26 PERSISTENT HOMOLOGY

Herbert Edelsbrunner and Dmitriy Morozov

INTRODUCTION

Persistent homology was introduced in [ELZ00] and quickly developed into the most influential method in computational topology. There were independent developments of this idea preceding this paper, and we mention the little known paper by Marston Morse [Mor40], the spectral sequences introduced by Jean Leray in [Ler46], the notion of prominence in mountaineering [Mun53], the size function introduced by Frosini [Fro90], and the study of fractal sets by Robins [Rob99]. Perhaps the fast algorithm described in [ELZ00] triggered the explosion of interest we currently observe because its availability as software facilitates the application to a broad collection of problems and datasets.

From the mathematical perspective, persistent homology is part of Morse theory [Mil63]. Functions come naturally, which explains the affinity to applications and to data. Persistent homology quantifies the critical points, which we illustrate with the resolution of a basic question in mountaineering: *what is a mountain?* We can identify a mountain with its peak, but clearly not every local maximum on Earth qualifies as a mountain. The summit of Mt. Everest is prominent but its South summit is not, despite being higher than any other mountain in the world. To identify peaks that are prominent, climbers measure how far they have to descend before they can climb to an even higher peak. This measurement, called the *topographic prominence*, assigns a significance to every local maximum of the elevation function on planet Earth, and peaks with prominence above 500 meters qualify as separate mountains. Persistence addresses the multi-scale aspects of natural phenomena by generalizing topographic prominence to more general functions and to higher-dimensional features and holes. It also applies to abstract and high-dimensional data, such as sound, documents, DNA sequences, and languages by equipping the data with a possibly discrete metric.

Section 26.1 explains persistence algebraically, as an extension of the classical notion of homology. Section 26.2 approaches persistence geometrically, focusing on the functions and complexes whose persistence we compute. Section 26.3 is analytic in nature, studying the space of persistence diagrams and the stability of persistence. Section 26.4 focuses on algorithms, explaining the connection to matrix reduction and how the computation can be made fast in important cases. Section 26.5 returns to the algebraic foundations, describing zigzag persistence as a unifying as well as generalizing concept. Section 26.6 illustrates the application of persistent homology by considering two mathematical questions.

26.1 ALGEBRA

We recall basic concepts from algebraic topology to explain homology as well as its recent extension, persistent homology. The first four of these concepts can also

be found in Chapter 18, and we repeat the definitions to be consistent with the notation adopted in this chapter.

GLOSSARY

Abstract simplex: a non-empty finite subset of a universal vertex set, $\alpha \subseteq \mathbf{U}$. Its *dimension* is one less than its cardinality, $\dim \alpha = \operatorname{card} \alpha - 1$. Setting $p = \dim \alpha$, we call α an *(abstract) p-simplex*. A *face* is a non-empty subset of α. An *ordering* of α is a sequence of its $p+1$ vertices, denoted $\alpha = [v_0, v_1, \ldots, v_p]$. An *orientation* is a class of orderings that differ by an even number of transpositions. Every simplex has two orientations, except for a 0-simplex, which has only one.

Abstract simplicial complex: a finite collection of abstract simplices, A, that is closed under the face relation: $\beta \in A$ and $\alpha \subseteq \beta$ implies $\alpha \in A$. Its *dimension* is the maximum dimension of any of its simplices. A *subcomplex* is an abstract simplicial complex $B \subseteq A$. The *p-skeleton* is the largest subcomplex whose dimension is p. For example, the 1-skeleton of A is a graph.

Geometric simplex: the convex hull of a non-empty and affinely independent set of points in $\mathbb{R}^n$. Assuming $\mathbf{U}$ is a set of points in $\mathbb{R}^n$, we define *dimension*, *(geometric) p-simplex*, *face*, *ordering*, and *orientation* as in the abstract case.

Geometric simplicial complex: a finite collection of geometric simplices, K, that is closed under the face relation such that $\sigma, \tau \in K$ implies $\sigma \cap \tau$ is either empty or a face of both simplices. We define *dimension* and *subcomplex* as in the abstract case. The *underlying space* of K, denoted $|K|$, is the set of points in $\mathbb{R}^n$ contained in simplices of K together with the Euclidean topology inherited from $\mathbb{R}^n$. While K is a combinatorial object, $|K|$ is a topological space.

Chain: a formal sum of ordered simplices of the same dimension, $c = \sum_i a_i \sigma_i$. The coefficients a_i are elements of an abelian group. If all simplices have dimension p, then we call c a *p-chain*. We *add* p-chains like polynomials. The set of p-chains of a simplicial complex, K, is the *p-chain group* of K, denoted $\mathsf{C}_p = \mathsf{C}_p(K)$. If the coefficients are elements of a field, then C_p is a vector space.

Boundary map: the linear map $\partial_p \colon \mathsf{C}_p \to \mathsf{C}_{p-1}$ defined by mapping an ordered p-simplex to the alternating sum of its $(p-1)$-dimensional faces. For $\sigma = [v_0, v_1, \ldots, v_p]$, we get $\partial_p \sigma = \sum_{i=0}^{p} (-1)^i \hat{\sigma}_i$, in which $\hat{\sigma}_i$ is σ without vertex v_i. Note that $\partial_0 [v_i] = 0$ for every i. It is not difficult to verify that $\partial_{p-1} \circ \partial_p = 0$ for all p.

Chain complex: an infinite sequence of chain groups connected by boundary maps, $\ldots \stackrel{\partial_{p+1}}{\to} \mathsf{C}_p \stackrel{\partial_p}{\to} \mathsf{C}_{p-1} \stackrel{\partial_{p-1}}{\to} \ldots$. The *$p$-cycle group* is the kernel of the p-th boundary map, $\mathsf{Z}_p = \ker \partial_p$. It is the subgroup of p-chains with zero boundary. The *p-boundary group* is the image of the $(p+1)$-st boundary map, $\mathsf{B}_p = \operatorname{img} \partial_{p+1}$. It is the subgroup of p-cycles that are the boundary of a $(p+1)$-chain.

Homology group: the quotient of the cycle group over the boundary group. Call two cycles *homologous* if they differ only by a boundary. A *homology class* is the maximal collection of homologous cycles in a complex. Fixing p, the set of homology classes of p-cycles is the *p-th homology group*, $\mathsf{H}_p = \mathsf{Z}_p / \mathsf{B}_p$. If the coefficients are elements in a field, then H_p is a vector space.

Relative homology: the generalization of homology to pairs $K_0 \subseteq K$. Here we distinguish two chains only if they differ within $K \setminus K_0$: $\mathsf{C}_p(K, K_0) = \mathsf{C}_p(K)/\mathsf{C}_p(K_0)$ for all p. The boundary maps $\partial_p \colon \mathsf{C}_p(K, K_0) \to \mathsf{C}_{p-1}(K, K_0)$ are well defined, and we set $\mathsf{Z}_p(K, K_0) = \ker \partial_p$, $\mathsf{B}_p(K, K_0) = \operatorname{img} \partial_{p+1}$, and $\mathsf{H}_p(K, K_0) = \mathsf{Z}_p(K, K_0)/\mathsf{B}_p(K, K_0)$.

Exact sequence: a sequence of abelian groups connected by homomorphisms, $\ldots \to V_i \to V_{i-1} \to \ldots$, in which the image of each map is equal to the kernel of the next map. In the case of vector spaces, the homomorphisms are linear maps. An exact sequence of five abelian groups that begins and ends with zero, $0 \to V_3 \to V_2 \to V_1 \to 0$, is called a *short exact sequence*.

Filtered simplicial complex: a simplicial complex, K, together with a function $f \colon K \to \mathbb{R}$ such that $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ. The *sublevel set* at a value $r \in \mathbb{R}$ is $f^{-1}(-\infty, r]$, which is a subcomplex of K. Letting $r_0 < r_1 < \ldots < r_m$ be the values of the simplices and writing $K_i = f^{-1}(-\infty, r_i]$, we call $K_0 \subseteq K_1 \subseteq \ldots \subseteq K_m$ the *sublevel set filtration* of f.

Persistence module: a sequence of vector spaces connected by linear maps, $\mathbf{f}_r^s \colon U_r \to U_s$, for every pair of values $r \leq s$, such that $\mathbf{f}_r^r = \mathrm{id}$ and $\mathbf{f}_r^t = \mathbf{f}_s^t \circ \mathbf{f}_r^s$ for all $r \leq s \leq t$. We denote such a module by $\mathcal{U} = (U_r, \mathbf{f}_r^s)$. For example, the persistence module of the above filtered simplicial complex is $\mathcal{H}(K) = (\mathsf{H}(K_i), \mathbf{f}_i^j)$ in which $\mathsf{H}(K_i)$ is the direct sum of the p-th homology groups of K_i, over all p, and $\mathbf{f}_i^j \colon \mathsf{H}(K_i) \to \mathsf{H}(K_j)$ is induced by the inclusion $K_i \subseteq K_j$.

HOMOLOGY

While we restricted the above definitions to simplicial complexes, homology groups can be defined in much greater generality. For example, we may use *singular simplices* (continuous maps of simplices) to construct the *singular homology groups* of a topological space. An example of such a space is $\mathbb{X} = |K|$, and importantly, the singular homology groups of $\mathbb{X}$ are isomorphic to the simplicial homology groups of K. We refer to the axiomatization of Eilenberg and Steenrod [ES52] for the main tool to reach a unified view of the many different constructions of homology in the literature.

Homology groups, Betti numbers, and Euler characteristic. Assuming coefficients in a field, $\mathbb{F}$, each homology group is a vector space, $\mathbb{F}^\beta$. Here, β is a non-negative integer, namely the *dimension* of the vector space or the *rank* of the group. Writing $\mathsf{H}_p(\mathbb{X}) = \mathbb{F}^{\beta_p}$, we call $\beta_p = \beta_p(\mathbb{X})$ the *p-th Betti number* of $\mathbb{X}$. Historically, Betti numbers preceded homology groups. In turn, the *Euler characteristic* of $\mathbb{X}$, which we can define as $\chi(\mathbb{X}) = \sum_{p \geq 0} (-1)^p \beta_p(\mathbb{X})$, preceded the Betti numbers. For example, it was known already to Leonhard Euler that χ of the boundary of any convex polytope in $\mathbb{R}^3$—then defined as the alternating sum of face numbers—is 2.

Relative and local homology. Instead of defining homology for a space that is partially open, it is more convenient to define it for a pair of closed spaces, $K_0 \subseteq K$. One motivation for introducing this concept is the desire to define homology locally, at a point $x \in \mathbb{X}$. We may capture the homology within an open neighborhood N of x by constructing $\mathsf{H}_p(\mathbb{X}, \mathbb{X} \setminus N)$. Shrinking the neighborhood towards x, the *local homology* of $\mathbb{X}$ at x is the limit of the relative homology.

Snake Lemma. As one of the main achievements of algebraic topology, this lemma is a recipe for the construction of long exact sequences. Specifically, if $0 \to \mathsf{C} \to \mathsf{D} \to \mathsf{E} \to 0$ is a short exact sequence of chain complexes, then

$$\ldots \to \mathsf{H}_{p+1}(\mathsf{E}) \to \mathsf{H}_p(\mathsf{C}) \to \mathsf{H}_p(\mathsf{D}) \to \mathsf{H}_p(\mathsf{E}) \to \mathsf{H}_{p-1}(\mathsf{C}) \to \ldots$$

is a long exact sequence of homology groups. This is useful because exact sequences provide a powerful language to compactly encode relationships between homology groups. We illustrate this with two examples.

Long exact sequence of a pair. Letting $K_0 \subseteq K$ be two simplicial complexes, we consider $0 \to \mathsf{C}(K_0) \to \mathsf{C}(K) \to \mathsf{C}(K, K_0) \to 0$, in which the middle two maps are inclusions. It is not difficult to see that the kernel of every map is the image of the preceding map. By the Snake Lemma,

$$\ldots \to \mathsf{H}_{p+1}(K, K_0) \to \mathsf{H}_p(K_0) \to \mathsf{H}_p(K) \to \mathsf{H}_p(K, K_0) \to \mathsf{H}_{p-1}(K_0) \to \ldots$$

is a long exact sequence.

Mayer–Vietoris long exact sequence. Letting $K = A \cup B$ be three simplicial complexes, we consider $0 \to \mathsf{C}(A \cap B) \to \mathsf{C}(A) \oplus \mathsf{C}(B) \to \mathsf{C}(A \cup B) \to 0$, in which the second map is the direct sum of two inclusions and the third map is the sum of two inclusions. By the Snake Lemma,

$$\ldots \to \mathsf{H}_{p+1}(K) \to \mathsf{H}_p(A \cap B) \to \mathsf{H}_p(A) \oplus \mathsf{H}_p(B) \to \mathsf{H}_p(K) \to \mathsf{H}_{p-1}(A \cap B) \to \ldots$$

is a long exact sequence. It has many applications within algebraic topology, including the construction of the zigzag pyramid as discussed in Section 26.5.

PERSISTENCE

While homology defines holes and counts them, persistent homology also measures them. This additional feature opened a floodgate of applications to the sciences and beyond and has established computational topology as a viable new field of mathematical inquiry; see e.g. Edelsbrunner and Harer [EH10]. The idea itself has several independent roots in the mathematical literature, the earliest of which is a little known paper by Marston Morse [Mor40]. The description in [ELZ00] together with the algorithms for computing persistence diagrams have initiated the current interest in the subject.

Persistence module of a function. Let $\mathbb{X}$ be a topological space and $f\colon \mathbb{X} \to \mathbb{R}$ a real-valued function. Writing $\mathbb{X}_r = f^{-1}(-\infty, r]$, the sublevel sets form a filtration of the topological space, $\mathbb{X}_r \subseteq \mathbb{X}_s$ for all $r \leq s$. Taking the homology of every sublevel set, we get a persistence module for each dimension, $\mathcal{H}_p(f) = (\mathsf{H}_p(\mathbb{X}_r), \mathbf{p}_r^s)$, in which the maps $\mathbf{p}_r^s$ are induced by the inclusions $\mathbb{X}_r \subseteq \mathbb{X}_s$. We often simplify the notation by taking direct sums of homology groups and maps, $\mathsf{H}(\mathbb{X}_r) = \bigoplus_{p \geq 0} \mathsf{H}_p(\mathbb{X}_r)$ and $\mathbf{f}_r^s = \bigoplus_{p \geq 0} \mathbf{p}_r^s$, and consider the persistence module $\mathcal{H}(f) = (\mathsf{H}(\mathbb{X}_r), \mathbf{f}_r^s)$ that simultaneously captures all dimensions.

Decomposition into summands. Under mild assumptions, a persistence module decomposes uniquely into elementary pieces. To describe this, we call a persistence module, $\mathcal{U} = (U_r, \mathbf{f}_r^s)$, *q-tame* ("$q$" for quadrant), if the rank of $\mathbf{f}_r^s$ is finite for all $r < s$. Given an interval $[b, d)$, the corresponding *interval module* is $\mathcal{I} = (I_r, \mathbf{i}_r^s)$, with vector spaces $I_r = \mathbb{F}$ whenever $b \leq r < d$ and $I_r = 0$ otherwise, as well as maps

$\mathbf{i}_r^s = \mathrm{id}$ whenever $b \le r \le s < d$ and $\mathbf{i}_r^s = 0$, otherwise. Every q-tame persistence module decomposes uniquely as a direct sum of interval modules, $\mathcal{U} = \bigoplus_j \mathcal{I}_j$. The interval modules are *indecomposable summands*. There is an alternative way to describe this decomposition: select elements u_i^r in each vector space U_r, such that the nonzero elements form a basis of U_r, and the maps diagonalize with respect to these bases, i.e., $\mathbf{f}_r^s(u_i^r) = u_i^s$ whenever $u_i^r \neq 0$. Note that the u_i^r are not unique, but the intervals they define are. If the decomposition of module $\mathcal{U}$ contains a module for interval $[b, d)$, we say that a class is born at U_b and dies entering U_d.

Persistence diagram. The information contained in a persistence module has intuitive combinatorial representations. The *persistence diagram* associated with $\mathcal{U} = \bigoplus_j \mathcal{I}_j$, denoted $\mathrm{Dgm}(\mathcal{U})$, is the multi-set of points (b_j, d_j) with $\mathcal{I}_j$ defined by $[b_j, d_j)$. For technical reasons that will become clear when we discuss the stability of persistence diagrams, we usually add infinitely many copies of the points (x, x) on the diagonal to the persistence diagram. When $\mathcal{U} = \mathcal{H}(f)$ is defined by the sublevel sets of a function, we abbreviate the notation to $\mathrm{Dgm}(f) = \mathrm{Dgm}(\mathcal{H}(f))$, or $\mathrm{Dgm}_p(f) = \mathrm{Dgm}(\mathcal{H}_p(f))$ if we wish to restrict the information to a single dimension p. Sometimes $\mathcal{U}$ is represented by the corresponding multi-set of intervals, $[b_j, d_j)$, which is referred to as the *barcode* of $\mathcal{U}$.

Equivalence of persistence modules. Two persistence modules, $\mathcal{U} = (U_r, \mathbf{f}_r^s)$ and $\mathcal{V} = (V_r, \mathbf{g}_r^s)$, are *isomorphic* if the vector spaces are pairwise isomorphic, $U_r \simeq V_r$ for all r, and these isomorphisms commute with the maps $\mathbf{f}_r^s$ and $\mathbf{g}_r^s$ in the modules. This situation can be graphically represented by the squares

$$\begin{array}{ccc} U_r & \xrightarrow{\mathbf{f}_r^s} & U_s \\ \simeq\downarrow & & \downarrow\simeq \\ V_r & \xrightarrow{\mathbf{g}_r^s} & V_s, \end{array} \tag{26.1.1}$$

which are required to commute for all $r \le s$.

THEOREM 26.1.1 *Persistence Equivalence Theorem*

Isomorphic persistence modules imply identical persistence diagrams, $\mathrm{Dgm}(\mathcal{U}) = \mathrm{Dgm}(\mathcal{V})$.

Extended persistence. Some applications of persistent homology require an extension of the filtration defining the module; see e.g. [AEHW06, EP16]. We explain this for the filtration of sublevel sets of a function $f\colon \mathbb{X} \to \mathbb{R}$ for which the extension consists of the pairs $(\mathbb{X}, \mathbb{X}^r)$, with $\mathbb{X}^r = f^{-1}[r, \infty)$ the *superlevel set* of f at r. Applying homology, we get vector spaces of the form $\mathsf{H}(\mathbb{X}, \mathbb{X}^r)$, and a persistence module with maps $\mathbf{f}_s^r\colon \mathsf{H}(\mathbb{X}, \mathbb{X}^s) \to \mathsf{H}(\mathbb{X}, \mathbb{X}^r)$ for all $r < s$. Since $\mathsf{H}(\mathbb{X}) = \mathsf{H}(\mathbb{X}, \emptyset)$, we can append the new persistence module to $\mathcal{H}(f)$ and get the *extended persistence module*, $\mathcal{H}^{\mathrm{ext}}(f)$, and the *extended persistence diagram*, $\mathrm{Dgm}(\mathcal{H}^{\mathrm{ext}}(f))$.

26.2 GEOMETRY

In many applications of persistent homology, the essential geometric information is encoded in the filtration. While this is not necessary, it is a convenient vehicle for measuring geometry with topology.

GLOSSARY

Star of a simplex: the set of simplices in a simplicial complex that contain the given simplex, $\operatorname{St}\sigma = \{\tau \in K \mid \sigma \subseteq \tau\}$. The *closed star* also contains the faces of the simplices that contain the given simplex, $\overline{\operatorname{St}}\,\sigma = \{\upsilon \in K \mid \upsilon \subseteq \tau \in \operatorname{St}\sigma\}$. The *link* contains all simplices in the closed star that avoid the given simplex, $\operatorname{Lk}\sigma = \{\upsilon \in \overline{\operatorname{St}}\,\sigma \mid \upsilon \cap \sigma = \emptyset\}$.

Piecewise-linear function: a real-valued function $f\colon |K| \to \mathbb{R}$ that is specified on the vertices of K and is interpolated linearly on the simplices. To stress the point that the complex is not refined before interpolation, f is sometimes referred to as a *simplexwise-linear function*.

Lower star of a vertex: the subset of the star in which the vertices of every simplex have function values smaller than or equal to the value at the given vertex, $\operatorname{St}_- u = \{\tau \in \operatorname{St} u \mid f(v) \leq f(u) \text{ for all } v \in \tau\}$. Similarly, the *lower link* of the vertex is $\operatorname{Lk}_- u = \{\tau \in \operatorname{Lk} u \mid f(v) \leq f(u) \text{ for all } v \in \tau\}$.

Barycentric subdivision: an abstract simplicial complex, $\operatorname{Sd} K$, which has the simplices $\sigma_i \in K$ as its vertices, and which has a simplex $\tau = [\sigma_0, \sigma_1, \ldots, \sigma_p]$ iff its vertices form a chain of faces in K, i.e., $\sigma_0 \subseteq \sigma_1 \subseteq \ldots \subseteq \sigma_p$.

Homotopy equivalence: an equivalence relation between topological spaces. To define it, we call a continuous map $h\colon \mathbb{X} \times [0,1] \to \mathbb{Y}$ a *homotopy* between the maps $a, b\colon \mathbb{X} \to \mathbb{Y}$ that satisfy $a(x) = h(x,0)$ and $b(x) = h(x,1)$ for all $x \in \mathbb{X}$. Now $\mathbb{X}$ and $\mathbb{Y}$ are *homotopy equivalent*, or they have the same *homotopy type*, denoted $\mathbb{X} \simeq \mathbb{Y}$, if there are maps $f\colon \mathbb{X} \to \mathbb{Y}$ and $g\colon \mathbb{Y} \to \mathbb{X}$ such that there is a homotopy between $g \circ f$ and $\mathrm{id}_{\mathbb{X}}$ as well as between $f \circ g$ and $\mathrm{id}_{\mathbb{Y}}$. A *contractible* set is homotopy equivalent to a point.

Deformation retraction from $\mathbb{X}$ to $\mathbb{Y}$: is a continuous map $D\colon \mathbb{X} \times [0,1] \to \mathbb{X}$ such that $D(x,0) = x$, $D(x,1) \in \mathbb{Y}$, and $D(y,t) = y$ for all $x \in \mathbb{X}$, $y \in \mathbb{Y}$, and $t \in [0,1]$. If such a D exists, then $\mathbb{Y}$ is a *deformation retract* of $\mathbb{X}$. Note that $\mathbb{Y}$ is then homotopy equivalent to $\mathbb{X}$. Indeed, setting $f\colon \mathbb{X} \to \mathbb{Y}$ and $g\colon \mathbb{Y} \to \mathbb{X}$ defined by $f(x) = D(x,1)$ and $g(y) = y$, we get D as a homotopy between $g \circ f$ and $\mathrm{id}_{\mathbb{X}}$, and we get the identity as a homotopy between $f \circ g$ and $\mathrm{id}_{\mathbb{Y}}$.

Hausdorff distance: an extension of a distance between points to a distance between point sets,

$$d_H(X,Y) = \max\{\sup_{x\in X}\inf_{y\in Y}\|x-y\|, \sup_{y\in Y}\inf_{x\in X}\|x-y\|\}. \tag{26.2.1}$$

Writing $X^r = X + B(0,r)$ for the Minkowski sum with the closed ball of radius $r \geq 0$, $d_H(X,Y)$ is the infimum radius r such that $X \subseteq Y^r$ and $Y \subseteq X^r$.

Nerve: a simplicial complex associated to a collection of sets. The sets are the vertices of the complex, and a simplex belongs to the complex iff its vertices have a non-empty intersection, $\operatorname{Nrv} S = \{\alpha \subseteq S \mid \bigcap_{A\in\alpha} A \neq \emptyset\}$. If the sets in S are convex, then the nerve is homotopy equivalent to the union of these sets. This results is known as the Nerve Theorem [Bor48, Ler46], and it generalizes to the case in which all non-empty common intersections are contractible.

Čech complex: the nerve of the ball neighborhoods of a set of points $X \subseteq \mathbb{R}^n$. Writing $B(x,r)$ for the closed ball of radius $r \geq 0$ centered at x, the *Čech complex* of X for radius r is $\check{\mathrm{C}}\mathrm{ech}_r(X) = \operatorname{Nrv}\{B(x,r) \mid x \in X\}$.

Alpha complex: the nerve of the clipped ball neighborhoods of a set of points $X \subseteq \mathbb{R}^n$. Here we clip $B(x,r)$ with the *Voronoi domain* of x, which consists of all points $a \in \mathbb{R}^n$ that are at least as close to x as to any other point in X. Writing $V(x) = \{a \in \mathbb{R}^n \mid \|a-x\| \leq \|a-y\| \text{ for all } y \in X\}$, the *alpha complex* of X for radius r is $\text{Alpha}_r(X) = \text{Nrv}\{B(x,r) \cap V(x) \mid x \in X\}$. To stress the connection to the dual of the Voronoi domains, $\text{Alpha}_r(X)$ is sometimes referred to as the *Delaunay complex* of X for radius r. It is a subcomplex of the *Delaunay triangulation*, $\text{Alpha}_\infty(X)$, which is the dual of the Voronoi domains.

Vietoris–Rips complex: the largest simplicial complex whose 1-skeleton is also the 1-skeleton of the Čech complex. Given a graph $G = (V,E)$, the *clique complex* contains all cliques of G, namely all simplices $\alpha \subseteq V$ for which E contains all edges of α. With this notation, the *Vietoris–Rips complex* of X for radius r is the clique complex of the 1-skeleton of the Čech complex of X and r, $\text{Rips}_r(X) = \{\alpha \subseteq X \mid \|u-v\| \leq 2r \text{ for all } u,v \in \alpha\}$.

PIECEWISE-LINEAR FUNCTIONS

Functions generally do not have finite descriptions, so we have to work with approximations. For example, we may fix the values at a finite set of points and make up the information in between by interpolation. Due to its simplicity, the description in which the interpolation is piecewise linear is popular in applications.

Sublevel sets and lower stars. Given a piecewise-linear function $f\colon |K| \to \mathbb{R}$ and a value $r \in \mathbb{R}$, we recall that the corresponding sublevel set is $|K|_r = f^{-1}(-\infty, r]$. It is not necessarily the underlying space of a subcomplex of K, but it is homotopy equivalent to one. Specifically, let $K_r \subseteq K$ contain all simplices whose vertices have values smaller than or equal to r. Equivalently, K_r is the union of the lower stars of all vertices with function values at most r. Clearly $|K_r| \subseteq |K|_r$, and it is not difficult to see that there is a deformation retraction from $|K|_r$ to $|K_r|$, implying that the two are homotopy equivalent.

Lower star filtration. Starting with the piecewise-linear function $f\colon |K| \to \mathbb{R}$, we construct a function $g\colon K \to \mathbb{R}$ that maps every simplex to the maximum function value of its vertices. Observe that K together with g is a filtered complex, and that $K_r = g^{-1}(-\infty, r]$ for every r. The resulting nested sequence of complexes is usually referred to as the *lower star filtration* of f or of g.

Hierarchy of equivalences between spaces. Recall that a *homeomorphism* $h\colon \mathbb{X} \to \mathbb{Y}$ is a continuous bijection whose inverse is continuous. If such an h exists, then $\mathbb{X}$ and $\mathbb{Y}$ are *topologically equivalent* or they have the same *topology type*. Topological equivalence is stronger than homotopy equivalence, which in turn is stronger than homology. In other words, if $\mathbb{X}$ and $\mathbb{Y}$ are homeomorphic, then they are homotopy equivalent, and if they are homotopy equivalent, then they have isomorphic homology groups. These implications cannot be reversed.

Equivalent persistence modules. Passing to homology, we get two persistence modules, $\mathcal{F} = (\mathsf{H}(|K|_r), \mathbf{f}_r^s)$ and $\mathcal{G} = (\mathsf{H}(K_r), \mathbf{g}_r^s)$. For any two values $r \leq s$, we get

a square,

$$\begin{array}{ccc} \mathsf{H}(|K|_a) & \longrightarrow & \mathsf{H}(|K|_b) \\ \uparrow & & \uparrow \\ \mathsf{H}(K_a) & \longrightarrow & \mathsf{H}(K_b), \end{array} \tag{26.2.2}$$

in which all four maps are induced by inclusion. It follows that the square commutes. Since $|K_r| \simeq |K|_r$, for all r, the vertical maps are isomorphisms. As explained earlier, this implies that the two modules have the same persistence diagram, $\mathrm{Dgm}(f) = \mathrm{Dgm}(g)$.

Filtration as a piecewise-linear function. Given a filtered simplicial complex, K with $f\colon K \to \mathbb{R}$, we define a piecewise-linear function on its barycentric subdivision, $g\colon |\operatorname{Sd} K| \to \mathbb{R}$, by setting the value at the vertices to $g(\sigma_i) = f(\sigma_i)$, for all $\sigma_i \in K$. The two persistence modules, $\mathcal{H}(f)$ and $\mathcal{H}(g)$, are isomorphic because the sublevel sets of g deformation retract onto the sublevel sets of f.

DISTANCE FUNCTIONS

Besides piecewise-linear functions, distance functions on the ambient space of given data are most popular in applications of persistent homology. Given $X \subseteq \mathbb{R}^n$, it is the function $d_X\colon \mathbb{R}^n \to \mathbb{R}$ defined by mapping every $a \in \mathbb{R}^n$ to the distance to X, $d_X(a) = \inf_{x\in X} \|a - x\|$. The sublevel set of d_X for $r \geq 0$ is the union of balls of radius r centered at the points of X.

Distance functions and Hausdorff distance. If two sets $X, Y \subseteq \mathbb{R}^n$ are close in Hausdorff distance, then their distance functions are close:

$$\sup_{a\in\mathbb{R}^n} |d_X(a) - d_Y(a)| \leq d_H(X, Y). \tag{26.2.3}$$

This motivates us to recover properties of a shape that are stable under Hausdorff perturbations from point samples of the shape. Indeed, we will see in the next section how the homology of a shape can be recovered from the homology of a point sample. Meanwhile, we study three filtered complexes that are commonly used to represent the sublevel sets of a distance function.

Persistence and nerves. Let S be a collection of convex sets and recall the Nerve Theorem, which asserts that $\operatorname{Nrv} S \simeq \bigcup S$.

THEOREM 26.2.1 *Persistence Nerve Theorem*

Let T be a second collection of convex sets and $b\colon S \to T$ a bijection such that $s \subseteq b(s)$ for every set $s \in S$. Then the inclusion between the two nerves induced by b and the inclusion between the two unions commute with the homotopy equivalences,

$$\begin{array}{ccc} \bigcup S & \longrightarrow & \bigcup T \\ \uparrow & & \uparrow \\ \operatorname{Nrv} S & \longrightarrow & \operatorname{Nrv} T. \end{array} \tag{26.2.4}$$

Čech and alpha complexes are homotopy equivalent. Indeed, both complexes are homotopy equivalent to the corresponding union of balls. Moreover, the

filtrations of Čech complexes, of alpha complexes, and of sublevel sets of the distance function have the same persistence diagram. For the Čech complexes and the sublevel sets, this statement follows from the Persistence Nerve Theorem 26.2.1 and the Persistence Equivalence Theorem 26.1.1. For the alpha complexes, we observe in addition that the union of balls and the union of clipped balls are the same.

Čech and Vietoris–Rips complexes are interleaved. By construction, the Vietoris–Rips complex of a set $X \subseteq \mathbb{R}^n$ for radius $r \geq 0$ contains the Čech complex of X for r. However, the holes left by $p+1$ balls that have pairwise non-empty intersections cannot be large. Indeed, if we grow each ball to twice its initial radius, each ball contains all $p+1$ centers. It follows that the $p+1$ balls of radius $2r$ have a non-empty common intersection. In summary, $\check{\mathrm{C}}\mathrm{ech}_r(X) \subseteq \mathrm{Rips}_r(X) \subseteq \check{\mathrm{C}}\mathrm{ech}_{2r}(X)$. For the special case of Euclidean distance, growing the balls to radius $\sqrt{2}r$ suffices to guarantee a non-empty common intersection, giving $\mathrm{Rips}_r(X) \subseteq \check{\mathrm{C}}\mathrm{ech}_{\sqrt{2}r}(X)$.

26.3 ANALYSIS

The step from functions to homology is a temporary excursion to algebra, and we are right back to analysis when we reason about the persistence diagrams that summarize the features of the function.

GLOSSARY

Interleaving distance: a notion of distance between persistence modules. Specifically, $\mathcal{U} = (U_r, \mathbf{f}_r^s)$ and $\mathcal{V} = (V_r, \mathbf{g}_r^s)$ are *ε-interleaved* if there are maps $\varphi_r \colon U_r \to V_{r+\varepsilon}$ and $\psi_r \colon V_r \to U_{r+\varepsilon}$, for all $r \in \mathbb{R}$, that commute with the maps inside the modules. We get the *interleaving distance* between $\mathcal{U}$ and $\mathcal{V}$ by taking the infimum of the $\varepsilon \geq 0$ for which the modules are ε-interleaved.

Bottleneck distance: a notion of distance between persistence diagrams. Finding a bijection, $\gamma \colon \mathrm{Dgm}(\mathcal{U}) \to \mathrm{Dgm}(\mathcal{V})$, we quantify it by taking the maximum L_∞-distance of any two corresponding points. The *bottleneck distance* between the persistence diagrams of $\mathcal{U}$ and $\mathcal{V}$ is the infimum over all bijections:

$$\mathrm{W}_\infty(\mathrm{Dgm}(\mathcal{U}), \mathrm{Dgm}(\mathcal{V})) = \inf_\gamma \sup_{x \in \mathrm{Dgm}(\mathcal{U})} \|x - \gamma(x)\|_\infty. \tag{26.3.1}$$

Wasserstein distances: a 1-parameter family of distances between persistence diagrams. Fixing a real number $q \geq 1$, we quantify a bijection $\gamma \colon \mathrm{Dgm}(\mathcal{U}) \to \mathrm{Dgm}(\mathcal{V})$ by taking the sum of the q-th powers of the L_∞-distances between corresponding points. The *q-Wasserstein distance* between the persistence diagrams of $\mathcal{U}$ and $\mathcal{V}$ is the infimum of the q-th roots of these sums over all bijections:

$$\mathrm{W}_q(\mathrm{Dgm}(\mathcal{U}), \mathrm{Dgm}(\mathcal{V})) = \inf_\gamma \left(\sum_{x \in \mathrm{Dgm}(\mathcal{U})} \|x - \gamma(x)\|_\infty^q \right)^{1/q}. \tag{26.3.2}$$

It approaches the bottleneck distance as q goes to infinity.

Moments: a 1-parameter family of summaries of a persistence diagram. Fixing a real number $q \geq 0$, the *q-th moment* of $\mathcal{U}$ is the sum of the q-th powers of the persistences:

$$M_q(\mathcal{U}) = \sum_{(b,d)\in \mathrm{Dgm}(\mathcal{U})} |d-b|^q. \tag{26.3.3}$$

It is sometimes referred to as the *q-th total persistence* of $\mathcal{U}$. There are variants, such as the *level set moment*, which is the alternating sum of persistences:

$$M_\chi(\mathcal{U}) = \sum_{p\geq 0} (-1)^p \sum_{(b,d)\in \mathrm{Dgm}_p(\mathcal{U})} (d-b). \tag{26.3.4}$$

We note that $M_\chi(\mathcal{U})$ is equal to the integral of the Euler characteristic of $f^{-1}(r)$, over all $r \in \mathbb{R}$, provided $\mathcal{U} = \mathcal{H}^{\mathrm{ext}}(f)$. Indeed, each point in the extended persistence diagram contributes its persistence to the level set moment, which is the contribution of the represented cycle to the integral of the Euler characteristic. We note that points below the diagonal contribute negative values because they represent level set homology classes of one dimension lower.

Local feature size at a point: the infimum distance of $x \in \mathbb{X}$ from a point $a \in \mathbb{R}^n$ that has two or more nearest points in $\mathbb{X} \subseteq \mathbb{R}^n$, denoted $\mathrm{lfs}(x)$. The closure of the set of points $a \in \mathbb{R}^n$ with two or more nearest points in $\mathbb{X}$ is commonly referred to as the *medial axis* of $\mathbb{X}$. The global version of the local feature size is the *reach* of $\mathbb{X}$ defined as the infimum local feature size over all points of $\mathbb{X}$, $\mathrm{reach}(\mathbb{X}) = \inf_{x\in\mathbb{X}} \mathrm{lfs}(x)$. For example, if $\mathbb{X}$ is a smoothly embedded compact manifold, then the reach is positive. The reach is at most one over the maximum curvature but it can be smaller.

Homological critical value of a real-valued function: any value $a \in \mathbb{R}$ such that the homology of the sublevel sets of the function changes at a, i.e., the map $\mathsf{H}\left(f^{-1}(-\infty, a-\varepsilon]\right) \to \mathsf{H}\left(f^{-1}(-\infty, a]\right)$, induced on homology by inclusion, is not an isomorphism for any sufficiently small ε.

Weak feature size of a set: the infimum positive homological critical value of the distance function, $d_{\mathbb{X}}\colon \mathbb{R}^n \to \mathbb{R}$, denoted $\mathrm{wfs}(\mathbb{X})$. It is larger than or equal to the reach.

STABILITY

Persistence diagrams are stable under perturbations of the function. This is perhaps the most important result in the theory. It was first proved for functions using the bottleneck distance between persistence diagrams by Cohen-Steiner, Edelsbrunner and Harer [CEH07]. Under some restrictions, stability holds also using Wasserstein distances, as proved in [CEHM10]. The stability under the bottleneck distance has been strengthened to a statement about persistence modules in [BL14, CDGO12].

Bottleneck stability for functions. Recall the bottleneck distance between two persistence diagrams.

THEOREM 26.3.1 *Bottleneck Stability Theorem*

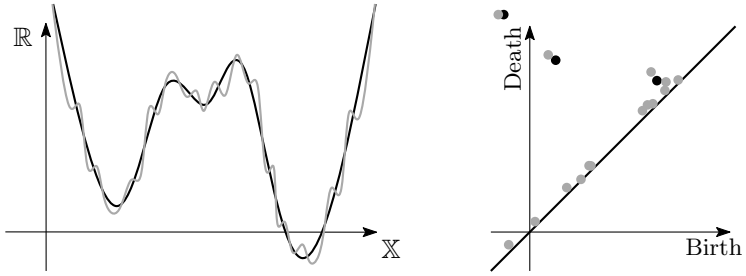
Given $f, g\colon \mathbb{X} \to \mathbb{R}$, the bottleneck distance between their persistence diagrams does

not exceed the L_∞-difference between the functions:

$$\mathrm{W}_\infty\left(\mathrm{Dgm}(f), \mathrm{Dgm}(g)\right) \leq \|f - g\|_\infty. \tag{26.3.5}$$

We refer to Figure 26.3.1 for an illustration of this inequality, which is true in great generality. For example, $\mathbb{X}$ is not constrained to manifolds or other special classes of topological spaces. Also, the requirement on f and g are mild: they need to be *tame*, which means that they have only finitely many homological critical values and every sublevel set has finite rank homology. The inequality can

FIGURE 26.3.1
Left: *the blue and red graphs of two scalar functions on $\mathbb{X} = \mathbb{R}$.* Right: *the persistence diagrams of the two functions. Each point marks the birth of a component at a local minimum and its death at a local maximum.*



be strengthened by composing g with a homeomorphism $h\colon \mathbb{X} \to \mathbb{X}$. In other words, the bottleneck distance between the two diagrams is bounded from above by $\inf_h \|f - g \circ h\|_\infty$.

Wasserstein stability for functions. For finite values of q, we get stability under the q-Wasserstein distance only if we impose conditions on the space and the function.

THEOREM 26.3.2 *Wasserstein Stability Theorem*

Let $\mathbb{X}$ be a triangulable compact metric space, let C be a constant such that the k-th moment of the persistence diagram of any function with Lipschitz constant 1 is bounded from above by C, and let $f, g\colon \mathbb{X} \to \mathbb{R}$ be two tame functions with Lipschitz constant 1. Then

$$\mathrm{W}_q\left(\mathrm{Dgm}(f), \mathrm{Dgm}(g)\right) \leq C^{\frac{1}{q}} \|f - g\|_\infty^{1 - \frac{k}{q}}, \tag{26.3.6}$$

for all $q \geq k$.

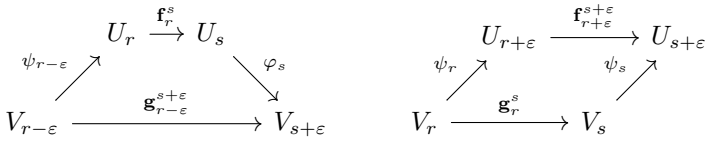
Provided the exponent is positive, $1 - \frac{k}{q} > 0$, the right-hand side vanishes as the L_∞-difference between f and g goes to zero. Hence, we have stability for $q > k$.

Interleaving isometry for persistence modules. The interleaving distance between persistence modules can be inserted between the two sides of the inequality in the Bottleneck Stability Theorem 26.3.1. Doing so, we get an equality on the left-hand side. But this equality holds more generally, namely also for persistence

modules that are not necessarily defined by functions:

$$\mathrm{W}_\infty\left(\mathrm{Dgm}(\mathcal{U}),\mathrm{Dgm}(\mathcal{V})\right) = \inf\{\varepsilon \mid \mathcal{U} \text{ and } \mathcal{V} \text{ are } \varepsilon\text{-interleaved}\}. \tag{26.3.7}$$

Traditionally, the upper bound implied by the equality is referred to as stability. For its proof it suffices to use the following two commuting diagrams and their symmetric versions:



(26.3.8)

Interleaving on log-scale. The interleaving of Čech and Vietoris–Rips complexes mentioned in the previous section suggests we draw the persistence diagrams for the two filtrations in log-scale. All critical values are non-negative, which we preserve by mapping $r \geq 0$ to $\log(r+1)$. Drawing the diagrams in log-scale is therefore the same as drawing the persistence diagrams for the Čech and Vietoris–Rips filtrations using $f = \log(d_{\mathbb{X}} + 1)$ instead of the distance function. The corresponding persistence modules are ε-interleaved for $\varepsilon = 1$, which implies $\mathrm{W}_\infty(\mathrm{Dgm}(\check{\mathrm{C}}\mathrm{ech}), \mathrm{Dgm}(\mathrm{Rips})) \leq 1$ for general metrics. The bound on the right-hand side improves to $\frac{1}{2}$ when we specialize to the Euclidean metric.

INFERENCE

The stability of persistent homology can be exploited to inferring the homology of a space from a finite point sample. We follow [CEH07] in how this is done and compare it with the inference result obtained without stability [NSW08].

Homology inference using stability. Given $\mathbb{X} \subseteq \mathbb{R}^n$, we recall that $d_{\mathbb{X}} \colon \mathbb{R}^n \to \mathbb{R}$ maps every point to its distance from the closest point in $\mathbb{X}$. Under mild conditions, it is possible to infer the homology of $\mathbb{X}$ from the distance function defined by a finite point set.

THEOREM 26.3.3 *Homology Inference Theorem*

Let $\mathbb{X} \subseteq \mathbb{R}^n$ be compact, with weak feature size $\mathrm{wfs}(\mathbb{X}) = 4\varepsilon > 0$, and let $X \subseteq \mathbb{R}^n$ be finite with Hausdorff distance $d_H(X, \mathbb{X}) \leq \varepsilon$ from $\mathbb{X}$. Then

$$\mathrm{rk}\,\mathsf{H}(\mathbb{X}) = \mathrm{rk}\left[\mathsf{H}(d_X^{-1}[0,\varepsilon]) \to \mathsf{H}(d_X^{-1}[0,3\varepsilon])\right]. \tag{26.3.9}$$

In words, every homology class of $\mathbb{X}$ can already be seen in the sublevel set at ε, and it can still be seen in the sublevel set at 3ε. Furthermore, under the given assumptions, no other homology classes can be seen in both these sublevel sets.

Homology inference without using stability. Here we let $\mathbb{X}$ be a manifold that is smoothly embedded in $\mathbb{R}^n$. Its reach is positive, and we let $X \subseteq \mathbb{X}$ have small Hausdorff distance, $d_H(X, \mathbb{X}) \leq \frac{\varepsilon}{2}$. Then for any $\varepsilon < \sqrt{3/5}\,\mathrm{reach}(\mathbb{X})$, the homology of the sublevel set of $d_X \colon \mathbb{R}^n \to \mathbb{R}$ for ε is the homology of the manifold:

$$\mathrm{rk}\,\mathsf{H}(\mathbb{X}) = \mathrm{rk}\,\mathsf{H}(d_X^{-1}[0,\varepsilon]). \tag{26.3.10}$$

This is so because the union of balls of radius ε centered at the points in X do not have any holes that are not also present in the manifold. Indeed, $d_X^{-1}[0, \varepsilon]$ is homotopy equivalent to $\mathbb{X}$, which is stronger than the claim about homology.

Comparison. The above two results differ in their assumptions on the space $\mathbb{X}$ whose homology is inferred. The first result is sensitive to the weak feature size, while the second result needs positive reach. We have $\text{reach}(\mathbb{X}) \leq \text{wfs}(\mathbb{X})$, which favors homology inference with stability. For example, if $\mathbb{X}$ is a manifold in $\mathbb{R}^n$ but not smoothly embedded, then its reach is zero—which prevents homology inference without using stability—while the weak feature size may very well be positive.

26.4 ALGORITHMS

Persistent homology owes a great deal of its popularity to the development of efficient algorithms. Edelsbrunner, Letscher, and Zomorodian [ELZ00] launched the current line of research with the introduction of a fast algorithm, which we review in this section together with various shortcuts, including optimizations of Chen and Kerber [CK11], fast updates for time-varying persistence [CEM06], and a sample of dualities in persistent homology [DMVJ11b].

GLOSSARY

Boundary matrix: a matrix representation of the boundary map. Given a simplicial complex with m_p p-simplices, the *p-th boundary matrix* is denoted $D_p[1..m_{p-1}, 1..m_p]$, with $D_p[i, j] = 0$ if σ_i is not a face of σ_j, and $D_p[i, j] = (-1)^k$ if $\sigma_j = [u_0, u_1, \ldots, u_p]$ and σ_i is σ_j with u_k removed.

Column operation: adding a multiple of one column to another. Similarly, a *row operation* adds a multiple of one row to another. Given a matrix, M, we write $M[i, \cdot]$ for its i-th row and $M[\cdot, j]$ for its j-th column.

Pivot: the lowest non-zero element in a column, $M[\cdot, j]$, denoted $\text{pvt}\, M[\cdot, j]$. The row of the pivot is denoted $\text{low}\, M[\cdot, j]$, with $\text{low}\, M[\cdot, j] = 0$ if the entire columns is zero.

Reduced matrix: a matrix M in which the pivots are in distinct rows; that is: $\text{low}\, M[\cdot, j] \neq \text{low}\, M[\cdot, k]$ or $\text{low}\, M[\cdot, j] = \text{low}\, M[\cdot, k] = 0$ whenever $j \neq k$.

REDUCTION ALGORITHM

As described in Munkres [Mun84], the homology groups of a simplicial complex can be computed by reduction of the boundary matrices. Similarly, we can compute the persistent homology by matrix reduction, but there are differences. In persistence, the ordering of the simplices is essential, and we do the reduction so it preserves order. It is therefore convenient to work with a single matrix that represents the boundary maps in all dimensions.

Filtered boundary matrix. Given a filtration, $K_0 \subseteq K_1 \subseteq \ldots \subseteq K_m$ of a simplicial complex, $K = K_m$, in which consecutive complexes differ by a single simplex, $K_i = K_{i-1} \cup \{\sigma_i\}$, we order the rows and the columns of the boundary

matrix of K by the index of their first appearance. We call the resulting matrix D, dropping the dimension subscript since the matrix combines all dimensions.

Column algorithm. Similar to Gaussian elimination, the following greedy algorithm reduces the filtered boundary matrix D: it computes a decomposition $R = DV$, in which the matrix R is reduced and the matrix V is invertible upper-triangular.

$R = D$; $V = I$;
for each column $R[\cdot, j]$ from 1 to m **do**
  while $\operatorname{low} R[\cdot, j] = \operatorname{low} R[\cdot, k] \neq 0$, with $k < j$ **do**
    $c = \operatorname{pvt} R[\cdot, j] / \operatorname{pvt} R[\cdot, k]$;
    $R[\cdot, j] = R[\cdot, j] - cR[\cdot, k]$;
    $V[\cdot, j] = V[\cdot, j] - cV[\cdot, k]$.

Observe that the algorithm works exclusively with left-to-right column operations. Each such operation is a loop, so we have three nested loops, which explains why the algorithm takes a constant times m^3 operations in the worst case.

Persistence pairing. The reduced matrix, R, contains the persistence information we are interested in. Specifically, the persistence module $\mathsf{H}(K_0) \to \mathsf{H}(K_1) \to \ldots \to \mathsf{H}(K_m)$ contains an indecomposable summand $\mathcal{I}[i, j)$ iff $\operatorname{low} R[\cdot, j] = i$. It contains a summand $\mathcal{I}[i, \infty)$ iff $R[\cdot, i] = 0$ and there is no column j with $\operatorname{low} R[\cdot, j] = i$.

Homology generators. While the indecomposable summands are encoded in the reduced matrix, the corresponding generators of the homology groups are sometimes stored in V and sometimes in R. Note that the algorithm maintains that $R[\cdot, j]$ is the boundary of the chain $V[\cdot, j]$, for every j. If $R[\cdot, i] = 0$, then $V[\cdot, i]$ is a cycle that first appears in complex K_i. If there is no column $R[\cdot, j]$ with $\operatorname{low} R[\cdot, j] = i$, then $V[\cdot, i]$ is a generator of homology in the final complex K. In contrast, if $\operatorname{low} R[\cdot, j] = i$, then $R[\cdot, j]$ records a cycle that appears in K_i and becomes a boundary in K_j. In other words, it is a homology generator in groups $\mathsf{H}(K_i)$ through $\mathsf{H}(K_{j-1})$.

Uniqueness of pairing. Given D, the decomposition $R = DV$ such that R is reduced and V is invertible upper-triangular is not unique. However, any two such decompositions have the same map, low, from the columns to the rows, including 0. When reducing the boundary matrix, we can therefore perform column operations in any order, as long as columns are added from left to right. Once the matrix is reduced, it gives the correct persistence pairing.

Row algorithm. Since the columns can be processed in any order, we can use an alternative reduction algorithm that processes the matrix row-by-row, from the bottom up. Despite its name, the algorithm reduces the matrix using column operations.

$R = D$; $V = I$;
for each row $R[i, \cdot]$ from m back up to 1 **do**
  $C = \{j \mid \operatorname{low} R[\cdot, j] = i\}$; $\mathit{leftmost} = \min C$;
  for $j \in C \setminus \{\mathit{leftmost}\}$ **do**
    $c = \operatorname{pvt} R[\cdot, j] / \operatorname{pvt} R[\cdot, \mathit{leftmost}]$;
    $R[\cdot, j] = R[\cdot, j] - cR[\cdot, \mathit{leftmost}]$;
    $V[\cdot, j] = V[\cdot, j] - cV[\cdot, \mathit{leftmost}]$.

Unlike the column algorithm, which produces the pairs in the order of deaths (from earliest to latest), the row algorithm produces the pairs in the order of births (from latest to earliest). If we are only interested in the persistence pairing, the row algorithm can discard columns once they have been used for the reduction. In

other words, column $R[\cdot, \mathit{leftmost}]$ can be dropped after completing its inner for-loop.

SHORTCUTS

For large complexes, even storing the full boundary matrix would be prohibitively expensive. To cope, we tacitly assume a sparse matrix representation that focuses on the non-zero elements. With this understanding, the above matrix reduction algorithms are surprisingly efficient, namely much faster than the worst case, which is a constant times m^3 operations. Needless to say that fast is never fast enough, and there is still much to be saved if we are clever about operations or exploit special properties that sometimes present themselves.

Compression optimization. If column $R[\cdot, j] \neq 0$ after reduction, then row $R[j, \cdot]$ can be set to zero. Indeed, the non-zero column witnesses that σ_j destroys a cycle, therefore it cannot create one. In the column algorithm, it is convenient to do this update on the fly: when processing column $R[\cdot, i]$, we can remove all those elements whose (already reduced) columns in R are not zero.

Clearing optimization. If $\operatorname{low} R[\cdot, j] = i \neq 0$ after reduction of the column, then $R[\cdot, i]$ is necessarily 0. This is helpful in speeding up the row algorithm by setting $R[\cdot, i] = 0$ immediately after processing $R[i, \cdot]$ and finishing with a non-zero row. This optimization produces a significant speed-up in practice, but it sacrifices matrix V. In other words, the optimized algorithm no longer computes the decomposition $R = DV$, but only matrix R, which is sufficient to recover the persistence pairing.

Fast algorithm for 0-dimensional persistence. Recall that both the row algorithm and the column algorithm require a constant times m^3 operations in the worst case, in which m is the number of simplices in K. If we are interested only in 0-dimensional persistence, we can take advantage of a faster algorithm that runs in time at most a constant times $m \log m$. The algorithm takes advantage of the standard union-find data structure, which maintains a collection of disjoint sets supporting the operation $\textsc{Find}(i)$, which finds the lowest-value representative of the set containing vertex i, and the operation $\textsc{Unite}(i, j)$, which unites the sets represented by i and j and, assuming $i < j$, makes i the new representative. The initial sorting dominates the running time.

Updates after a transposition. A filtration that changes continuously over time can be modeled as a sequence of transpositions of consecutive simplices. If we switch the order of simplices σ_i and σ_{i+1}, the boundary matrix changes by a transposition of the rows i and $i + 1$ and of the columns i and $i + 1$. Letting P be the corresponding permutation matrix, the new matrix $D' = PDP$. Performing the same transposition in matrices R and V, we get matrices $R' = PRP$ and $V' = PVP$, with $R' = D'V'$. But R' is not necessarily reduced and V' is not necessarily invertible upper-triangular. The latter condition can fail only if $V[i, i + 1] \neq 0$. In this case, we can subtract a multiple of column $V[\cdot, i]$ from $V[\cdot, i + 1]$, before the transposition, to ensure that $V[i, i + 1] = 0$. The former condition can fail if (1) $\operatorname{low} R[\cdot, i] = \operatorname{low} R[\cdot, i + 1]$ because of the update to matrix V, or (2) there are two columns with $\operatorname{low} R[\cdot, k] = \operatorname{low} R[\cdot, \ell] = i + 1$. Both cases can be fixed by subtracting the lower-index column from the higher-index column. In all cases, we perform only a constant number of column operations, which implies that a single

transposition takes time at most a constant times m.

Relative homology pairs and generators. Let $K_0 \subseteq K_1 \subseteq \ldots \subseteq K_m$ be a filtration of $K = K_m$. Mapping every complex, K_i, to the relative homology of the pair (K, K_i), we get a persistence module,

$$\mathsf{H}(K, K_0) \to \mathsf{H}(K, K_1) \to \ldots \to \mathsf{H}(K, K_{m-1}) \to \mathsf{H}(K, K_m), \tag{26.4.1}$$

which we denote as $\mathcal{H}^{\mathrm{rel}}(K)$. Its decomposition into interval summands is closely related to the decomposition of the module $\mathcal{H}(K)$. The intervals can be recovered from the decomposition $R = DV$ computed by either the row algorithm or the column algorithm. Specifically, $\mathcal{H}^{\mathrm{rel}}(K)$ has a summand $\mathcal{I}[i, j)$ in $(p+1)$-dimensional homology iff $\mathcal{H}(K)$ has the same summand in p-dimensional homology, which happens iff $\operatorname{low} R[\cdot, j] = i$. In this case, $V[\cdot, j]$ is the generator of the relative homology class. Module $\mathcal{H}^{\mathrm{rel}}(K)$ has a summand $\mathcal{I}[0, i)$ in p-dimensional homology iff module $\mathcal{H}(K)$ has a summand $\mathcal{I}[i, \infty)$ in p-dimensional homology, which happens iff $R[\cdot, i] = 0$ and there is no j with $\operatorname{low} R[\cdot, j] = i$. In this case, $V[\cdot, i]$ is the generator of the relative homology class.

26.5 ZIGZAG PERSISTENCE

Persistent homology fits naturally in the theory of quiver representations, where we view a discrete persistence module as a representation of the so-called A_n quiver, a path with n vertices. The theory implies that it is not important that all linear maps point in the same direction. Instead, they may alternate, and the resulting sequence still decomposes into interval summands. Carlsson and de Silva [CD10] introduced the connection between persistence modules and quiver representations to the computational topology community under the name of zigzag persistence. For technical reasons, we work with Steenrod homology throughout this section, which is a homology theory equipped with an axiom that ensures the Mayer–Vietoris sequences are exact for arbitrary spaces. Furthermore, we assume that all homology groups are finite, and that all functions have finitely many critical values. In the case of finite simplicial complexes, Steenrod homology agrees with simplicial homology, and all the piecewise-linear functions have a finite number of critical values.

GLOSSARY

Zigzag filtration: a discrete sequence of topological spaces, $(\mathbb{X}_i)_{i>0}$, such that $\mathbb{X}_i \subseteq \mathbb{X}_{i+1}$ or $\mathbb{X}_i \supseteq \mathbb{X}_{i+1}$ for every i. Its *type* is the sequence of symbols $(\tau_i)_{i>0}$, with $\tau_i \in \{\rightarrow, \leftarrow\}$, such that $\tau_i = \rightarrow$ implies $\mathbb{X}_i \subseteq \mathbb{X}_{i+1}$ and $\tau_i = \leftarrow$ implies $\mathbb{X}_i \supseteq \mathbb{X}_{i+1}$.

Zigzag persistence module: a sequence of vector spaces over a field $\mathbb{F}$ connected by linear maps $\mathbf{f}_i^{i+1} \colon U_i \to U_{i+1}$ if $\tau_i = \rightarrow$ and $\mathbf{f}_{i+1}^{i} \colon U_{i+1} \to U_i$ if $\tau_i = \leftarrow$. We denote such a module by $\mathcal{U} = (U_i, \mathbf{f})$, always assuming that the type is clear from the context. We call $\mathcal{U}$ a *zigzag interval module* if there exists an interval $[k, \ell]$ such that $U_i = 0$ whenever $i < k$ or $\ell < i$, $U_i = \mathbb{F}$ whenever $k \leq i \leq \ell$, and a linear map is zero if either the source or the target are zero, and it is the identity otherwise. We denote this zigzag interval module as $\mathcal{I}[k, \ell]$.

Zigzag interval decomposition: a zigzag persistence module decomposes as a direct sum of zigzag interval modules of the matching type, $\mathcal{U} = \oplus_j \mathcal{I}[k_j, \ell_j]$. As with ordinary persistence modules, this means we can select elements u_j^r in each vector space U_r, such that the nonzero elements form a basis of U_r, and the maps diagonalize with respect to the bases, i.e., $\mathbf{f}_r^s(u_j^r) = u_j^s$ whenever $u_j^r \neq 0$.

Relative Mayer–Vietoris long exact sequence: a long exact sequence associated to the relative homology groups of two nested pairs of spaces. Specifically, given two pairs of spaces, $(\mathbb{U}, \mathbb{A})$ and $(\mathbb{V}, \mathbb{B})$, with $\mathbb{A} \subseteq \mathbb{U}$ and $\mathbb{B} \subseteq \mathbb{V}$, the sequence

$$\cdots \to \mathsf{H}_{p+1}(\mathbb{U} \cup \mathbb{V}, \mathbb{A} \cup \mathbb{B}) \to \mathsf{H}_p(\mathbb{U} \cap \mathbb{V}, \mathbb{A} \cap \mathbb{B}) \to \mathsf{H}_p(\mathbb{U}, \mathbb{A}) \oplus \mathsf{H}_p(\mathbb{V}, \mathbb{B}) \to \mathsf{H}_p(\mathbb{U} \cup \mathbb{V}, \mathbb{A} \cup \mathbb{B}) \to \mathsf{H}_{p-1}(\mathbb{U} \cap \mathbb{V}, \mathbb{A} \cap \mathbb{B}) \to \ldots, \quad (26.5.1)$$

in which the middle two maps are induced by inclusions, is long exact.

LEVEL SET ZIGZAG

There is a natural zigzag associated to a real-valued function. One can view it as sweeping the level sets of the function from bottom to top. The construction was first introduced and studied by Carlsson et al. [CDM09].

Level set zigzag. Given a real-valued function $f\colon \mathbb{X} \to \mathbb{R}$ with critical values $s_0 < s_1 < \ldots < s_{m-1}$ and an interleaved sequence of regular values $-\infty = r_0 < s_0 < r_1 < \ldots < s_{m-1} < r_m = \infty$, we denote by $\mathbb{X}_i^{i+1} = f^{-1}[r_i, r_{i+1}]$ the pre-image of the interval $[r_i, r_{i+1}]$. We call the strictly alternating zigzag filtration,

$$\mathbb{X}_0^1 \supseteq \mathbb{X}_1^1 \subseteq \mathbb{X}_1^2 \supseteq \ldots \subseteq \mathbb{X}_{m-2}^{m-1} \supseteq \mathbb{X}_{m-1}^{m-1} \subseteq \mathbb{X}_{m-1}^m, \quad (26.5.2)$$

a *level set zigzag filtration* of f. Its type is $\leftarrow\rightarrow\leftarrow\rightarrow \ldots \leftarrow\rightarrow\leftarrow\rightarrow$. Passing to homology, we get the corresponding *level set zigzag persistence module*,

$$\mathsf{H}(\mathbb{X}_0^1) \leftarrow \mathsf{H}(\mathbb{X}_1^1) \to \ldots \leftarrow \mathsf{H}(\mathbb{X}_{m-1}^{m-1}) \to \mathsf{H}(\mathbb{X}_{m-1}^m), \quad (26.5.3)$$

which is unique up to isomorphism.

Mayer–Vietoris pyramid. We can arrange pairs of pre-images of the function in a grid, as shown in Figure 26.5.1, called the *Mayer–Vietoris pyramid*. The nodes of the grid are pairs of spaces: $(\mathbb{X}_k^\ell, \emptyset)$ in the bottom quadrant, $(\mathbb{X}_0^\ell, \mathbb{X}_0^k)$ in the left quadrant, $(\mathbb{X}_k^m, \mathbb{X}_\ell^m)$ in the right quadrant, $(\mathbb{X}, \mathbb{X}_0^k \cup \mathbb{X}_\ell^m)$ in the top quadrant.

Distinguished paths. Several paths through the Mayer–Vietoris pyramid have natural interpretations. The zigzag filtration along the bottom edge is the level set zigzag (26.5.2). The major diagonal, running from the bottom-left to the upper-right corner, and the minor diagonal, running from the bottom-right to the upper-left corner, have the following sequences of spaces:

$$\emptyset \subseteq \mathbb{X}_0^1 \subseteq \ldots \subseteq \mathbb{X}_0^m = \mathbb{X} \subseteq (\mathbb{X}, \mathbb{X}_{m-1}^m) \subseteq \ldots \subseteq (\mathbb{X}, \mathbb{X}_0^m) = (\mathbb{X}, \mathbb{X}), \quad (26.5.4)$$

$$\emptyset \subseteq \mathbb{X}_{m-1}^m \subseteq \ldots \subseteq \mathbb{X}_0^m = \mathbb{X} \subseteq (\mathbb{X}, \mathbb{X}_0^1) \subseteq \ldots \subseteq (\mathbb{X}, \mathbb{X}_0^m) = (\mathbb{X}, \mathbb{X}). \quad (26.5.5)$$

These are the extended filtrations of f and $-f$.

Sub-diagrams of extended persistence. We distinguish between three sub-diagrams of the extended persistence diagram. $\mathrm{Ord}_p(f)$ comprises the p-dimensional classes whose birth and death occur in absolute homology; $\mathrm{Rel}_p(f)$ comprises the

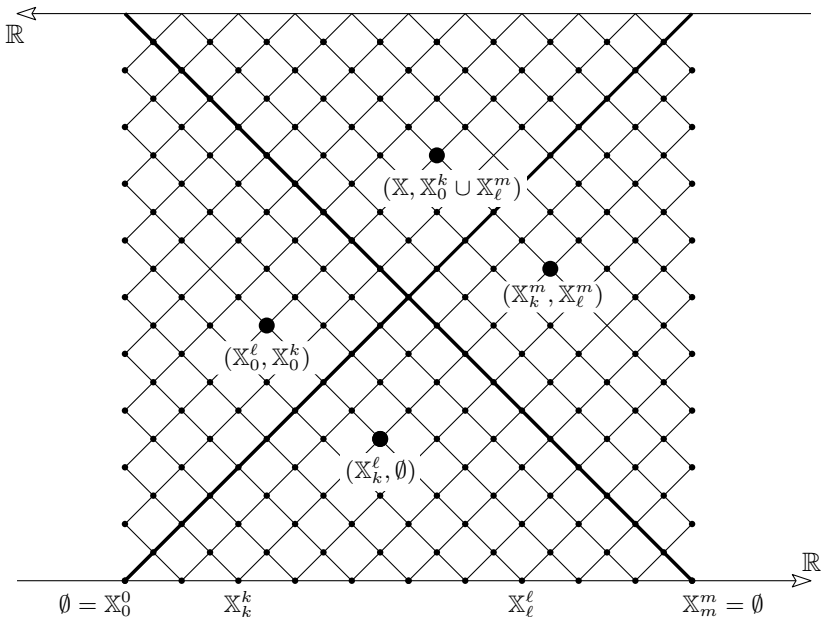


FIGURE 26.5.1

The Mayer–Vietoris pyramid. The two diagonals carry the extended persistence of f and $-f$. The pairs at the corners of every diamond satisfy the relative Mayer–Vietoris long exact sequence. The boundary map from the top row to the bottom row extends the square and turns the entire diagram into a Möbius strip.

classes whose birth and death occur in relative homology; $\mathrm{Ext}_p(f)$ comprises the classes born in absolute but dying in relative homology. They are referred to as the *ordinary*, the *relative*, and the *extended sub-diagrams* of $\mathrm{Dgm}(\mathcal{H}^{\mathrm{ext}}(f))$.

Mayer–Vietoris diamond. For every two pairs connected by a monotonically rising path in the grid, the lower pair includes into the upper pair. Moreover, any four pairs at the corners of a rectangle have a special property which we now explain; see Figure 26.5.1. The intersection of the left and right corners (as pairs of spaces) gives the bottom corner, and their union gives the top corner. For the spaces chosen in the figure, we have $(\mathbb{X}_k^\ell, \emptyset) = (\mathbb{X}_0^\ell \cap \mathbb{X}_k^m, \mathbb{X}_0^k \cap \mathbb{X}_\ell^m)$ and $(\mathbb{X}, \mathbb{X}_0^k \cup \mathbb{X}_\ell^m) = (\mathbb{X}_0^\ell \cup \mathbb{X}_k^m, \mathbb{X}_0^k \cup \mathbb{X}_\ell^m)$. This is true for all rectangles, independent of whether their corners lie in different quadrants or not. This property ensures that the homology groups of any four such spaces can be arranged in the relative Mayer–Vietoris long exact sequence, which lends its name to the pyramid.

Infinite strip. Passing to homology, the pyramid of spaces unrolls into an (infinite) strip of homology groups. The top edge of the pyramid in homological dimension p connects to the bottom edge in dimension $p-1$ via the boundary map, $\mathsf{H}_p(\mathbb{X}, \mathbb{X}_0^i \cup \mathbb{X}_{i+1}^m) \to \mathsf{H}_{p-1}(\mathbb{X}_i^i) \oplus \mathsf{H}_{p-1}(\mathbb{X}_{i+1}^{i+1})$.

DECOMPOSITION

An important property of the Mayer–Vietoris pyramid of homology groups is its decomposition into one-dimensional summands; see [CDM09, BEMP13]. It implies

that the full pyramid contains exactly the same information as the level set zigzag or extended persistence, and we can infer the homology of any interlevel set, as well as the rank of any map between any pair of interlevel sets, from the decomposition of the level set zigzag.

Flush diamonds. We distinguish a special shape inside the pyramid unrolled into an infinite strip, which we refer to as a *flush diamond*. It consists of all grid nodes inside a rectangle with sides running at 45° angles, and whose left and right corners are flush with pyramid sides. A *flush diamond indecomposable* is a pyramid that consists of zero vector spaces outside a flush diamond and of one-dimensional vector spaces, connected by identity maps, inside the diamond. Figure 26.5.2 illustrates the supports of four such indecomposables in the pyramid.

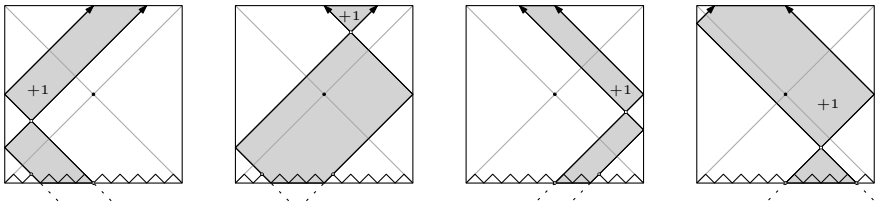


FIGURE 26.5.2
Flush diamonds in the Mayer–Vietoris pyramid that correspond to the four different types of intervals in the levelset zigzag.

Pyramid decomposition and restrictions. Any pyramid of homology groups with the Mayer–Vietoris structure decomposes as a direct sum of flush diamond indecomposables. Because flush diamond indecomposables restrict uniquely to the levelset zigzag, we can recover the decomposition of the entire pyramid from the decomposition of the zigzag. Therefore, we can read off the dimension of any (absolute or relative) homology group of the pre-image of any interval from the decomposition of the levelset zigzag. Furthermore, we can read off the rank of any map between any two such groups from the same decomposition. All zigzags that span the entire pyramid have this property, so we can recover all of the above structures from the decomposition of any other path through the pyramid—for example, from the extended persistence.

Symmetry of extended persistence. Because the two diagonals of the pyramid correspond to the extended persistence of functions f and $-f$, the decomposition of the pyramid into diamonds relates the indecomposables of the two functions. The following theorem relates the three sub-diagrams for f and $-f$; its statement and proof follow from Figure 26.5.2.

THEOREM 26.5.1 *Symmetry Theorem*

Given a function $f\colon \mathbb{X} \to \mathbb{R}$, the ordinary, relative, and extended persistence subdiagrams of f and $-f$ are related as follows:

$$\mathrm{Ord}_p(f) = \mathrm{Rel}^0_{p+1}(-f), \tag{26.5.6}$$

$$\mathrm{Ext}_p(f) = \mathrm{Ext}^R_p(-f), \tag{26.5.7}$$

$$\mathrm{Rel}_p(f) = \mathrm{Ord}^0_{p-1}(-f), \tag{26.5.8}$$

in which the superscript 0 denotes the central reflection through the origin, $(x, y) \mapsto$

$(-x,-y)$, in the diagram as drawn in Figure 26.3.1. Similarly, the superscript R denotes the reflection across the minor diagonal, $(x,y) \mapsto (-y,-x)$.

26.6 APPLICATIONS

Persistent homology is extraordinarily versatile, contributing to numerous questions in a variety of directions. In this section, we showcase the application of persistent homology to two particular topics. Following [DMVJ11a], we use persistence to write data in circular coordinates. Sketching results in [EP16], we show how persistent homology can be used to obtain converging Crofton-type formulas for the intrinsic volume of not necessarily convex shapes.

CIRCLE-VALUED COORDINATES

Machine learning offers a variety of methods to understand high-dimensional data. A popular class of techniques reduces the dimension by mapping a point set from high- to low-dimensional space, e.g., given $X \subseteq \mathbb{R}^n$, find $f\colon X \to \mathbb{R}^d$ with $d = 2$ or 3. How can we exploit topological constraints, such as the persistent homology of X in $\mathbb{R}^n$? Consider the group of homotopy classes of continuous maps from a space to the circle, denoted $[\mathbb{X}, \mathbb{S}^1]$. A classical equation in homotopy theory relates this group to the 1-dimensional cohomology group of the space, a dual of the 1-dimensional homology group, computed with integer coefficients:

$$[\mathbb{X}, \mathbb{S}^1] = \mathsf{H}^1(\mathbb{X}, \mathbb{Z}). \tag{26.6.1}$$

If we detect prominent 1-dimensional (co)cycles in the data, we can turn them into circle-maps. This idea is due to de Silva et al. [DMVJ11a], who propose to compute the persistence diagram of a filtration built on the input point sample, to select a persistent cohomology class, and to turn it into a circle-valued map. A distinctive feature of this method is that it can detect different homotopy classes of such maps and find the smoothest representative within any class. Specifically, de Silva et al. propose the following algorithm:

1. Build a Vietoris–Rips filtration, $\mathrm{Rips}(X)$, of the data. Let k be prime, use persistent cohomology to find a significant cohomology class in the filtration, and select its generator in a particular complex $[\alpha_k] \in \mathsf{H}^1(\mathrm{Rips}_r(X), \mathbb{Z}_k)$.
2. Lift $[\alpha_k]$ to a cohomology class with integer coefficients, $[\alpha] \in \mathsf{H}^1(\mathrm{Rips}_r(X), \mathbb{Z})$.
3. Smooth the integer cocycle to a harmonic cocycle in the same cohomology class, $\bar{\alpha} \in \mathsf{C}^1(\mathrm{Rips}_r(X), \mathbb{R})$.
4. Integrate the harmonic cocycle $\bar{\alpha}$ to a circle-valued function, $g\colon \mathrm{Rips}_r(X) \to \mathbb{S}^1$, which restricts to a circle-valued function on the data, $f\colon X \to \mathbb{S}^1$.

Figure 26.6.1 shows the two maps for a point sample of a figure-8.

INTRINSIC VOLUMES

We begin with a discussion of the convex case, which is well understood and contains some of the most beautiful theorems in geometry. Given a convex body, $\mathbb{K} \subseteq \mathbb{R}^n$,

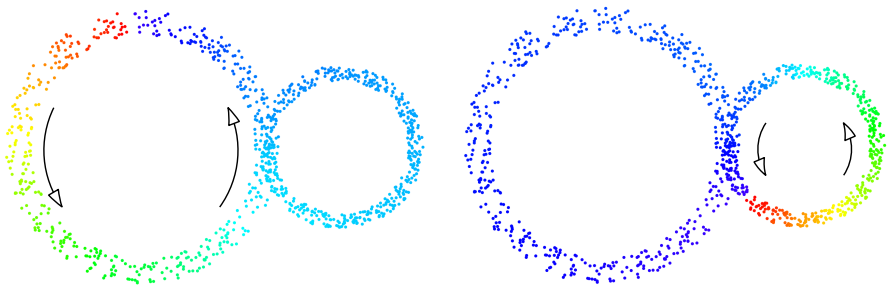


FIGURE 26.6.1
Two circle-valued parameterizations of a sample of a figure-8. The hue of the points is set to the circle coordinate. The arrows indicate the circle along which the coordinate varies smoothly.

we write $\mathbb{K}^r$ for the Minkowski sum with the ball of radius $r \geq 0$, and we recall the *Steiner polynomial* of degree n, which gives the n-dimensional volume of the thickened body:

$$\mathrm{Vol}(\mathbb{K}^r) = \sum_{p=0}^{n} b_p V_{n-p}(\mathbb{K}) \cdot r^p. \tag{26.6.2}$$

There are $n+1$ coefficients, each the product of the p-dimensional volume of the unit ball in $\mathbb{R}^p$, b_p, and the $(n-p)$-*th intrinsic volume* of $\mathbb{K}$, $V_{n-p}(\mathbb{K})$. In $\mathbb{R}^3$, the intrinsic volumes are a constant times the volume, surface area, total mean curvature, and total Gaussian curvature. Hadwiger's Characterization Theorem asserts that every measure on convex sets that is invariant under rigid motions, additive, and continuous is a linear combination of the intrinsic volumes [Had52]. The Crofton Formula provides an integral geometric representation of the intrinsic volume. Writing $\mathcal{L}_p^n \subseteq \mathcal{E}_p^n$ for the *linear* and *affine Grassmannians*, namely the p-dimensional linear and affine subspaces of $\mathbb{R}^n$, it asserts that

$$V_{n-p}(\mathbb{K}) = c_{p,n} \int_{E \in \mathcal{E}_p^n} \chi(\mathbb{K} \cap E)\, \mathrm{d}E, \tag{26.6.3}$$

for $0 \leq p \leq n$, in which $\chi(\mathbb{K} \cap E)$ is the Euler characteristic of the intersection, and $c_{p,n} = \binom{n}{p} \frac{b_n}{b_p b_{n-p}}$. The right-hand side of Crofton's Formula makes sense also for non-convex sets and can thus be used to generalize the concept of intrinsic volumes beyond convex bodies. Indeed, even the Steiner polynomial extends, albeit only to bodies with positive reach, for which it gives the correct volume for sufficiently small values of r; see e.g. the tube formulas of Weyl [Wey39].

While Crofton's Formula is invariant under rigid motion and additive also for non-convex bodies, it is not necessarily continuous. To see this, let $\mathbb{X} = B(0,1)$ be the unit disk in $\mathbb{R}^2$ and recall that the length of its boundary is 2π. Writing $r\mathbb{Z}^2$ for the scaled integer grid in the plane, we approximate $\mathbb{X}$ by the union of squares centered at grid points inside $\mathbb{X}$:

$$\mathbb{X}_r = \bigcup_{(ri,rj) \in \mathbb{X}} [ri - \tfrac{r}{2}, ri + \tfrac{r}{2}] \times [rj - \tfrac{r}{2}, rj + \tfrac{r}{2}]. \tag{26.6.4}$$

The total length of the left-facing edges bounding $\mathbb{X}_r$ varies between $2-r$ and $2+r$, and so do total lengths of the right-facing, up-facing, and down-facing edges. It

follows that the length of the boundary of $\mathbb{X}_r$ converges to 8 as r goes to 0. But 8 is not 2π. While the Crofton Formula does not converge to the correct value, we can modify it to do so. To this end, we rewrite the first intrinsic volume as an integral over level set moments:

$$V_1(\mathbb{X}) = c_{n-1,n} \int_{E \in \mathcal{E}^n_{n-1}} \chi(\mathbb{X} \cap E)\, \mathrm{d}E \tag{26.6.5}$$

$$= c_{n-1,n} \int_{L \in \mathcal{L}^n_{n-1}} \int_{y=-\infty}^{\infty} \chi(f_L^{-1}(y))\, \mathrm{d}y\, \mathrm{d}L \tag{26.6.6}$$

$$= c_{n-1,n} \int_{L \in \mathcal{L}^n_{n-1}} M_\chi(\mathcal{F}_L)\, \mathrm{d}L, \tag{26.6.7}$$

in which $f_L \colon \mathbb{X} \to \mathbb{R}$ is the height function in direction normal to L, $\mathcal{F}_L = \mathcal{H}^{\mathrm{ext}}(f_L)$, and $M_\chi(\mathcal{F}_L)$ is the level set moment of $\mathcal{F}_L$. For $\varepsilon \geq 0$, we define $M_\chi(\mathcal{F}_L, \varepsilon)$ by ignoring contributions of persistence ε or less. Finally, we define the *modified first intrinsic volume* of $\mathbb{X} \subseteq \mathbb{R}^n$ and $\varepsilon \geq 0$ as

$$V_1(\mathbb{X}, \varepsilon) = c_{n-1,n} \int_{L \in \mathcal{L}^n_{n-1}} M_\chi(\mathcal{F}_L, \varepsilon)\, \mathrm{d}L. \tag{26.6.8}$$

Choosing ε equal to the length of the cube diagonal, we can prove that the thus defined modified first intrinsic volume of the approximating body, $V_1(\mathbb{X}_r, r\sqrt{n})$, differs from $V_1(\mathbb{X})$ by at most some constant times r and therefore converges to the correct first intrinsic volume for compact bodies $\mathbb{X}$ whose boundaries are smoothly embedded $(n-1)$-manifolds in $\mathbb{R}^n$; see [EP16]. At the time of writing this article, such a convergence result was not known for all intrinsic volumes. The first open case is the surface area of bodies in $\mathbb{R}^3$, that is: $n = 3$ and $p = 1$.

26.7 SOURCES AND RELATED MATERIAL

The literature directly or indirectly related to persistent homology is extensive, and we had to make choices. Since this chapter does not focus on applications—which are numerous—we have touched the body of applied literature only tangentially.

BOOKS AND SURVEY ARTICLES

There are many textbooks in algebraic topology available, including Hatcher [Hat02] just to mention one. Of these, very few say anything about computations. In contrast, there are only three books in total that say anything about persistent homology, and the first two are heavily computational.

[Zom05]: a slightly modified version of the author's doctoral thesis.

[EH10]: an introductory text in computational topology, with a heavy focus on persistent homology.

[Ghr14]: a text on applied topology, which includes a few sections on persistent homology.

[Oud15]: a monograph on persistence, covering foundations as well as applications.

There are five survey articles on the topic of persistent homology available to aid in the introduction of non-experts into the field.

[Ghr07]: discusses the concept of the topology of data, focusing on the barcode as a concrete expression thereof.

[EH08]: surveys the state-of-the-art in 2008, giving weight to algorithms computing persistent homology.

[Car09]: takes a more abstract approach and conveys his vision of topology and data.

[Wei11]: gives a brief introduction to persistent homology addressing the mathematics community at large.

[EM12]: surveys the field in 2012, stressing the dichotomy and interplay between theory and practice.

RELATED CHAPTERS

Chapter 22: Topological methods in discrete geometry
Chapter 23: Random simplicial complexes
Chapter 24: Embedding and geometric realization
Chapter 25: Computational topology of graphs on surfaces
Chapter 27: High-dimensional topological data analysis

REFERENCES

[AEHW06] P.K. Agarwal, H. Edelsbrunner, J. Harer, and Y. Wang. Extreme elevation on a 2-manifold. *Discrete Comput. Geom.*, 36:553–572, 2006.

[BL14] U. Bauer and M. Lesnick. Induced matchings of barcodes and the algebraic stability of persistence. In *Proc. 30th Sympos. Comput. Geom.*, pages 355–364, ACM Press, 2014.

[BEMP13] P. Bendich, H. Edelsbrunner, D. Morozov, and A. Patel. Homology and robustness of level and interlevel sets. *Homology Homotopy Appl.*, 15:51–72, 2013.

[Bor48] K. Borsuk. On the imbedding of systems of compacta in simplicial complexes. *Fund. Math.*, 35:217–234, 1948.

[Car09] G. Carlsson. Topology and data. *Bull. Amer. Math. Soc.*, 46:255–308, 2009.

[CD10] G. Carlsson and V. de Silva. Zigzag persistence. *Found. Comput. Math.*, 10:367–405, 2010.

[CDGO12] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. The structure and stability of persistence modules. CGL Tech. Report, INRIA Saclay, 2012.

[CDM09] G. Carlsson, V. de Silva, and D. Morozov. Zigzag persistent homology and real-valued functions. In *Proc. 25th Sympos. Comput. Geom.*, pages 227–236, ACM Press, 2009.

[CEH07] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37:103–120, 2007.

[CEHM10] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and Y. Mileyko. Lipschitz functions have L_p-stable persistence. *Found. Comput. Math.*, 10:127–139, 2010.

[CEM06] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and vineyards by updating persistence in linear time. In *Proc. 22nd Sympos. Comput. Geom.*, pages 119–126, ACM Press, 2006.

[CK11] C. Chen and M. Kerber. Persistent homology computation with a twist. In *Proc. 27th European Workshop Comput. Geom.*, Morschach, 2011.

[DMVJ11a] V. de Silva, D. Morozov, and M. Vejdemo-Johansson. Persistent cohomology and circular coordinates. *Discrete Comput. Geom.*, 45:737–759, 2011.

[DMVJ11b] V. de Silva, D. Morozov. and M. Vejdemo-Johansson. Dualities in persistent (co)homology. *Inverse Problems*, 27:124003, 2011.

[EH08] H. Edelsbrunner and J. Harer. Persistent homology—a survey. In J.E. Goodman, J. Pach, and R. Pollack, editors, *Surveys on Discrete and Computational Geometry: Twenty Years Later*, vol.453 of *Contemp. Math.*, pages 257–282, AMS, Providence, 2008.

[EH10] H. Edelsbrunner and J. Harer. *Computational Topology. An Introduction.* AMS, Providence, 2010.

[ELZ00] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533, 2002.

[EM12] H. Edelsbrunner and D. Morozov. Persistent homology: theory and practice. In *Proc. Europ. Congress Math.*, pages 31–50, 2012.

[EP16] H. Edelsbrunner and F. Pausinger. Approximation and convergence of the intrinsic volume. *Adv. Math.*, 287:674–703, 2016.

[ES52] S. Eilenberg and N. Steenrod. *Foundations of Algebraic Topology.* Princeton University Press, 1952.

[Fro90] P. Frosini. A distance for similarity classes of submanifolds of a Euclidean space. *Bull. Australian Math. Soc.*, 42:407–416, 1990.

[Ghr07] R. Ghrist. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc.*, 45:61–75, 2007.

[Ghr14] R. Ghrist. *Elementary Applied Topology.* CreateSpace, Scotts Valley, 2014.

[Had52] H. Hadwiger. Additive Funktionale k-dimensionaler Eikörper I. *Arch. Math.*, 3:470–478, 1952.

[Hat02] A. Hatcher. *Algebraic Topology.* Cambridge University Press, 2002.

[Ler46] J. Leray. L'anneau d'homologie d'une représentation. *Les Comptes rendus de l'Académie des sciences*, 222:1366–1368, 1946.

[Mil63] J. Milnor. *Morse Theory.* Princeton University Press, 1963

[Mor40] M. Morse. Rank and span in functional topology. *Ann. Math.*, 41:419–454, 1940.

[Mun53] H.T. Munro et al. Munro's tables of the 3000-feet mountains of Scotland; and other tables of lesser heights. Scottish Mountaineering Club, 1953.

[Mun84] J.R. Munkres. *Elements of Algebraic Topology.* Perseus, Cambridge, 1984.

[NSW08] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39:419–441, 2008.

[Oud15] S.Y. Oudot. *Persistence Theory: from Quiver Representations to Data Analysis.* AMS, Providence, 2015.

[Rob99] V. Robins. Toward computing homology from finite approximations. *Topology Proc.*, 24:503–532, 1999.

[Wei11] S. Weinberger. What is . . . persistent homology? *Notices Amer. Math. Soc.*, 58:36–39, 2011.

[Wey39] H. Weyl. On the volume of tubes. *Amer. J. Math.*, 61:461–472, 1939.

[Zom05] A. Zomorodian. *Topology for Computing.* Cambridge University Press, 2005.