Topological Data Analysis

2022–2023

Lecture 7

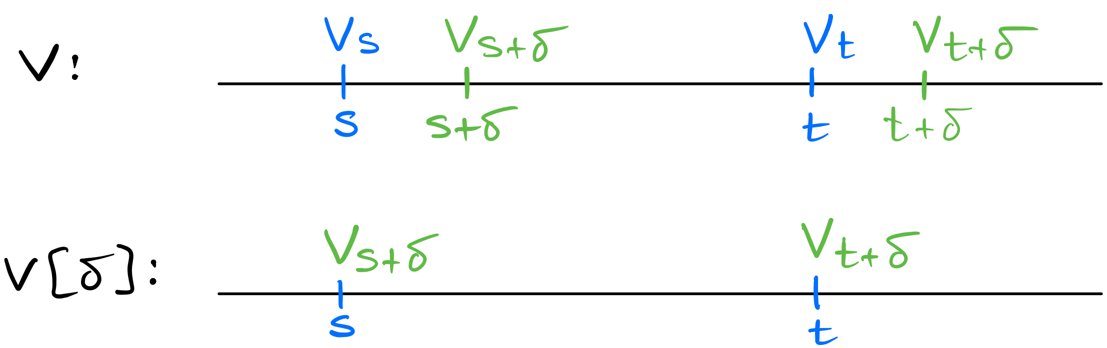
Interleaving Distance

24 November 2022

Let (V,π) be a persistence module and let $\delta \in \mathbb{R}$.
Let us define a persistence module $(V[\delta], \pi[\delta])$ as follows:

$$V[\delta]_t = V_{t+\delta}, \qquad \pi[\delta]_{s,t} = \pi_{s+\delta,\, t+\delta}.$$

This is called a <u>δ-shift</u> of (V,π). It is a backwards shift if $\delta > 0$.



Note that $\pi[\delta]_{s,t} \circ \pi[\delta]_{r,s} = \pi_{s+\delta,\, t+\delta} \circ \pi_{r+\delta,\, s+\delta} = \pi_{r+\delta,\, t+\delta} = \pi[\delta]_{r,t}$. ✓

Note also that, if (V,π) is of finite type, then $(V[\delta], \pi[\delta])$ is also of finite type. If the spectrum of V is $\{a_0, \dots, a_n\}$ then the spectrum of $V[\delta]$ is $\{a_0 - \delta, \dots, a_n - \delta\}$.

If $\delta \geq 0$, then there is a morphism $\sigma_\delta : V \to V[\delta]$ given by

$$(\sigma_\delta)_t = \pi_{t,t+\delta}.$$

Check that σ_δ is indeed a morphism:

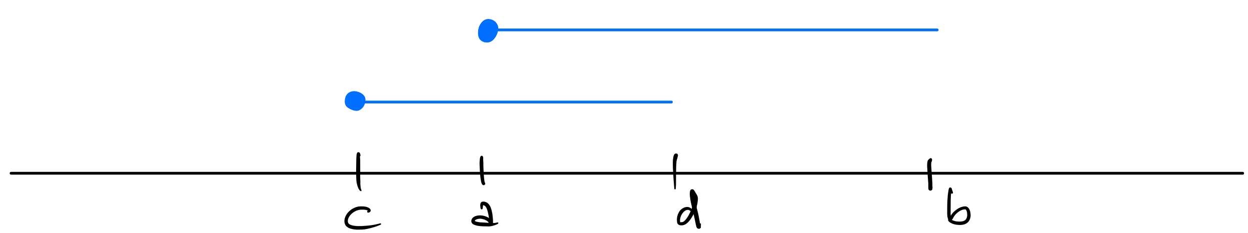
$$\left\{\begin{aligned} &(\sigma_\delta)_t \circ \pi_{s,t} = \pi_{t,t+\delta} \circ \pi_{s,t} = \pi_{s,t+\delta} \quad \text{since } s \leq t \leq t+\delta, \text{ as } \delta \geq 0 \\ &\pi[\delta]_{s,t} \circ (\sigma_\delta)_s = \pi_{s+\delta,t+\delta} \circ \pi_{s,s+\delta} = \pi_{s,t+\delta} \quad \checkmark \end{aligned}\right.$$

Moreover, each morphism $f : V \to V'$ of persistence modules yields a morphism $f[\delta] : V[\delta] \to V'[\delta]$ for all $\delta \in \mathbb{R}$, namely

$$f[\delta]_t = f_{t+\delta}.$$

$$f[\delta]_t \circ \pi[\delta]_{s,t} = f_{t+\delta} \circ \pi_{s+\delta,t+\delta} = \pi'_{s+\delta,t+\delta} \circ f_{s+\delta} = \pi'[\delta]_{s,t} \circ f[\delta]_s. \quad \checkmark$$

In what follows, recall that there is a nonzero morphism $\mathbb{F}[a,b) \to \mathbb{F}[c,d)$ if and only if $c \leq a < d \leq b$:



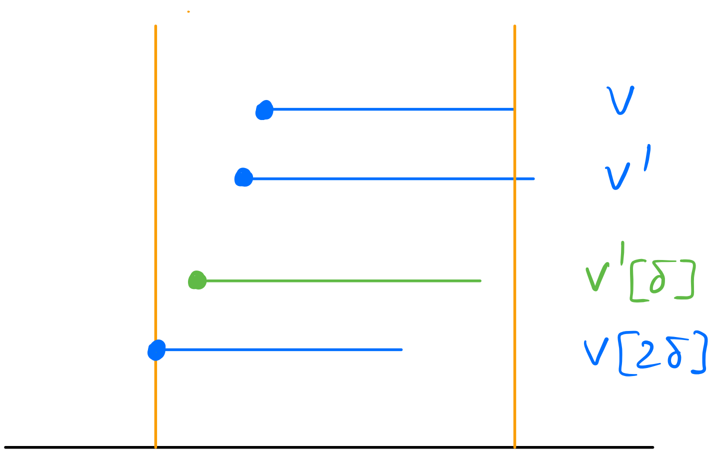
For $\delta > 0$, two persistence modules V and V' are <u>δ-interleaved</u> if there exist morphisms

$$V \xrightarrow{f} V'[\delta], \qquad V' \xrightarrow{g} V[\delta]$$

such that $g[\delta] \circ f = \sigma_{2\delta}$ and $f[\delta] \circ g = \sigma'_{2\delta}$.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[rr, bend right, "\sigma_{2\delta}"'] & V'[\delta] \arrow[r, "g[\delta]"] & V[2\delta] \end{tikzcd}$$

$$\begin{tikzcd} V' \arrow[r, "g"] \arrow[rr, bend right, "\sigma'_{2\delta}"'] & V[\delta] \arrow[r, "f[\delta]"] & V'[2\delta] \end{tikzcd}$$



$V'[\delta]$ cannot exceed the vertical lines if V and V' are δ-interleaved

V and V' are 0-interleaved $\Longleftrightarrow V \cong V'$

Suppose that V and V' are of finite type with ordered spectra $\{a_0, \dots, a_n\}$ and $\{a'_0, \dots, a'_m\}$. Denote $V_{a_n} = V_\infty$ and $V'_{a'_m} = V'_\infty$.

If V and V' are δ-interleaved for some $\delta > 0$, then $\dim V_\infty = \dim V'_\infty$.

Proof: Pick $t \in \mathbb{R}$ such that $t \geq \max\{a_n, a'_m\}$. Then $(\sigma_{2\delta})_t = \pi_{t,t+2\delta}$ is an isomorphism. Since $\sigma_{2\delta} = g[\delta] \circ f$, we infer that $f_t : V_t \to V'_{t+\delta}$ is injective and $g_{t+\delta} : V'_{t+\delta} \to V_{t+2\delta}$ is surjective. Hence $\dim V_t \leq \dim V'_{t+\delta}$ and this tells us that $\dim V_\infty \leq \dim V'_\infty$ since $V_t \cong V_\infty$ and $V'_{t+\delta} \cong V'_\infty$. Similarly $(\sigma'_{2\delta})_t$ is an isomorphism and it follows that $\dim V'_\infty \leq \dim V_\infty$. ✓

The interleaving distance between two persistence modules of finite type V and V' with $\dim V_\infty = \dim V'_\infty$ is defined as

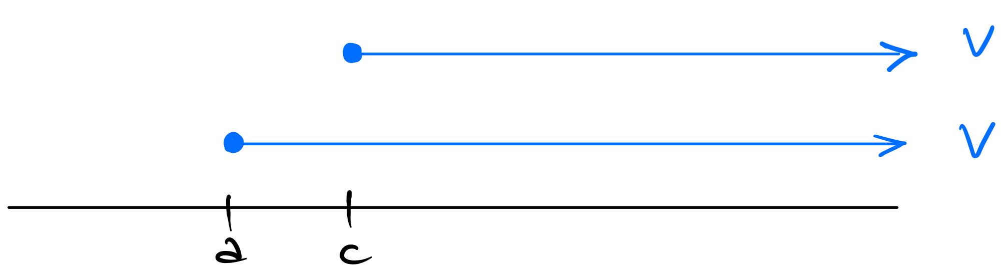
$$\boxed{d_{int}(V, V') = \inf\{\delta > 0 \mid V \text{ and } V' \text{ are } \delta\text{-interleaved}\}.}$$

We will next prove that V and V' are δ-interleaved for some $\delta > 0$ and therefore this number is well defined.

Examples:

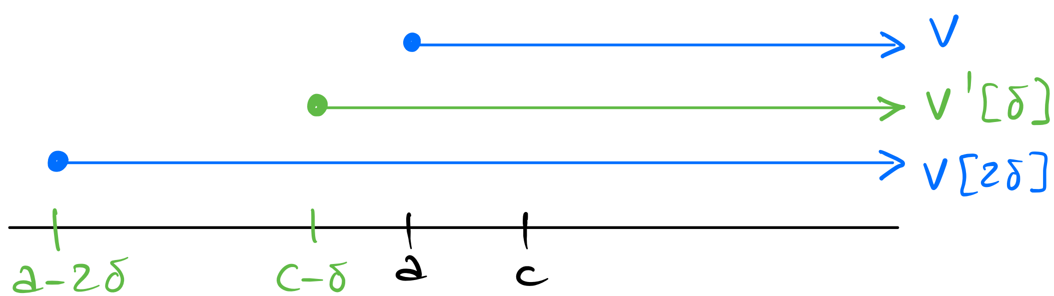
(1) $V = \mathbb{F}[a,\infty)$

$V' = \mathbb{F}[c,\infty)$

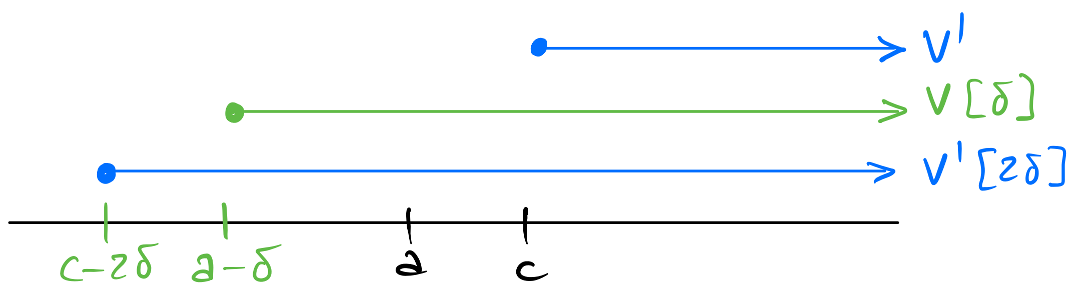


Suppose that $c > a$ without loss of generality.

First note that if $\delta \geq c-a$ then V and V' are δ-interleaved:

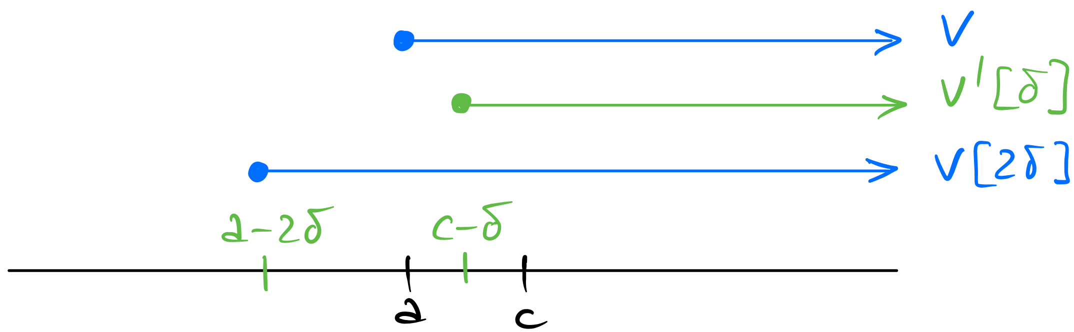


$$\left.\begin{array}{l} \delta \geq c-a \Rightarrow c-\delta \leq a \\ a-2\delta < a-\delta < c-\delta \end{array}\right\} \Rightarrow a-2\delta < c-\delta \leq a. \checkmark$$



$$\delta \geq c-a \Rightarrow c-\delta \leq a \Rightarrow c-2\delta \leq a-\delta \Rightarrow c-2\delta \leq a-\delta < a < c. \checkmark$$

We next check that if $\delta < c-a$ then V and V' are <u>not</u> δ-interleaved.



$\delta < c-a \Rightarrow a < c-\delta \Rightarrow$ Every morphism $f: V \to V'[\delta]$ is zero

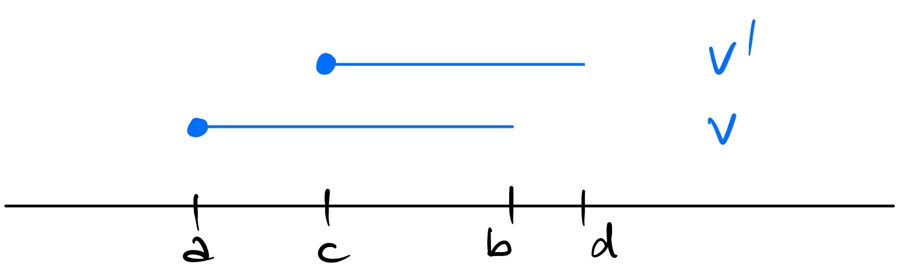
$\Rightarrow$ There are no f and g such that $g[\delta] \circ f = \sigma_{2\delta}$.

In conclusion, $d_{int}(\mathbb{F}[a,\infty), \mathbb{F}[c,\infty)) = |a-c|$.

(2)

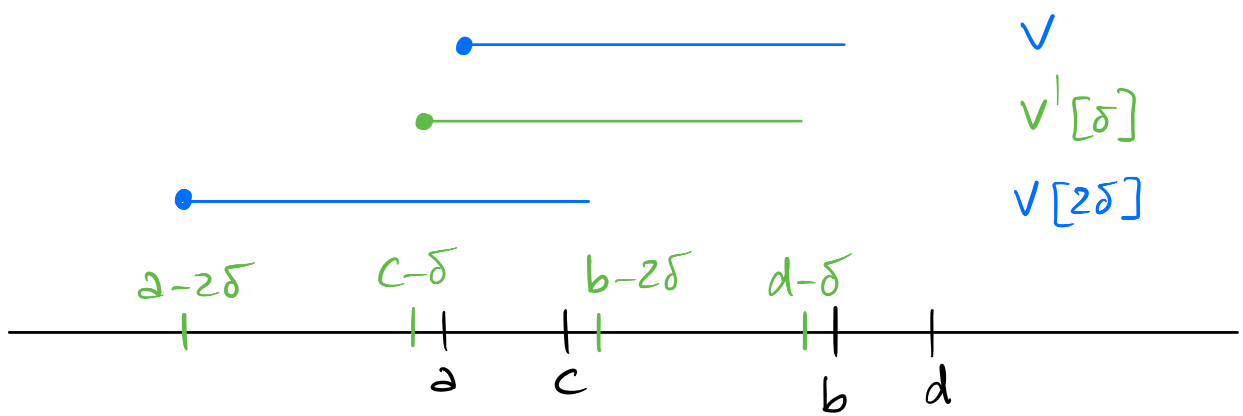
$V = \mathbb{F}[a,b)$

$V' = \mathbb{F}[c,d)$



Suppose first that $a \le c < b \le d$. Then V and V' are δ-interleaved in two cases:

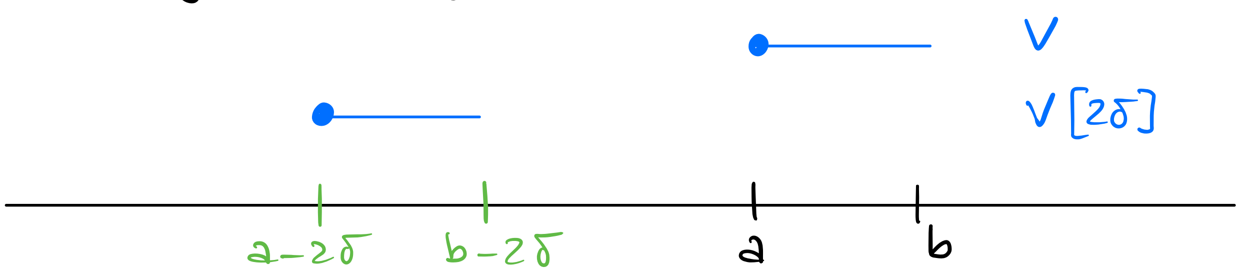
a) Suppose that $\sigma_{2\delta} \neq 0$. Then f and g exist with $g[\delta] \circ f = \sigma_{2\delta}$ if and only if
$$\begin{cases} a - 2\delta \leq c - \delta \leq a \\ b - 2\delta \leq d - \delta \leq b \end{cases}$$



The inequalities $a - 2\delta \leq c - \delta$ and $b - 2\delta \leq d - \delta$ are automatic.
Hence we need to impose that $\delta \geq c - a$ and $\delta \geq d - b$, that is,
$\delta \geq \max\{c - a, d - b\}$.

If we assume instead that $\sigma'_{2\delta} \neq 0$, then by symmetry f and g exist with $f[\delta] \circ g = \sigma'_{2\delta}$ if and only if $\delta \geq \max\{c - a, d - b\}$ as well.

b) It can also happen that $\sigma_{2\delta} = 0$. In this case, we can pick $f = 0$ and $g = 0$ and $g[\delta] \circ f = \sigma_{2\delta}$ holds. We say that V is <u>δ-short</u>.



The case $\sigma_{2\delta} = 0$ occurs when $b - 2\delta \leq a$, that is, $\delta \geq \frac{1}{2}(b-a)$.

In conclusion, V and V' are δ-interleaved if and only if either

- $\sigma_{2\delta} \neq 0$ and $\sigma'_{2\delta} \neq 0$ and $\delta \geq \max\{c-a, d-b\}$, or
- $\sigma_{2\delta} = 0$ and $\sigma'_{2\delta} \neq 0$ and $\delta \geq \max\{c-a, d-b\}$, or
- $\sigma_{2\delta} \neq 0$ and $\sigma'_{2\delta} = 0$ and $\delta \geq \max\{c-a, d-b\}$, or
- $\sigma_{2\delta} = 0$ and $\sigma'_{2\delta} = 0$.

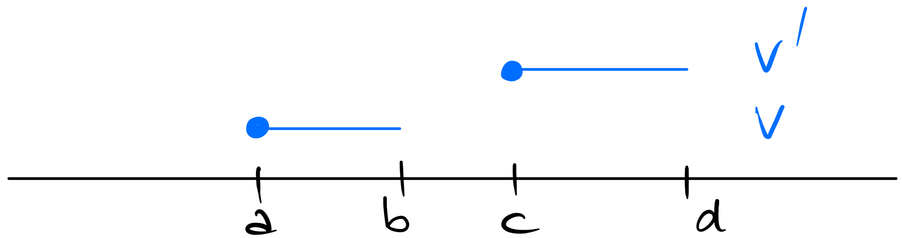
Equivalently, either

- $\delta \geq \max\{c-a, d-b\}$, or
- $\delta \geq \frac{1}{2}(b-a)$ and $\delta \geq \frac{1}{2}(d-c)$, i.e., $\delta \geq \max\left\{\frac{b-a}{2}, \frac{d-c}{2}\right\}$.

Therefore, if $a \leq c < b \leq d$ then

$$d_{int}(\mathbb{F}[a,b), \mathbb{F}[c,d)) = \min\{\max\{c-a, d-b\}, \max\{\frac{b-a}{2}, \frac{d-c}{2}\}\}.$$

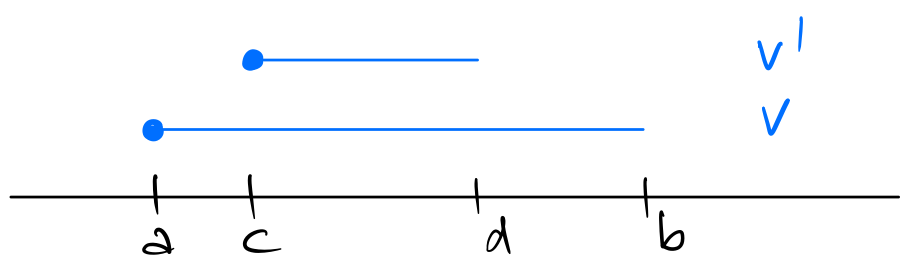
Now let us assume that $a < b \leq c < d$:



In this case $g: V' \to V[\delta]$ is necessarily zero and hence V and V' are δ-interleaved if and only if $\delta \geq \max\{\frac{b-a}{2}, \frac{d-c}{2}\}$. Thus,

$$d_{int}(\mathbb{F}[a,b), \mathbb{F}[c,d)) = \max\{\frac{b-a}{2}, \frac{d-c}{2}\} \quad \text{if } a < b \leq c < d.$$

Finally, suppose that $a \leq c < d \leq b$:



Then $d-c < b-a$, so $\max\{\frac{b-a}{2}, \frac{d-c}{2}\} = \frac{b-a}{2}$. In this case,

$$d_{int}(\mathbb{F}[a,b), \mathbb{F}[c,d)) = \max\{c-a, b-d\}.$$

To prove this claim, note that if $\frac{b-a}{2} < c-a$ then $b < 2c-a$ and hence $b < 2d-a$, from which it follows that $\frac{b-a}{2} > b-d$. Consequently $\frac{b-a}{2} > \max\{c-a, b-d\}$. By symmetry, if $c \leq a < b \leq d$ we find that $\frac{d-c}{2} > \max\{a-c, d-b\}$. In conclusion, in all cases,

$$\boxed{d_{int}(\mathbb{F}[a,b), \mathbb{F}[c,d)) = \min\left\{\max\{|c-a|, |d-b|\}, \max\left\{\frac{b-a}{2}, \frac{d-c}{2}\right\}\right\}}$$