Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Patrick Schnider

Introduction to Topological Data Analysis **Scribe Notes 17** **FS23**

Scribe notes by Simon Weber. Please contact me for corrections.
Lecture date: April 28, 2023
Last update: Tuesday 2nd May, 2023, 10:35

Interleaving Distance

Until now, we compared persistence diagrams. We will now introduce the interleaving distance, which instead compares persistence modules. Let us begin with a formal definition of persistence modules.

Definition 1. *A* persistence module $\mathbb{V}$ *over* $\mathbb{R}$ *is a collection* $\mathbb{V} = \{V_a\}_{a\in\mathbb{R}}$ *of vector spaces* V_a *together with linear maps* $v_{a,a'} : V_a \to V_{a'}$ *such that* $v_{a,a} = \text{id}$ *and* $v_{b,c} \circ v_{a,b} = v_{a,c}$ *for all* $a \le b \le c$.

You already know a few examples of persistence modules, e.g., the persistent homology of sublevel set filtrations or of Čech or Vietoris-Rips complexes (here one simply defines $V_a = 0$ for $a < 0$).

When are two persistence modules "the same"?

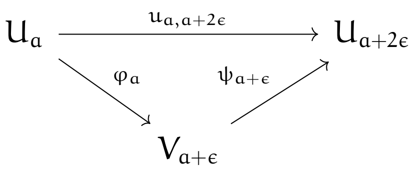
Definition 2. *We say that two persistence modules* $\mathbb{U}$ *and* $\mathbb{V}$ *are* isomorphic *if there are isomorphisms* $f_a : U_a \to V_a$ *such that*

$$\begin{array}{ccc} U_a & \xrightarrow{u_{a,a'}} & U_{a'} \\ \updownarrow{\scriptstyle f_a} & & \updownarrow{\scriptstyle f_{a'}} \\ V_a & \xrightarrow{v_{a,a'}} & V_{a'} \end{array}$$

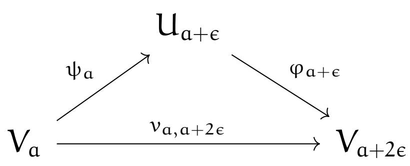
commutes both ways, i.e., $f_{a'} \circ u_{a,a'} = v_{a,a'} \circ f_a$*, and* $u_{a,a'} \circ f_a^{-1} = f_{a'}^{-1} \circ v_{a,a'}$.

Definition 3 (ϵ-interleaving persistence modules)**.** *Let* $\mathbb{U}$ *and* $\mathbb{V}$ *be persistence modules over* $\mathbb{R}$*. We say that* $\mathbb{U}$ *and* $\mathbb{V}$ *are* ϵ-interleaved *if there exist two families of maps,* $\varphi_a : U_a \to V_{a+\epsilon}$ *and* $\psi_a : V_a \to U_{a+\epsilon}$ *such that the following four diagrams are commutative:*

$$\begin{array}{ccccc} U_a & & \xrightarrow{u_{a,a'}} & & U_{a'} \\ & \searrow^{\varphi_a} & & & & \searrow^{\varphi_{a'}} \\ & & V_{a+\epsilon} & & \xrightarrow{v_{a+\epsilon,a'+\epsilon}} & & V_{a'+\epsilon} \end{array} \quad \textit{and} \quad \begin{array}{ccccc} & & U_{a+\epsilon} & & \xrightarrow{u_{a+\epsilon,a'+\epsilon}} & & U_{a'+\epsilon} \\ & \nearrow^{\psi_a} & & & & \nearrow^{\psi_{a'}} \\ V_a & & \xrightarrow{v_{a,a'}} & & V_{a'} \end{array}$$



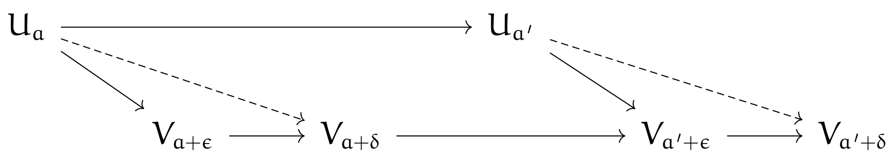
and



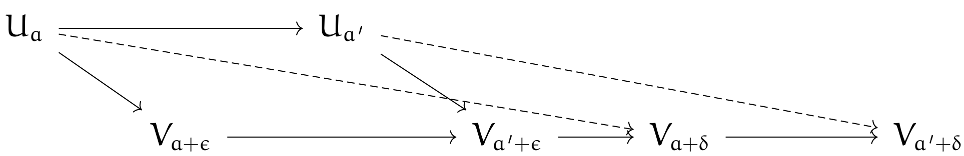
Theorem 4. *Assume $\mathbb{U}$ and $\mathbb{V}$ are ϵ-interleaving. Let $\delta > \epsilon$. Then $\mathbb{U}$ and $\mathbb{V}$ are also δ-interleaving.*

Proof. Given $\varphi'_a : U_a \to V_{a+\epsilon}$ we define $\varphi_a : U_a \to V_{a+\delta}$ simply as $\varphi_a := \nu_{a+\epsilon,a+\delta} \circ \varphi'_a$. Symmetrically, we define $\psi_a := u_{a+\epsilon,a+\delta} \circ \psi'_a$. To check that the correct diagrams commute, we only check the right of every pair of symmetric ones above. We have to distinguish two cases for the first diagram, $a + \delta < a' + \epsilon$ and $a + \delta > a' + \epsilon$.

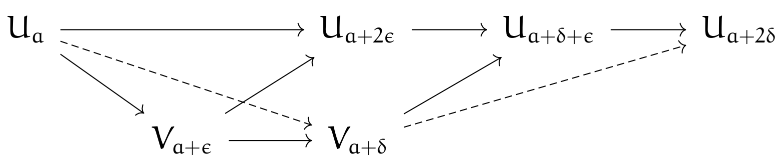
For the first case, we get the following diagram:



For the second case we get the diagram:



And finally, for the triangular diagram we get:



One can now verify that in all of these diagrams the correct paths commute. □

Thus, the following definition makes sense:

Definition 5 (Interleaving distance). $d_I(\mathbb{U}, \mathbb{V}) := \inf\{\epsilon \mid \mathbb{U}$ *and* $\mathbb{V}$ *are* ϵ*-interleaved* $\}$.

Definition 6 (Interleaving for Filtrations). *Let $\mathcal{F}, \mathcal{G}$ be filtrations over $\mathbb{R}$. $\mathcal{F}$ and $\mathcal{G}$ are ϵ-interleaved if there exist maps $\varphi_a : F_a \to G_{a+\epsilon}$ and $\psi_a : G_a \to F_{a+\epsilon}$ such that the same type of diagrams commute up to homotopy, that is, for example $\varphi_{a'} \circ \iota^F_{a,a'} \simeq \iota^G_{a+\epsilon,a'+\epsilon} \circ \varphi_a$ are homotopic (contiguous).*

We again define the interleaving distance (now between filtrations):

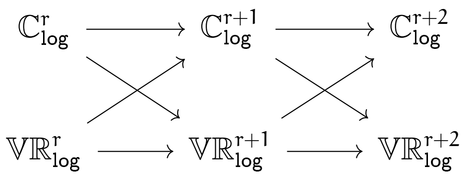
$$d_I(\mathcal{F}, \mathcal{G}) = \inf\{\epsilon \mid \mathcal{F} \text{ and } \mathcal{G} \text{ are } \epsilon\text{-interleaved } \}.$$

Observation 7. *For all $p \geq 0$, $d_I(H_p\mathcal{F}, H_p\mathcal{G}) \leq d_I(\mathcal{F}, \mathcal{G})$.*

The proof follows immediately from induced homology.

Recall that for a point cloud P and a radius r, we have the relationship between the Čech and Vietoris-Rips complexes as follows: $\mathbb{C}^r(P) \subseteq \mathbb{VR}^r(P) \subseteq \mathbb{C}^{2r}(P)$. Since this factor 2 is multiplicative, and we need an additive ϵ for interleaving, let us just take the logarithmic scale (base 2) for the radius, i.e., we define $\mathbb{C}^r_{\log} = \mathbb{C}^{2^r}$ and similarly $\mathbb{VR}^r_{\log} = \mathbb{VR}^{2^r}_{\log}$. Since $2^{(r+1)} = 2r$, we have $\mathbb{C}^r_{\log}(P) \subseteq \mathbb{VR}^r_{\log}(P) \subseteq \mathbb{C}^{r+1}_{\log}(P)$.

We thus have the following inclusions:



Since these are all inclusions, all relevant diagrams must commute, and thus we get that $d_I(\mathbb{C}_{\log}, \mathbb{VR}_{\log}) \leq 1$.

Definition 8. *A persistence module $\mathbb{V}$ is* q-tame *if the linear maps have finite rank.*

Note that in this definition, the q is not a parameter, just a name.

Theorem 9. *If $\mathbb{U}, \mathbb{V}$ are q-tame persistence modules over $\mathbb{R}$, then*

$$d_b(\mathrm{Dgm}\mathbb{U}, \mathrm{Dgm}\mathbb{V}) = d_I(\mathbb{U}, \mathbb{V}).$$

Thus, for every interleaving one can find between two persistence modules or between filtrations, one immediately gets a bound on the Bottleneck distance.

Let us consider an example where this theorem helps us a lot.

Interleaving of Čech Complexes

Consider two point clouds P, Q in the same metric space X.

Let us first consider the really simple case, where $P = \{p\}$, and $Q = \{q\}$ with $d(p, q) = d$. Then, $B(p, r) \subseteq B(q, r + d)$. Now, how does this generalize to larger point sets? To get the same kind of behaviour, we need that for every point in P, there exists some point in Q with distance at most d. This motivates the following distance measure:

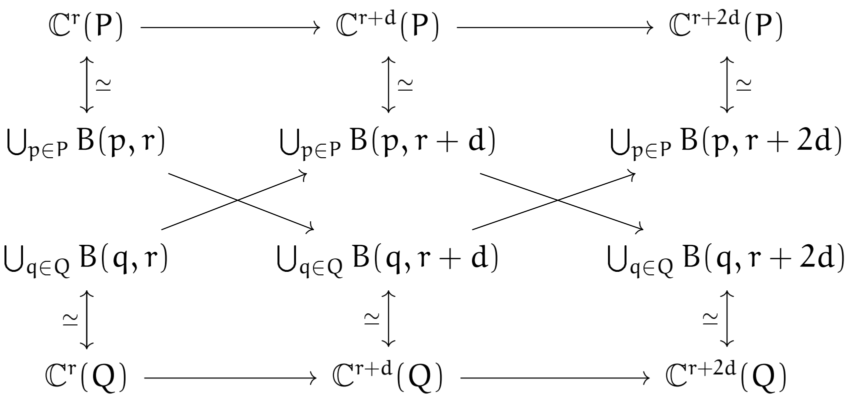
Definition 10 (Hausdorff distance). *Let $A, B \subseteq X$ be compact sets. Then the Hausdorff distance between A and B is defined as*

$$d_H(A, B) := \max\{\max_{a \in A} d(a, B), \max_{b \in B} d(b, A)\}.$$

Let $d_H(P, Q) = d$. Then, $\bigcup_{p \in P} B(p, r) \subseteq \bigcup_{q \in Q} B(q, r + d)$. From this, we get the following lemma:

Lemma 11. *The (filtration given by) the Čech complexes of P and Q are d-interleaved.*

Proof.



The relevant diagrams commute up to homotopy, since we only chain together homotopies and inclusion maps. □

Theorem 12. $d_b(\mathrm{Dgm}_p(\mathbb{C}(P)), \mathrm{Dgm}_p(\mathbb{C}(Q))) \leq d_H(P, Q)$ *for all* $p \geq 0$.

Proof. By Theorem 9, Observation 7, and finally Lemma 11, we have

$$d_b(\ldots) = d_I(H_p\mathbb{C}(P), H_p\mathbb{C}(Q)) \leq d_I(\mathbb{C}(P), \mathbb{C}(Q)) \leq d_H(P, Q).$$

□