PERSISTENCE SPECTRAL SEQUENCES

Álvaro Torras Casas

supervised by Dr. Ulrich Pennig

School of Mathematics

Cardiff University

A thesis submitted for the degree of

Doctor of Philosophy

2022

Dedicated to my grandparents

Feliu, Pilar, Josep i Teodora

Abstract

This Doctoral thesis is centered on connections between persistent homology and spectral sequences. We explain some of the approaches in the literature exploring this connection. Our main focus is on Mayer-Vietoris spectral sequences associated to filtered covers on filtered complexes. A particular case of this spectral sequence is used for measuring exact changes on barcode decompositions under small perturbations of the underlying data. On the other hand, these objects allow for a setup to parallelize persistent homology computations, while retaining useful information related to the chosen covers. We explore some generalizations of the traditional setup to diagrams of regular complexes consisting of regular morphisms; these become useful for working with non-sparse complexes. In addition, we explore stability results related to these new invariants, both with respect to local changes and with respect to changes on the chosen covering sets. Finally, we present some computational experiments by the use of PERMAVISS which illustrate some of these ideas.

Acknowledgements

I would like to thank my supervisor Dr. Ulrich Pennig for all his time and effort spent on this work. Also, Dr. Pennig suggested the thesis topic and has been very helpful and supportive in the development of these ideas; giving very good feedback in how to clarify and make this text as readable as possible. Also, I have been meeting on a regular basis with Dr. Pennig since October 2017. Without these regular meetings this work would not have been possible.

I would also like to thank my second supervisor Dr. Gandalf Lechner for giving some guidelines and counsel on mathematical writing. Also I would like to thank Dr. Padraig Corcoran and Dr. Thomas E. Woolley with whom we set up a TDA reading group during my first years of PhD which was really helpful for the development of these ideas. I would like to thank Dr. Vidit Nanda for his interest and suggestions on this work. Dr. Nanda suggested at the YTM2019 studying the problem that now constitutes chapter 5. Also, I would like to thank Dr. Umberto Lupo for his interest on this work, in particular for his kind help and suggestions to further speed up the PERMAVISS code.

My time at the Cardiff school of mathematics was a great experience, and I felt very welcome and supported. I am very grateful to all PhD colleagues, academics and staff members whose continual help made my PhD a much more lightweight experience. I would like to thank Dr. Hassan Izanloo for his support during my first years at Cardiff, and for proofreading some of the material contained now on chapter 6. Also, I would like to thank Dr. Henry Wilde, Dr. Geraint Palmer and Dr. Nikoleta Glynatsi for their help and support in answering many of my questions about python and general software design practices. I am very thankful to Dr. Christopher Seaman and Dr. Lorenzo de Biase for the interesting discussions about sheaves and algebraic geometry that we had during the first years of my PhD. Also, both Chris and Lorenzo where very helpful and supportive when it came to general questions about doing a PhD at Cardiff. Also, I would like to thank Tracey Lintern for all the invaluable help and advice related to PhD bureaucracy and paperwork. Also, I am very thankful to the GAPT (Geometry, Algebra, Mathematical Physics and Topology) group for inviting many applied topologists to give a seminar talk at Cardiff, which had a very positive impact on this work.

I am especially very thankful to Manuel Soriano-Trigueros and Rocío González-Díaz for their support and help while I was starting to research at *Uniersidad de Sevilla* and finishing the corrections which followed my VIVA.

A journey to a PhD is a long one, and this is why I am very thankful to everyone whose continual support gave me the opportunity to get to Cardiff and do a PhD. First, I would like to thank Prof. Carles Casacuberta who was my undergraduate project supervisor at *Universitat de Barcelona*, and who taught me several important lessons. I am also very thankful to Prof. Casacuberta for inviting me on numerous occasions to the Barcelona topology seminar to talk about this work and get invaluable feedback and suggestions. I am also very thankful to Prof. Frank Neumann who was my supervisor at the University of Leicester, where I started to work on spectral sequences. Also, I am thankful to Prof. Sarah A. Whitehouse with whom I had learned algebraic topology for the first time while I was on an Erasmus exchange at the *University of Sheffield*.

On a more personal level, I am very thankful to all friends that I made in Cardiff whose continual support and help has been invaluable. While I was in Cardiff, I spent two and a half years at Newman Hall, where I made many friends who helped balance out my drift to be thinking all the time about barcodes, such as going to visit the beautiful Welsh natural parks or going to the Opera. I am very thankful to May, Monisha, Tiala, Xavier and Sr. Clare Mary Obi DDL, for their continual support while in Cardiff. Also, I am very thankful to Very. Rev. Fr. Dr. Sebastian Jones for his continual help and advice, and for teaching me some important life lessons, such as how to take care of cats and chickens, or how to properly polish a table.

My friends and family from Barcelona have been an anchor for my continual development and this work success. I am very thankful to my friends Álvaro, Francisco-Javier, Luis and Josemaría from Barcelona for their continual longtime support. This thesis would not have been possible without the constant support, advice and help from my parents Francisco-Javier and Mireia, as well as to my brother Javier and to my sister Claudia. This thesis is also dedicated to them.

Finally, **I would like to express my gratitude to EPSRC for the grant EP/N509449/1 support with project number 1941653, without which I would not have been able to write this work.**

Dissemination of Work

Preprints

- A. Torras Casas. Distributing Persistent Homology via Spectral Sequences. *arXiv:1907.05228*, 2019 [115].
- A. Torras and U. Pennig. Interleaving Mayer-Vietoris spectral sequences. *arXiv:2105.03307*, 2021. [114].

Related Code

- A. Torras Casas. PerMaViss: Persistence Mayer Vietoris spectral sequences. *10.5281/zenodo.3613870*, 2020. [116].

Contents

Chapter 1

Introduction

In recent years Topological Data Analysis (TDA) methods have proven to be of use for problems where linear solutions are not good enough [58, 64]. One of the main and most studied tools in TDA is *persistent homology*, which is an application to data analysis of *homology*, which is widely used in Algebraic Topology. Persistent homology, however, has the disadvantage that sometimes it is computationally expensive and it might also miss important features of a given dataset. In this thesis, we will explore how an important tool from *homological algebra* called a *spectral sequence* relates to persistent homology. Indeed, it has been known for a while that the connections between both objects is an interesting topic to explore [58]. In this study we will review some of the approaches that one might find in the literature. Our emphasis will be mainly on the Mayer-Vietoris spectral sequence related to covering a filtered complex by filtered subcomplexes. We will also present methods and algorithms, as well as some computational implementations done in the python package PERMAVISS [116]. Through this thesis, we present as many examples as possible, so that all theory is directly linked to some application in mind.

1.1 Overview

Persistent Homology

Persistent homology has existed for about two decades [59]. This tool of applied topology has played a central role in applications, such as the study of geometric structure of sets of points lying in $\mathbb{R}^n$, see [49, 59]. This introduced the field of TDA which, very soon, was applied to a multitude of problems, see [21, 64] for a survey article and an introduction. Among many others, persistent homology has been applied to study coverage in sensor networks [46], pattern detection [99], classification and recovery of signals [100] and it has also had an impact on shape recognition using machine learning techniques, see [1, 55]. Other applications in machine learning

came about recently [25]. Another important application of persistence has been in medicine; in studying Chronic Obstructive Pulmonary Diseases [9], tumor vasculate networks [20] and neuroscience [97]. All these applications motivate the need for fast algorithms for computing persistent homology. The usual algorithm used for these computations was introduced in [59], with some later additions to speed up such as those of [32, 33, 47]. In [84] persistent homology is proven to be computable in matrix multiplication time. However, since these matrices become large very quickly, the computations are generally very expensive, both in terms of computational time and in memory required. In practice computing the persistent homology of a given filtered complex is equivalent to computing its matrices of differentials and perform successive Gaussian eliminations; see [58, 59]. In sections 3.2 and 3.3 we will briefly review some of the main properties and definitions related to persistent homology.

Updating Barcodes after small changes

Persistent homology enjoys very good stability properties [36], which is the main reason why it has been used in a multitude of applications. It is known that given a pair of filtered complexes X and Y, the distances between $\mathrm{PH}_*(X)$ and $\mathrm{PH}_*(Y)$ differ no more than X and Y differ to each other, see section 3.3 for details. An interesting question is to ask whether we can know exactly how much $\mathrm{PH}_*(X)$ changes whenever we perform some small local change on X. An example is the case when a pair of filtration values on some cells of X are interchanged, as explored in [39] where these changes are modelled in terms of *Vineyards*. In the case of simplicial complexes, one might use the link condition [51, 52, 29] to test whether removing some simplex will modify simplicial homology or not. An interesting direction is presented by B. Stolz [110], where the input point-cloud data is modified in a way that incurs small changes on persistent homology. We explain briefly this approach in section 3.4. As we will see in this thesis, one might also look at this problem from the point of view of covers and the Mayer-Vietoris spectral sequence.

Persistence Filtration Spectral Sequences

Right from the origins of persistent homology, it has been known that there are important connections between spectral sequences and persistent homology [124, 58]. When computing persistent homology, one starts with a finite filtered complex X. Breaking down this filtration into a finite sequence of filtration values $i_1 < i_2 < i_3 < \cdots < i_N$, we obtain a finite filtered complex

$$F_{i_1}X \subseteq F_{i_2}X \subseteq F_{i_3}X \subseteq \cdots \subseteq F_{i_N}X \subseteq X \ .$$

If $F_{i_N}X = X$, one might consider the *persistence filtration spectral sequence*

$$E^1_{p,q}(X,F) = \mathrm{PH}_q(F_{i_p}X / F_{i_{p-1}}X) \Rightarrow \mathrm{PH}_{p+q}(X) \ .$$

This spectral sequence was originally presented in [57, 58] and has been studied in [104, 2]. An application of this theory is the *spectral sequence method* from [58]. We will review this in section 3.5. In particular, Theorem 3.5.1 leads to an explicit connection between $\mathrm{PH}_*(X)$ and $E^1_{p,q}(X,F)$. This connection will determine that the rank of a given spectral sequence differential is equal to the number of persistence pairs associated to a particular pair of filtration values.

Distributed Computation

In recent years, some methods have been developed for the parallelization of persistent homology. The spectral sequence algorithm from [58] played an important role after it was successfully implemented in [6]. This consists in dividing the original matrix M into groups of rows, and sending these to different processors. These processors will, in turn, perform a local Gaussian Elimination and share the necessary information between them, see [6]. We review this approach in section 3.5, together with some minor modifications that keep the algorithm closer to the standard computation of a spectral sequence.

Perhaps more important for this thesis, one might also attempt to compute persistent homology of a filtered complex X by considering a cover $\mathscr{U}$ by filtered subcomplexes. In this case, one might consider the *Mayer-Vietoris blowup complex* $\Delta_{\mathscr{U}}X$ introduced in [123], where it is shown that there is an isomorphism $\mathrm{PH}_*(\Delta_{\mathscr{U}}X) \simeq \mathrm{PH}_*(X)$. This object was used in [75] for presenting the first algorithm that computed persistent homology by considering local information related to $\mathscr{U}$. This proceeds by computing the sparsified persistent homology for each cover, and then use this information to reduce the differential of $X^{\mathscr{U}}$ efficiently. Both of the presented parallelization methods have provided substantial speedups compared to the standard method presented in [59].

Locality of Persistent Homology

Following the ideas on [123], having an understanding of how persistence barcodes relate to a cover can help us obtain better representatives. In addition, one might also obtain an indication of *locality* of a given class in $\mathrm{PH}_*(X)$ with respect to the cover $\mathscr{U}$ as introduced in [123]. On this basis, it would be desirable to have a method that leads to the speedups from [6, 75], while still keeping cover information from [123]. Further, it would also be useful to consider several covers, such as usually done with functional covers as in the *mapper algorithm*, see [108]. This last point limits substantially the use of the blowup-complex, since the number of simplices grows

very quickly whenever we allow the intersections to grow. In fact, in the extreme case where a complex X is covered by n copies of X, the blowup complex $X^{\mathscr{U}}$ has the enormous complexity size of $2^n|X|$.

Persistence Mayer-Vietoris spectral sequence

Going back to the parallelization question, one might consider applying the Mayer-Vietoris spectral sequence to a filtered complex X together with a filtered cover $\mathscr{U}$. In this case, one is looking at the *persistence Mayer-Vietoris spectral sequence*

$$E^1_{p,q}(X,\mathscr{U}) = \bigoplus_{\sigma^p \in N_{\mathscr{U}}} \mathrm{PH}_q(\mathscr{U}^p_\sigma) \Rightarrow \mathrm{PH}_{p+q}(X) \ .$$

This approach was first presented in [76]. It is no surprise that these objects work in this context, since they have been employed for similar problems for a long time, see [15] or [83]. However, it was noticed that the extension problem was a stumbling block preventing an algorithm implementation; see example 6.1.1 for an intuitive explanation of this problem which we explain further in section 6.2.

The difficulties posed by the extension problem might be the reason why a number of approaches followed, where some hypotheses was taken about the cover $\mathscr{U}$. For example, the case of a cover by two filtered subcomplexes was studied in [54, 93]. It is worth noticing that this particular case corresponds to the *Mayer-Vietoris Theorem* adapted to the persistent case. Another approach was to consider a cover $\mathscr{U}$ such that the nerve $N_{\mathscr{U}}$ is one-dimensional. The reason behind this assumption is that the spectral sequence *collapses* on the second page. This idea was used for distributing computations of cohomology groups in a field in [43] by means of *discrete Morse theory* on *cellular sheaves*. Recently, in [119] and [120] spectral sequences are used for distributing persistent homology computations on these particular covers. In [119, Sec. 4.2] the extension problem was shown to be nontrivial in general and it was solved successfully for the case $\dim(N_{\mathscr{U}}) \leq 1$.

A last approach consists in allowing arbitrary cover nerves, while assuming that each set $\mathscr{U}_\sigma$ is ε-trivial; that is, $\mathrm{PH}_n(\mathscr{U}_\sigma) \sim_\varepsilon 0$ for all $n > 0$ while $\mathrm{PH}_0(\mathscr{U}_\sigma) \sim_\varepsilon \mathrm{I}(a_\sigma, \infty)$ for some values $a_\sigma \geq 0$. In this case one obtains the *approximate nerve theorem* [65] or the *generalized nerve theorem* [28] which puts an upper bound on the *interleaving distance* between $\mathrm{PH}_*(X)$ and $\mathrm{PH}_*(N_{\mathscr{U}})$. Notice that in the bounds obtained in [65] the extension problem plays an important role.

1.2 Contribution and Thesis Structure

In this thesis we study several of the topics introduced in section 1.1. The thesis starts with a pair of chapters 2 and 3, which contain the necessary background in Algebraic Topology and Topological Data Analysis respectively. We would like to point out that this background material contains some non-standard materias, some of which is our own work. In particular, the definition of regular morphism in section 2.6 as well as some of the material from sections 2.12, 2.10, 3.5 and also 3.6. Afterwards, we review a series of topics all of which are related to persistent Mayer-Vietoris spectral sequences as well as generalizations of these.

Persistence Algebra

The spectral sequences considered in this thesis are composed of persistence modules and persistence morphisms. As we would like to compute these spectral sequences, we need an efficient method to compute Images, Kernels and Cokernels of persistence morphisms. This problem was already studied in [38], although with some limitations which we detail at the end of section 3.7. We tackle this problem in an analogous way to the standard linear algebra methods. For this, we start defining *barcode basis* which store compatible persistent bases in a compact way; this is presented in Chapter 4. There, we define barcode bases and also we introduce an operation $\boxplus$ that allows to determine whether a group of *persistence vectors* are linearly independent or not. This machinery, although it might seem artificial, is the key to understanding what it really means to subtract columns from left to right in the Gaussian elimination outlined in Algorithm 4.1. This leads to an `image_kernel` procedure which computes bases for images and kernels, see Algorithm 4.2. Also, it helps us to encapsulate all the information related to a persistence morphism in a matrix that depends on the choice of two barcode bases. This is analogous to the case of linear algebra, where a linear morphisms is given in terms of a matrix relative to a domain and codomain basis. This approach has the advantage over [38] that `image_kernel` works for morphisms between any pair of *tame* persistence modules. In this chapter we also explain how to obtain barcode bases for the quotient of two persistence modules in section 4.4.

Barcode Shifts

In chapter 5, we study the impact of small local changes to persistent homology, with the aim of updating barcodes whenever it is possible. That is, given some filtered complex K and some subcomplex $V \subseteq K$ that we want to add or remove, we would like to know whether it is possible to obtain one of $\mathrm{PH}_*(K)$ and $\mathrm{PH}_*(K \setminus V)$ from the other. For this, we consider a subcomplex $M \subseteq K$ such that $V \subseteq M$ which determines the area where we are willing to compute the *local changes*

between $\mathrm{PH}_*(M)$ and $\mathrm{PH}_*(M \setminus V)$. As we will see in the *barcode shift lemma* 5.1.2, it is possible to solve this problem under some hypothesis that can be checked locally within M. This lemma follows from a particular case of the persistence Mayer-Vietoris spectral sequence. Eventually, one obtains an exact sequence of the form

$$0 \longrightarrow A \longrightarrow \mathrm{PH}_*(K \setminus V) \longrightarrow \mathrm{PH}_*(K) \longrightarrow B \longrightarrow 0\,,$$

where barcode decompositions for A and B have been obtained through examining M. We make heavy use of barcode bases to tackle the problem of deducing either of $\mathrm{PH}_*(K)$ or $\mathrm{PH}_*(K \setminus V)$ from the other; for this we introduce the concepts of embedding, projection and entanglement of barcodes in sections 5.3 and 5.4.

Persistent Mayer-Vietoris spectral sequences

In section 2.11, we give a detailed review of the Mayer-Vietoris spectral sequence in the homology case. Then, in chapter 6 we adapt the aforementioned exposition to the case of the persistence Mayer-Vietoris spectral sequence. The core of the chapter lies in section 6.2, where we give a solution to the extension problem. The solution is given by a careful consideration of the total complex homology, together with the use of barcode basis machinery developed in chapter 4. In section 6.3 we introduce PERMAVISS, an algorithm for computing the persistence Mayer-Vietoris spectral sequence and solving the extension problem. The advantage of this procedure is that all the simplicial information is enclosed within local matrices. This has one powerful consequence; this method consists in computing local Gaussian eliminations plus computing `image_kernel` on matrices whose order is that of homology classes. In particular, given enough processors and a 'good' cover of our data, one has that the complexity is about

$$\mathscr{O}(X^3) + \mathscr{O}(H^3),$$

where X is the order of the maximal local complex and H is the overall number of nontrivial persistence bars on the whole dataset. For more details on this, we refer the reader to section 6.3. We review briefly some example applications in section 6.5.

Regularly Filtered Diagrams Spectral Sequences

In chapter 7 we review a generalization of the Persistent Mayer-Vietoris spectral sequence. In this case, we consider regularly filtered CW-complexes X_*, which are functors $X : \mathbf{R} \to \mathbf{CW\text{-}cpx}$ such that each X_r is a regular complex for all $r \in \mathbf{R}$ and each morphism $X(r \leq s) : X_r \to X_s$ is a

regular morphism of CW-complexes for all pairs $r \leq s$ from $\mathbf{R}$; here $\mathbf{R}$ denotes the category of real numbers as a poset. These regularly filtered CW complexes are a generalization for the standard filtered CW-complexes, where it is assumed that the morphisms $X(r \leq s)$ are inclusions.

One might consider the classical cover of a complex as a *diagram of spaces* over the nerve, where the connecting morphisms are inclusions. We also generalize this point of view by considering diagrams of spaces over simplicial complexes whose morphisms are regular; we call these *Regularly Filtered Diagrams*. This setup is very natural for breaking down non-sparse complexes into pieces. For example, if one wants to parallelize computations over the Vietoris-Rips complex on a point cloud, it is much better to consider the *join diagram* than the traditional cover, as the information on the overlaps is reduced substantially. Given a regularly filtered diagram $\mathscr{F}$ over a simplicial complex K, we introduce the geometric realization $\Delta\mathscr{F}$, together with the associated spectral sequence

$$E^1_{p,q}(K,\mathscr{F}) = \bigoplus_{\sigma^p \in K} \mathrm{PH}_q(\mathscr{F}(\sigma^p)) \Rightarrow \mathrm{PH}_{p+q}(\Delta\mathscr{F}) \ .$$

Interleavings Mayer-Vietoris spectral sequences

In chapter 8 we look at the Mayer-Vietoris spectral sequence as an invariant in its own right. We obtain two stability results: one with respect to local changes and another with respect to changes on the chosen covers. For this we introduce the notion of ε-acyclic equivalences, which give a general framework for defining ε-interleavings between persistence modules. Let $\mathscr{D}$ and $\mathscr{L}$ be a pair of regularly filtered diagrams over a simplicial complex K. If $\mathscr{D}(\sigma)$ and $\mathscr{L}(\sigma)$ are ε-equivalent for all $\sigma \in K$, and assuming also that there are compatibility conditions between these equivalences, then there are ε-acyclic equivalences between the geometric realizations $\Delta\mathscr{D}$ and $\Delta\mathscr{L}$. Furthermore, the spectral sequences $E^*_{*,*}(K,\mathscr{D})$ and $E^*_{*,*}(K,\mathscr{L})$ are interleaved from the first page onwards.

In section 8.4 we consider a filtered complex X together with a refinement of covers $\mathscr{V} \preceq \mathscr{U}$. In this case there exists a morphism $E^*_{*,*}(X,\mathscr{V}) \to E^*_{*,*}(X,\mathscr{U})$ which is unique from the second page onwards. We study conditions as to when there exists an inverse to this refinement induced map. For this, we apply some results from coherent algebraic sheaves [105] to this particular situation. We adapt this result in section 8.5 in order to obtain local comparison conditions for interleavings between $E^*_{*,*}(X,\mathscr{V})$ and $E^*_{*,*}(X,\mathscr{U})$ for a pair of covers $\mathscr{U}$ and $\mathscr{V}$ of X that do not need to refine the other.

Code implementation: PERMAVISS

By using the ideas in this text we developed PERMAVISS, a Python3 library that computes the Persistence Mayer-Vietoris spectral sequence. In chapter 9 we outline the main ideas behind this

implementation. In addition, we include some code snippets illustrating how to use this library. In the examples section 9.6, one can see that nontrivial higher differentials come up and also the extension problem cannot be solved in a trivial way in general. This section is a first approach to implement many of the ideas included in this thesis.

Chapter 2

Background in Algebraic Topology

2.1 Simplicial Complexes

Simplicial complexes are the basic building blocks that will come up in various examples. Let $x_0,\dots,x_n$ be $n+1$ points in $\mathbb{R}^m$ such that they are in general position. Then the convex hull of these points is called a *n-simplex* and it is usually denoted by an ordered $n+1$-tuple $[x_0,\dots,x_n]$. Essentially 0-simplices are vertices, 1-simplices are edges, 2-simplices are faces, and so on, as illustrated on figure 2.1. We define combinations of simplices to form a *simplicial complex*, more formally:

Definition 2.1.1. Given a set X, a *simplicial complex* K is a subset of the *power set* $K \subseteq \mathscr{P}(X)$ such that if $\sigma \in K$, then for all subsets $\tau \subseteq \sigma$ we have that $\tau \in K$. Given a pair of simplicial complexes K and L, if $L \subseteq K$, then we say that L is a *subcomplex* of K.

For a simplicial complex $K \subseteq \mathscr{P}(X)$, we denote by $V(K)$ the subset from X containing an element $x \in X$ if and only if $\{x\} \in K$; this $V(K)$ will be called the *vertex set* of K. In particular, we must have that $K \subseteq \mathscr{P}(V(K))$. An element $\sigma \in K$ will be called a *n-simplex* whenever its number of vertices $|\sigma|$ is equal to $n+1$. If σ is an n-simplex, then we define its dimension to be n, and we indicate it by $\dim(\sigma) = n$. Given two simplices τ and σ from K, if all the vertices from τ are also vertices of σ, then we will say that τ is a *face* of σ, this will be denoted as $\tau \preceq \sigma$. Thus, if a simplex is contained in K all its faces must also be contained in K. Given a simplicial complex K, we will use the notation $\sigma^n \in K$ to mean that σ is a n-simplex from K. The union of simplices of dimension $\leq n$ from K is a subcomplex called the *n-skeleton* and denoted by K^n.

Definition 2.1.2 (Standard m-simplex)**.** Given $m > 0$, we define $\Delta^m = \mathscr{P}(\{0,1,\dots,m\})$, which will be called the *standard m-simplex*. Figure 2.1 illustrates these complexes. This space admits a *geometric realization*

$$|\Delta^m| = \left\{ (x_i)_{i=0}^m \in \mathbb{R}^{m+1} \;\middle|\; \sum_{i=0}^m x_i = 1 \right\},$$

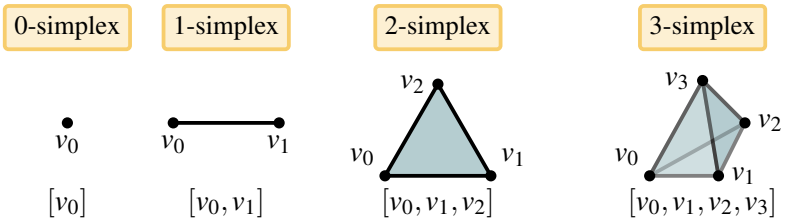


Figure 2.1: Standard simplexes of dimensions 0, 1, 2 and 3.

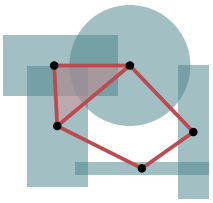


Figure 2.2: In red, the nerve of a cover whose covering regions are represented in gray.

where we have used the notation $(x_i)_{i=0}^m$ instead of the longer format $(x_0, x_1, \ldots, x_m)$.

Definition 2.1.3 (Nerve of a Cover)**.** Let X be a topological space together with a *finite covering set* $\mathscr{U} = \{U_i\}_{i=0}^N$ where each element is a topological subspace $U_i \subseteq X$ such that $X = \bigcup_{i=0}^N U_i$. For each simplex $\sigma \in \Delta^N$, we will use the notation $\mathscr{U}_\sigma := \bigcap_{i \in \sigma} U_i$. Altogether, we define the *nerve* of $\mathscr{U}$ as the simplicial complex

$$N_{\mathscr{U}} = \left\{ \sigma : \mathscr{U}_\sigma \neq \emptyset \right\} \subseteq \Delta^N \, .$$

See figure 2.2 for an example and illustration.

Definition 2.1.4. Given two simplicial complexes $K \subseteq \mathscr{P}(X)$ and $L \subseteq \mathscr{P}(Y)$, we call a function $f : K \to L$ a *simplicial morphism* whenever f respects the face structure, that is, if $\tau \preceq \sigma$ are two simplices from K, then $f(\tau) \preceq f(\sigma)$ in L. Notice that a simplicial morphism $f : K \to L$ is determined by a morphism on its vertex set $f : X \to Y$. That is, a simplex $\sigma = \{x_0, x_1, \ldots, x_n\}$ in K is sent to a simplex $f\sigma = \{f(x_0), f(x_1), \ldots, f(x_n)\}$ in L. This property makes working with simplicial complexes and simplicial morphisms very convenient.

Example 2.1.5. Let $K \subseteq \mathscr{P}(X)$ and $L \subseteq \mathscr{P}(Y)$ be two simplicial complexes. Consider their cartesian product $K \times L$, noticing that it is a subset of $\mathscr{P}(X \times Y)$. In fact, this is a simplicial complex by the following property: a simplex $\tau_1 \times \tau_2$ is a face of another $\sigma_1 \times \sigma_2$ if and only if $\tau_1 \preceq \sigma_1$ and also $\tau_2 \preceq \sigma_2$. We can take the iterative product of complexes $K_1 \times K_2 \times \cdots K_m$, which we will also denote sometimes as $\prod_{i=1}^m K_i$.

Given a simplex $\sigma \in K$, we will use the notation $\Delta^\sigma := \Delta^{\dim(\sigma)}$ for convenience. Given a pair $\tau \preceq \sigma$ in K, we have a corresponding inclusion $\imath(\tau \preceq \sigma) : |\Delta^\tau| \hookrightarrow |\Delta^\sigma|$ which is determined by the inclusion of vertices $V(\Delta^\tau) \hookrightarrow V(\Delta^\sigma)$. On the other hand, consider each set of n-simplices

$K^n \setminus K^{n-1}$ with the discrete topology. This allows to define the *geometric realization* $|K|$ of K as the topological space

$$|K| = \bigsqcup_{n\geq 0} |\Delta^n| \times (K^n \setminus K^{n-1}) \Big/ \sim$$

where $(\iota(\tau \preceq \sigma)(x), \sigma) \sim (x, \tau)$ whenever $\tau \preceq \sigma$ and $x \in |\Delta^\tau|$. We will later in definition 2.8.4 generalize the definition of geometric realization. We will abuse notation and often write Δ^m instead of $|\Delta^m|$. Notice that there is an inclusion $\Delta^\sigma \hookrightarrow |K|$ for any simplex $\sigma \in K$.

Example 2.1.6 (Triangulations)**.** Let M be a manifold embedded in $\mathbb{R}^m$. We define the triangulation T of M to be a simplicial complex such that there is a homeomorphism $|T| \simeq M$. Having a triangulation of a manifold is very convenient computationally, since we have turned an 'infinite' object into a finite combinatorial object.

2.2 Simplicial Homology

Consider a set of vertices $X = \{v_i\}_{i=1}^N$ together with a simplicial complex $K \subseteq \mathscr{P}(X)$. Following section 5 in [85], we pay attention to orientations on simplices. Given an n-simplex $\sigma = \{v_0, v_1, \ldots, v_n\}$, we might order its vertices to obtain a $n+1$-tuple $(v_0, v_1, \ldots, v_n)$. Suppose that we have ordered the vertices from σ by the tuples $(v_{\rho(0)}, v_{\rho(1)}, \ldots, v_{\rho(n)})$ and $(v_{\theta(0)}, v_{\theta(1)}, \ldots, v_{\theta(n)})$ for $\rho, \theta \in \Sigma_{n+1}$; where Σ_{n+1} denotes the *symmetric group* on $n+1$ elements. If ρ and θ differ by an even permutation of their vertices, then we set these to be equivalent and say that they have the same *orientation*. On the contrary, whenever two orderings differ by an odd permutation we say that their orientation is opposite. We will often write a simplex σ as the corresponding equivalence class of tuples up to even permutations, which we will denote by square brackets $[v_0, v_1, \ldots, v_n]$. Let $\mathbb{F}$ be a field. Then for each dimension $n \geq 0$ we define the free vector space over the n-simplices of K

$$S_n^\Delta(K) \coloneqq \bigoplus_{\sigma^n \in K} \mathbb{F}\,.$$

A set of generators for $S_n^\Delta(K)$ is given by ordered n-tuples $[v_0, v_1, \ldots, v_n]$ of vertices $v_i \neq v_j \in X$, and recall that by definition

$$[v_0, v_1, \ldots, v_n] = [v_{\rho(0)}, v_{\rho(1)}, \ldots, v_{\rho(n)}]$$

for any even permutation $\rho \in \Sigma_{n+1}$. On the other hand, we set

$$[v_0, v_1, \ldots, v_n] = -[v_{\rho(0)}, v_{\rho(1)}, \ldots, v_{\rho(n)}]$$

for any odd permutation $\rho \in \Sigma_{n+1}$. Elements $s \in S_n^\Delta(K)$ are usually called *simplicial n-chains* or just *chains* for short. We also consider linear maps $d_n : S_n^\Delta(K) \to S_{n-1}^\Delta(K)$ called *differentials*, defined by

$$d_n([v_0,\dots,v_n]) = \sum_{i=0}^{n}(-1)^i[v_0,\dots,\hat{v_i},\dots,v_n]; \tag{2.1}$$

where the hat notation, $\hat{v_i}$, is used to indicate omission of a vertex. As equation (2.1) is a bit long, we will usually write, for a n-simplex $\sigma^n = [v_0,\dots,v_n]$, the i^{th} face as σ_i^{n-1} instead of $[v_0,\dots,\hat{v_i},\dots,v_n]$. Setting $S_n^\Delta(K) = 0$ for all $n < 0$ we put all of these in a sequence

$$0 \xleftarrow{0} S_0^\Delta(K) \xleftarrow{d_1} S_1^\Delta(K) \xleftarrow{d_2} S_2^\Delta(K) \xleftarrow{d_3} \cdots \tag{2.2}$$

It follows from formula (2.1) that the composition of two consecutive differentials vanishes: $d_n \circ d_{n-1} = 0$ for all $n \geq 0$. In this case we say that (2.2) is a *chain complex*. As a consequence, we have that $\mathrm{Im}(d_{n+1}) \subseteq \mathrm{Ker}(d_n)$, and we can define the *simplicial homology* with coefficients in $\mathbb{F}$ to be the quotient of $\mathbb{F}$-vector spaces

$$\mathrm{H}_n^\Delta(K;\mathbb{F}) = \frac{\mathrm{Ker}(d_n)}{\mathrm{Im}(d_{n+1})},$$

for all $n \geq 0$. In general, $\mathbb{F}$ will be understood from the context and the notation $\mathrm{H}_n^\Delta(K)$ might be used instead. The dimensions of the homology groups encode topological information about the simplex K. Thus the dimension of the 0-homology is the number of connected components of K, the 1-homology group has as many copies of $\mathbb{F}$ as the dimension of the vector space generated by one dimensional cycles in K up to boundaries, and the same for their higher dimensional analogues.

Given two simplicial complexes K and T, is there a way to compare their respective homology groups $\mathrm{H}_n^\Delta(K)$ and $\mathrm{H}_n^\Delta(T)$? We can do this whenever we have a simplicial morphism $f\colon K \to T$. This induces a *chain morphism*

$$f_\# : S_n^\Delta(K) \longrightarrow S_n^\Delta(T)\ ,$$

where each n-simplex $\sigma^n \in K$ is sent to either an n-simplex $f(\sigma^n) \in S_n^\Delta(T)$ if $\dim(f(\sigma^n)) = n$ or, if $\dim(f(\sigma^n)) < n$, it is sent to 0. This $f_\#$ is a chain morphism in the sense that it commutes with the differentials $d_n \circ f_\# = f_\# \circ d_n$ for all $n \geq 0$. Thus, $f_\#$ induces a morphism of homology groups

$$f_*\colon \mathrm{H}_n^\Delta(K) \longrightarrow \mathrm{H}_n^\Delta(T)$$

and this is well defined since by commutativity $f_\#(\mathrm{Im}(d_n)) \subseteq \mathrm{Ker}(d_{n-1})$. This property is summarized by saying that H_n are *functors* from the *category* of simplicial complexes **SpCpx** to the

category of vector spaces **vect**; in section 2.3 we will explain these concepts.

Next, we consider the *augmentation map* $\varepsilon : S_0(K) \to \mathbb{F}$ generated by sending each point $s \in K^0$ to the unit $1_{\mathbb{F}} \in \mathbb{F}$. Then, we define the *reduced homology* by the quotients

$$\widetilde{\mathrm{H}}_0^{\Delta}(K;\mathbb{F}) = \frac{\mathrm{Ker}(\varepsilon)}{\mathrm{Im}(d_1)},$$

and $\widetilde{\mathrm{H}}_n^{\Delta}(K;\mathbb{F}) = \mathrm{H}_n^{\Delta}(K;\mathbb{F})$ for all $n > 0$. Consider the chain complex $\widetilde{S}_*(K)$, obtained by augmenting (2.2) by ε and a copy of $\mathbb{F}$ in degree -1:

$$0 \xleftarrow{0} \mathbb{F} \xleftarrow{\varepsilon} S_0(K) \xleftarrow{d_1} S_1(K) \xleftarrow{d_2} S_2(K) \xleftarrow{d_3} \cdots$$

Then one can see that computing reduced homology is the same as computing homology on $\widetilde{S}_*(K)$.

Example 2.2.1. Consider the standard m-simplex Δ^m. This leads to a chain complex $\widetilde{S}_*(\Delta^m)$

$$0 \xleftarrow{0} \mathbb{F} \xleftarrow{\varepsilon} S_0(\Delta^m) \xleftarrow{d_1} S_1(\Delta^m) \xleftarrow{d_2} S_2(\Delta^m) \xleftarrow{d_3} \cdots \xleftarrow{d_n} S_n(\Delta^m) \xleftarrow{} 0$$

By a standard result $\widetilde{S}_*(\Delta^m)$ is *exact*, that is, $\widetilde{\mathrm{H}}_n^{\Delta}(\Delta^m) = 0$ for all $n \geq 0$. For a proof, see Theorem 8.3 in [85].

Remark. For K being a finite simplicial complex, the computation of the homology of K becomes a standard calculation. As the differentials are linear morphisms between finite dimensional vector spaces, these can be represented by matrices. One can then compute homology by performing successive Gaussian eliminations. Of course, the difficulty lies both in performing such computations efficiently and in following an optimal information storage method.

2.3 Categories and Colimits

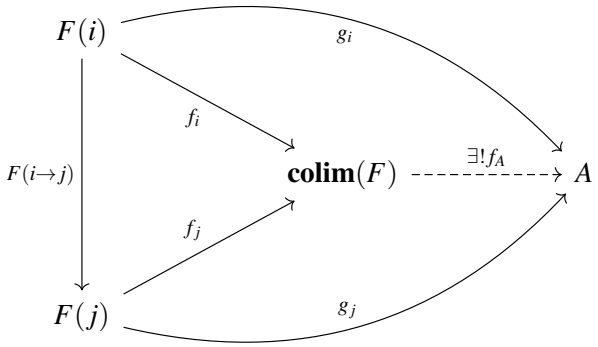
Throughout this text we will be working with basic notions from category theory, such as *categories*, *functors* and *natural transformations*. The reader can see [79, Chap. I] for an introduction to these concepts. Given a category $\mathscr{C}$, we will use the notation $a \in \mathscr{C}$ to indicate that a is an *object* in $\mathscr{C}$. Given a pair of objects $a, b \in \mathscr{C}$, we will write $\hom_{\mathscr{C}}(a,b)$ to indicate the *hom set* of arrows from a to b. A *small* category is such that the objects of $\mathscr{C}$ form a set. A *finite* category is such that it set of objects is a finite set. A *thin* category is a category whose hom sets contain at most one element. Given a category $\mathscr{C}$, we denote by $\mathscr{C}^{\mathrm{op}}$ its *opposite* category which is given by formally inverting all arrows from $\mathscr{C}$.

A *poset* is a small and thin category. Finite posets will play a key role in this thesis. An important poset is $\mathbf{R}$, the poset of real numbers with arrows corresponding to the order relation $\leq$.

Following, there is a list of relevant categories for this work:

- $\mathbf{vect}_{\mathbb{F}}$ the category of finite-dimensional vector spaces and linear morphisms over $\mathbb{F}$.
- **Top** the category of topological spaces and continuous maps.
- **SpCpx** the category of simplicial complexes and simplicial morphisms.
- **ChCpx** the category of chain complexes and chain morphisms.
- **CW-cpx** the category of finite regular CW-complexes and regular morphisms, see section 2.6.
- **RCW-cpx** the category of regularly filtered CW-complexes, that is functors $F : \mathbf{R} \to \mathbf{CW\text{-}cpx}$, this is developed in section 7.1.
- **FCW-cpx** the category of filtered CW-complexes, see section 7.1.
- **SpSq** the category of spectral sequences and spectral sequence morphisms, see section 2.12.
- $\mathbf{SpSq}^{[0,\infty)}$ the category of persistent spectral sequences and persistent spectral sequence morphisms, see section 2.12.

Let $\mathscr{J}$ be a small category; this will be the *indexing category*. Let a category $\mathscr{C}$ and let a functor $F : \mathscr{J} \to \mathscr{C}$. We define the *colimit* (if it exists) of F as an object $\mathbf{colim}(F) \in \mathscr{C}$ together with morphisms $f_j : F(j) \to \mathbf{colim}(F)$ for all $j \in J$ in such a way that for any pair $j, i \in \mathscr{J}$ with an arrow $i \to j$ in $\mathscr{J}$, the commutativity condition $f_j \circ F(i \to j) = f_i$ is satisfied. In addition, the definition of $\mathbf{colim}(F)$ requires that it has a *universal property* in the sense that whenever there exists an object $A \in \mathscr{C}$ together with morphisms $g_j : F(j) \to A$ for all $j \in J$ satisfying commutativity conditions $g_j \circ F(i \to j) = g_i$ for all arrows $i \to j$ from $\mathscr{J}$, there exists a unique morphism $f_A : \mathbf{colim}(F) \to A$ such that the following diagram commutes



for all arrows $i \to j$ from $\mathscr{J}$. Such object A is called a *cocone* of F. The aforementioned universal property ensures that $\mathbf{colim}(F)$ is unique up to isomorphism. In this work we will define colimits

only whenever we are working with categories where they exist. There is a *dual* definition for the limit $\mathbf{lim}(F)$, however, it is not very relevant for this work.

Example 2.3.1. Let K be a simplicial complex. One can think of K as a category with objects its simplices, together with a unique arrow $\tau \to \sigma$ for any pair $\tau \preceq \sigma$ from K. Then, define the *constant* functor $*_K : K^{\mathrm{op}} \to \mathbf{Top}$ given by sending each simplex $\sigma \in K$ to the one-point topological space $*$. Each arrow $\sigma \to \tau$ from K^{op} is sent to the constant morphism $* \to *$; where $\tau \preceq \sigma$. Notice that $\mathbf{colim}(*_K) = \pi_0(|K|)$, where $\pi_0(|K|)$ denotes the set of connected components from $|K|$ with the discrete topology.

Example 2.3.2. Consider a topological space X together with an open cover $\mathscr{U} = \{U_i\}_{i\in I}$, that is, for all $i \in I$ there are open sets $U_i \subseteq X$ such that $\bigcup_{i\in I} U_i = X$. In this case, one can consider this cover as a functor $X^{\mathscr{U}} : N_{\mathscr{U}}^{\mathrm{op}} \to \mathbf{Top}$ given by sending $\sigma \in N_{\mathscr{U}}$ to the corresponding intersection $X^{\mathscr{U}}(\sigma) = \mathscr{U}_\sigma = \bigcap_{i\in\sigma} U_i$. An arrow $\sigma \to \tau$ from $N_{\mathscr{U}}^{\mathrm{op}}$ is sent to the inclusion $X^{\mathscr{U}}(\tau \preceq \sigma) = \mathscr{U}_\sigma \hookrightarrow \mathscr{U}_\tau$. In this case the colimit is the covered space $\mathbf{colim}(X^{\mathscr{U}}) = X$. On the other hand, we consider a simplicial complex K and consider a cover by subcomplexes $\mathscr{A} = \{A_i\}_{i\in I}$. In this case, we can define the functor $K^{\mathscr{U}} : N_{\mathscr{U}}^{\mathrm{op}} \to \mathbf{Top}$ in an analogous way to $X^{\mathscr{U}}$. That is, for each $\sigma \in N_{\mathscr{U}}$ we define $K^{\mathscr{U}}(\sigma) = |\bigcap_{i\in\sigma} A_i|$ and the face maps $K^{\mathscr{U}}(\tau \preceq \sigma)$ are given by inclusions $|\bigcap_{i\in\sigma} A_i| \hookrightarrow |\bigcap_{i\in\tau} A_i|$. Now we have that $\mathbf{colim}(K^{\mathscr{U}}) = |K|$, where the geometric realization $|K|$ has the weak topology with respect to its simplices. That is, a set X is open (resp. closed) whenever $X \cap \iota_\sigma(|\Delta^\sigma|)$ is open (resp. closed) for all simplices $\sigma \in K$; where we have used the inclusion $\iota_\sigma : |\Delta^\sigma| \hookrightarrow |K|$.

Example 2.3.3. Let $(X^i)_{i\in\mathbb{N}}$ be a *filtration* of a topological space X, that is, a sequence of subspaces $X^j \subseteq X$ together with inclusions $X^i \hookrightarrow X^j$ for all integers $0 \leq i \leq j$. This can be thought as a functor $X^{\mathbf{N}} : \mathbf{N} \to \mathbf{Top}$ where $\mathbf{N}$ denotes the subcategory of $\mathbf{R}$ of natural numbers. In this case $\mathbf{colim}(X^{\mathbf{N}}) = \bigcup_{n\in\mathbf{N}} X^n \subseteq X$.

2.4 Singular Homology

In this section we review briefly *singular homology*, which unlike simplicial homology, can be applied to any topological space. Let $X \in \mathbf{Top}$, we define the *singular chain complex* $S_*(X)$ on the n term as

$$S_n(X) = \mathbb{F}\Big[\big\{f : |\Delta^n| \to X \mid f \text{ is continuous.}\big\}\Big]$$

for all $n \geq 0$, and 0 otherwise; where we use the notation $\mathbb{F}[\cdot]$ to indicate the vector space generated by a set. The elements from $S_n(X)$ are called *singular chains*; for ease, we will omit the vertical bars on the domain $|\Delta^n|$ and write Δ^n instead. The singular chain complex has differentials d_n :

$S_n(X) \to S_{n-1}(X)$ generated by the assignment

$$\big(f : \Delta^n \to X\big) \mapsto \sum_{i=0}^{n} (-1)^i (f \circ \iota_i^n : \Delta^{n-1} \hookrightarrow \Delta^n \to X)$$

where $\iota_i^n : \Delta^{n-1} \hookrightarrow \Delta^n$ denotes the inclusion of the $n-1$ standard simplex into the n standard simplex by skiping the i-th vertex from Δ^n. One can check that $d_{n-1} \circ d_n = 0$ for all $n \in \mathbb{N}$ and $S_*(X)$ is indeed a chain complex. Computing homology with respect to the chain complex $S_*(X)$ leads to the singular homology groups $\mathrm{H}_*(X;\mathbb{F})$. Given a continuous map $f : X \to Y$, there is an induced singular chain morphism $f_\# : S_*(X) \to S_*(Y)$ generated by sending each singular chain $\sigma : \Delta^n \to X$ to the singular chain on Y given by composition $f \circ \sigma : \Delta^n \to Y$. One can check that these homology groups define a functor $\mathrm{H}_*(-;\mathbb{F}) : \mathbf{Top} \to \mathbf{vect}_{\mathbb{F}}$. Another important property of homology groups is that these are homotopy invariants; that is, for homotopy equivalences $X \simeq Y$ there are isomorphisms $\mathrm{H}_n(X;\mathbb{F}) \simeq \mathrm{H}_n(Y;\mathbb{F})$ for all $n \geq 0$, see [68, Sec. 2.1.].

Example 2.4.1. Recall that in Section 2.2 we presented simplicial homology $\mathrm{H}_*^\Delta(K)$ for a simplicial complex K. One can show [85, §34] that there is an isomorphism of homology groups $\mathrm{H}_*^\Delta(K) \simeq \mathrm{H}_*(|K|)$ which is induced by a chain complex morphism $C_*^\Delta(K) \to C_*(|K|)$ sending an ordered tuple $[v_0, v_1, \ldots, v_n] \in C_n^\Delta(K)$ to the linear map $l : \Delta^n \to |K|$ given by $l((x_i)_{i=0}^n) = \sum_{i=0}^n x_i v_i$ for all points $(x_i)_{i=0}^n \in \Delta^n$. This is why we will usually talk about the "homology" of K without distinguishing between the two. Of course, when it comes to actual computations, one works with simplicial homology. Singular homology on the other hand has useful theoretical properties which we will be using from time to time.

As with simplicial homology, we can also consider an augmentation map $\varepsilon : S_0(X) \to \mathbb{F}$ sending each singular chain $p : \Delta^0 \to X$ to the unit $\varepsilon(p) = 1_{\mathbb{F}} \in \mathbb{F}$. This leads to the reduced singular chain complex $\widetilde{S}_n(X)$ whose homology groups are the reduced singular homology groups $\widetilde{\mathrm{H}}_n(X)$. A chain complex C_* with differentials $d_n : C_n \to C_{n-1}$ is said to be *exact* whenever all its homology groups $\mathrm{H}_*(C_*)$ vanish. An exact sequence of the form

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is called a *short exact sequence*.

Suppose that one has a pair of topological spaces $A \subseteq X$. Then, one can consider the singular chain complexes $S_*(A)$ and $S_*(X)$ and take their respective quotients, obtaining short exact sequences

$$0 \longrightarrow S_n(A) \longrightarrow S_n(X) \longrightarrow \frac{S_n(X)}{S_n(A)} \longrightarrow 0\,,$$

for all $n \geq 0$. Taking homology with respect with the *relative chain complexes* $S_n(X)/S_n(A)$ leads to *relative homology* groups, which are denoted by $\mathrm{H}_*(X,A)$. In this context, there is a *long exact sequence*

$$\cdots \longrightarrow \mathrm{H}_n(A) \xrightarrow{\ \iota\ } \mathrm{H}_n(X) \xrightarrow{\ q\ } \mathrm{H}_n(X,A) \xrightarrow{\ \partial_n\ } \mathrm{H}_{n-1}(A) \longrightarrow \cdots \tag{2.3}$$

which is obtained by the use of the *snake lemma* [118, §1.3]. In particular, notice that there is an isomorphism $\mathrm{H}_n(X,*) \simeq \widetilde{\mathrm{H}}_n(X)$ for all $n \in \mathbb{N}$. An important property of relative homology is given by *excision*. Suppose that there are three topological spaces $Z \subseteq A \subseteq X$ and such that the closure of Z is contained in the interior of A; in such a case we say that (A,Z) is an *excisive pair*. Then the inclusion $(X \setminus Z, A \setminus Z) \hookrightarrow (X,A)$ induces an isomorphism in homology $\mathrm{H}_*(X \setminus Z, A \setminus Z) \simeq \mathrm{H}_*(X,A)$. As a consequence, if (X,A) is a 'good pair' in the sense that A is a *deformation retract* of some neighbourhood of A in X, one can deduce the isomorphisms

$$\mathrm{H}_n(X,A) \simeq \mathrm{H}_n(X/A,*) \simeq \widetilde{\mathrm{H}}_n(X/A)\ ,$$

see for example [68, Prop. 2.22] (here X/A denotes the quotient topological space). Thus, the sequence (2.3) can be used for computing the quotient homology $\widetilde{\mathrm{H}}_n(X/A)$ from the factors $\mathrm{H}_n(X)$ and $\mathrm{H}_n(A)$.

One of the main advantages of homology is its computability compared to other homotopy invariants such as the homotopy groups. For simplicial complexes this becomes even more clear as one can use the formulas from section 2.2. Another advantage of homology comes from results such as the Mayer-Vietoris sequence (see section 2.11) or the Künneth formula. Such results assume that a space X is composed of simpler pieces, such as being the union of subspaces $\bigcup_{i\in I} U_i$ or the product of spaces $\prod_{i\in I} Y_i$. This is something we will discuss later in Section 2.8.

2.5 CW-complexes and Cellular Homology

There are some situations in which it is not practical to work with simplicial complexes. For instance, it is not always easy to find a triangulation of a space X. From the more computational perspective, a triangulation for X might contain too much unnecessary information leading to a waste of computational time and memory. To address both problems, we consider complexes where the building blocks are more general than simplices. Here we will closely follow the exposition in chapter IX from [80]. To start, let us define the closed disk

$$D^n = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$$

the open disk

$$U^n = \{x \in \mathbb{R}^n \mid |x| < 1\}$$

and the n-dimensional sphere

$$S^n = \left\{x \in \mathbb{R}^{n+1} \mid |x| = 1\right\} .$$

Definition 2.5.1. A *CW-complex* X is a Hausdorff space together with a filtration

$$X^0 \subseteq X^1 \subseteq \cdots \subseteq X^i \subseteq \cdots$$

such that

1. X^0 a disjoint union of points.

2. Each $X^n \subseteq X$ is a closed subset called the n-skeleton of X. Each complement set $X^n \setminus X^{n-1}$ is a disjoint union of open subsets e^n_λ in X^n, for all $\lambda \in \Lambda$. These e^n_λ are called *open n-cells* or just *n-cells* and are such that e^n_λ is homeomorphic to the open disk U_n for all $\lambda \in \Lambda$. These cells are such that there are characteristic maps $f_\lambda : D^n \to \overline{e^n_\lambda}$ sending U^n homeomorphically into e^n_λ for all $\lambda \in \Lambda$.

3. $X = \mathbf{colim}_{n\in\mathbb{N}} X^n$

For a CW-complex X, if c is a cell in X we will follow the notation from [41] and denote this by $c \in X$. Also, we will follow notation from [80] and use $\dot{c} = \overline{c} \setminus c$ for the boundary of any cell $c \in X$. We will call each X^n the *n-skeleton* from X. Given two CW-complexes X and Y, a continuous map $f : X \to Y$ is said to be a *cellular morphism* whenever it respects filtrations; that is, f restricts to morphisms $f^n : X^n \to Y^n$ for all $n \geq 0$. The acronym "CW" comes from the topology of X:

- **C** *closure finiteness*: a compact subset $K \subseteq X$ intersects only finitely many open cells from X. In particular, the closure $\overline{e^n_\lambda}$ of each n-cell intersects finitely many cells.

- **W** *weak topology*: the topology of X is induced by the topological colimit construction; a set A is open (resp. closed) if and only if $A \cap e$ is open (resp. closed) for all cells $e \in X$.

Let X be a CW-complex. The *cellular chain complex* $C^{\text{cell}}_*(X)$ is defined as the relative groups

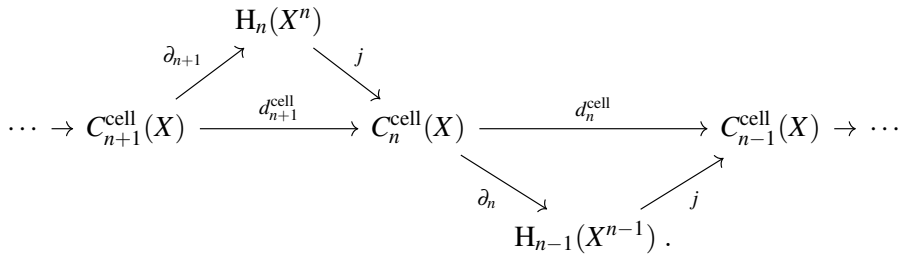
$$C^{\text{cell}}_n(X) = \mathrm{H}_n\left(X^n, X^{n-1}\right)$$

for all $n \geq 0$, with $C^{\text{cell}}_m(X) = 0$ for all $m < 0$. For each cell $e^n_\lambda \in X$, notice that the characteristic

map f_λ induces an isomorphism $\mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda) \simeq \mathrm{H}_n(D^n, S^{n-1})$. Thus, there are isomorphisms

$$\mathrm{H}_n\big(X^n, X^{n-1}\big) \simeq \bigoplus_{\lambda\in\Lambda} \mathrm{H}_n\big(\overline{e^n_\lambda}, \dot{e}^n_\lambda\big) \simeq \mathbb{F}\Big(\big\{e^n_\lambda \mid \lambda \in \Lambda\big\}\Big)\,,$$

and so the dimension of $C^{\mathrm{cell}}_n(X)$ is equal to the number of n-cells from X. On the other hand, the *cellular* differentials d^{cell}_* are defined by commutativity on the diagram:



The diagonal morphisms come from the relative homology long exact sequence (2.3), in particular, notice that $d^{\mathrm{cell}}_{n-1} \circ d^{\mathrm{cell}}_n = 0$. Computing homology on the cellular chain complex leads to cellular homology $\mathrm{H}^{\mathrm{cell}}_*(X)$, which is shown to be isomorphic to singular homology $\mathrm{H}_*(X)$ by carefully inspecting the involved relative long exact sequences, see [68, Thm. 2.35].

As with simplicial homology, we would like to use cellular chain complexes to obtain homology groups. In order to do this, we will need to define orientations for cells in an analogous way to the case of ordered tuples of vertices on a simplicial complex that we saw in Section 2.2. One chooses such orientations fixing a basis for each term $C^{\mathrm{cell}}_n(X)$, by choosing a generator $a^n_\lambda \in \mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda) \simeq \mathbb{F}$ which is usually called an *orientation* of the cell e^n_λ. However, for computational convenience we pay attention to homology coefficients when choosing such orientations. Usually homology is presented with $\mathbb{Z}$ coefficients instead of field coefficients, as we have done here. However, the *universal coefficient formula* relates the homology groups $\mathrm{H}_n(X;\mathbb{F})$ and $\mathrm{H}_n(X;\mathbb{Z})$ by means of short exact sequences

$$0 \longrightarrow \mathrm{H}_n(X;\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{F} \longrightarrow \mathrm{H}_n(X;\mathbb{F}) \longrightarrow \mathrm{Tor}(\mathrm{H}_{n-1}(X;\mathbb{Z}),\mathbb{F}) \longrightarrow 0$$

see [80, §6]. Thus, for our particular case there is an isomorphism

$$\mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda;\mathbb{F}) \simeq \mathbb{F} \simeq \mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda;\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{F}\,.$$

We pick a basis for each term $C^{\mathrm{cell}}_n(X;\mathbb{Z})$ by choosing a generator $a^n_\lambda \in \mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda;\mathbb{Z}) \simeq \mathbb{Z}$ which in turn determines a generator b^n_λ for $\mathrm{H}_n(X^n, X^{n-1};\mathbb{Z})$. By the universal coefficient formula, this determines a generator $b^n_\lambda \otimes_{\mathbb{Z}} 1_{\mathbb{F}}$ for $\mathrm{H}_n(X^n, X^{n-1};\mathbb{F})$, which we will also denote by b^n_λ for convenience. Once we have these local choices of orientations for the cells, we can define *incidence*

numbers $[b_\beta^{n-1} : b_\alpha^n] \in \mathbb{Z}$ which are such that

$$d_n^{\text{cell}}(b_\lambda^n) = \sum_{\mu \in \Lambda} [b_\lambda^n : b_\mu^{n-1}] b_\mu^{n-1} \ .$$

Of course, these incidence numbers depend on the particular choices of orientations that we have made over all cells of X. The incidence numbers have the following properties which are proved in [80, Lem. 5.1]:

(I. 1) For any cell e_λ^n, $[b_\lambda^n : b_\mu^{n-1}] = 0$ if e_μ^{n-1} is not contained in $\overline{e_\lambda^n}$.

(I. 2) For any n-cell e_λ^n and any $(n-2)$-cell e_λ^{n-2}

$$\sum_\mu [b_\lambda^n : b_\mu^{n-1}][b_\mu^{n-1} : b_\nu^{n-2}] = 0$$

(I. 3) For any 1-cell e_λ^1

$$\sum_\mu [b_\lambda^1 : b_\mu^0] = 0$$

(I. 4) $-[b_\lambda^n : b_\mu^{n-1}] = [-b_\lambda^n : b_\mu^{n-1}] = [b_\lambda^n : -b_\mu^{n-1}]$.

For a finite CW-complex, the differential d_n^{cell} can be seen as a $|X^n| \times |X^{n-1}|$-matrix with $\mathbb{Z}$-entries. A possible strategy to compute the incidence numbers $[b_\lambda^n : b_\mu^{n-1}]$ is by computing the *degree* $d_{\mu\lambda} \in \mathbb{Z}$ determining the bottom arrow on the diagram:

$$\begin{array}{ccccc}
\mathrm{H}_n(X^n) & & \longrightarrow & & \mathrm{H}_{n-1}(X^{n-1}) \\
\uparrow & & & & \uparrow \\
\mathbb{F} & & \xrightarrow{[b_\lambda^n : b_\mu^{n-1}]} & & \mathbb{F} \\
\downarrow \simeq & & & & \downarrow \simeq \\
\mathrm{H}_n(\overline{e_\lambda^n}, \dot{e}_\lambda^n) & & \longrightarrow & & \mathrm{H}_{n-1}(\overline{e_\mu^{n-1}}, \dot{e}_\mu^{n-1}) \\
\downarrow \simeq & & & & \downarrow \simeq \\
\widetilde{\mathrm{H}}_{n-1}(\dot{e}_\lambda^n) & \hookrightarrow & \mathrm{H}_{n-1}(X^{n-1}, X^{n-2}) & \twoheadrightarrow & \mathrm{H}_{n-1}(X^{n-1}, X^{n-1} \setminus e_\mu^{n-1}) \\
\uparrow \simeq & & & & \downarrow \simeq \\
\widetilde{\mathrm{H}}_{n-1}(S^{n-1}) & & \xrightarrow{d_{\mu\lambda}} & & \widetilde{\mathrm{H}}_{n-1}(S^{n-1}) \ .
\end{array}$$

For details of such an approach, as well as degree computations, check [68, §2.2.].

Example 2.5.2. Let K be a simplicial complex. In particular, notice that $|K|$ has the structure of a CW-complex, with filtration $|K^0| \subseteq |K^1| \subseteq \cdots \subseteq |K^i| \subseteq \cdots$. Recall that in example 2.4.1 we saw a morphism $l : S_n^\Delta(K) \to S_n(|K|)$ sending a simplex σ to the linear map $l(\sigma) : \Delta^\sigma \to K$ determined by the vertices from σ and which induces an isomorphism in homology. In particular,

$l(\sigma)$ is a generator of $\mathrm{H}_n(\overline{\sigma}, \dot{\sigma})$ and consequently determines an orientation. The singular boundary formula on $l(\sigma)$ and the simplicial boundary formula on σ coincide in the following sense: $d_n(l(\sigma)) = \sum_{i=0}^{n}(-1)^i l(\sigma)_{|\sigma_i} = \sum_{i=0}^{n}(-1)^i l(\sigma_i) = l(d_n(\sigma))$, in particular $[l(\sigma) : l(\sigma_i)] = (-1)^i$ for all $0 \leq i \leq n$.

Example 2.5.3. Consider a pair of finite CW-complexes X and Y. There is a very convenient description of the cellular chain complex of $X \times Y$ by means of an isomorphism (see [68, Thm 3.16])

$$C_k^{\mathrm{cell}}(X \times Y) \simeq \bigoplus_{i+j=k} C_i^{\mathrm{cell}}(X) \otimes C_j^{\mathrm{cell}}(Y)$$

sending a cell $a \times b \in X \times Y$ to the tensor product $a \otimes b$. The differential $d_n^{\otimes}$ is generated by

$$d^{\otimes}(a \otimes b) = d(a) \otimes b + (-1)^{\dim(a)} a \otimes d(b) \ .$$

In particular, if we have two finite simplicial complexes K and L and a product of simplices $\sigma \times \tau \in K \times L$ we can then describe the differential as

$$d^{\mathrm{cell}}(\sigma \times \tau) = \sum_{i=0}^{\dim(\sigma)} (-1)^i \sigma_i \times \tau + (-1)^{\dim(\sigma)} \sum_{j=0}^{\dim(\tau)} (-1)^j \sigma \times \tau_j \ .$$

Example 2.5.4. Let us consider complexes whose set of vertices V form a subset of a lattice $V \subseteq \mathbb{Z}^m$ for some $m \in \mathbb{N}$. Given a family I_i of copies of unit-length intervals $[n_i, n_i + 1]$ or singleton sets $\{n_i\}$ with $n_i \in \mathbb{Z}$ for all $1 \leq i \leq m$, we define a cube as the product $q = I_1 \times \cdots \times I_m$. The dimension of a cube q is equal to the sum of dimensions of its factors $\dim(q) = \sum_{i=1}^{m} \dim(I_i)$. For each $1 \leq i \leq m$, if I_i is an interval, then there are two faces of q given by

$$A^i(q) = I_1 \times \cdots \times I_{i-1} \times \{n_i + 1\} \times I_{i+1} \times \cdots \times I_n$$

and

$$B^i(q) = I_1 \times \cdots \times I_{i-1} \times \{n_i\} \times I_{i+1} \times \cdots \times I_n.$$

whose respective incidences are (by example 2.5.3)

$$[q : A^i(q)] = (-1)^{\sum_{i=1}^{i-1} \dim(I_i)} \qquad \text{and} \qquad [q : B^i(q)] = (-1)^{1+\sum_{i=1}^{i-1} \dim(I_i)} \ .$$

We say that a set of cubes $\mathscr{C}$ is a *cubical complex* if and only if for each cube contained in $\mathscr{C}$, all its faces are also contained in $\mathscr{C}$. Given such a $\mathscr{C}$, one can define a chain complex by setting $C_n(\mathscr{C})$ to be the free $\mathbb{F}$-vector space on the n-cubes of $\mathscr{C}$. We define the differential $\delta_n \colon C_n(\mathscr{C}) \to C_{n-1}(\mathscr{C})$

as

$$\delta_n(q) = \sum_{\substack{0\leq i\leq n \\ \dim(I_i)=1}} (-1)^{\sum_{i=1}^{i-1}\dim(I_i)}(A^i(q) - B^i(q)).$$

for each $q \in C_n(\mathscr{C})$. Without difficulties, one can check that $\delta_{n-1}\delta_n(q) = 0$. Thus one can define the *cubical homology* of $\mathscr{C}$ in an analogous way to simplicial homology, see [70, Chp. 2] for an introduction to cubical homology. Cubical complexes are a particular case of *regular CW-complexes*, which we will review through the following section.

2.6 Regular CW-complexes

We would like to consider now CW-complexes with properties analogous to those of simplicial complexes. A CW-complex X is said to be *regular* whenever all attaching maps $f_\lambda : D^n \to \overline{e^n_\lambda}$ are homeomorphisms onto their images. This avoids gluing a boundary multiple times to the same cell in the image. Regular complexes and their properties are presented thoroughly in [41] and in [80]. Here we will state three properties of regular CW-complexes, all of which are proved in [41, VIII, § 4].

(R. 1) For a pair of cells $a, b \in X$ if $\dim(b) < \dim(a)$ and if we have $\overline{a} \cap b \neq \emptyset$, then $b \subseteq \overline{a}$.

(R. 2) For any n-cell e^n, for $n \geq 0$, $\overline{e^n}$ and $\dot{e}^n$ are the underlying spaces of subcomplexes from X. In addition, $\dot{e}^n$ is the union of closures of $(n-1)$-cells forming a $(n-1)$-sphere. This follows from adapting the language from Theorem 4.1. in [41, VIII, § 4].

(R. 3) If e^n_λ and e^{n-2}_ρ are two cells from X such that $e^{n-2}_\rho \preceq e^n_\lambda$, then there are precisely two $(n-1)$-cells e^{n-1}_μ and e^{n-1}_ν such that $e^{n-2}_\rho \preceq e^{n-1}_\mu \preceq e^n_\lambda$ and $e^{n-2}_\rho \preceq e^{n-1}_\nu \preceq e^n_\lambda$.

Thanks to property **(R. 1)** it makes sense to speak about a cell b being a *face* of another cell a from X, which we will denote as $b \preceq a$. In particular, one can consider the poset of incidences where each object is a cell from X and there is an arrow $a \to b$ whenever $a \preceq b$. In fact this category can be seen as a simplicial complex $\mathrm{Bd}(X)$ called the *Barycentric subdivision* of X which has the same homotopy type than X. Using this fact one could have used the definition of simplicial homology to define alternatively the cellular chains for regular complexes; this approach is taken in [60, §2.1.]. Another consequence of **(R. 1)** together with property **(I. 1)** is that if $[b^n_\lambda : b^{n-1}_\mu] \neq 0$ then $e^{n-1}_\mu \preceq e^n_\lambda$.

Property **(R. 2)** implies that for any face $e^{n-1}_\mu \preceq e^n_\lambda$ one has $[b^n_\lambda : b^{n-1}_\mu] = \pm 1$. This follows from noticing that both $\overline{e^n}$ and $\dot{e}^n_\lambda \setminus e^{n-1}_\mu$ are contractible, and by examining the long exact sequences

for the pairs $(\overline{e^n_\lambda}, \dot{e}^n_\lambda)$ and $(\dot{e}^n_\lambda, \dot{e}^n_\lambda \setminus \dot{e}^{n-1}_\mu)$ there are isomorphisms in homology with $\mathbb{Z}$ coefficients:

$$\begin{array}{ccccccc}
\mathbb{Z} & & \xrightarrow{\quad [b^n_\lambda : b^{n-1}_\mu] \quad} & & & & \mathbb{Z} \\
\downarrow \simeq & & & & & & \uparrow \simeq \\
\mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda; \mathbb{Z}) & \xrightarrow{\simeq} & \widetilde{\mathrm{H}}_{n-1}(\dot{e}^n_\lambda; \mathbb{Z}) & \xrightarrow{\simeq} & \mathrm{H}_{n-1}(\dot{e}^n_\lambda, \dot{e}^n_\lambda \setminus e^{n-1}_\mu; \mathbb{Z}) & \xrightarrow{\simeq} & \mathrm{H}_{n-1}(\overline{e^{n-1}_\mu}, \dot{e}^{n-1}_\mu; \mathbb{Z}) \,,
\end{array}$$

where the last isomorphism on the bottom row follows from regarding e^{n-1}_μ as a cell of the complex $\dot{e}^n_\lambda$ and using the excision property of homology.

Using property (**I. 2)** together with (**R. 3)**, we deduce that

$$[b^n_\lambda : b^{n-1}_\mu][b^{n-1}_\mu : b^{n-2}_\rho] + [b^n_\lambda : b^{n-1}_\nu][b^{n-1}_\nu : b^{n-2}_\rho] = 0 \,.$$

By the same argument, if e is an edge with vertices A and B, then

$$[e : A] + [e : B] = 0 \,.$$

These results justify restricting our attention to such complexes, as these regular complexes have good computational properties. There is a converse for the properties of the incidence values which is in [80, Thm. 7.2]; that is, one can start with values satisfying the properties of incidences and determine an orientation of X from these. There is a GAP (Groups, Algorithms, and Programming) package called HAP (Homological Algebra Programming) which contains a command for computing incidences on regular complexes, see [112] and [61] as well as [60, §2.1].

We will say that a cellular morphism $f : X \to Y$ is a *regular morphism* if for all cells $e^n \in X$, the closure of their image $\overline{f(e^n)}$ is a subcomplex of Y such that either

- $\overline{f(e^n)}$ has dimension $< n$,
- f maps $\overline{e^n}$ homeomorphically into $\overline{f(e^n)}$.

We will write **CW-cpx** to denote the category of finite regular CW-complexes and regular morphisms. Given a regular morphism $f : X \to Y$, there is an induced chain map $f_\# : C^{\mathrm{cell}}_*(X) \to C^{\mathrm{cell}}_*(Y)$ such that for any cell $e^n_\lambda \in X$ one has that

$$f(b^n_\lambda) = \sum_{\mu \in \Lambda} [f(b^n_\lambda) : b^n_\mu] b^n_\mu$$

These incidences $[f(b^n_\lambda) : b^n_\mu]$ are defined by following the morphisms

$$\mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda) \longrightarrow \mathrm{H}_n(X^n, X^{n-1}) \xrightarrow{f} \mathrm{H}_n(Y^n, Y^{n-1}) \longrightarrow \mathrm{H}_n(Y^n, Y^n \setminus e^n_\mu) \xrightarrow{\simeq} \mathrm{H}_n(\overline{e^n_\mu}, \dot{e}^n_\mu) \,.$$

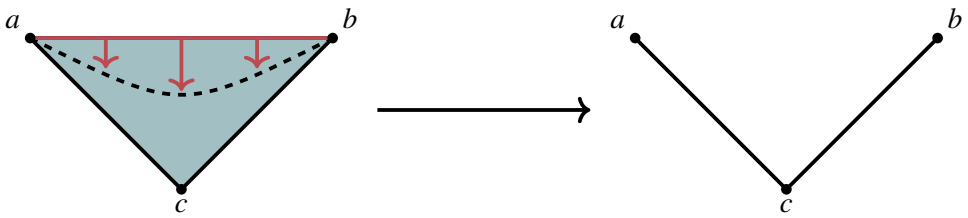


Figure 2.3: Vertical projection of a simplex $\overline{e^2}$ onto two of its faces. This is in fact a homotopy equivalence as indicated by the red arrows projection onto a middle arc on the left figure.

In particular, notice that if $\overline{f(e^n_\lambda)} \cap e^n_\mu = \emptyset$ then $[f(b^n_\lambda) : b^n_\mu] = 0$ as in this case $f(\overline{e^n_\lambda}) \subseteq Y^n \setminus e^n_\lambda$. Also, if $\dim(\overline{f(e^n_\lambda)}) < n$, by the fact that $\overline{f(e^n_\lambda)}$ is a subcomplex of Y^{n-1} we have that $[f(b^n_\lambda) : b^n_\mu] = 0$ for all $\mu \in \Lambda$. On the other hand, if $\dim(\overline{f(e^n_\lambda)}) = n$ and if $\overline{f(e^n_\lambda)} \cap e^n_\mu \neq \emptyset$, then we can follow the isomorphisms on $\mathbb{Z}$ coefficients

$$\mathrm{H}_n(\overline{e^n_\lambda}, \dot{e}^n_\lambda; \mathbb{Z}) \xrightarrow{\simeq} \mathrm{H}_n(\overline{f(e^n_\lambda)}, \partial f(e^n_\lambda); \mathbb{Z}) \xrightarrow{\simeq} \mathrm{H}_n(\overline{e^n_\mu}, \dot{e}^n_\mu; \mathbb{Z})$$

which determine $[f(b^n_\lambda) : b^n_\mu] = \pm 1$; here we have denoted by $\partial f(e^n_\lambda)$ the boundary of $f(e^n_\lambda)$ which is homeomorphic to the $n-1$-sphere by regularity hypotheses on f.

Example 2.6.1. Consider a product $X \times Y$ between two regular CW-complexes. And consider the projections $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ which of course are regular morphisms. For a cell $e^n_x \times e^m_y \in X \times Y$ with orientation $b^n_x \times b^m_y$ one can check by using the differential formula for the product that

$$[\pi_X(b^n_x \times b^m_y) : b^n_x] = \begin{cases} 1 \text{ if } m = 0, \\ 0 \text{ otherwise.} \end{cases}$$

and

$$[\pi_Y(b^n_x \times b^m_y) : b^n_y] = \begin{cases} 1 \text{ if } n = 0, \\ 0 \text{ otherwise.} \end{cases}$$

The case when X and Y are simplicial complexes will be relevant for this work.

Example 2.6.2. Consider a standard 2-simplex Δ^2 with vertices a, b and c as indicated on figure 2.3. One can consider the projection of Δ^2 onto two of its faces $L = \overline{[a,c]} \cup \overline{[b,c]}$, which is a regular morphism $\pi : \Delta^2 \to \overline{[a,c]} \cup \overline{[b,c]}$. In fact, one can build a homotopy equivalence between K and L by following a projection of the region enclosed between the face $[a,b]$ and the middle dashed arc on the left of figure 2.3. In this case, one can easily compute the coefficients defining a morphism $\pi_* : C^{\mathrm{cell}}_*(\Delta^n) \to C^{\mathrm{cell}}_*(L)$; notice in particular that

$$\pi_*([a,b]) = [a,c] - [b,c] \ .$$

This example, although naive, is related to Discrete Morse Theory [86], which is a machinery used for reducing simplicial complex computations. This also motivates the use of regular morphisms and regular complexes. Another important point from this example is to show that regular morphisms on simplicial complexes are more flexible than simplicial morphisms.

2.7 Chain Homotopies and Acyclic Carriers

In section 2.4 we mentioned that singular homology is invariant up to homotopy equivalences. This is because given a homotopy $h : X \times I \to Y$ between two continuous maps $f, g : X \to Y$, one might construct a *chain homotopy* between the corresponding singular chain maps $f_*, g_* : S_*(X) \to S_*(Y)$. We define now this.

Definition 2.7.1. Let (C_*, d^C) and (D_*, d^D) be a pair of chain complexes and consider a pair of chain morphisms $f_*, g_* : C_* \to D_*$. A *chain homotopy* between f_* and g_* is a sequence of morphisms $H_n : C_n \to D_{n+1}$ such that

$$f_n - g_n = H_{n-1} \circ d^C_n + d^D_{n+1} \circ H_n$$

for all degrees $n \in \mathbb{Z}$.

In particular, given a pair of chain homotopic morphisms $f_*, g_* : S_*(X) \to S_*(Y)$ their induced morphsims in homology are equal $H_*(f) = H_*(g)$.

Changing the topic, recall that on previous sections we have seen how a continuous map $f : X \to Y$ induces a morphism on chain complexes, e.g. on singular chains $f_* : S_*(X) \to S_*(Y)$, and how this induces in turn a morphism at the level of homology $f_* : \mathrm{H}_*(X) \to \mathrm{H}_*(Y)$. This allows us to compare the homology groups of X and Y respectively. However, what would we do if we wanted to compare $\mathrm{H}_*(X)$ and $\mathrm{H}_*(Y)$ but had no map $f : X \to Y$? Fortunately, homology theory contains a notion that allows to look at homology morphisms induced by "blurred" maps.

Let X and Y be two regular CW-complexes with their respective cellular chain complexes $\left(C^{\mathrm{cell}}_*(X), d^{\mathrm{cell}}_X\right)$ and $\left(C^{\mathrm{cell}}_*(Y), d^{\mathrm{cell}}_Y\right)$. Let $\langle \cdot, \cdot \rangle_X$ and $\langle \cdot, \cdot \rangle_Y$ denote the inner products on $C^{\mathrm{cell}}_*(X)$ and $C^{\mathrm{cell}}_*(Y)$ which are fixed by the fact that the oriented cells form an orthonormal basis. Thus, there is an equality $\langle b^m_\mu, d^{\mathrm{cell}}_X(b^n_\lambda) \rangle_X = [b^n_\lambda : b^m_\mu]$.

Definition 2.7.2. A *carrier* $F : X \rightrightarrows Y$ is a map from the set of cells of X to subcomplexes of Y that is semicontinuous in the sense that for any pair $\tau \prec \sigma$ in X, $F(\tau) \subseteq F(\sigma)$. A carrier $F : X \rightrightarrows Y$ is called *acyclic*, if for every $\sigma \in X$, $F(\sigma)$ is a nonempty acyclic subcomplex of Y.

Given either a chain map (resp. a chain homotopy) $w_p : C^{\mathrm{cell}}_n(X) \to C^{\mathrm{cell}}_{n+r}(Y)$ with $r = 0$ (resp.

r=1) and for all $n \geq 0$, we say that w_p is carried by F if for all cells $e^n_\lambda \in X^n$

$$\{e^{n+r}_\mu \in Y^{n+r} \mid \langle w_p(b^n_\lambda), b^{n+r}_\mu \rangle_Y \neq 0\} \subseteq F(\sigma) \ .$$

The next statement is an application of [85, Thm. 13.4]. In Proposition 8.1.2 we will prove a version of this statement that will apply to filtered CW-complexes.

Theorem 2.7.3. *Let $F : X \rightrightarrows Y$ be an acyclic carrier between CW-complexes X and Y. Then we have that*

- **existence:** *there is a chain map carried by F,*
- **equivalence:** *if F carries two chain maps f and g, then F carries a chain homotopy between f and g.*

Given two acyclic carriers $F, G : X \rightrightarrows Y$, we write $F \subseteq G$ whenever $F(\sigma) \subseteq G(\sigma)$ for all $\sigma \in X$. Given a pair of acyclic carriers $F : X \rightrightarrows Y$ and $H : Y \rightrightarrows Z$, we also define the composition carrier $H \circ F : X \rightrightarrows Z$, where each cell $\sigma \in X$ is sent to

$$H \circ F(\sigma) \coloneqq \bigcup_{\tau \in F(\sigma)} H(\tau) \ .$$

In particular, notice that if f is carried by F and g is carried by G, then $g \circ f$ is 'carried' by $G \circ F$. Note, however, that this composition does not need to be acyclic!

Example 2.7.4. Consider a regular morphism $f : X \to Y$. We can then define the (not necessarily acyclic) carrier $F_f : X \rightrightarrows Y$ induced by f as the assignment sending $\sigma \in X$ to $\overline{f(\sigma)}$. Notice that by continuity of f and property (**R. 1.**) we have that for any pair $\tau \prec \sigma$ in X, there is an inclusion $\overline{f(\tau)} \subseteq \overline{f(\sigma)}$. Also, $\overline{f(\sigma)} \neq \emptyset$ since it must contain at least a point. Given an acyclic carrier $G : Y \rightrightarrows Y$, we will denote by $G(f(\sigma))$ the composition of carriers $G \circ F_f(\sigma)$ for all $\sigma \in X$. This situation will come up very often in this text and whenever we are looking at the composition $G \circ F_f$ we will assume that it is acyclic. Note that F_ι is acyclic for inclusion maps $\iota : X \hookrightarrow Y$.

2.8 Homotopy Colimits and Geometric Realizations

Consider a poset P and let $F : P^{\mathrm{op}} \to \mathbf{Top}$ be a functor from the oposite category; we call F a *diagram of spaces* and denote the corresponding category by $\mathbf{Diag}(P)$; where the morphisms are given by natural transformations. Assume that a space X is given as the union of subspaces $\bigcup_{i \in I} U_i$. Following example 2.3.2 we can consider the diagram of spaces $X^{\mathscr{U}}$ over the oposite nerve $N^{\mathrm{op}}_{\mathscr{U}}$ and recall that $X = \mathbf{colim}(X^{\mathscr{U}})$. One might want to use this diagram structure to "deform" each

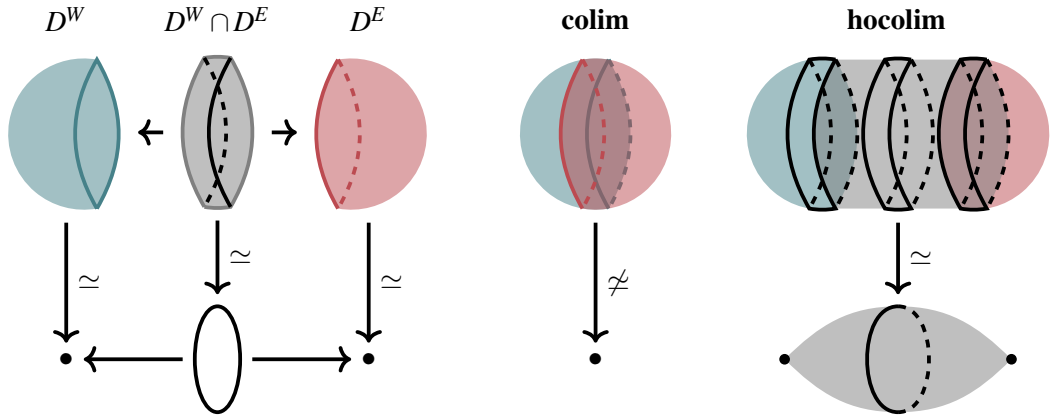


Figure 2.4: On the left, the natural transformation $F : X^{\mathscr{U}} \to \widetilde{X^{\mathscr{U}}}$ is illustrated. In the center we see that $\mathbf{colim}(X^{\mathscr{U}}) \not\simeq \mathbf{colim}(\widetilde{X^{\mathscr{U}}})$ while on the right $\mathbf{hocolim}(X^{\mathscr{U}}) \simeq \mathbf{hocolim}(\widetilde{X^{\mathscr{U}}})$.

piece $\mathscr{U}_\sigma$ for all $\sigma \in N^{\mathrm{op}}_{\mathscr{U}}$ in a compatible way so that the object X becomes simpler. Do the local homotopy equivalences carry over to the space X? In general the answer is negative as shown in the following example.

Example 2.8.1. Consider a sphere S^2 covered by an open disk containing the western hemisphere D^W and another containing eastern hemisphere D^E, together with a common intersection band $D^W \cap D^E \simeq S^1$, see figure 2.4. In this case $\mathscr{U} = \{D^W, D^E\}$ forms a cover for $X = S^2$, which can be seen as a functor $X^{\mathscr{U}} : (\Delta^1)^{\mathrm{op}} \to \mathbf{Top}$. Suppose that we collapse D^W and D^E to a point, while collapsing the band $D^W \cap D^E$ to the circle S^1. This gives another functor $\widetilde{X^{\mathscr{U}}} : (\Delta^1)^{\mathrm{op}} \to \mathbf{Top}$ in a way that there is a natural transformation $F : X^{\mathscr{U}} \to \widetilde{X^{\mathscr{U}}}$ where $F(\sigma) : X^{\mathscr{U}}(\sigma) \to \widetilde{X^{\mathscr{U}}}(\sigma)$ is a homotopy equivalence for all $\sigma \in \Delta^1$. One can see that unfortunately $\mathbf{colim}(X^{\mathscr{U}}) = S^2$ is not homotopy equivalent to $\mathbf{colim}(\widetilde{X^{\mathscr{U}}}) \simeq *$.

This lack of homotopy stability of the colimit motivates the definition of a *homotopy colimit*. As we are mainly interested in diagrams of spaces, we will only define homotopy colimits over posets. Consider diagram $F \in \mathbf{Diag}(P)$, we define the homotopy colimit as

$$\mathbf{hocolim}(F) = \coprod_{\substack{n \geq 0 \\ i_0^{(n)} \to i_1^{(n)} \to \cdots \to i_n^{(n)}}} \Delta^n \times F(i_0) \big/ \sim$$

where the disjoint union is taken over n-sequences of non-identity arrows from P^{op}. The different pieces are identified by the relation

$$\big((\Delta^k \hookrightarrow \Delta^n)(p), q\big) \sim \big(p, F(i_0^{(n)} \to i_0^{(k)})(q)\big)$$

for any pair of points $p \in \Delta^k$ and $q \in F(i_0)$ and for any pair of integers $k \leq n$ such that for all

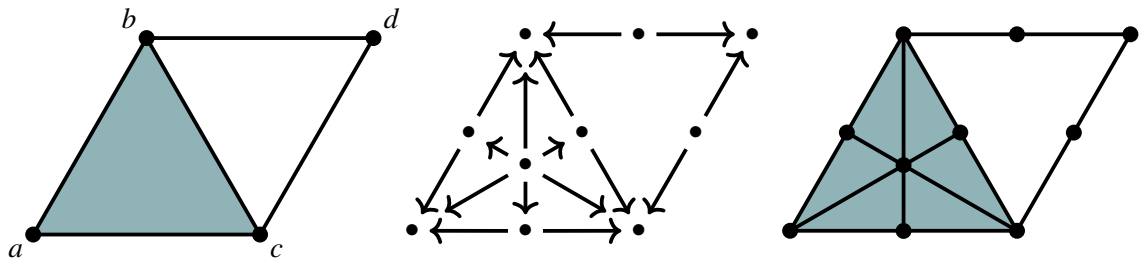


Figure 2.5: On the left, the geometric realization of a simplicial complex K. In the middle the constant diagram $*_K$ and on the right, the homotopy colimit $\mathbf{hocolim}\,(*_K)$.

$0 \leq r \leq k$ there is an equality $i_r^{(k)} = i_l^{(n)}$ for some $0 \leq l \leq n$.

Homotopy colimits enjoy quite general stability properties with respect to compatible homotopy equivalences. This will be the case for diagrams $\mathscr{D} : K^{\mathrm{op}} \to \mathbf{CW\text{-}cpx}$ over a simplicial complex K, as will follow from Theorem 2.8.6 and the remark at the end of this section; for further properties about homotopy colimits, see [73].

Example 2.8.2. Consider again example 2.8.1. In this case, the homotopy colimit is indeed stable under natural homotopy equivalences and $\mathbf{hocolim}\,(X^{\mathscr{U}}) \simeq \mathbf{hocolim}\,(\widetilde{X^{\mathscr{U}}}) \simeq S^2$ as depicted on figure 2.4.

Example 2.8.3. Consider a simplicial complex K together with the constant diagram $*_K : K^{\mathrm{op}} \to \mathbf{Top}$ as depicted on figure 2.5. Notice that the homotopy colimit $\mathbf{hocolim}\,(*_K)$ is homotopy equivalent to the geometric realization $|K|$. In fact, $\mathbf{hocolim}\,(*_K)$ has an underlying simplicial complex structure which is often called the barycentric subdivision of K and is denoted as $\mathrm{Bd}(K)$.

Example 2.8.3 shows that any simplicial complex can be seen as a homotopy colimit, at least up to homotopy equivalence. An inconvenience of using the homotopy colimit construction is that it contains too much redundant information. As one can see on figure 2.5, the homotopy colimit $\mathbf{hocolim}\,(*_K)$ is given by the barycentric subdivision $\mathrm{Bd}(K)$ which contains many more cells than $|K|$. This is why we turn our attention to objects that maintain the good properties of homotopy colimits, while being simpler from a combinatorial point of view. For this, we will restrict to diagrams $\mathscr{D} : K^{\mathrm{op}} \to \mathbf{CW\text{-}cpx}$ of regular CW-complexes over a given simplicial complex K, together with regular maps $\mathscr{D}(\tau \preceq \sigma)$ over K^{op}. Such an $\mathscr{D}$ will be called a *regular diagram of CW-complexes*, and we will denote the corresponding category by $\mathbf{RDiag}(K)$.

Definition 2.8.4. Let $\mathscr{D} \in \mathbf{RDiag}(K)$. The *geometric realization* $\Delta_K\mathscr{D}$ of $\mathscr{D}$ is the regular CW-complex defined as

$$\Delta_K\mathscr{D} = \bigsqcup_{\sigma \in K} \Delta^\sigma \times \mathscr{D}(\sigma) \big/ \sim ,$$

where, for any pair $\tau \preceq \sigma$ in K the relation identifies a pair of points

$$(\Delta^\tau \hookrightarrow \Delta^\sigma)(x) \times y \sim x \times \mathscr{D}(\tau \preceq \sigma)(y)$$

for each pair of points $x \in \Delta^\tau$ and $y \in \mathscr{D}(\sigma)$. This $\Delta_K\mathscr{D}$ has a natural filtration

$$F^0\Delta_K\mathscr{D} \subseteq F^1\Delta_K\mathscr{D} \subseteq \cdots \subseteq F^n\Delta_K\mathscr{D} \subseteq \cdots$$

given by $F^p\Delta_K\mathscr{D} = \bigcup_{\sigma\in K^p} \Delta^\sigma \times \mathscr{D}(\sigma)$ for all $p \geq 0$. A cell $\tau \times c$ is a face of another cell $\sigma \times a$ if and only if $\tau \preceq \sigma$ and also $c \in \overline{\mathscr{D}(\tau \preceq \sigma)(a)}$. If the underlying simplicial complex K is clear from the context, we will write $\Delta\mathscr{D}$ instead of $\Delta_K\mathscr{D}$.

Given a regular morphism $\mathscr{F} : \mathscr{D} \to \mathscr{L}$ of diagrams in $\mathbf{RDiag}(K)$, there is an induced morphism on the geometric realization which we denote $\Delta\mathscr{F}$. Denote by $*^{\mathscr{D}}$ the *constant diagram* given by

$$*^{\mathscr{D}}(\sigma) = \begin{cases} * & \text{if } \mathscr{D}(\sigma) \neq \emptyset \\ \emptyset & \text{else} \end{cases}$$

and note that there is a homotopy equivalence $\Delta(*^{\mathscr{D}}) \simeq |K^{\mathscr{D}}|$, where $K^{\mathscr{D}}$ is a subcomplex of K such that $\sigma \in K^{\mathscr{D}}$ if and only if $\mathscr{D}(\sigma) \neq \emptyset$. The projection onto the simplex coordinates gives a *base projection* $p_b : \Delta\mathscr{D} \to \Delta(*^{\mathscr{D}}) \simeq |K^{\mathscr{D}}|$.

Example 2.8.5. Let $\mathscr{D} \in \mathbf{RDiag}(K)$ and define by $\pi_0(\mathscr{D}) \in \mathbf{RDiag}(K)$ the diagram that sends each simplex $\sigma \in K$ to the set of connected components $\pi_0(\mathscr{D}(\sigma))$ with the discrete topology; also for any pair $\tau \preceq \sigma$ the morphism $\mathscr{D}(\tau \preceq \sigma)$ induces a morphism $\pi_0(\mathscr{D}(\sigma)) \to \pi_0(\mathscr{D}(\tau))$. We define the *multinerve* of $\mathscr{D}$ as

$$\mathrm{MNerv}(\mathscr{D}) = \Delta(\pi_0(\mathscr{D}))\ .$$

This object was first introduced in [40] in the case of $\pi_0(X^{\mathscr{U}})$ for a space X covered by $\mathscr{U}$, where it was defined as a simplicial poset. There are epimorphisms $\Delta\mathscr{D} \to \mathrm{MNerv}(\mathscr{D}) \to \Delta(*^{\mathscr{D}}) \simeq |K|$.

Theorem 2.8.6. *Let $\mathscr{F} : \mathscr{D} \to \mathscr{L}$ be a morphism of diagrams in $\mathbf{RDiag}(K)$. If $\mathscr{F}(\sigma)$ is a homotopy equivalence for all $\sigma \in K$, then $\Delta\mathscr{F} : \Delta\mathscr{D} \to \Delta\mathscr{L}$ is a homotopy equivalence.*

A proof of this result for more general diagrams of spaces (not just over a poset P) can be found in [68, Prop. 4G.1]. A proof in the context of semisimplicial sets can be found on [56, Thm. 2.2.].

Example 2.8.7. Let $X \in \mathbf{CW\text{-}cpx}$ covered by $\mathscr{U}$ and recall the diagram $X^{\mathscr{U}}$ from Example 2.3.2. In this case $\Delta(X^{\mathscr{U}})$ is the *Mayer-Vietoris blowup complex* associated to the pair $(X, \mathscr{U})$ and it

can be described as a subspace of the product $X \times |N_{\mathscr{U}}|$. This leads to the *fiber projection* $p_f : \Delta(X^{\mathscr{U}}) \to X$ and to the *base projection* $p_b : \Delta(X^{\mathscr{U}}) \to |N_{\mathscr{U}}|$. As shown in [68, Prop. 4G.2], p_f is a homotopy equivalence $\Delta(X^{\mathscr{U}}) \simeq X$. If each $X^{\mathscr{U}}(\sigma)$ is contractible for all $\sigma \in N_{\mathscr{U}}$, then p_b is also a homotopy equivalence by Theorem 2.8.6. This last result is usually known as the *nerve theorem*.

Remark. Let $\mathscr{D}$ be a regular diagram of CW-complexes over the simplicial complex K. We can extend $\mathscr{D}$ to a diagram $\mathscr{D}'$ on the barycentric subdivision $\mathrm{Bd}(K)$ by defining

$$\mathscr{D}'(\tau_0 \prec \cdots \prec \tau_n) = \mathscr{D}(\tau_n)$$

on an n-simplex $\tau_0 \prec \tau_1 \prec \cdots \prec \tau_n$ in $\mathrm{Bd}(K)$. A non-identity morphism in $\mathrm{Bd}(K)$ that has $\tau_0 \prec \tau_1 \prec \cdots \prec \tau_n$ as its codomain must have the same flag with one of the τ_k's left out as its domain. The diagram $\mathscr{D}'$ maps such a morphism to the identity in case $k \neq n$ or the morphism $\mathscr{D}(\tau_{n-1} \prec \tau_n)$ in case $k = n$. It is clear from the definition of the homotopy colimit that the geometric realization $\Delta(\mathscr{D}')$ is homotopy equivalent to $\mathbf{hocolim}\,\mathscr{D}$. A modified version of the homotopy equivalence $|K| \simeq |\mathrm{Bd}(K)|$ shows that $\Delta(\mathscr{D}) \simeq \Delta(\mathscr{D}')$. Hence, we could have worked with homotopy colimits all throughout, but we chose to work with the geometric realization since it is technically easier to handle and because in some instances it is the Mayer-Vietoris blowup complex, which has already appeared before in Topological Data Analysis [123].

2.9 The Mayer-Vietoris Long Exact Sequence

In this section we will review a classical result in homology theory. This will motivate the study of simplicial cosheaves, Čech homology and, ultimately, the study of the Mayer-Vietoris spectral sequence. Let X be a topological space, and let $\mathscr{U} = \{U_i\}_{i=0}^m$ be an open cover of X. Suppose that we want to compute the homology of X from the local information on each cover element. A naive approach to solving the problem would be to compute first the homology groups $\{\mathrm{H}_n(U_i)\}_{i=0}^m$, and proceed by adding all of them back together:

$$\mathrm{H}_n(X) = \bigoplus_{0 \leq i \leq m} \mathrm{H}_n(U_i). \tag{2.4}$$

This is hardly ever true and we will need to find other ways of dealing with this merging of information. Alternatively, one can notice that, as the space X is the colimit $\mathbf{colim}(X^{\mathscr{U}})$, the same would be true for $\mathrm{H}_n(X)$ being the colimit $\mathbf{colim}_{\sigma \in K^{\mathrm{op}}}(\mathrm{H}_n(X^{\mathscr{U}}(\sigma)))$. This approach is better than the former, but it is still not good enough as this does not happen in general.

Consider the torus $\mathbb{T}^2$ covered by two cylinders U and V, as illustrated in Figure 2.6. Then

one sees that equality (2.4) does not hold in dimensions 0 and 2:

$$\mathrm{H}_0(\mathbb{T}^2) = \mathbb{F} \ncong \mathbb{F} \oplus \mathbb{F} = \mathrm{H}_0(U) \oplus \mathrm{H}_0(V), \qquad \mathrm{H}_2(\mathbb{T}^2) = \mathbb{F} \ncong 0 = \mathrm{H}_2(U) \oplus \mathrm{H}_2(V).$$

In order to amend this, one has to look at the information given by the intersection $U \cap V$. This information comes as *identifications* and new *loops*. For example, U and V are connected through the intersection. Also, the loop going around each cylinder U and V is identified in the intersection. These identifications are performed by taking the quotient

$$\mathrm{I}_n \coloneqq \mathrm{coker}\Big(\mathrm{H}_n(U \cap V) \to \mathrm{H}_n(U) \oplus \mathrm{H}_n(V) \Big)$$

for all $n \geq 0$. Where the previous morphism is the Čech differential $\delta_1^n : S_n(U \cap V) \to S_n(U) \oplus S_n(V)$ given by the assignment on singular chains $s \mapsto (s, -s)$. Additionally, the 1-loops in the intersection merge to the same loop when included in each cylinder U or V. This situation creates a 2-loop or 'void', see Figure 2.6. Thus we have the n-loops detected by the kernel

$$\mathrm{L}_n \coloneqq \mathrm{Ker}\Big(\mathrm{H}_{n-1}(U \cap V) \to \mathrm{H}_{n-1}(U) \oplus \mathrm{H}_{n-1}(V) \Big)$$

for all $n \geq 0$. Notice that n-loops are found by $n-1$ information on the intersection. Putting all together, we have that

$$\mathrm{H}_0(\mathbb{T}^2) \cong \mathrm{I}_0 \cong \mathbb{F}, \qquad \mathrm{H}_1(\mathbb{T}^2) \cong \mathrm{I}_1 \oplus \mathrm{L}_1 \cong \mathbb{F} \oplus \mathbb{F}, \qquad \mathrm{H}_2(\mathbb{T}^2) \cong \mathrm{L}_2 \cong \mathbb{F}.$$

This leads to the expected result

$$\mathrm{H}_k(\mathbb{T}^2) \cong \begin{cases} \mathbb{F} & \text{for } k = 0, 2, \\ \mathbb{F} \oplus \mathbb{F} & \text{for } k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

On a more theoretical level, what we have presented here is commonly known as the Mayer-Vietoris theorem. This starts with short exact sequences

$$0 \longleftarrow S_n(X) \xleftarrow{\Sigma} S_n(U) \oplus S_n(V) \xleftarrow{\check{\delta}_1^n} S_n(U \cap V) \longleftarrow 0\,,$$

for all $n \geq 0$, where the morphism Σ is given by the assignment on singular chains $(a, b) \mapsto a + b$. One can check that the above short sequence is indeed exact. Using the snake lemma [118, §1.3]

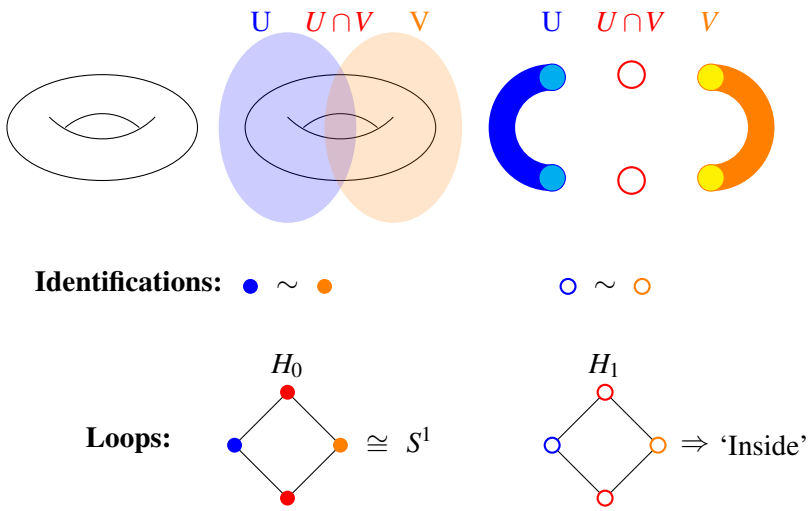


Figure 2.6: Torus covered by a pair of cylinders U and V.

one can show that there is a long exact sequence

$$\cdots \xleftarrow{\check{\delta}_1^{n-1}} \mathrm{H}_{n-1}(U\cap V) \xleftarrow{\partial_n} \mathrm{H}_n(X) \xleftarrow{\Sigma} \mathrm{H}_n(U)\oplus \mathrm{H}_n(V) \xleftarrow{\check{\delta}_1^{n}} \mathrm{H}_n(U\cap V) \xleftarrow{\partial_{n+1}} \cdots$$

which will be called the *Mayer-Vietoris* sequence. We can think of each homology group $\mathrm{H}_n(X)$ as a filtered $\mathbb{F}$-vector space,

$$\{0\} = F^{-1}\mathrm{H}_n(X) \subset F^0\mathrm{H}_n(X) \subset F^1\mathrm{H}_n(X) = \mathrm{H}_n(X).$$

Then, the Mayer-Vietoris sequence gives us the expressions for the different ratios between consecutive filtrations,

$$F^0\mathrm{H}_n(X) = \mathrm{I}_n, \qquad \frac{F^1\mathrm{H}_n(X)}{F^0\mathrm{H}_n(X)} = \mathrm{L}_n.$$

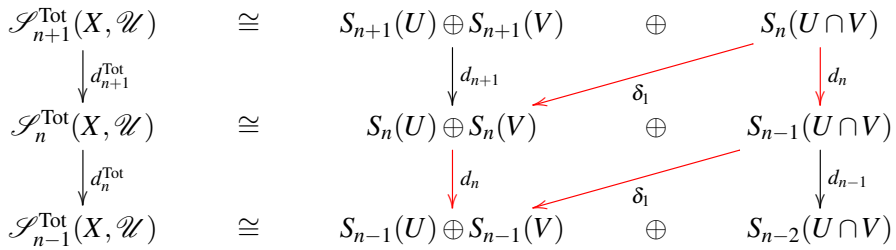
In particular, since we are working with vector spaces we obtain $\mathrm{H}_n(X) \cong \mathrm{I}_n \oplus \mathrm{L}_n$ for all $n \geq 0$.

The above discussion gives rise to the *total chain complex*,

$$\mathscr{S}_n^{\mathrm{Tot}}(X,\mathscr{U}) = S_n(V)\oplus S_n(U)\oplus S_{n-1}(U\cap V),$$

with morphism $d_n^{\mathrm{Tot}} = (d, d, d-\delta_1)$ for all $n \geq 0$. Notice that the first two morphisms do not change components, whereas the third encodes the 'merging' of information. This last morphism

is represented by the red diagonal arrows on the diagram:



where the rectangle of red arrows is commutative. In particular, this implies that $d^{\mathrm{Tot}}_n \circ d^{\mathrm{Tot}}_{n+1} = 0$ for all $n \geq 0$. Computing the homology with respect to the total differentials and using the previous characterization of I_n and L_n, one obtains

$$\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*(X,\mathscr{U})) \cong \mathrm{I}_n \oplus \mathrm{L}_n \cong \mathrm{H}_n(X). \tag{2.5}$$

This result will be further generalized in proposition 2.11.1. Equation (2.5) is a consequence of a result of a chain complex isomorphism $\mathscr{S}^{\mathrm{Tot}}_n(X,\mathscr{U}) \simeq S_n(\Delta X^{\mathscr{U}})$ together with the fact shown in example 2.8.7 that $\Delta X^{\mathscr{U}} \simeq X$.

2.10 Čech Homology

Notice that, although this section is included in the background chapter, the definitions of *simplicial precosheaves* and *simplicial cosheaves* differ slightly from the usual definition. Also, the proof of exactness of the sequence (2.7) by using transpositions is our own. We include this section here because it introduces relevant concepts and results for the introduction of the Mayer-Vietoris spectral sequence to be presented in section 2.11.

In section 2.9 we computed homology of a space X from the homology of a cover by two open sets $\mathscr{U} = \{U,V\}$. In this case notice that composing $X^{\mathscr{U}}$ with the n-homology functor leads to assignments

$$U \mapsto \mathrm{H}_n(U), \qquad V \mapsto \mathrm{H}_n(V), \quad \text{and} \quad [U,V] \mapsto \mathrm{H}_n(U\cap V)\,.$$

together with the homology maps induced by the inclusion of $U\cap V$ into U and V. These assignments can be considered as "systems of coefficients" over the oposite nerve $N^{\mathrm{op}}_{\mathscr{U}}$, a concept that we formalize now.

Recall that a simplicial complex K can be seen as a category, as explained in example 2.3.1; and from it we can consider the oposite category K^{op}. We define the *colimit cone* CK^{op} to be the category K^{op} with an extra object $*$ with unique arrows $\sigma \to *$ for all $\sigma \in K^{\mathrm{op}}$. Let $\mathscr{F} \in \mathbf{Diag}(K)$,

one can extend it to a functor $\mathscr{F} : CK^{\mathrm{op}} \to \mathbf{Top}$ sending $*$ to $\mathbf{colim}_K \mathscr{F}$ and any arrow $\sigma \to *$ to the unique morphism $\mathscr{F}(\sigma) \to \mathbf{colim}_{K^{\mathrm{op}}}(\mathscr{F})$ for all $\sigma \in K$; here notice that $\mathbf{colim}_{K^{\mathrm{op}}} \mathscr{F}$ exists since K^{op} is small, see [90].

Definition 2.10.1. Let $\mathscr{F} \in \mathbf{Diag}(K)$ be a diagram of spaces over a simplicial complex K and consider a functor $\mathscr{G} : \mathbf{Top} \to \mathbf{vect}_{\mathbb{F}}$. The composition $\mathscr{G}(\mathscr{F}) : CK^{\mathrm{op}} \to \mathbf{vect}_{\mathbb{F}}$ will be called a *simplicial precosheaf*. The maps $\mathscr{G}(\mathscr{F})(\tau \preceq \sigma)$ for $\tau, \sigma \in K$ are the *restriction maps*.

In fact, this is just a particular case of the more general concept of precosheaf, but for the present purposes this definition is good enough. For a good introduction to the matter see [16] or [44]. We will denote by **PreCosh** the category of simplicial precosheaves and by $\mathbf{PreCosh}(K)$ the subcategory of simplicial precosheaves over K. We will often omit mentioning the underlying diagram and functors of a simplicial precosheaf structure, and so we will say "let $\mathscr{L}$ be a simplicial precosheaf" or write "let $\mathscr{L} \in \mathbf{PreCosh}(K)$".

Example 2.10.2. Given $\mathscr{F} \in \mathbf{Diag}(K)$ for K being a simplicial complex and consider the n-homology functor $\mathscr{H}_n : \mathbf{Top} \to \mathbf{vect}_{\mathbb{F}}$. We obtain the simplicial precosheaf $\mathscr{H}_n(\mathscr{F}) : CK^{\mathrm{op}} \to \mathbf{vect}_{\mathbb{F}}$, so that $\mathscr{H}_n(\mathscr{F})(\sigma) = \mathrm{H}_n(\mathscr{F}(\sigma))$ for all $\sigma \in K$. The restriction maps are given by the induced maps in homology

$$\mathscr{H}_n(\mathscr{F})(\tau \preceq \sigma) = \mathscr{F}(\tau \preceq \sigma)_* : \mathrm{H}_n(\mathscr{F}(\sigma)) \longrightarrow \mathrm{H}_n(\mathscr{F}(\tau)) \,,$$

for all pairs $\tau \preceq \sigma$ from K. Of course, the colimit is sorted by setting $\mathscr{H}_n(\mathscr{F})(*) = \mathrm{H}_n(\mathbf{colim}_K \mathscr{F})$.

Now, let us go back to the simplicial cosheaves $\mathscr{H}_n(X^{\mathscr{U}})$ from the start of the section. Recall that computing homology with respect to the sequence

$$0 \longleftarrow \mathrm{H}_n(U) \oplus \mathrm{H}_n(V) \xleftarrow{\check{\delta}_1^n} \mathrm{H}_n(U \cap V) \longleftarrow 0 \,,$$

lead to the groups I_n and L_n which, as shown in section 2.9, determine the groups $\mathrm{H}_n(X)$ for all $n \geq 0$. We can generalize this idea and consider a simplicial precosheaf $\mathscr{L} \in \mathbf{PreCosh}(K)$ to define the *Čech chain complex* $\check{C}_*(K; \mathscr{L})$:

$$0 \longleftarrow \bigoplus_{\sigma^0 \in K} \mathscr{L}(\sigma^0) \xleftarrow{\check{\delta}_1^{\mathscr{L}}} \bigoplus_{\sigma^1 \in K} \mathscr{L}(\sigma^1) \xleftarrow{\check{\delta}_2^{\mathscr{L}}} \bigoplus_{\sigma^2 \in K} \mathscr{L}(\sigma^2) \longleftarrow \cdots \tag{2.6}$$

where $\check{\delta}_n^{\mathscr{L}}$ denotes the n-Čech differential. We will be using the notation $(s)_{\sigma^n}$ to denote the vector in $\bigoplus_{\sigma^n \in K} \mathscr{L}(\sigma^n)$ that is zero except on the component $\sigma^n \in K$ where it is equal to $s \in \mathscr{L}(\sigma^n)$. The

Čech differential is defined on each generator $(s)_{\sigma^n}$ by the assignment

$$\check{\delta}_n^{\mathscr{L}}\left((s)_{\sigma^n}\right) = \sum_{i=0}^{n}(-1)^i \mathscr{L}(\sigma_i^{n-1} \preceq \sigma^n)_*(s)\ ,$$

this definition extends to all $\bigoplus_{\sigma^n \in K} \mathscr{L}(\sigma^n)$ by linearity. It can be checked that $\check{\delta}_i^{\mathscr{L}} \circ \check{\delta}_{i+1}^{\mathscr{L}} = 0$ for all integers $i \geq 0$. Thus, we can define the *Čech Homology* groups as

$$\check{\mathrm{H}}_n(K;\mathscr{L}) = \frac{\mathrm{Ker}(\check{\delta}_n^{\mathscr{L}})}{\mathrm{Im}(\check{\delta}_{n+1}^{\mathscr{L}})}\ ,$$

for all $n \geq 0$.

Example 2.10.3. Notice that Čech homology is a generalization of simplicial homology. To see this, consider the constant diagram $*_K \in \mathbf{RDiag}(K)$, and consider the simplicial precosheaf $\mathscr{H}_0(*_K) : CK^{\mathrm{op}} \to \mathbf{vect}_{\mathbb{F}}$. This precosheaf $\mathscr{H}_0(*_K)$ sends all simplices of K to a field $\mathbb{F}$ and all face relations in K^{op} to the identity map in $\mathbb{F}$. The colimit is sent to $\mathrm{H}_0(K;\mathbb{F})$. Then there is an isomorphism of homology groups $\check{\mathrm{H}}_n(K;\mathscr{H}_0(*_K)) \simeq \mathrm{H}_n(K;\mathbb{F})$ for all $n \geq 0$.

Consider again the scenario of the precosheaves $\mathscr{H}_n(X^{\mathscr{U}})$ from the start of this section. Observe that the fact that some homology classes are detected in L_n makes one wonder whether $\mathscr{U}$ was a "good cover" for X; notice that L_n detects $n+1$ homology classes that are "broken down" by $\mathscr{U}$. Whenever $\mathrm{L}_{n+1} = 0$, then from the Mayer-Vietoris long exact sequence we obtain short exact sequences

$$0 \longleftarrow \mathrm{H}_n(X) \xleftarrow{\Sigma} \mathrm{H}_n(U) \oplus \mathrm{H}_n(V) \xleftarrow{\check{\delta}_1^n} \mathrm{H}_n(U \cap V) \longleftarrow 0$$

for all $n \geq 0$. In this case, the underlying precosheaf $\mathscr{H}_n(X^{\mathscr{U}})$ preserves the homology of the colimit, as we have isomorphisms:

$$\mathscr{H}_n(X^{\mathscr{U}})(*) = \mathrm{H}_n(\mathbf{colim}_{K^{\mathrm{op}}} X^{\mathscr{U}}) \simeq \mathrm{H}_n(X) \simeq \frac{\mathrm{H}_n(U) \oplus \mathrm{H}_n(V)}{\mathrm{Im}(\check{\delta}_1^n)} \simeq \mathbf{colim}_{K^{\mathrm{op}}}\left(\mathscr{H}_n(X^{\mathscr{U}})\right)\ .$$

Definition 2.10.4. Let $\mathscr{F} \in \mathbf{Diag}(K)$, and let some functor $\mathscr{G} : \mathbf{Top} \to \mathbf{vect}_{\mathbb{F}}$. Then we consider the simplicial precosheaf given by the composition of functors $\mathscr{G}(\mathscr{F}) \in \mathbf{PreCosh}(K)$. If the sequence

$$0 \longleftarrow \mathscr{G}(\mathbf{colim}_K(\mathscr{F})) \longleftarrow \bigoplus_{\sigma^0 \in K} \mathscr{G}(\mathscr{F})(\sigma^0) \longleftarrow \bigoplus_{\sigma^1 \in K} \mathscr{G}(\mathscr{F})(\sigma^1)$$

is exact, then $\mathscr{G}(\mathscr{F})$ is said to be a *simplicial cosheaf*. Notice that this is equivalent to the condition

$$\mathscr{G}(\mathbf{colim}_{K^{\mathrm{op}}}(\mathscr{F})) \simeq \mathbf{colim}_{K^{\mathrm{op}}}(\mathscr{G}(\mathscr{F}))\ .$$

We denote by $\mathbf{Cosh}(K)$ the category of simplicial cosheaves over K. Notice that in the literature usually a simplicial cosheaf is defined as a functor $\mathscr{L} : K^{\mathrm{op}} \to \mathbf{vect}_{\mathbb{F}}$, see for example [45, §5]. We take this alternative definition as it defines a distinction between simplicial precosheaves and simplicial sheaves.

Consider again the simplicial precosheaves $\mathscr{H}_n(X^{\mathscr{U}}) \in \mathbf{PreCosh}(N_{\mathscr{U}})$, for our particular case of X and $\mathscr{U}$. As pointed out earlier, in general these are not cosheaves. The exception comes in degree $n = 0$, where the Mayer-Vietoris sequence leads to the exact sequence

$$0 \longleftarrow \mathrm{H}_0(X) \xleftarrow{\Sigma} \mathrm{H}_0(U) \oplus \mathrm{H}_0(U) \xleftarrow{\check{\delta}_1^0} \mathrm{H}_0(U \cap V)$$

Thus $\mathscr{H}_0(X^{\mathscr{U}})$ is a simplicial cosheaf. We will show later that this is also the case for any space X and any cover $\mathscr{U}$.

Example 2.10.5. Continuing with example 2.10.3. In this case the sequence

$$0 \longleftarrow \mathrm{H}_0(K) \longleftarrow \bigoplus_{\sigma^0 \in K} \mathrm{H}_0(*_{\mathscr{U}}(\sigma^0)) \longleftarrow \bigoplus_{\sigma^1 \in K} \mathrm{H}_0(*_{\mathscr{U}}(\sigma^1))$$

is exact, and $\mathscr{H}_0(*_{\mathscr{U}})$ is a simplicial cosheaf. Notice that if K is connected then $\mathrm{H}_0(K) \simeq \mathbb{F}$. Here we would recover the notion of an augmented chain complex from section 2.2.

The previous example motivates the following definition.

Definition 2.10.6. Let $\mathscr{F} \in \mathbf{Diag}(K)$, and let some functor $\mathscr{G} : \mathbf{Top} \to \mathbf{vect}_{\mathbb{F}}$. We define the *augmented* Čech chain complex on the precosheaf $\mathscr{G}(\mathscr{F})$ as

$$0 \longleftarrow \mathscr{G}(\mathbf{colim}_{K^{\mathrm{op}}} \mathscr{F}) \xleftarrow{\Sigma_{\mathscr{F}}} \bigoplus_{\sigma^0 \in K} \mathscr{G}\mathscr{F}(\sigma^0) \xleftarrow{\check{\delta}_1^{\mathscr{G}\mathscr{F}}} \bigoplus_{\sigma^1 \in K} \mathscr{G}\mathscr{F}(\sigma^1) \xleftarrow{\check{\delta}_2^{\mathscr{G}\mathscr{F}}} \cdots$$

where the augmentation map $\Sigma_{\mathscr{F}}$ is generated by the morphisms $\mathscr{F}(\sigma^0) \to \mathbf{colim}_{K^{\mathrm{op}}} \mathscr{F}$ for all $\sigma^0 \in K$. For convenience, we will often denote the augmentation map and Čech differentials by Σ and $\check{\delta}$ respectively, omitting any references to $\mathscr{F}$ and $\mathscr{G}$.

Example 2.10.7. Consider now the case of a simplicial complex K together with a cover by subcomplexes $\mathscr{U}$. We consider the regular diagram $K^{\mathscr{U}} \in \mathbf{RDiag}(N_{\mathscr{U}})$ given by $K^{\mathscr{U}}(\sigma) = \mathscr{U}_\sigma$ for all $\sigma \in N_{\mathscr{U}}$. Then we consider the simplicial precosheaf given by taking simplicial chains $S_n^\Delta(K^{\mathscr{U}})$. Since $\mathscr{U}$ is a cover of K by subcomplexes, the sequence

$$0 \longleftarrow S_n^\Delta(K) \xleftarrow{\Sigma} \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} S_n^\Delta(\mathscr{U}_{\sigma^0}) \xleftarrow{\check{\delta}} \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} S_n^\Delta(\mathscr{U}_{\sigma^1})$$

is exact. Thus, $S_n^\Delta(K^{\mathscr{U}})$ is in fact a simplicial cosheaf.

Example 2.10.8. One could rewrite example 2.10.7 considering a space X with a cover $\mathscr{U}$, together with the singular precosheaf $S_n(X^{\mathscr{U}})$. However, in this case the colimit on the precosheaf leads to *local singular chains*

$$S_n^{\mathscr{U}}(X) = \mathbf{colim}_{N_{\mathscr{U}}^{\mathrm{op}}} S_n(X^{\mathscr{U}}) \ ,$$

which in general are not isomorphic to $S_n(X)$. Notice that there is an inclusion $\iota : S_n^{\mathscr{U}}(X) \hookrightarrow S_n(X)$, and in fact one can show that this inclusion is a chain homotopy equivalence, see [68, §2.1.]. Thus, even though $S_n(X^{\mathscr{U}})$ is not a cosheaf, it can be considered to be a "cosheaf up to chain homotopy".

We finish this section by focusing on the case of a simplicial complex K together with a finite cover $\mathscr{U}$ by subcomplexes. We consider the simplicial cosheaf $S_n^{\Delta}(K^{\mathscr{U}})$ from example 2.10.7. The aim will be to show that the augmented Čech chain complex $\check{C}_*(N_{\mathscr{U}}; S_n^{\Delta}(K^{\mathscr{U}}))$ expanded as

$$0 \longleftarrow S_n^{\Delta}(K) \xleftarrow{\Sigma} \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} S_n^{\Delta}(\mathscr{U}_{\sigma^0}) \xleftarrow{\check{\delta}} \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} S_n^{\Delta}(\mathscr{U}_{\sigma^1}) \xleftarrow{\check{\delta}} \cdots \tag{2.7}$$

is exact. In order to show this, consider a diagram of spaces $\mathscr{N}_{\mathscr{U}} \in \mathbf{RDiag}(K)$ given by sending a simplex $s \in K$ to $\mathscr{N}_{\mathscr{U}}(s) = \Delta^{\dim(\sigma(s))}$, where $\sigma(s) \in N_{\mathscr{U}}$ denotes the simplex of maximal dimension in $N_{\mathscr{U}}$ such that $s \in \mathscr{U}_{\sigma(s)}$. The restriction maps $\mathscr{N}_{\mathscr{U}}(t \preceq s)$ are given by the corresponding inclusions $\Delta^{\dim(\sigma(s))} \hookrightarrow \Delta^{\dim(\sigma(t))}$ for all simplices $t, s \in K$. Notice that $\mathbf{colim}_{K^{\mathrm{op}}} \mathscr{N}_{\mathscr{U}} = N_{\mathscr{U}}$ with the maps $\mathscr{N}_{\mathscr{U}}(\sigma) \to \mathbf{colim}_{K^{\mathrm{op}}} \mathscr{N}_{\mathscr{U}}$ given by inclusions $\Delta^{\dim(\sigma(s))} \hookrightarrow N_{\mathscr{U}}$. One can consider the simplicial precosheaf $S_n^{\Delta}(\mathscr{N}_{\mathscr{U}}) \in \mathbf{PreCosh}(K)$, with augmented Čech complex

$$0 \longleftarrow S_m^{\Delta}(N_{\mathscr{U}}) \xleftarrow{\Sigma} \bigoplus_{s^0 \in K} S_m^{\Delta}(\mathscr{N}_{\mathscr{U}}(s^0)) \xleftarrow{\check{\delta}} \bigoplus_{s^1 \in K} S_m^{\Delta}(\mathscr{N}_{\mathscr{U}}(s^1)) \xleftarrow{\check{\delta}} \cdots \tag{2.8}$$

There is an isomorphism

$$\begin{aligned} \psi : \qquad \textstyle\bigoplus_{s^n \in K} S_m^{\Delta}(N_{\mathscr{U}}(s^n)) &\xrightarrow{\simeq} \textstyle\bigoplus_{\sigma^m \in N_{\mathscr{U}}} S_n^{\Delta}(\mathscr{U}_{\sigma^m}) \\ (\sigma^m)_{s^n} &\longrightarrow (s^n)_{\sigma^m} \ . \end{aligned}$$

We will now use this "transposition" as a way to understand better the Čech chain complexes involved. First, we rewrite the chain (2.7) using this transposition:

$$0 \longleftarrow S_n^{\Delta}(K) \xleftarrow{\varepsilon^{\psi}} \bigoplus_{s^n \in K} S_0^{\Delta}(\mathscr{N}_{\mathscr{U}}(s^n)) \xleftarrow{d^{\psi}} \bigoplus_{s^n \in K} S_1^{\Delta}(\mathscr{N}_{\mathscr{U}}(s^n)) \xleftarrow{d^{\psi}} \cdots \tag{2.9}$$

Where the morphisms ε^{ψ} and d^{ψ} are induced by imposing commutativity of (2.7) with the transposition ψ. We claim that (2.9) corresponds to the direct sum of augmented simplicial chain complexes $\widetilde{S}_*^{\Delta}(\Delta^{\sigma(s^n)})$ for all $s^n \in K$; as this direct sum is exact by example 2.2.1, so the sequence (2.9)

would also be exact. To check this, we consider a chain $(\sigma^m)_{s^n} \in \bigoplus_{s^n \in K} S_m^{\Delta}(\mathscr{N}_{\mathscr{U}}(s^n))$, then we have

$$\begin{aligned} d^{\psi}((\sigma^m)_{s^n}) &= \psi^{-1} \circ \check{\delta} \circ \psi\big((\sigma^m)_{s^n}\big) \\ &= \psi^{-1} \circ \check{\delta}\big((s^n)_{\sigma^m}\big) \\ &= \psi^{-1}\Big(\sum_{i=0}^{m}(-1)^i (s^n)_{\sigma_i^m}\Big) \\ &= \Big(\sum_{i=0}^{m}(-1)^i \sigma_i^m\Big)_{s^n} \\ &= \big(d^{N^{\mathscr{U}}}(\sigma^m)\big)_{s^n} . \end{aligned}$$

Similarly, for the augmentation we have that

$$\varepsilon^{\psi}((\sigma^0)_{s^n}) = \Sigma \circ \psi((\sigma^0)_{s^n}) = s^n ,$$

and the result follows. Similarly, one could rewrite the sequence (2.8) by using the transposition ψ. In this case one obtains the direct sum of augmented chain complexes $\widetilde{S}_*^{\Delta}(\mathscr{U}_{\sigma})$ over all m-simplices $\sigma^m \in N_{\mathscr{U}}$. Consequently, (2.8) is exact if and only if each $\mathscr{U}_{\sigma^m}$ is acyclic for all $\sigma^m \in N_{\mathscr{U}}$.

Remark. Alternatively, one might consider the simplicial cosheaf $S_n^{\Delta}(K^{\mathscr{U}}) \in \mathbf{Cosh}(N_{\mathscr{U}})$ from example 2.10.7. This cosheaf is in fact a *flabby cosheaf*; that is, for each pair $V \subseteq U$, the morphism $\mathscr{S}_n(V \subseteq U) : \mathscr{S}_n(V) \to \mathscr{S}_n(U)$ is injective. Then, using 2.5, 4.3, and 4.4 from section VI. in [16], one has exactness of the Čech chain complex.

Remark. Recall that a regular CW complex X has incidence relations between its composing cells. For this reason, one could rewrite this section by considering cellular (pre)cosheaves over regular CW complexes. In particular, if we consider a covered regular CW-complex $(X, \mathscr{U})$, we might consider the corresponding cellular precosheaf given by $C_*^{\text{cell}}(X^{\mathscr{U}})$. The transposition construction should work to show that the augmented Čech sequence

$$0 \longleftarrow C_n^{\text{cell}}(K) \xleftarrow{\Sigma} \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} C_n^{\text{cell}}(\mathscr{U}_{\sigma^0}) \xleftarrow{\check{\delta}} \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} C_n^{\text{cell}}(\mathscr{U}_{\sigma^1}) \xleftarrow{\check{\delta}} \cdots$$

is exact for all $n \geq 0$. A motivation of working with Cosheaves is that there are fast algorithm techniques to compute their homology groups. For example, one could use Discrete Morse Theory, see [43] for a reference on this.

2.11 The Mayer-Vietoris spectral sequence

In this section, we give an introduction to the Mayer-Vietoris spectral sequence. These ideas come mainly from [15, 83, 62]. The reason for including this is because we think it beneficial to outline

a minimal, self-contained explanation of the procedure. Also, we will be using this as a necessary background for Chapter 6. For simplicity we will focus on ordinary homology over a field $\mathbb{F}$. Later on we will extend these ideas to the case of persistent homology over a field.

Consider a simplicial complex K with a covering $\mathscr{U} = \{U_i\}_{i=0}^m$ by subcomplexes. In this case, we need to take into account all the intersections between different subcomplexes. We can extend the intuition from section 2.9, by recalling the Čech chain complex (2.7) associated to the cosheaf $S_m^\Delta(K^{\mathscr{U}})$ for all $m \geq 0$. Stacking all these sequences on top of each other, and also multiplying differentials in odd rows by -1, we obtain a diagram:

$$\begin{array}{ccccccccc}
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
0 & \leftarrow & S_2^\Delta(K) & \xleftarrow{\Sigma} & \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} S_2^\Delta(\mathscr{U}_{\sigma^0}) & \xleftarrow{\check{\delta}_1} & \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} S_2^\Delta(\mathscr{U}_{\sigma^1}) & \xleftarrow{\check{\delta}_2} & \bigoplus_{\sigma^2 \in N_{\mathscr{U}}} S_2^\Delta(\mathscr{U}_{\sigma^2}) \leftarrow \cdots \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \downarrow d \\
0 & \leftarrow & S_1^\Delta(K) & \xleftarrow{-\Sigma} & \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} S_1^\Delta(\mathscr{U}_{\sigma^0}) & \xleftarrow{-\check{\delta}_1} & \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} S_1^\Delta(\mathscr{U}_{\sigma^1}) & \xleftarrow{-\check{\delta}_2} & \bigoplus_{\sigma^2 \in N_{\mathscr{U}}} S_1^\Delta(\mathscr{U}_{\sigma^2}) \leftarrow \cdots \\
 & & \downarrow d & & \downarrow d & & \downarrow d & & \downarrow d \\
0 & \leftarrow & S_0^\Delta(K) & \xleftarrow{\Sigma} & \bigoplus_{\sigma^0 \in N_{\mathscr{U}}} S_0^\Delta(\mathscr{U}_{\sigma^0}) & \xleftarrow{\check{\delta}_1} & \bigoplus_{\sigma^1 \in N_{\mathscr{U}}} S_0^\Delta(\mathscr{U}_{\sigma^1}) & \xleftarrow{\check{\delta}_2} & \bigoplus_{\sigma^2 \in N_{\mathscr{U}}} S_0^\Delta(\mathscr{U}_{\sigma^2}) \leftarrow \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0 & & 0
\end{array}$$

This leads to a *double complex* $(\mathscr{S}_{*,*}, \bar{\delta}, d)$ defined as

$$\mathscr{S}_{p,q} \coloneqq \bigoplus_{\sigma^p \in N_{\mathscr{U}}} S_q^\Delta(U_{\sigma^p})$$

for all $p,q \geq 0$, and also $\mathscr{S}_{p,q} \coloneqq 0$ otherwise. We denote $\bar{\delta} = (-1)^q \check{\delta}$, the Čech differential multiplied by a -1 on odd rows. The reason for this change of sign is because we want $\mathscr{S}_{*,*}$ to be a double complex, in the sense that the following equalities hold:

$$\bar{\delta} \circ \bar{\delta} = 0, \qquad d \circ d = 0, \qquad \bar{\delta} \circ d + d \circ \bar{\delta} = 0. \tag{2.10}$$

Since $\mathscr{S}_{*,*}$ is a double complex, we can study the associated chain complex $\mathscr{S}_*^{\mathrm{Tot}}$, commonly known as the *total complex*. This is formed by taking the sums of anti-diagonals

$$\mathscr{S}_n^{\mathrm{Tot}} \coloneqq \bigoplus_{p+q=n} \mathscr{S}_{p,q}$$

for each $n \geq 0$. The differentials on the total complex are defined by $d^{\mathrm{Tot}} = d + \bar{\delta}$, which satisfy

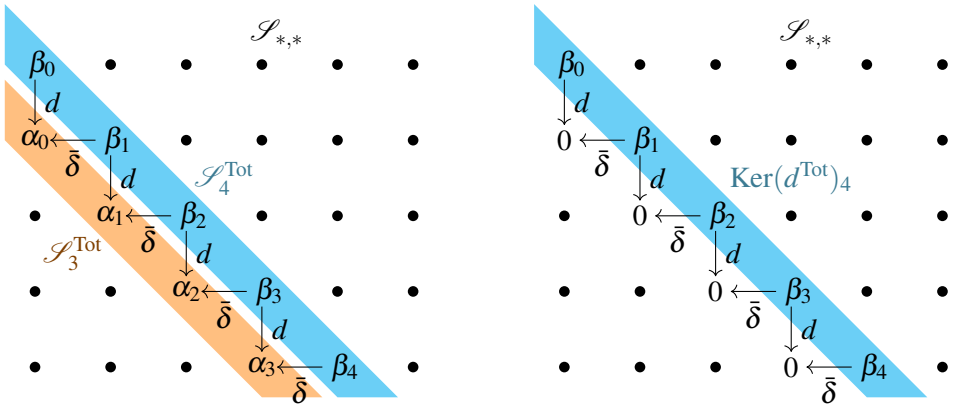


Figure 2.7: $\mathscr{S}_{*,*}$ represented as a lattice for convenience. On the left, the total complex $\mathscr{S}^{\mathrm{Tot}}$ associated to $\mathscr{S}_{*,*}$. Here $(\beta_0,\dots,\beta_4) \in \mathscr{S}_4^{\mathrm{Tot}}$ maps to $(\alpha_0,\dots,\alpha_3) \in \mathscr{S}_3^{\mathrm{Tot}}$, where $d(\beta_i)+\bar{\delta}(\beta_{i+1}) = \alpha_i$ for all $0 \le i \le 3$. On the right, the kernel $\mathrm{Ker}(d^{\mathrm{Tot}})_4$, where $d(\beta_i)+\bar{\delta}(\beta_{i+1}) = 0$ for all $0 \le i \le 3$.

$d^{\mathrm{Tot}} \circ d^{\mathrm{Tot}} = 0$ from equations (2.10), see Figure 2.7 for a depiction of this. Later, in proposition 2.11.1, we will prove that $\mathrm{H}_n(K) \cong \mathrm{H}_n(\mathscr{S}_*^{\mathrm{Tot}})$ for all $n \ge 0$. The problem still remains difficult, since computing $\mathrm{H}_n(\mathscr{S}_*^{\mathrm{Tot}})$ directly might be even harder than computing $\mathrm{H}_n(K)$. The key is that there is a divide and conquer method which allows us to break apart the calculation of $\mathrm{H}_n(\mathscr{S}_*^{\mathrm{Tot}})$ into small, computable steps.

Let us start by computing the kernel $\mathrm{Ker}(d_n^{\mathrm{Tot}})$, which is depicted in Figure 2.7. Recall that we will be working with vector spaces and linear maps all throughout. Let $s = (s_{k,n-k})_{0\le k\le n} \in \mathscr{S}_n^{\mathrm{Tot}}$ be in $\mathrm{Ker}(d_n^{\mathrm{Tot}})$. Then s will satisfy the equations $d(s_{k,n-k}) = -\bar{\delta}(s_{k+1,n-k-1})$ for all $0 \le k < n$. Thus, one can obtain kernel elements by considering subspaces $\mathrm{GK}_{p,q} \subseteq \mathscr{S}_{p,q}$. The subspace $\mathrm{GK}_{p,q}$ is composed of elements $s_{p,q} \in \mathscr{S}_{p,q}$ such that $d(s_{p,q}) = 0$, and there exists a sequence $s_{p-r,q+r} \in \mathscr{S}_{p-r,q+r}$ satisfying equations $d(s_{p-r,q+r}) = -\bar{\delta}(s_{p-r-1,q+r+1})$ for all $0 < r \le p$. Notice that $\mathrm{GK}_{p,q}$ is a subspace of $\mathscr{S}_{p,q}$ since both d and $\bar{\delta}$ are linear. We will see that one has (non-canonical) isomorphisms,

$$\mathrm{Ker}(d_n^{\mathrm{Tot}}) \cong \bigoplus_{p+q=n} \mathrm{GK}_{p,q}. \tag{2.11}$$

This is depicted in Figure 2.8. It turns out that this is true only when we are working with vector spaces. Later, we will work with a more general case where such isomorphisms do not hold. This will be known as the *extension problem*.

Hence, recovering the sets $\mathrm{GK}_{p,q}$ leads to the kernel of d_n^{Tot}. The problem with this approach is that each subspace $\mathrm{GK}_{p,q}$ still requires a large set of equations to be checked. A step-by-step way of computing these is by adding one equation at a time. For this we define the subspaces $\mathrm{GZ}_{p,q}^r \subseteq \mathscr{S}_{p,q}$ where we add the first r equations progressively. That is, we start setting $\mathrm{GZ}_{p,q}^0 = \mathscr{S}_{p,q}$. Then we define $\mathrm{GZ}_{p,q}^1$ to be elements $s_{p,q} \in \mathscr{S}_{p,q}$ such that $d(s_{p,q}) = 0$, or equivalently $\mathrm{GZ}_{p,q}^1 = \mathrm{Ker}(d)_{p,q}$. In an inductive way, for $r \ge 2$ we define $\mathrm{GZ}_{p,q}^r$ to be formed by elements $s_{p,q} \in$

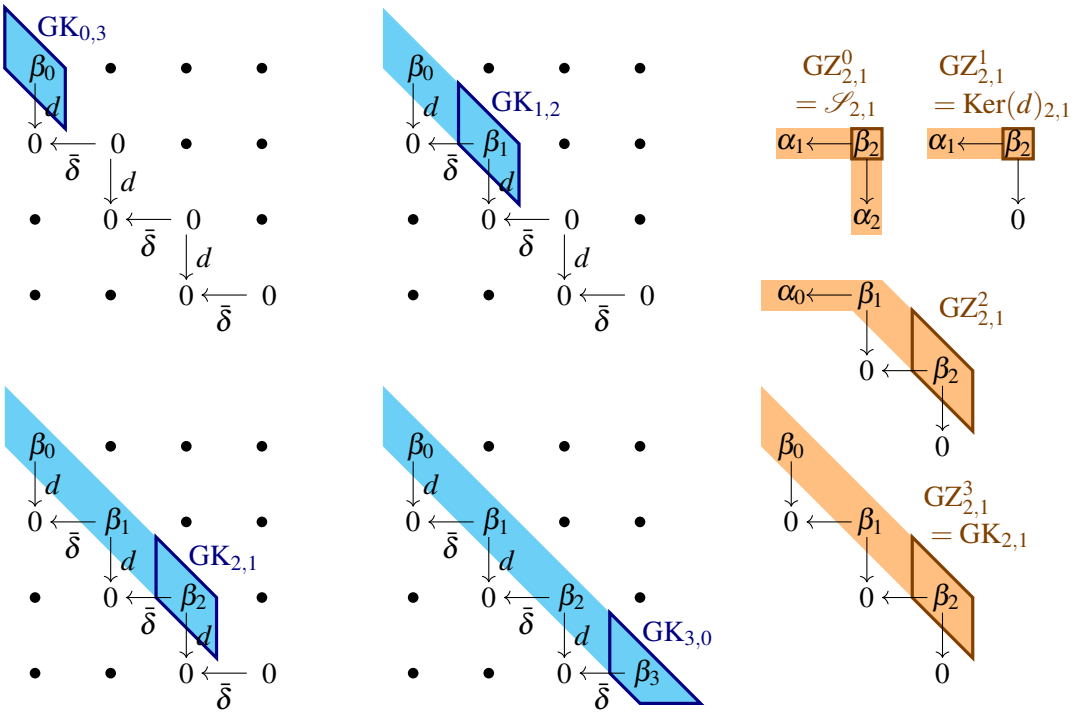


Figure 2.8: On the left, in cyan the four direct summands of $\mathrm{Ker}(d^{\mathrm{Tot}})_4$. The corresponding $\mathrm{GK}_{r,3-r}$ are framed to emphasize that they are respective subspaces of $\mathscr{S}_{r,3-r}$ for all $0 \leq r \leq 3$. On the right, in orange the subspaces $\mathrm{GZ}^r_{2,1}$, eventually shrinking to $\mathrm{GK}_{2,1}$. For convenience, we have labelled $\alpha_2 = d(\beta_2)$, $\alpha_1 = \bar{\delta}(\beta_2)$ and $\alpha_0 = \bar{\delta}(\beta_1)$.

$\mathrm{GZ}^{r-1}_{p,q}$ such that there exists a sequence $s_{p-k,q+k} \in \mathscr{S}_{p-k,q+k}$ satisfying equations $d(s_{p-k,q+k}) = -\bar{\delta}(s_{p-k+1,q+k-1})$ for all $1 \leq k < r$. Then, for all $p,q \geq 0$, we have a decreasing sequence

$$\mathrm{GK}_{p,q} = \mathrm{GZ}^{p+1}_{p,q} \subseteq \mathrm{GZ}^{p}_{p,q} \subseteq \cdots \subseteq \mathrm{GZ}^{0}_{p,q} = \mathscr{S}_{p,q}.$$

For intuition see Figure 2.8, and also Figure 2.10 for a depiction of $\mathrm{GZ}^2_{3,1}$ on a lattice. A very compact way of expressing that, is by the definition $\mathrm{GZ}^r_{p,q} = \mathrm{Ker}(d) \cap (\bar{\delta}^{-1} \circ d)^{r-1}(\mathscr{S}_{p-r+1,q+r-1})$ for all $r \geq 1$, where by $(\bar{\delta}^{-1} \circ d)^r$ we mean composing r-times the preimage $\bar{\delta}^{-1} \circ d$. In particular, since $\mathrm{GZ}^r_{p,q} = \mathrm{GZ}^{p+1}_{p,q}$ for all $r \geq p+1$, we sometimes use the convention $\mathrm{GZ}^{\infty}_{p,q} := \mathrm{GZ}^{p+1}_{p,q} = \mathrm{GK}_{p,q}$.

Now we explain the notation $\mathrm{GK}_{p,q}$ and the isomorphism (2.11). We start defining a *vertical filtration* F^*_V on $\mathscr{S}_{*,*}$ by the following subcomplexes for all $r \geq 0$:

$$F^r_V(\mathscr{S}_{*,*})_{p,q} := \begin{cases} \mathscr{S}_{p,q}, \text{ whenever } p \leq r, \\ 0, \text{ otherwise.} \end{cases}$$

Notice that this filtration increases with the index, so that we have inclusions $F^r_V(\mathscr{S}_{*,*}) \subseteq F^{r+1}_V(\mathscr{S}_{*,*})$ for all $r \geq 0$. Additionally, we obtain isomorphisms $F^p_V(\mathscr{S}_{*,*})/F^{p-1}_V(\mathscr{S}_{*,*}) \cong \mathscr{S}_{p,*}$ for all $p \geq 0$. The filtration F^*_V respects the morphisms in $\mathscr{S}_{*,*}$ in the sense that $d(F^t_V(\mathscr{S}_{*,*})) \subset F^t_V(\mathscr{S}_{*,*})$, and

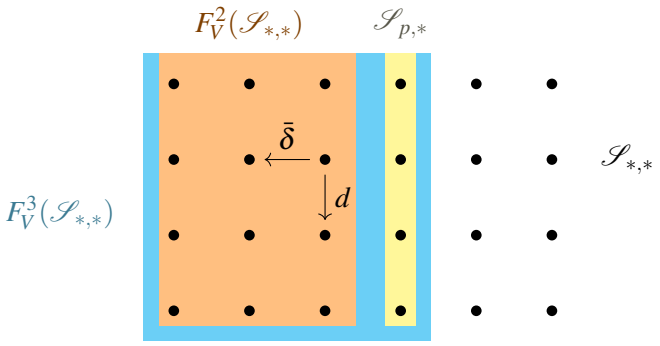


Figure 2.9: Note that $F_V^3(\mathscr{S}_{*,*})/F_V^2(\mathscr{S}_{*,*}) \cong \mathscr{S}_{3,*}$. Also notice that the differentials $\bar{\delta}$ and d respect the vertical filtration F_V^*.

also $\bar{\delta}(F_V^t(\mathscr{S}_{*,*})) \subset F_V^t(\mathscr{S}_{*,*})$. See Figure 2.9 for a depiction of F_V^*. Another point to notice is that F_V^* will filter the total complex $\mathscr{S}_*^{\text{Tot}}$, respecting its differential d^{Tot}. That is, $\mathscr{S}_n^{\text{Tot}}$ will be filtered by subcomplexes,

$$F_V^r \mathscr{S}_n^{\text{Tot}} := \bigoplus_{\substack{p+q=n \\ p \leq r}} \mathscr{S}_{p,q},$$

for all $r \geq 0$. In particular, notice that $\text{Ker}(d^{\text{Tot}})$ also inherits the filtration F_V^*, where we will have filtration sets $F_V^p \text{Ker}(d^{\text{Tot}})_n = F_V^p \mathscr{S}_n^{\text{Tot}} \cap \text{Ker}(d^{\text{Tot}})_n$. We define the *associated modules* of $\text{Ker}(d^{\text{Tot}})_n$ to be the quotients $G_V^p \text{Ker}(d^{\text{Tot}})_n = F_V^p \text{Ker}(d^{\text{Tot}})_n / F_V^{p-1} \text{Ker}(d^{\text{Tot}})_n$, which can be checked to be isomorphic to $\text{GK}_{p,q}$ for all $p+q=n$. This follows by considering morphisms

$$\begin{aligned} G_V^p \text{Ker}(d^{\text{Tot}})_n &\longrightarrow \text{GK}_{p,q}, \\ [(s_{0,n}, s_{1,n-1}, \ldots, s_{p,q}, 0, \ldots, 0)] &\longrightarrow s_{p,q}, \end{aligned} \tag{2.12}$$

which are well-defined since $s_{p,q}$ does not change for representatives of the same class. In fact, this morphism is injective since two classes with the same image will be equal by definition of $G_V^p \text{Ker}(d^{\text{Tot}})_n$. On the other hand, the definition of $\text{GK}_{p,q}$ ensures surjectivity. In particular, since we are working with vector spaces, we have that:

$$\text{Ker}(d_n^{\text{Tot}}) \cong \bigoplus_{p+q=n} G_V^p \text{Ker}(d^{\text{Tot}})_n \cong \bigoplus_{p+q=n} \text{GK}_{p,q}.$$

which justifies isomorphism (2.11).

Next, we explain the notation $\text{GZ}_{p,q}^r$. We introduce the objects

$$Z_{p,q}^r := \left\{ z \in F_V^p \mathscr{S}_{p+q}^{\text{Tot}} \ : \ d^{\text{Tot}}(z) \in F_V^{p-r} \mathscr{S}_{p+q-1}^{\text{Tot}} \right\}$$

for all $r \geq 0$. We can think of these as kernels of d^{Tot} up to some previous filtration. Then,

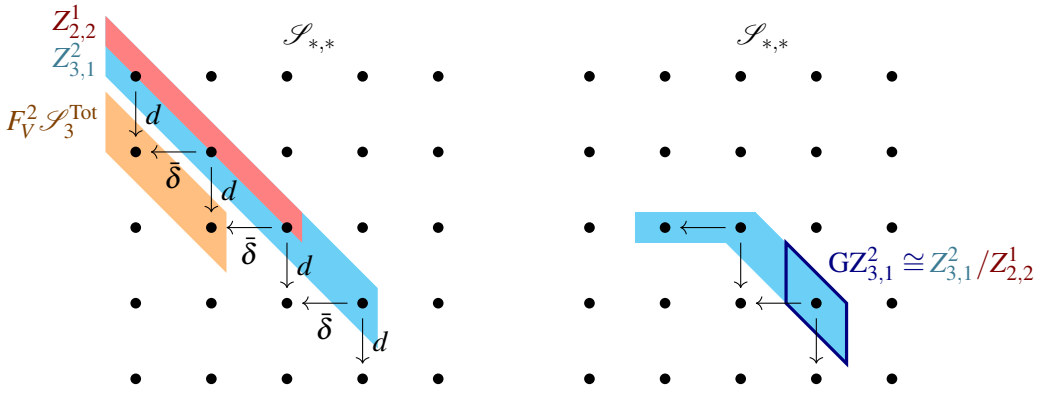


Figure 2.10: On the left the sets $Z^2_{3,1}$ and $Z^1_{2,2}$. On the right their respective quotient $\mathrm{GZ}^2_{3,1}$.

by definition, we have that $Z^0_{p,q} = F^p_V\mathscr{S}^{\mathrm{Tot}}_{p+q}$ and $Z^{p+1}_{p,q} = Z^\infty_{p,q} = F^p_V\mathrm{Ker}(d^{\mathrm{Tot}}_{p+q})$. Using a morphism analogous to (2.12), one can check that the quotients $Z^{r+1}_{p,q}/Z^r_{p-1,q+1}$ are isomorphic to $\mathrm{GZ}^{r+1}_{p,q}$ for all $p+q=n$. This is depicted in Figure 2.10. Thus, computing these quotients increasing $r \geq 0$ leads to the desired kernel $\mathrm{Ker}(d^{\mathrm{Tot}})$. With a little more work, we can do the same for computing the homology.

There is a procedure commonly known as a *spectral sequence* which leads to $\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$ after a series of small, computable steps. This is done in an analogous way as we did before for computing $\mathrm{Ker}(d^{\mathrm{Tot}})$. In this case we will need to take the extra steps of taking quotients by the images of d^{Tot}. First notice that the vertical filtration F^*_V transfers to homology $\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$ by the inclusions $F^p_V\mathscr{S}^{\mathrm{Tot}}_* \subseteq \mathscr{S}^{\mathrm{Tot}}_*$ for all $p \geq 0$. That is, we have filtered sets:

$$F^p_V\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*) := \mathrm{Im}\big(\mathrm{H}_n(F^p_V\mathscr{S}^{\mathrm{Tot}}_*) \longrightarrow \mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)\big)$$

which induce a filtration on $\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$. For this filtration the associated modules will be defined by the quotients $G^r_V\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*) = F^r_V\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)/F^{r-1}_V\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$ for all $r \geq 0$. Notice that in this case, since we are assuming that we are working over a field, we will recover the homology by taking direct sums:

$$\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*) \cong \bigoplus_{r=0}^{n} G^r_V\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*). \tag{2.13}$$

We will say that the direct sum in (2.13) is the *associated graded module* to the *homology of the total complex* $\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$. In Chapter 6, we will deal with filtered modules F^*A together with quotients $G^nA = F^nA/F^{n-1}A$ defining the associated graded module $GA = \bigoplus_{r=0}^n G^rA$; obtaining A from GA is known as the *extension problem*.

Previously, we defined the sets $Z^r_{p,q}$ which are kernels up to filtration. In an analogous way we

define boundaries up to filtration by setting

$$B^r_{p,q} \coloneqq \left\{ d^{\mathrm{Tot}}(c) \in F^p_V \mathscr{S}^{\mathrm{Tot}}_{p+q} \ : \ c \in F^{p+r}_V \mathscr{S}^{\mathrm{Tot}}_{p+q+1} \right\}$$

for all $r \geq 0$, and $p, q \geq 0$. These are images of d^{Tot} coming from a previous filtration. Notice that we will have relations $d^{\mathrm{Tot}}(Z^r_{p,q}) = B^r_{p-r,q+r-1}$ and also $d^{\mathrm{Tot}}(B^r_{p,q}) = 0$. Additionally there is a sequence of inclusions,

$$B^0_{p,q} \subset B^1_{p,q} \subset \cdots \subset B^{q+1}_{p,q} = B^\infty_{p,q} \subset Z^\infty_{p,q} = Z^p_{p,q} \subset \cdots Z^1_{p,q} \subset Z^0_{p,q},$$

for all $p, q \geq 0$.

From the previous discussion, we start defining the *first page* of the spectral sequence as the quotient

$$E^1_{p,q} \coloneqq \frac{Z^1_{p,q}}{Z^0_{p-1,q+1} + B^0_{p,q}} \cong \frac{\mathrm{GZ}^1_{p,q}}{\mathrm{Im}\left(B^0_{p,q} \to \mathrm{GZ}^1_{p,q}\right)},$$

for all $p, q \geq 0$. Recall that $\mathrm{Ker}(d)_{p,q} = \mathrm{GZ}^1_{p,q} = Z^1_{p,q}/Z^0_{p-1,q+1}$ and also $\mathrm{Im}\left(B^0_{p,q} \to \mathrm{GZ}^1_{p,q}\right)$ is isomorphic to $\mathrm{Im}(d)_{p,q}$. Therefore we deduce that $E^1_{p,q} \cong \mathrm{H}_q(\mathscr{S}_{p,*}, d)$. On this page d^{Tot} induces differentials $d^1 : E^1_{p,q} \to E^1_{p-1,q}$. That is, noticing that $d^{\mathrm{Tot}}(Z^1_{p,q}) = B^1_{p-1,q} \subset Z^1_{p-1,q}$ and also $d^{\mathrm{Tot}}(Z^0_{p-1,q+1} + B^0_{p,q}) = d^{\mathrm{Tot}}(Z^0_{p-1,q+1}) + 0 = B^0_{p-1,q}$ we will have that $d^1 : E^1_{p,q} \to E^1_{p-1,q}$ is well-defined. Notice that since $d^{\mathrm{Tot}} \circ d^{\mathrm{Tot}} = 0$ we will also have $d^1 \circ d^1 = 0$ and in particular one can define the homology on the first page $\mathrm{H}_{p,q}(E^1_{*,*}, d^1)$. Since

$$\mathrm{Ker}(d^1) = \frac{Z^2_{p,q}}{Z^2_{p,q} \cap (Z^0_{p-1,q+1} + B^0_{p,q})} = \frac{Z^2_{p,q}}{Z^1_{p-1,q+1} + B^0_{p,q}}, \text{ and } \mathrm{Im}(d^1) = \frac{B^1_{p,q}}{B^0_{p,q}}$$

then the second page will be

$$E^2_{p,q} \coloneqq \mathrm{H}_{p,q}(E^1_{*,*}, d^1) = \frac{\mathrm{Ker}(d^1)}{\mathrm{Im}(d^1)} = \frac{Z^2_{p,q}}{Z^1_{p-1,q+1} + B^1_{p,q}}.$$

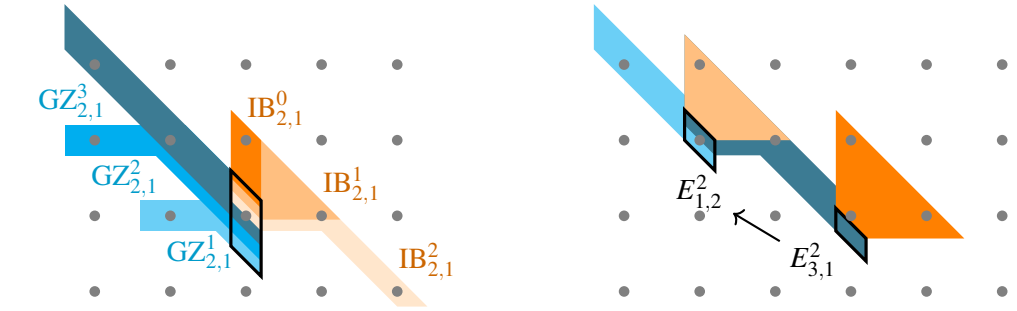
The second page has differential d^2 induced by the total complex differential d^{Tot}. Figure 2.11 illustrates this principle.

Doing the same for all pages we obtain the definition of the r-page:

$$E^r_{p,q} \coloneqq \mathrm{H}_{p,q}(E^{r-1}_{*,*}, d^{r-1}) = \frac{Z^r_{p,q}}{Z^{r-1}_{p-1,q+1} + B^{r-1}_{p,q}}$$

for all $r \geq 2$. Of course, we can express alternatively the r-page terms as:

$$E^r_{p,q} \coloneqq \frac{\mathrm{GZ}^r_{p,q}}{\mathrm{Im}\left(B^{r-1}_{p,q} \to \mathrm{GZ}^r_{p,q}\right)}.$$



$$IB^0_{2,1} \subset IB^1_{2,1} \subset IB^2_{2,1} \subset GZ^3_{2,1} \subset GZ^2_{2,1} \subset GZ^1_{2,1}$$

Figure 2.11: On the left, the different subspaces on $\mathscr{S}_{2,1}$. Here $IB^r_{2,1} = \mathrm{Im}\left(B^r_{2,1} \to GZ^{r+1}_{2,1}\right)$, for all $0 \leq r \leq 2$. The framed region represents $\mathscr{S}_{2,1}$. Brighter colours represent bigger regions than darker colours. Note that blue and orange colours have been assigned to $GZ^*_{2,1}$ and $IB^*_{2,1}$ respectively. On the right, the morphism $d^2 : E^2_{3,1} \to E^2_{1,2}$ on the second page. The two framed regions represent the codomain and domain of d^2, these have been assigned brighter and darker colours, respectively.

Thus, the ∞-page is:

$$E^\infty_{p,q} = \frac{Z^\infty_{p,q}}{Z^\infty_{p-1,q+1} + B^\infty_{p,q}} \cong \frac{GK_{p,q}}{\mathrm{Im}\left(B^\infty_{p,q} \to GK_{p,q}\right)}.$$

Then, for $n = p+q$ one has the equality

$$G^p_V H_n(\mathscr{S}^{\mathrm{Tot}}_*) = \frac{F^p_V H_n(\mathscr{S}^{\mathrm{Tot}}_*)}{F^{p-1}_V H_n(\mathscr{S}^{\mathrm{Tot}}_*)} = \frac{\mathrm{Im}\left(H_n(F^p_V \mathscr{S}^{\mathrm{Tot}}_*) \longrightarrow H_n(\mathscr{S}^{\mathrm{Tot}}_*)\right)}{\mathrm{Im}\left(H_n(F^{p-1}_V \mathscr{S}^{\mathrm{Tot}}_*) \longrightarrow H_n(\mathscr{S}^{\mathrm{Tot}}_*)\right)} \cong \frac{Z^\infty_{p,q}/B^\infty_{p,q}}{Z^\infty_{p-1,q+1}/B^\infty_{p-1,q+1}} \cong E^\infty_{p,q}$$

since $B^\infty_{p-1,q+1} \subseteq B^\infty_{p,q}$. Therefore, computing the spectral sequence is a way of approximating the associated module $G^p_V H_n(\mathscr{S}^{\mathrm{Tot}}_*)$. Thus adding up all of these leads to the result $H_n(\mathscr{S}^{\mathrm{Tot}}_*)$. By convention, since $E^\infty_{p,q} \cong G^p_V H_n(\mathscr{S}^{\mathrm{Tot}}_*)$ we say that $E^*_{p,q}$ *converges* to $H_n(\mathscr{S}^{\mathrm{Tot}})$ and we denote this as

$$E^*_{p,q} \Rightarrow H_n(\mathscr{S}^{\mathrm{Tot}}).$$

Remark. Here we have adopted the definition of $Z^r_{p,q}$ and $B^r_{p,q}$ that one can find in [83]. Other sources such as [15] and [76] use the same notation for other terms.

So far, we have studied spectral sequences for vertical filtrations. Similarly, there is a horizontal filtration,

$$F^r_H \mathscr{S}^{\mathrm{Tot}}_n := \bigoplus\nolimits_{q \leq r}^{p+q=n} \mathscr{S}_{p,q},$$

for all $r \geq 0$. We can apply the same argument to this filtration, to obtain a spectral sequence

$${}_H E^*_{*,*} \Rightarrow H_n(\mathscr{S}^{\mathrm{Tot}}).$$

An intuitive way of thinking of this is by applying a symmetry about the diagonal $x = y$ on the previous discussion. Thus the first page is computed with the homology with respect to horizontal differentials, the second with respect to vertical differentials, and so on. This leads easily to the following widely known result:

Proposition 2.11.1. $\mathrm{H}_n(\mathscr{S}_*^{\mathrm{Tot}}) \cong \mathrm{H}_n(K)$.

Proof. We will prove this by showing that ${}_HE^*_{*,*}$ *collapses* on the second page, which means that all differentials in the spectral sequence are trivial from the second page onwards. In order to turn to the first page, we need to compute homology with respect to the horizontal differentials δ. As shown in section 2.10, the Čech chain complexes associated to $S_*^{\Delta}(K^{\mathscr{U}})$ are exact, so that:

$$ {}_HE^1_{*,*} := H_{0,q}({}_HE^0_{*,*}, \bar{\delta}) = \begin{cases} S_q^{\Delta}(K) & \text{if } p = 0 \text{ and } q \geq 0 \\ 0 & \text{otherwise.} \end{cases} $$

After this one can compute the second page by the homology with respect to vertical differentials d induced on the first page,

$$ {}_HE^2_{*,*} := H_{0,q}({}_HE^1_{*,*}, d) = \begin{cases} H_q(K) & \text{if } p = 0 \text{ and } q \geq 0 \\ 0 & \text{otherwise.} \end{cases} $$

To proceed to the next page, we would need to consider homology with respect to diagonal differentials,

$$ d_2 : {}_HE^2_{p,q} \longrightarrow {}_HE^2_{p+1,q-2}. $$

Since the second page $E^2_{p,q}$ has only one non-zero column $p = 0$, computing homology with respect to d_2 leaves this page intact. The same happens when we consider for any $r > 2$ homology with respect to differentials

$$ d_r : {}_HE^r_{p,q} \longrightarrow {}_HE^r_{p+r-1,q-r}. $$

Thus, we say that ${}_HE^*_{p,q}$ has *collapsed* on the second page, which is usually denoted as ${}_HE^2_{p,q} = {}_HE^{\infty}_{p,q}$. Each diagonal has a unique nonzero entry ${}_HE^{\infty}_{0,q} \cong H^q(K)$. In particular, we have isomorphisms

$$ {}_HE^{\infty}_{0,n} \cong \mathrm{H}_n(\mathscr{S}_*^{\mathrm{Tot}}) \cong \mathrm{H}_n(K), $$

for all $n \geq 0$. Notice that there is no extension problem here, since only the 0-column is nontrivial.

□

Proposition 2.11.1 is a particular instance of the more general double complex spectral sequence theorem [83, Thm. 2.15].

Therefore, using proposition 2.11.1, we have that the spectral sequence converges to the wanted result

$$E^*_{p,q} \Rightarrow \mathrm{H}_n(K).$$

In particular, since we are in the category of vector spaces, there are no extension problems. Thus, we have an isomorphism

$$\mathrm{H}_n(K) \cong \bigoplus_{p+q=n} E^\infty_{p,q} \, .$$

Using the spectral sequence machinery, one can deduce a homology analogue of the Nerve theorem.

Corollary 2.11.2. *Suppose that K is a simplicial complex together with a cover by subcomplexes* $\mathscr{U} = \{U_i\}_{i=0}^N$. *Assume that* $\mathrm{H}_n(\mathscr{U}_\sigma) = 0$ *for all* $n \geq 0$ *and all simplices* $\sigma \in N_{\mathscr{U}}$. *Then there are isomorphisms* $\mathrm{H}_n(X) \simeq \mathrm{H}_n(N_{\mathscr{U}})$ *for all* $n \geq 0$.

Proof. Notice that the first page terms $E^*_{p,q}(X, \mathscr{U})$ are all zero for higher rows $q > 0$. Thus, the sequence $E^*_{p,q}(X, \mathscr{U})$ collapses at the first page and the only nonzero terms are $E^1_{p,0}(X, \mathscr{U}) \simeq S^\Delta_p(N_{\mathscr{U}})$ for all $p \geq 0$ (recall the sequence (2.8)). Thus the result follows. □

Now, instead of making assumptions about our covering elements, we assume that our nerve $N_{\mathscr{U}}$ is one dimensional. In such a case, the terms $E^*_{p,q}(X; \mathscr{U})$ are all zero whenever $p > 1$. The spectral sequence collapses on the second page, where the nonzero terms $E^*_{p,q}(X; \mathscr{U})$ are equal to the Čech homology groups $\check{\mathrm{H}}_p(N_{\mathscr{U}}; S^\Delta_q(K^{\mathscr{U}}))$ for $p = 0, 1$ and all $q \geq 0$.

Corollary 2.11.3. *Suppose that K is a simplicial complex together with a cover by subcomplexes* $\mathscr{U} = \{U_i\}_{i=0}^N$. *Assume that* $\dim(N_{\mathscr{U}}) < 2$. *Then there are isomorphisms*

$$\mathrm{H}_0(K) \simeq \check{\mathrm{H}}_0(N_{\mathscr{U}}; S^\Delta_0(K^{\mathscr{U}})), \qquad \mathrm{H}_n(K) \simeq \check{\mathrm{H}}_0(N_{\mathscr{U}}; S^\Delta_n(K^{\mathscr{U}})) \oplus \check{\mathrm{H}}_1(N_{\mathscr{U}}; S^\Delta_{n-1}(K^{\mathscr{U}})) \, ,$$

for all $n \geq 0$.

Remark. Following on the final remark from section 2.10, we could consider the Mayer-Vietoris spectral sequence associated to a covered regular CW-complex $(X, \mathscr{U})$

$$E^1_{p,q}(X, \mathscr{U}) = \bigoplus_{\sigma^p \in N_{\mathscr{U}}} \mathrm{H}_q(\mathscr{U}_{\sigma^p}) \Rightarrow \mathrm{H}_{p+q}(X) \, .$$

This whole section can be rewritten in this context.

2.12 Spectral Sequences and Twisted Double Complexes

In section 2.11 we saw an example of a very concrete spectral sequence. However, spectral sequences come up very naturally in many situations [34]. A more general setup [83] considers a graded R-module A_* for some ring R, together with differentials $d_n : A_n \to A_{n-1}$ for all $n \geq 0$ and a bounded filtration

$$F^0 A_* \subseteq F^1 A_* \subseteq \cdots \subseteq F^m A_* = A_* \ .$$

Additionally, we assume that the differentials are well behaved with respect to the filtration, in the sense that there are inclusions $d_*(F^p A) \subseteq F^p A$ for all $p \geq 0$. We say that A_* is a *filtered differential graded module* and denote this by the triple (A, d, F). In this case, we obtain the objects

$$Z^r_{p,q} := F^p A_{p+q} \cap d^{-1}(F^{p-r} A_{p+q})$$

and also

$$B^r_{p,q} := F^p A_{p+q} \cap d(F^{p+r} A_{p+q})$$

which are generalizations of the previously defined objects from section 2.11. In this case, the spectral sequence terms are given by the quotients

$$E^r_{p,q} = \frac{Z^r_{p,q}}{Z^{r-1}_{p,q} + B^{r-1}_{p,q}} \ ,$$

for all $r \geq 0$ and all $p, q \in \mathbb{Z}$. One might check [83, Thm. 2.6.] that the following properties hold:

- d induces differentials $d^r : E^r_{p,q} \to E^r_{p-r,q+r-1}$,
- $E^{r+1}_{*,*} \simeq \mathrm{H}_*(E^r_{*,*}, d^r_{*,*})$
- $E^1_{p,q} \simeq \mathrm{H}_{p+q}(F^p A_* / F^{p-1} A_*)$
- $E^\infty_{p,q} \simeq F^p \mathrm{H}_{p+q}(A_*) / F^{p-1} \mathrm{H}_{p+q}(A_*)$,

for all $r \geq 0$ and all $p, q \in \mathbb{Z}$. Then, $E^r_{p,q}$ converges to $\mathrm{H}_*(A)$, which is denoted as

$$E^1_{p,q} = \mathrm{H}_{p+q}(F^p A_* / F^{p-1} A_*) \Rightarrow F^p \mathrm{H}_{p+q}(A_*) / F^{p-1} \mathrm{H}_{p+q}(A_*) \ .$$

This means that we can use $E^*_{p,q}$ as a means to obtain the homology groups $\mathrm{H}_n(A)$ for all $n \geq 0$; notice that this is not exactly true, as we only recover the consecutive quotients induced by the filtration F^*. In this case, the $\mathrm{H}_*(A)$ has an *associated graded module* given by the direct sums

$$G\mathrm{H}_n(A) = \bigoplus_{p=0}^{n} F^p \mathrm{H}_n(A) / F^{p-1} \mathrm{H}_n(A) \simeq \bigoplus_{p=0}^{n} E^\infty_{p,n-p} \ ,$$

for all $n \geq 0$. The *extension problem* consists in obtaining $\mathrm{H}_n(A)$ from the various degrees from $G\mathrm{H}_n(A)$. In particular, we have an isomorphism $F^0\mathrm{H}_n(A) \simeq E^\infty_{0,n}$ as well as short exact sequences

$$0 \longrightarrow F^{r-1}\mathrm{H}_n(A) \longrightarrow F^r\mathrm{H}_n(A) \longrightarrow E^\infty_{r,n-r} \longrightarrow 0\,,$$

for all $1 \leq r \leq n$. Solving this problem is nontrivial in general.

A *morphism of spectral sequences* is a sequence of bigraded morphisms $f^r : E^r_{p,q} \to \overline{E}^r_{p,q}$ that commute with the spectral sequence differentials, ie. $d_r \circ f^r = d_r \circ f^r$ for all $r \geq 0$. Apart from that, these morphisms satisfy $f^{r+1} = H(f^r)$ for all $r \geq 0$. Suppose that $(\overline{A}_*, \overline{d}, \overline{F})$ is another filtered differential graded module together with its corresponding spectral sequence $\overline{E}^r_{p,q}$. Consider a morphism $f : A_* \to \overline{A}_*$ that intertwines with the differentials $f \circ d = \overline{d} \circ f$ and also preserves filtrations $f(F^pA_*) \subseteq \overline{F}^p(\overline{A}_*)$ for all $p \geq 0$. This induces a morphism of spectral sequences

$$f^r_{p,q} : E^r_{p,q} \to \overline{E}^r_{p,q}$$

by [83, Thm. 3.5]. We will denote by **SpSq** the category of spectral sequences, while we will denote by $\mathbf{SpSq}^{[0,\infty)}$ the category of functors $F : \mathbf{R} \to \mathbf{SpSq}$.

We apply this general theory of spectral sequences to a relevant spectral sequence for this thesis. Consider the situation of a finite filtered regular CW-complex

$$F^0X \subseteq F^1X \subseteq \cdots \subseteq F^nX = X\,, \tag{2.14}$$

whose differential respects filtration in the sense that $d(F^pX) \subseteq F^pX$ for all $p \geq 0$. For instance, this is the case with the filtration of the Mayer-Vietoris blowup complex associated to a space X with a cover $\mathscr{U}$

$$F^0\Delta X^{\mathscr{U}} \subseteq F^1\Delta_{\mathscr{U}}X^{\mathscr{U}} \subseteq \cdots \subseteq F^n\Delta_{\mathscr{U}}X^{\mathscr{U}}\,.$$

As we have seen in section 2.11, we might define a double complex $\mathscr{S}_{p,q}(X,\mathscr{U})$ whose associated spectral sequence converges to the target $\mathrm{H}_*(\Delta X^{\mathscr{U}}) \simeq \mathrm{H}_*(X)$. The key in the setup from section 2.11 is that the differential behaves very well with respect to the filtration F^*; for each cell $a \in F^pX \setminus F^{p-1}X$ we have that $d(a) \in F^pX \setminus F^{p-2}X$. This is what allows us to proceed as we did in section 2.11, by defining first a double complex and then studying the associated spectral sequence. In our current filtered complex (2.14), we have just the property $d(F^pX) \subseteq F^pX$. In this case, one might also define an associated spectral sequence [83, Thm. 2.6.], which is related to what is known as a *multicomplex* or a *twisted double complex* [35, 77] defined as

$$\mathscr{T}_{p,q}(X,F) = C^{\mathrm{cell}}_{p+q}(F^pX)/C^{\mathrm{cell}}_{p+q}(F^{p-1}X)$$

for all indices $p, q \in \mathbb{Z}$. Arguably, "twisted" is one of the most used words in mathematics admiting various different meanings. In this context, by "twisted" we will mean that there are differentials

$$t^r_{p,q} : \mathscr{T}_{p,q}(X,F) \to \mathscr{T}_{p-r,q+r-1}(X,F)$$

for all $r \geq 0$; the behaviour of the twisted complex is illustrated on figure 2.12. These twisted differentials $t^r_{p,q}$ are given by the restrictions of the cellular differential d^{cell}_{p+q} to the morphisms

$$C^{\text{cell}}_{p+q}(F^pX)/C^{\text{cell}}_{p+q}(F^{p-1}X) \to C^{\text{cell}}_{p+q-1}(F^{p-r}X)/C^{\text{cell}}_{p+q-1}(F^{p-r-1}X)$$

for all values $r \geq 0$. The twisted differentials satisfy the equalities

$$\sum_{i+j=n} t^i_{p-j,q+j-1} \circ t^j_{p,q} = 0$$

for all values $n \geq 0$ and all $p, q \geq 0$. Notice that this condition is satisfied by the fact that d^{cell} is a differential on X. Notice that if $t^i = 0$ for all $i > 1$, then we have that $\mathscr{T}_{p,q}(X,F)$ is just a double complex.

We define the *twisted total complex* as the direct sum

$$\mathscr{T}^{\text{Tot}}_n(X,F) = \bigoplus_{p+q=n} \mathscr{T}_{p,q}(X,F)$$

where given an element $a \in \mathscr{T}_{p,q}(X,F)$, its differential is computed by taking the sum

$$\sum_{i \geq 0} t^i_{p,q}(a) \ .$$

Of course, one might see that there are filtration preserving isomorphisms

$$F^p_V \mathscr{T}^{\text{Tot}}_n(X,F) \simeq C^{\text{cell}}_n(F^pX) \ ,$$

where the differentials coincide and where F_V denotes the vertical filtration, which is defined as that of the total complex on section 2.11. Also, one might obtain explicit formulations for the corresponding generalizations of the submodules $GK_{p,q} \subseteq \mathscr{T}_{p,q}(X,F)$ and $GB_{p,q} \subseteq \mathscr{T}_{p,q}(X,F)$ that we have seen in section 2.11, see [77] for details. Notice that in this case the spectral sequence $E^r_{p,q}$ is not necessarily restricted to the first quadrant. However, as we assumed that F_*X is a finite regular CW-complex, the spectral sequence will collapse eventually along each diagonal, see [12]. Thus, we have that

$$E^1_{p,q}(X,F) = \mathrm{H}_q\big(\mathscr{T}_{p,*}(X,F)\big) \Rightarrow \mathrm{H}_{p+q}(X) \ .$$

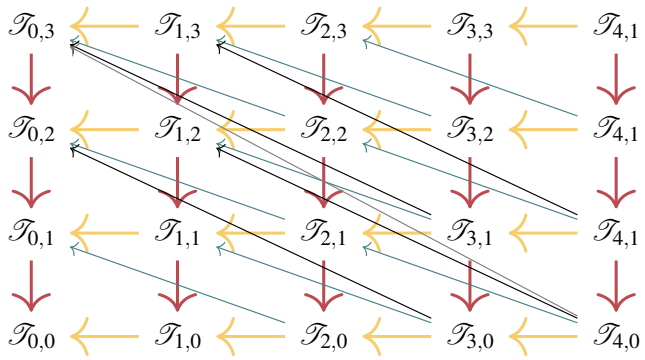


Figure 2.12: Twisted double complex

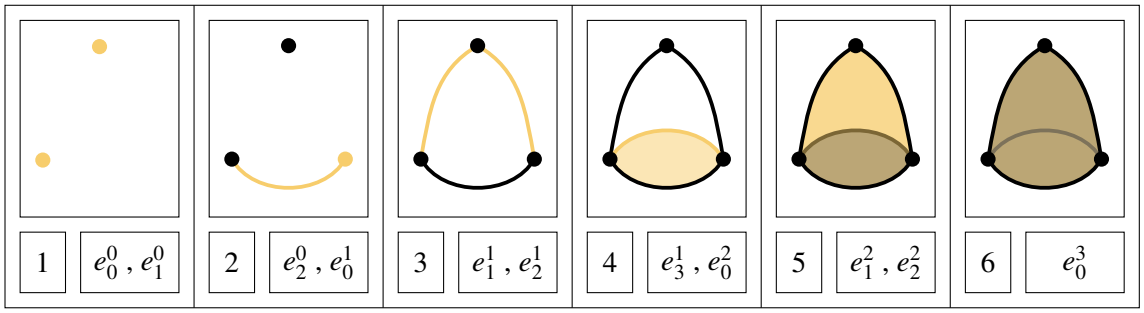


Figure 2.13: Filtered cone. New cells are indicated in yellow.

We proceed to illustrate the twisted double complex spectral sequence. Consider a filtered regular CW-complex structure on a cone as depicted on figure 2.13, F^*C. Notice that $F^0X = \emptyset$, the complement $F^5X \setminus F^4X$ contains the two "faces" from the cone, while $F^6X \setminus F^5X$ contains the 3-cell e_0^3 filling the empty cone. The twisted double complex $\mathscr{T}_{*,*}(C,F)$ is depicted on figure 2.14.

We define the *k-truncated filtration* F_k at level $k \geq 0$, by

$$F_k^n X = \begin{cases} F^k X \ , \text{ for all } n > k \\ F^n X \ , \text{ for } n \leq k \end{cases}$$

so that we have the convergence

$$E^1_{p,q}(X,F_k) = \mathrm{H}_q\big(\mathscr{T}_{p,*}(X,F_k)\big) \Rightarrow \mathrm{H}_{p+q}(F^kX) \ ,$$

where it is worth noticing that $\mathscr{T}_{p,*}(X,F_k) = 0$ for $p > k$. As we are working over a field, we have isomorphisms

$$\bigoplus_{r=0}^{k} E^\infty_{r,n-r}(X,F_k) \simeq \mathrm{H}_n(F^kX) \ ,$$

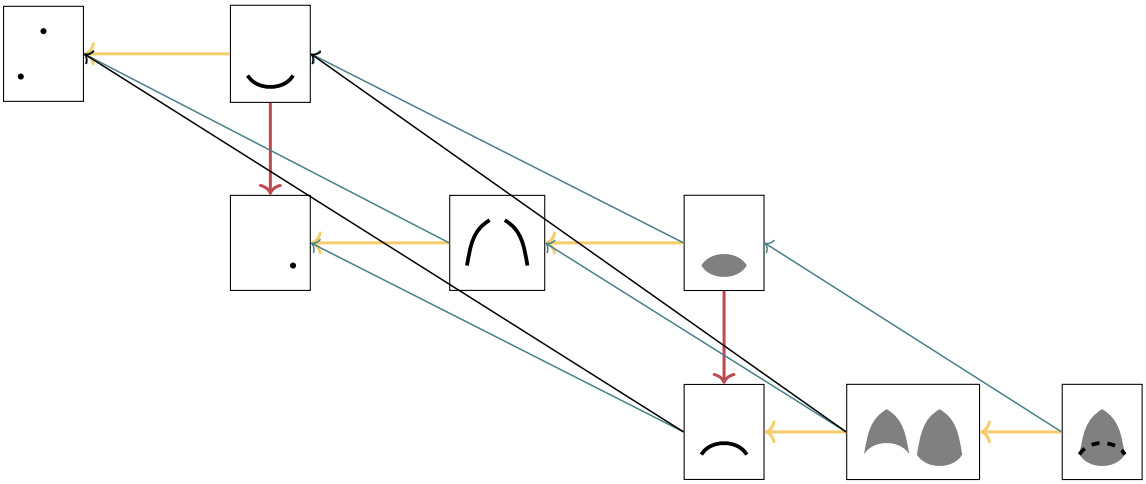


Figure 2.14: Twisted double complex for F^*C. Drawn are all nontrivial terms and nontrivial arrows. Generators on each term are drawn. The top left term corresponds to $\mathscr{T}_{1,-1}(C,F)$ while the bottom right term represents $\mathscr{T}_{6,-3}(C,F)$.

for all $k \geq 0$, where the infinity page terms are computed using the formula

$$E^{\infty}_{r,n-r}(X,F_k) = \frac{Z^{\infty}_{r,n-r}}{Z^{\infty}_{r-1,n-r+1} + B^{k-r}_{r,n-r}}$$

for all $0 \leq k$ and all $0 \leq r \leq k$. For all pairs of indices $i \leq k$, there is a morphism of twisted double complexes $\mathscr{T}_{*,*}(X,F_i) \to \mathscr{T}_{*,*}(X,F_k)$, meaning that there are morphisms $f_{p,q} : \mathscr{T}_{*,*}(X,F_i) \to \mathscr{T}_{*,*}(X,F_k)$ for all $p,q \in \mathbb{Z}$ such that these commute with the twisted morphisms $t^r_{p,q} \circ f_{p,q} = f^r_{p-r,q+r-1} \circ t^r_{p,q}$ for all $r \geq 0$. This morphism induces a morphism of filtered twisted total complexes $f : \mathscr{T}^{\mathrm{Tot}}_*(X,F_i) \to \mathscr{T}^{\mathrm{Tot}}_*(X,F_k)$ which respects vertical filtrations, in the sense that $f(F^r_V \mathscr{T}^{\mathrm{Tot}}_*(X,F_i)) \subseteq F^r_V \mathscr{T}^{\mathrm{Tot}}_*(X,F_k)$ for all $r \geq 0$. As mentioned earlier, by [83, Thm. 3.5] this induces a morphism of spectral sequences $f^r_{p,q} : E^r_{p,q}(X,F_i) \to E^r_{p,q}(X,F_k)$. Because of this, notice that the morphism $\mathrm{H}_n(F^iX) \to \mathrm{H}_n(F^kX)$ can be seen as a direct sum of morphisms:

$$\begin{array}{ccccccc}
E^{\infty}_{0,n}(X,F^i) & E^{\infty}_{1,n-1}(X,F^i) & \cdots & E^{\infty}_{i,n-i}(X,F^i) & 0 & \cdots & 0 \\
\downarrow{\scriptstyle f^{\infty}_{0,n}} & \downarrow{\scriptstyle f^{\infty}_{1,n-1}} & & \downarrow{\scriptstyle f^{\infty}_{i,n-i}} & & & \\
E^{\infty}_{0,n}(X,F^k) & E^{\infty}_{1,n-1}(X,F^k) & \cdots & E^{\infty}_{i,n-i}(X,F^k) & E^{\infty}_{i+1,n-i-1}(X,F^k) & \cdots & E^{\infty}_{k,n-k}(X,F^k)
\end{array}$$

where the $f^{\infty}_{r,n-r}$ are surjective for all $0 \leq r \leq i$. Consequently we can write the cokernel as

$$\mathrm{Coker}\big(\mathrm{H}(F^iX) \to \mathrm{H}(F^kX)\big) \simeq \bigoplus_{r=i+1}^{k} E^{\infty}_{r,n-r+1}(X,F^k)$$

while the kernel is

$$\begin{aligned}\ker\big(\mathrm{H}(F^iX)\to\mathrm{H}(F^kX)\big) &\simeq \bigoplus_{r=0}^{i}\ker\left(\frac{Z^\infty_{r,n-r}}{Z^\infty_{r-1,n-r+1}+B^{i-r}_{r,n-r}}\to\frac{Z^\infty_{r,n-r}}{Z^\infty_{r-1,n-r+1}+B^{k-r}_{r,n-r}}\right)\\ &\simeq \bigoplus_{r=0}^{i}\frac{B^{k-r}_{r,n-r}}{Z^\infty_{r-1,n-r+1}+B^{i-r}_{r,n-r}}\ .\end{aligned}$$

Chapter 3

Background in Topological Data Analysis

3.1 From Data to Complexes

In this section, we will start with a finite set $\mathbb{X} \subset \mathbb{R}^m$ and will try to understand the underlying shape of $\mathbb{X}$. The usual procedure consists in taking a positive real number r and define closed balls $B_r(x)$ of radius r around each point $x \in \mathbb{X}$. This leads to a family of covers $\mathscr{B}_r(\mathbb{X}) = \{B_r(x)\}_{x\in\mathbb{X}}$ for all real numbers $r \geq 0$. The union of the open balls $\bigcup_{x\in\mathbb{X}} B_r(x)$ is often denoted by $\mathbb{X}^\varepsilon$ and called the ε *offset* of $\mathbb{X}$. By looking at the nerves, we obtain a increasing family of simplicial complexes $N_{\mathscr{B}_r(\mathbb{X})}$ such that $N_{\mathscr{B}_r(\mathbb{X})} \subset N_{\mathscr{B}_s(\mathbb{X})}$ for all $r \leq s$. Each nerve $N_{\mathscr{B}_r(\mathbb{X})}$ is usually called the *Čech Complex* of radius r on $\mathbb{X}$, and is denoted as $\check{\mathrm{C}}_r(\mathbb{X})$. The main reason why we are choosing balls around each point is because of the following result commonly known as the Nerve Theorem [58, §III.2].

Theorem 3.1.1. *Let X be a compact space together with a finite cover $\mathscr{U} = \{U_i\}_{i\in I}$. If each element of the cover U_i is closed and convex, then $|N_{\mathscr{U}}|$ and $\cup_{i\in I} U_i$ have the same homotopy type.*

Proof. First, notice that a closed and convex set is contractible and any intersection $\mathscr{U}_\sigma$ for $\sigma \in N_{\mathscr{U}}$ will also be closed and convex, thus contractible. Then the result follows by Theorem 2.8.6

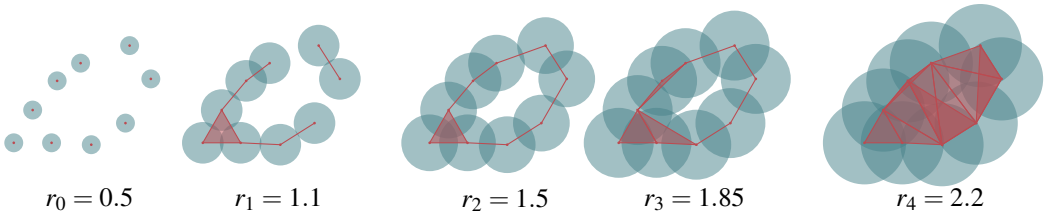


Figure 3.1: Čech complex at different radii.

using the fact that each intersection $\mathscr{U}_\sigma$ is contractible for all $\sigma \in N_{\mathscr{U}}$. More concisely, one has the chain of homotopy equivalences

$$X \simeq \Delta_{N_{\mathscr{U}}}\left(X^{\mathscr{U}}\right) \simeq \Delta_{N_{\mathscr{U}}}\left(*_{N_{\mathscr{U}}}\right) \simeq |N_{\mathscr{U}}| \,.$$

□

Thus, from this result we can ensure that the Čech complexes $\check{\mathrm{C}}_r(\mathbb{X})$ are homotopic to the union of balls $\bigcup_{x\in\mathbb{X}} B_r(x)$. Additionally, one might see $\check{\mathrm{C}}_r(\mathbb{X})$ as a filtered complex with *filtration values* $r \in \mathbf{R}$, so that $\check{\mathrm{C}}_r(\mathbb{X}) \subseteq \check{\mathrm{C}}_s(\mathbb{X})$ for all pairs $r \leq s$ from $\mathbf{R}$. See figure 3.1 for a depiction of a point cloud $\mathbb{X}$ together with the covers $\mathscr{B}_r(\mathbb{X})$ depicted in blue and the Čech complex depicted in red. An inconvenience about the Čech complex is that it might be expensive computationally to check whether a high intersection of balls in $\mathscr{B}_r(\mathbb{X})$ is empty or not. To facilitate computations, we consider the following definition.

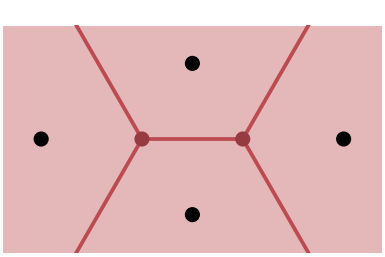
Definition 3.1.2. Let $r > 0$, we define the *Vietoris-Rips complex* of radius r, denoted by $\mathrm{VR}_r(\mathbb{X})$, to be the maximal complex in **SpCpx** such that the 1-simplex $[x,y]$ is contained in $\mathrm{VR}_r(\mathbb{X})$ if and only if $d^2(x,y) < r/2$, where d^2 is the Euclidean distance in $\mathbb{R}^n$.

Unwrapping definition 3.1.2, the Vietoris-Rips complex contains the same one-skeleton as the Čech complex. The differences start at higher simplices. Since we are taking the *maximal* simplicial complex, we include a simplex σ to $\mathrm{VR}_r(\mathbb{X})$ if and only if all the edges $\tau \preceq \sigma$ are contained in $\mathrm{VR}_r(\mathbb{X})$. This is much easier to compute than the Čech Complex, but at the expense of theoretical guarantees. Unfortunately, the nerve theorem does not hold in this case; in general the complex $\mathrm{VR}_r(\mathbb{X})$ is not isomorphic to the union $\bigcup_{x\in\mathbb{X}} B_r(x)$. There is still a good property for $\mathrm{VR}_r(\mathbb{X})$: for all radii $r > 0$ we have a chain of inclusions

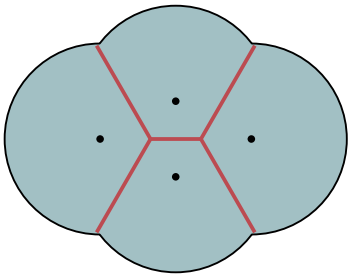
$$\check{\mathrm{C}}_r(X) \subseteq \mathrm{VR}_r(\mathbb{X}) \subseteq \check{\mathrm{C}}_{\sqrt{2}r}(\mathbb{X}) \,,$$

where the first inclusion is direct and the second inclusion can be checked, see [58, §III.2]. Thus, we might say that $\mathrm{VR}_r(\mathbb{X})$ approximates the shape of the point cloud $\mathbb{X}$. In fact, the Smale-Niyogi-Weinberger theorem [89] ensures that both the Čech and Vietoris Rips complexes on sampled points from a manifold preserve the homotopy type almost surely for some radii $r > 0$.

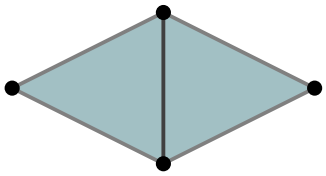
A difficulty with both the Čech complex and the Vietoris-Rips complex is that both become highly *non-sparse* as one increases the radius. To see this, notice that whenever r is big enough, then $\check{\mathrm{C}}_r(\mathbb{X})$ is isomorphic to the standard simplex $\Delta^{|\mathbb{X}|}$ whose size becomes impractical to work with for relatively small cardinalities $|\mathbb{X}|$. In order to avoid this problem, one might limit the dimension of the complexes that we are working with to obtain *sparse* simplicial complexes. One



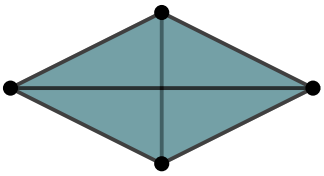
Voronoi diagram



Cover $\mathscr{R}_r(\mathbb{X})$



Alpha complex



Čech or Vietoris-Rips complex

Figure 3.2: Depiction of the Voronoi cells, the Alpha complex and the Čech complex for four fixed points in the plane. We have fixed some radius $r > 0$ which is big enough for $\check{C}_r(\mathbb{X})$ to be isomorphic to Δ^4.

strategy is to use a regular CW-structure on $\mathbb{R}^n$ known as the *Voronoi* diagram. To define this complex, we define the voronoi closed n-cells V_u about each point $u \in \mathbb{X}$ by setting these to be the closed regions

$$V_u := \left\{ x \in \mathbb{R}^n \mid d^2(x,u) \leq d^2(x,v) \text{ for all } v \in \mathbb{X} \right\} .$$

Consider the cover given by the Voronoi cells $\mathscr{V}(\mathbb{X}) = \{V_u \mid u \in \mathbb{X}\}$. We define the Delaunay complex $\mathscr{D}(\mathbb{X})$ as the nerve $N_{\mathscr{V}(\mathbb{X})}$ of the cover $\mathscr{V}(\mathbb{X})$. Next, we introduce a filtration into the Delaunay complex by considering the balls $\mathscr{B}_r(\mathbb{X})$, and intersect these with the Voronoi cells, defining the closed convex sets $V_u \cap B_r(u)$ for all $r \geq 0$. These sets form a cover $\mathscr{R}_r(\mathbb{X}) = \{V_u \cap B_r(u) \mid u \in X\}$. The nerve $N_{\mathscr{R}_r(\mathbb{X})}$ is known as the *alpha complex* $A_r(\mathbb{X})$. This has good properties since Theorem 3.1.1 applies, so that

$$\mathbb{X}^r = \bigcup_{x \in \mathbb{X}} B_r(x) = \bigcup_{x \in \mathbb{X}} V_x \cap B_r(x) \simeq A_r(\mathbb{X}) \subseteq \mathscr{D}(\mathbb{X}) ,$$

for all $r \geq 0$. Thanks to using the Delaunay complex, the alpha complex $A_r(\mathbb{X})$ is much smaller than either $\mathrm{VR}_r(\mathbb{X})$ or $\check{\mathrm{C}}_r(\mathbb{X})$ both in cardinality and dimension.

We have briefly reviewed three simplicial complexes $\check{\mathrm{C}}_r(\mathbb{X})$, $\mathrm{VR}_r(\mathbb{X})$ and $A_r(\mathbb{X})$ which all have $\mathbb{X}$ as the common vertex set. However, sometimes the problem is that the set $\mathbb{X}$ is far too big, and building a simplicial complex whose vertices are all points in $\mathbb{X}$ is not computationally

feasible. A construction to address this problem is the *strong Witness Complex* $W_\infty(\mathbb{X})$, which was introduced by G. Carlsson and Vin De Silva in [107]. Essentially we start by picking a subset $\mathscr{L} \subseteq \mathbb{X}$ of points, often called *landmark* points which will form the vertices of $W_\infty(\mathbb{X})$. These landmark points are usually selected at random or using the *min-max* selection method. Next for any pair of landmark points $u, v \in \mathscr{L}$, we include the edge $[u,v]$ between them if there exists a vertex $x \in \mathbb{X}$ such that v and u are the closest points from $\mathscr{L}$ to x. In this case we say that this x is a *witness* to the edge $[u,v]$. For higher dimensions, given an n-simplex σ with set of vertices $U = \{u_0, \ldots, u_n\} \subset \mathscr{L}$, we include this σ in $W_\infty(\mathbb{X})$ if all the faces from σ are included in $W_\infty(\mathbb{X})$ and if there exists a point (a witness) $x \in \mathbb{X}$ such that:

$$\max_{u \in U} \| x - u \| \leq \min_{p \in \mathscr{L} - U} \| x - p \| .$$

Checking whether a high dimensional simplex lies in the strong Witness complex might become computationally expensive. This is why one might look instead at the *Weak Witness complex* $W_1(\mathbb{X})$, whose one-skeleton coincides with the strong witness complex. The complex $W_1(\mathbb{X})$ is maximal in the sense that a simplex σ lies in $W_1(\mathbb{X})$ if and only if all its edges lie in $W_1(\mathbb{X})$. A problem of these Witness complexes is that they do not have much theoretical guarantees, in the sense that its homotopy type might differ much from the union $\bigcup_{x \in \mathbb{X}} B_r(x)$. When studying curves and surfaces, the Witness complexes can be used for approximating the homotopy type from sampled points on these objects [66]. Assuming some extra structure on Witness complexes, this approximation is also possible in high dimensions [13].

One could also encounter the problem that a dataset $\mathbb{X}$ is high dimensional. In this case one has to face the so called *curse of dimensionality* problem, which determines that it makes not much sense to use the distances between the different points in $\mathbb{X}$. In such a case, one might use other strategies to build simplicial complexes from the given dataset. One option introduced in [108] consists in using clustering algorithms for this purpose. The procedure works as follows; one starts with the point cloud $\mathbb{X} \subseteq \mathbb{R}^N$ together with a filtration function $f : \mathbb{R}^N \to \mathbb{R}$. Then a cover of $\mathbb{R}$ is taken which pulls back to a cover $\mathscr{U}$ for $\mathbb{R}^N$. Next, one chooses a clustering algorithm (DBSCAN, single-linkage, etc) and computes the clusters of each local pointcloud $\mathbb{X} \cap U$ for all covers $U \in \mathscr{U}$. These clusters determine the vertices of the *topological mapper* $\text{Mapper}(\mathbb{X}, \mathscr{U})$. Given two clusters, we connect them by an edge if they share at least a common point. See figure 3.3 for an illustration of this procedure. It is also possible to look at filtering functions $f : \mathbb{R}^N \to \mathbb{R}^M$ for values $M > 1$, in which case the object $\text{Mapper}(\mathbb{X}, \mathscr{U})$ is a simplicial complex [108]. Mapper has had an important role in exploring complex datasets, see for example [88]. The topological Mapper is not very stable with respect to small changes on the input data set $\mathbb{X}$, the filtration

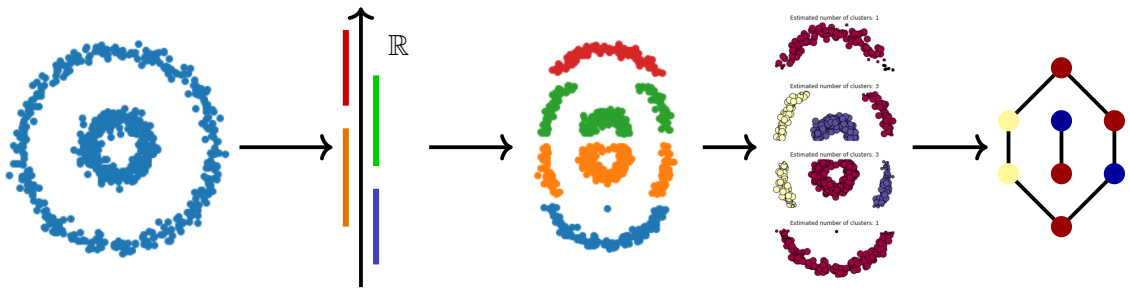


Figure 3.3: Illustration of the standard procedure for the construction of the Topological Mapper. On the left, the point cloud $\mathbb{X}$ which is projected into the real line. A cover of $\mathbb{R}$ is taken, and it is pulled back to a cover for $\mathbb{X}$. Then a clustering algorithm is used (in this case we used DBSCAN). Finally, by finding common points between clusters the Mapper complex is obtained.

function or the cover taken on $\mathbb{R}$. Fortunately, there are some positive results when comparing it to the *Reeb graph* which allows to measure stability of topological features, see [27] or also [17].

If the input data is a gray-scale image, one might consider cubical complexes [117]. Here, recall example 2.5.4 introducing these complexes. By a gray-scale image, we understand a real function on a lattice $f : \mathbb{Z}^N \to \mathbb{R}$; we can take corresponding cubical complexes $\mathscr{C}_r(\mathbb{Z}^N)$ for all real numbers $r \in \mathbb{R}$. This is done by including a vertex $v \in \mathscr{C}_r(\mathbb{Z}^N)$ whenever $f(v) \leq r$, while a cell a is contained in $\mathscr{C}_r(\mathbb{Z}^N)$ if all the vertices of a are contained in $\mathscr{C}_r(\mathbb{Z}^N)$. Alternatively, one could instead use the *lower star filtration* which leads to a simplicial complex [10].

Another possible scenario is when the input data comes as a graph G. In this case, one might take the *clique complex* [69] which is a simplicial complex Q_G containing a simplex σ if and only if all edges from σ are contained in G. Suppose that the edges in G have some *weight function* $w : E(G) \to \mathbb{R}$. In this case, the graph G has a natural filtration G_r for $r \in \mathbf{R}$, were an edge $e \in G$ is included in G_r if and only if $w(e) \leq r$. This filtration carries over to a filtration on the clique complex Q_{G_r} with $r \in \mathbf{R}$. If the edges in G are directed, then one might build the *flag complex* F_{G_r} for all $r \in \mathbf{R}$; this object is very relevant in neuroscience, see [97] and [78].

At the end of the day there is usually no 'best' complex to work with, instead their effectiveness depends on the particular problem that one is trying to solve.

3.2 The Persistence Viewpoint

In section 3.1 we briefly reviewed different methods for obtaining complexes from datasets. Notice that usually these complexes depend on at least one parameter, as for example the Čech complexes $\check{\mathrm{C}}_r(X)$ depend on the radius $r > 0$. In particular, the topology of the spaces $\check{\mathrm{C}}_r(X)$ changes dramatically as the parameter r is changed, see figure 3.1. On the other hand, there are inclusions $\check{\mathrm{C}}_r(X) \subseteq \check{\mathrm{C}}_s(X)$ for any pair of values $r < s$; in this case we say that $\check{\mathrm{C}}_*(X)$ is a *filtered complex*.

Within this context, instead of focusing on the topological features at a single filtration radius, we study how these *persist* along $\mathbf{R}$.

Definition 3.2.1. A *filtered complex* X, is a functor $X : \mathbf{R} \to \mathbf{CW\text{-}cpx}$ such that the morphisms $X(r \leq s) : X(r) \to X(s)$ are inclusions for all pairs $r \leq s$ from $\mathbf{R}$. We will usually write X_r instead of $X(r)$ for all $r \in \mathbf{R}$. We denote by **FCW-cpx** the category of filtered complexes with filtration preserving morphisms between them.

Given a filtered complex X and a value $\varepsilon > 0$, we denote by $X[\varepsilon]$ the ε-shifted complex, which is given by $X[\varepsilon](r) = X_{r+\varepsilon}$ for all $r \in \mathbf{R}$. We will denote by $\Sigma^{\varepsilon} : \mathbf{FCW\text{-}cpx} \to \mathrm{Hom}(\mathbf{FCW\text{-}cpx})$ the ε-shift functor, which sends a filtered complex X to the persistence morphism $\Sigma^{\varepsilon} X : X \to X[\varepsilon]$.

Definition 3.2.2. Let X be a filtered complex. We consider the composite functors $\mathrm{H}_n^{\mathrm{cell}}(X) : \mathbf{R} \to \mathbf{Vect}$ which we call the *persistent homology* of X, and which we denote by $\mathrm{PH}_n(X)$.

Definition 3.2.3. A *persistence module* $\mathbb{V}$ is a functor $\mathbb{V} : \mathbf{R} \to \mathbf{Vect}$. That is, to any $r \in \mathbf{R}$, $\mathbb{V}$ assigns a vector space in **Vect** which will be denoted either by $\mathbb{V}(r)$ or $\mathbb{V}^r$. That is, to any pair of real numbers $s \leq t$, there is a linear morphism $\mathbb{V}(s \leq t) : \mathbb{V}^s \to \mathbb{V}^t$. These morphisms satisfy $\mathbb{V}(s \leq s) = \mathrm{Id}_{\mathbb{V}^s}$ for any $s \in \mathbf{R}$, and the relation $\mathbb{V}(r \leq t) = \mathbb{V}(s \leq t) \circ \mathbb{V}(r \leq s)$ for all $r \leq s \leq t$ in $\mathbf{R}$. Given two persistence modules $\mathbb{V}$ and $\mathbb{W}$, a *morphism of persistence modules* is a natural transformation $f : \mathbb{V} \to \mathbb{W}$. Thus, for any pair of real numbers $s \leq t$, there is a commuting square

$$\begin{array}{ccc} \mathbb{V}^s & \xrightarrow{\mathbb{V}(s \leq t)} & \mathbb{V}^t \\ {\scriptstyle f^s}\downarrow & & \downarrow{\scriptstyle f^t} \\ \mathbb{W}^s & \xrightarrow{\mathbb{W}(s \leq t)} & \mathbb{W}^t. \end{array}$$

We denote by **PMod** the category of persistence modules and persistence morphisms.

We say that a persistence morphism $f : \mathbb{V} \to \mathbb{W}$ is an *isomorphism* whenever f_t is an isomorphism for all $t \in \mathbf{R}$. We write $\mathbb{V} \simeq \mathbb{W}$ to denote that $\mathbb{V}$ is isomorphic to $\mathbb{W}$. A point-wise finite dimensional (p.f.d.) persistence module is a functor $\mathbb{V} : \mathbf{R} \to \mathbf{vect}$, where **vect** is the category of finite-dimensional vector spaces.

Example 3.2.4. A special class of persistence modules will be the *interval modules*. For any pair of real numbers $s \leq t$, we denote by $\mathrm{I}(s,t)$ the interval module

$$\mathrm{I}(s,t)(r) = \begin{cases} \mathbb{F} & \text{for } r \in [s,t), \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

The morphisms $\mathrm{I}(s,t)(a \leq b)$ will be the identity for any two $a, b \in [s,t)$ and will be 0 otherwise.

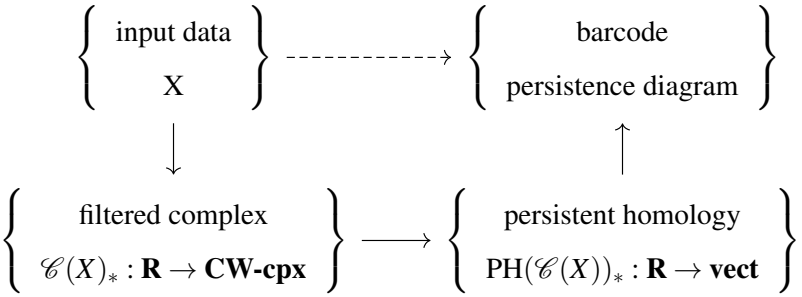
Notice that in an analogous way we could have defined modules $\mathrm{I}(s,t)$ over intervals of the form $[s,t]$, $(s,t]$ or (s,t), with $s \leq t$. For a given interval module $\mathrm{I}(s,t)$, the values s and t will be called respectively the *birth* and *death* values. Whenever $\mathbb{V}$ is a p.f.d persistence module, then it can be uniquely decomposed as a direct sum of barcodes $\bigoplus_{i\in J} \mathrm{I}(s_i,t_i)$, as shown in [42]. This means that there is an isomorphism $\mathbb{V} \simeq \bigoplus_{i\in J} \mathrm{I}(s_i,t_i)$ of persistence modules. This will be called the barcode decomposition of $\mathbb{V}$. Throughout this text, we will mainly be studying persistence modules that decompose into barcodes of the form (3.1). This leads to the following result, which by finiteness, follows also from the original decomposition stated in [122].

Theorem 3.2.5. *Suppose that X_* is a filtered CW-complex with a finite number of cells. Then*

$$\mathrm{PH}_m(X_*) \simeq \bigoplus_{i=0}^{N_m} \mathrm{I}(s_i,t_i)$$

for some $s_i < t_i \leq \infty$ for all indices $0 \leq i \leq N_m$ and for all dimensions $m \geq 0$.

Notice that a p.f.d. persistence module is uniquely determined by its interval decomposition, which is a multiset. This multiset is usually referred to as the *barcode* of the persistence module. Such a barcode can also be represented as a *persistence diagram*, where each bar $\mathrm{I}(a,b)$ is plotted as a point (a,b) in $\mathbb{R}^2$. The construction of persistent homology pipeline is as follows



In practice, computing persistent homology corresponds to computing a consecutive series of Gaussian eliminations, see [122] and also [124, §7] for a thorough explanation for the persistent homology algorithm. Let us review an example by considering a 'flower-shaped' filtered CW-complex X_* as depicted in figure 3.4. As the *filtration value* or *persistence value* increases, new cells are added into X_*, creating 'new' homology classes or deleting 'old' homology classes. This is usually expressed as classes being 'born' or 'dying'. If an added cell gives birth to a new class, such a cell is called *positive*, while if a cell kills a homology class it is called *negative*, see [122].

As X_* is a regular CW-complex, we can compute the incidences of adjacent cells as indicated

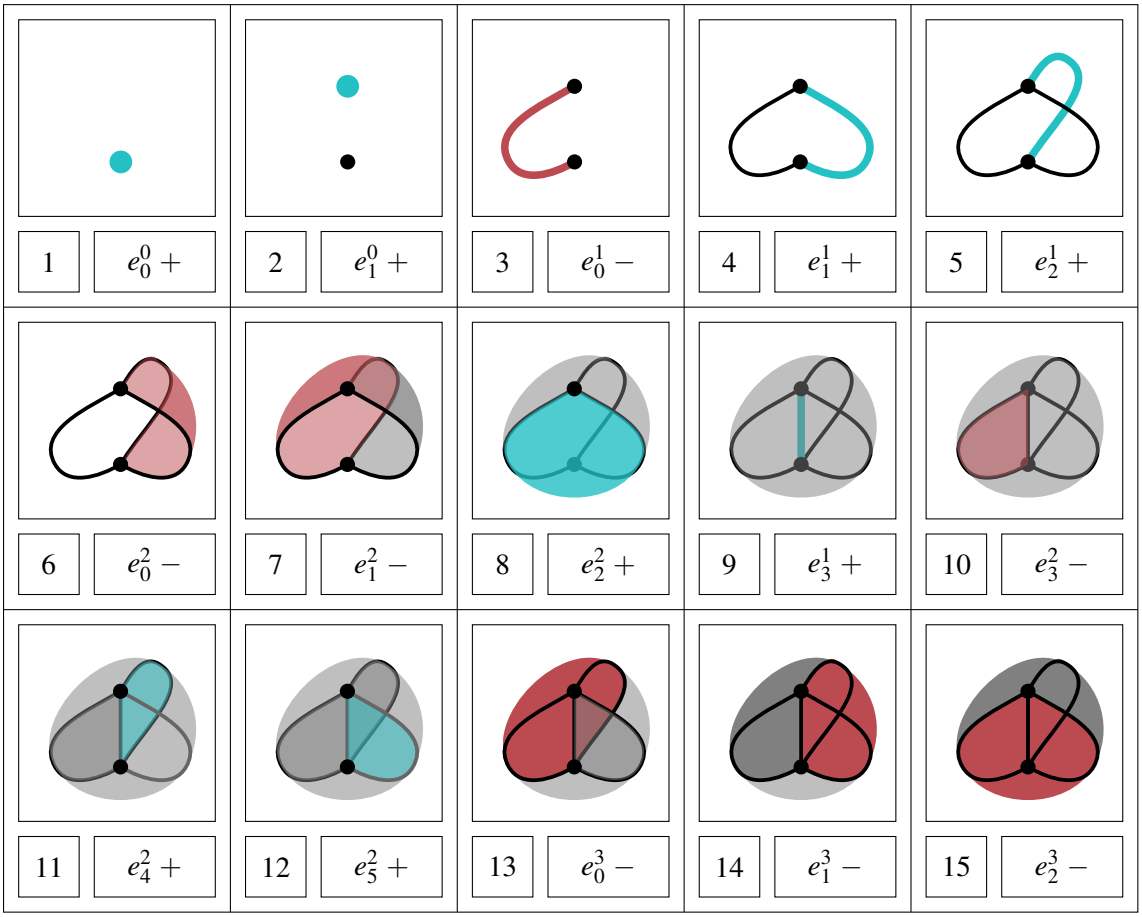


Figure 3.4: Depiction of a filtered CW-complex. New cells are colored in red or blue depending on whether these are negative or positive cells. Here we have used the notation e_i^j for cells, where the superindex j indicates the cell dimension, while the subindex i is used to identify the cell.

in section 2.6. In particular, we can obtain the following differentials (for some field $\mathbb{F}$)

$$D_1 = \left(\begin{array}{c|cccc} & e_0^1 & e_1^1 & e_2^1 & e_3^1 \\ \hline e_0^0 & 1 & 1 & 1 & 1 \\ e_1^0 & -1 & -1 & -1 & -1 \end{array}\right)$$

where we have labeled rows and columns by the corresponding cells. In addition, these rows and columns are ordered from lower filtration values to higher values from $\mathbf{R}$. For example, in D_1 notice that the columns are labeled by e_0^1, e_1^1, e_2^1 and e_3^1 whose respective filtration values are 3, 4, 5 and 9. Using the same notation, we write the differentials on the cellular chain complex in dimensions 2 and 3

$$D_2 = \left(\begin{array}{c|cccccc} & e_0^2 & e_1^2 & e_2^2 & e_3^2 & e_4^2 & e_5^2 \\ \hline e_0^1 & 0 & 1 & 1 & -1 & 0 & 0 \\ e_1^1 & 1 & 0 & -1 & 0 & 0 & -1 \\ e_2^1 & -1 & -1 & 0 & 0 & -1 & 0 \\ e_3^1 & 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right) \quad D_3 = \left(\begin{array}{c|ccc} & e_0^3 & e_1^3 & e_2^3 \\ \hline e_0^2 & 0 & 1 & 0 \\ e_1^2 & 1 & 0 & 0 \\ e_2^2 & 0 & 0 & -1 \\ e_3^2 & 1 & 0 & -1 \\ e_4^2 & -1 & -1 & 0 \\ e_5^2 & 0 & 1 & 1 \end{array}\right).$$

One can check that the entries on these differentials satisfy the incidence properties from section 2.6; an easy way to check this is to multiply the matrices associated to the compositions $d_1^{\text{cell}} \circ d_2^{\text{cell}}$ and $d_2^{\text{cell}} \circ d_3^{\text{cell}}$ and check that the resulting matrices vanish, as well as the fact that $[e_i^n : e_j^m] \neq 0$ if and only if $m = n-1$ and $e_j^m \preceq e_i^n$.

Next, we proceed to reduce the differential matrices by adding columns from left to right. This is done starting from D_3, keeping track of these additions in the upper labels

$$R_3 = \left(\begin{array}{c|ccc} & e_0^3 & e_1^3 & e_2^3 - e_1^3 + e_0^3 \\ \hline e_0^2 & 0 & 1 & -1 \\ e_1^2 & 1 & 0 & 1 \\ e_2^2 & 0 & 0 & -1 \\ e_3^2 & 1 & 0 & 0 \\ e_4^2 & -1 & -1 & 0 \\ e_5^2 & 0 & 1 & 0 \end{array}\right),$$

here we mark the pivot positions in yellow; as will be explained soon, these positions are very important when computing persistent homology. To facilitate computations, we use the *twist*

optimization from [32]. This is also called the *clear* optimization in [6], which consists in using pivots from R_3 to know combinations of columns in the matrix D_2 that vanish; in this case the third, fifth and sixth columns will vanish. The reason for this lies in the equation $d_2^{\text{cell}} \circ d_3^{\text{cell}} = 0$ which implies $D_2 D_3 = 0$, and consequently $D_2 R_3 = 0$. Each nontrivial column from R_3 gives "for free" a combination of column additions that will become zero.

$$R_2 = \left(\begin{array}{c|cccccc} & e_0^2 & e_1^2 - e_0^2 & -e_0^2 + e_1^2 - e_0^2 & e_3^2 & e_1^2 + e_3^2 - e_4^2 & e_0^2 - e_4^2 + e_5^2 \\ \hline e_0^1 & 0 & 1 & 0 & -1 & 0 & 0 \\ e_1^1 & 1 & -1 & 0 & 0 & 0 & 0 \\ e_2^1 & -1 & 0 & 0 & 0 & 0 & 0 \\ e_3^1 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

We can label the cleared rows by using the corresponding combinations from R_2. Using the same clear trick we get the following reduced matrix

$$R_1 = \left(\begin{array}{c|cccc} & e_0^1 & e_0^1 - e_1^1 & e_1^1 - e_2^1 & e_3^1 - e_0^1 \\ \hline e_0^0 & 1 & 0 & 0 & 0 \\ e_1^0 & -1 & 0 & 0 & 0 \end{array}\right).$$

Now, notice that the matrices of the differentials d_*^{cell} which we named D_1, D_2 and D_3 have associated filtration values for each column and row which depend on the respective "labelling" cells from X_*. The same values carry over to the matrices R_*; however, notice that now we have rows and columns which are "labelled" by a combination of cells which are born at different filtration values. To fix this, we always consider the higher filtration value. Thus, for the column labelled with $e_1^2 - e_0^2$ from R^2 we will associate the filtration value corresponding to e_1^2, which is 7; notice that the filtration value of e_0^2 is 6. We will formalize this idea later in section 4.1 through the use of *barcode bases* and the $\boxplus$ operation.

We can now read off the barcode decomposition of $\text{PH}_*(X)$ directly from the pivot positions in the matrices R_*, that is, the *birth* of a bar is equal to the filtration value of the row, while the *death* of the bar is equal to the filtration value of the column, see figure 3.5. One can check that pivot positions are independent of the particular order in which column additions are performed [58, §VII.1]. Notice that the barcode is a multiset of intervals, which we will denote by $\mathscr{B}_{\text{PH}_*(X)}$. We can plot these bars as points in the *persistence diagram* as well 3.6. This diagram is defined as the following subset of the extended plane of real numbers $\overline{\mathbb{R}}^2$:

$$\mathbf{Diag}(\text{PH}_*(X)) = \left\{(a_i, b_i) \mid [a_i, b_i) \in \mathscr{B}_{\text{PH}_*(X)}\right\} \cup \{(x,x) \mid x \in \mathbb{R}\},$$

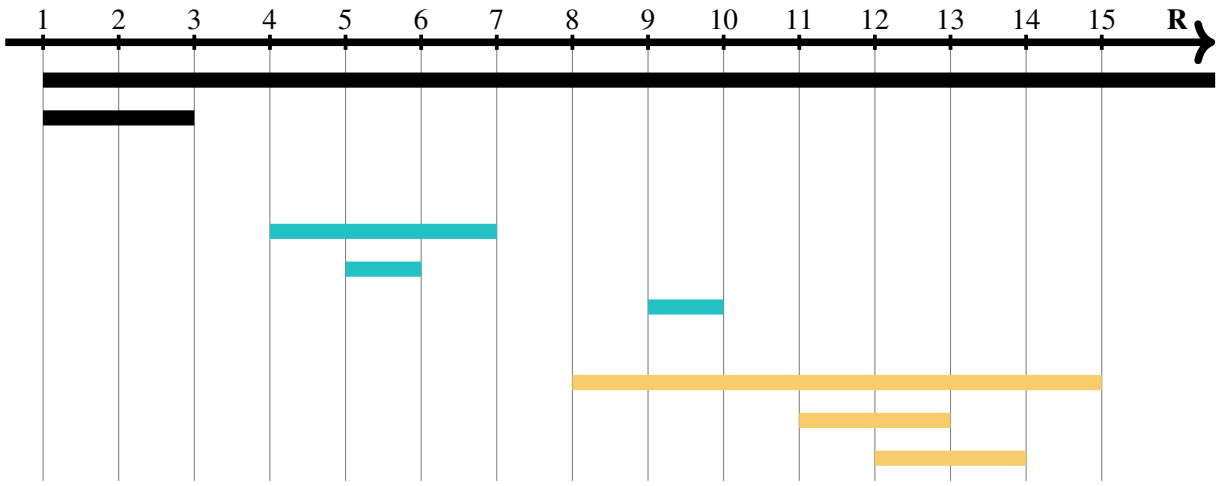


Figure 3.5: Persistence barcode decomposition of $\mathrm{PH}_*(X)$. Black, blue and yellow bars indicate the 0^{th}, 1^{st} and 2^{nd} homology dimensions respectively.

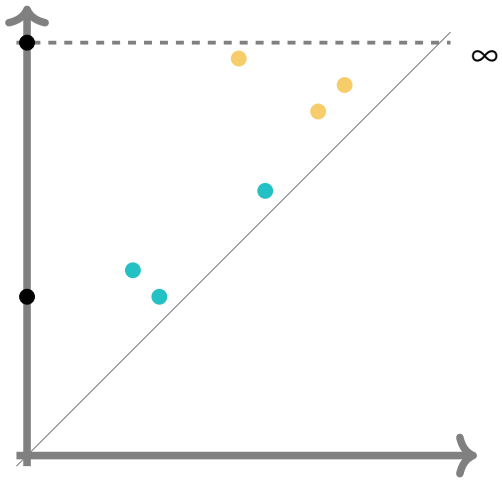


Figure 3.6: Persistence diagram. Notice that usually the y-axis is extended by the inclusion of an infinity line on top of the diagram.

where we have included the diagonal for technical reasons that will be clear after the definition of *bottleneck distance* on section 3.3. Usually, it is said that long bars correspond to relevant features of the dataset while short bars are regarded as noise; in the persistence diagram important features would correspond to points which are "far away" from the diagonal while the noise would lie close to the diagonal. However, this intuitive point of view is false in some situations; for example, the count of bars might be a relevant feature as in [96, §5.7].

We end this section by introducing a couple of notational conventions that we will adopt. Given a persistence module $\mathbb{V} \in \mathbf{PMod}$ and $\varepsilon \geq 0$, we define the ε-shifted persistence module $\mathbb{V}[\varepsilon] \in \mathbf{PMod}$ to be given by $\mathbb{V}[\varepsilon]_r = \mathbb{V}_{r+\varepsilon}$ for all $r \in \mathbf{R}$. Also, we will consider a functor $\Sigma^\varepsilon : \mathbf{PMod} \to \mathrm{Hom}(\mathbf{PMod})$ which sends $\mathbb{V} \in \mathbf{PMod}$ to $\Sigma^\varepsilon \mathbb{V} : \mathbb{V} \to \mathbb{V}[\varepsilon]$. Using this notation, a

persistence morphism $f : \mathbb{V} \to \mathbb{W}$ is such that one has the equality $f[\varepsilon] \circ \Sigma^{\varepsilon}\mathbb{V} = \Sigma^{\varepsilon}\mathbb{W} \circ f$, where we use the corresponding notation for shifted persistence morphisms $f[\varepsilon] : \mathbb{V}[\varepsilon] \to \mathbb{W}[\varepsilon]$.

3.3 Persistence Stability

The main reason why persistence is widely used is its stability with respect to perturbations or noise in the input data. In this section, we will review some of these results. To start, notice that we need to define a way of measuring both differences in the input as well as the output. Suppose that one starts with a pair of filtered regular CW-complexes X_* and Y_* with the same underlying CW-complex structure $\mathbf{colim}\, X_* = \mathbf{colim}\, Y_*$. An example could come from a couple of cubical complexes generated on two images with the same grid of pixels but with slightly different grayscale values. Further, given $\varepsilon > 0$, assume that X_* and Y_* are ε-close in the following way: there are inclusions $X_r \subseteq Y_{r+\varepsilon}$ and also $Y_r \subseteq X_{r+\varepsilon}$ for all values $r \in \mathbf{R}$. In this case, notice that there is a commutative diagram induced by inclusion maps

$$\begin{array}{ccccc} \mathrm{H}_k(X_r) & \longrightarrow & \mathrm{H}_k(X_r[\varepsilon]) & \longrightarrow & \mathrm{H}_k(X_r[2\varepsilon]) \\ & \searrow\nearrow & & \searrow\nearrow & \\ \mathrm{H}_k(Y_r) & \longrightarrow & \mathrm{H}_k(Y_r[\varepsilon]) & \longrightarrow & \mathrm{H}_k(Y_r[2\varepsilon]) \,. \end{array}$$

In this case, both $\mathrm{PH}_k(X_*)$ and $\mathrm{PH}_k(Y_*)$ are "ε-close to being isomorphic" persistence modules. This leads to the following definition.

Definition 3.3.1. Let $\mathbb{V}$ and $\mathbb{W}$ be two persistence modules. We say that $\mathbb{V}$ and $\mathbb{W}$ are ε-*interleaved* whenever there are persistence morphisms $\phi : \mathbb{V} \to \mathbb{W}[\varepsilon]$ and $\psi : \mathbb{W} \to \mathbb{V}[\varepsilon]$ such that the following diagram commutes

$$\begin{array}{ccccc} \mathbb{V} & \xrightarrow{\Sigma^{\varepsilon}\mathbb{V}} & \mathbb{V}[\varepsilon] & \xrightarrow{\Sigma^{\varepsilon}\mathbb{V}} & \mathbb{V}[2\varepsilon] \\ & \searrow^{\phi}\nearrow_{\psi} & & \searrow^{\phi[\varepsilon]}\nearrow_{\psi[\varepsilon]} & \\ \mathbb{W} & \xrightarrow[\Sigma^{\varepsilon}\mathbb{W}]{} & \mathbb{W}[\varepsilon] & \xrightarrow[\Sigma^{\varepsilon}\mathbb{W}]{} & \mathbb{W}[2\varepsilon] \,. \end{array}$$

Equivalently, the following equalities hold

$$\Sigma^{2\varepsilon}\mathbb{V} = \psi[\varepsilon] \circ \phi \qquad \text{and} \qquad \Sigma^{2\varepsilon}\mathbb{W} = \phi[\varepsilon] \circ \psi \,,$$

for all $r \in \mathbf{R}$.

This notion of interleaving leads to a natural way of comparing persistence modules. Given a couple of persistence modules $\mathbb{V}$ and $\mathbb{W}$, we define their *interleaving distance* as the following

infimum

$$d_I(\mathbb{V},\mathbb{W}) = \inf\{\varepsilon > 0 \mid \mathbb{V} \text{ and } \mathbb{W} \text{ are } \varepsilon\text{-interleaved}\} \ .$$

On the other hand, let us consider again the two regular CW-complexes X_* and Y_* with the same colimit $Z = \mathbf{colim}\, X_* = \mathbf{colim}\, Y_*$ and the same cell decomposition. Given a cell $e^n \in X$, we define its filtration degree to be

$$\mathbf{fdeg}(e^n, X) = \inf\{r \in \mathbf{R} \mid e^n \in X_r\} \ .$$

Then, we define the filtration cellular distance as the supremum

$$d_{\text{cell}}(X_*, Y_*) = \sup\{|\mathbf{fdeg}(e^n, X) - \mathbf{fdeg}(e^n, Y)| \mid \forall e^n \in Z\} \ .$$

Of course, if $\delta = d_{\text{cell}}(X_*, Y_*)$, then we have inclusions

$$\phi_r : X_r \hookrightarrow Y_r[\delta] \qquad \text{and} \qquad \psi_r : Y_r \hookrightarrow X_r[\delta]$$

for all values $r \in \mathbf{R}$. Thus, there are equalities $\Sigma^{2\delta} X_* = \psi_{r+\delta} \circ \phi_r$ and $\Sigma^{2\delta} Y_* = \phi_{r+\delta} \circ \psi_r$ for all values $r \in \mathbf{R}$, which induce an interleaving in persistent homology

$$d_I(\mathrm{PH}_*(X_*), \mathrm{PH}_*(Y_*)) \leq d_{\text{cell}}(X_*, Y_*) \ .$$

This result ensures stability of persistent homology on filtered regular CW-complexes as we change the filtration values but keep the underlying CW-structure.

A very strong property of persistent homology is that the stability mentioned above can be adapted to a situation when small perturbations lead to different CW-structures. This is the case for filtered simplicial complexes on point clouds $\mathbb{X} \subseteq \mathbb{R}^N$. Given a pair of finite subsets $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{R}^N$, we define the *Hausdorff* distance as the infimum

$$d_H(\mathbb{X}, \mathbb{Y}) = \inf\{r > 0 \mid \mathbb{X} \subseteq \mathbb{Y}^r \text{ and } \mathbb{Y} \subseteq \mathbb{X}^r\}$$

where $\mathbb{X}^r$ denotes the r offset of $\mathbb{X}$, as explained in section 3.1. If $\varepsilon = d_H(\mathbb{X}, \mathbb{Y})$, the triangle inequality implies the alternating inclusions $\mathbb{X}^r \subseteq \mathbb{Y}^{r+\varepsilon}$ and $\mathbb{Y}^r \subseteq \mathbb{X}^{r+\varepsilon}$ for all values $r \in \mathbf{R}$. As we have homotopy equivalences $\check{C}_r(\mathbb{X}) \simeq \mathbb{X}^r$ for all values $r \in \mathbf{R}$, we obtain the inequality

$$d_I(\mathrm{PH}_*(\check{C}_*(\mathbb{X})), \mathrm{PH}_*(\check{C}_*(\mathbb{Y}))) \leq d_H(\mathbb{X}, \mathbb{Y}) \ .$$

One might get the analogous result for Vietoris-Rips complexes

$$d_I(\mathrm{PH}_*(\mathrm{VR}_*(\mathbb{X})),\mathrm{PH}_*(\mathrm{VR}_*(\mathbb{Y}))) \leq 2d_H(\mathbb{X},\mathbb{Y}) \ ,$$

see [92, Prop. 7.8, Sec. 7.3]. We will show later in section 8.1 an alternative way to prove this stability by using ε-*acyclic equivalences*. Using these equivalences one might find stability results for filtered regular CW-complex, whenever there is some rule relating 'close' input data sets.

So far we have seen persistence stability. However, often the aim of persistence is to be used as a measure for comparing pairs of different inputs. Notice that given two persistence modules, it might be hard to find out how close these are by means of interleavings. Here we use the combinatorial nature of barcodes and persistence diagrams to define a metric on these. In particular, assume that we have a pair of p.f.d. persistence modules $\mathbb{V}$ and $\mathbb{W}$. These admit barcode decompositions, and in particular persistence diagrams. Thus, we can look at bijections between the points on the persistence diagrams $\gamma\colon \mathbf{Diag}(\mathbb{V}) \to \mathbf{Diag}(\mathbb{W})$. We define the *bottleneck distance* between the diagrams as the following infimum over bijections of persistence diagrams

$$d_B(\mathbf{Diag}(\mathbb{W}),\mathbf{Diag}(\mathbb{V})) = \inf_{\Gamma:\mathbf{Diag}(\mathbb{V})\to\mathbf{Diag}(\mathbb{W})} \left\{ \sup_{b\in\mathscr{B}_{\mathbb{V}}} \|\Gamma(b)-b\|_\infty \right\}$$

This definition justifies the inclusion of the diagonals into the persistence diagrams, as we can pair off-diagonal points from $\mathbf{Diag}(\mathbb{V})$ to points in the diagonal if there are no "close" off-diagonal points from $\mathbf{Diag}(\mathbb{W})$. The *isometry theorem* [8] implies that for a pair of p.f.d. persistence modules $\mathbb{V}$ and $\mathbb{W}$, the interleaving and bottleneck distances are equal

$$d_B(\mathbf{Diag}(\mathbb{V}),\mathbf{Diag}(\mathbb{W})) = d_I(\mathbb{V},\mathbb{W}) \ ,$$

and thus we can use the bottleneck distance to deduce interleavings for two different persistence modules. The space of persistence diagrams is a metric space and it is in fact convex, where a line between two different persistence diagrams might be traced by constructing interpolating diagrams [30, §3.4.].

Very often stability of persistence modules is stated in terms of sub-level sets of continuous functions [37]. It is worth noticing that the bottleneck distance is very sensible to outliers; so that other distances are computed instead, such as the p-Wasserstein distance [109]. Persistent homology has had a number of applications such as in detecting chronic obstructive pulmonary disease [9], analysis of signals [100] through the use of the sliding window embedding [95], it has also been used in pattern recognition [22, 99], shape comparison [55] and many other applications. Also, it is worth mentioning that the interpretation of persistence diagrams is a big

subject. One might interpret them as *persistence landscapes* [18, 19], *persistence images* [1], kernels for machine learning [98], and many more.

There are variations for standard persistent homology, such as *Zig-Zag persistence* [24, 23, 111], where the focus is on persistence of homology classes over sequences of vector spaces and linear morphisms that are distributed in a *zig-zag* fashion, such as the following:

$$A_0 \longrightarrow A_1 \longleftarrow A_2 \longrightarrow \cdots \longleftarrow A_n \longrightarrow \cdots .$$

There are also generalizations of persistence diagrams [102, 94], together with stability measures formulated in terms of interleavings. In recent years, stronger invariants than persistent homology have been explored. One of these is *multipersistent* homology [26, 1], where the persistence of homology classes is studied over general posets, rather than the concrete poset $\mathbf{R}$. One might use interleavings as a way to measure stabilities in this context [74], although computing interleaving distances directly is an NP-hard problem [11]. In fact, interleavings apply in a general context of *categories with a flow* [48], an example of which is the category of persistent spectral sequences that we will see later 8.

3.4 Computing Persistent Homology

We will proceed to briefly review some computational techniques for computing persistent homology. There are a number of books and surveys on this topic such as [122, 92, 71, 91]. In [91] one might find references to several software packages, together with a comparison of their capabilities and benchmark estimates.

The original setup [59, 122] starts by constructing a filtered simplicial complex, compute its boundary matrices and perform a series of Gaussian eliminations to obtain the barcodes. For example, one might start from a point cloud $\mathbb{X} \subseteq \mathbb{R}^N$, compute its filtered Vietoris-Rips complex $\mathrm{VR}_r(\mathbb{X})$, obtain the matrices of the simplicial complex differentials and look for persistence pairs to obtain the barcode decomposition of $\mathrm{PH}_n(\mathrm{VR}_r(\mathbb{X}))$. There are a number of steps towards making such computations efficient. For example, when computing $\mathrm{VR}_r(\mathbb{X})$ there are computational strategies [121] as well as efficient data structures for storing the simplicial information [14]. For the particular case of persistent homology of Vietoris-Rips complexes, there is a highly efficient C++ implementation called *Ripser* [4]; part of the success of this implementation relies on using the clear optimization [6] together with computing persistent cohomology [47] instead of persistent homology, as well as avoiding the explicit construction of the coboundary matrix. There are efficient libraries implementing differential and matrix manipulations, such as PHAT [7] or GUDHI [113]. These libraries implement persistence functions from various complexes, such as

Alpha, Witness, etc. A user-friendly python package called scikit-tda can be found in [87].

The traditional setup for computing persistent homology has a fundamental computational difficulty. This comes from the fact that, given a filtered complex

$$F^0X \subseteq F^1X \subseteq F^2X \subseteq \cdots \subseteq F^NX \subseteq \cdots ,$$

the higher filtration levels F^NX for N being "large enough", the combinatorial data in F^NX is far too large, while the underlying topological structure might not necessarily increase at the same rate. This is the case with the Vietoris Rips complex $\mathrm{VR}_*(\mathbb{X})$, which eventually becomes a high-dimensional standard simplex on the set $\mathbb{X}$. A solution for this problem was studied in [106, 29], where the authors introduced a *sparse rips filtration* $\mathscr{S}(\mathrm{VR}(\mathbb{X}))_*$ which is a filtered simplicial complex satisfying the inclusions

$$\mathscr{S}(\mathrm{VR}(\mathbb{X}))_r \subseteq \mathrm{VR}_r(\mathbb{X}) \subseteq \mathscr{S}(\mathrm{VR}(\mathbb{X}))_{r(1+\varepsilon)}$$

for all $r \geq 0$. In this case, one says that there is a *multiplicative* $(1+\varepsilon)$*-interleaving* between the persistent homologies of $\mathscr{S}(\mathrm{VR}(\mathbb{X}))_*$ and $\mathrm{VR}_*(\mathbb{X})$, see [106]. This approach has lead to sequences of simplicial complexes and simplicial morphisms

$$X^0 \to X^1 \to X^2 \to \cdots \to X^N \to \cdots ,$$

whose persistent homology is equivalent or approximately close to a filtered complex that might contain too much information, see [52, 72, 53]. In [52] a barcode decomposition for $\mathrm{PH}_*(X^*)$ is obtained by means of tracking a *persistent cohomology* basis through *annotations*. Notice that the barcode decomposition of persistent cohomology and persistent homology coincide [47], thanks to the *Universal Coefficient Theorem*. Currently, the only software package implementing persistent homology on sequences of simplicial maps is SimPers [50]. Combining this approach together with *elementary simplicial collapses*, an efficient implementation for computing persistent homology of Vietoris-Rips filtrations was obtained in [53]. In section 4.1 we will introduce *barcode bases*, which allow to compute persistent homology of sequences of regular CW-complexes with regular morphisms, which we will call *regularly filtered CW-complexes*.

Another possible way to go consists in using *Discrete Morse Theory* [63]. This approach has the added advantage that it works on the more general framework of filtered CW-complexes [3, 86, 43]. Discrete Morse Theory lies at the heart of several highly efficient algorithms in topological data analysis, such as in Ripser [4]. This allows to perform substantial reductions on the underlying filtered complex, so that the subsequent computations are less burdensome. In the

similar direction, one could opt to reduce a filtered cellular complex. This is the approach taken in *effective homology*, which has the advantage that it might handle computations on complexes that are not necessarily finitely generated [104].

Yet another promising direction is to make small changes in the input dataset so that the following computations are less heavy. An example of this approach is taken by finding choices of landmark points which take into account local computations of persistent homology [110]. In this particular case, the focus is on classifying the impact of removing vertices on a filtered simplicial complex, which is measured by means of the Mayer-Vietoris exact sequence. In order to explain this, we introduce *stars* and *links*.

Definition 3.4.1. Let K be a simplicial complex and let $\sigma \in K$ be a simplex of K. We define $\mathrm{St}(\sigma)$, the *open star* of σ, to be the set given by the union of simplices $\tau \in K$ such that $\sigma \subseteq \tau$. One can see that $\mathrm{St}(\sigma)$ is open in K since its complement $K \setminus \mathrm{St}(\sigma)$ is a subcomplex of K. On the other hand we define $\overline{\mathrm{St}(\sigma)}$, the *closed star* of σ, to be the closure of $\mathrm{St}(\sigma)$. Finally, we define the *link* of σ as $\mathrm{Lk}(\sigma) = \overline{\mathrm{St}(\sigma)} \setminus \mathrm{St}(\sigma)$.

Now, consider a simplicial complex K and let v be a vertex from K. Then, we might consider the cover of K given by two subcomplexes $\mathscr{U} = \left\{K \setminus \mathrm{St}(v), \overline{\mathrm{St}(v)}\right\}$ noticing that the intersection corresponds to the link: $\mathrm{Lk}(v) = \left(K \setminus \mathrm{St}(v)\right) \cap \overline{\mathrm{St}(v)}$. Applying the Mayer-Vietoris Theorem, we obtain the exact sequence

$$\cdots \longrightarrow \mathrm{H}_n(\mathrm{Lk}(v)) \longrightarrow \mathrm{H}_n(K \setminus \mathrm{St}(v)) \oplus \mathrm{H}_n(\overline{\mathrm{St}(v)}) \longrightarrow \mathrm{H}_n(K) \longrightarrow \mathrm{H}_{n-1}(\mathrm{Lk}(v)) \longrightarrow \cdots$$

The closed star $\overline{\mathrm{St}(v)}$ is contractible, and if we assume that $\mathrm{H}_n(\mathrm{Lk}(v)) = \mathrm{H}_{n-1}(\mathrm{Lk}(v)) = 0$, then, there is an exact sequence

$$0 \to \mathrm{H}_n(K \setminus \mathrm{St}(v)) \to \mathrm{H}_n(K) \to 0 \,,$$

so that $\mathrm{H}_n(K \setminus \mathrm{St}(v)) \simeq \mathrm{H}_n(K)$ for all $n \geq 0$. We want to apply this to persistence: if K is a filtered complex, then the idea is that if both $\mathrm{PH}_n(\mathrm{Lk}(v))$ and $\mathrm{PH}_{n-1}(\mathrm{Lk}(v))$ are small enough, then $\mathrm{PH}_n(K \setminus \mathrm{St}(v)) \approx \mathrm{PH}_n(K)$. This idea is applied to classify points on a given point cloud $\mathbb{X}$ as having large or small *outlier values* [110]. Notice that for a point cloud $\mathbb{X}$, one considers a radius $r > 0$ and pays attention to the bounded link $\mathrm{Lk}^r(p)$ which is the link of p in the bounded Vietoris-Rips complex $\mathrm{VR}_s^r(\mathbb{X}) = \mathrm{VR}_s(\mathbb{X})$ for all $s \leq r$. In section 5.1 we will present a related result which we call the *Barcode Shift Lemma*.

The last computational approach that we would like to mention is related to *localized homology* [123]. This approach relies on a covered space $(X, \mathscr{U})$, together with the fact that the Mayer-Vietoris blowup complex $\Delta_{\mathscr{U}} X$ is homotopy equivalent to X. The filtration on $\Delta_{\mathscr{U}} X$ is

used to determine the locality of a homology class in $\mathrm{H}_*(X)$ with respect to the cover $\mathscr{U}$ [123]. On the computational side, in [75] the blowup complex $\Delta_{\mathscr{U}}X$ was used for parallelizing persistent homology computations, in the case when X is a filtered complex. The idea consists in first computing local persistent homologies on each cover element from $\mathscr{U}$. Then, using this local information, the authors proceeded to reduce the matrix of the differential of $\Delta_{\mathscr{U}}X$ by paying attention to the important rows and columns, discarding non-relevant information. We will present in chapter 6 a method that also computes the persistent homology of $\Delta_{\mathscr{U}}X$, although through the use of spectral sequences.

3.5 Persistence Filtration Spectral Sequences

Persistent homology associates to a filtration of topological spaces a filtration of vector spaces.. However, spectral sequences have been used for a long time on filtered topological spaces. This is why a number of authors have studied the intricate connections between both situations. The setup starts with a finite filtered regular CW-complex

$$F^0X \subseteq F^1X \subseteq \cdots \subseteq F^nX \ , \tag{3.2}$$

whose differential respects filtration in the sense that $d(F^pX) \subseteq F^pX$ for all $p \geq 0$. As explained in section 2.12, one obtains a twisted double complex together with a spectral sequence associated to it. In the applied topology literature, such spectral sequences have been handled by the use of *exact couples* [2] and [104]. Although, we prefer the point of view from [83, Thm. 2.6.] which is in line with the Mayer-Vietoris spectral sequence exposition from section 2.11. Connections between spectral sequences and persistent homology have been thoroughly studied in the literature [58, 2, 102, 60]. We now state an important connection [102, Thm. 4.4.]

Theorem 3.5.1. *Denote by* $\mathrm{BD}_n^{i,k}$ *the following quotient*

$$\frac{\ker\big(\mathrm{H}_n(F^iX) \to \mathrm{H}_n(F^kX)\big) \cap \mathrm{Coker}\big(\mathrm{H}_n(F^{i-1}X) \to \mathrm{H}_n(F^iX)\big)}{\ker\big(\mathrm{H}_n(F^iX) \to \mathrm{H}_n(F^{k-1}X)\big) \cap \mathrm{Coker}\big(\mathrm{H}_n(F^{i-1}X) \to \mathrm{H}_n(F^iX)\big)}$$

which corresponds to classes in $\mathrm{PH}_n(F^*X)$ *being born at value i and dying at value k. There is an isomorphism:*

$$\mathrm{BD}_n^{i,k} \simeq \mathrm{Im}\big(d_{k,n-k+1}^{k-i}\big)$$

Proof. This follows by using the kernel and cokernel formulas from the end of section 2.12, as

we obtain that

$$\mathrm{Coker}\big(\mathrm{H}_n(F^{i-1}X) \to \mathrm{H}_n(F^iX)\big) \simeq E^{\infty}_{i,n-i}(X,F^i) \simeq \frac{Z^{\infty}_{i,n-i}}{Z^{\infty}_{i-1,n-i+1} + B^0_{i,n-i}}$$

so that

$$\ker\big(\mathrm{H}_n(F^iX) \to \mathrm{H}_n(F^kX)\big) \cap \mathrm{Coker}\big(\mathrm{H}_n(F^{i-1}X) \to \mathrm{H}_n(F^iX)\big)$$

is isomorphic to

$$\ker\left(\frac{Z^{\infty}_{i,n-i}}{Z^{\infty}_{i-1,n-i+1} + B^0_{i,n-i}} \longrightarrow \frac{Z^{\infty}_{i,n-i}}{Z^{\infty}_{i-1,n-i+1} + B^{k-i}_{i,n-i}}\right) \simeq \frac{B^{k-i}_{i,n-i}}{Z^{\infty}_{i-1,n-i+1} + B^0_{i,n-i}} \,.$$

Putting everything together we have the isomorphisms

$$\mathrm{BD}^{i,k}_n \simeq \frac{B^{k-i}_{i,n-i+1}}{Z^{\infty}_{i-1,n-i+1} + B^0_{i,n-i}} \Big/ \frac{B^{k-1-i}_{i,n-i+1}}{Z^{\infty}_{i-1,n-i+1} + B^0_{i,n-i}} \simeq \frac{B^{k-i}_{i,n-i+1}}{Z^{\infty}_{i-1,n-i+1} + B^{k-1-i}_{i,n-i}} \simeq \mathrm{Im}(d^{k-i}_{k,n-k+1}) \,.$$

□

As explained in [102], the notation $\mathrm{BD}^{i,j}_n$ stands for *birth* and *death*. The proof of Theorem 3.5.1 by using the k-truncated filtration as done here is our contribution and differs from the original exposition of this fact [102].

Example 3.5.2. Let us illustrate theorem 3.5.1 by computing the spectral sequence from the example filtered cone from section 2.12 with coefficients in some field $\mathbb{F}$. See figure 3.7 to view the consecutive pages of the spectral sequence $E^*_{p,q}(C,F)$, where it might be useful to see again figure 2.14 to understand how the 0-th page is generated. In particular, using theorem 3.5.1 we pay attention to nontrivial differentials to read off barcodes from there; notice that the differentials on the 0-th column correspond to non-persistent cycles, and thus we ignore them. We obtain the following isomorphisms:

$$\mathrm{BD}^{5,6}_2 \simeq \mathrm{Im}(d^1_{6,-3}) \simeq \mathbb{F} \,, \qquad \mathrm{BD}^{3,5}_1 \simeq \mathrm{Im}(d^2_{5,-3}) \simeq \mathbb{F} \,, \qquad \mathrm{BD}^{1,3}_0 \simeq \mathrm{Im}(d^2_{3,-2}) \simeq \mathbb{F} \,,$$

which lead to the barcode decomposition of $\mathrm{PH}_*(FC)$, the infinite bar in dimension 0 is recovered from reading the copy of $\mathbb{F}$ from the page $E^3_{p,q}(C,F)$.

In [2], the authors found some other formulas connecting both persistent homology and spectral sequences, such as

$$\dim(E^r_{p,q}(X,F)) = (b^{p,p+r-1}_n - b^{p-1,p+r-1}_n) + (b^{p-r,p-1}_n - b^{p-r,p}_n)$$

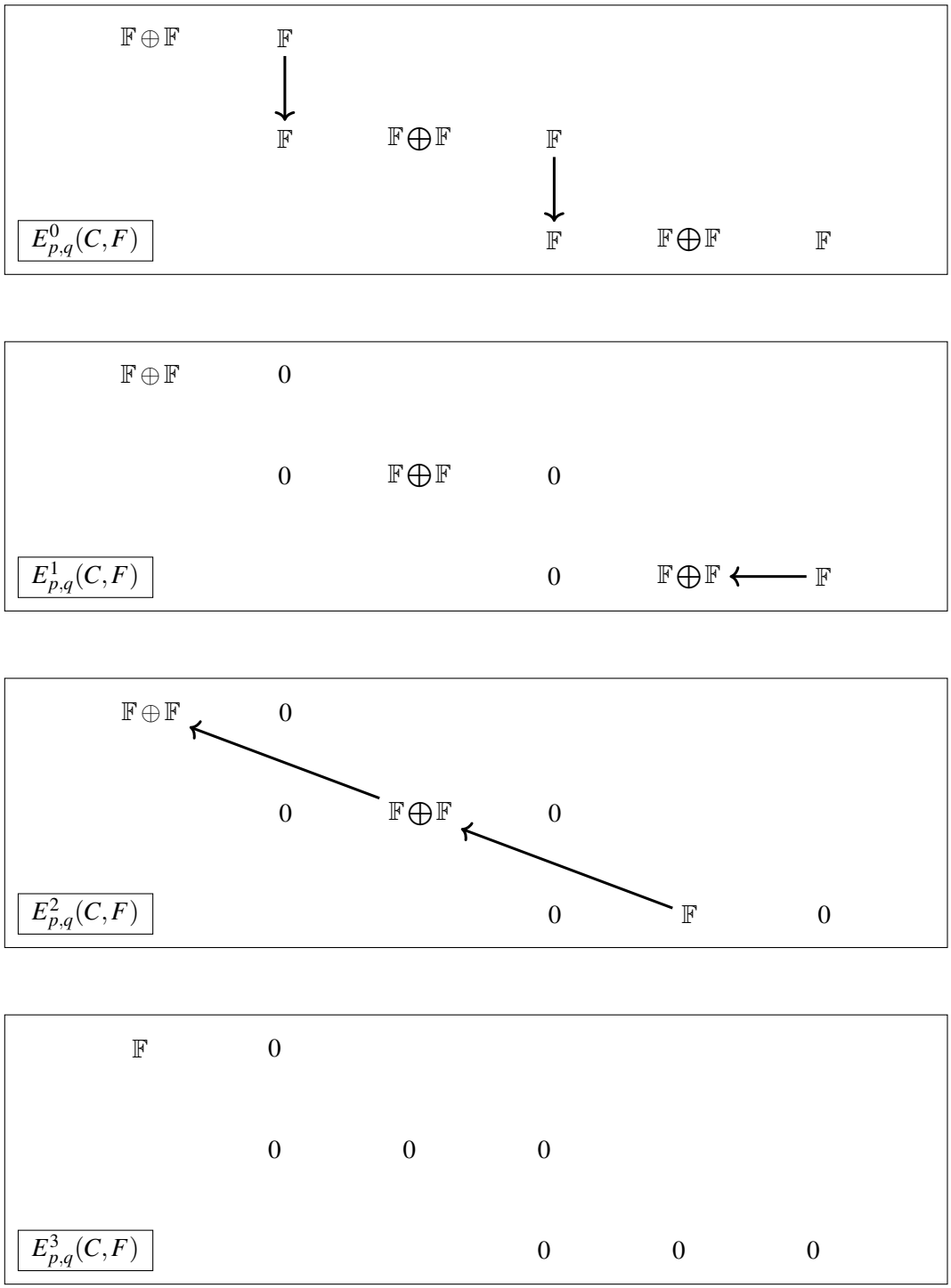


Figure 3.7: Pages of the spectral sequence associated to the filtered cone F^*C.

where we define the *persistence betti numbers*

$$b_n^{i,j} = \dim\big(\mathrm{Im}\big(\mathrm{H}_n(F^iX) \longrightarrow \mathrm{H}_n(F^jX)\big)\big) \ ,$$

for all $n, i, j \in \mathbb{Z}$. The multiple connections between spectral sequences and persistent homology have been explored computationally as well. In particular, there is a Kenzo package [103] for computing spectral sequences associated to a filtered CW-complex. Notice that in [103] the authors were able to handle computations in $\mathbb{Z}$ coefficients where the extension problem associated to the corresponding spectral sequence is nontrivial. A recent promising direction [77] consists in studying the same idea for *multidimensional persistent homology* [26]. This approach explores recent advances in *spectral systems*, which are spectral sequences associated to chain complexes filtered over a poset [81]. Notice that this approach has also an associated module implemented in Kenzo [67].

3.6 The Spectral Sequence Method

One of most widely known applications of spectral sequences in applied topology is the *spectral sequece method* introduced by Edelsbrunner and Harer [57, 58]. This approach was successfully applied to parallel computation of persistent homology in [5]. This point of view consists in computing the successive pages from $E^*_{*,*}(X,F)$ as a block reduction of a block matrix containing the boundaries of all cells. We will briefly review this approach, drawing connections between both the block matrix reductions and the twisted double complex spectral sequence from section 2.12 and section 3.5. In doing so, we would like to point out that both in [58] and [5], it is assumed that the underlying boundary matrix has been constructed. Each column in the boundary matrix leads directly to a cycle in the twisted total complex. This, however, differs from the spectral sequence "spirit", where cycles are found sequentially. Our exposition of this algorithm will be closer to the usual computation of spectral sequences. First, consider a filtered complex X_*, and take a sequence of real numbers

$$r_0 < r_1 < \cdots < r_j < \cdots < r_n$$

so that a filtration F_* for X is obtained by setting $F_iX = X_{r_i}$ for all $0 \leq i \leq n$, and where we assume that $F_nX = X$. Thus, we obtain the spectral sequence associated to this "block" filtration

$$E^0_{p,q}(X,F) = C^{\mathrm{cell}}_{p+q}\big(F_pX, F_{p-1}X\big) \Rightarrow \mathrm{H}^{\mathrm{cell}}_{p+q}(X) \ .$$

We proceed to compute the first page of the spectral sequence by obtaining the block matrices corresponding to the zero page differentials

$$d^0_{p,q} : C^{\text{cell}}_{p+q}\big(F_pX, F_{p-1}X\big) \longrightarrow C^{\text{cell}}_{p+q-1}\big(F_pX, F_{p-1}X\big) \ ,$$

which we denote by $M^{p+q}_{p,p}$, where the superscript denotes the differential dimension, while the subscripts lead to a position on a $\mathbb{Z}\times\mathbb{Z}$-grid. Next, we reduce the matrices $M^{p+q}_{p,p}$ from higher to lower values of q, denoting the resulting matrices by $\mathscr{R}(M^{p+q}_{p,p})$ and using the clear optimization. Here we pay attention to the zero columns from the reduced matrix $\mathscr{R}(M^{p+q}_{p,p})$ and such that these are not cleared by a pivot from $\mathscr{R}(M^{p+q+1}_{p,p})$. The corresponding preimage chains from $C^{\text{cell}}_{p+q}(F_pX, F_{p-1}X)$ represent generators of a basis for $E^1_{p,q}(X,F)$.

Next, we describe the procedure inductively on $r \geq 1$. Assume that some generating chains are given in $C^{\text{cell}}_{p+q}(F_pX, F_{p-r}X)$ which represent classes on the page $E^r_{p,q}(X,F)$. For ease, we will write $E^r_{p,q}$ instead of $E^r_{p,q}(X,F)$. First we compute a matrix $\widetilde{M}^{p+q}_{p,p+r}$ with $\dim\big(E^0_{p-r,q+r-1}\big)$ rows and $\dim\big(E^r_{p,q}\big)$ columns; whose columns contain representative coordinates in $E^0_{p-r,q+r-1}$ for the image of the r-differential $d^r_{p,q} : E^r_{p,q} \longrightarrow E^r_{p-r,q+r-1}$. Then we ignore all the rows from $\widetilde{M}^{p+q}_{p,p-r}$ which correspond to nonzero columns from the matrices $M^{p+q-1}_{p-r,p-r+s}$ for all $0 \leq s \leq r-1$. We end up with a block matrix $M^{p+q}_{p,p-r}$ which has $\dim\big(Z^r_{p-r,q+r-1}\big)$ rows and $\dim\big(E^r_{p,q}\big)$ columns.

Then, we reduce $M^{p+q}_{p,p-r}$ by using all reduced matrix blocks $\mathscr{R}(M^{p+q}_{p-r+j,p-r})$ from $s=0$ up to $r-1$; notice that before adding a column from some reduced block we need to first ignore some rows. These reductions correspond to *lifting* our zero page representatives to the r-page, see figure 3.8. The resulting reduced matrix $\mathscr{R}(M^{p+q}_{p,q-r})$ has nonzero columns which correspond to generators of $\mathrm{Im}(d^r_{p,q})$, which further clarifies the connection of Theorem 3.5.1 with the usual interpretation of persistence barcodes coming from pivots that we saw on section 3.2. Furthermore, we can use the nonzero columns from $\mathscr{R}(M^{p+q}_{p,p-r})$ to clear columns from $M^{p+q-1}_{p-r,p-2r}$. The zero columns from $\mathscr{R}(M^{p+q}_{p,p-r})$ that have not been cleared lead to representatives in $C^{\text{cell}}_{p+q}(F_pX, F_{p-r-1}X)$ for a basis of $E^{r+1}_{p,q}$; i.e. we track the preimages corresponding to the column additions.

In both [58] and [5] all blocks along the same column have the same number of columns and rows, unlike the blocks $M^*_{*,*}$ presented on the previous paragraphs. However, one might object a couple of points about this proposed approach:

1. representative cycles are "lost" as successive pages are computed,

2. the clear optimization is not totally used, as new boundaries are added as successive pages are computed.

The answer for both objections lies in the hope that the blocks $M^{p+q}_{p,p-r}$ should get dramatically smaller as r increases. In such a case, one might start following all the columns from a reduced

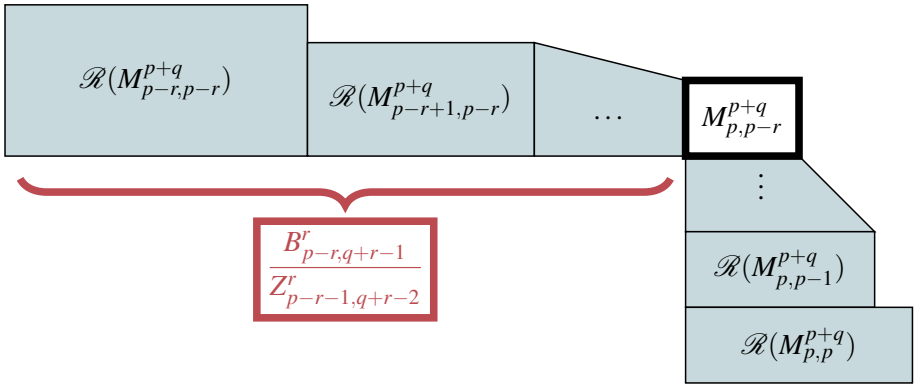


Figure 3.8: The block matrix $M^{p+q}_{p,p-r}$ together with the blocks that have been reduced on the left and down. Notice that the blocks along the column p decrease in the number of columns.

stage $\mathscr{R}(M^{p+q}_{p,p-r})$ independently of whether these vanish or not (and the same for the rows). That is, one follows exactly the algorithm from [58] and [5] from some page $r > 0$ onwards. This allows to obtain only those cycle representatives whose corresponding bars are "long enough". A future direction of interest would be to adapt the approach of [5] to compute spectral sequences of filtrations with the slight modification presented here.

3.7 Persistent Mayer-Vietoris Spectral Sequences

Perhaps more interesting for this thesis, we now proceed to review an idea proposed for the first time in [76]. This consists in using the Mayer-Vietoris spectral sequence as a tool to parallelize persistent homology computations. That is, consider a filtered complex X together with a cover $\mathscr{U}$ by filtered subcomplexes. As we have seen in section 2.11, for each $r \in \mathbf{R}$ we can compute the Mayer-Vietoris spectral sequence

$$E^1_{p,q}(X,\mathscr{U})(r) = \bigoplus_{\sigma^p \in N_{\mathscr{U}}} \mathrm{H}_q\big(\mathscr{U}_{\sigma^p}(r)\big) \Rightarrow \mathrm{H}_{p+q}\big(\mathscr{S}^{\mathrm{Tot}}_*(X,\mathscr{U})(r)\big)\ .$$

For each pair $r \leq s$ in $\mathbf{R}$, there are morphisms $\Sigma^{s-r} : \mathscr{S}^{\mathrm{Tot}}_*(X,\mathscr{U})(r) \to \mathscr{S}^{\mathrm{Tot}}_*(X,\mathscr{U})(s)$ which respect the vertical filtration F^*_V. These induce spectral sequence morphisms $\Sigma^{s-r} : E^*_{p,q}(X,\mathscr{U})(r) \to E^*_{p,q}(X,\mathscr{U})(s)$, which in turn induce morphisms between the targets of convergence. Then, we may define the *persistent Mayer-Vietoris spectral sequence*

$$E^1_{p,q}(X,\mathscr{U}) = \bigoplus_{\sigma^p \in N_{\mathscr{U}}} \mathrm{PH}_q(\mathscr{U}_{\sigma^p}) \Rightarrow \mathrm{PH}_{p+q}(\mathscr{S}^{\mathrm{Tot}}_*(X,\mathscr{U}))\ . \tag{3.3}$$

By repeating the argument in proposition 2.11.1, we obtain an isomorphism of persistence modules

$$\mathrm{PH}_n(\mathscr{S}_*^{\mathrm{Tot}}(X,\mathscr{U})) \simeq \mathrm{PH}_n(X)$$

for all dimensions $n \geq 0$.

An inconvenience with the persistent Mayer-Vietoris spectral sequence is that it has an associated extension problem which is nontrivial to solve in general, see section 6.2 for an explanation of this problem. The first instance in the literature that tackled this problem in the context of persistence appeared with the study of an *approximate nerve theorem* [65]. Essentially, one assumes that $\mathscr{U}$ is a ε-*acyclic cover* for some value $\varepsilon > 0$, in the sense that $d_I(\mathrm{PH}_n(\mathscr{U}_\sigma),0) < \varepsilon$ for all dimensions $n > 0$ and also $d_I(\mathrm{PH}_0(\mathscr{U}_\sigma), I(a_\sigma,\infty)) < \varepsilon$ for some value $a_\sigma \in \mathbf{R}$. On the other hand, one considers the filtered nerve $N_{\mathscr{U}} : \mathbf{R} \to \mathbf{SpCpx}$, where a simplex σ is in $N_{\mathscr{U}}(r)$ whenever $r \in [a_\sigma,\infty)$. Now we proceed to state the important result from [65].

Theorem 3.7.1. *Let X be a filtered complex, together with a cover by subcomplexes $\mathscr{U}$, and let $Q = \min(\dim(X), \dim(N_{\mathscr{U}}))$. If $\mathscr{U}$ is a ε-acyclic cover, then*

$$d_I(\mathrm{PH}_m(X), \mathrm{PH}_m(N_{\mathscr{U}})) < 2(Q+1)\varepsilon$$

for all $m \geq 0$.

If the parameter Q from theorem 3.7.1 is bounded: $Q < \infty$, then the spectral sequence $E^r_{p,q}(X,\mathscr{U})$ collapses at the $Q+1$ page. The factor 2 in the upper bound amounts to the successive page computations as well as tracking possible nontrivial extension problems. In chapter 8 we will obtain similar bounds for comparing persistence Mayer-Vietoris spectral sequences associated to two different covers.

Another instance of the Mayer-Vietoris spectral sequence in applied topology appeared in the context of cellular cosheaves [44, 43]. In [43] an example is presented for parallelizing homology computations on a space X, together with a one dimensional cover $\mathscr{U}$. These computations are further improved by using discrete Morse theory on cellular cosheaf homology. As we have seen in the corollary at the end of section 2.11, this bound on the nerve leads to a description of the homology of the covered space from the Čech homology of the corresponding cosheaves. Following this direction, Yoon and Ghrist [119, 120] studied the persistence Mayer-Vietoris spectral sequence for the particular case of one-dimensional nerves. The second page terms in the persistence spectral sequence can be described as

$$E^2_{p,q}(K,\mathscr{U}) = \check{\mathscr{H}}_p\Big(N_{\mathscr{U}}, \mathrm{PH}_q(\mathscr{U})\Big)$$

for $p = 0, 1$ and zero otherwise. Here we write $\mathrm{PH}_q(\mathscr{U})$ for the persistence precosheaf over $N_{\mathscr{U}}$ given by $\mathrm{PH}_q(\mathscr{U})(\sigma) = \mathrm{PH}_q(\mathscr{U}_\sigma)$ for all $\sigma \in N_{\mathscr{U}}$. Instead of recovering $\mathrm{PH}_*(X)$ directly from direct sums, as one does in the homology case, here there is a nontrivial extension problem to be solved. A solution to this problem was presented in [120], which we will later reproduce and extend to the case of unbounded dimension of the nerve in section 6.2. In [120] it was pointed out that computing the persistent Mayer-Vietoris spectral sequence leads to extra information about multiscale features of the dataset. In a similar direction, we will study stability properties of this spectral sequence on chapter 8.

Bringing back attention to the spectral sequence (3.3), notice that for effectively computing the successive pages, we need to be able to compute images, kernels and quotients of persistence modules and morphisms. Needless to say, these should be computed in an optimal way. This question has already been studied in [38], where the authors give a very efficient algorithm. However, there are couple of problems that come up when using [38] in spectral sequences:

1. In [38] the authors consider the case of a morphism is induced by the inclusion $X \subseteq Y$; for a simplicial morphism $X \to Y$ the authors consider the *mapping cylinder*. One problem with working with the mapping cylinder is that it introduces many unnecesary cells. However, the main issue comes from the fact that when working with the Mayer-Vietoris spectral sequence we consider maps in the second, third and higher pages which are not induced by a simplicial morphism at all. Furthermore, even when computing the first page we cannot adapt the work of [38] to compute the differentials. Indeed, the Čech differentials are not induced by simplicial morphisms; as a simplex from an intersection is sent to several copies along lower degree intersections.

2. A key assumption in [38] is that the filtrations in X and Y are both general in the sense that a simplex in either X or Y is born at a time. However, as we are considering persistence morphisms $f : \mathbb{V} \to \mathbb{W}$ which might not be induced by simplicial maps some problems happen as the interval decomposition from $\mathbb{V}$ and $\mathbb{W}$ might admit various bars with common startpoints and endpoints. In spectral sequences generality *hardly ever* holds. Indeed, this follows from the fact that a simplex might be contained in various overlapping covers.

Thus, if we want to compute images, kernels and cokernels, we will need to be able to overcome these two difficulties first. Also, notice that a good solution should lead to the representatives, as these are needed for computing spectral sequences.

Chapter 4

Persistence Algebra

4.1 Barcode Bases

In this section we will use the result from [30] to introduce barcode bases. Notice that all persistence modules that we will consider here are assumed to be tame. Our aim will be to come up with an efficient way of computing homology in the category of persistence modules. Notice that, while there is a uniquely determined barcode decomposition of $\mathbb{V}$, the particular choice of a basis will not be unique. This is analogous to the case of vector spaces, where a vector space can admit multiple bases but has always the same dimension. The main reason why we are introducing barcode bases is because we would like to work with morphisms between persistence modules $f : \mathbb{V} \to \mathbb{W}$. At the end we will introduce an algorithm for computing images and kernels, and we will evaluate its computational complexity. In addition, we will show in section 4.4 how to obtain barcode bases for quotients of persistence modules.

Definition 4.1.1 (Barcode Basis)**.** A barcode basis $\mathscr{B}$ of a tame persistence module $\mathbb{V}$ is a choice of an isomorphism, $\beta : \bigoplus_{i=1}^{N} \mathrm{I}(a_i, b_i) \to \mathbb{V}$. Each direct summand of β defines a restricted morphism from an interval module $\beta_i : \mathrm{I}(a_i, b_i) \to \mathbb{V}$, and will be called a *barcode generator*. We will usually denote a barcode basis $\mathscr{B}$ by the set of barcode generators $\mathscr{B} = \{\beta\}_{i=1}^{N}$.

Within the context of definition 4.1.1, we would like to make some notational remarks.

- Given a barcode generator $\beta_i \in \mathscr{B}$, we write $\beta_i \sim [a_i, b_i)$ to denote that β_i is a natural transformation $\beta_i : \mathrm{I}(a_i, b_i) \to \mathbb{V}$. In this case we say that β_i is *associated* to the interval $[a_i, b_i)$.

- Notice that if we choose $\beta_i \in \mathscr{B}$ with $\beta_i \sim [a_i, b_i)$ and $r \in \mathbf{R}$, we have a linear transformation $\beta_i(r) : \mathrm{I}(a_i, b_i)(r) \to \mathbb{V}(r)$. In particular, since $\mathrm{I}(a_i, b_i)(r)$ is either 0 or $\mathbb{F}$, the morphism $\beta_i(r)$ is uniquely determined by the image $\beta_i(r)(1_{\mathbb{F}}) \in \mathbb{V}(r)$ for $r \in [a_i, b_i)$. In addition, notice that $\beta_i(r) \neq 0$ for all $r \in [a_i, b_i)$ since otherwise β would not be injective.

- For any given $r \in \mathbf{R}$, we define the subset of $\mathscr{B}$

$$\mathscr{B}^r = \left\{ \beta_i : 0 \leq i \leq N,\ \ \beta_i(r) \neq 0 \right\}.$$

In this case, if $\beta_i \in \mathscr{B}^r$ and $\beta_i \sim [a_i, b_i)$, then $a_i \leq r < b_i$ by naturality of β_i. Also, evaluating all the elements from $\mathscr{B}^r$ on $1_{\mathbb{F}}$ leads to a vector base $\mathscr{B}^r(1_{\mathbb{F}})$ for $\mathbb{V}(r)$, where

$$\mathscr{B}^r(1_{\mathbb{F}}) = \left\{ \beta_i(r)(1_{\mathbb{F}}) : \text{ with } \beta_i \in \mathscr{B}^r \right\}.$$

Remark. We can think of a persistence module $\mathbb{V}$ as a sheaf over $\mathbb{R}$, where $\mathbb{R}$ is endowed with the topology where the open sets are either the intervals $[a,\infty)$ or (a,∞), for any $a \in \mathbb{R}$. Thus the restriction morphism $\rho_{b,a} : [a,\infty) \to [b,\infty)$ with values $a \leq b$ in $\mathbb{R}$, correspond to $\mathbb{V}(a \leq b) : \mathbb{V}^a \to \mathbb{V}^b$. A barcode base is determined by a set of global sections of the sheaf $\mathbb{V}$, such that they restrict to bases of the vector spaces $\mathbb{V}^r$, for all $r \in \mathbb{R}$. That is, $\mathscr{B} \subset \mathrm{PVect}(\mathbb{V})$ forms a barcode base for $\mathbb{V}$ if and only if $\mathscr{B}^r(1_{\mathbb{F}})$ forms a base of $\mathbb{V}^r$ for all $r \in \mathbb{R}$.

Let us formalize this remark on the following proposition.

Proposition 4.1.2. *Given a persistence morphism $\beta : \bigoplus_{i=1}^N \mathrm{I}(a_i, b_i) \to \mathbb{V}$ consider the set of direct summands $\mathscr{B} = \{\beta_i\}_{i=1}^N$. Then $\mathscr{B}$ is a barcode basis for $\mathbb{V}$ if and only if $\mathscr{B}^r(1_{\mathbb{F}})$ is a basis for $\mathbb{V}(r)$ for all $r \in \mathbf{R}$.*

Proof. Since $\mathbf{vect}_{\mathbb{F}}$ is an abelian category and $\mathbf{R}$ is a small category we have that β is an isomorphism if and only if $\beta(r)$ is an isomorphism for all $r \in \mathbf{R}$. That is, we consider the submodule $\mathrm{Ker}(\beta) \hookrightarrow \bigoplus_{i=1}^N \mathrm{I}(a_i, b_i)$ and notice that $\mathrm{Ker}(\beta) = 0$ iff $\mathrm{Ker}(\beta)(r) = 0$ for all $r \in \mathbf{R}$; a similar argument is done for surjectivity. Then, $\beta(r)$ is an isomorphism iff $\mathscr{B}^r(1_{\mathbb{F}})$ is a base for $\mathbb{V}(r)$ and the result follows. □

Suppose that $\mathscr{B}$ is a barcode base for a persistence module $\mathbb{V}$; there is a natural order for $\mathscr{B}$. For any pair of barcode generators $\alpha \sim [a,b)$ and $\beta \sim [c,d)$, we will write $\alpha < \beta$ whenever $a < c$ or when we have that $a = c$ and $d < b$. Assuming that $\mathbb{V}$ is tame and making some additional choices we might extend this order to a total order for $\mathscr{B}$.

Our next aim will consist in using barcode bases to understand persistence morphisms $f : \mathbb{V} \to \mathbb{W}$. In particular, given a pair of barcode bases $\mathscr{A}$ and $\mathscr{B}$ for $\mathbb{V}$ and $\mathbb{W}$ respectively, we will be able to obtain a single matrix F, not depending on the filtration parameter, which will be associated to f. Again, this is analogous to the case of linear algebra where a choice of bases leads to a single matrix associated to a linear map. As in linear algebra, we would like to develop a gaussian elimination method which consists in adding columns from the left to the right of an associated

matrix; here we will assume that both $\mathscr{A}$ and $\mathscr{B}$ are totally ordered. As we add columns from the left to right on F, we need to consider the addition of the corresponding barcode generators. However, when trying to add barcode generators we run into some problems which we illustrate on the following example.

Example 4.1.3. Consider $\mathbb{V} \simeq \mathrm{I}(0,2) \oplus \mathrm{I}(1,3)$ together with the canonical basis $\mathscr{B}$ given by generators $\beta_1 \sim [0,2)$ and $\beta_2 \sim [1,3)$. Then we cannot add β_1 and β_2 since at some points the domains differ; for example at filtration value 0 we have that $\beta_1(0)$ has $\mathbb{F}$ as domain but $\beta_2(0)$ has 0 as domain.

Thus, we cannot add barcode generators to obtain further barcode generators. To fix this, we consider the following set of pairs

$$\mathrm{PVect}(\mathbb{V}) = \{(\beta,(a,b)) \mid -\infty \leq a \leq b \leq \infty \text{ with } \beta : \mathrm{I}(a,b) \to \mathbb{V} \text{ where } \beta(r) \neq 0 \text{ iff } r \in [a,b)\} \ .$$

Given some pair $(\beta,(a,b)) \in \mathrm{PVect}(\mathbb{V})$, we will call the first component β a *persistence vector* and we will say that β is associated to the pair (a,b), which we will denote as $\beta \sim [a,b)$. Notice that in our definition of $\mathrm{PVect}(\mathbb{V})$ we include the pairs $(Z_a,(a,a))$ where $Z_a : 0 \to \mathbb{V}$ is the zero morphism; here we distinguish the zero persistence vectors Z_a and Z_b by the fact that these are associated to different pairs for $a \neq b$. We will define $\mathscr{Z} \subseteq \mathrm{PVect}(\mathbb{V})$ to be the subset of zero element pairs $(Z_a,(a,a))$ for all $a \in \mathbf{R}$.

We will proceed to show that the space $\mathrm{PVect}(\mathbb{V})$ has many properties analogous to those of vector spaces. To start, we define an operation to add persistence vectors.

Definition 4.1.4 (Barcode Sum)**.** We define the *barcode sum* as the assignment

$$\boxplus : \mathrm{PVect}(\mathbb{V}) \times \mathrm{PVect}(\mathbb{V}) \longrightarrow \mathrm{PVect}(\mathbb{V})$$

which sends $\big((\alpha,(a,b)),(\beta,(c,d))\big)$ to the pair $(\alpha \boxplus \beta,(\max(a,c),B(b,d,\alpha,\beta)))$ where

$$B(b,d,\alpha,\beta) = \begin{cases} \max(b,d) & \text{if } b \neq d \\ S(b,d,\alpha,\beta) & \text{if } b = d \ , \end{cases}$$

with

$$S(b,d,\alpha,\beta) = \begin{cases} \min(b,d) & \text{if } b \neq d \\ \sup\{r \in [\max(a,c),b) \mid \alpha(r)+\beta(r) \neq 0\} & \text{if } b = d \ , \end{cases}$$

and where for each $r \in [\max(a,c), B(b,d,\alpha,\beta))$ we define

$$\alpha \boxplus \beta(r) = \begin{cases} \alpha(r)+\beta(r) & \text{for } r \in [\max(a,c), S(b,d,\alpha,\beta)) \\ \alpha(r) & \text{for } r \in [d,b) \text{ if } d < b \\ \beta(r) & \text{for } r \in [b,d) \text{ if } b < d\ . \end{cases}$$

From checking the different cases one can see that $\alpha \boxplus \beta : \mathrm{I}(\max(a,c), B(b,d,\alpha,\beta)) \to \mathbb{V}$ is a well defined persistence morphism.

Notice that $\alpha \boxplus \beta(r) \neq 0$ if and only if $r \in [\max(a,c), B(b,d,\alpha,\beta))$, and that $\boxplus$ is commutative. For brevity, we will omit the pair notation $(\alpha,(a,b))$ of elements form $\mathrm{PVect}(\mathbb{V})$ and refer only to the first component α, the persistence vectors. Thus, we will abuse notation and say, "given a persistence vector $\alpha \in \mathrm{PVect}(\mathbb{V})$" or "given a subset of persistence vectors $\mathscr{B} \subseteq \mathrm{PVect}(\mathbb{V})$". In fact, the reason why we keep track of persistence vectors as pairs is to make sure we can distinguish zero elements within $\mathscr{Z} \subseteq \mathrm{PVect}(\mathbb{V})$. Also, notice that the elements $Z_a \in \mathscr{Z}$ behave nontrivially with respect to $\boxplus$; for example, given a persistence vector $\alpha \sim [a_\alpha, b_\alpha)$ and considering $c > b_\alpha$ we have that $\alpha \boxplus Z_c = Z_c$. We will now prove associativity.

Proposition 4.1.5. $\boxplus$ *is associative in* $\mathrm{PVect}(\mathbb{V})$.

Proof. Consider three persistence vectors $\alpha \sim [a_\alpha, b_\alpha)$, $\beta \sim [a_\beta, b_\beta)$ and $\gamma \sim [a_\gamma, b_\gamma)$ from $\mathrm{PVect}(\mathbb{V})$. We would like to show $L = R$ for $L := (\alpha \boxplus \beta) \boxplus \gamma$ and $R := \alpha \boxplus (\beta \boxplus \gamma)$. Notice that both L and R will share the same startpoint $A = \max(a_\alpha, a_\beta, a_\gamma)$; thus L and R are associated to a pair of intervals $[A, B_L)$ and $[A, B_R)$ respectively. Additionally, we have that $L(A) = \alpha(A) + \beta(A) + \gamma(A) = R(A)$. Thus, by naturality of persistence morphisms, we might deduce that for all $r \in \mathbf{R}$ with $A \leq r$

$$L(r) = \mathbb{V}(A \leq r)(L(A)) = \mathbb{V}(A \leq r)(R(A)) = R(r)\ .$$

Since $L(r) \neq 0$ iff $r \in [A, B_L)$ and $R(r) \neq 0$ iff $r \in [A, B_R)$ we must have $B_L = B_R$ and the equality $L = R$ holds. □

Another ingredient that we will need for developing algebraic manipulations on $\mathrm{PVect}(\mathbb{V})$ is scalar multiplication. We will define $\lambda : \mathbb{F} \times \mathrm{PVect}(\mathbb{V}) \to \mathrm{PVect}(\mathbb{V})$ to send a pair (c, β) to

$$\lambda(c,\beta) = \begin{cases} c\beta & \text{if } c \neq 0\ , \\ Z_a & \text{if } c = 0\ , \end{cases}$$

where $c\beta(r) = c \cdot \beta(r)$ for all $r \in \mathbf{R}$. With the definition of $\boxplus$ we can introduce the concept of *linear independence* of a subset $\mathscr{B} \subseteq \mathrm{PVect}(\mathbb{V})$. This will be key for characterizing barcode bases.

Definition 4.1.6. Given a subset $\mathscr{B} \subset \mathrm{PVect}(\mathbb{V}) \setminus \mathscr{Z}$, we say that $\mathscr{B}$ is *linearly independent* or that the vectors within $\mathscr{B}$ are *linearly independent* iff for any nonempty subset $\mathscr{S} \subseteq \mathscr{B}$ and any coefficients $k_\beta \in \mathbb{F} \setminus \{0\}$ with $\beta \in \mathscr{S}$, the sum $\boxplus_{\beta\in\mathscr{S}} k_\beta\beta$ is associated to the interval $[\max_{\beta\in\mathscr{S}}(a_\beta), \max_{\beta\in\mathscr{S}}(b_\beta))$, where $\beta \sim [a_\beta, b_\beta)$ for all $\beta \in \mathscr{S}$.

Let us illustrate linear independence with an example.

Example 4.1.7. Suppose that $\{\alpha \sim [0,2), \beta \sim [0,1)\}$ is a basis of $\mathbb{V} \simeq \mathrm{I}(0,2) \oplus \mathrm{I}(0,1)$. Then, one can see that $\{\alpha, \beta\}$ forms a set of linearly independent persistence vectors, this will follow from proposition 4.1.10. On the other hand, α and $\alpha \boxplus \beta$ are not linearly independent since the sum $(-\alpha) \boxplus (\alpha \boxplus \beta) = \beta$ is associated to $[0,1)$ but $-\alpha \sim [0,2)$ and $\alpha \boxplus \beta \sim [0,2)$ and the maximum of the endpoints is 2.

In proposition 4.1.10 we will show that a barcode base is linearly independent. However, we would like that a barcode base also generates the set $\mathrm{PVect}(\mathbb{V})$. For this, we need to introduce a further ingredient.

Definition 4.1.8. We define the *barcode cut* operation $\mathbf{1}_s : \mathrm{PVect}(\mathbb{V}) \to \mathrm{PVect}(\mathbb{V})$ for all $s \in \mathbf{R}$ as

$$\mathbf{1}_s(\alpha) = \alpha \boxplus Z_s \ ,$$

for all $s \in \mathbf{R}$ and all $\alpha \in \mathrm{PVect}(\mathbb{V})$.

Notice that λ and $\{\mathbf{1}_s\}_{s\in\mathbf{R}}$ are compatible, in the sense that $\mathbf{1}_r(k\beta) = k\mathbf{1}_r(\beta)$ for all $\beta \in \mathrm{PVect}(\mathbb{V})$, all $r \in \mathbf{R}$ and all $k \in \mathbb{F}$. Also, given $\beta, \gamma \in \mathrm{PVect}(\mathbb{V})$ and parameters $s, r \in \mathbf{R}$ it can be checked that

$$\mathbf{1}_s(\beta) \boxplus \mathbf{1}_r(\gamma) = \mathbf{1}_{\max(s,r)}(\beta \boxplus \gamma) \ .$$

Thus, one might think of persistence vectors on $\mathbb{V}$ as a tuple $\big(\mathrm{PVect}(\mathbb{V}), \boxplus, \lambda, \{\mathbf{1}_s\}_{s\in\mathbf{R}}\big)$.We can now define the concept of generation.

Definition 4.1.9. Given a nonempty subset $\mathscr{B} \subseteq \mathrm{PVect}(\mathbb{V})$, we say that $\mathscr{B}$ *generates* $\mathrm{PVect}(\mathbb{V})$ if and only if for any element $\alpha \in \mathrm{PVect}(\mathbb{V}) \setminus \mathscr{Z}$ there exist a subset $\mathscr{S} \subseteq \mathscr{B}$ together with coefficients $k_\beta \in \mathbb{F} \setminus \{0\}$ and some $s \in \mathbf{R}$ such that

$$\alpha = \mathbf{1}_s \left(\bigboxplus_{\beta\in\mathscr{S}} k_\beta \beta \right) \ .$$

We can now characterize barcode bases with the following result.

Proposition 4.1.10. *A subset $\mathscr{B} \subseteq \mathrm{PVect}(\mathbb{V})$ is a barcode basis for $\mathbb{V}$ if and only if $\mathscr{B}$ generates $\mathrm{PVect}(\mathbb{V})$ and it is linearly independent.*

Proof. Assume first that $\mathscr{B}$ is a barcode basis. By proposition 4.1.2 we know that $\mathscr{B}^r(1_\mathbb{F})$ is a basis for the vector space $\mathbb{V}(r)$ for all persistence values $r \in \mathbb{F}$. Then, for any persistence vector $\alpha : \mathrm{I}(a_\alpha, b_\alpha) \to \mathbb{V}$ we must have that $\alpha(a_\alpha)(1_\mathbb{F}) = \sum_{\beta \in \mathscr{B}^a} k_\beta \beta(a_\alpha)(1_\mathbb{F})$ for some coefficients $k_\beta \in \mathbb{F}$ for all $\beta \in \mathscr{B}^{a_\alpha}$. By naturality of α, this implies $\alpha = \mathbf{1}_{a_\alpha}\big(\coprod_{\beta \in \mathscr{B}^{a_\alpha}} k_\beta \beta\big)$. In particular, $\mathscr{B}$ generates PVect($\mathbb{V}$). On the other hand, assume that $\mathscr{B}$ is not linearly independent. Then there exist some nonempty subset $\mathscr{S} \subseteq \mathscr{B}$ together with coefficients $k_\beta \in \mathbb{F} \setminus \{0\}$ such that $\coprod_{\beta \in \mathscr{S}} k_\beta \beta$ is associated to an interval $[\max_{\beta \in \mathscr{S}}(a_\beta), B)$ with $B < \max_{\beta \in \mathscr{S}}(b_\beta)$. However this implies that $\mathscr{B}^B(1_\mathbb{F})$ is not linearly independent in $\mathbb{V}(B)$, since $\sum_{\beta \in \mathscr{S} \cap \mathscr{B}^B} k_\beta \beta(B)(1_\mathbb{F}) = 0$, but $k_\beta \in \mathbb{F} \setminus \{0\}$ for all $\beta \in \mathscr{S} \cap \mathscr{B}^B \neq \emptyset$. This would break linear independence of $\mathscr{B}^B(1_\mathbb{F})$, reaching a contradiction. Thus $\mathscr{B}$ must be linearly independent.

Now suppose that $\mathscr{B}$ generates PVect($\mathbb{V}$) and is linearly independent. We will prove that $\mathscr{B}$ is a barcode base by using proposition 4.1.2, so that all we need to show is that $\mathscr{B}^r(1_\mathbb{F})$ is a basis for the vector space $\mathbb{V}(r)$ for all filtration values $r \in \mathbf{R}$. First, we will show that $\mathscr{B}^r(1_\mathbb{F})$ generates $\mathbb{V}(r)$ for all $r \in \mathbf{R}$. Consider any nonzero vector $g \in \mathbb{V}(r)$ and define the persistence vector $\gamma : I(r,s) \to \mathbb{V}$ by setting $\gamma(r)(1_\mathbb{F}) = g$, where $s = \sup\big\{a : a \in \mathbf{R} \text{ and } \mathbb{V}(r \leq a)(g) \neq 0\big\}$. Thus, by generation of $\mathscr{B}$ there exists some subset $\mathscr{S} \subseteq \mathscr{B}$ together with some coefficients $k_\beta \in \mathbb{F}$ such that $\gamma = \mathbf{1}_r\big(\coprod_{\beta \in \mathscr{S}} k_\beta \beta\big)$. In particular we have that $g = \gamma(r)(1_\mathbb{F}) = \sum_{\beta \in \mathscr{S}} k_\beta \beta(r)(1_\mathbb{F})$. Therefore the claim follows. Next, let us show that $\mathscr{B}^r(1_\mathbb{F})$ is linearly independent. To see this, consider any non-empty subset $\mathscr{S} \subseteq \mathscr{B}^r$ together with coefficients $k_\beta \in \mathbb{F} \setminus \{0\}$ for all $\beta \in \mathscr{S}$. Then $\Gamma = \coprod_{\beta \in \mathscr{S}} k_\beta \beta$ is associated to $[\max_{\beta \in \mathscr{S}}(a_\beta), \max_{\beta \in \mathscr{S}}(b_\beta))$ which must contain r, and so $\Gamma(r)(1_\mathbb{F}) = \sum_{\beta \in \mathscr{S}} k_\beta \beta(r)(1_\mathbb{F}) \neq 0$. Thus, $\mathscr{B}^r(1_\mathbb{F})$ is indeed linearly independent and it forms a basis for $\mathbb{V}(r)$. □

Let $f : \mathbb{V} \to \mathbb{W}$ be a morphism of tame persistence modules and consider two bases $\mathscr{A}$ and $\mathscr{B}$ for $\mathbb{V}$ and $\mathbb{W}$ respectively. Given a persistence vector $\gamma \in \mathrm{PVect}(\mathbb{V})$ with $\gamma \sim [a_\gamma, b_\gamma)$, we define the image $f(\gamma) : \mathrm{I}(a_\gamma, b_{f(\gamma)}) \to \mathbb{W}$ where $b_{f(\gamma)} = \sup\big\{r \in [a_\gamma, b_\gamma) \mid f(r) \circ \gamma(r) \neq 0\big\}$ and $f(\gamma)(r) = f(r) \circ \gamma(r)$ for all $r \in [a_\gamma, b_{f(\gamma)})$.

Now, for each barcode generator $\alpha \sim [a_\alpha, b_\alpha)$ in $\mathscr{A}$, as $\mathscr{B}$ is a barcode base, there exist some subset $\mathscr{S} \subseteq \mathscr{B}$ together with coefficients $k_{\beta,\alpha} \in \mathbb{F} \setminus \{0\}$ for all $\beta \in \mathscr{S}$ such that

$$f(\alpha) = \mathbf{1}_{a_\alpha}\left(\coprod_{\beta \in \mathscr{S}} k_{\beta,\alpha}\beta\right).$$

One might assume that $\mathscr{S} \subset \mathscr{B}^{a_\alpha}$ since adding elements from $\mathscr{B} \setminus \mathscr{B}^{a_\alpha}$ would have no effect or would cut the startpoint to a value greater than a_α. Also notice that if $\beta(b_\alpha) \neq 0$, then $k_{\beta,\alpha} = 0$, since otherwise f would not be natural as a persistence morphism. Thus we can define the subset

of $\mathscr{B}^{a_\alpha}$ associated to α:

$$\mathscr{B}(\alpha) \coloneqq \left\{ \beta : \beta \in \mathscr{B},\ \beta(a_\alpha) \neq 0,\ \beta(b_\alpha) = 0 \right\} \subseteq \mathscr{B}^{a_\alpha} \subseteq \mathscr{B}$$

The set $\mathscr{B}(\alpha)$ contains the barcode generators $\beta \in \mathscr{B}$ such that the coefficients $k_{\beta,\alpha}$ might be non-zero; thus $\mathscr{S} \subseteq \mathscr{B}(\alpha)$. By pointwise-linearity and naturality of f, we have that for any subset $\mathscr{S} \subseteq \mathscr{A}$ the equality

$$f\left(\coprod_{\alpha \in \mathscr{S}} k_\alpha \alpha \right) = \coprod_{\alpha \in \mathscr{S}} k_\alpha f(\alpha)$$

holds, where $k_\alpha \in \mathbb{F}$ for all $\alpha \in \mathscr{S}$.

Corollary 4.1.11. *Consider a pair of persistence modules $\mathbb{V}$ and $\mathbb{W}$ together with a pair of respective barcode bases $\mathscr{A}$ and $\mathscr{B}$. If there is a persistence morphism $f : \mathbb{V} \to \mathbb{W}$, then there is a unique associated matrix $F = \left(k_{\beta,\alpha}\right)_{\beta \in \mathscr{B}, \alpha \in \mathscr{A}}$ which is well-defined in the sense that whenever $k_{\beta\alpha} \neq 0$ then $\beta \in \mathscr{B}(\alpha)$. Conversely, assume that F is well-defined, then there exists a unique persistence morphism $f : \mathbb{V} \to \mathbb{W}$ whose associated matrix is F.*

Proof. By the reasoning above, we only need to prove the converse statement. First, notice that for each $\alpha \in \mathscr{A}$ such that $\alpha \sim [a_\alpha, b_\alpha)$, we can define $f(\alpha) \coloneqq \mathbf{1}_{a_\alpha}\left(\coprod_{\beta \in \mathscr{B}(\alpha)} k_{\beta,\alpha}\beta \right)$. By linear independence of $\mathscr{B}$, $f(\alpha)$ is associated to the interval $[a_\alpha, B)$ for $B = \max_{\beta \in \mathscr{S}}(b_\beta)$ with $\mathscr{S}$ being the set $\{\beta \in \mathscr{B}(\alpha) | k_{\beta,\alpha} \neq 0\}$. We can extend the definition of f by the linear formula $f(\coprod_{\alpha \in \mathscr{A}} c_\alpha \alpha) = \coprod_{\alpha \in \mathscr{A}} c_\alpha f(\alpha)$ for any coefficients $c_\alpha \in \mathbb{F}$ for all $\alpha \in \mathscr{A}$. This implies the claim as f is then natural and pointwise linear. □

4.2 Computing Kernels and Images

Consider two finite barcode bases $\mathscr{A} = \{\alpha_i\}_{i=1}^n$ and $\mathscr{B} = \{\beta_j\}_{j=1}^m$ for $\mathbb{V}$ and $\mathbb{W}$ respectively. Additionally, suppose that both $\mathscr{A}$ and $\mathscr{B}$ have total orderings. That is, even if two barcode generators are associated to the same interval $\alpha_r, \alpha_s \sim [a,b)$ for $r \neq s$, we have already made a choice $\alpha_r < \alpha_s$. Then we consider $M_I = (f(\alpha_1), \ldots, f(\alpha_n))$ the matrix of f in the bases $\mathscr{A}$ and $\mathscr{B}$. The aim will be to transform M_I performing left to right column additions until obtaining the reduced matrix

$$\mathscr{I} = \left(f(\alpha_1) \middle| f(\alpha_2) \boxplus k_{2,1} f(\alpha_1) \middle| \ldots \middle| f(\alpha_n) \boxplus \coprod_{j=1}^{n-1} k_{n,j} f(\alpha_j) \right) \tag{4.1}$$

for suitable $k_{i,j} \in \mathbb{F}$ and $1 \leq j < i \leq n$. This $\mathscr{I}$ will have the property that its non-zero columns form a basis $\widetilde{\mathscr{I}}$ for $\mathrm{Im}(f)$. The key to check this property will be, as in traditional linear algebra, to pay close attention to the column pivots.

Definition 4.2.1. Given some subset $\mathscr{S} \subseteq \mathscr{A}$, we might consider a vector $V = (k_\alpha)_{\alpha\in\mathscr{A}}$ such that $k_\alpha \neq 0$ iff $\alpha \in \mathscr{S}$. We call the *pivot* of V, to be the element $\tau \in \mathscr{S}$ whose associated interval $[a_\tau, b_\tau)$ is such that for any other element $\alpha \in \mathscr{S}$ associated to $[a_\alpha, b_\alpha)$ we have that $b_\alpha \leq b_\tau$, and if $b_\alpha = b_\tau$ then $\alpha < \tau$ in the chosen basis order from $\mathscr{A}$. We will also refer to the *pivot* of $\mathscr{S}$ or the *pivot* of $\coprod_{\alpha\in\mathscr{S}} k_\alpha \alpha$.

In particular, notice that our definition of a pivot differs from the traditional "last nonzero entry" definition. A matrix $\mathscr{I}$ that has unique pivots yields a set of linearly independent persistence vectors; we will show this in proposition 4.2.3.

Now, let us go back to the reduced matrix $\mathscr{I}$ from (4.1). By linearity of persistence morphisms we have that the j column from $\mathscr{I}$ is $f(\alpha_j \boxplus \coprod_{i=1}^{j-1} k_{j,i}\alpha_i)$; thus, its preimage p_j is given as $\alpha_j \boxplus \coprod_{i=1}^{j-1} k_{j,i}\alpha_i$; we define by $\mathscr{P}\mathscr{I}$ the set of preimages of $\mathscr{I}$. Given some preimage $p_j \in \mathscr{P}\mathscr{I}$, notice that it is associated to some interval $[a_j, b_j)$ while its image $f(p_j)$ is associated to $[a_j, c_j)$ for filtration values $c_j \leq b_j$. Thus, we must have that $\mathbf{1}_{c_j}(p_j) \in \ker(f)$ for all $1 \leq j \leq n$. We consider $\mathscr{G}\mathscr{K} = \{\mathbf{1}_{c_j}(p_j)\}_{1\leq j\leq n}$ which generates $\ker(f)$, as it is shown later in proposition 4.3.1. Then, we order $\mathscr{G}\mathscr{K}$ by choosing a permutation $\sigma : \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ such that it is consistent with the standard order of barcode vectors. Using the order from $\mathscr{G}\mathscr{K}$ we consider the matrix $M_K = \left(\mathbf{1}_{c_{\sigma(1)}}(p_{\sigma(1)}), \ldots, \mathbf{1}_{c_{\sigma(n)}}(p_{\sigma(n)})\right)$ noticing that it is the associated matrix for the composition of $\bigoplus_{j=1}^n \mathbf{1}_{c_{\sigma(j)}}(p_{\sigma(j)}) : \bigoplus_{j=1}^n \mathrm{I}(c_{\sigma(j)}, b_{\sigma(j)}) \twoheadrightarrow \ker(f)$ with the inclusion $\ker(f) \hookrightarrow \mathbb{V}$. Reducing columns we find some coefficients $q_{i,j} \in \mathbb{F}$ so that the resulting matrix has unique pivots:

$$\mathscr{K} = \left(\mathbf{1}_{c_{\sigma(1)}}\left(p_{\sigma(1)}\right) \middle| \mathbf{1}_{c_{\sigma(2)}}\left(p_{\sigma(1)} \boxplus q_{2,2}p_{\sigma(2)}\right) \middle| \ldots \middle| \mathbf{1}_{c_{\sigma(n)}}\left(p_{\sigma(n)} \boxplus \coprod_{i=1}^{n-1} q_{j,i}p_{\sigma(i)}\right)\right) .$$

Taking the non-zero columns from $\mathscr{K}$ leads to a basis $\widetilde{\mathscr{K}}$ for $\mathrm{Ker}(f)$. In the following we will present an algorithm obtaining such bases $\mathrm{Im}(f)$ and $\ker(f)$. First we will go through an illustrative example encoding some of the basic principles of the procedure.

Example 4.2.2. Consider two persistence modules

$$\mathbb{V} \simeq \mathrm{I}(1,5) \oplus \mathrm{I}(1,4) \oplus \mathrm{I}(2,5), \qquad \mathbb{W} \simeq \mathrm{I}(0,5) \oplus \mathrm{I}(0,3) \oplus \mathrm{I}(1,4)$$

with canonical barcode bases $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ respectively. Let the morphism $f : \mathbb{V} \to \mathbb{W}$ be given by the $|\mathscr{B}| \times |\mathscr{A}|$ matrix M_I that we depict on the left:

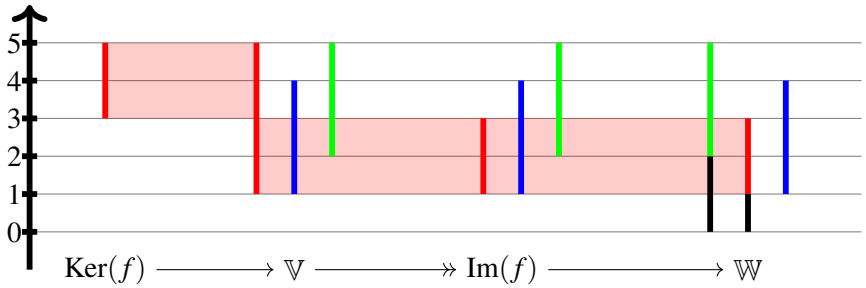


Figure 4.1: Decomposition of barcodes in image, kernel, domain and codomain of $f : \mathbb{V} \to \mathbb{W}$. The colors correspond to the different generators associated to $\widetilde{\mathscr{I}}$ and $\widetilde{\mathscr{K}}$.

$$\left(\begin{array}{c|c|c|c} M_I & \alpha_1 & \alpha_2 & \alpha_3 \\ \hline \beta_1 & 0 & 0 & 1 \\ \beta_2 & 1 & 0 & 0 \\ \beta_3 & 1 & 1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{c|c|c|c} \mathscr{I} & \alpha_1 & \alpha_2 \boxplus (-\alpha_1) & \alpha_3 \\ \hline \beta_1 & 0 & 0 & 1 \\ \beta_2 & 1 & -1 & 0 \\ \beta_3 & 1 & 0 & 1 \end{array}\right).$$

Then notice that the first two columns from M_I share the same pivot β_3, while the third's column pivot is β_1. We add the first column to the second multiplying by -1, so that we obtain the matrix $\mathscr{I}$ above which has unique pivots for each column. From the matrix $\mathscr{I}$ we obtain:

$$f(\alpha_1) = \mathbf{1}_1(\beta_2 \boxplus \beta_3) \qquad f(\alpha_2 \boxplus (-\alpha_1)) = -\mathbf{1}_1(\beta_2) \qquad f(\alpha_3) = \mathbf{1}_2(\beta_1 \boxplus \beta_3)\,.$$

In particular, we obtain a basis for the image $\widetilde{\mathscr{I}} = \{\mathbf{1}_1(\beta_2 \boxplus \beta_3),\ -\mathbf{1}_1(\beta_2),\ \mathbf{1}_2(\beta_1 \boxplus \beta_3)\}$, which leads to the interval decomposition $\mathrm{Im}(f) \simeq \mathrm{I}(1,4) \oplus \mathrm{I}(1,3) \oplus \mathrm{I}(2,5)$. At the same time, we obtain a corresponding set of preimages $\mathscr{PI} = \{\alpha_1, \alpha_2 \boxplus (-\alpha_1), \alpha_3\}$. From this, we deduce the ordered set of kernel generators $\mathscr{GK} = \{\mathbf{1}_3(\alpha_2 \boxplus (-\alpha_1)),\ \mathbf{1}_4(\alpha_1),\ \mathbf{1}_5(\alpha_3)\}$. Thus, we consider the matrix M_K for the kernels and reduce it

$$\left(\begin{array}{c|c|c|c} M_K & \mathbf{1}_3(\alpha_2 \boxplus (-\alpha_1)) & \mathbf{1}_4(\alpha_1) & \mathbf{1}_5(\alpha_3) \\ \hline \alpha_1 & -1 & 1 & 0 \\ \alpha_2 & 1 & 0 & 0 \\ \alpha_3 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{c|c|c|c} \mathscr{K} & \mathbf{1}_3(\alpha_2 \boxplus (-\alpha_1)) & \mathbf{1}_4(\alpha_2) & \mathbf{1}_5(\alpha_3) \\ \hline \alpha_1 & -1 & 0 & 0 \\ \alpha_2 & 1 & 1 & 0 \\ \alpha_3 & 0 & 0 & 1 \end{array}\right)$$

Since $\mathbf{1}_4(\alpha_2) = Z_4$ and $\mathbf{1}_5(\alpha_3) = Z_5$ we have that the second and third columns from $\mathscr{K}$ vanish. Thus, we obtain a basis for the kernel $\widetilde{\mathscr{K}} = \{\mathbf{1}_3(\alpha_2 \boxplus (-\alpha_1))\}$. which will be such that $\ker(f) \simeq \mathrm{I}(3,5)$. This is illustrated in Figure 4.1.

We end this section with useful result to obtain linear independence.

Proposition 4.2.3. *Let $\mathbb{V}$ be a persistence module with a barcode base $\mathscr{A}$. Consider a set of persistence vectors $\mathscr{M} \subset \mathrm{PVect}(\mathbb{V})$ and suppose that their pivots on the basis $\mathscr{A}$ are all different. Then $\mathscr{M}$ is linearly independent.*

Proof. First, for each element $m \in \mathscr{M}$ we write it in terms of $\mathscr{A}$ as the sum $m = \mathbf{1}_{r_m}\left(\boxplus_{\alpha\in\mathscr{S}_m} k_\alpha \alpha\right)$ for some subset $\mathscr{S}_m \subseteq \mathscr{A}$ and coefficients $k_\alpha \in \mathbb{F}\setminus\{0\}$ for all $\alpha \in \mathscr{S}_m$ and some filtration value $r_m \in \mathbf{R}$. We will write $[a_m, b_m)$ for the associated interval of m and α_m for its pivot from $\mathscr{S}_m$. In particular, by the definition of pivot we must have $\alpha_m \sim [c_m, b_m)$ for some filtration value $c_m \leq a_m$. Now, consider a nonempty subset $\mathscr{R} \subseteq \mathscr{M}$ together with coefficients $q_m \in \mathbb{F}\setminus\{0\}$ for all $m \in \mathscr{R}$ and take the sum $V = \boxplus_{m\in\mathscr{R}} q_m m$. We claim that $V \sim [\max_{m\in\mathscr{R}}(a_m), \max_{m\in\mathscr{R}}(b_m))$. Consider the element $P \in \mathscr{R}$ whose pivot α_P is the 'highest' according to definition 4.2.1; i.e. the pivot from $\bigcup_{m\in\mathscr{R}} \mathscr{S}_m$. Consequently, we have that $b_P = \max_{m\in\mathscr{R}}(b_m)$. Notice that the element P is unique, since otherwise there would be two elements from $\mathscr{M}$ with the same pivots, contradicting our assumption. This implies that V is written in terms of $\mathscr{A}$ with a nonzero coefficient for α_P. By linear independence of $\mathscr{A}$ the claim follows. □

4.3 `image_kernel` algorithm

Here, we present an algorithm performing the above procedure. Suppose that $f : \mathbb{V} \to \mathbb{W}$ is a morphism between two tame persistence modules. Let $\mathscr{A}$ and $\mathscr{B}$ be barcode bases for $\mathbb{V}$ and $\mathbb{W}$ respectively. Suppose also that we know $f(\mathscr{A})_{\mathscr{B}}$, the matrix associated to f with respect to barcode bases $\mathscr{A}$ and $\mathscr{B}$, and store it into M_I. Performing left to right column additions on M_I will eventually lead to the reduced matrix $\mathscr{I}$; we will call this method `box_gauss_reduce` [1], and it is summarized in algorithm 4.1. Notice that this procedure is a standard Gaussian elimination using $\boxplus$. Here we use a method `get_pivot`, which obtains the pivot of a given nontrivial column from a matrix. Once we have run `box_gauss_reduce`, we obtain the reduced matrix for the images $\mathscr{I}$ and the corresponding preimages $\mathscr{P}\mathscr{I}$. Then we obtain the set of generators $\mathscr{G}\mathscr{K}$ for $\ker(f)$, order it according to the standard order and define a matrix $\mathscr{G}\mathscr{K}_{\mathscr{A}}$ which expresses the elements from $\mathscr{G}\mathscr{K}$ in terms of $\mathscr{A}$. Applying `box_gauss_reduce` to the triple $(\mathscr{G}\mathscr{K}, \mathscr{A}, \mathscr{G}\mathscr{K}_{\mathscr{A}})$ we recover the reduced matrix $\mathscr{K}$. An outline of this procedure is shown in Algorithm 4.2. Reading out the nonzero elements from $\mathscr{I}$ and $\mathscr{K}$ leads to the barcode bases $\widetilde{\mathscr{I}}$ and $\widetilde{\mathscr{K}}$.

Proposition 4.3.1. *Algorithm 4.2 computes $\widetilde{\mathscr{K}}$ and $\widetilde{\mathscr{I}}$ bases for the kernel and image of f. Furthermore, it takes at most $\mathscr{O}(N|\mathscr{A}|^2)$ time, where $N = \max(|\mathscr{A}|, |\mathscr{B}|)$.*

[1] Here we use the Python notation for matrices, where for a matrix M, the (i,j)-entry is denoted by `M[i][j]` and the j^{th} column is denoted by `M[:,j]`. Also, we implement persistence vectors as ordinary vectors, however, one needs to make sure that ordinary vector addition is adapted to the operation $\boxplus$, also, one needs to keep track of the associated intervals of the entries.

Algorithm 4.1 `box_gauss_reduce`

Input: $\mathscr{A}$, $\mathscr{B}$, $f(\mathscr{A})_{\mathscr{B}}$
Output: $\mathscr{I}$, $\mathscr{P}\mathscr{I}$

Set $\mathscr{I} = \emptyset$ and $\mathscr{P}\mathscr{I} = \emptyset$.
Set $M_I = f(\mathscr{A})_{\mathscr{B}}$, $T = \mathrm{Id}_{|\mathscr{A}|}$, $\texttt{pivots} = (-1,-1,\cdots,-1) \in \mathscr{M}(\mathbb{F})_{1\times|\mathscr{A}|}$
for $1 \le i \le |\mathscr{A}|$ **do**
    $\texttt{pivots}[i] \leftarrow \texttt{get\_pivot}(\mathscr{I}[:,i])$
    for $1 \le j \le i-1$ **do**
        if $\texttt{pivots}[i] == \texttt{pivots}[j]$ **then**
            $M_I[:,i] \leftarrow M_I[:,i] \boxplus (-M_I[\texttt{pivots}[i]][i]M_I[:,j])$
            $T[:,i] \leftarrow T[:,i] - M_I[\texttt{pivots}[i]][i]T[:,j]$
            $\texttt{pivots}[i] \leftarrow \texttt{get\_pivot}(M_I[i])$
        end if
    end for
    Compute $\mathscr{S}_I^i = \{\beta_j \in \mathscr{B} \,|\, M_I[j][i] \neq 0\} \subseteq \mathscr{B}$ and also $\mathscr{S}_P^i = \{\alpha_j \in \mathscr{A} \,|\, T[j][i] \neq 0\} \subseteq \mathscr{A}$.
    Add $f(p_i) = \boxplus_{\beta_j \in \mathscr{S}_I^i} M_I[j][i]\beta_j$ to $\mathscr{I}$.
    Add $p_i = \boxplus_{\alpha_j \in \mathscr{S}_P^i} T[j][i]\alpha_j$ to $\mathscr{P}\mathscr{I}$.
end for
return $\mathscr{I}$ and $\mathscr{P}\mathscr{I}$

Algorithm 4.2 `image_kernel`

Input: $\mathscr{A}$, $\mathscr{B}$, $f(\mathscr{A})_{\mathscr{B}}$
Output: $\widetilde{\mathscr{K}}$, $\widetilde{\mathscr{I}}$, $\mathscr{P}\mathscr{I}$

$\mathscr{I}, \mathscr{P}\mathscr{I} \leftarrow \texttt{box\_gauss\_reduce}(\mathscr{A}, \mathscr{B}, f(\mathscr{A})_{\mathscr{B}})$
Define $\mathscr{G}\mathscr{K}$ to be the set containing all $\mathbf{1}_{c_i}(p_i)$ for all $p_i \in \mathscr{P}\mathscr{I}$ with $f(p_i) \sim [a_i, c_i)$.
Order $\mathscr{G}\mathscr{K}$, also define the matrix $\mathscr{G}\mathscr{K}_{\mathscr{A}}$ of coordinates from $\mathscr{G}\mathscr{K}$ in terms of $\mathscr{A}$.
$\mathscr{K}, \_ \leftarrow \texttt{box\_gauss\_reduce}(\mathscr{G}\mathscr{K}, \mathscr{A}, \mathscr{G}\mathscr{K}_{\mathscr{A}})$
We get rid of zero elements to obtain $\widetilde{\mathscr{K}}$ and $\widetilde{\mathscr{I}}$ from $\mathscr{K}$ and $\mathscr{I}$.
return $\widetilde{\mathscr{K}}$ and $\widetilde{\mathscr{I}}$ (optionally return $\mathscr{P}\mathscr{I}$ for preimages)

Proof. First of all, notice that by proposition 4.2.3 we know that $\widetilde{\mathscr{I}}$ and $\widetilde{\mathscr{K}}$ are linearly independent, as these are both sets of persistence vectors with different pivots. Thus, all we need to show is that both sets generate $\mathrm{PVect}(\mathrm{Im}(f))$ and $\mathrm{PVect}(\mathrm{Ker}(f))$ respectively.

Let us prove that $\widetilde{\mathscr{I}}$ generates $\mathrm{PVect}(\mathrm{Im}(f))$. First, we will show that $\mathscr{P}\mathscr{I}$ generates $\mathrm{PVect}(\mathbb{V})$. Consider $\gamma \in \mathrm{PVect}(\mathbb{V})$ and write $\gamma = \mathbf{1}_a(\boxplus_{i\in I} k_i\alpha_i)$ for coefficients $k_i \in \mathbb{F}\setminus\{0\}$ with $i \in I$ for some subset $I \subseteq \{1,2,\ldots,|\mathscr{A}|\}$. Then, consider the maximum index m from I and compute $\widetilde{\gamma} = \gamma \boxplus (-k_m p_m)$. Here it is key to recall that the preimage $p_m \in \mathscr{P}\mathscr{I}$ is written as $\alpha_m \boxplus \boxplus_{i=1}^{m-1} k_{m,i}\alpha_i$ for coefficients $k_{m,i} \in \mathbb{F}$ for $1 \leq i < m$. Now, $\widetilde{\gamma} = \mathbf{1}_a\left(\boxplus_{i\in J} \widetilde{k_i}\alpha_i\right)$ for coefficients $\widetilde{k_i} \in \mathbb{F}\setminus\{0\}$ with $i \in J$ for some subset $J \subseteq \{1,2,\ldots,m-1\}$. Repeating this argument, eventually we write γ in terms of $\mathscr{G}\mathscr{I}$. This implies that $f(\gamma)$ can be expressed in terms of $\widetilde{\mathscr{I}}$. Thus, $\widetilde{\mathscr{I}}$ generates $\mathrm{PVect}(\mathrm{Im}(f))$.

Now, let us show that $\widetilde{\mathscr{K}}$ generates $\mathrm{PVect}(\mathrm{Ker}(f))$. In fact, it will be enough to show that $\mathscr{G}\mathscr{K}$ generates $\mathrm{PVect}(\ker(f))$. This is because $\widetilde{\mathscr{K}}$ is obtained from reducing $\mathscr{G}\mathscr{K}_{\mathscr{A}}$ in a similar manner as $\widetilde{\mathscr{I}}$ was obtained by reducing $f(\mathscr{A})_{\mathscr{B}}$. Consequently, by replicating the argument which proved that $\mathscr{P}\mathscr{I}$ generates $\mathrm{PVect}(\mathbb{V})$, we can show that $\widetilde{\mathscr{K}}$ generates $\mathrm{PVect}(\ker(f))$. So let us prove our claim. Suppose that $\gamma : \mathrm{I}(a,b) \to \mathbb{V}$ lies in the kernel; i.e. $f(\gamma) = Z_a$. As $\mathscr{P}\mathscr{I}$ generates $\mathrm{PVect}(\mathbb{V})$, we have that $\gamma = \mathbf{1}_a(\boxplus_{i\in I} k_i p_i)$ for coefficients $k_i \in \mathbb{F}\setminus\{0\}$ with $i \in I$ for some subset $I \subseteq \{1,2,\ldots,|\mathscr{A}|\}$. Applying f, we obtain the equality $f(\gamma) = Z_a = \boxplus_{i\in I} k_i f(p_i)$ and notice that $f(p_i)(a) = 0$ for all $i \in I$; otherwise linear independence of $\widetilde{\mathscr{I}}$, and in particular that of $\widetilde{\mathscr{I}}(a)(1_{\mathbb{F}})$, would be contradicted. However, if $f(p_i)(a) = 0$ for all $i \in I$, then $\mathbf{1}_{c_i}(p_i) \in \mathscr{G}\mathscr{K}$ for some $c_i \leq a$ and all $i \in I$. Altogether we obtain that γ must be generated by $\mathscr{G}\mathscr{K}$.

The computational bound follows from computing `box_gauss_reduce` on lines 1 and 4 of algorithm 4.2. For the image, it takes about $\mathscr{O}(|\mathscr{A}|^2)$ time to run the pair of nested "for loops" from algoritmh 4.1 and about $\mathscr{O}(\max(|\mathscr{A}|,|\mathscr{B}|))$ time to perform the column additions from lines 7-8. Reducing the matrix associated to the kernel should take about $\mathscr{O}(|\mathscr{A}|^3)$ time. □

Remark. The procedure for computing images and kernels described here relies heavily on proposition 4.1.10. However, the early versions of PERMAVISS [116] used an `image_kernel` procedure which mainly relied on proposition 4.1.2. This original `image_kernel` procedure, together with its improved version, is described in appendix A. However, `box_gauss_reduce` algorithm will outperform these methods as our estimates suggest. Nonetheless, for completeness, we include all these procedures in this thesis.

4.4 Computing Quotients

Now we consider the problem of computing quotients. Suppose that we have inclusions $\mathbb{H} \subseteq \mathbb{G} \subseteq \mathbb{V}$ of finite dimensional persistence modules, together with totally ordered barcode bases $\mathscr{H}$, $\mathscr{G}$ and $\mathscr{A}$ respectively. The aim will be to find a barcode base for $\mathbb{G}/\mathbb{H}$. Consider the inclusions $\imath^{\mathbb{H}} : \mathbb{H} \hookrightarrow \mathbb{V}$ and $\imath^{\mathbb{G}} : \mathbb{G} \hookrightarrow \mathbb{V}$, together with their respective associated matrices $\imath^{\mathbb{H}}(\mathscr{H})_{\mathscr{A}} \in \mathscr{M}_{|\mathscr{A}|\times|\mathscr{H}|}(\mathbb{F})$ and $\imath^{\mathbb{G}}(\mathscr{G})_{\mathscr{A}} \in \mathscr{M}_{|\mathscr{A}|\times|\mathscr{G}|}(\mathbb{F})$. We consider the module $\mathbb{H}\oplus\mathbb{G}$ together with a barcode base given by the pair $(\mathscr{H}|\mathscr{G})$; here we extend the orders from $\mathscr{H}$ and $\mathscr{G}$ with the rule $h < g$ for any pair of generators $h \in \mathscr{H}$ and $g \in \mathscr{G}$; of course, this might break the standard persistence vector order. Then, we consider the morphism given by the addition of the inclusions $\imath = \imath^{\mathbb{H}} + \imath^{\mathbb{G}} : \mathbb{H}\oplus\mathbb{G} \to \mathbb{V}$ which will have the associated block matrix $(\imath^{\mathbb{H}}(\mathscr{H})_{\mathscr{A}}|\imath^{\mathbb{G}}(\mathscr{G})_{\mathscr{A}})$. We send the triple $((\mathscr{H}|\mathscr{G}), \mathscr{A}, (\imath^{\mathbb{H}}(\mathscr{H})_{\mathscr{A}}|\imath^{\mathbb{G}}(\mathscr{G})_{\mathscr{A}}))$ to the `box_gauss_reduce` algorithm and obtain the corresponding set of reduced columns $\mathscr{I}$. Even though the standard barcode order was not respected in the domain, notice that $\mathscr{I}$ will be linearly independent by proposition 4.2.3. A key observation will be that the subset $\mathscr{I}[\mathscr{H}]$ containing the first $|\mathscr{H}|$ elements from $\mathscr{I}$ forms a barcode base for $\mathbb{H}$ by proposition 4.3.1.

We will focus on the subset $\mathscr{I}[\mathscr{G}] = \mathscr{I} \setminus \mathscr{I}[\mathscr{H}]$ from $\mathscr{I}$. Recall that the `box_gauss_reduce` algorithm adds columns from $\imath^{\mathbb{H}}(\mathscr{H})_{\mathscr{A}}$ and $\imath^{\mathbb{G}}(\mathscr{G})_{\mathscr{A}}$ to eventually obtain each element $\Gamma \in \mathscr{I}[\mathscr{G}]$. For each $\Gamma \in \mathscr{I}[\mathscr{G}]$, we will write it as $\Gamma = \Gamma^{\mathbb{H}} \boxplus \Gamma^{\mathbb{G}}$ where $\Gamma^{\mathbb{H}}$ and $\Gamma^{\mathbb{G}}$ denote the respective combinations from elements from $\imath^{\mathbb{H}}(\mathscr{H})_{\mathscr{A}}$ and $\imath^{\mathbb{G}}(\mathscr{G})_{\mathscr{A}}$ that we obtain in `box_gauss_reduce`. Given $\Gamma \in \mathscr{I}[\mathscr{G}]$, we will use the notation $\Gamma^{\mathbb{G}} \sim [a_\Gamma, d_\Gamma)$ and $\Gamma \sim [b_\Gamma, c_\Gamma)$ for the corresponding associated intervals; in particular, notice that $a_\Gamma \leq b_\Gamma$. Then we define the persistence vector

$$\overline{\Gamma^{\mathbb{G}}} : \mathrm{I}(a_\Gamma, c_\Gamma) \to \frac{\mathbb{V}}{\mathbb{H}} ,$$

which is defined by $\overline{\Gamma^{\mathbb{G}}}(r) = p_{\mathbb{H}}(r) \circ \Gamma^{\mathbb{G}}(r)$ for all $r \in [a_\Gamma, c_\Gamma)$, where we use the projection $p_{\mathbb{H}} : \mathbb{V} \twoheadrightarrow \mathbb{V}/\mathbb{H}$. We claim that $\overline{\Gamma^{\mathbb{G}}}$ is well-defined, i.e. $\overline{\Gamma^{\mathbb{G}}}(r) \neq \overline{0}$ iff $r \in [a_\Gamma, c_\Gamma)$. First, notice that $\overline{\Gamma^{\mathbb{G}}}(c_\Gamma) = \overline{0}$ since by definition $\Gamma(c_\Gamma) = 0$, which implies $\Gamma^{\mathbb{G}}(c_\Gamma) = -\Gamma^{\mathbb{H}}(c_\Gamma)$. Next, we need to show that $\overline{\Gamma^{\mathbb{G}}}(r) \neq 0$ for all $r \in [a_\Gamma, c_\Gamma)$. In fact, we will prove the stronger statement that

$$\widetilde{\mathscr{Q}} = \left\{ \overline{\Gamma^{\mathbb{G}}} \;\middle|\; \Gamma \in \mathscr{I}[\mathscr{G}] \text{ such that } a_\Gamma < c_\Gamma \right\}$$

is linearly independent. Take a subset $\mathscr{S} \subseteq \mathscr{I}[\mathscr{G}]$ such that $\{\overline{\Gamma^{\mathbb{G}}}\}_{\Gamma\in\mathscr{S}}$ is a nonempty subset of $\widetilde{\mathscr{Q}}$; notice that this leads to any nonempty subset of $\widetilde{\mathscr{Q}}$. Also, take some coefficients $k_\Gamma \in \mathbb{F} \setminus \{0\}$ for all $\Gamma \in \mathscr{S}$. We want to show that $\overline{V^{\mathbb{G}}} := \coprod_{\Gamma\in\mathscr{S}} k_\Gamma \overline{\Gamma^{\mathbb{G}}}$ is associated to the interval $[A, C)$, where we use the notation $A = \max_{\Gamma\in\mathscr{S}}(a_\Gamma)$, $B = \max_{\Gamma\in\mathscr{S}}(b_\Gamma)$, and $C = \max_{\Gamma\in\mathscr{S}}(c_\Gamma)$. By contradiction,

suppose that $\overline{V^{\mathbb{G}}}$ is associated to $[A,r)$ for some value $r \in [A,C)$. This implies that $\mathbf{1}_r(V^{\mathbb{G}})$ is in PVect$(\mathbb{H})$ and in particular $\mathbf{1}_{\max(r,B)}(V)$ is in PVect$(\mathbb{H})$, where we define $V^{\mathbb{G}} \coloneqq \coprod_{\Gamma\in\mathscr{S}} k_\Gamma \Gamma^{\mathbb{G}}$ and $V \coloneqq \coprod_{\Gamma\in\mathscr{S}} k_\Gamma\Gamma$. We consider two possible cases:

- Assume that $B < C$. There exists $\mathscr{T} \subseteq \mathscr{I}[\mathscr{H}]$ together with $t_h \in \mathbb{F}\setminus\{0\}$ for all $h \in \mathscr{T}$ such that

$$\mathbf{1}_{\max(r,B)}\Big(\coprod_{h\in\mathscr{T}} t_h h\Big) = \mathbf{1}_{\max(r,B)}(V)$$

  However, this implies that $V \boxplus \big(-\coprod_{h\in\mathscr{T}} t_h h\big)\big)$ is associated to an interval whose endpoint is $\max(r,\ B) < C$; but this breaks linear independence of $\mathscr{I}$, as $\mathscr{S}\cup\mathscr{T} \subseteq \mathscr{I}$ and $C = \max_{\Gamma\in\mathscr{S}}(c_\Gamma)$.

- Assume that $B = C$, in this case $V = Z_B$ as $V \sim [B,C)$. Since we assume that the elements from $\widetilde{\mathscr{Q}}$ are nonzero, we must have $a_\Gamma < c_\Gamma$ for all $\Gamma \in \mathscr{S}$ and in particular $A < B = C$. This implies that $V^{\mathbb{H}} \coloneqq \coprod_{\Gamma\in\mathscr{S}} k_\Gamma\Gamma^{\mathbb{H}}$ is associated to an interval whose startpoint is B, as this is the case for $V = V^{\mathbb{G}} \boxplus V^{\mathbb{H}}$ but $V^{\mathbb{G}} \sim [A,x)$ for some $x \geq A$. Consequently, $V^{\mathbb{H}}$ is written in terms of $\mathscr{I}[\mathscr{H}]$ with a nontrivial coefficient on some element $h \in \mathscr{I}[\mathscr{H}]$ with $h \sim [B,y)$ for some $y \in \mathbf{R}$ with $B < y$. Recall that $\mathbf{1}_r(V^{\mathbb{G}}) \in$ PVect$(\mathbb{H})$. Then $V^{\mathbb{H}} \boxplus \mathbf{1}_r(V^{\mathbb{G}})$ is written in terms of $\mathscr{I}[\mathscr{H}]$ with h as a nontrivial summand, as it cannot be a summand from $\mathbf{1}_r(V^{\mathbb{G}})$, since $\mathbf{1}_r(V^{\mathbb{G}}) \sim [A,x)$ with $A < B$. By linear independence of $\mathscr{I}[\mathscr{H}]$, the sum $V^{\mathbb{H}} \boxplus \mathbf{1}_r(V^{\mathbb{G}})$ must be associated to an interval containing $[B,y)$, here recall that $r < C = B$. However, by hypotheses $V^{\mathbb{H}} \boxplus \mathbf{1}_r(V^{\mathbb{G}}) = V = Z_B$ and we reach a contradiction.

It can be shown that $\widetilde{\mathscr{Q}}$ generates $\mathbb{G}/\mathbb{H}$ by a similar reasoning as used in proposition 4.3.1. Consequently, $\widetilde{\mathscr{Q}}$ is a barcode base for the quotient.

4.5 Homology of Persistence Modules

Consider a chain of tame persistence modules:

$$0 \longleftarrow \mathbb{V}_0 \xleftarrow{d_1} \mathbb{V}_1 \xleftarrow{d_2} \mathbb{V}_2 \longleftarrow \cdots \xleftarrow{d_n} \mathbb{V}_n, \tag{4.2}$$

where each term has basis $\mathscr{B}_j$ for $0 \leq j \leq n$. Then applying `image_kernel` we will obtain bases $\mathscr{I}_{j-1}$ and $\mathscr{K}_j$ for the image and kernel of d_j for all $0 \leq j \leq n$. Proceeding as on the previous section, we send triples $((\mathscr{I}_j|\mathscr{K}_j), \mathscr{B}_j, ((\mathscr{I}_j)_{\mathscr{B}_j}|(\mathscr{K}_j)_{\mathscr{B}_j})$ to `box_gauss_reduce` for all $0 \leq j \leq n$. This leads to bases $\mathscr{Q}_j$ for the homology for all $0 \leq j \leq n$.

Example 4.5.1. Suppose that we have a filtered simplicial complex $(X_r)_{r\in\mathbf{R}}$ in such a way that $X_r \subset X_s$ for all $r \leq s$. Thus, taking free chains on the k-simplexes of these leads to tame persistence

modules $S_k^{\Delta}(X_*)$ for every positive integer k. This induces a sequence of tame persistence modules,

$$0 \longleftarrow S_0^{\Delta}(X_*) \xleftarrow{d_1} S_1^{\Delta}(X_*) \xleftarrow{d_2} S_2^{\Delta}(X_*) \xleftarrow{d_3} S_3^{\Delta}(X_*) \longleftarrow \cdots$$

with the property that barcodes 'never die'. This setting would be the classical case of persistence homology over a filtered space presented in [59]. In this context, we can use the *clear optimization*, as presented in section 3.2. From the point of view of barcode bases, this optimization is because of the composition rule $d \circ d = 0$ and also because of the fact that barcodes never die in the modules $S_*^{\Delta}(X_*)$. This property allows us to create a more efficient algorithm. Thus, performing the homology of the chain complex above we recover the persistent homology $\mathrm{PH}_n(X_*)$ for all degrees $n \geq 0$.

Remark. An interesting question arises here. Let $f : \mathbb{V} \to \mathbb{W}$ be a persistence morphism and choose a base $\mathscr{A}$ for $\mathbb{V}$. Supose that we have computed a base $\mathscr{I}$ for $\mathrm{Im}(f)$ in terms of $\mathscr{A}$. Can we obtain some information about a base $\mathscr{K}$ for $\mathrm{Ker}(f)$ from $\mathscr{I}$? The simple answer is no (this is sometimes known as *clear optimization* and when working with persistence module homology we will not be able to use it in general). In chapter 5 we will see in proposition 5.3.6 some conditions on the relations between the interval decomposition of an inclusion, and in particular this applies to $\mathrm{Im}(f) \hookrightarrow \mathrm{Ker}(f)$. Finding information of $\mathrm{Ker}(f)$ from $\mathrm{Im}(f)$ is an open question for future research. A possible direction is to use the fact that the pivots from the generators from $\mathrm{Ker}(f)$ are hinted by the obtained generators from $\mathrm{Im}(f)$.

Chapter 5

Barcode Shifts

In this chapter we will explore the impact of changing data on persistent homology. We will continue in the direction started by B. Stolz [110], although from a different perspective. The main result of this section will be the *barcode shift lemma*, which gives some conditions as to when global changes in barcodes can be measured by computing barcode changes in a neighborhood. The rest of this chapter will be geared towards the idea of obtaining a precise description of these changes. The introduced theory on the Mayer-Vietoris spectral sequence from section 2.11 will play a key role, while the barcode bases machinery developed in chapter 4 will be crucial.

5.1 Barcode Shift Lemma

Let K be a filtered simplicial complex and consider a small subset $V \subset K$; e.g. V could be the subset $\mathrm{St}(\sigma)$ for some simplex $\sigma \in K$. We would like to know by how much $\mathrm{PH}_*(K \setminus V)$ and $\mathrm{PH}_*(K)$ differ. Something one could do is to compute straight away $\mathrm{PH}_*(K)$ and $\mathrm{PH}_*(\overline{K \setminus V})$, comparing both results. The problem with this approach is that it defeats the whole purpose of complex reduction, where we want to avoid computing $\mathrm{PH}_*(K)$ as it is very expensive. Additionally, one might want to check this for different subsets V, but computing $\mathrm{PH}_*(\overline{K \setminus V})$ might become as expensive as computing $\mathrm{PH}_*(K)$. Instead, we will consider a subset $M \subset K$ which is a neighborhood for V. In this case, we would like to know the differences between $\mathrm{PH}_*(K)$ and $\mathrm{PH}_*(K \setminus V)$ relying uniquely on information from M.

We consider a sequence of filtered subsets of K

$$V \subseteq N \subseteq U \subseteq M \subseteq K.$$

such that M and N are subcomplexes of K. Then we have that $\overline{K \setminus U}$, $\overline{M \setminus V}$ and N form a cover for K; see figure 5.1 for an illustration for this covering. The following will be the main example

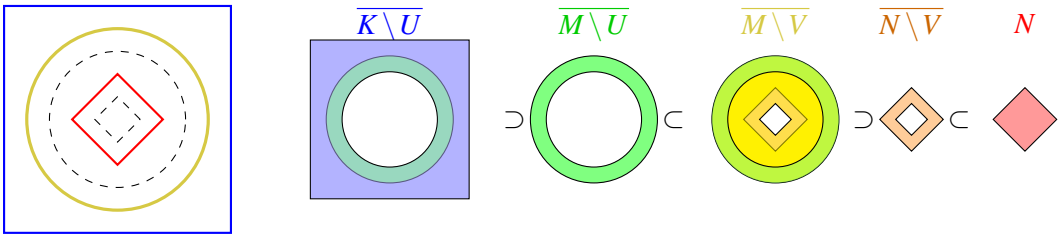


Figure 5.1: Illustration of a simplicial complex K covered by subcomplexes $\overline{K \setminus U}$, $\overline{M \setminus V}$ and N. Notice that there will be only two double intersections $\overline{M \setminus U}$ and $\overline{N \setminus V}$.

that we will have in mind.

Example 5.1.1. Recall the definition of star $\mathrm{St}(p)$ and closed star $\overline{\mathrm{St}(p)}$ of a vertex p from definition 3.4.1. Let $\mathbb{X} \subseteq \mathbb{R}^n$ be a finite point cloud, and consider the *truncated Vietoris Rips complex* $\mathrm{VR}^a_*(\mathbb{X})$ at $a > 0$, so that $\mathrm{VR}^a_r(\mathbb{X}) = \mathrm{VR}_r(\mathbb{X})$ for all $r \in [0,a)$ and it is undefined for $r \notin [0,a)$. The choice of the parameter $a > 0$ will depend on the maximum persistence value that we will perform our computations on. Then, taking a pair of radii $r, R \in \mathbb{R}$ such that $a \leq r < R$, and also $a \leq R - r$ we have the covering by filtered subcomplexes

$$\overline{\mathrm{VR}^a_*(\mathbb{X}) \setminus \mathrm{VR}^a_*(B_r(p) \cap \mathbb{X})} \qquad \mathrm{VR}^a_*(B_R(p) \cap \mathbb{X}) \setminus \mathrm{St}(p) \qquad \overline{\mathrm{St}(p)}\,.$$

In this case, $V = \mathrm{St}(p)$, $N = \overline{\mathrm{St}(p)}$, $U = \mathrm{VR}^a_*(B_r(p) \cap \mathbb{X})$ and $M = \mathrm{VR}^a_*(B_R(p) \cap \mathbb{X})$.

For ease of notation, we use the naming convention $A = \mathrm{PH}(\overline{K \setminus U})$, $B = \mathrm{PH}(\overline{M \setminus V})$ and $C = \mathrm{PH}(N)$, as well as $p_B \colon A \oplus B \oplus C \to B$ for the projection onto the second summand. On the other hand, we consider the morphisms

$$\delta_1 \colon \mathrm{PH}(\overline{M \setminus U}) \longrightarrow A \oplus B \oplus C \qquad \delta_2 \colon \mathrm{PH}(\overline{N \setminus V}) \longrightarrow A \oplus B \oplus C$$

which come from the chain maps

$$\begin{aligned} \delta_1 \colon \quad & S^\Delta_k(\overline{M \setminus U}) \longrightarrow S^\Delta_k(\overline{K \setminus U}) \oplus S^\Delta_k(\overline{M \setminus V}) \oplus S^\Delta_{k-1}(N) \\ & s \longrightarrow (s, -s, 0) \end{aligned}$$

and

$$\begin{aligned} \delta_2 \colon \quad & S^\Delta_k(\overline{N \setminus V}) \longrightarrow S^\Delta_k(\overline{K \setminus U}) \oplus S^\Delta_k(\overline{M \setminus V}) \oplus S^\Delta_{k-1}(N) \\ & s \longrightarrow (0, s, -s). \end{aligned}$$

These form the Čech differential $\delta = \delta_1 + \delta_2$. For ease of notation, from now on we will omit the closure overline sign, assuming that we are working with complexes.

Recall the definition of total complex from section 2.11. For a space K covered by $\{K \setminus U, M \setminus$

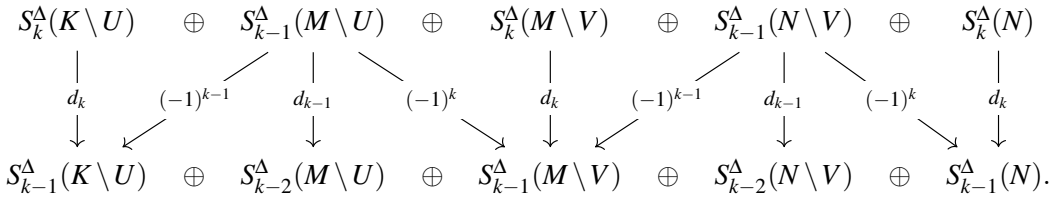
$V,N\}$, we define the total complex in degree $k \geq 0$ by

$$\mathscr{S}_k^{\mathrm{Tot}}(K) = S_k^{\Delta}(K \setminus U) \oplus S_k^{\Delta}(M \setminus V) \oplus S_k^{\Delta}(N) \oplus S_{k-1}^{\Delta}(M \setminus U) \oplus S_{k-1}^{\Delta}(N \setminus V) \ .$$

The differential of the total complex is given by $d^{\mathrm{Tot}} := d_k + (-1)^{k-1}\check{\delta}$, which sends $(a,b,c,\gamma_1,\gamma_2)$ to

$$(d_k(a) + (-1)^{k-1}\gamma_1,\ d_k(b) + (-1)^k\gamma_1 + (-1)^{k-1}\gamma_2,\ d_k(c) + (-1)^k\gamma_2,\ d_{k-1}(\gamma_1),\ d_{k-1}(\gamma_2)) \ ,$$

where we have used the definition of δ_1 and δ_2 above. One can easily check that $d^{\mathrm{Tot}} \circ d^{\mathrm{Tot}} = 0$. Such a differential can be summarized by the following diagram



Notice that, by convenience, here we wrote the summands in a different order from the definition of $\mathscr{S}_k^{\mathrm{Tot}}(K)$ above. Additionally, from section 3.7, recall that one has an isomorphism $\mathrm{PH}_*(\mathscr{S}_*^{\mathrm{Tot}}(K)) \cong \mathrm{PH}_*(K)$. On the other hand, we could also consider the total complex associated to the space $K \setminus V$ with the cover $\{K \setminus U, M \setminus V\}$ given by

$$\mathscr{S}_k^{\mathrm{Tot}}(K \setminus V) = S_k^{\Delta}(K \setminus U) \oplus S_k^{\Delta}(M \setminus V) \oplus S_{k-1}^{\Delta}(M \setminus U)$$

which satisfies $\mathrm{PH}_*(\mathscr{S}_k^{\mathrm{Tot}}(K \setminus V)) \cong \mathrm{PH}_*(K \setminus V)$. Notice that there is an inclusion of covers $\mathscr{V} \subset \mathscr{U}$, which induces a morphism $\phi_K \colon \mathscr{S}_k^{\mathrm{Tot}}(K \setminus V) \to \mathscr{S}_k^{\mathrm{Tot}}(K)$ by the assignment

$$(a,b,\gamma_1) \mapsto (a,b,0,\gamma_1,0).$$

This induces a morphism of persistence modules $\phi_K \colon \mathrm{PH}(K \setminus V) \to \mathrm{PH}(K)$.

In an analogous way we can define the total complex associated to the space M with the cover by two subcomplexes $\{M \setminus V, N\}$

$$\mathscr{S}_k^{\mathrm{Tot}}(M) = S_k^{\Delta}(M \setminus V) \oplus S_k^{\Delta}(N) \oplus S_{k-1}^{\Delta}(N \setminus V)$$

and also $\mathscr{S}_k^{\mathrm{Tot}}(M \setminus V) = S_k^{\Delta}(M \setminus V)$. These come with a chain complex morphism $\phi_M \colon \mathscr{S}_k^{\mathrm{Tot}}(M \setminus V) \to \mathscr{S}_k^{\mathrm{Tot}}(M)$ induced by the inclusion of complexes $M \setminus V \subseteq M$. On the other hand, the inclusions $\varphi \colon M \to K$ and $\psi \colon M \setminus V \to K \setminus V$ induce respective morphism in persistent homology.

Lemma 5.1.2 (Barcode Shift Lemma)**.** *Consider the following commutative square of persistence modules and persistence morphisms*

$$\begin{array}{ccc} \mathrm{PH}_k(M\setminus V) & \xrightarrow{\phi_M} & \mathrm{PH}_k(M) \\ \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \varphi} \\ \mathrm{PH}_k(K\setminus V) & \xrightarrow{\phi_K} & \mathrm{PH}_k(K). \end{array} \tag{5.1}$$

for any $k\geq 0$ in $\mathbb{Z}$. Then the induced morphisms $\varphi_{|\mathrm{Coker}}\colon \mathrm{Coker}(\phi_M)\to\mathrm{Coker}(\phi_K)$ and $\psi_{|\mathrm{Ker}}\colon \mathrm{Ker}(\phi_M)\to \mathrm{Ker}(\phi_K)$ are injective and surjective respectively. Furthermore, if we assume that

$$\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)=0,$$

then $\varphi_{|\mathrm{Coker}}$ and $\psi_{|\mathrm{Ker}}$ are isomorphisms.

Remark. Recall the Excision Theorem for homology that we briefly discussed on section 2.4. In [93] an Excision Theorem was obtained in the context of persistence. Thus, suppose that a persistence version is used for the excision isomorphism $\mathrm{H}_k(K\setminus V, M\setminus V)\simeq \mathrm{H}_k(K,M)$, then we can fit a pair of isomorphisms on the top and bottom of the commutative square (5.1). However, it is not clear how we could use this for proving the barcode shift lemma.

Proof. First recall that from section 3.7 we have isomorphisms of persistence modules

$$\mathrm{PH}_k(M)\cong \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(M)),\qquad \mathrm{PH}_k(K\setminus V)\cong \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K\setminus V))$$

and also

$$\mathrm{PH}_k(K)\cong \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K)).$$

Additionally, recall that the induced map $\varphi\colon \mathscr{S}_k^{\mathrm{Tot}}(M)\to\mathscr{S}_k^{\mathrm{Tot}}(K)$ respects the filtrations, and thus it induces $\varphi\colon \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(M))\to\mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K))$. Similarly, ψ induces a morphism between the respective persistence modules. Altogether, the square (5.1) is isomorphic to the middle square in the following commutative diagram

$$\begin{array}{ccccccc} \mathrm{Ker}(\phi_M) & \longrightarrow & \mathrm{PH}_k\left(S_*(M\setminus V)\right) & \xrightarrow{\phi_M} & \mathrm{PH}_k\left(\mathscr{S}_*^{\mathrm{Tot}}(M)\right) & \longrightarrow & \mathrm{Coker}(\phi_M) \\ \downarrow{\scriptstyle \psi_{|\mathrm{Ker}}} & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi_{|\mathrm{Coker}}} \\ \mathrm{Ker}(\phi_K) & \longrightarrow & \mathrm{PH}_k\left(\mathscr{S}_*^{\mathrm{Tot}}(K\setminus V)\right) & \xrightarrow{\phi_K} & \mathrm{PH}_k\left(\mathscr{S}_*^{\mathrm{Tot}}(K)\right) & \longrightarrow & \mathrm{Coker}(\phi_K) \end{array} \tag{5.2}$$

A comment on notation. Throughout this proof we will look at persistent homology classes for different sets. In particular, for a given cochain $(a,b,c,\gamma_1,\gamma_2)\in\mathscr{S}_k^{\mathrm{Tot}}(K)$, we will denote by $[a,b,c,\gamma_1,\gamma_2]$ the homology class in $\mathrm{PH}_k(K)$. Additionally, we will denote by $[a]_A$, $[b]_B$, $[c]_C$, $[\gamma_1]_1$

and $[\gamma_2]_2$ the respective homology classes in A, B, C, $\mathrm{PH}_*(M\setminus U)$ and $\mathrm{PH}_*(N\setminus V)$ respectively.

Surjectivity of $\psi_{|\mathrm{Ker}}$: Suppose that $z\in \mathrm{Ker}(\phi_K)_k$ and there is $(a,b,\gamma_1)\in \mathscr{S}_k^{\mathrm{Tot}}(K\setminus V)$, so that $z=[a,b,\gamma_1]$. By hypotheses, we have

$$\phi_K(z)=[a,b,0,\gamma_1,0]=0.$$

This implies that there exists a chain $(a',b',c',\beta_1,\beta_2)\in \mathscr{S}_{k+1}^{\mathrm{Tot}}(K)$ such that

$$\begin{aligned}0&=(a,b,0,\gamma_1,0)+d^{\mathrm{Tot}}(a',b',c',\beta_1,\beta_2)\\ &=(a,b,0,\gamma_1,0)+(d_{k+1}a'+(-1)^k\beta_1,d_{k+1}b'+(-1)^{k-1}\beta_1+(-1)^k\beta_2,d_{k+1}c'+(-1)^{k-1}\beta_2,d_k(\beta_1),d_k(\beta_1))\end{aligned}$$

In particular, one has that $[a+(-1)^k\beta_1]_A=0$, $[b+(-1)^{k-1}\beta_1+(-1)^k\beta_2]_B=0$, $[(-1)^{k-1}\beta_2]_C=0$, and also $[\gamma_1]_1=0$. Thus, by adding the Čech differential image $\check{\delta}(\beta_1)$, we get identities between classes from $\mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K\setminus V))$

$$\begin{aligned}[a,b,\gamma_1]&=[a,b,0]+d^{\mathrm{Tot}}([0,0,\beta_1])\\ &=[0,b+(-1)^{k-1}\beta_1,0]\\ &=\psi[b+(-1)^{k-1}\beta_1].\end{aligned}$$

It remains to be shown that $b+(-1)^{k-1}\beta_1$ is in $\mathrm{Ker}(\phi_M)$. This follows from the equalities

$$\begin{aligned}\phi_M[b+(-1)^{k-1}\beta_1]&=[b+(-1)^{k-1}\beta_1,0,0]+d^{\mathrm{Tot}}[0,0,\beta_2]\\ &=[b+(-1)^{k-1}\beta_1,0,0]+[(-1)^k\beta_2,(-1)^{k-1}\beta_2,0]\\ &=[b+(-1)^{k-1}\beta_1+(-1)^k\beta_2,(-1)^{k-1}\beta_2,0]\\ &=0\end{aligned}$$

since the B and C components vanish from our previous computations.

Injectivity of $\varphi_{|\mathrm{Coker}}$: Consider $0\neq q\in\mathrm{Coker}(\phi_M)$ and take a representative $(b,c,\gamma_2)\in\mathscr{S}_k^{\mathrm{Tot}}(M)$ such that $q=[b,c,\gamma_2]+\mathrm{Im}(\phi_M)$, and suppose $\varphi(q)\in\mathrm{Im}(\phi_K)$. This implies that there must exist $(a',b',\gamma_1')\in\mathscr{S}_k^{\mathrm{Tot}}(K\setminus V)$ so that

$$\varphi[b,c,\gamma_2]=[0,b,c,0,\gamma_2]=\phi_K([a',b',\gamma_1'])=[a',b',0,\gamma_1',0]$$

Consequently $[-a',b-b',c,-\gamma_1',\gamma_2]=0$ which in turn implies $[\gamma_1']_1=0$ and $[\gamma_2]_2=0$. This also implies that there exists $(\beta_1,\beta_2)\in S_k(M\setminus U)\oplus S_k(N\setminus V)$ so that $[-a'+(-1)^k\beta_1]_A=0$,

$[b-b'+(-1)^{k-1}\beta_1+(-1)^k\beta_2]_B=0$ and $[c+(-1)^{k-1}\beta_2]_C=0$. But then, in $\mathrm{PH}_k(M)$ one has

$$\begin{aligned} q=[b,c,\gamma_2] &= [b,c,0]+d^{\mathrm{Tot}}([0,0,\beta_2]) \\ &= [b+(-1)^k\beta_2,0,0]=\phi_M[b+(-1)^k\beta_2]\ , \end{aligned}$$

reaching a contradiction, since we assumed $0\neq q\in\mathrm{Coker}(\phi_M)$.

Now, assume that condition $\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)=0$ holds.

Injectivity of $\psi_{|\mathrm{Ker}}$: By contradiction, suppose that an element $0\neq z\in\mathrm{Ker}(\phi_M)$ is such that $\psi_{|\mathrm{Ker}}(z)=0$. Let $s\in S_k(M\setminus V)$ be a representative so that $[s]_B=z$. Since $z\in\mathrm{Ker}(\phi_M)$, one has that in $\mathrm{PH}_k(M)$

$$\phi_M(z)=\phi_M([s]_B)=[s,0,0]=0.$$

Consequently, there exists some $\beta_2\in S_{k+1}(N\setminus V)$ such that $[s+(-1)^k\beta_2]_A=0$ and $[(-1)^{k-1}\beta_2]_B=0$. This leads to $z\in\mathrm{Im}(\delta_2)$. On the other hand, since $\psi_{|\mathrm{Ker}}(z)=0$ one has that

$$\psi(z)=[0,s,0]=0$$

in $\mathrm{PH}_k(K\setminus V)$. Thus, there exists $\beta_1\in S_{k+1}(M\setminus U)$ such that $[(-1)^k\beta_1]_A=0$ and $[s+(-1)^{k-1}\beta_1]_B=0$. Which implies $z\in\mathrm{Im}(\delta_1)$, but then $\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)\neq 0$ contradicting our assumption.

Surjectivity of $\varphi_{|\mathrm{Coker}}$: Let $q\in\mathrm{Coker}(\phi_K)$ and let $(a,b,c,\gamma_1,\gamma_2)\in\mathscr{S}_k^{\mathrm{Tot}}(K)$ so that $q=[a,b,c,\gamma_1,\gamma_2]+\mathrm{Im}(\phi_K)$. Using that $d^{\mathrm{Tot}}(q)=0$ together with the equalities

$$\begin{aligned} d^{\mathrm{Tot}}q &= d^{\mathrm{Tot}}[a,b,c,\gamma_1,\gamma_2] \\ &= [d_ka+(-1)^k\gamma_1, d_kb+(-1)^{k-1}\gamma_1+(-1)^k\gamma_2, d_kc+(-1)^{k-1}\gamma_2, d_{k-1}\gamma_1, d_{k-1}\gamma_2] \\ &= [(-1)^k\gamma_1,(-1)^{k-1}\gamma_1+(-1)^k\gamma_2,(-1)^{k-1}\gamma_2,0,0] \end{aligned}$$

we obtain $[(-1)^k\gamma_1]_A=0$, $[(-1)^{k-1}\gamma_1+(-1)^k\gamma_2]_B=0$ and $[(-1)^{k-1}\gamma_2]_C=0$. Thus by the hypotheses that $\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)=0$, we deduce that $[\gamma_1]_B=[\gamma_2]_B=0$. Choosing a chain $\widetilde{b}_1\in S_k(M\setminus V)$ such that $d_k(\widetilde{b}_1)=\gamma_1$, one has that

$$d^{\mathrm{Tot}}(b+(-1)^{k-1}\widetilde{b}_1,c,\gamma_2)=(d_kb+(-1)^{k-1}\gamma_1+(-1)^k\gamma_2,c+(-1)^{k-1}\gamma_2,d_{k-1}\gamma_2)=(0,0,0)$$

and so $(b+(-1)^{k-1}\widetilde{b}_1,c,\gamma_2)\in\mathrm{Ker}(d_k^{\mathrm{Tot}}:\mathscr{S}_k^{\mathrm{Tot}}(K\setminus V)\to\mathscr{S}_{k-1}^{\mathrm{Tot}}(K\setminus V))$. On the other hand, one

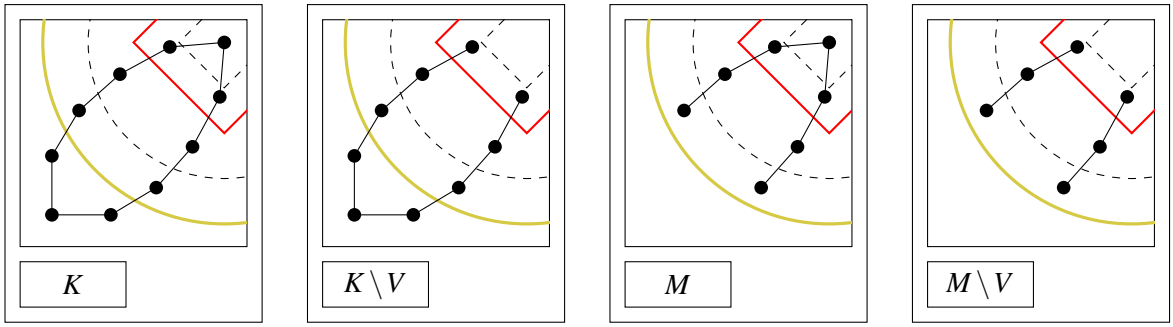


Figure 5.2: Example when $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \neq 0$. In this case one can see that in dimension 0, $\mathrm{Ker}(\phi_M)$ will contain a class that is not in $\mathrm{Ker}(\phi_K)$. As in figure 5.1, from the inner dashed rhombus to the outer circle, we denote the boundaries of the domains of the nested sets $V \subset N \subset U \subset M$.

can also check that $(a, (-1)^k \widetilde{b}_1, \gamma_1) \in \mathrm{Ker}(d_k^{\mathrm{Tot}} : \mathscr{S}_k^{\mathrm{Tot}}(M) \to \mathscr{S}_{k-1}^{\mathrm{Tot}}(M))$. Altogether we have that

$$\begin{aligned}\varphi[b + (-1)^{k-1}\widetilde{b}_1, c, \gamma_2] =& [0, b + (-1)^{k-1}\widetilde{b}_1, c, 0, \gamma_2] \\ =& [a, b, c, \gamma_1, \gamma_2] - \phi_K[a, (-1)^k \widetilde{b}_1, \gamma_1].\end{aligned}$$

Thus, $\varphi_{|\mathrm{Coker}}([b + (-1)^{k-1}\widetilde{b}_1, c, \gamma_2] + \mathrm{Im}(\phi_M)) = q + \mathrm{Im}(\phi_K)$, and the result follows. □

This lemma tells us how much barcodes will change whenever we remove some subset V from a complex K. In general we will compute:

$$p_B(\mathrm{Im}(\delta_1)) \cap \mathrm{Im}(\delta_2)$$

instead of $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2)$. This is because we want to check these conditions locally, without computing the whole persistent homology of K. Since $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2)$ is a persistence submodule of $p_B(\mathrm{Im}(\delta_1)) \cap \mathrm{Im}(\delta_2)$, this will give us an upper bound on the shifts of the barcode diagram. It is worth noticing that in dimension 0, one can have that $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) = 0$. Indeed, if this did not hold, then there exist $s_1 \in S_0(M \setminus U)$ and $s_2 \in S_0(N \setminus V)$ such that $[s_1]_A = 0$ and $[s_2]_C = 0$. This last statement does not need to hold in general.

Example 5.1.3. Consider diagram (5.2). In this case, we consider homology in dimension 0. One can see that the two points in $M \setminus U$ are equivalent to the pair of points in $N \setminus V$ up to boundaries in $M \setminus V$. At the same time, the previous pairs of points are zero up to boundaries in both $K \setminus U$ and N. Thus, one will have that $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \neq 0$, and a new 0-homology class will appear in $M \setminus V$, which does not appear in $K \setminus V$. Consequently $\mathrm{Ker}(\phi_M) \ncong \mathrm{Ker}(\phi_K)$, and the barcode shift lemma does not hold in this case, see figure 5.2 for an illustration.

Example 5.1.4. Consider a point cloud $\mathbb{X}$ together with a point $p \in \mathbb{X}$ as depicted on figure 5.3,

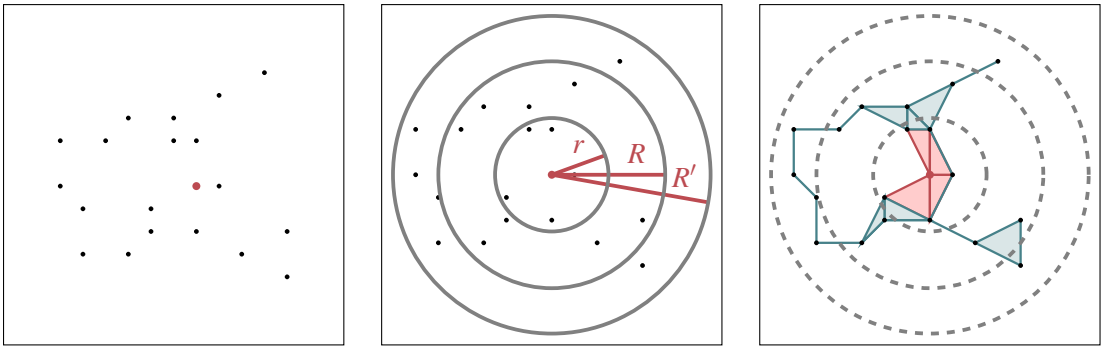
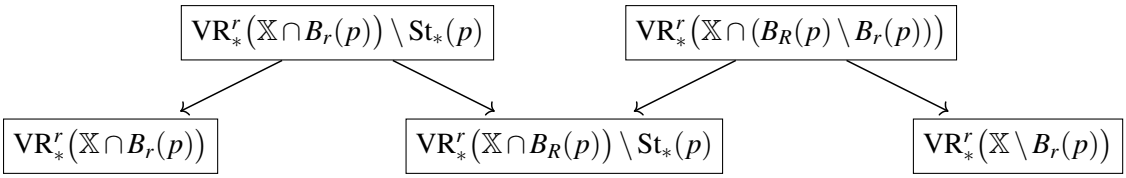


Figure 5.3: We consider removing the red point p from $\mathbb{X}$ (left). Depiction of circles with radii $r < R < R'$ around p (center). The complex $\mathrm{VR}^r_*(\mathbb{X})$ with the open star $\mathrm{St}(p)$ in red (right).

where p is marked in red. We consider the truncated Vietoris-Rips complex $\mathrm{VR}^r_*(\mathbb{X})$ together with the cover that we considered on example 5.1.1 with parameter $a = r$. Suppose that one wishes to take out the point p. For this, we take the maximum persistence value $r > 0$ together with an extra radius $R > r$ which constitutes the "local" region of $\mathrm{VR}^r_*(\mathbb{X})$ which we are willing to compute. This leads to the diagram of our cover



which we depict in figure 5.4. From example 5.1.1, recall that we must have $a \leq R - r$ in order to have a well-defined cover for $\mathrm{VR}^a_*(\mathbb{X})$; and so $2r \leq R$, as we choosed $a = r$. Notice that $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \simeq \mathrm{I}(s,t)$, where $s < t < r$ and t is the value at which we add the dashed red lines in figure 5.4. In this case, the morphism

$$\phi_M : \mathrm{PH}_*\big(\mathrm{VR}^r_*(\mathbb{X}\cap B_R(p))\setminus \mathrm{St}_*(p)\big) \longrightarrow \mathrm{PH}_*\big(\mathrm{VR}^r_*(\mathbb{X}\cap B_R(p))\big)\ ,$$

has a bar $\mathrm{I}(s,t)$ in $\mathrm{Ker}(\phi_M)$ for homology dimension 0, while such a bar does not appear in $\mathrm{Ker}(\phi_K)$, but instead it appears in $\mathrm{Coker}(\phi_K)$ for homology dimension 1. Of course, this is due to the fact that $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \neq 0$. To correct this, one might substitute R by a bigger radius. For instance, $R' > 0$ depicted on figure 5.3 solves the problem trivially as $\mathbb{X} \subseteq B_{R'}(p)$.

5.2 Stability to Local Changes

In this section, we will review a few stability results that can be easily deduced from the barcode shift lemma.

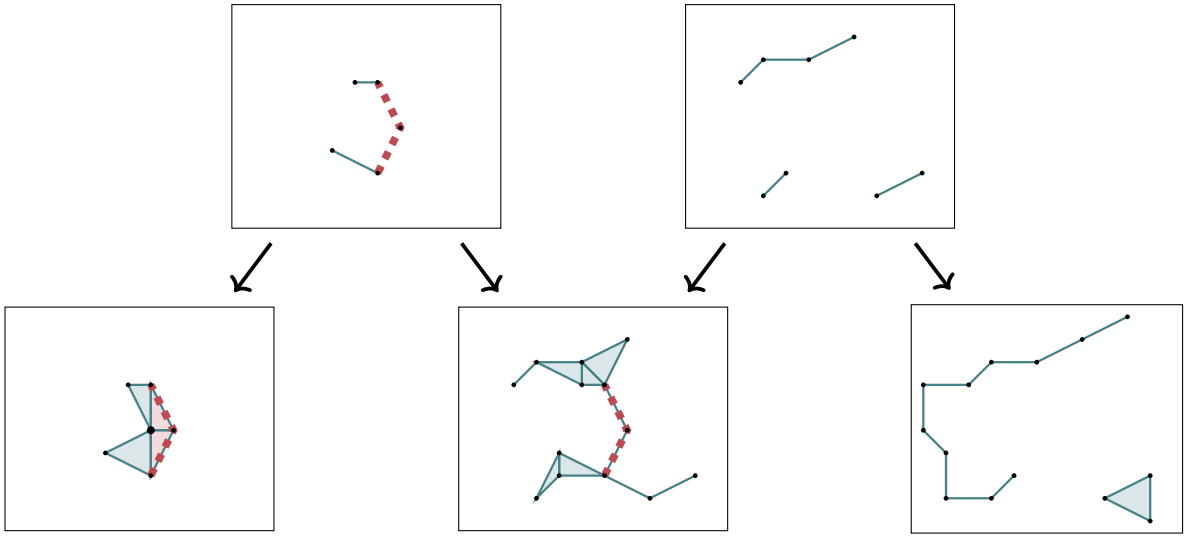


Figure 5.4: Depiction of the diagram of a cover for $\mathrm{VR}_*^r(\mathbb{X})$. In gray, we have drawn the complex associated to $\mathrm{VR}_*^r(\mathbb{X})$ just before the dashed red lines are added; in this case $\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)\neq 0$. Straight after the dashed lines are added, this intersection of images becomes zero.

Proposition 5.2.1. *We have equalities and isomorphisms*

$$\mathrm{Ker}(\psi_{|\mathrm{Ker}(\phi_M)}) = \mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2) \simeq \mathrm{Coker}(\varphi_{|\mathrm{Coker}(\phi_M)}) \ ,$$

in particular, we obtain an exact sequence

$$0 \longrightarrow \mathrm{Coker}(\phi_M) \xrightarrow{\varphi_{|\mathrm{Coker}}} \mathrm{Coker}(\phi_K) \longrightarrow \mathrm{Ker}(\phi_M) \xrightarrow{\psi_{|\mathrm{Ker}}} \mathrm{Ker}(\phi_K) \longrightarrow 0$$
$$\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2) \ .$$

Proof. The statement follows from the Barcode Shift Lemma. For the first equality, notice that we have that $\mathrm{Ker}(\psi_{|\mathrm{Ker}(\phi_M)}) \subseteq \mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)$ from the argument showing that $\psi_{|\mathrm{Ker}(\phi_M)}$ is injective within the proof of lemma 5.1.2. The other inclusion follows directly from the commutative diagram (5.2), as well as the description of $\mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(M))$ and $\mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K\setminus V))$ for all $k \geq 0$. Next, we proceed to show the isomorphism $\mathrm{Coker}(\varphi_{|\mathrm{Coker}}) \simeq \mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)$. Recall the observation that if $[a,b,c,\gamma_1,\gamma_2] \in \mathrm{PH}(K)$, then $[\gamma_1]_B = [\gamma_2]_B$ lies in $\mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2)$. Then, we define the morphism

$$\begin{aligned} f\colon \quad \mathrm{Coker}(\varphi_{|\mathrm{Coker}(\phi_M)}) &\longrightarrow \mathrm{Im}(\delta_1)\cap\mathrm{Im}(\delta_2) \\ [a,b,c,\gamma_1,\gamma_2]+\mathrm{Im}(\varphi)+\mathrm{Im}(\phi_K) &\longrightarrow [\gamma_1]_B \ , \end{aligned}$$

which is well defined; given a chain $[a,b,\gamma_1] \in \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(K\setminus V))$, we have that $[(-1)^{k-1}\gamma_1]_B = 0$ and so we obtain $f(\phi_K([a,b,\gamma_1])) = 0$, while if $[b,c,\gamma_2] \in \mathrm{PH}_k(\mathscr{S}_*^{\mathrm{Tot}}(M))$ then $[(-1)^k\gamma_2]_B = 0$ and

$f(\varphi([b,c,\gamma_2])) = 0$.

We claim that f is an isomorphism. First, given some class $\Gamma \in \mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2)$, there exist a pair of cycles $\alpha_1 \in S^{\Delta}_{k-1}(M \setminus U)$ and $\alpha_2 \in S^{\Delta}_{k-1}(N \setminus V)$ such that $\Gamma = [(-1)^{k-1}\alpha_1]_B = [(-1)^k \alpha_2]_B$ and also $[\alpha_1]_A = 0$ and $[\alpha_2]_C = 0$. Then, there must exist chains $a' \in S^{\Delta}_k(K \setminus U)$, $b' \in S^{\Delta}_k(M \setminus V)$ and $c' \in S^{\Delta}_k(N)$, such that

$$\begin{aligned} d_A(a') + (-1)^k \alpha_1 &= 0 \\ d_B(b') + (-1)^{k-1}\alpha_1 + (-1)^k \alpha_2 &= 0 \\ d_C(c') + (-1)^{k-1}\alpha_2 &= 0 \end{aligned}$$

and so $[a', b', c', \alpha_1, \alpha_2]$ is a class in $\mathrm{PH}_k(K)$. Thus, $f([a', b', c', \alpha_1, \alpha_2]) = \Gamma$, and so f is surjective. Now, assume that $f([a,b,c,\gamma_1,\gamma_2]) = [\gamma_1]_B = [0]_B$. Then, we obtain the equality

$$[a,b,c,\gamma_1,\gamma_2] = \phi_K[a,b,\gamma_1] + \varphi[0,c,\gamma_2]$$

and $[a,b,c,\gamma_1,\gamma_2]$ is trivial in $\mathrm{Coker}(\varphi_{|\mathrm{Coker}(\phi_M)})$. □

We will proceed to use the short exact sequence from proposition 5.2.1 to deduce a stability result for local changes. For now, we recall proposition 4.6 and 4.14 from [65]. Before we do this, we need to introduce left and right interleavings. For convenience, we adopt the notation $M \sim^{\varepsilon} N$ to mean that M and N are ε-interleaved modules.

Definition 5.2.2 (Left Interleaving)**.** Two modules M and N are ε-left interleaved, denoted by $M \sim^{\varepsilon}_L N$, whenever there is a short exact sequence $0 \to M \to N \to P \to 0$, where $P \sim^{\varepsilon} 0$.

Definition 5.2.3 (Right Interleaving)**.** Two modules M and N are ε-right interleaved, denoted by $M \sim^{\varepsilon}_R N$, whenever there is a short exact sequence $0 \to Q \to M \to N \to 0$, where $Q \sim^{\varepsilon} 0$.

Notice that both notions of left and right interleavings are not symmetric; i.e. $M \sim^{\varepsilon}_L N$ does not imply $N \sim^{\varepsilon}_L M$. For more properties, we refer the reader to section 4 from [65]. Here we recall two propositions from that manuscript.

Proposition 5.2.4 (4.6 from [65])**.** *Suppose we are given an exact sequence*

$$0 \to M \to N \to P \to 0$$

where $M \sim^{\varepsilon_1} 0$ *and* $P \sim^{\varepsilon_2} 0$. *Then* $N \sim^{\varepsilon_1+\varepsilon_2} 0$.

Proposition 5.2.5 (4.14 from [65], part 4.)**.** *If* $M \sim^{\varepsilon}_R P$ *and* $P \sim^{\varepsilon}_L N$. *Then* $M \sim^{\varepsilon} N$.

Additionally, we add the following simple result:

Proposition 5.2.6. *Given an exact sequence* $M \xrightarrow{f} N \longrightarrow 0$ *, and suppose* $M \sim^{\varepsilon} 0$*. Then* $N \sim^{\varepsilon} 0$*.*

Proof. Denote by Σ^{ε} the shift functor. This is proven by a similar argument to proposition 5.2.4. That is, we have to prove that $\Sigma^{2\varepsilon} N = 0$; i.e. for all persistence values $r \in \mathbf{R}$ the shift morphism $\Sigma^{2\varepsilon}_r N : N_r \to N_{r+2\varepsilon}$ is zero. For $n \in N_r$, by surjectivity of f there exists $m \in M_r$ such that $f(m) = n$. Thus, since f is a persistence morphism we have

$$\Sigma^{2\varepsilon} N(n) = \Sigma^{2\varepsilon} N(f(m)) = f(\Sigma^{2\varepsilon} M(m)) = f(0) = 0 \,,$$

and the result follows; since $M \sim^{\varepsilon} 0$ iff $\Sigma^{2\varepsilon} M = 0$. □

Now we can use this theory to prove a useful result from the barcode shift lemma.

Proposition 5.2.7. *Suppose that we are in the situation from lemma 5.1.2. Further, assume that* $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \sim^{\rho} 0$*, and* $\mathrm{Ker}(\phi_M) \sim^{\varepsilon_1} 0$ *and* $\mathrm{Coker}(\phi_M) \sim^{\varepsilon_2} 0$*. Then* $\mathrm{PH}(K \setminus V)$ *and* $\mathrm{PH}(K)$ *are* $\max(\varepsilon_1, \rho + \varepsilon_2))$*-interleaved.*

Proof. From proposition 5.2.1 we have exact sequences

$$\mathrm{Ker}(\phi_M) \to \mathrm{Ker}(\phi_K) \to 0 \,,$$

and

$$0 \to \mathrm{Coker}(\phi_M) \to \mathrm{Coker}(\phi_K) \to \mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \to 0 \,.$$

By hypothesis, using proposition 5.2.6 and 5.2.4 we obtain $\mathrm{Ker}(\phi_K) \sim^{\varepsilon_1} 0$ and $\mathrm{Coker}(\phi_K) \sim^{\varepsilon_1+\rho} 0$. Next, consider two short exact sequences with the obvious maps

$$0 \to \mathrm{Ker}(\phi_K) \to \mathrm{PH}(K \setminus V) \to \frac{\mathrm{PH}(K \setminus V)}{\mathrm{Ker}(\phi_K)} \to 0$$

and

$$0 \to \frac{\mathrm{PH}(K \setminus V)}{\mathrm{Ker}(\phi_K)} \to \mathrm{PH}(K) \to \mathrm{Coker}(\phi_K) \to 0 \,.$$

These imply that there are interleavings

$$\mathrm{PH}(K \setminus V) \sim^{\varepsilon_1}_{R} \frac{\mathrm{PH}(K \setminus V)}{\mathrm{Ker}(\phi_K)} \text{ and } \frac{\mathrm{PH}(K \setminus V)}{\mathrm{Ker}(\phi_K)} \sim^{\rho+\varepsilon_2}_{L} \mathrm{PH}(K)$$

and by proposition 5.2.5 we obtain the result $\mathrm{PH}(K \setminus V) \sim^{\max(\varepsilon_1, \rho+\varepsilon_2)} \mathrm{PH}(K)$. □

Notice that when applying proposition 5.2.7 one can use $p_B\big(\mathrm{Im}(\delta_1)\big) \cap \mathrm{Im}(\delta_2) \sim^{\rho} 0$ instead of the weaker, but harder to check condition $\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) \sim^{\rho} 0$.

Remark. The barcode shift lemma and proposition 5.2.7 can be replicated in the context of filtered regular CW-complexes X_* and cellular chains $C_*^{\mathrm{cell}}(X)$. Thus, it could also be applied to complexes such as the filtered cubical complex.

The barcode shift lemma, together with proposition 5.2.7 lead to a good comparison between $\mathrm{PH}_*(K)$ and $\mathrm{PH}_*(K \setminus V)$ from just checking local information. In particular, if we assumed the strong condition

$$\mathrm{Im}(\delta_1) \cap \mathrm{Im}(\delta_2) = \mathrm{Ker}(\phi_M) = \mathrm{Coker}(\phi_M) = 0$$

then $\mathrm{PH}_*(K) \cong \mathrm{PH}_*(K \setminus V)$. The problem with this is that the nullity condition above might be too restrictive. Instead, consider the case when $p_B(\mathrm{Im}(\delta_1)) \cap \mathrm{Im}(\delta_2) = 0$ while allowing both $\mathrm{Ker}(\phi_M)$ and $\mathrm{Coker}(\phi_M)$ to be nontrivial. Then, using the Barcode Shift Lemma, we have isomorphisms $\mathrm{Ker}(\phi_M) \cong \mathrm{Ker}(\phi_K)$ and $\mathrm{Coker}(\phi_M) \cong \mathrm{Coker}(\phi_K)$. Next we consider the following exact sequence

$$0 \longrightarrow \mathrm{Ker}(\phi_K) \longrightarrow \mathrm{PH}_*(K \setminus V) \longrightarrow \mathrm{PH}_*(K) \longrightarrow \mathrm{Coker}(\phi_K) \longrightarrow 0,,$$

which can be broken down into two short exact sequences

$$0 \longrightarrow \mathrm{Ker}(\phi_K) \longrightarrow \mathrm{PH}_*(K \setminus V) \longrightarrow \mathrm{Im}(\phi_K) \longrightarrow 0\,,$$

and

$$0 \longrightarrow \mathrm{Im}(\phi_K) \longrightarrow \mathrm{PH}_*(K) \longrightarrow \mathrm{Coker}(\phi_K) \longrightarrow 0\,.$$

Assuming that we already know a barcode decomposition for $\mathrm{PH}(K \setminus V)$, is it possible to obtain $\mathrm{PH}(K)$? More generally, consider a short exact sequence of persistence modules

$$0 \to M \hookrightarrow N \twoheadrightarrow P \to 0\,,$$

is it posible to obtain the barcode decomposition of one of the terms given the barcode decompositions for the other two? The next two sections will be dedicated to study this problem.

5.3 Embedding and projection of barcodes

Let us consider a short exact sequence of persistence modules

$$0 \longrightarrow M \xrightarrow{\ \iota\ } N \xrightarrow{\ q\ } P \longrightarrow 0\,. \tag{5.3}$$

In this section we will consider the *embedding problem*, which consists in the following: given $\mathscr{B}_M$ and $\mathscr{B}_N$ two barcode decompositions for M and N respectively, we would like to obtain a barcode decomposition for P without having any explicit knowledge of the morphism ι. Dually, we will also consider the *projection problem*, where we are given barcode decompositions for N and P and look for a barcode decomposition for M, without having any explicit knowledge of q.

We will start by revising proposition 6.1 from [8] which sheds some light on the considered problem; this result is best known as the *structure theorem of persistence submodules and quotients*. Here we give an alternative proof using the `box_gauss_reduce` Algorithm 4.1, and in particular the pivots from definition 4.2.1.

Proposition 5.3.1. *Let M and N be a pair of tame persistence modules, and suppose $\mathscr{A}$ and $\mathscr{B}$ are barcode bases for M and N respectively.*

(i) If $\iota : M \hookrightarrow N$ is an injection, then ι induces an injection $j_\iota : \mathscr{A} \to \mathscr{B}$ sending each barcode generator $\alpha \sim [a_\alpha, b_\alpha)$ to $j_\iota(\alpha) \sim [a_{j_\iota(\alpha)}, b_{j_\iota(\alpha)})$, with $a_\alpha \geq a_{j_\iota(\alpha)}$ and $b_\alpha = b_{j_\iota(\alpha)}$. In such a case, we say that j_ι is induced by ι.

(ii) If $q : M \twoheadrightarrow N$ is a surjection, then it induces an injection $i_q : \mathscr{B} \to \mathscr{A}$, sending each barcode generator $\beta \sim [a_\beta, b_\beta)$ to $i_q(\beta) \sim [a_{i_q(\beta)}, b_{i_q(\beta)})$ with $a_\beta = a_{i_q(\beta)}$ and $b_\beta \leq b_{i_q(\beta)}$. We say that i_q is induced by q.

Proof. (i) We consider the setup from Algorithm 4.2 in section 4.3 for $\iota : M \hookrightarrow N$. Eventually we obtain a reduced matrix

$$\mathscr{I} = \left(\begin{array}{c|c|c|c} \iota(\alpha_1) & \iota(\alpha_2) \boxplus k_{2,1}\iota(\alpha_1) & \cdots & \iota(\alpha_n) \boxplus \coprod_{i=1}^{n-1} k_{n,i}\iota(\alpha_i) \end{array} \right)$$

with unique pivots and whose columns $\widetilde{\mathscr{I}}$ form a basis for $\mathrm{Im}(\iota)$ by proposition 4.3.1. Suppose that each $\alpha_r \in \mathscr{A}$ is associated to $[a_{\alpha_r}, b_{\alpha_r})$. Then, the corresponding column $\mathscr{I}_r$ contains a pivot $j(\alpha_r) \in \mathscr{B}$ which is associated to $[a_{j(\alpha_r)}, b_{j(\alpha_r)})$ with $a_{j(\alpha_r)} \leq a_{\alpha_r}$. We claim that $b_{j(\alpha_r)} = b_{\alpha_r}$. Now, the image $\iota(\alpha_r)$ must have a pivot whose associated interval has b_{α_r} as an endpoint by injectivity and naturality of ι. Recall that the `box_gauss_reduce` algorithm leads to this column by adding previous columns sharing the same pivot, and so $b_{j(\alpha_r)} \leq b_{\alpha_r}$. On the other hand, the preimage $\mathscr{P}\mathscr{I}_r$ of $\mathscr{I}_r$ must be associated to an interval containing $[a_{\alpha_r}, b_{\alpha_r})$ by linear independence of $\mathscr{A}$. Since ι is injective, and since $\mathscr{I}_r = \iota(\mathscr{P}\mathscr{I}_r)$, $\mathscr{I}_r$ is associated to an interval whose endpoint is greater or equal to b_{α_r}. This proves the claim, as $\alpha_r \sim [a_{\alpha_r}, b_{\alpha_r})$ is sent to $j_\iota(\alpha_r) \sim [a_{j(\alpha_r)}, b_{j(\alpha_r)})$ with $b_{\alpha_r} \leq b_{j(\alpha_r)}$. Thus, this defines $j_\iota(\alpha_r)$ for all $0 \leq r \leq |\mathscr{A}|$, and since the pivots are unique, this defines an injection $j_\iota : \mathscr{A} \to \mathscr{B}$ with the desired properties.

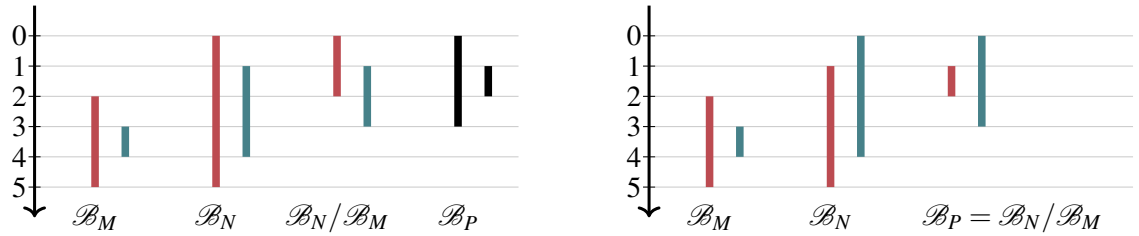
(ii) Again, we consider the matrix $\mathscr{I}$, although this time it is associated to $q : M \twoheadrightarrow N$; now $\mathscr{A}$ and $\mathscr{B}$ denote bases for M and N respectively. Recall that by proposition 4.3.1 the nonzero columns from $\mathscr{I}$ form a basis $\widetilde{\mathscr{I}}$ for N. Consequently, for any $\beta \in \mathscr{B}$ with $\beta \sim [a_\beta, b_\beta)$, we might find $S^\beta \subseteq \{1, 2, \ldots, |\mathscr{A}|\}$ together with coefficients $c_i^\beta \in \mathbb{F} \setminus \{0\}$ for all $i \in S^\beta$ such that $\beta = \mathbf{1}_{a_\beta}\left(\coprod_{i \in S^\beta} c_i^\beta \mathscr{I}_r\right)$. In particular, by proposition 4.2.3, β must be the pivot of some column $\mathscr{I}_s \sim [a_s, b_\beta)$. Also, notice that $a_s \geq a_\beta$ by naturality of q and on the other hand $a_s \leq a_\beta$ since we assume that $s \in S^\beta$; altogether $a_s = a_\beta$. Considering the particular generator $\alpha_s \in \mathscr{A}$, whose index in the order of $\mathscr{A}$ corresponds to the column index of $\mathscr{I}_s$, we define $i_q(\beta) = \alpha_s$. Also, notice that i_q must be injective since $\mathscr{I}$ has unique pivots. □

Let $\iota : M \hookrightarrow N$ be an injection together with a pair of barcode decompositions $\mathscr{B}_M$ and $\mathscr{B}_N$ for M and N respectively. Given two barcode bases $\mathscr{A}$ and $\mathscr{B}$ for M and N respectively, we can use the induced injection $j_\iota : \mathscr{A} \to \mathscr{B}$ from proposition 5.3.1. In turn, this induces an injection $j_\iota : \mathscr{B}_M \to \mathscr{B}_N$ sending each interval $[a,b) \in \mathscr{B}_M$ to some interval $[j_\iota(a), j_\iota(b))$ with $j_\iota(a) \leq a$ and $b = j_\iota(b)$. Unless we had some prior knowledge about the barcode bases $\mathscr{A}$ and $\mathscr{B}$ as well as the morphism ι, there might be multiple possibilities for j_ι and it is difficult to see which is the "best choice" from the barcode decompositions $\mathscr{B}_M$ and $\mathscr{B}_N$ alone. This is why we need to assume that barcode bases are *general*, in the sense that for each persistence value $r \in \mathbf{R}$, either a single bar is being born, a single bar is dying or nothing happens. For example, this holds whenever we consider a simplicial complex with a different filtration value for each simplex. In such a case, uniqueness of j_ι is guaranteed and we can obtain it directly from $\mathscr{B}_M$ and $\mathscr{B}_N$; all we need to do is examine the common endpoints. Now, going back to the embedding problem and sequence 5.3, one could try to 'quotient out' the intervals from $\mathscr{B}_N$ modulo those from $\mathscr{B}_M$, to obtain a barcode decomposition for P. The following definition describes this construction.

Definition 5.3.2. Consider two multisets of intevals $\mathscr{B}_M$ and $\mathscr{B}_N$ and assume there exists an injection $j : \mathscr{B}_M \to \mathscr{B}_N$ such that an interval $[a,b) \in \mathscr{B}_M$ is sent to some interval $[j(a), j(b))$ with $j(a) \leq a$ and $b = j(b)$. We define the *quotient barcode* denoted as $\mathscr{B}_N / \mathscr{B}_M$ as the interval decomposition computed by computing the 'quotient' of bars. That is

$$\mathscr{B}_N / \mathscr{B}_M \coloneqq \{[j(a), a) \mid [a,b) \in \mathscr{B}_M\} \cup (\mathscr{B}_N \setminus j(\mathscr{B}_M)) \,.$$

The natural question to ask is whether $\mathscr{B}_N / \mathscr{B}_M$ leads always to a barcode decomposition for P. As the following example illustrates, in general this is not the case.



(a) $\mathscr{B}_M$ does not embed strongly into $\mathscr{B}_N$ (b) $\mathscr{B}_M$ embeds strongly into $\mathscr{B}_N$

Figure 5.5: The persistence direction is from top to bottom, going through values 0 up to 5. Induced barcode assignments are indicated by the barcode coloring.

Example 5.3.3. Suppose that the short exact sequence (5.3) is realized as

$$0 \to I(2,5) \oplus I(3,4) \to I(0,5) \oplus I(1,4) \to I(0,3) \oplus I(1,2) \to 0\,, \tag{5.4}$$

where the interval decomposition corresponding to this exact sequence corresponds to the barcodes labelled $\mathscr{B}_M$, $\mathscr{B}_N$ and $\mathscr{B}_P$ on subfigure 5.5a. Denote by $\{\alpha_1,\alpha_2\}$, $\{\beta_1,\beta_2\}$ and $\{\gamma_1,\gamma_2\}$, the corresponding canonical bases from the terms in the short exact sequence (5.4). We define ι and p by the assignments:

$$\alpha_1 \mapsto \mathbf{1}_2(\beta_1 \boxplus (-\beta_2)) \qquad \alpha_2 \mapsto \mathbf{1}_3(\beta_2) \qquad \beta_1 \mapsto \gamma_1 \qquad \beta_2 \mapsto \gamma_1 \boxplus \gamma_2\,.$$

These lead to

$$p \circ \iota(\alpha_1) = p(\mathbf{1}_2(\beta_1 \boxplus (-\beta_2))) = \mathbf{1}_2(\gamma_1 \boxplus (-(\gamma_1 \boxplus \gamma_2))) = \mathbf{1}_2(\gamma_2) = Z_2\,,$$

since $\gamma_2 \sim [1,2)$, and also

$$p \circ \iota(\alpha_2) = p(\mathbf{1}_3(\beta_2)) = \mathbf{1}_3(\gamma_1 \boxplus \gamma_2) = Z_3$$

One might also check that ι and p are injective and surjective respectively, and so (5.4) is indeed a short exact sequence. Assume that we do not know the rightmost barcode decomposition $\mathscr{B}_P$. Then, we inspect the injection j_ι which we illustrate by using the same colors in figure 5.5a. However, by computing $\mathscr{B}_N/\mathscr{B}_M$ we obtain $\big\{[0,2),[1,3)\big\}$, which is different from the original decomposition $\mathscr{B}_P$. In general, unless we have some more information, this problem will have no solution. However, it seems that there are some cases when $\mathscr{B}_P$ can only be $\mathscr{B}_N/\mathscr{B}_M$. For example, we will show in proposition 5.3.6 that for a short exact sequence whose interval decomposition is that from figure 5.5b, this is indeed the case.

Consider a persistence module M together with bases $\mathscr{A}$. We, define a *change of basis* as a

matrix $T \in \mathscr{M}_{|\mathscr{A}|\times|\mathscr{A}|}(\mathbb{F})$ such that

- $T_{i,i} \neq 0_{\mathbb{F}}$ for all indices $1 \leq i \leq |\mathscr{A}|$, and
- if $T_{i,j} \neq 0$, then $a_i \leq a_j$ and $b_i \leq b_j$ for indices $1 \leq i, j \leq |\mathscr{A}|$ with $\alpha_i \sim [a_i, b_i)$ and $\alpha_j \sim [a_j, b_j)$.

In particular, notice that the set

$$\widetilde{\mathscr{A}} = \left\{ \widetilde{\alpha}_j \mid \widetilde{\alpha}_j = \coprod_{i \in S_j} T_{i,j}\alpha_i \text{ where } S_j = \left\{ i \;\middle|\; 1 \leq i \leq |\mathscr{A}| \text{ such that } T_{i,j} \neq 0 \right\} \text{ for all } 1 \leq j \leq |\mathscr{A}| \right\} ,$$

has unique pivots, and so by proposition 4.2.3, $\widetilde{\mathscr{A}}$ is linearly independent. In fact, one can see that $\widetilde{\mathscr{A}}$ defines a new barcode base for M, as $\mathscr{A}$ is a barcode base and each generator $\widetilde{\alpha}_i$ has α_i as a pivot and both share the same associated interval. In this case we will say that T defines a *change of basis* from $\widetilde{\mathscr{A}}$ to $\mathscr{A}$.

Lemma 5.3.4. *Consider a pair of barcode bases $\mathscr{A}$ and $\widetilde{\mathscr{A}}$ for M. There exists a change of basis from $\widetilde{\mathscr{A}}$ to $\mathscr{A}$.*

Proof. Denote by $\mathscr{B}_M$ the interval decomposition of M, and suppose that both $\mathscr{A}$ and $\widetilde{\mathscr{A}}$ are given by the isomorphisms:

$$\alpha : \bigoplus_{1\leq i\leq|\mathscr{B}_M|} \mathrm{I}(a_i, b_i) \to M \text{ and } \widetilde{\alpha} : \bigoplus_{1\leq i\leq|\mathscr{B}_M|} \mathrm{I}(\widetilde{a}_i, \widetilde{b}_i) \to M ,$$

where we index the generators over the ordered bars from $\mathscr{B}_M$. Then, we consider the composition $\alpha^{-1} \circ \widetilde{\alpha}$, which is an isomorphism, and denote by T its associated matrix; where we take the canonical base for $\bigoplus_{1\leq i\leq|\mathscr{B}_M|} \mathrm{I}(a_i, b_i)$. One can check that, up to reordering of the columns and rows associated to the same intervals, T must satisfy the conditions of a change of base, since otherwise $\alpha^{-1} \circ \widetilde{\alpha}$ could not possibly be an isomorphism. □

Within the context of lemma 5.3.4, notice that for each index $1 \leq i \leq |\mathscr{B}_M|$ we have

$$\widetilde{\alpha}_j = \coprod_{i \in S_j} T_{i,j}\alpha_j ,$$

where $S_j = \left\{ i \;\middle|\; 1 \leq i \leq |\mathscr{A}| \text{ such that } T_{i,j} \neq 0 \right\}$. Thus, given some coordinates $x = (x_i)_{1\leq i\leq|\mathscr{A}|}$ in terms of $\widetilde{\mathscr{A}}$, we can "turn them" to coordinates in terms of $\mathscr{A}$ by multiplying Tx^T; where x^T denotes the transpose of x. Here notice that one always needs to take care of which elements we add when actually performing computations with $\boxplus$.

Definition 5.3.5. Let $f : M \to N$ be a persistence morphism. We say that f admits a *strong normal form* whenever there exits a pair of barcode bases $\mathscr{A}$ and $\mathscr{B}$ such that the associated matrix of f has a unique nonzero entry for each row and column, and such an entry is equal to $1_{\mathbb{F}}$. Suppose that f admits a strong normal form, then we say that

(i) if f is an injection, then M *embeds strongly* into N, or

(ii) if f is a projection, then we say that M *projects strongly* onto N.

The choice of the word *strong* on definition 5.3.5 is due to the fact that these are strong conditions. Now, given the short exact sequence (5.3), if we know that M embeds strongly into N, then it is straightforward to compute a barcode basis for P. Indeed, there will exist barcode bases $\mathscr{A}$ for M and $\mathscr{B}$ for N such that the associated matrix F of ι is in strong normal form. This means that for each generator $\alpha \in \mathscr{A}$ with associated interval $[a_\alpha, b_\alpha)$, one has that $\iota(\alpha) = \mathbf{1}_{a_\alpha}(j_\iota(\alpha))$. Then the barcode decomposition for the quotient module $N/\iota(M)$ coincides with the quotient of barcodes $\mathscr{B}_N/\mathscr{B}_M$ from definition 5.3.2. The next proposition will give a straightforward way to check when there is a strong embedding or projection.

Consider a persistence morphism $f : M \to N$ together with bases $\mathscr{A}$ and $\mathscr{B}$ for M and N respectively as well as its associated matrix F. If f admits a strong normal form, by lemma 5.3.4 there exist a pair of matrices T_A and T_B such that $T_B F T_A$ is in strong normal form. Conversely, given a pair of such matrices T_A and T_B, these determine a pair of barcode bases $\widetilde{\mathscr{A}}$ and $\widetilde{\mathscr{B}}$ such that F is written in strong normal form.

Proposition 5.3.6. *Let M and N be two tame persistence modules with associated bases $\mathscr{A}$ and $\mathscr{B}$ respectively.*

(i) Let $\iota : M \hookrightarrow N$ be an injection, and let $j_\iota : \mathscr{A} \to \mathscr{B}$ be the injection induced by ι. Suppose that for every generator $\alpha \sim I_\alpha = [a_\alpha, b_\alpha)$ in $\mathscr{A}$, we have that

- *$\forall \gamma \in \mathscr{A} \setminus \{\alpha\}$ with $\gamma \sim I_\gamma$, if $I_\gamma \subset I_\alpha$ then either $I_{j_\iota(\gamma)} \not\subset I_{j_\iota(\alpha)}$ or $a_\alpha \notin I_{j_\iota(\gamma)}$, and*
- *$\forall \beta \in \mathscr{B} \setminus j_\iota(\mathscr{A})$ with $\beta \sim I_\beta$, we have that either $I_\beta \not\subset I_{j_\iota(\alpha)}$ or $a_\alpha \notin I_\beta$.*

Then M embeds strongly into N.

(ii) Dually, let $q : M \twoheadrightarrow N$ be a surjection, and consider the associated injection $i_q : \mathscr{B} \to \mathscr{A}$. Suppose that for every $\beta \sim I_\beta = [a_\beta, b_\beta)$ in $\mathscr{B}$, we have that

- *$\forall \gamma \in \mathscr{B} \setminus \{\beta\}$ with $\gamma \sim I_\gamma$, if $I_\gamma \subset I_\beta$ then either $I_{i_q(\gamma)} \not\subset I_{i_q(\beta)}$ or $b_\beta \notin I_{i_q(\gamma)}$, and*
- *$\forall \alpha \in \mathscr{A} \setminus i_q(\mathscr{B})$ with $\alpha \sim I_\alpha$, we have that either $I_\alpha \not\subset I_{i_q(\beta)}$ or $b_\beta \notin I_{i_q(\alpha)}$.*

Then M projects strongly onto N.

Proof. (i) Consider the matrix $\mathscr{I}$ resulting from algorithm 4.2. For each $1 \leq r \leq |\mathscr{A}|$, we consider the column

$$\mathscr{I}_r = \iota(\alpha_r) \boxplus \left(\bigboxplus_{i=1}^{r-1} k_{r,i} \iota(\alpha_i) \right) = \mathbf{1}_{a_r} \left(\bigboxplus_{s \in S_r} q_{r,s} \beta_s \right)$$

where $S_r \subseteq \{1, 2, \ldots, |\mathscr{B}|\}$ and $q_{r,s} \in \mathbb{F} \setminus \{0\}$ for all $s \in S_r$. For each column $\mathscr{I}_r$ we must have a pivot $\beta_{p(r)} \in \mathscr{B}$ with $p(r) \in S_r$ such that $j_\iota(\alpha_r) = \beta_{p(r)}$. Assuming that $\alpha_r \sim [a_r, b_r)$, $\bigboxplus_{s \in S_r} q_{r,s} \beta_s \sim [A_r, B_r)$ and $\beta_{p(r)} \sim [c_{p(r)}, d_{p(r)})$, there are inequalities $c_{p(r)} \leq A_r \leq a_r$. Also, since ι is injective, using proposition 5.3.1 we obtain $b_r = B_r = d_{p(r)}$. Next we consider two separate cases corresponding to whether $c_{p(r)} = A_r$ or $c_{p(r)} < A_r$.

- If $c_{p(r)} = A_r$ then we define $\tilde{\beta}_{p(r)} = \bigboxplus_{s \in S_r} q_{r,s} \beta_s$ and we have that $\tilde{\beta}_{p(r)}$ is associated to $[c_{p(r)}, d_{p(r)})$. Also, if we define $\tilde{\alpha}_r = \alpha_r \boxplus \left(\bigboxplus_{i=1}^{r-1} k_{r,i} \alpha_i \right)$ then $\tilde{\alpha}_r$ is associated to $[a_r, b_r)$. Thus, replacing α_r by $\tilde{\alpha}_r$ in $\mathscr{A}$, and replacing $\beta_{p(r)}$ by $\tilde{\beta}_{p(r)}$ in $\mathscr{B}$, as we keep the pivots on each expression we still have barcode bases. Furthermore, we have that $\mathbf{1}_{a_r}(\tilde{\beta}_{p(r)}) = \iota(\tilde{\alpha}_r)$.

- Now, assume that $c_{p(r)} < A_r$. Since $\mathscr{B}$ is finite, there exists $\beta_{s(r)} \in \mathscr{B}$ with $q_{r,s(r)} \neq 0$ and $\beta_{s(r)} \sim [A_r, d_{s(r)})$, and the endpoint $d_{s(r)}$ satisfies $a_r \leq d_{s(r)} \leq d_{p(r)} = b_r$. Consequently, $a_r \in [A_r, d_{s(r)}) \subseteq [c_{p(r)}, b_r)$ and $\beta_{s(r)} \in j_\iota(\mathscr{A})$ since otherwise we contradict our second hypotheses. Thus, there is some column $\mathscr{I}_{u(r)}$ with $1 \leq u(r) \leq |\mathscr{A}|$ and $u(r) \neq r$ and with pivot $j(\alpha_{u(r)}) = \beta_{s(r)}$. In this case we write

$$\mathscr{I}_{u(r)} = \mathbf{1}_{a_{u(r)}} \left(\bigboxplus_{s \in S_{u(r)}} q_{u(r),s} \beta_s \right)$$

  and the coefficient of the pivot satisfies $q_{u(r),s(r)} \neq 0$. This implies that $\alpha_{u(r)}$ will be associated to $[a_{u(r)}, d_{s(r)})$ with $A_r \leq a_{u(r)}$, as we know that $\beta_{s(r)} \sim [A_r, d_{s(r)})$. Additionally, we have that $a_{u(r)} < a_r$, since otherwise $[a_{u(r)}, d_{s(r)}) \subset [a_r, b_r)$, contradicting our first assumption. Thus, we have that $u(r) < r$, and we can redefine the r-column in $\mathscr{I}$ by the left to right column addition $\tilde{\mathscr{I}}_r = \mathscr{I}_r \boxplus (-q_{r,s(r)}/q_{u(r),s(r)}) \mathscr{I}_{u(r)}$. Notice that $\tilde{\mathscr{I}}_r \sim [a_r, b_r)$ and that its pivot will still be $\beta_{p(r)}$. Now, we look into the new expression $\tilde{\mathscr{I}}_r = \mathbf{1}_{a_r} \left(\bigboxplus_{s \in \tilde{S}_r} \tilde{q}_{r,s} \beta_s \right)$ and write $[\tilde{A}_r, b_r)$ for the associated interval of $\bigboxplus_{s \in \tilde{S}_r} \tilde{q}_{r,s} \beta_s$. We must have that $c_{p(r)} \leq \tilde{A}_r \leq A_r$, and we consider two cases depending on whether $c_{p(r)} = \tilde{A}_r$ or $c_{p(r)} < \tilde{A}_r$. By repeating the above procedure as necessary, since $\mathscr{B}$ is finite, eventually $c_{p(r)} = \tilde{A}_r$. This takes us back to the previous case.

Eventually, we have managed to perform a series of changes of basis such that the associated

matrix is in strong normal form, and the claim holds.

(ii) Let us now show the second claim. In fact, we will use duality so that the first result applies. We consider the matrix Q associated to the morphism $q : M \twoheadrightarrow N$ with respective barcode bases $\mathscr{A}$ and $\mathscr{B}$ for M and N. Recall that from corollary 4.1.11 the matrix $Q = \left(k_{\beta,\alpha}\right)_{\beta\in\mathscr{B},\alpha\in\mathscr{A}}$ is compatible, in the sense that whenever $k_{\beta,\alpha} \neq 0$, then we have inequalities $a_\beta \leq a_\alpha \leq b_\beta \leq b_\alpha$. We define the dual persistence modules

$$M^* = \bigoplus_{\alpha\in\mathscr{A}} \mathrm{I}(-b_\alpha, -a_\alpha) \qquad\qquad N^* = \bigoplus_{\beta\in\mathscr{B}} \mathrm{I}(-b_\beta, -a_\beta)$$

together with the canonical barcode bases $\mathscr{A}^*$ and $\mathscr{B}^*$ for M^* and N^* respectively. We consider the dual matrix $Q^* = \left(k_{\alpha^*,\beta^*}\right)_{\alpha^*\in\mathscr{A}^*,\beta^*\in\mathscr{B}^*}$; if we order $\mathscr{A}^*$ and $\mathscr{B}^*$ by choosing the reverse order of $\mathscr{A}$ and $\mathscr{B}$, then $Q^* = Q^T$. Then Q^* is well-defined in relation to $\mathscr{A}^*$ and $\mathscr{B}^*$, in the sense that whenever $k_{\alpha^*,\beta^*} \neq 0$, then we obtain the inequalities $-b_\alpha \leq -b_\beta \leq -a_\alpha \leq -a_\beta$; where recall that $\alpha^* \sim [-b_\alpha, -a_\alpha)$ and $\beta^* \sim [-b_\beta, -a_\beta)$. Thus, there exists a persistence morphism $q^* : N^* \to M^*$ by corollary 4.1.11. Next, we need to show that q^* is injective. By contradiction, assume that there exists a persistence value $r \in \mathbf{R}$ together with a nonzero vector $v \in N^*(r)$ such that $q^*(r)(v) = 0$. As N^* is tame, there exists $s \in \mathbf{R}$ such that $r < s$, the shifted image satisfies $\Sigma^{s-r}N^*(v) \neq 0$ and s is not an endpoint or startpoint of any of the associated intervals to the generators from $\mathscr{A}^*$ and $\mathscr{B}^*$. By naturality of q^*, we must have $q^*(s)\left(\Sigma^{s-r}N^*(v)\right) = \Sigma^{s-r}M^*(q^*(r)(v)) = 0$, implying that $q^*(s)$ is not injective. Notice that $q^*(s)$ is associated to the matrix $Q^*(s)$, which denotes the submatrix of Q^* whose rows and columns are active at value s; that is those from $\mathscr{A}^{*s}$ and $\mathscr{B}^{*s}$, see definition 4.1.1. Then the kernel of $Q^*(s)$ is nontrivial. On the other hand, by hypotheses $q(-s) : M(-s) \to N(-s)$ is a linear morphism whose associated matrix is a submatrix of Q which we denote by $Q(-s)$; this submatrix is obtained by taking the rows and columns of *active* barcode generators at $-s$. Thus, by standard linear algebra, the standard dual morphism $q(-s)^* : \hom(N(-s), \mathbb{F}) \to \hom(M(-s), \mathbb{F})$ is injective and its associated matrix $Q(-s)^T$ has a trivial kernel. By hypothesis, s is not on the extremes of any associated intervals from $\mathscr{A}^*$ and $\mathscr{B}^*$, and therefore $Q^*(s) = Q(-s)^T$. However, this leads to a contradiction, as the kernel of $Q^*(s)$ is nontrivial while the kernel of $Q(-s)^T$ has to be zero. Thus, q^* must be injective.

Now, notice that the barcode conditions from $\mathscr{A}$ and $\mathscr{B}$ in part (ii) translate to the needed conditions in $\mathscr{A}^*$ and $\mathscr{B}^*$ from part (i). Consequently, there exist a pair of basis changes T_U

and T_L such that $T_L Q^T T_U$ is in strong normal form. In particular, we have that the product

$$\left(T_L Q^T T_U\right)^T = T_U^T Q^{TT} T_L^T = T_U^T Q T_L^T \ ,$$

is in strong normal form. Now, notice that both T_U^T and T_L^T are well defined changes of bases on $\mathscr{B}$ and $\mathscr{A}$ respectively, and the result follows.

□

5.4 Glued and Entangled barcodes

We consider again the short exact sequence (5.3). Assuming that one has barcode decompositions $\mathscr{B}_M$ and $\mathscr{B}_P$ for M and P respectively, can we find a barcode decomposition for N? In order to solve this question, we will proceed very similarly to section 5.3, starting with the hypotheses that barcodes are tame and general. Given an interval $I = [a,b)$, we define the *start point* and *endpoint* as the respective singleton sets $s(I) = \{a\}$ and $e(I) = \{b\}$.

Definition 5.4.1. We define a *gluing* $G : \mathscr{B}_M^g \to \mathscr{B}_P^g$ between $\mathscr{B}_M$ and $\mathscr{B}_P$ to consist of:

- a pair of subsets $\mathscr{B}_M^b \subseteq \mathscr{B}_M$ and $\mathscr{B}_P^b \subseteq \mathscr{B}_P$, together with a bijection $B : \mathscr{B}_M^b \to \mathscr{B}_P^b$ such that each interval $I \in \mathscr{B}_M^b$ is sent to an interval $B(I) \in \mathscr{B}_P^b$, with $s(I) = e(B(I))$,
- a pair of sets defined by $\mathscr{B}_M^g = \mathscr{B}_M \cup e(\mathscr{B}_P \setminus \mathscr{B}_P^b)$ and $\mathscr{B}_P^g = \mathscr{B}_P \cup s(\mathscr{B}_M \setminus \mathscr{B}_M^b)$,
- a *gluing morphism* $G : \mathscr{B}_M^g \to \mathscr{B}_P^g$ sending each $I \in \mathscr{B}_M^g$ to

$$G(I) = \begin{cases} B(I) \ \text{ if } \ I \in \mathscr{B}_M^b, \\ s(I) \ \text{ if } \ I \in \mathscr{B}_M \setminus \mathscr{B}_M^b, \\ J \text{ if } I = e(J) \text{ for } J \in \mathscr{B}_P \setminus \mathscr{B}_P^b \ . \end{cases}$$

Of course, G defines a bijection.

Using generality of barcodes, we can find the "canonical" gluing $G : \mathscr{B}_M^g \to \mathscr{B}_P^g$ by defining $\mathscr{B}_M^b$ to be the set of intervals $[b,c)$ from $\mathscr{B}_M$ such that there exists an interval $[a,b)$ in $\mathscr{B}_P$. This correspondence leads to a uniquely determined bijection $B : \mathscr{B}_M^b \to \mathscr{B}_P^b$. The other elements from $G : \mathscr{B}_M^g \to \mathscr{B}_P^g$ are uniquely determined by definition 5.4.1.

Definition 5.4.2 (Join of barcodes)**.** Suppose $\mathscr{B}_M$ and $\mathscr{B}_P$ are two barcode decompositions, together with a gluing G: $\mathscr{B}_M^g \to \mathscr{B}_P^g$. Then, we define the *G-join* of $\mathscr{B}_M$ and $\mathscr{B}_P$ by setting

$$\mathscr{B}_M \vee_G \mathscr{B}_P := \left\{ I \cup G(I) \mid I \in \mathscr{B}_M^b \right\} \cup (\mathscr{B}_M \setminus \mathscr{B}_M^b) \cup (\mathscr{B}_P \setminus \mathscr{B}_P^b).$$

Notice that each element in the G-join is a right-open interval.

Given the canonical gluing $G\colon \mathscr{B}_M^g \to \mathscr{B}_P^g$, does $\mathscr{B}_M \vee_G \mathscr{B}_P$ lead to a barcode decomposition for N? As in section 5.3 we will need to make some assumptions about the known barcode decompositions $\mathscr{B}_M$ and $\mathscr{B}_P$.

Definition 5.4.3. Let $G\colon \mathscr{B}_M^g \to \mathscr{B}_P^g$ be a gluing. We say that $\mathscr{B}_M^g$ and $\mathscr{B}_P^g$ are *G-entangled* whenever there exists a pair of different elements I and J in $\mathscr{B}_M^g$, such that $I \subseteq J$ and $G(J) \subseteq G(I)$.

The following example illustrates G-entanglement.

Example 5.4.4. Consider a short exact sequence of $\mathbb{Z}_2$ persistence modules

$$0 \longrightarrow \mathrm{I}(1,5) \oplus \mathrm{I}(3,4) \longrightarrow \mathrm{I}(0,5) \oplus \mathrm{I}(1,4) \longrightarrow \mathrm{I}(0,3) \longrightarrow 0$$

where we consider the canonical bases, together with the associated matrices

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

One might check that the interval decomposition of the aforementioned short exact sequence corresponds to the barcodes $\mathscr{B}_M$, $\mathscr{B}_N$ and $\mathscr{B}_P$ depicted on figure 5.6a. Then, we define canonical gluing sets as $\mathscr{B}_M^g = \{[1,5),[3,4)\}$ and $\mathscr{B}_P^g = \{[0,3),\{1\}\}$ since $s([1,5)) = \{1\}$, and the canonical gluing $G\colon \mathscr{B}_M^g \to \mathscr{B}_P^g$ is given by the assignments

$$[1,5) \mapsto \{1\} \text{ and } [3,4) \mapsto [0,3)\ .$$

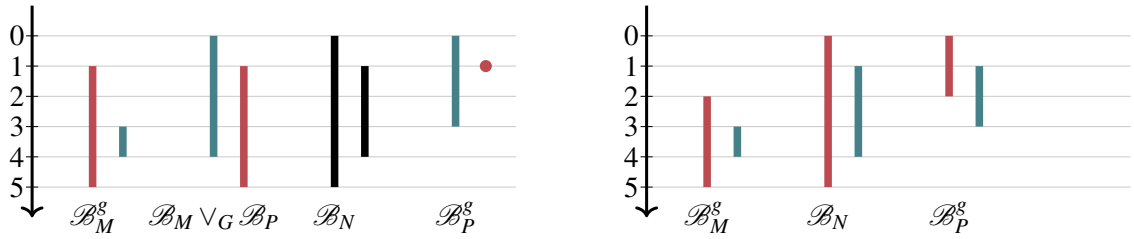
Then we have that $\mathscr{B}_M^g$ and $\mathscr{B}_P^g$ are G-entangled. In this case the G-join is $\mathscr{B}_M \vee_G \mathscr{B}_P = \{[0,4),[1,5)\}$ while we have that $\mathscr{B}_N = \{[0,5),[1,4)\}$. This is depicted on figure 5.6a.

On the other hand, we consider the case presented in figure 5.6b. Here we have that $\mathscr{B}_M^g = \{[2,5),[3,4)\}$ and $\mathscr{B}_P^g = \{[0,2),[1,3]\}$ with the canonical gluing G defined by the assignment

$$[2,5) \mapsto [0,2) \quad \text{and} \quad [3,4) \mapsto [1,3).$$

In this case $\mathscr{B}_M$ and $\mathscr{B}_P$ are not G-entangled. We will see soon that $\mathscr{B}_M \vee_G \mathscr{B}_P$ is the only possible interval decomposition for the middle term.

We will now proceed to show that non-G-entangled pairs of barcodes lead to barcode decompositions of the missing term through the G-join construction. The key observation is to study Algorithm 4.2 from section 4.3 as an inverse problem. That is, we see what barcode decomposition the missing term needs to have in order to be compatible with the others.



(a) $\mathscr{B}^g_M$ is G-entangled with $\mathscr{B}^g_P$ (b) $\mathscr{B}^g_M$ is not G-entangled with $\mathscr{B}^g_P$

Figure 5.6: Illustration for G-entanglement of barcodes, where G is the canonical gluing.

Proposition 5.4.5. *Let a short exact sequence of tame persistent modules*

$$0 \longrightarrow M \xrightarrow{\iota} N \xrightarrow{q} P \longrightarrow 0\,,$$

where M, N and P decompose into the barcodes $\mathscr{B}_M$, $\mathscr{B}_N$ and $\mathscr{B}_P$. Suppose that $\mathscr{B}_M$, $\mathscr{B}_N$ and $\mathscr{B}_P$ are all general, in the sense that no two bars coincide at any endpoints, and take the canonical gluing $G\colon \mathscr{B}^g_M \to \mathscr{B}^g_P$. If $\mathscr{B}^g_M$ and $\mathscr{B}^g_P$ are not G-entangled, then the join $\mathscr{B}_M \vee_G \mathscr{B}_P$ is an interval decomposition for N.

Proof. We consider a pair of barcode bases $\mathscr{B}$ and $\mathscr{C}$ for N and P respectively, and assume that these are ordered. Then we take the matrix $q(\mathscr{B})$ with the i-column being $q(\beta_i)$ in terms of $\mathscr{C}$. Next, we reduce $q(\mathscr{B})$ by following the `box_gauss_reduce` algorithm, denoting the resulting reduced matrix by $\mathscr{I}$. In particular, as q is surjective, one might take the set of columns $\{\mathscr{I}_i\}_{i=1}^n$ as a new barcode basis for P. Suppose that the r^{th} generator $\beta_r \in \mathscr{B}$ is associated to the interval $[a_r, b_r)$ while the r^{th} column $\mathscr{I}_r$ is associated to the interval $[a_r, c_r)$; notice that $c_r \leq b_r$ by naturality of q. Meanwhile, we also pay attention to the preimages $\mathscr{P}\mathscr{I}_i \in N$ for all $1 \leq i \leq n$, which are given by the corresponding linear combinations of generators from $\mathscr{B}$, so that

$$\mathscr{I} = (\mathscr{I}_1 \mid \mathscr{I}_2 \mid \cdots \mid \mathscr{I}_n) = (q(\mathscr{P}\mathscr{I}_1) \mid q(\mathscr{P}\mathscr{I}_2) \mid \cdots \mid q(\mathscr{P}\mathscr{I}_n))\ .$$

Assuming that the corresponding preimage $\mathscr{P}\mathscr{I}_r \in N$ is associated to $[a_r, B_r)$, we have the inequality $b_r \leq B_r$, as β_r is a nontrivial summand from $\mathscr{P}\mathscr{I}_r$. Now, take indices $i_1, i_2, \ldots, i_n$ so that the intervals on the set $\{[c_{i_j}, B_{i_j})\}_{i=1}^n$ are ordered in the standard order. As done in `image_kernel`, we consider the matrix $\mathscr{G}\mathscr{K}$ given by appending the columns

$$\mathscr{G}\mathscr{K} = \left(\mathbf{1}_{c_{i_1}}(\mathscr{P}\mathscr{I}_{i_1}) | \mathbf{1}_{c_{i_2}}(\mathscr{P}\mathscr{I}_{i_2}) | \cdots | \mathbf{1}_{c_{i_n}}(\mathscr{P}\mathscr{I}_{i_n})\right)\,,$$

and we compute its reduced form

$$\mathscr{K} = \left(\mathbf{1}_{c_{i_1}}(\mathscr{P}\mathscr{I}_{i_1}) \mid \cdots \mid \mathbf{1}_{c_{i_r}}(\mathscr{P}\mathscr{I}_{i_r}) \boxplus \left(\coprod_{l=1}^{r-1} k_{r,l}\mathbf{1}_{c_{i_l}}(\mathscr{P}\mathscr{I}_{i_l}) \right) \mid \cdots \right),$$

as done in the `image_kernel` procedure. This leads to a barcode basis for $\mathrm{Ker}(q) \simeq M$. Now, the s-column from $\mathscr{K}$, $\mathscr{K}_s$, has associated interval $[c_{i_s}, d_{i_s})$, where we must have that $d_{i_s} \leq B_{i_s}$. Therefore, $\mathscr{K}_s \sim [c_{i_s}, d_{i_s})$ is born just when $\mathscr{I}_{i_s} \sim [a_{i_s}, c_{i_s})$ dies and we can use this correspondence to define a gluing $\widetilde{G} : \mathscr{B}_M^{\widetilde{g}} \to \mathscr{B}_P^{\widetilde{g}}$; where the endpoint and startpoint cases correspond to whenever $c_{i_s} = d_{i_s}$ or when $a_{i_s} = c_{i_s}$.

Notice that in principle $\widetilde{G}$ does not need to be equal to the canonical gluing G. By contradiction, consider an interval $[c_{i_s}, d_{i_s}) \in \mathscr{B}_M$ and assume that $G([c_{i_s}, d_{i_s})) \neq \widetilde{G}([c_{i_s}, d_{i_s}))$. By generality of $\mathscr{B}_P$, the only option is that there exists an interval $[a_{i_t}, c_{i_s}) \in \mathscr{B}_P$ with $i_t \neq i_s$ and $a_{i_t} < c_{i_s}$, such that $G([c_{i_s}, d_{i_s})) = [a_{i_t}, c_{i_s})$ while $\widetilde{G}([c_{i_s}, d_{i_s})) = \{c_{i_s}\}$ and also $\widetilde{G}(\{c_{i_s}\}) = [a_{i_t}, c_{i_s})$. Thus, on the index i_s we have a generator $\beta_{i_s} \in \mathscr{B}$ associated to an interval with startpoint c_{i_s}; and there must exist a bar in $\mathscr{B}_N$ with startpoint c_{i_s}. On the other hand, we pay attention to the column $\mathscr{I}_{i_t} \sim [a_{i_t}, c_{i_s})$ and consider its preimage $\mathscr{P}\mathscr{I}_{i_t}$. By assumption, $\mathscr{K}_t = Z_{c_{i_s}}$ which implies that

$$\mathbf{1}_{c_{i_s}}(\mathscr{P}\mathscr{I}_{i_t}) = \mathbf{1}_{c_{i_s}} \left(\coprod_{r \in \mathscr{S}} k_r \mathscr{K}_r \right)$$

for some subset $\mathscr{S} \subseteq \{1, 2, \ldots, t-1\}$ and some coefficients $k_r \in \mathbb{F} \setminus \{0\}$. Notice that the sum $\coprod_{r\in\mathscr{S}} k_r\mathscr{K}_r$ must have a startpoint C such that $C < c_{i_s}$ by generality of the barcode decomposition from M. This implies that

$$\Gamma_{i_t} = \mathscr{P}\mathscr{I}_{i_t} - \coprod_{r \in \mathscr{S}} k_r \mathscr{K}_r$$

is associated to the nontrivial interval $[C, c_{i_s})$. Also, by hypotheses $p(\Gamma_{i_t}) = \mathbf{1}_{a_{i_t}}(\mathscr{I}_{i_t})$ and thus the pivot from Γ_{i_t} in the basis $\mathscr{B}$ must have endpoint c_{i_s}; and so an interval from $\mathscr{B}_N$ must have c_{i_s} as an endpoint. However, by hypothesis we assumed that the barcode decomposition of N must be general, and so we reach a contradiction. Thus, G and $\widetilde{G}$ must coincide.

Let us show that $\mathscr{B}_M \vee_G \mathscr{B}_P$ is a barcode decomposition for N. This is equivalent to showing that $[a_{i_s}, b_{i_s})$ is contained in $\mathscr{B}_M \vee_G \mathscr{B}_P$ for all $1 \leq s \leq n$. If $d_{i_s} = b_{i_s}$, then since $[a_{i_s}, c_{i_s}) \cup [c_{i_s}, d_{i_s}) = [a_{i_s}, b_{i_s})$, the claim follows. Otherwise we have two cases: either $d_{i_s} < b_{i_s}$ or $d_{i_s} > b_{i_s}$.

- Suppose $d_{i_s} < b_{i_s}$. Recall that before reducing the matrix $\mathscr{G}\mathscr{K} = \left(\mathbf{1}_{c_{i_r}}(\mathscr{P}\mathscr{I}_{i_r})\right)_{r=1}^{n}$ the column i_s had β_{i_s} as a pivot. Therefore, there must exist a minumum index $t < s$, such that $\mathscr{G}\mathscr{K}_t$ and also $\mathscr{K}_t$ contain a nonzero coefficient of β_{i_s}, which implies that $b_{i_s} \leq d_{i_t}$. As $t < s$, we have $c_{i_t} \leq c_{i_s}$ and thus $[c_{i_s}, d_{i_s}) \subseteq [c_{i_t}, d_{i_t})$, see figure 5.7. On the other hand, recall that

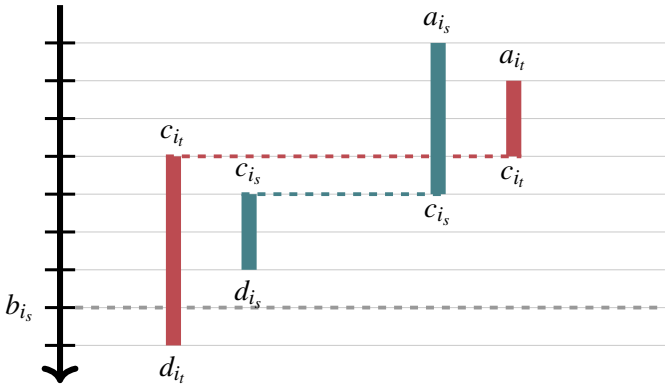


Figure 5.7: If $d_{i_s} < b_{i_s}$ then $\mathscr{B}_M$ and $\mathscr{B}_P$ are G-entangled.

$G([c_{i_s}, d_{i_s})) = [a_{i_s}, c_{i_s})$ and also $G([c_{i_t}, d_{i_t})) = [a_{i_t}, c_{i_t})$. As β_{i_t} must be a nonzero summand in $\mathscr{G}\mathscr{K}_t$, we have that it is a summand from $\mathscr{P}\mathscr{I}_{i_t}$ and so $a_{i_s} \leq a_{i_t}$, since $\mathscr{P}\mathscr{I}_{i_t}$ was obtained by adding columns from left to right in $q(\mathscr{B})$. Consequently $[a_{i_t}, c_{i_t}) \subseteq [a_{i_s}, c_{i_s})$, reaching a contradiction, since then $\mathscr{B}_M$ and $\mathscr{B}_P$ are G-entangled, see again figure 5.7. Notice that if $\mathscr{K}_{i_s}$ is zero, then we consider the interval $[a_{i_s}, c_{i_s})$ together with its endpoint $\{c_{i_s}\}$; the same argument can be applied to find $t < s$ such that $\{c_{i_s}\} \subseteq [c_{i_t}, d_{i_t})$ and $[a_{i_t}, c_{i_t}) \subseteq [a_{i_s}, c_{i_s})$. On the other hand, if $\mathscr{I}_{i_s}$ is zero, then we consider $[c_{i_s}, d_{i_s})$ together with its startpoint $\{c_{i_s}\}$; we reach a contradiction by following a similar argument as above.

- Now, assume instead $b_{i_s} < d_{i_s}$. Then $\mathscr{K}_{i_s}$ has a pivot β_{q_1} with associated interval $[a_{q_1}, b_{q_1})$ and with $b_{q_1} = d_{i_s}$. Then we consider the column $\mathscr{K}_{q_1}$, whose associated interval will be $[c_{q_1}, d_{q_1})$. Now if $d_{q_1} < b_{q_1}$, we are back in the previous case and reach a contradiction. Hence we must have $d_{q_1} > b_{q_1} = d_{i_s}$, as by generality of the barcode decomposition $\mathscr{B}_M$ we cannot have $d_{q_1} = b_{q_1} = d_{i_s}$. Thus, we consider the pivot $\beta_{q_2} \sim [a_{q_2}, b_{q_2})$ with $b_{q_2} = d_{q_1}$. In turn, this leads to a column $\mathscr{K}_{q_1} \sim [c_{q_2}, d_{q_2})$ and look for the pivot $\beta_{q_2} \sim [a_{q_2}, d_{q_2})$. Repeating the same argument we did with β_{q_1} successively, we find an infinite series $i_s, q_1, q_2, \ldots$ of different indices, as we have that $d_{i_s} < d_{q_1} < d_{q_2} < \cdots$, from the finite set $\{1, 2, \ldots, n\}$, reaching therefore a contradiction.

Thus, $b_{i_s} = d_{i_s}$ for all $1 \leq s \leq n$, so that $\mathscr{B}_M \vee_{\widetilde{G}} \mathscr{B}_P$ defines an interval decomposition for N. $\square$

Chapter 6

Persistent Mayer-Vietoris Spectral Sequences

6.1 Introduction

In this chapter we will study the spectral sequence that results from adapting the setup of Mayer-Vietoris spectral sequences from section 2.11 to persistence modules, as briefly discussed on section 3.7. The deep reason as to why we can do this adaptation is because **PMod**, the category of persistence modules, is an *abelian category* as **Vect** is an abelian category and $\mathbf{R}$ is a small category. The theory of spectral sequences can be developed for arbitrary abelian categories. For an introduction to this, see chapter 5 in [118].

Suppose that we have covered a filtered simplicial complex X with filtered subcomplexes $\mathscr{U} = \{U_i\}_{i\in I}$, so that $X = \bigcup_{i\in I} U_i$. Then, we can compute the spectral sequence

$$E^1_{p,q}(X,\mathscr{U}) = \bigoplus_{\sigma\in\Delta^m_p} \mathrm{PH}_q(U_\sigma) \Rightarrow \mathrm{PH}_n(X),$$

where $p+q=n$. However, unlike the case of vector spaces, we might have that

$$\bigoplus_{p+q=n} E^\infty_{p,q}(X,\mathscr{U}) \ncong \mathrm{PH}_n(X).$$

All we know is that $E^\infty_{p,q}(X,\mathscr{U}) \cong G^p\mathrm{PH}_{p+q}(X)$ for all $p,q \geq 0$. This is the *extension problem*, which we will solve in section 6.2. After solving this problem we will obtain the persistent homology for X. We will even recover more information. Notice that as pointed out in [119], the knowledge of which subset $J \subset I$ detects a feature from $\mathrm{PH}_n(X)$ can potentially add insight into the information given by ordinary persistent homology. The following example illustrates this.

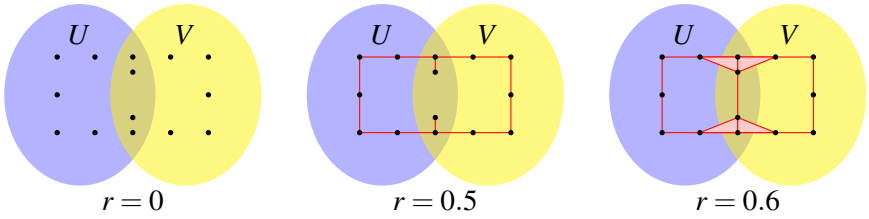


Figure 6.1: As the radius increases, more edges are added. At radius $r = 0.5$ a circle will be across the two covers U and V. Later on, at radius $r = 0.6$ this circle will be split into two.

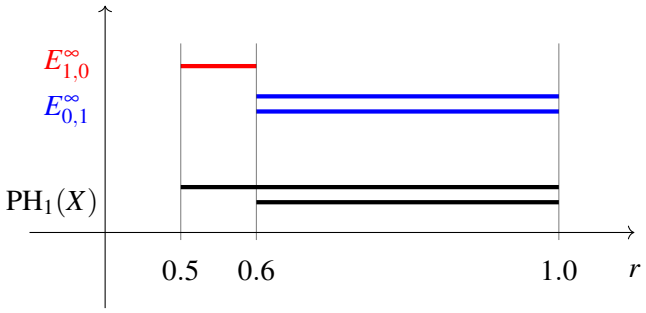


Figure 6.2: Barcode on associated module.

Example 6.1.1. Consider the case of a point cloud X covered by two open sets as in Figure 6.1. From section 2.11 and chapter 4, we know how to compute the ∞-page $(E^{\infty}_{*,*})^r$ associated to any value $r \in \mathbf{R}$. In particular, when we take $r = 0.5$, then the combination of U and V detects a 1-cycle. On the other hand, when $r = 0.6$ this cycle splits into two smaller cycles which are detected by U and V individually. Notice that if we want to come up with a persistent Mayer-Vietoris method then we need to be able to track this behavior. That is, we need to know how cycles develop as r increases. In particular, the barcode $\mathrm{I}(0.5,1)$ from $\mathrm{PH}_1(X)$ will be broken down into some smaller barcodes, see diagram 6.2. These will be $E^{\infty}_{1,0} \cong \mathrm{I}(0.5,0.6)$ and also $E^{\infty}_{0,1} \cong \mathrm{I}(0.6,1.0) \oplus \mathrm{I}(0.6,1.0)$. The way we will solve this problem is by using the barcode basis machinery developed in chapter 4.

6.2 The Extension Problem

Recall the definition of the total complex, vertical filtrations and associated modules from section 2.11. Through this section we study the extension problem. That is, we will recover $\mathrm{H}_n(\mathscr{S}^{\mathrm{Tot}}_*)$ from the associated modules $G^p_V\left(\mathrm{H}_n\left(\mathscr{S}^{\mathrm{Tot}}_*\right)\right)$ for all integers $n \geq 0$. Also, we will assume that the spectral sequence collapses after a finite number of pages. Consider the persistence module

$$\mathbb{V} = \mathbb{V}(n) := \mathrm{H}_n\left(\mathscr{S}^{\mathrm{Tot}}_*\right),$$

together with the corresponding filtration

$$0 = F_V^{-1}\mathbb{V} \subset F_V^0\mathbb{V} \subset \cdots \subset F_V^n\mathbb{V} = \mathbb{V}. \tag{6.1}$$

We define the *associated modules* of $(\mathbb{V}, F^*)$ as the quotients $\mathbb{G}^k = F^k\mathbb{V}/F^{k-1}\mathbb{V}$ for all $0 \leq k \leq n$. This gives rise to short exact sequences,

$$0 \longrightarrow F^{k-1}\mathbb{V} \xrightarrow{\iota} F^k\mathbb{V} \xrightarrow{p^k} \mathbb{G}^k \longrightarrow 0, \tag{6.2}$$

for all $0 \leq k \leq n$. On the sequences (6.2) we consider successive extension problems where we know the first and last term and want to know the middle term for $0 \leq k \leq n$; notice that this differs from the usual group extension as we are extending persistence modules instead. This leads to $\mathbb{V}$; however, in this work we will obtain directly a persistence module isomorphic to $\mathbb{V}$ (see proposition 6.2.1). Adding up all associated modules we obtain a persistence module $\mathbb{G} \coloneqq \bigoplus_{i=0}^n \mathbb{G}^i$ with an additional filtration given by $F^k\mathbb{G} = \bigoplus_{i=0}^k \mathbb{G}^i$ for all $0 \leq k \leq n$. Since $\mathbb{G}^k \cong E^\infty_{k,n-k}$ for all $0 \leq k \leq n$, a spectral sequence algorithm will lead to a barcode basis for $\mathbb{G}$. We formulate the extension problem as computing a basis for $\mathbb{V}$ from the obtained basis $\mathscr{G}$ of $\mathbb{G}$.

To start, notice that for each $r \in \mathbf{R}$ the sequence (6.2) splits, leading to morphisms

$$\mathscr{F}^k(r) : \mathbb{G}^k(r) \to F^k\mathbb{V}(r), \tag{6.3}$$

such that $p^k(r) \circ \mathscr{F}^k(r) = \mathrm{Id}_{\mathbb{G}^k(r)}$ for all $0 \leq k \leq n$. In particular, $\mathscr{F}^k(r)$ is injective for all $0 \leq k \leq n$. On the other hand, for any class $[\beta_k]^\infty_{k,n-k}$ of $E^\infty_{k,n-k}$ with representative $\beta_k \in E^0_{k,n-k}$, since $\beta_k \in \mathrm{GK}_{k,n-k}r$, we have that $d(\beta_k) = 0$ and there exists a sequence of $\beta_i \in \mathscr{S}_{i,n-i}r$ such that $d(\beta_i) = -\bar{\delta}(\beta_{i+1})$ for all $0 \leq i < k$. The choice of this sequence determines $\mathscr{F}^k(r)$, so that

$$\mathscr{F}^k(r)([\beta_k(r)]^\infty_{k,n-k}) = [(\beta_0(r), \beta_1(r), \ldots, \beta_k(r), 0, \ldots, 0)]^{\mathrm{Tot}}_n.$$

Notice that if we already computed $\mathscr{G}$ from the Mayer-Vietoris spectral sequence, then there is no need to do any extra computations to obtain these morphisms $\mathscr{F}^k(r)$. All we need to do is to store our previous results. Adding over all $0 \leq k \leq n$ we obtain the isomorphism $\mathscr{F}(r) = \bigoplus_{k=0}^n \mathscr{F}^k(r) : \bigoplus_{k=0}^n \mathbb{G}^k(r) \to \mathbb{V}(r)$. This last morphism is an isomorphism since all its summands are injective, their images have mutual trivial intersection, and the dimensions of the domain and codomain coincide.

Recall that $\mathbb{G}$ has induced morphisms $\mathbb{G}(r \leq s)$ from $\mathbb{V}(r \leq s)$ for all values $r \leq s$ in $\mathbf{R}$. Given a basis $\mathscr{G}$ for $\mathbb{G}$, we would like to compute a basis $\mathscr{B}$ for $\mathbb{V}$ from this information. Notice that this is not a straightforward problem since (6.3) does not imply that one has an isomorphism

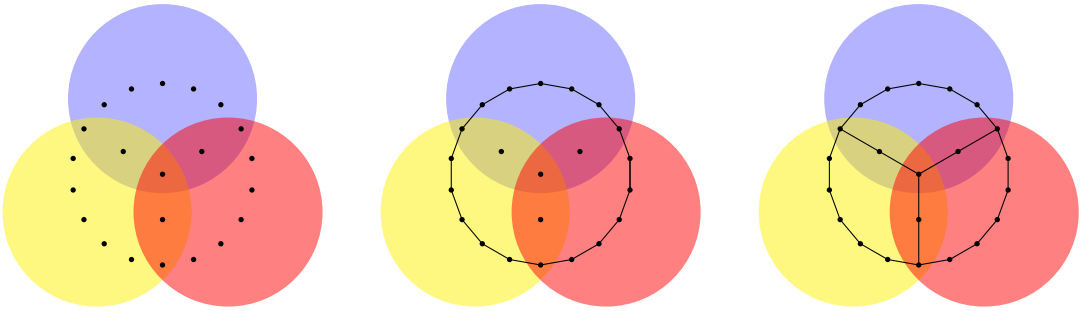


Figure 6.3: A one loop is detected at value $r \sim 0.208$ which goes through three covers. Later, at radius $r = 0.5$, this loop splits into three loops, each included in one of the three covers.

$\mathscr{F} : \mathbb{G} \to \mathbb{V}$. A point to start is to define the image along each generator in $\mathscr{G}$. That is, for each barcode generator $g_i \sim [a_i, b_i)$ in $\mathscr{G}$, we choose an image at the start $\mathscr{F}(a_i)(g_i(a_i))$. After, we set $\mathscr{F}(r)(g_i(r)) \coloneqq \mathbb{V}(a_i < r) \circ \mathscr{F}(a_i)(g_i(a_i))$ for all $a_i < r < b_i$. This leads to commutativity of $\mathscr{F}$ along each generator g_i. Nevertheless this is still far from even defining a morphism $\mathscr{F} : \mathbb{G} \to \mathbb{V}$. This si because $g_i \sim [a_i, b_i)$ but its image might be associated to a barcode which dies after b_i, breaking therefore naturality.

The solution to the problem above is to define a new persistence module $\widetilde{\mathbb{G}}$. We define $\widetilde{\mathbb{G}}(s) \coloneqq \mathbb{G}(s)$ for all $s \in \mathbf{R}$. Then, if $\mathscr{G} = \{g_i\}_{i=1}^N$ is a barcode basis for $\mathbb{G}$, by proposition 4.1.2 $\mathscr{G}^s(1_\mathbb{F})$ will be a basis of $\widetilde{\mathbb{G}}(s)$ for all $s \in \mathbf{R}$. Now, given $g_i \sim [a_i, b_i)$ a generator in $\mathscr{G}$, we define the morphism $\widetilde{\mathbb{G}}(r \leq s)$ by the recursive formula

$$\widetilde{\mathbb{G}}(r \leq s)(g_i(r)(1_\mathbb{F})) \coloneqq \begin{cases} \sum_{g_j \in \mathscr{G}^{b_i}} c_{i,j} \widetilde{\mathbb{G}}(b_i \leq s)(g_j(b_i)(1_\mathbb{F})) & \text{if } r \in [a_i, b_i) \text{ and } b_i \leq s, \\ g_i(s)(1_\mathbb{F}) & \text{if } r, s \in [a_i, b_i), \\ 0 & \text{otherwise,} \end{cases}$$

where $c_{i,j} \in \mathbb{F}$ for all $1 \leq i, j \leq |\mathscr{G}|$. We want to define $c_{i,j}$ in such a way that $\widetilde{\mathbb{G}}$ is isomorphic to $\mathbb{V}$. For this we impose the commutativity condition

$$\widetilde{\mathbb{G}}(a_i \leq b_i)(g_i(a_i)(1_\mathbb{F})) = \mathscr{F}(b_i)^{-1} \circ \mathbb{V}(a_i \leq b_i) \circ \mathscr{F}(a_i)(g_i(a_i)(1_\mathbb{F})),$$

which leads to the equation

$$\sum_{g_j \in \mathscr{G}^{b_i}} c_{i,j} g_j(b_i)(1_\mathbb{F}) = \mathscr{F}(b_i)^{-1} \circ \mathbb{V}(a_i \leq b_i) \circ \mathscr{F}(a_i)(g_i(a_i)(1_\mathbb{F})). \tag{6.4}$$

This determines uniquely the coefficients $c_{i,j}$ for all $1 \leq i, j \leq |\mathscr{G}|$; here notice that we define $c_{i,j} = 0$ for all $g_j \notin \mathscr{G}^{b_i}$. Notice that $\widetilde{\mathbb{G}}$ respects the filtration on $\mathbb{V}$, since the right hand side

in (6.4) is a composition of filtration preserving morphisms. In particular, if $g_i \in \mathrm{PVect}(F^k\widetilde{\mathbb{G}})$, then $c_{i,j} = 0$ for all $1 \leq j \leq |\mathscr{G}|$ such that $g_j \notin \mathrm{PVect}(F^k\mathbb{G})$.

Fix a generator $g_i \in \mathscr{G}$ such that $g_i \in \mathrm{PVect}(\mathbb{G}^k)$ and such that $g_i \sim [a_i, b_i)$. Let us calculate the coefficients $c_{i,j}$. Suppose that we have a representative $\widetilde{g}_j = (\beta_0^j, \beta_1^j, \ldots, \beta_k^j, 0, \ldots, 0) \in \mathscr{S}_n^{\mathrm{Tot}}$ for each generator $g_j \in \mathscr{G}$, with $g_j = [\beta_k^j]_{k,n-k}^{\infty}$. Also, for all $0 \leq q \leq n$ we define the subset $I_{b_i}^q \subseteq \{1, \ldots, |\mathscr{G}|\}$ of indices $1 \leq j \leq |\mathscr{G}|$ such that $g_j \in \mathrm{PVect}(\mathbb{G}^q)$ and also $g_j(b_i) \neq 0$. Then the coefficients $c_{i,j}$ for $j \in I_{b_i}^k \setminus \{i\}$ are determined by the following equality in $\mathbb{G}^k(b_i)$ (where we use p^k from the sequence (6.2))

$$p^k(b_i)\left([\widetilde{g}_i(b_i)(1_{\mathbb{F}})]_n^{\mathrm{Tot}}\right) = \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j} g_j(b_i)(1_{\mathbb{F}}).$$

Thus, we have

$$p^k(b_i)\left(\left[\widetilde{g}_i(b_i)(1_{\mathbb{F}}) - \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j}\widetilde{g}_j(b_i)(1_{\mathbb{F}})\right]_n^{\mathrm{Tot}}\right) = 0$$

where $[\cdot]_n^{\mathrm{Tot}}$ denotes the n-homology class of the total complex. Hence, by (6.2) there must exist some chain $\Gamma \in \mathscr{S}_{n+1}^{\mathrm{Tot}}(b_i)$ such that

$$\widetilde{g}_i(b_i)(1_{\mathbb{F}}) - \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j}\widetilde{g}_j(b_i)(1_{\mathbb{F}}) - d^{\mathrm{Tot}}\Gamma \tag{6.5}$$

is contained in $F_V^{k-1}\mathscr{S}_n^{\mathrm{Tot}}(b_i)$. How do we compute Γ? We start by searching for the first page $r \geq 2$ such that

$$\left[\beta_k^i(b_i)(1_{\mathbb{F}}) - \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j}\beta_k^j(b_i)(1_{\mathbb{F}})\right]_{k,n-k}^{r} = 0 \tag{6.6}$$

where $[\cdot]_{k,n-k}^r$ denotes the class in the r-page in position $(k, n-k)$. Notice that this r must exist since we assumed that (6.6) vanishes on the ∞-page. In fact, there exists $\Gamma_{k+r-1} \in E_{k+r-1,n-k-r+2}^{r-1}(b_i)$ such that

$$\left[\beta_k^i(b_i)(1_{\mathbb{F}}) - \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j}\beta_k^j(b_i)(1_{\mathbb{F}})\right]_{k,n-k}^{r-1} - d^{r-1}(\Gamma_{k+r-1}) = 0$$

on $E_{k,n-k}^{r-1}(b_i)$. Repeating for all pages leads to $\Gamma_{k+t} \in E_{k+t,n-k-t+1}^{t}(b_i)$ for all $0 \leq t \leq r-1$, such that

$$\beta_k^i(b_i)(1_{\mathbb{F}}) - \sum_{j \in I_{b_i}^k \setminus \{i\}} c_{i,j}\beta_k^j(b_i)(1_{\mathbb{F}}) - \sum_{t=0}^{r-1} \widetilde{d^t(\Gamma_{k+t})} = 0, \tag{6.7}$$

where $\widetilde{d^t(\Gamma_{k+t})} \in \mathscr{S}_{k,n-k}(b_i)$ is a representative for the class $d^t(\Gamma_{k+t}) \in E_{k,n-k}^t(b_i)$. Notice that equation (6.7) holds independently of the representatives, since if we changed some term, then

the other representatives would adjust to the change. In particular, we have that the k component of (6.5) vanishes, whereas the $k-1$ component will be equal to

$$\beta^i_{k-1}(b_i)(1_{\mathbb{F}}) - \sum_{j\in I^k_{b_i}\setminus\{i\}} c_{i,j}\beta^j_{k-1}(b_i)(1_{\mathbb{F}}) - \bar{\delta}(\Gamma_k).$$

Next we proceed to find coefficients $c_{i,j}\in\mathbb{F}$ so that in $\mathbb{G}^{k-1}(b_i)$ we get the equality

$$\left[\beta^i_{k-1}(b_i)(1_{\mathbb{F}}) - \sum_{j\in I^k_{b_i}\setminus\{i\}} c_{i,j}\beta^j_{k-1}(b_i)(1_{\mathbb{F}}) - \bar{\delta}(\Gamma_k)\right]^{\infty}_{k-1,n-k+1} = \sum_{j\in I^{k-1}_{b_i}} c_{i,j}g_j(b_i)(1_{\mathbb{F}})\ .$$

Then we proceed as we did on $\mathbb{G}^k$. Doing this for all parameters $0\le r\le k$, there are coefficients $c_{i,j}\in\mathbb{F}$, and an element $\bar{\Gamma}\in\mathscr{S}^{\mathrm{Tot}}_n(b_i)$ so that

$$\widetilde{g}_i(b_i)(1_{\mathbb{F}}) = \sum_{0\le r\le k}\left(\sum_{j\in I^r_{b_i}} c_{i,j}\widetilde{g}_j(b_i)(1_{\mathbb{F}})\right) + d^{\mathrm{Tot}}\bar{\Gamma}.$$

Thus,

$$\widetilde{\mathbb{G}}(a_i\le b_i)(g_i(a_i)(1_{\mathbb{F}})) = \sum_{g_j\in\mathscr{G}^{b_i}} c_{i,j}g_j(b_i)(1_{\mathbb{F}}).$$

Proposition 6.2.1. $\widetilde{\mathbb{G}}\cong\mathbb{V}$.

Proof. Since each $\mathscr{F}(s)$ is an isomorphism, and also we have commutative squares:

$$\begin{array}{ccc}\widetilde{\mathbb{G}}(r) & \xrightarrow{\widetilde{\mathbb{G}}(r\le s)} & \widetilde{\mathbb{G}}(s)\\ {\scriptstyle\mathscr{F}(r)}\downarrow & & \downarrow{\scriptstyle\mathscr{F}(s)}\\ \mathbb{V}(r) & \xrightarrow[\mathbb{V}(r\le s)]{} & \mathbb{V}(s)\end{array}$$

for all $r\le s$, then $\mathscr{F}$ must be an isomorphism of persistence modules. □

This gives $\widetilde{\mathbb{G}}\cong\mathbb{V}$, but we still need to compute a barcode basis. In fact, this can be done by considering a quotient. Define $\mathbb{A}\simeq\bigoplus_{g_i\in\mathscr{G}}\mathrm{I}(a_i,\infty)$ where $g_i\sim[a_i,b_i)$ for all $g_i\in\mathscr{G}$; here the $\mathscr{A}=\{\alpha_i\}_{0\le i\le|\mathscr{G}|}$ will denote the canonical base for $\mathbb{A}$. Consider the coefficients $c_{i,j}$ for $0\le i,j\le|\mathscr{G}|$ from the construction of $\widetilde{\mathbb{G}}$ and define the sets of indices $S_i=\{0\le j\le|\mathscr{G}||c_{i,j}\neq 0\}$ for all $0\le i\le|\mathscr{G}|$. We consider a submodule $\mathbb{B}\subseteq\mathbb{A}$ such that $\mathrm{PVect}(\mathbb{B})$ is generated by

$$\left\{\mathbf{1}_{b_i}\left((-\alpha_i)\boxplus\coprod_{j\in S_i}c_{i,j}\alpha_j\right)\right\}_{0\le i\le|\mathscr{G}|}.$$

Also, notice that $\mathbb{B}\simeq\bigoplus_{0\le i\le|\mathscr{G}|}\mathrm{I}(b_i,\infty)$. One might see that by construction $\widetilde{\mathbb{G}}\simeq\mathbb{A}/\mathbb{B}$. Now, we

pick up the canonical base for $\mathbb{B}$ and consider the inclusion $\iota : \mathbb{B} \hookrightarrow \mathbb{A}$; this will lead to an associated matrix $(\iota(\mathscr{B}))_{\mathscr{A}}$. Thus, we send $((\mathscr{B}|\mathscr{A}), \mathscr{A}, ((\iota(\mathscr{B}))_{\mathscr{A}}|\mathrm{Id}_{|\mathscr{A}|}))$ to `box_gauss_reduce` and obtain the result.

6.3 Efficient implementation: PERMAVISS

Here we outline a procedure for implementing the persistence Mayer-Vietoris spectral sequence. In section 6.2 we worked with $\mathrm{GZ}^r_{p,q}$ and $\mathrm{IB}^r_{p,q}$, which is very intuitive from a mathematical perspective. However, it is more efficient to work directly with the sets $Z^r_{p,q}$ and $B^r_{p,q}$. By storing representatives in $Z^r_{p,q}$, we avoid repeating computations on each page and in the extension problem. Furthermore, this approach allows to easily track the complexity of the algorithm. Additionally, we will use more extensively barcode bases through this section, which enables to use a more compact notation.

0**-Page**. We start by defining the 0-page as the quotient

$$E^0_{p,q} = \frac{F^p_V \mathscr{S}^{\mathrm{Tot}}_{p+q}}{F^{p-1}_V \mathscr{S}^{\mathrm{Tot}}_{p+q}} \cong \mathscr{S}_{p,q} = \bigoplus_{\sigma \in N^{\mathscr{U}}_p} S_q(U_\sigma)$$

for all pair of integers $p, q \geq 0$. The 0 differential d^0, is isomorphic to the standard chain differential

$$d^0_{p,q} \cong d_q : \mathscr{S}_{p,q} \to \mathscr{S}_{p,q-1}.$$

In particular, for each simplex $\sigma \in N^{\mathscr{U}}_q$, the morphism $d^0_{p,q}$ restricts to a local differential

$$d^\sigma_q : S_q(U_\sigma) \to S_{q-1}(U_\sigma).$$

Thus, we can compute persistent homology to obtain a local base for the image $\mathrm{Im}(d^\sigma_{q+1})$ and the homology $\mathscr{E}^1_{\sigma,q}$. Putting all of these together, we get a basis for $E^1_{p,q}$ as the union $\mathscr{E}^1_{p,q} = \bigcup_{\sigma \in N^{\mathscr{U}}_p} \mathscr{E}^1_{\sigma,q}$. Additionally, we will denote by $\mathscr{K}^1_{\sigma,q}$ the subset of representatives of the classes from $\mathscr{E}^1_{\sigma,q}$ contained in $\mathrm{PVect}(E^0_{\sigma,q})$, and the same for $\mathscr{K}^1_{p,q} \subseteq \mathrm{PVect}(E^0_{p,q})$. Here the notation $\mathscr{K}^n_{*,*}$ comes from the fact that these are persistence vectors contained on the kernels, and that these represent n-page generators. Further, for each generator $\alpha \in \mathscr{E}^1_{p,q} \subseteq \mathrm{PVect}(E^1_{p,q})$, we store a chain $\alpha_p \in \mathscr{S}_{p,q}$ so that $\alpha = [(0, \ldots, 0, \alpha_p, 0, \ldots, 0)]^1_{p,q}$. Where we denote by $[\,\cdot\,]^r_{p,q}$ a class in $E^r_{p,q}$ for all $r \geq 0$. We define $\widetilde{\mathscr{E}}^1_{p,q}$ to be the set of representatives in $\mathscr{S}^{\mathrm{Tot}}_{p+q}$ given by $(0, \ldots, 0, \alpha_0, 0, \ldots, 0)$.

1**-Page**. Recall that the first page elements are given as classes in the quotient

$$E^1_{p,q} = \frac{Z^1_{p,q}}{Z^0_{p-1,q+1} + B^0_{p,q}}.$$

Therefore, for each generator $\alpha \in \mathscr{E}^1_{p,q}$, with $\alpha \sim [a_\alpha, b_\alpha)$, there is a chain $\alpha_p \in \mathscr{S}_{p,q}$, so that $\alpha = [(0,\dots,0,\alpha_p,0,\dots,0)]^0_{p,q}$. Then we compute

$$d^1_{p,q}(\alpha) = \Big[d^{\mathrm{Tot}}(0,\dots,0,\alpha_p,0,\dots,0) \Big]^1_{p-1,q} = \Big[\big(0,\dots,0,\bar{\delta}_p(\alpha_p),0,\dots,0\big) \Big]^1_{p-1,q} .$$

Now, for each simplex $\tau \in N^{\mathscr{U}}_{p-1}$, we have local coordinates $\big(\bar{\delta}_p(\alpha_p)(a_\alpha)(1_{\mathbb{F}})\big)_\tau \in S_q(U_\tau)(a_\alpha)$. We proceed to solve the linear equation at $a_\alpha \in \mathbf{R}$

$$\big(\mathrm{Im}\big((d_{q+1})_\tau\big) \,\big|\, \mathscr{K}^1_{\tau,q}\big)^{a_\alpha} X = \big(\bar{\delta}_p(\alpha_p)(a_\alpha)(1_{\mathbb{F}})\big)_\tau ,$$

where the vector X has as many entries as needed for the equation to make sense. Also, we have used

$$\big(\mathrm{Im}\big((d_{q+1})_\tau\big) \,\big|\, \mathscr{K}^1_{\tau,q}\big)^{a_\alpha} = \big(\mathrm{Im}\big((d_{q+1})_\tau\big)(a_\alpha) \,\big|\, \mathscr{K}^1_{\tau,q}(a_\alpha)\big)$$

for denoting the matrix on value a_α, and whose rows correspond to a basis of $S_q(U_\tau)(a_\alpha)$. The solution X leads to a subset $\mathscr{T}^1_{\tau,q} \subseteq \mathscr{K}^1_{\tau,q}$ together with $c^1_\beta \in \mathbb{F}$ for all $\beta \in \mathscr{T}^1_{\tau,q}$ and an element $\Gamma_\tau \in S_{q+1}(U_\tau)$ such that

$$\bar{\delta}_p(\alpha_\tau) \boxplus d_{q+1}(\Gamma_\tau) = \mathbf{1}_{a_\alpha} \left(\bigboxplus_{\beta \in \mathscr{T}^1_{\tau,q}} c^1_\beta \beta \right) .$$

Repeating this for all $\tau \in N^{\mathscr{U}}_{p-1}$, we get a subset $\mathscr{T}^1_{p-1,q} \subseteq \mathscr{K}^1_{p-1,q}$ together with coefficients $c^1_\beta \in \mathbb{F} \setminus \{0\}$ for all $\beta \in \mathscr{J}^1_{p-1,q}$ and an element $\Gamma_p \in E^1_{p,q+1}$ so that

$$\bar{\delta}_p(\alpha_p) \boxplus d^0_{p,q+1}(\Gamma_p) = \mathbf{1}_{a_\alpha} \left(\bigboxplus_{\beta \in \mathscr{T}^1_{p-1,q}} c^1_\beta \beta \right) .$$

This leads to $d^1_{p,q}(\alpha) = \mathbf{1}_{a_\alpha}\big(\boxplus_{\beta \in \mathscr{T}^1_{p-1,q}} c^1_\beta [\beta]^1_{p-1,q}\big)$ where $[\beta]^1_{p-1,q} \in \mathscr{E}^1_{p-1,q}$ for all $\beta \in \mathscr{T}^1_{p-1,q}$. We will also denote by $\mathscr{T}^1_{p-1,q}$ the subset of $\mathscr{E}^1_{p-1,q}$ formed by elements $[\beta]^1_{p-1,q}$ for all $\beta \in \mathscr{T}^1_{p-1,q}$ (it should be clear by the context when we refer to the corresponding base). Repeating this for all generators $\alpha \in \mathscr{E}^1_{p,q}$ leads to an associated matrix $D^1_{p,q}$ for $d^1_{p,q}$. Using `image_kernel`, we compute bases for the kernel and image. Additionally, we store a basis for $\mathrm{Im}(d^1_{p,q})$ together with the corresponding preimages, as enabled by the `image_kernel` algorithm. Thus, we might compute a base $\mathscr{E}^2_{p,q}$ for the second page term $E^2_{p,q}$ by applying `box_gauss_reduce` to compute the quotient $\mathrm{Ker}(d^1_{p,q})/\mathrm{Im}(d^1_{p+1,q})$. We also store first page representatives $\mathscr{K}^2_{p,q}$ for the generators from $\mathscr{E}^2_{p,q}$. Additionally, for each generator $\alpha \in \mathscr{E}^2_{p,q}$, we proceed to find a good representative by using $\widetilde{\mathscr{E}}^1_{p,q}$ together with the computed coordinates of α in terms of $\mathscr{E}^1_{p,q}$. This leads to $\widetilde{\alpha} = (0,\dots,0,\alpha_p,0,\dots,0) \in \mathscr{S}^{\mathrm{Tot}}_{p+q}$, so that $\alpha = [\widetilde{\alpha}]^2_{p,q}$. Since by hypotheses $[\widetilde{\alpha}]^1_{p,q} \in$

$\mathrm{Ker}(d^1_{p,q})$, we might find $\alpha_{p-1} \in E^0_{p-1,q+1}$ such that $d^0_{p-1,q}(\alpha_{p-1}) = -\overline{\delta}(\alpha_p)$. Altogether, we set $\widetilde{\alpha} \leftarrow (0,\dots,0,\alpha_{p-1},\alpha_p,0,\dots,0)$ and store it in $\widetilde{\mathscr{E}}^2_{p,q}$; notice that it has the property that $d^{\mathrm{Tot}}(\widetilde{\alpha}) \in F_V^{p-2}\mathscr{S}^{\mathrm{Tot}}$.

2-**Page**. Now, we proceed to compute the third page. We start from $\alpha \in \mathscr{E}^2_{p,q}$, with $\alpha \sim [a_\alpha, b_\alpha)$ together with a total complex representative $\widetilde{\alpha} \in \widetilde{\mathscr{E}}^2_{p,q}$, so that

$$d^2_{p,q}(\alpha) = \left[d^{\mathrm{Tot}}(\widetilde{\alpha})\right]^2_{p-2,q+1} = \left[(0,\dots,0,\overline{\delta}_{p-1}(\alpha_{p-1}),0,\dots,0)\right]^2_{p-2,q+1}.$$

As before, by solving local linear equations, we obtain $\mathscr{T}^1_{p-2,q+1} \subseteq \mathscr{E}^1_{p-2,q+1}$ together with coefficients $c^1_\beta \in \mathbb{F}$ for all $\beta \in \mathscr{T}^1_{p-2,q+1}$, so that

$$\left[d^{\mathrm{Tot}}(\widetilde{\alpha})\right]^1_{p-2,q+1} = \mathbf{1}_{a_\alpha}\left(\bigsqcup_{\beta\in\mathscr{T}^1_{p-2,q+1}} c^1_\beta\beta\right).$$

Now, we solve the linear equation on X

$$\left(\mathrm{Im}(d^1_{p-1,q+1}) \,\middle|\, \mathscr{K}^2_{p-2,q+1}\right)^{a_\alpha} X = (c^1_\beta)_{\beta\in\mathscr{E}^1_{p-2,q+1}},$$

where $c^1_\beta = 0$ for $\beta \notin \mathscr{T}^1_{p-2,q+1}$. The solution X leads to a subset $\mathscr{T}^2_{p-2,q+1} \subseteq \mathscr{K}^2_{p-2,q+1}$ together with coefficients $c^2_\beta \in \mathbb{F}$ for all $\beta \in \mathscr{T}^2_{p-2,q+1}$, and an element $\Gamma_{p-1} \in E^1_{p-1,q+1}$ so that

$$\left[d^{\mathrm{Tot}}(\widetilde{\alpha})\right]^1_{p-2,q+1} = \left[\mathbf{1}_{a_\alpha}\left(\bigsqcup_{\beta\in\mathscr{T}^1_{p-2,q+1}} c^1_\beta\beta\right)\right]^1_{p-2,q+1} + d^1_{p-1,q+1}(\Gamma_{p-1}) = \mathbf{1}_{a_\alpha}\left(\bigsqcup_{\beta\in\mathscr{T}^2_{p-2,q+1}} c^2_\beta\beta\right).$$

This leads to an expression of $d^2_{p,q}(\alpha)$ in terms of $\mathscr{E}^2_{p-2,q+1}$. Repeating this for all generators $\alpha \in \mathscr{E}^2_{p,q}$ leads to a matrix $D^2_{p,q}$ associated to $d^2_{p,q}$. Then applying `image_kernel` we obtain bases for the kernel, images and preimages. Then, applying `box_gauss_reduce` we obtain a base $\mathscr{E}^3_{p,q}$ for the term $E^3_{p,q}$; store second page representatives on $\mathscr{K}^3_{p,q}$. Now, for each $\kappa \in \mathscr{E}^3_{p,q} \subseteq E^2_{p,q}$ we proceed to prepare a nice representative on the total complex. For this, we use the stored representatives from $\widetilde{\mathscr{E}}^2_{p,q}$ together with the expression of κ in terms of $\mathscr{E}^2_{p,q}$ to obtain $\widetilde{\kappa} \in F^p_V\mathscr{S}^{\mathrm{Tot}}_{p+q}$ such that $\kappa = [\widetilde{\kappa}]^3_{p,q}$ and also $d^{\mathrm{Tot}}(\widetilde{\kappa}) \in F^{p-2}_V\mathscr{S}^{\mathrm{Tot}}$. Next, by hypotheses $[d^{\mathrm{Tot}}(\widetilde{\kappa})]^2_{p-2,q+1} = 0$, and so there exists some $\gamma_{p-1} \in E^1_{p-1,q+1}$ such that $d^1_{p-1,q+1}(\gamma_{p-1}) = [d^{\mathrm{Tot}}(\widetilde{\kappa})]^1_{p-2,q+1}$. From the stored information, one might obtain $\widetilde{\gamma}_{p-1} \in \mathscr{S}^{\mathrm{Tot}}_{p+q}$ such that $[\widetilde{\gamma}_{p-1}]^1_{p-1,q+1} = \gamma_{p-1}$ and also $d^{\mathrm{Tot}}(\widetilde{\gamma}_{p-1}) \in F^2_V\mathscr{S}^{\mathrm{Tot}}_{p+q-1}$. Then, we set $\widetilde{\kappa} \leftarrow \widetilde{\kappa} - \widetilde{\gamma}$ and obtain that $[d^{\mathrm{Tot}}(\widetilde{\kappa})]^1_{p-2,q+1} = 0$ while $d^{\mathrm{Tot}}(\widetilde{\kappa}) \in F^2_V\mathscr{S}^{\mathrm{Tot}}_{p+q-1}$. Consequently, we might lift the $p-2$ component of $d^{\mathrm{Tot}}(\widetilde{\kappa})$ by the vertical differential. Modifying $\widetilde{\kappa}$ using this information leads to a new redefinition of $\widetilde{\kappa}$ such that $d^{\mathrm{Tot}}(\widetilde{\kappa}) \in F^{p-3}_V\mathscr{S}^{\mathrm{Tot}}_{p+q-1}$. These new representatives are stored in $\widetilde{\mathscr{E}}^3_{p,q}$.

***k*-Page**. Suppose that we have computed generators $\mathscr{E}^k_{p,q} \subseteq \mathrm{PVect}(E^{k-1}_{p,q})$, together with total complex representatives $\widetilde{\mathscr{E}}^k_{p,q}$ for some $k \geq 3$. Notice that if $p-k+1 \leq 0$, then $d^k_{p,q} = 0$. Thus, we will focus on the case that $p-k+1 > 0$. Let a generator $\alpha \in \mathscr{E}^k_{p,q}$ with $\alpha \sim [a_\alpha, b_\alpha)$ together with a representative $\widetilde{\alpha} \in \widetilde{\mathscr{E}}^k_{p,q}$ with $\widetilde{\alpha} = (0,\ldots,0,\alpha_{p-k+1},\ldots,\alpha_p,0,\ldots,0)$ so that

$$d^k(\alpha) = [d^{\mathrm{Tot}}(\widetilde{\alpha})]^k_{p-k,q+k-1} = [(0,\ldots,0,\bar{\delta}_{p-k+1}(\alpha_{p-k+1}),0,\ldots,0)]^k_{p-k,q+k-1} \, .$$

We will proceed by 'lifting' $d^{\mathrm{Tot}}(\widetilde{\alpha})$ to the k-page. As before, by solving local linear equations, we obtain a subset $\mathscr{T}^1_{p-k,q+k-1} \subseteq \mathscr{E}^1_{p-k,q+k-1}$ together with coefficients $c^1_\beta \in \mathbb{F} \setminus \{0\}$ for all $\beta \in \mathscr{T}^1_{p-k,q+k-1} \subseteq \mathscr{S}^{\mathrm{Tot}}_{p+q-1}$ giving us an expression in terms of the first page basis $[d^{\mathrm{Tot}}(\widetilde{\alpha})]^1_{p-k,q+k-1} = \mathbf{1}_{a_\alpha}\big(\coprod_{\beta\in\mathscr{T}^1_{p-k,q+k-1}} c^1_\beta \beta\big)$. Next, for each $2 \leq r \leq k$, we solve the linear equation on X

$$\Big(\mathrm{Im}\big(d^{r-1}_{p-k+r-1,q+k-r+1}\big) \,\Big|\, \mathscr{K}^r_{p-k,q+k-1}\Big)^{a_\alpha} X = (c^{r-1}_\beta)_{\beta\in\mathscr{E}^{r-1}_{p-k,q+k-1}} \, ,$$

where $c^{r-1}_\beta = 0$ for $\beta \notin \mathscr{T}^{r-1}_{p-k,q+k-1}$. This leads to a subset $\mathscr{T}^r_{p-k,q+k-1} \subseteq \mathscr{K}^r_{p-k,q+k-1}$ together with coefficients $c^r_\beta \in \mathbb{F}\setminus\{0\}$ for all $\beta \in \mathscr{T}^r_{p-k,q+k-1}$, and an element $\Gamma_{p-k+r-1} \in E^{r-1}_{p-k+r-1,q+k-r}$ such that

$$\left[\mathbf{1}_{a_\alpha}\left(\coprod_{\beta\in\mathscr{T}^{r-1}_{p-k,q+k-1}} c^{r-1}_\beta\beta\right)\right]^{r-1}_{p-k,q+k-1} \coprod d^{r-1}_{p-k+r-1,q+k-r}(\Gamma_{p-k+r-1}) = \mathbf{1}_{a_\alpha}\left(\coprod_{\beta\in\mathscr{T}^r_{p-k,q+k-1}} c^r_\beta\beta\right) \, .$$

Thus, we deduce an expression of $\big[d^{\mathrm{Tot}}\widetilde{\alpha}\big]^r_{p-k,q+k-1}$ in terms of $\mathscr{E}^r_{p-k,q+k-1}$. In particular, this will hold for $r=k$, which leads to the associated matrix of $d^k_{p,q}$, and then we can compute `image_kernel`, etc. Now, consider a generator $\kappa \sim [a_\kappa, b_\kappa)$ from $\mathscr{E}^{k+1}_{p,q}$. There exists $\mathscr{T}(\kappa) \subseteq \mathscr{E}^k_{p,q}$ together with coefficients c^κ_β for all $\beta \in \mathscr{T}(\kappa)$ such that $\kappa = \mathbf{1}_{a_\kappa}\big(\coprod_{\beta\in\mathscr{T}(\kappa)} c^\kappa_\beta\beta\big)$. Then we define $\widetilde{\kappa} = \mathbf{1}_{a_\kappa}\Big(\coprod_{\beta\in\mathscr{T}(\kappa)} c_\beta\widetilde{\beta}\Big)$, and notice that $\kappa = [\widetilde{\kappa}]^{k+1}_{p,q}$ as well as $d^{\mathrm{Tot}}(\widetilde{\kappa}) \in F^{p-k}_V(\mathscr{S}^{\mathrm{Tot}})$. We will use $\widetilde{\kappa} = (0,\ldots,0,\kappa_{p-k+1},\ldots,\kappa_p,0,\ldots,0)$. Now, by hypotheses

$$d^k_{p,q}(\kappa) = [d^{\mathrm{Tot}}(\widetilde{\kappa})]^k_{p-k,q+k-1} = [(0,\ldots,0,\bar{\delta}_{p-k+1}(\kappa_{p-k+1}),0,\ldots,0)]^k_{p-k,q+k-1} = 0 \, .$$

Then, there exist $\gamma_{p-1} \in E^{k-1}_{p-1,q+1}$ such that $d^{p-1}_{p-1,q+1}(\gamma_{p-1}) = [d^{\mathrm{Tot}}(\widetilde{\kappa})]^{k-1}_{p-k,q+k-1}$. By writing γ_{p-1} in terms of $\mathscr{E}^{k-1}_{p-1,q+1}$ and using their stored representatives, we may get $\widetilde{\gamma}_{p-1} \in \mathscr{S}^{\mathrm{Tot}}_{p+q}$ such that $\gamma_{p-1} = [\widetilde{\gamma}_{p-1}]^{k-1}$ and also $d^{\mathrm{Tot}}(\widetilde{\gamma}_{p-1}) \in F^{p-k}\mathscr{S}^{\mathrm{Tot}}_{p+q-1}$. In particular, $[d^{\mathrm{Tot}}(\widetilde{\kappa} - \widetilde{\gamma}_{k-1})]^{k-1} = 0$ and we set $\widetilde{\kappa} \leftarrow \widetilde{\kappa} - \widetilde{\gamma}_{k-1}$. Hence, by induction, we can repeat this procedure for all $1 \leq r \leq k$. Eventually, we should obtain a representative $\widetilde{\kappa} = (0,\ldots,0,\kappa_{p-k},\ldots,\kappa_p,0,\ldots,0)$ such that $d^{\mathrm{Tot}}(\widetilde{\kappa}) \in F^{p-k-1}\mathscr{S}^{\mathrm{Tot}}_{p+q-1}$. We denote the new set of representatives as $\widetilde{\mathscr{E}}^{k+1}_{p,q}$.

Extension Problem

After computing all pages of the spectral sequence, we still have to solve the extension problem. It turns out that the procedure is almost exactly the same as for when computing a page on the spectral sequence. We start from a basis $\mathscr{E}^{\infty}_{p,q}$, with total complex representatives $\widetilde{\mathscr{E}}^{\infty}_{p,q}$. Since we assume that the spectral sequence is bounded, it collapses at an $L>0$ page. Then, for each generator $\alpha \in \mathscr{E}^{L}_{p,q}$, with $\alpha \sim [a_\alpha, b_\alpha)$, we have a corresponding representative $\widetilde{\alpha} = (\alpha_0, \dots, \alpha_p, 0, \dots, 0) \in \mathscr{S}^{\mathrm{Tot}}_{p+q}$ in $\widetilde{\mathscr{E}}^{L}_{p,q}$. The main procedure consists in lifting α_p to the L-page. We do this by means of local linear equations as done on the 1-page. However, this time, instead of using the value a_α we use b_α. This leads to a subset $\mathscr{T}^1_{p,q} \subseteq \mathscr{K}^1_{p,q}$ together with coefficients $c^1_\beta \in \mathbb{F}\setminus\{0\}$ for all $\beta \in \mathscr{T}^1_{p,q}$ and $\Gamma_p \in \mathscr{S}_{p,q+1}$, so that $\mathbf{1}_{b_\alpha}(\alpha_p) \boxplus d_{q+1}(\Gamma_p) = \mathbf{1}_{b_\alpha}\left(\boxplus_{\beta\in\mathscr{T}^1_{p,q}} c^1_\beta \beta\right)$. The same happens for all the pages $1 \le r \le L$; we find $\mathscr{T}^r_{p,q} \subseteq \mathscr{K}^r_{p,q}$ together with coefficients $c^r_\beta \in \mathbb{F}\setminus\{0\}$ for all $\beta \in \mathscr{T}^r_{p,q}$ and $\Gamma_{p+r} \in \mathscr{S}_{p+r,q-r+1}$, so that

$$\mathbf{1}_{b_\alpha}([\widetilde{\alpha}]^r_{p,q}) \boxplus d^r_{p+r,q-r+1}(\Gamma_{p+r}) = \mathbf{1}_{b_\alpha}\left(\bigboxplus_{\beta\in\mathscr{T}^r_{p,q}} c^1_\beta \beta\right)$$

Now, using the same notation of $\mathscr{T}^L_{p,q}$ for the corresponding subset from $\mathscr{E}^L_{p,q}$, we define

$$\widetilde{\alpha}^{p-1} = \mathbf{1}_{b_\alpha}(\widetilde{\alpha}) \boxplus \bigboxplus_{r=0}^{L-1}\left(\bigboxplus_{\Gamma_{p+r}\in\mathscr{E}^r_{p+r,q-r+1}} d^r_{p+r,q-r+1}(\Gamma_{p+r})\right) - \bigboxplus_{\beta\in\mathscr{E}^L_{p,q}} c^L_\beta \widetilde{\beta}\ .$$

In particular, notice that $[\widetilde{\alpha}^{p-1}]^L_{p,q} = 0$. In fact, for all integers $L-1 \ge r \ge 0$ one has that $[\widetilde{\alpha}^{p-1}]^r_{p,q} = 0$, since both the adding and substracting terms are a sum of elements in $\mathscr{E}^r_{p,q}$ with the same coefficients. As a consequence the p-component of $\widetilde{\alpha}^{p-1}$ vanishes, so $\widetilde{\alpha}^{p-1} \in F_V^{p-1}\mathscr{S}^{\mathrm{Tot}}_{p+q}$. Then, one can repeat this process with $\widetilde{\alpha}^r$ for all $p-1 \ge r \ge 0$. This leads to all coefficients $(c^L_\beta)_{\beta\in\mathscr{E}^L_{p-r,q+r}}$ for all $0 \le r \le p$. These can be used to define a basis for the submodule $\mathbb{B}$ from end of section 6.2; this solves the extension problem by computing the corresponding quotient.

6.4 Complexity Analysis

Let D_s be the maximum simplex dimension in K, and $\dim(N^{\mathscr{U}})$ the dimension of the nerve. Let L be the number of pages. Denote $N^{\mathscr{U}}_{\ge 1} = \bigcup_{k\ge 1} N^{\mathscr{U}}_k$. Let

$$X = \max_{q\ge 0,\,\sigma\in N^{\mathscr{U}}} \{\,\#\ q\text{-simplices in } \mathscr{U}_\sigma\,\} \qquad \text{and} \qquad Y = \max_{q\ge 0,\,\tau\in N^{\mathscr{U}}_{\ge 1}} \{\,\#\ q\text{-simplices in } \mathscr{U}_\tau\,\}\ .$$

Notice that $X \geq Y$. On the other hand, we define

$$H = \max_{p,q \geq 0} \left\{ |\mathscr{E}^1_{p,q}| \mid \text{ where } \mathscr{E}^1_{p,q} \text{ is a base for } E^1_{p,q} \right\}.$$

Assume P is the number of processors.

0**-page**. When computing the first page, all we need to do is calculate persistent homology in parallel. Then, the complexity is

$$\left\lceil \frac{|\mathscr{U}|}{P} \right\rceil \mathscr{O}(X^3) + \left\lceil \frac{|N^{\mathscr{U}}_{\geq 1}|}{P} \right\rceil \mathscr{O}(Y^3)$$

This leads to generators for the first page.

1**-page**. For the first page, recall that we start from a generator $\alpha \in \mathscr{E}^1_{p,q}$ with $\alpha \sim [a_\alpha, b_\alpha)$ and proceed to solve $|N^{\mathscr{U}}_{p-1}|$ linear equations. Notice that this can be done for all generators from $\mathscr{E}^1_{p,q}$ simultaneously. This is because as the value a_α changes, only columns are added and removed to the local linear equations, leaving the rows intact. On the other hand, we need to execute `image_kernel` on at most $\dim(N^{\mathscr{U}})D_s$ elements on the first page. Notice that for each of these, we first compute a basis for the images and kernels, and afterwards we perform the quotients using `box_gauss_reduce`. Each of these takes a complexity of at most $\mathscr{O}(H^3)$. Also, we need to add the complexity of the Čech differential. An option for computing this, is to compare simplices in different covers by their vertices; two simplices are the same iff they share the same vertex set. This would take less than $\mathscr{O}(|N^{\mathscr{U}}|D_s X^2 H)$ operations. Thus the overall complexity becomes

$$\left\lceil \frac{|\mathscr{U}|}{P} \right\rceil \mathscr{O}(X^2 H) + \left\lceil \frac{|N^{\mathscr{U}}_{\geq 1}|}{P} \right\rceil \mathscr{O}(Y^2 H) + \left\lceil \frac{\dim(N^{\mathscr{U}})D_s}{P} \right\rceil \left(\mathscr{O}(|N^{\mathscr{U}}|D_s X^2 H) + \mathscr{O}(H^3) \right)$$

k**-page**. Now, we proceed for the complexity of the page $k \geq 2$. This is the same as for the 1 page, with the addition of Gaussian eliminations of higher pages. These take at most $\mathscr{O}(H^2)$ time for each generator in $\mathscr{E}^r_{p,q}$. If we do these for all generators simultaneously, since we need to update both rows and columns in a matrix, we might use `image_kernel` and the complexity becomes $\mathscr{O}(H^3)$. Denoting by L the infinity page, we have the new term

$$\left\lceil \frac{\dim(N^{\mathscr{U}})D_s}{P} \right\rceil \mathscr{O}(H^3)$$

which added to the complexity of the 1-page, we obtain

$$\left\lceil \frac{|\mathscr{U}|}{P} \right\rceil \mathscr{O}(X^2H) + \left\lceil \frac{|N_{\geq 1}^{\mathscr{U}}|}{P} \right\rceil \mathscr{O}(Y^2H) + \left\lceil \frac{\dim(N^{\mathscr{U}})D_s}{P} \right\rceil \left(\mathscr{O}(|N^{\mathscr{U}}|D_sX^2H) + \mathscr{O}(H^3) + \mathscr{O}(LH^3) \right)$$
$$= \left\lceil \frac{\dim(N^{\mathscr{U}})D_s}{P} \right\rceil \left(\mathscr{O}(|N^{\mathscr{U}}|D_sX^2H) + \mathscr{O}(H^3) \right).$$

Extension problem. If the spectral sequence collapses at $L > 0$, then the complexity of extending all generators in $\mathscr{E}^L_{p,q}$ is bounded by that of computing the L page about D_s times.

Overall complexity. Altogether, we have a complexity bounded by that of computing the first page plus that of computing the L page $L + D_s$ times. Here the L comes from computing the L page L times and D_s from the extension problem. Thus, the overal complexity is bounded by

$$\left\lceil \frac{|\mathscr{U}|}{P} \right\rceil \mathscr{O}(X^3) + \left\lceil \frac{|N_{\geq 1}^{\mathscr{U}}|}{P} \right\rceil \mathscr{O}(Y^3) + \left(L + D_s \right) \left\lceil \frac{\dim(N^{\mathscr{U}})D_s}{P} \right\rceil \left(\mathscr{O}(|N^{\mathscr{U}}|D_sX^2H) + \mathscr{O}(H^3) \right).$$

Notice that in general D_s, L and $\dim(N^{\mathscr{U}})$ are much smaller than H and X. Thus, for covers such that $Y \ll X$ and $|N^{\mathscr{U}}| \ll X$, and assuming we have enough processors, the complexity can be simplified to the two dominating terms

$$\mathscr{O}(X^3) + \mathscr{O}(H^3).$$

Notice that this last case is satisfied for those covers whose mutual intersections are generally smaller than each cover. Also, in this case H is approximately of the order of nontrivial barcodes over all the input complex. This shows that PERMAVISS isolates simplicial data, while only merging homological information. It is worth to notice that in general H, being the number of nontrivial bars, is much smaller than the size of the whole simplicial complex. However, in some cases this might not be true. Nevertheless our complexity estimates are very generous, leaving plenty of space for improvement on concrete applications.

6.5 Applications

We study three applications for our results. We will start analysing Vietoris-Rips complexes on point clouds. Afterwards we will work with cubical complexes, showing the great advantage these have on analysing images composed by pixels. Finally, we will study a mixture of both, the α-complexes. These will allow us to analyse the topology of point clouds, but with similar theoretical advantages than cubical complexes.

Vietoris-Rips complexes

Whenever X is finite, this defines a tame persistent module $\mathscr{S}_n(\mathrm{VR}_r(X))\colon \mathbf{R} \to \mathbf{vect}$, for each degree $n \geq 0$. Given a cover of P into sub point clouds $\mathscr{P} = \{P_i\}_{i=1}^N$, we would like to recover $\mathscr{S}_n(\mathrm{VR}_r(P))$ from the persistence modules $\{\mathscr{S}_n(\mathrm{VR}_r(P_i))\}_{i=1}^N$. This is possible along an interval $[0,R)$. The following proposition and corollary generalize Lemma 13 from [119].

Proposition 6.5.1. *Let $\mathscr{P} = \{P_i\}_{i=1}^N$ be a cover of P with $P = \bigcup_{i=1}^N P_i$, and define*

$$R \coloneqq \min_{0\leq i<j\leq N} \{d(P_i \setminus P_j, P_j \setminus P_i)\}.$$

Then $VR_r(P) = \bigcup_{i=1}^N VR_r(P_i)$ for all $r \in [0,R)$.

Proof. Suppose that $0 < r < R$ and we take a k-simplex $\sigma \in \mathrm{VR}_r(P)$. We want to prove that σ must be contained in some $\mathrm{VR}_r(P_j)$. For the sake of contradiction, we assume this is not the case. Also, assume that $\max_{1\leq i\leq N}\{|\sigma \cap P_i|\} = M$, and take i such that $|\sigma \cap P_i| = M$. Then, since σ is not contained in $\mathrm{VR}_r(P_i)$, there must exist some point $p \notin P_i$. On the other hand, since $\mathscr{P}$ is a cover of P, there exists some P_k such that $p \in P_k$. Now since $|\sigma \cap P_k| \leq M$, there must exist some point $q \in P_i$ such that $q \notin P_k$. Thus we have $p \in P_k \setminus P_i$ and $q \in P_i \setminus P_k$. Consequently, $d(p,q) \geq R$, and σ cannot be in $\mathrm{VR}_r(P)$, reaching a contradiction. □

In particular, we obtain the following convergence result:

Corollary 6.5.2. *Let $\mathscr{P}$ be a cover of P and let R be as in proposition 6.5.1. Then we have a spectral sequence*

$$E^2_{p,q} = \mathrm{H}_p(\mathscr{P}\mathscr{H}_q(\mathscr{P}), \bar{\delta}) \Rightarrow \mathrm{PH}_n(P)$$

which converges to $\mathrm{PH}_n(P)$ along the interval $[0,R)$. Here $\mathscr{P}\mathscr{H}_q$ is the q-persistent homology precosheaf on $N_\mathscr{V}$, which is defined by setting

$$\mathscr{P}\mathscr{H}_q(\sigma) \coloneqq \mathrm{PH}_q(P_\sigma)$$

for all $q \geq 0$, and all simplices $\sigma \in N_\mathscr{V}$. For any inclusion $\tau \subseteq \sigma$ of simplices in $N_\mathscr{V}$, we have an induced morphism:

$$\mathscr{P}\mathscr{H}_q(\tau \subseteq \sigma) \coloneqq \mathrm{PH}_q(\tau \subseteq \sigma)\colon \mathrm{PH}_q(P_\tau) \to \mathrm{PH}_q(P_\sigma).$$

Notice that exactly computing R might be expensive. In practice, we will divide P into convex regions from $\mathbb{R}^n$ with some overlap between mutually adjacent regions. More precisely, consider a finite covering $\mathscr{U} = \{U_i\}_{i=1}^N$ of $\mathbb{R}^n$ in such a way that each U_i is convex. One can, for instance,

consider hypercubes spanned by pairs of points $p,q \in \mathbb{R}^n$. With such a cover it is not difficult to compute the value $K \coloneqq \min_{0\leq i<j\leq N}\{d(U_i \setminus U_j, U_j \setminus U_i)\}$. Now considering the cover of point clouds defined as $\mathscr{P}_{\mathscr{U}} \coloneqq \{U_i \cap P\}_{i=1}^N$, one has that $K \leq R$. Thus, in this case one can use the previous results to recover $\mathrm{PH}_n(P)$ from $E^2_{p,q}$ along the intervals $[0,K)$. Vietoris-Rips complexes grow very quickly in size and this is the main reason why we need an upper bound for recovering its global information from local. Later on, we will see that α-complexes are much more well-behaved for parallelization.

Cubical Complexes over Lattices

In this section we review Mayer-Vietoris spectral sequences on cubical complexes; these are formed over a lattice subset $V \subseteq \mathbb{Z}^m$ as presented in example 2.5.4. Our motivation for studying these objects comes from image processing on grayscale images, as briefly mentioned on section 3.1. Suppose that one has a greyscale image, e.g. a two dimensional or three dimensional image. Since the image is greyscale, we can assign to each pixel (or voxel) a real number. Then we fix a threshold parameter $\rho > 0$ and proceed to define a cubical complex $\mathscr{C}_\rho$ on top of $V \subset \mathbb{Z}^m$. $\mathscr{C}_\rho$ will be defined to be the maximal cubical complex subject to two conditions:

1. All points in V with value less than ρ are exactly the vertices of $\mathscr{C}_\rho$.

2. A cube $q \in \mathscr{C}$ is contained in $\mathscr{C}_\rho$ if and only if all its vertices are also contained in $\mathscr{C}_\rho$.

Notice that in example 2.5.4 the cubes from $\mathscr{C}$ have unit-length factors. This is very useful, since it allows us to cover a cubical complex with a very small overlap between adjacent regions. This presents a great advantage in comparison to the case of Vietoris-Rips complexes, where we had to be careful with the maximum radius R. Given a pair of lattice points $p,q \in \mathbb{Z}^m$, we define the rectangular set

$$R(p,q) = \{x \in \mathbb{Z}^m \mid p_i \leq x_i \leq q_i, \text{ for all } 1 \leq i \leq m\}.$$

We will call $\mathbb{I} \subset \mathbb{Z}^m$ a rectangle if $\mathbb{I} = R(p,q)$ for some pair $p < q \in \mathbb{Z}^m$. Here the poset relation in $\mathbb{Z}^m$ is induced by the coordinatewise order. That is, we have that $p \leq q$ if and only if $p_i \leq q_i$ for all $1 \leq i \leq m$.

Proposition 6.5.3. *Let $\mathbb{I} \subset \mathbb{Z}^m$ be a filtered rectangle (if we want to consider a cubical complex that is not a rectangle we might assume that some of the cells from $\mathbb{I}$ take the infinity filtration value). Next, consider a cover $\mathscr{J} = \{\mathbb{I}_k = R(p^k,q^k)\}_{k=0}^N$ for $\mathbb{I}$ by rectangles with the same filtration. Suppose also that for any pair $1 \leq k,r \leq N$, one has that $(p^k)_l - (q^r)_l \neq 1$ for all $1 \leq l \leq m$. Then $\mathscr{C}(\mathbb{I}) = \bigcup_{k=1}^N \mathscr{C}(\mathbb{I}_k)$.*

Proof. By contradiction, suppose that a cube $c \in \mathscr{C}_n(\mathbb{I})$ is such that $c \notin \mathscr{C}_n(\mathbb{I}_k)$ for every $1 \leq k \leq N$. By definition, this means that not all vertices from c lie in some $\mathscr{C}(\mathbb{I}_k)$ for all $1 \leq k \leq N$. Since $\mathscr{J}$ forms a cover of $\mathbb{I}$, the vertex $c_{(0,0,\cdots,0)}$ from c lies in $\mathbb{I}_{k_1}$, for some $1 \leq k_1 \leq N$. Let $c_{j^1} = c_{(j^1_1, j^1_2, \ldots, j^1_m)}$ be the maximal vertex from c lying in $\mathbb{I}_{k_1}$, and $\mathscr{S}_1 \subseteq \{1,2,\ldots,m\}$ the subset of indices l satisfying $j^1_l = 0$. Notice that $(j^1_1, j^1_2, \ldots, j^1_m) \neq (1,1,\ldots,1)$ since otherwise c would be in $\mathbb{I}_{k_1}$. Then we consider $c_{1(\mathscr{S}_1)}$, where $1(\mathscr{S}_1)_l = 1$ if $l \in \mathscr{S}_1$ and $1(\mathscr{S}_1)_l = 0$ for $l \notin \mathscr{S}_1$. Since $(c_{1(\mathscr{S}_1)})_l > q^{k_1}_l$ for all $l \in \mathscr{S}_1$, we have that $c_{1(\mathscr{S}_1)}$ lies in some other $\mathbb{I}_{k_2}$, for $1 \leq k_2 \leq N$ with $k_2 \neq k_1$. Therefore we obtain $p^{k_2} \leq c_{1(\mathscr{S}_1)}$, and since $(p^{k_2})_l - (q^{k_1})_l \neq 1$ for all $1 \leq l \leq n$, we must have $(p^r)_l \leq (q^k)_l$ for all $l \in \mathscr{S}_1$. In particular, c_0 must be in $\mathbb{I}_{k_2}$. Let c_{j_2} be the maximal vertex from c lying in $\mathbb{I}^{k_2}$, again this cannot be c_1 since otherwise c would be contained in $\mathbb{I}_{k_2}$. Hence, we define the subset $\mathscr{S}_2 \subsetneq \{1,2,\ldots,m\}$ of indices l such that $(c_{j_2})_l = 0$. Notice that $\mathscr{S}_1 \cap \mathscr{S}_2 = \emptyset$ and we have that $\mathscr{S}_1 \subsetneq \mathscr{S}_1 \cup \mathscr{S}_2$. Then we consider the vertex $c_{1(\mathscr{S}_1 \cup \mathscr{S}_2)}$ and repeat the argument. Eventually there will be an index $t \geq 1$ such that $\mathscr{S}_1 \cup \cdots \cup \mathscr{S}_n = \{1,2,\ldots,m\}$. Then we consider $\mathbb{I}_r$ for some $1 \leq r \leq N$ containing c_1. For each coordinate index $1 \leq l \leq m$ we will find $(c_1)_l - (q^{k_i})_l = 1$ for some $1 \leq i \leq t$. Then c_0 must be contained in $\mathscr{I}_r$. Nevertheless this implies c is in $\mathbb{I}_r$, reaching a contradiction. □

Corollary 6.5.4. *Let $\mathbb{I} \subset \mathbb{Z}^m$ be a filtered rectangle, together with a cover $\mathscr{J} = \{\mathbb{I}_i \subset \mathbb{Z}^m\}_{i=0}^N$ by rectangles with the same filtration. Suppose also that for any pair $1 \leq k, r \leq N$, one has that $(p^k)_l - (q^r)_l \neq 1$ for all $1 \leq l \leq m$. Then there is a converging spectral sequence,*

$$E^2_{p,q} = \mathrm{H}_p(\mathscr{P}\mathscr{H}_q(\mathscr{J}), \bar{\delta}) \Rightarrow \mathrm{PH}_n(\mathbb{I}),$$

where $\mathscr{P}\mathscr{H}_q$ denotes the cubical q-persistent homology precosheaf on $N_{\mathscr{J}}$.

α-complexes

Let P be a point cloud contained in the Euclidean space $\mathbb{R}^n$. Given a point $q \in P$, recall from section 3.1 the definition of Voronoi cell V_q about the point q. Also, recall the definition of alpha complex $A_r(P)$ of radius $r > 0$. As mentioned in section 3.1, an advantage of using $A_r(P)$ instead of $\mathrm{VR}_r(P)$ is that the number of simplices gets reduced substantially. Given a subset $Q \subseteq P$ we define $A^P_r(Q)$ to be the maximal subcomplex of $A_r(P)$ with vertices contained in Q.

Proposition 6.5.5. *Consider two subpointclouds Q and S of a point cloud P. If the mutual complements of Voronoi cells are non-adjacent, that is, if*

$$\overline{V^P_{Q \setminus S}} \cap \overline{V^P_{S \setminus Q}} = \emptyset$$

then

$$A_r^P(Q \cup S) = A_r^P(Q) \cup A_r^P(S)$$

for all radii $r \geq 0$. Where we have used the notation $V_Q^P = \bigcup_{p \in Q} V_p^P$.

Proof. The inclusion $A_r^P(Q) \cup A_r^P(S) \subseteq A_r^P(Q \cup S)$ is direct by definition. Let us consider the opposite inclusion. By contradiction, suppose a simplex $\sigma \in A_r^P(Q \cup S)$ is not contained in $A_r^P(Q)$ or $A_r^P(S)$. Then there exists a pair of vertices $v, w \in \sigma$ lying on the mutual complements $v \in Q \setminus S$ and $w \in S \setminus Q$. By hypothesis, $V_v^P \cap V_w^P = \emptyset$ but then $[v, w] \subset \sigma$ cannot be an edge in $A_r^P(Q \cup S)$, and so $\sigma \notin A_r^P(Q \cup S)$, reaching a contradiction. □

Corollary 6.5.6. *Let $\mathscr{P} = \{P_i\}_{i=1}^N$ be a cover of P such that $\overline{V_{Q \setminus S}^P} \cap \overline{V_{S \setminus Q}^P} = \emptyset$ for all pairs $1 \leq i < j \leq N$, then*

$$A_r(P) = \bigcup_{i=1}^N A_r^P(P_i).$$

Furthermore, there is a convergent spectral sequence

$$E_{p,q}^2 = \check{\mathrm{H}}_p(N_{\mathscr{P}}, \mathscr{P}\mathscr{H}_q^{\alpha;P}(\mathscr{P})) \Rightarrow \mathrm{PH}_n^{\alpha}(P).$$

Where $\mathscr{P}\mathscr{H}_q^{\alpha,P}$ denotes the α-complex persistent homology precosheaf over $N_{\mathscr{P}}$. That is, for all $\sigma \in N_{\mathscr{P}}$ we have

$$\mathscr{P}\mathscr{H}_q^{\alpha,P}(\sigma) = \mathrm{PH}_q\left(A_*^P\left(\bigcap_{i \in \sigma} P_i\right)\right)$$

Proof. The first statement follows inductively from proposition 6.5.5. Applying the persistent Mayer-Vietoris spectral sequence leads to the result. □

Chapter 7

Regularly Filtered Diagrams

7.1 Regularly filtered CW-complexes

There are a number of situations in topological data analysis where it is not practical to work with filtered complexes. For example, the Vietoris-Rips complex on a point cloud $\mathrm{VR}_*(\mathbb{X})$ grows very quickly. Following on the collapsing idea from [52], we introduce a general object that will be of interest.

Definition 7.1.1. A functor $X : \mathbf{R} \to \mathbf{CW\text{-}cpx}$ where the connecting morphisms $X(r \leq s) : X_r \to X_s$ are regular for all $r \leq s$ from $\mathbf{R}$ will be called a *regularly filtered CW complex*. Given a pair of regular CW-complexes X and Y, we will consider natural transformations $f : X \to Y$ such that f_r is a regular morphism for all $r \in \mathbf{R}$. We denote the corresponding category by **RCW-cpx**.

The category of regularly filtered CW-complexes has similar definitions to that of persistence modules, as both categories contain functors from $\mathbf{R}$. We say that an object $X \in \mathbf{RCW\text{-}cpx}$ is *tame*, whenever X is constant along a finite number of right open intervals decomposing the poset $\mathbf{R}$. For $X \in \mathbf{RCW\text{-}cpx}$, we will write X_r for the regular CW-complex $X(r)$ for all $r \in \mathbf{R}$. We will also use the ε-shift notation $X[\varepsilon]$ and the shift functor $\Sigma^\varepsilon : \mathbf{RCW\text{-}cpx} \to \mathrm{Hom}(\mathbf{RCW\text{-}cpx})$ for all $\varepsilon \geq 0$.

If the morphisms $X(r \leq s) : X_r \to X_s$ are injections preserving the cellular structure for all $r \leq s$ in $\mathbf{R}$, then X is a filtered CW-complex, denoting by **FCW-cpx** the corresponding subcategory of **RCW-cpx**. Notice that objects in **FCW-cpx** can be seen as a pair $(\mathbf{colim}\, X_*, f)$ where $\mathbf{colim}\, X_*$ is a regular CW-complex and $f : \mathbf{colim}\, X_* \to \mathbb{R}$ is a *filtration function* such that $X_r = f^{-1}(r)$ for all values $r \in \mathbf{R}$.

Definition 7.1.2. Let $X \in \mathbf{RCW\text{-}cpx}$. We define the n-persistent homology of X as the composed functor $\mathrm{H}_n^{\mathrm{cell}}(X) : \mathbf{R} \to \mathbf{Vect}$. We will also denote this by $\mathrm{PH}_n(X)$ for all integers $n \geq 0$.

Notice that by finiteness of X_r, the vector space $\mathrm{PH}_n(X)_r$ is finite dimensional for all $r \in \mathbf{R}$. If in addition X is tame, $\mathrm{PH}_n(X)$ only changes at a finite number of points $r \in \mathbf{R}$.

Remark. Notice that the standard algorithm for the computation of persistent homology cannot be applied to objects in **RCW-cpx**. However, there are some cases which can be easily computed by using the results from chapter 4. For example, consider $X_* \in$ **RCW-cpx** such that for any value $r \in \mathbf{R}$ and any cell $e^n_r \in X_r$, the structure morphisms $X_*(r \leq s)$ send e^n_r to an n-cell $e^n_s \in X_s$ or to a lower dimensional subcomplex from X_s, for all pairs $r < s$ from $\mathbb{R}$. In this case, these "consistent" cells e^n_* determine a barcode basis for the chain complex $C_n(X)$ associated to X_*. Notice that if we allow more general regular morphisms $X_*(r \leq s)$ then we might need to do more work, such as using the quotients constructed on section 4.4.

7.2 Filtered Diagrams and Filtered Geometric Realizations

Recall that in section 2.8 we saw regular diagrams as well as their associated geometric realizations. We might adapt this to diagrams of regularly filtered complexes.

Definition 7.2.1. A *regularly filtered regular diagram of CW-complexes* $\mathscr{D}$ over K is a functor $\mathscr{D} : K^{\mathrm{op}} \to$ **RCW-cpx**, we will denote this category by **RRDiag**(K). A morphism $f : \mathscr{D} \to \mathscr{L}$ between a pair of diagrams over K consists of natural transformations $f : \mathscr{D}(\sigma) \to \mathscr{L}(\sigma)$ for all $\sigma \in K$ such that $f \circ \mathscr{D}(\tau \prec \sigma) = \mathscr{L}(\tau \prec \sigma) \circ f$ for all $\tau \prec \sigma$ in K. On the other hand if we restrict to functors $\mathscr{D} : K^{\mathrm{op}} \to$ **FCW-cpx** we will call $\mathscr{D}$ a *filtered regular diagram of CW-complexes* denoting the corresponding category by **FRDiag**(K). If for a diagram $\mathscr{D} \in$ **FRDiag**(K) the maps $\mathscr{D}(\tau \prec \sigma)$ are inclusions respecting the cellular structures for all $\tau \prec \sigma$ from K, then we call $\mathscr{D}$ a *fully filtered diagram of CW-complexes* denoting the corresponding category by **FFDiag**(K). We have embeddings of categories

$$\mathbf{FFDiag}(K) \subset \mathbf{FRDiag}(K) \subset \mathbf{RRDiag}(K)$$

for all simplicial complexes K.

Example 7.2.2. Consider a filtered CW-complex X covered by filtered subcomplexes $\mathscr{U} = \{U_i\}_{i \in I}$. For each $r \in \mathbf{R}$, one might consider a regular diagram $X^{\mathscr{U}}(r) : N^{\mathrm{op}}_{\mathscr{U}} \to$ **CW-cpx** as explained on example 2.3.2. This leads to a fully filtered diagram $X^{\mathscr{U}} : N^{\mathrm{op}}_{\mathscr{U}} \to$ **FCW-cpx**. On the other hand, there is also a constant diagram $*^{\mathscr{U}} \in$ **FFDiag**$(N_{\mathscr{U}})$ given by $*^{\mathscr{U}}(J)_r = *$ if $X^{\mathscr{U}}(J)_r \neq \emptyset$ or $*^{\mathscr{U}}(J)_r = \emptyset$ otherwise; for all $J \in N_{\mathscr{U}}$ and all $r \in \mathbf{R}$. Then, there is an obvious epimorphism of diagrams $X^{\mathscr{U}} \to *^{\mathscr{U}}$. Continuing with the same example, we consider composing $X^{\mathscr{U}}$ with π_0 as in example 2.8.5. This leads to $\pi_0(X^{\mathscr{U}}) \in$ **RRDiag**(K). Each $\pi_0(\mathscr{U}_J)$ is a disjoint union of points that

are identified with each other as the filtration value increases. Altogether we have a sequence of epimorphisms of diagrams $X^{\mathscr{U}} \to \pi_0(X^{\mathscr{U}}) \to *^{\mathscr{U}}$.

Recall definition 2.8.4 about the geometric realizations for regular diagrams. Notice that this can be extended to a geometric realization $\Delta_K\mathscr{D}$ of a diagram $\mathscr{D} \in \mathbf{RRDiag}(K)$, by setting $(\Delta_K\mathscr{D})_r \coloneqq \Delta_K(\mathscr{D}_r)$ for all $r \in \mathbf{R}$. The corresponding gluing conditions from definition 2.8.4 are consistent in this case as

$$\mathscr{D}(\tau \preceq \sigma) \circ \Sigma^t\mathscr{D}(\sigma)(y) = \Sigma^t\mathscr{D}(\tau) \circ \mathscr{D}(\tau \preceq \sigma)(y)$$

for any pair $\tau \preceq \sigma$ from K and all $t > 0$ and all points $y \in \mathscr{D}(\sigma)$. Altogether we obtain $\Delta_K(\mathscr{D}) \in$ **RCW-cpx**.

Example 7.2.3. Consider a diagram $\mathscr{D} \in \mathbf{RRDiag}(K)$. As with regular diagrams, we can also define the constant diagram, which is given by

$$*^{\mathscr{D}}(\sigma)_r = \begin{cases} * & \text{if } \mathscr{D}(\sigma)_r \neq \emptyset \\ \emptyset & \text{else .} \end{cases}$$

Note that there is a homotopy equivalence $\Delta(*^{\mathscr{D}})_r \simeq |K_r^{\mathscr{D}}|$, where $K^{\mathscr{D}}$ is the filtered simplicial complex with the same underlying vertex set as K and $\sigma \in K_r^{\mathscr{D}}$ if and only if $\mathscr{D}(\sigma)_r \neq \emptyset$. On the other hand, we define the *filtered multinerve* of $\mathscr{D}$ as

$$\mathrm{MNerv}(\mathscr{D}) = \Delta(\pi_0(\mathscr{D})) \ .$$

There are epimorphisms $\Delta\mathscr{D} \to \mathrm{MNerv}(\mathscr{D}) \to \Delta(*^{\mathscr{D}}) \simeq |K|$.

One might also deduce analogous results to Theorem 2.8.6. An interesting direction of research would be to use this result to define compatible complex simplifications, such as in Discrete Morse Theory (see [86] and [3]) and end up with a diagram of regular CW-complexes. This motivates the study of spectral sequences associated to such diagrams. We will see further reasons in Section 7.3. On the other hand, given the importance of Theorem 2.8.6, we would like to adapt it to an approximate version in the context of diagrams in $\mathbf{RRDiag}(K)$. Instead of studying homotopy equivalences, we will consider equivalences induced by acyclic carriers. This will be done in Section 8.2.

7.3 Spectral Sequences for Geometric Realizations

Recall the persistent Mayer-Vietoris spectral sequence from chapter 6 associated to a pair $(X, \mathscr{U})$ of a space with a cover:

$$E^1_{p,q}(X,\mathscr{U}) = \bigoplus_{\sigma^p \in N^{\mathscr{U}}} \mathrm{PH}_q(X^{\mathscr{U}}(\sigma^p)) \Rightarrow \mathrm{PH}_{p+q}(\Delta X^{\mathscr{U}}) \simeq \mathrm{PH}_{p+q}(X)\ . \tag{7.1}$$

As pointed out in section 6.5, there are some limitations to the applicability of this spectral sequence to the computation of persistent homology on Vietoris-Rips complexes; it is only possible to recover $\mathrm{PH}_*(X)$ up to some value $R > 0$ determined by the cover overlaps. In section 7.4 we will present an alternative regular diagram of CW-complexes that avoids this upper limit problem completely, see example 7.4.4.

We will proceed by first defining spectral sequences associated to regular diagrams, leaving the case of filtrations for later; in particular, we will first define the cellular complex associated to a geometric realization. Given a diagram $\mathscr{D}$ in $\mathbf{RDiag}(K)$, we will denote by $\mathscr{D}(\tau \preceq \sigma)_*$ the induced morphism of cellular chain complexes $C^{\mathrm{cell}}_*(\mathscr{D}(\sigma)) \to C^{\mathrm{cell}}_*(\mathscr{D}(\tau))$. The cellular chain complex $C^{\mathrm{cell}}_*(\Delta\mathscr{D}, \delta^\Delta)$ associated to $\Delta\mathscr{D}$ is defined as follows: For all $m \geq 0$ we have that $C^{\mathrm{cell}}_m(\Delta\mathscr{D})$ is a vector space generated by products of cells $\sigma^p \times c^q$ with $\dim(\sigma^p) = p$ and $c^q \in \mathscr{D}(\sigma)_q$ so that $p+q = m$. On such a product cell $\sigma^p \times c^q$ the differential δ^Δ is given by

$$\delta^\Delta(\sigma^p \times c^q) = \sum_{\sigma^{p-1}_i \prec \sigma^p} (-1)^i \left(\sum_{a \in \overline{\mathscr{D}(\sigma^{p-1}_i \preceq \sigma^p)(c^q)}} [a : \mathscr{D}(\sigma^{p-1}_i \preceq \sigma^p)(c^q)] \sigma^{p-1}_i \times a \right) + (-1)^p \sum_{b \in \overline{c^q} \setminus c^q} [b : c^q] \sigma^p \times b$$

where d_q is the differential $d_q : C_q(\mathscr{D}(\sigma^p)) \to C_{q-1}(\mathscr{D}(\sigma^p))$ and the sum runs over the faces σ^{p-1}_i of σ^p. As we will see in the proof of Lemma 7.3.1 the map δ^Δ is indeed a differential. In addition, notice that the filtration of $\Delta_K(\mathscr{D})$ carries over to $C^{\mathrm{cell}}_*(\Delta_K\mathscr{D})$ by taking $F^pC_*(\Delta_K\mathscr{D}) \coloneqq C_*(F^p\Delta_K\mathscr{D})$ for all $p \geq 0$.

Now, consider the double complex $(C_{p,q}(\mathscr{D}), d^V, d^H)$ given by

$$C_{p,q}(\mathscr{D}) = \bigoplus_{\sigma^p \in K} C^{\mathrm{cell}}_q\big(\mathscr{D}(\sigma^p)\big)$$

for all $p, q \geq 0$. The vertical differential is defined by the direct sum of chain differentials $d^V_{p,q} = (-1)^p \bigoplus_{\sigma^p \in K} d^{\sigma^p}_q$ where $d^{\sigma^p}_*$ denotes the differential from $C^{\mathrm{cell}}_*(\mathscr{D}(\sigma^p))$ for all $\sigma^p \in K$. The horizontal differential is given by the Čech differential $d^H_{p,q}$ which is defined for a cell $a^q \in \mathscr{D}(\sigma)_q$ as $\sum_{\sigma^{p-1}_i \prec \sigma^p} (-1)^i \mathscr{D}(\sigma^{p-1}_i \prec \sigma^p)^*(a^q)$, where $\mathscr{D}(\sigma^{p-1}_i \prec \sigma^p)^*$ denotes the induced chain mor-

phism $C_*^{\text{cell}}(\mathscr{D}(\sigma^p)) \to C_*^{\text{cell}}(\mathscr{D}(\sigma_i^{p-1}))$ for all faces σ_i^{p-1} from σ^p. Of course $d^V \circ d^V = 0$ and $d^H \circ d^H = 0$ by functoriality of $C_*^{\text{cell}}(\cdot)$ and the fact that $\mathscr{D}(\rho \prec \tau)\mathscr{D}(\tau \prec \sigma) = \mathscr{D}(\rho \prec \sigma)$ for any three simplices $\rho \prec \tau \prec \sigma$ from K. On the other hand, anticommutativity $d^V \circ d^H = -d^H \circ d^V$ follows since $\mathscr{D}(\tau \prec \sigma)^*$ is a chain morphism for all cell pairs $\tau \prec \sigma$ from K.

Now, we consider the double complex spectral sequence from [83, Section 2.4]. Given $\mathscr{D}$ in $\mathbf{RDiag}(K)$ there is a spectral sequence

$$E^1_{p,q}(\mathscr{D}) = \bigoplus_{\sigma^p \in K} \mathrm{H}_q(\mathscr{D}(\sigma^p)) \Rightarrow \mathrm{H}_{p+q}(S_*^{\text{Tot}}(\mathscr{D}))$$

where $S^{\text{Tot}}(\mathscr{D})$ is the *total complex* defined as $S_n^{\text{Tot}}(\mathscr{D}) = \bigoplus_{p+q=n} C_{p,q}(\mathscr{D})$ together with a differential $d^{\text{Tot}} = d^V + d^H$. Also, recall from section 2.11 that the total complex has a filtration induced by the vertical filtration on $C_{p,q}(\mathscr{D})$ given by

$$F^m S_*^{\text{Tot}}(\mathscr{D}) = \bigoplus_{\substack{p+q=n \\ p \leq m}} C_{p,q}(\mathscr{D})$$

for all integers $m \geq 0$. We will now relate this total complex to the geometric realization from Definition 2.8.4.

Lemma 7.3.1. *There is an isomorphism $C_*^{\text{cell}}(\Delta\mathscr{D}, \delta^\Delta) \simeq S_*^{\text{Tot}}(\mathscr{D})$ which preserves filtration. That is, $F^p C_*^{\text{cell}}(\Delta\mathscr{D}, \delta^\Delta) \simeq F^p S_*^{\text{Tot}}(\mathscr{D})$ for all $p \geq 0$.*

Proof. First we define a chain morphism $\psi : C_m^{\text{cell}}(\Delta\mathscr{D}) \to S_m^{\text{Tot}}(\mathscr{D})$ generated by the assignment: a cell $\sigma^p \times c^q \in (\Delta\mathscr{D})_m$ with $p+q=m$, is sent to $\psi(\sigma^p \times c^q) = (c^q)_{\sigma^p} \in S_m^{\text{Tot}}(\mathscr{D})$. Here we denote by $(c^q)_{\sigma^p}$ the vector in $S_m^{\text{Tot}}(\mathscr{D})$ that is zero everywhere except at the entry $C_q^{\text{cell}}(\mathscr{D}(\sigma^p))$, where it contains the chain c^q. On the other hand, ψ is a chain morphism since we have the equality

$$\begin{aligned}
\psi\big(\delta^\Delta(\sigma^p \times c^q)\big) &= \sum_{\sigma_i^{p-1} \prec \sigma^p} (-1)^i \Big(\sum_{a \in \overline{\mathscr{D}(\sigma_i^{p-1} \prec \sigma^p)(c^q)}} ([a : \mathscr{D}(\sigma_i^{p-1} \prec \sigma^p)(c^q)]a)_{\sigma_i^{p-1}} \Big) \\
+ (-1)^{\dim(\sigma^p)} \sum_{b \in \overline{c^q} \setminus c^q} ([b : c^q]b)_{\sigma^p} &= \sum_{\sigma_i^{p-1} \prec \sigma^p} (-1)^i (\mathscr{D}(\sigma_i^{p-1} \prec \sigma^p)^*(c^q))_{\sigma_i^{p-1}} + (-1)^p (d_q^{\sigma^p}(c^q))_{\sigma^p} \\
&= (d^H + d^V)((c^q)_{\sigma^p}) = d^{\text{Tot}}((c^q)_{\sigma^p}) = d^{\text{Tot}}(\psi(\sigma^p \times c^q)).
\end{aligned}$$

One can see that ψ is injective, and admits an inverse $\psi^{-1} : S_m^{\text{Tot}}(\mathscr{D}) \to C_m^{\text{cell}}(\Delta\mathscr{D})$ that sends $(\sigma^p)_{c^q}$ to $\sigma^p \times c^q$. Notice that by definition ψ sends a chain in $F^p C_n^{\text{cell}}(\Delta\mathscr{D})$ to a chain in $F^p S_n^{\text{Tot}}(\mathscr{D})$ for all $p \geq 0$ and in particular it preserves filtration. □

7.4 The $(K, \mathscr{P})$-Join Diagram Spectral Sequence

In this section we will present an important spectral sequence related to filtered simplicial complexes. For this, we will start by reviewing the join construction.

Definition 7.4.1. We will denote by $K * L$ the *join* of two simplicial complexes K and L. This is given by a quotient on the space $K \times L \times \Delta^1$, collapsing the subspace $K \times L \times \{0\}$ to K and collapsing also $K \times L \times \{1\}$ to L. Similarly, given simplicial complexes K_i for $0 \leq i \leq m$, the iterative join is defined as the quotient space

$$K_0 * K_1 * \cdots * K_m \coloneqq K_0 \times K_1 \times \cdots \times K_m \times \Delta^m \Big/ \sim$$

where the relation $\sim$ identifies two elements

$$(a_0, a_1, \ldots, a_m, \sum_{i=0}^{m} x_i) \sim (b_0, b_1, \ldots, b_m, \sum_{i=0}^{m} y_i)$$

whenever $x_i = y_i$ for all $0 \leq i \leq m$ and if $a_i = b_i$ for all indices $0 \leq i \leq m$ such that $x_i \neq 0$.

The join will be useful for constructing a diagram of spaces whose geometric realization is homeomorphic to $|K|$ for any finite simplicial complex K. In order to proceed, we take a finite partition $\mathscr{P}$ of the vertex set $V(K)$ and denote by $K(U)$ the maximal subcomplex of K with vertices in $U \in \mathscr{P}$. We will denote by $\Delta^{\mathscr{P}}$ the standard simplex with vertices in $\mathscr{P}$. For a simplex $\tau \in K$, we define $\mathscr{P}(\tau) \in \Delta^{\mathscr{P}}$ to be the simplex consisting of all partitioning sets $U \in \mathscr{P}(\tau)$ such that $\tau \cap U \neq \emptyset$. In particular if $U \in \mathscr{P}(\tau)$, then it determines a standard simplex $\tau(U) \in K(U)$ of dimension $|\tau \cap U| - 1 \geq 0$. For a vertex $v \in K$, we will denote by $\mathscr{P}(v)$ the partitioning set from $\mathscr{P}$ which contains v.

We define the $(K, \mathscr{P})$-*join* diagram $\mathscr{J}^K_{\mathscr{P}} : \Delta^{\mathscr{P}} \to$ **FCW-cpx** for all $\sigma \subseteq \mathscr{P}$ by assigning the subspace

$$\mathscr{J}^K_P(\sigma) = \bigcup_{\substack{\rho \in K \\ \mathscr{P}(\rho) = \sigma}} \operatorname{Im}\left(\prod_{U \in \sigma} \Delta^{\rho(U)} \hookrightarrow \prod_{U \in \sigma} |K(U)| \right),$$

for all $\sigma \in \Delta^{\mathscr{P}}$. For any pair $\tau \preceq \sigma$ in $\Delta^{\mathscr{P}}$, the projection $\pi_{\tau \preceq \sigma} : \prod_{U \in \sigma} |K(U)| \to \prod_{U \in \tau} |K(U)|$ induces a regular morphism $\mathscr{J}^K_{\mathscr{P}}(\tau \preceq \sigma) : \mathscr{J}^K_{\mathscr{P}}(\sigma) \to \mathscr{J}^K_{\mathscr{P}}(\tau)$.

Lemma 7.4.2. *Let K be a simplicial complex together with a partition $\mathscr{P}$ of its vertex set $V(K)$. There is a CW-complex homeomorphism $\Delta(\mathscr{J}^K_{\mathscr{P}}) \simeq |K|$.*

Proof. Consider the continuous map $f : \Delta(\mathscr{J}^K_{\mathscr{P}}) \to |K|$ defined by mapping a point

$$\Big(\sum_{U \in \mathscr{P}(\tau)} y_U U, \Big(\sum_{v \in U} x_v v \Big)_{U \in \mathscr{P}(\tau)} \Big) \in \Delta^{\mathscr{P}(\tau)} \times \prod_{U \in \mathscr{P}(\tau)} \Delta^{\tau(U)} \Big/ \sim$$

to $\sum_{v\in\tau} y_{\mathscr{P}(v)} x_v v$ in Δ^τ for all $\tau \in K$; where we have values $0 \le y_U \le 1$ and $0 \le x_v \le 1$ for all $U \in \mathscr{P}(\tau)$ and all $v \in U$, and such that $\sum_{U\in\mathscr{P}(\tau)} y_U = 1$ and $\sum_{v\in U} x_v = 1$ for all $U \in \mathscr{P}$. On the other hand, let $\sum_{v\in\tau} x_v v \in \Delta^\tau$ be a point such that $0 \le x_v \le 1$ for all $v \in \Delta^\tau$ and such that $\sum_{v\in\tau} x_v = 1$. Then we can define the inverse continuous map

$$f^{-1}\Big(\sum_{v\in\tau} x_v v\Big) = \Big(\sum_{U\in\mathscr{P}(\tau)} \Big(\sum_{v\in U} x_v\Big) U, \Big(\psi_U\Big(\sum_{v\in\tau} x_v v\Big)\Big)_{U\in\mathscr{P}(\tau)}\Big)$$

where

$$\psi_U\Big(\sum_{v\in\tau} x_v v\Big) = \begin{cases} \sum_{v\in\tau}\Big(\dfrac{x_v}{\sum_{v\in\tau(U)} x_v}\Big) v \text{ if } \sum_{v\in\tau(U)} x_v \neq 0 \\ * \in \Delta^{\tau(U)} \text{ otherwise (see below),} \end{cases}.$$

If $\sum_{v\in\tau(U)} x_v = 0$, then $x_v = 0$ for all $v \in \tau(U)$ and the U-coordinate of the simplex τ in $\Delta^{\mathscr{P}(\tau)}$ is 0, which means that $\tau(U)$ is collapsed to a single point by the equivalence relation used to define $\Delta(\mathscr{J}^K_{\mathscr{P}})$. It is straightforward to check that f and f^{-1} are well defined and consistent along K. □

Example 7.4.3. Consider a simplicial complex K depicted in the top left part of Figure 7.1. We consider a partition of the vertex set of K into the two subsets $\mathscr{P} = \{U,V\}$, where points in U are indicated by black circles and points in V are indicated by red squares. In the top right of Figure 7.1, we depict the standard 1-simplex $\Delta^{\mathscr{P}}$ together with the diagram $\mathscr{J}^K_{\mathscr{P}}$ over it. In particular, notice that $\mathscr{J}^K_{\mathscr{P}}([U,V])$ is a subset of the product $|K(U)| \times |K(V)|$ and that the morphisms $\mathscr{J}^K_{\mathscr{P}}([U,V]) \to \mathscr{J}^K_{\mathscr{P}}(V) = |K(V)|$ and $\mathscr{J}^K_{\mathscr{P}}([U,V]) \to \mathscr{J}^K_{\mathscr{P}}(U) = |K(U)|$ are both projections. Finally, the bottom left of Figure 7.1 shows the geometric realization $\Delta\mathscr{J}^K_{\mathscr{P}}$, where each green line and each red dashed line gets "squashed" to a single point.

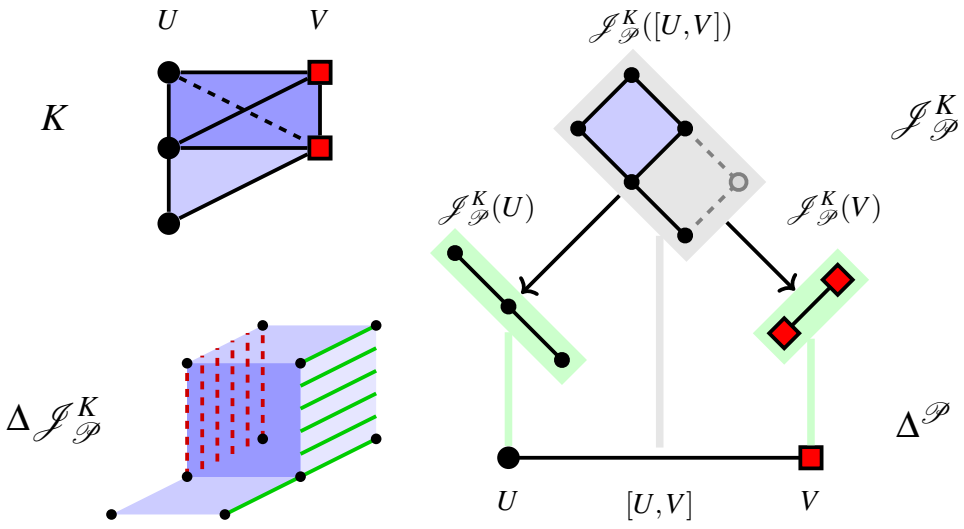


Figure 7.1: Illustration of an example of a $\mathscr{J}^K_{\mathscr{P}}$ diagram.

Note that $\mathscr{J}^K_{\mathscr{P}}$ is not a diagram of simplicial complexes, but of *prodsimplicial* complexes as in [73, Def. 2.43]. See Figure 7.2 for an example of $\mathscr{J}^K_{\mathscr{P}}$ where $\mathscr{P}$ partitions the vertex set of a 7-simplex. In particular, one can consider the spectral sequence

$$E^1_{p,q}(\mathscr{J}^K_{\mathscr{P}}) = \bigoplus_{\sigma\in\Delta^{\mathscr{P}}} \mathrm{H}_q(\mathscr{J}^K_{\mathscr{P}}(\sigma)) \Rightarrow \mathrm{H}_{p+q}(K)\ .$$

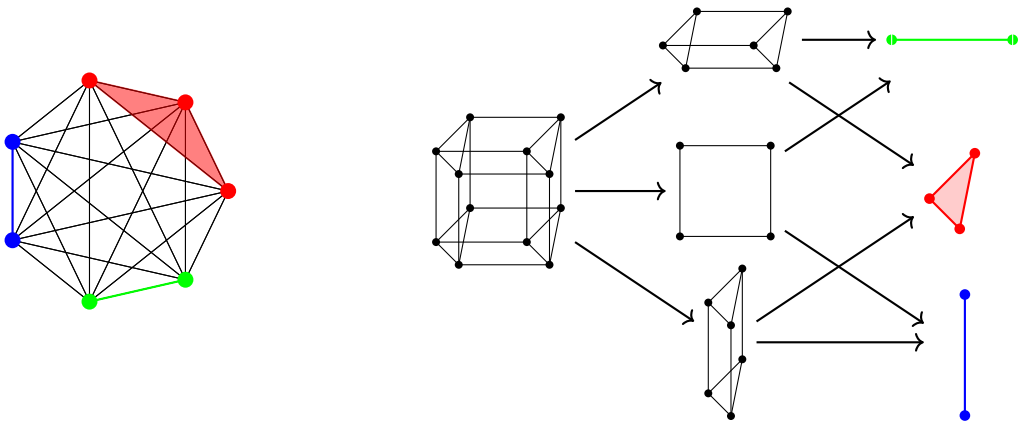


Figure 7.2: On the left the 7-simplex together with a partition P of its vertex set. On the right, the associated diagram of spaces $\mathscr{J}^K_P$.

Now, let us consider a filtered simplicial complex $K_* \in \mathbf{FCW\text{-}cpx}$ such that its vertex set $V(K_*)$ is fixed throughout all values of $\mathbf{R}$. Let $\mathscr{P}$ be a partition of $V(K_*)$. We define the filtered regular diagram $\mathscr{J}^K_{\mathscr{P}} \in \mathbf{FRDiag}(\mathscr{P})$ by sending $r \in \mathbf{R}$ to $\mathscr{J}^{K_r}_{\mathscr{P}}$. These diagrams inherit the shift morphisms ΣK_* from K_* in the following way: Let $\sigma \in \Delta^{\mathscr{P}}$ and notice that we have restrictions $\Sigma^{s-r}K_{|U} : |K_r(U)| \to |K_s(U)|$ for all $U \in \sigma$ and all $s \geq r$, so that we have induced morphisms

$$\prod_{U\in\sigma} \Sigma^{s-r}K_{|U} : \mathscr{J}^{K_r}_{\mathscr{P}}(\sigma) \to \mathscr{J}^{K_s}_{\mathscr{P}}(\sigma)$$

for all $\sigma \in \Delta^{\mathscr{P}}$. In turn, these induce a shift morphism on $\Delta J^K_{\mathscr{P}}$ which respect filtrations, so that we have a commutative diagram

$$\begin{array}{ccccc}
E^*_{p,q}(J^{K_r}_{\mathscr{P}}) & \Longrightarrow & \mathrm{PH}_{p+q}\left(\Delta\mathscr{J}^{K_r}_{\mathscr{P}}\right) & \xrightarrow{\simeq} & \mathrm{PH}_{p+q}\left(K_r\right) \\
\downarrow & & \downarrow & & \downarrow \\
E^*_{p,q}(J^{K_s}_{\mathscr{P}}) & \Longrightarrow & \mathrm{PH}_{p+q}\left(\Delta\mathscr{J}^{K_s}_{\mathscr{P}}\right) & \xrightarrow{\simeq} & \mathrm{PH}_{p+q}\left(K_s\right)
\end{array}$$

and thus $\mathrm{PH}_*(\Delta\mathscr{J}^K_{\mathscr{P}}) \simeq \mathrm{PH}_*(K_*)$. For each simplex $\sigma \in \Delta^{\mathscr{P}}$ one can see $\mathscr{J}^K_{\mathscr{P}}(\sigma)$ as a filtered

simplicial complex, so that

$$E^1_{p,q}(\mathscr{J}^K_{\mathscr{P}}) = \bigoplus_{\sigma \in (\Delta^{\mathscr{P}})^p} \mathrm{PH}_q(\mathscr{J}^K_{\mathscr{P}}(\sigma)) \Rightarrow \mathrm{PH}_{p+q}(K) \ .$$

Example 7.4.4. Consider a point cloud $\mathbb{X}$, a partition $\mathscr{P}$ and consider its Vietoris Rips complex $\mathrm{VR}_*(\mathbb{X}) \in$ **FCW-cpx**. In this case we have a fixed partition of the vertex set of $\mathrm{VR}_*(\mathbb{X})$, which allows us to consider the spectral sequence:

$$E^1_{p,q}\big(\mathscr{J}^{\mathrm{VR}_*(\mathbb{X})}_{\mathscr{P}}\big) = \bigoplus_{\sigma \in \Delta^{\mathscr{P}}} \mathrm{PH}_q(\mathscr{J}^{\mathrm{VR}_*(\mathbb{X})}_{\mathscr{P}}(\sigma)) \Rightarrow \mathrm{PH}_{p+q}(\mathrm{VR}_*(\mathbb{X})) \ .$$

This is very convenient as it avoids the main difficulties with the Mayer-Vietoris blowup complex associated to a cover. Namely, one recovers $\mathrm{PH}_*(K)$ completely without any bounds depending on the cover overlaps. In addition, notice that $\Delta\mathscr{J}^{\mathrm{VR}_*(\mathbb{X})}_{\mathscr{P}}$ does not 'blowup' more than $\mathrm{VR}_*(\mathbb{X})$, as both are homeomorphic by Lemma 7.4.2.

The $(K, \mathscr{P})$-join diagram is related to [101, Example 4]. There the motivation behind the filtrations is given by a consistency radius and a filtration based on the differences between local measurements. A similar example appears (without a filtration) as one of the opening examples in [68, Appendix 4.G].

Chapter 8

Interleaving Mayer-Vietoris spectral sequences

In this chapter we will explore persistence spectral sequences as invariants. First we will introduce in section 8.1 a general framework to obtain interleavings in the category **RCW-cpx**. Then, we will focus on the case of regularly filtered regular diagrams, and present which compatibility conditions lead to stability of the geometric realization, see section 8.2. These conditions carry over to the stability results on the first page of spectral sequences that are presented on section 8.3. The next two sections contain a study of stability of the persistence Mayer-Vietoris spectral sequence for changing the chosen covers.

8.1 ε-acyclic carriers

The following definition will encode our notion of 'noise'.

Definition 8.1.1. Let $X, Y \in$ **RCW-cpx**. An *ε-acyclic carrier* $F^{\varepsilon}_{*} : X_{*} \rightrightarrows Y[\varepsilon]_{*}$ is a family of acyclic carriers $F^{\varepsilon}_{a} : X_{a} \rightrightarrows Y_{a+\varepsilon}$ for all $a \in \mathbf{R}$ such that

$$Y(a+\varepsilon \leq b+\varepsilon)F^{\varepsilon}_{a}(c) \subseteq F^{\varepsilon}_{b}(X(a \leq b)(c))$$

for all cells c of X_a and $a, b \in \mathbf{R}$ with $a \leq b$.

The proposition below is an adaptation of [85, Thm. 13.4] or [41, Lem. 2.4] to the context of tame filtered CW-complexes.

Proposition 8.1.2. *Let $X_{*}, Y_{*} \in$ **FCW-cpx** be tame, and assume that there exists an ε-acyclic carrier*

$$F^{\varepsilon}_{*} : X_{*} \rightrightarrows Y[\varepsilon]_{*} \ .$$

Then there exist chain morphisms $f_a^\varepsilon : C_*(X_a) \to C_*(Y_{a+\varepsilon})$ carried by F_a^ε for all $a \in \mathbf{R}$, so that $Y(a+\varepsilon \le b+\varepsilon) \circ f_a^\varepsilon = f_b^\varepsilon \circ X(a \le b)$. Furthermore, given another such sequence of morphisms $g_a^\varepsilon : C_*(X_a) \to C_*(Y_{a+\varepsilon})$, there exist chain homotopy equivalences $H_a^\varepsilon : g_a^\varepsilon \simeq f_a^\varepsilon$ which are carried by F_a^ε for all $a \in \mathbf{R}$.

Proof. Let $b \in \mathbf{R}$ and assume that f_a^ε has already been defined for all values $a < b$, where we allow for $b = -\infty$. We first define f_b^ε on all cells which are in the image of $X(a < b)$ for any $a < b$ using the definition

$$f_b^\varepsilon \circ X(a < b) = Y(a+\varepsilon < b+\varepsilon) \circ f_a^\varepsilon \ .$$

Notice that the assumption that $X_a \subseteq X_b$ is crucial for this to work. By hypotheses, given a cell $c \in \mathrm{Im}(X(a < b))$, its image $f_b^\varepsilon(c)$ is then contained in

$$Y(a+\varepsilon < b+\varepsilon) F_a^\varepsilon(\widetilde{c}) \subseteq F_b^\varepsilon(X(a < b)(\widetilde{c})) \ ,$$

where $\widetilde{c} \in X_a$ is such that $c = X(a < b)(\widetilde{c})$. Hence, f_b^ε satisfies the carrier condition. Next we define f_b^ε on the remaining cells in

$$\widetilde{X_b} = X_b \setminus \big(\bigcup_{a<b} X(a < b) \big) \ .$$

We proceed to prove this by induction. First, choose a 0-cell $f_b^\varepsilon(v) \in F_b^\varepsilon(v)$ for each remaining 0-cell $v \in \widetilde{X_b}$, and notice that $d_* f_b^\varepsilon(v) = 0 = f_b^\varepsilon(d_* v)$, where we use d_* for the chain complex differentials. Next, by induction, assume that for a fixed $p \ge 0$, the p-cells $s \in X_b$ have image $f_b^\varepsilon(s)$ carried by $F_b^\varepsilon(s)$ and such that $d_* \circ f_b^\varepsilon(s) = f_b^\varepsilon \circ d_*(s)$. We would like to extend f_b^ε to the $(p+1)$-cells. By semicountinuity, given such a cell $c \in X_b$, its boundary $d_* c$ will be contained in $F_b^\varepsilon(c)$. On the other hand, notice that by linearity and the induction hypotheses $d_* f_b^\varepsilon(d_* c) = f_b^\varepsilon(d_* d_* c) = 0$, thus $f_b^\varepsilon(d_* c)$ is a cycle in $F_b^\varepsilon(c)$. By acyclicity, there exists $h \in F_b^\varepsilon(c)$ such that $d_* h = f_b^\varepsilon(d_* c)$ and thus we can define $f_b^\varepsilon(c) = h$. Altogether, we have defined a chain morphism f_b^ε which is carried by F_b^ε. Since X is tame, the statement holds for all filtration values on $\mathbf{R}$.

Now, assume that g_b^ε is also carried by F_b^ε for all $b \in \mathbb{R}$. Following [82, Sec. 12.3], we define the chain complex $\mathscr{I}$ given by $\mathscr{I}_0 = \langle [0], [1] \rangle$ and $\mathscr{I}_1 = \langle [0,1] \rangle$ and $\mathscr{I}_k = 0$ for $k > 0$. This is the cellular chain complex of the unit interval I decomposed into two 0-cells and one 1-cell. A chain homotopy $h_b^\varepsilon : f_b^\varepsilon \simeq g_b^\varepsilon$ corresponds to a chain map $h_b^\varepsilon \colon C_*^{\mathrm{cell}}(X_b) \otimes \mathscr{I} \to C_*^{\mathrm{cell}}(Y_b)$ such that $h_b^\varepsilon(x, [0]) = f_b^\varepsilon(x)$ and $h_b^\varepsilon(x, [1]) = g_b^\varepsilon(x)$ for all $x \in X_b$. Let $H_b^\varepsilon(c, i) = F_b^\varepsilon(c)$ for a cell $(c, i) \in X \times I$. By assumption $H^\varepsilon \colon X \times I \rightrightarrows Y$ is an ε-acyclic carrier. Note that $C_*^{\mathrm{cell}}(X_b) \otimes \mathscr{I} \cong C_*^{\mathrm{cell}}(X_b \times I)$. Replicating the first part of the proof we can now extend any map $h_b^\varepsilon \colon C_*^{\mathrm{cell}}(X_b) \otimes \mathscr{I}_0 \to C_*^{\mathrm{cell}}(Y_b)$ with the above properties to all cells of $X \times I$. □

Definition 8.1.3. Let $X_*, Y_* \in$ **RCW-cpx**. We will call an ε-acyclic carrier $I_X^\varepsilon : X_* \rightrightarrows X_{*+\varepsilon}$ carrying the standard shift $\Sigma^\varepsilon X_*$ a *shift* carrier. Suppose that there are ε-acyclic carriers

$$F^\varepsilon : X_* \rightrightarrows Y_{*+\varepsilon} \ ,$$
$$G^\varepsilon : Y_* \rightrightarrows X_{*+\varepsilon} \ .$$

together with shift carriers $I_X^{2\varepsilon}$ and $I_Y^{2\varepsilon}$. We say that X_* and Y_* are ε*-acyclic equivalent* whenever we have inclusions $G^\varepsilon \circ F^\varepsilon \subseteq I_X^{2\varepsilon}$ and $F^\varepsilon \circ G^\varepsilon \subseteq I_Y^{2\varepsilon}$.

The motivation for the definition of ε-acyclic equivalences is the following lemma:

Proposition 8.1.4. *Let X_* and Y_* be two tame elements from* **FCW-cpx** *which are* ε-acyclic equivalent. *Then* $\mathrm{PH}(X_*)$ *and* $\mathrm{PH}(Y_*)$ *are ε-interleaved.*

Proof. By Prop. 8.1.2 we know that there exist two chain maps $f_*^\varepsilon \colon C_*(X_*) \to C_*(Y_{*+\varepsilon})$ and $g^\varepsilon \colon C_*(Y_*) \to C_*(X_{*+\varepsilon})$ carried by F^ε and G^ε respectively. By hypothesis the compositions $g^\varepsilon \circ f^\varepsilon$ and $f^\varepsilon \circ g^\varepsilon$ are carried by corresponding shift carriers $I_X^{2\varepsilon}$ and $I_Y^{2\varepsilon}$. Thus, using the second part of Prop. 8.1.2 we obtain chain homotopies $g^\varepsilon \circ f^\varepsilon \simeq \Sigma^{2\varepsilon} C_*(X)$ and $f^\varepsilon \circ g^\varepsilon \simeq \Sigma^{2\varepsilon} C_*(Y)$. Altogether, in homology these compositions are equal to the corresponding shifts, and $\mathrm{PH}_*(X_*)$ and $\mathrm{PH}_*(Y_*)$ are ε-interleaved. □

Example 8.1.5. Consider two finite metric spaces $\mathbb{X}$ and $\mathbb{Y}$. Let $d_H(\mathbb{X},\mathbb{Y})$ be their Hausdorff distance and set $\varepsilon = 2d_H(\mathbb{X},\mathbb{Y})$. Given a subcomplex $K \subseteq \mathrm{VR}(\mathbb{X})$, we denote its vertex set by $\mathbb{X}(K) \subseteq \mathbb{X}$. Likewise for a simplex $\sigma \in \mathrm{VR}(\mathbb{X})$, we write $\mathbb{X}(\sigma) \subseteq \mathbb{X}$ for the vertices spanning σ. Define a carrier $F^\varepsilon : \mathrm{VR}(\mathbb{X}) \rightrightarrows \mathrm{VR}(\mathbb{Y})$ by mapping a simplex $\sigma \in \mathrm{VR}(\mathbb{X})_a$ to

$$F^\varepsilon(\sigma) = |\sup\{K \subseteq \mathrm{VR}(\mathbb{Y})_{a+\varepsilon} \mid d_{\mathrm{H}}(\mathbb{X}(\sigma), \mathbb{Y}(K)) \leq \varepsilon/2\}|$$

This is clearly semicontinuous. If $v_0, \ldots, v_n$ are vertices in $F^\varepsilon(\sigma)$, then by definition $\{v_0, \ldots, v_n\}$ is an n-simplex in $F^\varepsilon(\sigma)$. Therefore we have $F^\varepsilon(\sigma) \simeq \Delta^N$ for some $N \in \mathbb{N}$, which is acyclic. In particular, F^ε is an ε-acyclic carrier. Interchanging the roles of $\mathbb{X}$ and $\mathbb{Y}$ we also obtain an ε-acyclic carrier $G^\varepsilon : \mathrm{VR}(\mathbb{Y}) \rightrightarrows \mathrm{VR}(\mathbb{X})$. Similarly, we define for a simplex $\sigma \in \mathrm{VR}(\mathbb{X})_a$ the shift carrier

$$I_{\mathbb{X}}^{2\varepsilon}(\sigma) = |\sup\{K \subseteq \mathrm{VR}(\mathbb{X})_{a+2\varepsilon} \mid d_{\mathrm{H}}(\mathbb{X}(\sigma), \mathbb{X}(K)) \leq \varepsilon\}|$$

Analogously one defines $I_{\mathbb{Y}}^{2\varepsilon}$. Since $G^\varepsilon \circ F^\varepsilon \subseteq I_{\mathbb{X}}^{2\varepsilon}$ and $F^\varepsilon \circ G^\varepsilon \subseteq I_{\mathbb{Y}}^{2\varepsilon}$, Prop. 8.1.4 implies that $\mathrm{PH}_*(\mathrm{VR}(\mathbb{X}))$ and $\mathrm{PH}_*(\mathrm{VR}(\mathbb{Y}))$ are ε-interleaved. This is similar to the proof using *correspondences*, see [92, Prop. 7.8, Sec. 7.3].

Example 8.1.6. Consider $\mathbb{R}^N$ together with a 1-Lipschitz function $f : \mathbb{R}^N \to \mathbb{R}$ with constant $\varepsilon > 0$. On the other hand consider the lattices $\mathbb{Z}^N$ and $r\mathbb{Z}^N + l$ for a pair of vectors $r, l \in \mathbb{R}^N$ such that the coordinates of r satisfy $0 < r_i \leq 1$ for all $1 \leq i \leq N$. Then we take their corresponding cubical complexes $\mathscr{C}(\mathbb{Z}^N)$ and $\mathscr{C}(r\mathbb{Z}^N + l)$ thought as embedded in $\mathbb{R}^N$. The function f induces a natural filtration for these cubical complexes: a vertex $v \in \mathscr{C}(\mathbb{Z}^N)$ is contained in $\mathscr{C}(\mathbb{Z}^N)_{f(v)}$, while a cell $a \in \mathscr{C}(\mathbb{Z}^N)$ appears at the maximum filtration value on its vertices. There is a ε-acyclic carrier $F^\varepsilon : \mathscr{C}(\mathbb{Z}^N) \rightrightarrows \mathscr{C}(r\mathbb{Z}^N + l)$ sending each cell $a \in \mathscr{C}(\mathbb{Z}^N)$ to the smallest subcomplex $F^\varepsilon(a)$ containing all $b \in \mathscr{C}(r\mathbb{Z}^N + l)$ such that $\overline{b} \cap a \neq \emptyset$. In an analogous way the inverse acyclic carrier can be defined, and the compositions $F^\varepsilon \circ G^\varepsilon$ and $G^\varepsilon \circ F^\varepsilon$ define the shift carriers. Thus, using Proposition 8.1.4, one shows that $\mathrm{PH}_*(\mathscr{C}(\mathbb{Z}^N))$ and $\mathrm{PH}_*(\mathscr{C}(r\mathbb{Z}^N + l))$ are ε-interleaved.

Here an important assumption of Proposition 8.1.2 is that we are dealing with tame filtered CW-complexes. However, what if we considered instead a pair of elements $X_*, Y_* \in \mathbf{RCW\text{-}cpx}$? In this context, we notice that given an ε-acyclic carrier $F^\varepsilon : X_* \to Y_*[\varepsilon]$, it is not necessarily true that the compositions

$$Y(a + \varepsilon \leq b + \varepsilon) F^\varepsilon_a(c) \text{ and } F^\varepsilon_b(X(a \leq b)(c))$$

are still acyclic for all pairs $a \leq b$ from $\mathbf{R}$. Thus, whenever we talk about ε-acyclic carriers $F^\varepsilon : X_* \to Y_*[\varepsilon]$ in this context we will assume that $F^\varepsilon_b(X(a \leq b)(c))$ is acyclic for all pairs $a, b \in \mathbf{R}$ with $a \leq b$ and all cells $c \in X(a)$.

Lemma 8.1.7. *Let $X_*, Y_* \in \mathbf{RCW\text{-}cpx}$ be a pair of elements such that both are ε-acyclic equivalent in the above sense. Then $d_I(\mathrm{PH}_*(X_*), \mathrm{PH}_*(Y_*)) \leq \varepsilon$.*

Proof. For each persistence value $a \in \mathbf{R}$, we use Theorem 2.7.3 twice to obtain a pair of chain morphisms $f_a : C^{\mathrm{cell}}_a(X) \to C^{\mathrm{cell}}_{a+\varepsilon}(Y)$ and $g_{a+\varepsilon} : C^{\mathrm{cell}}_{a+\varepsilon}(Y) \to C^{\mathrm{cell}}_{a+2\varepsilon}(X)$. In a similar way we obtain a pair of chain homotopies $g_{a+\varepsilon} \circ f_a \simeq (\Sigma^{2\varepsilon} C^{\mathrm{cell}}_*(X))_a$ and $f_{a+\varepsilon} \circ g_a \simeq (\Sigma^{2\varepsilon} C^{\mathrm{cell}}_*(Y))_a$ so that we have equalities between the induced homology morphisms $[g_{a+\varepsilon}] \circ [f_a] = [(\Sigma^{2\varepsilon} C^{\mathrm{cell}}_*(X))_a]$ and $[f_{a+\varepsilon}] \circ [g_a] = [(\Sigma^{2\varepsilon} C^{\mathrm{cell}}_*(Y))_a]$ for all $a \in \mathbf{R}$. Now, for a pair of values $a \leq b$ from $\mathbf{R}$, it is not necessarily true that $Y(a + \varepsilon \leq b + \varepsilon) \circ f_a = f_b \circ X(a \leq b)$. However, since $Y(a + \varepsilon \leq b + \varepsilon) \circ f_a$ and $f_b \circ X(a \leq b)$ are both included in $F^\varepsilon_b(X(a \leq b)(c))$ by hypotheses, then by applying Theorem 2.7.3 again there is a chain homotopy equivalence $Y(a + \varepsilon \leq b + \varepsilon) \circ f_a \simeq f_b \circ X(a \leq b)$, which implies

$$[Y(a + \varepsilon \leq b + \varepsilon)] \circ [f_a] = [f_b] \circ [X(a \leq b)] \, ,$$

and we have defined a persistence morphism $[f_*] : \mathrm{PH}_*(X_*) \to \mathrm{PH}_*(Y_*[\varepsilon])$. Similarly we can also put together the g_a for all $a \in \mathbf{R}$ so that we obtain a morphism $[g_*] : \mathrm{PH}_*(Y_*) \to \mathrm{PH}_*(X_*[\varepsilon])$. This

leads to the claimed ε-interleaving. □

Example 8.1.8. Consider a point cloud $\mathbb{X}$ and a Vietoris Rips complex $\mathrm{VR}_*(\mathbb{X})$ on top of it. For this example, we will look at this applying an exponential transformation $X_r = \mathrm{VR}_{\exp(r)}(\mathbb{X})$ for all $r \in \mathbf{R}$. We will look for an element $Y \in \mathbf{RCW\text{-}cpx}$ that is ε-equivalent to $\mathrm{VR}_{\exp(r)}(\mathbb{X})$ while being of much less complexity. We let $\mathscr{P}_r$ be a partition of $\mathbb{X}$ in such a way that for any partitioning set $U \in \mathscr{P}_r$, for any pair of points $p,q \in U$, we have that $d(p,q) \leq \exp(\varepsilon)/2$. For a point $p \in \mathbb{X}$, we will denote by $\mathscr{P}_r(p)$ the partitioning set from $\mathscr{P}_r$ containing p. Then we define a simplicial complex Y_r over the set $\mathscr{P}_r$, and where we include a simplex $[\mathscr{P}_r(p_0), \mathscr{P}_r(p_1), \ldots, \mathscr{P}_r(p_N)]$ whenever there exists a simplex $[p_0, p_1, \ldots, p_N]$ in $\mathrm{VR}_{\exp(r)}(\mathbb{X})$; here we ignore degenerate simplices. It is easy to see that in general Y_* is not a filtered complex but rather a regularly filtered CW complex. We define the ε-acyclic carrier $F_r^\varepsilon : X_r \to Y_{r+\varepsilon}$ by sending a simplex $[p_0, p_1, \ldots, p_N] \in X_r$ to the standard simplex $\Delta^{[\mathscr{P}_r(p_0), \mathscr{P}_r(p_1), \ldots, \mathscr{P}_r(p_N)]}$. On the other way, we start from a simplex $[U_0, U_1, \ldots, U_M] \in Y_r$ and send it to the subcomplex of $X_{r+\varepsilon}$ given by the join $\Delta(U_0) * \Delta(U_1) * \cdots * \Delta(U_M)$, where by $\Delta(U_i)$ we mean the standard simplex with vertices on the elements from $U_i \in \mathscr{P}_r$. These assignements lead to a ε-acyclic equivalence between X_r and Y_r, where the shift carriers are given by composition. Now, by Lemma 8.1.7 there is a ε-interleaving between $\mathrm{PH}_*(X)$ and $\mathrm{PH}_*(Y)$.

Remark. Notice that our notion of acyclicity is different from that in [28] and [65]. In [65] a filtered complex K_* is called ε-acyclic whenever the induced homology maps $H_*(K_r) \to H_*(K_{r+\varepsilon})$ vanish for all $r \in \mathbb{R}$. In this case one can still (trivially) define acyclic carriers between $*$ and K_*. The problem arises when defining the shift carrier $I_K^{A\varepsilon}$ for some constant $A > 0$, which does not exist in general. One can however, adapt the proof of Proposition 8.1.2 so that there is a chain morphism $\psi^{\varepsilon(\dim(K_r)+1)} : C_*(K_r) \to C_*(K_{r+\varepsilon(\dim(K_r)+1)})$; and that this coincides up to chain homotopy with the composition through $C_*(*)$. One does this by following the same proof as in Proposition 8.1.2, but increasing the filtration value by ε each time we assume that some cycle lies in an acyclic carrier. Thus, if we have $\dim(K) = \sup_{r\in\mathbb{R}}(\dim(K_r)) < \infty$, then one could say that there is an $\varepsilon(\dim(K)+1)$-*approximate* chain homotopy equivalence between $C(*)$ and $C(K_*)$.

8.2 Interleaving Geometric Realizations

Next, we focus on acyclic carrier equivalences between a pair of diagrams $\mathscr{D}, \mathscr{L} \in \mathbf{RRDiag}(K)$. We start by taking ε-acyclic carriers $F_\sigma^\varepsilon : \mathscr{D}(\sigma) \rightrightarrows \mathscr{L}(\sigma)$ for all $\sigma \in K$ which have to be compatible in the following sense: For any pair $\tau \preceq \sigma$ and any cell $c \in \mathscr{D}(\sigma)$, there is an inclusion

$$\mathscr{L}(\tau \preceq \sigma)(F_\sigma^\varepsilon(c)) \subseteq F_\tau^\varepsilon(\mathscr{D}(\tau \preceq \sigma)(c)) \tag{8.1}$$

and we assume in addition that $F^{\varepsilon}_{\tau}(\mathscr{D}(\tau \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c))$ is acyclic for all $r \geq 0$. This compatibility leads to 'local' diagrams of spaces. That is, given a pair of values $a \in \mathbf{R}$ and $r \geq 0$ and a cell $c \in \mathscr{D}(\sigma)_a$, we consider an object $F^{r,\varepsilon}_{\sigma\times c} \in \mathbf{RDiag}(\Delta^{\sigma})$. It is given by the space $F^{r,\varepsilon}_{\sigma\times c}(\tau) = F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c)\big)$ for all faces $\tau \preceq \sigma$. For any sequence $\rho \preceq \tau \preceq \sigma$ in K, there are morphisms in $F^{r,\varepsilon}_{\sigma\times c}$ given by restricting morphisms from $\mathscr{L}$

$$\begin{array}{ccccc} \tau & \longrightarrow & F^{r,\varepsilon}_{\sigma\times c}(\tau) & = & F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c)\big) \\ \preceq\uparrow & & \downarrow & & \downarrow{\scriptstyle \mathscr{L}(\rho\preceq\tau)} \\ \rho & \longrightarrow & F^{r,\varepsilon}_{\sigma\times c}(\rho) & = & F^{\varepsilon}_{\rho}\big(\mathscr{D}(\rho \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c)\big)\ . \end{array}$$

Using condition (8.1) repeatedly on the cells from $L = \mathscr{D}(\tau \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c)$, we see that we have an inclusion

$$\mathscr{L}(\rho \preceq \tau)\big(F^{\varepsilon}_{\tau}(L)\big) \subseteq F^{\varepsilon}_{\rho}\big(\mathscr{D}(\rho \preceq \tau)(L)\big)\ .$$

Thus the diagram $F^{r,\varepsilon}_{\sigma\times c}$ is indeed well defined, and we may consider the geometric realization $\Delta F^{r,\varepsilon}_{\sigma\times c}$. By hypothesis each $F^{r,\varepsilon}_{\sigma\times c}(\tau)$ is acyclic for all $\tau \preceq \sigma$, so that the first page of the spectral sequence $E^{*}_{p,q}(F^{r,\varepsilon}_{\sigma\times c}) \Rightarrow \mathrm{H}_{p+q}(\Delta F^{r,\varepsilon}_{\sigma\times c})$ is equal to

$$E^{1}_{p,q}(F^{r,\varepsilon}_{\sigma\times c}) = \bigoplus_{\tau^p \in \Delta^{\sigma}} \mathrm{H}_q(F^{r,\varepsilon}_{\sigma\times c}(\tau^p)) = \begin{cases} \bigoplus_{\tau^p \in \Delta^{\sigma}} \mathbb{F} \text{ if } q = 0, \\ 0 \text{ otherwise.} \end{cases}$$

where $\mathbb{F}$ denotes our coefficient field for homology. In fact, computing the homology with respect to the horizontal differentials on the first page corresponds to computing the homology of Δ^{σ}. Thus, $E^{2}_{p,q}(F^{r,\varepsilon}_{\sigma\times c})$ is zero everywhere except at $p = q = 0$ where it is equal to $\mathbb{F}$. Thus, the spectral sequence collapses on the second page, and $\Delta F^{r,\varepsilon}_{\sigma\times c}$ is acyclic. We will use the notation $F^{\varepsilon}_{\sigma\times c} = F^{0,\varepsilon}_{\sigma\times c}$.

Definition 8.2.1. Let $\mathscr{D}$ and $\mathscr{L}$ be two diagrams in $\mathbf{RRDiag}(K)$. Suppose that there are ε-acyclic carriers $F^{\varepsilon}_{\sigma} : \mathscr{D}(\sigma) \rightrightarrows \mathscr{L}(\sigma)$ for all $\sigma \in K$, and such that

$$\mathscr{L}(\tau \preceq \sigma)\big(F^{\varepsilon}_{\sigma}(c)\big) \subseteq F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)(c)\big)$$

for all $c \in \mathscr{D}(\sigma)$ and in addition $F^{\varepsilon}_{\tau}(\mathscr{D}(\tau \preceq \sigma)\Sigma^r \mathscr{D}(\sigma)(c))$ is acyclic for all $r \geq 0$. Then we say that the set $\{F^{\varepsilon}_{\sigma}\}_{\sigma\in K}$ is a (ε, K)-*acyclic carrier* between $\mathscr{D}$ and $\mathscr{L}$. We denote this by $F^{\varepsilon} : \mathscr{D} \rightrightarrows \mathscr{L}$.

Theorem 8.2.2. *Let $\mathscr{D}$ and $\mathscr{L}$ be two diagrams in* $\mathbf{RRDiag}(K)$. *Suppose that there are (ε, K)-acyclic carriers $F^{\varepsilon} : \mathscr{D} \rightrightarrows \mathscr{L}$ and $G^{\varepsilon} : \mathscr{L} \rightrightarrows \mathscr{D}$, together with a pair of shift (ε, K)-acyclic carriers*

$I^{2\varepsilon}_{\mathscr{D}} : \mathscr{D} \rightrightarrows \mathscr{D}$ and $I^{2\varepsilon}_{\mathscr{L}} : \mathscr{L} \rightrightarrows \mathscr{L}$, and such that these restrict to acyclic equivalences

$$G^{\varepsilon}_{\tau} \circ F^{\varepsilon}_{\tau} \subseteq (I^{2\varepsilon}_{\mathscr{D}})_{\tau} \text{ and } F^{\varepsilon}_{\tau} \circ G^{\varepsilon}_{\tau} \subseteq (I^{2\varepsilon}_{\mathscr{L}})_{\tau}$$

for each simplex $\tau \in K$. Then there is an ε-acyclic equivalence $F^{\varepsilon} : \Delta\mathscr{D} \rightrightarrows \Delta\mathscr{L}$ which preserves filtrations. That is, there are ε-acyclic equivalences $F^{p}F^{\varepsilon} : F^{p}\Delta\mathscr{D} \rightrightarrows F^{p}\Delta\mathscr{L}$ for all $p \geq 0$.

Proof. Let $\sigma \times c \in \Delta\mathscr{D}$ be a cell, where c is an m-cell in $\mathscr{D}(\sigma)$. Define an acyclic carrier $F^{\varepsilon} : \Delta\mathscr{D} \rightrightarrows \Delta\mathscr{L}$ by sending $\sigma \times c$ to the acyclic carrier $\Delta F^{\varepsilon}_{\sigma\times c}$, which is a subcomplex of $\Delta\mathscr{L}$. Let us first check semicontinuity. For any pair of cells $\tau \times a \preceq \sigma \times c$ in $\Delta\mathscr{D}$, the cell a is contained in the subcomplex $\overline{\mathscr{D}(\tau \preceq \sigma)(c)}$, and by continuity of $\mathscr{D}(\rho \preceq \tau)$ we have that $\mathscr{D}(\rho \preceq \tau)(a) \subseteq \overline{\mathscr{D}(\rho \preceq \sigma)(c)}$. Thus there are inclusions

$$F^{\varepsilon}_{\rho}(\mathscr{D}(\rho \preceq \tau)(a)) \subseteq F^{\varepsilon}_{\rho}(\overline{\mathscr{D}(\rho \preceq \sigma)(c)}) = F^{\varepsilon}_{\rho}(\mathscr{D}(\rho \preceq \sigma)(c))$$

for all $\rho \preceq \tau$. More concisely, $F^{\varepsilon}_{\tau\times a}(\rho) \subseteq F^{\varepsilon}_{\sigma\times c}(\rho)$ for all $\rho \preceq \tau$. As a consequence $\Delta F^{\varepsilon}_{\tau\times a} \subseteq \Delta F^{\varepsilon}_{\sigma\times c}$ and semicontinuity holds.

Next, notice that $F^{\varepsilon}\big(\Sigma^{r}\Delta\mathscr{D}(\sigma \times c)\big) = F^{\varepsilon}\big(\sigma \times \Sigma^{r}\mathscr{D}(\sigma)(c)\big) = \Delta F^{r,\varepsilon}_{\sigma\times c}$ which is an acyclic carrier. In order for F^{ε} to be an ε-acyclic carrier, it remains to show the inclusion $\Sigma^{r}\Delta\mathscr{L} \circ F^{\varepsilon} \subseteq F^{\varepsilon} \circ \Sigma^{r}\Delta\mathscr{D}$ for all $r \geq 0$. For this, take $\sigma \times c \in \Delta\mathscr{D}$ and see that

$$\begin{aligned}
\Sigma^{r}\Delta\mathscr{L} \circ F^{\varepsilon}(\sigma \times c) &= \Sigma^{r}\Delta\mathscr{L}\Big(\bigcup_{\tau\preceq\sigma} \tau \times F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)(c)\big)\Big) \\
&= \bigcup_{\tau\preceq\sigma} \tau \times \Sigma^{r}\mathscr{L}(\tau)\big(F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)(c)\big)\big) \subseteq \bigcup_{\tau\preceq\sigma} \tau \times F^{\varepsilon}_{\tau}\big(\Sigma^{r}\mathscr{D}(\tau)\mathscr{D}(\tau \preceq \sigma)(c)\big) \\
&= \bigcup_{\tau\preceq\sigma} \tau \times F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)\Sigma^{r}\mathscr{D}(\sigma)(c)\big) = F^{\varepsilon}\big(\sigma \times \Sigma^{r}\mathscr{D}(\sigma)(c)\big) = F^{\varepsilon} \circ \Sigma^{r}\Delta\mathscr{D}(\sigma \times c)\ .
\end{aligned}$$

Similarly, one can define an ε-acyclic carrier $G^{\varepsilon} : \Delta\mathscr{L} \rightrightarrows \Delta\mathscr{D}$ sending $\sigma \times c \in \Delta\mathscr{L}$ to $\Delta G^{\varepsilon}_{\sigma\times c}$. In addition, we define respective shift ε-acyclic carriers $I^{2\varepsilon}_{\mathscr{D}} : \Delta\mathscr{D} \rightrightarrows \Delta\mathscr{D}$ and $I^{2\varepsilon}_{\mathscr{L}} : \Delta\mathscr{L} \rightrightarrows \Delta\mathscr{L}$, sending respectively $\sigma \times c \in \Delta\mathscr{D}$ to $\Delta(I^{2\varepsilon}_{\mathscr{D}})_{\sigma\times c}$ and $\tau \times a \in \Delta\mathscr{L}$ to $\Delta(I^{2\varepsilon}_{\mathscr{L}})_{\tau\times a}$. Then we have

$$\begin{aligned}
G^{\varepsilon} \circ F^{\varepsilon}(\sigma \times c) = G^{\varepsilon}(\Delta F^{\varepsilon}_{\sigma\times c}) &= G^{\varepsilon}\Big(\bigcup_{\tau\preceq\sigma} \tau \times F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)(c)\big)\Big) \\
&= \bigcup_{\rho\preceq\tau\preceq\sigma} \rho \times G^{\varepsilon}_{\rho}\Big(\mathscr{L}(\rho \preceq \tau)F^{\varepsilon}_{\tau}\big(\mathscr{D}(\tau \preceq \sigma)(c)\big)\Big) \\
&\subseteq \bigcup_{\rho\preceq\sigma} \rho \times G^{\varepsilon}_{\rho}F^{\varepsilon}_{\rho}\big(\mathscr{D}(\rho \preceq \sigma)(c)\big) \subseteq \Delta(I^{2\varepsilon}_{\mathscr{D}})_{\sigma\times c} = I^{2\varepsilon}_{\mathscr{D}}(\sigma \times c),
\end{aligned}$$

where we have used the commutativity condition and equivalence of F^{ε}_{ρ} and G^{ε}_{ρ}. Consequently

$G^\varepsilon \circ F^\varepsilon \subseteq I^{2\varepsilon}_{\mathscr{D}}$; the other inclusion $F^\varepsilon \circ G^\varepsilon \subseteq I^{2\varepsilon}_{\mathscr{L}}$ follows by symmetry. Altogether, we have obtained an ε-equivalence $F^\varepsilon : \Delta\mathscr{D} \rightrightarrows \Delta\mathscr{L}$. Finally, notice that for all $p \geq 0$ and for each cell $\sigma \times c \in F^p\Delta\mathscr{D}$, its carrier $\Delta F^\varepsilon_{\sigma\times c}$ is contained in $F^p\Delta\mathscr{D}$ and so it preserves filtration. The same follows for the other acyclic carriers. □

Let X be a filtered simplicial complex, together with a cover $\mathscr{U}$ by filtered subcomplexes. Recall the definitions of the diagrams $X^{\mathscr{U}}$ and $\pi_0^{\mathscr{U}}$ over $N_{\mathscr{U}}$ from example 7.2.2. Consider the case when $d_I\big(\mathrm{PH}_*(X^{\mathscr{U}}(\sigma)), \mathrm{PH}_*(\pi_0^{\mathscr{U}}(\sigma))\big) \leq \varepsilon$ for all $\sigma \in N_{\mathscr{U}}$. This example has been of interest before, see for example [65] or [28]. As mentioned in the remark at the end of Section 8.1, our notion of ε-acyclicity is much stronger than that from [65]. This is why we obtain a result closer to the *Persistence Nerve Theorem* from [31] than to the *Approximate Nerve Theorem* from [65].

Corollary 8.2.3 (Strong Approximate Multinerve Theorem)**.** *Consider a diagram $\mathscr{D}$ in* $\mathbf{FRDiag}(K)$. *Assume that there are is a (ε, K)-acyclic carrier $F^\varepsilon : \pi_0\mathscr{D} \rightrightarrows \mathscr{D}$. Then, there is a ε-acyclic equivalence $F^\varepsilon : \mathrm{MNerv}(\mathscr{D}) \rightrightarrows \Delta\mathscr{D}$. Consequently,*

$$d_I(\mathrm{PH}_*(\mathrm{MNerv}(\mathscr{D})), \mathrm{PH}_*(\Delta\mathscr{D})) \leq \varepsilon\ .$$

Proof. Notice that $\pi_0\mathscr{D}$ is a well defined element from $\mathbf{RRDiag}(K)$ and there is an obvious choice for the (ε, K)-acyclic carrier $G^\varepsilon : \mathscr{D} \rightrightarrows \pi_0\mathscr{D}$ where we send cells to their corresponding connected component classes. The compatibility condition

$$\pi_0(\mathscr{D}(\tau \preceq \sigma))(G^\varepsilon(\mathscr{D}(\sigma))) \subseteq G^\varepsilon(\mathscr{D}(\tau))$$

also follows. The shift $(2\varepsilon, K)$-carrier $I^{2\varepsilon}_{\pi_0\mathscr{D}}$ sends points to points, while the other $I^{2\varepsilon}_{\mathscr{D}}$ is defined as the composition $F^\varepsilon \circ G^\varepsilon$, which can be checked to define a $(2\varepsilon, K)$-acyclic carrier. Altogether, we can use Proposition 8.2.2 and there exists a ε-acyclic equivalence $F^\varepsilon : \mathrm{MNerv}(\mathscr{D}) \rightrightarrows \Delta(\mathscr{D})$. □

Example 8.2.4. Consider a point cloud $\mathbb{X}$ together with a partition $\mathscr{P}$. Assume that the $(\mathrm{VR}_*(\mathbb{X}), \mathscr{P})$-join diagram $\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}$ is such that there are compatible ε-acyclic equivalences $\pi_0(\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}(\sigma)) \rightrightarrows \mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}(\sigma)$ for all $\sigma \in \Delta^P$. Then there is an ε-acyclic equivalence $\Delta\pi_0\big(\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}\big) \rightrightarrows \Delta\,\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}$ so that

$$d_I\big(\mathrm{PH}_*(\mathrm{MNerv}(\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})})), \mathrm{PH}_*(\mathrm{VR}_*(\mathbb{X}))\big) \leq \varepsilon\ .$$

Acyclic carriers have been used in [70] and in [86] for approximating continuous morphisms by means of simplicial maps. Here we have used the same tools to obtain an approximate "homotopy colimit theorem", i.e. an approximate version of Theorem 2.8.6. The acyclic carrier theorem is an instance of the more general acyclic Model theorem. An interesting future research direction

would be to see how that general result can bring new insights into applied topology.

8.3 Interleaving Spectral Sequences

Definition 8.3.1. Let $\mathscr{A}$ and $\mathscr{B}$ from **SpSq**. A *n-spectral sequence morphism* $f : \mathscr{A} \to \mathscr{B}$ is a spectral sequence morphism $f : \mathscr{A} \to \mathscr{B}$ which is defined from page n.

Definition 8.3.2. Given two objects $\mathscr{A}$ and $\mathscr{B}$ in $\mathbf{SpSq}^{[0,\infty)}$. We say that $\mathscr{A}$ and $\mathscr{B}$ are (ε, n)-*interleaved* whenever there exist two n-morphisms $\psi : \mathscr{A} \to \mathscr{B}[\varepsilon]$ and $\varphi : \mathscr{B} \to \mathscr{A}[\varepsilon]$ such that the following diagram commutes

$$\begin{array}{ccccc} \mathscr{A} & & & & \mathscr{B} \\ {\scriptstyle \Sigma^{\varepsilon}\mathscr{A}}\downarrow & \searrow^{\psi} & & \swarrow_{\varphi} & \downarrow{\scriptstyle \Sigma^{\varepsilon}\mathscr{B}} \\ \mathscr{A}[\varepsilon] & & & & \mathscr{B}[\varepsilon] \\ {\scriptstyle \Sigma^{\varepsilon}\mathscr{A}[\varepsilon]}\downarrow & \searrow^{\psi[\varepsilon]} & & \swarrow_{\varphi[\varepsilon]} & \downarrow{\scriptstyle \Sigma^{\varepsilon}\mathscr{B}[\varepsilon]} \\ \mathscr{A}[2\varepsilon] & & & & \mathscr{B}[2\varepsilon] \end{array} \tag{8.2}$$

for all pages $r \geq n$. This interleaving defines a pseudometric in $\mathbf{SpSq}^{[0,\infty)}$

$$d_I^n\left(\mathscr{A}, \mathscr{B}\right) := \inf\{\varepsilon \mid \mathscr{A} \text{ and } \mathscr{B} \text{ are } (\varepsilon, n)\text{-interleaved}\} \ .$$

Proposition 8.3.3. *Suppose that $\mathscr{A}$ and $\mathscr{B}$ are (ε, n)-interleaved. Then these are (ε, m)-interleaved for all $m \geq n$. In particular, we have that*

$$d_I^m\left(\mathscr{A}, \mathscr{B}\right) \leq d_I^n\left(\mathscr{A}, \mathscr{B}\right)$$

for any pair of integers $m \geq n$.

Proof. Follows directly from the definitions. □

We start now by considering Mayer-Vietoris spectral sequences. Under some conditions which are a special case of Theorem 8.2.2, one can obtain one page stability. In fact this stability is due to morphisms directly defined on the underlying double complexes, which is a very strong property.

Proposition 8.3.4. *Let $(X, \mathscr{U})$ and $(Y, \mathscr{V})$ be two tame elements in* **FCW-cpx** *together with covers by subcomplexes, both having the same finite nerve $K = N_{\mathscr{U}} = N_{\mathscr{V}}$. Suppose that there are (ε, K)-acyclic carriers $F^{\varepsilon} : X^{\mathscr{U}} \rightrightarrows Y^{\mathscr{V}}$ and $G^{\varepsilon} : Y^{\mathscr{V}} \rightrightarrows X^{\mathscr{U}}$, together with a pair of shift (ε, K)-acyclic carriers $I^{2\varepsilon}_{X^{\mathscr{U}}} : X^{\mathscr{U}} \rightrightarrows X^{\mathscr{U}}$ and $I^{2\varepsilon}_{Y^{\mathscr{V}}} : Y^{\mathscr{V}} \rightrightarrows Y^{\mathscr{V}}$, and such that these restrict to acyclic equivalences*

$$G^{\varepsilon}_{\tau} \circ F^{\varepsilon}_{\tau} \subseteq (I^{2\varepsilon}_{X^{\mathscr{U}}})_{\tau} \text{ and } F^{\varepsilon}_{\tau} \circ G^{\varepsilon}_{\tau} \subseteq (I^{2\varepsilon}_{Y^{\mathscr{V}}})_{\tau}$$

for each simplex $\tau \in K$. Then there are a pair of double complex morphisms $\phi^\varepsilon : C_{*,*}(X, \mathscr{U}) \to C_{*,*}(Y, \mathscr{V})[\varepsilon]$ and $\psi^\varepsilon : C_{*,*}(Y, \mathscr{V}) \to C_{*,*}(X, \mathscr{U})[\varepsilon]$ inducing a first page interleaving between $E^*_{*,*}(X, \mathscr{U})$ and $E^*_{*,*}(Y, \mathscr{V})$).

Proof. Unpacking the definitions this means we have to give chain homomorphisms

$$(\phi^\varepsilon_\sigma)_r \colon C_*(X^{\mathscr{U}}(\sigma)_r) \to C_*(Y^{\mathscr{V}}(\sigma)_{r+\varepsilon}) \,,$$
$$(\psi^\varepsilon_\sigma)_r \colon C_*(Y^{\mathscr{V}}(\sigma)_r) \to C_*(X^{\mathscr{U}}(\sigma)_{r+\varepsilon})$$

that are natural in $\sigma \in K$ and in $r \in \mathbf{R}$. Since K is a poset category, these can be constructed inductively as follows: As in Prop. 8.1.2 we may define ϕ^ε_σ on all simplices $\sigma \in K$ of dimension $\dim(\sigma) = \dim(K)$. Note that $(\phi^\varepsilon_\sigma)_r$ is carried by $(F^\varepsilon_\sigma)_r$ for all $r \in \mathbf{R}$. Assume by induction that ϕ^ε_τ are defined and carried by F^ε_τ for all $\tau \in K$ with $n \leq \dim(\tau) \leq \dim(K)$ in such a way that for all cofaces $\tau \preceq \sigma$ the naturality condition $\phi^\varepsilon_\tau \circ X^{\mathscr{U}}(\tau \prec \sigma) = Y^{\mathscr{V}}(\tau \prec \sigma)[\varepsilon] \circ \phi^\varepsilon_\sigma$ holds. Now let $\tau \in K$ have dimension $\dim(\tau) = n-1 \geq 0$. The naturality condition on the simplices fixes ϕ^ε_τ on the filtered subcomplex $X^\tau = \bigcup_{\tau \prec \sigma} \mathrm{Im}\big(X^{\mathscr{U}}(\tau \prec \sigma)\big)$, where the union is taken over all cofaces σ of τ. Here notice that we can assume that ϕ^ε_τ is well defined since the previous choices of ϕ^ε_σ for all cofaces $\tau \prec \sigma$ are consistent due to the fact that for each cell $c \in X_\tau$ there exists a unique maximal simplex $\sigma \in N_{\mathscr{U}}$ such that $c \in X_\sigma$. In addition, notice that by hypotheses $Y^{\mathscr{V}}(\tau \prec \sigma)((F^\varepsilon_\sigma)(c)) \subseteq F^\varepsilon_\tau(X^{\mathscr{U}}(\tau \prec \sigma)(c))$ for all $a \in \mathbf{R}$ and $c \in X^{\mathscr{U}}(\sigma)$, so that our definition of ϕ^ε_τ in X^τ is indeed carried by F^ε_τ. We then proceed as in Prop. 8.1.2 to define $(\phi^\varepsilon_\tau)_a$ on all simplices in the subset $X^{\mathscr{U}}(\tau)_a \setminus X^\tau_a$ for all $a \in \mathbf{R}$. The resulting chain map $(\phi^\varepsilon_\tau)_a$ is carried by $(F^\varepsilon_\tau)_a$ for all $a \in \mathbf{R}$. Since $X^{\mathscr{U}}$ is tame, we only need finitely many steps to obtain a morphism $\phi^\varepsilon_\tau : C_*(X^{\mathscr{U}}(\tau)) \to C_*(Y^{\mathscr{V}}(\tau)[\varepsilon])$ that satisfies the induction hypotheses.

Thus, we obtain double complex morphisms $\phi^\varepsilon_{p,q} : C_{p,q}(X, \mathscr{U}) \to C_{p,q}(Y, \mathscr{V})[\varepsilon]$ for all $p, q \geq 0$ by adding up our defined local morphisms

$$\phi^\varepsilon_{p,q} : \bigoplus_{\sigma^p \in K} \phi^\varepsilon_{\sigma^p} : \bigoplus_{\sigma^p \in K} C_q(X^{\mathscr{U}}(\sigma^p)) \longrightarrow \bigoplus_{\sigma^p \in K} C_q(Y^{\mathscr{V}}(\sigma^p))[\varepsilon] \,.$$

Notice that $\phi^\varepsilon_{p,q}$ commute both with horizontal and vertical differentials since we assumed that each ϕ^ε_σ is a chain morphism and these satisfy a naturality condition with respect to K. Thus, this double complex morphism induces a spectral sequence morphism $\phi^\varepsilon_{p,q} : E^*_{p,q}(X^{\mathscr{U}}) \to E^*_{p,q}(Y^{\mathscr{V}})[\varepsilon]$. By doing the same construction, we can obtain $\psi^\varepsilon_\sigma : C_*(Y^{\mathscr{V}}(\sigma)) \to C_*(X^{\mathscr{U}}(\sigma))[\varepsilon]$ so that by Proposition 8.1.2 we have $[\psi^\varepsilon_\sigma] \circ [\phi^\varepsilon_\sigma] = [\Sigma^{2\varepsilon} C_*(X^{\mathscr{U}}(\sigma))]$ and also $[\phi^\varepsilon_\sigma] \circ [\psi^\varepsilon_\sigma] = [\Sigma^{2\varepsilon} C_*(Y^{\mathscr{V}}(\sigma))]$ for all simplices $\sigma \in K$. Then we get a double complex morphism $\psi^\varepsilon_{p,q} : C_{p,q}(Y, \mathscr{V}) \to C_{p,q}(X, \mathscr{U})[\varepsilon]$ inducing an "inverse" spectral sequence morphism $\psi^\varepsilon_{p,q} : E^*_{p,q}(Y, \mathscr{V}) \to E^*_{p,q}(X, \mathscr{U})[\varepsilon]$. These are

such that from the first page, $\phi^{\varepsilon}_{*,*}$ and $\psi^{\varepsilon}_{*,*}$ form a $(\varepsilon,1)$-interleaving of spectral sequences. □

Notice that the proof of Proposition 8.3.4 relies heavily on the fact that the diagrams we are considering come from a cover. This allows us to define a pair of double complex morphisms that are compatible along the common indexing nerve. However, in Theorem 8.2.2 we observed that, under some conditions, the geometric realizations of regularly filtered regular diagrams are stable. Does this stability carry over to the associated spectral sequences? The next theorem shows that this is indeed the case.

Theorem 8.3.5. *Let $\mathscr{D}$ and $\mathscr{L}$ be two diagrams in* $\mathbf{RRDiag}(K)$. *Suppose that there are (ε,K)-acyclic carriers $F^{\varepsilon}:\mathscr{D}\rightrightarrows\mathscr{L}$ and $G^{\varepsilon}:\mathscr{L}\rightrightarrows\mathscr{D}$, together with a pair of shift (ε,K)-acyclic carriers $I^{2\varepsilon}_{\mathscr{D}}:\mathscr{D}\rightrightarrows\mathscr{D}$ and $I^{2\varepsilon}_{\mathscr{L}}:\mathscr{L}\rightrightarrows\mathscr{L}$, and such that these restrict to acyclic equivalences*

$$G^{\varepsilon}_{\tau}\circ F^{\varepsilon}_{\tau}\subseteq (I^{2\varepsilon}_{\mathscr{D}})_{\tau} \text{ and } F^{\varepsilon}_{\tau}\circ G^{\varepsilon}_{\tau}\subseteq (I^{2\varepsilon}_{\mathscr{L}})_{\tau}$$

for each simplex $\tau\in K$. Then

$$d_I^1(E(\mathscr{D},K),E(\mathscr{L},K))\leq\varepsilon\ .$$

Proof. Recall from Theorem 8.2.2 that there is a filtration-preserving acyclic carrier $F^{\varepsilon}:\Delta_K\mathscr{D}\rightrightarrows\Delta_K\mathscr{L}[\varepsilon]$. This implies that there are chain complex morphisms $f^{\varepsilon}_a:C_*(\Delta\mathscr{D})_r\to C_*(\Delta\mathscr{L})_{a+\varepsilon}$ which respect filtrations in the sense that $f^{\varepsilon}_r(F^pC_*(\Delta\mathscr{D})_r)\subseteq F^pC_*(\Delta\mathscr{L})_{r+\varepsilon}$ for all $p\geq 0$. By Lemma 7.3.1 this defines a morphism $f^{\varepsilon}_r:S^{\mathrm{Tot}}_*(\mathscr{D})_r\to S^{\mathrm{Tot}}_*(\mathscr{L})_{r+\varepsilon}$ which respects filtrations. Altogether we deduce that f^{ε}_r determines a morphism of spectral sequences $f^{\varepsilon}_r:E^*_{p,q}(\mathscr{D})_r\to E^*_{p,q}(\mathscr{L})_{r+\varepsilon}$. Similarly as in Lemma 8.1.7 the commutativity

$$\Sigma^{\varepsilon}E^*_{p,q}(\mathscr{L})_{r+\varepsilon}\circ f^{\varepsilon}_r=f^{\varepsilon}_{r+\varepsilon}\circ\Sigma^{\varepsilon}E^*_{p,q}(\mathscr{D})_r \tag{8.3}$$

does not need to hold for all $r\in\mathbf{R}$. However, recall from Theorem 8.2.2 that there is an inclusion $\Sigma^r\Delta\mathscr{L}\circ F^{\varepsilon}\subseteq F^{\varepsilon}\circ\Sigma^r\Delta\mathscr{D}$ where the superset is acyclic, so that $\Sigma^rC_*(\Delta\mathscr{D})_{a+\varepsilon}\circ f^{\varepsilon}_a$ and $f^{\varepsilon}_{a+r}\circ\Sigma^rC_*(\Delta\mathscr{D})_a$ are both carried by the filtration preserving acyclic carrier $F^{\varepsilon}\circ\Sigma^r\Delta\mathscr{D}$. This implies that there exist chain homotopies $h^{\varepsilon}_r:C_n(\Delta\mathscr{D})_r\to C_{n+1}(\Delta\mathscr{D})_{r+\varepsilon}$ which respect filtrations and such that

$$f^{\varepsilon}_{a+r}\circ\Sigma^rC_*(\Delta\mathscr{D})_a-\Sigma^rC_*(\Delta\mathscr{D})_{a+\varepsilon}\circ f^{\varepsilon}_a=\delta^{\Delta}\circ h^{\varepsilon}_a+h^{\varepsilon}_a\circ\delta^{\Delta}\ .$$

for all $a\in\mathbf{R}$ and all $r\geq 0$. Recall that the zero page terms are given as quotients on successive filtration terms $E^0_{p,q}(\mathscr{D})_a=F^pS^{\mathrm{Tot}}_{p+q}(\mathscr{D})_a/F^{p-1}S^{\mathrm{Tot}}_{p+q}(\mathscr{D})_a$, for all $a\in\mathbf{R}$ and all integers $p,q\geq 0$. Thus, by Lemma 7.3.1 these chain homotopies carry over to $S^{\mathrm{Tot}}_*(\mathscr{D})_a$ and the commutativity

relation from equation (8.3) holds from the first page onwards.

Similarly, we can define spectral sequence morphisms $g_r^\varepsilon : E^*_{p,q}(\mathscr{D})_a \to E^*_{p,q}(\mathscr{D})_{a+\varepsilon}$ for all $a \in \mathbf{R}$ which commute with the shift morphisms from the first page. Also, by inspecting the shift carriers, we can obtain equalities of 1-spectral sequence morphisms $g^\varepsilon_{r+\varepsilon} \circ f^\varepsilon_r = \Sigma^{2\varepsilon} E^*_{p,q}(\mathscr{D})_r$ and also $f^\varepsilon_{r+\varepsilon} \circ g^\varepsilon_r = \Sigma^{2\varepsilon} E^*_{p,q}(\mathscr{L})_r$ for all $r \in \mathbf{R}$, and the result follows. □

Example 8.3.6. Consider a pair of point clouds $\mathbb{X}, \mathbb{Y} \in \mathbb{R}^N$, together with partitions $\mathscr{P}$ and $\mathscr{Q}$ for $\mathbb{X}$ and $\mathbb{Y}$ respectively. Also, assume that there is an isomorphism $\phi : \Delta^{\mathscr{P}} \to \Delta^{\mathscr{Q}}$ such that $d_H(\mathbb{X} \cap V, \mathbb{Y} \cap \phi(V)) < \varepsilon/2$ for all $V \in \mathscr{P}$. As defined in example 8.1.5 there are ε-acyclic carrier equivalences $F^\varepsilon_V : \mathrm{VR}(\mathbb{X} \cap V) \rightrightarrows \mathrm{VR}(\mathbb{Y} \cap V)$ for all $V \in \mathscr{U}$. For any $\sigma \in \Delta^{\mathscr{P}}$, one can define ε-acyclic equivalences $F^\varepsilon_\sigma : \mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}(\sigma) \rightrightarrows \mathscr{J}_{\mathscr{Q}}^{\mathrm{VR}_*(\mathbb{Y})}(\sigma)$ by sending a cell $\prod_{V \in \sigma} \tau_V \in \mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}(\sigma)$ to $\prod_{V \in \sigma} F^\varepsilon_V(\tau_V) \in \mathscr{J}_{\mathscr{Q}}^{\mathrm{VR}_*(\mathbb{Y})}(\sigma)$. These acyclic equivalences are compatible so that applying theorem 8.3.5 we obtain the result:

$$d_I^1(E(\mathscr{J}_{\mathscr{P}}^{\mathrm{VR}_*(\mathbb{X})}, \Delta^{\mathscr{P}}), E(\mathscr{J}_{\mathscr{Q}}^{\mathrm{VR}_*(\mathbb{Y})}, \Delta^{\mathscr{Q}})) \leq \varepsilon \, .$$

8.4 Refinement Induced Interleavings

In the previous sections we considered general diagrams in $\mathbf{RRDiag}(K)$ for some simplicial complex K. We will now focus on the situation where we have a filtered simplicial complex X together with a cover $\mathscr{U}$ by filtered subcomplexes, which provides a diagram $X^{\mathscr{U}} : N_{\mathscr{U}} \to \mathbf{FCW\text{-}cpx}$. The associated spectral sequence will be denoted by $E^*_{*,*}(X, \mathscr{U})$, as done at the start of section 7.3. We want to measure how $E^*_{*,*}(X, \mathscr{U})$ changes depending on $\mathscr{U}$ and follow ideas from [105] to achieve this. First we consider a refinement $\mathscr{V} \prec \mathscr{U}$, which means that for all $V \in \mathscr{V}$, there exists $U \in \mathscr{U}$ such that $V \subseteq U$. In particular, one can choose a morphism $\rho^{\mathscr{U},\mathscr{V}} : N_{\mathscr{V}} \to N_{\mathscr{U}}$ such that $\mathscr{V}_\sigma \subseteq \mathscr{U}_{\rho\sigma}$ for all $\sigma \in N_{\mathscr{V}}$. This choice is of course not necessarily unique. We would like to compare the Mayer-Vietoris spectral sequences of both covers. For this, we recall the definition of the Čech chain complex outlined on section 2.10 paying attention to the remark at the end. This leads to the following isomorphism on the terms from the 0-page

$$E^0_{p,q}(X, \mathscr{U}) := \bigoplus_{\sigma^p \in N_{\mathscr{U}}} C^{\mathrm{cell}}_q(\mathscr{U}_{\sigma^p}) \simeq \bigoplus_{s^q \in X} C^{\mathrm{cell}}_p(\mathscr{N}_{\mathscr{U}}(s^q)). \tag{8.4}$$

Here, $\mathscr{N}_{\mathscr{U}} \in \mathbf{RDiag}(X)$ denotes the nerve diagram over the regular CW complex X, so that $C^{\mathrm{cell}}_n(\mathscr{N}_{\mathscr{U}}) \in \mathbf{PreCosh}(X)$ for all $n \geq 0$. The isomorphism in (8.4) is given by sending a generator $(a^q)_{\sigma^p} \in \bigoplus_{\sigma^p \in N_{\mathscr{U}}} C^{\mathrm{cell}}_q(\mathscr{U}_\sigma)$ to its transpose $(\sigma^p)_{a^q}$, for all cells $a^q \in X$ and all $\sigma^p \in N_{\mathscr{U}}$.

Returning to a refinement $\mathscr{V} \prec \mathscr{U}$ and a morphism $\rho^{\mathscr{U},\mathscr{V}} : N_{\mathscr{V}} \to N_{\mathscr{U}}$, there is an induced

double complex morphism $\rho_{p,q}^{\mathscr{U},\mathscr{V}} : C_{p,q}(X,\mathscr{U}) \to C_{p,q}(X,\mathscr{V})$ given by

$$\rho_{p,q}^{\mathscr{U},\mathscr{V}}((\sigma^p)_{a^q}) = \begin{cases} (\rho^{\mathscr{U},\mathscr{V}}\sigma^p)_{a^q} & \text{if } \dim(\rho^{\mathscr{U},\mathscr{V}}\sigma^p) = p, \\ 0 & \text{otherwise,} \end{cases}$$

for all generators $(\sigma^p)_{a^q} \in C_{p,q}(X,\mathscr{U})$ with cells $\sigma^p \in N_{\mathscr{V}}$ and $a^q \in X$.

Lemma 8.4.1. *$\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ is a morphism of double complexes. Thus, it induces a morphism of spectral sequences*

$$\rho_{p,q}^{\mathscr{U},\mathscr{V}} : E_{p,q}^*(X,\mathscr{V}) \to E_{p,q}^*(X,\mathscr{U})$$

dependent on the choice of $\rho^{\mathscr{U},\mathscr{V}}$.

Proof. Let $\delta_{\mathscr{V}}$ and $\delta_{\mathscr{U}}$ denote the respective Čech differentials from $\check{C}_p(\mathscr{V};C_q^{\text{cell}})$ and $\check{C}_p(\mathscr{U};C_q^{\text{cell}})$. As we have the refinement chain morphism $\rho_*^{\mathscr{U},\mathscr{V}} : C_*^{\text{cell}}(N_{\mathscr{V}}) \to C_*^{\text{cell}}(N_{\mathscr{U}})$ we also have commutativity $\rho_{*,*}^{\mathscr{U},\mathscr{V}} \circ \delta^{\mathscr{V}} = \delta^{\mathscr{U}} \circ \rho_{*,*}^{\mathscr{U},\mathscr{V}}$. This implies that $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ commutes with the horizontal differential d^H. For commutativity with d^V, we consider a generating chain $(\sigma^p)_{a^q} \in E_{p,q}^0(X,\mathscr{V})$ with $\sigma^p \in N_{\mathscr{V}}$ and $a^q \in X$. Then if $\dim(\rho^{\mathscr{U},\mathscr{V}}\sigma^p) = p$ we have

$$\begin{aligned} \rho_{p,q-1}^{\mathscr{U},\mathscr{V}} \circ d^V\left((\sigma^p)_{a^q}\right) = \rho_{p,q-1}^{\mathscr{U},\mathscr{V}}\Big((-1)^p \sum_{b \preceq a^q} ([b:a^q]\sigma^p)_b\Big) &= (-1)^p \sum_{b \preceq a^q} ([b:a^q]\rho^{\mathscr{U},\mathscr{V}}\sigma^p)_b \\ &= (-1)^p d_q^{\text{cell}}\big((\rho^{\mathscr{U},\mathscr{V}}\sigma^p)_{a^q}\big) = d^V \circ \rho_{p,q}^{\mathscr{U},\mathscr{V}}\left((\sigma^p)_{a^q}\right) \end{aligned}$$

and for $\dim(\rho^{\mathscr{U},\mathscr{V}}\sigma^p) < p$ commutativity follows since both terms vanish.

A morphism of double complexes gives rise to a morphism of the vertical filtrations. By [83, Thm. 3.5] this induces a morphism of spectral sequences $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$. □

Since $\rho^{\mathscr{U},\mathscr{V}} : N_{\mathscr{V}} \to N_{\mathscr{U}}$ is not unique, the induced morphism $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ on the 0-page will in general not be unique either. We have, however, the following:

Proposition 8.4.2. *The 2-morphism obtained by restricting $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ is independent of the particular choice of refinement map $\rho^{\mathscr{U},\mathscr{V}} : N_{\mathscr{V}} \to N_{\mathscr{U}}$.*

Proof. We have to show that $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ is independent of the particular choice of the refinement morphism. First, define a carrier $R : N_{\mathscr{V}} \rightrightarrows N_{\mathscr{U}}$ by the assignment

$$\sigma \mapsto R(\sigma) = \left\{ \nu \in N_{\mathscr{U}} \mid V_\sigma \subseteq U_\nu \right\}.$$

The geometric realisation of the subcomplex $R(\sigma)$ is homeomorphic to a standard simplex, in particular contractible, so R is acyclic. Note that $\rho_{*,*}^{\mathscr{U},\mathscr{V}}$ is carried by R. Hence, by Thm. 2.7.3

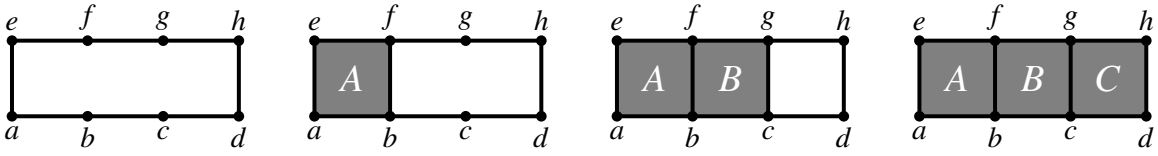


Figure 8.1: Cubical complex $\mathscr{C}_*$ at values 0,1,2 and 3.

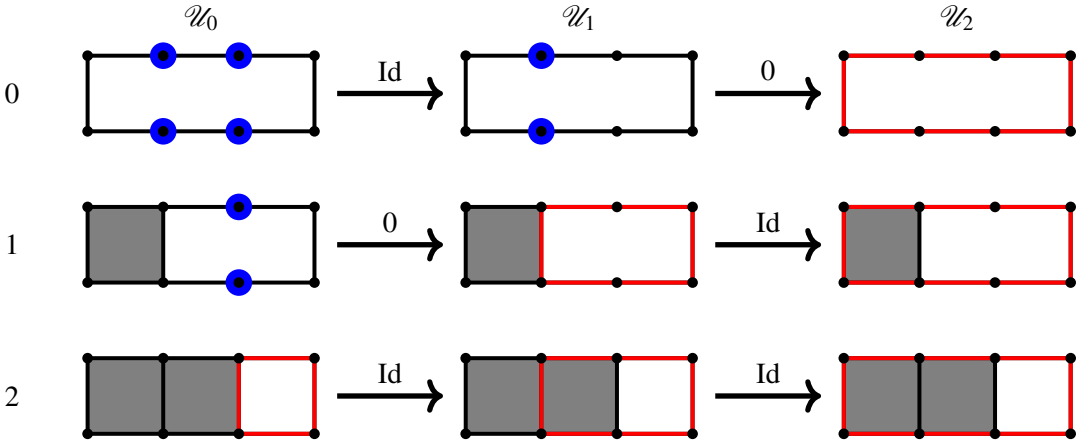


Figure 8.2: Cubical complex $\mathscr{C}_*$ with covers $\mathscr{U}_0$, $\mathscr{U}_1$ and $\mathscr{U}_2$, and with filtration values 0,1 and 2. Blue dots represent classes in $E^2_{1,0}(\mathscr{C},\mathscr{U}_i)$ and red loops represent classes on $E^2_{0,1}(\mathscr{C},\mathscr{U}_i)$, for $i=0,1,2$.

for any pair of refinement maps $\rho^{\mathscr{U},\mathscr{V}},\tau^{\mathscr{U},\mathscr{V}}: N_{\mathscr{V}} \to N_{\mathscr{U}}$, there exists a chain homotopy k_* : $C_n(N_{\mathscr{V}}) \to C_{n+1}(N_{\mathscr{U}})$ carried by R, so that

$$k_*\delta^{\mathscr{V}} + \delta^{\mathscr{U}} k_* = \tau_*^{\mathscr{U},\mathscr{V}} - \rho_*^{\mathscr{U},\mathscr{V}}$$

for all $n \geq 0$ and where $\tau_*^{\mathscr{U},\mathscr{V}}$ and $\rho_*^{\mathscr{U},\mathscr{V}}$ are induced morphisms of chain complexes $C_*(N_{\mathscr{V}}) \to C_*(N_{\mathscr{U}})$. In particular, using the same notation, this translates into chain homotopies $k_* : E^0_{p,q}(X,\mathscr{V}) \to E^0_{p+1,q}(X,\mathscr{U})$ on the 0-page such that

$$k_*\delta^{\mathscr{V}} + \delta^{\mathscr{U}} k_* = \tau_{*,*}^{\mathscr{U},\mathscr{V}} - \rho_{*,*}^{\mathscr{U},\mathscr{V}}$$

Thus, $\tau_{*,*}^{\mathscr{U},\mathscr{V}} = \rho_{*,*}^{\mathscr{U},\mathscr{V}}$ from the second page onward. □

Example 8.4.3. Consider a filtered cubical complex $\mathscr{C}_*$. At value 0, $\mathscr{C}_*$ is given by the vertices on $\mathscr{R}^2$ at the coordinates $a=(0,0), b=(1,0), c=(2,0), d=(3,0), e=(0,1), f=(1,1), g=(2,1), h=(3,1)$, together with all edges contained in the boundary of the rectangle $adhe$. Then, at value 1 there appears the edge bf with the face $abfe$. At value 2 the edge gc with the face $fgcb$, and finally at value 3 the face $ghdc$ appears. This is depicted on figure 8.1. Denote the three squares from $\mathscr{C}_*$ by $A=(a,b,f,e)$, $B=(b,c,g,f)$ and $C=(c,d,h,g)$, and consider the following

three covers of $\mathscr{C}_*$:

$$\mathscr{U}_0 = \left\{\overline{A}, \overline{B}, \overline{C}\right\}, \qquad \mathscr{U}_1 = \left\{\overline{A}, \overline{B \cup C}\right\}, \qquad \mathscr{U}_2 = \left\{\mathscr{C}_*\right\}.$$

The induced morphisms on second page terms at different filtration values are either null or the identity, as illustrated on figure 8.2. For example, on the top left of figure 8.2 a class in $E^2_{1,0}(\mathscr{C}, \mathscr{U}_0)(0)$ is indicated by the four blue dots which are contained in the respective intersections $\overline{A} \cap \overline{B}$ and $\overline{B} \cap \overline{C}$; applying the first page differential to these points leads to trivial classes in the covers from $\mathscr{U}_0$ (assuming that we take $\mathbb{Z}_2$ coefficients). This same class is still maintained at the term $E^2_{1,0}(\mathscr{C}, \mathscr{U}_0)(1)$, where it is represented by a pair of points from the intersection $\overline{B} \cap \overline{C}$. At value 2 we can see that this cycle appears on the term $E^2_{0,1}(\mathscr{C}, \mathscr{U}_0)(2)$ as a red loop; see bottom left of figure 8.2. On the other hand, the refinement induced morphism $E^2_{0,1}(\mathscr{C}, \mathscr{U}_0)(0) \to E^2_{0,1}(\mathscr{C}, \mathscr{U}_1)(0)$ sends the aforementioned class represented by the four blue dots to the class represented by two dots contained in $\overline{B} \cap \overline{C}$. However, notice that $E^2_{0,1}(\mathscr{C}, \mathscr{U}_0)(1) \to E^2_{0,1}(\mathscr{C}, \mathscr{U}_1)(1)$ is the zero morphism since the class that previously appeared in $E^2_{0,1}(\mathscr{C}, \mathscr{U}_1)(0)$ now has been "extended" to a class from $E^2_{1,0}(\mathscr{C}, \mathscr{U}_1)(1)$.

A consequence of Prop. 8.4.2 is that if we have a space X together with covers $\mathscr{U} \prec \mathscr{V} \prec \mathscr{U}$, then by uniqueness the morphism on the second page induced by the consecutive inclusions coincides with the identity. This happens for example when we consider two covers $\mathscr{U}$ and $\mathscr{V}$ from X and we consider the pair of covers which result from adding the whole space as a covering element: $\widetilde{\mathscr{U}} = \mathscr{U} \cup \{X\}$ and $\widetilde{\mathscr{V}} = \mathscr{V} \cup \{X\}$; in this case $\widetilde{\mathscr{U}}$ and $\widetilde{\mathscr{V}}$ refine one another. This gives rise to the next result.

Proposition 8.4.4. *Suppose a pair of covers $\mathscr{U}$ and $\mathscr{V}$ of X are a refinement of one another. Then there is a 2-spectral sequence isomorphism $E^2_{*,*}(X, \mathscr{U}) \simeq E^2(X, \mathscr{V})$.*

This corollary implies that for any cover $\mathscr{U}$ of X, the cover $\mathscr{U} \cup X$ obtained by adding the extra covering element X is such that the second page $E^2_{p,q}(X, \mathscr{U} \cup X)$ has only the first column nonzero.

Lemma 8.4.5. *Consider a cover $\mathscr{U}$ of a space X, and suppose that $X \in \mathscr{U}$. Then $E^2_{p,q}(X, \mathscr{U}) = 0$ for all $p > 0$.*

Proof. This follows from the observation that the cover $\{X\}$ consisting of a single element satisfies $\{X\} \prec \mathscr{U} \prec \{X\}$. Using Prop. 8.4.4 we therefore obtain isomorphisms $E^2_{p,q}(X, \mathscr{U}) \simeq E^2_{p,q}(X, \{X\})$, and the result follows. □

Suppose that none of the two covers $\mathscr{V}$ and $\mathscr{U}$ refines the other. One can still compare them using the common refinement $\mathscr{V} \cap \mathscr{U} = \left\{V \cap U\right\}_{V \in \mathscr{V}, U \in \mathscr{U}}$ which is a cover of X. Thus, there are

two refinement morphisms

$$E^2_{p,q}(X,\mathscr{U}) \xleftarrow{\rho^{\mathscr{U},\mathscr{V}\cap\mathscr{U}}_{p,q}} E^2_{p,q}(X,\mathscr{V}\cap\mathscr{U}) \xrightarrow{\rho^{\mathscr{V},\mathscr{V}\cap\mathscr{U}}_{p,q}} E^2_{p,q}(X,\mathscr{V}). \tag{8.5}$$

Following [105, Sec. 28] we can now build the double complex $C_{p,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)$ which is given by

$$\begin{array}{ccc}
\bigoplus_{\substack{\sigma^{p+1}\in N_{\mathscr{V}}\\ \tau^q\in N_{\mathscr{U}}}} \mathrm{PH}_k(\mathscr{V}_{\sigma^{p+1}}\cap\mathscr{U}_{\tau^q}) & \xleftarrow{(-1)^{p+1}\delta^{\mathscr{U}}} & \bigoplus_{\substack{\sigma^{p+1}\in N_{\mathscr{V}}\\ \tau^{q+1}\in N_{\mathscr{U}}}} \mathrm{PH}_k(\mathscr{V}_{\sigma^{p+1}}\cap\mathscr{U}_{\tau^{q+1}}) \\
\downarrow{\scriptstyle \delta^{\mathscr{V}}} & & \downarrow{\scriptstyle \delta^{\mathscr{V}}} \\
\bigoplus_{\substack{\sigma^{p}\in N_{\mathscr{V}}\\ \tau^q\in N_{\mathscr{U}}}} \mathrm{PH}_k(\mathscr{V}_{\sigma^{p}}\cap\mathscr{U}_{\tau^q}) & \xleftarrow{(-1)^{p}\delta^{\mathscr{U}}} & \bigoplus_{\substack{\sigma^{p}\in N_{\mathscr{V}}\\ \tau^{q+1}\in N_{\mathscr{U}}}} \mathrm{PH}_k(\mathscr{V}_{\sigma^{p}}\cap\mathscr{U}_{\tau^{q+1}})
\end{array}$$

for any pair of integers $p,q\geq 0$ and for each $k\geq 0$. From this double complex we can study the two associated spectral sequences

$$\begin{aligned}
{}^{\mathrm{I}}E^1_{p,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k) &= \bigoplus_{\sigma^p\in N_{\mathscr{V}}} \check{\mathscr{H}}_q\big((-1)^p\delta^{\mathscr{U}}|_{\mathscr{V}_{\sigma^p}\cap\mathscr{U}};\mathrm{PH}_k\big)\,, \\
{}^{\mathrm{II}}E^1_{p,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k) &= \bigoplus_{\tau^q\in N_{\mathscr{U}}} \check{\mathscr{H}}_p\big(\delta^{\mathscr{V}}|_{\mathscr{V}\cap\mathscr{U}_{\tau^q}};\mathrm{PH}_k\big)\,,
\end{aligned}$$

whose common target of convergence is $\check{\mathscr{H}}_n(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k)$ with $p+q=n$. For details about the spectral sequence associated to a double complex, the reader is recommended to look at [83, Thm. 2.15].

Example 8.4.6. Consider the cubical complex $\mathscr{C}_*$ from example 8.4.3. Set $\mathscr{U}=\mathscr{U}_1$, that is, $\mathscr{U}$ is the cover by the sets $U_1=A$ and $U_2=A\cup B$. On the other hand, consider $\mathscr{V}$ to be formed of $V_1=A\cup B$ and $V_2=C$. The double complex $C_{p,q}(\mathscr{C},\mathscr{V},\mathscr{U},\mathrm{PH}_k)$ is illustrated on figure 8.3 for filtration values 0 and 1, and for $k=0$. One can see that the refinement morphisms from (8.5) are actually projections.

Consider the nerve $N_{\mathscr{U}\cap\mathscr{V}}$ as a subset of the product of nerves $N_{\mathscr{U}}\times N_{\mathscr{V}}$. We have then two projections $\pi^{\mathscr{U}}: N_{\mathscr{U}\cap\mathscr{V}}\to N_{\mathscr{U}}$ and $\pi^{\mathscr{V}}: N_{\mathscr{U}\cap\mathscr{V}}\to N_{\mathscr{V}}$, both of which induce chain morphisms $\pi^{\mathscr{U}}_*: C_*(N_{\mathscr{U}\cap\mathscr{V}})\to C_*(N_{\mathscr{U}})$ and $\pi^{\mathscr{V}}_*: C_*(N_{\mathscr{U}\cap\mathscr{V}})\to C_*(N_{\mathscr{V}})$. These induce a pair of morphisms

$$\bigoplus_{\sigma^p\in N_{\mathscr{U}}} C^{\mathrm{cell}}_k(\mathscr{V}_{\sigma^p}) \xleftarrow{\pi^{\mathscr{V}}_{p,k}} \bigoplus_{\substack{\sigma^p\in N_{\mathscr{U}}\\ \tau^q\in N_{\mathscr{V}}}} C^{\mathrm{cell}}_k(\mathscr{V}_{\sigma^p}\cap\mathscr{U}_{\tau^q}) \xrightarrow{\pi^{\mathscr{U}}_{q,k}} \bigoplus_{\tau^q\in N_{\mathscr{V}}} C^{\mathrm{cell}}_k(\mathscr{U}_{\tau^q})\,,$$

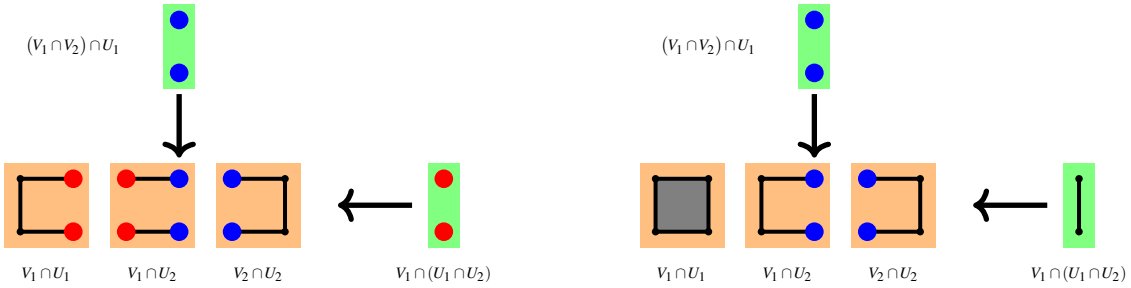


Figure 8.3: $C_{p,q}(\mathscr{C}, \mathscr{V}, \mathscr{U}, \mathrm{PH}_k)$ at filtration values 0 and 1.

for any pair of integers $p, q \geq 0$. The induced map $\pi^{\mathscr{V}}_{p,k}$ on $C_k(\mathscr{V}_{\sigma^p} \cap \mathscr{U}_{\tau^q})$ satisfies

$$\pi^{\mathscr{V}}_{p,k}\left((\sigma^p \times \tau^q)_s\right) = \begin{cases} (\sigma^p)_s & \text{if } \dim(\tau^q) = q = 0, \\ 0 & \text{else,} \end{cases}$$

for all $\sigma^p \in N_{\mathscr{V}}, \tau^q \in N_{\mathscr{U}}$ and a cell $a^k \in \mathscr{V}_\sigma \cap \mathscr{U}_\tau$. The map $\pi^{\mathscr{U}}_{*,*}$ acts similarly. By definition $\pi^{\mathscr{U}}_{*,*}$ and $\pi^{\mathscr{V}}_{*,*}$ both commute with the Čech differentials $\delta^{\mathscr{U}}$ and $\delta^{\mathscr{V}}$ respectively. Let $\sigma^p \in N_{\mathscr{V}}$ and $\tau^0 \in N_{\mathscr{U}}$. Then we have a commutative square

$$\begin{array}{ccc} \left(\sigma^p \times \tau^0\right)_{a^k} & \xrightarrow{\pi^{\mathscr{V}}_{*,*}} & \left(\sigma^p\right)_{a^k} \\ \downarrow d_n & & \downarrow d_n \\ \sum_{b \in \overline{a^k}} \left([b : a^k]\sigma^p \times \tau^0\right)_b & \xrightarrow{\pi^{\mathscr{V}}_{*,*}} & \sum_{b \in \overline{a^k}} \left([b : a^k]\sigma^p\right)_b \end{array}$$

for all cells $a^k \in \mathscr{V}_\sigma \cap \mathscr{U}_\tau$. This implies that $\pi^{\mathscr{V}}_{*,*}$ commutes with d_n and the same holds for $\pi^{\mathscr{U}}_{*,*}$. We obtain a morphism

$$\pi^{\mathscr{V}}_{p,k} \colon \bigoplus_{\substack{\sigma^p \in N_{\mathscr{V}} \\ \tau^0 \in N_{\mathscr{U}}}} C_k(\mathscr{V}_{\sigma^p} \cap \mathscr{U}_{\tau^0}) \to \bigoplus_{\sigma^p \in N_{\mathscr{V}}} C_k(\mathscr{V}_{\sigma^p}) \,,$$

commuting with d_* and $\delta^{\mathscr{V}}$ and $\delta^{\mathscr{U}}$. Thinking of the 0-th column ${}^{\mathrm{I}}E^0_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k)$ as a chain complex with Čech differential $\delta^{\mathscr{V}}$, one has a chain morphism

$$\pi^{\mathscr{V}}_{p,k} : {}^{\mathrm{I}}E^1_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k) \to \check{\mathscr{C}}_p(\mathscr{V}; \mathrm{PH}_k) = E^1_{p,k}(X, \mathscr{V})$$

for all $p \geq 0$. By the same argument there is another chain morphism

$$\pi^{\mathscr{U}}_{q,k} : {}^{\mathrm{II}}E^1_{0,q}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k) \to \check{\mathscr{C}}_q(\mathscr{U}; \mathrm{PH}_k) = E^1_{q,k}(X, \mathscr{U}) \,.$$

for all $q \geq 0$. There is a very natural way of understanding how much $\pi^{\mathscr{V}}_{p,k}$ fails to be an isomor-

phism. To do this, we take for each simplex $\sigma^p \in N_{\mathscr{V}}$, the Mayer-Vietoris spectral sequence for $\mathscr{V}_{\sigma^p}$ covered by $\mathscr{V}_{\sigma^p} \cap \mathscr{U}$

$$M^2_{q,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U}) \Rightarrow \mathrm{PH}_{q+k}(\mathscr{V}_{\sigma^p}),$$

where we changed the notation from $E^2_{q,k}(\mathscr{V}_{\sigma^p}, \mathscr{V}_{\sigma^p} \cap \mathscr{U})$ to $M^2_{q,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U})$ as it helps distinguishing this spectral sequence from ${}^{I}E^*_{p,q}$. Then

$${}^{\mathrm{I}}E^1_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k) = \bigoplus_{\sigma^p \in N_{\mathscr{V}}} M^2_{0,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U}) \,.$$

Here we notice that the restriction of $\pi^{\mathscr{V}}_{p,k}$ to the summand $M^2_{0,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U})$ is given by the composition

$$M^2_{0,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U}) \twoheadrightarrow M^\infty_{0,k}(\mathscr{V}_{\sigma^p} \cap \mathscr{U}) \hookrightarrow \mathrm{PH}_k(\mathscr{V}_{\sigma^p}).$$

In consequence there is an induced morphism on the second page 0^{th} column $\pi^{\mathscr{V}}_{p,k} : {}^{\mathrm{I}}E^2_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k) \to \check{\mathscr{H}}_p(\mathscr{V}; \mathrm{PH}_k)$. Notice that PH_0 is a cosheaf, and in this case $M^2_{0,0}(\mathscr{V}_{\sigma^p} \cap \mathscr{U}) = \mathrm{PH}_0(\mathscr{V}_{\sigma^p})$ for all $\sigma^p \in N_{\mathscr{V}}$. This implies that $\pi^{\mathscr{V}}_{p,0}$ is an isomorphism for all $p \geq 0$.

Now we turn to the morphism $\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k}$ defined by the composition

$$\check{\mathscr{H}}_p(\mathscr{V} \cap \mathscr{U}; \mathrm{PH}_k) \twoheadrightarrow {}^{\mathrm{I}}E^\infty_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k) \hookrightarrow {}^{\mathrm{I}}E^2_{p,0}(\mathscr{V}, \mathscr{U}, \mathrm{PH}_k) \xrightarrow{\pi^{\mathscr{V}}_{p,k}} \check{\mathscr{H}}_p(\mathscr{V}; \mathrm{PH}_k)$$

and notice that $\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k}$ is carried by the acyclic carrier sending $\sigma \times \tau \in N_{\mathscr{V}\cap\mathscr{U}}$ to $\Delta^\sigma \subseteq N_{\mathscr{V}}$. In particular, if $\mathscr{V} \prec \mathscr{U}$ then by using Lemma 8.4.5 we have ${}^{\mathrm{I}}E^1_{p,q} = 0$ for all $q > 0$ and the first two arrows in the definition of $\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k}$ are isomorphisms. Similarly, in this case we obtain $M^2_{q,k} = 0$ for all $q > 0$, and $\pi^{\mathscr{V}}_{p,k}$ becomes an isomorphism. Altogether, the inverse $(\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k})^{-1}$ is well-defined, and by composition we define morphisms $\theta^{\mathscr{U}, \mathscr{V}}_{p,k} = \theta^{\mathscr{U}, \mathscr{V}\cap\mathscr{U}}_{p,k} \circ (\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k})^{-1}$. Here notice that $\theta^{\mathscr{U}, \mathscr{V}\cap\mathscr{U}}_{p,k}$ is defined in an analogous way to $\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k}$, but using ${}^{\mathrm{II}}E^*_{0,p}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k)$ instead of ${}^{\mathrm{I}}E^*_{p,0}(\mathscr{V}, \mathscr{U}; \mathrm{PH}_k)$. The following proposition should also follow from applying an appropriate version of the universal coefficient theorem to [105, Prop. 4.4]. Instead we will prove the dual statement of this proposition by means of acyclic carriers.

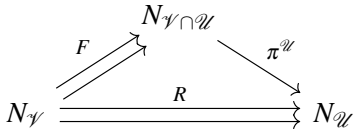
Proposition 8.4.7. *Suppose that $\mathscr{V} \prec \mathscr{U}$, and let $\rho^{\mathscr{U}, \mathscr{V}}$ denote a refinement map. The morphism $\theta^{\mathscr{U}, \mathscr{V}}_{p,k} : E^2_{p,k}(X, \mathscr{V}) \to E^2_{p,k}(X, \mathscr{U})$ coincides with the standard morphism induced by $\rho^{\mathscr{U}, \mathscr{V}}$.*

Proof. Since $\mathscr{V} \prec \mathscr{U}$, the morphism $\theta^{\mathscr{V}, \mathscr{V}\cap\mathscr{U}}_{p,k} : \check{\mathscr{H}}_p(\mathscr{V} \cap \mathscr{U}, \mathrm{PH}_k) \to \check{\mathscr{H}}_p(\mathscr{V}, \mathrm{PH}_k)$ is an isomor-

phism. Now consider the diagram

$$\begin{array}{ccccc}
\check{\mathscr{H}}_p(\mathscr{V};\mathrm{PH}_k) & & \xrightarrow{\qquad\qquad\rho^{\mathscr{U},\mathscr{V}}_{p,k}\qquad\qquad} & & \check{\mathscr{H}}_p(\mathscr{U};\mathrm{PH}_k) \\
\uparrow\simeq & & & & \uparrow\pi^{\mathscr{U}}_{p,k} \\
\check{\mathscr{H}}_p(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k) & \twoheadrightarrow & {}^{\mathrm{II}}E^{\infty}_{0,p}(\mathscr{V},\mathscr{U};\mathrm{PH}_k) & \hookrightarrow & {}^{\mathrm{II}}E^{2}_{0,p}(\mathscr{V},\mathscr{U};\mathrm{PH}_k).
\end{array}$$

To check that it commutes we study the following triangle of acyclic carriers



where R is defined in Prop. 8.4.2. The carrier F is given for every $\sigma \in N_{\mathscr{V}}$ by $F(\sigma) = \Delta^{\sigma} \times |R(\sigma)|$. Since F is acyclic, there exists $f_* : C_*(N_{\mathscr{V}}) \to C_*(N_{\mathscr{V}\cap\mathscr{U}})$ inducing a chain morphism $f_* : \check{\mathscr{C}}_p(\mathscr{V}, S_k) \to \check{\mathscr{C}}_p(\mathscr{V}\cap\mathscr{U}, S_k)$ by the assignment $(\sigma)_s \mapsto (f_*(\sigma))_s$ for all simplices $s \in S_k(X)$ and all $\sigma \in N_{\mathscr{V}}$. In fact, F defines an acyclic equivalence by considering the inverse carrier $P : N_{\mathscr{V}\cap\mathscr{U}} \rightrightarrows N_{\mathscr{V}}$ sending $\sigma \times \tau$ to Δ^{σ}. In this case the shift carrier $I_{\mathscr{V}} : N_{\mathscr{V}} \rightrightarrows N_{\mathscr{V}}$ is given by the assignment $\sigma \mapsto \Delta^{\sigma}$, and $I_{\mathscr{V}\cap\mathscr{U}} : N_{\mathscr{V}\cap\mathscr{U}} \rightrightarrows N_{\mathscr{V}\cap\mathscr{U}}$ is given by $\sigma \times \tau \mapsto \Delta^{\sigma} \times |R(\sigma)|$. As $\theta^{\mathscr{V},\mathscr{V}\cap\mathscr{U}}_{p,k}$ is carried by P, this implies that $f_* = \left(\theta^{\mathscr{V},\mathscr{V}\cap\mathscr{U}}_{p,k}\right)^{-1}$ as morphisms $\check{\mathscr{H}}_p(\mathscr{V}, \mathrm{PH}_k) \to \check{\mathscr{H}}_p(\mathscr{V}\cap\mathscr{U}, \mathrm{PH}_k)$. Consequently, $\theta^{\mathscr{U},\mathscr{V}}_{p,k}$ is carried by $\pi^{\mathscr{U}} F = R$. Altogether, we obtain the equality $\theta^{\mathscr{U},\mathscr{V}}_{p,k} = \rho^{\mathscr{U},\mathscr{V}}_{p,k}$ as morphisms $\check{\mathscr{H}}_p(\mathscr{V}, \mathrm{PH}_k) \to \check{\mathscr{H}}_p(\mathscr{U}, \mathrm{PH}_k)$ since both are carried by R. □

Still assuming that $\mathscr{V} \prec \mathscr{U}$, we now look for conditions for the existence of an inverse $\varphi^{\mathscr{V},\mathscr{U}}_{p,k}$: $E^2_{p,k}(X,\mathscr{U}) \to E^2_{p,k}(X,\mathscr{V})$ of $\theta^{\mathscr{U},\mathscr{V}}_{p,k}$.

Proposition 8.4.8. *Suppose that $\mathscr{V} \prec \mathscr{U}$. If $M^2_{p,k}(\mathscr{V}\cap\mathscr{U}_{\tau^q}) = 0$ for all $p > 0$, $k \geq 0$ and all $\tau^q \in N_{\mathscr{U}}$, then the maps $\theta^{\mathscr{U},\mathscr{V}}_{*,*}$ induce a 2-isomorphism of spectral sequences*

$$E^{\geq 2}_{*,*}(X,\mathscr{U}) \simeq E^{\geq 2}_{*,*}(X,\mathscr{V}).$$

Proof. By Prop. 8.4.2 and Prop.8.4.7 we can choose a refinement map $\rho^{\mathscr{U},\mathscr{V}} : N_{\mathscr{V}} \to N_{\mathscr{U}}$ giving a morphism of spectral sequences

$$\rho^{\mathscr{U},\mathscr{V}}_{*,*} : E^{\geq 2}_{*,*}(X,\mathscr{V}) \to E^{\geq 2}_{*,*}(X,\mathscr{U})$$

that coincides with $\theta^{\mathscr{U},\mathscr{V}}_{*,*}$. Our assumption about $M^2_{p,k}$ implies ${}^{\mathrm{II}}E^2_{p,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k) = 0$ for all $p > 0$, which in turn gives

$$\mathrm{Ker}\Big(\check{\mathscr{H}}_q(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k) \twoheadrightarrow {}^{\mathrm{II}}E^{\infty}_{0,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k)\Big) = 0 \tag{8.6}$$

and

$$\mathrm{Coker}\Big({}^{\mathrm{II}}E^{\infty}_{0,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k)\hookrightarrow {}^{\mathrm{II}}E^{2}_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)\Big)=0. \tag{8.7}$$

Now note that $\pi^{\mathscr{U}}_{q,k}$ yields an isomorphism ${}^{\mathrm{II}}E^{2}_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)\simeq \check{\mathscr{H}}_q(\mathscr{U},\mathrm{PH}_k)$. This shows that $\theta^{\mathscr{U},\mathscr{V}}_{q,k}$ is a composition of isomorphisms; thus the statement follows. □

We will now relax the conditions in Prop. 8.4.8 and use the relations of *left-interleaving* and *right-interleaving* of persistence modules (denoted by $\sim^{\varepsilon}_{L}$ and $\sim^{\varepsilon}_{R}$, respectively) to achieve this (see section 5.2 or [65, § 4]). We have to adapt [65, Prop. 4.14].

Lemma 8.4.9. *Suppose that we have persistence modules A, B and C, and a parameter $\varepsilon\geq 0$ such that $A\sim^{\varepsilon}_{R}B$ and $B\sim^{\varepsilon}_{L}C$. Denote by Φ the morphism $\Phi: A\to C$ given by the composition $A\twoheadrightarrow B\hookrightarrow C$. Then there exists $\Psi: C\to A[2\varepsilon]$ such that Φ and Ψ define a 2ε-interleaving $A\sim^{2\varepsilon}C$.*

Proof. By hypothesis, we have a sequence

$$\mathscr{E}_1\longrightarrow A\xrightarrow{\ f\ }B\xhookrightarrow{\ g\ }C\longrightarrow\mathscr{E}_2$$

which is exact in A and C and where $\mathscr{E}_1\sim^{\varepsilon}0$ and $\mathscr{E}_2\sim^{\varepsilon}0$. Then, let $v\in C$ and notice that $\Sigma^{\varepsilon}C(v)\in\mathrm{Im}(g)$. Thus, there exists a unique vector $w\in B$ such that $g(w)=\Sigma^{\varepsilon}C(v)$. On the other hand, there exists $z\in A$, not necessarily unique, such that $f(z)=w$. This defines a unique element $\Sigma^{\varepsilon}A(z)\in A$. To see this, suppose that another $z'\in A$ is such that $f(z')=w$. Then $f(z-z')=0$ and $z-z'\in\mathrm{Ker}(f)$, which implies $0=\Sigma^{\varepsilon}A(z-z')=\Sigma^{\varepsilon}A(z)-\Sigma^{\varepsilon}A(z')$, and then $\Sigma^{\varepsilon}A(z)=\Sigma^{\varepsilon}A(z')$. Altogether, we set $\Psi=\Sigma^{\varepsilon}A\circ\Phi^{-1}\circ\Sigma^{\varepsilon}C$, which is well-defined. □

Recall that for $\mathscr{V}\prec\mathscr{U}$ we have that $\check{\mathscr{H}}_q(\mathscr{V};\mathrm{PH}_k)\simeq\check{\mathscr{H}}_q(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k)$ for all $k\geq 0$ and $q\geq 0$. There is a natural way to relax (8.6) and (8.7) to the persistent case. We assume that for $\varepsilon\geq 0$, there are right and left interleavings

$$\check{\mathscr{H}}_q(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k)\sim^{\varepsilon}_{R}{}^{\mathrm{II}}E^{\infty}_{0,q}(\mathscr{V},\mathscr{U};\mathrm{PH}_k)\sim^{\varepsilon}_{L}{}^{\mathrm{II}}E^{2}_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k). \tag{8.8}$$

If we define $\Phi_{q,k}:\check{\mathscr{H}}_q(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k)\to{}^{\mathrm{II}}E^{2}_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)$ to be the composition of the associated persistence morphisms as in Lem. 8.4.9, then there exists

$$\Psi_{q,k}:{}^{\mathrm{II}}E^{2}_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)\to\check{\mathscr{H}}_q(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_k)[2\varepsilon],$$

such that $\Phi_{q,k}$ and $\Psi_{q,k}$ define a 2ε-interleaving. We repeat this argument for the local Mayer-

Vietoris spectral sequences. Assume that for some $\nu \geq 0$ there are interleavings

$$ {}^{\mathrm{II}}E^1_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k) \sim^{\nu}_R \bigoplus_{\tau^q \in N_{\mathscr{U}}} M^{\infty}_{k,0}(\mathscr{V}\cap\mathscr{U}_{\tau^q}) \sim^{\nu}_L \bigoplus_{\tau^q \in N_{\mathscr{U}}} \mathrm{PH}_k(\mathscr{U}_{\tau^q}). \tag{8.9} $$

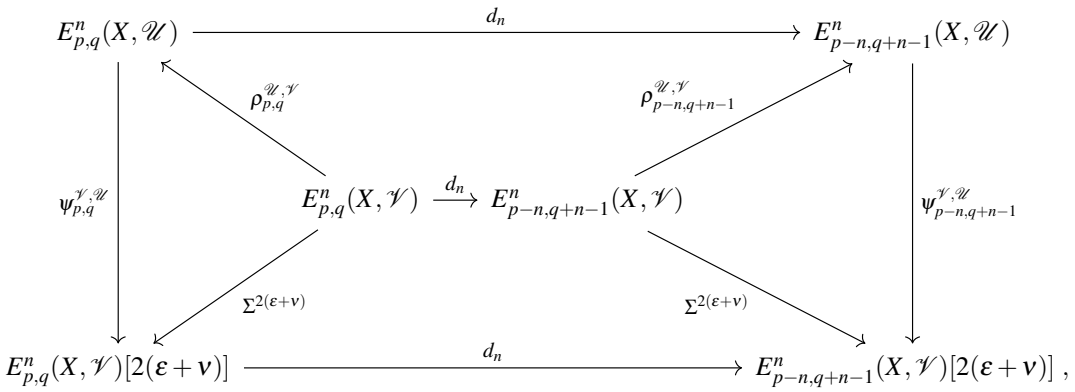
Let $\Pi_{q,k} : {}^{\mathrm{II}}E^1_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k) \to \bigoplus_{\tau^q\in N_{\mathscr{U}}} \mathrm{PH}_k(\mathscr{U}_{\tau^q})$ be the composition of the associated morphisms. By Lem. 8.4.9 there exists $\Xi_{q,k}$ such that $\Pi_{q,k}$ and $\Xi_{q,k}$ define a 2ν-interleaving. By slight abuse of notation we will continue to denote the induced 2ν-interleaving between ${}^{\mathrm{II}}E^2_{0,q}(\mathscr{V},\mathscr{U},\mathrm{PH}_k)$ and $\check{\mathscr{H}}_q(\mathscr{U};\mathrm{PH}_*)$ by $\Pi_{q,k}$ and $\Xi_{q,k}$. Altogether we have that $\theta^{\mathscr{U},\mathscr{V}}_{q,k} = \Pi_{q,k}\circ\Phi_{q,k}\circ(\theta^{\mathscr{V},\mathscr{V}\cap\mathscr{U}}_{q,k})^{-1}$ and in this situation there is indeed an 'inverse' $\varphi^{\mathscr{V},\mathscr{U}}_{q,k} = \theta^{\mathscr{V},\mathscr{V}\cap\mathscr{U}}_{q,k}\circ\Psi_{q,k}\circ\Xi_{q,k}$, which increases the persistence values by $2(\varepsilon+\nu)$.

Theorem 8.4.10. *Suppose that $\mathscr{V}\prec\mathscr{U}$ and for $\varepsilon\geq 0$ and $\nu\geq 0$ the interleavings in (8.8) and (8.9) hold. Then*

$$ \varphi^{\mathscr{V},\mathscr{U}}_{p,q} : E^*_{p,q}(X,\mathscr{U}) \to E^*_{p,q}(X,\mathscr{V})[2(\varepsilon+\nu)] $$

defines a 2-morphism of spectral sequences such that $\theta^{\mathscr{U},\mathscr{V}}_{p,q}$ and $\varphi^{\mathscr{V},\mathscr{U}}_{p,q}$ is a 2-page $2(\varepsilon+\nu)$-interleaving between $E^*_{p,q}(X,\mathscr{U})$ and $E^*_{p,q}(X,\mathscr{V})$.

Proof. The only thing that remains to be proved is that $\psi^{\mathscr{V},\mathscr{U}}_{p,q}$ commutes with the spectral sequence differentials d_n for all $n\geq 2$. Since these differentials commute with the shift morphisms $\Sigma^{2(\varepsilon+\nu)}$, this follows from considering the diagram



in which the two trapeziums and the two triangles commute. □

Example 8.4.11. Consider a cubical complex $\mathscr{C}_*$ as shown in Fig. 8.4, together with the covers $\mathscr{V}=\{\overline{A},\overline{B},\overline{C},\overline{D}\}$ and $\mathscr{U}=\{\overline{A\cup B},\overline{C\cup D}\}$, see Fig. 8.4 for the cells A,B,C and D. In this case we have

$$ \check{\mathscr{H}}_1(\mathscr{V};\mathrm{PH}_0) \simeq \check{\mathscr{H}}_1(\mathscr{V}\cap\mathscr{U};\mathrm{PH}_0) \simeq \mathrm{I}(0,1+\varepsilon)\oplus\mathrm{I}(1,1+\varepsilon) \sim^{\varepsilon} \mathrm{I}(0,1) \simeq {}^{\mathrm{II}}E^2_{0,1}(\mathscr{V},\mathscr{U},\mathrm{PH}_0) $$

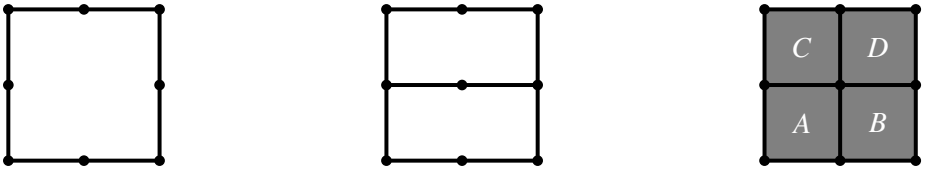


Figure 8.4: Cubical complex $\mathscr{C}_*$ at values 0,1 and $1+\varepsilon$.

Figure 8.5: Morphisms $\theta_{1,0}^{\mathscr{U},\mathscr{V}}$ along $[0,1)$ and along $[1,1+\varepsilon)$.

and also

$$^{\mathrm{II}}E^1_{0,0}(\mathscr{V},\mathscr{U},\mathrm{PH}_1) \simeq 0 \sim^{\varepsilon} \mathrm{I}(1,1+\varepsilon) \oplus \mathrm{I}(1,1+\varepsilon) \simeq \bigoplus_{\tau^0 \in N_{\mathscr{U}}} \mathrm{PH}_1(\mathscr{U}_{\tau^0}).$$

These interleavings are shown in Fig. 8.5. Thm. 8.4.10 implies that there is a 4ε-interleaving between $E^*_{p,q}(X,\mathscr{U})$ and $E^*_{p,q}(X,\mathscr{V})$. Notice that in this example, the nontrivial interleaved terms are in different positions of the spectral sequences. Therefore we can improve the upper bound to 2ε. We will use this observation later in Prop. 8.5.1.

8.5 Interleavings with respect to different covers

Consider $X \in$ **FCW-cpx**, together with a pair of covers $\mathscr{W}$ and $\mathscr{U}$ so that $\mathscr{W} \prec \mathscr{U}$. Motivated by the interleaving constructed in Thm. 8.4.10 we take a closer look at the following finite sequence of covers interpolating between $\mathscr{W}$ and a cover that both refines and is refined by $\mathscr{U}$. Let the strict r-th intersections of $\mathscr{U}$ be the family of sets $\mathscr{U}^r = \{\mathscr{U}_{\tau^r}\}_{\tau^r \in N_{\mathscr{U}}}$ for all $r \geq 0$. We define the $(r,\mathscr{W},\mathscr{U})$-*interpolation* as the covering set $\mathscr{W}^r = \mathscr{W} \cup \mathscr{U}^r$. In particular, note that the $(0,\mathscr{W},\mathscr{U})$-interpolation has the property that $\mathscr{W}^0 \prec \mathscr{U} \prec \mathscr{W}^0$, and consequently $E^2_{p,q}(X,\mathscr{U}) \simeq E^2_{p,q}(X,\mathscr{W}^0)$. In addition if $\mathscr{U}$ is a finite cover, then we will have $\mathscr{U}^N = \emptyset$ for $N \geq 0$ sufficiently large and consequently $\mathscr{W}^N = \mathscr{W}$.

Proposition 8.5.1 (Local Checks)**.** *Let $\mathscr{W} \prec \mathscr{U}$ be a pair of covers for X, where $\mathscr{U}$ is finite. Let $N \geq 0$ be such that $\mathscr{U}^N = \emptyset$. For every $0 \leq r \leq N$, we assume that there exist $\varepsilon_r \geq 0$ and $\nu_r \geq 0$ such that for all $\tau^r \in N_{\mathscr{U}}$*

$$E^2_{0,q}\big(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{|\mathscr{U}_{\tau^r}}\big) \sim^{\nu_r}_R E^{\infty}_{0,q}\big(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{|\mathscr{U}_{\tau^r}}\big) \sim^{\nu_r}_L \mathrm{PH}_q(\mathscr{U}_{\tau^r})$$

and also

$$d_I(E^2_{p,q}(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{|\mathscr{U}_{\tau^r}}), 0) \leq \varepsilon_r \ .$$

for all $p > 0$, and $q \geq 0$. Then we have that

$$d_I^2\big(E^*_{p,q}(X, \mathscr{W}^k), E^*_{p,q}(X, \mathscr{W}^{k+1})\big) \leq 2\max(\varepsilon_r, \nu_r).$$

Therefore, by using the triangle inequality, we obtain

$$d_I^2\big(E^*_{p,q}(X, \mathscr{U}), E^*_{p,q}(X, \mathscr{W})\big) \leq \sum_{k=0}^{N} 2\max(\varepsilon_r, \nu_r) \ .$$

Proof. We need to consider the spectral sequence ${}^{\mathrm{II}}E^2_{p,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k)$. Note that by construction of the covers $\mathscr{W}^r$ we have that for each $\tau^r \in N_{\mathscr{U}}$ with $\dim(\tau^r) = r > 0$ the set $\mathscr{W}^r_{\tau^r}$ is contained in one of the open sets from $\mathscr{W}^{r+1}$. By Lemma 8.4.5 this implies that ${}^{\mathrm{II}}E^1_{p,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) = 0$ for all $p > 0$ and $q > 0$ and $k \geq 0$. Moreover, we have that ${}^{\mathrm{II}}E^1_{0,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) = \bigoplus_{\tau^r \in N^q_{\mathscr{W}^r}} \mathrm{PH}_k(\mathscr{W}^r_{\tau^r})$ for all $q > 0$ and $k \geq 0$. The resulting spectral sequence is shown in Fig. 8.6.

As a consequence of these observations condition (8.9) holds for these indices with $\nu = 0$. In addition, ${}^{\mathrm{II}}E^2_{0,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) = E^2_{q,k}(X, \mathscr{W}^r)$ holds for all $q \geq 2$ and $k \geq 0$ (see Fig. 8.6 and 8.7). In particular, there is only one possible non-trivial differential for each entry in the bottom row as indicated in Fig. 8.7. Note that our hypothesis $d_I(E^2_{p,q}(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{|\mathscr{U}_{\tau^r}}), 0) \leq \varepsilon_r$ applies to the entries in the first column with $p > 0$ and gives left and right interleavings of the form

$$\check{\mathscr{H}}_q(\mathscr{W}^{r+1} \cap \mathscr{W}^r; \mathrm{PH}_k) \sim^{\varepsilon_r}_R {}^{\mathrm{II}}E^\infty_{0,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) \sim^{\varepsilon_r}_L {}^{\mathrm{II}}E^2_{0,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k)$$

for all $q > 0$ and $k \geq 0$. Hence, condition (8.8) holds with value ε_r.

$$\begin{array}{ccccccc}
{}^{\mathrm{II}}E^1_{2,0}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) & & 0 & & 0 & & \ddots \\
{}^{\mathrm{II}}E^1_{1,0}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) & & 0 & & 0 & & 0 \\
{}^{\mathrm{II}}E^1_{0,0}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) & \xleftarrow{d_1} & \bigoplus_{\tau^1 \in N_{\mathscr{W}^r}} \mathrm{PH}_k(\mathscr{W}^r_{\tau^1}) & \longleftarrow & \bigoplus_{\tau^2 \in N_{\mathscr{W}^r}} \mathrm{PH}_k(\mathscr{W}^r_{\tau^2}) & \longleftarrow & \bigoplus_{\tau^3 \in N_{\mathscr{W}^r}} \mathrm{PH}_k(\mathscr{W}^r_{\tau^3})
\end{array}$$

Figure 8.6: First page of ${}^{\mathrm{II}}E^*_{p,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k)$.

Let us look now at the case $q = 0$. Here we have $\check{\mathscr{H}}_0(\mathscr{W}^{r+1} \cap \mathscr{W}^r; \mathrm{PH}_k) = {}^{\mathrm{II}}E^2_{0,0}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k)$ and consequently (8.8) holds with value $\varepsilon = 0$. Next, by hypothesis, for all $k \geq 0$ we have right

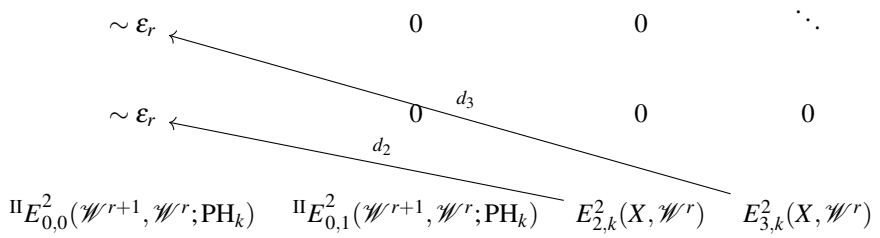


Figure 8.7: Second page of ${}^{\mathrm{II}}E^*_{p,q}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k)$ together with higher differentials.

and left interleavings

$$M^2_{0,k}\big(\mathscr{U}_{\tau^r} \cap \mathscr{W}^{r+1}\big) \sim^{\nu_r}_R M^\infty_{0,k}\big(\mathscr{U}_{\tau^r} \cap \mathscr{W}^{r+1}\big) \sim^{\nu_r}_L \mathrm{PH}_k(\mathscr{U}_{\tau^r}) \,,$$

for all $\tau^r \in N_{\mathscr{U}}$. Thus by taking the direct sum of these interleavings we obtain

$${}^{\mathrm{II}}E^1_{0,0}(\mathscr{W}^{r+1}, \mathscr{W}^r; \mathrm{PH}_k) \sim^{\nu_r}_R \bigoplus_{\tau^r \in N_0^{\mathscr{W}^r}} M^\infty_{0,k}(\mathscr{W}^r_{\tau^r} \cap \mathscr{W}^{r+1}) \sim^{\nu_r}_L E^1_{0,k}(X, \mathscr{W}^r) \,.$$

and condition (8.9) also holds for $q = 0$. The result now follows from Thm. 8.4.10.

Notice that we can slightly improve the statement of Thm. 8.4.10 here: For each term in the bottom row of the spectral sequence in this particular example only one of the two conditions (8.8) and (8.9) is nontrivial, and the proof of Thm. 8.4.10 carries over with $2\max(\varepsilon_r, \nu_r)$ replacing $2(\varepsilon_r + \nu_r)$. □

Remark. Notice that for reasonable cases the parameters ν_r are bounded above by $K\varepsilon_r$ for some constant $K > 0$ by a result from [65]. Nevertheless, we would like to keep ν_r and ε_r separated here, since we hope to compute it from $M^*_{p,k}\big(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{|\mathscr{U}_{\tau^r}}\big)$ for $\tau^r \in N_{\mathscr{U}}$ hereby get more accurate estimates. Intuitively, asking for ε_r and ν_r to be small is equivalent to asking for cycle representatives in covers from $\mathscr{W}^r$ to be approximately contained in covering sets from $\mathscr{W}^{r+1}$.

Finally, we would like to compare two separate covers $\mathscr{U}$ and $\mathscr{V}$ and have an estimate for the interleaving distance between the associated spectral sequences. The main idea of Prop. 8.5.1 is to translate this comparison problem into a few local checks that can be run in parallel. We formalize this in the following Corollary.

Corollary 8.5.2 (Stability of Covers)**.** *Consider two pairs $(X, \mathscr{U})$ and $(X, \mathscr{V})$, where X is a space and $\mathscr{U}$ and $\mathscr{V}$ are covers. Let $\mathscr{W} = \mathscr{U} \cap \mathscr{V}$ and denote by $\mathscr{W}^r_{\mathscr{U}}$ and $\mathscr{W}^r_{\mathscr{V}}$ the respective $(r, \mathscr{W}, \mathscr{U})$ and $(r, \mathscr{W}, \mathscr{V})$ interpolations. For every $0 \leq r \leq N$, we assume that there exist $\varepsilon_r, \varepsilon'_r \geq 0$ and*

$\nu_r, \nu'_r \geq 0$ such that for all $\tau^r \in N_{\mathscr{U}}$ and $\sigma^r \in N_{\mathscr{V}}$

$$E^2_{0,q}\big(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{\mathscr{U}}\big) \sim^{\nu_r}_R E^\infty_{0,q}\big(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{\mathscr{U}}\big) \sim^{\nu_r}_L \mathrm{PH}_q(\mathscr{U}_{\tau^r}),$$
$$E^2_{0,q}\big(\mathscr{V}_{\sigma^r}, \mathscr{W}^{r+1}_{\mathscr{V}}\big) \sim^{\nu'_r}_R E^\infty_{0,q}\big(\mathscr{V}_{\sigma^r}, \mathscr{W}^{r+1}_{\mathscr{V}}\big) \sim^{\nu'_r}_L \mathrm{PH}_q(\mathscr{V}_{\sigma^r}),$$

for all $r \geq 0$, and also

$$d_I(E^2_{p,q}(\mathscr{U}_{\tau^r}, \mathscr{W}^{r+1}_{\mathscr{U}}), 0) \leq \varepsilon_r \qquad , \qquad d_I(E^2_{p,q}(\mathscr{V}_{\sigma^r}, \mathscr{W}^{r+1}_{\mathscr{V}}), 0) \leq \varepsilon'_r$$

for all $p > 0$, and $q \geq 0$. Then we have that

$$d_I^2\big(E^*_{p,q}(X, \mathscr{U}), E^*_{p,q}(X, \mathscr{V})\big) \leq R(\mathscr{U}, \mathscr{V})$$

where $R(\mathscr{U}, \mathscr{V}) = \max\Big(\sum_{r=0}^N 2\max(\varepsilon_r, \nu_r), \sum_{r=0}^N 2\max(\varepsilon'_r, \nu'_r)\Big)$.

Proof. By Lemma 8.4.1 there are double complex morphisms given by the refinement maps

$$\check{\mathscr{C}}_p(\mathscr{U}, C^{\mathrm{cell}}_q) \xleftarrow{\rho^{\mathscr{U},\mathscr{W}}_{p,q}} \check{\mathscr{C}}_p(\mathscr{W}, C^{\mathrm{cell}}_q) \xrightarrow{\rho^{\mathscr{V},\mathscr{W}}_{p,q}} \check{\mathscr{C}}_p(\mathscr{V}, C^{\mathrm{cell}}_q)\,.$$

In turn, these induce 2-morphisms of spectral sequences

$$E^2_{p,q}(X, \mathscr{U}) \xleftarrow{\rho^{\mathscr{U},\mathscr{W}}_{p,q}} E^2_{p,q}(X, \mathscr{W}) \xrightarrow{\rho^{\mathscr{V},\mathscr{W}}_{p,q}} E^2_{p,q}(X, \mathscr{V})\,.$$

Let $\psi^{\mathscr{U},\mathscr{W}}_{p,q}$ and $\psi^{\mathscr{V},\mathscr{W}}_{p,q}$ be the 'inverses' of $\rho^{\mathscr{U},\mathscr{W}}_{p,q}$ and $\rho^{\mathscr{V},\mathscr{W}}_{p,q}$, respectively, witnessing the interleavings of the two spectral sequences (see Thm. 8.4.10 and Prop. 8.5.1). The result follows from considering the commutative diagram

$$\begin{array}{ccccc}
E^2_{p,q}(X, \mathscr{U}) & \xleftarrow{\rho^{\mathscr{U},\mathscr{W}}_{p,q}} & E^2_{p,q}(X, \mathscr{W}) & \xrightarrow{\rho^{\mathscr{V},\mathscr{W}}_{p,q}} & E^2_{p,q}(X, \mathscr{V}) \\
\downarrow{\scriptstyle \Sigma^{R(\mathscr{V},\mathscr{U})}} & \searrow^{\psi^{\mathscr{W},\mathscr{U}}_{p,q}} & \downarrow{\scriptstyle \Sigma^{R(\mathscr{V},\mathscr{U})}} & \swarrow_{\psi^{\mathscr{W},\mathscr{V}}_{p,q}} & \downarrow{\scriptstyle \Sigma^{R(\mathscr{V},\mathscr{U})}} \\
E^2_{p,q}(X, \mathscr{U})[R(\mathscr{V}, \mathscr{U})] & \xleftarrow{\rho^{\mathscr{U},\mathscr{W}}_{p,q}} & E^2_{p,q}(X, \mathscr{W})[R(\mathscr{V}, \mathscr{U})] & \xrightarrow{\rho^{\mathscr{V},\mathscr{W}}_{p,q}} & E^2_{p,q}(X, \mathscr{V})[R(\mathscr{V}, \mathscr{U})]
\end{array}$$

where all arrows are 2-morphisms of spectral sequences. □

Chapter 9

PerMaViss

The aim of this chapter is to explain how one implements in practice the persistence spectral sequence outlined in chapter 6. In doing so, we will review some examples that will illustrate how these objects behave. Currently, our code can be found in the python module PERMAVISS [116], which stands for **Per**sistence **Ma**yer **Vi**etoris **s**pectral **s**equence. Further, we will explain and describe the main PERMAVISS pipeline, but avoiding getting too much focused on the details: for these, see the PERMAVISS documentation. As a programming language, we have chosen Python as its syntax is easy to read. This module handles the particular case of Vietoris Rips complexes with spatial covers taken from point clouds. However, the code implementation has been kept as abstract as possible so that other examples, such as the $(K, \mathscr{P})$ join diagram, can be implemented in the future. Just before we start, a couple of comments about some of the special notation from this chapter:

- We will write commands after the signs `>>>`
- We will denote by "`np`" the NumPy python library.
- Matrices will be encoded using the `np.array` object, see the NumPy documentation.
- Given a numpy array `M`, we denote the i^{th} row by `M[i]` while we denote the i^{th} column by `M[:,i]`.
- We will often use python lists. These objects are given by square brackets containing some elements, such as `[a, 1, 64]`. We will denote by `[]` the empty list. Given a list `A`, we will use the notation $a \in$ `A` to indicate that a is an element on the list `A`.

9.1 Quick Start Guide to PERMAVISS

We start by revising how to run the code. First, before installing PERMAVISS, make sure that you have python3 updated, as well as the modules `scipy` and `numpy` installed. Then, run the command

```
>>> pip3 install permaviss
```

or alternatively you can download and install locally:

```
>>> git clone https://github.com/atorras1618/PerMaViss
>>> cd PerMaViss
>>> pip3 install -e .
```

As a first explanation of PERMAVISS, we will review a simple example using spectral sequences to compute persistent homology on a noisy circle. We start by taking 200 points in a noisy circle of radius 1. Next, we use the `maxmin` procedure to take a subample of 50 points, which is implemented in the function `take_sample`

```
>>> from permaviss.sample_point_clouds.examples import (random_circle, take_sample)
>>> point_cloud = random_circle(200, 1, epsilon=0.3)
>>> point_cloud = take_sample(point_cloud, 50)
```

Now we set the parameters for generating a `spectral_sequence` object instance. These are

- a prime number `p`,
- the maximum dimension of the Rips Complex `max_dim`,
- the maximum radius of the Rips filtration `max_r`,
- the number of divisions `max_div` along the maximum range in `point_cloud`,
- and the `overlap` between different covering regions.

In our case, we set the parameters to cover our circle with 4 covering square regions. Notice that in order for the algorithm to give the correct result we need `overlap` $>$ `max_r`.

```
>>> p = 5
>>> max_dim = 3
>>> overlap = 1
>>> max_div = 2
>>> max_r = overlap*0.99
```

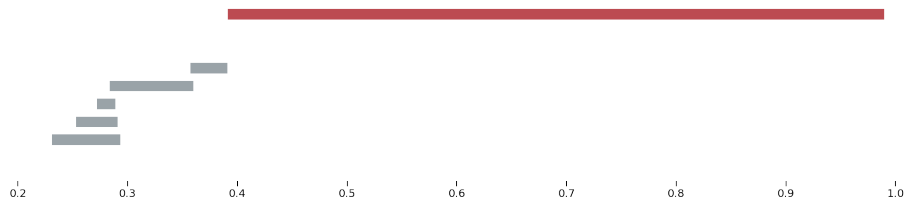
Then, we compute the spectral sequence, noticing that the method prints the successive page ranks.

```
>>> from permaviss.spectral_sequence.MV_spectral_seq import create_MV_ss
>>> MV_ss = create_MV_ss(point_cloud, max_r, max_dim, max_div, overlap, p)
PAGE: 1
[[ 0  0  0  0]
 [ 7  2  0  0]
 [79 29  0  0]]
PAGE: 2
[[ 0  0  0  0]
 [ 5  0  0  0]
 [50  1  0  0]]
PAGE: 3
[[ 0  0  0  0]
 [ 5  0  0  0]
 [50  1  0  0]]
Total PerMaViss time:12.526841640472412
```

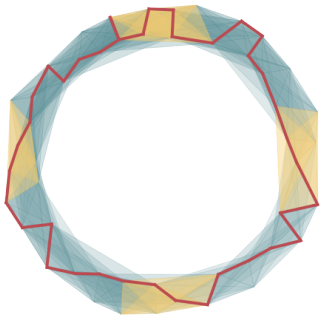
We can inspect the obtained barcodes on the 1st dimension

```
>>> MV_ss.persistent_homology[1].barcode
array([[0.23081607, 0.29335407],
       [0.25273638, 0.29106543],
       [0.27194715, 0.28877474],
       [0.28346206, 0.35987622],
       [0.35728549, 0.39131679],
       [0.39122743, 0.99      ]])
```

an we can plot this barcode, indicating in red the bar comming from the $(1,0)$ position on the spectral sequence:



Notice that in this case, there was no need to solve the extension problem. Also, we can plot the representative relating to the red bar



where the yellow regions indicate overlaps between consecutive covering subcomplexes. These plots were created reading the information stored in the spectral sequence object `MV_ss` and then using Matplotlib.

9.2 Barcode Basis object

We created an object `barcode_basis` that implements some information and methods relevant when working with barcode bases as shown in section 4.1. A `barcode_basis` object `A` is created using the following information

- A matrix of dimensions $(\texttt{A.dim}, 2)$ storing the birth and death values for all bars of `A`,
- (Optional) a reference to another `barcode_basis` `B`, together with coordinates of generators of `A` in terms of `B`.
- (Optional) Whether we want to store the indices of which bars are well defined. Here, an interval $[a,b)$ from the barcode decomposition is "well-defined" whenever $a < b$.
- (Optional) Whether the `barcode_basis` is broken[1] and the corresponding associated broken base coefficients.

In this section we will review some of the methods that are used by the barcode basis object. For a full documentation reference, see [116]. When a `barcode_basis` object is created, the first thing to check is that the given barcode is well-defined, that is, whether all birth values are strictly less than all death values for each given interval. Following is an example where we create a barcode basis object with four bars:

```
>>> import numpy as np
>>> from permaviss.persistence_algebra.barcode_bases import barcode_basis
```

[1] This was a former concept that was used for modelling the extension problem. Ultimately, these broken barcodes where a way to deal with quotients of persistence morphisms.

```
>>> bars = np.array([[0, 1], [-1, 2], [1, 3], [2, 2]])
>>> base1 = barcode_basis(bars, store_well_defined=True)
>>> base1.dim
3
>>> base1.well_defined
array([ True,  True,  True, False], dtype=bool)
>>> print(base1)
Barcode basis
[[ 0  1]
 [-1  2]
 [ 1  3]]
```

By default, the undefined bars that we have sent are not stored on the barcode basis object; thus, since one of the given bars is trivial, the dimension of the resulting `barcode_basis` object is 3 instead of 4. We can sort the bars according to the standard barcode order by using the `sort()` method:

```
>>> base1.sort()
>>> print(base1)
Barcode basis
[[-1  2]
 [ 0  1]
 [ 1  3]]
```

Next, we will create a barcode basis which has coordinates in terms of the previously created `base1`. These coordinates correspond to the columns of the matrix we need to send. Following is an example:

```
>>> bars = np.array([[0, 2], [2, 3]])
>>> coordinates = np.array([[1, 0], [1, 0], [0, 2]])
>>> base2 = barcode_basis(bars, base1, coordinates)
>>> base2.dim
2
>>> print(base2)
Barcode basis
[[0 2]
 [2 3]]
[[1 0]
```

```
[1 0]
[0 2]]
```

To better understand `base2` we will use some of the notation from chapter 4. First, `base1` stores the barcode basis $\mathscr{A} = \{\alpha_0, \alpha_1, \alpha_2\}$ with $\alpha_0 \sim [-1,2)$, $\alpha_1 \sim [0,1)$ and $\alpha_2 \sim [1,3)$, whereas `base2` stores the barcode basis $\mathscr{B} = \{\beta_0, \beta_1\}$ with $\beta_0 \sim [0,2)$ and $\beta_1 \sim [2,3)$. The relations between both barcode bases are given by the equations:

$$\beta_0 = \alpha_0 \boxplus \alpha_1 \qquad\qquad \beta_1 = \mathbf{1}_2(\alpha_2)\ .$$

In particular, notice that the persistence module generated by $\mathscr{B}$ is a submodule of that generated by $\mathscr{A}$.

Next, we give a list including the most relevant methods related to a `barcode_basis` object:

- `sort(self, precision=7, send_order=False)):`

  This method sorts a `barcode_basis` object acording to the standard order of barcodes. A precision parameter is required to round the endpoints of bars. Also, one can set `send_order` to `True` if one needs to recover how the original barcodes have been ordered. Following is an example with 4 bars:

```
>>> bars = np.array([[1,2],[0,4],[-1,3],[1,4]])
>>> base4 = barcode_basis(bars)
>>> base4.sort(send_order=True)
array([2, 1, 3, 0])
>>> base4 = barcode_basis(bars)
>>> print(base4)
Barcode basis
[[ 1  2]
 [ 0  4]
 [-1  3]
 [ 1  4]]
>>> base4.sort(send_order=True)
array([2, 1, 3, 0])
>>> print(base4)
Barcode basis
[[-1  3]
 [ 0  4]
```

```
 [ 1  4]
 [ 1  2]]
```

- `active(self, rad)`:

  Given a persistence value `rad`, it returns the indices of the generators that are alive.

```
>>> base4.active(0.5)
array([0, 1])
```

- Similar to the `active` method, there are methods `birth` and `death` for getting the indices of the generators being born or dying at a persistence value `rad`.

```
>>> base4.birth(1)
array([2, 3])
>>> base4.death(4)
array([1, 2])
```

9.3 Image Kernel

Next, we present a brief example computing the image and kernel barcode basis from example 4.2.2 from section 4.2. To start, we do the necessary imports

```
>>> import numpy as np
>>> from permaviss.persistence_algebra.barcode_bases import barcode_basis
>>> from permaviss.persistence_algebra.image_kernel import image_kernel
```

Now, we can execute example 4.2.2, obtaining the expected result:

```
>>> p = 5 # computations to be performed mod 5
>>> A = barcode_basis([[1, 5], [1, 4], [2, 5]])
>>> B = barcode_basis([[0, 5], [0, 3], [1, 4]])
>>> F = np.array([[0, 0, 1], [1, 0, 0], [1, 1, 1]])
>>> Im, Ker, PreIm = image_kernel(A, B, F, p)
>>> print(Im)
Barcode basis
[[1. 4.]
 [1. 3.]
 [2. 5.]]
[[0 0 1]
 [4 4 0]
```

```
 [4 0 0]]
>>> print(Ker)
Barcode basis
[[3. 5.]]
[[4.]
 [1.]
 [0.]]
>>> print(PreIm)
[[4. 4. 0.]
 [0. 1. 4.]
 [0. 0. 1.]]
```

Notice that the coordinate matrix on the barcode bases `Im` and `Ker` coincide with the respective generators from the sets $\mathscr{I}$ and $\mathscr{K}$ from example 4.2.2. The preimages matrix `PreIm` contains the combinations of generators from $\mathscr{A}$ whose images generate barcode bases for $\mathrm{Im}(f)$. This is very useful when computing the spectral sequence.

9.4 Quotients and Persistence Module Homology

One can also adapt `image_kernel` to implement the computation of quotients, as briefly explained in section 4.4. Essentially, suppose that we have a pair of barcode bases and we want to compute the quotient of one by the other. For instance, consider again example 4.2.2, and assume that we want to compute the quotient $\mathtt{B}/\mathrm{Im}(\mathtt{F})$. Thus, we proceed as in section 4.4 by first defining the barcode bases and the quotient matrix:

```
>>> A = barcode_basis([[1, 5], [1, 4], [2, 5]])
>>> B = barcode_basis([[0, 5], [0, 3], [1, 4]])
>>> A_B_barcode = np.concatenate((A.barcode, B.barcode, axis=0))
>>> A_B = barcode_basis(A_B_barcode)
>>> matrix_A_B = np.concatenate((F, np.identity(3)), axis=1)
>>> print(matrix_A_B)
[[0. 0. 1. 1. 0. 0.]
 [1. 0. 0. 0. 1. 0.]
 [1. 1. 1. 0. 0. 1.]]
```

Then we execute `image_kernel`with the aim of obtaining the quotient:

```
>>> Q, _ = image_kernel(
...   A_B, B, matrix_A_B, 5, start_index=3, prev_basis=B)
```

```
>>> print(Q)
Barcode basis
[[0. 2.]
 [0. 1.]]
[[1. 0.]
 [0. 1.]
 [0. 0.]]
```

What the result means, using notation from example 4.2.2, is that the quotient module $\mathtt{B}/\mathrm{Im}(\mathtt{F})$ is generated by the classes $[\beta_1]$ and $[\beta_2]$ with associated barcodes $[0,2)$ and $[0,1)$ respectively. Figure 4.1 might help in visualizing this quotient.

We want to be able to compute quotients so that we are able to compute persistence module homology, as explained on section 4.5. In the following example we will do this for a short exact sequence of persistence modules

$$M \xrightarrow{f} N \xrightarrow{g} P\,.$$

First, we define the respective barcode bases and morphisms:

```
>>> A_bas = barcode_basis([[1, 4], [2, 4], [3, 5]])
>>> B_bas = barcode_basis([[0, 4], [0, 2], [1, 4], [1, 3]])
>>> C_bas = barcode_basis([[0, 3], [0, 2], [0, 2]])
>>> Base = [C_bas, B_bas, A_bas]
>>> f = np.array([
...   [0, 1, 1],
...   [1, 0, 0],
...   [0, 4, 1],
...   [0, 0, 0]])
>>> g = np.array([
...   [0, 0, 0, 1],
...   [4, 0, 1, 3],
...   [1, 0, 1, 4]])
>>> D = [None, f,g]
```

where the entries in `Base` and `D` are included in reverse order to the arrows of the short exact sequence. Then, we use the `module_persistence_homology` method to obtain barcode bases for the homology groups, and we print the results.

```
>>> from permaviss.persistence_algebra.module_persistence_homology import (
... module_persistence_homology)
>>> p = 5 # computations modulo 5
>>> Hom, Im, PreIm = module_persistence_homology(D, Base, p)
>>> print(Hom[2])
Barcode basis
[[2. 4.]
 [4. 5.]]
[[1. 0.]
 [0. 1.]
 [0. 1.]]

>>> print(Hom[1])
Barcode basis
[[0. 1.]
 [2. 3.]]
[[0. 1.]
 [1. 0.]
 [0. 0.]
 [0. 0.]]

>>> print(Hom[0])
Barcode basis
[[0. 1.]
 [0. 1.]]
[[1. 0.]
 [0. 1.]
 [0. 0.]]
```

9.5 Computing PERMAVISS

We will now present the Mayer-Vietoris spectral sequence setup, with the particular case of covers taken on Vietoris-Rips complexes in mind. Let $\mathbb{X} \subseteq \mathbb{R}^n$ be a point cloud with n small. We first divide $\mathbb{X}$ into cubical regions that depend on two parameters: `overlap` and `max_div`. Essentially, this is done in three steps:

- We start by taking the hypercube C that contains all data points from $\mathbb{X}$.
- Next, we divide C into smaller hypercubes, so that all of them have equal length along all their dimensions and the number of stacked hypercubes along any dimension is less than or equal to `max_div`.
- Afterwards, the dividing hypercubes are 'thickened' so that the overlap between adjacent cubes is greater or equal to `overlap`.

Of course, one could come up with more clever ways to divide $\mathbb{X}$. However, since the focus here is on developing an algorithm for computing persistent spectral sequences, these simple covers are good enough. Once we have chosen a cover $\mathscr{U} = \{U_i\}_{i\in I}$, it is straightforward to obtain the nerve $N_{\mathscr{U}}$.

As in section 9.1, we use the function `create_MV_ss` to introduce a `spectral_sequence` object with all relevant information filled in. This contains the cover taken for $\mathbb{X}$, subcomplexes of $\mathrm{VR}_*(\mathbb{X})$ arising from this cover as well as information related to all pages on the spectral sequence $E^*_{*,*}(\mathbb{X},\mathscr{U})$. An important aspect of the PERMAVISS algorithm is to treat differently class representatives contained on the total complex $\mathscr{S}^{\mathrm{Tot}}_*(\mathbb{X},\mathscr{U})$ from those representatives contained on the first or higher pages; the reason being that the former is much larger than the latter. Thus, the terms on the total complex are stored using a class which we call `local_chains`, which consists of a compact way for storing sparse vectors where the sparsity depends on the location of each chain with respect $\mathscr{U}$. On the other hand, we use `barcode_basis` objects for storing bases for the successive page terms, as well as bases for the images of the differentials.

Once `create_MV_ss` is called, a cover is taken for $\mathbb{X}$. Then, the first computational step consists in computing the 0-th and 1-st pages in a parallel way. Essentially, for each simplex $\sigma \in N_{\mathscr{U}}$ we have a local point cloud X_σ, and we compute $\mathrm{PH}_*(X_\sigma)$ by using the persistent homology method included in PERMAVISS. The results are stored using the `local_chains` class whose information is 'localized', in the sense that representatives of first page classes are given as chains in some open cover $\mathscr{U}_\sigma$. This is not the case for higher pages, where information is mixed up from several covers. Even though there are various implementations of persistent homology computations, as explained in section 3.4, we opted to implement our own since the spectral sequence algorithm requires some information that the usual algorithms do not return. This has the unfortunate effect that our resulting algorithm is not very efficient compared to standard algorithms. However, this opens the door to further improvements by implementing some of the speedups from the other approaches.

Let us now describe how we compute the second page from the first page terms. The first step consists in computing the matrices associated to the horizontal differentials, which are induced by

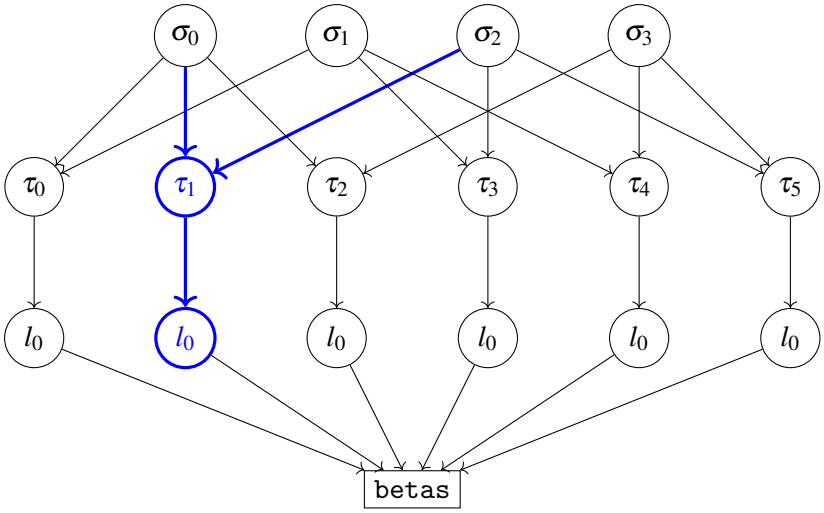


Figure 9.1: For each $\tau_i \in N_{\mathscr{U}}^{\texttt{n\_deg}-1}$, in parallel, we read information from `1_representatives` over the simplices in $N_{\mathscr{U}}^{\texttt{n\_deg}}$ that are cofaces of τ_i, followed by a lift to first page l_i. We highlight in blue the parallel computations over τ_1.

the Čech differentials. Let us describe this in detail for a pair of integers $\texttt{n\_deg} > 0$ and $\texttt{dim} \geq 0$ indicating respectively the horizontal and vertical position of a term on the first page of the spectral sequence. From the previous step, we have first page representatives `1_representatives` in `local_chains` format for the position $(\texttt{n\_deg}, \texttt{dim})$. Next, we compute the Čech differential applied to `1_representatives`, followed by a lift to first page. We do this in a parallel way over the simplices of $N_{\mathscr{U}}^{\texttt{n\_deg}-1}$. At the end, we can join all the information to obtain the matrix `betas` associated to the Čech differential induced on the first page. In the PERMAVISS code, within the `spectral_sequence` class, this is done using a function named `cech_diff_and_lift` that calls in parallel `cech_diff_and_lift_local`; see figure 9.1.

Using the procedure mentioned in the previous paragraph, we compute the Čech differentials of a given dimension `dim` by going over all indices $\texttt{n\_deg} > 0$. This gives us the information of a chain $E^1_{*,\texttt{dim}}$ corresponding to the `dim` row from the spectral sequence. Once that is done, we call `module_persistent_homology` to compute the persistent homology of $E^1_{*,\texttt{dim}}$ and obtain barcode bases generating the second page terms. We end the second page computation by storing 'nice' total complex representatives for classes. First we compute the representatives in $(\texttt{n\_deg}, \texttt{dim})$ by using the combinations from $E^2_{\texttt{n\_deg},\texttt{dim}}$, which are stored in the `local_chains` format, let us name this $\alpha_{\texttt{n\_deg}}$. Next, we compute the Čech differential and lift by using `cech_diff_and_lift`. Notice that in this case the lift to the first page should be zero and the interesting data is given by the preimages of the vertical differential, which we store on $\alpha_{\texttt{n\_deg}-1}$. Thus one has that

$$d_{\texttt{dim}+1}(\alpha_{\texttt{n\_deg}-1}) + \check{\delta}_{\texttt{n\_deg}}(\alpha_{\texttt{n\_deg}}) = 0\,,$$

and the total complex chain is given by $(\alpha_{\texttt{n\_deg}-1}, \alpha_{\texttt{n\_deg}})$.

Computation of higher page representatives is slightly more complex, although it follows exactly the steps outlined on section 6.3. These steps use `cech_diff_and_lift` together with the improved `image_kernel`. The same follows for the extension problem.

Remark. We would like to notice that the computation of persistence module homology of the bottom row from the first page $(E^1_{*,0}, \check{\delta})$ becomes particularly expensive. However, in this case the barcode bases of all terms have trivial birth values. This is why we think it is possible to speed up computations by calling a special function in this case. However, as this is the part of speeding up the code, we leave it for future research.

9.6 Examples

In this section we present some examples of the use of PERMAVISS.

Example 9.6.1. Consider 600 points from a ring torus of major radius 3 and minor radius 1, which have been sub-sampled using the min max algorithm out of a randomly sample of 2300 points:

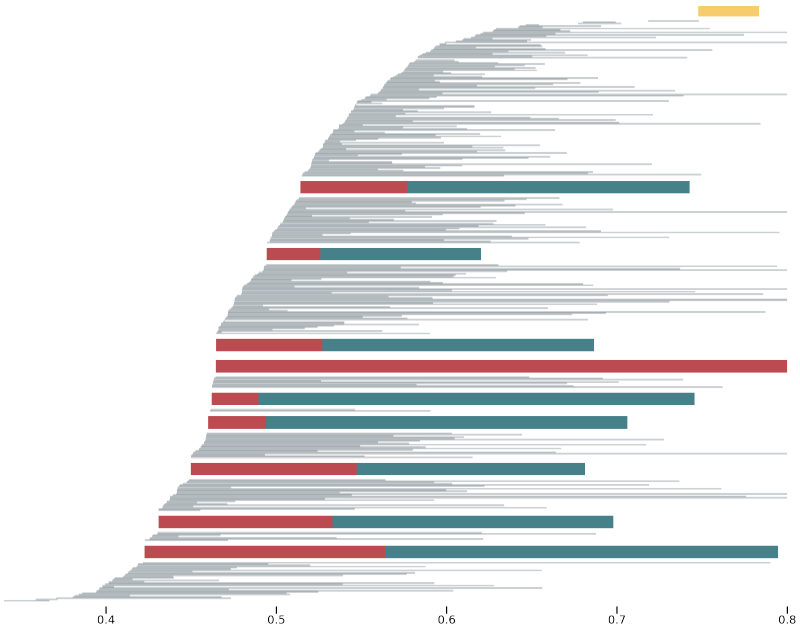
```
>>> X = torus3D(2500, min_rad=1, max_rad=3)
>>> X = take_sample(X,600)
```

Suppose that we want to compute persistent homology of the Vietoris Rips complex up to radius 0.8.

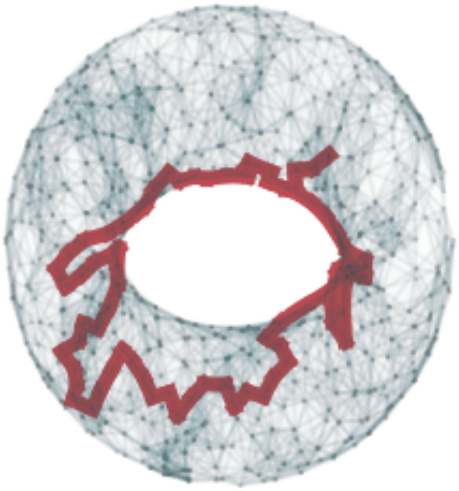
```
>>> max_r = 0.8
>>> max_dim = 4
>>> p = 5
>>> max_div = 2
>>> overlap = max_r * 1.01
>>> MV_ss = create_MV_ss(X, max_r, max_dim, max_div, overlap, p)
PAGE: 1
[[  0   0   0   0   0   0   0   0]
 [  1   0   0   0   0   0   0   0]
 [338  22   0   0   0   0   0   0]
 [705 105   0   0   0   0   0   0]]
PAGE: 2
[[  0   0   0   0   0   0   0   0]
 [  1   0   0   0   0   0   0   0]
 [316   0   0   0   0   0   0   0]
```

```
  [600   9   0   0   0   0   0   0]]
Total PerMaViss time:178.7223949432373
```

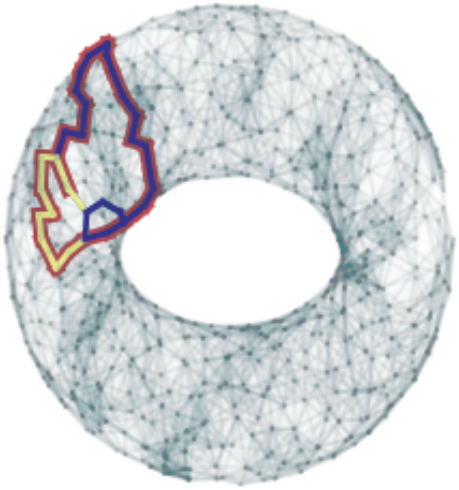
Under the same conditions, it took our implementation of persistent homology without spectral sequences about 737.7 seconds; thus, by using the spectral sequence we have managed to speedup substantially the persistence algorithm. However, there is still work to do to make PERMAVISS efficient; for example, Ripser.py barely blinks at this example. Let us now have a look at the generated barcode on dimension 1. We plotted it using matplotlib, together with the information contained in `MV_ss`:



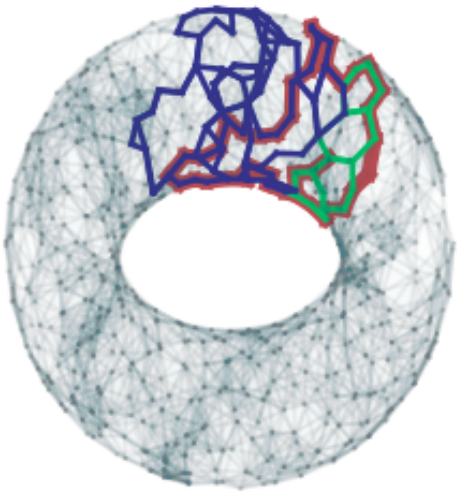
In this barcode we show the one-dimensional bars plus the two-dimensional interval in yellow at the top. The splitted bars contain a red portion coming from the $(1,0)$ position of the spectral sequence and a blue portion from the $(0,1)$ position. Next, we might plot some cycle representatives as, for example, the long red bar which indicates that a cycle is going across various covers along all considered persistence values:



Examining some of the extended bars, we might use the extension coefficients, together with total complex representatives. The following diagram shows a representative for a cycle in the $(1,0)$ position (red) which "breaks down" into smaller cycles that extend it which are contained on the $(0,1)$ position. We indicate these extending cycles with different tones depending on the cover these are contained in:



Although in the previous visualization the relation between cycle representatives seems quite straightforward, sometimes this is not the case:



Example 9.6.2. In this example we will see that higher nontrivial differentials come up in fairly general situations. Consider a uniform sample of 3000 points from a unit cube, and assume that we take a subsample of 300 points through the min-max algorithm

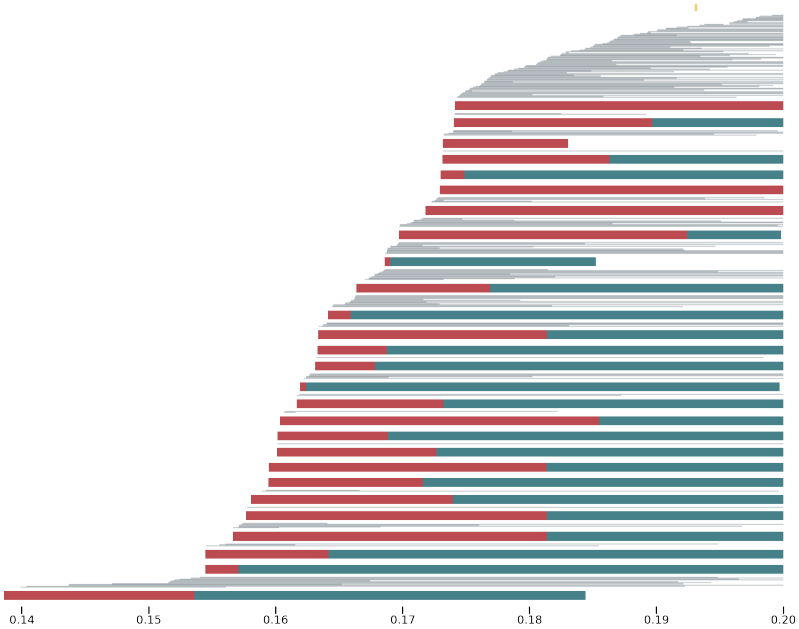
```
>>> X = random_cube(3000,3)
>>> X = take_sample(X,300)
```

Then, we compute PERMAVISS setting appropriate parameters

```
>>> max_r = 0.2
>>> max_dim = 4
>>> p = 5
>>> max_div = 2
>>> overlap = max_r*1.01
>>> MV_ss = create_MV_ss(X, max_r, max_dim,  max_div, overlap, p)
PAGE: 1
[[  0   0   0   0   0   0   0   0]
 [  1   0   0   0   0   0   0   0]
 [257  38   0   0   0   0   0   0]
 [503 350 280 238 168  84  24   3]]
```

```
PAGE: 2
[[  0   0   0   0   0   0   0   0]
 [  1   0   0   0   0   0   0   0]
 [221   0   0   0   0   0   0   0]
 [300  27   2   0   0   0   0   0]]
PAGE: 3
[[  0   0   0   0   0   0   0   0]
 [  1   0   0   0   0   0   0   0]
 [219   0   0   0   0   0   0   0]
 [300  27   0   0   0   0   0   0]]
Total PerMaViss time:118.49028873443604
```

In this case, notice that there is a nontrivial differential on the second page starting on position $(2,0)$ and ending in $(0,1)$. We proceed to plot the first dimensional and second dimensional barcodes as done on example 9.6.1, obtaining the following barcode:



In this case there are 27 extended bars in dimension 1. We can plot the representatives coming from the $(1,0)$ position in red, while plotting the corresponding extended representatives in position $(0,1)$ in different colors depending on the cover region these come from, see figure 9.2. In particular, notice that in general the extended cycle representatives differ substantially from the cycle representatives from position $(1,0)$. Although here it is important to notice that we are looking at cycle representatives from a particular projection angle.

The last example will differ from the previous two in that we will take a custom cover.

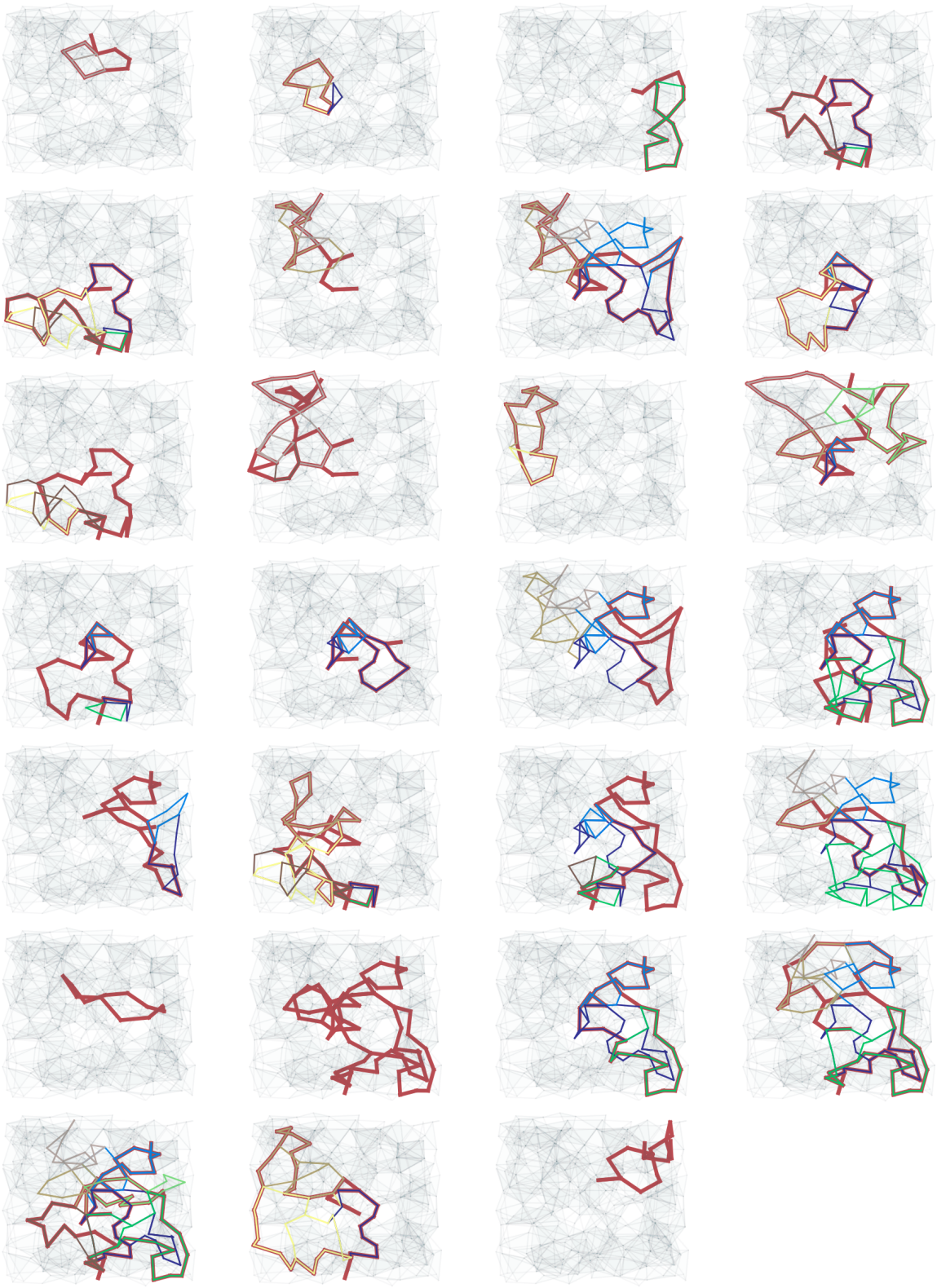


Figure 9.2: Depiction of cycle representatives on the $(1,0)$ spectral sequence position, together with the representatives of their respective extension.

Example 9.6.3. We consider a sample of 500 points from a sphere with two circles attached. The sphere is centered at the origin and has radius 1, while the two circles have radius 2 and touch the sphere at a single point.

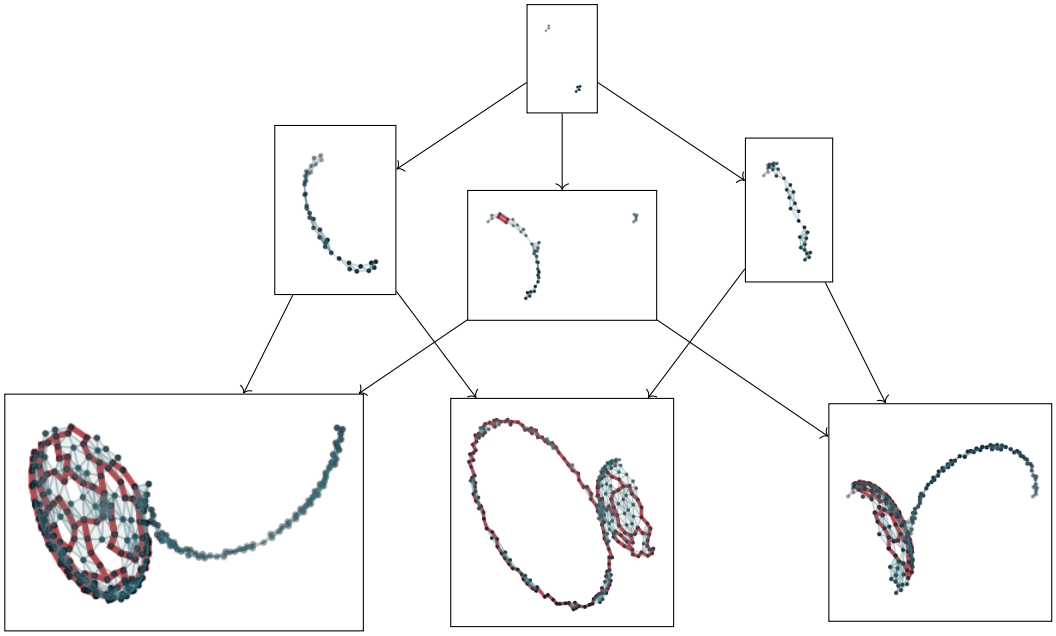
```
>>> circle1 = take_sample(random_circle(500,2,0.08,[3.3,0]),100)
>>> circle1 = np.c_[circle1, 0.3*circle1[:,0]]
>>> circle2 = take_sample(random_circle(500,2,0.07,[-3.3,0]),100)
>>> circle2 = np.c_[circle2[:,0], np.zeros(len(circle2)), circle2[:,1]]
>>> sphere = take_sample(random_sphere(2400, 1, 3), 300)
>>> X = np.r_[circle1, circle2, sphere]
```

For this example, instead of using cubical covers, we specified our own cover into three regions with mutual overlaps whose diameters are bounded by 0.4, and which are determined as:

- $A = \{(x,y,z) \in X \mid z > -0.2 \text{ and } x > -0.2 \}$
- $B = \{(x,y,z) \in X \mid z > -0.2 \text{ and } x < 0.2 \}$
- $C = \{(x,y,z) \in X \mid z < 0.2 \}$.

Now we plot X by taking different colors for the different covering regions:

We aim to compute the persistent homology associated to the Vietoris-Rips filtration by paying attention to the cover $\{A,B,C\}$. We can depict the mutual intersections by the diagram:



Notice that from the triple intersection on the top, one might suspect that a nontrivial barcode will appear on position $(2,0)$ from the spectral sequence; as there are two connected components. This is indeed the case, as we check when we compute the spectral sequence. As here we used a custom cover, the computations where done by modifying some of the code from `create_MV_ss`. Here we write the parameters and the output obtained:

```
>>> max_r = 0.
>>> max_dim = 4
>>> p = 5
...
PAGE: 1
[[  0   0   0   0]
 [  2   0   0   0]
 [195  15   0   0]
 [596 103   7   0]]
PAGE: 2
[[  0   0   0   0]
 [  2   0   0   0]
 [180   0   0   0]
 [500  10   1   0]]
 PAGE: 3
 [[  0   0   0   0]
  [  2   0   0   0]
```

```
[179   0   0   0]
[500  10   1   0]]
```

In addition, notice that there is a nontrivial second page differential.

Chapter 10

Conclusions and Future Work

In this thesis we have explored connections between persistent homology and spectral sequences. Following is an outline of the main contributions and possible directions for future research:

- We surveyed briefly the persistence filtration spectral sequence that is associated to a filtered cellular complex on section 3.5 and in doing so we gave some links to the twisted double complex from section 2.12. These lead to a straightforward way for proving Theorem 3.5.1 and also were useful in drawing further connections between the spectral sequence algorithm and the persistence spectral sequence. It would be worth exploring this topic further.

- We have developed the concept of a barcode basis together with the $\boxplus$-operation. This lead to a characterization of these bases in terms of generation and linear independence as outlined on proposition 4.1.10. Eventually we obtained an `image_kernel` algorithm which allows to compute images and kernels of persistence morphisms. Thanks to using barcode bases many proofs where done in a compact way, such as those from chapter 5. Also, we showed in section 4.4 that one might use the `box_gauss_reduce` algorithm 4.1 for computing quotients of persistence modules. All these tools are central for the persistence Mayer-Vietoris spectral sequence computations. The early implementations of PERMAVISS use the `image_kernel` version described on appendix A. We hope that the current work developed in chapter 4 will lead to further improvements for PERMAVISS.

- In chapter 5 we have payed close attention to the problem of detecting changes in persistent homology from local information. The barcode shift lemma 5.1.2 captures this idea very precisely while Proposition 5.2.7 gives a direct way to measure these variations. The results in sections 5.3 and 5.4 aim at tracking these changes precisely, leading to exact conditions on the barcode decompositions of the involved short exact sequences. It would be worth seeing how this works in practical situations and how these results could be adapted to be used on natural datasets.

- The Persistence Mayer-Vietoris spectral sequence is undoubtedly the main object of interest in this work. Here we have presented how to compute persistent homology from this spectral sequence on chapter 6, giving algorithmic descriptions which were later implemented on the PERMAVISS module. In chapter 9 we have given an outline of how to parallelize computations, presenting some illustrative examples of various principles such as higher differentials, extension problems, etc. This approach, however has limitations which are overcome by considering geometric realization constructions, such as the $(K, \mathscr{P})$-join diagram from chapter 7. Implementing spectral sequences for join diagrams, while speeding up the overall algorithm is a matter for future research.

- In chapter 8 we explored the Persistence Mayer-Vietoris spectral sequence as an invariant. We gave stability results in terms of small changes on the diagram elements while also in terms of different covers taken. These stability results are summarized by Theorem 8.3.5 and the results from section 8.5. Notice that this spectral sequence has two parameters involved: the filtration of the covered complex and the simplicial complex of the corresponding diagram. This is why one might think that there are some relations with 2-parameter persistent homology. It would be worth exploring further the value of the extra information contained on this spectral sequence.

Appendix A

(Original) `image_kernel` Algorithm

Here, we present an algorithm which obtains barcode bases for the image and kernel of a persistence morphism. This was already covered by the `image_kernel` procedure explained in section 4.3, however, here we present the original procedure which appeared in PERMAVISS v.0.0.1. The core idea is to work based on proposition 4.1.2. Notice that this algorithm can be adapted to obtain barcode bases for quotients of persistence modules. Also, we include a descripition of an improved version of the `image_kernel` procedure which was implemented in PERMAVISS version 0.2.

Suppose that $f : \mathbb{V} \to \mathbb{W}$ is a morphism between two tame persistence modules. Let $\mathscr{A}$ and $\mathscr{B}$ be barcode bases for $\mathbb{V}$ and $\mathbb{W}$ respectively. Suppose also that we know $f(\mathscr{A})_{\mathscr{B}}$, the matrix associated to f with respect to barcode bases $\mathscr{A}$ and $\mathscr{B}$. We want to find a barcode basis for the image $\mathscr{I}$, and a barcode basis for the kernel $\mathscr{K}$. In order to achieve this, $\mathscr{I}$ will start being set to be equal to the $|\mathscr{B}| \times |\mathscr{A}|$ matrix $f(\mathscr{A})_{\mathscr{B}}$. Performing left to right column additions will lead to the nonzero columns of $\mathscr{I}$ forming a basis for the image. On the other hand, $\mathscr{K}$ will be a matrix with $|\mathscr{A}| + 1$ rows and whose number of columns will 'grow' as the computations develop. The extra row will be used for storing the parameter of the multiplying step function. Notice that $\mathscr{K}$ will have at most $|\mathscr{A}|$ columns, which is useful to know if we wanted to preallocate space for speed.

Notice that there exist values $-\infty = a_0 < a_1 < \cdots < a_{n+1} = \infty$ such that f is constant along $[a_i, a_{i+1})$ for each $0 \leq i \leq n$. We start by computing the values a_i for all $0 \leq i \leq n$. We will denote by $\mathscr{A}^{a_i}(j)$ the index $1 \leq \mathscr{A}^{a_i}(j) \leq |\mathscr{A}|$ of the j-element from $\mathscr{A}^{a_i}$. Also given a matrix A, we will denote by $A[j]$ the j^{th} column of A. The matrices R^i will denote the successive Gaussian reductions as we increase the parameter $0 \leq i \leq n+1$. That is, we start with $\widetilde{R}^0$ which will be the $|\mathscr{B}^{a_0}| \times |\mathscr{A}^{a_0}|$-matrix of f along the interval $a_0 < a_1$, then we reduce it to R^0. Simultaneously, we perform exactly the same transformations to $\mathscr{I}$. In order to track these additions performed, we will use a $|\mathscr{A}| \times |\mathscr{A}|$ matrix T. This T will be the identity matrix $\mathrm{Id}_{|\mathscr{A}|}$. Thus, whenever we

add columns in R^0 we perform the same additions in T. On the other hand, if some column $R^0[j]$ becomes zero, where $1 \leq j \leq |\mathscr{A}^{a_0}|$, we add $T[\mathscr{A}^{a_0}(j)]$ at the right end of the matrix of kernels K^0. Additionally, we append $T[\mathscr{A}^{a_0}(j)]$ to $\mathscr{K}$, with associated step function coefficient a_0. Since we require $\mathscr{K}$ to be linearly independent, we will introduce a set `pivots` for tracking the pivots of the elements in $\mathscr{K}$. For each $T[\mathscr{A}^{a_0}(j)]$ that we add into $\mathscr{K}$, we add $\mathscr{A}^{a_0}(j)$ into `pivots`. Note that in this first step there will be no repeated elements in `pivots` and the matrix $\mathscr{K}$ will be already reduced. Once we finish, we jump to the next parameter a_1.

Let us go through the procedure for a_1. For this, we add or take out rows and columns from R^0 and K^0 according to the life of each generator in $\mathscr{A}$ and $\mathscr{B}$; these changes are stored into $\widetilde{R}^1$ and $\widetilde{K}^1$, respectively. Observe that $\widetilde{K}^1$ might not be reduced. Since we would like to obtain a basis for the kernel of f, we reduce it further to $K^1 = R(\widetilde{K}^1)$, performing the same additions on $\mathscr{K}$. Next we proceed to reduce $\widetilde{R}^1$. There is a trick we can use here to speed up the computations. For each j-column in K^1, if the pivot p of the column is such that $\mathscr{A}^{a_1}(p)$ is not in `pivots`, this means that the p column in $\widetilde{R}^1$ will become zero after reducing. Then we set $\widetilde{R}^1[p]$ to zero directly, substitute the column $\mathscr{I}[\mathscr{A}^{a_1}(p)]$ by $f(K^1[j])$, and add $\mathscr{A}^{a_1}(p)$ into `pivots`. Here by $f(K^1[j])$ we mean the result after adding the columns from $f(\mathscr{A})_{\mathscr{B}}$ with coefficients given by $K^1[j]$. Notice that this is the same as performing left to right column additions to the column $\mathscr{I}[\mathscr{A}^{a_1}(p)]$, although we also permit this column be multiplied by a non-zero coefficient $t \in \mathbb{F} \setminus \{0\}$. After performing these preprocessing tasks, we reduce $\widetilde{R}^1$ into R^1, repeating the same transformations to T and $\mathscr{I}$. Then we examine R^1, and look for columns $1 \leq j \leq |\mathscr{A}^{a_1}|$ of R^1, such that $R^1[j] = 0$ and also $\mathscr{A}^{a_1}(j)$ is not in `pivots`. For each such column j, we append $T[\mathscr{A}^{a_1}(j)]$ at the right end of K^1, and also into $\mathscr{K}$ with birth value a_1. Finally, we add $\mathscr{A}^{a_1}(j)$ into `pivots`. This finishes the iteration for a_1.

We repeat the previous step again for parameters $a_1 < a_2 < \cdots < a_n$. On the i iteration, where $2 \leq i \leq n$, we assume that we have well defined matrices R^{i-1} and K^{i-1}. As before, we update these matrices into a $\mathscr{B}^{a_i} \times \mathscr{A}^{a_i}$-matrix $\widetilde{R}^i$, and a matrix with $|\mathscr{A}^{a_i}|$ columns $\widetilde{K}^i$. These updates are performed by adding and deleting columns as the barcodes from $\mathscr{A}$ and $\mathscr{B}$ are born or die respectively. The rest of the procedure for a_i is exactly as we outlined for a_1 earlier. Notice that while we are on the ith step, both K^i and $\mathscr{K}$ will have the same number of columns. An outline of this procedure is shown in Algorithm A.1.

Proposition A.0.1. *Algorithm A.1 computes $\mathscr{K}$ and $\mathscr{I}$ bases for the kernel and image of f. Furthermore, it takes at most $\mathscr{O}(nM|\mathscr{A}|^2)$ time, where $M = \max(|\mathscr{A}|, |\mathscr{B}|)$.*

Proof. The key observation is that $\mathscr{K}$ forms a barcode basis for $\mathrm{Ker}(f)$ if and only if $\mathscr{K}^r$ is a basis for $\mathrm{Ker}(f)^r$ for all $r \in \mathbb{R}$. Now, notice that $\mathscr{K}^r$ generates $\mathrm{Ker}(f)^r$ since all kernel elements were sent to $\mathscr{K}$. On the other hand, each $\mathscr{K}^r$ is a linearly independent set, since we have performed

Algorithm A.1 `image_kernel`

Input: $\mathscr{A}$, $\mathscr{B}$, $f(\mathscr{A})_{\mathscr{B}}$
Output: $\mathscr{K}$, $\mathscr{I}$

Find values $a_0 < a_1 < \cdots < a_n$ where a barcode generator dies, or is born in $\mathscr{A}$ or $\mathscr{B}$
Set $\mathscr{I} = f(\mathscr{A})_{\mathscr{B}}$, $\mathscr{K} = \emptyset$, $T = \mathrm{Id}_{|\mathscr{A}|}$, $R^{-1} = \emptyset$, $K^{-1} = \emptyset$, `pivots` $= \emptyset$
for $0 \leq i \leq n$ **do**
    Update $\widetilde{R}^i$ and $\widetilde{K}^i$ from R^{i-1}, and K^{i-1} respectively
    Reduce $\widetilde{K}^i$ obtaining K^i. Perform the same reductions to $\mathscr{K}$
    for each j-column of K^i with pivot p such that $\mathscr{A}^{a_i}(p) \notin$ `pivots` **do**
        $\mathscr{I}[\mathscr{A}^{a_i}(p)] \leftarrow f(K^i[j])$
        $\widetilde{R}^i[p] \leftarrow 0$
        Add $\mathscr{A}^{a_i}(p)$ into `pivots`
    end for
    Reduce $\widetilde{R}^i$ into R^i. Perform the same reductions to T and $\mathscr{I}$
    for $1 \leq j \leq |\mathscr{A}^{a_i}|$ **do**
        if $R^i[j] = 0$ & $\mathscr{A}^{a_i}(j) \notin$ `pivots` **then**
            Append $T[\mathscr{A}^{a_i}(j)]$ at end of K^i, and also at $\mathscr{K}$ with step coefficient a_i
            Add $\mathscr{A}^{a_i}(j)$ into `pivots`
        end if
    end for
end for
return $\mathscr{K}$ and $\mathscr{I}$ (optionally return T for preimages)

Gaussian eliminations that ensured this. Similarly, for any $r \in \mathbb{R}$ we have that $\mathscr{I}^r$ generates all the columns from $f(\mathscr{A})^r_{\mathscr{B}}$, and thus it generates $\mathrm{Im}(f)^r$. We have also ensured linear independence of $\mathscr{I}^r$ by the Gaussian elimination process. Thus, $\mathscr{I}$ is a barcode basis for $\mathrm{Im}(f)$.

Let us compute the complexity of the algorithm. We start noticing that n comes from the outer loop. Then the Gaussian reduction of $\widetilde{K}^i$ might take at most $\mathscr{O}(|\mathscr{A}|^3)$ time. On the other hand the reduction of $\widetilde{R}^i$ might take $\mathscr{O}(|\mathscr{B}||\mathscr{A}|^2)$ time. The first inner loop will take less than $\mathscr{O}(|\mathscr{A}|(\log(|\mathscr{A}|) + |\mathscr{A}||\mathscr{B}|))$ time, where the multiplying $|\mathscr{A}|$ comes from the iteration. Within round brackets, the first term comes from checking `pivots` by a hash table or similar, whereas the second comes from computing $f(K^i[j])$. The second inner loop takes $\mathscr{O}(|\mathscr{A}|\log(|\mathscr{A}|))$ time, where $|\mathscr{A}|$ is for the iteration and $\log(|\mathscr{A}|)$ for checking `pivots`. Putting all together we obtain the following complexity:

$$n\Big(\mathscr{O}\big(|\mathscr{A}|^3\big) + \mathscr{O}\big(|\mathscr{B}||\mathscr{A}|^2\big) + \mathscr{O}\big(|\mathscr{A}|(\log(|\mathscr{A}|) + |\mathscr{A}||\mathscr{B}|)\big) + \mathscr{O}\big(|\mathscr{A}|\log(|\mathscr{A}|)\big)\Big)$$
$$= n\mathscr{O}\big(M|\mathscr{A}|^2\big) = \mathscr{O}\big(nM|\mathscr{A}|^2\big),$$

where $M = \max\big(|\mathscr{A}|, |\mathscr{B}|\big)$. □

In the first release of PERMAVISS v.0.0.1 we implemented `image_kernel` as it is shown in 4.3. However, the resulting method becomes slow very quickly. The main reason is that too much information is updated about how the matrices of images and kernels are changed as the

persistence values increase. However, most of this information is redundant. This is why we decided to implement a version of the method that tracks only the necessary information; in this case, the pivots on each column. This improved substantially the computational speed of the `image_kernel` method. This feature is included on PERMAVISS version 0.2.

We will now proceed to explain how the improved `image_kernel` procedure works. As usual, the setup will be a persistence morphism $f : M \to N$ together with fixed barcode bases $\mathscr{A}$ and $\mathscr{B}$ for M and N respectively. Here $\mathscr{B}$ can be given as a broken base, although for the moment we will assume that $\mathscr{B}$ is not broken; we will look at this case at the end of this section. The following will be the main variables to consider:

- Barcode basis objects `A` and `B` for the respective domain and codomain bases $\mathscr{A}$, $\mathscr{B}$.
- Matrix `F` storing $f(\mathscr{A})_{\mathscr{B}}$.
- Lists of pivots on image and kernel `pivots_Im` and `pivots_Ker`.
- We track the dimensioin of the kernel by using `kernel_dim`, initially set to 0.
- Matrices for storing the barcode decompositions for $\mathrm{Ker}(f)$ and $\mathrm{Im}(f)$ which are called `Im_barcodes` and `Ker_barcodes` respectively; both have dimensions $(2, \texttt{A.dim})$. The first column of `Im_barcodes` is equal to the birth values of the barcodes in $\mathscr{A}$. That is, we run the following command

  ```
  >>> Im_barcodes[:, 0] = A.barcode[:, 0]
  ```

- Matrices for storing coordinates `Im_coordinates` and `Ker_coordinates` of respective dimensions $(\texttt{B.dim}, \texttt{A.dim})$ and $(\texttt{A.dim}, \texttt{A.dim})$.
- $(\texttt{A.dim}, \texttt{A.dim})$ matrix `T` for tracking column addition operations. This starts being equal to the identity matrix.
- Empty list `new_pivots`.
- A pair of vectors `pivots_Im` and `pivots_Ker` which are of respective length `A.dim` and `B.dim` and with all entries equal to -2.

The improved image kernel method consists of executing the `image_kernel` procedure in an implicit, rather than explicit way. For this, instead of updating kernel and image matrices, as done in the standard procedure from section 4.3, we pay attention to the column pivots and update a column only when it is strictly necessary. That is, we loop over all persistence values `rad` at which either a bar is being born or dies in `A` or `B`, and we review the vectors `pivots_Ker` and `pivots_Im`; a sequence of checks and updates are done depending on some conditions. These are very similar

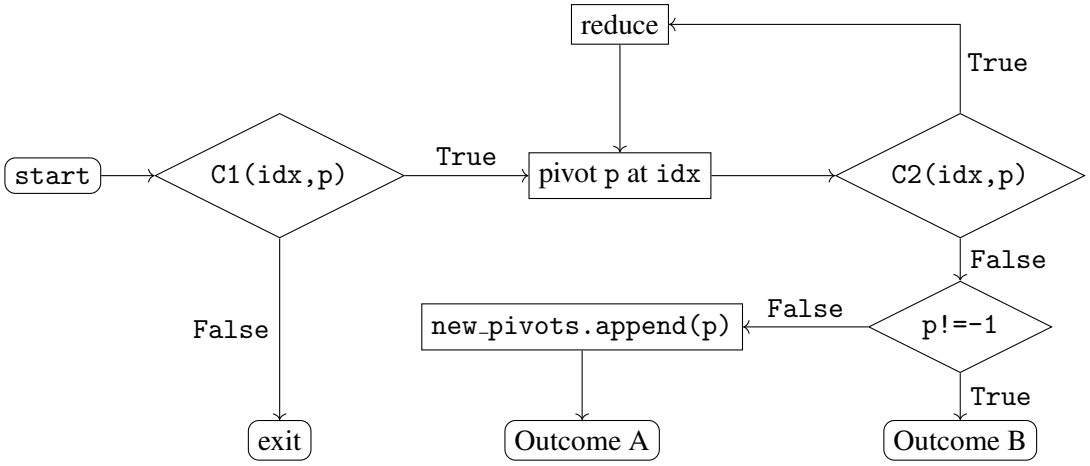


Figure A.1: Improved `image_kernel`.

for both images and kernels and are summarized on figure A.1. Following, we list the instructions for the Kernel, which are executed for an index `idx` going from 0 up to `kernel_dim`:

1. `p` $\leftarrow$ `pivots_Ker[idx]` and reset `new_pivots` $=$ `[]`.

2. We check a condition which we will name `C1(idx, p)`. This consists of checking whether one of the following is `True`:

   - `p` $\in$ `A.death(rad)`,
   - `p` $\in$ `new_pivots`.

3. If `C1(idx,p)` is `False`, then we `exit`. On the other hand, if the condition is `True`, we go on to step 4.

4. Take the `idx` column from `Ker_coordinates` restricted to the rows given by `A.active(rad)`; in NumPy notation this is `Ker_coordinates[A.active(rad), idx]`. We compute the pivot index of this column and store it on the variable `p_active`. Then, we assign to the variable `p` the value `A.active(rad)[p_active]`, however, if `A.active(rad)` is the empty list `[]`, we assign -1 to `p`.

5. Check `C2(idx,p)`, which is whether $p \in$ `pivots_Ker[:idx]` holds. That is, whether p is contained in the list `pivots_Ker` up to the entry with index `idx`; the subsequent entries are ignored. If `True` go to step 6, otherwise go to step 7.

6. Take the index `j` at which `pivots_Ker[j]` equals `p`. Then, we substract with a convenient coefficient the column `Ker_coordinates[j]` from `Ker_coordinates[idx]` to change the pivot of the `idx` column. Afterwards, go back to step 4.

7. Check whether p $\neq -1$. If False, store p into new_pivots (e.g. new_pivots.append(p)) and execute the outcome A. Otherwise execute the outcome B:

   - **Outcome A:** Set the death value of an image bar:

     ```
     >>> Im_barcodes[p, 1] = rad
     ```

     Store respective coordinates for the image, by multiplying the active rows of F times the active entries in the idx column from Ker_coordinates:

     ```
     >>> Im_coordinates[:, p] = np.matmul(
     ... F[:, A.active(rad)], Ker_coordinates[A.active(rad), idx])
     ```

   - **Outcome B:** Set the death value of a bar in the kernel:

     ```
     >>> Ker_barcodes[idx, 1] = rad
     ```

     In addition, we set pivots_Ker[idx] to be equal to -2, this being done mainly to avoid conflicts when checking C2(idx,p).

Next, for the image, we execute for all indices idx from 0 up to A.dim the following set of instructions:

1. p $\leftarrow$ pivots_Ker[idx] and reset new_pivots $=$ [].

2. We check a condition which we will name C1(idx, p). This consists in checking whether one of the following is True:

   - p $\in$ B.death(rad),
   - p $\in$ new_pivots,
   - or idx $\in$ A.birth(rad).

3. If C1(idx,p) is False, then we exit. On the other hand, if the condition is True, we go on to step 4.

4. We select the column Im_coordinates[B.active(rad), idx] and compute the pivot which denote by p_active. Next, we assign p the value B.active(rad)[p_active] while if the list B.active(rad) is empty, we assign -1 to p.

5. Check C2(idx,p), which is whether $p \in$ pivots_Im[:idx] holds. If True go to step 6, otherwise go to step 7.

6. Take the index j at which pivots_Im[j] equals p. Then, we subtract with a convenient coefficient the column Im_coordinates[j] to Im_coordinates[idx]; this way the pivot of the idx column is changed. We repeat the same operation by subtracting the column T[j] to the column T[idx]. Afterwards, go back to step 4.

7. Check whether `p` $\neq -1$. If `False` execute outcome A, otherwise execute outcome B:

   - **Outcome A:** Store `p` into `new_pivots` (e.g. `new_pivots.append(p)`).
   - **Outcome B:** Set death and birth values for bars in the image and kernel:

     ```
     >>> Im_barcodes[p, 1] = rad
     >>> Ker_barcodes[kernel_dim, 0] = rad
     ```

     Set coordinates for the kernel, add a new pivot to `pivots_Ker` and increase the dimension:

     ```
     >>> Ker_coordinates[:, kernel_dim] = T[:, idx]
     >>> pivots_Ker[kernel_dim] = idx
     >>> kernel_dim += 1
     ```

     To end, we set `pivots_Im[idx]` to be equal to -2.

Bibliography

[1] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier. Persistence images: a stable vector representation of persistent homology. *J. Mach. Learn. Res.*, 18:Paper No. 8, 35, 2017.

[2] S. Basu and L. Parida. Spectral sequences, exact couples and persistent homology of filtrations. *Expositiones Mathematicae*, 35(1):119–132, 2017.

[3] U. Bauer. *Persistence in discrete Morse theory*. PhD thesis, Niedersächsische Staats-und Universitätsbibliothek Göttingen, 2011.

[4] U. Bauer. Ripser: efficient computation of Vietoris-Rips persistence barcodes, 2021.

[5] U. Bauer, M. Kerber, and J. Reininghaus. Distributed Computation of Persistent Homology. *Proceedings of the Workshop on Algorithm Engineering and Experiments*, 2013.

[6] U. Bauer, M. Kerber, and J. Reininghaus. Clear and compress: Computing persistent homology in chunks. In *Topological methods in data analysis and visualization III*, pages 103–117. Springer, 2014.

[7] U. Bauer, M. Kerber, J. Reininghaus, and H. Wagner. Phat—persistent homology algorithms toolbox. *J. Symbolic Comput.*, 78:76–90, 2017.

[8] U. Bauer and M. Lesnick. Induced matchings of barcodes and the algebraic stability of persistence. *Proceedings of the Annual Symposium on Computational Geometry*, 6(2):355–364, 2014.

[9] F. Belchi, M. Pirashvili, J. Conway, M. Bennett, R. Djukanovic, and J. Brodzki. Lung Topology Characteristics in patients with Chronic Obstructive Pulmonary Disease. *Scientific Reports*, 8(1):5341, 2018.

[10] P. Bendich, H. Edelsbrunner, and M. Kerber. Computing Robustness and Persistence for Images. *IEEE transactions on visualization and computer graphics*, 16:1251–1260, 2011.

[11] H. B. Bjerkevik, M. B. Botnan, and M. Kerber. Computing the Interleaving Distance is NP-Hard. *Foundations of Computational Mathematics*, 2019.

[12] J. M. Boardman. Conditionally convergent spectral sequences. *Contemporary Mathematics*, 239:49–84, 1999.

[13] J.-D. Boissonnat, L. J. Guibas, and S. Y. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. *Discrete Comput. Geom.*, 42(1):37–70, 2009.

[14] J.-D. Boissonnat and C. Maria. The simplex tree: an efficient data structure for general simplicial complexes. *Algorithmica*, 70(3):406–427, 2014.

[15] R. Bott and L. W. Tu. *Differential forms in algebraic topology*, volume 82 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[16] G. E. Bredon. *Sheaf Theory*. Springer-Verlag New York, 1997.

[17] A. Brown, O. Bobrowski, E. Munch, and B. Wang. Probabilistic convergence and stability of random mapper graphs. *Journal of Applied and Computational Topology*, 5(1):99–140, 2021.

[18] P. Bubenik. Statistical topological data analysis using persistence landscapes. *The Journal of Machine Learning Research*, 16(1):77–102, 2015.

[19] P. Bubenik and P. Dłotko. A persistence landscapes toolbox for topological statistics. *Journal of Symbolic Computation*, 78:91–114, 2017.

[20] H. Byrne, H. Harrington, R. Muschel, G. Reinert, B. Stolz, and U. Tillmann. Topology characterises tumour vasculature. *Mathematics Today*, 5(55), 2019.

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

[22] G. Carlsson. Topological pattern recognition for point cloud data. *Acta Numerica*, 2014.

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

[24] G. Carlsson, V. de Silva, and D. Morozov. Zigzag persistent homology and real-valued functions. *Proceedings of the 25th annual symposium on Computational geometry*, 2009.

[25] G. Carlsson and R. B. Gabrielsson. Topological approaches to deep learning. *arXiv:1811.01122*, 2018.

[26] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. *Discrete Comput. Geom.*, 42(1):71–93, 2009.

[27] M. Carrière and S. Oudot. Structure and stability of the one-dimensional mapper. *Found. Comput. Math.*, 18(6):1333–1396, 2018.

[28] N. J. Cavanna. *Methods in Homology Inference*. Doctoral dissertations. 2118., University of Connecticut, 2019.

[29] N. J. Cavanna, M. Jahanseir, and D. R. Sheehy. A geometric perspective on sparse filtrations. *arXiv:1506.03797*, 2015.

[30] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. *The structure and stability of persistence modules*. SpringerBriefs in Mathematics. Springer, [Cham], 2016.

[31] F. Chazal and S. Y. Oudot. Towards persistence-based reconstruction in Euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 232–241, 2008.

[32] C. Chen and M. Kerber. Persistent homology computation with a twist. *27th European Workshop on Computational Geometry . . .*, 45(3):28–31, 2011.

[33] C. Chen and M. Kerber. An output-sensitive algorithm for persistent homology. *Computational Geometry: Theory and Applications*, 46(4):435–447, 2013.

[34] T. Y. Chow. You could have invented spectral sequences. *Notices of the AMS*, pages 15–19, 2006.

[35] J. Cirici, D. Egas Santander, M. Livernet, and S. Whitehouse. Derived A-infinity algebras and their homotopies. *Topology Appl.*, 235:214–268, 2018.

[36] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 37(1):103–120, 2007.

[37] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 2007.

[38] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Persistent homology for kernels, images, and cokernels. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 1011–1020, USA, 2009. Society for Industrial and Applied Mathematics.

[39] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and Vineyards by Updating Persistence in Linear Time. *Proceedings of the twenty-second annual symposium on Computational geometry*, pages 119–126, 2006.

[40] É. Colin de Verdière, G. Ginot, and X. Goaoc. Helly numbers of acyclic families. *Advances in Mathematics*, 253:163–193, 2014.

[41] G. E. Cooke and R. L. Finney. *Homology of cell complexes*. Based on lectures by Norman E. Steenrod. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1967.

[42] W. Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *J. Algebra Appl.*, 14(5):8,1550066, 2015.

[43] J. Curry, R. Ghrist, and V. Nanda. Discrete Morse theory for computing cellular sheaf cohomology. *Foundations of Computational Mathematics*, 16(4):875–897, 2016.

[44] J. M. Curry. *Sheaves, Cosheaves and Applications*. PhD thesis, The University of Pennsylvania, 2014.

[45] J. M. Curry. Topological data analysis and cosheaves. *Japan Journal of Industrial and Applied Mathematics*, 32(2):333–371, 2015.

[46] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic and Geometric Topology*, 7(1):339–358, 2007.

[47] V. De Silva, D. Morozov, and M. Vejdemo-Johansson. Dualities in persistent (co)homology. *Inverse Problems*, 27(12):1–16, 2011.

[48] V. de Silva, E. Munch, and A. Stefanou. Theory of Interleavings on Categories with a Flow. *Theory and Applications of Categories*, 33(21):583–607, 2018.

[49] C. J. A. Delfinado and H. Edelsbrunner. An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere. *Computer Aided Geometric Design*, 12(7):771–784, nov 1995.

[50] T. Dey, F. Fan, D. Shi, and Y. Wang. SimPers: Software for Topological Persistence under Simplicial Maps.

[51] T. K. Dey, H. Edelsbrunner, S. Guha, and D. V. Nekhayev. Topology preserving edge contraction. *Publ. Inst. Math.(Beograd)(NS)*, 66(80):23–45, 1999.

[52] T. K. Dey, F. Fan, and Y. Wang. Computing topological persistence for simplicial maps. In *Proceedings of the thirtieth annual symposium on Computational geometry*, pages 345–354, 2014.

[53] T. K. Dey, D. Shi, and Y. Wang. SimBa: An Efficient Tool for Approximating Rips-Filtration Persistence via Simplicial Batch Collapse. *ACM J. Exp. Algorithmics*, 24, 2019.

[54] B. Di Fabio and C. Landi. A Mayer–Vietoris Formula for Persistent Homology with an Application to Shape Recognition in the Presence of Occlusions. *Foundations of Computational Mathematics*, 11:499–527, 2011.

[55] B. Di Fabio and C. Landi. Persistent homology and partial similarity of shapes. *Pattern Recognition Letters*, 33(11):1445–1450, aug 2012.

[56] J. Ebert and O. Randal-Williams. Semisimplicial spaces. *Algebraic and Geometric Topology*, 19(4):2099–2150, 2019.

[57] H. Edelsbrunner and J. Harer. Persistent homology—a survey. In *Surveys on discrete and computational geometry*, volume 453 of *Contemp. Math.*, pages 257–282. Amer. Math. Soc., Providence, RI, 2008.

[58] H. Edelsbrunner and J. Harer. *Computational topology.* American Mathematical Society, Providence, R.I., 2013.

[59] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.

[60] G. Ellis. *An invitation to computational homotopy*. Oxford University Press, Oxford, 2019.

[61] G. Ellis. *HAP – Homological Algebra Programming, Version 1.29*, 2021.

[62] A. Fomenko and D. Fuchs. Homotopical topology, volume 273 of Graduate Texts in Mathematics, 2016.

[63] R. Forman. A user's guide to discrete Morse theory. *Sém. Lothar. Combin.*, 48:Art. B48c, 35, 2002.

[64] R. W. Ghrist. *Elementary applied topology*, volume 1. Createspace Seattle, 2014.

[65] D. Govc and P. Skraba. An Approximate Nerve Theorem. *Foundations of Computational Mathematics*, 18(5):1245–1297, sep 2018.

[66] L. J. Guibas and S. Y. Oudot. Reconstruction using witness complexes. *Discrete Comput. Geom.*, 40(3):325–356, 2008.

[67] A. Guidolin and A. Romero. Effective computation of generalized spectral sequences. In *ISSAC'18—Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, pages 183–190. ACM, New York, 2018.

[68] A. Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[69] D. Horak, S. Maletić, and M. Rajkovic. Persistent Homology of Complex Networks. *J. Stat. Mech.: Theory Experiment*, 2009, 2008.

[70] T. Kaczynski, K. Mischaikow, and M. Mrozek. *Computational homology*, volume 157. Springer Science and Business Media, 2004.

[71] M. Kerber. Persistent Homology – State of the art and challenges 1 Motivation for multi-scale topology. *Internat. Math. Nachrichten Nr*, 231(231):15–33, 2016.

[72] M. Kerber and H. Schreiber. Barcodes of Towers and a Streaming Algorithm for Persistent Homology. *Discrete and Computational Geometry*, pages 1–21, 2018.

[73] D. Kozlov. *Combinatorial algebraic topology*, volume 21 of *Algorithms and Computation in Mathematics*. Springer, Berlin, 2008.

[74] M. Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Found. Comput. Math.*, 15(3):613–650, 2015.

[75] R. Lewis and D. Morozov. Parallel computation of persistent homology using the blowup complex. *Annual ACM Symposium on Parallelism in Algorithms and Architectures*, 2015-June:323–331, 2015.

[76] D. Lipsky, P. Skraba, and M. Vejdemo-Johansson. A spectral sequence for parallelized persistence. *arXiv:1112.1245*, 2011.

[77] M. Livernet, S. Whitehouse, and S. Ziegenhagen. On the spectral sequence associated to a multicomplex. *Journal of Pure and Applied Algebra*, 224(2):528–535, 2020.

[78] D. Lütgehetmann, D. Govc, J. P. Smith, and R. Levi. Computing Persistent Homology of Directed Flag Complexes. *Algorithms*, 13(1), 2020.

[79] S. Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.

[80] W. S. Massey. *A basic course in algebraic topology*, volume 127 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[81] B. Matschke. Successive Spectral Sequences. *arXiv:1308.3187*, 2013.

[82] P. May. *A Concise Course in Algebraic Topology*. University of Chicago Press, 1999.

[83] J. McCleary. *A User's Guide to Spectral Sequences*. Cambridge University Press, 2000.

[84] N. Milosavljević, D. Morozov, and P. Škraba. Zigzag persistent homology in matrix multiplication time. *Proceedings of the Annual Symposium on Computational Geometry*, pages 216–225, 2011.

[85] J. R. Munkres. *Elements of algebraic topology*. Addison-Wesley, 1984.

[86] V. Nanda. *Discrete Morse theory for filtrations*. PhD thesis, Rutgers University-Graduate School-New Brunswick, 2012.

[87] C. T. Nathaniel Saul. Scikit-TDA: Topological Data Analysis for Python, 2019.

[88] M. Nicolau, A. J. Levine, and G. Carlsson. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences*, 108(17):7265–7270, apr 2011.

[89] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1-3):419–441, 2008.

[90] NLab authors. Top. `http://ncatlab.org/nlab/show/Top`, 2021.

[91] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(1), 2017.

[92] S. Oudot. *Persistence Theory: From Quiver Representations to Data Analysis*, volume 209. American Mathematical Society, 2015.

[93] M. Palser. An Excision Theorem for Persistent Homology. *arXiv:1910.03348*, 2019.

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

[95] J. A. Perea and J. Harer. Sliding windows and persistence: an application of topological methods to signal analysis. *Found. Comput. Math.*, 15(3):799–838, 2015.

[96] R. Rabadan and A. J. Blumberg. *Topological Data Analysis for Genomics and Evolution: Topology in Biology*. Cambridge University Press, 2020.

[97] M. W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram. Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function. *Frontiers in Computational Neuroscience*, 11:48, 2017.

[98] J. Reininghaus, S. Huber, U. Bauer, and R. Kwitt. A stable multi-scale kernel for topological machine learning. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 07-12-June, pages 4741–4748, 2015.

[99] V. Robins and K. Turner. Principal component analysis of persistent homology rank functions with case studies of spatial point patterns, sphere packing and colloids. *Physica D: Nonlinear Phenomena*, 334:99–117, nov 2016.

[100] M. Robinson. *Topological Signal Processing*. Springer-Verlag, Berlin Heidelberg, 2014.

[101] M. Robinson. Assignments to sheaves of pseudometric spaces. *Compositionality*, 2(2), 2020.

[102] A. Romero, J. Heras, J. Rubio, and F. Sergeraert. Defining and computing persistent Z-homology in the general case. *arXiv:1403.7086*, 2014.

[103] A. Romero, J. Rubio, and F. Sergeraert. Computing spectral sequences. *Journal of Symbolic Computation*, 41(10):1059–1079, 2006.

[104] A. Romero, J. Rubio, and F. Sergeraert. Effective homology of filtered digital images. *Pattern Recognition Letters*, 83:23–31, 2016.

[105] J.-P. Serre. Faisceaux algébriques cohérents. (French). *Annals of Mathematics*, 61:197–278, 1955.

[106] D. R. Sheehy. Linear-size approximations to the Vietoris-Rips filtration. *Discrete Comput. Geom.*, 49(4):778–796, 2013.

[107] V. D. Silva and G. Carlsson. Topological estimation using witness complexes. *Proc. Sympos. Point-Based Graphics*, 2004.

[108] G. Singh, F. Mémoli, and G. Carlsson. Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. *Eurographics Symposium on Point-Based Graphics*, pages 91–100, 2007.

[109] P. Skraba and K. Turner. Wasserstein Stability for Persistence Diagrams. *arXiv:2006.16824*, 2021.

[110] B. J. Stolz. Outlier-robust subsampling techniques for persistent homology. *arXiv:2103.14743*, 2021.

[111] A. Tausz and G. Carlsson. Applications of zigzag persistence to topological data analysis. *arXiv preprint arXiv:1108.3545*, 2011.

[112] The GAP group. *GAP – Groups, Algorithms, and Programming, Version 4.11.1*, 2021.

[113] The GUDHI Project. *GUDHI User and Reference Manual*. GUDHI Editorial Board, 3.4.1 edition, 2021.

[114] A. Torras and U. Pennig. Interleaving Mayer-Vietoris spectral sequences. *arXiv:2105.03307*, 2021.

[115] A. Torras Casas. Distributing Persistent Homology via Spectral Sequences. *arXiv:1907.05228*, 2019.

[116] A. Torras Casas. PerMaViss: Persistence Mayer Vietoris spectral sequence. 10.5281/zenodo.3613870, jan 2020.

[117] H. Wagner, C. Chen, and E. Vuçini. *Efficient Computation of Persistent Homology for Cubical Data*, pages 91–106. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[118] C. A. Weibel. *An Introduction to Homological Algebra*. Cambridge University Press, Cambridge, 1994.

[119] H. R. Yoon. *Cellular Sheaves And Cosheaves For Distributed Topological Data Analysis*. PhD thesis, University of Pennsylvania, 2018.

[120] H. R. Yoon and R. Ghrist. Persistence by Parts: Multiscale Feature Detection via Distributed Persistent Homology. *arXiv:2001.01623*, 2020.

[121] A. Zomorodian. Fast construction of the Vietoris-Rips complex. *Computers and Graphics (Pergamon)*, 34(3):263–271, 2010.

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

[123] A. Zomorodian and G. Carlsson. Localized homology. *Computational Geometry: Theory and Applications*, 41(3):126–148, 2008.

[124] A. J. Zomorodian. *Topology for computing*, volume 16 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.

Index