Math 583B: Topological Data Analysis

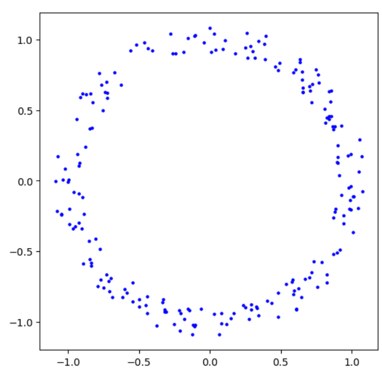
Kyle Ormsby

Spring 2024

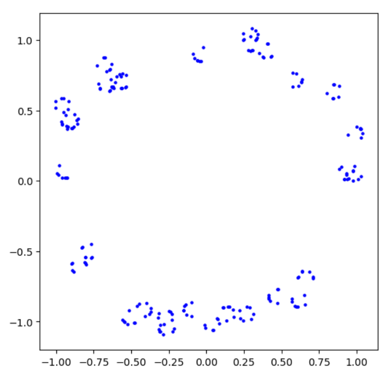
1 *Inferring the shape of data* — 25 March 2024

Imagine that you're running an experiment in which you measure a large number — say N — of real-valued variables with each observation. Each observation is then a point in $\mathbb{R}^N$, and if you make k total observations, then the data associated with your experiment is a *point cloud* $P = \{x_1, x_2, \ldots, x_k\} \subseteq \mathbb{R}^N$.

If the system being observed is not purely random, then — up to issues of noise and accuracy — we expect P to be sampled from a subspace $M \subseteq \mathbb{R}^N$. How might we infer the structure and shape of M from P, at least under the assumption that k is relatively large? This is one of the questions that topological data analysis (TDA) aims to answer, at least for particular notions of "structure" and "shape". In the figure presented here, we see a point cloud P in $\mathbb{R}^2$ sampled with noise from the unit circle $S^1 \subseteq \mathbb{R}^2$, and we seek algorithmic methods that will recognize (features of) S^1 as the underlying space from which P is sampled. Of course, in practice, N might be very large, and it is unlikely that your visual cortex will rise to the challenge of guessing M.



But even for small N, we can still ask more from our methods. Consider the displayed point cloud $Q \subseteq \mathbb{R}^2$ which exhibits strikingly different structure at different scales. At small scales, points seem to be sampled from disjoint circles. After zooming out (so at a larger scale), those small circles seem to assemble into one big copy of S^1. The tools we will develop are *scale independent* and do not depend on parameter tuning. We will ultimately produce concise, interpretable summaries that capture the nature of data at all scales.



Our first and primary tool will be the *persistent homology* of the *Čech* or *Vietoris–Rips filtered complex* associated with a point cloud $P \subseteq \mathbb{R}^N$. We may view this as a two-step process:

$$\{\text{point clouds}\} \overset{\check{\mathbb{C}}}{\underset{\mathbb{VR}}{\rightrightarrows}} \{\text{filtered spaces}\} \xrightarrow{\text{PH}} \{\text{persistence modules}\}.$$

A *filtered space* $\mathscr{X} = \{X_s\}_{s \in \mathbb{R}}$ is a collection of spaces[1] X_s indexed by scales $s \in \mathbb{R}$ such that

$$s \leq t \implies X_s \subseteq X_t.$$

For the purposes of this introduction, we will focus on the Čech filtered complex $\check{\mathbb{C}}(P)$ of our point cloud $P \subseteq \mathbb{R}^N$. At scale $s \in \mathbb{R}$, $\check{\mathbb{C}}_s(P)$ is

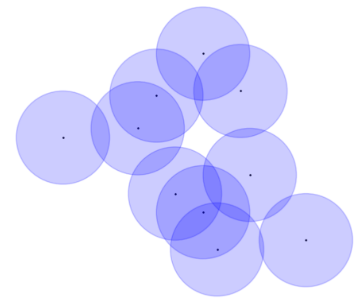
[1] By *space* we might mean topological space or (abstract) simplicial complex. If working with complexes, we take $\subseteq$ to mean subcomplex.

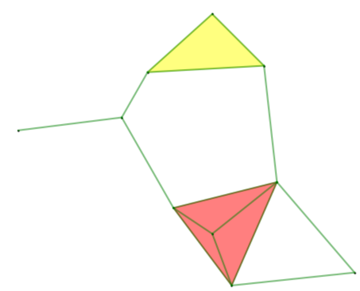
the simplicial complex with one n-simplex for each subset $A \subseteq P$ with $|A| = n+1$ and

$$\bigcap_{x \in A} \overline{B}_s(x) \neq \varnothing.$$

In other words, we get an n-simplex for each $(n+1)$-subset of P for which the closed Euclidean balls of radius s centered at points of A have nonempty common intersection. Since the intersection condition becomes less stringent as s gets larger, we have that $\check{C}_s(P)$ is a subcomplex of $\check{C}_t(P)$ when $s \leq t$. Later, we will encounter the Nerve Lemma, which roughly says that $\check{C}_s(P)$ is homotopy equivalent to $\bigcup_{x \in P} \overline{B}_s(P)$ in reasonable scenarios. Note that the combinatorial nature of $\check{C}_s(P)$ makes it much better adapted to computation than the filtered topological space $\{\bigcup_{x \in P} \overline{B}_s(P)\}_{s \in \mathbb{R}}$.

We write $\overline{B}_s(x)$ for the closed ball of radius s centered at x.





Black points are 0-simplices, green edges are 1-simplices, yellow shading is a 2-simplex, and red shading is a 3-simplex. Note that the bottom right triangle is not filled in yellow because the triple intersection of the balls around those vertices is empty.

Now that we have a filtered space $\mathscr{X} = \check{C}(P)$, we aim to capture features of each space $X_s := \check{C}_s(P)$ and how these features are related as the filtration parameter changes. Taking a cue from algebraic topology, we view $\mathrm{H}_*(X_s; \mathbb{F})$ — the homology[2] of X_s with coefficients in a field $\mathbb{F}$ — as a good summary of the features of X_s. Functoriality of homology then provides us with $\mathbb{F}$-linear transformations

$$(\iota_s^t)_* \colon \mathrm{H}_*(X_s; \mathbb{F}) \longrightarrow \mathrm{H}_*(X_t; \mathbb{F})$$

for $s \leq t$ and $\iota_s^t \colon X_s \subseteq X_t$, and these maps $(\iota_s^t)_*$ provide our comparisons of features. Packaging all of the homologies and comparisons maps together produces a *persistence module* $\mathrm{PH}_*(\mathscr{X}; \mathbb{F})$, the $\mathbb{F}$-*persistent homology* of $\mathscr{X}$, which is our scale independent summary of the shape of our data.

The miracle here is that persistence modules admit a convenient and complete invariant called a *barcode* or (after a mild but tremendously beneficial transformation) *persistence diagram*. To give the flavor of barcodes, we will consider a simplified scenario in which we have $\mathbb{F}$-vector spaces $\{V_i\}_{i \in \mathbb{N}}$ and linear transformations $\iota_i^j \colon V_i \to V_j$ for $0 \leq i \leq j$ such that

(1) $\iota_i^i = \mathrm{id}_{V_i}$ for all i, and

(2) for $0 \leq i \leq j \leq k$, $\iota_j^k \circ \iota_i^j = \iota_i^k$.

The essential data here is of the form

$$V_0 \to V_1 \to \cdots \to V_i \to V_{i+1} \to \cdots$$

and we may view the persistence module $(\{V_i\}_{i \in \mathbb{N}}, \{\iota_i^j\}_{i \leq j})$ as a functor $\mathscr{V} = \{V_i\}_{i \in \mathbb{N}}$ from the category associated with the partially ordered set $(\mathbb{N}, \leq)$ to the category of $\mathbb{F}$-vector spaces and linear transformations. Such a persistence module might arise from a point cloud by considering Čech complexes at scales $s_0 < s_1 < \cdots$.

Let $\mathbb{F}[t]$ denote the ring of polynomials in variable t over $\mathbb{F}$, graded so that $|t| = 1$, and set

$$\Theta(\mathscr{V}) := \bigoplus_{i \in \mathbb{N}} V_i.$$

[2] We will review homology theory next lecture. It is a lie in the direction of truth to say that the dimension of the $\mathbb{F}$-vector space $\mathrm{H}_n(X_s; \mathbb{F})$ measures the number of n-dimensional “holes” in X_s.

Then we may endow $\Theta(\mathscr{V})$ with the structure of a graded $\mathbb{F}[t]$-module by setting the action of the polynomial generator t to be

$$t \cdot (v_i)_{i\in\mathbb{N}} := (\iota_{i-1}^{i} v_{i-1})_{i\in\mathbb{N}}$$

where $v_{-1} := 0$. In fact, Θ is an equivalence of categories between $\mathbb{N}$-persistence modules and graded $\mathbb{F}[t]$-modules.[3]

[3] The inverse functor takes M_* to $\{M_i\}_{i\in\mathbb{N}}$ with ι_i^j given by multiplication by t^{j-i}.

A common capstone theorem of a first course in algebra is the classification of finitely generated modules over a principal ideal domain. A graded version of this theorem holds *mutatis mutandis*, and so it behooves us to understand which persistence modules correspond to finitely generated graded $\mathbb{F}[t]$-modules. Call a persistence module $\mathscr{V} = \{V_i\}_{i\in\mathbb{N}}$ *tame* when every V_i is finite-dimensional and ι_i^{i+1} is an isomorphism for sufficiently large i. One may prove that $\mathscr{V}$ is tame if and only if $\Theta(\mathscr{V})$ is finitely generated over $\mathbb{F}[t]$.

By the classification theorem for finitely generated graded modules over a PID, if $\mathscr{V}$ is tame then there are (essentially unique) integers $i_1, \ldots, i_m, j_1, \ldots, j_n, \ell_1, \ldots, \ell_n$ and an isomorphism

$$\Theta(\mathscr{V}) \cong \bigoplus_{s=1}^{m} \Sigma^{i_s}\mathbb{F}[t] \oplus \bigoplus_{t=1}^{n} \Sigma^{j_t}\mathbb{F}[t]/(t^{\ell_t})$$

where Σ^r denotes a grading shift upwards by r.[4] Translating this into the world of persistence modules, we learn that every tame persistence module decomposes (essentially uniquely) as

[4] That is, $(\Sigma^r M_*)_s = M_{s-r}$

$$\mathscr{V} \cong \bigoplus_{j=0}^{N} \mathbb{I}[b_j, d_j]$$

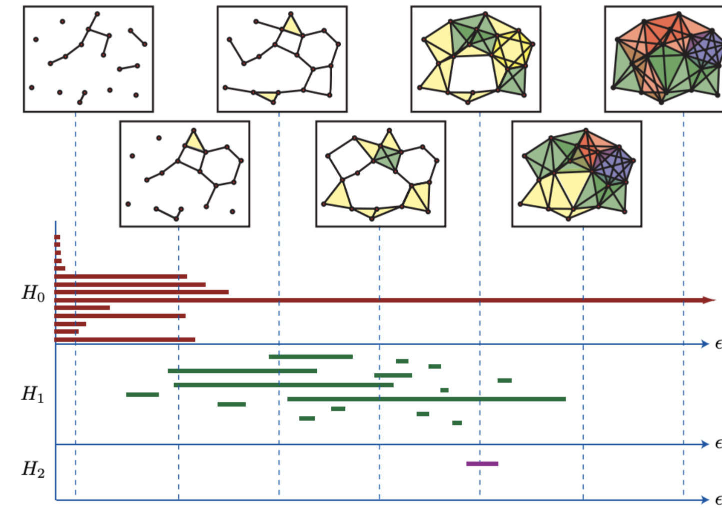
where each b_j is a nonnegative integer, $d_j \in \mathbb{N} \cup \{\infty\}$, and $\mathbb{I}[b_j, d_j]$ is the *interval persistence module* with

$$\mathbb{I}[b_j, d_j]_i = \begin{cases} \mathbb{F} & \text{if } b_j \le i \le d_j, \\ 0 & \text{otherwise,} \end{cases}$$

and $\iota_i^{i'} = \mathrm{id}_{\mathbb{F}}$ for $b_j \le i \le i' \le d_j$.

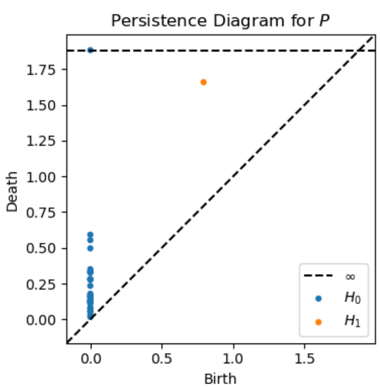
For an interval persistence module $\mathbb{I}[b, d]$, we refer to b as the *birth* and d as the *death* scale. We may then visualize the decomposition of $\mathscr{V}$ as a multiset of intervals $[b_j, d_j]$ called the *barcode* of $\mathscr{V}$. The following illustration is taken from Ghrist.[5] Beware, though, that it uses the Vietoris–Rips filtration instead of the Čech filtration; we will study $\mathbb{VR}$ in detail later.

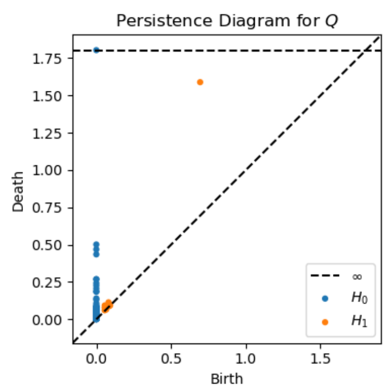
[5] Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc. (N.S.)*, 45(1):61–75



While barcodes prevailed in the early days of TDA, experience has shown that *persistence diagrams* are better suited to statistical analysis. The persistence diagram of $\mathscr{V}$ consists of the multiset of points (b_j, d_j) lying on or above the diagonal of $\mathbb{N} \times (\mathbb{N} \cup \{\infty\})$.

The Vietoris–Rips filtered complexes of the data sets P and Q from our initial discussion have the following persistence diagrams (with PH_0 in blue and PH_1 in orange):





Focusing on the blue PH_0 classes, we see that in both cases all connected components are born at time 0, and at scales above ≈ 0.7 there is a single connected component that persists to $+\infty$. This last class is analogous to the red bar of infinite length in the previous diagram.

Looking at orange PH_1 classes, we can readily observe significant differences between the point clouds. In each, there is a highly persistent class born around scale 0.75, but Q detects the small scale structure as well, giving a cluster of short-lived PH_1 classes born around scale 0.1. These classes witness the small radii circles (arranged around the unit circle) from which Q is sampled.

It is often claimed that classes with large persistence $d - b$ (*i.e.*, those high above the diagonal) represent the “true” topology of the data, while small persistence classes correspond to noise. The point clouds P and Q illustrate that this is not necessarily the case.

1.1 Future topics

One of our primary tasks will be the development of pseudometrics allowing us to compare persistence diagrams. We leave this to future development, along with the many foundational details elided or overlooked in this introduction. Once the foundations are established, the rest of the course will focus on the following:

(1) applications of persistent homology to particular data modalities,

(2) extending persistent homology to filtrations indexed by more exotic partially ordered sets, and

(3) refining PH_0 via hierarchical clustering.

See the syllabus for a detailed (but flexible) schedule of topics.

1.2 Notes

The content of this introduction was primarily drawn from the Oudot's textbook[6] and Carlsson's survey article.[7] The original images were produced in Python using the Ripser persistent homology package.[8] We will use Ripser extensively when exploring examples and applications, and you should follow the installation instructions at https://ripser.scikit-tda.org/ to get it working on your personal computer. You can find the Jupyter notebook used to produce diagrams from this and future lectures at https://github.com/kyleormsby/math583.

[6] Oudot, S. Y. (2015). *Persistence theory: from quiver representations to data analysis*, volume 209 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI

[7] Carlsson, G. (2009). Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308

[8] Tralie, C., Saul, N., and Bar-On, R. (2018). Ripser.py: A lean persistent homology library for python. *The Journal of Open Source Software*, 3(29):925

1.3 Exercises

(1) Install the necessary software and run the demos from today's class on your personal computer.

(2) Determine the smallest[9] point cloud in $\mathbb{R}^3$ whose Čech filtered complex exhibits nonzero PH_2 as some scale. What about in $\mathbb{R}^2$?

[9] Smallest in terms of cadinality — the least number of points.

2 *Spaces, complexes, and homology* — 27 March 2024

2.1 *Topology*

Felix Klein (1849–1925)

From the Kleinian perspective, geometry is the study of properties invariant under isometries, that is, distance-preserving transformations. Indeed, when a geometer says that two triangles are the same (or *congruent* or *isometric*), they do not mean that each triangle consists of exactly the same points, but rather that one may translate, rotate, and reflect one triangle until it matches the other.

Topology plays a similar game, but with a much coarser notion of "sameness". We say that two spaces — the objects of topology — are *homeomorphic* when there are continuous functions between them that are mutually inverse. In this sense, topology is the study of properties that are invariant under homeomorphism. Such properties include such notions as connectivity and compactness, but exclude more rigid properties such as angle, distance, or volume.

We will generally assume that the reader is familiar with point-set topology, but will quickly recall some of the basic definitions.

The standard reference for point-set topology is Munkres; see also the recent graduate text of Bradley–Bryson–Terilla.

Munkres, J. R. (2000). *Topology*. Prentice Hall, Inc., Upper Saddle River, NJ, second edition; and Bradley, T.-D., Bryson, T., and Terilla, J. (2020). *Topology—a categorical approach*. MIT Press, Cambridge, MA

Definition 2.1. A *topological space* is a pair $(X, \mathscr{U})$ consisting of a set X and a collection of subsets $\mathscr{U} \subseteq 2^X$ called *open sets* such that

(1) $\varnothing$ and X are in $\mathscr{U}$,

(2) $\mathscr{U}$ is closed under arbitrary unions: $U_\alpha \in \mathscr{U}$ for $\alpha \in A$ implies $\bigcup_{\alpha \in A} U_\alpha \in \mathscr{U}$, and

(3) $\mathscr{U}$ is closed under finite intersections: $U_i \in \mathscr{U}$ for i in a finite set I implies $\bigcap_{i \in I} U_i \in \mathscr{U}$.

We will write X for $(X, \mathscr{U})$ when the topology $\mathscr{U}$ is clear from context. A subset $U \subseteq X$ is called *open* when it belongs to $\mathscr{U}$, and a subset $C \subseteq X$ is called *closed* when $X \smallsetminus C$ is open. These properties are not mutually exclusive, as exhibited by the *clopen* sets $\varnothing$ and X.

Example 2.2. In the *standard topology* on Euclidean space $\mathbb{R}^n$, a subset $U \subseteq \mathbb{R}^n$ is open if and only if it is a union of *open balls* $B_r(x) := \{y \in \mathbb{R}^n \mid |y - x| < r$. This is equivalent to saying that U is open if and only if for each $x \in U$ there exists $r > 0$ such that $B_r(x) \subseteq U$.

Example 2.3. Suppose X is a topological space and Y is a subset of X. We may endow Y with the *subspace* topology (relative to X) by declaring that the open sets of Y are exactly those sets of the form $U \cap Y$ for $U \subseteq X$ open.

As a subexample of subspaces, consider the interval $[0,1] \subseteq \mathbb{R}$, where $\mathbb{R}$ carries the standard topology. Then $(1/2, 1] = (1/2, 3/2) \cap [0,1]$ is open in $[0,1]$, but not in $\mathbb{R}$.

Definition 2.4. A function $f\colon X \to Y$ between topologyical spaces is *continuous* when the preimage $f^{-1}U$ over every open set $U \subseteq Y$ is open in X. A continuous function $f\colon X \to Y$ is a *homeomorphism* when it admits a continuous inverse $g\colon Y \to X$. In this case, we say that X and Y are *homeomorphic* and write $X \cong Y$.

The categorically inclined reader will note that topological spaces and continuous functions form a category, and the isomorphisms in this category are exactly the homeomorphisms.

Example 2.5. If X is a topological space and $f\colon X \to \mathbb{R}$ is continuous, then the *sublevel set* $f^{-1}(-\infty, u) = \{x \in X \mid f(x) < u\}$ is open in X since the interval $(-\infty, u) = \{t \in \mathbb{R} \mid t < u\}$ is open in $\mathbb{R}$. Similarly, $f^{-1}(-\infty, u]$ is closed.[10]

[10] *Exercise*: For $f\colon X \to Y$ any continus map and $C \subseteq Y$ closed, check that $f^{-1}C$ is closed in X.

It will frequently be important to study a yet weaker notion of "sameness" in topology call *homotopy*. This is a two-step definition that first identifies when two continuous functions are homotopic, and then proceeds to spaces.

Definition 2.6. Continuous functions $f, g\colon X \to Y$ are *homotopic* when there exists a continuous function $H\colon X \times [0,1] \to Y$ such that the restriction of H to $X \times 0$ agrees with f and the restion to $X \times 1$ agrees with g. Such a map H is called a *homotopy* between f and g and we write $H\colon f \simeq g$.

We may think of H is a "movie" continuously interpolating between f and g.

Recall that spaces X and Y are homeomorphic when there are continuous functions $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. The following definition formalizes the notion a map admitting an inverse "up to homotopy".

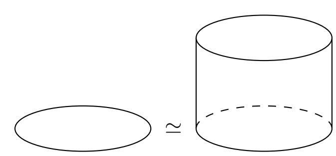
Definition 2.7. Two spaces X and Y are *homotopic* when there exist continuous functions $f\colon X \to Y$ and $g\colon Y \to X$ such that $g \circ f \simeq \mathrm{id}_X$ and $f \circ g \simeq \mathrm{id}_Y$.

Example 2.8. We check that the circle S^1 is homotopic to the cylinder $S^1 \times [0,1]$. Take

$$\begin{aligned} f\colon S^1 &\longrightarrow S^1 \times [0,1] \\ z &\longmapsto (z, 0) \end{aligned}$$

and

$$\begin{aligned} g\colon S^1 \times [0,1] &\longrightarrow S^1 \\ (z,t) &\longmapsto z. \end{aligned}$$



Then $g \circ f = \mathrm{id}_{S^1}$ (which is clearly homotopic to id_{S^1}) and $f \circ g\colon (z,t) \mapsto (z, 0)$. We define

$$\begin{aligned} H\colon (S^1 \times [0,1]) \times [0,1] &\longrightarrow S^1 \times [0,1] \\ ((z,t), s) &\longmapsto (z, ts). \end{aligned}$$

We have $((z,t),0) \mapsto (z,0)$, which is $f \circ g$, while $((z,t),1) \mapsto (z,t)$, which is $\mathrm{id}_{S^1 \times [0,1]}$. Thus $H\colon f \circ g \simeq \mathrm{id}_{S^1 \times [0,1]}$, as needed to verify that $S^1 \simeq S^1 \times [0,1]$.

In fact, there was nothing special about S^1 in this argument. Any space X is homotopic to $X \times [0,1]$, the cylinder on X.

2.2 Geometric and abstract simplicial complexes

Let $P = \{x_0, \ldots, x_k\} \subseteq \mathbb{R}^N$ be a point cloud in $\mathbb{R}^N$. An *affine combination* of P is a sum of the form

$$\sum_{i=0}^{k} \lambda_i x_k$$

where $\lambda_i \in \mathbb{R}$ and $\sum \lambda_i = 1$. The collection of all affine combinations of P is called the *affine hull* of P; it is always an affine linear subspace of $\mathbb{R}^N$.

The point cloud P is *affinely independent* if no $x \in P$ is an affine combination of $P \setminus \{P\}$. This is equivalent to the set $\{x_1 - x_0, \ldots, x_k - x_0\}$ being a linearly independent set of vectors.

Recall that a subset M of $\mathbb{R}^N$ is *convex* when the line segment joining any two points of M is a subset of M; the *convex hull* of M is the intersection $\mathrm{Conv}(M)$ of all convex sets containing M. When $P = \{x_0, \ldots, x_k\} \subseteq \mathbb{R}^N$ is affinely independent, we get *barycentric coordinates* on $\mathrm{Conv}(P)$. The barycentric coordinates of a point $x \in \mathrm{Conv}(P)$ are the unique $\lambda_i \in [0, 1]$ such that

$$x = \sum_{i=0}^{k} \lambda_i x_i \qquad \text{and} \qquad \sum_{i=0}^{k} \lambda_i = 1.$$

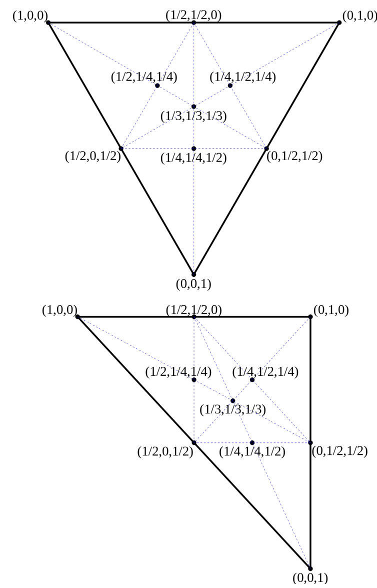


Image by user Rubybrian, CC BY-SA 3.0.

We can now define geometric simplicial complexes, whose basic building blocks are geometric simplices:

Definition 2.9. Suppose $k, N \in \mathbb{N}$ with $k \leq N$. A *geometric k-simplex* σ is the convex hull of an affinely independent point cloud $P = \{x_0, \ldots, x_k\} \subseteq \mathbb{R}^N$ with $k + 1$ elements, *i.e.*,

$$\sigma = \mathrm{Conv}(P).$$

The *dimension* of σ is k, and we will sometimes write $\sigma = \sigma^k$ to express its dimension. The points $x_0, \ldots, x_k$ are the *vertices* of σ, its *edges* are the convex hulls of pairs of vertices of σ, and, more generally, the convex hull τ of any subset of P is called a *face* of σ. A face τ of σ is a *facet* when $\dim \tau = \dim \sigma - 1$.

Definition 2.10. Let $N \in \mathbb{N}$. A (finite) *geometric simplicial complex* $K \subseteq \mathbb{R}^N$ is a (finite) collection of geometric simplices K that is closed under taking faces ($\sigma \in K$ and τ a face of K implies $\tau \in K$) and compatible with intersection (if $\sigma, \tau \in K$, then $\sigma \cap \tau$ is either empty or a common face of both σ and τ).

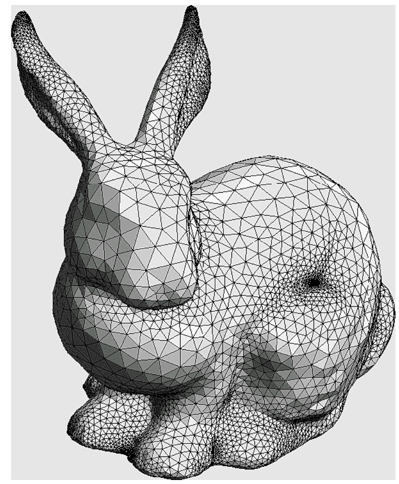
The *dimension* of a geometric simplicial complex is the maximal dimension of its simplices. The *body* of K is

$$|K| = \bigcup_{\sigma \in K} \sigma.$$

In a standard act of laziness, we will often blur the distinction between K and $|K|$.

Definition 2.11. A *triangulation* of a subspace $X \subseteq \mathbb{R}^N$ is a geometric simplicial complex K in $\mathbb{R}^N$ such that $|K| \cong X$.

We warn the reader that not every subspace of $\mathbb{R}^N$ admits a triangulation. But "reasonable" spaces do, and there is a tremendous amount of interesting topology one can do with geometric simplicial complexes. Triangulations are also essential in computer graphics, as illustrated by the Stanford bunny pictured here.



While geometric simplicial complexes nicely match our intuition for how spaces might be chopped up into simplicial pieces, they are very inefficient as data structures. By working with abstract simplicial complexes, we can recover the homeomorphism type of (the body of) a geometric simplicial complex far more efficiently.

Definition 2.12. Let P be a finite set. An *abstract simplicial complex* L on P is a family of nonempty subsets of P that is closed under taking nonempty subsets: $\sigma \in L$ and $\varnothing \neq \tau \subseteq \sigma$ implies $\tau \in L$.

Example 2.13. Every geometric simplicial complex K on vertex set P determines an abstract simplicial complex L on P by declaring that $\sigma \in L$ if and only if $\mathrm{Conv}(\sigma)$ is a face of K.

We may also create a geometric simplicial complex from any abstract simplicial complex, a process called *geometric realization*. The defining feature of a geometric realization K of an abstract simplicial complex L is that the abstract simplicial complex L' associated with K (as in the above example) is L again up to relabeling of vertices. When P is finite, constructing geometric realizaitons is fairly straightforward:

Proposition 2.14. *Every abstract simplicial complex K with k vertices admits a geometric realization in $\mathbb{R}^{k-1}$.*

Proof. The complex K is a subcomplex of the full $(n-1)$-simplex on P, which may be geometrically realized as the convex hull of $0, e_1, \ldots, e_{n-1} \in \mathbb{R}^{n-1}$. □

It is actually the case that every abstract simplicial complex on k vertices admits a geometric realization in $\mathbb{R}^{2k+1}$, but we will not go into the proof. One can also geometrically realize infinite abstract simplicial complexes via a colimit construction.

In order to compare simplicial complexes, we need an appropriate class of maps.

Definition 2.15. A *simplicial map* between abstract simplicial complexes K and L is a function $f\colon K^{(0} \to L^{(0)}$ on vertices such that the image of every abstract simplex in K is an abstract simplex in L.

Definition 2.16. A *simplicial map* between geometric simplicial complexes K and L is a function $f\colon |K| \to |L|$ such that the restriction of f to $K^{(0)}$

induces a simplicial map between associated abstract simplicial complexes and which is linear on geometric simplices (in terms of barycentric coordinates), *i.e.*, if $t_0, \ldots, t_k \in [0,1]$ with $\sum t_i = 1$ and $v_0, \ldots, v_k \in K^{(0)}$, then

$$f\left(\sum t_i v_i\right) = \sum t_i f(v_i). \tag{2.17}$$

Tracing through the definitions, one may check that simplicial maps between geometric simplicial complexes induce abstract simplicial maps, and conversely, each simplicial map between abstract simplicial complexes extends in a unique way to a simplicial map between geometric realizations via (2.17). It is also the case that simplicial maps induce continuous functions between bodies. Perhaps surprising, though, is that continuous maps between (bodies of) geometric simplicial complexes can be approximated by simplicial maps.

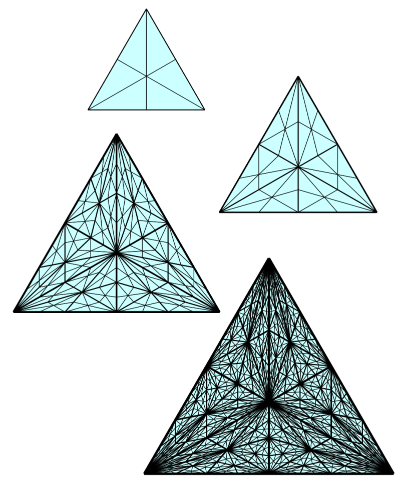


Image by user Studentofrationality showing successive barycentric subdivisions of an equilateral triangle, CC BY-SA 4.0.

Theorem 2.18 (Simplicial approximation)**.** *Suppose $f\colon K \to L$ is a continuous function between geometric simplicial complexes. Then there exist sufficiently fine subdivisions K' of K and L' of L, and a simplicial map $f'\colon K' \to L'$ such that $f \simeq f'$.*

Here a subdivision of K is a geometric simplicial complex K' such that every face of K is a union of simplices of K'. The proof of the simplicial approximation theorem is covered in many standard combinatorial or algebraic topology texts and we won't attempt it here.

We need one more crucial definition before we move on to simplicial homology, namely that of *orientation*. This amounts to a choice of ordering on vertices, taken up to a certain equivalence relation.

Definition 2.19. An *oriented simplex* on vertices $x_0, \ldots, x_k$ is an ordered $(k+1)$-tuple $\sigma = \langle x_0, x_1, \ldots, x_k \rangle$ subject to the rule

$$\sigma = \operatorname{sgn}(\pi)\langle x_{\pi(0)}, x_{\pi(1)}, \ldots, x_{\pi(k)} \rangle$$

where σ is a permutation of $\{0, 1, \ldots, k\}$ and $\operatorname{sgn}(\pi)$ is the *signature* of π:

$$\operatorname{sgn}(\pi) = (-1)^m$$

where m is the number of transpositions in a decomposition of σ as a composite of transpositions.[11]

We also give each 0-dimensional simplex two orientations, $\langle x \rangle$ and $-\langle x \rangle$.

Finally, two k-simplices sharing a $(k-1)$-dimensional face σ are *consistently oriented* when they induce opposite orientations on σ.

[11] Such decompositions always exist; the number m is not unique, but all such m have the same parity.

2.3 Simplicial homology

We are now going to "measure" the "holes" in a simplicial complex with a tool called homology. The slogan for homology is "cycles mod boundaries".

Both cycles and boundaries are special types of "chains", which are formal linear combinations of oriented simplices. The cycles are those chains with trivial boundary (so they properly "enclose" part of the complex), and the boundaries are boundaries of chains. Crucially, the boundary of a boundary is trivial, so every boundary chain is a cycle. The word "mod" means we take a quotient, an algebraic operation which identifies cycles when they differ by a common boundary. In particular, every boundary becomes 0. The idea here is that when the region enclosed by a cycle can be filled in (*i.e.*, the cycle is a boundary), it is *not* a hole. Meanwhile, cycles that don't bound a lower dimensional chain are essential and do get counted as holes.

We now formalize these ideas. Let $\mathbb{F}$ be a field,[12] let K be an abstract simplicial complex of dimension n, and let k be a natural number between 0 and n.

[12] That is, a number system in which you can add, multiply, and divide (by nonzero numbers). Examples include the rational, real, and complex numbers. We will extensively use the finite fields $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ which encode "clock arithmetc" on a clock with hours $0, 1, \ldots, p-1$ for p prime. Our favorite case will be $\mathbb{F}_2 = \{0, 1\}$ in which $2 = 0$ and $1 = -1$. Make sure you know what the addition and scalar multiplication rules for $C_k(K; \mathbb{F})$ are.

Definition 2.20. A *k-chain* is a formal $\mathbb{F}$-linear combination $\sum \lambda_i \sigma_i^k$ of oriented k-dimensional simplices in K subject to the rule $(-1) \cdot \sigma = -\sigma$, where the latter term indicates σ with the opposite orientation. The *chain group* $C_k(K; \mathbb{F})$ is the $\mathbb{F}$-vector space of all k-chains.

Note that if K has n_k many k-simplices, then $C_k(K; \mathbb{F}) \cong \mathbb{F}^{n_k}$ with basis given by the k-simplices.

We next formalize the notion of boundary. The rough idea is that the boundary of an k-simplex consists of all its $(k-1)$-dimensional facets added together (as a $(k-1)$-chain) with consistent orientations. We then extend this assignment linearly to all of $C_k(K; \mathbb{F})$. Given an oriented simplex $\sigma = \langle x_0, x_1, \ldots, x_k \rangle$ and $0 \leq i \leq k$, let

$$\hat{\sigma}_i := \langle x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k \rangle$$

be the $(k-1)$-dimensional facet of σ with x_i removed with the induced orientation.

Definition 2.21. Let $k \in \mathbb{N}$. The *boundary map*

$$\partial = \partial_k \colon C_k(K; \mathbb{F}) \longrightarrow C_{k-1}(K; \mathbb{F})$$

is the unique $\mathbb{F}$-linear transformation such that

$$\partial_k \sigma = \sum_{i=0}^{k} (-1)^i \hat{\sigma}_i.$$

If $k = 0$, we set $\partial_0 \colon C_0(K; \mathbb{F}) \to 0$ to be the trivial map.

If our combinatorial algebra correctly captures geometric intuition, then the boundary of a boundary should be trivial. This brings us to what Dennis Sullivan[13] calls the most important equation in mathematics: $\partial^2 = 0$.

[13] Sullivan won the 2022 Abel Prize for his contributions to algebraic topology, geometric topology, and dynamics.

Theorem 2.22. *For* $k \geq 1$,

$$\partial^2 := \partial_{k-1} \circ \partial_k = 0.$$

Proof. It suffices to show that $\partial^2\sigma = 0$ for $\sigma = \langle x_0, \ldots, x_k\rangle$ an oriented k-simplex. For $0 \le i < j \le k$, let $\hat{\sigma}_{ij}$ be the $(k-2)$-dimensional simplex with x_i and x_j removed. This term appears twice in the expansion of $\partial^2\sigma$, once with sign $(-1)^i(-1)^j$ and once with sign $(-1)^i(-1)^{j-1}$. These terms cancel so the total sum for $\partial^2\sigma$ is equal to 0. □

It follows that the image of ∂_k is contained in the kernel of ∂_{k-1}. This make the chain groups and ∂ into a *chain complex*:

$$\cdots \xrightarrow{\partial} C_n(K;\mathbb{F}) \xrightarrow{\partial} C_{n-1}(K;\mathbb{F}) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_1(K;\mathbb{F}) \xrightarrow{\partial} C_0(K;\mathbb{F}) \xrightarrow{\partial} 0.$$

[TODO: Add example with matrices representing boundary maps; observe that consecutive matrices multiply to 0.]

We are now ready to implement our slogan of "homology equals cycles mod boundaries".

Definition 2.23. Let K be an abstract simplicial complex, let $\mathbb{F}$ be a field, and let k be a natural number. Then

» the group of *k-cycles*, $Z_k(K;\mathbb{F})$, is the kernel of ∂_k, which is a subspace of $C_k(K;\mathbb{F})$,

» the group fo *k-boundaries*, $B_k(K;\mathbb{F})$, is the image of ∂_{k+1}, which is a subspace of $Z_k(K;\mathbb{F})$, and

» the *k-th homology* group of K is the quotient vector space[14]

$$\mathrm{H}_k(K;\mathbb{F}) := Z_k(K;\mathbb{F})/B_k(K;\mathbb{F}).$$

We call $\dim H_k(K;\mathbb{F})$ the q-th $\mathbb{F}$-Betti number of K, denoted $b_q(K;\mathbb{F})$.

One amazing feature of homology is that it is both a homemorphism and homotopy invariant. Note that this is wildly false for chains, cycles, and boundaries. Implicit in this claim is that homology is also *functorial*: given a simplicial map $f\colon K \to L$, the assignment

$$\begin{aligned} f_*\colon H_k(K;\mathbb{F}) &\longrightarrow H_k(L;\mathbb{F}) \\ [\sigma] &\longmapsto \begin{cases} [f(\sigma)] & \text{if } f(x_0), \ldots, f(x_k) \text{ are distinct,} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is a well-defined linear transformation. Moreover, $(\mathrm{id}_K)_* = \mathrm{id}_{H_k(K;\mathbb{F})}$ and if $g\colon L \to M$ is another simplicial map, then $(g \circ f)_* = g_* \circ f_*$. Homotopy invariance is now the following statement:

Theorem 2.24. *If $k \in \mathbb{N}$ and $f\colon K \simeq L$ is a simplicial homotopy equivalence, then $f_*\colon H_k(K;\mathbb{F}) \cong H_k(L;\mathbb{F})$.*

In particular, if K is *contractible*,[15] then $H_0(K;\mathbb{F}) \cong \mathbb{F}$ and $H_k(K;\mathbb{F}) = 0$

[14] If U is a vector supspace of V, then V/U is the vector space of U-*cosets* of the form $v + U = \{v + u \mid u \in U\}$ for $v \in V$. Note that if $v - w \in U$, then $v + U = w + U$. Addition is given by $(v + U) + (w + U) = (v + w) + U$ and scalar multiplication by $\lambda(v + U) = (\lambda v) + U$. If you're not familiar with quotient vector spaces, you should check that these operations are well-defined. Observe that $U = 0 + U$ is the trivial (or zero) element of V/U. This is the sense in which V/U "kills" the subspace U.
The quotient space V/U also enjoys a universal property. First note that there is a canonical quotient map $q\colon V \to V/U$ taking v to $v + U$. If $f\colon V \to W$ is a linear transformation such that $f(U) = 0$ – *i.e.*, $U \le \ker f$ – then there is a unique linear transformation $\tilde{f}\colon V/U \to W$ such that $f = \tilde{f} \circ q$. We say that linear transformation killing U *factor* uniquely through V/U (via q).

[15] *I.e.*, K is homotopy equivalent to a point, written $K \simeq *$.

for $k > 0$ simply by a quick computation of chains, cycles, and boundaries for a point.

We now list some of the many properties one would prove about simplicial homology in a full development of the subject:

» $b_0(K; \mathbb{F})$ is the number of connected components of K.

» Let S^n denote the unit sphere in $\mathbb{R}^{n+1}$. For $n \geq 1$,

$$H_k(S^n; \mathbb{F}) \cong \begin{cases} \mathbb{F} & \text{if } k = 0, n, \\ 0 & \text{otherwise.} \end{cases}$$

» Let $T^n := S^1 \times \cdots \times S^1$ be the n-fold Cartesian product of the circle with itself. Then

$$b_k(T^n; \mathbb{F}) = \binom{n}{k}$$

for all fields $\mathbb{F}$.

» Let $K \amalg L$ denote the disjoint union of simplicial complexes K and L. Then

$$H_k(K \amalg L; \mathbb{F}) \cong H_k(K; \mathbb{F}) \oplus H_k(L; \mathbb{F}).$$

Crucially, homology also has excellent properties with respect to decompositions into subcomplexes. The so-called *Mayer–Vietoris sequence* is a powerful method for computing the homology of $A \cup B$ in terms of the homology of A, B, and $A \cap B$ when A and B are subcomplexes of $A \cup B$. We can think of this tool as a "derived" version of the inclusion-exclusion theorem. Developing and even stating the theorem requires a fair bit of homological algebra, so we will point the reader to Section 8.2 of Virk's notes[16] for details.

[16] Virk, Ž. (2022). Introduction to persistent homology. https://zalozba.fri.uni-lj.si/virk2022.pdf. Accessed on 19 March 2024

2.4 Notes

The presentation of simplicial complexes and homology is a compressed version of Chapters 3 and 7 along with Section 4.2 of Virk's notes. I strongly recommend this text for those new to the subject!

2.5 Exercises

(1) Use chains, cycles, and boundaries to compute the homology of a circle, modeled as the simplicial set $\{a, b, c, ab, bc, ca\}$. (Here we are writing x for $\{x\}$ and xy for $\{x, y\}$.)

(2) Triangulate the Klein bottle K and prove that

$$H_k(K; \mathbb{F}_2) \cong \begin{cases} \mathbb{F}_2 & \text{if } k = 0, 2, \\ \mathbb{F}_2^2 & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

while

$$H_k(K;\mathbb{F}) \cong \begin{cases} \mathbb{F} & \text{if } k = 0, 1, \\ 0 & \text{otherise} \end{cases}$$

if $\mathbb{F}$ is a field in which $2 \neq 0$. This demonstrates that homology is sensitive to the arithmetic of the field of coefficients!

(3) Define the *Euler characteristic* of a finte simplicial complex K with n_k many k-simplices to be the alternating sum

$$\chi(K) = \sum_{k\geq 0} (-1)^k n_k.$$

Use a rank-nullity argument to prove that $\chi(K)$ is also equal to the alternating sum of Betti numbers $\sum_{k\geq 0} b_k(K;\mathbb{F})$. Conclude (a) that Euler characteristic is a homotopy invariant and (b) that the alternating sum of Bett numbers does not depend on $\mathbb{F}$. (Observe that this is consistent with the $\mathbb{F}$-Betti numbers of the Klein bottle, which has Euler characteristic 0.)

3 *Persistence modules* — 1 April 2024

Note: We didn't cover homology during our last meeting, so the 1 April lecture started with oriented simplices.

Recall fromt he first lecture the fundamental pipeline for persistent homology of point clouds:

$$\{\text{point clouds}\} \underset{\mathbb{VR}}{\overset{\check{\mathrm{C}}}{\rightrightarrows}} \{\text{filtered spaces}\} \xrightarrow{\mathrm{PH}} \{\text{persistence modules}\}.$$

Given a filtered abstract simplicial complex $\mathscr{X} = \{X_s\}_{s\in\mathbb{R}}$ (with $\iota_s^t \colon X_s \subseteq X_t$ for $s \leq t$), we may now apply the k-th homology functor $H_k(-;\mathbb{F})$ to get a persistence module

$$\mathrm{PH}_k(\mathscr{X};\mathbb{F}) := \{H_k(X_s;\mathbb{F}), (\iota_s^t)_* \mid s \leq t \in \mathbb{R}\}.$$

Our current goal is to understand the structure of such persistence modules.

Definition 3.1. An $\mathbb{R}$-persistence module $\mathscr{V}$ over a field $\mathbb{F}$ is a collection of $\mathbb{F}$-vector spaces V_s, $s \in \mathbb{R}$, along with linear transition maps $\iota_s^t \colon V_s \to V_t$ for each pairs $s \leq t$ such that

(1) for $s \in \mathbb{R}$, $\iota_s^s = \mathrm{id}_{V_s}$, and

(2) for $r \leq s \leq t$, $\iota_s^t \circ \iota_r^s = \iota_r^t$.

Here's a categorical take on persistence modules: For any poset $(P, \leq)$, also write P for the associated category. A P-persistence module is a functor over $\mathbb{F}$ is a functor $P \to \mathrm{Vect}_{\mathbb{F}}$. Maps between P-persistence modules are natural transformations.

A *map* of persistence modules $f \colon \mathscr{V} \to \mathscr{W}$ is a collection of linear transformations $f_s \colon V_s \to W_s$ such that the diagram

$$\begin{array}{ccc} V_s & \xrightarrow{f_s} & W_s \\ {\scriptstyle \iota_s^t}\downarrow & & \downarrow{\scriptstyle \iota_s^t} \\ V_t & \xrightarrow[f_t]{} & W_t \end{array}$$

commutes for all $s \leq t$. (Here we are abusing notations and writing ι_s^t for the transition maps for both $\mathscr{V}$ and $\mathscr{W}$.) Two persistence modules are *isomorphic*, written $\mathscr{V} \cong \mathscr{W}$, when there is a map of persistence modules $f \colon \mathscr{V} \to \mathscr{W}$ which admits a two-sided inverse or, equivalently, has f_s a bijection for each $s \in \mathbb{R}$.

We now follow the presentation in Section 4.5 of Carlsson–Vejdemo-Johansson[17] in order to classify $\mathbb{R}$-persistence modules.[18]

[17] Carlsson, G. and Vejdemo-Johansson, M. (2022). *Topological data analysis with applications*. Cambridge University Press, Cambridge

[18] Later, when we study multipersistence, we will see that for most posets P, P-persistence modules do not admit a classification (in the sense of being "wild type" problems in the language of representation theory).

We begin with the free $\mathbb{R}$-persistence modules. Let X be a set and consider a function $\rho \colon X \to \mathbb{R}$. We may view (X, ρ) as an $\mathbb{R}$-filtered set via the sublevel filtration with $X_s := \rho^{-1}(-\infty, s] = \{x \in X \mid \rho(x) \leq s\}$. We write $\mathscr{V}(X, \rho)$ for the $\mathbb{R}$-persistence module with

$$\mathscr{V}(X,\rho)_s := \mathbb{F} \cdot X_s.$$

Here $\mathbb{F} \cdot X_s$ is the $\mathbb{F}$-vector space with basis the set X_s; its elements are formal linear combinations of elements of X_s. The linear transformations $\iota_s^t \colon \mathscr{V}(X,\rho)_s \to \mathscr{V}(X,\rho)_t$ are the subspace inclusions induced by $X_s \subseteq X_t$.

Definition 3.2. We say that an $\mathbb{R}$-persistence module is *free* when it is isomorphic to an $\mathbb{R}$-persistence module of the form $\mathscr{V}(X,\rho)$; it is additionally *finitely generated* when X is a finite set.

Importantly, we can take quotients of $\mathbb{R}$-persistence modules. If $\mathscr{U} \le \mathscr{V}$ is a sub-$\mathbb{R}$-persistence module (meaning $U_s \le V_s$ for all s with compatible transition maps), then $\mathscr{V}/\mathscr{U}$ is the $\mathbb{R}$-persistence module with s-th vector space V_s/U_s and

$$\iota_s^t(v_s + U_s) = \iota_s^t(v_s) + U_t.$$

Given a map of a persistence modules $f \colon \mathscr{V} \to \mathscr{W}$, we define the *cokernel* of f to be

$$\operatorname{coker} f := \mathscr{W} / \operatorname{im} f$$

where $\operatorname{im} f$ is the sub-persistence module of $\mathscr{W}$ with $(\operatorname{im} f)_s = \operatorname{im} f_s$.

Definition 3.3. We say that an $\mathbb{R}$-persistence module is *finitely presented* when it is isomprhic to an $\mathbb{R}$-persistence module of the form $\operatorname{coker} f$ where f is a map between finitely generated free $\mathbb{R}$-persistence modules.

Recall from linear algebra that choosing bases for vector spaces V, W allows us to write down a matrix A_f[19] for every linear transformation $f \colon V \to W$ representing f. If the bases for V, W are $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$, respectively, and $f(v_j) = \sum_i \lambda_i w_i$, then the v_j-column of A_f is $(\lambda_1, \ldots, \lambda_m)$. If U is another vector space with basis $u_1, \ldots, u_\ell$ and $g \colon U \to V$ is another linear transformation, then $A_{g \circ f} = A_g \cdot A_f$. This is the purpose of matrix multiplication.

[19] Importantly, A_f depends on the choice of bases, but we will not include the bases in our notation. Also note that we are being coy about the role that ordering of bases plays.

For a pair of finite sets X, Y, define an (X,Y)-matrix to be an array $[a_{xy}]_{(x,y) \in X \times Y}$ of elements of $\mathbb{F}$. Write $r(x)$ for the row associated with x, and $c(y)$ for the column associated with y. Given a finitely generated free persistence module $\mathscr{V}(X,\rho)$, note that

$$\mathscr{V}(X,\rho)_s = \mathbb{F} \cdot X_s = \mathbb{F} \cdot X \text{ for } s \gg 0$$

since X is finite. Thus any map $f \colon \mathscr{V}(Y,\sigma) \to \mathscr{V}(X,\rho)$ of finitely generated free persistence modules induces a linear transformation $f_\infty \colon \mathbb{F} \cdot Y \to \mathbb{F} \cdot X$ with an associated (X,Y)-matrix $[a_{xy}] := A_{f_\infty}$ where we work with the bases Y, X. This gives us an encoding of maps between finitely generated free $\mathbb{R}$-persistence modules:

Theorem 3.4. *Given f as above, the (X,Y)-matrix $A_{f_\infty} = [a_{xy}]$ satisfies $a_{xy} = 0$ whenever $\rho(x) > \sigma(y)$. Furthermore, any (X,Y)-matrix $A = [a_{xy}]$ satisfying this condition induces a map of $\mathbb{R}$-persistence modules $f_A \colon \mathscr{V}(Y,\sigma) \to \mathscr{V}(X,\rho)$ and the correspondences $f \mapsto A_{f_\infty}$ and $A \mapsto f_A$ are mutually inverse.*

Proof. First suppose $\rho(x) > \sigma(y)$. We have $f_\infty(y) = \sum_{x \in X} a_{xy} x$. Such a linear combination lies in $\mathscr{V}(X, \rho)_s$ if and only if $a_{xy} = 0$ for $\rho(x) > s$. Specializing to $s = \sigma(y)$ gives the first claim. We leave the second and third statements to the reader. □

Definition 3.5. Given $\mathbb{R}$-filtered finite sets (X, ρ) and (Y, σ), call an (X, Y)-matrix satisfying the condition of the theorem (ρ, σ)-*adapted*. Given a (ρ, σ)-adapted (X, Y)-matrix A, define

$$\theta(A) := \operatorname{coker} f_A.$$

We see straightaway that $\theta(A)$ is a finitely presented $\mathbb{R}$-persistence module, and the theorem implies that every finitely presented $\mathbb{R}$-persistence module is isomorphic to one of the form $\theta(A)$. Thus to classify finitely presented $\mathbb{R}$-persistence modules, it suffices to understand for which (X, Y)-matrices A, A' we have $\theta(A) \cong \theta(A')$.

Lemma 3.6. *Let (X, ρ) be a finite $\mathbb{R}$-filtered set. Then the automorphisms*[20] *of $\mathscr{V}(X, \rho)$ can be identified with the invertible (ρ, ρ)-adapted (X, X)-matrices.*[21] □

[20] An *automorphism* is an isomorphism with the same source and target.

[21] Unpacking this, we have an invertible (X, X)-matrix $[a_{xx'}]_{(x,x') \in X \times X}$ satisfying $a_{xx'} = 0$ for $\rho(x) > \rho(x')$. If X is ordered by increasing ρ values, this is an invertible upper triangular matrix, so upper triangular with no 0's on the diagonal.

The following proposition now follows by abstract nonsense:

Proposition 3.7. *Let (X, ρ) and (Y, σ) be finite $\mathbb{R}$-filtered sets, and let A be a (ρ, σ)-adapted (X, Y)-matrix. Let B be an invertible (ρ, ρ)-adapted (X, X)-matrix and let C be an invertible (σ, σ)-adapted (Y, Y)-matrix. Then BAC is a (ρ, σ)-adapted (X, Y)-matrix and*

$$\theta(A) \cong \theta(BAC).$$

We now establish notation that will allows us to state our classification theorem. For $a \in \mathbb{R}$ and $b \in \mathbb{R} \cup \infty$ with $a < b$, define

$$\mathbb{I}[a, b)_s := \begin{cases} \mathbb{F} & \text{if } a \leq s < b \\ 0 & \text{otherwise} \end{cases}$$

and $\iota_s^t \colon \mathbb{I}[a, b)_s \to \mathbb{I}[a, b)_t$ to be $\mathrm{id}_{\mathbb{F}}$ if $a \leq s \leq t < b$ and 0 otherwise. This is the *interval* $\mathbb{R}$-*persistence module* associated with $[a, b)$.

Lemma 3.8. *The interval $\mathbb{R}$-persistence module $\mathbb{I}[a, b)$ is finitely presented.*

Proof. For $b < \infty$, let $(X, \rho) = (\{x\}, \rho(x) = a)$ and $(Y, \sigma) = (\{y\}, \sigma(y) = b)$. One may check that the 1×1 matrix $[1]$ is (ρ, σ)-adapted with $\mathbb{I}[a, b) = \theta([1])$. If $b = \infty$, then $\mathbb{I}[a, \infty) = \mathscr{V}(\{x\}, \rho(x) = a) = \theta(\varnothing)$ where $\varnothing$ is the 1×0 matrix representing the map $0 \to V(\{x\}, \rho(x) = a)$. □

This brings us to our classification theorem for finitely presented $\mathbb{R}$-persistence modules over $\mathbb{F}$:

Theorem 3.9. *Every finitely presented $\mathbb{R}$-persistence module over $\mathbb{F}$ is isomorphic to a finite direct sum of the form*

$$\mathbb{I}[a_1, b_1) \oplus \mathbb{I}[a_2, b_2) \oplus \cdots \oplus \mathbb{I}[a_n, b_b)$$

for some $a_i \in \mathbb{R}$, $b_i \in \mathbb{R} \cup \infty$, and $a_i < b_i$ for all i. Moreover,

$$\bigoplus_{i \in I} \mathbb{I}[a_i, b_i) \cong \bigoplus_{j \in J} \mathbb{I}[c_j, d_j)$$

for I, J finite sets if and only if $|I| = |J|$ and the multiset[22] *of intervals $[a_i, b_i)$ equals the multiset of intervals $[c_j, d_j)$.*

[22] A *multiset* is a set "with multiplicity". This can be formalized as a set X together with a function $m\colon X \to \mathbb{N}$ counting multiplicity.

Proof. First note that if A is the (ρ, σ)-adapted (X, Y) matrix with 1's in positions $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ and 0's elsewhere and there is at most one 1 in every row and column, then

$$\theta(A) \cong \bigoplus_{i=1}^{n} \mathbb{I}[\rho(x_i), \sigma(y_i)) \oplus \bigoplus_{x \in X \setminus \{x_1, \ldots, x_n\}} \mathbb{I}[\rho(x), \infty).$$

Thus given an arbitrary (ρ, σ)-adapted (X, Y)-matrix A, it suffices to construct an invertible (ρ, ρ)-adapted (X, X)-matrix B and invertible (σ, σ)-adapted (Y, Y)-matrix C such that every row and column of BAC has at most one 1 with all other entries 0.

We accompish this task via (ρ, σ)*-adapted row and column operations*. These are scalings of rows or columns by a nonzero element of $\mathbb{F}$, adding a multiple of $r(x)$ to $r(x')$ when $\rho(x) \geq \rho(x')$, and adding a multiple of $c(y)$ to $c(y')$ when $\sigma(y) > \sigma(y')$. The reader should check that these may be accomplished by multiplying A on the left or right by the appropriate kind of invertible matrix.

Now find y maximizing $\sigma(y)$ over all y with $c(y) \neq 0$. Then find x maximizing $\rho(x)$ over x such that $a_{xy} \neq 0$. We are free to add multiples of $r(x)$ to all other rows to cancel out $c(y)$ except for a_{xy}. We are then further free to add multiples of $c(y)$ to cancel out $r(x)$ except for a_{xy}. Multiplying $r(x)$ by $1/a_{xy}$ we get 1 in the xy-position. Now keep repeating the process with the next largest $\sigma(y)$ and $\rho(x)$ with $a_{xy} \neq 0$, and we eventually get the desired form.

The example following this proof implements the algorithm of this paragraph. The reader may wish to read the example in parallel with the proof.

We leave the uniqueness statement to the reader, who can also look at Proposition 4.52 of Carlsson–Vejdemo-Johansson. □

Example 3.10. Let us demonstrate the (ρ, σ)-adapted Gaussian elimination algorithm from the above proof. Let $X = \{x_1, \ldots, x_6\}$ with

$$\rho(x_1) = \rho(x_2) = 0, \quad \rho(x_3) = \rho(x_4) = \rho(x_5) = 1, \quad \rho(x_6) = 2$$

and $Y = \{y_1, \ldots, y_5\}$ with

$$\sigma(y_1) = 1, \quad \sigma(y_2) = \sigma(y_3) = 2, \quad \sigma(y_4) = \sigma(y_5) = 3.$$

Using the natural orderings of X and Y, we see that

$$A = \begin{bmatrix} 0 & 4 & 16 & 12 & 4 \\ 6 & 14 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 2 & 0 & 2 & 2 \\ 0 & 2 & 4 & 2 & 4 \\ 0 & 1 & 2 & 1 & 2 \end{bmatrix}$$

is (ρ, σ)-adapted. (In fact, the only condition is that the (x_6, y_1)-entry in the bottom left is 0.) We begin with the algorithm with the bottom right (x_6, y_5)-entry and use the bottom row to cancel out the rest of the rightmost column. This results in

$$\begin{bmatrix} 0 & 2 & 12 & 10 & 0 \\ 6 & 14 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 2 \end{bmatrix}.$$

We then clear out the bottom row with the rightmost column and multiply the bottom row by $1/2$ to get

$$\begin{bmatrix} 0 & 2 & 12 & 10 & 0 \\ 6 & 14 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We now work at the (x_3, y_4)-entry and clear out the rest of the y_4-column by adding $-5/2r(x_3)$ to $r(x_1)$ to get

$$\begin{bmatrix} 0 & 2 & 7 & 0 & 0 \\ 6 & 14 & 2 & 0 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now using $c(y_4)$ to zero out the rest of $r(x_3)$ and then multiplying $r(x_3)$ by

1/4 we get

$$\begin{bmatrix} 0 & 2 & 7 & 0 & 0 \\ 6 & 14 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Moving on to the (x_4, y_3)-entry the next round of row and column operations produces

$$\begin{bmatrix} 0 & 2 & 0 & 0 & 0 \\ 6 & 14 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now working in position (x_2, y_2) we finally arrive at

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This implies that the interval decomposition of $\theta(A)$ is

$$\begin{aligned} \theta(A) &\cong \mathbb{I}[\rho(x_1), \infty) \oplus \mathbb{I}[\rho(x_2), \sigma(y_2)) \oplus \mathbb{I}[\rho(x_3), \sigma(y_4)) \oplus \mathbb{I}[\rho(x_4), \sigma(y_3)) \oplus \mathbb{I}[\rho(x_5), \infty) \oplus \mathbb{I}[\rho(x_6), \sigma(y_5)) \\ &= \mathbb{I}[0, \infty) \oplus \mathbb{I}[0, 2) \oplus \mathbb{I}[1, 3) \oplus \mathbb{I}[1, 2) \oplus \mathbb{I}[1, \infty) \oplus \mathbb{I}[2, 3) \end{aligned}$$

As described in the first lecture, we will record the isomorphism type of an $\mathbb{R}$-persistence module with a *barcode* or *persistence diagram*. The barcode of

$$\mathbb{I}[a_1, b_1) \oplus \mathbb{I}[a_2, b_2) \oplus \cdots \oplus \mathbb{I}[a_n, b_b)$$

is the multiset of intervals $[a_i, b_i)$, typically drawn as stacked horizontal intervals. The persistence diagram of the same $\mathbb{R}$-persistence module is the multiset of points (a_i, b_i) in the extended plane $\mathbb{R} \times (\mathbb{R} \cup \infty)$. Since $a_i \leq b_i$, all these points lie on or above the line of slope 1 through the origin. The persistence diagram for the above example is displayed in the margin.

3.1 Notes

In the first lecture, we classified tame $\mathbb{N}$-persistence modules via the theory of finitely generated graded modules over a graded PID. Here we classified finitely presented $\mathbb{R}$-persistence modules essentially through a modified Gaussian elimination algorithm. Gaussian elimination requires $O(n^3)$ arithmetic operations, so the algorithm implicit in our proof is polynomial time but still computationally expensive. Zomorodian[23] provides methods for speeding up these algorithms when working with the persistent homology of particular types of simplicial complexes — called tidy sets — that include Vietoris–Rips complexes of point cloud data. This is one reason for going beyond Čech complexes when considering filtered spaces induced by point clouds, as we shall in the next lecture.

[23] Zomorodian, A. (2010). The tidy set: a minimal simplicial set for computing homology of clique complexes. In *Proceedings of the Twenty-Sixth Annual Symposium on Computational Geometry*, SoCG '10, page 257–266, New York, NY, USA. Association for Computing Machinery

3.2 Exercises

(1) Let $\mathbb{F} = \mathbb{Q}$, $X = \{0, 2, 4\}$, $Y = \{1, 2, 3, 4\}$, $\rho(x) = x$, and $\sigma(y) = y$. Check that

$$A = \begin{bmatrix} 2 & 1 & 2 & 0 \\ 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

is a (ρ, σ)-adapted (X, Y)-matrix (with the natural order of rows and columns) and then implement the algorithm from the proof of Theorem 3.9 to determine the persistence diagram associated with $\theta(A)$.

(2) Your answer to the previous question should have a 1 in the $(4, 4)$-position (so third row, fourth column), so one of the summands of the interval decomposition of $\theta(A)$ is $\mathbb{I}[4, 4)$. How do we interpret this $\mathbb{R}$-persistence module?

(3) Suppose A is a (ρ, σ)-adapted (X, Y)-matrix and let $x_{\max}$ and $y_{\max}$ be elements of X and Y maximizing ρ and σ, respectively. Assume that the $y_{\max}$ column of A is nonzero. Prove that the interval decomposition of $\theta(A)$ includes $\mathbb{I}[\rho(x_{\max}), \sigma(y_{\max}))$.

(4) Suppose $|X| = |Y|$ and the determinant of a (ρ, σ)-adapted matrix A is nonzero. What can you say about the associated interval decomposition of $\theta(A)$?

4 Čech and Vietoris–Rips filtered complexes — 3 April 2024

Today we will analyze the two main filtered spaces (abstract simplicial complexes) that arise from point clouds, the Čech and Vietoris–Rips filtered complexes. In preparation, we will discuss the nerve construction and lemma, a crucial tool guaranteeing that, under certain hypotheses, the Čech complex recovers the homotopy type of a space from which points are sampled. We will also recast point clouds as finite metric spaces, a framework that will allow us to apply persistence homology to additional data types.

4.1 The nerve of an open cover

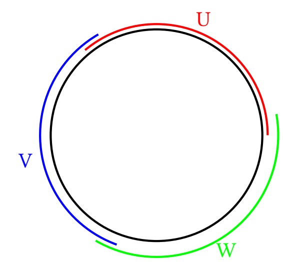
Let X be a topological space. An *open cover* of X is a collection $\mathfrak{U} = \{U_\alpha \mid \alpha \in A\}$ of open sets of X such that

$$X = \bigcup_{\alpha \in A} U_\alpha.$$

We have followed tradition in using mathfrak U, $\mathfrak{U}$, for an open cover. When we need another open cover we'll use mathfrak V, *i.e.*, $\mathfrak{V}$. We leave the reason for the incscrutability of this font as an exercise for the reader.

Given an open cover $\mathfrak{U}$, its *nerve* $N(\mathfrak{U})$ is the abstract simplicial complex with vertex set $\mathfrak{U}$ and a subset $\{U_\beta \mid \beta \in B \subseteq A\}$ in $N(\mathfrak{U})$ if and only if

$$\bigcap_{\beta \in B} U_\beta \neq \varnothing.$$



Example 4.1. Suppose $X = S^1$ and $\mathfrak{U} = \{U, V, W\}$ as illustrated. Then the vertices of $N(\mathfrak{U})$ are U, V, W (now viewed as 0-dimensional points), the edges of $N(\mathfrak{U})$ are $\{U, V\}$, $\{V, W\}$, and $\{W, U\}$. There are no 2- or higher-dimensional simplices as $U \cap V \cap W = \varnothing$.

The following Nerve Lemma tells us that nice open covers of nice spaces have nerves homotopy equivalent to the original space.

Lemma 4.2 (Nerve Lemma)**.** *Suppose $\mathfrak{U}$ is an open cover of a paracompact*[24] *space X. Suppose further that every nonempty intersection of finitely many elements of $\mathfrak{U}$ is contractible. Then*

$$X \simeq N(\mathfrak{U}).$$

Proof. This is Corollary 4G.3 of Hatcher.[25] The entirety of Section 4G is a fun read which is secretly about homotopy colimits. □

[24] A space is *paracompact* when every open cover has an open refinement that is locally finite. Here $\mathfrak{V}$ is a *refinement* of $\mathfrak{U}$ when every set in $\mathfrak{V}$ is a subset of a set in $\mathfrak{U}$. A cover is *locally finite* when every point in X has an open neighborhood that intersects only finitely many sets in the cover. Note that every subspace of $\mathbb{R}^N$ and every metric space is paracompact.

[25] Hatcher, A. (2002). *Algebraic topology*. Cambridge University Press, Cambridge

4.2 The Čech filtered complex

We now recast[26] the Čech filtered complex from the first lecture in terms of nerves. Fix a point cloud $P \subseteq \mathbb{R}^N$ and a scale $s \in \mathbb{R}$. Let $\mathfrak{U}_s(P) := \{B_r(x) \mid x \in P\}$ be the collection of radius r open balls centered at points in P; this is (tautologically) a cover of the subspace

$$X_s := \bigcup_{x \in P} B_r(x) \subseteq \mathbb{R}^N.$$

[26] In the first lecture, I used closed balls instead of open balls to define the Čech complex. We are now switching to the more standard convention of using open balls.

The *Čech complex* of P at scale s is the abstract simplicial complex

$$\check{\mathbb{C}}_s(P) := N(\mathfrak{U}_s(P)).$$

By labeling the vertices of $\check{\mathbb{C}}_s(P)$ by elements of P (rather than by open balls $B_r(x)$ for $x \in P$), we see that there is a natural simplicial inclusion $\check{\mathbb{C}}_s(P) \subseteq \check{\mathbb{C}}_t(P)$ for $s \leq t$. This makes

$$\check{\mathbb{C}}(P) := \{\check{\mathbb{C}}_s(P) \mid s \in \mathbb{R}\}$$

into an $\mathbb{R}$-filtered abstract simplicial complex which we call the *Čech filtered complex* of P.

Proposition 4.3. *For* $s > 0$,

$$\check{\mathbb{C}}_s(P) \simeq \bigcup_{x \in P} B_s(x).$$

Proof. This follows from the Nerve Lemma (Lemma 4.2) because the intersection of Euclidean balls is a convex open subset of $\mathbb{R}^N$ and hence contractible. □

One of your exercises is to check that convex open subsets of $\mathbb{R}^N$ are contractible, *i.e.*, homotopy equivalent to a point.

4.3 The Vietoris–Rips filtered complex

As demonstrated by the Nerve Lemma (Lemma 4.2), the Čech construction has very desirable theoretical properties. It is not, though, always the best tool for the job. First, it is computationally expensive to determine when k-fold intersections of open balls are nonempty. Second, our "data" might not come from a point cloud in $\mathbb{R}^N$ or indeed be sampled from any ambient metric space. For these reasons, we introduce the Vietoris–Rips complex of a (usually finite) metric space P.

Recall that a *metric space* is a set P equipped with a function $d\colon P \times P \to \mathbb{R}_{\geq 0}$ such that

» $d(x,y) = 0$ if and only if $x = y$ (identity of indiscernibles),

» $d(x,y) = d(y,x)$ (symmetry), and

» $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality).

When P is a point cloud in $\mathbb{R}^N$, we may endow it with the *induced metric*, where $d(x,y) = \|y - x\|$, but there are many other examples of metric spaces, many of which do not embed isometrically in Euclidean space. (See Morgan[27] for precise criteria on when this is possible.)

[27] Morgan, C. L. (1974). Embedding matric spaces in Euclidean space. *J. Geom.*, 5:101–107

Definition 4.4. The *Vietoris–Rips complex* of a metric space (P,d) at scale $s \in \mathbb{R}$, denoted $\mathbb{VR}_s(P)$, is the abstract simplicial complex with vertex set P and for which $\{x_0, \ldots, x_k\} \subseteq P$ is a k-simplex in $\mathbb{VR}_s(P)$ if and only if

$$d(x_i, x_j) < s$$

for $0 \leq i < j \leq k$.

It is clear that $\mathbb{VR}_s(P) \subseteq \mathbb{VR}_t(P)$ for $s \leq t$. Further, we see that $\{x_0, \dots, x_k\}$ is in $\mathbb{VR}_s(P)$ if and only if each $\{x_i, x_j\}$ is in $\mathbb{VR}_s(P)$. This makes $\mathbb{VR}_s(P)$ a *flag complex* — a simplicial complex determined by its 1-skeleton.

Slightly more subtle is the relation between Vietoris–Rips and Čech complexes:

Proposition 4.5. *For a point cloud $P \subseteq \mathbb{R}^N$ and scale $s \geq 0$, we have*

$$\mathbb{VR}_{s/2}(P) \subseteq \check{\mathbb{C}}_s(P) \subseteq \mathbb{VR}_{2s}(P).$$

Furthermore, the flag complex on $\check{\mathbb{C}}_s(P)$ is exactly $\mathbb{VR}_{2s}(P)$.

Every abstract simplicial complex K induces a flag complex $\mathscr{F}K$ determined by the 1-skeleton of K. The set $\{x_0, \dots, x_k\}$ is a k-simplex of $\mathscr{F}K$ if and only if each $\{x_i, x_j\}$ is a 1-simplex of K. Of course, we always have $K \subseteq \mathscr{F}K$.

Proof. We begin by checking that $\check{\mathbb{C}}_s(P) \subseteq \mathbb{VR}_{2s}(P)$. We must verify that $\{x_0, \dots, x_k\} \in \check{\mathbb{C}}_s(P)$ implies $d(x_i, x_j) < 2s$ for all i, j. Since the intersection of all the $B_s(x_i)$ is nonempty, we may choose some z in their intersection. By the triangle inequality,

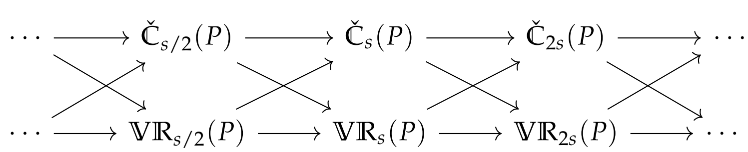
$$d(x_i, x_j) \leq d(x_i, z) + d(z, x_j) < 2s,$$

as desired.

Next, suppose that $\{x_0, \dots, x_k\}$ is a k-simplex of $\mathbb{VR}_{s/2}(P)$. Since $d(x_0, x_1) < s/2$, we may choose $z \in B_{s/4}(x_0) \cap B_{s/4}(x_1)$, in which case $d(z, x_0) < s/4$. For $0 \leq i \leq k$ we have $d(z, x_i) \leq d(z, x_0) + d(x_0, x_i) < 3s/4 < s$. Thus $z \in \bigcap_{i=0}^k B_s(x_i) \neq \varnothing$, so $\{x_0, \dots, x_k\}$ is a k-simplex of $\check{\mathbb{C}}_s(P)$. This shows that $\mathbb{VR}_{s/2}(P) \subseteq \check{\mathbb{C}}_s(P)$.

For the final assertion, we already know that $\mathbb{VR}_{2s}(P)$ is flag and contains $\check{\mathbb{C}}_s(P)$, so we automatically have $\mathscr{F}\check{\mathbb{C}}_s(P) \subseteq \mathbb{VR}_{2s}(P)$. It remains to show that if $\{x_0, \dots, x_k\}$ is a k-simplex of $\mathbb{VR}_{2s}(P)$, then each $\{x_i, x_j\}$ is a 1-simplex of $\check{\mathbb{C}}_s(P)$. We know that $d(x_i, x_j) < 2s$, so the midpoint of the line segment connecting x_i and x_j witnesses that $B_s(x_i) \cap B_s(x_j) \neq \varnothing$. This completes the argument. □

We may visualize the multiplicative interleaving of the Čech and Vietoris–Rips filtered complexes with the following diagram:



4.4 Notes

The presentation here follows Section 4.3 of Carlsson–Vejdemo-Johansson.[28] There are other filtered complexes popular in TDA such as the alpha complex, witness complex, and Delaunay complex, but we will not cover them here.

[28] Carlsson, G. and Vejdemo-Johansson, M. (2022). *Topological data analysis with applications.* Cambridge University Press, Cambridge

4.5 Exercises

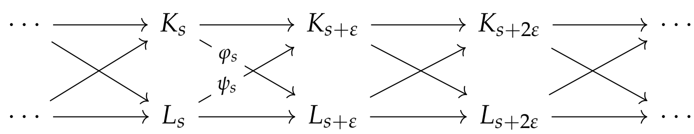
(1) Suppose that $\mathfrak{U}$ is an open cover of X that includes the open set X. Prove that $N(\mathfrak{U}) \simeq *$.

(2) Prove the assertion in the proof of Proposition 4.3 that any convex open subset of $\mathbb{R}^N$ is contractible.

5 *Distance and stability for persistent homology* — 8 April 2024

We now investigate the stability properties of persistent homology. Our goal is to demonstrate — in an appropriately precise fashion — that small changes to point clouds result in only small changes to persistent homology (with respect to Čech or Vietoris–Rips filtered complexes).

Let $\mathscr{K} = \{K_s\}_{s\in\mathbb{R}}$ and $\mathscr{L} = \{L_s\}_{s\in\mathbb{R}}$ be $\mathbb{R}$-filtered abstract simplicial complexes. Fix $\varepsilon > 0$. We say that $\mathscr{K}$ and $\mathscr{L}$ are *ε-interleaved* if there exist simplicial map $\varphi_s\colon K_s \to L_{s+\varepsilon}$ and $\psi_s\colon L_s \to K_{s+\varepsilon}$ for all $s \in \mathbb{R}$ such that the diagrams



commute for all $s \in \mathbb{R}$.

Example 5.1. Given a finite metric space P, the log-Čech and log-Vietoris–Rips filtered complexes $\{\check{\mathbb{C}}_{\log_2(s)}(P)\}_{s>0}$ and $\{\mathbb{VR}_{\log_2(s)}(P)\}$ are 1-interleaved.

The *interleaving distance* between $\mathscr{K}$ and $\mathscr{L}$, denoted $d_I(\mathscr{K},\mathscr{L})$, is the infimum of all values $\varepsilon > 0$ such that $\mathscr{K}$ and $\mathscr{L}$ are ε-interleaved. The reader may check that d_I is a metric on isomorphism classes of $\mathbb{R}$-filtered simplicial complexes.

Given an $\mathbb{R}$-filtered simplicial complex $\mathscr{K}$, let $K_\infty = \bigcup_{s\in\mathbb{R}} K_s$ and define the *filtration function* of $\mathscr{K}$ to be $f\colon K_\infty \to \mathbb{R}$, the simplicial function which assigns to a simplex $\sigma \in K_\infty$ the smallest scale at which σ appears:

$$f(\sigma) = \inf\{s \in \mathbb{R} \mid \sigma \in K_s\}.$$

Then under a mild hypothesis[29] on $\mathscr{K}$, we have that $\mathscr{K}$ is the sublevel filtration for f:

$$K_s = f^{-1}(-\infty, s].$$

[29] We require that for each scale $s \in \mathbb{R}$ there exists $s' > s$ such that $K_{s'} = K_s$. This means that simplices have a definite birth scale, as opposed to only existing on an open half-infinite interval (s, ∞).

We may also define a *filtration function* on a simplicial complex K to be a function $f\colon K \to \mathbb{R}$ such that $\sigma \subseteq \tau \implies f(\sigma) \le f(\tau)$. Then each filtration function gives rise to an $\mathbb{R}$-filtered simplicial complex via the sublevel construction.

Proposition 5.2. *Let K be a simplicial complex and suppose $f, g\colon K \to \mathbb{R}$ are filtration functions. Then the sublevel filtrations of K corresponding to f and g are $\|f - g\|_\infty$ interleaved.*

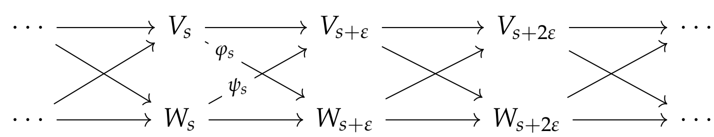
Here $\|f-g\|_\infty = \max_{\sigma\in K} |f(\sigma) - g(\sigma)|$ measures the maximal difference between f and g.

Proof. Let $\varepsilon = \|f-g\|_\infty$, $K_s = f^{-1}(-\infty, s]$, and $L_s = g^{-1}(-\infty, s]$. By the definition of $\|\ \|_\infty$, we know that $\sigma \in K_s$ implies $\sigma \in L_{s+\varepsilon}$ and similarly $\sigma \in L_s$ implies $\sigma \in K_{s+\varepsilon}$. This allows us to define the necessary interleaving functions. □

Corollary 5.3. *The interleaving distance between $\mathbb{R}$-filtered simplicial complexes is bounded above by the $\| \|_\infty$-norm of the difference between their filtration functions.*

Are they in fact equal? I don't know.

There is an analogous notion of interleaving distance between $\mathbb{R}$-persistence modules. Fix $\varepsilon > 0$ and suppose $\mathscr{V} = \{V_s\}_{s\in\mathbb{R}}$ and $\mathscr{W} = \{W_s\}_{s\in\mathbb{R}}$ are $\mathbb{R}$-persistence modules. We say that $\mathscr{V}$ and $\mathscr{W}$ are *ε-interleaved* if there exist linear transformations $\varphi_s\colon V_s \to W_{s+\varepsilon}$ and $\psi_s\colon W_s \to W_{s+\varepsilon}$ for all $s \in \mathbb{R}$ such that the diagrams



commute for all s. The *interleaving distance* between $\mathscr{V}$ and $\mathscr{W}$ is

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

By functoriality of homology, we automatically have that ε-interleaved filtered complexes have ε-interleaved persistence modules, resulting in the following proposition:

Proposition 5.4. *If $\mathscr{K}, \mathscr{L}$ are $\mathbb{R}$-filtered complexes, then*

$$d_I(\mathrm{PH}_p(\mathscr{K};\mathbb{F}), \mathrm{PH}_p(\mathscr{L};\mathbb{F})) \le d_I(\mathscr{K},\mathscr{L}).$$

While nice, this result is insufficient for our purposes since interleaving of simplicial complexes is too stringent a condition to demand. To remedy this, we introduce the *Hausdorff* and *Gromov–Hausdorff* distances.

Definition 5.5. If (X,d) is a metric space and $A, B \subseteq X$, then the *Hausdorff distance* between A and B is

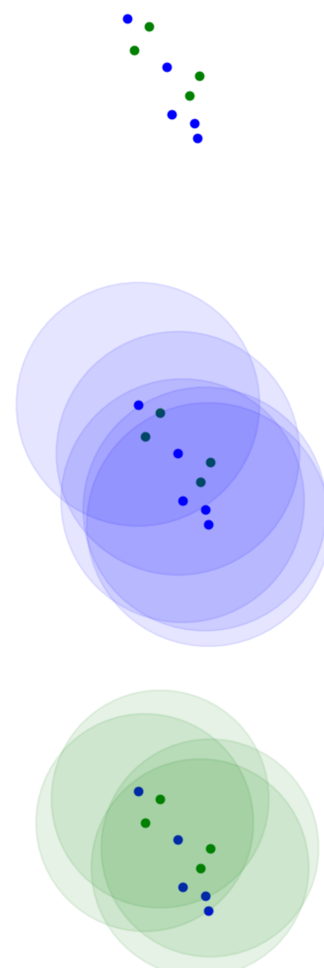
$$d_H(A,B) := \max\left\{\sup_{a\in A} d(a,B), \sup_{b\in B} d(A,b)\right\}$$

where $d(a,B) := \inf_{b\in B} d(a,b)$ and similarly for $d(A,b)$.

The reader may check that this is in fact a metric on subsets of X.

Remark 5.6. Write $N(A,s) = \bigcup_{a\in A} B_s(a)$ for the s-*inflation* of A. Then

$$d_H(A,B) = \inf\{s > 0 \mid A \subseteq N(B,s) \text{ and } B \subseteq N(A,s)\}.$$



Suppose the blue points are $A \subseteq \mathbb{R}^2$ and the green points are $B \subseteq \mathbb{R}^2$. The radius of each blue disk is $\sup_{a\in A} d(a,B)$ and the radius of each green disk is $\sup_{b\in B} d(A,b)$. The larger of these values is $d_H(A,B)$.

The Gromov–Hausdorff distance bootstraps the Hausdorff distance into a metric that compares compact metric spaces.

Definition 5.7. Suppose A and B are compact[30] metric spaces. The *Gromov–Hausdorff distance* between A and B is

$$d_{GH}(A,B) := \inf_{\mu,\nu} d_H(\mu(A),\nu(B))$$

where the infimum is taken over all isometric embeddings $\mu\colon A \to X$, $\nu\colon B \to X$ into a metric space X.

[30] A space is *compact* when every open cover of that space has a finite subcover. We will be most interested in the case where A and B are finite.

While we will not give a proof, it turns out that d_{GH} is a metric on isometry classes of compact metric spaces, and the infimum in its definition is always attained.

Proposition 5.8. *Let A, B be finite metric spaces. Then for each $p \in \mathbb{N}$,*

$$d_I(\mathrm{PH}_p(\check{\mathbb{C}}(A), \check{\mathbb{C}}(B)) \leq d_{GH}(A,B) \quad \textit{and} \quad d_I(\mathrm{PH}_p(\mathbb{VR}(A)), \mathrm{PH}_p(\mathbb{VR}(B))) \leq 2d_{GH}(A,B).$$

Proof idea. Set $\varepsilon = d_{GH}(A,B)$. It suffices to produce an ε-interleaving of $\mathrm{PH}_p(\check{\mathbb{C}}(A))$ and $\mathrm{PH}_p(\check{\mathbb{C}}(B))$, and a 2ε-interleaving of $\mathrm{PH}_p(\mathbb{VR}(A))$ and $\mathrm{PH}_p(\mathbb{VR}(B))$. The idea is to set $\varphi_s(a)$ equal to some $b \in B$ with $d(a,b) < \varepsilon$ and similarly $\psi_s(b)$ equal to some $a \in A$ with $d(a,b) < \varepsilon$. This won't induce an interleaving of filtered complexes, but will induce appropriate interleavings of persistence modules. See Proposition 10.2.9 of Virk[31] for more details, at least in the Vietoris–Rips, $p = 1$ case. □

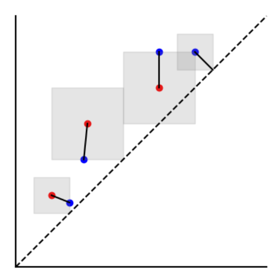
[31] Virk, Ž. (2022). Introduction to persistent homology. https://zalozba.fri.uni-lj.si/virk2022.pdf. Accessed on 19 March 2024

While interleaving distance of persistence modules has excellent formal properties, it is challenging to compute. This is remedied by *bottleneck distance*, which is equal to interleaving distance and computed purely in terms of persistence diagrams.

Let D and E denote persistence diagrams, by which we mean multisets of points in $\overline{\mathbb{R}}^2 := \mathbb{R} \times (\mathbb{R} \cup \infty)$ that lie above the diagonal $\Delta = \{(x,x) \mid x \in \mathbb{R}\}$. For $v = (x,y) \in \mathbb{R}^2$, set $\overline{v} := ((x+y)/2, (x+y)/2)$. A *partial matching* between D and E is a bijection $\varphi \colon D' \to E'$ where $D' \subseteq D$ and $E' \subseteq E$.[32] The *matching length* of φ is defined to be

[32] It is helpful to think of the points v in $D \setminus D'$ and $E \setminus E'$ being paired with $\overline{v}$. Recall that $\|(x,y)\|_\infty = \max\{|x|,|y|\}$.

$$\ell_M(\varphi) := \max\left\{\sup_{v \in D'} \|v - \varphi(v)\|_\infty, \sup_{v \in D \setminus D'} \|v - \overline{v}\|_\infty, \sup_{v \in E \setminus E'} \|v - \overline{v}\|_\infty\right\}.$$



A partial matching between red and blue persistence diagrams is indicated. The matching length is half of the side length of the largest gray box. To get the bottleneck distance, you would consider all partial matchings and find the one with the smallest largest gray box.

Definition 5.9. Let D and E be persistence diagrams and write $\mu(D,E)$ for the set of partial matchings from D to E. The *bottleneck distance* between D, E is

$$d_B(D,E) := \inf_{\varphi \in \mu(D,E)} \ell_M(\varphi).$$

Theorem 5.10. *Suppose D is the persistence diagram for $\mathrm{PH}_p(\mathscr{K})$ and E is the persistence diagram for $\mathrm{PH}_p(\mathscr{L})$. Then*

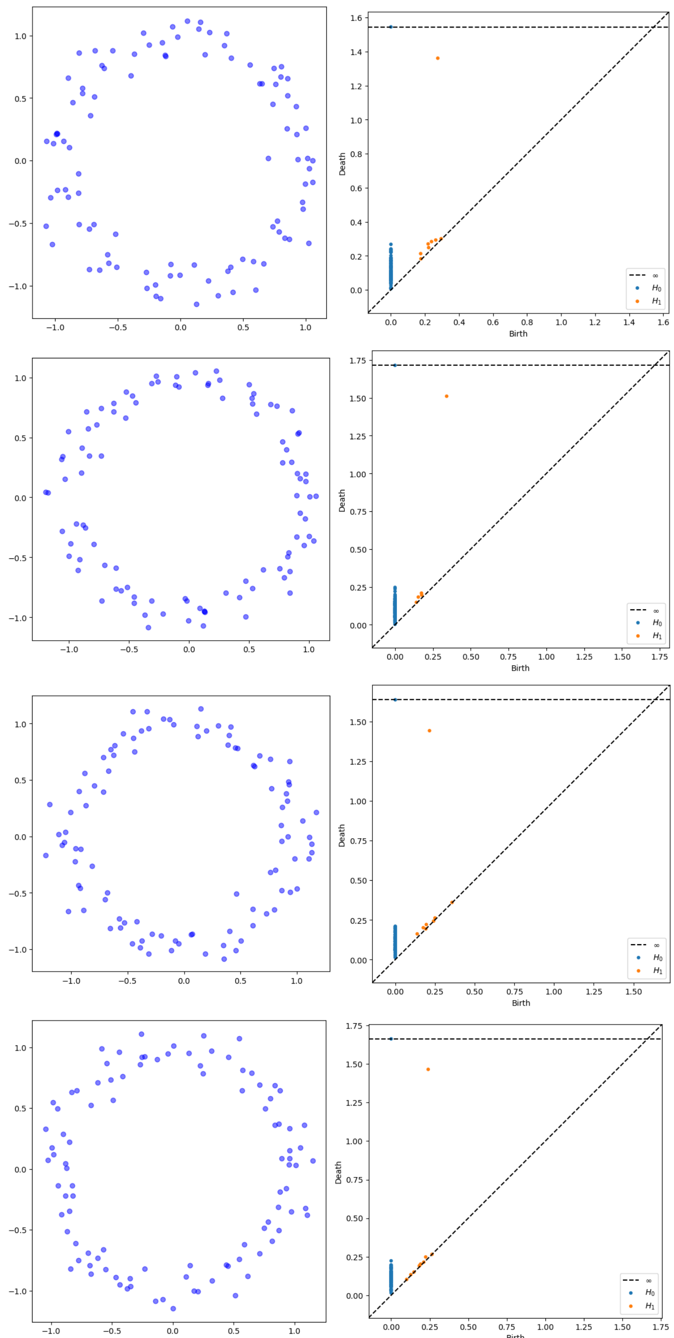
$$d_B(D,E) = d_I(\mathrm{PH}_p(\mathscr{K}), \mathrm{PH}_p(\mathscr{L})).$$

We point the reader to Chazal *et al*[33] and Bauer–Lesnick[34] for a proof of this theorem. Our classification of finitely presented $\mathbb{R}$-persistence modules allows one to reduce to interval modules, and then the key fact is that $\mathbb{I}[a,b)$ and $\mathbb{I}[a',b')$ are $\|(a,b)-(a',b')\|_\infty$-interleaved.

[33] Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L. J., and Oudot, S. Y. (2009). Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, SCG '09, page 237–246, New York, NY, USA. Association for Computing Machinery

[34] Bauer, U. and Lesnick, M. (2014). Induced matchings of barcodes and the algebraic stability of persistence. In *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, page 355–364, New York, NY, USA. Association for Computing Machinery

We conclude by exploring stability with an example. The following four pictures show different collections of 100 points sampled with noise from the same circle, along with their persistence diagrams (simultaneously depicting PH_0 and PH_1 of Vietoris–Rips complexes). Since the point clouds are close to each other in terms of Gromov–Hausdorff distance, the bottleneck distance between the persistence diagrams is small.



5.1 Notes

The presentation here follows Chapter 10 of Virk[35] and Section 5.1 of Carlsson–Vejdemo-Johansson.[36]

[35] Virk, Ž. (2022). Introduction to persistent homology. https://zalozba.fri.uni-lj.si/virk2022.pdf. Accessed on 19 March 2024

[36] Carlsson, G. and Vejdemo-Johansson, M. (2022). *Topological data analysis with applications*. Cambridge University Press, Cambridge

5.2 Exercises

(1) What does the open ball of radius s centered at a given persistence diagram look like in the bottleneck distance?

(2) Emulate the examples at the end of this section but with points sampled with noise from a sphere. Make sure to include PH_2 by setting `maxdim` equal to 2.

6 *Morse functions and zigzag persistence* — 10 April 2024

6.1 *Zigzag diagrams and their persistence*

A *zigzag diagram* of spaces is a sequence

$$\mathscr{X}: \quad X_1 \leftrightarrow X_2 \leftrightarrow \cdots \leftrightarrow X_{n-1} \leftrightarrow X_n$$

where each X_i is a space and each arrow $\leftrightarrow$ represents a continuous map $X_i \to X_{i+1}$ or $X_i \leftarrow X_{i+1}$. Applying homology $H_p(\ ;\mathbb{F})$ to this diagram results in a *zigzag module*

There is exactly one arrow between X_i and X_{i+1} for $1 \leq i < n$. The directions of arrows for different indices can change.

$$\mathrm{H}_p(\mathscr{X};\mathbb{F}): \quad \mathrm{H}_p(X_1;\mathbb{F}) \leftrightarrow \mathrm{H}_p(X_2;\mathbb{F}) \leftrightarrow \cdots \leftrightarrow \mathrm{H}_p(X_{n-1};\mathbb{F}) \leftrightarrow \mathrm{H}_p(X_n;\mathbb{F}).$$

A theorem of Gabriel[37] classifies the finite-dimensional zigzag modules as direct sums of *interval modules*

[37] Gabriel, P. (1972). Unzerlegbare Darstellungen. I. *Manuscripta Math.*, 6:71–103; correction, ibid. **6 (1972), 309**,

$$\mathbb{I}^{zz}[b,d]: \quad I_1 \leftrightarrow I_2 \leftrightarrow \cdots \leftrightarrow I_n$$

where $I_i = \mathbb{F}$ for $b \leq i \leq d$, $I_i = 0$ otherwise, and every map $\mathbb{F} \to \mathbb{F}$ or $\mathbb{F} \leftarrow \mathbb{F}$ is the identity. The list of summands in a decomposition $\mathscr{V} \cong \bigoplus_i \mathbb{I}^{zz}[b_i,d_i]$ is unique up to reordering.

The p-th *zigzag persistence diagram* of $\mathscr{X}$ is the multiset of intervals $[b_i,d_i]$ appearing in the interval decomposition of $\mathrm{H}_p(\mathscr{X};\mathbb{F})$.

6.2 *Morse type functions and levelset zigzags*

Suppose X is a space and $f\colon X \to \mathbb{R}$ is a continuous function. For $t \in \mathbb{R}$, write $X_t := f^{-1}\{t\}$; for $I \subseteq \mathbb{R}$ an interval, write $X_I := f^{-1}I$. We say that the pair (X,f) is *Morse type* when there are finitely many real numbers $a_1 < a_2 < \cdots < a_n$ — called *critical values* of f — such that for $I = (-\infty,a_1),(a_1,a_2),\ldots,(a_{n-1},a_n),(a_n,\infty)$, there is a homeomorphism $Y_I \times I \xrightarrow{\cong} X_I$ for some space Y_I via which f becomes projection onto the second factor; further, we require that the homeomorphism extends to a continuous function $Y_I \times \overline{I} \to X_{\overline{I}}$ where $\overline{I}$ is the closure of I in $\mathbb{R}$. We also assume that each X_t has finitely generated homology.

Example 6.1. Before investigating how Morse type functions produce an important case of zigzag persistence, we briefly recall their archetype, namely *Morse functions* on a smooth manifold.

Let M be a differentiable manifold and $f\colon M \to \mathbb{R}$ be a smooth function. If $df_p = 0$, then we call p a *critical point* and $f(p)$ a *critical value*; non-critical points and values are called *regular*. If t is a regular value of f, then $M_t := f^{-1}\{t\}$ is a differentiable submanifold of M, but the fibers over critical values have singularities.

When $df_p = 0$, the *Hessian* of f at p is well-defined. In local coordinates, this is the symmetric matrix of second order partial derivatives. When the Hessian is nonsingular, we call p a *non-degenerate critical point*. The smooth function f is called a *Morse function* when it has no degenerate critical points and all of its critical values are distinct. In this case, each critical point has a chart U with coordinates $(x_1, \dots, x_n)$ such that each $x_i(p) = 0$ and

$$f(x) = f(p) - x_1^2 - \cdots - x_\gamma^2 + x_{\gamma+1}^2 + \cdots + x_n^2$$

on U. We call γ the *index* of f at p. Note that index 0 corresponds to a local minimum of f, while index n (the dimension of M) is a local maximum; when $0 < \gamma < n$ the critical point is a saddle with γ decreasing directions and $n - \gamma$ increasing directions.

When $[a, b]$ is an interval in $\mathbb{R}$ containing no critical values of f, we have that $M_{[a,b]} := f^{-1}[a, b]$ is diffeomorphic to $M_a \times [a, b] \cong M_b \times [a, b]$ in a manner compatible with f.[38] If $[a, b]$ contains a single critical value $t = f(p)$ and p has index λ, then M_b is diffeomorphic to M_a with a λ-handle attached; this means that $M_b \cong M_a \cup_\varphi H$ where $H = D^\lambda \times D^{n-\lambda}$ and $\varphi\colon \partial D^\lambda \times D^{n-\lambda} \to M_a$ is an embedding. We also have $M_b \simeq M_a \cup_\psi D^\lambda$ for some $\psi\colon S^{\lambda-1} \to M_a$.[39]

[38] One proves this by choosing a Riemannian metric on M (smoothly varying choice inner product on tangent spaces) and employing *gradient flow*.

[39] While only true up to homotopy — instead of diffeomorphism — this final statement is useful for producing CW decompositions and making computations on manifolds.

Given a space X and function $f\colon X \to \mathbb{R}$ of Morse type with critical values $a_1 < \cdots < a_n \in \mathbb{R}$, we may choose $s_0, \dots, s_n \in \mathbb{R}$ such that

$$s_0 < a_1 < s_1 < a_2 < \cdots < s_{n-1} < a_n < s_n.$$

Setting $X_i^j := X_{[s_i, s_j]}$ we get a zigzag diagram

$$\mathscr{X}: \quad X_0^0 \to X_0^1 \leftarrow X_1^1 \to X_1^2 \leftarrow \cdots \to X_{n-1}^n \leftarrow X_n^n.$$

We define the *level set zigzag persistent homology* of (X, f) to be $\mathrm{H}_*(\mathscr{X}; \mathbb{F})$. Due to the product structure between critical values, it does not depend on the choice of s_i. As such, we make the convention that X_{i-1}^i is labeled by a_i, and each X_i^i is labeled by the interval (a_i, a_{i+1}) containing it (where $a_0 = -\infty$ and $a_{n+1} = \infty$). The zigzag persistence intervals of $\mathscr{X}$ are then labeled by the union of the labels of the supporting X_i^j:

$$\begin{aligned}
[X_{i-1}^i, X_{j-1}^j] &\longleftrightarrow [a_i, a_j] \\
[X_{i-1}^i, X_{j-1}^{j-1}] &\longleftrightarrow [a_i, a_j) \\
[X_i^i, X_{j-1}^j] &\longleftrightarrow (a_i, a_j] \\
[X_i^i, X_{j-1}^{j-1}] &\longleftrightarrow (a_i, a_j).
\end{aligned}$$

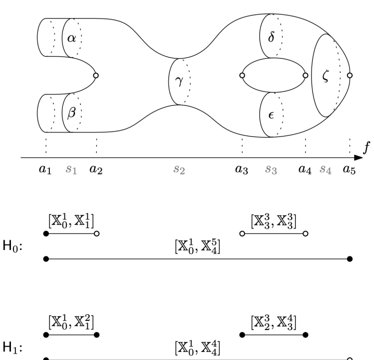


Image from (Carlsson–de Silva–Morozov, 2009).

Example 6.2. The diagram on the right depicts a Morse function on a surface with boundary in which f is projection onto the horizontal copy of the real line. Let's consider the zigzag persistence associated with

$$X_0^0 \to X_0^1 \leftarrow X_1^1 \to X_1^2 \leftarrow X_2^2 \to \cdots$$

when we apply H_1. This gives the zigzag of vector spaces

$$0 \to \mathbb{F}\{\alpha, \beta\} \leftarrow \mathbb{F}\{\alpha, \beta\} \to \mathbb{F}\{\alpha, \beta\} \xleftarrow{g} \mathbb{F}\{\gamma\} \to \cdots$$

where α, β, γ are the 1-cycles indicated in the diagram and $g(\gamma) = \alpha + \beta$. This has direct sum decomposition as

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & \mathbb{F}\{\alpha\} & \longleftarrow & \mathbb{F}\{\alpha\} & \longrightarrow & \mathbb{F}\{\alpha\} & \longleftarrow & 0 & \longrightarrow & \cdots \\ \oplus & & \oplus & & \oplus & & \oplus & & \oplus & & \\ 0 & \longrightarrow & \mathbb{F}\{\alpha+\beta\} & \longleftarrow & \mathbb{F}\{\alpha+\beta\} & \longrightarrow & \mathbb{F}\{\alpha+\beta\} & \longleftarrow & \mathbb{F}\{\gamma\} & \longrightarrow & \cdots . \end{array}$$

The upper summand gives interval $[X_0^1, X_1^2] \leftrightarrow [a_1, a_2]$ in the decomposition of $H_1(\mathscr{X}; \mathbb{F})$. The full collection of intervals for $H_1(\mathscr{X}; \mathbb{F})$ is

$$\{[X_0^1, X_4^4], [X_0^1, X_1^2], [X_2^3, X_3^4]\} \leftrightarrow \{[a_1, a_5), [a_1, a_2], [a_3, a_4]\}.$$

6.3 *Time series analysis via zigzag persistence*

Imagine that we have a simplicial complex evolving in time and we are able to sample what the complex looks like at discrete time steps resulting in a sequence of simplicial complexes $K_0, K_1, K_2, \ldots, K_N$.[40] We will additionally assume that each K_i is a subcomplex of some ambient complex $\mathbf{K}$, whence it makes sense to consider the unions of consecutive complexes $K_i \cup K_{i+1}$.

[40] It seems more likely that we might sample a point cloud at discrete time steps. After committing the cardinal sin of choosing a scale, we could get a time series of simplicial complexes. In later sections we will encounter two potential remedies to this embarassing state of affairs. First, it seems quite plausible that we might be sampling a network (graph) at discrete time steps. In this case, the associated sequence of *clique complexes* will have the desired data type. Alternatively, we might generate a *generalized persistence diagram* from a time series of point clouds: each discrete time could have an associated $\mathbb{R}$-filtered complex. We will consider the inherent complexities of this situation when we discuss *multipersistence*.

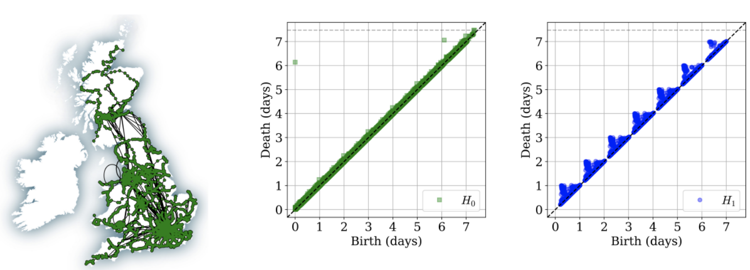
Zigzag persistence offers a method for detecting when the sequence (K_i) undergoes topological changes. We consider the zigzag diagram

$$\mathscr{K}: \quad K_0 \hookrightarrow K_0 \cup K_1 \hookleftarrow K_1 \hookrightarrow K_1 \cup K_2 \hookleftarrow \cdots \hookrightarrow K_{N-1} \cup K_N \hookleftarrow K_n$$

and the associated zigzag modules $H_p(\mathscr{K}; \mathbb{F})$. By indexing K_i by i and $K_i \cup K_{i+1}$ by $i + 1/2$, the intervals at play are of the form $[b, d]$ for $0 \leq b < d \leq N$ with b, d both integers or half-integers. We may also take b, d natural numbers between 0 and N and then work with open and half-open intervals where, *e.g.*, $[b, d) = [b, d - 1/2]$.

The following image from Myers *et al*[41] depicts a zigzag analysis of a time series of simplicial complexes extracted from the British rail network.

[41] Myers, A., Muñoz, D., Khasawneh, F. A., and Munch, E. (2023). Temporal network analysis using zigzag persistence. *EPJ Data Science*, 12(1):6



(a) Full Rail Travel Network. (b) Zero-dimensional zigzag persistence. (c) One-dimensional zigzag persistence.

The highly persistent H_0 class with birth time 0 and death 6 indicates that some component of the network remains connected Monday–Saturday. Meanwhile, the daily periodicity of the H_1 persistence diagram with midday “peaks” indicates that many loops exist in the network midday but are destroyed overnight when the trains are snuggled into bed.

6.4 Notes

The presentation here is primarily based on Carlsson–de Silva–Morozov.[42] We did not discuss one of the most remarkable pieces of their paper, namely the Pyramid Theorem which relates zigzag persistence to *extended persistent homology* in the sense of Cohen-Steiner–Edelsbrunner–Harer[43] via a diagram of Mayer–Vietoris diamonds.

Computing zigzag persistence is not supported by Ripser, but is available in Dionysus; see https://mrzv.org/software/dionysus2/.

[42] Carlsson, G., de Silva, V., and Morozov, D. (2009). Zigzag persistent homology and real-valued functions. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, SCG '09, page 247–256, New York, NY, USA. Association for Computing Machinery

[43] Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2009). Extending persistence using Poincaré and Lefschetz duality. *Found. Comput. Math.*, 9(1):79–103

6.5 Exercises

(1) Work out the rest of the H_1 persistence diagram for the Morse function discussed in Example 6.2.

7 *Topological analysis of weighted graphs* — 15 April 2024

7.1 *Undirected graphs*

A *(simple) graph* is a pair $G = (V, E)$, where V is the set of *vertices* and $E \subseteq \binom{V}{2}$ is a collection of edges. A *weighted (simple) graph* is a triple $G = (V, E, w)$ where (V, E) is a (simple) graph and $w\colon E \to \mathbb{R}$ assigns a *weight* to each edge.

Simple graphs have no self-loops, no multiple edges between vertices, and no direction on the edges.

Topologically, a graph is a 1-dimensional abstract simplicial complex. Thus, if the graph is connected, it is homotopic to a bouquet of circles.[44] There is another simplicial complex, though, that we may assign to a graph, namely its *clique* or *flag complex* CG. This complex has vertex set V and contains a k-simplex $\{v_0, \ldots, v_k\}$ if and only if each edge $\{v_i, v_j\}$ is in G; in other words, the k-simplices of CG are exactly the $(k+1)$-cliques of G.

[44] *I.e.*, a wedge (one-point union), $\bigvee S^1$.

A *clique* is a complete induced subgraph.

Given a weighted graph $G = (V, E, w)$, for each $s \in \mathbb{R}$ we define a simple graph

$$G(s) := (V, w^{-1}(-\infty, s]) = (V, \{e \in E \mid w(e) \leq s\}).$$

This produces an $\mathbb{R}$-filtered graph $\{G(s)\}_{s\in\mathbb{R}}$ and $\mathbb{R}$-filtered simplicial complex

$$\mathscr{C}G := \{CG(s)\}_{s\in\mathbb{R}}$$

called the *clique filtered complex* of G. The following proposition follows directly from definitions:

Proposition 7.1. *If G is a finite weighted graph, then for s sufficiently negative, $CG(s)$ is the 0-dimensional complex with vertices V, and there exists $S \in \mathbb{R}$ such that for $s \geq S$, $CG(s) = \bigcup_{t\leq S} CG(t)$.*

Remark 7.2. If the weight function w, viewed as a symmetric function $V \times V \to \mathbb{R}$, $(u, v) \mapsto w(\{u, v\})$, is a metric, then the clique filtered complex of G is identical to the Vietoris–Rips filtered complex of the metric space (V, w).

By applying $\mathrm{H}_p(\ ;\mathbb{F})$ to $\mathscr{C}G$, we get an $\mathbb{R}$-persistence module which we denote $\mathrm{PH}_p(G;\mathbb{F})$ and call the *p-th persistent homology group of G* (with coefficients in $\mathbb{F}$). We automatically have that $\mathrm{PH}_p(G;\mathbb{F}) = 0$ for $p \geq |V|$.

7.2 *Experiments*

See the github repository for code generating the following table. What do you notice? What do you wonder?

7.3 *Directed graphs*

A *directed graph* is a pair $G = (V, \vec{E})$ where $\vec{E} \subseteq V \times V \setminus \Delta$ is the directed edge set. We view $(u, v) \in \vec{E}$ as directed from u to v and may draw it as

We are removing the diagonal $\Delta = \{(v, v) \mid v \in V\}$ to avoid self-loops.

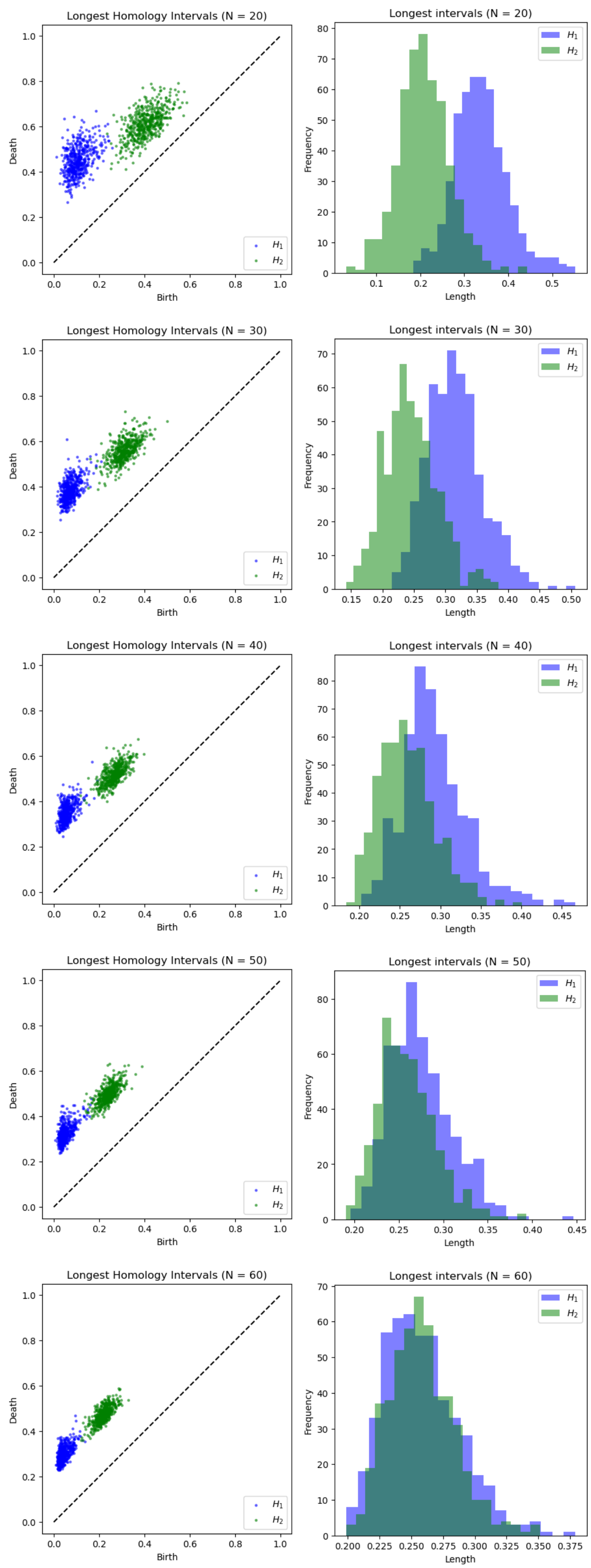


Table 1: Plots of longest (most persistent) (birth, death) scales of PH_1 and PH_2 classes for 500 complete weighted graphs with weights sampled uniformly randomly from $[0, 1]$.

$u \to v$. A *weighted directed graph* is a triple $G = (V, \vec{E}, w)$ where $(V, \vec{E})$ is a directed graph and $w\colon \vec{E} \to \mathbb{R}$ asigns a weight to each directed edge.

Define the *in-degree* of a vertex v to be

$$\deg_{\text{in}}(v) := |\{u \in V \mid (u, v) \in \vec{E}\}|$$

and define its *out-degree* to be

$$\deg_{\text{out}}(v) := |\{u \in V \mid (v, u) \in \vec{E}\}|.$$

We call a vertex a *source* when its in-degree is 0, and a *sink* when its out-degree is 0.

We would like to produce a directed version of the of the clique complex, and thus need a notion of directed clique. To build up this notion, first recall that a *directed cycle* is a sequence $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k), (v_k, v_0)$ of directed edged in G. We call G a *directed acyclic graph* (DAG) when it contains no directed cycles. A *directed k-clique* of G is an induced subgraph on k vertices $\{v_1, \ldots, v_k\}$ that is a DAG and such that there is exactly one edge between any unordered pair of vertices. By reordering the vertices, we may write a directed k-clique as an ordered k-tuple of vertices $(v_1, \ldots, v_k)$ such that the edges between the v_i are exactly $v_i \to v_j$ for $1 \le i < j \le n$.

Definition 7.3. If $G = (V, \vec{E})$ is a directed graph, then the *directed clique complex* of G is the abstract simplicial complex $\vec{C}G$ whose k-simplices are directed $(k+1)$-cliques of G.

If $G = (V, \vec{E}, w)$ is a weighted directed graph, then we may once again threshhold to get an $\mathbb{R}$-filtered directed graph $\{G(s)\}_{s \in \mathbb{R}}$ and $\mathbb{R}$-filtered simplicial complex

$$\vec{\mathscr{C}}G := \{\vec{C}G(s)\}_{s \in \mathbb{R}}.$$

Applying $\mathrm{H}_p(\ ; \mathbb{F})$, we get the persistence module $\mathrm{PH}_p(G; \mathbb{F})$.

7.4 Application: the Blue Brain Project

The work of Reimann *et al.*[45] with EPFL's Blue Brain Project (https://www.epfl.ch/research/domains/bluebrain/) uses directed clique complexes to analyze the neuronal connectivity of a digitally reconstructed mouse brain. *In vitro* experiments on the brain connectome are limited by the number of neurons that can be recorded simultaneously. The Blue Brain Project synthesizes a mouse connectome *in silica* in order to explore large scale structure in a synthetic connectome with $\sim 30,000$ neurons. A directed graph of neurons is determined by pre- and post-synaptic structure of chemical (as opposed to electric) synapses. In a functional network, the authors plot the Euler characteristic and Betti numbers of the network to detect time-variation in the network.

[45] Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H. (2017). Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in Computational Neuroscience*, 11

Using these methods, the authors discovered directed simplices up to dimension 7 in the reconstructed connectome, and they also demonstrate that the topology of the network organizes spike correlations. They do not employ zigzag persistence to analyze their time-varying network; it would be interesting to see how this tool captures the evolving electrical activity of the connectome.

The following figure is copied from Reimann *et al.* (2017):

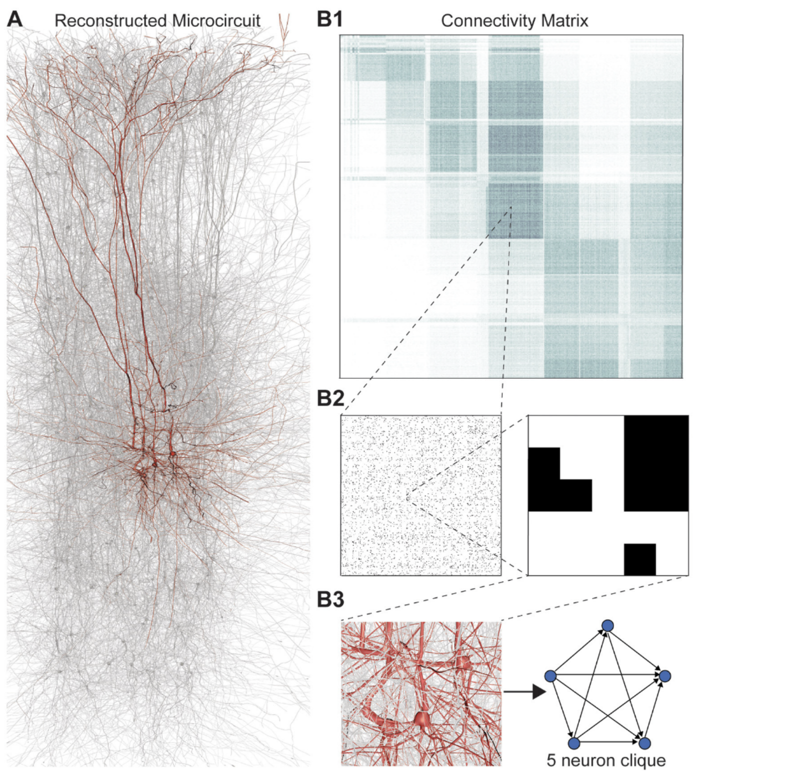


FIGURE 1 | (A) Thin (10 μm) slice of *in silico* reconstructed tissue. Red: A clique formed by five pyramidal cells in layer 5. **(B1)** Full connection matrix of a reconstructed microcircuit with 31,146 neurons. Neurons are sorted by cortical layer and morphological type within each layer. Pre-/postsynaptic neurons along the vertical/horizontal axis. Each grayscale pixel indicates the connections between two groups of 62 neurons each, ranging from white (no connections) to black ($\geq$8% connected pairs). **(B2)** Zoom into the connectivity between two groups of 434 neurons each in layer 5, i.e., 7 by 7 pixels in **(A)**, followed by a further zoom into the clique of 5 neurons shown in **(A)**. Black indicates presence, and white absence of a connection. **(B3)** Zoom into the somata of the clique in **(A)** and representation of their connectivity as a directed graph.

7.5 Notes

Some of this material is covered in Chapter 8 of Dey–Wang.[46] That source also describes path homology, an interesting alternative to simplicial homology of directed clique complexes when working with data naturally encoded as a weighted directed graph.

[46] Dey, T. K. and Wang, Y. (2022). *Computational topology for data analysis.* Cambridge University Press, Cambridge

7.6 Exercises

(1) Adapt the code used to generate Table 1 so that it also plots results for PH_3.

(2) Keep playing with the code and generate some conjectures.

(3) Prove your conjectures, write a paper, and get it published.

8 *Poset-filtered objects and generalized persistence* — 17 April 2024

We will spend the next couple of lectures delving into generalized and multiparameter persistence. This is a rapidly developing research area that considers situations in which data is filtered by multiple parameters or, more generally, a partially ordered set. The structure theory for generalized persistence modules is wildly[47] more complicated, and the research community is actively searching for better invariants and hypotheses that will permit the fruitful application of generalized persistence.

[47] In a technical sense.

Today's lecture will define generalized persistence and then explore several methods for producing poset-filtered simplicial complexes from data.

Definition 8.1. A *partially ordered set* (or *poset*) is a pair $(P, \leq)$ of a set P and a relation $\leq$ which is

We will commonly just write P for a poset when the relation $\leq$ is understood by context.

» reflexive ($x \leq x$),

» antisymmetric ($x \leq y$ and $y \leq x$ implies $x = y$), and

» transitive ($x \leq y$ and $y \leq z$ implies $x \leq z$).

A partial order is *total* when every pair of elements $x, y \in P$ is comparable ($x \leq y$ or $y \leq x$). Examples of total orders are $(\mathbb{Z}, \leq)$ and $(\mathbb{R}, \leq)$, but there are many partial orders which are not total. For instance, consider the set $\mathbb{N} = \{0, 1, 2, \ldots\}$ with the relation $x \preccurlyeq y$ if and only if x divides y (*i.e.*, there exists a natural number z such that $xz = y$). There are many incomparable numbers under the divisibility relation, but the reader may check that $\preccurlyeq$ is still a partial order.

This is often denoted $x \mid y$, but the lack of directionality in the pipe symbol can lead to confusion.

Example 8.2. Given posets P and Q there is an induced partial order on $P \times Q$ given by $(x, y) \leq (x', y')$ if and only if $x \leq x'$ and $y \leq y'$. This will be the standard partial order we consider on $\mathbb{Z}^2$ and $\mathbb{R}^2$, and similarly for higher Cartesian powers of these sets.

There is a category of partially ordered sets in which the morphisms $f\colon P \to Q$ are the *monotonic* (also called *nondecreasing*) functions, *i.e.*, those satisfying $x \leq y \implies f(x) \leq f(y)$. An *isomorphism* of posets is a monotonic function admitting a monotonic inverse; this is equivalent to being a monotonic bijection.

Example 8.3. Write $\underline{n} := \{1, 2, \ldots, n\}$ for the finite linearly ordered chain with $1 < 2 < \cdots < n$. By the construction of Example 8.2, we get an induced partial order on $\underline{m} \times \underline{n}$ which we refer to as a *rectangular grid* poset. It is amusing to note that $\underline{m} \times \underline{n}$ is isomorphic to $\{p^a q^b \mid 0 \leq a \leq m,\ 0 \leq b \leq n\}$ under divisibility.

In general we write $x < y$ for $x \leq y$ and $x \neq y$.

Every poset P induces a category with objects the underlying set P and a unique morphism $x \to y$ if and only if $x \leq y$. We will abuse notation and write P for the category induced by P.

Definition 8.4. Let P be a poset. A *P-filtered object* in a category C is a functor[48] $P \to \mathsf{C}$. A *morphism* of P-filtered objects is a natural transformation.[49] For any category C and poset P, we write C^P for the category of P-filtered objects in C and morphisms.

[48] A functor F between categories C, D, denoted $F\colon \mathsf{C} \to \mathsf{D}$ is a function from the objects of C to the objects of D (also denoted F) along with a function from morphisms of C to morphisms of D (also denoted F) such that $f\colon c \to c'$ in C implies $Ff\colon Fc \to Fc'$ in D that also takes identity morphisms to identity morphisms and respects composition.

[49] For functors $F, G\colon \mathsf{C} \to \mathsf{D}$, a natural transformation $\alpha\colon F \implies G$ is the data of morphisms $\alpha_c\colon Fc \to Gc$ in D for each object c of C such that for each morphism $f\colon c \to c'$ in C the diagram

$$\begin{array}{ccc} Fc & \xrightarrow{\alpha_c} & Gc \\ {\scriptstyle Ff}\downarrow & & \downarrow{\scriptstyle Gf} \\ Fc' & \xrightarrow[\alpha_{c'}]{} & Gc' \end{array}$$

commutes.

We will often write $\{K_x\}_{x \in P}$ for a P-filtered object $K\colon P \to \mathsf{C}$. This omits the morphisms $K_x \to K_y$ for $x \le y$ from notation, but we should not forget them!

We will be particularly interested in P-filtered spaces and P-filtered vector spaces. Write sCpx for the category of abstract simplicial complexes and simplicial maps, and write $\mathsf{Vec}_{\mathbb{F}}$ for the category of $\mathbb{F}$-vector spaces and $\mathbb{F}$-linear transformations; the full subcategory of finite-dimensional $\mathbb{F}$-vector spaces will be denoted $\mathsf{vec}_{\mathbb{F}}$.

Definition 8.5. For a poset P, a *P-persistence module* over $\mathbb{F}$ is a P-filtered $\mathbb{F}$-vector space. If a persistence module lies in the subcategory $\mathsf{vec}_{\mathbb{F}}^P$ we call it *pointwise finite-dimensional.*

By taking p-th homology levelwise, we get a functor

$$\mathsf{sCpx}^P \xrightarrow{\qquad \mathrm{H}_p \qquad} \mathsf{Vec}_{\mathbb{F}}^P$$

$$\{K_x\}_{x \in P} \longmapsto \{\mathrm{H}_p(K_x; \mathbb{F})\}_{x \in P}.$$

We would like to study data via P-persistence modules. To do so, we need methods that produce topologically meaningful P-filtered simplicial complexes from data, and we also need good invariants for P-persistence modules. We'll look at the first step today, and investigate invariants of generalized persistence modules during our next meeting.

While it is interesting to study poset-filtrations in general, the most common posets arising in practice are $\mathbb{Z}^n$ and $\mathbb{R}^n$ with $\vec{x} \le \vec{y}$ if and only $x_i \le y_i$ for $1 \le i \le n$. We will refer to both $\mathbb{Z}^n$- and $\mathbb{R}^n$-filtered objects as *multifiltered.* Among these, the $n = 2, 3$ cases of *bifiltered* and *trifiltered* objects are most common. We will also use these terms when one or more of the coordinate orders $\le$ is replaced with $\ge$.

Definition 8.6. For a finite metric space (X, d), $s \in \mathbb{R}$, and $t \in \mathbb{N}$, we define $\mathbb{VR}^{\mathrm{deg}}_{s,t}(X)$ to be the maximal subcomplex of $\mathbb{VR}_s(X)$ whose vertices have degree at least $t - 1$ in the 1-skeleton of $\mathbb{VR}_s(X)$. Then $\mathbb{VR}^{\mathrm{deg}}(X)$ is an $\mathbb{R} \times \mathbb{N}^{\mathrm{op}}$-filtered simplicial complex, where $\mathbb{N}^{\mathrm{op}} := (\mathbb{N}, \ge)$. We call $\mathbb{VR}^{\mathrm{deg}}(X)$ the *degree-Vietoris–Rips bifiltration.*

Definition 8.7. Suppose X is a finite metric space and $\delta\colon X \to \mathbb{R}$ is a function. We might think of δ as a *density* on the points of X (presumably under the hypothesis that $\delta \ge 0$), and we might view denser points as those which are more significant or those in which we have more confidence. For $s, t \in \mathbb{R}$, we define

$$\mathbb{VR}^{\downarrow}_{s,t}(\delta) := \mathbb{VR}_s(\delta^{-1}[t, \infty))$$

to be the radius s Vietoris–Rips complex on points in X of density at least t. Then $\mathbb{VR}^{\downarrow}(\delta)$ is an $\mathbb{R} \times \mathbb{R}^{\mathrm{op}}$-filtered simplicial complex called the *density-Vietoris–Rips bifltration* of (X, δ).

Definition 8.8. Given a finite metric metric space X, define

$$\begin{aligned} \gamma \colon X &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{|X|} \sum_{y \in X} d(x,y) \end{aligned}$$

to be the average distance from a point $x \in X$ to the other points of X; we call γ an *eccentricity function* and may define $\mathbb{VR}^{\downarrow}(\gamma)$ in the same way as the previous example; it is called the *eccentricity-Vietoris–Rips bifiltration* of X.

Of course, both the density- and eccentricity-Vietoris–Rips bifiltrations may be modified to use sublevel instead of superlevel sets. In the following example, we demonstrate how to build a direction-agnostic multifiltration. Define $\overline{U}$ to be the subposet of $\mathbb{R}^{\mathrm{op}} \times \mathbb{R}$ consisting of pairs (a, b) with $a \leq b$. Note that $(a, b) \leq (c, d)$ in $\overline{U}$ if and only $a \geq c$ and $b \leq d$; we can think of this as the interval $[a, b]$ being a subinterval of $[c, d]$.

Definition 8.9. For a finite metric space X and function $f \colon X \to \mathbb{R}$, we define the *interlevel-Vietoris–Rips trifiltration* $\overline{U} \times \mathbb{R} \to \mathsf{sCpx}$ by

$$\mathbb{VR}(f)_{(a,b),s} := \mathbb{VR}_s(f^{-1}[a, b]).$$

We conclude with a multifiltered generalization of clique complexes.

Definition 8.10. Supoose $G = (V, E, w \colon E \to \mathbb{R}^n)$ is a multiweighted graph. For $s_1, \ldots, s_n \in \mathbb{R}$, we may set

$$\mathscr{C}_{s_1,\ldots,s_n}(G) := C(G(s_1, \ldots, s_n))$$

where $G(s_1, \ldots, s_n)$ is the simple graph with vertex set V and edges

$$E(s_1, \ldots, s_n) = w^{-1}((-\infty, s_1] \times \cdots \times (-\infty, s_n]).$$

Then $\mathscr{C}(G)$ is an $\mathbb{R}^n$-filtered simplicial complex called the *multifiltered clique complex* of G.

8.1 Notes

Some of the material here — especially the definitions of various multifiltrations — is drawn from Botnan–Lesnick.[50]

[50] Botnan, M. B. and Lesnick, M. (2023). An introduction to multiparameter persistence. arXiv:2203.14289

8.2 Exercises

(1) Find a monotonic injection of the divisibility poset $(\mathbb{N}, \preccurlyeq)$ into the product of countably many copies of $\{0 < 1 < 2 < \cdots < \infty\}$. (*Hint*: Prime factorization.)

(2) For a poset P, check that $H_p\colon \mathsf{sCpx}^P \to \mathsf{Vec}_{\mathbb{F}}^P$ is really a functor. In particular, check that it respects so-called *vertical composition* of natural transformations.

9 *Structure and invariants for generalized persistence modules —* 22 April 2024

In the case of finitely presented $\mathbb{Z}$- and $\mathbb{R}$-persistence modules, we observed that barcodes / persistence diagrams provide a complete invariant in the sense that two such persistence modules are isomorphic if and only if they have the same barcode (up to permuatation). The same was true for zigzag persistence modules. The situation is far less well-behaved for more general persistence modules, but a few pleasant properties persist to this level of generality. Throughout, let P be a poset.

Definition 9.1. A P-persistence module $M\colon P \to \mathsf{Vec}_{\mathbb{F}}$ is *indecomposable* when it cannot be expressed up to isomorphism as the direct sum of two nontrivial P-persistence modules.

Definition 9.2. A P-persistence module is *Noetherian* when every ascending chain of submodules stabilizes; it is *Artinian* when ever descending chain of submodules stabilizes; it has *finite length* when it is both Noetherian and Artinian.

Theorem 9.3 (Krull–Schmidt)**.** *Let P be a poset. If a P-persistence module $M\colon P \to \mathsf{Vec}_{\mathbb{F}}$ has finite length, then it has a decomposition into indecomposables*

$$M \cong \bigoplus_{\lambda \in \Lambda} M^{\lambda}$$

which is unique up to permuatation.

If $P = \mathbb{Z}$ or $\mathbb{R}$ or is of type A_n (a zigzag), then the indecomposables are exactly the interval modules $\mathbb{I}[b,d)$. But — alas and alack — almost all other posets have *wild type* representation theory. This is a consequence of Gabriel's theorem[51] on quiver representations. For the sake of simplicity, we limit the ensuing discussion to finite posets, but the results adapt infinite posets in a natural way.

[51] Gabriel, P. (1972). Unzerlegbare Darstellungen. I. *Manuscripta Math.*, 6:71–103; correction, ibid. **6 (1972), 309**,

Definition 9.4. Let P be a finite poset. Then P has *finite representation type* when $\mathsf{vec}_{\mathbb{F}}^{P}$ has a finite number of indecomposables up to isomorphism.

We say P has *tame representation type* when, up to isomorphism, all but finitely many indecomposables in $\mathsf{vec}_{\mathbb{F}}^{P}$ appear in one of finitely many one-parameter families[52] of representations.

We say P has *wild representation type* when, up to isomorphism, $\mathsf{vec}_{\mathbb{F}}^{P}$ contains a two-parameter family of indecomposables.

[52] We won't define this term precisely, but you'll get a feeling for it as soon as we look at some examples.

Drozd's theorem implies that for $\mathbb{F}$ algebraically closed, every P has finite, tame, or wild representation type.

For brevity, let's start saying P-module instead of P-persistence module. A common class of indecomposable P-module is an interval module.

Definition 9.5. An *interval*[53] I in a poset P is a nonempty subset of P such that

[53] We caution the reader that this is *not* the most common definition of intervals in posets. The standard version is either (i) a pair $x \le y$ in P or, equivalently, (ii) $[x,y] := \{z \in P \mid x \le z \le y\}$ for $x \le y$. Standard version (ii) is a special case of the intervals considerered in TDA's boutique definition.

(1) $x, y \in I$ and $x \leq z \leq y$ implies $z \in I$, and

(2) if $x, y \in I$ then there is a zigzag $x = u_0, u_1, \ldots, u_m = y$ of elements of I such that each u_i, u_{i+1} are comparable in P.

For $P = \mathbb{R}$, we previously used the notation $\mathbb{I}[b, d)$ for $\mathbb{F}_{[b,d)}$.

Definition 9.6. If I is an interval in a poset P, then the *interval module* $\mathbb{F}_I$ is defined by

$$(\mathbb{F}_I)_x = \begin{cases} \mathbb{F} & \text{if } x \in I, \\ 0 & \text{otherwise} \end{cases}$$

with transition maps $\mathrm{id}_{\mathbb{F}}$ whenever source and target are $\mathbb{F}$ and otherwise 0.

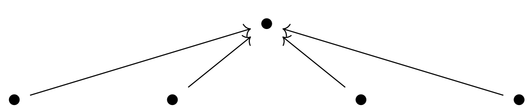
Definition 9.7. A P-module M is *interval-decomposable* when it splits as a direct sum of interval modules.

Previously, we learned that every finitely presented $\underline{n}$-persistence module is interval indecomposable. Since $\underline{n}$ contains finitely many intervals (as does any finite poset), this means that $\underline{n}$ has finite representation type with indecomposables completely classified by intervals. The same is true for finite zigzags, and it is also the case that $\underline{2} \times \underline{2}$ has finite type representations with interval-decomposition.

Example 9.8. The poset $\underline{3} \times \underline{2}$ has finite type representations but some of its indecomposables do not correspond to intervals. See Example 8.3 of Botnan–Lesnick.[54]

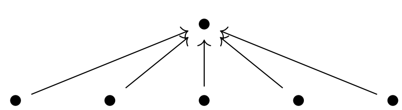
[54] Botnan, M. B. and Lesnick, M. (2023). An introduction to multiparameter persistence. arXiv:2203.14289

Example 9.9. The poset with Hasse diagram



has tame representation type. See Example 8.4 of *ibid.*

Example 9.10. The poset with Hasse diagram



has wild representation type. See Example 8.5 of *ibid.* Indeed, for $\lambda \in \mathbb{F}$, let $J(\lambda)$ denote the Jordan block

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}$$

with λ on the diagonal, 1 on the superdiagonal, and 0 elsewhere. Now for $\lambda, \mu \in \mathbb{F}$, consider the indecomposable persistence module $M^{\lambda,\mu}$ given by

$$\begin{array}{ccccc} & & \mathbb{F}^n \oplus \mathbb{F}^n & & \\ \nearrow^{A} & \nearrow^{B} & \uparrow^{C} & \nwarrow^{D} & \nwarrow^{E} \\ \mathbb{F}^n & \mathbb{F}^n & \mathbb{F}^n & \mathbb{F}^n & \mathbb{F}^n \end{array}$$

where $A = \binom{I_n}{0}$, $B = \binom{0}{I_n}$, $C = \binom{I_n}{J(\lambda)}$, $D = \binom{I_n}{J(\mu)}$, and $E = \binom{I_n}{I_n}$. This is what a *2-parameter family* of representations and makes this "5-star poset" wild type.

Here is why wild representation type is so, well, wild:

Theorem 9.11. *A finite poset P is wild type if and only if it satisfies any of the following equivalent conditions:*

(1) There exists a 2-parameter family of indecomposable P-modules.

(2) There exists an n-parameter family of indecomposable P-modules for any $n \in \mathbb{N}$.

(3) For every finite poset Q, there exists an exact functor $F\colon \mathsf{vec}_{\mathbb{F}}^{Q} \to \mathsf{vec}_{\mathbb{F}}^{P}$ such that

(i) if M is indecomposable, then so is $F(M)$;

(ii) $F(M) \cong F(N)$ if and only if $F \cong N$; and

(iii) $F(\lambda f + \mu g) = \lambda F(f) + \mu F(g)$ for all $f, g\colon M \to N$ and $\lambda, \mu \in \mathbb{F}$.

We see, thus, that a full classificaiton of indecomposable P-modules for P wild type would yield a full classificaiton of indecomposable Q-modules for all finite posets Q, which is simply too much to hope for.

We can embed the above 2-parameter family of indecomposable 5-star-modules into the grid $\underline{5} \times \underline{5}$ in the following manner:

$$\begin{array}{ccccccccc}
\mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n \\
\uparrow & & \uparrow{\scriptstyle B} & & \uparrow{\scriptstyle \mathrm{id}} & & \uparrow{\scriptstyle \mathrm{id}} & & \uparrow{\scriptstyle \mathrm{id}} \\
0 & \longrightarrow & \mathbb{F}^n & \xrightarrow{B} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n \\
\uparrow & & \uparrow & & \uparrow{\scriptstyle C} & & \uparrow{\scriptstyle \mathrm{id}} & & \uparrow{\scriptstyle \mathrm{id}} \\
0 & \longrightarrow & 0 & \longrightarrow & \mathbb{F}^n & \xrightarrow{C} & \mathbb{F}^n \oplus \mathbb{F}^n & \xrightarrow{\mathrm{id}} & \mathbb{F}^n \oplus \mathbb{F}^n \\
\uparrow & & \uparrow & & \uparrow & & \uparrow{\scriptstyle D} & & \uparrow{\scriptstyle \mathrm{id}} \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{F}^n & \xrightarrow{D} & \mathbb{F}^n \oplus \mathbb{F}^n \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow{\scriptstyle E} \\
0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{F}^n.
\end{array}$$

This settles matters. Nearly all finite posets have wild type representations, and we cannot hope to fully classify them.

Continuing on undeterred, though, we may still seek robust and informative invariants of generalized persistence modules. Here are a few common examples. For each, consider what advantages and disadvantages the invariant poses:

Definition 9.12. The *Hilbert function* of $M \in \mathsf{vec}_{\mathbb{F}}^{P}$ is

$$\begin{aligned} \mathrm{HF}_M \colon P &\longrightarrow \mathbb{N} \\ x &\longmapsto \dim_{\mathbb{F}} M_x. \end{aligned}$$

Definition 9.13. The *rank invariant* $M \in \mathsf{vec}_{\mathbb{F}}^{P}$ is the function

$$\begin{aligned} \mathrm{rk}_M \colon P^{\underline{2}} &\longrightarrow \mathbb{N} \\ (x \leq y) &\longmapsto \dim_{\mathbb{F}} \mathrm{im}(M_x \to M_y). \end{aligned}$$

For $P = \mathbb{Z}$ or $\mathbb{R}$, the rank invariant is complete. In particular, you can recover barcodes from the rank invariant.

Definition 9.14. For $M \in \mathsf{vec}_{\mathbb{F}}^{\mathbb{R}^n}$, there is a minimal free resolution

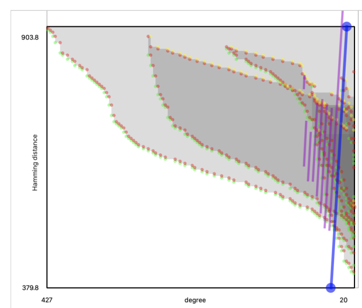
$$0 \to F_n \to F_{n-1} \to \cdots \to F_1 \to F_0 \to M \to 0.$$

For $x \in \mathbb{R}^n$ and $0 \leq i \leq n$, set

$$\beta_i^M(x) := \dim_{\mathbb{F}}(F_i)_x$$

where $(F_i)_x$ is the x-graded part of F_i. The functions $\beta_i^M \colon \mathbb{R}^n \to \mathbb{N}$ are the *multi-graded Betti numbers* of M.

Definition 9.15. Call an affine line $L \subseteq \mathbb{R}^n$ *admissible* when the product partial order on $\mathbb{R}^n$ restricts to a total order on L. The *fibered barcode* of $M \in \mathsf{vec}_{\mathbb{F}}^{P}$ is the function taking admissible lines L to the barcode of $M \circ \gamma_L$ where $\gamma_L \colon \mathbb{R} \to L$ is an isometric parametrization of L.



The software RIVET (https://rivet.readthedocs.io/) allows one to interactively explore fibered barcodes and other invariants of bipersistence modules.

There is a notion of interleaving distance on $\mathbb{R}^n$-persistence modules that is both stable and universal. Unfortunately, its computation is also NP-hard.

In the below discussion, for $M \in \mathsf{Vec}_{\mathbb{F}}^{\mathbb{R}^n}$ and $u \in P$, write $M(u)$ for the *u-shift* of M with $M(u)_x = M_{u+x}$ and structure maps shifted accordingly. Also, given $\varepsilon \geq 0$, define $\vec{\varepsilon} := (\varepsilon, \varepsilon, \dots, \varepsilon) \in \mathbb{R}^n$

Definition 9.16. Given $\varepsilon \geq 0$, we say that $M, N \in \mathsf{Vec}_{\mathbb{F}}^{\mathbb{R}^n}$ are *ε-interleaved* when there exist morphisms $f \colon M \to N(\vec{\varepsilon})$ and $g \colon N \to M(\vec{\varepsilon})$ such that $g(\vec{\varepsilon}) \circ f = \phi_M^{2\vec{\varepsilon}}$ and $f(\vec{\varepsilon}) \circ g = \phi_N^{2\vec{\varepsilon}}$.

The *interleaving distance* between M, N, denoted $d_I(M, N)$ is the infimum of $\varepsilon \geq 0$ such that M, N are ε-interleaved.

There is a precise notion of stability for (pseudo-) distances on persistence modules. Not only is d_I stable, but when the base field is prime ($\mathbb{Q}$

or $\mathbb{F}_p$), it is also known that d_I is maximal among stable distances; see Theorem 6.7 of Botnan–Lesnick.

But alas, the following theorem of Bjerkevik–Botnan–Kerber[55] tells us that interleaving distance is quite hard to compute.

[55] Bjerkevik, H. B., Botnan, M. B., and Kerber, M. (2020). Computing the interleaving distance is np-hard. *Foundations of Computational Mathematics*, 20(5):1237–1271

Theorem 9.17. *For $\mathbb{F}$ a finite field, approximating the interleaving distance on bipersistence modules within a factor of 3 is NP-hard.*

Despite these challenges, research persists in further understanding the structure and application of generalized persistence modules.

9.1 Notes

Everything in this lecture is a hasty and condensed permutation of material from Botnan–Lesnick.[56]

[56] Botnan, M. B. and Lesnick, M. (2023). An introduction to multiparameter persistence. arXiv:2203.14289

9.2 Exercises

(1) For an interval I in a poset P, prove that the interval module $\mathbb{F}_I$ is indecomposable.

(2) What shape do intervals in $\mathbb{R}^2$ have?

(3) Download RIVET and use it to explore some bipersistence modules.

References

Bauer, U. and Lesnick, M. (2014). Induced matchings of barcodes and the algebraic stability of persistence. In *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, page 355–364, New York, NY, USA. Association for Computing Machinery.

Bjerkevik, H. B., Botnan, M. B., and Kerber, M. (2020). Computing the interleaving distance is np-hard. *Foundations of Computational Mathematics*, 20(5):1237–1271.

Botnan, M. B. and Lesnick, M. (2023). An introduction to multiparameter persistence. arXiv:2203.14289.

Bradley, T.-D., Bryson, T., and Terilla, J. (2020). *Topology—a categorical approach.* MIT Press, Cambridge, MA.

Carlsson, G. (2009). Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308.

Carlsson, G., de Silva, V., and Morozov, D. (2009). Zigzag persistent homology and real-valued functions. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, SCG '09, page 247–256, New York, NY, USA. Association for Computing Machinery.

Carlsson, G. and Vejdemo-Johansson, M. (2022). *Topological data analysis with applications.* Cambridge University Press, Cambridge.

Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L. J., and Oudot, S. Y. (2009). Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, SCG '09, page 237–246, New York, NY, USA. Association for Computing Machinery.

Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2009). Extending persistence using Poincaré and Lefschetz duality. *Found. Comput. Math.*, 9(1):79–103.

Dey, T. K. and Wang, Y. (2022). *Computational topology for data analysis.* Cambridge University Press, Cambridge.

Gabriel, P. (1972). Unzerlegbare Darstellungen. I. *Manuscripta Math.*, 6:71–103; correction, ibid. **6 (1972), 309**,.

Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc. (N.S.)*, 45(1):61–75.

Hatcher, A. (2002). *Algebraic topology.* Cambridge University Press, Cambridge.

Morgan, C. L. (1974). Embedding matric spaces in Euclidean space. *J. Geom.*, 5:101–107.

Munkres, J. R. (2000). *Topology*. Prentice Hall, Inc., Upper Saddle River, NJ, second edition.

Myers, A., Muñoz, D., Khasawneh, F. A., and Munch, E. (2023). Temporal network analysis using zigzag persistence. *EPJ Data Science*, 12(1):6.

Oudot, S. Y. (2015). *Persistence theory: from quiver representations to data analysis*, volume 209 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI.

Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., and Markram, H. (2017). Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in Computational Neuroscience*, 11.

Tralie, C., Saul, N., and Bar-On, R. (2018). Ripser.py: A lean persistent homology library for python. *The Journal of Open Source Software*, 3(29):925.

Virk, Ž. (2022). Introduction to persistent homology. `https://zalozba.fri.uni-lj.si/virk2022.pdf`. Accessed on 19 March 2024.

Zomorodian, A. (2010). The tidy set: a minimal simplicial set for computing homology of clique complexes. In *Proceedings of the Twenty-Sixth Annual Symposium on Computational Geometry*, SoCG '10, page 257–266, New York, NY, USA. Association for Computing Machinery.