Morfismos, Vol. 23, No. 1, 2019, pp. 1–16

A Brief History of Persistence *

Jose A. Perea [1]

Abstract

Persistent homology is currently one of the more widely known tools from computational topology and topological data analysis. We present in this note a brief survey of how the subject has evolved over the last few years.

2010 Mathematics Subject Classification: 55N99, 68W05, 55U99
Keywords and phrases: Persistent homology, barcodes, stability, zigzags, quivers

1 The Early Years: Computer Science Meets Geometry, Algebra and Topology

Our current view of persistent homology can be traced back to work of Patrizio Frosini (1992) on size functions [29], and of Vanessa Robins (1999) [44] on using experimental data to infer the topology of attractors in dynamical systems. Both approaches rely on singular homology as a shape descriptor, which leads to what is known today as the *"homology inference problem"*: Given a finite set X (the data) sampled from/around a topological space $\mathbb{X}$ (e.g., the attractor), how can one infer the homology of $\mathbb{X}$ from X with high confidence? See for instance [41] for the case when $\mathbb{X}$ is a compact Riemannian submanifold of Euclidean space, and $X \subset \mathbb{X}$ is sampled according to the intrinsic uniform distribution. From here on out it will be useful to think of X and $\mathbb{X}$ as subspaces of a bounded metric space $(\mathbb{M}, \rho)$. In this case, one can

***Invited Paper.**
[1]This work was partially supported by the NSF (DMS-1622301) and DARPA (HR0011-16-2-003).

formalize the statement "X approximates $\mathbb{X}$" by saying that if $Z \subset \mathbb{M}$, $\epsilon \geq 0$, and $Z^{(\epsilon)} := \{x \in \mathbb{M} : \rho(x, Z) \leq \epsilon\}$, then the Hausdorff distance

$$d_H(X, \mathbb{X}) := \inf\left\{\epsilon > 0 : X \subset \mathbb{X}^{(\epsilon)} \text{ and } \mathbb{X} \subset X^{(\epsilon)}\right\}$$

is small. The goal is then to approximate the topology of $\mathbb{X}$ by that of $X^{(\epsilon)}$. Below in Figure 1 we illustrate the evolution of $X^{(\epsilon)}$ as ϵ increases.

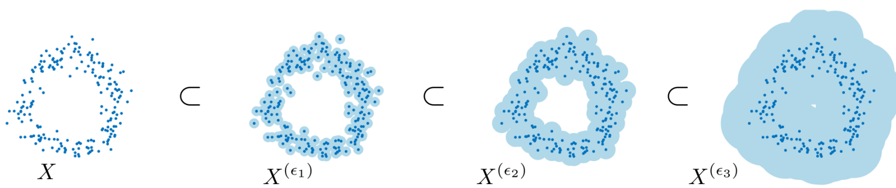

Figure 1: Some examples of $X^{(\epsilon)}$ for $X \subset \mathbb{R}^2$ sampled around the unit circle, and ϵ values $0 < \epsilon_1 < \epsilon_2 < \epsilon_3$.

In order to capture the multiscale nature of $\mathcal{X} = \{X^{(\epsilon)}\}_\epsilon$, and deal with the instability of topological features in $X^{(\epsilon)}$ as ϵ changes, Frosini and Robins introduced (independently) the idea of *homological persistence*: for $\epsilon, \delta \geq 0$ let

$$\iota^{\epsilon,\delta} : X^{(\epsilon)} \hookrightarrow X^{(\epsilon+\delta)}$$

be the inclusion, and consider the induced linear map in homology with coefficients in a field $\mathbb{F}$

$$\iota_*^{\epsilon,\delta} : H_n\left(X^{(\epsilon)}; \mathbb{F}\right) \longrightarrow H_n\left(X^{(\epsilon+\delta)}; \mathbb{F}\right)$$

The image of $\iota_*^{\epsilon,\delta}$ is the δ-persistent n-th homology group of the filtered space $\mathcal{X}$ at ϵ, denoted $H_n^{\epsilon,\delta}(\mathcal{X}; \mathbb{F})$; and

$$\mathsf{rank}\left(\iota_*^{\epsilon,\delta}\right)$$

is the persistent Betti number $\beta_n^{\epsilon,\delta}(\mathcal{X}; \mathbb{F})$.

The design of algorithms to efficiently compute/approximate these integers is of course predicated on first replacing the spaces $X^{(\epsilon)}$ by finite, combinatorial models of their topology. Fortunately there is a vast literature on how to do this. Take for instance the Vietoris-Rips complex, first introduced by Leopold Vietoris in the nineteen-twenties in an attempt to define a homology theory for general metric spaces [47]. It is defined, for $Z \subset \mathbb{M}$ and $\epsilon \geq 0$, as the abstract simplicial complex

$$R_\epsilon(Z) := \big\{\{z_0, \ldots, z_k\} \subset Z : \rho(z_i, z_j) \leq \epsilon \;\text{ for all }\; 0 \leq i, j \leq k\big\}$$

Below in Figure 2 we show an example of how $R_\epsilon(Z)$ evolves as ϵ increases, for $Z \subset \mathbb{R}^2$ sampled around the unit circle, and for ϵ values $0 < \epsilon_1 < \epsilon_2 < \epsilon_3$.

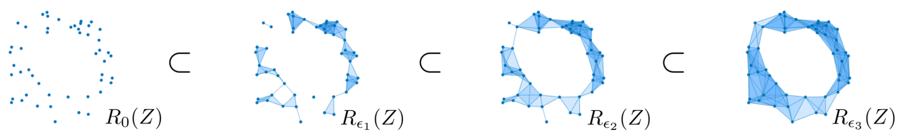

Figure 2: Some examples of the Rips complex, for points sampled around the unit circle in $\mathbb{R}^2$.

Notice that $R_\epsilon(Z) \subset R_{\epsilon+\delta}(Z)$ whenever $\delta \geq 0$; in other words, $\mathcal{R}(Z) = \{R_\epsilon(Z)\}_\epsilon$ is a filtered simplicial complex. Janko Latschev shows in [35] that when $\mathbb{X}$ is a closed Riemannian manifold, there is an $\epsilon_0 > 0$, so that if $0 < \epsilon \leq \epsilon_0$, then there exists $\delta > 0$ for which $d_H(X, \mathbb{X}) < \delta$ implies that the geometric realization of $R_\epsilon(X)$ is homotopy equivalent to $\mathbb{X}$. Discarding the manifold hypothesis — which is not expected to hold in general applications — highlights the value of persistence as a homology inference tool. Indeed, in [17] Chazal, Oudot and Yan show that if $\mathbb{X} \subset \mathbb{R}^d$ is compact with positive *weak feature size*[2] [16], and $X \subset \mathbb{R}^d$ is finite with $d_H(X, \mathbb{X})$ small, then there exists a range for $\epsilon > 0$ where $H_n^{\epsilon,3\epsilon}(\mathcal{R}(X); \mathbb{F})$ is isomorphic to $H_n\left(\mathbb{X}^{(\epsilon)}; \mathbb{F}\right)$. It is worth noting that while these theorems deal with small ϵ, far less is known about the large-scale regime. Indeed, aside from trivial examples, the circle is (essentially) the only space Z for which the homotopy type of $R_\epsilon(Z)$ is known explicitly for all $\epsilon > 0$ [1, 2].

The efficient computation of the persistent Betti numbers of a finite filtered simplicial complex

$$\mathcal{K} = \{K_0 \subset K_1 \subset \cdots \subset K_J = K\},$$

was addressed by Edelsbrunner, Letscher and Zomorodian in (2000) [27], for subcomplexes of a triangulated 3-sphere and homology with coefficients in $\mathbb{F}_2 = \{-1, 1\}$. This restriction was a tradeoff between

[2] this is a notion of how complex the embedding of $\mathbb{X}$ into Euclidean space is.

generality and speed: the algorithm was based on previous work of Delfinado and Edelsbrunner [22] to compute (standard) Betti numbers incrementally in time $O(N\alpha(N))$, where N is the number of simplices of K and α is the inverse of the Ackermann function [19]. Since the Ackermann function grows very rapidly, its inverse α grows very slowly. Though limited in generality, the approach by Delfinado and Edelsbrunner highlights the following idea: If K_j is obtained from K_{j-1} by adding a single simplex $\tau \in K$, and

$$H_n(K_{j-1};\mathbb{F}) \longrightarrow H_n(K_j;\mathbb{F})$$

is not surjective, then either τ is an n-simplex creating a new homology class, or it is an $n+1$-simplex eliminating a class from K_{j-1}. Thus, simplices in K that either create or annihilate a given persistent homology class can be put in pairs (τ,σ) of the form creator-annihilator. These pairings are in fact a byproduct of the incremental algorithm of Delfinado and Edelsburnner. The *barcode* is also introduced in [27] as a visualization tool for persistence: each pair (τ,σ) yields an interval $[j,\ell)$, where j (birth time) is the smallest index so that $\tau \in K_j$, and $\ell > j$ (death time) is the smallest index for which $\sigma \in K_\ell$. Thus, long intervals indicate stable homological features throughout $\mathcal{K}$, while short ones reflect topological noise. The resulting multiset of intervals (as repetitions are allowed) is called a barcode. The notation is $\mathsf{bcd}_n(\mathcal{K})$. Moreover, the barcode subsumes the persistent Betti numbers, since $\beta_n^{\epsilon,\delta}(\mathcal{K};\mathbb{F})$ is the number of intervals $[j,\ell) \in \mathsf{bcd}_n(\mathcal{K})$ with $j \leq \epsilon$ and $\ell > \epsilon + \delta$. Below in Figure 3 we show an example of a filtered simplicial complex, the simplicial pairings (τ,σ), and the resulting barcodes.

2 Here Comes the Algebra

The developments up to this point can be thought of as the computational and geometric era of persistent homology. Around 2005 the focus started to shift towards algebra. Zomorodian and Carlsson introduced in [50] the persistent homology

$$PH_n(\mathcal{K};\mathbb{F}) := \bigoplus_{j\in\mathbb{Z}} H_n(K_j;\mathbb{F}) \quad , \quad \mathcal{K} = \{K_j\}_{j\in\mathbb{Z}}$$

of a filtered complex $\mathcal{K}$, as the graded module over $\mathbb{F}[t]$ with left multiplication by t on j-homogeneous elements given by the linear map

$$\phi_j : H_n(K_j;\mathbb{F}) \longrightarrow H_n(K_{j+1};\mathbb{F})$$

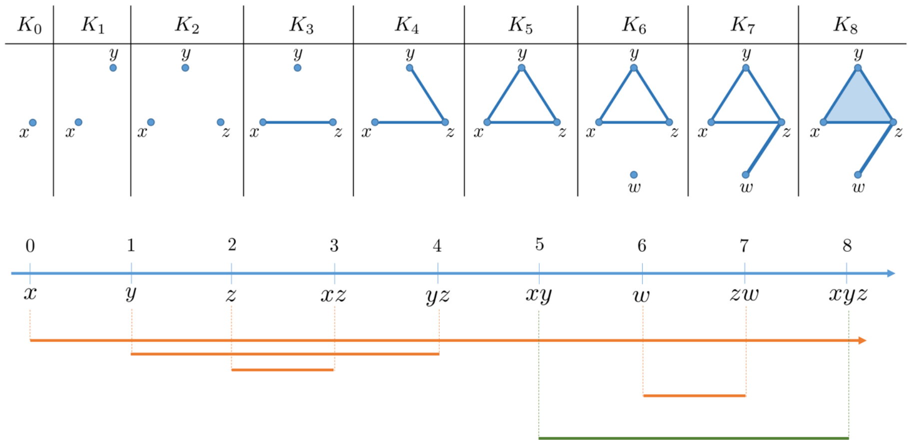

Figure 3: A filtered simplicial complex $\mathcal{K} = \{K_0 \subset \cdots \subset K_8\}$, along with the simplicial pairings (τ, σ), and the resulting barcodes for homology in dimensions 0 (orange) and 1 (green).

induced by the inclusion $K_j \hookrightarrow K_{j+1}$. Since then, $PH_n(\mathcal{K}; \mathbb{F})$ is referred to in the literature as a persistence module. More generally [6, 7], let J and C be categories with J small (i.e., so that its objects form a set). The category of J-indexed persistence objects in C is defined as the functor category $\mathsf{Fun}(\mathsf{J}, \mathsf{C})$; its objects are functors $F : \mathsf{J} \to \mathsf{C}$, and its morphisms are natural transformations $\varphi : F_1 \Rightarrow F_2$. The typical indexing category comes from having a partially ordered set $(P, \preceq)$, and letting $\underline{P}$ denote the category with objects $\mathsf{Obj}(\underline{P}) = P$, and a unique morphism from p_1 to p_2 whenever $p_1 \preceq p_2$. We'll abuse notation and denote this morphism by $p_1 \preceq p_2$, instead of the categorical notation $p_1 \to p_2$.

It can be readily checked that if Mod_R denotes the category of (left) modules over a commutative ring R with unity, and $g\mathsf{Mod}_{R[t]}$ is the category of $\mathbb{Z}$-graded modules over the polynomial ring $R[t]$, then

$$
\begin{array}{rcl}
\mathsf{Fun}(\underline{\mathbb{Z}}, \mathsf{Mod}_R) & \longrightarrow & g\mathsf{Mod}_{R[t]} \\
M \ , \ \varphi & \mapsto & \bigoplus_{j \in \mathbb{Z}} M(j) \ , \ \bigoplus_{j \in \mathbb{Z}} \varphi_j
\end{array}
\tag{1}
$$

is an equivalence of categories. On the graded $R[t]$-module side, multiplication by t on j-homogeneous elements is given by $M(j \leq j+1) : M(j) \longrightarrow M(j+1)$. This equivalence shows why/how the evolution of homological features in a $\mathbb{Z}$-filtered complex $\mathcal{K}$, can be encoded as the algebraic structure of the persistence module $PH_n(\mathcal{K};\mathbb{F})$.

2.1 Persistence Modules and Barcodes

When $PH_n(\mathcal{K};\mathbb{F})$ is finitely generated as an $\mathbb{F}[t]$-module — e.g. if $K_j = \emptyset$ for $j < 0$ and $\bigcup\limits_{j\in\mathbb{Z}} K_j$ is finite — then one has a graded isomorphism

$$PH_n(\mathcal{K};\mathbb{F}) \cong \left(\bigoplus_{q=1}^{Q} t^{n_q}\cdot\mathbb{F}[t]\right) \oplus \left(\bigoplus_{\ell=1}^{L} \left(t^{m_\ell}\cdot\mathbb{F}[t]\right)/\left(t^{m_\ell+d_\ell}\right)\right) \tag{2}$$

for $n_q, m_\ell \in \mathbb{Z}$ and $d_\ell \in \mathbb{N}$ [48]. The decomposition (2) is unique up to permutations, and thus the intervals

$$\begin{aligned}&[n_1,\infty),[n_2,\infty),\ldots,[n_Q,\infty),\\ &\qquad\qquad [m_1,m_1+d_1),[m_2,m_2+d_2),\ldots,[m_L,m_L+d_L)\end{aligned}$$

provide a complete discrete invariant for (i.e., they uniquely determine) the $\mathbb{F}[t]$-isomorphism type of $PH_n(\mathcal{K};\mathbb{F})$. Moreover, this multiset recovers the barcode $\mathsf{bcd}_n(\mathcal{K})$ of Edelsbrunner, Letscher and Zomorodian [27].

Carlsson and Zomorodian also observe that $PH_n(\mathcal{K};\mathbb{F})$ is in fact the homology of an appropriate chain complex of graded $\mathbb{F}[t]$-modules. Hence, a graded version of the Smith Normal Form [24] computes the barcode decomposition (2), providing a general-purpose algorithm. This opened the flood gates; barcodes could now be computed as a linear algebra problem for any finite filtered simplicial complex $K_0 \subset \cdots \subset K_J = K$, over any (in practice finite) field of coefficients, and up to any homological dimension. The resulting matrix reduction algorithm, implemented initially in the JPlex library (now javaPlex) [3], runs in polynomial time: its worst time complexity is $O(N^3)$, where N is the number of simplices of K. In fact Dmitriy Morozov exhibits in [40] a finite filtered complex of dimension 2, attaining the worst-case. This shows that the cubic bound is tight for general barcode computations.

While this sounds potentially slow, specially compared to the time complexity $O(N\cdot\alpha(N))$ of the sequential algorithm, Morozov's example should be contrasted with filtrations arising from applications. In practice the matrices to be reduced are sparse, and computing their associated barcode decomposition takes at worst matrix-multiplication time $O(N^\omega)$ [38], where $\omega \approx 2.373$ [49]. Over the last ten years or so there has been a flurry of activity towards better implementations and faster persistent homology computations. A recent survey [42] compares several leading open source libraries for computing persistent homology. All of them implement different optimizations, exploit new theoretical developments and novel heuristics/approximations. For instance, one improvement is to first simplify the input filtered complex without changing its persistent homology (e.g., using discrete Morse theory [39]); or to compute persistent cohomology, since it is more efficient than persistent homology and gives the same answer [21].

The shift towards algebra has had other important consequences; specifically: 1) a better understanding of stability for barcodes, and 2) several theorems describing how the choice of categories J and C impacts the computability of isomorphism invariants for objects in $\mathsf{Fun}(\mathsf{J}, \mathsf{C})$. Let me say a few words about stability.

2.2 The Stability of Persistence

Let $\mathbb{X}$ be a triangulable topological space (e.g., a smooth manifold) and let $f : \mathbb{X} \longrightarrow \mathbb{R}$ be a tame function (this is a generalization of being Morse). The prototypical example in TDA arises from a compact set $X \subset \mathbb{R}^d$, and letting $f_X : \mathbb{R}^d \longrightarrow \mathbb{R}$ be

$$f_X(y) = \inf_{x \in X} \|x - y\|.$$

Thus $f_X^{-1}(-\infty, \epsilon] = X^{(\epsilon)}$. Given $f : \mathbb{X} \longrightarrow \mathbb{R}$, let $\mathsf{bcd}_n(f)$ denote the barcode for the n-th persistent homology of $\left\{f^{-1}(-\infty, \epsilon]\right\}_{\epsilon \in \mathbb{R}}$. Drawing inspiration from Morse theory, Cohen-Steiner, Edelsbrunner and Harer introduced in (2007) [18] two foundational ideas: (1) the bottleneck distance $d_B(\,\cdot\,,\,\cdot\,)$ between barcodes, and (2) the stability theorem asserting that for tame $f, g : \mathbb{X} \longrightarrow \mathbb{R}$ one has that[3]

$$d_B(\mathsf{bcd}_n(f), \mathsf{bcd}_n(g)) \leq \|f - g\|_\infty$$

In particular, if $X, Y \subset \mathbb{R}^d$ are compact and $\mathcal{X} = \left\{X^{(\epsilon)}\right\}_\epsilon$, $\mathcal{Y} = \left\{Y^{(\epsilon)}\right\}_\epsilon$, then $d_B(\mathsf{bcd}_n(\mathcal{X}), \mathsf{bcd}_n(\mathcal{Y})) \leq d_H(X, Y)$. This inequality implies that slight changes to the input data change the barcodes slightly, which is key for applications where (Hausdorff) noise plays a role.

[3] A similar result was established in [25] for $n = 0$.

Towards the end of 2008 Chazal et. al. solidified the idea of stability with the introduction of interleavings for $\mathbb{R}$-indexed persistence modules [13]. The construction is as follows. For $\delta \geq 0$ let $T_\delta : \underline{\mathbb{R}} \longrightarrow \underline{\mathbb{R}}$ be the translation functor $T_\delta(\epsilon) = \epsilon + \delta$. An δ-interleaving between two persistence vector spaces $V, W : \underline{\mathbb{R}} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ is a pair (φ, ψ) of natural transformations

$$\varphi : V \Rightarrow W \circ T_\delta \qquad \text{and} \qquad \psi : W \Rightarrow V \circ T_\delta$$

so that $\psi_{\epsilon+\delta} \circ \varphi_\epsilon = V(\epsilon \leq \epsilon + 2\delta)$ and $\varphi_{\epsilon+\delta} \circ \psi_\epsilon = W(\epsilon \leq \epsilon + 2\delta)$ for all $\epsilon \in \mathbb{R}$. The interleaving distance between V and W, denoted $d_I(V, W)$, is defined as the infimum over all $\delta \geq 0$ so that V and W are δ-interleaved, if interleavings exist. If there are no interleavings, the distance is defined as ∞. It readily follows that d_I is an extended (since infinity can be a value) pseudometric on $\mathsf{Fun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$, and that $d_I(V, W) = 0$ whenever $V \cong W$. The converse, however, is false in general (more on this later).

Chazal et. al. [13] show that if $V : \underline{\mathbb{R}} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ satisfies $\mathsf{rank}\big(V(\epsilon < \epsilon')\big) < \infty$ for all pairs $\epsilon < \epsilon'$, this is called being q-tame, then V has a well-defined barcode $\mathsf{bcd}(V)$ (see [20] for a shorter proof when $\mathsf{dim}_{\mathbb{F}} V(\epsilon) < \infty$ for all ϵ; this is called being pointwise finite). Moreover, if V, W are q-tame, then one has the algebraic stability theorem $d_B(\mathsf{bcd}(V), \mathsf{bcd}(W)) \leq d_I(V, W)$. This turns out to be an equality:

$$d_B(\mathsf{bcd}(V), \mathsf{bcd}(W)) = d_I(V, W)$$

which nowadays is referred to as the Isometry Theorem; the first known proof is due to Lesnick [36].

As I said earlier, $d_I(V, W)$ can be zero for V and W nonisomorphic, and thus $\mathsf{bcd}(V)$ is not a complete invariant in the q-tame $\mathbb{R}$-indexed case. This can be remedied as follows. Let $\mathsf{qFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$ denote the full subcategory of $\mathsf{Fun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$ comprised of q-tame persistence modules. The ephemeral category $\mathsf{eFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$, is the full subcategory of $\mathsf{qFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$ with objects $V : \underline{\mathbb{R}} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ satisfying $V(\epsilon < \epsilon') = 0$ for all $\epsilon < \epsilon'$. The observable category $\mathsf{oFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$ is the quotient category

$$\mathsf{qFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}}) / \mathsf{eFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}})$$

As shown by Chazal et. al. in [14], d_I descends to an extended metric on the observable category, and taking barcodes

$$\mathsf{bcd} : \big(\mathsf{oFun}(\underline{\mathbb{R}}, \mathsf{Mod}_{\mathbb{F}}), d_I\big) \longrightarrow \big(\mathsf{Bcd}, d_B\big)$$

is an isometry. Hence, the barcode is a complete invariant for the isomorphism type of observable $\mathbb{R}$-indexed persistence vector spaces. In summary, the modern view of stability is algebraic; persistence modules are compared via interleavings, which one then tries to relate to the bottleneck distance between the associated barcodes if they exist.

2.3 Changing Indexing Categories: Multi-d Persistence, Quivers and Zigzags

One of the early realizations in TDA was the usefulness of having filtrations indexed by more than one parameter (1999) [30]. For instance, given a data set $X \subset \mathbb{R}^d$ one might want to focus on densely-populated regions [9], or portions with high/low curvature [12]. This leads naturally to $\mathbb{Z}^n$-filtered complexes: $\{K_{\mathbf{u}}\}_{\mathbf{u}\in\mathbb{Z}^n}$, $\mathbf{u} = (u_1, \ldots, u_n)$, where $K_{\mathbf{u}} \subset K_{\mathbf{v}}$ whenever $\mathbf{u} \preceq \mathbf{v}$ (i.e., $u_1 \leq v_1, \ldots, u_n \leq v_n$). In this multi-filtered complex, each filtering direction $u_1, \ldots, u_n$ is meant to capture an attribute: e.g. distance/scale, density, curvature, etc. Taking homology with coefficients in $\mathbb{F}$ yields objects in $\mathsf{Fun}(\underline{\mathbb{Z}^n}, \mathsf{Mod}_{\mathbb{F}})$, and just like before, $\mathbb{Z}^n$-indexed persistence $\mathbb{F}$-vector spaces correspond to n-graded modules over the n-graded polynomial ring $P_n := \mathbb{F}[t_1, \ldots, t_n]$. Parameterizing the isomorphism classes of said modules, for $n \geq 2$, turns out to be much more involved than the barcodes from $n = 1$. Indeed, around 2009 Carlsson and Zomorodian [11] showed that the isomorphism type of a finitely generated n-graded P_n-module is uniquely determined by the following data: two finite multisets $\xi_0, \xi_1 \subset \mathbb{Z}^n$ encoding the location and multiplicity of birth-death events, and a point in the quotient of an algebraic variety $\mathcal{RF}(\xi_0, \xi_1)$ by the algebraic action of an algebraic group G_{ξ_0}. The multisets ξ_0, ξ_1 are the discrete portions of the resulting isomorphism invariant, while $\mathcal{RF}(\xi_0, \xi_1)/G_{\xi_0}$ parameterizes the (potentially) continuous part. Here is an example due to Carlsson and Zomorodian [11] illustrating how complicated this quotient can be. For $n = 2$, consider the isomorphism classes of P_2-modules having $\xi_0 = \{(0,0), (0,0)\}$ and $\xi_1 = \{(3,0), (2,1), (1,2), (0,3)\}$. If $\mathsf{Gr}_1(\mathbb{F}^2)$ denotes the Grassmannian of lines in $\mathbb{F}^2$, then

$$\mathcal{RF}(\xi_0, \xi_1) = \mathsf{Gr}_1(\mathbb{F}^2) \times \mathsf{Gr}_1(\mathbb{F}^2) \times \mathsf{Gr}_1(\mathbb{F}^2) \times \mathsf{Gr}_1(\mathbb{F}^2)$$

and G_{ξ_0} turns out to be the degree 2 general linear group $\mathsf{GL}_2(\mathbb{F})$ acting diagonally on $\mathsf{Gr}_1(\mathbb{F}^2)^4$. Since $\mathsf{Gr}_1(\mathbb{F}^2)^4/\mathsf{GL}_2(\mathbb{F})$ contains a copy of $\mathbb{F} \smallsetminus \{0,1\}$, and each point in this set yields a distinct isomorphism class of

P_2-modules, it follows that there is no complete discrete invariant for (finite!) multi-d persistence.

The vast majority of recent results from multidimensional persistence focus on computable descriptors/visualizations of its intricate algebraic structure. Besides introducing the parametrization

$$\xi_0, \xi_1, \mathcal{RF}(\xi_0, \xi_1)/G_{\xi_0},$$

Carlsson and Zomorodian also propose the rank invariant: For a q-tame module $V : \underline{\mathbb{Z}^n} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$, it is defined as the function ρ_V sending each pair $\mathbf{u} \preceq \mathbf{v}$ in $\mathbb{Z}^n$ to the integer $\mathsf{rank}\, V(\mathbf{u} \preceq \mathbf{v})$. ρ_V is computable (see [10] for a polynomial-time algorithm), it is discrete, and an invariant of the isomorphism type of V. When $n = 1$ one can recover $\mathsf{bcd}(V)$ from ρ_V and viceversa, and thus ρ_V is complete in the 1-dimensional case. Knudson notes in [34] that $\xi_0(V)$ and $\xi_1(V)$ are in fact the locations/multiplicities of birth events in the torsion modules $\mathsf{Tor}_0^{P_n}(V, \mathbb{F})$ and $\mathsf{Tor}_1^{P_n}(V, \mathbb{F})$, respectively; here $\mathbb{F}$ is identified with the P_n-module $\mathbb{F}[t_1, \ldots, t_n]/(t_1, \ldots, t_n)$. The higher-dimensional analogs $\mathsf{Tor}_j^{P_n}(V, \mathbb{F})$, $j \geq 2$, lead to a family of finite multisets $\xi_j(V) \subset \mathbb{Z}^n$, each with its own geometric interpretation, serving as isomorphism invariants for V. Other approaches to invariants for multidimensional persistence include the Hilbert Series of Harrington et. al. [33], the extended algebraic functions of Skryzalin and Carlsson [46], and the feature counting invariant of Scolamiero et. al. [45]. Lesnick and Wright have recently released RIVET, the Rank Invariant Visualization and Exploration Tool [37]. Put simply, RIVET uses the fact that if $V : \underline{\mathbb{R}^2} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ is q-tame and $L \subset \mathbb{R}^2$ is a line with nonnegative slope (hence a totally ordered subset of $(\mathbb{R}^2, \preceq)$), then $V^L : \underline{L} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$, the 1-dimensional persistence vector space obtained by restricting V to L, has a well-defined barcode $\mathsf{bc}\left(V^L\right)$. The key feature in RIVET is a graphical interface which, for finite bi-filtrations, displays $\mathsf{bc}\left(V^L\right)$ interactively as the user varies L. This is particularly useful for parameter selection and the exploratory analysis of data sets with filtering functions.

Multidimensional persistence is a great example of how a seemingly innocuous change in indexing category, say from $\underline{\mathbb{Z}}$ to $\underline{\mathbb{Z}^2}$, can lead to a widely different and much more complicated classification problem. With this in mind, one would like to have a systematic approach to address the ensuing complexity. The representation theory of Quivers [23] offers one such avenue. It turns out that the classification of finite J-indexed persistence vector spaces $V : \mathsf{J} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ can be studied directly from the shape of the indexing category J. Indeed, let $G(\mathsf{J})$ be the finite directed (multi)graph with the objects of J as vertices, and one arrow for every morphism that is neither an identity nor a composition. Also, let $\widetilde{G}(\mathsf{J})$ be the undirected graph obtained from $G(\mathsf{J})$ by forgetting arrow directions. When $\widetilde{G}(\mathsf{J})$ is acyclic, Gabriel's theorem [31] implies that pointwise finite objects in $\mathsf{Fun}(\mathsf{J}, \mathsf{Mod}_{\mathbb{F}})$ can be classified via complete discrete invariants, if and only if the connected components of $\widetilde{G}(\mathsf{J})$ are Dynkin diagrams of the types described in Figure 4 below.

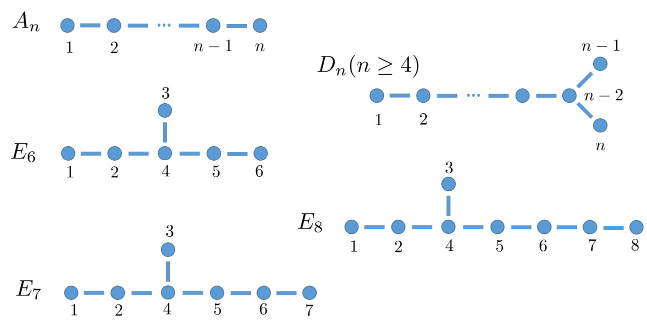

Figure 4: Dynkin diagrams of type A_n for $n \geq 1$, D_n for $n \geq 4$, and E_n for $n = 6, 7, 8$.

Here is an example of how this result can be used to avoid unpleasant surprises: Suppose that $G(\mathsf{J})$ is the graph with vertices $x_0, \ldots, x_N$ and $N \geq 5$ edges $x_n \to x_0$, $n = 1, \ldots, N$ (see Examples 3 and 8 in [23]). While the resulting J-indexed persistence vector spaces $V : \mathsf{J} \longrightarrow \mathsf{Mod}_{\mathbb{F}}$ may look simple (just star-shaped, right?), the connected graph $\widetilde{G}(\mathsf{J})$ is not a Dynkin diagram, and the ensuing classification problem is in fact of "wild type": complete invariants must include continuous high-dimensional pieces, just like in multidimensional persistence.

These ideas entered the TDA lexicon around 2010 with the definition of Zigzag persistence by Carlsson and de Silva [8]. Regular persistence addresses the problem of identifying stable homological features in a monotone system of spaces and continuous maps

$$X_1 \to X_2 \to \cdots \to X_J.$$

Zigzag persistence, on the other hand, is a generalization to the non-monotone case. Here is a practical example: suppose one has an ordered sequence of spaces $X_1, \ldots, X_J$ (e.g., from time varying data), but no obvious maps $X_j \to X_{j+1}$. The need to track topological features as j

varies leads one to consider the system

$$X_1 \hookrightarrow X_1 \cup X_2 \hookleftarrow X_2 \hookrightarrow \cdots \hookleftarrow X_j \cup X_{j+1} \hookrightarrow \cdots \hookleftarrow X_J$$

and the resulting *zigzag diagram*

$$V_1 \rightarrow V_2 \leftarrow V_3 \rightarrow \cdots \leftarrow V_n$$

at the homology level. More generally, a (finite) zigzag is a sequence of vector spaces $V_1, \ldots, V_n$ and linear maps $V_j \rightarrow V_{j+1}$ or $V_j \leftarrow V_{j+1}$. The sequence of arrow directions, e.g. $\tau = (\mathsf{left}, \mathsf{left}, \mathsf{right}, ..., \mathsf{right}, \mathsf{left})$, is the zigzag type. Since in this case any choice of τ forces the indexing category J_τ to satisfy $\widetilde{G}(\mathsf{J}_\tau) = A_n$ (one of the aforementioned Dynkin diagrams), then Gabriel's theorem implies that finite zigzags

$$V : \mathsf{J}_\tau \longrightarrow \mathsf{Mod}_{\mathbb{F}}$$

are completely classified by a discrete invariant. Just as for regular 1-dimensional persistence the invariant turns out to be a barcode, which can be efficiently computed [38], and for which there is a zigzag stability theorem [5] recently established by Botnan and Lesnick.

When the graph $\widetilde{G}(\mathsf{J})$ has cycles, the functoriality of objects in $\mathsf{Fun}(\mathsf{J}, \mathsf{Mod}_{\mathbb{F}})$ is captured by the notion of a quiver with relations. The taxonomy from Gabriel's theorem no longer applies, but one can still find some answers in the representation theory of associative algebras. A particularly important instance is when the cycles of $\widetilde{G}(\mathsf{J})$ are not oriented cycles in $G(\mathsf{J})$; in this case the algebras of interest are finite dimensional (hence Artinian) and Auslander-Rieten theory [4] becomes relevant. Escolar and Hiraoka [28] have recently put these ideas to use in the context of persistent objects indexed by commutative ladders; that is, the persistence of a morphism between two zigzags of the same type:

$$\begin{tikzcd}
\bullet \arrow[r] \arrow[d] & \bullet \arrow[d] & \bullet \arrow[l] \arrow[d] & \cdots \arrow[l] \arrow[r] & \bullet \arrow[r] \arrow[d] & \bullet \arrow[d] & \bullet \arrow[l] \arrow[d] \\
\bullet \arrow[r] & \bullet & \bullet \arrow[l] & \cdots \arrow[l] \arrow[r] & \bullet \arrow[r] & \bullet & \bullet \arrow[l]
\end{tikzcd}$$

The resulting theory sits somewhere between zigzag persistence and multi-dimensional persistence: short ladders (length ≤ 4) have complete discrete invariants, but longer ones do not. Escolar and Hiraoka present an algorithm for computing these invariants, and also an interesting application to computational chemistry.

I think this is a good place for me to stop; hopefully it is also a good starting point for the reader interested in persistent homology. There are several books covering many of the ideas I presented here, as well as many others. The interested reader would certainly benefit from these resource [32, 15, 26, 43].

Jose A. Perea
Department of Computational Mathematics
Science & Engineering,
Department of Mathematics,
Michigan State University,
East Lansing, MI, USA
joperea@msu.edu

References

[1] Adamaszek, M., and Adams, H. The vietoris–rips complexes of a circle. *Pacific Journal of Mathematics 290*, 1 (2017), 1–40.

[2] Adamaszek, M., Adams, H., and Reddy, S. On vietoris–rips complexes of ellipses. *Journal of Topology and Analysis* (2017), 1–30.

[3] Adams, H., Tausz, A., and Vejdemo-Johansson, M. Javaplex: A research software package for persistent (co) homology. In *International Congress on Mathematical Software* (2014), Springer, pp. 129–136. Software available at https://appliedtopology.github.io/javaplex.

[4] Auslander, M., Reiten, I., and Smalo, S. O. *Representation theory of Artin algebras*, vol. 36. Cambridge university press, 1997.

[5] Botnan, M. B., and Lesnick, M. Algebraic stability of zigzag persistence modules. *arXiv preprint arXiv:1604.00655* (2016).

[6] Bubenik, P., de Silva, V., and Scott, J. Metrics for Generalized Persistence Modules. *Foundations of Computational Mathematics 15*, 6 (2015), 1501–1531.

[7] Bubenik, P., and Scott, J. A. Categorification of persistent homology. *Discrete & Computational Geometry 51*, 3 (2014), 600–627.

[8] Carlsson, G., and De Silva, V. Zigzag persistence. *Foundations of computational mathematics 10*, 4 (2010), 367–405.

[9] Carlsson, G., Ishkhanov, T., De Silva, V., and Zomorodian, A. On the local behavior of spaces of natural images. *International journal of computer vision 76*, 1 (2008), 1–12.

[10] Carlsson, G., Singh, G., and Zomorodian, A. Computing multidimensional persistence. In *International Symposium on Algorithms and Computation* (2009), Springer, pp. 730–739.

[11] Carlsson, G., and Zomorodian, A. The theory of multidimensional persistence. *Discrete & Computational Geometry 42*, 1 (2009), 71–93.

[12] Carlsson, G., Zomorodian, A., Collins, A., and Guibas, L. J. Persistence barcodes for shapes. *International Journal of Shape Modeling 11*, 02 (2005), 149–187.

[13] Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L. J., and Oudot, S. Y. Proximity of persistence modules and their diagrams. In *Proceedings of the twenty-fifth annual symposium on Computational geometry* (2009), ACM, pp. 237–246.

[14] Chazal, F., Crawley-Boevey, W., and De Silva, V. The observable structure of persistence modules. *arXiv preprint arXiv:1405.5644* (2014).

[15] Chazal, F., De Silva, V., Glisse, M., and Oudot, S. *The structure and stability of persistence modules.* Springer, 2016.

[16] Chazal, F., and Lieutier, A. Weak feature size and persistent homology: computing homology of solids in $\mathbb{R}^n$ from noisy data samples. In *Proceedings of the twenty-first annual symposium on Computational geometry* (2005), ACM, pp. 255–262.

[17] Chazal, F., and Oudot, S. Y. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry* (2008), ACM, pp. 232–241.

[18] Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. Stability of persistence diagrams. *Discrete & Computational Geometry 37*, 1 (2007), 103–120.

[19] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. *Introduction to algorithms.* MIT press, 2009.

[20] Crawley-Boevey, W. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications 14*, 05 (2015), 1550066.

[21] De Silva, V., Morozov, D., and Vejdemo-Johansson, M. Dualities in persistent (co) homology. *Inverse Problems 27*, 12 (2011), 124003.

[22] Delfinado, C. J. A., and Edelsbrunner, H. An incremental algorithm for betti numbers of simplicial complexes on the 3-sphere. *Computer Aided Geometric Design 12*, 7 (1995), 771–784.

[23] Derksen, H., and Weyman, J. Quiver representations. *Notices of the AMS 52*, 2 (2005), 200–206.

[24] Dumas, J.-G., Heckenbach, F., Saunders, D., and Welker, V. Computing simplicial homology based on efficient smith normal form algorithms. In *Algebra, Geometry and Software Systems.* Springer, 2003, pp. 177–206.

[25] dAmico, M., Frosini, P., and Landi, C. Optimal matching between reduced size functions. *DISMI, Univ. di Modena e Reggio Emilia, Italy, Technical report*, 35 (2003).

[26] Edelsbrunner, H., and Harer, J. *Computational topology: an introduction.* American Mathematical Soc., 2010.

[27] Edelsbrunner, H., Letscher, D., and Zomorodian, A. Topological persistence and simplification. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on* (2000), IEEE, pp. 454–463.

[28] Escolar, E. G., and Hiraoka, Y. Persistence modules on commutative ladders of finite type. *Discrete & Computational Geometry 55*, 1 (2016), 100–157.

[29] Frosini, P. Measuring shapes by size functions. In *Intelligent Robots and Computer Vision X: Algorithms and Techniques* (1992), International Society for Optics and Photonics, pp. 122–133.

[30] Frosini, P., Mulazzani, M., et al. Size homotopy groups for computation of natural size distances. *Bulletin of the Belgian Mathematical Society-Simon Stevin 6*, 3 (1999), 455–464.

[31] Gabriel, P. Unzerlegbare darstellungen i. *Manuscripta mathematica 6*, 1 (1972), 71–103.

[32] Ghrist, R. W. *Elementary applied topology*, vol. 1. Createspace Seattle, 2014.

[33] Harrington, H. A., Otter, N., Schenck, H., and Tillmann, U. Stratifying multiparameter persistent homology. *arXiv preprint arXiv:1708.07390* (2017).

[34] Knudson, K. A. A refinement of multidimensional persistence. *Homology, Homotopy & Applications 10*, 1 (2008), 259–281.

[35] Latschev, J. Vietoris-rips complexes of metric spaces near a closed riemannian manifold. *Archiv der Mathematik 77*, 6 (2001), 522–528.

[36] Lesnick, M. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics 15*, 3 (2015), 613–650.

[37] Lesnick, M., and Wright, M. Interactive visualization of 2-d persistence modules. *arXiv preprint arXiv:1512.00180* (2015). Software available online at http://rivet.online.

[38] Milosavljević, N., Morozov, D., and Skraba, P. Zigzag persistent homology in matrix multiplication time. In *Proceedings of the twenty-seventh annual symposium on Computational geometry* (2011), ACM, pp. 216–225.

[39] Mischaikow, K., and Nanda, V. Morse theory for filtrations and efficient computation of persistent homology. *Discrete & Computational Geometry 50*, 2 (2013), 330–353.

[40] Morozov, D. Persistence algorithm takes cubic time in worst case. *BioGeometry News, Dept. Comput. Sci., Duke Univ 2* (2005).

[41] Niyogi, P., Smale, S., and Weinberger, S. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry 39*, 1-3 (2008), 419–441.

[42] Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., and Harrington, H. A. A roadmap for the computation of persistent homology. *EPJ Data Science 6*, 1 (2017), 17.

[43] Oudot, S. Y. *Persistence theory: from quiver representations to data analysis*, vol. 209. American Mathematical Society Providence, RI, 2015.

[44] Robins, V. Towards computing homology from finite approximations. In *Topology proceedings* (1999), vol. 24, pp. 503–532.

[45] Scolamiero, M., Chachólski, W., Lundman, A., Ramanujam, R., and Öberg, S. Multidimensional persistence and noise. *Foundations of Computational Mathematics 17*, 6 (2017), 1367–1406.

[46] Skryzalin, J., and Carlsson, G. Numeric invariants from multidimensional persistence. *Journal of Applied and Computational Topology 1*, 1 (2017), 89–119.

[47] Vietoris, L. Über den höheren zusammenhang kompakter räume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen 97*, 1 (1927), 454–472.

[48] Webb, C. Decomposition of graded modules. *Proceedings of the American Mathematical Society 94*, 4 (1985), 565–571.

[49] Williams, V. V. Multiplying matrices faster than coppersmith-winograd. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing* (2012), ACM, pp. 887–898.

[50] Zomorodian, A., and Carlsson, G. Computing persistent homology. *Discrete & Computational Geometry 33*, 2 (2005), 249–274.