Analyzing Stratified Spaces Using Persistent Versions of Intersection and Local Homology

by

Paul Bendich

Department of Mathematics
Duke University

Ph.D. Dissertation
2009

Abstract

This dissertation places intersection homology and local homology within the framework of persistence, which was originally developed for ordinary homology by Edelsbrunner, Letscher, and Zomorodian. The eventual goal, begun but not completed here, is to provide analytical tools for the study of embedded stratified spaces, as well as for high-dimensional and possibly noisy datasets for which the number of degrees of freedom may vary across the parameter space.

Specifically, we create a theory of persistent intersection homology for a filtered stratified space and prove several structural theorems about the pair groups associated to such a filtration. We prove the correctness of a cubic algorithm which computes these pair groups in a simplicial setting. We also define a series of intersection homology elevation functions for an embedded stratified space and characterize their local maxima in dimension one.

In addition, we develop a theory of persistence for a multi-scale analogue of the local homology groups of a stratified space at a point. This takes the form of a series of local homology vineyards which allow one to assess the homological structure within a one-parameter family of neighborhoods of the point. Under the assumption of dense sampling, we prove the correctness of this assessment at a variety of radius scales.

Contents

List of Tables

List of Figures

Acknowledgements

Due to the efforts great and small of a wide variety of people, these past five years have been wonderful for me. If there is anyone I have forgotten to thank here, I sincerely apologize.

My education has been made possible by the love, encouragement, and financial support of my parents, Steve and Susie, my aunt Naomi, my grandmothers Ginny, Clare, and Sally, my aunt Judy, and my uncle Arnie. I am very grateful for all they have given me.

My advisor, John Harer, has been a tremendous help and inspires me by the generous and collaborative spirit in which he works. It has been a great deal of fun to do research within such an environment and I have been left with many valuable colleagues and friends. I thank him for all the time and effort he has put in with me. I am especially grateful for the hospitality shown me by him, and of course by his wife Anne and daughter Aoife, during my extended time in Paris. I would also like to thank Danielle Zelic for housing me during that time.

I am indebted to Herbert Edelsbrunner for his time and patience in introducing me to many new ideas. For similar reasons, I thank his student and my good friend, Dmitriy Morozov, who has put up with far too many of my trivial questions. I would also like to thank Simon Lunagomez, Mauro Maggioni, Jonathan Mattingly, Yuriy Mileyko, Sayan Mukherjee, Amit Patel, and Bei Wang, for many stimulating research conversations that I hope will continue in the future.

I also thank the many interesting people, and hopefully potential collaborators, that I have met at the DARPA-TDA and DARPA-FUNBIO grant meetings. For setting up these meetings, and for generous financial support during some of my graduate semesters, I am extremely grateful to DARPA and in particular to the project manager Ben Mann. I would also like to thank Kevin Knudson and Mississippi State University for setting up an excellent summer research program, and Henry King for his hospitality and valuable research advice during my visit to Maryland.

I greatly appreciate the work of past and present members of the departmental staff during my time here. In particular, I would like to thank Carolyn Sessoms and Shannon Holder for patiently dragging me out of several self-inflicted administrative jams, and Georgia Barnes and Katy Wilson for interesting lunchtime conversation. I am also grateful to Gail Smith from the Computer Science Department for help with travel and grant purchases.

Teaching at Duke has been an extremely valuable experience for me, both professionally and personally. The training, advice, and encouragement provided by Jack Bookman and Lewis Blake has been wonderful. I would especially like to thank Lewis for his help in navigating through several particularly treacherous grading issues, and

Jack for his teaching seminar and for his insightful and oft-shared perspective on academic life. In addition, I would like to thank these two gentlemen, as well as Mike Reed and the many graduate student members of the Graduate Student Calculus Curriculum Committee, for their extensive efforts in designing and supporting the new course Math 41L. I have learned a great deal from this process and it has been surprisingly enjoyable.

My fellow graduate students, as well as many past and present postdocs, have been a great source of support and light entertainment over the past five years. In no particular order:

- Rann Bar-On, for friendship, well-calibrated culinary snobbery, and invaluable computer assistance. To enhance my future electability, I hereby distance myself from all of his political beliefs.
- Arya Roy, for being a perfect English gentleman and taking smoke breaks at the same frequency that I took "go outside and don't work" breaks.
- Mike Gratton, for a well-timed gift of corn whiskey. Also, the whole dukedissertation.cls creation thing (

  ```
  \do{research}
  ```

) was very nice, as was the heroic hand-holding during my format check induced mental breakdown.
- Abe Smith, for therapeutic assistance with LaTex and differential geometry.
- Andrea Watkins, for being generally sane.
- Eric Katz, for being a great deal of help with things algebraic, as well as serving as a cautionary example.
- Graham Cox, for being Canadian.
- David Rose, ditto.
- Nick Robbins, Joseph Spivey(!), and Tim Lucas, for their solidarity during the fairly traumatic jobs festival in San Diego.
- Aubrey HB and Mike Jenista, for giving hope to the younger generation.
- Dave Cesa, for his sunny disposition and inspirational outlook on life.

Finally, I would like to thank the owners, staff, and habitual customers of The Green Room, Dain's, Bean Traders, Broad Street Cafe, and the now sadly extinct Liberty Diner, where much of this dissertation was researched and written.

1

Introduction

This dissertation takes two algebraic tools which have proven very useful in the study of stratified spaces and adapts them into the persistent homology framework. More specifically, it concerns:

- the theoretical development of persistent intersection homology for a triangulated stratified space, along with an algorithm for its computation and the proof of several structural results. The eventual aim, only partially completed here, is to furnish a series of measurements which illuminate important features of an embedded stratified space and the relationship among its strata.
- a treatment of persistent local homology for point cloud data, together with proofs of several stability results. The future goal is to provide a series of possible stratifications for a point cloud which reflect its local structure at a number of scale levels.

These two topics will be outlined at greater length later in this introduction. However, it seems wise to first briefly discuss the still fairly new context in which they fit.

1.1 Persistent Homology

The main definitions and theorems of persistent homology will be discussed in great detail in Chap. 2.

A key motivating principle is the following: in a large dataset or other uncertain scientific application, it is not really possible to distinguish between artifacts caused by noise and actual phenomena which happen to exist on an extremely small scale. By using topological methods, and persistent homology in particular, one hopes to

provide a consistent framework for measuring data on a variety of scales, allowing the user to decide what is relevant to whatever problem is under study.

For a toy example of this idea, consider an annulus and a disc. A traditional topological tool, the first homology group, distinguishes these objects. On the other hand, suppose we do not have the spaces themselves, but only a set of points sampled from each, possibly with noise. Depending on the density and accuracy of the sample, and on the relative sizes of the inner and outer radii of the annulus, these point clouds might look very similar, or quite distinct, or somewhere in between. Persistent homology will provide a multiscale measurement of the difference between the two clouds. Depending upon the particular application, one might then set a threshold beyond which the clouds are considered adequately distinct.

Persistent homology was first developed in [17]. As its name might suggest, persistent homology is a method for measuring how long a homology class persists along a filtration of a topological space (in the example above, the topological space would be the plane, and the filtration would be provided by the gradually thickening point clouds).

It has already been a important theoretical tool in a wide variety of applications, including image compression ([8]) and segmentation ([16]), speech pattern analysis ([7]), neuroscience ([9]), effective coverage in sensor networks ([21], [20]), and gene expression analysis [2].

Elevation One example of persistence along a particular filtration is the following. Suppose that X is a manifold and let f be a real-valued Morse ([29]) function on X. For all real numbers α, consider the sublevel set $X_{\leq\alpha} = f^{-1}((-\infty, \alpha])$. As α increases, these sublevel sets filter the manifold. According to Morse Theory, the homology will change only when we pass a critical point of f: in that case, either a single class will be added or a single class will be destroyed. By tracking which critical point creates a class and then which critical point destroys it, we get a pairing between the critical points of f. These pairings allow us not only to compute the homology of X itself, but also to measure the size of various transient homological classes which exist only within certain ranges of sublevel set. Importantly, these measurements are stable under small perturbations of the function ([11]).

More specifically, suppose X is a manifold embedded in Euclidean space of one dimension higher and that $f = f_v$ measures height in a particular unit direction v; in this case, a point $x \in X$ will be critical for f_v iff v is normal to X at the point. So, letting v vary over the entire unit sphere, we see that x will be paired for precisely two such height functions. As a consequence of Poincaré Duality, it can be shown ([10]) that the same persistence measurement is obtained at x for each of these two directions. This allows the definition of an Elevation ([1]) function E on X, where $E(x)$ is defined to be this persistence measurement. The local maxima of this function then locate interesting features of the manifold, such as the deepest point in a pocket; in [1] and [34], these ideas are applied to protein docking.

1.2 Intersection Homology

On the other hand, there are many applications in which X might not be a manifold: for example, the protein-protein interface surface for three or more proteins ([3]), is a 2-dimensional stratified space embedded in three-space. In such a case, restricting to measurements derived from homology alone will miss important relationships between X and its singularities. By computing an analogue of the Elevation function discussed above, but this time using intersection homology ([22]) groups, we aim to give a fuller picture.

The development of the theory needed for this goal makes up the bulk of the original work in this dissertation. It takes place over three chapters:

- Chap. 3 describes the intersection homology groups and discusses some of their properties; as will be explained in that chapter, we also need to make a few adjustments to the definition from [22] in order to make the theory more effective in our context.
- Chap 4 defines persistent intersection homology for a filtered stratified space and gives an algorithm for its computation along with a proof of correctness. In addition, two structural theorems, Intersection Pair Group Duality and Symmetry, are discussed and proven.
- In Chap. 5, we compare the information gained by homology and intersection homology persistence on an embedded stratified space. We also define a series of intersection homology elevation functions and characterize some of their local maxima.

1.3 Local Homology

The rest of this dissertation concerns the theoretical development of a multi-scale analogue of the local homology groups of a space X at a point z. As explained in Chap. 6, the result will be a series of persistence diagrams which are stacked up into *vineyards*([13]); these local homology vineyards illustrate the persistent homological features of X within a series of neighborhoods of z.

The eventual goal, begun but not completed here, is to use these vineyards to approximate a stratification, or number of stratifications, for a high-dimensional point cloud. We now explain briefly why this might be useful.

Dimension Reduction Much scientific data takes the form of a dataset P which lies in some high-dimensional ambient space. Often the first step in analysis is an attempted reduction in dimension: one tries to find a space X, of hopefully much smaller dimension than the ambient space, such that P lies on or near X.

One way to do this is to search for dependency among variables, hoping thereby to recognize that the dataset really lies on a smaller dimensional manifold (or a noisy version thereof) embedded in a higher dimensional space. This approach goes under the name *manifold learning*. Techniques in manifold learning include using principal component analysis to identify the direction to project the data, local versions of PCA that do the same at each point, and many other more innovative methods that use, for example, Graph Laplacians, to try to account for the curvature of the underlying manifold (see, for example, [5] and [33]). Implicit in most of this work is the assumption that the dataset lies on a manifold and therefore that the dimensionality of the dependence is constant. Of course, it is well known in the manifold learning community that this is not true in general, but so far there are no formal methods for discovering the structure of the dimension changes that may occur.

It is our hope that that the local homology vineyards can eventually make a contribution to the development of "manifold-learning"-type methods which can be applied more effectively to datasets where the number of degrees of freedom varies considerably across the parameter space.

These vineyards are defined in Chap. 6, portions of which have already appeared in print as [6]. In the last section of that Chapter, we begin the discussion of how to use the vineyards, or variants thereof, to find multi-scale stratified structure.

2

Basics and Background

In this chapter, we give a detailed discussion of persistent homology. Sec. 2.1 contains the basic definitions, as well as a description of the simplicial persistence algorithm. Certain specific examples of persistence, upon which we will later build using other types of homology, are described in detail in Sec. 2.2. The chapter closes with a description of extended persistence and the algorithm for its computation.

Throughout this chapter, and in the rest of this dissertation, all homology groups are computed using $\mathbb{Z}/2\mathbb{Z}$ coefficients.

2.1 Ordinary Persistence

Persistent homology is meant to provide a measure of the "importance" of homology classes which appear, and sometimes subsequently disappear, at certain levels of a filtration. In this section, we make this idea precise, starting in the most general setting and then restricting to the case of a filtered simplicial complex. The section closes with an algorithm to compute the persistent homology of the finest possible filtration of a simplicial complex: one in which we add one simplex at a time. More immediately meaningful examples, such as Morse persistence on a manifold or alpha-shape persistence for a point set, will be discussed in the next section.

2.1.1 General Definition

Let X be a topological space along with a filtration F by finitely many closed subspaces:

$$\varnothing = X_0 \subseteq X_1 \subseteq X_2 \ldots \subseteq X_n = X \tag{2.1}$$

If $i < j$, we will often say that X_i is earlier or older than X_j, and that X_j is later or younger than X_i. To simplify notation, we let H_r^i denote the homology group $H_r(X_i)$. For all integers i, j, r with $0 \leq i < j \leq n$ and $r \geq 0$, the inclusion $X_i \hookrightarrow X_j$ induces a map on homology $f_r^{i,j} : H_r^i \to H_r^j$. Let $\beta_r^{i,j}$ denote the rank of the image of this map.

Birth, Death, and Pair Groups A class $\alpha \in H_r^i$ is said to be *born* at level i iff $\alpha \notin \mathrm{Im}(f_r^{i-1,i})$; in other words, iff α has no chain representative from earlier in the filtration. Observe that if α is born at level i, then every class in the coset $[\alpha] = \alpha + \mathrm{Im}(f_r^{i-1,i})$ is also born at the same level. We define a *birth event* (often just called a birth) at level i to be the appearance of such a coset. Note that there may of course be several birth events, or none at all, at a particular level.

Now suppose an r-class α is born at the ith level. Then α is said to *die* at level j iff:

- $f_r^{i,j}(\alpha) \in \mathrm{Im}(f_r^{i-1,j})$
- $f_r^{i,j-1}(\alpha) \notin \mathrm{Im}(f_r^{i-1,j-1})$

Note that if α dies at level j, then every class in the coset $[\alpha]$ will also die at the same level. In other words, the disapperance of this entire coset represents a single death event at the jth level. Furthermore, there will be some class $\tilde{\alpha}$ in this coset such that $f_r^{i,j}(\tilde{\alpha}) = 0$. To see this last fact formally, choose a $\beta \in H_r^{i-1}$ such that $f_r^{i-1,j}(\beta) = f_r^{i,j}(\alpha)$. Then just put $\tilde{\alpha} = \alpha + f_r^{i-1,i}(\beta)$. It is then easy to see:

- $\tilde{\alpha} \in [\alpha]$
- $f_r^{i,j-1}(\alpha) \in \mathrm{Ker}(f_r^{j-1,j})$
- $f_r^{i,j-1}(\alpha) \notin \mathrm{Im}(f_r^{i-1,j-1})$

Then the number of r-dimensional birth events at level i which subsequently become death events at the jth level is the rank $\mu_r^{i,j}$ of the following group, which we call a *pair group*:

$$P_r^{i,j} = \frac{\mathrm{Im}(f_r^{i,j-1}) \cap \mathrm{Ker}(f_r^{j-1,j})}{\mathrm{Im}(f_r^{i-1,j-1}) \cap \mathrm{Ker}(f_r^{j-1,j})} \tag{2.2}$$

If we need to specify a particular filtration F of a particular space X, we will call these groups $P_r^{i,j}(X, F)$.

A class α born at level i which then later dies at some other level will be called an *inessential* class; otherwise, it will be called an *essential* class. These latter classes, and their eventual death along a descending filtration, are the main subject of Sec. 2.3.

Persistence and Persistence Diagrams Given a class $\alpha \in P_r^{i,j}$, one could say that its *persistence* is simply the number of levels in which it lives; in other words, one might define $pers(\alpha) = j - i$.

However, many filtrations of a topological space come equipped with some meaningful nondecreasing value $h(i)$ placed on each level i of the filtration; one important class of such examples, the sublevel set filtration given by a Morse function, will be discussed in detail in the next section. In such a case, we obviously would gain a richer type of information by defining $pers(\alpha) = h(j) - h(i)$.

These persistence values are then encoded into several *persistence diagrams*, one for each dimension. For each nonzero pair group $P_r^{i,j}$, we place a point of multiplicity $\mu_r^{i,j}$ at $(h(i), h(j))$ in the extended plane. And for each coset of essential classes born at the ith level, we place a point $(h(i), \infty)$. Note then that the persistence of a class is seen via the vertical distance of its corresponding point from the diagonal in the diagram.

We also place a point of infinite multiplicity at each point on the major diagonal; this is so that we can later define a notion of distance between persistence diagrams which will be stable in many important situations. The multiset of all such points is then called the rth persistence diagram for this filtration. See Fig. 2.2 for an example of such a diagram.

The interpretation of these diagrams will of course depend on the meaning attached to a given filtration of a topological space. However, we can make two general observations:

- The rank $\beta_r^{i,j}$ of $f_r^{i,j}$ is just the number of points, counted with the appropriate multiplicity, in the upper left quadrant defined by the point $(h(i), h(j))$.
- The longer a class persists in the filtration, the farther its corresponding point will be from the diagonal. In many applications, the further away a point is from the diagonal, the less likely it is to be interpreted as "noise," although it it of course impossible to attach a precise meaning to this concept.

We close this subsection with a proposition which will be needed later for computational purposes; its proof is obvious.

1 (Proposition)**.** *Suppose that $F = \{X_i\}$ and $G = \{Y_i\}$ are filtrations of the topological spaces X and Y respectively. Suppose further that there are maps $\phi_i : X_i \to Y_i$, such that:*

- *ϕ_i induces isomorphisms on homology, for each i and in each dimension.*

- *For each $i < j$, the following diagram commutes:*

$$\begin{array}{ccc} X_i & \longrightarrow & X_j \\ {\scriptstyle \phi_i}\downarrow & & \downarrow{\scriptstyle \phi_j} \\ Y_i & \longrightarrow & Y_j \end{array}$$

Then $P_r^{i,j}(X, F) \cong P_r^{i,j}(Y, G)$, for all integers r and all $1 \leq i < j \leq n$.

2.1.2 Simplicial Setting

We can also define persistent homology for the case of a filtered simplicial complex. Let K be a finite simplicial complex and suppose that

$$\varnothing = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K \tag{2.3}$$

is a filtration of K by a nested sequence of subcomplexes. We then make exactly the same definitions as before for birth, death, and persistence of homology classes, as well as for persistence diagrams.

One Simplex at a Time Suppose we restrict to the case where K has n simplices which come with some ordering $\sigma_1, \sigma_2, \ldots, \sigma_n$, and that we filter K by adding one simplex at a time; that is, $K_i = K_{i-1} \cup \{\sigma^i\}$. In this case, we can say very precisely what happens between successive levels in the filtration. Suppose that $\dim(\sigma_i) = r$. There are then only two possibilities:

- The $(r-1)$-cycle α represented by $\partial\sigma_i$ is not 0 in H_r^{i-1}: After the addition of σ_i, it becomes trivial in H_r^i. Hence all classes in the coset of α die at level i, and $\beta_{r-1}(K_i) = \beta_{r-1}(K_{i-1}) - 1$, while all other Betti numbers remain unchanged. In this case, we say that σ_i is a **negative** r-simplex.

- $\alpha = 0$ in H_r^i. In other words, $\exists\gamma \in C_r(K_i)$ such that $\partial\gamma = \partial\sigma_i$. Then the r-class represented by $\gamma + \sigma_i$ gives a coset of r-classes born at the ith level. So $\beta_r(K_i) = \beta_r(K_{i-1}) + 1$; we say that σ_i is a **positive** r-simplex.

In other words, there is at most one coset of classes born at any K_i; we identify σ_i with this coset if it exists. Similarly, there is at most one coset of classes which may die at any K_j, and this coset gets identified with σ_j. Using these identifications, we can say loosely that σ_j kills σ_i, and pair the two simplices together; by the persistence of the pair (σ_i, σ_j), we will just mean the persistence of any class in the unique coset born at level i and dying at the jth level. If the coset born at level i consists of essential classes, we leave σ_i unpaired.

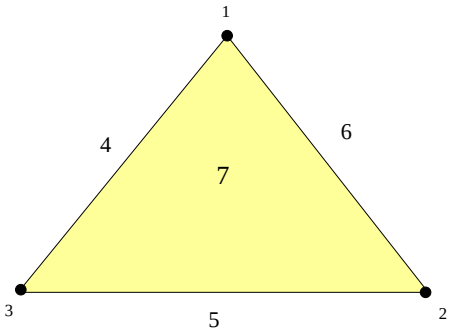

FIGURE 2.1: The simplices of the triangle are added in increasing numerical order. The addition of edge 4 merges the component formed by vertex 3, so these two simplices are paired. Similarly, we pair vertex 2 and edge 5. Edge 6 and triangle 7 are paired, while vertex 1, which represents the only essential homology class, goes unpaired.

There is then a one-to-one correspondence between the pairs (σ_i, σ_j) and the nonzero, and thus necessarily rank one, pair groups $P_r^{i,j}$, where $r = \dim(\sigma_i)$. Similarly, there is a one-to-one correspondence between the unpaired r-simplices and the rank of $H_r(K)$. Fig. 2.1 illustrates this correspondence for a filtered triangle.

The simplicial persistence algorithm, described below, computes these pairs of simplices and also identifies the unpaired ones.

2.1.3 Algorithm

The algorithm takes as input a simplicial complex K, along with an ordering $\sigma_1, \sigma_2, \ldots, \sigma_n$ on its simplices; all faces of a given simplex must precede it in the ordering.

We form the $n \times n$ $0-1$ matrix D by setting $D[i,j] = 1$ iff σ_i is a codimension-one face of σ_j. The basic idea of the algorithm is that certain column operations are performed to transform D into a *reduced* matrix; the pairings and essential classes are then read off from the result. We now explain what this means in detail.

Reduced Matrices Let M be an arbitrary $n \times n$ $0-1$ matrix. We define a "lowest-one" function $low_M : \{1, 2, \ldots, n\} \to \{0, 1, 2, \ldots, n\}$ as follows. If the jth column of M is nonzero, then we set $low_M(j) = i$, where i is the index of the lowest nonzero entry in the column. Otherwise, we set $low_M(j) = 0$.

A matrix M is said to be *reduced* iff low_M is injective on the complement of the preimage of 0; in other words, iff $low_M(j) = low_M(k) \neq 0$ implies $j = k$.

Reduction Process and Interpretation The algorithm reduces D by performing column-operations left-to-right. Precisely, we perform:

for j = 1 to n **do**
  while $\exists j' < j$ with $low(j') = low(j) \neq 0$ **do**

add column j' to column j
end while
end for.

This produces a reduced matrix R. The pairs are then given directly by the associated lowest-one function low_R as follows:

- if $low_R(k) = 0$, then σ_k is a positive simplex. It will either be paired later or remain unpaired.
- if $low_R(j) = i$, then we pair the positive simplex σ_i with the negative simplex σ_j.

The column operations above can of course be encoded by an $n \times n$ elementary matrix V: $R = DV$. The columns of V give additional information. For each chain group $C_i(K)$, we take the standard basis consisting of the i-simplices of K. Then we read off:

- If $low_R(k) = 0$, then the entries of the kth column of V give a chain representative for one of the classes in the unique coset of classes born after the addition of σ_k.
- If $low_R(j) = i$, then the entries in the jth column of V display a chain representative for one of the classes in the unique coset which die after the addition of σ_j.

Pairing Uniqueness Lemma Of course, there are many different sequences of column operations that one can perform on a given matrix in order to reduce it. Put another way, there could certainly be another elementary matrix $V' \neq V$ and another reduced matrix $R' \neq R$ such that $R' = DV'$. However, the lowest-one function of these reduced matrices will be identical and can be characterized entirely in term of ranks of certain minors of the original matrix D; this is the content of the Pairing Uniqueness Lemma, proven in [13]. Thus the pairs (and unpaired simplices) that we read off the reduced matrix will not depend on which reduction process we use. This fact will prove useful in the proof of correctness for the intersection homology persistence algorithm, which appears at the end of Chap. 4.

2.2 Important Special Cases

In the previous section, we defined persistence in its most general context: a filtered space along with some sort of value placed on each level of the filtration. In this section, we discuss several important examples for which the persistence values have immediate and meaningful interpretations; crucially, these interpretations are stable under perturbation.

2.2.1 Tame Functions

Let X be a topological space and let $f : X \to \mathbb{R}$ be a real-valued function. For each $a \in \mathbb{R}$, $X_{\leq a}$ denotes the sublevel set $f^{-1}((-\infty, a])$. Whenever $a < b$, there are homomorphisms $H_*(X_{\leq a}) \to H_*(X_{\leq b})$ induced by inclusion.

The real number a is called a *homological regular value* of f iff the maps $H_*(X_{\leq a-\epsilon}) \to H_*(X_{\leq a+\epsilon})$ are isomorphisms for all small enough positive epsilons. Otherwise a is a *homological critical value*.

We then say that a function is *tame* iff it has finitely many homological critical values and if the homology groups of each sublevel set have finite rank. Examples of tame functions include, but are not limited to, Morse functions ([29]) on compact manifolds and Stratified Morse functions ([24]) on Whitney stratified spaces.

Sublevel Set Filtration Suppose f is a tame function on X, with finitely many homological critical values $t_1 < t_2 < \ldots < t_m$. We choose $m+1$ regular values $s_1, \ldots s_m$ such that $s_{i-1} < t_i < s_i$, for $1 \leq i \leq m$; we also set $s_{-1} = t_0 = -\infty, s_{m+1} = t_{m+1} = \infty$. Letting X_i be shorthand for the sublevel set $X_{\leq s_i}$, this defines a filtration of X:

$$\varnothing = X_{-1} \subseteq X_0 \subseteq X_1 \ldots \subseteq X_{m+1} = X \tag{2.4}$$

Now the actual filtration in (2.4) depends on the particular choice of interleaved regular values s_i, whereas the only interesting homological events will happen at the critical values t_i. So we define the value of each level using the critical values. That is, we create a nondecreasing function $h : \{-1, 0, 1, \ldots, m, m+1\} \to \mathbb{R}$ by the rule $h(i) = t_i$. The persistences of homology classes along the filtration in (2.4) are then measured using these values; we let $Dgm_r(f)$ denote the resulting persistence diagrams in each dimension.

Morse Functions Consider a smooth compact d-manifold M along with a Morse ([29]) function $f : M \to \mathbb{R}$. This means that all critical points of f are nondegenerate and have distinct critical values $t_1 < t_2 < \ldots < t_m$. As before, we choose interleaved regular values s_i and define $X_i = X_{\leq s_i}$ and $h(i) = t_i$.

Recall the two fundamental theorems of Morse Theory, rephrased here in our terminology:

- For any two regular values $a < b$ such that $[a, b]$ contains no critical values, $X_{\leq b}$ deformation retracts onto $X_{\leq a}$; in fact the two sublevel sets are diffeomorphic.
- If t_i is an index r critical point of f, then X_i is homotopically equivalent to the space obtained by attaching an r-cell along its boundary to X_{i-1}.

The first fact implies that the critical values and *homological* critical values of f are identical. The second means that in the case of a Morse function the homological change at each level of the filtration is particularly simple. Assuming t_i is index r with associated critical point x_i, an analysis of the Mayer-Vietorris sequence for the union $X_{i-1} \cup \{r - \text{cell}\}$ shows that exactly one of two things may happen:

- A single coset of r-dimensional homology classes is born. In this case, we say that x_i is positive.
- A single coset of $(r-1)$-dimensional classes dies. Here x_i is called negative.

We can thus pair the critical points in the same manner as we paired the simplices of a filtered simplicial complex in which one simplex is added at each level.

One-Dimensional Example Suppose that X is the 1-dimensional manifold embedded in the plane as shown in Fig. 2.2. We let $f : X \to \mathbb{R}$ measure height in the vertical direction; more precisely, $f(x) =< x, v >$, where $v = (0, 1)$. The critical points of f are then just the local minima and maxima of this function. Each non-infinite point in $Dgm_0(f)$ then corresponds to a min-max pair, where the minimum creates a new component which is then subsequently merged by the maximum.

Intuitively, we imagine that each such point in $Dgm_0(f)$ represents a topological feature of the space; the farther away the point is from the diagonal, the less prominent the feature, and also the more sensitive to a perturbation in the Morse function, in a sense which we make precise below. Note that the global minimum is unpaired: it creates an essential component corresponding to the infinite point in $Dgm_0(f)$. In Sec. 2.3, we will see a theoretical justification for pairing this global minimum with the global maximum; the persistence of this pair will then measure the size of the manifold in the vertical direction.

Stability The persistence diagrams which arise from the sublevel set filtrations given by a tame function are stable under small perturbations of the function.

In Fig. 2.3, we see the graphs of two tame (in fact, Morse) functions $f, g : [0, 1] \to \mathbb{R}$ which are small perturbations of each other. We also see their respective persistence diagrams $Dgm_0(f), Dgm_0(g)$, which are also "close" in some sense: the two red points are very near each other, and the blue point in $Dgm_0(g)$ is almost touching the diagonal. We make this idea precise by defining a distance measure between any two persistence diagrams.

Suppose that D, D' are two persistence diagrams and consider all possible bijections $\gamma : D \to D'$; recall that each diagram contains points of infinite multiplicity all along the diagonal. The *bottleneck distance* between the two diagrams is then defined to be:

$$d_B(D, D') = \inf_\gamma \sup_{p \in D} ||p - \gamma(p)||_\infty \tag{2.5}$$

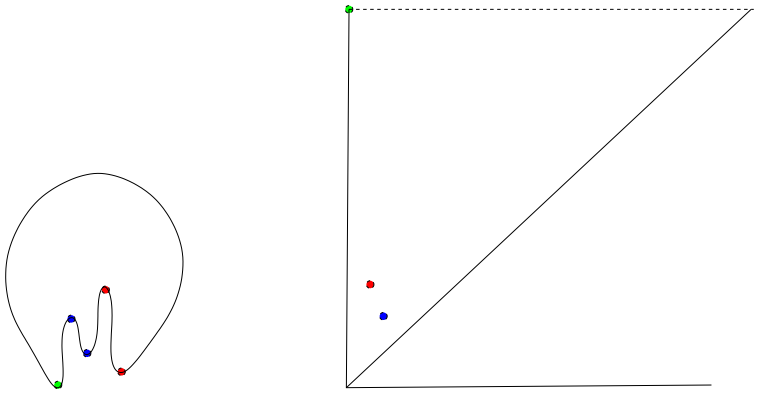

FIGURE 2.2: The origin is located at the green point and f measures height in the vertical direction. The persistence diagram $Dgm_0(f)$ appears on the right: each colored point in the diagram corresponds to a min-max pair of colored points on the loop, except for the green point which corresponds to the global minimum.

For example, one might define a bijection $\gamma : Dgm_0(f) \to Dgm_0(g)$ by mapping the red point to the other red point, and the blue point to its closest diagonal neighbor.

In [11], the authors prove that persistence diagrams are stable under this notion of distance:

2 (Stability of Persistence Diagrams)**.** *Let X be a topological space and $f, g : X \to \mathbb{R}$ tame functions. Then for each nonnegative integer r, $d_B(Dgm_r(f), Dgm_r(g)) \leq ||f - g||_\infty$.*

Lower Star Filtration Let K be a finite simplicial complex along with an injective real-valued function g on its n vertices. One may extend g via linear interpolation to a function $\tilde{g} : |K| \to \mathbb{R}$. Note that $\tilde{g}$ will be tame, for its homological critical values will be some subset of its values at the vertices. The persistence diagrams $Dgm_r(\tilde{g})$ are then defined as above. Here we describe a particular filtration of the simplicial complex K which permits the computation of these persistence diagrams.

We order the vertices of K by putting $v < w$ iff $g(v) < g(w)$; in this case, we say the v is *older* or *earlier* than w. Then for $1 \leq i \leq n$, we let K_i be the full subcomplex of K spanned by the oldest i vertices $v_1, v_2, \ldots, v_i$; we also set $K_0 = \varnothing$. and set the value of the ith level to $g(v_i)$.

Note that K_i is obtained from K_{i-1} by attaching the (closed) lower star of v_i along its lower link. For this reason, we call

$$\varnothing = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K \tag{2.6}$$

the lower star filtration $LS(g)$ of K.

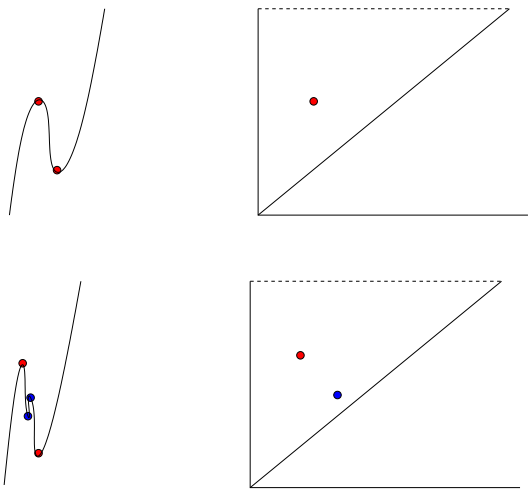

FIGURE 2.3: The graphs of two "L_∞-close" tame functions f, g on a closed interval are shown, along with their respective persistence diagrams. Note that the two diagrams are also "close" in bottleneck distance.

The pair groups $P_r^{i,j}(K, LS(g))$ can then be computed by the algorithm described in Subsection 2.1.2. The input ordering is defined with the simplices in $St_-(v_1)$ coming first, then $St_-(v_2)$ and so on; within each lower star, the simplices can be ordered arbitrarily, with $g(v_i)$ being the value of each simplex in the ith lower star.

Since $\tilde{g}$ is defined by linear interpolation on the vertices of K, the set of its homological critical values must be a subset of the values of g at the vertices. Thus, for any a such that $\tilde{g}(v_i) < a <, \tilde{g}(v_{i+1})$, the sublevel set $\tilde{g}^{-1}((-\infty, a])$ deformation retracts onto K_i.

Hence, appealing to Prop. 1 above, we see that the persistence diagrams computed from the lower star filtration of K will be identical to the diagrams $Dgm_r(\tilde{g})$ coming from the tame function on $|K|$.

Now if f is a tame function on the topological space X, we can choose a triangulation K of X and piecewise linear approximation of f on K. Using the lower star filtration above, we can compute the persistence diagrams for this approximation. The Stability Theorem, coupled with the Simplicial Approximation Theorem ([30]), tells us that we can thus get as precise an approximation of $Dgm(f)$ as we wish.

2.2.2 Distance Functions

Let X be a topological space embedded in $\mathbb{R}^d$ and consider the distance function $d_X : \mathbb{R}^d \to \mathbb{R}$ defined by $d_X(y) = \inf_{x \in X} ||x - y||$. The sublevel sets $\mathbb{R}^d_{\leq \alpha}$, which we abbreviate by X_α, should be thought of as X "thickened by α". They provide a filtration of the ambient space which leads to the persistence diagrams $Dgm_r(d_X)$.

.

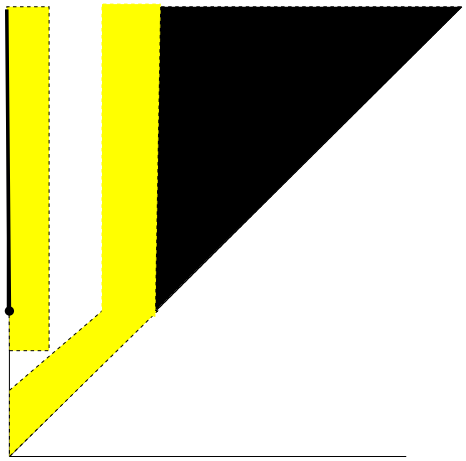

FIGURE 2.4: The persistence diagram for d_X is contained within the black regions, with the "true" homology of X corresponding to the points on the black stick at left. The expanded region, with new territory marked in yellow, contains all possible points in $Dgm(d_P)$. If the sampling is dense enough, then there will be two distinct yellow regions separated by white space; we can infer the "true" homology of X from the points in the leftmost yellow region.

We define the *homological feature size* $hfs(X)$ of X to be the smallest nonzero homological critical value of d_X. A finite set of points P in $\mathbb{R}^d$ will be called an *ϵ-sample* of X if X and P are at most epsilon apart in Hausdorff distance, or equivalently, if $||d_X - d_P|| < \epsilon$.

With these definitions, we can state the Homology Inference Theorem, which was proven in [11] as a corollary to the stability theorem:

3 (Homology Inference Theorem)**.** *If $\epsilon < \frac{hfs(X)}{4}$, then for all sufficiently small δ, the ranks of $H_r(X_\delta)$ and $Im(H_r(P_\epsilon) \to H_r(P_{3\epsilon}))$ are either both infinite or are the same.*

Fig 2.4 illustrates this result in terms of persistence diagrams. Briefly, if P is a fine enough sample of X, then the actual homology of X can be inferred from $Dgm_r(d_P)$; this latter diagram can be computed via the *alpha-shape* filtration([14]) of the *Delaunay* triangulation of P.

2.3 Extended Persistence

Let X be a topological space along with two filtrations F, G by closed subspaces:

$$\varnothing = X_0 \subseteq X_1 \subseteq \ldots \subseteq X_n = X \tag{2.7}$$

and

$$\varnothing = Y_0 \subseteq Y_1 \subseteq \ldots \subseteq Y_n = X. \tag{2.8}$$

For reasons that become apparent later, we call the first filtration the *ascending* filtration and the latter the *descending* one. Fixing a nonnegative integer r, we recall the earlier shorthand notation $H_r^i = H_r(X_i)$ and extend this notation to define $H_r^{n+i} = H_r(X, Y_i)$.

For $i < j$, the inclusion $Y_i \hookrightarrow Y_j$ induces a map on relative homology groups $H_r(X, Y_i) \to H_r(X, Y_j)$. So we obtain an extended sequence, beginning and ending with the zero group, of homology groups and homomorphisms:

$$0 \to H_r^1 \to H_r^2 \to \ldots \to H_r^n = H_r(K) \to H_r^{n+1} \to \ldots \to 0 \tag{2.9}$$

The notions of birth and death of homology classes along this extended sequence are then defined exactly as in Sec. 2.1. In the same manner, we extend the definition of the pair group $P_r^{i,j}$ to allow $1 \leq i < j \leq 2n$. If we need to specify the particular space along with the two filtrations, we write $P_r^{i,j}(X, F, G)$.

These pair groups come in three distinct types. If $j < n$, the *ordinary* pair group $P_r^{i,j}$ contains classes which appear and disappear during the ascending filtration.

For $i > n$, we have the *relative* pair groups. In [10], the authors prove a relationship between these groups and the pair groups $P_r^{i,j}(X, G)$ for the descending sequence (Eqn 2.8). Namely, by using the long exact sequences of the pairs (X, Y_i), they prove:

$$P_r^{n+i,n+j}(X, F, G) \cong P_{r-1}^{i,j}(X, G), \quad , 1 \leq i < j \leq n \tag{2.10}$$

In other words, a relative pair in the extended sequence corresponds to the appearance and disappearance of an inessential class, of one lower dimension, along the descending filtration.

Finally, if $i \leq n < j$, we have an *extended* pair group. In this case, an element $\alpha \in P_r^{i,j}(X, F, G)$ corresponds to an essential r-class of X which is born at the i*th* level during the ascending filtration. During the descending filtration, this same class is born at the jth level. Reversing this argument, there is then a relationship between the extended pairs obtained by first ascending and then descending, and those computed by swapping the order of the filtrations:

$$P_r^{i,n+j}(X, F, G) \cong P_r^{j,n+i}(X, G, F) \;\; 1 \leq i < j \leq n. \tag{2.11}$$

As in Sec. 2.1, these same definitions can be made in the context of a simplicial complex along with two filtrations by subcomplexes. In the special case that these two filtrations arise from adding just one simplex at a time, then we may pair the simplices which create and destroy a class, exactly as before.

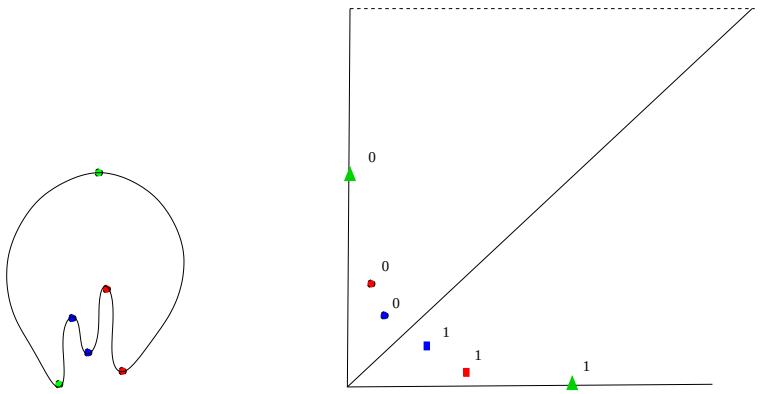

FIGURE 2.5: Here we have overlaid the diagrams $Dgm_0(f)$ and $Dgm_1(f)$. Circles correspond to ordinary points, boxes to relative, and triangles to extended. Note the obvious symmetry in the diagram.

2.3.1 Tame Functions

Let f be a real-valued tame function on X, with homological critical values $t_1 < t_2 < \ldots < t_n$, and interleaved regular values s_i defined as before. The sublevel lets $X_i = X_{\leq s_i}$, with associated value t_i, give an ascending filtration of X, while the superlevel sets $Y_{n-i} = X^{\geq s_i} = f^{-1}([s_i, \infty))$, with value t_{i+1}, form the descending filtration.

By $Dgm_r(f)$, we now mean the r-persistence diagram for the full extended sequence given by these two filtrations. The Stability Theorem (2) extends in the obvious way ([10]).

The ordinary subdiagram $Ord_r(f)$ consists of points corresponding to nonzero ordinary pair groups; these points all lie above the major diagonal. The relative subdiagram $Rel_r(f)$ lies below the major diagonal, while the extended subdiagram $Ext_r(f)$ can lie on either side.

This is illustrated in Fig. 2.5, where f is the vertical height function from Fig. 2.2.

Duality and Diagram Symmetries Consider the special case where X is a d-manifold equipped with a smooth tame real-valued function f. Then, for each i, the sublevel set X_i is a d-manifold with boundary. And so Lefschetz Duality provides a perfect pairing, for each $0 \leq r \leq d$, between $H_r(X_i)$ and $H_{d-r}(X_i, \partial X_i)$.

But $\partial X_i = \partial Y_{n-i}$. So by excision, we in fact have a perfect pairing:

$$H_r^i \otimes H_{d-r}^{2n-i} \to \mathbb{Z}/2\mathbb{Z} \tag{2.12}$$

Furthermore ([10]), these pairings are compatible with the maps on absolute and relative homology induced by the inclusions among the subspaces in the ascending and descending filtrations. This leads to:

4 (Pair Group Duality). *For $1 \leq i < j \leq 2n$, $0 \leq r \leq d$, there is a perfect pairing*

$$P_r^{i,j} \otimes P_{d-r}^{2n-j+1,2n-i+1} \to \mathbb{Z}/2\mathbb{Z}$$

This result can also be phrased in terms of the subdiagrams discussed above. Let ^{T} mean reflection in the major diagonal; that is, $(x,y)^T = (y,x)$. Then the subdiagrams are related by $Ord_r(f) = [Rel_{d-r}(f)]^T$ and $Ext_r(f) = [Ext_{d-r}(f)]^T$. See Fig. 2.5, for example.

On the other hand, consider the negative function $-f : X \to \mathbb{R}$. Pair Group Duality, coupled with Eqn. 2.10 and Eqn. 2.11, leads to the following relationships between the subdiagrams of $Dgm(f)$ and $Dgm(-f)$:

- $Ord_r(f) = [Rel_{d-r-1}(-f)]^R$
- $Rel_r(f) = [Ord_{d-r+1}(-f)]^R$
- $Ext_r(f) = [Ext_{d-r}(-f)]^N$,

where the superscripts R and N indicate the involutions of the plane defined by $(x,y)^R = (-y,-x)$ and $(x,y)^N = (-x,-y)$.

In Chap. 4, we prove similar pair group duality and diagram symmetry results for persistent intersection homology.

Simplicial Analogue Consider a simplicial complex K along with its n vertices ordered via an injective real-valued function g; as before, we extend g by linear interpolation to $\tilde{g} : |K| \to \mathbb{R}$. Also as before, we have the lower star filtration $\{K_i\}$ of K, which we now consider to be the ascending filtration.

For $1 \leq i \leq n$ let L_{n-i} be the full subcomplex spanned by the youngest i vertices $v_{i+1}, v_{i+2}, \ldots, v_n$; we let $g(v_{i+1})$ be the value of L_{n-i}. The nested sequence of these subcomplexes defines the upper star filtration of K, which we take as the descending filtration.

Note that for any value a such that $g(v_{i-1}) < a < g(v_i)$, the superlevel set $g^{-1}([a,\infty))$ deformation retracts onto L_{n-i}. Hence, $Dgm(\tilde{g})$ will be the same as that for the extended filtration of the simplicial complex K.

In the case that K triangulates a d-manifold, the subcomplexes K_i and L_{n-i} will each be homotopically equivalent to d-manifolds with a shared boundary, and so the above duality and symmetry results apply.

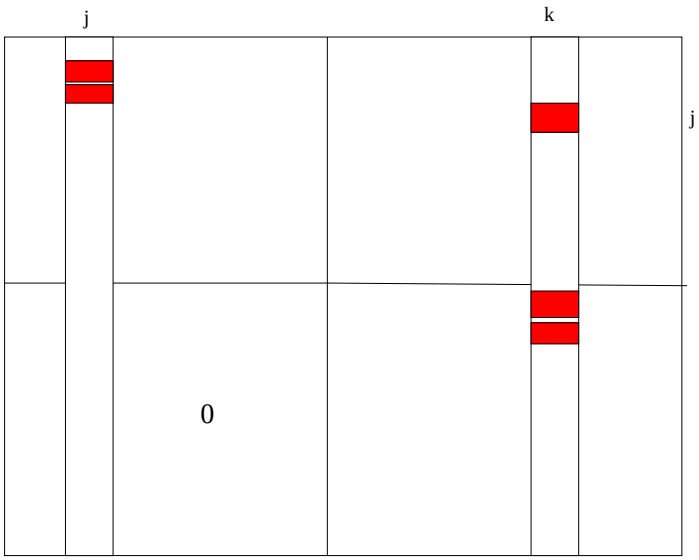

FIGURE 2.6: The leftmost column stores the boundary of the edge σ_j. The lower two entries in the rightmost column indicate the boundary of the edge τ_k, while the first entry in the rightmost column tells us that $\sigma_j = \tau_k$.

2.3.2 Extended Persistence Algorithm

Here we describe the algorithm for computing extended persistence when we have a simplicial complex K along with two sets of orderings on its simplices.

Suppose that K has n simplices, and that $\sigma_1, \ldots, \sigma_n$ is an ordering which defines the ascending filtration, while the descending filtration is given by the ordering $\tau_1, \ldots, \tau_n$.

We define the $2n \times 2n$ binary matrix M in 4 blocks of equal size as follows (see Fig. 2.6).

The upper left block simply stores the boundary matrix for the ascending filtration: for $1 \leq i \leq j \leq n$, $M[i, j] = 1$ iff σ_i is a boundary face of σ_j. The lower left block is, and will remain, all zeroes.

The lower right block is the boundary matrix for the descending filtration, with indices adjusted upwards by n; in other words, $M[n+i, n+j] = 1$ iff τ_i is a boundary face of τ_j.

Finally, the upper right block is a permutation matrix which encodes the relationship between the two filtrations. $M[i, n + j] = 1$ iff $\sigma_i = \tau_j$.

The matrix M is then reduced exactly as in SubSec. 2.1.3, producing a reduced matrix R. The associated lowest-one function, low_R, gives the ordinary, relative, and extended pairings, as follows. Suppose $low_R(j) = i$.

- $j \leq n$: σ_j is paired with σ_i in an ordinary pair.
- $i > n$: τ_{j-n} is paired with τ_{i-n} in a relative pair.

- $i \leq n < j$: this gives the extended pair (σ_i, τ_{j-n}).

3

Intersection Homology

3.1 Introduction

This chapter gives an detailed exposition of manifold stratified spaces and intersection homology as a prelude to the definitions of persistent intersection homology in the next chapter. Before we delve into any great detail, it seems wise to give a brief historical overview of the subject and discuss the context into which our version fits.

Historical Discussion In general, a manifold stratified space X is a space that, while it may not itself be a manifold, is decomposable into manifold pieces of different dimensions. Often one makes further requirements to ensure that these pieces fit together in some nice uniform fashion. There are several different versions of these requirements; for an extensive survey, see [26]. Whatever the precise definition, these spaces do not in general satisfy Poincaré Duality for ordinary homology.

The original treatment of intersection homology was given by Goresky and Macpherson in [22]. Initially, their intersection homology groups were defined only for a PL-pseudomanifold X, which was assumed to have no stratum of co-dimension one. The basic idea was that they started with the group of geometric chains on X, and then defined a series of subgroups corresponding to chains whose dimension of intersection with the singular strata of X did not exceed certain limits. These limits were given by perversities $\bar{p}$, which were sequences of integers $(p_2, p_3, \ldots, p_d)$ subject to the condition that $p_2 = 0$ and $p_i \leq p_{i+1} \leq p_i + 1$. Using this last condition, the authors proved that their intersection homology groups were topologically invariant; in particular, they were independent of the particular choice of stratification for the PL-pseudomanifold. For the groups corresponding to dual perversities, they defined a perfect intersection pairing which restored an analogue of Poincaré Duality to these spaces. In [22], the authors exploited this duality to define a signature on singular

spaces with even codimensional strata. Many further applications of the theory have since been made; for a discussion, see [4].

In a later paper ([23]), the same authors used sheaf theory to extend the definition of their groups, and the topological invariance thereof, to a wider class of spaces. In [27], King gave a singular chain version of intersection homology and used this version to provide another proof of invariance; this version later turned out to be isomorphic to the sheaf formulation. Finally, the authors in [25] extended the topological invariance result to the broader category of locally-conelike topological spaces.

Our Context We wish to apply intersection homology theory in at least two, and possibly many more, distinct contexts. The work done in [6], and described in detail in Chap. 6, begins a program with the aim to provide a series of approximate stratifications of a point cloud in some high-dimensional Euclidean space. Eventually, we envision using intersection homology theory to assess these stratifications and perhaps measure the difference between them. In this context, we most certainly do not desire stratification independence. For this reason, we drop the requirement that $p_i \leq p_{i+1} \leq p_i + 1$.

However, we also foresee applications of persistent intersection homology to spaces where the stratification is already clear. In Chap. 5, we begin the definition of a series of intersection homology elevation functions on an embedded stratified space; the critical points of these function help locate interesting features of the space and illuminate the relationship between its singularities. We envision many applications of these functions. For example, the PL-version of the protein-protein interface surface, as defined and studied in [3], is a two-dimensional stratified space embedded in Euclidean space; the singular strata correspond to interfaces between three or more molecules.

For both of these application types, we require that everything be actually computable; hence we wish to work entirely within a simplicial theory. Fortunately, the original geometric chain formulation in [22] can trivially be recast as a simplicial theory, as long as the triangulation is 'flaglike' [28].

A more serious issue concerns the fact that we do not wish to impose any precondition on the possible codimensions of strata for our spaces; in particular, we allow strata of codimension one. On the other hand, we wish to preserve Poincaré Duality so that our intersection homology elevation functions will enjoy the same symmetry properties that ordinary homology elevation does for manifolds. In [19], Friedman defines a sheaf-theoretic version of intersection homology which preserves Poincaré Duality, as well as giving something very close to stratification independence, for a wide class of stratified spaces which can have codimension-one strata.

We take a far more prosaic approach. Letting Σ be the singular set of X, we simply work within the relative chain group $C_*(X, \Sigma)$ and define our groups using perversities $\bar{p} = (p_1, p_2, \ldots, p_d)$, with no restrictions put on the values. It seems quite

likely that our definition can be shown to be equivalent to Friedman's in our far more restricted context. Nonetheless, at the end of this chapter we provide a direct proof that our version satisfies a suitably redefined statement of Poincaré Duality for dual perversities. This proof uses formulae concerning the intersection homology of a cone to do a local computation and then patches together the result via a Mayer-Vietorris argument; the inspiration comes entirely from a suggestion in [28] for a direct proof technique in the pseudomanifold case.

3.2 Stratified Spaces

In this section we review the definition of a topologically stratified space. In addition, we give several examples of these objects and discuss the fact that they do not in general satisfy Poincaré Duality.

Definition A *d-dimensional topologically stratified space* is a topological space $X \subset \mathbb{R}^n$ together with a chain of closed subsets:

$$X = X_d \supseteq X_{d-1} \supseteq X_{d-2} \supseteq \dots X_1 \supseteq X_0 \supseteq X_{-1} = \varnothing$$

so that $X_d - X_{d-1}$ is dense in X and so that the following condition is satisfied: For each $x \in X_i - X_{i-1}$ there is a stratified space V

$$V = V_d \supseteq \dots \supseteq V_i = \{point\}$$

where $V_k - V_{k-1}$ has dimension $k - i$, and a map

$$\phi_x : B^i \times V_d \to X$$

such that $B^i \times V_k$ maps PL-homeomorphically onto a closed neighborhood of x in X_k, for all $k \geq i$. Here B^i is a closed i-dimensional PL-ball.

A few remarks on this definition may help to clarify.

1. By taking $k = i$ in the above condition, we see that $X_i - X_{i-1}$ must be an i-manifold. We denote this subspace S_i and call it the ith stratum of X. The connected components of the strata are called pieces. The union of lower strata X_{d-1} is also called Σ, the "singular set" of X.

2. The existence of ϕ in the above definition is often referred to as "local normal triviality"; indeed, the space V in the above condition may be thought of as a sort of "normal slice" at $x \in S_i$. To make this more precise, let N be a subspace of X which is transverse to each stratum and intersects S_i in the single point x and let B_δ be a small enough ball in X centered at x. Then V will be homeomorphic to $N \cap B_\delta$, which we denote N_x. One can show([24]) that the homeomorphism type of the normal slice N_x depends neither on the choice of sufficiently small δ nor of N, nor indeed on the choice of x within a

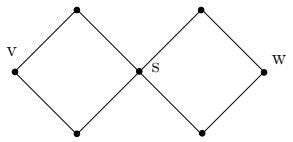

FIGURE 3.1: a 1-dimensional stratified simplicial complex

particular piece of S_i. Hence the pieces, themselves manifolds, fit uniformly into the larger space. An example of this construction is shown in Fig. 3.2.

Note: The above definition concerns a space X *along* with its stratification by the sets X_i. There are of course many different ways to stratify a given space. The intersection homology groups defined later will in general be sensitive to this choice.

If X has a closed subset ∂X such that each S_i is an i-manifold with (possibly empty) boundary $\partial S_i = S_i \cap \partial X$, then we call X a topologically stratified space with boundary ∂X.

By a triangulation of a stratified space, we will mean a simplicial complex K which triangulates a topologically stratified space X in such a way that all X_i are triangulated by subcomplexes of K.

Finally, by a *stratified subspace* Y of X, we will mean a closed subspace $Y \subseteq X$ which is itself a stratified space under the stratification inherited from that of X. This means that the ith strata of Y is $S_i \cap Y$. One way to ensure this is to demand that Y intersect each S_i transversely. For instance, the normal slice at a point is a stratified subspace. We now give several examples of stratified spaces.

Example: The wedge of two circles. If L triangulates a d-manifold, then L is trivially a stratified simplicial complex with only one nonempty stratum. Perhaps the simplest possible non-trivial example of a stratified simplicial complex is shown in Fig. 3.1.

Here $X_1 = K$, $X_0 = \{s\}$. Note that S_1 consists of two open arcs, which is a disconnected 1-manifold, while S_0 is simply the point s, which is certainly a 0-manifold. Homologically, K consists of one component and two 1-cycles. Since the Betti numbers of complementary dimension are not equal, Poincaré Duality cannot hold.

It is instructive to see more explicitly why the intersection pairing is not even well-defined for this space. Let A represent the left circle. Then A intersects v in precisely one point, v itself. On the other hand, v is homologous to w, which is not contained in A, nor in any circle or sum of circles which is homologous to A. Hence the intersection of the homology class of A and the homology class of v cannot be well-defined. Note, however, that any path connecting v and w must intersect the wedge point s, which is in a singular stratum. If we were to simply forbid any such path, then v and w would no longer be homologous and the problem outlined above could not occur. It is this intuition which informs the definition of the intersection

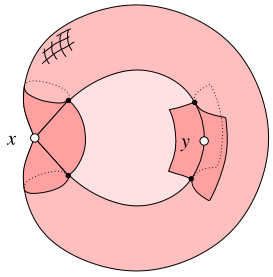

FIGURE 3.2: A stratified space with more than one singular stratum: the normal slices at $p \in S_0$ and at $y \in S_1$ are highlighted.

homology groups, although the situation gets considerably more involved for spaces with more complicated singularities.

Example: Pinched Torus with Disc Our next example Y is shown in Fig. 3.2. It is a torus which has had one of its boundary circles pinched to a point (which we'll call p) and then had a disc stretched tight across the hole. Let us call the remaining circle C.

If we remove C from Y, we obtain a disconnected 2-manifold which forms S_2. Note that C itself is a one-manifold. However, not all points on C are singularities of the same kind. If $y \in C, y \neq p$, then y has a neighborhood homeomorphic to three sheets glued together along a line; in terms of the definition, this neighborhood is the product of a 1-ball in C and a cone on three points, one from the disc and two from the torus. On the other hand, p has no such neighborhood; in fact all of its sufficiently small neighborhoods consist of a cone on two circles (in the torus) joined by a line (in the disc). Hence the "local normal triviality" condition demands that we place p in its own individual stratum, which leads us to the following stratification of X:

$$Y = X_2 \supseteq X_1 = C \supseteq X_0 = \{p\}$$

.

Example: Suspended Torus Our final example is a stratified space with more complicated homology. Let ΣT denote the suspended torus, defined to be the result of collapsing each end of the product $T \times [-1, 1]$ to a point. This space does not embed in $\mathbb{R}^3$, so we picture it in $\mathbb{R}^4$ as the union of two cones. The middle section $T \times 0$ is the usual embedding of the torus in $\mathbb{R}^3$. The cone points a and b are the points $(0, 0, 0, \pm 1)$. The cones are then the collection of straight line segments in $\mathbb{R}^4$ from the torus to the cone points.

ΣT is a three-dimensional stratified space, with stratification:

$$\Sigma T = X^3 \supseteq X^2 = X^1 = X^0 = \{a, b\}$$

.

Since ΣT is connected, we have $\beta_0 = 1$. Now T itself had two non-bounding one-cycles, represented by the two circles C_1, C_2. Within ΣT, these cycles become boundaries: for example, C_1 bounds the cone $C_1 * a$. Thus, $\beta_1 = 0$. On the other hand, we also see that $C_1 = \partial(C_1 * b)$. Thus, we obtain a 2-cycle, represented by $C_1 * a + C_1 * b$, which we will denote by ΣC_1. Similarly, ΣC_2 is a 2-cycle, and we find $\beta_2 = 2$. Finally, $\beta_3 = 1$, a three-cycle formed by suspending the fundamental 2-cycle of T. Since the Betti numbers in complementary dimension are not equal, this object cannot satisfy Poincaré Duality. However, as was the case with the wedge of two circles, all of the trouble was caused by chains (like $C_1 * a$) which intersected the singular set.

Stratum-Preserving Homotopies Let X, Y be two stratified spaces, of dimensions d, m, respectively. A map $f : X \to Y$ is said to be stratum-preserving ([18]) if the image under f of each stratum of X is contained within the stratum of the same codimension in Y; in other words, if $f(X_{d-k} - X_{d-k-1}) \subseteq Y_{m-k} - Y_{m-k-1}$.

A stratum-preserving homotopy $F : X \times I \to Y$ is then just a stratum-preserving map, where one stratifies $X \times I$ by $(X \times I)_k = X_k \times I$. Finally, a map $f : X \to Y$ is called a stratum-preserving homotopy equivalence if there exists a map $g : Y \to X$ such that gf and fg are homotopic to the identity via a stratum-preserving homotopy.

3.3 Intersection Homology

In this section, we define the intersection homology groups, with $\mathbb{Z}/2\mathbb{Z}$-coefficients, for a stratified simplicial complex. For the reasons explained in the introduction, our definition will be slightly different from the one originally given in [22] by Goresky & MacPherson.

3.3.1 Definition

Perversities A *perversity* is a sequence of integers $\bar{p} = (p_1, p_2, \ldots p_d)$. Let $\bar{t} = (-1, 0, 1, \ldots, d-2)$ Two perversities $\bar{p}$ and $\bar{q}$ will be called *dual* if they add to $\bar{t}$. For example, if $d = 2$, then $(-1, 0)$ and $(0, 0)$ are dual perversities. We use these perversities to provide a measure of how much intersection between simplices and lower-dimensional strata we will accept.

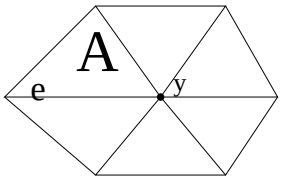

FIGURE 3.3: Triangle A is proper, but its boundary edge e is not.

Proper Simplices An i-simplex σ in K is said to be $\bar{p}$-proper if the following condition holds for all $k = 1, \ldots, d$:

$$\dim \bar{\sigma} \cap X_{d-k} \leq i - k + p_k$$

where $\bar{\sigma}$ denotes the closure of the open simplex σ. Here we are intentionally confusing σ and $\bar{\sigma}$ with their underlying topological space.

The intuition behind this inequality is as follows: if an i-dimensional subspace intersects a codim-k subspace *transversely*, the dimension of the intersection will be $i - k$. A non-transverse intersection will result in a higher dimension. Thus, if $p_k = 0$, we are requiring that for σ to be proper, $\bar{\sigma}$ must intersect the codim-k stratum transversely. Higher values of p_k give more tolerant intersection conditions.

To prove the Duality results in the presence of a codimension-one stratum, we will need to work within the relative chain group $C_i(K, \Sigma) = C_i(K)/C_i(\Sigma)$. Thus, an i-chain ξ will be a sum of i-simplices which do not lie entirely within Σ; furthermore, the boundary $\partial\xi$ of this i-chain will be the sum of those $(i-1)$-simplices in the boundary of ξ which also do not lie entirely within Σ.

Allowable Chains To define a homology theory, we need a sequence of groups in each dimension and a sequence of boundary maps ∂_i between them that satisfy the condition $\partial_i \circ \partial_{i+1} = 0$. In ordinary homology, one defines $C_i(K)$ to be the $\mathbb{Z}/2\mathbb{Z}$-vector space with the i-simplices of K as basis. The obvious boundary map ∂_i has the useful property that it takes a sum of i-simplices to a sum of $i-1$ simplices; in other words, it gives a well-defined homomorphism $\partial_i : C_i K \to C_{i-1}(K)$.

The most natural construction then, given a perversity $\bar{p}$, would be to define a 'good' i-chain to be a sum of proper i-simplices; that is, to let the ith chain group take the proper i-simplices as basis. Unfortunately, this cannot work: our boundary map would be ill-defined, as there is no guarantee that the boundary of a proper i-simplex will be the sum of proper $(i-1)$ simplices.

This problem is illustrated in Fig. 3.3, where all of the triangles are filled in.

Suppose that y is a singular point of codimension two and that $\bar{p} = (-1, 0)$. Then triangle A is itself a proper 2-simplex. On the other hand, ∂A is *not* a sum of proper 1-simplices: for example, the edge e is not a proper 1-simplex. Thus A by itself cannot be part of our 2-dimensional chain group. On the other hand, the sum of all

the triangles drawn could be: it is a sum of proper 2-simplices whose boundary is itself a sum of proper 1-simplices.

This picture illustrates the general definition: we say that a chain $\xi \in C_i(K)/C_i(\Sigma)$ is $\bar{p}$-allowable if ξ is a sum of proper i-simplices **and** if $\partial\xi$ is a sum of proper $(i-1)$-simplices. In the above example, A would not be allowable, but the sum of all the triangles would be. The $\mathbb{Z}/2\mathbb{Z}$-vector space of allowable i-chains for perversity $\bar{p}$ is denoted $I^pC_i(K)$.

Now suppose ξ is an allowable i-chain. Note that $\partial\partial\xi = 0$. In particular, $\partial\partial\xi$ may (trivially) be written as a sum of proper $(i-2)$-simplices. In other words, $\partial\xi$ is itself an allowable $(i-1)$-chain; therefore, the boundary maps ∂_i give a sequence of well-defined homomorphisms $\partial_i : I^pC_i(K) \to I^pC_{i-1}(K)$; we also have $\partial_i \circ \partial_{i+1} = 0$, which means that we have a chain complex. The intersection homology groups are then just the homology of this chain complex.

Intersection Homology Groups Define $I^pH_i(K)$, the ith intersection homology group, with perversity $\bar{p}$, of K, to be the kernel of the map $\partial{:}I^pC_i(K) \to I^pC_{i-1}(K)$ modulo the image of the map $\partial{:}I^pC_{i+1}(K) \to I^pC_i(K)$. This is the $\mathbb{Z}/2\mathbb{Z}$-vector space with basis consisting of those *allowable* i-cycles which are not the boundary of an *allowable* $(i+1)$-chain.

Restriction on Perversities With these definitions, we can see that the restriction to perversities satisfying $-1 \leq p_k \leq k-1$ for all k will not actually reduce the possibilities for intersection homology groups. To see this, observe that $p_k = k$ simply allows a proper i-simplex σ to lie entirely within X_k: for $i \leq i-k+k$. But we are computing mod Σ and thus such a simplex cannot affect the calculation in any case. On the other side of the inequality, note that any $p_k < 0$ will give equivalent results to $p_k = -1$.

Proper/Improper Decision Procedures Suppose that the symbol $\bar{q}$ does not stand for a perversity, but instead is simply a procedure which decides, for a given simplex σ, whether or not σ is to be considered "proper." Via the exact same algebraic machinery as above, one can define the "intersection homology groups for procedure $\bar{q}$." Of course, these groups need not satisfy Duality nor indeed have any relation to topological properties of any particular space. Nonetheless, in the next chapter, we provide an algorithm to compute the intersection homology pair groups for a simplicial complex equipped with any such procedure; this general setting will prove useful in the proof of correctness for the algorithm.

Singular Intersection Homology If X is a stratified space, we may wish to make reference to its intersection homology without bothering to triangulate it. This can be done by considering its singular intersection homology groups ([27]), defined as follows.

Let $\sigma \in S_i(X)$. In other words, $\sigma : \Delta^i \to X$ is a map from the standard i-simplex into the stratified space. One says that σ is $\bar{p}$-proper if $\sigma^{-1}(X_{d-k})$ is contained within the $i-k+p_k$ skeleton of Δ^i, for each k, and then proceeds exactly as above to define the singular intersection homology groups.

Relative Intersection Homology Suppose K is a stratified complex with strata $\{S_i\}$ and that Y is a subcomplex of K which intersects each S_i transversely. The intersections $Y \cap S_i$ stratify Y in a natural way. The assumption of transversality guarantees that the codimension of $Y \cap S_i$ in Y will be equal to the codimension of S_i in K. This means that we have an inclusion $I^pC_i(Y) \hookrightarrow I^pC_i(K)$. We can then define $I^pH_r(K, Y)$ to be the rth homology of the chain complex $\frac{I^pC_*(K)}{I^pC_*(Y)}$.

Duality Proofs for the following two theorems in our context will be given at the end of this chapter.

5 (Poincaré Duality)**.** *If K is a stratified complex of dimension d and if $\bar{p}$ and $\bar{q}$ are dual perversities, then there is a well-defined and perfect pairing:*

$$I^pH_r(K) \otimes I^qH_{d-r}(K) \to \mathbb{Z}/2\mathbb{Z}$$

This pairing is given by intersecting representatives of the respective homology classes and counting the number mod 2 of points in their intersection.

6 (Lefschetz Duality)**.** *Let K be a stratified complex of dimension d with a boundary ∂K which is transverse to the strata of K. Then for dual perversities $\bar{p}$ and $\bar{q}$ there is a well-defined and perfect intersection pairing:*

$$I^pH_r(K) \otimes I^qH_{d-r}(K, \partial K) \to \mathbb{Z}/2\mathbb{Z}$$

Note that the former claim is clearly implied by the latter, by taking $\partial X = \varnothing$.

Example: Wedge of two circles Consider the wedge K of two circles (Fig. 3.1). and the two perversities $\bar{p} = (-1)$ and $\bar{q} = (0)$. Notice that these perversities are dual.

For $\bar{p}$, any edge e which contains s in its closure will be improper: $\dim(\bar{e} \cap X_0) = 0 > 1-1+p_1 = -1$. Thus, any one-chain which connects v to w cannot be allowable: this shows that $I^pH_0(K) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis the components represented by v and by w. By the same logic, neither of the two loops in K are allowable: $I^pH_1(K) = 0$.

For $\bar{q}$, all edges are proper: if e contains s in its closure, then $\dim(\bar{e} \cap X_0) = 0 \leq 1 - 1 + q_1 = 0$. Since v may be connected to s by an allowable chain of edges, and since we are computing boundaries mod $\Sigma = \{s\}$, we find that v is itself a boundary. Similarly, w is a boundary. Hence $I^qH_0(K) = 0$. On the other hand, since all edges are proper, we regain the two loops: $I^qH_1(K) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis the homology class of the left loop A and of the right loop B.

Note that A contains v. Crucially, A also contains any point which is $\bar{p}$-homologous to v; similarly with B and w. Thus, we may unambiguously define a perfect intersection pairing between these two groups, which are of complementary dimension and defined using dual perversities:

$$I^pH_0(K) \otimes I^qH_1(K) \to \mathbb{Z}/2\mathbb{Z}$$

Example: Suspended Torus Next we calculate the intersection homology groups of the suspended torus using the two dual perversities $\bar{p} = (-1, 0, 0)$ and $\bar{q} = (0, 0, 1)$. Any edge whose closure contains the codim-three singularity a (or b) cannot be proper for either perversity, since we require $\dim(\bar{e} \cap X_0) \leq 1-3+q_1 = -1$. Thus, no single point in ΣT is a boundary. On the other hand, any two vertices in the smooth part of ΣT can be connected via a path which entirely avoids the two singular points. Hence, $I^pH_0(\Sigma T) = I^qH_0(\Sigma T) = \mathbb{Z}/2\mathbb{Z}$.

The sum ξ of all three-simplices in any triangulation of ΣT necessarily contains the singular points. If σ is one such three-simplex, then from the computation $\dim(\bar{\sigma} \cap X_0) = 0 \leq 3 - 3 + p_3 = 0$, we see that ξ is a sum of proper simplices. Since $\partial\xi = 0$ and thus trivially a sum of proper simplices, ξ is allowable. Hence we have: $I^pH_3(\Sigma T) = I^qH_3(\Sigma T) = \mathbb{Z}/2\mathbb{Z}$.

In dimensions 1 and 2, the two perversities give different answers. For $\bar{p}$, the 2-simplices which we obtain by coning the circles of the torus to either one of the singular points are not proper: for example, $\dim(C_1 * a \cap X_3) = 0 > 2 - 3 + p_3 = -1$. Hence the boundary circles C_1 and C_2 are allowable 1-cycles which are not the boundary of an *allowable* 2-chain, from which we see that $I^pH_1(\Sigma T) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis elements the homology classes of C_1 and of C_2. On the other hand, $I^pH_2(\Sigma T) = 0$.

Replacing $p_3 = 0$ with $q_3 = 1$ in the above discussion shows that $I^qH_1(\Sigma T) = 0$, while $I^qH_2(\Sigma T) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis elements the homology classes of ΣC_1 and ΣC_2. Note that ΣC_1 and C_2 intersect in precisely one point, as do ΣC_2 and C_1. So once again intersection gives a perfect pairing $I^pH_1(\Sigma T) \otimes I^qH_2(\Sigma T) \to \mathbb{Z}/2\mathbb{Z}$.

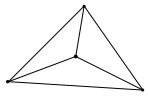

FIGURE 3.4: A 'bad' triangulation of the stratified sphere.

3.3.2 Invariances and Induced Maps

In this section we address the dependence of these intersection homology groups on the choice of stratification and triangulation of a stratified space. In addition, we discuss the types of homotopy under which intersection homology is invariant.

Dependence on Stratification The definition of a stratified space in Sec. 3.2 includes both a space X and the filtration of X by closed subsets which leads to the stratification. Now we may of course stratify a given space in different ways. There are certain assumptions ([22]) which can be made that guarantee independence of intersection homology from this choice; for example, one might require that $S_{d-1} = \varnothing$ and that $p_i \leq p_{i+1} \leq p_i + 1$. However, in our general context, the intersection homology groups will in fact be dependent on choice of stratification.

As an example, consider again the wedge of two circles (Fig 3.1), but this time we take $X_0 = \{v, s\}$. We find in this case that $I^{\bar{p}}H_0(K) \cong H_0(K - \{s, v\}) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$. A change to the stratification produced a change in the intersection homology groups: two components became three.

Dependence on Triangulation Recall that the ordinary homology groups of a space do not depend on its triangulation; in particular, they are unchanged when we subdivide. This does *not* hold for intersection homology: two different triangulations of the same stratified space may well give different groups. Here is an example.

Suppose we take a 2-sphere with codimension-2 stratum consisting of 4 points. Further suppose we triangulate this stratified sphere as the boundary of a tetrahedron with the four singular points as vertices (Fig. 3.4).

Let $\bar{p} = (-1, 0)$. The $0th$-dimensional simplicial intersection homology group for this stratified space, using our chosen triangulation, will be trivial: there are no proper vertices at all, and thus no allowable components.

On the other hand, suppose we subdivide one of the edges. This creates one proper vertex x. Since $p_2 = 0$, the edges connecting x to any of the singular vertices are not proper. Thus, x represents a nontrivial component. The subdivision changed the zero-dimensional intersection homology.

Flaglike Triangulations This above problem is easily fixed by a simple requirement on the relationship between the triangulation and the stratification. A triangulation is called 'flaglike' if for every simplex $\bar{\sigma}$ and for every X_k, the intersection $\sigma \cap X_k$ is a *single* face of $\bar{\sigma}$. Note that the example above is *not* a flaglike triangulation: in fact, every edge violates the condition.

Flaglike triangulations are, in essence, the 'correct' ones to use for the computation of intersection homology groups: if K is a flaglike triangulation, then any further subdivision will give isomorphic intersection homology groups [28]. In what follows, we will therefore always assume that our spaces have flaglike triangulations. This is not that restrictive a condition since the first barycentric subdivision of *any* triangulation will already be flaglike.

Induced Maps and Homotopy Invariance Let $f : X \to Y$ be a map between stratified spaces. Unlike the case of homology, f need not in general induce maps on intersection homology; briefly, a problem arises if f maps an allowable chain in X to a non-allowable one in Y.

The additional requirement that $f : X \to Y$ be a stratum-preserving map guarantees that f will induce maps on intersection homology. Furthermore, a stratum-preserving homotopy equivalence will induce isomorphisms on all intersection homology groups ([18]).

3.4 Proof of Duality

In this section, we prove the following:

7. *Let $X = X_d \supseteq X_{d-1} \supseteq X_{d-2} \supseteq \ldots \supseteq X_0$ be a d-dimensional stratified space with boundary ∂X. Suppose that $\bar{p}$ and $\bar{q}$ are dual perversities. Then, for each j, there is a well-defined and non-degenerate pairing*

$$I^{p}H_j(X) \otimes I^{q}H_{d-j}(X, \partial X) \to \mathbb{Z}/2\mathbb{Z}$$

given by choosing representatives for the respective classes which intersect transversely and counting the number (modulo 2) of points in their intersection.

3.4.1 Proof Overview

The proof of the theorem is a bit complicated and requires several interlocking steps. We first provide a brief overview.

The definition of a stratified space is an inductive one: locally, a stratified space looks like the product of a ball with the cone on a lower-dimensional stratified space, the link. Below, we indicate several formulae which relate the intersection homology (and relative intersection homology mod the boundary) of such a product to the intersection homology of the link. This would suggest that we attempt to prove the

claim via induction on dimension: if we assume that the claim holds for the link, we may then use the formulae to prove the existence of a perfect pairing on the product.

However, a stratified space is only locally such a product. In order to prove the claim in general, we must perform a further induction on the number of such sets needed in order to cover our stratified space. More formally, we exploit the following lemma, the proof of which we will defer until the end of this section:

8. *Any d-dimensional stratified space has a cover consisting of closed sets, each of which are homeomorphic to $B^i \times c(L)$, where L is a stratified space with $dim(L) + 1 + i = d$. Furthermore, for any collection of sets $\{B^{i_k} \times c(L_k)\}$ in this cover, the intersection $\bigcap_k B^{i_k} \times c(L_k)$ is either empty or is homeomorphic to $B^{i_j} \times c(L_j)$, where $i_j = max_k i_k$.*

To complete the proof, we induct on the number of sets needed for such a "good cover" of our stratified space. The base case can be proven using the formulae alluded to above; for the inductive step, we appeal to the Mayer-Viettoris sequences for absolute and relative intersection homology.

In what follows, we will refer to the induction on dimension as the "outer" induction, while the "inner" induction is the one on the number of sets needed for a good cover.

The outer base case is immediate, since a 0-dimensional stratified space is just a 0-dimensional manifold without boundary; hence the theorem reduces in this case to the trivial statement that each point intersects itself exactly once.

3.4.2 Inner Base Case

Suppose $X = B^i \times c(L)$ and that $\bar{p}, \bar{q}$ are dual perversities. The link L is a stratified space of dimension $d - i - 1 < d$. So by the outer inductive assumption, L satisfies the theorem.

Now consider $c(L)$, which is a $(d-i)$-dimensional stratified space with a boundary, $L \times \{0\}$, that we identify with L; we let v denote the cone vertex. In the next several paragraphs, we prove two formulae which show how the intersection homology of $c(L)$ and the relative intersection homology of $(c(L), L)$ depend on the intersection homology of L; as a consequence, we can then infer that the perfect Poincaré Pairing on L induces a perfect Lefschetz Pairing on $(c(L), L)$.

It is easy to see that a collection of simplices in $c(L)$ which contains v may not form a cycle; hence any non-bounding cycle in $c(L)$ is in fact homologous to one in L. On the other hand, not every non-bounding cycle in L remains non-bounding when considered as a cycle in $c(L)$.

To see this, suppose $\gamma \in I^p H_j(c(L))$ is represented by a chain, which we also call γ, in $L \times \{0\}$. Clearly, $\gamma = \partial(c(\gamma))$, where $c(\gamma)$ is the $(j+1)$-chain formed by coning all of the simplices in γ to v. Thus, $\gamma = 0 \in I^p H_J(c(L))$ iff $c(\gamma)$ is $\bar{p}$-allowable. We now investigate when that happens and when it does not happen.

Recall that $c(L)$ is stratified in such a way that, for all $k < d-i$, the codimension k stratum of $c(L)$ is simply the cone, $c(S_{d-i-1-k})$, on the codimension k stratum, $S_{d-i-1-k}$, of L. Since $dim(c(\gamma))\cap c(S_{d-i-1-k})) = dim(\gamma \cap S_{d-i-1-k})$, the only potential obstacle to $c(\gamma)$ being $\bar{p}$-allowable is the dimension of its intersection with v, the codimension $(d-i)$-stratum of $c(L)$. Thus, $c(\gamma)$ is allowable iff $0 = dim(c(\gamma) \cap v) \leq (j+1) - (d-i) + p_{d-i}$.

Hence, some algebraic manipulation gives:

$$I^pH_j(c(L) \cong \begin{cases} I^pH_j(L) & j < d-i-p_{d-i}-1 \\ 0 & else \end{cases}$$

Note: if we worked instead with absolute chain groups, this formula would be invalid for $c(L)$ 1-dimensional and $\bar{p} = (0)$: individual vertices cannot be boundaries, unless we work within $C_*(X, \Sigma)$.

On the other hand, examining the long exact sequence in intersection homology for the pair $(c(L), L)$ gives us the following:

$$I^pH_j(c(L), L) \cong \begin{cases} I^pH_{j-1}(L) & j \geq d-i-p_{d-i} \\ 0 & else \end{cases}$$

Now suppose that $\bar{p}, \bar{q}$ are two dual perversities. There is then, for each j, a perfect pairing:

$$T_j : I^pH_j(L) \otimes I^qH_{d-i-1-j}(L) \to \mathbb{Z}/2\mathbb{Z}$$

.

First suppose that $j \geq d-i-p_{d-i}-1$. From the equation $p_{d-i} + q_{d-i} = t_{d-i} = d-i-2$, we derive the further inequality $(d-i)-j \leq d-i-q_{d-i}$. Hence, for such j, the groups $I^pH_j(c(L))$ and $I^qH_{d--i-j}(c(L), L)$ are both trivial.

Now suppose $j < d-i-p_{d-i}-1$ and let $\gamma \in I^pH_j(c(L)), \delta \in I^qH_{d-i-j}(c(L), L)$. By the logic above, γ can be thought of as belonging to $I^pH_j(L)$. On the other hand, there is a class $\delta_1 \in I^pH_{d-i-j-1}(L)$ such that $\gamma = c(\delta)$. Using these classes, we define a pairing

$$\tilde{T}_j : I^pH_j(c(L)) \otimes I^qH_{d-i-j}(L) \to \mathbb{Z}/2\mathbb{Z}$$

via the rule $\tilde{T}_j(\gamma, \delta) = T_j(\gamma, \delta_1)$. By its definition, this pairing is also perfect. Hence we know that $c(L)$ satisfies Lefschetz Duality.

Now consider the actual set in our inner base case, $B^i \times c(L)$. The ordinary and relative Künneth formulae are valid for intersection homology in the special case that one of the sets in the product is a ball([27]). Hence we see that, for any j and any perversity $\bar{p}$, $I^pH_j(B^i \times c(L)) \cong I^pH_j(c(L))$, while $I^pH_j(B^i \times c(L), \partial) \cong$

$I^pH_{j-i}(c(L), L)$. Via these isomorphisms, we can derive a perfect Lefschetz pairing on our set using the one from $(c(L), L)$. This proves the inner base case.

3.4.3 Inner Inductive Step

We now assume that the theorem is true for any d-dimensional stratified space which requires less than n closed sets in a "good cover." Suppose X is a d-dimensional stratified space, along with a good cover consisting of n sets $U_1, U_2, \ldots, U_n$, and again let $\bar{p}, \bar{q}$ be two dual perversities. Define Y to be the subspace of X formed by taking the union of the first $n-1$ closed sets and let $Z = Y \cap U_n$.

Now U_n is a d-dimensional stratified space which has a good cover consisting of 1 set, while Y, by construction, is a d-dimensional stratified space which is well-covered by $n-1$ sets. Hence, by the inner inductive assumption, there is a perfect pairing: $I^pH_i(Y) \otimes I^qH_{d-i}(Y, \partial Y) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$ and a similar pairing for U_{n+1}.

On the other hand, Z is also a d-dimensional stratified space which has a good cover with $n-1$ sets: namely, $U_1 \cap U_n, U_2 \cap U_n, \ldots, U_{n-1} \cap U_n$. So again by the inner induction we have a perfect pairing $I^pH_i(Z) \otimes I^qH_{d-i}(Z, \partial Z) \to \frac{\mathbb{Z}}{2\mathbb{Z}}$.

From these pairings, we infer duality for the larger space X by way of the following diagram, where the top row is the Mayer-Vietorris Sequence for intersection homology and the bottom row is the corresponding sequence for relative intersection homology:

$$
\begin{array}{ccccccc}
I^pH_i(Z) & \longrightarrow & I^pH_i(Y) \oplus I^pH_i(U_n) & \longrightarrow & I^pH_i(X) & \longrightarrow & I^pH_{i-1}(Z) \\
\otimes & & \otimes & & \otimes & & \otimes \\
I^qH_{d-i}(Z, \partial) & \leftarrow & I^qH_{d-i}(Y, \partial) \oplus I^qH_{d-i}(U_n, \partial) & \leftarrow & I^qH_{d-i}(X, \partial) & \longleftarrow & I^qH_{d-i+1}(Z, \partial) \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}} & & \frac{\mathbb{Z}}{2\mathbb{Z}}
\end{array}
$$

3.5 Covering Lemma

In this section, we prove the following lemma, which asserts the existence of a "good cover" for any stratified space:

Lemma: Any d-dimensional stratified space has a cover consisting of closed sets, each of which are homeomorphic to $B^i \times c(L)$, where L is a stratified space with $dim(L) + 1 + i = d$. Furthermore, for any collection of such sets $\{B^{i_k} \times c(L_k)\}$, the intersection $\bigcap_k B^{i_k} \times c(L_k)$ is either empty or is homeomorphic to $B^{i_j} \times c(L_j)$, where $i_j = \max_k i_k$.

3.5.1 Control Data

It suffices to prove the lemma for Whitney stratified spaces, since any triangulable stratified space is Whitney and vice-versa ([24]). So suppose X is a Whitney stratified space embedded in Euclidean space. Then X possesses "control data" ([31] . More specifically, for each piece S of X, there exists a tube T_S, (which is the intersection of a tubular neighborhood of S with X), a locally trivial fibre bundle $\pi_S : T_S \to S$ with fibre $c(L_S)$, and a "tubular distance function" $\rho_S : T_S \to [0, \epsilon)$. The map $\pi_S \times \rho_S : T_S - S \to S \times (0, \epsilon)$, when restricted to each piece of the domain, is a smooth submersion. Furthermore, for any pair of pieces S, R such that $S \leq R$, the following functional equalities hold on $T_S \cap T_R$

- $\pi_S \circ \pi_R = \pi_S$
- $\rho_S \circ \pi_R = \rho_S$

3.5.2 Construction of the Covering

For simplicity, we assume that all strata are connected; if this fails to hold, the covering argument can just be repeated piece-by-piece. Let S be the stratum of lowest dimension, and put $i = dim(S)$. We let M_S be the subset of T_S defined by $\rho_S^{-1}([0, \frac{3\epsilon}{4}])$. Note that M_S is still a fibre bundle, with fibre homeomorphic to $c(L_S)$, over the i-manifold S. Hence, by taking a local trivialization of this bundle, we cover M_S by closed sets U_S^j, each of which is homeomorphic to the set $B^i \times c(L_s)$; furthermore, the intersection of any collection of these sets is either empty or homeomorphic to the sets themselves.

Now let R be any piece of X. We define $M_R = \rho_R^{-1}([0, \frac{3\epsilon}{4}])$. We now "cut" M_R to produce a smaller subspace N_R in the following manner:

$$N_R = M_R - \bigcup_{S \leq R} [M_R - \rho_S^{-1}([0, \frac{\epsilon}{2}))]$$

.

Since $\rho_S \circ \pi_R = \rho_S$, our blade has not cut any fibres in two; instead, we have cut some fibres entirely off of M_R, while leaving others intact. Hence the remaining

set N_R is still a fibre bundle, with each fibre still homeomorphic to $c(L_R)$, over the subspace $R \cap N_R = R - \bigcup_{S \leq R}[R - \rho_S^{-1}([0, \frac{\epsilon}{2}))]$. The interior of this last set is a smooth dim(R)-manifold. On the other hand its boundary is formed by the union of the sublevel sets $R \cap \rho_S^{-1}(\frac{\epsilon}{2})$, as S ranges over the pieces in the closure of R.

These sublevel sets must intersect transversely. To see this, let $S \leq W$ be two pieces in the closure of R. Then $R \cap \rho_S^{-1}(\frac{\epsilon}{2}) \cap \rho_W^{-1}(\frac{\epsilon}{2})$ is contained within $R \cap T_W$. The first sublevel set is comprised of one single fibre of T_W (since the control data condition gives us that $\rho_S \circ \pi_W = \rho_S$). On the other hand, $\pi_W \times \rho_W$ is a smooth submersion on every piece of $T_W - W$. Therefore, the sublevel set $\rho_W^{-1}(\frac{\epsilon}{2})$ must intersect each fibre of T_W transversely.

Consider a local trivialization of M_R. By the above reasoning, our "cuts" turn the sets in this local trivialization into a covering of the space N_R by closed sets U_R^j, each of which is homeomorphic to $B^{\dim \mathrm{R}} \times c(L_R)$; as before, the intersection of any collection of these sets is either empty or homeomorhpic to any one of the sets.

To complete the proof, we consider two sets U_S, U_R, for two pieces $S \leq R$. More specifically, let B_S, B_R be two closed balls in S and R such that $U_S \cong B_S \times c(L_S)$, and a similar statement with R. Suppose these sets intersect non-trivially, and let V denote this intersection.

Suppose we have a point $r \in B_R \cap U_S$. Then r belongs to a fibre $\pi_S^{-1}(s)$, for some $s \in B_S$. However, since $\pi_S \circ \pi_R = \pi_S$, and since $\{s\} \times \pi_S^{-1}(s) \subseteq U_S$, it follows that the entire fibre $\pi_R^{-1}(r)$ also lies in U_S. In other words, V is itself a fibre bundle, with fibre $c(L_R)$, over the set $U_S \cap B_R$. By shrinking the balls B_S, B_R if necessary, we can assume this bundle will be trivial, and hence homeomorphic to $B_R \times c(L_S)$. This proves the claim.

4

Persistence for Intersection Homology

This chapter concerns persistence, and extended persistence, for intersection homology. We first provide a general definition in the case of any simplicial complex along with an ordering on its simplices and a proper/improper decision procedure. The chapter closes with an algorithm, and proof of correctness, which computes the intersection homology pair groups for such an equipped complex.

In the interim, we give a detailed motivating example of intersection homology persistence for an embedded stratified space filtered by a Stratified Morse function. We then prove some duality and symmetry properties for the subdivided star filtration of a simplicial complex; this filtration is designed to mimic the intersection homology persistence values obtained in the Stratified Morse context.

4.1 Simplicial Persistent Intersection Homology

In this section, we define persistent intersection homology in its most general simplicial setting. We begin by discussing ordinary persistent intersection homology and then quickly define its extended version. These definitions are all formally identical to those found in Chap. 2 in the case of homology persistence for a general filtered simplicial complex.

Suppose that

$$\varnothing = K_0 \subseteq K_1 \subseteq K_2 \ldots \subseteq K_m = K. \tag{4.1}$$

is a filtration by subcomplexes of a simplicial complex K, and that $\bar{q}$ is a proper/improper decision procedure on K. We restrict this procedure to each K_i and let $I^q H_r^i$ denote the rth intersection homology group with procedure $\bar{q}$ of the subcomplex K_i.

Persistence and Extended Persistence For $0 < i < j \leq m$, and for each nonnegative integer r, the inclusion maps $K_i \hookrightarrow K_j$ induce maps $f_r^{i,j} : I^q H_r^i \to I^q H_r^j$. We say that a class $\alpha \in I^q H_r^i$ is *born* at level i iff $\alpha \notin \mathrm{Im} f_r^{i-1,1}$. Such a class *dies* at level j iff $f_r^{i,j}(\alpha) \in \mathrm{Im} f_r^{i-1,j}$ and $f_r^{i,j-1}(\alpha) \notin \mathrm{Im} f_r^{i-1,j-1}$. We then define the rth intersection pair group for these two levels to be the $\mathbb{Z}/2\mathbb{Z}$-vector space with basis the r-classes which are born at K_i and die at K_j; we denote this group $I^q P_r^{i,j}$.

Now suppose we have another filtration by subcomplexes of K:

$$\varnothing = L_0 \subseteq L_1 \subseteq L_2 \ldots \subseteq L_n = K$$

.

Using the same decision procedure $\bar{q}$ as before, we put $I^q H_r^{m+k} = I^q H_r(K, L_k)$, and then simply extend the definition of our maps $f_r^{i,j}$ and intersection pair groups to allow $0 < i < j \leq m + n$.

Persistence Values and Diagrams Suppose that a real-number value $h(i)$ is assigned to each level i of the extended filtration. Then we measure the persistence of a class α which is born at i and dies at j as $pers(\alpha) = |h(j) - h(i)|$. Intersection persistence diagrams, along with their ordinary, relative, and extended subdiagrams, are then defined in identical fashion to their counterparts in the usual homology case.

4.2 Stratified Morse Persistence

Let X be a d-dimensional stratified space embedded in some smooth manifold M (in the next chapter, we restrict to the case where $M = \mathbb{R}^{d+1}$) and suppose we have a real-valued function $f : M \to \mathbb{R}$. Consider the restriction, also denoted f, of this function to X. As outlined in Sec. 2.2, f gives rise to two filtrations of X: one by the sublevel sets $X_{\leq a} = f^{-1}((-\infty, a])$, and one by the superlevel sets $X^{\geq a} = f^{-1}([a, \infty))$. Fixing a perversity $\bar{p}$, we wish to do intersection homology persistence on the resulting extended filtration.

Let us suppose that f is the restriction to X of a smooth function on M. We say that f is critical at a point $x \in X$ iff f is critical at x when restricted to the manifold piece containing x. Further, we suppose that f is a Stratified Morse function([24]); in particular, this means that f is Morse when restricted to each piece and that its critical values, $a_1 < a_2 < \ldots < a_m$, are all distinct. We choose one set of m regular values $t_1, \ldots, t_m$, such that $a_i < t_i < a_i + \epsilon$. And we pick another set $s_1, \ldots, s_m$, such that $a_i - \epsilon < s_i < a_i$.

Consider two regular values $b < c$ that are between two adjacent critical values. Then, as $X_{\leq b}$ is stratum-preserving homeomorphic to $X_{\leq c}$ ([24]), the inclusion induces isomorphisms on their respective intersection homology groups.

Hence we need consider only the maps $I^p H_*(X_{\leq t_i}) \to I^p H_*(X_{\leq t_j})$, and $I^p H_*(X, X^{\geq s_i}) \to I^p H_*(X, X^{\geq s_j})$, for $i < j$. Fixing a dimension r, recall the shorthand definitions $I^p H_r^i = I^p H_r(X_{\leq t_i})$, and $I^p H_r^{m+j} = I^p H_r(X, X_{\geq s_j})$, for $0 < i, j \leq m$.

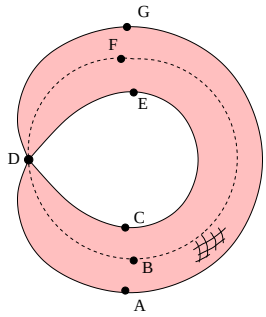

FIGURE 4.1: A disc (not pictured) is attached along the dashed circle.

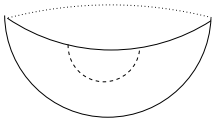

FIGURE 4.2: There are three $\bar{p}$-components.

We then have the following sequence of groups and homomorphisms, beginning and ending with the zero group, with which we compute intersection homology persistence:

$$0 \to I^p H_r^1 \to I^p H_r^2 \to \ldots \to I^p H_r^m \to I^p H_r^{m+1} \to \ldots \to 0$$

Example For purposes of illustration, we follow a particular example in some detail.

Let X be a pinched torus to which a disc has been attached along a closed curve which goes through the pinch point. We embed X into $\mathbb{R}^3$ as pictured in Fig. 4.1; for simplicity of viewing, the disc itself is not pictured, although its attached boundary is given by the dashed circle. Let v be the unit vector in the vertical direction and let $f_v(x) =< x, v >$; in other words, f_v measures height in the vertical direction. The critical points of f_v are as labeled in the picture; we assume the disc is attached in such a way that its interior contains no critical points of f_v.

Ascending Past the Critical Points We fix our perversity $\bar{p} = (-1, 0)$ and begin the ascent through the filtration. Before we pass point A, we have the empty set. Immediately afterwards, the sublevel set is homeomorphic to a bowl: there is one component born at A. This component lives to the top of the filtration and will be paired on the way down by extended persistence.

Fig. 4.2 depicts the sublevel set immediately after passing point B. It consists of a bowl and a portion of a disc (again, not pictured) attached along the dashed line. As $p_1 = -1$, one-chains may not touch the dashed line. Hence two components are born at B: one, which we call B_1, is represented by any point on the front portion of

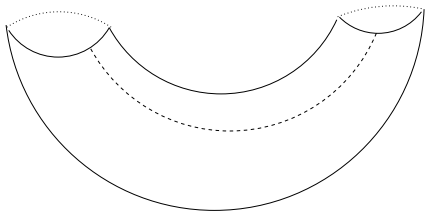

FIGURE 4.3: One of the components born at B merges here.

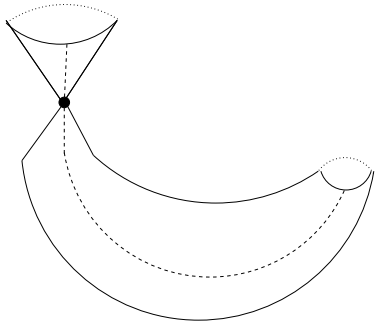

FIGURE 4.4: A new component is born.

the bowl cut off by the dashed line, while the other, B_2, is the class of any point on the disc. Note a key difference here from Morse Persistence for ordinary homology on a manifold: more than one birth (or death) may occur when passing a critical point. We identify all of these events with the single point B for pairing purposes and distance measurement. As we discuss in the next chapter, these multiple events can only happen when the critical point lies in a singular stratum.

In Fig. 4.3, we see the result of passing the third critical point. There are now only two components, as the dashed line no longer separates the torus portion. Hence we pair point B with point C; more precisely, we have a point $(f_v(B), f_v(C)$ in the 0-dimensional intersection homology persistence diagram. Note there is no allowable 1-cycle here, since any closed non-bounding loop would have to make non-allowable contact with the dashed line.

A new component is born when passing D (Fig. 4.4). As this component dies after passing the next critical point (Fig. 4.5), we pair D with E.

Upon passing point F (Fig. 4.6), an allowable one-cycle is born, represented by the circle at the top of the figure. We pair F with G, as the cycle is capped off by the descending manifold of the index-two smooth global maximum of the space.

Descending Two components survived to the top of the filtration. As we descend past point G, the component representing the pinched torus becomes homologous to zero; hence we have the extended pair (A, G). The other essential component,

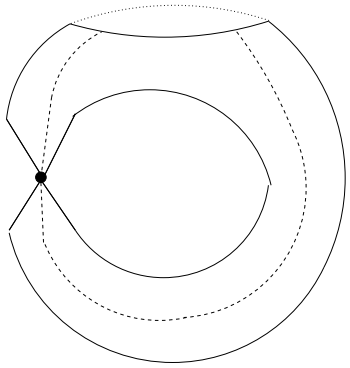

FIGURE 4.5: The youngest component dies here.

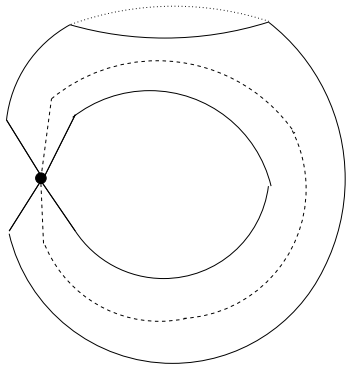

FIGURE 4.6: The circle on top is non-bounding.

represented by any point on the disc, becomes a boundary upon descending past point F, leading to the extended pair (B, F).

We also have the relative pair (F, E), reflecting the new relative one-cycle with one possible representative drawn in red in Fig. 4.7. Similarly, Fig. 4.8 demonstrates the relative pair (D, C). Finally, a relative 2-cycle is born upon descending past point B; this cycle lasts all the way to the bottom, leading to the relative pair (B, A).

4.3 Filtration by Subdivided Stars

In this section, we connect the two above notions of intersection homology persistence by showing how to construct a filtration on a simplicial complex which mimics the stratified morse version.

Vertex Ordering Consider a real-valued function $h : K \to \mathbb{R}$ which is given by linear interpolation of its values on the vertices of K. We order those vertices by their

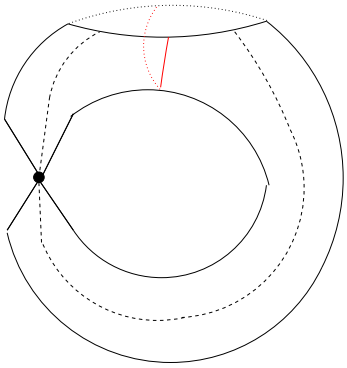

FIGURE 4.7: The new relative one-cycle is drawn in red.

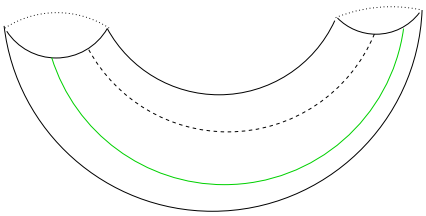

FIGURE 4.8: The green arc represents a relative one-cycle which had earlier not been allowable: it would have been required to go through the pinch point.

function values: $v < w$ iff $h(v) < h(w)$. Let $h(v_i) = r_i$, pick some t_i just larger than r_i, and consider the sublevel set $K_{\leq t_i} = h^{-1}((-\infty, t_i])$.

Recall the lower-star filtration defined in Chap. 2: one let K_i be the full subcomplex of K spanned by the lowest i vertices. Then the sublevel set $K_{\leq t_i}$ was homotopically equivalent to K_i, and thus the persistence homology pairings computed via the K_i matched perfectly with those from the topological space.

The same technique fails for intersection homology, for the simple reason that intersection homology is invariant only under *stratum-preserving* homotopy. As an extremely simple example of this failure, let K be the triangulation of a wedge of two circles shown in Fig. 4.9, where h measures height in the vertical direction.

Note that $K_{\leq t_3}$ looks like a circle with two line segments sticking out, while K_3 is a circle. These spaces are homotopically equivalent, but not via a homotopy that preserves codimension of strata. If we compute 0th-dimensional intersection homology for the perversity $\bar{p} = (-1)$, we find three and one components, respectively. In particular, neither of the two components born at $K_{\leq t_3}$ in the stratified morse context are born at K_3.

Filtrations This problem is easily fixed, however. Let K' denote the first barycentric subdivision of K. For each $i < n$, we define:

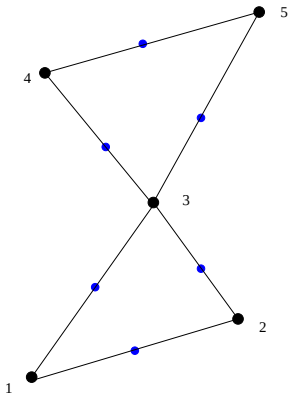

FIGURE 4.9: The first barycentric subdivision of a triangulation of a wedge of two circles. The original vertices of K are drawn in black and ordered by height in the vertical direction

$$\bar{K}_i = \bigcup_{j \leq i} \bar{st}(v_j, K') \tag{4.2}$$

and

$$\bar{L}_{n-i} = \bigcup_{s > i} \bar{st}(v_s, K') \tag{4.3}$$

Letting i run from 1 to n, these definitions give us an ascending and a descending filtration of K. Furthermore, for any i, there is indeed a stratum-preserving deformation retraction between $K_{\leq t_i}$ and $\bar{K}_i$, and hence the intersection pair groups will be identical.

Note that $\bar{K}_i$ and $\bar{L}_{n-i}$ are stratified subcomplexes, with boundary, of K. Furthermore, these boundaries are actually equal: their shared boundary $\partial \bar{K}_i = \partial \bar{L}_{n-i}$ consists of the full subcomplex of K' spanned by the barycentres of simplices in K which themselves are spanned by at least one vertex lower than or equal to v_i and and at least one vertex higher than v_i.

4.4 Pair Group Duality and Symmetry

Recall the proofs of pair group duality and symmetry for ordinary persistent homology on a simplicial complex which triangulates a d-manifold.

The first step, as outlined in Sec. 2.3, was to observe that Lefschetz Duality, along with excision, provided a perfect pairing between the groups $H_r(K_i)$ and $H_{d-r}(K, L_{n-i})$, where n was the number of vertices in K. Then some algebraic manipulation allowed us to obtain the needed results.

The same procedure will give us the analogous results for persistent intersection homology. Let K be a d-dimensional stratified simplicial complex with an ordering on its n vertices and consider a pair $\bar{p}, \bar{q}$ of dual perversities.

For each $i \leq n$, the subcomplex $\bar{K}_i$ is a stratified subcomplex with boundary $\partial \bar{K}_i = \partial \bar{L}_{n-i}$. Lefschetz Duality for intersection homology then gives a perfect pairing between $I^p H_r(\bar{K}_i)$ and $I^q H_{d-r}(\bar{K}_i, \partial \bar{L}_{n-i})$. By excision, this latter group is isomorphic to $I^q H_{d-r}(K, \bar{L}_{n-i})$. So, adopting the shorthand defined above, we have, for each i and each r, a perfect pairing:

$$I^p H_r^i \otimes I^q H_{d-r}^{2n-i} \to \mathbb{Z}/2\mathbb{Z} \tag{4.4}$$

By the same logic as in Sec. 2.3, the result above leads to the following two results, where $(x, y)^R = (-y, -x)$ and $(x, y)^N = (-x, -y)$.

9 (Intersection Pair Group Duality)**.** *For $0 < i < j \leq 2n$ and for a pair of dual perversities, we have perfect pairings:*

$$I^p P_r^{i,j} \otimes I^q P_{d-r}^{2n-j+1, 2n-i+1} \to \mathbb{Z}/2\mathbb{Z}$$

10 (Intersection Homology Diagram Symmetry)**.** *Let f be a real-valued function on the vertices of of a d-dimensional stratified complex K and let $\bar{p}, \bar{q}$ be two dual perversities. Then the $\bar{p}$ pairings for the filtration defined by f and the $\bar{q}$ pairings for the one defined by $-f$ are related in the following manner:*

$$\begin{aligned} I^p Ord_r(f) &= [I^q Ord_{d-r-1}(f)]^R \\ I^p Rel_r(f) &= [I^q Rel_{d-r+1}(-f)]^R \\ I^p Ext_r(f) &= [I^q Ext_{d-r}^R(-f)]^N \end{aligned}$$

Hence if all we care about is *persistence* of classes, rather than order or dimension, the information gained by the two filtrations will be identical.

Example Recall the stratified space in Fig. 4.1 and let $\bar{p} = (-1, 0), \bar{q} = (0, 0)$, and v be the vertical direction. Recall that when we did I^p-persistence for the filtration defined by f_v, there was an ordinary 1-class born at point F which was then capped off at point G; this corresponds to the point $(f_v(F), f_v(G)) \in I^p Ord_r(f)$.

Suppose we now do I^q-persistence using the filtration defined by $f_{-v} = -f_v$. There will then be a component born at G which dies as soon as we pass F, since any point which can be allowably connected to the singular set will become a boundary. In other words, there will be a point $(-f_v(G), -f_v(F)) \in I^q Ord_{d-r-1}(-f)$.

The complete persistence diagrams for the two functions are shown in Fig. 4.10.

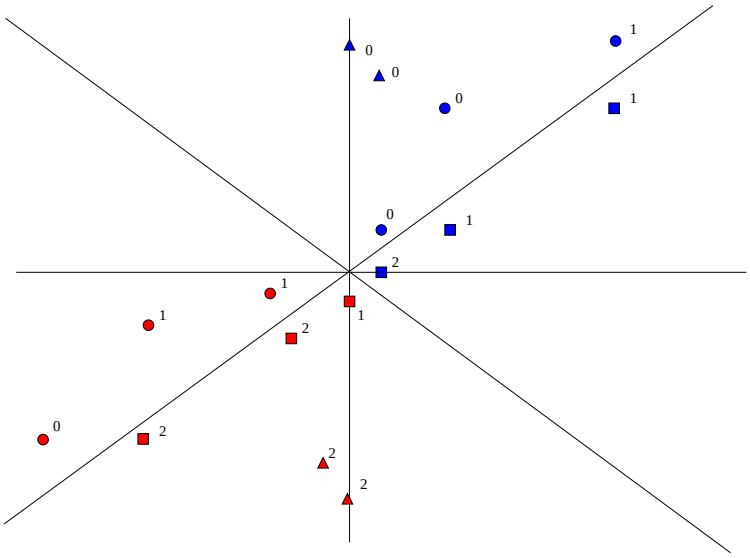

FIGURE 4.10: The diagrams $I^pDgm(f), I^qDgm(-f)$, in all three relevant dimensions, are superimposed, with the former diagrams in blue and the latter in red. Circles, boxes, and triangles indicate ordinary, relative, and extended points, respectively. Each point is labelled by dimension.

4.5 Algorithm for Computing Persistence

Suppose that we have a simplicial complex K with a "singular" subcomplex Σ. Further suppose this complex comes equipped with both an ordering on its simplices and a procedure $\bar{q}$ for deciding whether or not a given simplex is proper.

In this section, we first give an algorithm to compute the ordinary intersection pair groups for such an equipped complex. We then give an algorithm for computing the extended intersection pair groups for a simplicial complex K equipped with two such orderings, with only one proper/improper decision procedure, on its simplices. Proofs of correctness for these two algorithms are provided in the next section.

4.5.1 Neutral Simplices

The algorithm is quite similar in form to the ordinary persistence algorithm ([17], [15]), in that it reduces a boundary matrix in such a way that an injective "lowest-one" function may be defined on the columns of this matrix. The chief difference for our algorithm lies in the interpretation of this "lowest-one" function, as well as in the initial ordering of the columns and rows of the boundary matrix.

These changes are necessitated by the fact that we can no longer partition the set of simplices into "positive" and "negative", as we did for ordinary homology in

SubSec. 2.1.2. Instead, there is a third category, "neutral," which requires special attention. Before describing the algorithm, it seems wise to first address this distinction in more detail.

Case Analysis Suppose that we add a simplex σ to a simplicial complex L. We consider the possible differences between the intersection homology of L and that of $L \cup \{\sigma\}$. First of all, an improper simplex σ will effect no change whatsoever, as it can not form part of an allowable chain. So we assume σ is a proper simplex, recalling that in general σ can have improper faces in its boundary.

For ease of notation, we make the following definitions. For each i, let $P_i(K)$ be the $\mathbb{Z}/2\mathbb{Z}$ vector space with basis the proper i-simplices of K. For each chain $\gamma \in P_i(K)$, we denote by $Pr(\gamma)$ and $Imp(\gamma)$ the sets of proper and improper simplices, respectively, in its boundary. For example, consider Fig. 3.3. If Δ is any triangle in that picture, then $Imp(\Delta)$ has two edges in it. On the other hand, if γ represents the sum of all the triangles, then $Imp(\gamma) = \varnothing$.

So suppose the newly added σ is a proper i-simplex. There are two cases, the first of which has two sub-cases:

- There exists $\gamma \in P_i(K)$ such that $Imp(\gamma) = Imp(\sigma)$ (this includes the possibility that $Imp(\sigma) = \varnothing$, of course). In this case, the sum $\gamma + \sigma$ is an *allowable* chain since the addition cancels out all improper simplices. Put $\alpha = \partial(\gamma + \sigma)$. Note that α is an allowable cycle. There are then two cases:

  1. α was not the boundary of an allowable i-chain in K. But now it is. Thus, the addition of σ lowered the $(i-1)$st intersection Betti number by one. In this case, we call σ **negative**.
  2. α was the boundary of an allowable i-chain β in K. Then $\sigma + \gamma + \beta$ represents a new non-bounding allowable i-cycle. The ith intersection Betti number increases by one and we say that σ is **positive**.

- For every $\gamma \in P_i(K)$, we have $Imp(\gamma) \neq Imp(\sigma)$. In this case, the addition of σ cannot create any new allowable chains. All intersection Betti numbers remain the same and we label σ as **neutral**.

Sometimes we will wish to stress only that a particular simplex is not neutral, without specifying whether it is positive or negative. In this case, we call the simplex **active**.

4.5.2 Algorithm Description

Values Fix a symbol $\bar{q}$, which might be a perversity or just any procedure that decides whether or not a given simplex is proper. The input to the algorithm is a list of n simplices ordered $\tau_1, \tau_2, \ldots, \tau_n$, such that a simplex comes after its boundary faces, along with a real-number value associated to each simplex; for example, each simplex might inherit a value from a real-valued function on the vertices.

We filter K one simplex at a time: let K_i be the first i simplices in the above ordering and put $I^{\bar{q}}H^i_* = I^{\bar{q}}H_*(K_i)$. Our algorithm will then track classes which are born, for example, at K_i and die at, say, K_j. However, we always define the actual numerical persistence of our pairs using these values. Hence, although our algorithm adds one simplex at each level and indeed starts by re-ordering the input simplices, this will not affect the numerical persistence of the pairs that it computes.

Re-Ordering Let $Prop(K), Imp(K)$ denote the subsets of proper and improper simplices; assume they are of size $s, n-s$, respectively. We reorder the input simplices so that all simplices in $Prop(K)$ come first, while otherwise preserving the input ordering among them. That is, we list the proper simplices as $\sigma_1, \sigma_2, \ldots, \sigma_s$, where for each $i < j$, σ_i came before σ_j in the input ordering. Since we are working within the relative chain groups $C_*(X, \Sigma)$, we discard all improper simplices which lie entirely within the singular set; suppose this leaves us with m total simplices. The remaining improper simplices $\sigma_{s+1}, \sigma_{s+2}, \ldots, \sigma_m$, ordered in the same manner as above, are then appended to the end of the list.

We will later need to refer back to the original ordering when formulating the proof of correctness for our algorithm. To make this easier, we define an injective and order-preserving function $g : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, n\}$ by $g(i) = k$ where $\sigma_i = \tau_k$.

Reduction Algorithm The $m \times s$ binary matrix D (Fig. 4.11) is constructed as follows. The s proper simplices index the columns, while the rows are indexed first by the s proper simplices and then by the $m - s$ improper ones. We define $D[i, j]$ to be 1 iff σ_i is a codimension one face of σ_j.

For each nonzero column j of an arbitrary matrix M, we define $low_M(j) = i$ iff the ith row contains the lowest nonzero entry in the jth column; if column j is zero, we set $low_M(j) = 0$. The matrix M is called *reduced* if low_M is an injective function when restricted to the nonzero columns. Our algorithm reduces D by adding columns left-to-right. The reduction procedure $Reduce - Matrix$ is identical to that found in Chap. 2, but we include it here for completeness:

for j = 1 to s **do**
  while $\exists j' < j$ with $low(j') = low(j) \neq 0$ **do**
    add column j' to column j
  end while
end for.

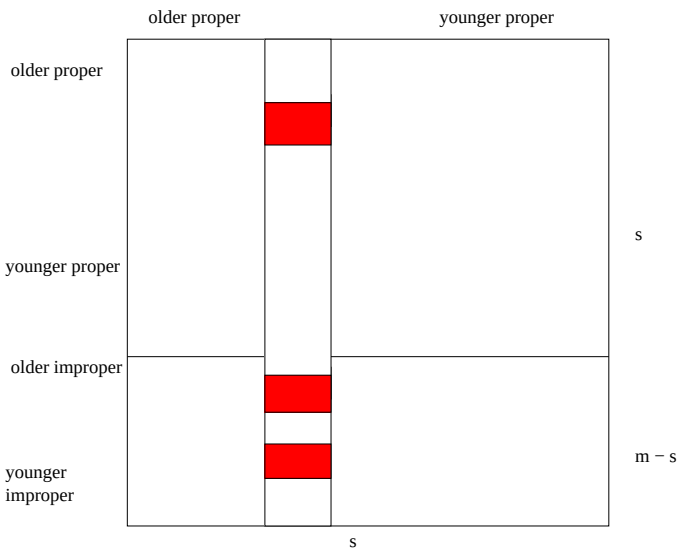

FIGURE 4.11: The initial (unreduced) boundary matrix D for the IH algorithm. The white column is indexed by a proper triangle which has one proper edge and two improper edges in its boundary.

Interpretation and Pairings The algorithm $ReduceMatrix$ produces a reduced matrix R. We read the pairings from low_R as follows (See Fig. 4.12):

- if $low_R(j) = 0$, then the addition of the simplex σ_j created an allowable $\dim(\sigma_j)$-cycle. This simplex will either be paired with a later one or will go unpaired.
- if $low_R(j) = i \leq s$, then the addition of σ_j killed the allowable $\dim(\sigma_i)$ cycle that was born when adding σ_i. We obtain the pair (σ_i, σ_j).
- if $low_R(j) = k > s$, then the addition of σ_j had no immediate effect on the intersection homology of the filtration. σ_j is neutral.

Intuition Consider the original simplex ordering and imagine adding one simplex at a time in sequence. An *allowable* chain, whether it is a cycle or not, must necessarily be a sum which consists entirely of *proper* simplices. Hence the addition of an *improper* simplex can neither create an allowable cycle nor destroy one via an allowable chain. For this reason, we imagine that we are adding only the proper simplices in sequence and we index the columns accordingly.

However, the boundary of a proper simplex σ need not itself consist of a sum of proper simplices; see, for example, any one of the triangles in Fig. 3.3. In this case, the simplex is not, by itself, an allowable chain. However, this does not mean that σ is neutral, since we might hope to add older proper simplices to σ in an attempt to cancel off the improper simplices in its boundary. For this reason, we include the improper simplices at the bottom of the row listing.

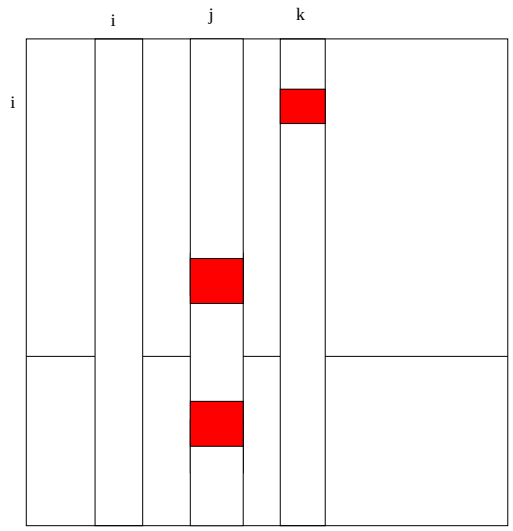

FIGURE 4.12: The reduced matrix from the algorithm. We read that σ_i created an allowable class which was then destroyed by the addition of σ_k, but σ_j neither created nor destroyed an allowable class.

4.5.3 Extended Persistence

We now suppose that we have *two* simplex orderings on K: the ascending $\tau_1, \tau_2, \ldots, \tau_n$, and the descending $\lambda_1, \lambda_2, \ldots, \lambda_n$. We prune and re-order both sets as before: we discard all entirely singular simplices and then place the proper simplices first in each ordering. Let $\sigma_1, \ldots, \sigma_s, \sigma_{s+1}, \ldots, \sigma_m$, and $\kappa_1, \ldots, \kappa_s, \kappa_{s+1}, \kappa_m$ be the two new orderings.

Extended Persistence Algorithm We fill the entries of a $2m \times 2s$ $0-1$ matrix D (Fig. 4.13 as follows. Divide D up into 4 $m \times s$ submatrices. The lower left submatrix will be all zeroes. The upper left submatrix, which we call A, is just the boundary matrix from the ordinary intersection homology persistence algorithm, defined for the ascending filtration; in other words, $A[i,j] = 1$ iff σ_i is a codimension 1 face of σ_j. The lower right submatrix, labeled B, will be the analogous matrix for the descending filtration; that is, $B[i,j] = D[m+i, s+j] = 1$ iff κ_i is a boundary face of κ_j.

Finally, we call the upper right submatrix P. For $i \leq s$, we set $P[i,j] = D[i, s+j] = 1$ iff $\sigma_i = \kappa_j$, while for $i > s$, we set $P[i,j] = D[i, s+j] = 0$.

The algorithm now proceeds exactly as above: we add columns in D left-to-right until the lowest one in each column, if it exists, is in a unique row. Notice that if we ever need to add a column from the left half-matrix to the right one, that column will already have had its lowest one raised above the proper/improper dividing line. Hence the all-zero submatrix of P will remain so during the entire process.

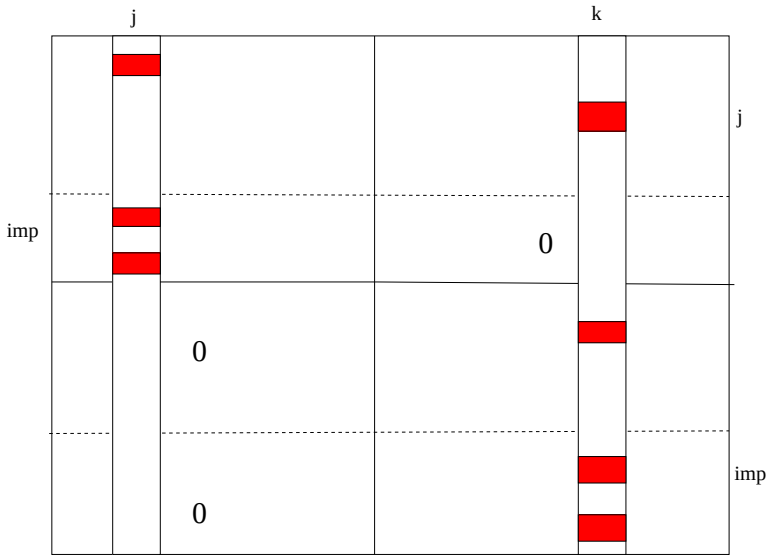

FIGURE 4.13: The initial boundary matrix for the extended persistence algorithm. The triangle $\sigma_j = \kappa_k$ and has one proper and two improper simplices in its boundary.

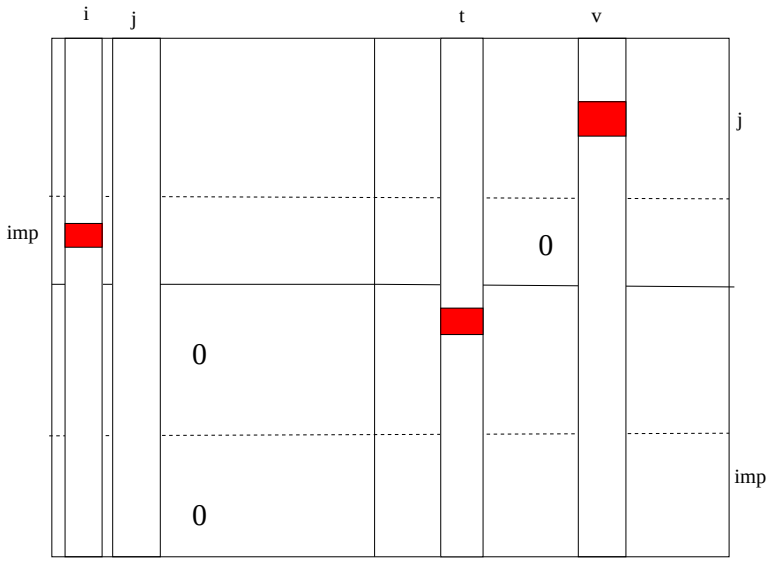

FIGURE 4.14: The reduced matrix from the extended IH algorithm. We see that σ_i is neutral σ_j created an essential class and is then paired with κ_v, while κ_t destroyed a relative class.

Interpretation and Pairings The resulting reduced matrix will be called R, and low_R will be the corresponding "lowest-one" function. The values of low_R will give us ordinary pairings, relative pairings, extended pairings, and simplices which are not paired at all, as follows (see Fig. 4.14). Suppose $low_R(i) = j$. There are then several possible cases, the first two of which are the same as before:

- $i, j \leq s$: we have (σ_i, σ_j) as an ordinary pair.
- $i \leq s, s < j \leq m$: σ_j is neutral.
- $i \leq s, j > s$: we have the extended pair (σ_i, κ_{j-s}).
- $i > m, s < j < m + s$: this gives the relative pair $(\kappa_{i-m}, \kappa_{j-s})$.
- $i > m, j > m + s$: κ_j is neutral.

4.6 Proofs of Correctness

In this section, we prove the correctness of the two algorithms presented above. We start by providing a precise statement of correctness for the first algorithm.

11 (Correctness of Ordinary IH Algorithm)**.** *For $0 < i, j \leq s$, (σ_i, σ_j) is computed by our algorithm iff $I^q P_r^{g(i),g(j)} = \mathbb{Z}/2\mathbb{Z}$, where $r = dim(\sigma_i)$.*

To prove this statement, we start by constructing another reduction algorithm for D for which the associated *low*-function clearly computes the correct persistence pairs. We then employ the Pairing Uniqueness Lemma ([13]) which states that any such *low* function must depend only on D and hence that our given algorithm computes the same pairs as the provably correct ones.

Finally, we prove the correctness of the intersection homology extended persistence algorithm in identical fashion to the proof for ordinary homology extended persistence given in [10]: namely, we construct a "cone complex" on which the ordinary persistence algorithm computes the same pairings as our extended persistence algorithm does on the original complex.

4.6.1 Proof for First Algorithm

The outline of the proof is as follows. We define a function low_M on our proper simplices that will give us a set of pairs of simplices. To define this function, we first construct a recursive procedure $Make - Active_i$ to decide if a given simplex σ_i is or is not active. The resulting active simplices are then input into the $Pair - Simplices$ procedure, also constructed below, which completes the definition of the function. We then prove that the analogous statement to Statement 11 above is true for this new procedure. Finally we appeal to the Pairing Uniqueness Lemma to complete the proof.

Before giving these algorithms, we first define a function $n : P_i(K) \to \{s+1, s+2, \ldots, m\}$ via $n(\gamma) = i$ where σ_i is the youngest simplex in $Imp(\gamma)$. If $Imp(\gamma) = \varnothing$, which means that γ is an allowable chain, we set $n(\gamma) = 0$.

Finding Neutral Simplices Here is pseudocode for the procedure $Make - Active_i$:

$\gamma_i = \sigma_i$
for $j = 1$ to $i - 1$ **do**
  $\gamma_j = Make - Active_j(\sigma_j)$
end for
while $\exists j < i$ such that $n(\gamma_j) = n(\gamma_i)$ **do**
  $\gamma_i \leftarrow \gamma_j + \gamma_i$
end while
Return γ_i

For each i, we define

$$a(\sigma_i) = \partial(\gamma_i) \tag{4.5}$$

Notice that σ_i is neutral iff $n(a(\sigma_i)) \neq 0$. If σ_i is indeed neutral, we define $low_M(\sigma_i) = n(a(\sigma_i))$. Otherwise, we leave $low_M(\sigma_i)$ undefined for the moment.

The rest of the proof now very closely follows the proof of correctness for the usual homology persistence algorithm given in [17]. As a few details are different, we will write a full description here, while maintaining some of the notation found therein.

Basis Construction We start by constructing, for each positive r-simplex σ_i, an r-cycle c_i which contains σ_i as its only positive simplex. To see that this is possible, observe that the addition of σ_i created a new class in $I^qH_r^{g(i)}$. Let γ be a chain representative of this class and suppose that γ has more than one positive simplex. These positive simplices come earlier in the filtration than σ_i; hence by induction they are contained within cycles which contain no other positive simplices. Adding these cycles to γ, we cancel all positive simplices while adding no new ones, thus constructing our cycle c_i. Let h_i denote the homology class of c_i.

Now suppose $\alpha \in I^qH_r^{g(i)}$. Then α was born at some level $g(k) \leq g(i)$. Hence α can be written as a sum

$$\alpha = h_k + \sum_{j \in I(\alpha)} h_j, \tag{4.6}$$

where $I(\alpha)$ is a set of indices all less than i. In other words, for some subset of indices $I(i)$ taken from $\{1, 2, \ldots i\}$, the classes h_j, or more precisely $f_r^{g(j),g(i)}(h_j)$, for $j \in I(i)$, form a basis for the intersection homology group $I^qH_r^{g(i)}$.

Using this fact, we define a function, $y : I^qH_r^{g(i)} \to \{1, 2, \ldots i\}$ by $y(\alpha) = k$, where k is defined as in Eqn. 4.6.

Pair-Simplices Algorithm We now give an algorithm which pairs some of the active simplices. These pairings will complete the definition of the low_M function, the values of which have already been given for the neutral simplices. The algorithm maintains, for each r and k, a list P_r^k of the paired simplices at stage k of the algorithm. Here is the pseudocode:

$\forall r, P_r^0 = \varnothing$
for $j = 1$ to s **do**
  if σ_j non-negative **then**
    $\forall r, P_r^{g(j)} = P_r^{g(j-1)}$
  else
    $i = y([a(\sigma_j)])$
    $k = \dim(\sigma_j)$
    $P_k^{g(j)} = P_k^{g(j-1)} \cup \{(\sigma_i, \sigma_j)\}$
    $\forall r \neq k, P_r^{g(j)} = P_r^{g(j-1)}$
  end if
end for

We then finish the definition of low_M by defining $low_M(j) = i$ iff (σ_i, σ_j) is produced by $Pair - Simplices$.

Matrix Formulation As with our original algorithm, the procedures above can all be accomplished by performing column operations on the original boundary matrix D, or alternatively, by multiplying on the right by a product of elementary matrices V.

Let us call a simplex σ_k *potentially neutral* if $Imp(\sigma) \neq \varnothing$. The $Make - Active_k$ procedure manifests itself by adding columns from the left which correspond to actually neutral simplices, to column j. If the procedure succeeds in raising the lowest one in column j above the proper/improper demarcation line, then σ_j is in fact active. Otherwise, it is actually neutral and we never use its corresponding column again in the reduction. The rest of the algorithm just completes the reduction of the matrix, starting all over again from left to right, but this time only employing columns corresponding to active simplices.

At the end of the process, we have a reduced matrix M and a corresponding lowest one function low_M. If $low_M(j) = i > s$, then column j of V stores a chain of neutral simplices; in other words, a non-allowable chain. If column j of M is empty, then column j of V stores a cycle consisting of positive, negative, and neutral simplices; this cycle is precisely the representative of the basis element d_j described above. Finally, if $low_M(j) = i \leq s$, then column j of V stores an allowable chain which destroys the class created by σ_i.

Hence we have one matrix D and two reduced matrices R, M which result from performing column operations on D. By the Pairing Uniqueness Lemma ([13]) then, we conclude $low_M = low_R$.

The correctness proof will therefore be complete after we prove the following, which is the analogue to Statement 11 for our algorithm.

12. *Let $r = dim(\sigma_i)$. Then $low_M(j) = i < s$ iff $I^q P_r^{g(i),g(j)} = \mathbb{Z}/2\mathbb{Z}$*

Proof. We prove the forward direction; the other direction is essentially just a restatement of what follows below.

Let c_i be the cycle containing σ_i as constructed above (4.6), and let d_i be its homology class in $I^q H_r^{g(i)}$. We show that d_i is born at level $g(i)$ and dies at level $g(j)$:

First, we show $d_i \notin \mathrm{Im} f_r^{g(i-1),g(i)}$: Suppose it was. Then $\exists \alpha \in I^p H_r^{g(i-1)}$ such that $f_r^{g(i-1),g(i)}(\alpha) = d_i$. Writing α as in (4.6), we arrive at a contradiction.

Next we show $f_r^{g(i),g(j)}(d_i) \in \mathrm{Im} f_r^{g(i-1),g(j)}$: By construction, $y(a(\sigma_j)) = i$. This means that, dropping the maps induced by inclusion for the moment, we can write:

$$a(\sigma_j) = c_i + \sum_{k \in I} c_k \tag{4.7}$$

where I is some set of indices less than i. We then pass this equation to homology and push it forward to level $g(j)$, where $[a(\sigma_j)] = 0$, since by construction (see (4.5), $a(\sigma_j) = \partial(\gamma_j)$. Hence at level $g(j)$, we see that the image of d_i is equal to the image of the sum of classes on the right. But all of these classes existed at levels lower than $g(i)$. Hence d_i died at least by level $g(j)$.

Finally, we show that $f_r^{g(i),g(j-1)}(d_i) \notin \mathrm{Im} f_r^{g(i-1),g(j-1)}$. Suppose it were. Then at level $g(j-1)$, the image of d_i is homologous to the image of a class coming from before level $g(i)$. Hence, using the basis defined in (4.6) and dropping maps, we can find an allowable chain $\eta \in I^q H_{r+1}^{g(j-1)}$ such that

$$\partial(\eta) = c_i + \sum_{t \in J} c_t$$

,

for some set of indices less than i. Notice that $\partial(\eta)$ is homologous to zero at level $g(j-1)$. So we can add this equation to (4.7) and pass to homology to obtain:

$$[\partial(\eta)] + [a(\sigma_j)] = [a(\sigma_j)] = \sum_{k \in I} d_k + \sum_{t \in J} d_t.$$

But all the indices on the right hand side are less than i, and so this contradicts the definition of $y(a(\sigma_j))$. Therefore, d_i dies at level $g(j)$.

□

4.6.2 Proof of Correctness for Extended Algorithm

This can be reduced to the correctness of the original algorithm as follows. Let $w * K$ be the simplicial complex obtained by coning every simplex in K to a new vertex ω.

We then define $Prop(w * K) = Prop(K) \cup w * (Prop(K))$; that is, $w * \tau$ is proper iff τ is proper. Note that this proper/improper decision process need not be able to come from a choice of perversity; but our algorithm will work on any list of simplices along with a sublist of proper simplices.

Suppose K has ascending and descending simplex orderings $\tau_1, \ldots, \tau_n$, and $\lambda_1, \ldots, \lambda_n$. We convert these to two orderings on $Prop(K)$: $\sigma_1, \ldots, \sigma_m$, and $\kappa_1, \ldots, \kappa_n$. Then we use these orderings to define an ordering on $Prop(w * K)$ as follows: $\sigma_1, \ldots, \sigma_m, w * \kappa_1, \ldots, w * \kappa_m$. And we do ordinary intersection homology persistence on this filtration.

The proof is complete when we note that an ordinary pairing (σ_i, σ_j) corresponds to the same pairing in the cone complex, a relative pairing (κ_i, κ_j) corresponds to $(w * \kappa_i, w * \kappa_j)$, and an extended pairing (σ_i, κ_j), gives $(\sigma_i, w * \kappa_j)$.

5

Elevation Functions and Comparison with Persistent Homology

This chapter concerns two distinct but interconnected topics.

First we compare and contrast the measurements obtainable from the different types of persistence which could be used on a given embedded stratified space. We will show that, in certain cases, intersection homology persistence can give a richer description of a stratified space than can standard homology persistence. On the other hand, we will show that the reverse is true in other cases.

The second object of discussion is a series of intersection homology elevation functions which we will define on different parts of an embedded stratified space. The last part of this chapter will give a complete characterization of these functions in dimension one.

Both topics make extensive use of the fundamental theorems of Stratified Morse Theory ([24]). So we begin the chapter with a review of these theorems. We also make some definitions and set some notation.

5.1 Background and Definitions

Stratified Morse Theory Let

$$X = X_d \supseteq X_{d-1} \supseteq \dots X_1 \supseteq X_0$$

be a d-dimensional Whitney stratified space embedded within some smooth manifold M, and let $\tilde{f} : M \to \mathbb{R}$ be a smooth function. The restriction f of this function to X, is defined to be critical at a point $x \in X$ iff it is critical when restricted to the particular manifold piece which contains that point; note this forces an isolated

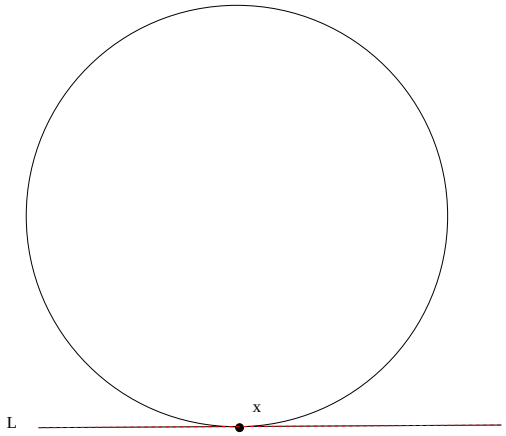

FIGURE 5.1: Suppose this loop is stratified so that x is a singular point. Then the function f which measures height in the vertical direction is not Stratified Morse: $d\tilde{f}_x$ annihilates the generalized tangent line L.

singular point to be critical for any function. By a critical value of f, we then just mean its value at a critical point.

To express one of the conditions in the definition of a Stratified Morse function, we will need the notion of a *generalized tangent space*. Let x be in a stratum S and let $x_i \in R$ be a sequence of points converging to x, where R is some stratum which contains S; by $T_{x_i}R$, we mean the tangent space to the manifold piece R at x_i. Then the limit $\lim_{x_i \to x} T_{x_i}R$ is called a generalized tangent space at x.

Finally, f is said to be Stratified Morse iff:

- f is a Morse function when restricted to each manifold piece.
- All critical values of f are distinct.
- For each critical point $x \in S_i$ the differential $d\tilde{f}_x$ does not annihilate any generalized tangent space to x other than $T_x S_i$.

See Fig. 5.1 for an example of this last condition being violated. In such a situation, we say that f is *overly critical* at x.

As in the previous chapter, we imagine that some Stratified Morse function f filters X via the sublevel sets $X_{\leq a} = f^{-1}((-\infty, a])$. If $b < c$ are two regular values such that no critical value lies between them, then the first fundamental theorem of Stratified Morse Theory tells us that $X_{\leq b}$ is stratum-preserving diffeomorphic to $X_{\leq c}$.

On the other hand, suppose that $x \in S_i$ is a critical point of index j, with $a = f(x)$ its associated critical value. Then the second fundamental theorem of Stratified Morse Theory tells us what happens as we move past the sublevel set $X_{\leq a}$:

the change is stratum-preserving homotopically equivalent to the result of gluing in the topological product of tangential and normal *Morse Data*.

To see what this means, let $N(x)$ denote the normal slice to x. Picking a small enough ϵ, we define $N^-(x,f) = N(x) \cap f^{-1}(a-\epsilon)$ and $N(x,f) = N(x) \cap f^{-1}([a-\epsilon, a+\epsilon])$. Finally, let $G(x,f) = (N^-(x,f) \times B^i) \cup (N(x,f) \times \partial B^j)$, and $D(x,f) = N(x,f) \times B^i$. Then there is a stratum-preserving homotopy equivalence between $X_{\leq a+\epsilon}$ and the space obtained by gluing in $D(x,f)$, along $G(x,f)$, to $X_{\leq a-\epsilon}$. In other words:

$$X_{\leq a+\epsilon} \backsimeq X_{\leq a-\epsilon} \cup_{G(x,f)} D(x,f) \tag{5.1}$$

Kernels, Cokernels, Birth, and Death Now we fix a stratified morse function f and consider the persistent homology along the filtration it defines. As always, we let $Dgm_r(f)$, $r = 0, \ldots, d$ denote the persistence diagram, with the usual notations for ordinary, relative, and extended subdiagrams.

Upon passing a given critical point x, with $a = f(x)$, one or more homology classes may be born, one or more classes may die, or nothing at all may happen. Consider the maps $\phi_*(x,f) : G(x,f) \to D(x,f)$ and $\psi_*(x,f) : G(x,f) \to X_{\leq a-\epsilon}$, both induced by inclusion.

By examining the Mayer-Vietorris sequence which comes from the union of spaces in Eqn. 5.1, one can easily show:

- each basis element of $\mathrm{CoKer}(\phi_r)$ corresponds to an r-class born at x. Hence $f(x)$ will appear as the abscissa in an ordinary or an extended pair in $Dgm_r(f)$.
- each basis element of $\mathrm{Ker}(\phi_r)$ corresponds to either:
  - an r-class which dies at x. This is when the basis element is not in $\mathrm{Ker}(\psi_r)$. In this case, $f(x)$ will appear as the ordinate in an ordinary pair in $Dgm_r(f)$.
  - an $(r+1)$-class which is born at x. This is when the basis element is in $\mathrm{Ker}(\psi_r)$. In this case, $f(x)$ will be the abscissa of an ordinary or extended pair in $Dgm_{r+1}(f)$.

Fixing a perversity $\bar{p}$, we let $I^p\phi_*(x,f), I^p\psi_*(x,f)$ denote the corresponding maps on intersection homology induced by inclusion. The same relationships hold between the intersection homology births and deaths at x, and the kernels and cokernels of these maps.

In the next section, we will discuss the relationships between homology and intersection homology persistence which hold in certain special cases. To ease exposition, we adopt the following convention: if an intersection homology class α is born at x, we say that it is $\bar{p}$-born at x, with a similar definition for $\bar{p}$-death; if a distinction needs to be drawn, we will say that a standard homology class is H-born or H-dies at a critical point.

Height Functions For the rest of this chapter, we will restrict to the case when X is embedded in Euclidean space of one higher dimension; that is, when $M = \mathbb{R}^{d+1}$. For each unit vector $v \in S^d$, we define the height function in direction v via $f_v(x) =< x, v >$. The height functions corresponding to a dense open subset of S^d will be Stratified Morse ([24], p. 54).

To ease notation, we will often let v play the role of f_v in the definitions above. So for example, we will write $G(x, v)$ in place of $G(x, f_v)$.

5.2 Intersection Homology vs. Homology

For a given space X and direction vector v, one could either do standard homology persistence on the filtration of x defined by f_v, or one could do intersection homology persistence for any number of perversities. The persistence pairings for these different homology theories will then, hopefully, illuminate different features of the space and its relationship to its singularities.

In this section, we discuss the differences between these types of persistence. First, we give an example, involving a codimension-one stratum, in which intersection homology gives more information. Then we prove that in the case of a pseudomanifold, any information obtained from intersection homology could have also been gleaned from standard homology. Finally, we give a higher-dimensional example, with no codimension-one stratum, wherein intersection homology again provides distinct information.

Pair Sets The notion of "information" above is obviously a little vague. To make it precise, we suppose that we are given an ascending and descending filtration of X via the sublevel and superlevel sets of f_v. By doing extended persistence with these filtrations, we get a multi-set of ordered pairs of critical points. Let $H(V)$ and $I^pH(V)$ denote the multisets which result after doing extended persistence for homology and intersection homology with perversity $\bar{p}$, respectively.

As there could well be repeated pairs in these multi-sets, we also let $H(v)$ be the *set* of *unordered* pairs of paired critical points, and we make the same definition for $I^pH(v)$. We call these latter sets the *pair sets*. Essentially, these pair sets contain all possible distance measurements that may be obtained from persistence in this direction. As an obvious consequence of Intersection Pair Group Duality the equality $I^pH(v) = I^qH(v)$ holds for any two dual perversities.

As an illustration of these definitions, consider again the pinched torus with disc attached, embedded as in Fig. 5.2; let $\bar{p} = (-1, 0)$ and v be the vertical direction $(0, 0, 1)$. Then from the detailed discussion in the last chapter, we find:

$$I^pH(v) = \{\{A, G\}, \{B, C\}, \{B, F\}, \{D, E\}, \{F, G\}, \{E, F\}, \{C, D\}, \{A, B\}\}. \quad (5.2)$$

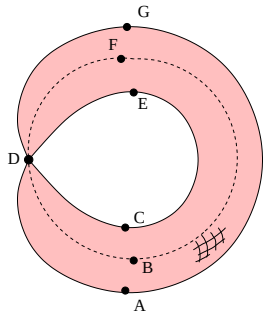

FIGURE 5.2: The Homology pair set in the vertical direction does not include the pairs $\{A, B\}$ or $\{F, G\}$

On the other hand, we could do a similar calculation for homology persistence on this space and direction. We would then find that the homology pair set $H(v)$ fails to include either $\{A, B\}$ or $\{F, G\}$. These two pairs measure the distance between the global minimum (maximum) of the entire space and the global minimum (maximum) of the attached disc; in other words, we gain information ,about the placement of the disc on the torus, that we would have missed using only standard homology persistence.

5.2.1 PseudoSurfaces

A reasonable objection could well be raised to the previous example. Namely, the only new information gained involved a distance measurement between a point on the codimension-one stratum and one on the smooth part. And in fact this same measurement could have been obtained by doing persistence using $H_*(X - S_1)$; that is, standard homology persistence on the stratified space with boundary obtained by removing the 1-stratum. So one might well ask for an example which did not involve codimension-one strata. It turns out that there is no such example in dimension two. There are, however, examples in higher dimensions; one such will be seen later in this chapter.

13 (Pseudosurface Pair Set Inclusion)**.** *Let $X = X_2 \supseteq X_1 = X_0$ be a two-dimensional pseudosurface embedded in $\mathbb{R}^3$, and let $v \in S^2$ be a unit direction vector such that $f_v : X \to \mathbb{R}$ is Stratified Morse.*

Then for the two available perversities $\bar{p} = (-1, 0), \bar{q} = (0, 0)$, we have

$$I^{p}H(v) = I^{q}H(v) \subseteq H(v)$$

.

Before proving the above claim, we first provide an example to show that the above inclusion can certainly be strict. In Fig. 5.3, we see a pinched torus, with no disc attached this time. If we let v be the horizontal direction along the page, then

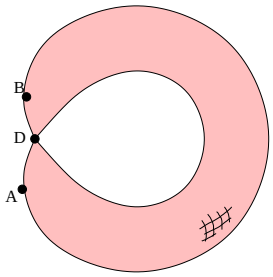

FIGURE 5.3:

point D will simply not be paired by intersection homology persistence: before *and* after passing D, there are two allowable components and no other $\bar{p}$-classes. On the other hand, $H(v)$ contains the pair $\{B, D\}$; hence $I^{\bar{p}}H(v) \neq H(v)$ in this case.

Proof Overview Key to the proof is the fact that $\bar{p}$ will be self-dual for a pseudosurface. Hence by Intersection Pair Group Duality, every $\bar{p}$-pairing occurs an even number of times; that is, the multiset $I^{\bar{p}}H(V)$ contains no element of odd multiplicity. In what follows, we show that $H(V)$ must contain at least one element for each group of repeated elements in $I^{\bar{p}}H(V)$; passing to the sets itself, this will prove the claim.

To do this, we make a list of the relationship between $\bar{p}$-births and $\bar{p}$-deaths, and standard homology births and deaths, at all possible critical points of f_v.

Kernels and Cokernels For any space Y, there are maps $I^{\bar{p}}H_*(Y) \to H_*(Y)$, induced by the chain group inclusions $I^{P}C_*(Y) \to C_*(Y)$. Any inclusion $Y \subseteq Z$ will induce maps on (intersection) homology which commute with the maps above. So in our context, for any critical point w, there are maps $\mathrm{Ker}(I^{\bar{p}}\phi_*(w,v)) \to \mathrm{Ker}(\phi_*(w,v))$, $\mathrm{Ker}(I^{\bar{p}}\psi_*(w,v)) \to \mathrm{Ker}(\psi_*(w,v))$

If w is a smooth point, then $G(w,v), D(w,v)$ are both manifolds. So the maps on kernels above are in this case isomorphisms in each dimension.

So let us assume w is singular, which in this case means $w \in X_0$. Since X is a pseudosurface, $N(w)$ is simply a cone on one or more disjoint circles. Hence $G(w,v)$, which is the intersection of $N(w)$ and the plane $f^{-1}(f(w)-\epsilon)$, will be a disjoint union of circles and arcs; in other words, a 1-manifold, possibly with boundary. Thus $I^{\bar{p}}H_*(G(w,v)) \cong H_*(G(w,v))$, and hence the intersection homology kernels inject into the homology kernels. We now analyze their relationship in each dimension:

- $* = 2$: since $G(w,v)$ is 1-dimensional, it has neither homology nor intersection homology 2-cycles. So all kernels are zero.

- $* = 1$: Suppose $\alpha \in \mathrm{Ker}(\phi_1)$. Then α has a cycle representative which is the boundary of some 2-chain γ in $D(w, v)$. But $D(w, v)$ has only codimnesion-2 singularities and $p_2 = 0$; hence any such γ will be $\bar{p}$-allowable, and thus $\mathrm{Ker}(I^p\phi_1) \cong \mathrm{Ker}(\phi_1)$. The same isomorphism holds between the ψ's.

- $* = 0$: A 1-chain in $D(w, v)$ which intersects the codim-2 singularity will *not* be $\bar{p}$-allowable. Thus, there is an injection, but not necessarily an isomorphism, from the 0-dimensional intersection homology kernels to their homology counterparts.

Of course, an identical discussion holds for the descending filtration defined by $f_{-v} = -f_v$.

Births and Deaths We now give a case-by-case analysis of the different possibilities for births and deaths at a fixed critical point w for the two types of persistence. These statements follow directly from the kernel discussion above.

- If a component is H-born at w, then it is certainly also $\bar{p}$-born at w; the converse need not hold.

- If a component $\bar{p}$-dies at w, then it also dies at w using standard homology; again, the converse need not hold.

- All two-cycle births are identical for the two types of persistence. Of course, these will all be essential two-cycles, for dimensional reasons.

- All 1-cycle deaths are identical.

- If a 1-cycle is $\bar{p}$-born at w, then it will also be born using standard homology. The converse fails. However, we do have the following important lemma concerning birth locations of *essential* 1-cycles.

14 (Lemma)**.** *Let w be a singular point and suppose an essential 1-cycle $\alpha \in H_1(X)$ is born at w. Then $\alpha \notin Im[I^pH_1(X) \to H_1(X)]$.*

Proof:
Let $\tilde{X} \to X$ be the topological normalization of X. For any pseudomanifold, we have $I^pH_*(X) \cong I^pH_*(\tilde{X})$ [22]. But for a pseudosurface, $\tilde{X}$ is just a manifold, and so $I^pH_*(X) \cong H_*(\tilde{X})$. Let N denote the union of the normal slices of all singular points. Then, by a standard result, there is an exact sequence

$$0 \to H_1(\tilde{X}) \to H_1(X) \to H_1(X, \partial N) \to 0 \tag{5.3}$$

Note that ∂N is just the disjoint union of the links of all the singular points. Using the Mayer-Vietorris Sequence for $X_{\leq f_v(w)+\epsilon} = D(w, v) \cup_{G(w,v)} X_{\leq f_v(w)-\epsilon}$, we construct a chain representative for α as follows. There is a 0-cycle $\delta \in C_0(G(w, v))$

which does not bound in $C_0(G(w,v))$. But this 0-cycle *is* the boundary of both a 1-chain $\gamma \in C_1(D(w,v)$ and a 1-chain β in the lower sublevel set. We take $\tau = \gamma + \beta$ as a representative for the 1-cycle born at w.

Suppose τ became bounding when considered as an element of the relative chain group $C_1(X,U)$; put another way, suppose α mapped to zero via the last nonzero map in Seq. 5.3. Then there would be a 2-chain $\sigma \in C_2(X)$ such that $\partial(\sigma) = \tau + \eta$, where $\eta \in C_1(\partial N)$. And so

$$\partial^2(\sigma) = 0 \Longrightarrow \partial\tau = \partial\eta = \delta$$

.

But this contradicts the fact that δ did not bound in $G(w,v)$. The lemma then follows from exactness.

Completion of Proof Suppose that $\{A,B\} \in I^pH(v)$. Without loss of generality, we assume that in fact $(A,B) \in I^pH(V)$; in other words, the point $(f_v(A), f_v(B))$ appears in the intersection homology persistence diagram. We finish the proof by considering every possible subdiagram in which this pair could appear, showing that in each case the same two points will be paired by homology persistence. Note that the subdiagram $I^pOrd_2(v)$ is necessarily empty, since X is only 2-dimensional.

- $(f_v(A), f_v(B)) \in I^pOrd_1(v)$. By the birth-death discussion above, we also must have $(f_v(A), f_v(B)) \in Ord_1(v)$.
- $(f_v(A), f_v(B)) \in I^pOrd_0(v)$: Then by Intersection Pair Group Symmetry, there is a 1 cycle α $\bar{p}$-born at B and $\bar{p}$-dying at A, for the filtration of X defined by $-v$. For the same reasons as above, these events also take place for homology persistence. Translating between the descending filtration and the extended filtration, this means we have $(f_v(B), f_v(A)) \in Rel_2(v)$.
- $(f_v(A), f_v(B)) \in I^pExt_0(v)$. Then by Intersection Pair Group Duality, a 2-dimensional class is also $\bar{p}$-born at A and $\bar{p}$-dies at B, and the same things must happen in homology: $(f_v(A), f_v(B)) \in Ext_2(v)$. Of course, the argument concerning $I^pExt_2(v)$ is even simpler.
- $(f_v(A), f_v(B)) \in I^PExt_1(v)$. Using Lemma 14, we infer that the same point appears in $Ext_1(v)$. This completes the proof.

5.2.2 Example

As promised, we now give an example of an embedded 4-dimensional stratified space, along with a direction v and a perversity $\bar{p}$, such that $I^pH(v) \nsubseteq H(v)$.

Space and Embedding Consider $X = \Sigma(\Sigma(T^2))$, the double suspension of the 2-torus, and put $\bar{p} = (-1, 0, 0, 2), v = (0, 0, 0, 0, 1) \in S^4$.

We embed X in $\mathbb{R}^5$ as follows. First, we place the torus in $\{0\} \times \{0\} \times \mathbb{R}^3$ in the standard manner, making sure to keep it below the hyperplane $x_5 = 1$. This torus is then coned off to the two points $a = (0, -1, 0, 0, 2), b = (0, 1, 0, 0, 3)$, thus forming a suspended torus, which is in turn coned off to $y = (-1, 0, 0, 0, 4), z = (1, 0, 0, 0, 5)$.

The perversity we have chosen does not guarantee stratification independence of the intersection homology groups. So to be more precise, we consider X along with its coarsest stratification:

$$X = X_4 \supseteq X_3 = X_2 = X_1 = \{\Sigma(a) \cup \Sigma(b)\} \supseteq X_0 = \{y, z\} \tag{5.4}$$

Pair Sets We now consider intersection homology persistence for the filtration defined by f_v. For $1 < c < 2$, the sublevel set $X_{\leq c}$ retracts to the torus. Thus, there must be two critical points A, B, both with critical values less than 1, which give birth to the two 1-cycles α, β of the torus; these are both index-1 smooth critical points. Note that these 1-cycles cannot die at any codimension-3 critical point: since $0 > 2 - 3 + p_3$, the cones on these cycles would not be $\bar{p}$-allowable.

On the other hand, consider the codimension-4 critical point y, which is the lower of the two final cone points. The normal slice $N(y) \cong c(\Sigma(T^2))$, while $G(y, v) \cong \Sigma(T^2)$. Since $2 - 4 + p_4 = 0$, the two 1-cycles in $G(y, v)$ can now be allowably coned off to y. Hence the intersection pair set $I^{\bar{p}}H(v)$ contains the two elements $\{A, y\}, \{B, y\}$.

Claim: Neither of these elements appears in $H(v)$.

We show that $\{A, y\}$ can appear neither as an ordinary nor relative nor extended pair; the argument for the other element is identical.

- Ordinary: For the same reasons as given in the pseudosurface discussion, a 1-dimensional homology cycle is also born at A; as A is smooth, this is the only event which happens there. On the other hand, no 1-cycle can die at y, for the simple reason that $H_1(\Sigma(T^2)) = 0$.

- Relative: For the descending filtration defined by $-f_v$, there is only one component born at y. On the other hand, A is an index-3 critical point for $-f_v$, and so it cannot kill a component.

- Extended: Other than the one component at the bottom, all essential classes in X are born at the highest cone point z.

5.3 Elevation Functions

In this section, we define the elevation functions on the smooth part of a d-dimensional stratified space embedded in $\mathbb{R}^{d+1}$. For $d = 1$, we then give a complete characterization of its discontinuities and local maxima. Finally, we also define and characterize a series of "normal circle" functions for each singular point of a 1-dimensional stratified space embedded in the plane.

Upper and Lower Hemisphere As above, we define the height function in direction $v \in S^d$ via $f_v(x) =< x, v >$. Then each $y \in S_{d-k}$ will be critical for the height functions corresponding to a particular k-subsphere of S^d; we call this the *normal sphere* of y and denote it NS_y. In particular, observe that each nonsingular x will be critical for precisely two height functions: the ones defined by the two unit normals to the nonsingular stratum S_d at x.

To differentiate between these two normals, we make an arbitrary division of the unit d-sphere S^d into two antipodal parts, as follows. Thinking of S^0 as the points $1, -1$ on the real line, we define $S^0_+ = \{1\}$. Then viewing S^d as the set of vectors in $\mathbb{R}^{d+1}$ of unit length, we embed S^{d-1} into S^d as the set of points with $x_d = 0$. Finally, we make the inductive definition $S^d_+ = \{x \in S^d | x_d > 0\} \cup S^{d-1}_+$.

Then it is easy to see that when we restrict v to the upper part S^d_+, each nonsingular x will be critical for precisely one height function f_{v_x}. Fix a perversity $\bar{p}$ and assume for the moment that this particular f_{v_x} is Stratified Morse; for almost all unit vectors, it will be. We then consider the extended I^p-persistence diagram for the filtration of X defined by f_{v_x}. The value $f_{v_x}(x)$ will appear as a member of a pair in this diagram exactly twice. In precisely one of these pairs, which we term the ascending $\bar{p}$-pair of x, $f_{v_x}(x)$ will either be the abscissa in an extended pair, or the abscissa or ordinate of an ordinary pair. In either case, let $f_{v_x}(y)$ be the other coordinate of the $\bar{p}$-ascending pair for x. We call y the ascending $\bar{p}$-antipode of x; in symbols, $y = A^p_+(x)$. Note that y could certainly be a singular point. We make a similar definition for the descending $\bar{p}$-antipode of x.

Finally, we define $E^p : X - \Sigma \to \mathbb{R}$ by $E^p(x) = |f_{v_x}(x) - f_{v_x}(A^p_+(x))|$; we call this number the $\bar{p}$-elevation of the nonsingular point x. When the perversity is clear from context, we simply write $E(x)$ and call it the elevation of x.

Suppose that $\bar{q}$ is the dual perversity to $\bar{p}$. Then by using Res. 10 from Chap. 4, one can see that $A^p_-(x) = A^q_+(x)$, and so $E^q(x)$ will measure the distance to the descending $\bar{p}$-antipode of x. Hence we lose no information by restricting to unit vectors in S^d_+. This situation is illustrated in Fig. 5.4, where v_x is the unit vertical direction.

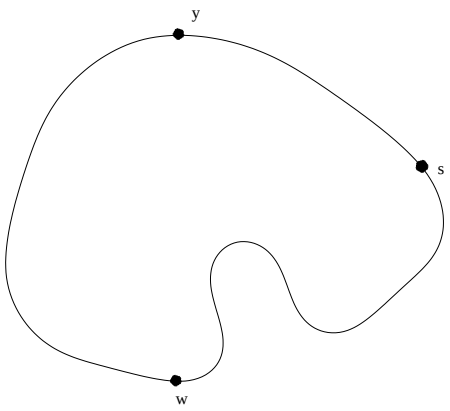

FIGURE 5.4: s is put into the singular 0-stratum, $\bar{p} = (-1), \bar{q} = (0)$. Then $A^p_+(x) = s = A^q_-(x)$, and $A^q_+(x) = w = A^p_-(x)$.

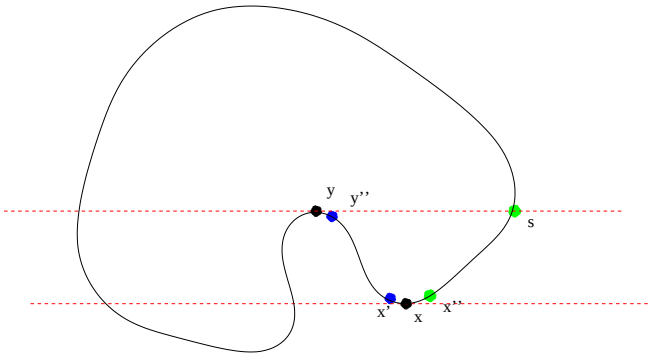

FIGURE 5.5: For perversity $\bar{q} = (0)$, point x is ambiguously paired with the smooth point y and the singular point s. Then the $\bar{q}$-elevation of x is just the vertical distance between x and either of those points. The $\bar{q}$-elevation is not defined at either y or s.

Non-Morse Height Functions and Ambiguous Pairings For certain nonsingular points x, the height function f_{v_x} may fail to be Stratified Morse. For example (see Fig. 5.5), there may be two or more critical points $y_1, y_2, \ldots y_m$, either singular or nonsingular, which are all at the same height. In this case, it may be unclear how to pair x for the filtration defined by f_{v_x}.

But there exists ([24], p.53) an open set of unit vectors v near v_x such that the functions $f_v x$ are Stratified Morse; we imagine perturbing v_x in all directions within this open set. Suppose that, for every such perturbation, a point near x is $\bar{p}$-paired with a point near one of $y_1, y_2, \ldots, y_m$. Then we say that x is *ambiguously paired* with the set of y-points and we set $E^p(x) = |f_{v_x}(x) - f_{v_x}(y_1)|$. Such a point x will be called *m-legged*. On the other hand, we leave the elevation of the legs undefined.

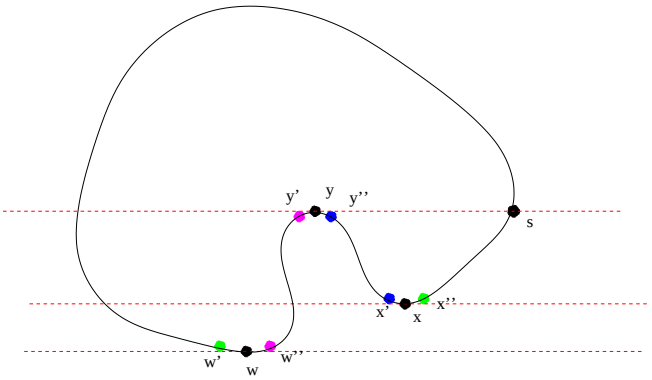

FIGURE 5.6: s is in the singular 0-stratum, $\bar{p} = (-1), \bar{q} = (0)$. $A^q_+(y') = w''$ while $A^q_+(y'') = x'$, and so E^q is not defined at y. On the other hand, E^p *is* defined (and regular) at y.

5.3.1 Dimension One

In this subsection, we give a complete analysis, including a characterization of all possible local maxima, of the elevation functions in the simple case of a 1-dimensional stratified space X embedded in the plane. In addition, we define and then discuss a series of "normal circle" elevation functions for each singular point of X.

Possible Problems When we vary the unit vector v within S^1_+, f_v may fail to be Stratified Morse in the three different ways discussed at the beginning of this chapter. Here we analyze the relationship between these failures and possible discontinuities in elevation.

- Two (or more) critical points share the same height. This happens when there is a line normal to v which contains all the points. Generically, three points should not be collinear. So at worst we expect to find certain pairs of points for which the elevation is not defined. (see Fig. 5.6).

- f_v has a degenerate critical point for its restriction to a particular stratum. In our case, the only problem could arise when this stratum is the nonsingular 1-stratum and f_v is critical at a flat point of X. In this case, varying v in any direction will produce a pair of paired smooth critical points near x. Hence the elevation goes to zero near an inflexion point and so we can simply define the elevation of that point to be zero.

- f_v is overly critical at x. This does not lead to a problem for the elevation functions that we have so far defined. However, it may well necessitate special treatment in the "normal sphere" elevation functions that we will soon define; we defer that discussion until later.

Local Maxima For a generic (no three points cotangent) one dimensional stratified space, there will be two types of local maxima: one-legged and two-legged. Necessary and sufficient conditions for both types are given below.

We fix a perversity and let ρ_x be the radius of the osculating circle at a nonsingular point x.

15 (One-Legged Maxima)**.** *Let x be unambiguously paired with y. Then x is a local maximum for the elevation function iff*

- *y nonsingular $\Rightarrow$ $y - x$ parallel to v_x, and $E(x) \geq \rho_x + \rho_y$.*
- *y singular $\Rightarrow$ $y - x$ parallel to v_x, and $E(x) \geq \rho_x$.*

Proof

Without loss of generality, let $x = (0, 0)$, $v = v_x = (0, 1)$, and $y = (a, b)$. Then $E(x) = \pm b$, depending on whether x is lower than y or vice-versa; we might as well assume x is lower and hence $E(x) = b$.

Let $B(x)$ be a small neighborhood around x and consider the Gauss map $\phi_x : B(x) \to S^1_+$ which takes each point $z \in B_x$ to its normal vector v_z. Flat points have elevation zero and so cannot be local maxima. Thus we can assume that ϕ_x maps $B(x)$ homeomorphically onto an arc neighborhood of $v_x = (0, 1)$: using a parameter $\alpha \in [-\epsilon, \epsilon]$, this arc neighborhood then consists of points $v(\alpha) = (\sin\alpha, \cos\alpha)$. Finally, we denote by $x(\alpha)$ the point $\phi_x^{-1}(v(\alpha))$.

In the case that y is nonsingular, we do an identical construction for $y(\alpha)$. Since x is one-legged, $x(\alpha)$ will be paired with $y(\alpha)$ for each arc parameter value. Using the osculating circle, we get the following approximations, which get better as α gets smaller, for the coordinates of the paired points:

- $x(\alpha) \approx (-\rho_x \sin\alpha, \rho_x(1 - \cos\alpha))$
- $y(\alpha) \approx (a + \rho_y \sin\alpha, b - \rho_y(1 - \cos\alpha))$

Note that this approximation assumes that x and y are index 0 and 1 smooth critical points, respectively. They are paired with each other; so either this assumption or its reverse must be true. Putting $g(\alpha) = E(x(\alpha)) - E(x)$, we investigate under which conditions $g(\alpha) \leq 0$ for all small α; in other words, when x is a local elevation maximum.

Computing and then simplifying, we find $g(\alpha) \approx$

$$< y(\alpha) - x(\alpha), v(\alpha) > -b =$$

$$(1 - \cos\alpha)\left[a\cot\frac{\alpha}{2} + \rho_x + \rho_y - b\right] \tag{5.5}$$

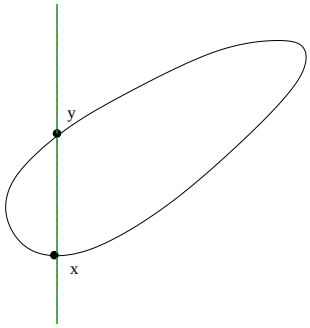

FIGURE 5.7: The normal to x goes through the singular point y. Assuming the curvature condition is satisfied, x will be a one-legged local maximum for $\bar{q}$-elevation, where $\bar{q} = (0)$.

where we have used the half-angle formula for cotangents. For values of α near zero, the term with the cotangent dominates the expression in 5.5 and changes sign at zero. Hence $g(\alpha) \leq 0$ for *all* small values of α requires that $a = 0$, or, in other words, that $y - x = (0, b)$ be parallel to $v_x = (0, 1)$. In which case, we also need the condition $b = E(x) \geq \rho_x + \rho_y$ for x to be an elevation maximum.

On the other hand, suppose that y is singular (see, for example, Fig. 5.7) Then for all values of α, y will be critical for $f_{v(\alpha)}$; for all small values, $x(\alpha)$ will again be paired with y. A similar computation for $h(\alpha) =< x(\alpha) - y, v(\alpha) > -b$ completes the claim.

16 (Two-Legged Maxima)**.** *Suppose x is ambiguously paired with y and z. Then x is a local maximum iff the projection of x onto the line spanned by y and z lies between the two points.*

Proof:

As before, we assume $x = (0, 0), v_x = (0, 1)$, and x is an index-0 smooth critical point. Then since y and z are cotangent and their mutual tangent line is normal to v_x, we have $y = (a, b), z = (d, b)$ and $E(x) = b$.

Define $v(\alpha), x(\alpha)$ as before, as well as $y(\alpha), z(\alpha)$ in case $y.z$ are nonsingular. By definition of a two-legged point, if we move α from zero in the positive direction, $x(\alpha)$ will be paired either with $y(\alpha)$ or $z(\alpha)$; moving α in the negative direction, we get the opposite pairing. So without loss of generality, assume that $x(\alpha)$ is paired with $z(\alpha)$ for all small *positive* α.

Then exactly as above, we compute $E(x(\alpha)) - b \approx$

$$(1 - \cos\alpha)\left[d\cot\frac{\alpha}{2} + \rho_x + \rho_y - b\right] \tag{5.6}$$

Requiring this expression to be negative for all small positive values of α forces $d < 0$. A similar argument concerning $y(\alpha)$ and negative α forces $a > 0$. In case either y or z is singular (see, for example, Fig. 5.6, where x is a local max for the elevation defined by $\bar{q} = (0)$) , a similar argument completes the claim.

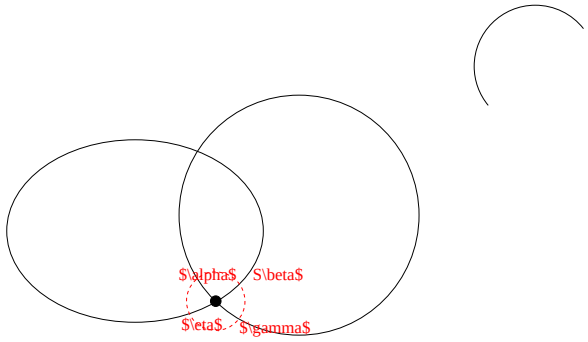

FIGURE 5.8: The four components in the link of a singular point. On the right is the arc $S(\alpha)$.

5.3.2 Normal Circle Elevation Functions

Thus far we have defined a notion of intersection homology elevation only for points in the nonsingular stratum. We now address the 0-dimensional singularities in a 1-dimensional stratified space.

Upper and Lower Links All isolated singularities will be critical points for any height function f_v. For each singular point s, we fix a small enough $\delta > 0$ and observe that the intersection $\partial B_\delta(s) \cap X$, where $B_\delta(s)$ is a ball in the plane, will consist of $m = m(s)$ points; we call these points $P_1, P_2, \ldots, P_m$. In fact, $B_\delta(s) \cap X$ can really just be thought of as $N(s)$, the normal slice at s.

We can then define a series of m elevation functions, each on a different closed arc of S^1, as follows. Consider some $v \in S^1$ and assume our perversity is $\bar{p} = (-1)$ (a very similar discussion can be given for the other perversity). Then when passing s in direction v, one or more components may be born: within $N(s)$, each such component can be represented by P_i, for one and only one $i \in \{1, \ldots, m\}$. In this case, we say that P_i is born at s in direction v. Notice that if P is born in direction v, then it will also be born for all directions in some closed arc neighborhood within S^1 of v; we denote this neighborhood by $S(P)$.

For each $v \in S(P)$, the component represented by P will be killed, either in an ordinary or extended persistence sense, by some other point y, which may be singular or nonsingular. We then define $E_P : S(P) \to \mathbb{R}$, via $E(v) = |f_v(s) - f_v(y)|$; that is, we measure the distance, in direction v, between the two points s and y. Ambiguities in pairing are resolved in the same manner as before; also as before, we say that a vector $v \in S(P)$ is k-legged if the component P is ambiguously paired in direction v with k points. And as before, the elevation of certain vectors must go undefined.

Local Maxima The classification of local maxima for the functions E_P now proceeds as in the nonsingular elevation function case. For one-legged local maxima $v \in S(P)$, the vector $y - s$ must be parallel to v, where y is the paired point; if y is nonsingular, we must also add the same curvature condition as before. The same remarks apply to the two-legged case.

Overly Critical Vectors As promised above, we now discuss what happens to the nonsingular elevation functions when a height function is overly critical. Suppose that x is a *nonsingular* point such that f_{v_x} is overly critical. Then there exists some singular point s such that v_x is normal to a generalized tangent line L at s. If x is paired with a point other than s, then there is nothing to discuss.

Otherwise, suppose that x is paired with some component P born at s. It is then easy to see that such a v must lie in the boundary of the closed arc $S(P)$. So we can assume that $v = (0, 1)$ and that $S(P)$ is an open arc parametrized by small positive vales of α; we define $x(\alpha)$ as before.

Then for positive α, P is still born at s. Hence $E(x(\alpha)) =< x(\alpha) - s, v(\alpha) >$ On the other hand, suppose α is small but negative. Then the component P is in fact born slightly earlier than s; it will be born at some smooth local minimum $s(\alpha)$ near s with normal vector $v(\alpha)$. In this case, $E(x(\alpha)) =< x(\alpha) - s(\alpha), v(\alpha) >$. Hence a nonsingular point with an overly critical normal vector will not cause any problems for the nonsingular elevation function.

6

Local Homology

This chapter puts the concept of local homology into a multi-scale persistence framework. As explained in the Introduction, this will hopefully be the first step towards providing a series of approximate stratifications of an input high-dimensional data set for which the number of degrees of freedom seems to vary across its parameter space.

In Sec. 6.1, we first review the traditional definition of the local homology groups of a space at a point and then proceed to show how the presence of some uncertainty in the space creates a need for a change in this definition. The section closes with the definition of persistent local homology, and the resulting persistence diagrams, within a fixed ball. An algorithm for their computation is detailed in [6], although not included here.

Sec. 6.2 discusses a variety of ways in which these diagrams are stable:

- Within a fixed radius, they are stable under small perturbations of the space. This allows us, under certain conditions, to infer the local homological structure of a space from a finite point sample.
- The diagrams are also stable under small changes in radius. This permits the construction of a series of local homology *vineyards*([13]) which give a multi-scale picture at all radii.
- They are stable under small movement of the center point. We hope this will prove useful in any future algorithm which reconstructs strata.

Finally, Sec. 6.3 gives one possible path towards finding stratifications of a point cloud.

6.1 Local homology

6.1.1 Adapting the Traditional Definition

Let $\mathbb{X}$ be a topological space, along with the subspace topology coming from its embedding in $\mathbb{R}^n$. Fixing a point $z \in \mathbb{R}^n$ for the rest of this section, we let $d_z : \mathbb{R}^n \to \mathbb{R}$ be the function which measures the distance from z: $d_z(x) = ||x - z||$. We write $B_r(z) = d_z^{-1}([0, r])$ and $B^r(z) = d_z^{-1}([r, \infty)$ for the sublevel and superlevel sets defined by r; if z is clear from context, we will drop it from this notation and simply write B_r and B^r.

Traditionally, the kth-dimensional local homology group of $\mathbb{X}$ at z is defined to be the relative homology group $H_k(\mathbb{X}, \mathbb{X} - \{z\})$ [30]. In other words, a local k-cycle γ at z is a k-chain whose boundary misses z. Of course, the boundary of such a γ must also miss some small open set $\mathbb{X} \cap \mathrm{Int}(B_r)$ containing z, that is, γ belongs to $H_k(\mathbb{X}, \mathbb{X} - (\mathbb{X} \cap \mathrm{Int}(B_r))) = H_k(\mathbb{X}, \mathbb{X} \cap B^r)$.

Now for any $s < r$, there are relative homology maps induced by inclusion on the second factor: $H_k(\mathbb{X}, \mathbb{X} \cap B^r) \to H_k(\mathbb{X}, \mathbb{X} \cap B^s)$ Our local cycle γ must lie in the image of these maps for all possible choices of $s < r$.

As a consequence of this discussion, we see that the above definition of local homology at a point z can be recast as the *direct limit* of relative homology groups: $\lim_{r \to 0} H_k(\mathbb{X}, \mathbb{X} \cap B^r)$. Alternatively, we can use excision to see that this is the same as the direct limit:

$$\lim_{r \to 0} H_k(\mathbb{X} \cap B_r, \mathbb{X} \cap \partial B_r), \tag{6.1}$$

where now the maps are given by intersection on the chain level.

Whitney Stratified Spaces If $\mathbb{X}$ happens to be an embedded Whitney Stratified Space, then z has a "small enough" neighborhood $\mathbb{X} \cap B_r$ such that shrinking r any further causes no topological change (Using the notation from Chap. 4, this neighborhood would be $\phi_z(B^i \times c(L_z))$, assuming z is in the i-dimensional stratum). In other words, there is some fixed radius $R = R(z)$ such that Eqn. 6.1 reduces to the single relative homology group $H_k(\mathbb{X} \cap B_R, \mathbb{X} \cap \partial B_R)$; for example, see Fig. 6.1.

Uncertainty On the other hand, suppose there is some uncertainty in our space $\mathbb{X}$. For example, consider the black point z in Fig. 6.2, one of the singular points in the 1-dimensional stratified space $\mathbb{X}$; also shown in the figure is a noisy point sample U, drawn from the space. How are we to define the notion of local homology groups at z in such a way that we can reasonably infer these groups from the point sample alone? Here we discuss this informally; below we give a more rigorous treatment.

The local homology of $\mathbb{X}$ at z is the same as that for two lines crossing at a point; the intersection with $\mathbb{X}$ of the smaller ball B_R in Fig. 6.2 accurately reflects this local homological structure. However, if we thicken $\mathbb{X}$ just a little (alternatively, if

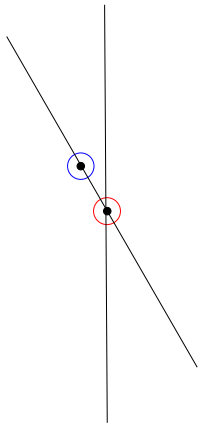

FIGURE 6.1: A 1-dimensional Whitney Stratified Space. The blue and red balls are "small enough" balls to calculate the local homology groups at the smooth point and the singular point, respectively.

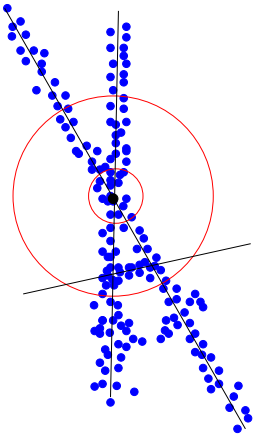

FIGURE 6.2: The inner ball is small enough to reflect the local structure at the black point. However, the noisily sampled points almost fill this ball completely, making it hard to "see" this structure. The point cloud more faithfully represents the homology of the space within the larger ball. But this ball is too large to capture the actual local homological structure at the black point.

we try to use the thickened sample points found in $U \cap B_R$ as a proxy), we lose all resolution: within the small ball, the thickened space, or the thickened points, will be indistinguishable from a disc.

On the other hand, consider the larger ball B_S and the homological structure of the space $\mathbb{X} \cap B_S$. If we thicken $\mathbb{X}$ a small amount as above, we do not lose sight of this structure; in terms of the point cloud, it is quite reasonable to hope to infer the groups $H_*(\mathbb{X} \cap B_S, \mathbb{X} \cap \partial B_S)$ using the available set of points. But there is an obvious tradeoff. Namely, these groups are of course not the traditional local homology groups of $\mathbb{X}$ at z: there are two extra relative 1-cycles which don't exist in the "small enough" ball.

To adapt to this context, we discard the "small enough" radius of the traditional definition, choosing instead to consider the groups $H_*(\mathbb{X} \cap B_r, \mathbb{X} \cap \partial B_r)$ for the entire family of r-balls around the center point. Within each such ball, we allow $\mathbb{X}$ to thicken and look for persistent homological features during this process. We now describe this more rigorously.

6.1.2 Persistent Local Homology within a Fixed Ball

Let Y be either the space $\mathbb{X}$ or the point cloud U. For the moment, we fix a positive radius r and attempt to turn the homological structure of Y within B_r into a multi-scale concept. In other words, we want to capture the important features of the relative homology group $H_k(Y \cap B_r, Y \cap \partial B_r)$, while allowing Y to thicken to make up for uncertainty.

By Y_α and Y^α, we mean the sublevel and superlevel sets, respectively, of the distance function d_Y. For $\alpha < \beta$, the inclusion of pairs $(Y_\alpha \cap B_r, Y_\alpha \cap \partial B_r) \hookrightarrow (Y_\beta \cap B_r, Y_\beta \cap \partial B_r)$ induces a map:

$$H_k(Y_\alpha \cap B_r, Y_\alpha \cap \partial B_r) \to H_k(Y_\beta \cap B_r, Y_\beta \cap \partial B_r), \tag{6.2}$$

while inclusion on the first factor induces:

$$H_k(Y_\alpha \cap B_r) \to H_k(Y_\beta \cap B_r). \tag{6.3}$$

Extended Sequences As α increases, the two types of map above capture the structure of the thickening Y within the fixed r-ball. We now show how they both fit within the theoretical framework of extended persistence (Sec. 2.3); in the next section, this will allow us to prove a variety of inference and stability results.

First consider the restriction of d_Y to B_r. Its sublevel and superlevel sets define an ascending and descending filtration of the r-ball, leading to the following extended sequence, where we first go up with α from 0 to ∞ and then down from ∞ to 0:

$$\begin{aligned} 0 \to H_k(Y_\alpha \cap B_r) \to \ldots \to H_k(B_r) \\ \to H_k(B_r, Y^\alpha \cap B_r) \to \ldots \to 0. \end{aligned} \tag{6.4}$$

The kth persistence diagram for this extended filtration will be denoted $\mathrm{Dgm}_k(d_Y|B_r)$. Note that the maps in (6.3) fit within the top half of Seq. 6.4. On the other hand, the maps in (6.2) form the top half of the following sequence, the persistence diagram for which we denote $\mathrm{Dgm}_k(d_Y|(B_r, \partial B_r))$. Again we first go up with α from 0 to ∞ and then down from ∞ to 0:

$$\begin{aligned} 0 \to H_k(Y_\alpha \cap B_r, Y_\alpha \cap \partial B_r) \to \ldots \to H_k(B_r, \partial B_r) \\ \to H_k(B_r, \partial B_r \cup (Y^\alpha \cap B_r)) \to \ldots \to 0. \end{aligned} \tag{6.5}$$

Equivalence of diagrams. The persistence diagrams corresponding to the two extended sequences above contain the same information. Specifically, one may establish isomorphisms between the homology groups in Seq. 6.4 and Seq. 6.5 which show that the corresponding pairings are dual and thus give the same diagrams.

The top half of Seq. 6.4 and the bottom half of Seq. 6.5 are paired by Lefschetz Duality, after excision. These pairings are perfect and are compatible with the maps. The remaining halves require some more effort.

To shorten the notation and clarify the relations we set $X = Y_\alpha \cap B_r$, decompose its boundary $\partial X = F \cup G$ where $F = Y_\alpha \cap Y^\alpha \cap B_r$ and $G = \partial B_r \cap Y_\alpha$, and set $A = F \cap G$. Generically, X is an n-manifold with boundary, F and G are $(n-1)$-manifolds with boundary, and A is an $(n-2)$-manifold without boundary. We use excision to rewrite (6.4) and (6.5), running them anti-parallel against each other:

$$\begin{array}{ccccc} \to & H_{n-k}(X) & \to & H_{n-k}(X, F) & \to \\ & \otimes & & \otimes & \\ \leftarrow & H_k(X, \partial X) & \leftarrow & H_k(X, G) & \leftarrow \\ & \downarrow & & \downarrow & \\ & \mathsf{Z}/2\mathsf{Z} & & \mathsf{Z}/2\mathsf{Z} & \end{array}$$

The first vertical pairing was addressed above. The other vertical pairing is also perfect, as we now show.

17 (Isomorphism Lemma)**.** *For every dimension k, the intersection pairing on X induces a perfect pairing*

$$H_{n-k}(X, F) \otimes H_k(X, G) \to \mathsf{Z}/2\mathsf{Z}.$$

Proof. First notice that by excision, the relative homology groups $H_k(\partial X, G)$ and $H_k(F, A)$ are isomorphic. Next consider the exact cohomology sequence of the pair (X, F), shown in the bottom row in Table 6.1, and the exact homology sequence of the triple $(X, \partial X, G)$,

$$\begin{aligned} \to H_{k+1}(X, \partial X) \to H_k(\partial X, G) \to H_k(X, G) \\ \to H_k(X, \partial X) \to H_{k-1}(\partial X, G) \to \end{aligned}$$

$$\begin{array}{ccccccccccc}
\to & H_{k+1}(X,\partial X) & \to & H_k(F,A) & \to & H_k(X,G) & \to & H_k(X,\partial X) & \to & H_{k-1}(F,A) & \to \\
 & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & \\
\to & H^{n-k-1}(X) & \to & H^{n-k-1}(F) & \to & H^{n-k}(X,F) & \to & H^{n-k}(X) & \to & H^{n-k}(F) & \to
\end{array}$$

Table 6.1: Commuting diagram with isomorphisms between the terms in the exact homology sequence of the triple $(X, \partial X, G)$ at the top and the exact cohomology sequence of the pair (X, F) on the bottom.

Replacing $H_k(\partial X, G)$ by $H_k(F, A)$ we get the diagram in Table 6.1. Here each vertical arrow is the Poincaré-Lefschetz duality map defined by $\gamma \to f_\gamma$ where $f_\gamma(\delta) = \gamma \cdot \delta$, the intersection number between the two classes. It is not difficult to check that this diagram commutes. The two vertical maps on the left and the two vertical maps on the right are isomorphisms by Poincaré-Lefschetz duality. The Steenrod Five-Lemma then tells us that the center vertical map is also an isomorphism [30]. Finally we note that the Poincaré duality map $H_k(X, G) \to H^{n-k}(X, F)$ being an isomorphism implies that the intersection pairing is perfect. The claim follows. ◻

We see that the pairings between the groups in (6.4) and (6.5) are perfect and all diagrams commute. It follows that if we use the superscript T to denote reflection across the diagonal we have $\mathrm{Dgm}_k(d_Y|B_r) = \mathrm{Dgm}^T_{n-k}(d_Y|(B_r, \partial B_r))$ for all dimensions k and all radii r.

6.2 Stabilities, Inference, and Vineyard

6.2.1 Local Homology Inference

In this subsection, we prove that even with rather mild assumptions on the sampling of a space it is possible to infer its local homology within certain fixed radii. Perhaps more important than the guaranteed recognition is the interpretation of our result as describing the set of spaces that can possibly give rise to the sample.

Assumptions on Space and Sample The data we consider is a finite set of points, $U \subseteq \mathbb{R}^n$. It will be convenient to index the points in this set as u_i. We assume that U is sampled from or near a compact space $\mathbb{X} \subseteq \mathbb{R}^n$. For example, $\mathbb{X}$ may be a compact Whitney stratified space but the existence of a stratification will play no role in what we prove in this subsection. It will, however, be important that the diagram of the restricted distance functions of $\mathbb{X}$ be stable. We therefore assume that $d_{\mathbb{X}}|B_r$ is tame for every $z \in \mathbb{R}^n$ and every ball B_r centered at z. Recall that the space $\mathbb{X}$ is unknown and the main question we ask is how much we can find out about $\mathbb{X}$ under what assumptions relating U with $\mathbb{X}$.

We will assume that U is an ε-sample of X; recall this means the Hausdorff distance between the space and the point set is smaller than ϵ, or equivalently, that $||d_X - d_U||_\infty < \varepsilon$.

Homological critical Values Consider $d_{\mathbb{X}}|B_r$, the restriction to the r-ball of the distance function from the space $\mathbb{X}$, with its sublevel and superlevel sets defined as above. As in Chap. 2, we say that α is an *absolute homological regular value* if there exists $\delta > 0$ such that the inclusion

$$(\mathbb{X}_{\alpha-\delta} \cap B_r) \hookrightarrow (\mathbb{X}_{\alpha+\delta} \cap B_r)$$

induces homology isomorphisms in all dimensions; otherwise α is an absolute homological *critical* value.

In a similar fashion, we say that α is a *relative* homological regular value if there exists $\delta > 0$ such that the inclusion of pairs

$$(B_r, \mathbb{X}^{\alpha+\delta} \cap B_r) \hookrightarrow (B_r, \mathbb{X}^{\alpha-\delta} \cap B_r)$$

induces isomorphisms on relative homology in all dimension; otherwise, it is a relative homological critical value.

Resolution. Fixing a point z, a radius r, and a dimension k, we consider the sequence (Seq. 6.4), which we reprint here:

$$\begin{aligned} 0 \to H_k(Y_\alpha \cap B_r) \to \ldots \to H_k(B_r) \\ \to H_k(B_r, Y^\alpha \cap B_r) \to \ldots \to 0, \end{aligned}$$

where Y is either U or $\mathbb{X}$. Recall that the bottom half of this sequence captures the persistence information from Seq. 6.2.

For each radius $r > 0$ we thus consider the series of persistence diagrams $\mathrm{Dgm}(d_Y|B_r)$. The only non-trivial homology group of B_r is in dimension zero and this group has rank one. There is therefore only one extended point in this series tracking the first component that appears in the filtration.

To determine the local homological structure of $\mathbb{X} \cap B_r$ at a point z from the sample U it is necessary that the points sample all relevant features of the space within the ball finely enough to be recognized. To make this precise, we require:

18 (Definition)**.** *A radius* r resolves $\mathbb{X}$ at z to ε *if the smallest positive absolute or relative homological critical value of* $d_{\mathbb{X}}|B_r$ *exceeds* 3ε*.*

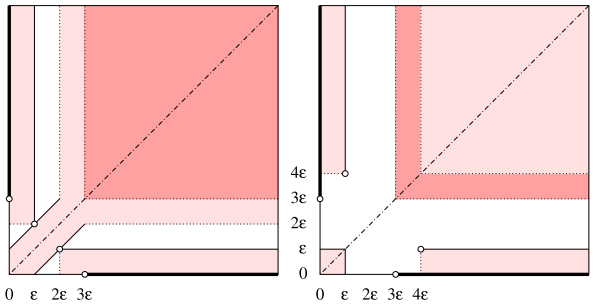

FIGURE 6.3: Left: the dark regions contain the persistence diagram of $d_{\mathbb{X}}$ for every radius $r \in R_{\mathbb{X}}(\varepsilon)$ and the light regions expand them to contain the persistence diagram of d_U for every radius $r \in R'_U(\varepsilon)$. Right: the light regions contain the persistence diagram of d_U for every radius $r \in R''_U(\varepsilon)$ and the dark regions contain the persistence diagrams of the distance function of U_ε.

For a radius r that resolves $\mathbb{X}$ to ε there are no births and no deaths in the interval $(0, 3\varepsilon]$. In other words, the corridors separating the two boldface segments from the dark square in Figure 6.3, left, are empty. It follows that everything born at $\alpha = 0$ lives for a while and if it dies on the way up, as α increases, then it dies strictly after 3ε. Symmetrically, everything that dies at $\alpha = 0$ must have lived for a while and if it was born on the way down, as α decreases, then it was born strictly before 3ε. Radii that have this property are of special interest, so we define $R_{\mathbb{X}}(\varepsilon)$ as the set of radii r for which the points in $\mathrm{Dgm}(d_{\mathbb{X}}|B_r)$ all lie in the dark portion of Figure 6.3, left, which includes the vertical segment with lower endpoint $(0, 3\varepsilon)$, the horizontal segments with left endpoint $(3\varepsilon, 0)$, and the quadrant $(3\varepsilon, \infty) \times (3\varepsilon, \infty)$.

Inference The inference theorem below will be phrased in terms of the number of points in certain regions of persistence diagrams. To ease exposition, we first define some shorthand. Let $a = (\xi, \zeta)$, with $0 \leq \xi \leq \zeta$. Consider the shaded regions in Fig. 6.4. With Y being either $\mathbb{X}$ or U, we set $\#_k^a(d_Y|B_r)$ to be the number of ordinary and relative points in the dark shaded region, plus the number of extended points in the light shaded region, of the kth persistence diagram $\mathrm{Dgm}_k(d_U|B_r)$; of course, we count points with their appropriate multiplicity. Note that these points correspond to homology classes which live throughout the entire interval (ξ, ζ).

For example, if a is the origin, 0, then $\#_k^a(d_{\mathbb{X}}|B_r)$ counts the points on the horizontal Birth-axis and the vertical Death-axis of $\mathrm{Dgm}_k(d_{\mathbb{X}}|B_r)$. These points correspond to the actual homological structure, in dimension k, of the space within the ball.

19 (Local Homology Inference Theorem)**.** *Let $\varepsilon > 0$, $\mathbb{X}$ a compact space, U an ε-approximation of $\mathbb{X}$, z a point in $\mathbb{R}^n$, and k a nonnegative integer. Then $\#_k^0(d_{\mathbb{X}}|B_r) = \#_k^{(\varepsilon,2\varepsilon)}(d_U|B_r)$ for every radius $r \in R_{\mathbb{X}}(\varepsilon)$.*

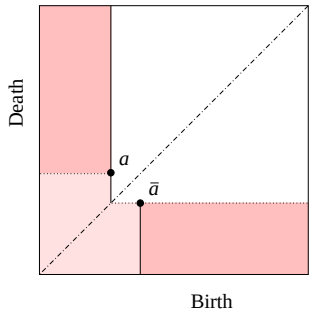

FIGURE 6.4: $a = (\xi, \zeta)$, and $\bar{a} = (\zeta, \xi)$

Proof. We will prove $R_{\mathbb{X}}(\varepsilon) \subseteq R'_U(\varepsilon)$, where the latter set consists of all radii r for which the points in $\mathrm{Dgm}(d_U|B_r)$ all lie in the shaded portion of Figure 6.3, left, which expands the dark regions and the diagonal by ε in the vertical as well as the horizontal direction. We will see that this containment of sets implies the claimed equation.

Since $r \in R_{\mathbb{X}}(\varepsilon)$ we have $\#^0(d_{\mathbb{X}}|B_r) = \#^a(d_{\mathbb{X}}|B_r)$ for every $a \in [0, 3\varepsilon]^2$. Since $\|d_U - d_{\mathbb{X}}\|_\infty \leq \varepsilon$, the Stability Theorem of extended persistence implies a bijection such that each point in $\mathrm{Dgm}(d_{\mathbb{X}}|B_r)$ lies within L_∞-distance ε from its corresponding point in $\mathrm{Dgm}(d_U|B_r)$. This implies that all points of $\mathrm{Dgm}(d_U|B_r)$ lie inside the ε-expanded region depicted in Figure 6.3, left. This region consists of three disjoint subregions, one expanding the vertical segment, one expanding the horizontal segment, and the third expanding the quadrant that contains the remaining points of $\mathrm{Dgm}(d_{\mathbb{X}}|B_r)$ as well as the diagonal. By disjointness of the three subregions, the points of $\mathrm{Dgm}(d_{\mathbb{X}}|B_r)$ in the two segments cannot map to any points other than the ones in the subregions that expand them. The points of $\mathrm{Dgm}(d_U|B_r)$ in these two subregions are counted by $\#^a(d_U|B_r)$ with $a = (\varepsilon, 2\varepsilon)$. This implies the claimed equality. ◻

Inverse Recall that U is known but $\mathbb{X}$ is not. Given an actual data set U, we envision using the Local Homology Inference Theorem by identifying radii r for which the white corridors in the left portion of Fig. 6.3 are empty. For each such r there is a chance that it belongs to $R_{\mathbb{X}}(\varepsilon)$ and if it does we know the homological structure of $\mathbb{X}$ within the ball of this radius r. The trouble is that we can generally not be sure that r really belongs to $R_{\mathbb{X}}(\varepsilon)$. However, we can further restrict the regions that contain the points of $\mathrm{Dgm}(d_U|B_r)$ so that they imply the existence of a space $\mathbb{X}$ for which U is an ε-approximation and r is in $R_{\mathbb{X}}(\varepsilon)$. Let $R''_U(\varepsilon)$ be the set of radii r for which the points in $\mathrm{Dgm}(d_U|B_r)$ are contained in the light shaded region in Figure 6.3, right.

20 (Inverse LHI Theorem)**.** *Let $\varepsilon > 0$, U a subset of $\mathbb{R}^n$, and z a point in $\mathbb{R}^n$. Then there exists a compact space $\mathbb{X} \subseteq \mathbb{R}^n$ for which U is an ε-approximation and $R''_U(\varepsilon) \subseteq R_{\mathbb{X}}(\varepsilon)$.*

Proof. Set $\mathbb{X} = U_\varepsilon$ and note that U is an ε-approximation of $\mathbb{X}$. The distance function defined by $\mathbb{X}$ is $d_{\mathbb{X}}(x) = \max\{0, d_U(x) - \varepsilon\}$. It follows that each birth and each death happens at 0 or ε earlier than before. The corresponding transformation of persistence diagrams is a shift by ε down and a shift by ε to the left, except that a movement stops before the point enters the negative regions of birth or of death. If $r \in R''_U(\varepsilon)$ then all points in the diagrams of $\mathbb{X}$ lie on the two segments and the quadrant that define $R_{\mathbb{X}}(\varepsilon)$. ◻

6.2.2 r-stability and the Local Homology Vineyard

The discussion above fixes a particular radius r and then attempts to use the persistent homology of $U \cap B_r$ to infer the structure of $\mathbb{X} \cap B_r$ On the other hand, we have no way of knowing in advance which radius to fix. Nor, given the sample but not the space, will we really have any idea which radii are resolving and which are not. Thus it is important to consider the diagrams $\mathrm{Dgm}(d_Y|B_r)$ for the entire family of radii, where again Y can be either $\mathbb{X}$ or U. Fortunately, these diagrams vary continuously with r:

21 (r-Stability Lemma)**.** *Let $Y \subseteq \mathbb{R}^n$ and $z \in \mathbb{R}^n$ such that the restriction of $d_Y : \mathbb{R}^n \to \mathbb{R}$ to any ball centered at z is tame. Then the bottleneck distance between the series of persistence diagrams for two radii $r \leq r'$ is bounded by the difference between the radii:*

$$d_B(\mathrm{Dgm}(d_Y|B_r), \mathrm{Dgm}(d_Y|B_{r'})) \leq r' - r$$

Proof. Without loss of generality, we assume that z is located at the origin; note that this change of coordinates does not affect the persistence diagrams. Letting $f, g : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = d_Y(rx)$ and $g(x) = d_Y(r'x)$, the restrictions of d_Y correspond to the restrictions of f and g to the unit ball B_1. We now show that the L_∞ distance between these two functions is bounded from above by $r' - r$. The extension of the Stability Theorem ([11]) to extended persistence, as described in [10], will then imply the claim.

Fix an $x \in B_1$, and let $y, y' \in Y$ be the nearest neighbors in Y to $rx, r'x$, respectively. Then we have:

$$f(x) = d_Y(rx) = ||y - rx|| \leq ||y' - rx||$$

and

$$g(x) = d_Y(r'x) = ||y' - r'x|| \leq ||y - r'x||$$

Consider the difference $|f(x)-g(x)|$. Using either the first or the second inequality above, depending on which of $f(x)$ or $g(x)$ is greater, we find:

$$|f(x) - g(x)| \leq ||x||(r' - r)) \tag{6.6}$$

But $||x|| \leq 1$. Hence the claim follows by maximizing Eqn. 6.6 over the unit ball. ◻

Local Homology Vineyards The r-stability result suggests we vary r within $[0, \infty)$ and describe the homology in the neighborhood of $z \in \mathbb{R}^n$ by the resulting 1-parameter family of persistence diagrams. Stacking up the diagrams in $\mathbb{R}^3$ using r as the third coordinate, each point sweeps out a curve which we refer to as a *vine*. Together the vines form a collection of curves which we refer to as the *vineyard* of the two distance functions d_Y, d_z; see [13]. Specifically, we denote the vineyard obtained by stacking up the dimension k persistence diagrams by $\mathrm{Vnrd}_k(d_Y|d_z)$ and the series of vineyards by $\mathrm{Vnrd}(d_Y|d_z)$. On occasion we call this the series of $(\alpha|r)$-*vineyards* thus emphasizing that the each diagram is obtained by varying the threshold α for the distance to Y while fixing the threshold r for the distance to z, and that the vines are then obtained by varying r.

Multi-scale example. Observe that the Local Homology Inference Theorem describes the relationship between the persistence diagrams of $\mathbb{X}$ and of U for a fixed radius r. It is difficult to know ahead of time which value of r is most appropriate and in many situations it is not even desirable to make a choice. We cope with this difficulty by examining the persistent behavior across all radii. We use the example in Figure 6.5 to illustrate what we have in mind.

Here $\mathbb{X}$ is a one-dimensional space embedded in $\mathbb{R}^2$. It consists of a string of loops, each connected to the loop before and the loop after. Its dimension 1 vineyard at the point z contains a prominent vine that has high persistence across all values of r. This vine tracks a dimension 1 relative homology class and corresponds to the chain itself which, from a distance, may be seen as a single curve. It can be detected even for rather sparse samples. Furthermore, the vineyard contains two small vines per loop, one tracking a relative and the other an absolute homology class. The relative class emerges at the moment the ball B_r first intersects the loop. It attains its largest persistence when B_r reaches the maximum near the center of the loop after which time the corresponding point in the diagram stops moving and sweeps out a vertical vine. At the same moment the absolute class emerges and attains its largest persistence when B_r reaches the other end of the loop after which time the corresponding point stops moving and sweeps out a vertical vine, as before.

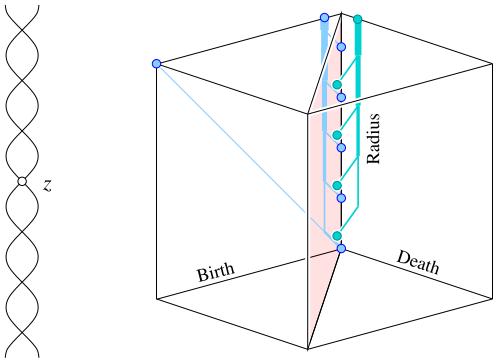

FIGURE 6.5: Left: the one-dimensional chain of loops, $\mathbb{X}$. Right: the dimension 1 $(\alpha|r)$-vineyard of $\mathbb{X}$ at z. The only significantly persistent vine runs roughly diagonally in the Birth-Radius plane and tracks a relative 1-cycle. All other vines run near the Radius-axis and track classes caused by the loops in the chain.

As the word "local" suggests, we are primarily interested in small values of r, that is, the lower portion of the vineyard. Of course, what small means is in the eye of a beholder. On the other hand, the Local Homology Inference Theorem and its inverse can be used to make informed guesses. If the space $\mathbb{X}$ in Figure 6.5 is sampled sufficiently densely, then small values of r resolve it, and we are able to detect the three dimension 1 cycles in the local homology of z. Specifically, there are three vines emerging from the origin, each tracking a relative homology class. If the sampling is not sufficiently dense then we cannot distinguish $\mathbb{X}$ from a 1-manifold. Indeed, an arc passing through the vertices joining the loops could conceivably produce the same sample.

z-stability We close this section with a proof that the local homology persistence diagrams are stable under small movements of the center point z.

Fix a radius r and consider two points $z, z' \in \mathbb{R}^n$. The restrictions $d_Y|B_r(z)$ and $d_Y|B_r(z')$ are identical to the restrictions to the unit ball of $f(x) = d_Y(z + rx)$ and $g(x) = d_Y(z' + rx)$, respectively. By the same reasoning as in the proof of (21) above, we find $\|f - g\|_\infty \leq ||z - z'||$, which implies:

22 (z-Stability)**.** *Let $Y \subseteq \mathbb{R}^n$ and $z, z' \in \mathbb{R}^n$ such that the restriction of $d_Y : \mathbb{R}^n \to \mathbb{R}$ to any ball centered either at z or z' is tame. Then, for each r, we have:*

$$d_B(\mathrm{Dgm}(d_Y|B_r(z)), \mathrm{Dgm}(d_Y|B_r(z')) \leq \|z - z'\|.$$

6.3 Towards Finding Strata

Finally, we put down here some very preliminary thoughts as to how to stratify a point cloud.

Strata on Different Scales If $\mathbb{X}$ is any compact space, then it has a coarsest stratification (see [32], for example) given in the following way. First define $p \sim q$ iff each point has a local neighborhood N_p, N_q in $\mathbb{X}$ and there is a homeomorphism $N_p \cong N_q$ taking p to q; for example, N_p might be the "small enough" neighborhood $\mathbb{X} \cap B_R(p)$ alluded to at the end of SubSec. 6.1.1. Then the pieces in the coarsest stratification of $\mathbb{X}$ will then just be the connected components of the $\sim$-equivalence classes.

Of course, we do not dare to compute homeomorphism type, so we must content ourselves with a weaker definition that uses only homology. Furthermore, we want a definition which allows for uncertainty at small radii.

Thus, taking inspiration from the idea of a homology stratification ([32]), we might make the following attempt at a definition. For each radius $r > 0$, define an equivalence relation $\sim_r$ on $\mathbb{R}^n$ as follows: $p \sim_r q$ iff there exists a chain of points $p = z_0, z_1, \ldots, z_n = q$ such that the maps:

$$\begin{array}{c} H_*(\mathbb{X} \cap B_r(z_i), \mathbb{X} \cap \partial B_r(z_i)) \\ \cong \downarrow \\ H_*(\mathbb{X} \cap B_r(z_i) \cap B_r(z_{i+1}), \mathbb{X} \cap \partial B_r(z_i) \cap \partial B_r(z_{i+1})) \\ \cong \uparrow \\ H_*(\mathbb{X} \cap B_r(z_{i+1}), \mathbb{X} \cap \partial B_r(z_{i+1})) \end{array}$$

induced by chain intersection, as well as the reverse maps on absolute homology induced by inclusion on the first factor, are isomorphisms in all dimensions for all $0 \leq i \leq n-1$. We might then define the "pieces at scale r" to be the $\sim_r$ equivalence classes. Fig. 6.6 illustrates this idea, as well as the kernel/cokernel persistence idea below.

Kernel/Cokernel Persistence Now suppose U is a point sample taken from $\mathbb{X}$. How are we to infer the pieces at scale r from this sample? One preliminary attempt might be to use the kernel/cokernel persistence algorithms, recently developed in [12], to mimic the maps in the diagram above.

More precisely, for each α and each i, consider the diagram:

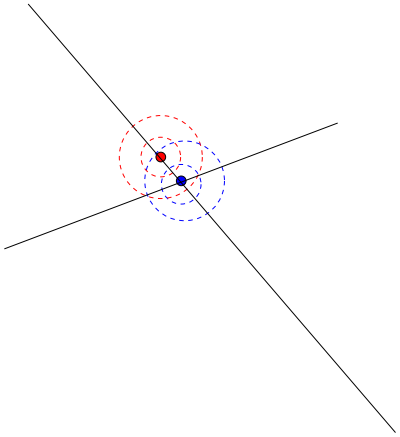

FIGURE 6.6: We compare the red point to the blue point. At the scale of the larger dashed balls, they seem to be in the same piece, while this is not the case at the smaller scale shown. On the other hand, there will be an early birth and fairly late death cokernel class for the map coming from the larger red ball.

$$
\begin{array}{c}
H_*(U_\alpha \cap B_r(z_i), U_\alpha \cap \partial B_r(z_i)) \\
\downarrow{\scriptstyle \phi_\alpha} \\
H_*(U_\alpha \cap B_r(z_i) \cap B_r(z_{i+1}), U_\alpha \cap \partial B_r(z_i) \cap \partial B_r(z_{i+1})) \\
\uparrow{\scriptstyle \psi_\alpha} \\
H_*(U_\alpha \cap B_r(z_{i+1}), U_\alpha \cap \partial B_r(z_{i+1}))
\end{array}
$$

For any $\alpha < \beta$, there are maps $\mathrm{Ker}(\phi_\alpha) \to \mathrm{Ker}(\phi_\beta)$, $\mathrm{Coker}(\phi_\alpha) \to \mathrm{Coker}(\phi_\beta)$, and similar maps for the ψ's. Hence we can do persistence using these maps, as α increases. In this context, an "isomorphism" would correspond to any kernel/cokernel classes having very low persistence.

Of course, it is not all clear what persistence threshold to set for these classes, but it certainly should depend in some way on the radius. The $z - \mathit{Stability}$ result (Res. 22) seems to imply that any such threshold need not vary too much as we move the center point from p to q.

Bibliography

[1] P.K. Agarwal, H. Edelsbrunner, J. Harer, and Y. Wang. Extreme elevation on a 2-manifold. In *Proc. 20th Ann. Sympos. Comput. Geom.*, pages 357–365, 2004.

[2] S. Ahnert, M. Dequeant, H. Edelsbrunner, T. Fink, E. Glynn, G. Hattem, H. Kudlicki, Y. Mileyko, J. Morton, A. Mushegian, L. Pachter, O. Pourquie, M. Rowicka, A. Shiu, and B. Sturmfels. Comparision of pattern detection models applied to the temporal transcription program of the mouse segmenation clock. In preparation.

[3] Y.-H. A. Ban, H. Edelsbrunner, and J. Rudolph. Interface surfaces for protein-protein complexes. In *Proc. 8th Intl. Conf. Res. Comput. Mol. Bio.*, pages 205–212, 2004.

[4] M. Banagl. Extending intersection homology type invariants to non-Witt spaces. In *Memoirs of the Amer. Math. Soc.*, volume **160**, 2002.

[5] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. In *Neural Computation*, pages 347–356, 2003.

[6] P. Bendich, D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Inferring local homology from sampled stratified spaces. In *Proc. 48th Ann. Sympos. Found. Comp. Sci.*, pages 536–546, 2007.

[7] K. Brown and K. Knudson. Non-linear statistics of human speech data. Preprint.

[8] G. Carlsson, T. Ishkhanov, V. De Silva, and A. Zomorodian. On the local behavior of spaces of natural images. *Internat. J. Comput. Vision*, 2006.

[9] G. Carlsson, T. Ishkhanov, F. Memoli, D. Ringach, and G. Sapiro. Topological analysis of the responses of neurons in V1. (in preparation).

[10] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Extending persistence using Poincare and Lefschez duality. *Found. Comput. Math.* to appear.

[11] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. In *Proc. 21st Ann. Sympos. Comput. Geom.*, pages 263–271, 2005.

[12] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Persistent homology for kernel, images, and cokernels. In preparation.

[13] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and vineyards by updating persistence in linear time. In *Proc. 22nd Ann. Sympos. Comput. Geom.*, pages 119–126, 2006.

[14] H. Edelsbrunner. The union of balls and its dual shape. In *Proc. 9th Ann. Sympos. Comput. Geom.*, pages 218–231, 1993.

[15] H. Edelsbrunner and J. Harer. Persistent homology-a survey. In J.E. Goodman, J. Pach, and R. Pollack, editors, *Twenty Years After*. AMS. to appear.

[16] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci. Morse-Smale complexes for piecewise linear 3-manifolds. In *Proc. 19th Ann. Sympos. Comput. Geom.*, pages 361–370, 2003.

[17] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, **28**:511–533, 2002.

[18] G. Friedman. Straified fibrations and the intersection homology of regular neighborhoods of bottom strata. *Topology and its Applications*, **134**, 2003.

[19] G. Friedman. Intersection homology and Poincaré duality for homotopically stratified spaces. *Arxiv*, 2007.

[20] R. Ghrist and V. De Silva. Coordinate-free coverage in sensor networks with controlled boundaries via homology. submitted.

[21] R. Ghrist and V. De Silva. Coverage in sensor networks via persistent homology. *Alg. Geo. Topology*, **7**:339–358, 2007.

[22] M. Goresky and R. MacPherson. Intersection homology I. *Topology*, **19**:135–162, 1980.

[23] M. Goresky and R. MacPherson. Intersection homology II. *Inv. Math.*, **71**:77–129, 1983.

[24] M. Goresky and R. Macpherson. *Stratified Morse Theory*. Springer-Verlage, Berlin, 1987.

[25] N. Habegger and L. Saper. Intersection cohomology of cs-spaces and Zeeman's filtration. *Inv. Math.*, **105**:247–272, 1991.

[26] B. Hughes and S. Weinberger. *Surveys on Surgery Theory*, chapter Surgery and Stratified Spaces, pages 311–342. Princeton Univ. Press, 2000.

[27] H. King. Topological invariance of intersection homology without sheaves. *Topology Appl.*, **20**:149–160, 1985.

[28] F. Kirwan and J. Woolf. *An introduction to intersection homology theory*. CRC Press, 2006.

[29] J. Milnor. *Morse Theory*. Princeton Univ. Press, Princeton, New Jersey, 1963.

[30] J.R. Munkres. *Elements of algebraic topology*. Addison-Wesley, Redwood City, California, 1984.

[31] M. Pflaum. *Analytic and geometric study of stratified spaces*. Springer-Verlage, Berlin, 2001.

[32] C. Rourke and B. Sanderson. Homology stratifications and intersection homology. *Geometry and Topology*, **2**:455–472, 1999.

[33] J. Tenenbaum, V. De Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, **290**:2319–2323, 2000.

[34] Y. Wang, P.K. Agarwal, P. Brown, H. Edelsbrunner, and J. Rudolph. Coarse and reliable geometric alignment for protein docking. In *Proc. Pacific Sympos. Biocomput.*, pages 65–75, 2005.

Biography

Paul Bendich was born in New York City on May 17, 1979. He received an M.A in Mathematics from Duke University in 2005, a B.A. in Physics from Grinnell College in 2001, and two General Equivalency Diplomae, one from New York and one from New Jersey, in 1997.

During his time at Duke, he was coauthor of the paper "Inferring Local Homology from Sampled Stratified Spaces" and the conference paper "A Probabilistic Perspective on Persistence Homologies." In addition, he received the L.P. and Barbara Smith Award for Teaching Excellence, as well as first place in the weekly shufflepuck tournament at the Green Room, a local dive bar. In August 2008, he will begin a job at Pennsylvania State University.