array}{ccccccccccc}
\cdots \longrightarrow & \mathbf{0} & \longrightarrow & \mathbb{F} & \xrightarrow{1} & \mathbb{F} & \xrightarrow{1} & \mathbb{F} & \longrightarrow & \mathbf{0} & \longrightarrow \cdots \\
& t=0 & & t=1 & & t=2 & & t=3 & & t=4 &
\end{array}
$$

Another reason for the importance of category $\mathbf{vect}^T_\star(\mathbb{F})$ is that it includes several types of vector space diagrams as objects that are important to persistent homology. A *filtered persistence space* ${}_\bullet F$ is a persistence vector space such that ${}_sF \to {}_tF$ is given by subspace inclusion ${}_sF \subseteq {}_tF$ for every $s \leqslant t$; that is,

$$\cdots \hookrightarrow {}_{j-1}F \hookrightarrow {}_jF \hookrightarrow {}_{j+1}F \hookrightarrow \cdots$$

Provided that for a large enough level $m^+ \in T$ the sequence "stabilizes" in the sense of ${}_tF \simeq {}_pF$ if $t \geqslant m^+$, then we have that ${}_\bullet F$ is tempered. In our usage, filtered objects will also be required to satisfy ${}_sF = \mathbf{0}$ for any $s \leqslant m^-$, for some $m^- \in T$.

Another example, *chain complexes* are diagrams of form $C_\bullet = (C_k, \partial_k)$, where

$$\cdots \longleftarrow C_{j-1} \xleftarrow{\partial_j} C_j \xleftarrow{\partial_{j+1}} C_{j+1} \longleftarrow \cdots$$

and $\partial_j \circ \partial_{j+1} = 0$ for all j. Recalling that $\mathbf{vect}(\mathbb{F})$ is equivalent to its dual (see example 3.4), we can use the opposite functor $\mathbf{vect}(\mathbb{F}) \to \mathbf{vect}^{op}(\mathbb{F})$ on this diagram to obtain a persistence vector space $(C^*_\bullet, \partial^*_\bullet)$ with the condition $\partial^*_{j+1} \circ \partial^*_j = 0$ for all j; usually, $(C^*_\bullet, \partial^*_\bullet)$ is called a cochain complex. Provided that there exist $m^-, m^+ \in T$ such that $C^*_s \simeq C^*_{m^-}$ and $C^*_t \simeq C^*_{m^+}$ for all $s \leqslant m^- \leqslant m^+ \leqslant t$, then we have that this persistence vector space is tempered. One may note from section 2.1 that typically $C_{m^-} \simeq \mathbf{0}$ in classical computations of homology.

Note that since tempered persistence vector spaces can be represented by finite diagrams in $\mathbf{vect}(\mathbb{F})$, which is abelian, then Theorem 3.10 shows that tempered objects have limits and colimits. These are actually pretty simple to describe: the *limit* of ${}_\bullet V$ is represented by ${}_{m^-}V$ for a sufficiently small $m^- \in T$, while its *colimit* is represented by ${}_{m^+}V$ for a sufficiently large $m^+ \in T$. Strictly speaking (going by the discussion in section 3.1.2), these are the underlying objects of the categorical limit and colimit of the diagram ${}_\bullet V$ in $\mathbf{vect}(\mathbb{F})$.

3.3.1 Categories of Persistence Vector Spaces

Let ${}_\bullet V$ and ${}_\bullet W$ be objects in $\mathbf{vect}^T(\mathbb{F})$. Given morphisms ${}_\bullet f, {}_\bullet g : {}_\bullet V \to {}_\bullet W$ and scalars $\alpha, \beta \in \mathbb{F}$, it is elementary to verify that $\alpha \cdot {}_\bullet f + \beta \cdot {}_\bullet g$ is a morphism of ${}_\bullet V \to {}_\bullet W$ and that composition in $\mathbf{vect}^T(\mathbb{F})$ is bilinear (since composition is a bilinear operation in the category of vector spaces). Hence, $\mathbf{vect}^T(\mathbb{F})$ is an enriched category over $\mathbf{Vect}(\mathbb{F})$; it is generally not true that $\mathrm{Hom}({}_\bullet V, {}_\bullet W)$ is finite-dimensional. It possesses a zero object represented by ${}_\bullet \mathbf{0}$.

For any ${}_\bullet V, {}_\bullet W$, the category also possess a biproduct given by ${}_\bullet V \oplus {}_\bullet W$ and every morphism ${}_\bullet f : {}_\bullet V \to {}_\bullet W$ has a kernel $\ker({}_\bullet f) \to {}_\bullet V$ and a cokernel ${}_\bullet W \to \mathrm{coker}({}_\bullet f)$. To see this, note that for these objects to satisfy the required universal properties it is necessary (and sufficient) that the universal properties hold true for every level $t \in T$, allowing us to derive these properties from $\mathbf{vect}(\mathbb{F})$. Lastly, the objects $\mathrm{coim}({}_\bullet f)$ and $\mathrm{im}({}_\bullet f)$ are isomorphic, since the canonical morphism ${}_\bullet \check{f} : \mathrm{coim}({}_\bullet f) \to \mathrm{im}({}_\bullet f)$ restricts to an isomorphism at every $t \in T$. This proves the following:

Proposition 3.18. *The category* $\mathbf{vect}^T(\mathbb{F})$ *is abelian.*

By Theorem 3.13, objects in $\mathbf{vect}^T(\mathbb{F})$ have a finite direct sum decomposition in terms of objects that have local endomorphism rings; however, these indecomposable objects are hard to classify.

In $\mathbf{vect}^T(\mathbb{F})$, we can look at the full subcategory that contains only tempered persistence objects. It inherits a zero object from $\mathbf{vect}^T(\mathbb{F})$ and is closed with respect to finite direct sums. It is also closed with respect to the taking of kernels and cokernels; to see this, note that the equation ${}_t f \circ ({}_s V \to {}_t V) = ({}_s W \to {}_t W) \circ {}_s f$ for morphisms implies that ${}_s f$ and ${}_t f$ uniquely determine each other when the structure maps of ${}_\bullet V$ and ${}_\bullet W$ are isomorphisms between levels $s \leqslant t$. We then have the following:

Proposition 3.19. *The category* $\mathbf{vect}_\star^T(\mathbb{F})$ *of tempered persistence vector spaces is linear abelian. This category is Krull-Schmidt: every object has a unique (up to isomorphism) direct sum decomposition in terms of a finite collection of interval complexes.*

Of course, the last statement is a direct consequence of Theorem 3.13 and Corollary 2.14. Recall that the set of intervals determining the interval decomposition of ${}_\bullet V$ is called its *barcode* and is denoted $\mathcal{B}_V$.

Filtered persistence spaces ${}_\bullet F$ whose colimit is a finite-dimensional vector space ${}_pF = F$ (for some $p \in T$) comprise a strictly full subcategory $\mathbf{filt}(T, \mathbb{F}) = \mathbf{filt}(T)$ of the tempered persistence objects; note that the colimit of ${}_\bullet F$ is $\mathbf{0}$. Note that any summand of a filtered object is necessarily isomorphic to a filtered object. Combining these facts with Proposition 3.19 yields:

Lemma 3.20. *The category* $\mathbf{filt}(T, \mathbb{F})$ *of (tempered) filtered objects is Krull-Schmidt. A filtered object* ${}_\bullet F$ *is indecomposable if and only if its colimit* F *is indecomposable object in* $\mathbf{vect}(\mathbb{F})$.

This is important because the category $\mathbf{filt}(T)$ is not abelian; in particular, the kernel/cokernel of morphisms in $\mathbf{vect}_\star^T(\mathbb{F})$ may not be filtered objects. This forbids us from simply using Theorem 3.13 to claim the second statement of Lemma 3.20, despite it being true.

One needs to remember that we are tasked with computing the homology of these objects, which requires some differential structure inherent to these objects. It thus becomes necessary to look at a slight generalization of these results. First note that chain complexes $C_\bullet = (C_k, \partial_k)$ comprise a category that is dually equivalent to the full subcategory of cochain complexes in $\mathbf{vect}_\star^T(\mathbb{F})$, which is abelian [59]. Following the simple observation that a category is abelian if and only if its opposite category is abelian, we arrive at the following result.

Proposition 3.21. *The category of tempered chain complexes is linear abelian, and therefore Krull-Schmidt. A complex is indecomposable if and only if it is isomorphic to an interval complex.*

Now, a *filtered chain complex* ${}_\bullet C_\bullet$ is a diagram of tempered chain complexes:

$$\cdots \longrightarrow ({}_{t-1}C_\bullet, {}_{t-1}\partial_\bullet) \longrightarrow ({}_tC_\bullet, {}_t\partial_\bullet) \longrightarrow ({}_{t+1}C_\bullet, {}_{t+1}\partial_\bullet) \longrightarrow \cdots$$

that is indexed over T and every arrow is monic; in particular, at every position (*degree*) $k \in \mathbb{Z}$:

- the range of the filtration on C_{k+1} under the map $\partial_{k+1} : C_{k+1} \to C_k$ is a is a subfiltration of the filtration on C_k, and
- the diagram restricts to a filtered persistence space ${}_\bullet C_k$.

Every filtered chain complex ${}_\bullet C_\bullet$ has a colimit – a chain complex, denoted $C_\bullet$, where the vector space having degree $k \in \mathbb{Z}$ is the colimit of that particular filtration, and the differential map ∂_k is obtained from the commutative squares formed by a pair of "adjacent" filtrations in the diagram (it is well-defined since the filtrations mutually stabilize for a large enough level in T). Recalling the discourse in section 2.1, we can compute the k^{th} homology of this complex by finding the quotient of the subspace $\ker(\partial_k)$ by the range $\text{im}(\partial_{k+1})$. However, the filtration also induces a filtered basis on each vector space C_k, and as per section 2.3.2 we can compute the generators of the homology that are consistent across all levels in T (aka, the persistent features).

There are several different ways to perform operations with differential matrices that take into account organization both with respect to the $\mathbb{Z}$-degree (in the chain complex) and T-level (in the filtration) of the bases. One can simply take a basis β of $\bigoplus_{k\in\mathbb{Z}} C_k$ whose elements are ordered lexicographically first by the filtration level and then by their degree, then write the differential matrix D with respect to this basis; this "prioritizes" the (filtration) level over the (complex) degree. Another way was described in section 2.3.2, where the separate matrices D_k are used to represent the differential operators ∂_k with respect to bases ordered by the T-level. These matrices represent blocks in the block-superdiagonal differential matrix D representing the boundary operator $\partial_\bullet = \bigoplus_{k\in\mathbb{Z}} \partial_k$ with respect to a basis β that prioritizes degree of the elements over their filtration level. We focus here on the second approach.

Call a filtered chain complex *basic* if its colimit $C_\bullet$ is isomorphic to an interval complex; equivalently, a filtered complex is basic if its boundary operator $\partial_\bullet$ can be represented by differential matrix consisting of a single 2×2 Jordan block. Then, Theorem 1.6 in [43] states the following:

Theorem 3.22 (Categorical Structural Theorem)**.** *The category of filtered chain complexes is Krull-Schmidt. A filtered complex is indecomposable iff it is basic.*

The proof involves nonconstructive categorical methods similar to those used in the analysis for Lemma 3.20. The main difference is a generalization of Proposition 3.19, where the tempered diagrams are considered in any linear abelian category, not just $\mathbf{vect}(\mathbb{F})$. However, the authors in [43] also provide a constructive proof of this result via the Matrix Structural Theorem (not stated here); in fact, the two theorems are shown to be equivalent.

The importance that is given to filtrations and filtered complexes becomes unveiled in the next section.

3.3.2 Categorical Factorization of Persistent Homology

Categorically, the procedure that calculates the n^{th}-degree homology of a complex with coefficients in the field $\mathbb{F}$ can be described as a functor H_n from the category of chain complexes to $\mathbf{vect}(\mathbb{F})$. To be specific, any chain complex $C_\bullet = (C_k, \partial_k)$ is assigned to the vector space $H_n(C_\bullet) = \ker(\partial_n)/\operatorname{im}(\partial_{n+1})$, and any chain morphism $f_\bullet : C_\bullet \to C'_\bullet$ is assigned to the induced quotient map $H_n(f_\bullet)$. More generally, when chain complexes and chain morphisms are only assumed to be $\mathbb{Z}$-linear, the functor H_n assigns to objects in $\mathbf{mod}(\mathbb{Z})$. This finally elaborates on the closing idea of section 2.1.4.

In the persistent setting, these objects possess an intrinsic filtration. Let $\mathcal{P} = \mathbf{vect}_\star^T(\mathbb{F})$ be the category of tempered persistence vector spaces and $\mathcal{F} = \mathbf{filt}(T)$ be the category of filtered chain complexes, with level indices in T. Here, the n^{th}-degree persistent homology functor $P_n : \mathcal{F} \to \mathcal{P}$ assigns

$$({}_\bullet F_\bullet, {}_\bullet\partial_\bullet) \mapsto P_n({}_\bullet F_\bullet) = \ker({}_\bullet\partial_n)/\operatorname{im}({}_\bullet\partial_{n+1})$$

whose structure maps are induced by canonical inclusions $\ker({}_s\partial_n) \to \ker({}_t\partial_n)$ and $\operatorname{im}({}_s\partial_{n+1}) \to \operatorname{im}({}_t\partial_{n+1})$, while morphisms ${}_\bullet f_\bullet : {}_\bullet F_\bullet \to {}_\bullet G_\bullet$ between complexes are assigned to morphisms $P_n({}_\bullet F_\bullet) \to P_n({}_\bullet G_\bullet)$ given by induced quotient maps at every level $t \in T$.

The standard framework for studying the persistent homology functors $P_n : \mathcal{F} \to \mathcal{P}$ is clarified by the structural theorem for the category $\mathcal{P}$, Proposition 3.19. It suffices to work with an appropriate Krull-Schmidt subcategory of the Krull-Schmidt category $\mathcal{P}$. A filtered chain complex ${}_\bullet F_\bullet$ is studied by decomposing the persistence vector space $P_n({}_\bullet F_\bullet)$ as a sum of indecomposables. Note that in our usage of filtrations, there is a level $q \in T$ such that the diagram ${}_qF_n = \mathbf{0}$, so q^{th} level of the diagram $P_n({}_\bullet F_\bullet)$ is zero; it follows necessarily that all of its indecomposables ${}_\bullet I$ must satisfy ${}_qI = \mathbf{0}$. Denote by $\operatorname{im} P_n$ the full subcategory of $\mathcal{P}$ consisting of persistence vector spaces that satisfy this property.

The notation for the category $\operatorname{im} P_n$ is slightly misleading for it doesn't really depend on n; it is always the same subcategory of $\mathcal{P}$. Anyhow, it is easy to verify that $\operatorname{im} P_n$ is a linear abelian subcategory of $\mathcal{P}$, which completes the proof of the following:

Theorem 3.23. *The persistent homology functor $P_n : \mathcal{F} \to \mathcal{P}$ factors as*

$$\mathcal{F} \to \operatorname{im} P_n \to \mathcal{P}$$

The category $\operatorname{im} P_n$ is Krull-Schmidt. An object in $\operatorname{im} P_n$ is indecomposable if and only if it is isomorphic to a interval complex ${}_\bullet[i, j)$ for some $i \neq -\infty$.

This result may be strengthened by factoring $\mathcal{F} \to \operatorname{im} P_n$ further. To do so, we will work with an appropriate quotient category of the Krull-Schmidt category $\mathcal{F}$; recall the notion of quotient categories from section 3.1.3. Define the category $\operatorname{coim} P_n$ to be be quotient of $\mathcal{F}$ subject to the following equivalence relation on morphisms: two morphisms f and f' in $\mathcal{F}$ are equivalent iff the morphisms $P_n(f)$ and $P_n(f')$ in $\mathcal{P}$

are equal. Initially, it begs to question whether if the tools of abelian categories that were used so far are still applicable here. Firstly, we have:

Lemma 3.24. *A quotient of a Krull-Schmidt category $\mathcal{C}$ is Krull-Schmidt.*

Proof. Since every object X has a finite direct sum decomposition in $\mathcal{C}$, the same is clearly true for the quotient category. Suppose then that X is indecomposable in the quotient category. As an indecomposable object in $\mathcal{C}$, it is either a zero object or has a local endomorphism ring in $\mathcal{C}$. It follows that one of these must also be true in the quotient category; see e.g. page 431 of [39]. ∎

The classification of indecomposables in category $\operatorname{coim} P_n$ now easily follows from the Categorical Structural Theorem. This is independent of the well-known classification of indecomposables in $\mathbf{vect}_\star^T(\mathbb{F})$, a la Theorem 3.19. Using the classification of indecomposables in both (Krull-Schmidt) categories $\operatorname{coim} P_n$ and $\operatorname{im} P_n$, it becomes easy to verify that the functor $\operatorname{coim} P_n \to \operatorname{im} P_n$ is full, faithful, and dense. Invoking Theorem 3.5, we obtain the following result (Theorem 3.3 in [43]):

Theorem 3.25. *The persistent homology functor $P_n : \mathcal{F} \to \mathcal{P}$ factors as*

$$\mathcal{F} \to \operatorname{coim} P_n \to \operatorname{im} P_n \to \mathcal{P}$$

where the functor $\operatorname{coim} P_n \to \operatorname{im} P_n$ is an equivalence of categories.

Note that (unlike the case with $\operatorname{im} P_n$) the category $\operatorname{coim} P_n$ now does depend on the integer n; each $\operatorname{coim} P_n$ is a different quotient category of $\mathcal{F}$. Nonetheless, it is striking that the equivalence of categories holds, suggesting that it is possible to solely work in the category of filtered complexes to obtain information about their barcode decomposition. One may argue that we actually obtain more information this way.

The implicit role of filtrations in computing homology will be thoroughly explained in chapter 4, where a variant persistent homology algorithm is proposed operating on forthright methods of working with (tempered) filtrations.

3.3.3 Induced Matchings of Indecomposables

Following the discussion in this section, a full categorical understanding of persistence vector spaces is achieved if we have a description of how morphisms between them behave. Specifically, we are interested in knowing whether morphisms ${}_\bullet f : {}_\bullet V \to {}_\bullet W$ somehow "associate", or perhaps "match", persistent features of the source space ${}_\bullet V$ with those of the target space ${}_\bullet W$. Since persistent features of a space are associated to intervals in the barcode decomposition of a space, this settles down to whether ${}_\bullet f$ induces some kind of a "function" $\mathcal{B}_V \to \mathcal{B}_W$.

This inquiry is the stepping stone to uncovering the Isometry Theorem (see Theorem 2.19) and the mechanism behind it. A concise categorical approach is outlined in a recent paper [4], whose main results are briefly summarized below, in the language of persistence vector spaces.

Given a morphism of persistence vector spaces ${}_\bullet f : {}_\bullet V \to {}_\bullet W$, begin by writing ${}_\bullet V$ and ${}_\bullet W$ in terms of their interval decompositions:

$${}_\bullet V = \bigoplus_{I \in \mathcal{B}_V} {}_\bullet I \qquad \text{and} \qquad {}_\bullet W = \bigoplus_{J \in \mathcal{B}_W} {}_\bullet J$$

Here, $\mathcal{B}_V$ and $\mathcal{B}_W$ are the barcodes of ${}_\bullet V$ and ${}_\bullet W$ (respectively), and are treated as objects in the category **Mch**; note that barcodes are represented as sets in a way that allows each to contain multiple isomorphic copies of an interval. The idea is then to produce a partial matching $\chi_f : \mathcal{B}_V \to \mathcal{B}_W$ for the morphism ${}_\bullet f$, which is a type of relation on sets that was defined in section 3.1.3.

Noting that the category $\mathbf{vect}^T(\mathbb{Z})$ is abelian, we can factor the morphism ${}_\bullet f$ as ${}_\bullet V \xrightarrow{\bullet e} {}_\bullet R \xrightarrow{\bullet m} {}_\bullet W$ for some monic ${}_\bullet m$ and epic ${}_\bullet e$. Bauer and Lesnick found the following for any monic ${}_\bullet m$ and any epic ${}_\bullet e$:

1. There exists a partial matching $\chi_m : \mathcal{B}_R \to \mathcal{B}_W$ such that if $\chi_m\Big([b,d)\Big) = [b',d')$ then $d = d'$ and $b \geqslant b'$.

2. There exists a partial matching $\chi_e : \mathcal{B}_V \to \mathcal{B}_R$ such that if $\chi_e\Big([b,d)\Big) = [b',d')$ then $b = b'$ and $d \leqslant d'$.

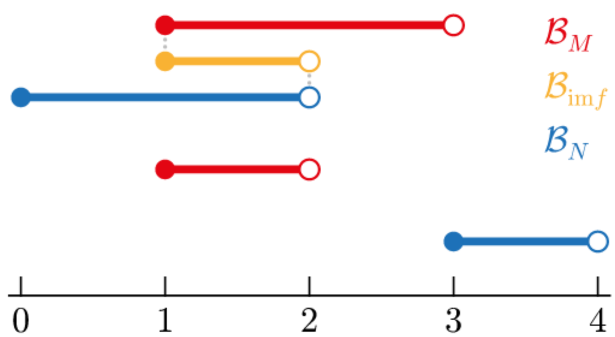
This can also be interpreted by using the framework in [57], where Skraba and Vejdemo-Johansson study the computational aspects of persistent homology using techniques from commutative algebra – particularly the existence of explicit presentations for persistence modules in terms of free generators, calculated by reducing graded matrices to find their Smith normal form.

These are referred to as *canonical matchings* assigned to m and e, respectively. The matching χ_m is constructed by first partitioning each barcode $\mathcal{B}_R$ and $\mathcal{B}_W$ into subsets $\langle \cdot, d\rangle$ of intervals with the same "death" time d. Then, χ_m is defined to be the union of the partial matchings $\langle \cdot, d\rangle_R \to \langle \cdot, d\rangle_W$ where every i^{th}-longest interval in $\langle \cdot, d\rangle_R$ is assigned to the i^{th}-longest interval in $\langle \cdot, d\rangle_W$. Similarly, the matching χ_e is constructed by: partitioning each barcode $\mathcal{B}_V$ and $\mathcal{B}_R$ into subsets $\langle b, \cdot\rangle$ of intervals with the same "birth" time b, taking partial matchings $\langle b, \cdot\rangle_R \to \langle b, \cdot\rangle_W$ where every i^{th}-longest interval in $\langle b, \cdot\rangle_R$ gets assigned to from the i^{th}-longest interval in $\langle b, \cdot\rangle_V$, and finally defining χ_e to be the union of all these matchings. Bauer and Lesnick show that χ_m is necessarily monic and χ_e is necessarily epic in **Mch**.

The partial matching χ_f for the morphism ${}_\bullet f$ is then defined as $\chi_f = \chi_m \circ \chi_e$.

Following the description above, we have $\chi_f\Big([b,d)\Big) = [b',d')$ only if $b \geqslant b'$ and $d \geqslant d'$; furthermore, there has to exist an interval $[b,d')$ in $\mathcal{B}_R$ such that $\chi_e\Big([b,d)\Big) = [b,d')$ and $\chi_m\Big([b,d')\Big) = [b',d')$. That is, a partial matching $\chi_f : \mathcal{B}_V \to \mathcal{B}_W$ is determined constructively by $\mathcal{B}_V$, $\mathcal{B}_W$, $\mathcal{B}_R$ (since $\mathcal{B}_R$ is the barcode of $\mathrm{Im}({}_\bullet f) \simeq \mathrm{Coim}({}_\bullet f)$, with canonical morphisms from/to ${}_\bullet V$ and ${}_\bullet W$), and vice versa.

Example 3.26. The following is a simple example used by the authors of [4] to calculate the induced matching from a morphism of persistence modules.



Let ${}_\bullet M = {}_\bullet[1,2) \oplus {}_\bullet[1,3)$ and ${}_\bullet N = {}_\bullet[0,2) \oplus {}_\bullet[3,4)$. Let ${}_\bullet f : {}_\bullet M \to {}_\bullet N$ be a morphism which maps the summand ${}_\bullet[1,2)$ injectively into the summand ${}_\bullet[0,2)$ and maps the summand ${}_\bullet[1,3)$ to ${}_\bullet\mathbf{0}$. Then $\mathrm{Im}({}_\bullet f) \simeq {}_\bullet[1,2)$. The barcodes $\mathcal{B}_M$, $\mathcal{B}_{\mathrm{Im}(f)}$, and $\mathcal{B}_N$ are plotted. We have $\chi_f : [1,3) \mapsto [0,2)$.

In more concrete terms, we have ${}_\bullet f = \mathrm{im}({}_\bullet f) \circ ({}_\bullet \check{f}) \circ \mathrm{coim}({}_\bullet f)$ for the image morphism $\mathrm{im}({}_\bullet f) : \mathrm{Im}({}_\bullet f) \to {}_\bullet W$, the coimage morphism $\mathrm{coim}({}_\bullet f) : {}_\bullet V \to \mathrm{Coim}({}_\bullet f)$, and the canonical isomorphism ${}_\bullet\check{f} : \mathrm{Coim}({}_\bullet f) \xrightarrow{\sim} {}_\bullet\mathrm{Im}({}_\bullet f)$. Since ${}_\bullet\check{f}$ is both monic and epic it induces an identity matching on $\mathcal{B}_{\mathrm{Coim}(f)} \simeq \mathcal{B}_{\mathrm{Im}(f)}$, while $m = \mathrm{im}({}_\bullet f)$ and $e = \mathrm{coim}({}_\bullet f)$ behave in the same manner as prescribed above.

A note of caution should be taken, as the authors of [4] trustily point out. It may naively appear that assigning any ${}_\bullet f : {}_\bullet V \to {}_\bullet W$ to a matching $\chi_f : \mathcal{B}_V \to \mathcal{B}_W$ specifies a functor $\mathbf{vect}_\star^T(\mathbb{F}) \to \mathbf{Mch}$, but this is not the case; the authors give a short example showing that this assignment does not preserve compositions of morphisms. In fact, the authors of [4] prove that there cannot exist a functor $\mathbf{vect}_\star^T(\mathbb{F}) \to \mathbf{Mch}$ sending every object ${}_\bullet V$ to its barcode $\mathcal{B}_V$, generalizing on a simpler claim where $T = \{1\}$; further examples will be shown in chapter 4. Nonetheless, the assignment ${}_\bullet V \mapsto \mathcal{B}_V$ and ${}_\bullet f \mapsto \chi_f$ does characterize a functor on two specific subcategories of $\mathbf{vect}_\star^T(\mathbb{F})$, namely: (1) the subcategory whose morphisms are monics, and (2) the subcategory whose morphisms are epics.

The Induced Matching Theorem of [4] then states that the kernel and cokernel of ${}_\bullet f : {}_\bullet V \to {}_\bullet W$ allow to assess the size of the matching χ_f as well as restrictions on any pair intervals for which $\chi_f : [b, d) \mapsto [b', d')$, summarized by the following points.

- If $\mathrm{Coker}({}_\bullet f)$ is ε-trivial and $|b' - d'| > \varepsilon$, then χ_f matches some $[b, d') \in \mathcal{B}_{\mathrm{Im}(f)}$ to $[b', d') \in \mathcal{B}_W$ with $b' \leqslant b \leqslant b' + \varepsilon$.

- If $\mathrm{Ker}({}_\bullet f)$ is ε-trivial and $|b - d| > \varepsilon$, then χ_f matches $[b, d) \in \mathcal{B}_V$ to some $[b, d') \in \mathcal{B}_{\mathrm{Coim}(f)}$ with $d' - \varepsilon \leqslant d \leqslant d'$.

Referring to section 2.2.4, the kernel and cokernel of an ε-interleaving are necessarily 2ε-trivial, which by the Induced Matching Theorem guarantees that the matching $r_\varepsilon \circ \chi_f : \mathcal{B}_V \to \mathcal{B}_W$ is an ε-matching; here, $r_\varepsilon : [b, d) \mapsto [b - \varepsilon, d - \varepsilon)$ is a iso matching between $\mathcal{B}_W$ and a multiset of intervals formed from $\mathcal{B}_W$ by shifting the intervals endpoints by ε to the left. This establishes the Algebraic Stability Theorem (part of the Isometry Theorem 2.19), whose converse is also proved near the end of [4].

These appear to have some connection to the canonical short exact sequences of ${}_\bullet f$, shown below; a similar observation is made by Skraba and Turner in [56].

$$ {}_\bullet\mathbf{0} \longrightarrow \operatorname{Ker}({}_\bullet f) \longrightarrow {}_\bullet V \xrightarrow{\operatorname{coim}({}_\bullet f)} \operatorname{Coim}({}_\bullet f) \longrightarrow {}_\bullet\mathbf{0} $$

$$ {}_\bullet\mathbf{0} \longrightarrow \operatorname{Im}({}_\bullet f) \xrightarrow{\operatorname{im}({}_\bullet f)} {}_\bullet W \longrightarrow \operatorname{Coker}({}_\bullet f) \longrightarrow {}_\bullet\mathbf{0} $$

The morphisms $\operatorname{Ker}(f) \to V$ and $W \to \operatorname{Coker}(f)$ are monic and epic, respectively, which thus induce a monic partial matching $\mathcal{B}_{\operatorname{Ker}(f)} \to \mathcal{B}_V$ and an epic partial matching $\mathcal{B}_W \to \mathcal{B}_{\operatorname{Coker}(f)}$. Likewise, morphisms $\operatorname{Im}(f) \to W$ and $V \to \operatorname{Coim}(f)$ induce respective monic and epic matchings $\mathcal{B}_{\operatorname{Im}(f)} \to \mathcal{B}_W$ and $\mathcal{B}_V \to \mathcal{B}_{\operatorname{Coim}(f)} \simeq \mathcal{B}_{\operatorname{Im}(f)}$. Thus, we produce diagrams:

$$ \varnothing \longrightarrow \mathcal{B}_{\operatorname{Ker}(f)} \longrightarrow \mathcal{B}_V \xrightarrow{\chi_{\operatorname{coim}(f)}} \mathcal{B}_{\operatorname{Im}(f)} \longrightarrow \varnothing $$

$$ \varnothing \longrightarrow \mathcal{B}_{\operatorname{Im}(f)} \xrightarrow{\chi_{\operatorname{im}(f)}} \mathcal{B}_W \longrightarrow \mathcal{B}_{\operatorname{Coker}(f)} \longrightarrow \varnothing $$

However, since the assignment $\mathbf{vect}^T_\star(\mathbb{F}) \to \mathbf{Mch}$ is not a functor, we should not expect that these diagrams of partial matchings are complexes, let alone short exact sequences. Indeed, we shall see an explicit example later for which these sequences exhibit this behavior, showing that the category **Mch** cannot be entirely used to represent the structure of objects in $\mathbf{vect}^T_\star(\mathbb{F}) \to \mathbf{Mch}$.

Hence, it is not so straightforward that the Induced Matching Theorem (and so by extension, the Isometry Theorem) should follow from considerations of barcode matchings alone, suggesting that reference to the original structure of persistence vector spaces is necessary to arrive at the desired results. In the next chapter, we discuss how induced matchings may be derived from explicit manipulations of filtered objects in $\operatorname{coim} P_n$, by which the persistent homology of complexes and their mappings shall be computed. This can be called *algorithmic matching* procedure and is contingent on the proposed variant of the persistent homology algorithm (as outlined in section 4.2).

Chapter 4

Calculating Persistent Homology with Filtration Quotients

This chapter demonstrates an algorithm by which persistent homology can be computed by filtration quotients, primarily based on the work in [44] by Killian Meehan, Andrei Pavlichenko, and Jan Segert. Ultimately, it allows transparent extension to (co)kernel and (co)image of a persistence map of filtered chain complexes.

Fundamentally, it works by looking at inclusions of some of naturally-induced filtrations by filtered chain complexes; a background on these notions was previously developed in section 3.3 and more broadly in the paper [43]. In particular, the calculation involves using reduction to find bases of vector spaces that "trivialize" the inclusion map for a pair of filtered spaces, leading to a well-defined matching of generating bases for these filtrations, not unlike how this was done with creator and destroyers of persistent features in section 2.3.2.

A related approach was demonstrated by Skraba and Vejdemo-Johansson in [57]. Here, persistence vector spaces ${}_\bullet V$ are viewed as quotients of finitely-generated free modules by some submodule of "relations" among the selected generators, and persistent homology is then calculated by finding the graded version of the Smith normal form for that inclusion. This viewpoint and ours both enable to formulate an algebraic treatment of (co)kernels and (co)images assigned to morphisms via universal constructions and decomposition results that exist in the category $\mathbf{vect}_{\star}^{T}(\mathbb{F})$. However, our approach using filtration quotients allows everything to be handled the same way: persistent homology of a complex is stated by Proposition 4.12, while (co)kernels and (co)images are defined by diagrams in Propositions 4.16 and 4.17.

Both approaches differ from the conventional way by which persistent homology is computed, where reduction is used to "trivialize" adjacent blocks matrices from the representation of the differential map ${}_{\bullet}C_{\bullet}$. While this is an efficient procedure for finding the creator-destroyer pairs whose birth-death times are summarized by the barcode of the complex, it does not easily generalize to a method for finding out the persistent homology of other data associated with the complexes. Of note, morphisms ${}_{\bullet}f_{\bullet}$ between filtered chain modules are not so easily described; finding the interval decomposition associated to the kernel, image, and cokernel persistence vector spaces of ${}_{\bullet}f_{\bullet}$ require the construction and reduction of matrices whose meaning is not immediately clear. This is evident by studying the elaborate and well-defined algorithm that is presented in [16] by Cohen-Steiner, Edelsbrunner, Harer, and Morozov. Note that it only works with certain types of morphisms – namely, those induced by inclusions of simplicial filtrations.

Nevertheless, all of the algorithms are rooted in elementary linear algebraic operations, where matrix reduction is used to compute appropriate bases of vector spaces. Ultimately, the main distinction between the two is which class of vector space diagrams in $\mathbf{vect}^{T}(\mathbb{F})$ are to be used, on which the construction of representative matrices depends. It is also germane to note that whichever algorithm is used, an important principle to adhere to is for these constructions and the ensuing reduction(s) to be performed in way independent of the choice of vector space basis.

4.1 Elementary notions and operations

In this short section, we briefly discuss some linear-algebraic properties of column-reduction. Doing column-reduction easily yields ordered bases for the kernel and also the image of a matrix. Another common form of matrix reduction is Gaussian elimination, which by contrast easily yields a basis for the image a matrix, but requires additional back-substitution to produce a basis for the kernel. Column-reduction algorithms are therefore a convenient alternative to Gaussian elimination for matrix computations in general. Perhaps this fact seems to be generally underappreciated.

4.1.1 Column Reduction and Bruhat Factorization

Let $f : X \to Y$ be a linear map between finite-dimensional vector spaces (over some fixed field $\mathbb{F}$). We can select a basis β_X and β_Y for respective X and Y. Recall that any other basis for X (resp. Y) can be found from $\beta = \beta_X$ (resp. $\beta = \beta_Y$) by a sequence of *elementary operations*, which consist of:

- swapping the positions of two elements z_i and z_j in β;
- replacing any $z_k \in \beta$ by $\alpha \cdot z_k$, where $\alpha \in \mathbb{F}$ is nonzero;
- replacing any $z_i \in \beta$ by $z_i + \alpha \cdot z_j$, where z_j is in β and $\alpha \in \mathbb{F}$.

For each $z_i \in \beta_X$, the element $f(z_i)$ can be written as a linear combination of elements in β_Y; recall that these equations are unique due to linear independence. The *matrix representation* of the map f is a matrix $M = [f]_{\beta_X}^{\beta_Y}$ whose column (in position) k

consists of coefficients appearing in the linear combination of $f(z_i)$ (including all zero entries) with respect to the ordered basis β_Y; these are ordered from top of the column downwards. For any $z \in X$, this determines the unique linear combination of elements in β_Y needed to produce $f(z)$ by taking appropriate linear combinations of the columns in its matrix.

Applying a sequence of elementary operations on β_X is equivalent to applying *column operations* on M:

[Type 1] some columns i and j of M are swapped;

[Type 2] some column k of M is multiplied by a nonzero scalar;

[Type 3] add a scalar multiple of some column l to some column k.

Concretely, if elementary operations are used to transform β_X into a basis β'_X then M transforms into a matrix $M' = [f]^{\beta_Y}_{\beta'_X}$ such that $M' = M \cdot [1_V]^{\beta'_X}_{\beta_X}$, where $1_X : X \to X$ is the identity map. Because the same column operations are used to compute both M' and $[1_X]^{\beta'_X}_{\beta_X}$, it is efficient to apply them on the augmented matrix $\left[\frac{M}{I}\right]$ to get the result (where $I = [1_X]^{\beta_X}_{\beta_X}$ is the identity matrix).

Similarly, applying a sequence elementary basis operations on β_Y is equivalent to doing *row operations* on M; infer these from above. Concretely, the elementary operations used to transform β_W into β''_W also transform M into a matrix $M'' = [f]^{\beta''_Y}_{\beta_X}$ such that $M'' = [1_Y]^{\beta''_Y}_{\beta_Y} \cdot M$. It is efficient to apply the row operations on the augmented matrix $[I|M]$ to get the result.

A major technique that uses column (and/or row) operations is matrix reduction. In section 2.3.2, it was already shown how find the column-reduced form of the matrix for the differential map can be used to compute the persistent homology associated with a filtration. Building on that, our usage shall be restricted to elementary column operations of Type 3 where a scalar multiple of column l can be added to column k only if $l < k$. It readily follows that this transforms basis β_V to a basis β'_X such that the change-of-basis matrix $V = [1_X]_{\beta_X}^{\beta'_X}$ is upper triangular.

Column reduction is an inductive procedure: start at column $i = 1$, and follow the inductive step, which goes as follows:

```
(1) Check if current column i has only zero entries; if yes, go to (4).
(2) If no, find the ``lowest'' row p where column i has a nonzero
    entry; entries in rows p+1, p+2, ... of column i should be zero.
(3) Check if there is some column j to the left of column i such that
    their lowest nonzero entries are in the same row.
   (3a) If yes, subtract a multiple of column j from column i so that
        updated column i has a zero entry in row p; then go to (1).
   (3b) If no, declare column i to have a pivot in row p.
(4) Continue to column i+1 and repeat.
```

The resulting matrix M' will have the property that every column is either a zero column or a pivoted column, and each pivot entry lies in a unique row of M; we thus proclaim M' to be a *reduced matrix*.

Example 4.1 (see Appendix B in [43]). Let M be a matrix given by

$$M = \begin{bmatrix} 1 & -2 & 0 & -8 \\ 2 & -4 & 6 & 2 \\ 1 & -2 & 2 & -2 \end{bmatrix}$$

We reduce the matrix by applying the inductive algorithm detailed above. We will work with the augmented matrix $\left[\frac{M}{I}\right]$, but only consider the nonzero entries from top three rows (ie., those in M itself) to answer conditional statements in the algorithm. For emphasis, all column pivots will boldened.

$$\frac{M}{I} = \left[\begin{array}{cccc} 1 & -2 & 0 & -8 \\ 2 & -4 & 6 & 2 \\ 1 & -2 & 2 & -2 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \mapsto \left[\begin{array}{cccc} 1 & 0 & -2 & -6 \\ 2 & 0 & 2 & 6 \\ \mathbf{1} & 0 & 0 & 0 \\ \hline 1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \mapsto \left[\begin{array}{cccc} 1 & 0 & -2 & 0 \\ 2 & 0 & \mathbf{2} & 0 \\ \mathbf{1} & 0 & 0 & 0 \\ \hline 1 & 2 & -2 & 8 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{array}\right] = \frac{R}{V}$$

Clearly, we have $R = M \cdot V$ or equivalently $M = R \cdot V^{-1}$. Furthermore, we can reduce the rows of R. The nonzero entries of every column in R (except for its pivot) can be eliminated by Type 3 row operations using the row that contains its pivot. Working from the bottom-most row upwards, this results in the following equation:

$$[I|R] = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 2 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 1 & \mathbf{1} & 0 & 0 & 0 \end{array}\right] \mapsto \left[\begin{array}{ccc|cccc} 1 & 1 & -3 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 1 & \mathbf{1} & 0 & 0 & 0 \end{array}\right] = [W|Q]$$

Hence, $Q = WR$ or $R = W^{-1}Q = UQ$. Combining these results together:

$$M = \begin{bmatrix} 1 & -2 & 0 & -8 \\ 2 & -4 & 6 & 2 \\ 1 & -2 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 & 2 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} = U \cdot Q \cdot V^{-1}$$

Alternatively, equation $R = W^{-1}Q$ can be found by "re-shuffling" pivoted columns of R so their pivots lie on the diagonal of an upper triangular matrix.

The process can be repeated for any matrix M, resulting in the equation

$$M = UQV^{-1} \tag{4.1}$$

where U, V are *(upper) uni-triangular* and Q is a *quasi-monomial* matrix – that is, U and V are upper-triangular with 1 on their diagonal, and Q is such that each row and column has at most one non-zero entry. The change-of-bases matrices U and V

are generally non-unique in this factorization. However, the matrix Q is essentially unique as the following quick lemma suggests:

Lemma 4.2 (see Lemma A.1 in [43]). *Suppose $MV = UN$, where M and N are quasi-monomial and U and V are upper uni-triangular. Then $M = N$.*

Proof. It suffices to show that every nonzero entry of N is also an entry of M; it will follow that every nonzero entry of M is also an entry of N by rewriting the matrix equation as $U^{-1}M = NV^{-1}$. Since U is uni-triangular, the nonzero entries of N are exactly the (column) pivots of UN, and hence the pivots of MV. On the other hand, since V is uni-triangular, every nonzero entry of M is its row's "leading" nonzero entry in MV; in particular, the leading nonzero entries of rows in MV lie in distinct columns. We claim: every pivot of $S = UN$ is also its row's leading entry in $S = MV$, thus proving the initial proposition.

This is proved by contradiction. Suppose (to the contrary) that the pivot of some column j_1 of S is not the leading nonzero entry of its row i_1 in S. Then there exists an entry in column $j_2 < j_1$ of row i_1; since (column) pivots are in distinct rows of S, then column j_2 has a pivot in row $i_2 > i_1$. The argument repeats: this nonzero entry cannot be the leading one of row i_2 (since row i_1 already has a leading entry in column j_2), hence there a column $j_3 < j_2$ where that leading entry is; but then, there must be a row $i_3 > i_2$ where the pivot of column j_3 is located (since column j_2 already has a pivot in row i_2). It follows that there is an infinite sequence of pivots in matrix S, which contradicts its finite dimensionality. ∎

When M is invertible, equation 4.1 can be called the *Bruhat factorization* of M. In this case, it is necessarily true that Q is invertible, and can be obtained from a permutation matrix P by multiplying its columns by appropriate nonzero scalars. Actually, equation 4.1 can be rewritten to reflect this if we admit Type 2 column operations into our repertoire, although this comes at the cost of the matrices V and W no longer being uni-triangular.

4.1.2 Filtrations and Adapted Bases

Filtrations of vector spaces were previously introduced in section 3.3, with Lemma 3.20 covering the importance of their colimit. A filtration ${}_\bullet U$ specifically refers to a filtration of subspaces ${}_t U \subseteq F$, for some finite-dimensional vector space F and index set $T = \mathbb{Z}$; the space F can be called a *reference vector space* for the filtration ${}_\bullet U$. To any such filtration ${}_\bullet U$, we associate the ${}_\bullet U$*-level* $\lambda(z) = \lambda(z, {}_\bullet U)$ for any $z \in F$ that is defined via the set $q(z) = \{s \in \mathbb{Z} | z \in {}_s U\}$:

- $\lambda(z) = j$ if $q(z)$ is nonempty and has smallest element $j \in \mathbb{Z}$.
- $\lambda(z) = \infty$ if $q(z)$ is empty.
- $\lambda(z) = -\infty$ if $q(z)$ is nonempty and has no smallest element.

We now discuss how to represent filtrations by adapted bases of F.

Definition 4.3. Let ${}_\bullet U$ be a filtration of an n-dimensional reference vector space F. A basis $\beta = \{z^j\}_{j=1}^n$ of F is *adapted* to the filtration ${}_\bullet U$ if for every $t \in \mathbb{Z}$ the subspace ${}_t U$ equals the span of $\{z^1, \ldots, z^{\dim {}_t U}\}$.

Given an basis β of F adapted to a filtration ${}_\bullet U$, we can easily construct other bases adapted to ${}_\bullet U$; we can do this by performing a sequence of any Type 3 elementary operations on β where $z^k \in \beta$ is replaced by $z^k + \alpha \cdot z^j$ such that $\lambda(z^j) \leqslant \lambda(z^k)$. It is also clear that Type 2 operations can be used to "modify" an adapted basis, while Type 1 operations can only be used with basis elements that the same ${}_\bullet U$-level.

Conversely, a choice of basis β of F and a choice of level $\lambda(v)$ for all $v \in \beta$ together naturally specify a filtration ${}_\bullet U$ to which β is adapted; for every $t \in \mathbb{Z}$, ${}_t U$ is the span of the set $\{v \in \beta : \lambda(v) \leqslant t\}$. This suggests a slightly more general definition, yet one that proves very crucial.

Definition 4.4. Given a filtration ${}_\bullet U$ of an n-dimensional reference vector space F with limit $\mathbf{0}$ and colimit F, a basis β of F is *almost-adapted* to filtration ${}_\bullet U$ if every subspace ${}_t U$ is spanned by the set of elements in β that are contained in ${}_t U$.

Equivalently, there exists a permutation π of $\{1, \ldots, n\}$ such that every ${}_t U$ equals the span of $\{z^{\pi(j)}\}_{j=1}^{\dim {}_t U}$. So an almost-adapted basis is adapted if and only if the level $\lambda(v^k)$ is a nondecreasing function of the index variable k.

Take any two bases β and μ of F, such that μ is adapted to a filtration ${}_t U$ on F. The identity function $1_F : F \to F$ has a matrix representation $M = [1_F]_\beta^\mu$ with respect to these bases; then calculate the Bruhat factorization of M (refer to section 4.1.1). The upper uni-triangular matices establish bases β' and μ' of F, with μ' being adapted to ${}_\bullet U$. Of interest however is that the quasi-monomial matrix $[1_F]_{\beta'}^{\mu'}$, transforming β' into μ' by reordering the basis (and possibly applying Type 2 elementary operations on its elements), results in an almost-adapted basis β' to ${}_\bullet U$.

Remark 4.5. In finding the almost-adapted basis β' for ${}_\bullet U$ it suffices to perform column reduction on M, which already describes the transformation of β into β'.

Now if β was adapted to another filtration ${}_\bullet W$ of vector space F, then β' is also adapted ${}_\bullet W$ by construction. That is, the basis β' is *simultaneously* (!) almost-adapted to both filtrations ${}_\bullet U$ and ${}_\bullet W$. This simple observation is nonetheless a grand underpinning of things to come, so it is restated separately below:

Lemma 4.6. *Given two filtrations ${}_\bullet U$ and ${}_\bullet U'$ of a reference vector space F, there exists a basis β of F that is almost-adapted to both.*

Alternatively, one may view Lemma 4.6 as a geometric manifestation of Bruhat factorization: this represents every action $g \in G$ of (complete) flags F. For more discussion about the induced action by the general linear group G on flags, see section 5 of [3]. On another note, there generally do not exist bases that are adapted to more than any chosen two filtrations of a reference vector space, as calculations in section 4.3 demonstrate.

4.1.3 Elementary Operations on Filtrations

In a given vector space, there are many useful operations that can be performed on its linear subspaces; many of which can be performed on filtrations of the vector space. These were first presented in section 2.2.2, which we will briefly describe here. Throughout, let F and G be finite-dimensional vector spaces.

Here, we will be using our extended definition of level (including $-\infty$ and ∞) as it is necessary to discuss range and preimage filtration. Equivalently, the colimit of the range may not by given by F while the limit of the preimage filtration may not be given by $\mathbf{0}$. Because the level λ of a basis element in a filtration is conventionally stated by a function into T, this is presumably the reason these objects are not generally discussed in most sources.

Suppose that $h : F \to G$ is a linear map. If there exists a filtration ${}_\bullet U$ on F, then we have the *range filtration* $h\left({}_\bullet U\right)$ whose reference vector space is G. An adapted basis of this filtration can be found via the following procedure.

1. `Find bases` β `of` F `and` γ `of` G `such that` β `is adapted to` ${}_\bullet U$.
2. `Apply` h `to every element` z^k `in` β`, and write matrix representation` $[h]_\beta^\gamma$`: every column` k `represents element` $h(z^k)$ `with respect to` γ.
3. `Do the reduction procedure on this matrix; the pivoted columns of the reduced matrix represent elements in the basis, denoted` $\rho_\times$`, of the range subspace` $h(F) \subseteq G$.
4. `Extend` $\rho_\times$ `to a basis` ρ `of` G `by appending elements` g^k `of` γ `to` $\rho_\times$ `for any row` k `of the reduced matrix that does not contain any column pivots.`

It checks out that the $h\left({}_\bullet U\right)$-level of any element in $\rho_\times$ equals the ${}_\bullet U$-level of $z^k \in \beta$, where k is the column in the reduced matrix that contains the γ-representation of ρ_0. By definition, elements in $\rho \setminus \rho_\times$ have $h\left({}_\bullet U\right)$-level equal to ∞.

On the other hand, if there exists a filtration ${}_\bullet V$ on G, then we have the *preimage filtration* $h^{-1}\left({}_\bullet V\right)$ whose reference vector space is F. An adapted basis of this filtration can be found via a modification of the previous procedure.

1. `Find a bases` β `of` F `and` γ `of` G `such that` γ `is adapted to` ${}_\bullet V$.
2. `Apply` h `to every element` z^k `in` β`, and write the matrix` $[h]_\beta^\gamma$.
3. `Do the reduction procedure on this matrix; the change-of-basis matrix transforms` β `into basis` κ `of` F.

Denote the set of elements $w \in \kappa$ such that $h(w) \neq 0$ by $\kappa_\times$. It checks out that the $h^{-1}\left({}_\bullet V\right)$-level of any element $w \in \kappa_\times$ equals the ${}_\bullet V$-level of $h(w)$, while (by definition) elements in $\kappa \setminus \kappa_\times$ have $h^{-1}\left({}_\bullet V\right)$-level equal to $-\infty$.

Lemma 4.7. *The bases κ of F and ρ of G produced by the procedures above are almost-adapted to the preimage filtration $h^{-1}({}_{\bullet}V)$ and adapted to the range filtration $h({}_{\bullet}U)$, respectively. Therefore, reordering the basis with respect to the associated filtration level makes it adapted to the filtration.*

Proof. It is a well-known fact from linear algebra that the column space of a matrix generates the range of its operator; column reduction merely uncovers a linearly independent subset $\rho_\times$ of these elements. The range-level of elements in G is trivial to determine.

On the other hand, basis κ was constructed such that $\kappa \setminus \kappa_\times$ is a basis for $h^{-1}(\mathbf{0})$ and the ordered set $\{h(w) : w \in \kappa_\times\}$ is a basis of $h(F)$ that is almost-adapted to the range filtration $h(h^{-1}({}_{\bullet}V))$. By definition of the preimage, the basis κ is almost-adapted to $h^{-1}({}_{\bullet}V)$. ∎

Therefore, computing adapted bases of the preimage and range filtrations for linear map $h : F \to G$ involves the reduction of the matrix $[h]$, whose pivots then specify finite levels for the range (determined from their column index) and the preimage (determined from their row index). Specifically, one just needs to do a look-up of the filtration level using these indices.

We also have operations for working with filtrations ${}_{\bullet}U$ and ${}_{\bullet}V$ of a single reference space F. Here we can take their *sum filtration* ${}_{\bullet}U + {}_{\bullet}V$ and *intersection filtration* ${}_{\bullet}U \cap {}_{\bullet}V$. The adapted bases for these filtrations can be easily found, as the next lemma demonstrates.

Lemma 4.8. *If basis β of F is simultaneously almost-adapted to filtrations ${}_{\bullet}U$ and ${}_{\bullet}V$, then β is almost-adapted to ${}_{\bullet}U + {}_{\bullet}V$ and ${}_{\bullet}U \cap {}_{\bullet}V$. The level of an element $z \in \beta$*

with respect to these filtrations is determined by its ${}_\bullet U$*-level* $u(z)$ *and* ${}_\bullet V$*-level* $v(z)$*:*

- *the* $({}_\bullet U + {}_\bullet V)$*-level of* z *equals* $\min(u(z), v(z))$*;*
- *the* $({}_\bullet U \cap {}_\bullet V)$*-level of* z *equals* $\max(u(z), v(z))$.

Proof. It is readily follows that for any $t \in \mathbb{Z}$, the subspaces ${}_tU + {}_tV$ and ${}_tU \cap {}_tV$ are spanned by some elements of β. Conversely, the criteria for determining the level of a basis element ensure that $z \in \beta$ is in ${}_tU + {}_tV$ only if $z \in {}_tU$ or $z \in {}_tV$, and similarly for the intersection. ∎

Remark 4.9. A classical way to find the basis of the sum and intersection subspaces is by considering the short exact sequence involving the direct sum of submodules in section 3.2.2. Simply put, we take an matrix M whose columns consist first of elements in the basis adapted to subspace $U \subseteq F$ and then elements in the basis adapted to subspace $V \subseteq F$; the columns are then reordered so that those at level ∞ (in their respective subspace) are put at the end. Matrix M represents the linear operator $F \oplus F \to F + F = F$ with respect to these adapted bases. After reducing this matrix, we look at the pivoted columns and the kernel of the matrix – the pivoted columns determine a basis adapted to the sum $U + V$, while the kernel helps find a basis adapted to the intersection $U \cap V$. The last step is slightly involved, which may appear discouraging for those attempting explicit calculations.

Another important operation, *quotients* of filtrations, deserves its own section.

4.1.4 Quotients of Filtered Spaces

Suppose ${}_\bullet U \subseteq {}_\bullet W$ are filtrations of the reference vector space F . The *quotient* ${}_\bullet W/{}_\bullet U$ is the persistence vector space

$$\cdots \longrightarrow {}_{t-1}W/{}_{t-1}U \longrightarrow {}_{t}W/{}_{t}U \longrightarrow {}_{t+1}W/{}_{t+1}U \longrightarrow \cdots$$

where each arrow ${}_sW/{}_sU \to {}_{s+1}W/{}_{s+1}U$ is the map induced from the identity $F \to F$. Such a quotient is a tempered persistence vector space. The type and multiplicity of indecomposable summands in persistent homology are commonly described in terms of the barcode invariant(s).

Now, let's consider the decomposition of the filtration quotient ${}_\bullet W/{}_\bullet U$. For $s \in \mathbb{Z}$, an element of any quotient vector space ${}_sW/{}_sU$ is represented as an equivalence class of a vector $w \in {}_sW$,

$${}_s[w] := w + {}_sU = \{w + u \in {}_sW | u \in {}_sU\}$$

where ${}_s[w'] = {}_s[w]$ if and only if $w' - w \in {}_sU$. We will extend this notation to any $z \in F$ by defining ${}_s[z] := 0 \in {}_sW/{}_sU$ whenever $z \notin {}_sW$. Then to any z in the reference vector space F we associate the persistence vector space $\langle {}_\bullet[z] \rangle$ defined by the diagram of vector spaces

$$\cdots \longrightarrow \langle {}_{t-1}[z] \rangle \longrightarrow \langle {}_{t}[z] \rangle \longrightarrow \langle {}_{t+1}[z] \rangle \longrightarrow \cdots$$

where each span $\langle {}_s[z] \rangle$ is a subspace of ${}_sW/{}_sU$ and each arrow between (nonzero) vector spaces is induced from the identity $F \to F$.

Proposition 4.10. *Suppose ${}_\bullet U \subseteq {}_\bullet W$ are filtrations of the reference vector space F and $z \in F$ is nonzero. Let $u(z)$ denote the ${}_\bullet U$-level of z, and let $w(z)$ denote the ${}_\bullet W$-level of z. Then $w(z) \leqslant u(z)$, and exactly one of the following cases holds:*

1. *If $w(z) < u(z)$, then $\langle_\bullet [z]\rangle$ is isomorphic to the interval persistence vector space:*

$$\langle_\bullet [z]\rangle \simeq {}_\bullet\Big[w(z), u(z)\Big)$$

2. *If $w(z) = u(z)$, then $\langle_\bullet [z]\rangle$ is the zero persistence vector space ${}_\bullet\mathbf{0}$.*

Proof. It is necessarily true that $w(z) \leqslant u(z)$, for otherwise the inclusion relation ${}_\bullet U \subseteq {}_\bullet W$ would be violated. The case $w(z) = u(z)$ is trivial, so assume $w(z) < u(z)$. For every $s \in \mathbb{Z}$, the vector space $\langle_s [z]\rangle$ is at most 1-dimensional; one can check that this happens precisely when $w(z) \leqslant s < u(z)$. ∎

For the case when $w(z) = u(z) = p$, the zero persistence vector space is often denoted by an "empty interval" ${}_\bullet[p, p) := {}_\bullet\mathbf{0}$. Note that provided a filtration ${}_\bullet U$ satisfies ${}_{m^-}U = \mathbf{0}$ for some $m^- \in \mathbb{Z}$, the level of a nonzero vector cannot be $-\infty$.

Following Lemma 4.6, we obtain a clean decomposition result for filtration quotients.

Theorem 4.11. *Suppose ${}_\bullet U \subseteq {}_\bullet W$ are filtrations of the reference vector space F. Suppose a basis $\beta = \{z^k | 1 \leqslant k \leqslant \dim F\}$ of F is simultaneously almost-adapted to ${}_\bullet U$ and to ${}_\bullet W$. Then:*

$${}_\bullet W / {}_\bullet U = \bigoplus_{k=1}^{\dim V} \langle_\bullet [z^k]\rangle$$

where the zero summands may be discarded.

Proof. The fact that β is almost adapted to both filtrations ensures that for each $s \in \mathbb{Z}$ the quotient ${}_sW / {}_sU$ is isomorphic to the span of some subset of β. But then, the transition maps of ${}_\bullet W / {}_\bullet U$ simply take elements of β to themselves. ∎

Considering that the interval decomposition of a persistence vector space is what allows us to find persistent homology, this theorem is a crucial component of our calculations. One must again point out that this is possible in part due to Bruhat factorization, given by equation 4.1.

4.2 Homology of Filtered Complexes

We now demonstrate our algorithm to explicitly compute persistent features of filtered chain complexes based on the mechanics of simultaneously almost-adapted bases. Sections 4.2.1 and 4.2.2 each contain an example of a filtered chain complex for which the H_1 homology is computed. Then, in section 4.3 we explore the persistent homology of a morphism between them.

Recall filtered chain complexes $({}_\bullet C_\bullet, {}_\bullet\partial_\bullet)$, which were discussed formally in sections 3.3.1 and 3.3.2, and let $(C_\bullet, \partial_\bullet)$ be its colimit complex. These come equipped with what we call *canonical filtrations*, given by the filtration specified for each vector space C_k in the complex. We further define:

- The *preimage-canonical filtration* on C_k is the preimage ${}_\bullet P_k = \partial_k^{-1}({}_\bullet C_{k-1})$ of the canonical filtration on C_{k-1} under the differential $\partial_k : C_k \to C_{k-1}$.

- The *range-canonical filtration* on C_k is the range ${}_\bullet R_k = \partial_{k+1}({}_\bullet C_{k+1})$ of the canonical filtration on C_{k+1} under the differential $\partial_{k+1} : C_{k1+1} \to C_k$.

A computation with a filtered chain complex ${}_\bullet C_\bullet$ starts by identifying bases adapted to the canonical filtrations of the relevant vector spaces C_k. An *intrinsic basis* is chosen that is adapted to the canonical filtration: it is necessary to choose an ordering of the simplices such that the filtration level is nondecreasing. The choice of an intrinsic basis is not unique, except in the special case where each k-simplex has a distinct filtration level. Note that any ordering of the k-simplices of a filtered complex yields a basis of C_k that is almost-adapted to the canonical filtration.

Now, we utilize the machinery we developed in section 4.1 to explain the variant persistent homology algorithm. Given a filtered chain complex ${}_\bullet C_\bullet$, let ${}_\bullet Z$ be the restriction of $Z = \ker(\partial_n)$ to the canonical filtration on C_n and ${}_\bullet B$ is the restriction of of $B = \mathrm{im}(\partial_{n+1})$ to the range-canonical filtration on C_n.

Proposition 4.12. *With the filtrations ${}_\bullet Z$ and ${}_\bullet B$ of reference vector space Z as described above, the persistent homology of a filtered chain complex ${}_\bullet C_\bullet$ can be expressed as the quotient of filtrations:*

$$H_n({}_\bullet C_\bullet) = {}_\bullet Z / {}_\bullet B$$

The algorithm is based on factorizing the differential ∂_{n+1} to provide an alternative description of the filtration ${}_\bullet B$. Let $i_n : Z \to C_n$ be the subspace inclusion of $Z = \ker(\partial_n) \subseteq C_n$. Since $\partial_n \circ \partial_{n+1} = 0$ and $i_n : Z \to C_n$ is a categorical kernel of $\partial_n : C_n \to C_{n-1}$, its universal property ensures that there exists a unique linear map

$$\delta_{n+1} : C_{n+1} \to Z$$

that satisfies $\partial_{n+1} = i_n \circ \delta_{n+1}$. Consequently, we have the following:

Lemma 4.13. *The range filtration of the canonical filtration on C_{n+1} under the linear map $\delta_{n+1} : C_{n+1} \to Z$ produces exactly the filtration ${}_\bullet B$ on $Z \subseteq C_n$.*

The variant persistent homology (**PH**) algorithm can be summarized in two steps:

PH1: *Express δ_{n+1} as a matrix Δ_{n+1} relative to appropriately adapted bases.*

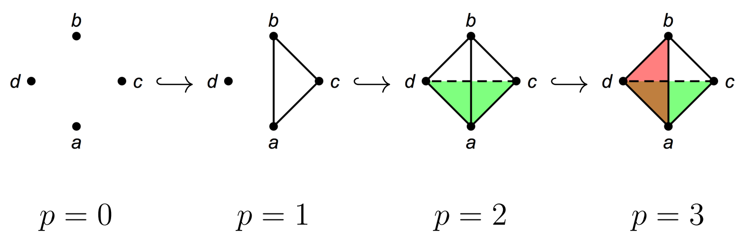
PH2: *Apply matrix reduction to Δ_{n+1} to obtain a basis of Z that is adapted to ${}_\bullet B$ and simultaneously almost-adapted to ${}_\bullet Z$.*

We draw attention to the fact that reduction of the single matrix Δ_{n+1} outputs the desired simultaneously almost-adapted basis. But before this step, additional computation is required to construct the matrix Δ_{n+1}. By contrast, the standard algorithm achieves better computational efficiency by cleverly combining the outputs from reduction of the two matrices, the expressions of ∂_n and of ∂_{n+1}, to construct a simultaneously almost-adapted basis.

Remark 4.14. In what follows, elements of a given basis β will be denoted ${}_{\lambda}\mathbf{e}^k$. Here, k is the order of $\mathbf{e}$ within β, while λ refers to the level $\lambda(\mathbf{e})$ of $\mathbf{e}$ in the filtration to which β is adapted. This filtration should be inferred from context.

4.2.1 First computation

This is a detailed demonstration of the variant persistent homology algorithm, applied to compute the dimension $n = 1$ persistent homology of the filtered simplicial complex shown in Figure 4.1. See sections 2.1.1 and 2.3.1 for the necessary background.



$$C_0 = \left\langle {}_0a, {}_0b, {}_0c, {}_0d \right\rangle$$
$$C_1 = \left\langle {}_1ab, {}_1ac, {}_1bc, {}_2ad, {}_2bd, {}_2cd \right\rangle$$
$$C_2 = \left\langle {}_2acd, {}_3abd \right\rangle$$

Figure 4.1: Shown on the right are the intrinsic bases of C_0, C_1, C_2; the prescript on the elements denotes their level in the canonical filtration.

PH1a: *Construct a basis* $\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^{\dim Z}$ *of* $Z \subseteq C_1$ *that is adapted to the restriction* ${}_{\bullet}Z$ *of the canonical filtration on* C_1.

To do this, we reduce the matrix $\mathbf{D}_1$ representing the differential operator $\partial_1 : C_1 \to C_0$ relative to intrinsic bases of C_0 and C_1. Doing so constructs a basis $\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^{\dim C_1}$ of C_1 that is simultaneously: adapted to the canonical filtration ${}_{\bullet}C_1$ and almost-adapted to the preimage-canonical filtration ${}_{\bullet}P_1$.

Having the background on differential operators of simplicial complexes as established in section 2.1.2, construct matrix $\mathbf{D}_1$ with reference to Figure 4.1:

$$\mathbf{D}_1 = \begin{array}{c} \\ {}_0a \\ {}_0b \\ {}_0c \\ {}_0d \end{array} \begin{array}{c} \begin{array}{cccccc} {}_1ab & {}_1ac & {}_1bc & {}_2ad & {}_2bd & {}_2cd \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Reduce $\mathbf{D}_1$ according to the procedure in section 4.1.1. This yields the reduced matrix $\underline{\mathbf{D}}_1$ and the upper uni-triangular change of basis matrix $\mathbf{V}_1$ such that $\underline{\mathbf{D}}_1 = \mathbf{D}_1\, \mathbf{V}_1$. The prescript λ of ${}_{\lambda}\mathbf{c}^k$ denotes its ${}_{\bullet}C_1$-level.

$$\underline{\mathbf{D}}_1 = \begin{array}{c} \\ {}_0a \\ {}_0b \\ {}_0c \\ {}_0d \end{array} \begin{array}{c} \begin{array}{cccccc} {}_1\mathbf{c}^1 & {}_1\mathbf{c}^2 & {}_1\mathbf{c}^3 & {}_2\mathbf{c}^4 & {}_2\mathbf{c}^5 & {}_2\mathbf{c}^6 \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & -1 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \end{bmatrix} \end{array} \qquad \mathbf{V}_1 = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} \begin{array}{cccccc} {}_1\mathbf{c}^1 & {}_1\mathbf{c}^2 & {}_1\mathbf{c}^3 & {}_2\mathbf{c}^4 & {}_2\mathbf{c}^5 & {}_2\mathbf{c}^6 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Let ρ be the level function on C_1 with respect to filtration ${}_{\bullet}P_1$. For each basis element $\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^{\dim C_1}$ we can read off the ${}_{\bullet}P_1$-level from the rows of $\underline{\mathbf{D}}_1$:

$$\rho({}_1\mathbf{c}^1) = 0,\ \ \rho({}_1\mathbf{c}^2) = 0,\ \ \rho({}_1\mathbf{c}^3) = -\infty,\ \ \rho({}_2\mathbf{c}^4) = 0,\ \ \rho({}_2\mathbf{c}^5) = -\infty,\ \ \rho({}_2\mathbf{c}^6) = -\infty.$$

Since the basis $\mathbf{c}^1, \mathbf{c}^2, \dots, \mathbf{c}^{\dim C_1}$ is almost-adapted to ${}_{\bullet}P_1$, discarding elements that do not have ${}_{\bullet}P_1$-level $-\infty$ yields a basis $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^{\dim Z}$ of $Z = {}_{-\infty}P_1$:

$${}_1\mathbf{p}^1 = {}_1\mathbf{c}^3, \quad {}_2\mathbf{p}^2 = {}_2\mathbf{c}^5, \quad {}_2\mathbf{p}^3 = {}_2\mathbf{c}^6.$$

The prescript λ of ${}_{\lambda}\mathbf{p}^k$ denotes its ${}_{\bullet}C_1$-level, same as of ${}_{\lambda}\mathbf{c}$ (adapted to ${}_{\bullet}C_1$), maintaining that ${}_{\bullet}Z$ is the restriction of $Z \subseteq C_1$ to canonical filtration ${}_{\bullet}C_1$.

We otherwise discard unused columns of the matrix $\mathbf{V}_1$, yielding a matrix $\mathbf{P}$ that expresses each basis element $\mathbf{p}$ in terms of the intrinsic basis of C_1:

$$\mathbf{P} = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} \begin{array}{ccc} {}_1\mathbf{p}^1 & {}_2\mathbf{p}^2 & {}_2\mathbf{p}^3 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The basis $\mathbf{p}^1, \mathbf{p}^2, \ldots, \mathbf{p}^{\dim Z}$ is adapted to the canonical filtration ${}_\bullet Z$ on Z.

Remark 4.15. The equations

$${}_{-\infty}\mathbf{q}^1 = {}_1\mathbf{c}^3, \quad {}_{-\infty}\mathbf{q}^2 = {}_2\mathbf{c}^5, \quad {}_{-\infty}\mathbf{q}^3 = {}_2\mathbf{c}^6, \quad {}_0\mathbf{q}^4 = {}_1\mathbf{c}^1, \quad {}_0\mathbf{q}^5 = {}_1\mathbf{c}^2, \quad {}_0\mathbf{q}^6 = {}_2\mathbf{c}^4$$

determine a basis of C_1 adapted to the preimage filtration, where we simply reorder the almost-adapted basis $\mathbf{c}^1, \mathbf{c}^2, \ldots, \mathbf{c}^{\dim C_1}$ with respect to the level function ρ. The permutation matrix $\mathbf{Y}$ below describes these equations. A neat way to construct this matrix is by taking the reduced matrix $\underline{\mathbf{D}}_1$, removing any rows that have no (column) pivot entry, and *row-completing* it to a square permutation matrix by adding rows at the top (appropriately-chosen from the identity matrix). The elements $\mathbf{q}^1, \mathbf{q}^2, \ldots, \mathbf{q}^{\dim C_1}$ can be represented by the intrinsic basis by multiplying the inverse of $\mathbf{Y}$ on the left by $\mathbf{V}_1$.

$$\mathbf{Y} = \begin{array}{c} \\ {}_{-\infty}\mathbf{q}^1 \\ {}_{-\infty}\mathbf{q}^2 \\ {}_{-\infty}\mathbf{q}^3 \\ {}_0\mathbf{q}^4 \\ {}_0\mathbf{q}^5 \\ {}_0\mathbf{q}^6 \end{array} \begin{array}{c} \begin{array}{cccccc} {}_1\mathbf{c}^1 & {}_1\mathbf{c}^2 & {}_1\mathbf{c}^3 & {}_2\mathbf{c}^4 & {}_2\mathbf{c}^5 & {}_2\mathbf{c}^6 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array} \qquad \mathbf{V}_1\mathbf{Y}^T = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} \begin{array}{cccccc} {}_{-\infty}\mathbf{q}^1 & {}_{-\infty}\mathbf{q}^2 & {}_{-\infty}\mathbf{q}^3 & {}_0\mathbf{q}^4 & {}_0\mathbf{q}^5 & {}_0\mathbf{q}^6 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Matrix completion techniques shall be revisited later in section 4.3.1.

PH1b: *Express $\delta_2 : C_2 \to Z$ as a matrix Δ_2 relative to bases adapted to the canonical filtrations: the intrinsic basis of C_2 and the basis $\mathbf{p}^k$ of Z found in* **PH1a**.

Recall that δ_2 is a unique linear map that satisfies the equation $i_1 \circ \delta_2 = \partial_2$, where i_1 is the inclusion $Z \subseteq C_1$. Hence, the goal is to find the unique solution to the matrix equation:

$$\mathbf{P} \cdot \Delta_2 = \mathbf{D}_2$$

where matrix $\mathbf{D}_2$, representing the differential operator ∂_2 with respect to intrinsic bases of C_1 and C_2, is constructed with reference to Figure 4.1:

$$\mathbf{D}_2 = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} {}_2acd \quad {}_3abd \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ -1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}.$$

This constitutes a system of linear equations that is conventionally solved by Gaussian elimination. However, the calculation is made simple by the fact that $\mathbf{P}$ is already reduced. In particular, it suffices to solve the equation

$$\mathbf{P}' \cdot \Delta_2 = \mathbf{D}_2'$$

where the matrices

$$\mathbf{P}' = \begin{array}{c} \\ {}_1bc \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} {}_1\mathbf{p}^1 \quad {}_2\mathbf{p}^2 \quad {}_2\mathbf{p}^3 \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \qquad \mathbf{D}_2' = \begin{array}{c} \\ {}_1bc \\ {}_2bd \\ {}_2cd \end{array} \begin{array}{c} {}_2acd \quad {}_3abd \\ \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

are obtained from $\mathbf{P}$ and $\mathbf{D}_2$ respectively by removing like-indexed rows which contain no (column) pivots in $\mathbf{P}$; here, those would be rows 1, 2, and 4 from both. But this makes $\mathbf{P}'$ an invertible matrix, so the solution is:

$$\Delta_2 = (\mathbf{P}')^{-1}\,\mathbf{D}_2' = \begin{array}{c} \\ {}_1\mathbf{p}^1 \\ {}_2\mathbf{p}^2 \\ {}_2\mathbf{p}^3 \end{array} \begin{array}{c} \begin{array}{cc} {}_2acd & {}_3abd \end{array} \\ \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{array}\right] \end{array}.$$

This completes Step **PH1**, and we proceed to Step **PH2**.

PH2: *Reduce the matrix Δ_2 to construct a basis $\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^{\dim Z}$ of $Z \subseteq C_1$ that is simultaneously: adapted to the restriction ${}_\bullet B$ of the range-canonical filtration and almost-adapted to the restriction ${}_\bullet Z$ of the canonical filtration. Then read off the barcodes and persistent homology cycles.*

Generally, matrix reduction of Δ_2 in this step will produce the reduced matrix $\underline{\Delta}_2$ and a upper triangular change-of-basis matrix $\mathbf{V}_2$, such that $\underline{\Delta}_2 = \Delta_2\,\mathbf{V}_2$; the change-of-basis V_2 specifies a basis ${}_\lambda\mathbf{a}^k$ of C_2, where λ denotes its ${}_\bullet C_2$-level. In this case, reduction is trivial:

$$\underline{\Delta}_2 = \begin{array}{c} \\ {}_1\mathbf{p}^1 \\ {}_2\mathbf{p}^2 \\ {}_2\mathbf{p}^3 \end{array} \begin{array}{c} \begin{array}{cc} {}_2\mathbf{a}^1 & {}_3\mathbf{a}^2 \end{array} \\ \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{array}\right] \end{array} \qquad \mathbf{V}_2 = \begin{array}{c} \\ {}_2acd \\ {}_3abd \end{array} \begin{array}{c} \begin{array}{cc} {}_2\mathbf{a}^1 & {}_3\mathbf{a}^2 \end{array} \\ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \end{array}$$

The basis $\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^{\dim Z}$ of Z should be adapted to the range-canonical filtration of C_2; this is equivalent to finding a basis of Z adapted to range filtration induced by δ_2 of the canonical filtration. This is neatly done by working with the reduced matrix $\underline{\Delta}_2$; namely, we *complete* $\underline{\Delta}_2$ to a square matrix *by columns.* After removing any zero columns (none in this case), we append columns (appropriately-chosen from the identity matrix) to get an invertible square matrix $\mathbf{S}$. The columns of $\mathbf{S}$ that came from $\underline{\Delta}_2$ retain their levels (in

${}_\bullet C_2$), while the columns that were appended are assigned level ∞.

$$\mathbf{S} = \begin{array}{c c} & \begin{array}{ccc} {}_2\mathbf{b}^1 & {}_3\mathbf{b}^2 & {}_\infty\mathbf{b}^3 \end{array} \\ \begin{array}{c} {}_1\mathbf{p}^1 \\ {}_2\mathbf{p}^2 \\ {}_2\mathbf{p}^3 \end{array} & \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right] \end{array} \qquad \mathbf{N} = \begin{array}{c c} & \begin{array}{ccc} {}_2\mathbf{b}^1 & {}_3\mathbf{b}^2 & {}_\infty\mathbf{b}^3 \end{array} \\ \begin{array}{c} {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} & \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right] \end{array}$$

Also exhibited above is the pivot matrix $\mathbf{N}$ of $\mathbf{S}$; that is, every column pivot of $\mathbf{S}$ is replaced with 1 and every other (nonzero) entry with 0. This makes $\mathbf{N}$ a permutation matrix because $\mathbf{S}$ is reduced and invertible, and it represents the identity map $1_Z : Z \to Z$ with respect to basis $\mathbf{b}^k$ adapted to ${}_\bullet B$ and some basis $\mathbf{z}^k$ (given by the columns of $\mathbf{S}$) adapted to ${}_\bullet Z$. Here it happens that $\mathbf{N} = \mathbf{S}$ and $\mathbf{z}^k = \mathbf{p}^k$, although this is generally not the case.

The significance of the permutation matrix $\mathbf{N}$ is that it cleanly encodes the barcodes of the simplicial complex. Following Theorem 4.11, because the basis $\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^{\dim Z}$ of Z is simultaneously almost-adapted to ${}_\bullet B$ and ${}_\bullet Z$ the interval indecomposables of the complex can be represented by $\langle {}_\bullet[\mathbf{b}^k] \rangle$. It remains to measure the ${}_\bullet Z$-level of each element $\mathbf{b}^k$, which is exactly what $\mathbf{N}$ does: letting ζ be the level function of filtration ${}_\bullet Z$, we see from the pivot entries of $\mathbf{N}$:

$\zeta({}_2\mathbf{b}^1) = 2$, yielding interval $[2, 2)$ with cycle ${}_2\mathbf{b}^1 = ac - ad + cd$;

$\zeta({}_3\mathbf{b}^2) = 2$, yielding interval $[2, 3)$ with cycle ${}_3\mathbf{b}^2 = ab - ad + bd$;

$\zeta({}_\infty\mathbf{b}^3) = 1$, yielding interval $[1, \infty)$ with cycle ${}_\infty\mathbf{b}^3 = ab - ac + bc$.

The corresponding persistent cycles (written as combination of the intrinsic basis of C_1) themselves were determined by transforming $\mathbf{b}^k$ to basis $\mathbf{p}^k$, which is quickly done by calculating the matrix $\mathbf{B} = \mathbf{PS}$. Concurrently, we can calculate the matrix $\mathbf{Z} = \mathbf{BN}^T$ representing the basis $\mathbf{z}^k$ in terms of the intrinsic basis;

however, this step is not necessary.

$$\mathbf{B} = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \!\!\begin{array}{c} \begin{array}{ccc} {}_2\mathbf{b}^1 & {}_3\mathbf{b}^2 & {}_\infty\mathbf{b}^3 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array} \qquad \mathbf{Z} = \begin{array}{c} \\ {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} \!\!\begin{array}{c} \begin{array}{ccc} {}_1\mathbf{z}^1 & {}_2\mathbf{z}^2 & {}_2\mathbf{z}^3 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The algorithm deduces $\zeta(\mathbf{b}^j)$ directly from $\mathbf{B}$, as the prescript of the intrinsic basis element corresponding the pivot entry of the column. We choose to take the extra step of constructing the permutation matrix $\mathbf{N}$ in order to retain the flexibility to work with filtrations other than canonical filtrations.

This completes Step **PH2** and thus finishes the computation of persistent homology, which is summarized below.

$$H_1({}_\bullet C_\bullet) \simeq [1, \infty) \oplus [2, 3)$$

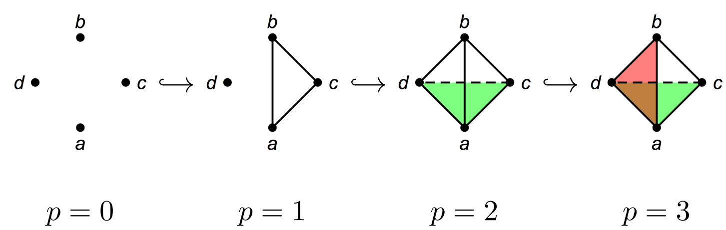
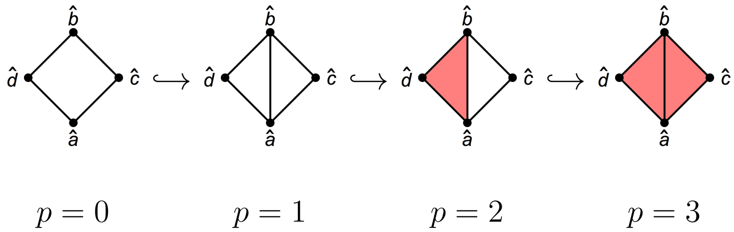


Figure 4.2: The barcode invariants corresponding to the decomposition of the persistent homology $H_1({}_\bullet C_\bullet)$ as computed in this section.

4.2.2 Second computation

We repeat the variant persistent homology algorithm to compute the dimension $n = 1$ persistent homology of the filtered simplicial complex shown in Figure 4.3. Because the steps are identical, less detail is presented for brevity.



$$\hat{C}_0 = \left\langle {}_0\hat{a},\, {}_0\hat{b},\, {}_0\hat{c},\, {}_0\hat{d} \right\rangle$$
$$\hat{C}_1 = \left\langle {}_0\hat{a}\hat{c},\, {}_0\hat{a}\hat{d},\, {}_0\hat{b}\hat{c},\, {}_0\hat{b}\hat{d},\, {}_1\hat{a}\hat{b} \right\rangle$$
$$\hat{C}_2 = \left\langle {}_2\hat{a}\hat{b}\hat{d},\, {}_3\hat{a}\hat{b}\hat{c} \right\rangle$$

Figure 4.3: Shown on the right are the intrinsic bases of $\hat{C}_0$, $\hat{C}_1$, $\hat{C}_2$; the prescript on the elements denotes their level in the canonical filtration.

PH1a: *Construct a basis $\hat{\mathbf{p}}^1, \hat{\mathbf{p}}^2, \ldots, \hat{\mathbf{p}}^{\dim \hat{Z}}$ of $\hat{Z} \subseteq \hat{C}_1$ that is adapted to the restriction ${}_\bullet\hat{Z}$ of the canonical filtration on $\hat{C}_1$.*

Construct matrix $\hat{\mathbf{D}}_1$ with reference to Figure 4.3:

$$\hat{\mathbf{D}}_1 = \begin{array}{c} \\ {}_0\hat{a} \\ {}_0\hat{b} \\ {}_0\hat{c} \\ {}_0\hat{d} \end{array} \begin{array}{c} \begin{array}{ccccc} {}_0\hat{a}\hat{c} & {}_0\hat{a}\hat{d} & {}_0\hat{b}\hat{c} & {}_0\hat{b}\hat{d} & {}_1\hat{a}\hat{b} \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & 0 & -1 \\ 0 & 0 & -1 & -1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Reduce $\hat{\mathbf{D}}_1$ according to the procedure in section 4.1.1. This construct a basis $\hat{\mathbf{c}}^1, \hat{\mathbf{c}}^2, \ldots, \hat{\mathbf{c}}^{\dim C_1}$ of $\hat{C}_1$ that is simultaneously adapted to the canonical filtration ${}_\bullet\hat{C}_1$ and almost-adapted to the preimage-canonical filtration ${}_\bullet\hat{P}_1$.

$$\underline{\hat{\mathbf{D}}}_1 = \begin{array}{c} \\ {}_0\hat{a} \\ {}_0\hat{b} \\ {}_0\hat{c} \\ {}_0\hat{d} \end{array} \begin{array}{c} \begin{array}{ccccc} {}_0\hat{\mathbf{c}}^1 & {}_0\hat{\mathbf{c}}^2 & {}_0\hat{\mathbf{c}}^3 & {}_0\hat{\mathbf{c}}^4 & {}_1\hat{\mathbf{c}}^5 \end{array} \\ \begin{bmatrix} -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & \mathbf{-1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 \end{bmatrix} \end{array} \qquad \hat{\mathbf{V}}_1 = \begin{array}{c} \\ {}_0\hat{a}\hat{c} \\ {}_0\hat{a}\hat{d} \\ {}_0\hat{b}\hat{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{ccccc} {}_0\hat{\mathbf{c}}^1 & {}_0\hat{\mathbf{c}}^2 & {}_0\hat{\mathbf{c}}^3 & {}_0\hat{\mathbf{c}}^4 & {}_1\hat{\mathbf{c}}^5 \end{array} \\ \begin{bmatrix} 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Let $\hat{\rho}$ be the level function on $\hat{C}_1$ with respect to filtration ${}_\bullet\hat{P}_1$. For each basis element $\hat{\mathbf{c}}^1, \hat{\mathbf{c}}^2, \ldots, \hat{\mathbf{c}}^{\dim \hat{C}_1}$, read off the ${}_\bullet\hat{P}_1$-level from the rows of $\underline{\hat{\mathbf{D}}}_1$:

$$\hat{\rho}({}_0\hat{\mathbf{c}}^1) = 0, \quad \hat{\rho}({}_0\hat{\mathbf{c}}^2) = 0, \quad \hat{\rho}({}_0\hat{\mathbf{c}}^3) = 0, \quad \hat{\rho}({}_0\hat{\mathbf{c}}^4) = -\infty, \quad \hat{\rho}({}_1\hat{\mathbf{c}}^5) = -\infty.$$

Discarding basis elements that do not have ${}_{\bullet}\hat{P}_1$-level equal to $-\infty$ yields a basis $\hat{\mathbf{p}}^1, \ldots, \hat{\mathbf{p}}^{\dim \hat{Z}}$ of $\hat{Z} = {}_{-\infty}\hat{P}_1$:

$$ {}_{0}\hat{\mathbf{p}}^1 = {}_{0}\hat{\mathbf{c}}^4, \quad {}_{1}\hat{\mathbf{p}}^2 = {}_{1}\hat{\mathbf{c}}^5. $$

The prescript λ of ${}_{\lambda}\hat{\mathbf{p}}^k$ denotes its ${}_{\bullet}\hat{C}_1$-level, same as of ${}_{\lambda}\hat{\mathbf{c}}$ (adapted to ${}_{\bullet}\hat{C}_1$). We otherwise discard unused columns of the matrix $\hat{\mathbf{V}}_1$, yielding a matrix $\hat{\mathbf{P}}$ that expresses each basis element $\hat{\mathbf{p}}$ in terms of the intrinsic basis of $\hat{C}_1$:

$$ \hat{\mathbf{P}} = \begin{array}{c} \\ {}_{0}\hat{a}\hat{c} \\ {}_{0}\hat{a}\hat{d} \\ {}_{0}\hat{b}\hat{c} \\ {}_{0}\hat{b}\hat{d} \\ {}_{1}\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_{0}\hat{\mathbf{p}}^1 & {}_{1}\hat{\mathbf{p}}^2 \end{array} \\ \begin{bmatrix} 1 & -1 \\ -1 & 0 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} $$

This basis is adapted to the canonical filtration ${}_{\bullet}\hat{Z}$ on $\hat{Z}$.

PH1b: *Express $\hat{\delta}_2 : \hat{C}_2 \to \hat{Z}$ as a matrix $\hat{\Delta}_2$ relative to bases adapted to the canonical filtrations: the intrinsic basis of $\hat{C}_2$ and the basis $\hat{\mathbf{p}}^k$ of $\hat{Z}$ found in* **PH1a**.

The goal is to find the unique solution to the matrix equation:

$$ \hat{\mathbf{P}} \cdot \hat{\Delta}_2 = \hat{\mathbf{D}}_2 $$

where matrix $\hat{\mathbf{D}}_2$, representing the differential operator $\hat{\partial}_2$ with respect to intrinsic bases of $\hat{C}_1$ and $\hat{C}_2$, is constructed with reference to Figure 4.3:

$$ \hat{\mathbf{D}}_2 = \begin{array}{c} \\ {}_{0}\hat{a}\hat{c} \\ {}_{0}\hat{a}\hat{d} \\ {}_{0}\hat{b}\hat{c} \\ {}_{0}\hat{b}\hat{d} \\ {}_{1}\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_{2}\hat{a}\hat{b}\hat{d} & {}_{3}\hat{a}\hat{b}\hat{c} \end{array} \\ \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \end{array} $$

Since $\hat{\mathbf{P}}$ is already reduced, it suffices to solve the equation

$$\hat{\mathbf{P}}' \cdot \hat{\Delta}_2 = \hat{\mathbf{D}}_2'$$

where the matrices

$$\hat{\mathbf{P}}' = \begin{array}{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_0\hat{\mathbf{p}}^1 & {}_1\hat{\mathbf{p}}^2 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \quad \hat{\mathbf{D}}_2' = \begin{array}{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{a}\hat{b}\hat{d} & {}_3\hat{a}\hat{b}\hat{c} \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \end{array}$$

are obtained from $\hat{\mathbf{P}}$ and $\hat{\mathbf{D}}_2$ respectively by removing like-indexed rows which contain no (column) pivots in $\hat{\mathbf{P}}$; here, those would be rows 1, 2, and 3 from both. Since $\hat{\mathbf{P}}'$ an invertible matrix, the solution is:

$$\hat{\Delta}_2 = \left(\hat{\mathbf{P}}'\right)^{-1} \hat{\mathbf{D}}_2' = \begin{array}{c} \\ {}_0\hat{\mathbf{p}}^1 \\ {}_1\hat{\mathbf{p}}^2 \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{a}\hat{b}\hat{d} & {}_3\hat{a}\hat{b}\hat{c} \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \end{array}$$

This completes Step **PH1**, and we proceed to Step **PH2**.

PH2: *Reduce the matrix $\hat{\Delta}_2$ to construct a basis $\hat{\mathbf{b}}^1, \hat{\mathbf{b}}^2, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$ of $\hat{Z} \subseteq \hat{C}_1$ that is simultaneously: adapted to the restriction ${}_\bullet\hat{B}$ of the range-canonical filtration and almost-adapted to the restriction ${}_\bullet\hat{Z}$ of the canonical filtration. Then read off the barcodes and persistent homology cycles.*

Matrix reduction of $\hat{\Delta}_2$ yields the reduced matrix $\underline{\hat{\Delta}}_2$ and the upper triangular change of basis matrix $\hat{\mathbf{V}}_2$, where $\underline{\hat{\Delta}}_2 = \hat{\Delta}_2 \, \hat{\mathbf{V}}_2$:

$$\underline{\hat{\Delta}}_2 = \begin{array}{c} \\ {}_0\hat{\mathbf{p}}^1 \\ {}_1\hat{\mathbf{p}}^2 \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{\mathbf{a}}^1 & {}_3\hat{\mathbf{a}}^2 \end{array} \\ \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \end{array} \quad \hat{\mathbf{V}}_2 = \begin{array}{c} \\ {}_2\hat{a}\hat{b}\hat{d} \\ {}_3\hat{a}\hat{b}\hat{c} \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{\mathbf{a}}^1 & {}_3\hat{\mathbf{a}}^2 \end{array} \\ \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \end{array}$$

To calculate $\hat{\mathbf{b}}^1, \hat{\mathbf{b}}^2, \ldots, \hat{\mathbf{b}}^{\dim Z}$, we find a basis of $\hat{Z}$ adapted to range filtration induced by $\hat{\delta}_2$ of the canonical filtration on $\hat{C}_2$. Complete $\underline{\hat{\Delta}}_2$ to a square matrix by columns: remove any zero columns (none in this case), and append columns (appropriately-chosen from the identity matrix) to get an invertible square matrix $\hat{\mathbf{S}}$. Let $\hat{\mathbf{N}}$ be the pivot matrix of $\mathbf{S}$: all nonzero entries of $\hat{\mathbf{S}}$ are set to 0, except for the (column) pivots which are set to 1.

The columns of $\hat{\mathbf{S}}$ that came from $\underline{\hat{\Delta}}_2$ retain their levels (in ${}_\bullet\hat{C}_2$), while the columns that were appended are assigned to level ∞. The resulting basis $\hat{\mathbf{b}}^1, \hat{\mathbf{b}}^2, \ldots, \mathbf{b}^{\dim Z}$ of $\hat{Z}$ is then adapted to the range filtration ${}_\bullet\hat{B}$, and simultaneously almost-adapted to the canonical filtration ${}_\bullet\hat{Z}$; on the other hand, basis $\hat{\mathbf{z}}^1, \hat{\mathbf{z}}^2, \ldots, \hat{\mathbf{z}}^{\dim \hat{Z}}$ of $\hat{Z}$ is adapted ${}_\bullet\hat{Z}$.

$$\hat{\mathbf{S}} = \begin{array}{c} \\ {}_0\hat{\mathbf{p}}^1 \\ {}_1\hat{\mathbf{p}}^2 \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \end{array} \\ \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \end{array} \qquad \hat{\mathbf{N}} = \begin{array}{c} \\ {}_0\hat{\mathbf{z}}^1 \\ {}_1\hat{\mathbf{z}}^2 \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

The matrices $\hat{\mathbf{B}} = \hat{\mathbf{P}}\hat{\mathbf{S}}$ and $\hat{\mathbf{Z}} = \hat{\mathbf{B}}\hat{\mathbf{N}}^T$ represent their respective adapted bases $\hat{\mathbf{b}}^1, \hat{\mathbf{b}}^2, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$ and $\hat{\mathbf{z}}^1, \hat{\mathbf{z}}^2, \ldots, \hat{\mathbf{z}}^{\dim \hat{Z}}$ of $\hat{Z}$ as linear combinations of the intrinsic basis elements in $\hat{C}_1$. Technically, calculating matrix $\hat{\mathbf{Z}}$ is unnecessary.

$$\hat{\mathbf{B}} = \begin{array}{c} \\ {}_0\hat{a}\hat{c} \\ {}_0\hat{a}\hat{d} \\ {}_0\hat{b}\hat{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \end{array} \\ \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \end{array} \qquad \hat{\mathbf{Z}} = \begin{array}{c} \\ {}_0\hat{a}\hat{c} \\ {}_0\hat{a}\hat{d} \\ {}_0\hat{b}\hat{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{array}{cc} {}_0\hat{\mathbf{z}}^1 & {}_1\hat{\mathbf{z}}^2 \end{array} \\ \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \end{array}$$

Following Theorem 4.11, the interval indecomposables of the complex can be represented by $\langle {}_\bullet[\hat{\mathbf{b}}^k] \rangle$. The ${}_\bullet\hat{Z}$-level $\hat{\zeta}$ of each basis element $\hat{\mathbf{b}}^1, \hat{\mathbf{b}}^2, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$ is the level of the pivot entry of the corresponding column of $\hat{\mathbf{N}}$:

$\hat{\zeta}({}_2\hat{\mathbf{b}}^1) = 1$, yielding interval $[1, 2)$ with cycle ${}_2\hat{\mathbf{b}}^1 = -\hat{a}\hat{d} + \hat{b}\hat{d} + \hat{a}\hat{b}$;

$\hat{\zeta}({}_3\hat{\mathbf{b}}^2) = 0$, yielding interval $[0, 3)$ with cycle ${}_3\hat{\mathbf{b}}^2 = -\hat{a}\hat{c} + \hat{a}\hat{d} + \hat{b}\hat{c} - \hat{b}\hat{d}$.

This completes Step **PH2** and thus the computation of persistent homology.

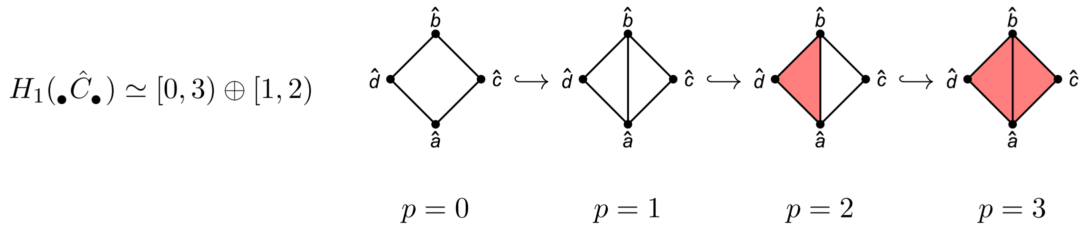


Figure 4.4: The barcode invariants corresponding to the decomposition of the persistent homology $H_1({}_\bullet\hat{C}_\bullet)$ as computed in this section.

4.3 Homology of Filtered Chain Maps

The framework of our algorithm has an advantage that it can be readily adapted to study (co)kernels and (co)images of filtered chain maps. This is demonstrated on the example in Figure 4.5, originally from [44]; the interval decompositions of ${}_\bullet C_\bullet$ and ${}_\bullet\hat{C}_\bullet$ were already found in section 4.2.

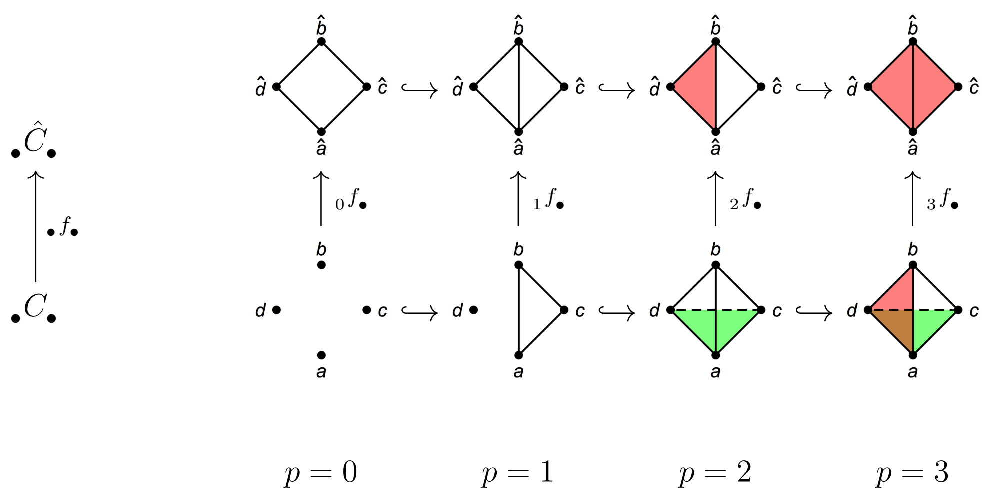


Figure 4.5: The morphism ${}_\bullet f_\bullet$ will be called the *filtered folding map*, and is defined by the action on 0-simplices given by: $a \mapsto \hat{a}$, $b \mapsto \hat{b}$, $c \mapsto \hat{c}$, $d \mapsto \hat{c}$.

The practical importance of knowing the kernel and cokernel was already seen in section 3.3.3; namely, in [4] the authors show that the ε-triviality of the $\ker H_n({}_\bullet f_\bullet)$ and $\operatorname{coker} H_n({}_\bullet f_\bullet)$ leads to ε-tight bounds on the endpoints of the matched intervals to the coimage $\operatorname{coim} H_n({}_\bullet f_\bullet)$ and image $\operatorname{im} H_n({}_\bullet f_\bullet)$, thus establishing the algebraic stability theorem. Knowing the interval decomposition of $\operatorname{coim} H_n({}_\bullet f_\bullet) \simeq \operatorname{im} H_n({}_\bullet f_\bullet)$ is itself a crucial component in some other applications, such as computing the localized persistent homology [61].

One of the earliest attempts to provide an explicit algorithm for calculating these was conferred in [16] by Cohen-Steiner, Edelsbrunner, Harer, and Morozov. In this setting, $L \subset K$ are filtered simplicial complexes, each having a level function $f : K \to \mathbb{R}$ and $g : L \to \mathbb{R}$ such that $f \leqslant g$ on L; although by means of a mapping cylinder construction, the authors were able to reduce their problem to the case where $f = g$ on L. This an important historical precedent for the manual presented further, and their algorithm will briefly be analyzed in the Appendix. However, this algorithm relies on the rigidity of the inclusion $L \subseteq K$ to arrive at a consistent basis between these complexes on which the rigorous matrix constructions are based. On the other hand, the algorithm presented further functions in a basis-invariant way to provide such a consistent basis, which easily allows to generalize the results to the case where $L \to K$ is not monic. Ultimately, a more flexible picture began to develop with the presentation in [57] of a theory on persistent (co)kernels and (co)images based on their categorical foundations, where the main computational device used is the graded Smith Normal Form representation of inclusion maps. We continue this development by suggesting a uniform language for expressing all these objects' persistent homology.

Our variant algorithm demonstrated ahead is founded on the following reasoning, familiar to anyone with a background in linear algebra. Recall that the kernel and image of a morphism $\phi : P \to \hat{P}$ of R-modules are given by $\phi^{-1}(\mathbf{0}) \subseteq P$ and $\phi(P) \subseteq \hat{P}$, respectively; their quotients are the coimage $P \to P/\phi^{-1}(\mathbf{0})$ and cokernel $\hat{P} \to \hat{P}/\phi(P)$. In our setup we are working the quotient modules $P = Z/B$ and $\hat{P} = \hat{Z}/\hat{B}$ with $\phi = H_n(f)$ being the quotient morphism induced by $f : Z \to \hat{Z}$; the kernel and image of ϕ are represented by inclusions $M = K \cap Z \subseteq Z$ (modulo B) and $\hat{M} = \hat{K} + \hat{B} \subseteq Z$ (modulo $\hat{B}$), where

$$K = f^{-1}(B) \qquad \text{and} \qquad \hat{K} = f(\hat{Z})$$

The third isomorphism theorem from algebra then allows us to represent the coimage and cokernel as $P \to Z/M$ and $\hat{P} \to \hat{Z}/\hat{M}$. This is further underscored by the nature of the canonical short exact sequences associated to the morphism $H_n f$, which Propositions 4.16 and 4.17 will demonstrate for the case of filtered complexes.

As before, we will use a basis notation that is elaborated by Remark 4.14.

4.3.1 Restriction to a Map of n-Cycles

Because of our focus on homology of dimension $n = 1$, we begin by expressing the linear map $f_1 : C_1 \to \hat{C}_1$ in Figure 4.5 as a matrix $\mathbf{F}_1$ with respect to the intrinsic bases on C_1 and $\hat{C}_1$:

$$\mathbf{F}_1 = \begin{array}{c} \\ {}_0\hat{a}\hat{c} \\ {}_0\hat{a}\hat{d} \\ {}_0\hat{b}\hat{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} \begin{array}{c} \begin{matrix} {}_1ab & {}_1ac & {}_1bc & {}_2ad & {}_2bd & {}_2cd \end{matrix} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Going further, the property $f_1(Z) \subseteq \hat{Z}$ allows us to restrict to $f : Z \to \hat{Z}$. Our immediate goal is to express that as a matrix relative to previously-constructed bases.

Specifically, the morphism will be represented by a matrix $\mathbf{F}$ relative to bases $\mathbf{z}^1, \ldots, \mathbf{z}^{\dim Z}$ of Z and $\hat{\mathbf{b}}^1, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$ of $\hat{Z}$. Recall from section 4.2 that the matrices $\mathbf{Z}$ and $\hat{\mathbf{B}}$ each express these elements in terms of the intrinsic bases on C_1 and $\hat{C}_1$, shown below:

$$\mathbf{Z} = \begin{array}{cc} & \begin{array}{ccc} {}_1\mathbf{z}^1 & {}_2\mathbf{z}^2 & {}_2\mathbf{z}^3 \end{array} \\ \begin{array}{c} {}_1ab \\ {}_1ac \\ {}_1bc \\ {}_2ad \\ {}_2bd \\ {}_2cd \end{array} & \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \qquad \hat{\mathbf{B}} = \begin{array}{cc} & \begin{array}{cc} {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \end{array} \\ \begin{array}{c} {}_0\hat{a}\hat{c} \\ {}_0\hat{a}\hat{d} \\ {}_0\hat{b}\hat{c} \\ {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} & \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 0 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \end{array}$$

Then the sought-for matrix $\mathbf{F}$ can be found as the the unique solution of the following matrix equation:

$$\hat{\mathbf{B}} \cdot \mathbf{F} = \mathbf{F}_1 \cdot \mathbf{Z}.$$

Once again, we are presented with a system of linear equations that can be quickly solved by using the fact that $\hat{\mathbf{B}}$ is reduced. Remove all the like-indexed rows of $\hat{\mathbf{B}}$ and $\mathbf{F}_1$ that do not contain a column pivot in $\hat{\mathbf{B}}$; here, that constitutes rows 1-3 of both matrices. Then the original equation reduces to $\hat{\mathbf{B}}'\mathbf{F} = \mathbf{F}_1'\mathbf{Z}$, where:

$$\hat{\mathbf{B}}' = \begin{array}{cc} & \begin{array}{cc} {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \end{array} \\ \begin{array}{c} {}_0\hat{b}\hat{d} \\ {}_1\hat{a}\hat{b} \end{array} & \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \end{array} \qquad \mathbf{F}_1' = \begin{array}{cc} & \begin{array}{cccccc} ab & ac & bc & ad & bd & cd \end{array} \\ \begin{array}{c} \hat{b}\hat{d} \\ \hat{a}\hat{b} \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

The invertibility of $\hat{\mathbf{B}}'$ then allows us to finally solve for $\mathbf{F}$ as:

$$\mathbf{F} = \left(\hat{\mathbf{B}}'\right)^{-1} \mathbf{F}_1'\mathbf{Z} = \begin{array}{cc} & \begin{array}{ccc} {}_1\mathbf{z}^1 & {}_2\mathbf{z}^2 & {}_2\mathbf{z}^3 \end{array} \\ \begin{array}{c} {}_2\hat{\mathbf{b}}^1 \\ {}_3\hat{\mathbf{b}}^2 \end{array} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

Now, we apply the usual reduction procedure on matrix $\mathbf{F}$ to obtain $\underline{\mathbf{F}} = \mathbf{F}\,\mathbf{W}$, where $\underline{\mathbf{F}}$ is a reduced matrix and $\mathbf{W}$ is a upper uni-triangular change-of-basis matrix, shown below:

$$\underline{\mathbf{F}} = \begin{array}{c} \\ {}_2\hat{\mathbf{b}}^1 \\ {}_3\hat{\mathbf{b}}^2 \end{array} \begin{array}{c} \begin{array}{ccc} {}_1\mathbf{z}_+^1 & {}_2\mathbf{z}_+^2 & {}_2\mathbf{z}_+^3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array} \qquad \mathbf{W} = \begin{array}{c} \\ {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} \begin{array}{c} \begin{array}{ccc} {}_1\mathbf{z}_+^1 & {}_2\mathbf{z}_+^2 & {}_2\mathbf{z}_+^3 \end{array} \\ \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The matrix $\underline{\mathbf{F}}$ represents the linear map $f : Z \to \hat{Z}$ with respect to a basis $\mathbf{z}_+^1, \ldots, \mathbf{z}_+^{\dim Z}$ of Z (adapted to the ${}_\bullet Z$ filtration) and basis $\hat{\mathbf{b}}^1, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$ of $\hat{Z}$ (adapted to the ${}_\bullet \hat{B}$ filtration). Recalling Lemma 4.7, we can use $\underline{\mathbf{F}}$ to compute bases adapted to the range filtration ${}_\bullet \hat{K} := f({}_\bullet Z)$ and the preimage filtration ${}_\bullet K := f^{-1}({}_\bullet \hat{B})$.

- Remove any zero columns from $\underline{\mathbf{F}}$, and append columns (appropriately-chosen from the identity matrix) to get an invertible square matrix $\hat{\mathbf{K}}$. the kept columns retain their level in ${}_\bullet\mathbf{Z}$, while the appended columns get level ∞.

- Remove any rows from the pivot matrix of $\underline{\mathbf{F}}$ that do not contain a (column) pivot, and prepend columns (appropriately-chosen from the identity matrix) to get an invertible square matrix $\mathbf{K}_+$; the kept rows retain their level in ${}_\bullet\hat{\mathbf{B}}$, while the prepended rows get level $-\infty$. Then calculate $\mathbf{K} = \mathbf{W}\mathbf{K}^T$.

We thus procure:

$$\hat{\mathbf{K}} = \begin{array}{c} \\ {}_2\hat{\mathbf{b}}^1 \\ {}_3\hat{\mathbf{b}}^2 \end{array} \begin{array}{c} \begin{array}{cc} {}_1\hat{\mathbf{k}}^1 & {}_\infty\hat{\mathbf{k}}^2 \end{array} \\ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array} \qquad \mathbf{K} = \begin{array}{c} \\ {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} \begin{array}{c} \begin{array}{ccc} {}_{-\infty}\mathbf{k}^1 & {}_{-\infty}\mathbf{k}^2 & {}_3\mathbf{k}^3 \end{array} \\ \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

The matrix $\hat{\mathbf{K}}$ represents some basis $\hat{\mathbf{k}}^1, \ldots, \hat{\mathbf{k}}^{\dim \hat{Z}}$ relative to basis $\hat{\mathbf{b}}^1, \ldots, \hat{\mathbf{b}}^{\dim \hat{Z}}$; the basis $\hat{\mathbf{k}}^j$ is adapted to the range filtration $f({}_\bullet Z)$ by construction, and is almost-adapted to filtration ${}_\bullet \hat{B}$ (since $\hat{\mathbf{K}}$ is reduced). The matrix $\mathbf{K}$ represents some basis

$\mathbf{k}^1, \ldots, \mathbf{k}^{\dim Z}$ relative to the basis $\mathbf{z}^1, \ldots, \mathbf{z}^{\dim Z}$; the basis $\mathbf{k}^j$ is adapted to the preimage filtration $f^{-1}({}_{\bullet}\hat{B})$ by construction, and is almost-adapted to filtration ${}_{\bullet}Z$.

We can draw a parallel between the utility of the matrix $\mathbf{F}$ and the matrices $\Delta_2, \hat{\Delta}_2$ from section 4.2. In both cases we represent a linear operator in matrix form relative to a choice of bases on the source and target reference vector spaces, so as to prepare for finding an almost-adapted basis for either: its induced range filtration and the chosen filtration on the target space, or its induced preimage filtration and the chosen filtration on the source space. In this case, we are able to achieve both.

4.3.2 Image and Cokernel

Following the discussion in the introduction, we have the following proposition that allows us to calculate the image and cokernel of the morphism ${}_{\bullet}f_{\bullet}$.

Proposition 4.16 (see [44]). *Let* ${}_{\bullet}\hat{M} := {}_{\bullet}\hat{B} + {}_{\bullet}\hat{K}$ *be the sum of filtrations* ${}_{\bullet}\hat{B}$ *and* ${}_{\bullet}\hat{K} := f({}_{\bullet}Z)$ *of* $\hat{Z}$*. Then* ${}_{\bullet}\hat{B} \subseteq {}_{\bullet}\hat{M} \subseteq {}_{\bullet}\hat{Z}$ *as filtrations of* $\hat{Z}$*, and:*

1. *The map* ${}_{\bullet}\hat{M}/{}_{\bullet}\hat{B} \to {}_{\bullet}\hat{Z}/{}_{\bullet}\hat{B}$ *induced from the identity* $\hat{Z} \to \hat{Z}$ *is the categorical image,* $\operatorname{im} H_n({}_{\bullet}f_{\bullet}) \to H_n({}_{\bullet}\hat{C}_{\bullet})$.
2. *The map* ${}_{\bullet}\hat{Z}/{}_{\bullet}\hat{B} \to {}_{\bullet}\hat{Z}/{}_{\bullet}\hat{M}$ *induced from the identity* $\hat{Z} \to \hat{Z}$ *is the categorical cokernel,* $H_n({}_{\bullet}\hat{C}_{\bullet}) \to \operatorname{coker} H_n({}_{\bullet}f_{\bullet})$.

Proof. This readily follows from the fact that these relations hold at every $s \in \mathbb{Z}$. ∎

The following diagram summarizes the short exact sequence of vector spaces arising from quotients of filtrations ${}_\bullet\hat{B} \subseteq {}_\bullet\hat{M} \subseteq {}_\bullet\hat{Z}$ of the reference vector space $\hat{Z}$:

$$\begin{tikzcd}
{}_\bullet\hat{B} \arrow[r, "\sim"] \arrow[d, hook] & {}_\bullet\hat{B} \arrow[r, "\sim"] \arrow[d, hook] & {}_\bullet\hat{B} \arrow[r, hook] \arrow[d, hook] & {}_\bullet\hat{M} \arrow[r, hook] \arrow[d, hook] & {}_\bullet\hat{Z} \arrow[d, hook] \\
{}_\bullet\hat{B} \arrow[r, hook] \arrow[d, two heads] & {}_\bullet\hat{M} \arrow[r, hook] \arrow[d, two heads] & {}_\bullet\hat{Z} \arrow[r, "\sim"] \arrow[d, two heads] & {}_\bullet\hat{Z} \arrow[r, "\sim"] \arrow[d, two heads] & {}_\bullet\hat{Z} \arrow[d, two heads] \\
{}_\bullet\mathbf{0} \arrow[r] & \operatorname{im} H_n({}_\bullet f_\bullet) \arrow[r] & H_n({}_\bullet\hat{C}_\bullet) \arrow[r] & \operatorname{coker} H_n({}_\bullet f_\bullet) \arrow[r] & {}_\bullet\mathbf{0}
\end{tikzcd}$$

The key ingredient is the basis $\hat{\mathbf{k}}^1, \ldots, \hat{\mathbf{k}}^{\dim \hat{Z}}$ given by matrix $\hat{\mathbf{K}}$, which was found to be almost-adapted to each of the filtrations ${}_\bullet\hat{B}$ and ${}_\bullet\hat{K} = f({}_\bullet Z)$. This basis is therefore also almost-adapted the sum of the filtrations ${}_\bullet\hat{M} = {}_\bullet\hat{K} + {}_\bullet\hat{B}$; letting $\hat{\mu}$ denote the level function of ${}_\bullet\hat{M}$, we have

$$\hat{\mu}({}_1\hat{\mathbf{k}}^1) = \min(1,3) = 1, \quad \hat{\mu}({}_\infty\hat{\mathbf{k}}^2) = \min(\infty, 2) = 2$$

Hence we have the basis $\hat{\mathbf{m}}^1, \ldots, \hat{\mathbf{m}}^{\dim \hat{Z}}$ of $\hat{Z}$ that is adapted to the ${}_\bullet\hat{M}$ filtration:

$${}_1\hat{\mathbf{m}}^1 = \hat{\mathbf{k}}^1, \quad {}_2\hat{\mathbf{m}}^2 = \hat{\mathbf{k}}^2$$

which we represent by columns of the matrix $\hat{\mathbf{M}}$.

$$\hat{\mathbf{M}} = \begin{array}{c c} & \begin{array}{cc} {}_1\hat{\mathbf{m}}^1 & {}_2\hat{\mathbf{m}}^2 \end{array} \\ \begin{array}{c} {}_0\hat{\mathbf{k}}^1 \\ {}_1\hat{\mathbf{k}}^2 \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array}, \quad \hat{\mathbf{M}}_- = \hat{\mathbf{K}}\hat{\mathbf{M}} = \begin{array}{c c} & \begin{array}{cc} \hat{\mathbf{m}}^1 & \hat{\mathbf{m}}^2 \end{array} \\ \begin{array}{c} {}_2\hat{\mathbf{b}}^1 \\ {}_3\hat{\mathbf{b}}^2 \end{array} & \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \end{array}, \quad \hat{\mathbf{M}}_+ = \hat{\mathbf{N}}\hat{\mathbf{M}}_- = \begin{array}{c c} & \begin{array}{cc} \hat{\mathbf{m}}^1 & \hat{\mathbf{m}}^2 \end{array} \\ \begin{array}{c} {}_0\hat{\mathbf{z}}^1 \\ {}_1\hat{\mathbf{z}}^2 \end{array} & \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \end{array}.$$

Following Proposition 4.16 and Theorem 4.11, finding the image requires calculating a basis $\hat{\mathbf{m}}_-^1, \ldots, \hat{\mathbf{m}}_-^{\dim \hat{Z}}$ that is adapted to ${}_\bullet\hat{M}$ and simultaneously almost-adapted to ${}_\bullet\hat{B}$, while finding the cokernel requires calculating a basis $\hat{\mathbf{m}}_+^1, \ldots, \hat{\mathbf{m}}_+^{\dim \hat{Z}}$ that is adapted to ${}_\bullet\hat{M}$ and simultaneously almost-adapted to ${}_\bullet\hat{Z}$.

Since basis $\hat{\mathbf{m}}^k$ is simultaneously almost-adapted to ${}_\bullet\hat{M}$ and ${}_\bullet\hat{B}$ (by construction), it suffices to set ${}_1\hat{\mathbf{m}}^1_- = {}_1\hat{\mathbf{m}}^1$ and ${}_2\hat{\mathbf{m}}^2_- = {}_2\hat{\mathbf{m}}^2$. Their ${}_\bullet\hat{B}$-levels are inferred from the pivots of the (already reduced) matrix $\hat{\mathbf{M}}_-$ above. Hence:

Image cycle ${}_1\hat{\mathbf{m}}^1_- = -\hat{a}\hat{c} + \hat{b}\hat{c} + \hat{a}\hat{b}$, representing interval $[1, 3)$;

Image cycle ${}_2\hat{\mathbf{m}}^2_- = -\hat{a}\hat{d} + \hat{b}\hat{d} + \hat{a}\hat{b}$, representing interval $[2, 2)$.

Respectively, basis $\hat{\mathbf{m}}^k_+$ of $\hat{Z}$ is found by working with the matrix $\hat{\mathbf{M}}_+$ above; generally, this will actually require a round of reduction.

$$\underline{\hat{\mathbf{M}}}_+ = \begin{array}{c} \\ {}_0\hat{\mathbf{z}}^1 \\ {}_1\hat{\mathbf{z}}^2 \end{array} \!\!\begin{array}{c} \begin{array}{cc} \hat{\mathbf{m}}^1_+ & \hat{\mathbf{m}}^2_+ \end{array} \\ \left[\begin{array}{cc} 1 & -1 \\ 1 & 0 \end{array}\right] \end{array}, \qquad \hat{\mathbf{U}} = \begin{array}{c} \\ \hat{\mathbf{m}}^1 \\ \hat{\mathbf{m}}^2 \end{array} \!\!\begin{array}{c} \begin{array}{cc} \hat{\mathbf{m}}^1_+ & \hat{\mathbf{m}}^2_+ \end{array} \\ \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \end{array}\right] \end{array}$$

The ${}_\bullet\hat{Z}$-levels are inferred from the pivots of $\underline{\hat{\mathbf{M}}}_+$ above, and so we summarize:

Cokernel cycle ${}_1\hat{\mathbf{m}}^1_+ = -\hat{a}\hat{c} + \hat{b}\hat{c} + \hat{a}\hat{b}$, representing interval $[1, 1)$;

Cokernel cycle ${}_2\hat{\mathbf{m}}^2_+ = \hat{a}\hat{c} - \hat{a}\hat{d} - \hat{b}\hat{c} + \hat{b}\hat{d}$, representing interval $[0, 2)$.

Figure 4.6 represents the (image-cokernel) canonical short exact sequence of $H_1({}_\bullet f_\bullet)$ by morphisms of filtered complexes, in the category coim P_1. This sequence necessarily does not split, since the sequence of induced matchings is not short exact.

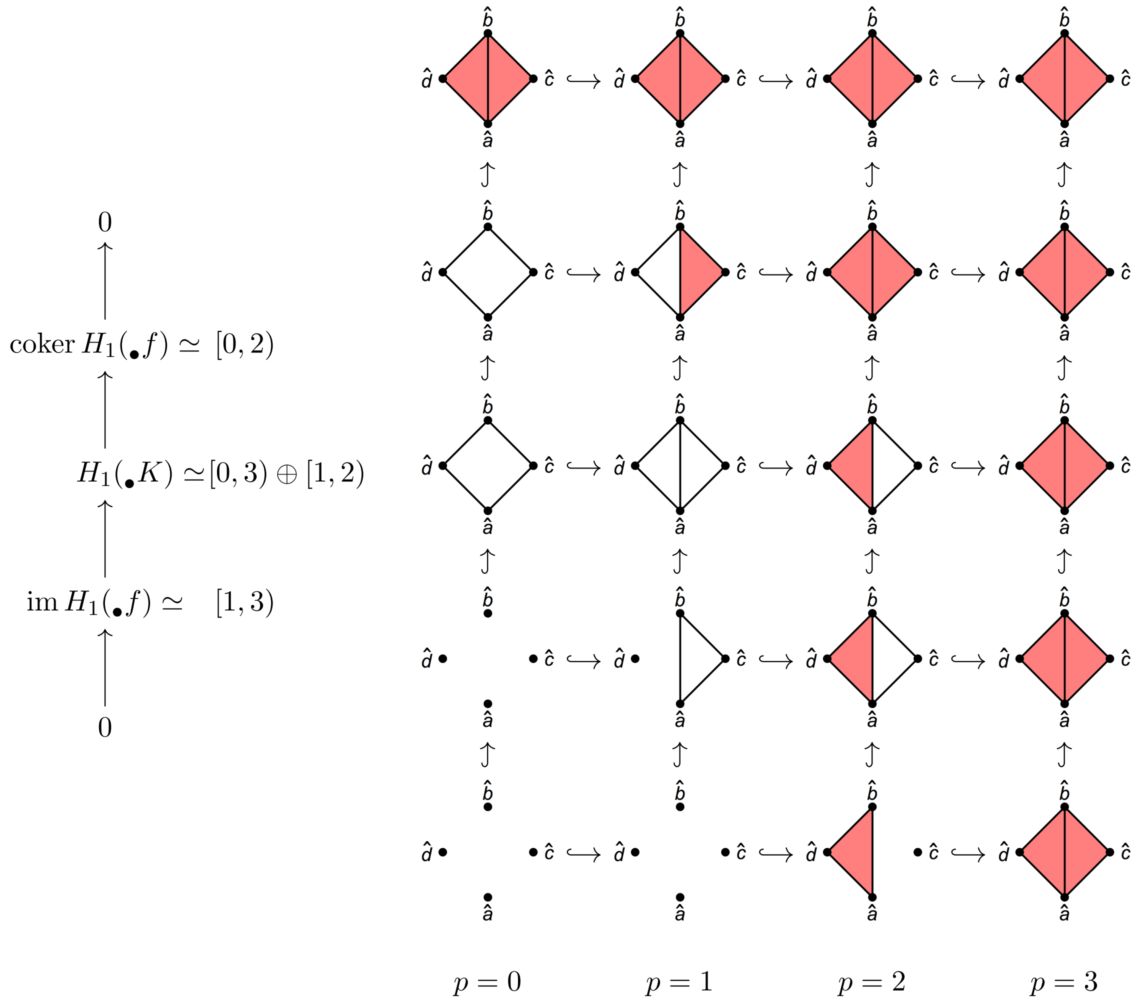


Figure 4.6: Note that the filtered complexes in the top two rows have non-intrinsic 2-simplices (relative to ${}_\bullet\hat{C}$) to emphasize their homology, while the filtered complexes in the bottom two rows are subcomplexes of ${}_\bullet\hat{C}_\bullet$.

4.3.3 Kernel and Coimage

Following the discussion in the introduction, we have the following proposition that allows us to calculate the image and cokernel of the morphism ${}_\bullet f_\bullet$.

Proposition 4.17 (see [44]). *Let ${}_\bullet M := {}_\bullet Z \cap {}_\bullet K$ be the intersection filtration of ${}_\bullet Z$ and $K := f^{-1}({}_\bullet \hat{B})$ be the preimage of ${}_\bullet \hat{B} \subseteq {}_\bullet \hat{Z}$. Then ${}_\bullet B \subseteq {}_\bullet M \subseteq {}_\bullet Z$ as subspaces of Z, and:*

1. *The map* ${}_\bullet M/{}_\bullet B \to {}_\bullet Z/{}_\bullet B$ *induced from the identity* $Z \to Z$ *is the categorical kernel,* $\ker H_n({}_\bullet f_\bullet) \to H_n({}_\bullet C_\bullet)$.

2. *The map* ${}_\bullet Z/{}_\bullet B \to {}_\bullet Z/{}_\bullet M$ *induced from the identity* $Z \to Z$ *is the categorical coimage,* $H_n({}_\bullet C_\bullet) \to \operatorname{coim} H_n({}_\bullet f_\bullet)$.

Proof. This readily follows from the fact that these relations hold at every $s \in \mathbb{Z}$. ∎

The following diagram summarizes the short exact sequence of vector spaces arising from quotients of the subspaces ${}_\bullet B \subseteq {}_\bullet M \subseteq {}_\bullet Z$ of the reference vector space Z:

$$\begin{tikzcd}
{}_\bullet B \arrow[r, "\sim"] \arrow[d, hook] & {}_\bullet B \arrow[r, "\sim"] \arrow[d, hook] & {}_\bullet B \arrow[r, hook] \arrow[d, hook] & {}_\bullet M \arrow[r, hook] \arrow[d, hook] & {}_\bullet Z \arrow[d, hook] \\
{}_\bullet B \arrow[r, hook] \arrow[d, two heads] & {}_\bullet M \arrow[r, hook] \arrow[d, two heads] & {}_\bullet Z \arrow[r, "\sim"] \arrow[d, two heads] & {}_\bullet Z \arrow[r, "\sim"] \arrow[d, two heads] & {}_\bullet Z \arrow[d, two heads] \\
{}_\bullet \mathbf{0} \arrow[r] & \ker H_n({}_\bullet f_\bullet) \arrow[r] & H_n({}_\bullet C_\bullet) \arrow[r] & \operatorname{coim} H_n({}_\bullet f_\bullet) \arrow[r] & {}_\bullet \mathbf{0}
\end{tikzcd}$$

The key ingredient is the basis $\mathbf{k}^1, \ldots, \mathbf{k}^{\dim Z}$ given by matrix $\mathbf{K}$, which was found to be almost-adapted to each of the filtrations ${}_\bullet Z$ and ${}_\bullet K = f^{-1}({}_\bullet \hat{B})$. This basis is therefore also almost-adapted the intersection of the filtrations ${}_\bullet M = {}_\bullet K \cap {}_\bullet Z$; letting μ denote the level function of ${}_\bullet M$, we have

$$\mu({}_{-\infty}\mathbf{k}^1) = \max(-\infty, 2) = 2, \;\; \mu({}_{-\infty}\mathbf{k}^2) = \max(-\infty, 2) = 2, \;\; \mu({}_3\mathbf{k}^3) = \max(3, 1) = 3$$

Hence we have the basis $\mathbf{m}^1, \ldots, \mathbf{m}^{\dim Z}$ of Z that is adapted to the ${}_\bullet M$ filtration:

$${}_2\mathbf{m}^1 = \mathbf{k}^1, \quad {}_2\mathbf{m}^2 = \mathbf{k}^2, \quad {}_3\mathbf{m}^3 = \mathbf{k}^3$$

represented by columns of the matrix $\mathbf{M}$. We also find $\mathbf{M}_+ = \mathbf{K}\mathbf{M}$ and $\mathbf{M}_- = \mathbf{N}^T\mathbf{M}_+$.

$$\mathbf{M} = \begin{array}{cc} & \begin{array}{ccc} {}_2\mathbf{m}^1 & {}_2\mathbf{m}^2 & {}_3\mathbf{m}^3 \end{array} \\ \begin{array}{c} {}_{-\infty}\mathbf{k}^1 \\ {}_{-\infty}\mathbf{k}^2 \\ {}_3\mathbf{k}^3 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}, \quad \mathbf{M}_+ = \begin{array}{cc} & \begin{array}{ccc} \mathbf{m}^1 & \mathbf{m}^2 & \mathbf{m}^3 \end{array} \\ \begin{array}{c} {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} & \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}, \quad \mathbf{M}_- = \begin{array}{cc} & \begin{array}{ccc} \mathbf{m}^1 & \mathbf{m}^2 & \mathbf{m}^3 \end{array} \\ \begin{array}{c} {}_2\mathbf{b}^1 \\ {}_3\mathbf{b}^2 \\ {}_\infty\mathbf{b}^3 \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \end{array}.$$

Following Proposition 4.16 and Theorem 4.11, finding the coimage requires calculating a basis $\mathbf{m}_+^1, \ldots, \mathbf{m}_+^{\dim Z}$ that is adapted to ${}_\bullet M$ and simultaneously almost-adapted to ${}_\bullet Z$, while finding the kernel requires calculating a basis $\mathbf{m}_-^1, \ldots, \mathbf{m}_-^{\dim Z}$ that is adapted to ${}_\bullet M$ and simultaneously almost-adapted to ${}_\bullet B$.

Since basis $\mathbf{m}^k$ is simultaneously almost-adapted to ${}_\bullet M$ and ${}_\bullet Z$ (by construction), it suffices to set: ${}_2\mathbf{m}_+^1 = {}_2\mathbf{m}^1$, ${}_2\mathbf{m}_+^2 = {}_2\mathbf{m}^2$, and ${}_3\mathbf{m}_+^3 = {}_3\mathbf{m}^3$. Their ${}_\bullet Z$-levels are inferred from the pivots of the (already reduced) matrix $\hat{\mathbf{M}}_+$ above. Hence:

Coimage cycle ${}_2\mathbf{m}_+^1 = ac - bc - ad + bd$, representing interval $[2, 2)$;

Coimage cycle ${}_2\mathbf{m}_+^2 = ac - ad + cd$, representing interval $[2, 2)$.

Coimage cycle ${}_3\mathbf{m}_+^3 = ab - ac + bc$, representing interval $[1, 3)$.

Respectively, basis $\mathbf{m}_-^k$ of Z is found by working with the matrix $\mathbf{M}_-$ above; generally, this will actually require a round of reduction.

$$\underline{\mathbf{M}}_- = \begin{array}{cc} & \begin{array}{ccc} \mathbf{m}_-^1 & \mathbf{m}_-^2 & \mathbf{m}_-^3 \end{array} \\ \begin{array}{c} {}_2\mathbf{b}^1 \\ {}_3\mathbf{b}^2 \\ {}_\infty\mathbf{b}^3 \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \end{array}, \quad \mathbf{U} = \begin{array}{cc} & \begin{array}{ccc} \mathbf{m}_-^1 & \mathbf{m}_-^2 & \mathbf{m}_-^3 \end{array} \\ \begin{array}{c} \mathbf{m}^1 \\ \mathbf{m}^2 \\ \mathbf{m}^3 \end{array} & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The ${}_\bullet B$-levels are inferred from the pivots of $\underline{\mathbf{M}}_-$ above, and so we summarize:

Kernel cycle ${}_2\mathbf{m}_-^1 = ab - ac + bd - cd$, representing interval $[2, \infty)$;

Kernel cycle ${}_2\mathbf{m}_-^2 = ab - ac + bc$, representing interval $[2, 2)$;

Kernel cycle ${}_3\mathbf{m}_-^3 = ab - ad + bd$, representing interval $[3, 3)$.

Figure 4.6 represents the (coimage-kernel) canonical short exact sequence of $H_1({}_{\bullet}f_{\bullet})$ by morphisms of filtered complexes, in the category coim P_1. This sequence necessarily does not split, since the sequence of induced matchings is not short exact.

$$
\begin{tikzcd}
0 \\
\operatorname{coim} H_1({}_{\bullet}f) \simeq [1,3] \arrow[u] \\
H_1({}_{\bullet}L) \simeq [1,\infty) \oplus [2,3) \arrow[u] \\
\ker H_1({}_{\bullet}f) \simeq [2,\infty) \arrow[u] \\
0 \arrow[u]
\end{tikzcd}
$$

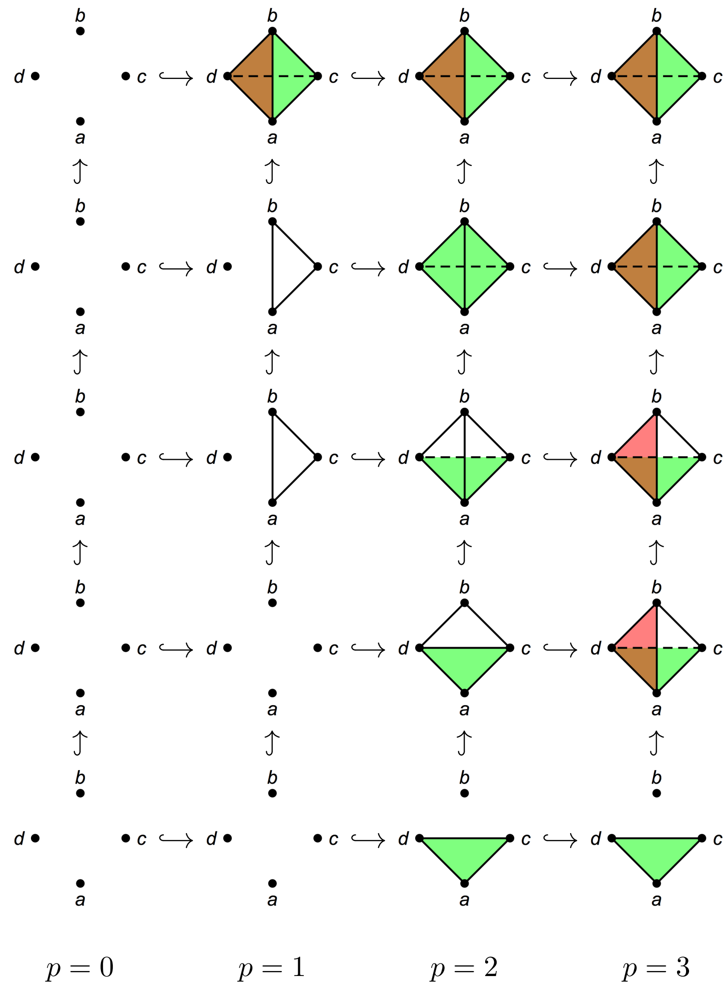


Figure 4.7: Note that the filtered complexes in the top two rows have non-intrinsic 2-simplices (relative to ${}_{\bullet}C$) to emphasize their homology, while the filtered complexes in the bottom two rows are subcomplexes of ${}_{\bullet}C_{\bullet}$.

4.4 Algorithmic Matching of Cycles and Barcodes

In section 3.3.3 we discussed the work of Bauer and Lesnick [4] on their "canonical" and "induced" matchings of barcodes. Their work elucidates the discussion of the fundamental stability and isomorphism theorems of persistent homology. It suffices to look into barcode invariants, which are entirely independent of the choices of decompositions into indecomposable summands, that is choices of a "persistent cycle" corresponding to each indecomposable summand.

Because our algorithm for (co)kernels and (co)images produces such representative cycles (not necessarily independent of the original choice of basis), it is good to elaborate on how barcode matchings behave in this formulation of persistent homology. We will make a distinction between the *canonical matching* scheme that was studied in [4] (which is purely combinatorial in nature) and an "algorithmic" matching scheme that hinges on the mechanics of how filtrations can be made conformal with respect to an almost-adapted choice of basis. A preliminary concept for discussion is that of an overlap matching, taken from [44].

Definition 4.18. Suppose $\mathbf{v}_{-}^{i}$ and $\mathbf{v}_{+}^{j}$ are bases of a vector space F . We say that the basis $\mathbf{v}_{-}^{i}$ *overlaps* with basis $\mathbf{v}_{+}^{i}$ if the change of basis matrix between them is (upper) triangular. In this case, the term *overlap matching* can be used to describe the following bijection of these ordered basis elements:

$$\mathbf{v}_{-}^{1} \leftrightarrow \mathbf{v}_{+}^{1} \quad , \quad \mathbf{v}_{-}^{2} \leftrightarrow \mathbf{v}_{+}^{2} \quad , \quad \dots \quad , \quad \mathbf{v}_{-}^{\dim V} \leftrightarrow \mathbf{v}_{+}^{\dim V}$$

Because invertible upper-triangular matrices form a multiplicative group, overlap matching is an equivalence relation on the set of all bases of F. Fundamentally, an overlap matching between bases $\mathbf{v}_-^i$ and $\mathbf{v}_+^j$ is determined from the pivots of the change-of-basis matrix $[1_F]_{\mathbf{v}_-^i}^{\mathbf{v}_+^j}$ located on its diagonal.

In our case, we use ordered bases adapted to filtrations on F so we are interested in how an overlap matching conforms to that notion.

Lemma 4.19. *Suppose that $\mathbf{v}_-^i$ and $\mathbf{v}_+^j$ are bases of a vector space F that overlap, and let ${}_\bullet U$ be a filtration on F with level function u.*

1. *If $\mathbf{v}_+^j$ is adapted to a filtration ${}_\bullet U$, then $u(\mathbf{v}_+^k) = u(\mathbf{v}_-^k)$ for all k. In particular, $\mathbf{v}_-^i$ is also adapted to ${}_\bullet U$.*

2. *If $\mathbf{v}_+^j$ is almost-adapted to a filtration ${}_\bullet U$, then $u(\mathbf{v}_+^k) \leqslant u(\mathbf{v}_-^k)$ for all k.*

Proof. Let $\mathbf{M}$ be the change of basis matrix from $\mathbf{v}_-^i$ to $\mathbf{v}_+^j$; it is upper-triangular by hypothesis, so every $\mathbf{v}_-^k$ is generated by $\{\mathbf{v}_+^1, \ldots, \mathbf{v}_+^k\}$.

If $\mathbf{v}_+^j$ is almost-adapted to ${}_\bullet U$, then ${}_t\beta = \{\mathbf{v}_+^j : u(\mathbf{v}_+^j) \leqslant t\}$ is a basis for ${}_tU$; note that $t = \max_{\mathbf{v} \in {}_t\beta} u(\mathbf{v})$. Clearly $\mathbf{v}_-^k \in {}_tU$ if and only if $\{\mathbf{v}_+^1, \ldots, \mathbf{v}_+^k\} \subseteq {}_t\beta$, so $t \geqslant \max\{u(\mathbf{v}_+^j) : j = 1, \ldots, k\} \geqslant u(\mathbf{v}_+^k)$. This proves the second statement.

If $\mathbf{v}_+^j$ is adapted to ${}_\bullet U$, then $u(\mathbf{v}_+^1) \leqslant \cdots \leqslant u(\mathbf{v}_+^{\dim F})$; hence, any element $\mathbf{v}$ in the span of $\{\mathbf{v}_+^1, \ldots, \mathbf{v}_+^k\}$ must satisfy $u(\mathbf{v}) \leqslant u(\mathbf{v}_+^k)$. Setting $\mathbf{v} = \mathbf{v}_-^k$, equality in the first statement follows since $\mathbf{v}_+^j$ is (trivially) almost-adapted to ${}_\bullet U$. ∎

As a direct consequence, we have the following inequalities that play a role in describing barcode matchings arising from overlap matchings of cycles.

Corollary 4.20. *Suppose ${}_\bullet B \subseteq {}_\bullet M \subseteq {}_\bullet Z$ are filtrations of the vector space Z, with filtration levels denoted by λ_B, λ_M, and λ_Z respectively.*

1. *Let $\mathbf{b}^i$ and $\mathbf{b}_-^j$ be overlapping bases of Z. If $\mathbf{b}^i$ is almost-adapted to ${}_\bullet Z$, then $\lambda_Z(\mathbf{b}^k) \leqslant \lambda_M(\mathbf{b}_-^k)$ for each k.*

2. *Let $\mathbf{z}^i$ and $\mathbf{z}_+^j$ be overlapping bases of Z. If $\mathbf{z}_+^i$ is almost-adapted to ${}_\bullet M$, then $\lambda_M(\mathbf{z}_+^k) \leqslant \lambda_B(\mathbf{z}^k)$ for each k.*

Proof. The first statement follows since Lemma 4.19 implies $\lambda_Z(\mathbf{b}^k) \leqslant \lambda_Z(\mathbf{b}_-^k)$ and Proposition 4.10 implies $\lambda_Z(\mathbf{b}_-^k) \leqslant \lambda_M(\mathbf{b}_-^k)$. The second statement is similar. ∎

4.4.1 Overlap Matching for Image and Cokernel

Recall from section 4.2.2 the derivation of the following matrices.

$$\hat{\mathbf{N}} = \begin{array}{c|cc} & {}_2\hat{\mathbf{b}}^1 & {}_3\hat{\mathbf{b}}^2 \\ \hline {}_0\hat{\mathbf{z}}^1 & 0 & 1 \\ {}_1\hat{\mathbf{z}}^2 & 1 & 0 \end{array}, \qquad \hat{\mathbf{N}}^T = \begin{array}{c|cc} & {}_0\hat{\mathbf{z}}^1 & {}_1\hat{\mathbf{z}}^2 \\ \hline {}_2\hat{\mathbf{b}}^1 & 0 & 1 \\ {}_3\hat{\mathbf{b}}^2 & 1 & 0 \end{array}$$

Also recall from section 4.3.1 the following change-of-basis matrices.

$$\hat{\mathbf{M}}_- = \begin{array}{c|cc} & {}_1\hat{\mathbf{m}}_-^1 & {}_2\hat{\mathbf{m}}_-^2 \\ \hline {}_2\hat{\mathbf{b}}^1 & 1 & 1 \\ {}_3\hat{\mathbf{b}}^2 & 1 & 0 \end{array}, \qquad \underline{\hat{\mathbf{M}}}_+ = \begin{array}{c|cc} & {}_1\hat{\mathbf{m}}_+^1 & {}_2\hat{\mathbf{m}}_+^2 \\ \hline {}_0\hat{\mathbf{z}}^1 & 1 & -1 \\ {}_1\hat{\mathbf{z}}^2 & 1 & 0 \end{array}.$$

These display the representation of bases $\hat{\mathbf{m}}_-^i$ and $\hat{\mathbf{m}}_+^j$ with respect to the basis $\hat{\mathbf{b}}^i$ adapted to the range-canonical filtration and the basis $\hat{\mathbf{z}}^j$ adapted the the preimage-canonical filtration. Use $\hat{\mathbf{M}}_-$ to find a basis $\hat{\mathbf{b}}_-^i$ adapted to ${}_\bullet\hat{B}$ by looking at the matrix pivots and assigning levels appropriately:

$${}_2\hat{\mathbf{b}}_-^1 = \hat{\mathbf{m}}_-^2, \quad {}_3\hat{\mathbf{b}}_-^2 = \hat{\mathbf{m}}_-^1$$

Use $\underline{\hat{\mathbf{M}}}_+$ to find a basis $\hat{\mathbf{z}}_-^i$ adapted to ${}_\bullet\hat{Z}$ by looking at the matrix pivots and assigning levels appropriately:

$${}_0\hat{\mathbf{z}}_+^1 = \hat{\mathbf{m}}_+^2, \quad {}_1\hat{\mathbf{z}}_+^2 = \hat{\mathbf{m}}_+^1$$

Induced matching of $\operatorname{im} H_1({}_\bullet f_\bullet) \to H_1({}_\bullet \hat{C}_\bullet)$ is achieved via the overlap matching $\hat{\mathbf{b}}^i_- \leftrightarrow \hat{\mathbf{b}}^i$ (where the one on the left is almost-adapted to ${}_\bullet \hat{M}$ and the one on the right is almost-adapted to ${}_\bullet \hat{Z}$). This induces a matching of interval complexes $\langle {}_\bullet[\hat{\mathbf{b}}^i_-]\rangle \leftrightarrow \langle {}_\bullet[\hat{\mathbf{b}}^i]\rangle$, corresponding to an interval matching

$$\Big[\lambda_{\hat{M}}(\hat{\mathbf{b}}^i_-), \lambda_{\hat{B}}(\hat{\mathbf{b}}^i_-)\Big) \longleftrightarrow \Big[\lambda_{\hat{Z}}(\hat{\mathbf{b}}^i), \lambda_{\hat{B}}(\hat{\mathbf{b}}^i)\Big)$$

which by Corollary 4.20 specifies a monic matching of barcodes:

$[2,2) \leftrightarrow [1,2)$, representing the overlap $\hat{\mathbf{b}}^1_- \leftrightarrow \hat{\mathbf{b}}^1$

$[1,3) \leftrightarrow [0,3)$, representing the overlap $\hat{\mathbf{b}}^2_- \leftrightarrow \hat{\mathbf{b}}^2$

Induced matching of $H_1({}_\bullet \hat{C}_\bullet) \to \operatorname{coker} H_1({}_\bullet f_\bullet)$ is achieved via the overlap matching $\hat{\mathbf{z}}^i \leftrightarrow \hat{\mathbf{z}}^i_+$ (where the one on the left is almost-adapted to ${}_\bullet B$ and the one on the right is almost-adapted to ${}_\bullet M$). This induces a matching of interval complexes $\langle {}_\bullet[\hat{\mathbf{z}}^i]\rangle \leftrightarrow \langle {}_\bullet[\hat{\mathbf{z}}^i_+]\rangle$, corresponding to an interval matching

$$\Big[\lambda_{\hat{Z}}(\hat{\mathbf{z}}^i), \lambda_{\hat{B}}(\hat{\mathbf{z}}^i)\Big) \longleftrightarrow \Big[\lambda_{\hat{Z}}(\hat{\mathbf{z}}^i_+), \lambda_{\hat{M}}(\hat{\mathbf{z}}^i_+)\Big)$$

which by Corollary 4.20 specifies an epic matching of barcodes:

$[0,3) \leftrightarrow [0,2)$, representing the overlap $\hat{\mathbf{z}}^1 \leftrightarrow \hat{\mathbf{z}}^1_+$

$[1,2) \leftrightarrow [1,1)$, representing the overlap $\hat{\mathbf{z}}^2 \leftrightarrow \hat{\mathbf{z}}^2_+$

4.4.2 Overlap Matching for Kernel and Coimage

Recall from section 4.2.1 the derivation of the following matrices.

$$\mathbf{N} = \begin{array}{c} \\ {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} \begin{array}{c} \begin{array}{ccc} {}_2\mathbf{b}^1 & {}_3\mathbf{b}^2 & {}_\infty\mathbf{b}^3 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}, \qquad \mathbf{N}^T = \begin{array}{c} \\ {}_2\mathbf{b}^1 \\ {}_3\mathbf{b}^2 \\ {}_\infty\mathbf{b}^3 \end{array} \begin{array}{c} \begin{array}{ccc} {}_1\mathbf{z}^1 & {}_2\mathbf{z}^2 & {}_2\mathbf{z}^3 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

Also recall from section 4.3.2 the calculation of the following change-of-basis matrices.

$$\mathbf{M}_+ = \begin{array}{c c} & \begin{array}{ccc} {}_2\mathbf{m}_+^1 & {}_2\mathbf{m}_+^2 & {}_3\mathbf{m}_+^3 \end{array} \\ \begin{array}{c} {}_1\mathbf{z}^1 \\ {}_2\mathbf{z}^2 \\ {}_2\mathbf{z}^3 \end{array} & \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}, \quad \underline{\mathbf{M}}_- = \begin{array}{c c} & \begin{array}{ccc} {}_2\mathbf{m}_-^1 & {}_2\mathbf{m}_-^2 & {}_3\mathbf{m}_-^3 \end{array} \\ \begin{array}{c} {}_2\mathbf{b}^1 \\ {}_3\mathbf{b}^2 \\ {}_\infty\mathbf{b}^3 \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \end{array}.$$

These display the representation of bases $\mathbf{m}_+^i$ and $\mathbf{m}_-^j$ with respect to the basis $\mathbf{z}^i$ adapted to the preimage-canonical filtration and the basis $\hat{\mathbf{b}}^j$ adapted the the range-canonical filtration. Use $\underline{\mathbf{M}}_-$ to find a basis $\mathbf{b}_-^i$ adapted to ${}_\bullet B$ by looking at the matrix pivots and assigning levels appropriately:

$${}_2\mathbf{b}_-^1 = \mathbf{m}_-^2, \quad {}_3\mathbf{b}_-^2 = \mathbf{m}_-^3, \quad {}_\infty\mathbf{b}_-^3 = \mathbf{m}_-^1$$

Use $\mathbf{M}_+$ to find a basis $\mathbf{z}_+^i$ adapted to ${}_\bullet Z$ by looking at the matrix pivots and assigning levels appropriately:

$${}_1\mathbf{z}_+^1 = \hat{\mathbf{m}}_+^3, \quad {}_2\mathbf{z}_+^2 = \hat{\mathbf{m}}_+^1, \quad {}_2\mathbf{z}_+^3 = \hat{\mathbf{m}}_+^2$$

Induced matching of $\ker H_1({}_\bullet f_\bullet) \to H_1({}_\bullet C_\bullet)$ is achieved via the overlap matching $\mathbf{b}_-^i \leftrightarrow \mathbf{b}^i$ (where the one on the left is almost-adapted to ${}_\bullet M$ and the one on the right is almost-adapted to ${}_\bullet Z$). This induces a matching of interval complexes $\langle {}_\bullet[\mathbf{b}_-^i]\rangle \leftrightarrow \langle {}_\bullet[\mathbf{b}^i]\rangle$, corresponding to an interval matching

$$\Big[\lambda_M(\mathbf{b}_-^i), \lambda_B(\mathbf{b}_-^i)\Big) \longleftrightarrow \Big[\lambda_Z(\mathbf{b}^i), \lambda_B(\mathbf{b}^i)\Big)$$

which by Corollary 4.20 specifies a monic matching of barcodes:

$[2,2) \leftrightarrow [2,2)$, representing the overlap $\mathbf{b}_-^1 \leftrightarrow \mathbf{b}^1$

$[3,3) \leftrightarrow [2,3)$, representing the overlap $\mathbf{b}_-^2 \leftrightarrow \mathbf{b}^2$

$[2,\infty) \leftrightarrow [1,\infty)$, representing the overlap $\mathbf{b}_-^3 \leftrightarrow \mathbf{b}^3$

Induced matching of $H_1({}_\bullet C_\bullet) \to \operatorname{coim} H_1({}_\bullet f_\bullet)$ is achieved via the overlap matching $\mathbf{z}^i \leftrightarrow \mathbf{z}^i_+$ (where the one on the left is almost-adapted to ${}_\bullet B$ and the one on the right is almost-adapted to ${}_\bullet M$). This induces a matching of interval complexes $\langle {}_\bullet[\mathbf{z}^i] \rangle \leftrightarrow \langle {}_\bullet[\mathbf{z}^i_+] \rangle$, corresponding to an interval matching

$$\Big[\lambda_Z(\mathbf{z}^i), \lambda_B(\mathbf{z}^i)\Big) \longleftrightarrow \Big[\lambda_Z(\mathbf{z}^i_+), \lambda_M(\mathbf{z}^i_+)\Big)$$

which by Corollary 4.20 specifies an epic matching of barcodes:

$[1, \infty) \leftrightarrow [1, 3)$, representing the overlap $\mathbf{z}^1 \leftrightarrow \mathbf{z}^1_+$

$[2, 3) \leftrightarrow [2, 2)$, representing the overlap $\mathbf{z}^2 \leftrightarrow \mathbf{z}^2_+$

$[2, 2) \leftrightarrow [2, 2)$, representing the overlap $\mathbf{z}^3 \leftrightarrow \mathbf{z}^3_+$

4.4.3 Matchings Compatible with the Canonical Isomorphism

The isomorphism $\operatorname{coim} H_1({}_\bullet f_\bullet) \to \operatorname{im} H_1({}_\bullet f_\bullet)$ is induced by the morphism $f : {}_\bullet Z \to {}_\bullet \hat{M}$ that restricts to ${}_\bullet M \to {}_\bullet \hat{B}$. Since we used the same reduced matrix for coimage and image, the decompositions of the coimage and image into (appropriately) adapted bases are compatible via the canonical isomorphism.

Recall section 4.3.1 the matrix $\underline{\mathbf{F}}$, which represents the chain map f_1.

$$\underline{\mathbf{F}} = \begin{array}{c c} & \begin{array}{ccc} {}_1\mathbf{z}^1_+ & {}_2\mathbf{z}^2_+ & {}_2\mathbf{z}^3_+ \end{array} \\ \begin{array}{c} {}_2\hat{\mathbf{b}}^1 \\ {}_3\hat{\mathbf{b}}^2 \end{array} & \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right] \end{array}$$

It sends ${}_3\mathbf{m}^3_+ = {}_3\mathbf{k}^3 = {}_1\mathbf{z}^1 \in Z$ to ${}_1\hat{\mathbf{m}}^1_- = {}_1\hat{\mathbf{k}}^1 = {}_2\hat{\mathbf{b}}^1 + {}_3\hat{\mathbf{b}}^2 \in \hat{Z}$.

The factorization of the morphism through coimage and image is

$$\mathbf{[1, \infty)} \oplus [2, 3) \to \mathbf{[1, 3)} \to \mathbf{[1, 3)} \to \mathbf{[0, 3)} \oplus [1, 2)$$

The (nontrivial) matching of the epic coimage map is the epic $[1, \infty) \to [1, 3)$. The nontrivial matching of the monic image map is the monic $[1, 3) \to [0, 3)$. These pairwise matchings are indicated in the above diagram in bold, where the matching in the middle is an isomorphism and is shown below:

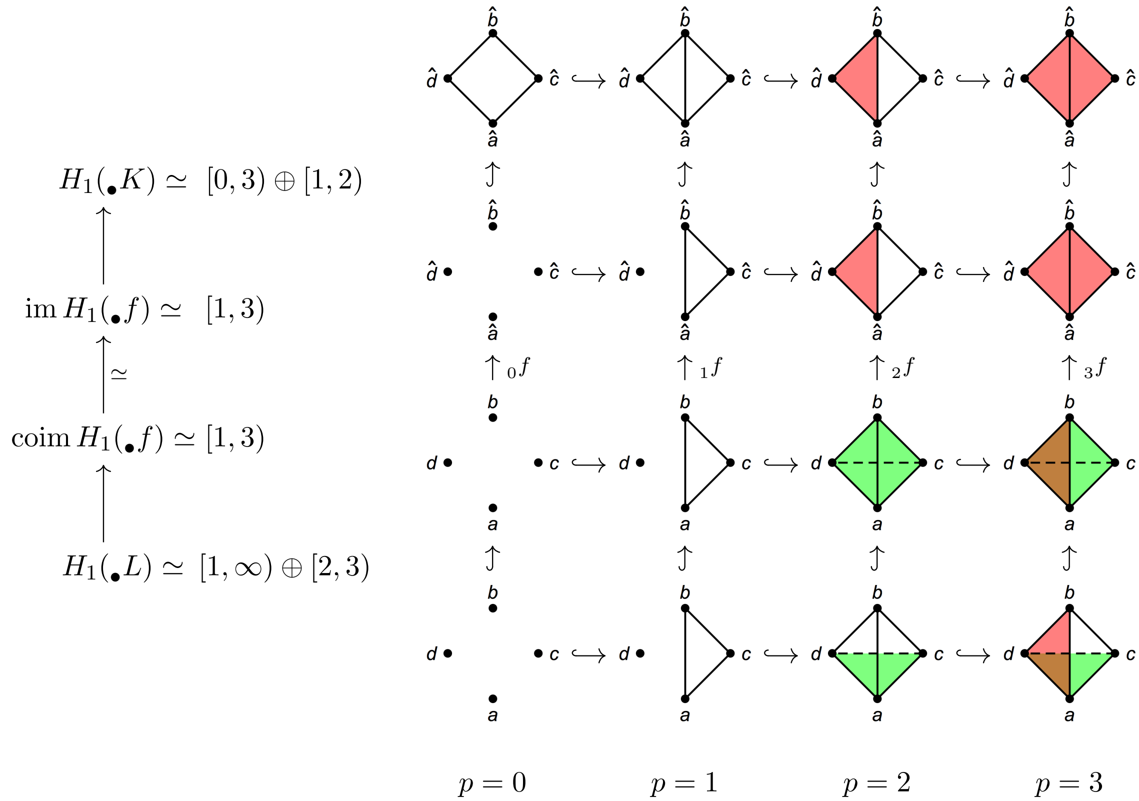


Figure 4.8: The canonical factorization of the original persistence vector space map $H_1({}_\bullet f) : H_1({}_\bullet C) \to H_1({}_\bullet \hat{C})$ through its coimage and image can also be described as the homology of a grid diagram of filtered chain complexes. Note that each row is a filtered complex, but a column is not a filtered complex in general.

Note that the nonempty overlap necessarily satisfies:

$$[1,3) = [1,\infty) \cap [0,3).$$

If the map $H_n(f)$ of persistence vector spaces is monic, then the coimage map is the identity and the image map is $H_n(f)$. So the matched cycles have same deaths. By permutation of bases, one recovers the existence of the canonical matching for monic barcodes as given in [4].

If the map $H_n(f)$ of persistence vector spaces is epic, then the coimage map is $H_n(f)$ and the image map is the identity. So the matched cycles have the same births. By permutation of bases, one recovers the existence of the canonical matching for epic barcodes as given in [4].

This shows that the properties of the algorithmic matching scheme are consistent with the those exhibited by canonical matchings originally presented in [4]. Every cycle matching induces a barcode matching in the obvious way by forgetting the cycles but remembering their barcodes.

Then these induce a matching of barcodes as an association of barcodes of C and $\hat{C}$ via the isomorphism matching of the coimage to image. Some intervals are not matched, represented by an overlap matching of $\mathbf{v}_-^1 \leftrightarrow \mathbf{v}_+^1$ where at least one of the elements belongs to the trivial class in the quotient of filtrations. The induced matching in [4] is an "optimal" matching of barcodes of C and $\hat{C}$, ultimately giving the proof of the algebraic stability theorem. Note that this is purely combinatorial, independent of any choices of cycles/indecomposables on C or $\hat{C}$.

Any decomposition of C and $\hat{C}$ as well as image and coimage of $H_n(f)$, such as those provided by our variant algorithm, can be used to do an algorithmic cycle matching. Of course we can rearrange orders of bases and get different results, so in general these representative bases are not unique. Then our *algorithmic cycle matching* algorithm produces these cycle matchings via the composition of morphisms in $C \to \operatorname{coim}_\bullet f_\bullet \to \operatorname{im}_\bullet f_\bullet \to \hat{C}$, resulting in a cycle matching from C to $\hat{C}$.

An algorithmic matching exists by default, after choices of interval representatives for C and $\hat{C}$ have been made – this matching is procedurally generated using the mechanics of overlap matchings stated in this section. We can state it as an immediate corollary of the existence of algorithmic cycle matchings guarantees the existence of an “optimal” barcode matching, which may be the induced barcode matching of [4]. Additionally, our stated algorithmic cycle matching results, together with the inequalities of Corollary 4.20, yield an alternate proof of the canonical barcode matchings of Theorem 4.2 in [4].

However, algorithmic cycle matchings do not in general agree with canonical barcode matchings: it is not too arduous to devise examples of some monic or epic map $H_n({}_\bullet C_\bullet) \to H_n({}_\bullet C_\bullet)$ such that the induced barcode matching from overlapping bases is not a canonical matching.

Bibliography

[1] Jiří Adámek, Horst Herrlich, George E. Strecker. "Abstract and Concrete Categories: The Joy of Cats"; 512 pages. *Dover Publications*, New York (2009). ISBN 978-0-486-46934-8. [http://katmat.math.uni-bremen.de/acc].

[2] Michael Atiyah (1956). "On the Krull-Schmidt theorem with application to sheaves". *Bulletin de la S.M.F.* **84**, pages 307-317. [http://www.numdam.org/item?id=BSMF_1956__84__307_0]

[3] Jonathan L. Alperin, Rowen B. Bell. "Groups and Representations"; 206 pages. *Springer-Verlag*, New York (1995). ISBN 978-0387945262.

[4] Ulrich Bauer, Michael Lesnick (2013). "Induced Matchings and the Algebraic Stability of Persistence Barcodes". *Journal of Computational Geometry* **6**(2), pages 162–191. [arXiv:1311.3681]

[5] Ulrich Bauer, Michael Lesnick (2020). "Persistence Diagrams as Diagrams: A Categorification of the Stability Theorem". *Topological Data Analysis. Abel Symposia* **15**. [arXiv:1610.10085]

[6] Paul Bendich, Peter Bubenik, Alexander Wagner (2015). "Stabilizing the unstable output of persistent homology computations." *Journal of Applied and Computational Topology* **4**, pages 309-338.

[7] Jean-Daniel Boissonnat, Ramsay Dyer, Arijit Ghosh (2013). "Delaunay stability via perturbations". *International Journal of Computational Geometry & Applications* **24**(2), pages 125–152. [arXiv:1310.7696]

[8] Raoul Bott, Loring W. Tu. "Differential Forms in Algebraic Topology"; 1st edition, 352 pages. *Springer-Verlag*, New York (1982). ISBN 978-0387906133.

[9] Peter Bubenik, Vin de Silva, Jonathan Scott (2013). "Metrics for generalized persistence modules". *Foundations of Computational Mathematics* **15**(6), pages 1501–1531.

[10] Gunnar Carlsson (2020). "Persistent Homology and Applied Homotopy Theory". Handbook of Homotopy Theory, CRC Press. [arXiv:2004.00738]

[11] Gunnar Carlsson, Vin de Silva (2008). "Zigzag Persistence". *Foundations of Computational Mathematics* **10**, pages 367–405. [arXiv:0812.0197]

[12] Ana Carpio, Luis Bonilla, James Mathews, Allen Tannenbaum (2019). "Fingerprints of cancer by persistent homology".

[13] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas Guibas, Steve Oudot (2008). "Proximity of Persistence Modules and their Diagrams." *Proceedings of the 25th Annual Symposium on Computational Geometry*, pages 237-246. [https://geometrica.saclay.inria.fr/team/Marc.Glisse/publications/futur/stable.pdf]

[14] Frédéric Chazal, Vin de Silva, Marc Glise, Steve Oudot (2012). "The structure and stability of persistence modules." Springer Briefs in Mathematics, *Springer-Verlag*. ISBN 978-3-319-42545-0. [arXiv:1207.3674]

[15] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer (2007). "Stability of persistence diagrams." *Discrete Comput. Geom.* **37**(1), pages 103–120.

[16] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, Dmitriy Morozov."Persistent homology for kernels, images, and cokernels". *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* (Jan. 2009), pages 1011-1020.

[17] Jean Dieudonné. "A History of Algebraic and Differential Topology, 1900–1960". Birkhäuser Basel (1989). ISBN 978-3764333881.

[18] Herbert Edelsbrunner, John Harer. "Computational Topology: an introduction". Department of Computer Science and Mathematics, Duke University. [https://www.maths.ed.ac.uk/ v1ranick/papers/edelcomp.pdf]

[19] Herbert Edelsbrunner, Grzegorz Jabłoński, Marian Mrozek (2015). "The Persistent Homology of a Self-Map". *Foundations of Computational Mathematics* **15**, pages 1213–1244.

[20] Herbert Edelsbrunner, David Letscher, Afra Zomorodian (2000). "Topological persistence and simplification". *Foundations of Computer Science*, Proceedings 41st Annual Symposium IEEE 2000, pages 454463.

[21] Herbert Edelsbrunner, Dmitriy Morozov (2013). "Persistent Homology: Theory and Practice". [https://mrzv.org/publications/persistent-homology-theory-practice/ecm/]

[22] Herbert Edelsbrunner, Nimish R. Shah (1994). "Triangulating Topological Spaces". *International Journal of Computational Geometry & Applications* **7**(4), pages 365–378. [http://pub.ist.ac.at/ edels/Papers/1997-J-01-TriangulatingTopologicalSpaces.pdf]

[23] Samuel Eilenberg, Max G. Kelly (1966). "Closed Categories". *Proceedings of the Conference on Categorical Algebra* – La Jolla 1965 , Springer Heidelberg, pages 421–562.

[24] Samuel Eilenberg, Norman E. Steenrod. "Foundations of algebraic topology"; 15+328 pages. *Princeton University Press*, Princeton NJ (1952).

[25] Pavel Etingof, Shlomo Gelaki, Dmitri Nikshych, Victor Ostrik. "Tensor Categories"; 344 pages. AMS, *Mathematical Surveys and Monographs* **205**, 2015. ISBN: 978-1-4704-3441-0. [http://www-math.mit.edu/ etingof/egnobookfinal.pdf]

[26] Leonard Euler (1758). "Elementa doctrinae solidorum". *Novi Commentarii academiae scientiarum Petropolitanae* **4**, pages 109–140.

[27] Peter Freyd. "Abelian Categories – An Introduction to the Theory of Functors"; 1st edition, 164 pages. *Harper and Row*, New York (1964). ASIN B0000CMCLZ.

[28] Greg Friedman (2012). "An elementary illustrated introduction to simplicial sets". *Rocky Mountain J. Math.* **42**(2), pages 353-423. [arXiv:0809.4221]

[29] J. Gamble, G. Heo (2010). "Exploring uses of persistent homology for statistical analysis of landmark-based shape data". *Journal of Multivariable Analysis* **101**(9), pages 2184–2199.

[30] Sergei Gelfand, Yuri Manin. "Methods of Homological Algebra". *Springer Berlin Heidelberg*, 1996.

[31] Rocio Gonzalez-Diaz, Eduardo Paluzo-Hidalgo, José F. Quesada (2019). "Towards Emotion Recognition: A Persistent Entropy Application". *7th International Workshop on Computational Topology in Image Context*, pages 96–109.

[32] Allen Hatcher. "Algebraic Topology"; 556 pages. *Cambridge University Press*, 2001. ISBN 9780521795401. [https://pi.math.cornell.edu/ hatcher/AT/AT.pdf]

[33] Christoph Hofer, Roland Kwitt, Marc Niethammer, Andreas Uhl (2017). "Deep Learning with Topological Signatures". *Advances in Neural Information Processing Systems* **30**. [arXiv:1707.04041]

[34] Li Juanjuan, Wei Wen, Liu Zhongfen, Chen Chuang, Cheng Jing, Gong Yiping, Wang Changhua, Yu Dehua, Sun Shengrong. (2015). "Clinical pathological characteristics of breast cancer patients with secondary diabetes after systemic therapy: a retrospective multicenter study". *Tumour biology : the Journal of the International Society for Oncodevelopmental Biology and Medicine* **36**(9).

[35] Johannes Kepler. "The Harmony of the World" (translated by Dr. Juliet Field). *The American Philosophical Society*, 1997. ISBN 0-87169-209-0.

[36] Henning Krause (2014). "Krull-Schmidt categories and projective covers". [arXiv:1410.2822]

[37] Serge Lang. "Algebra"; 993 pages. *Springer*, 2002. ISBN 9780387953854.

[38] Michael Lesnick (2015). "The Theory of the Interleaving Distance on Multidimensional Persistence Modules." *Foundations of Computational Mathematics* **15**(3), pages 613-650. [arXiv:1106.5305]

[39] Shiping Liu (2010). "Auslander-Reiten theory in a Krull-Schmidt category". *Sao Paulo Journal of Mathematical Sciences* **4**(3), pages 425-472 [https://www.ime.usp.br/ spjm/articlepdf/432.pdf]

[40] Saunders Mac Lane. "Categories for the Working Mathematician"; 2nd edition, 329 pages. *Springer-Verlag*, New York (1978). ISBN 978-1441931238.

[41] Yukio Matsumoto. "An Introduction to Morse Theory"; 219 pages. AMS, *Translations of Mathematical Monographs*, Providence RI (1997). ISBN 9780821810224.

[42] Killian Meehan, David Meyer (2017). "An Isometry Theorem for Generalized Persistence Modules". [arXiv:1710.02858]

[43] Killian Meehan, Andrei Pavlichenko, Jan Segert (2019). "On the Structural Theorem of Persistent Homology". *Discrete Comput Geom* **62**, pages 945989. [arXiv:1701.02055]

[44] Killian Meehan, Andrei Pavlichenko, Jan Segert (2019). "On Persistent Homology (Co)Kernels, (Co)Images, and Matchings". Manuscript in progress.

[45] Ezra Miller (2015). "Fruit Flies and Moduli: Interactions between Biology and Mathematics". *Notices of the American Mathematical Society* **62**(10). [arXiv:1508.05381]

[46] James R. Munkres. "Elements of Algebraic Topology"; 1st edition, 468 pages. *CRC Press* (1993). ISBN 978-0367091415.

[47] Steve Oudot. "Persistence Theory: From Quiver Representations to Data Analysis". AMS, *Mathematical Surveys and Monographs* **209** (2015). ISBN 9781470425466.

[48] A. Piron, F. van der Heijden, M. F. Destain (2011). "Weed detection in 3D images". *Precision Agriculture* **12**(5), pages 607 – 622.

[49] Chi Seng Pun, Brandon Yung Sin Yong, Kelin Xia (2020). "Weighted-persistent-homology-based machine learning for RNA flexibility analysis". *PLoS One* **15**(8):e0237747.

[50] Emily Riehl. "Categorical Homotopy Theory". *Cambridge University Press* (2014). ISBN 9781107261457. [http://www.math.jhu.edu/ eriehl/cathtpy.pdf]

[51] Christian Rüschoff (2017). "Lecture Notes: Simplicial Sets". Ruprecht Karl University of Heidelberg. [https://www.mathi.uni-heidelberg.de/ rueschoff/ss17sset/sset.pdf]

[52] Katsuro Sakai (2010). "Simplicial Homology – A Short Course". Institute of Mathematics, University of Tsukuba. [https://sites.google.com/site/ksakaiidtopology/home/homepage-of-katsuro-sakai/alg-top].

[53] Ralf Schiffler. "Quiver Representations"; 244 pages. CMS Books in Mathematics, *Springer-Verlag* (2014). ISBN 978-3319092034.

[54] Vin de Silva, Gunnar Carlsson (2004). "Topological estimation using witness complexes". *Symposium on Point Based Graphics* (The Eurographics Association), pages 157-166.

[55] Vin de Silva, Dmitriy Morozov, Mikael Vejdemo-Johansson (2011). "Dualities in persistent (co)homology". *Inverse Problems* **27**(12). [arXiv:1107.5665]

[56] Primoz Skraba, Katharine Turner (2020). "Wasserstein Stability for Persistence Diagrams". [arXiv:2006.16824]

[57] Primoz Skraba, Mikael Vejdemo-Johansson (2013). "Persistence modules: Algebra and algorithms". [arXiv:1302.2015]

[58] Edwin H. Spanier. "Algebraic topology"; xiv+528 pages. McGraw Series in Higher Mathematics, *McGraw-Hill*, New York (1966). ISBN 978-0070598836.

[59] Stacks Project Authors, Homological Algebra, The Stacks Project. [http://stacks.math.columbia.edu/download/homology.pdf]

[60] Afra Zomorodian, Gunnar Carlsson (2005). "Computing Persistent Homology." *Discrete & Computational Geometry* **33**, pages 249–274.

[61] Afra Zomorodian, Gunnar Carlsson (2008). "Localized Homology" *Computational Geometry* **41**(3), pages 126–148.

VITA

Andrei Pavlichenko was born in Moscow, Russia in 1991 where he lived before moving to the United States at a young age. He received his Bachelors degree in Mathematics from the University of Rochester in 2013. That same year, he joined the graduate program at the University of Missouri working on his Masters and Doctoral degrees. During this time, he has worked with Professor Jan Segert studying the theory and applied methods of persistent homology.