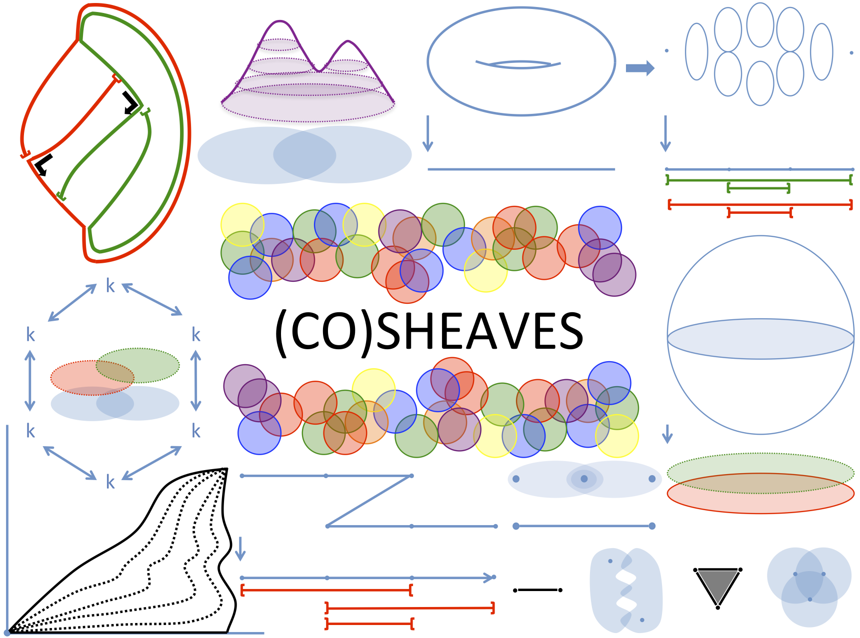
Persistent Homology via Cellular (Co)Sheaves

Justin Curry

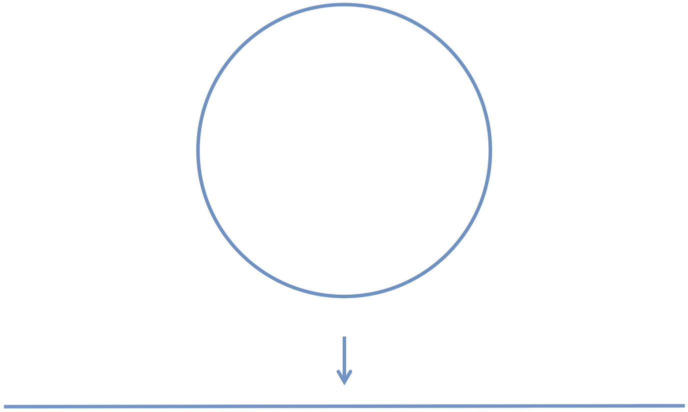
University of Pennsylvania

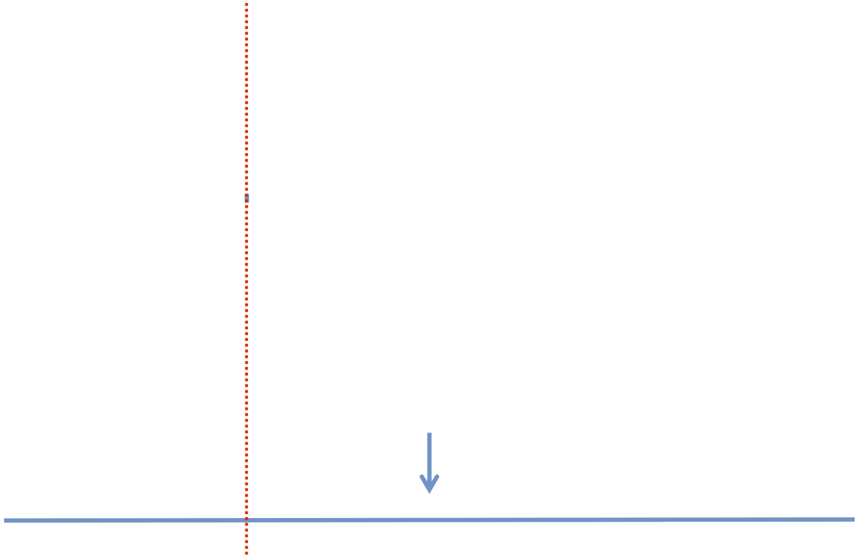
ACAT 2013 - Bremen

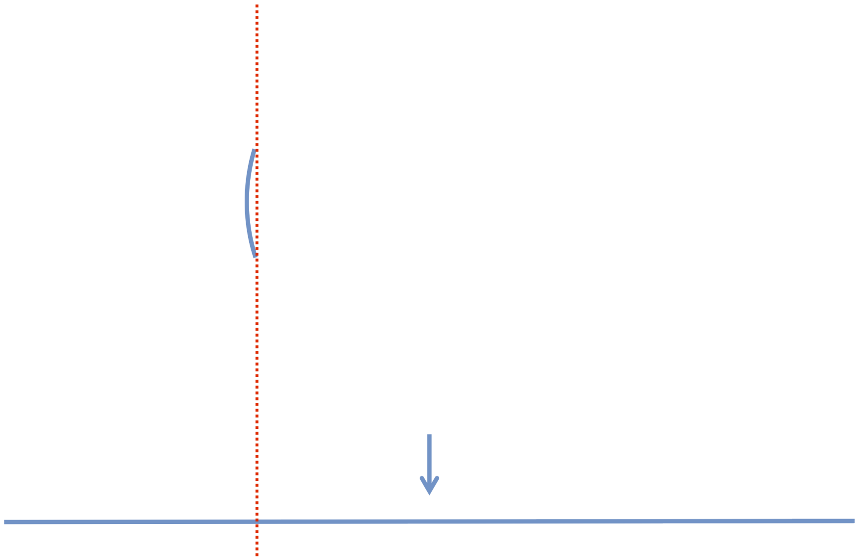

(CO)SHEAVES
k
k
k
k
k
k

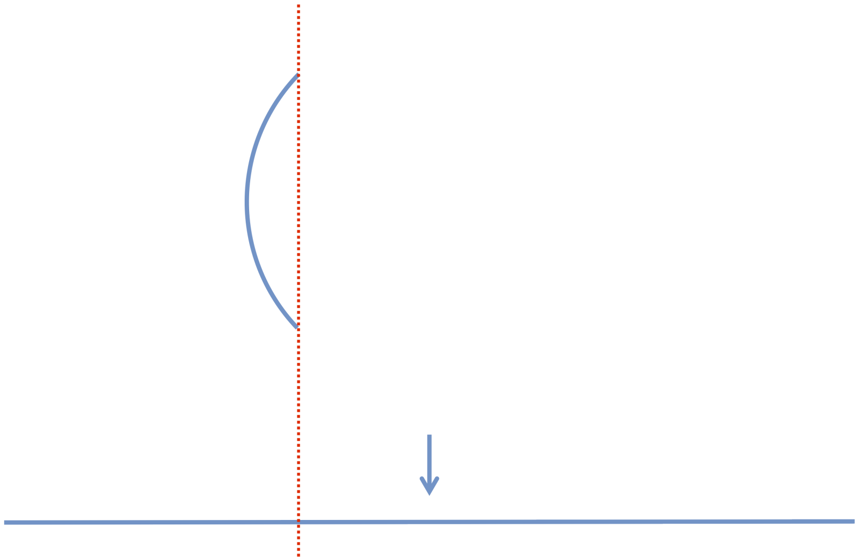
Families of Spaces

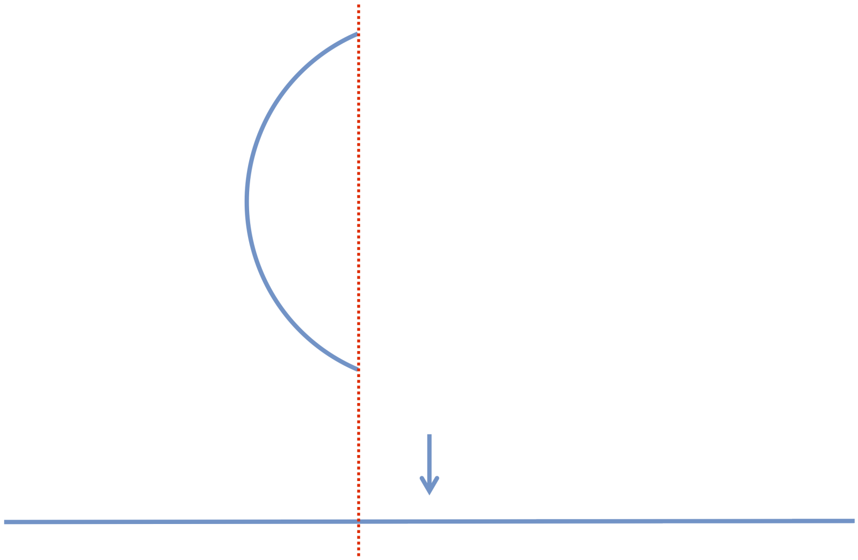
- Given $h : Y \to \mathbb{R}$ we can study the sub-level sets $Y_{\leqslant t} := h^{-1}(-\infty, t]$.

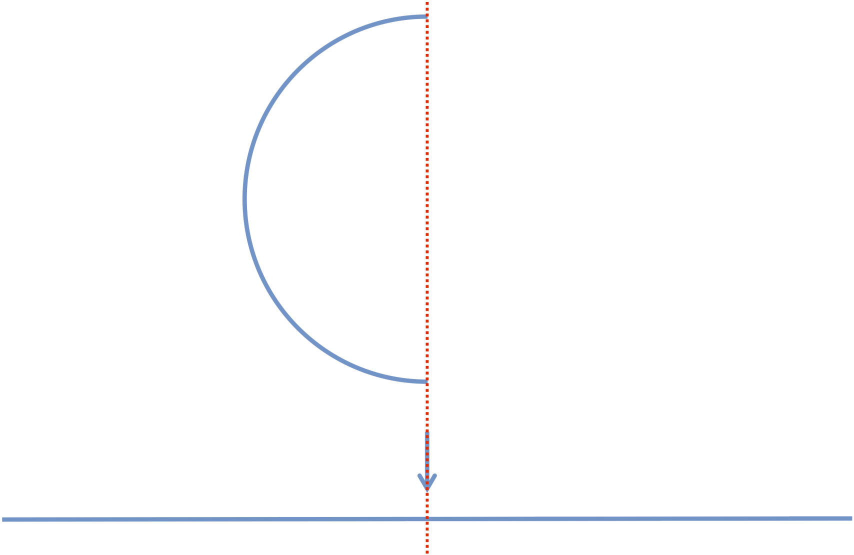


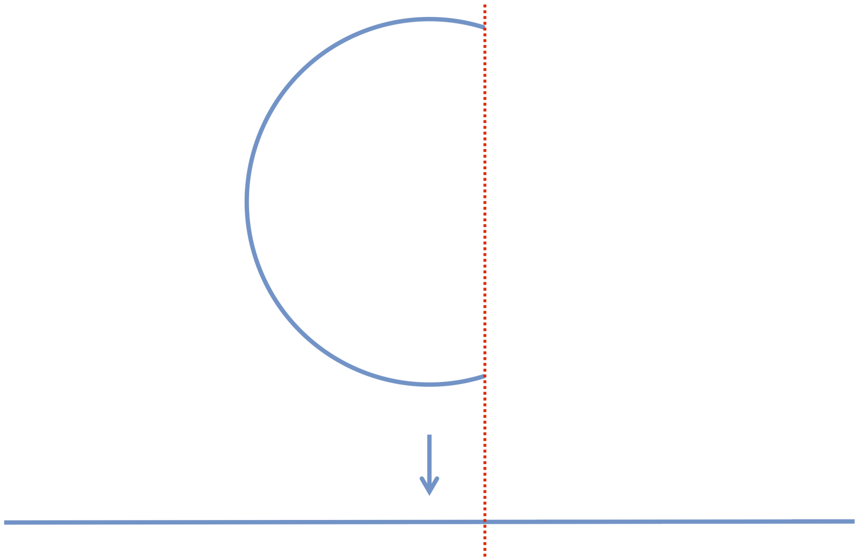


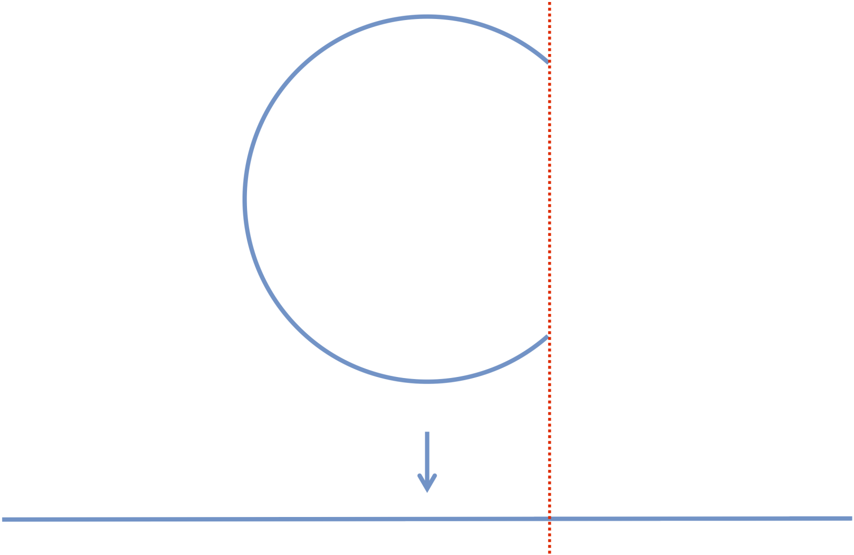


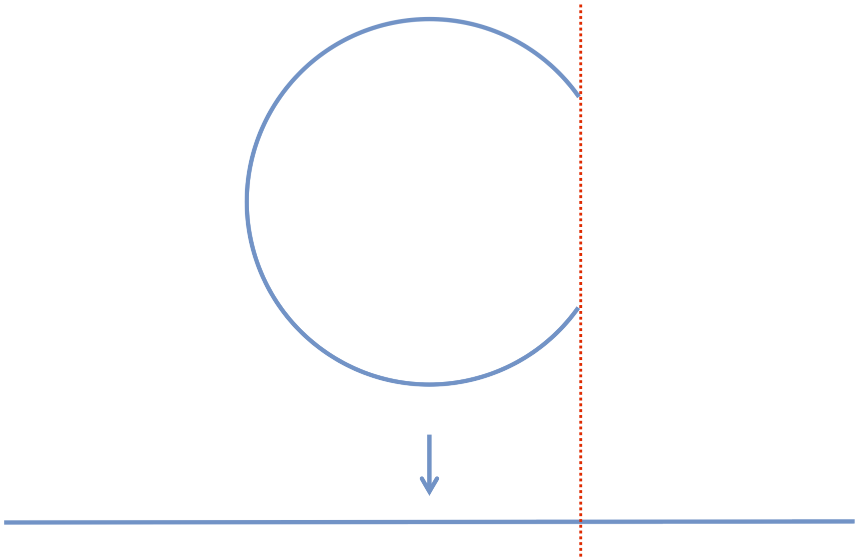


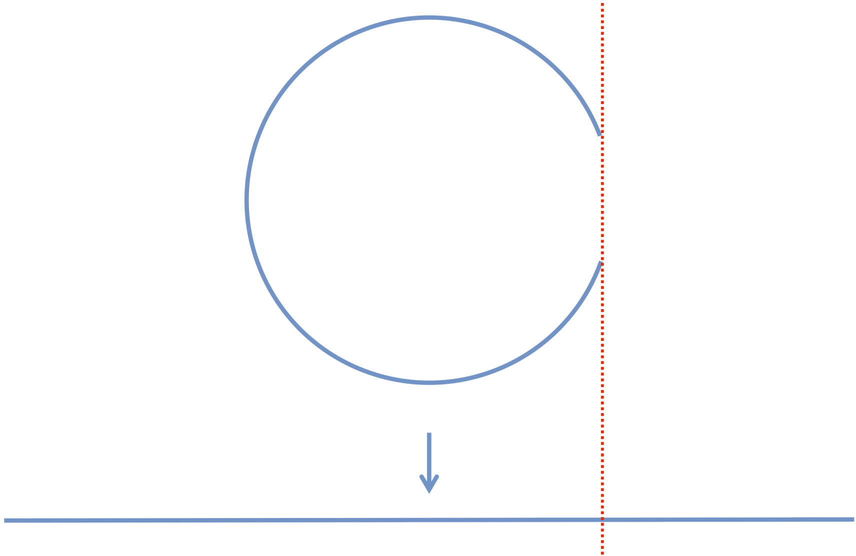


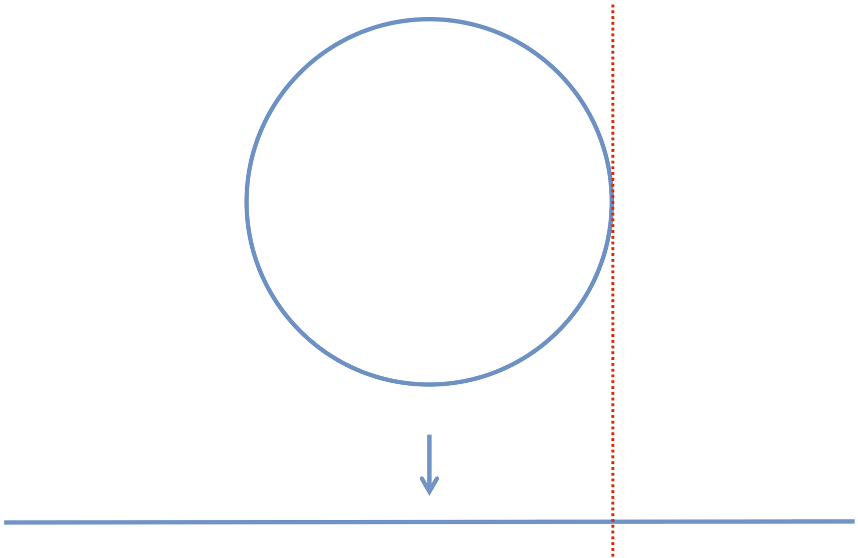












Families of Spaces and Data

- To each number $t \in \mathbb{R}$, we have a space $Y_{\leqslant t} := h^{-1}(-\infty, t]$:

$$\begin{tikzcd} t \arrow[r, squiggly] \arrow[d, "\leqslant"'] & Y_{\leqslant t} \arrow[d, "f_{s,t}"] \\ s \arrow[r, squiggly] & Y_{\leqslant s} \end{tikzcd}$$

- For $t \leqslant r \leqslant s$, we have $f_{s,t} = f_{s,r} \circ f_{r,t}$, i.e.

Families of Spaces and Data

- To each number $t \in \mathbb{R}$, we have a space $Y_{\leqslant t} := h^{-1}(-\infty, t]$:

$$
\begin{tikzcd}
t \arrow[r, rightsquigarrow] \arrow[d, "\leqslant"'] & Y_{\leqslant t} \arrow[d, "f_{s,t}"] \\
s \arrow[r, rightsquigarrow] & Y_{\leqslant s}
\end{tikzcd}
$$

- For $t \leqslant r \leqslant s$, we have $f_{s,t} = f_{s,r} \circ f_{r,t}$, i.e.

$F : (\mathbb{R}, \leqslant) \to \mathbf{Top}$ is a **functor**.

Families of Spaces and Data

- To each number $t \in \mathbb{R}$, we have a space $Y_{\leqslant t} := h^{-1}(-\infty, t]$:

$$\begin{tikzcd} t \arrow[r, rightsquigarrow] \arrow[d, "\leqslant"'] & Y_{\leqslant t} \arrow[d, "f_{s,t}"] \\ s \arrow[r, rightsquigarrow] & Y_{\leqslant s} \end{tikzcd}$$

- For $t \leqslant r \leqslant s$, we have $f_{s,t} = f_{s,r} \circ f_{r,t}$, i.e.

$F : (\mathbb{R}, \leqslant) \to \mathbf{Top}$ is a **functor**.

- **Homology** in degree i with field coefficients is a functor

$$H_i(-; k) : \mathbf{Top} \to \mathbf{Vect}_k$$

Families of Spaces and Data

- To each number $t \in \mathbb{R}$, we have a space $Y_{\leqslant t} := h^{-1}(-\infty, t]$:

$$\begin{tikzcd} t \arrow[r, rightsquigarrow] \arrow[d, "\leqslant"'] & Y_{\leqslant t} \arrow[d, "f_{s,t}"] \\ s \arrow[r, rightsquigarrow] & Y_{\leqslant s} \end{tikzcd}$$

- For $t \leqslant r \leqslant s$, we have $f_{s,t} = f_{s,r} \circ f_{r,t}$, i.e.

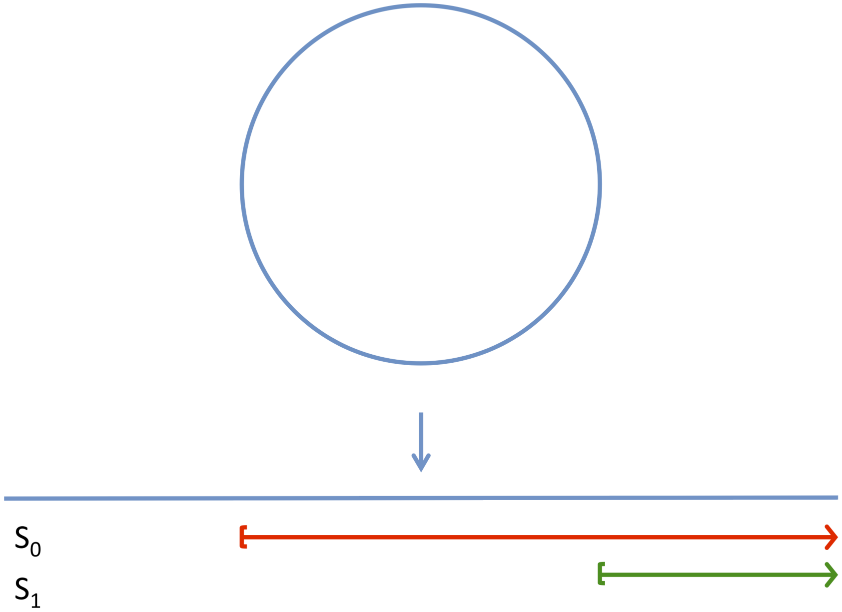
$$F : (\mathbb{R}, \leqslant) \to \mathbf{Top} \text{ is a } \textbf{functor}.$$

- **Homology** in degree i with field coefficients is a functor

$$H_i(-; k) : \mathbf{Top} \to \mathbf{Vect}_k$$

- (Sub-level set) **Persistent Homology** is the composition of these functors

$$S_i := H_i(-; k) \circ F$$

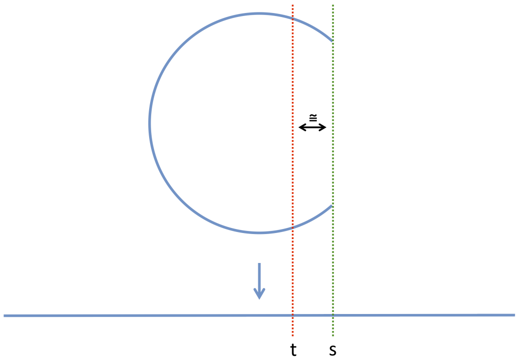

S_0
S_1

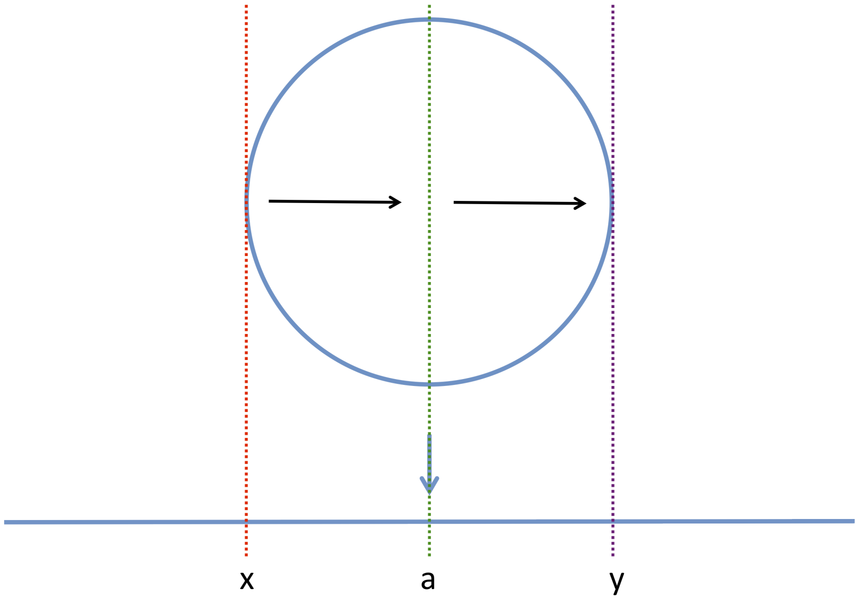
Compressed Representation

- **Question:** What is the smallest poset that contains all the information of the map h?

Compressed Representation

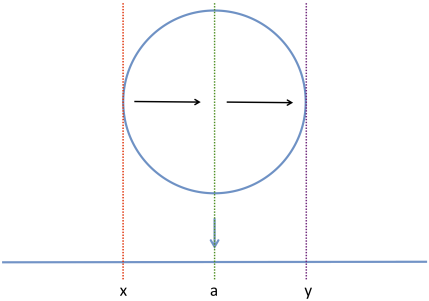
- **Question:** What is the smallest poset that contains all the information of the map h?
- **Morse Theory Tells Us:** If the interval $[t, s]$ contains no critical values, then $Y_{\leqslant t} \cong Y_{\leqslant s}$




x
a
y

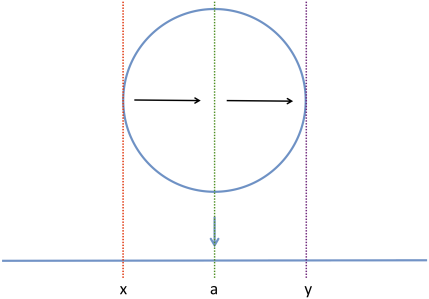
Compressed Representation

- Define a quotient poset $q : \mathbb{R} \to P = \mathbb{R} / \sim$ where $t \sim s$ iff for every $r \in [t, s]$, $Y_{\leqslant t} \to Y_{\leqslant r}$ is a homeomorphism, i.e. is an invertible continuous map.
- $\{x \leqslant a \leqslant y\} \cong P$
- $F : \mathbb{R} \to \mathbf{Top}$ is actually $G : P \to \mathbb{R}$ precomposed with q, i.e. $F = q^* G$.



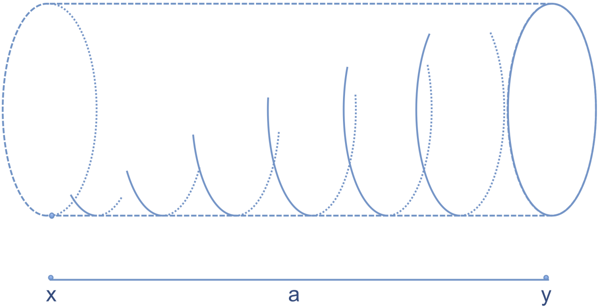
Compressed Representation

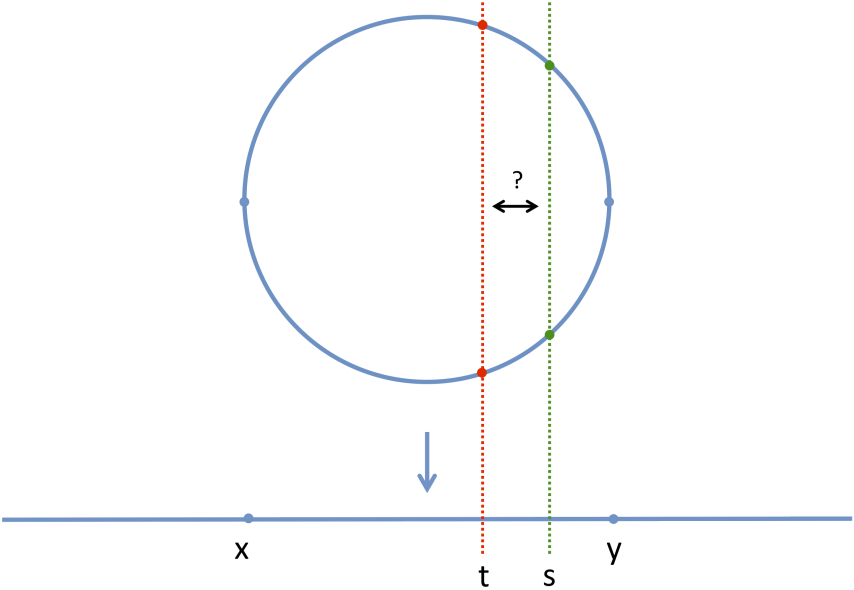
- **Moral:** The Morse condition allowed us to work with a smaller poset in a loss-free way.



Motivating Level-Set Persistence

- **Problem:** Sub-level persistence $h : Y \to \mathbb{R}$ depends on order of $\mathbb{R}$, which doesn't generalize to (multi-dimensional) persistence over $\mathbb{R}^2$, for example.
- **Solution:** Do level-set persistence!



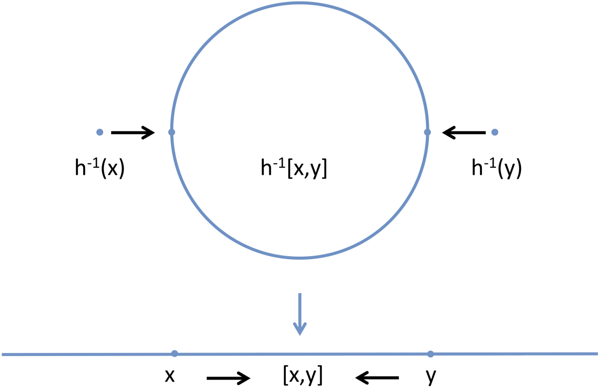


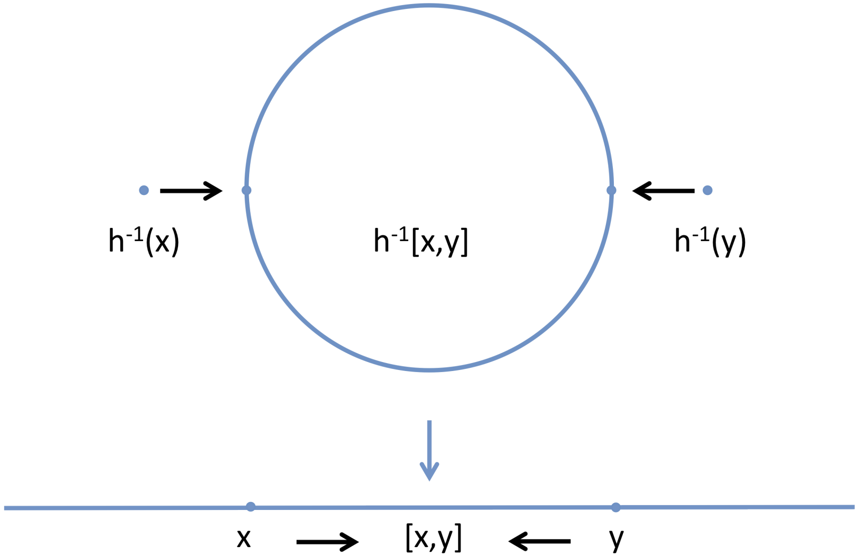
HOW DO WE RELATE THE FIBERS?

3 Ways to "Connect the Fibers"

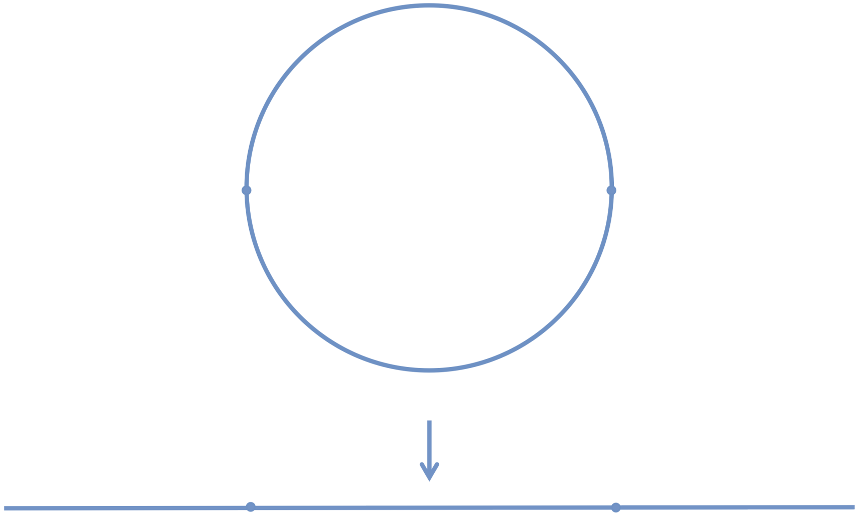
(1) **Closed Cells:** For $h : Y \to \mathbb{R}$, pick a mesh $\cdots < x_{i-1} < x_i < x_{i+1} < \cdots$, then get a **zigzag** of spaces

$$\cdots \leftarrow h^{-1}(x_i) \to h^{-1}([x_i, x_{i+1}]) \leftarrow h^{-1}(x_{i+1}) \to \cdots$$





$h^{-1}(x)$
$h^{-1}[x,y]$
$h^{-1}(y)$
x
[x,y]
y



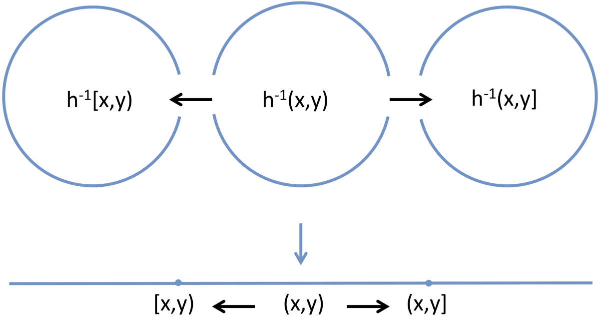
$$H_0(-;k) \qquad k \longrightarrow k \longleftarrow k$$

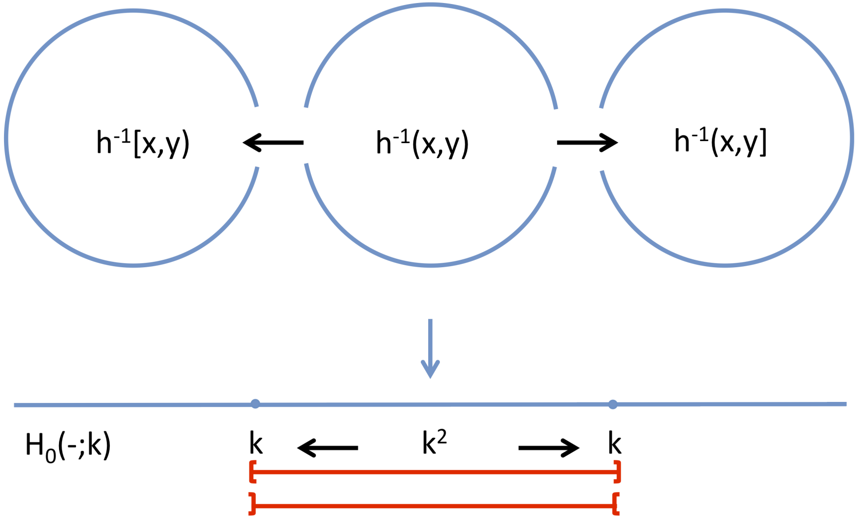
$$H_1(-;k) \qquad 0 \longrightarrow k \longleftarrow 0$$

3 Ways to "Connect the Fibers"

(2) **Open Stars:** For $h : Y \to \mathbb{R}$, pick a mesh $\cdots < x_{i-1} < x_i < x_{i+1} < \cdots$, then get a **zigzag** of spaces

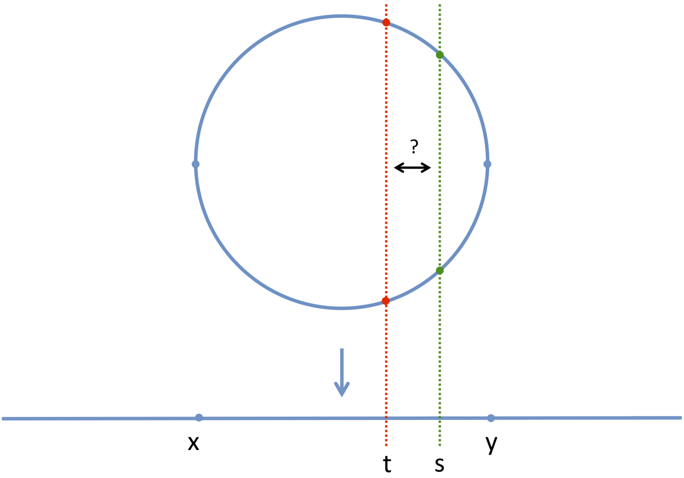
$$\cdots \leftarrow h^{-1}((x_{i-1}, x_i)) \to h^{-1}((x_{i-1}, x_{i+1})) \leftarrow h^{-1}((x_i, x_{i+1})) \to \cdots$$

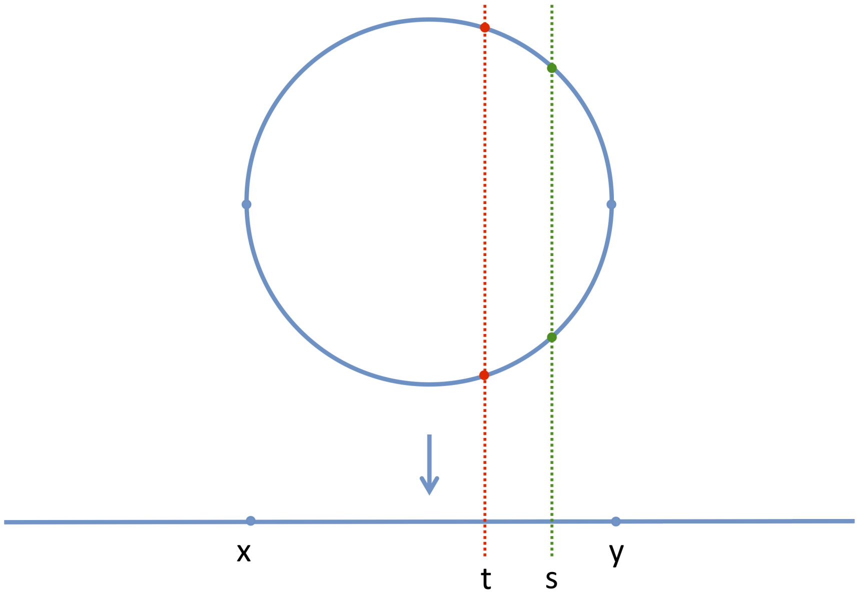


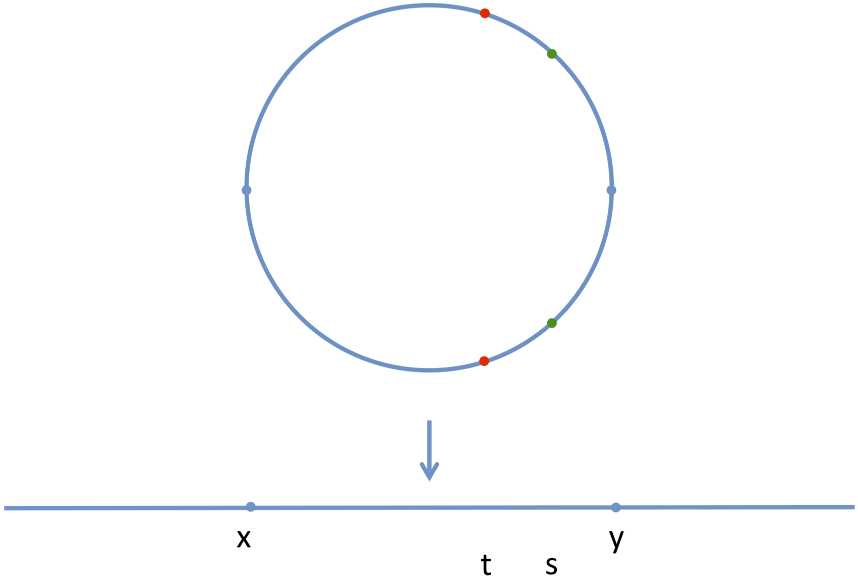

$h^{-1}[x,y)$
$h^{-1}(x,y)$
$h^{-1}(x,y]$
$H_0(-;k)$
k
k^2
k

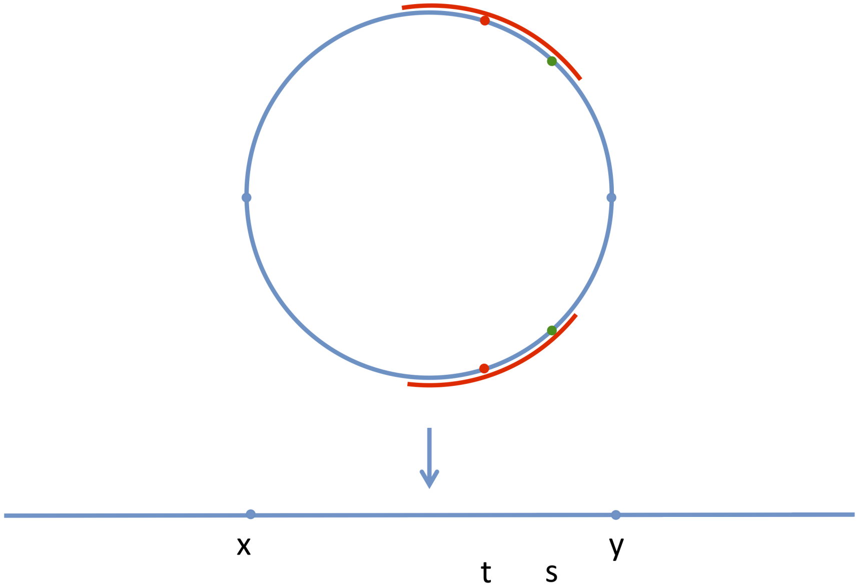
3 Ways to "Connect the Fibers"

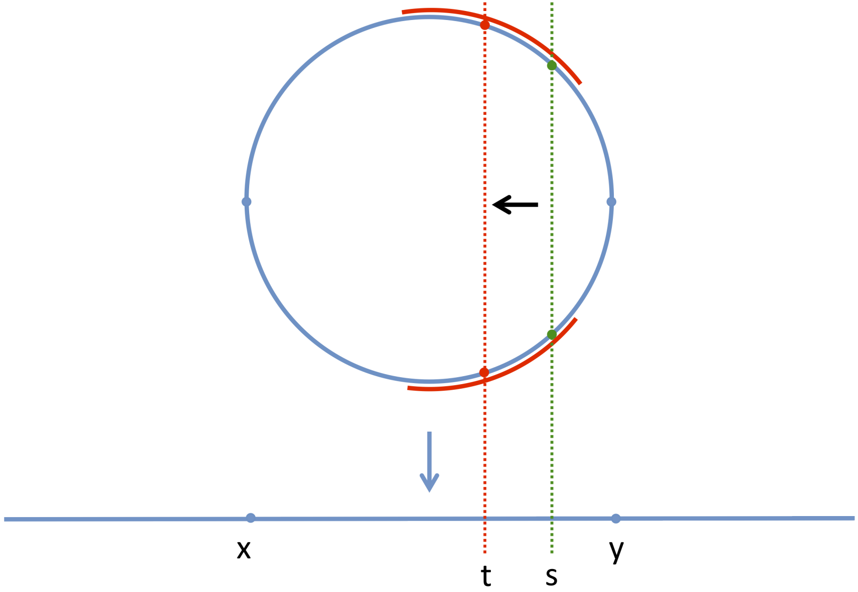
(3) **Actually do level-set persistence!**



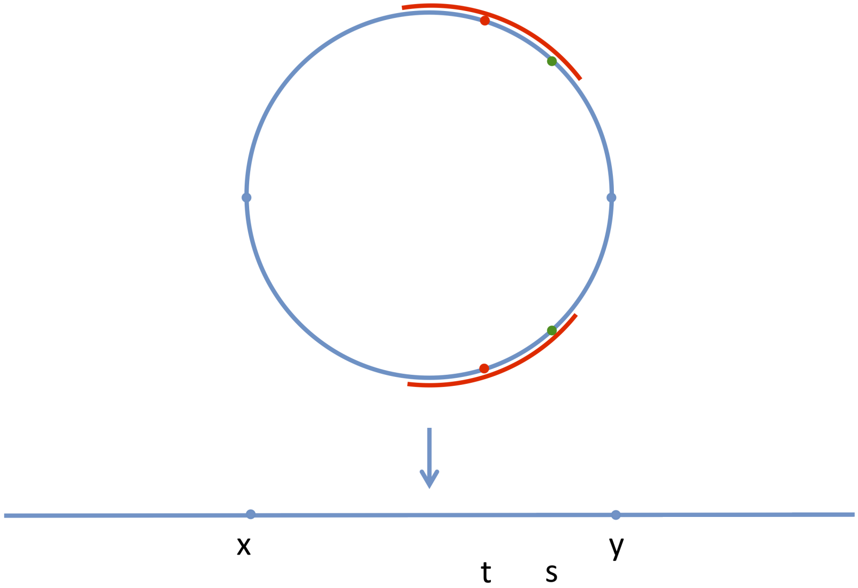

x
t
s
y

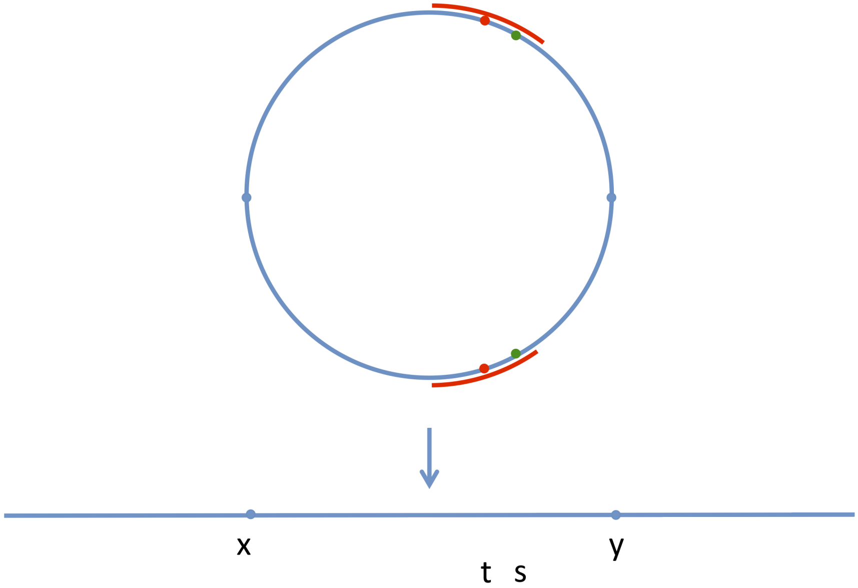

x
t
s
y

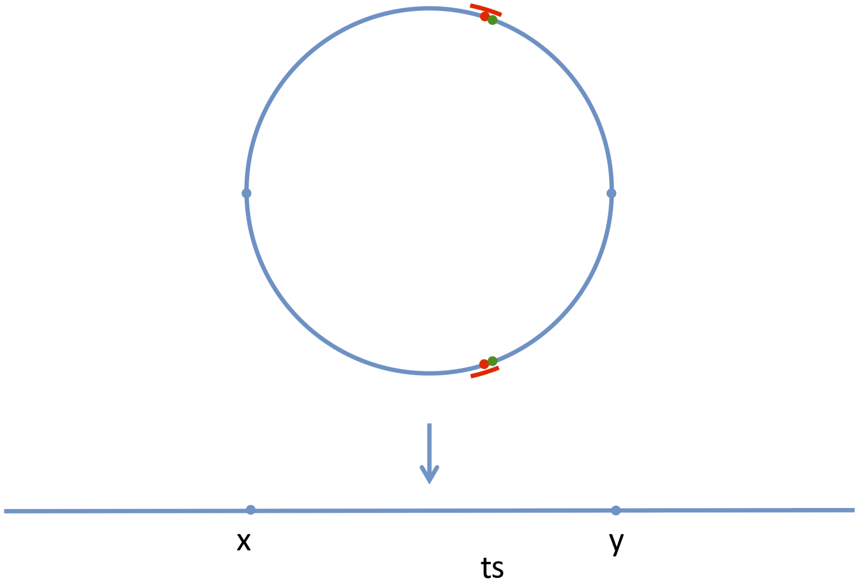

x
t
s
y



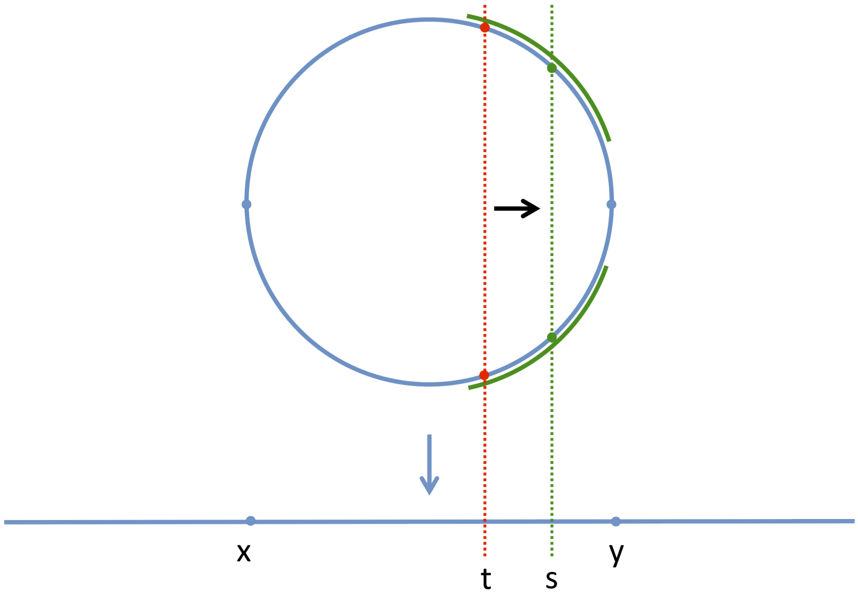
x
t
s
y

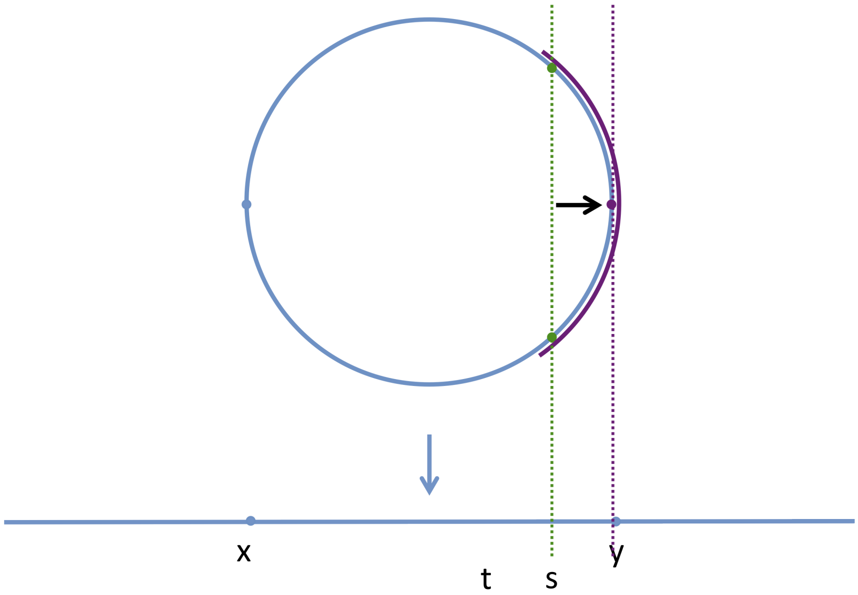

x
t
s
y


x
y
t s



x
y
ts


x
t
s
y


x
t
s
y

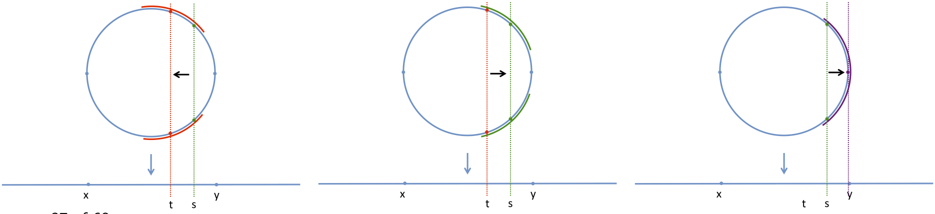
Level-set Persistence

What's the right indexing set (category)?

- For each $t \in (x, y)$ we have a space Y_t
- For each $s \in (x, y)$ there is a neighborhood U_t of Y_t that contains Y_s

$$Y_t \xhookrightarrow{\simeq} U_t \hookleftarrow Y_s$$

- Allows us to define an invertible map on homology between the fibers Y_s and Y_t
- But, there is only a map from Y_t to Y_y and from Y_t to Y_x.



Level-set Persistence

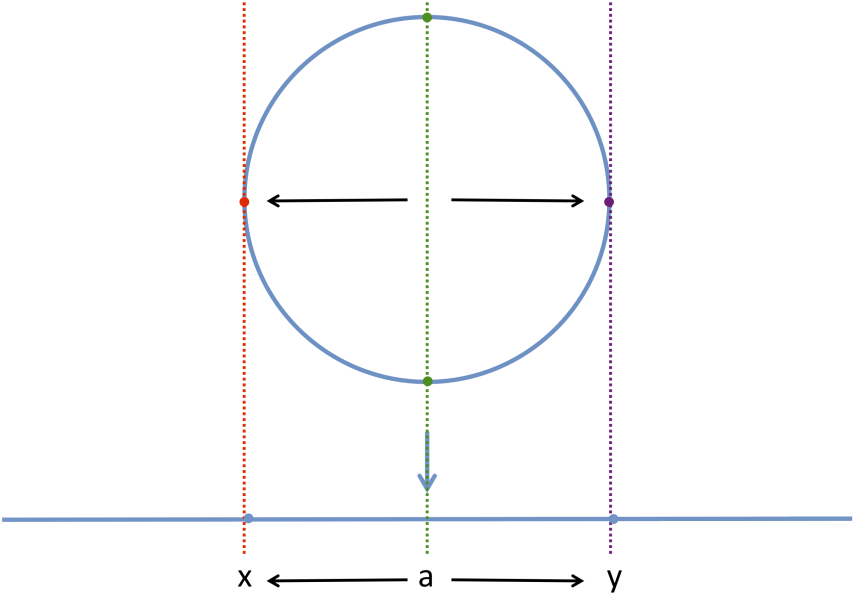
The Entrance Path Category

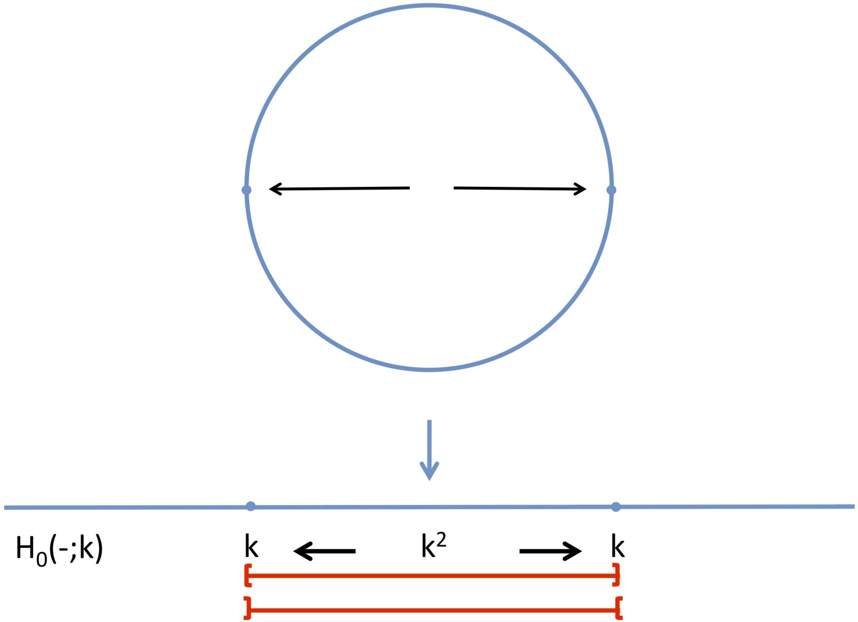
- To the cell complex $X := [x, y]$ with cells x, y and $a = (x, y)$ we associate a pre-ordered set $\mathfrak{Entr}(X)$ (poset w/o anti-symmetry)
- This set has an element for every point in X, but with relations $t \rightsquigarrow s$ and $s \rightsquigarrow t$ for $t, s \in a$

$$x \leftsquigarrow t \rightsquigarrow y$$

- Defining an equivalence relation $t \sim s$ for all $t, s \in a$ yields the opposite of the **face relation poset**, i.e. X^{op}:

$$\begin{tikzcd} & a \arrow[dl] \arrow[dr] & \\ x & & y \end{tikzcd}$$


x
a
y



$H_0(-;k)$
k
k^2
k

Level-set Persistence

- Approach (3) and approach (2) are actually equivalent.

Level-set Persistence

- Approach (3) and approach (2) are actually equivalent.
- Approach (3) gives a definition for level-set persistence. Given $f : Y \to X$, for each i we have an assignment

$$\hat{\mathsf{F}}_i : \mathbf{Entr}(X) \simeq X^{op} \to \mathbf{Vect} \qquad t \in \sigma \subset X \mapsto H_i(f^{-1}(t); k)$$

Level-set Persistence

- Approach (3) and approach (2) are actually equivalent.
- Approach (3) gives a definition for level-set persistence. Given $f : Y \to X$, for each i we have an assignment

$$\hat{\mathsf{F}}_i : \mathbf{Entr}(X) \simeq X^{op} \to \mathbf{Vect} \qquad t \in \sigma \subset X \mapsto H_i(f^{-1}(t); k)$$

- Approach (3) generalizes to arbitrary dimensions and maps $f : Y \to X$ as long as f is a proper **stratified map**, i.e. $f : Y \to X$ admits a decomposition $X = \cup X_\sigma$ into connected manifolds (a **stratification**) where $f_\sigma : f^{-1}(X_\sigma) \to X_\sigma$ is a fiber bundle.

Level-set Persistence

- Approach (3) and approach (2) are actually equivalent.
- Approach (3) gives a definition for level-set persistence. Given $f: Y \to X$, for each i we have an assignment

$$\hat{\mathsf{F}}_i : \mathbf{Entr}(X) \simeq X^{op} \to \mathbf{Vect} \qquad t \in \sigma \subset X \mapsto H_i(f^{-1}(t); k)$$

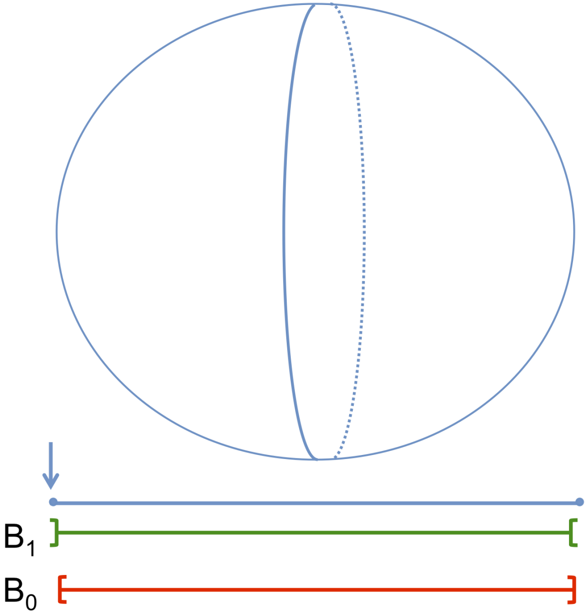
- Approach (3) generalizes to arbitrary dimensions and maps $f: Y \to X$ as long as f is a proper **stratified map**, i.e. $f: Y \to X$ admits a decomposition $X = \cup X_\sigma$ into connected manifolds (a **stratification**) where $f_\sigma : f^{-1}(X_\sigma) \to X_\sigma$ is a fiber bundle.
- Stratified maps are more general than triangulable maps, e.g.

$$f: \mathbb{R}^2 \to \mathbb{R}^2 \qquad (x, y) \mapsto (x, xy)$$

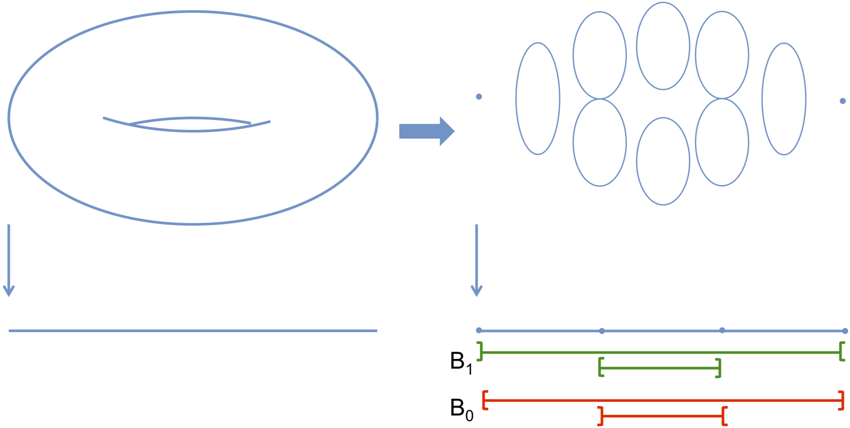
is NOT triangulable.

THESE ARE EXAMPLES OF

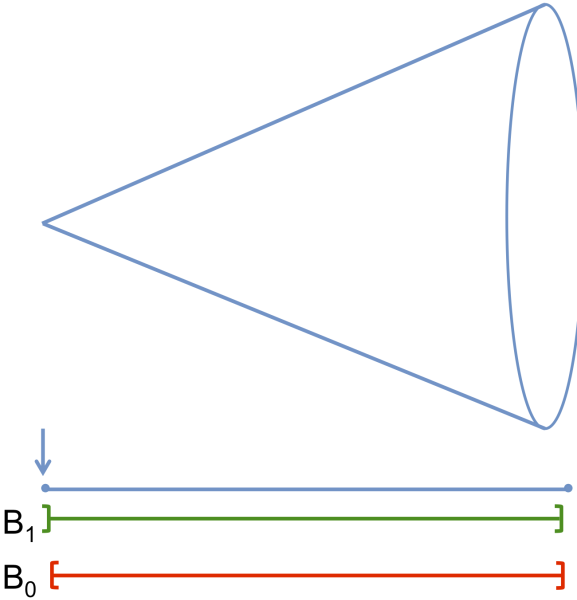
(CO)SHEAVES



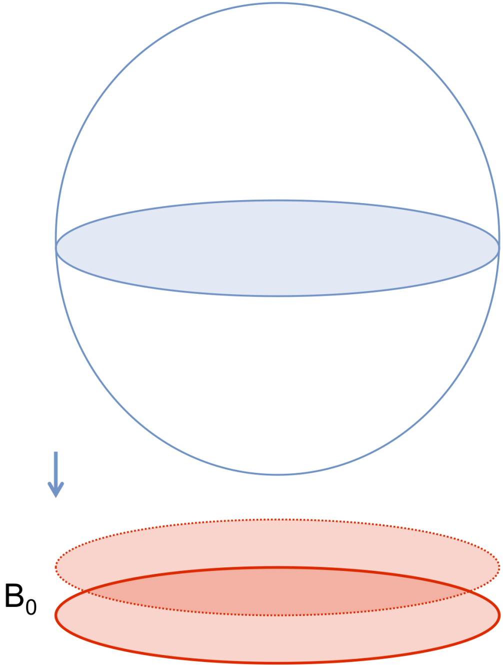
B_1
B_0



B_1
B_0



B_1
B_0


B0

Cosheaves have a homology theory

- Can compute **homology of X with coefficients in** $\hat{\mathsf{F}} : X^{op} \to \mathbf{Vect}$ by choosing local orientations $[\sigma : \tau] = < \sigma, \partial\tau > = \{\pm 1, 0\}$ and setting

$$\partial_i = \sum [\sigma^i : \tau^{i+1}] r_{\sigma,\tau}$$

$$\cdots \to \oplus \hat{\mathsf{F}}(\text{faces}) \to \oplus \hat{\mathsf{F}}(\text{edges}) \to \oplus \hat{\mathsf{F}}(\text{vertices}) \to 0$$

Cosheaves have a homology theory

- Can compute **homology of X with coefficients in** $\hat{\mathsf{F}} : X^{op} \to \mathbf{Vect}$ by choosing local orientations $[\sigma : \tau] = < \sigma, \partial\tau > = \{\pm 1, 0\}$ and setting

$$\partial_i = \sum [\sigma^i : \tau^{i+1}] r_{\sigma,\tau}$$

$$\cdots \to \oplus \hat{\mathsf{F}}(\text{faces}) \to \oplus \hat{\mathsf{F}}(\text{edges}) \to \oplus \hat{\mathsf{F}}(\text{vertices}) \to 0$$

- So What?

Cosheaves have a homology theory

- Can compute **homology of X with coefficients in** $\hat{\mathsf{F}} : X^{op} \to \mathbf{Vect}$ by choosing local orientations $[\sigma : \tau] = < \sigma, \partial\tau > = \{\pm 1, 0\}$ and setting

$$\partial_i = \sum [\sigma^i : \tau^{i+1}] r_{\sigma,\tau}$$

$$\cdots \to \oplus \hat{\mathsf{F}}(\text{faces}) \to \oplus \hat{\mathsf{F}}(\text{edges}) \to \oplus \hat{\mathsf{F}}(\text{vertices}) \to 0$$

- So What?
  - Every cosheaf over the real line has a **barcode decomposition** and cosheaf homology gives homology of barcodes as a special case.

Cosheaves have a homology theory

- Can compute **homology of X with coefficients in** $\hat{\mathsf{F}} : X^{op} \to \mathbf{Vect}$ by choosing local orientations $[\sigma : \tau] =< \sigma, \partial\tau >= \{\pm 1, 0\}$ and setting

$$\partial_i = \sum [\sigma^i : \tau^{i+1}] r_{\sigma,\tau}$$

$$\cdots \to \oplus \hat{\mathsf{F}}(\text{faces}) \to \oplus \hat{\mathsf{F}}(\text{edges}) \to \oplus \hat{\mathsf{F}}(\text{vertices}) \to 0$$

- So What?
  - Every cosheaf over the real line has a **barcode decomposition** and cosheaf homology gives homology of barcodes as a special case.
  - By additivity, computing **homology of barcodes** gives cosheaf homology over one-dimensional cell complexes X, e.g. $X = [0, 1]$.

Cosheaves have a homology theory

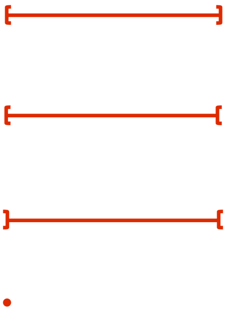
- Can compute **homology of X with coefficients in** $\hat{\mathsf{F}} : X^{op} \to \mathbf{Vect}$ by choosing local orientations $[\sigma : \tau] =< \sigma, \partial\tau >= \{\pm 1, 0\}$ and setting

$$\partial_i = \sum [\sigma^i : \tau^{i+1}] r_{\sigma,\tau}$$

$$\cdots \to \oplus \hat{\mathsf{F}}(\text{faces}) \to \oplus \hat{\mathsf{F}}(\text{edges}) \to \oplus \hat{\mathsf{F}}(\text{vertices}) \to 0$$

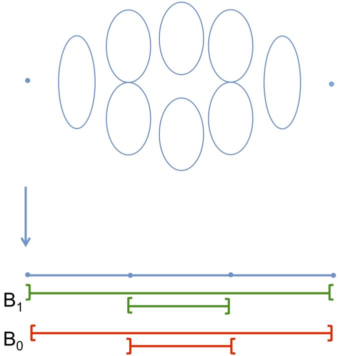
- So What?
  - Every cosheaf over the real line has a **barcode decomposition** and cosheaf homology gives homology of barcodes as a special case.
  - By additivity, computing **homology of barcodes** gives cosheaf homology over one-dimensional cell complexes X, e.g. $X = [0, 1]$.
  - Cosheaf homology well-defined in absence of barcodes.

$F(x) \leftarrow F(a) \rightarrow F(y)$	$H_0(X;F)$	$H_1(X;F)$
$k \leftarrow k \rightarrow k$	k	0
$k \leftarrow k \rightarrow 0$	0	0
$0 \leftarrow k \rightarrow 0$	0	k
$k \leftarrow 0 \rightarrow 0$	k	0



Borel-Moore homology $H_*^{BM}(-)$ (similar to compactly supported cohomology) is a topological homology theory that agrees with the homology of the barcodes.

Cosheaves on the Line

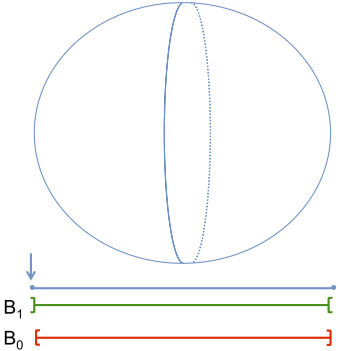


$$H_0(X;\hat{\mathsf{F}}_1) = k \qquad H_1(X;\hat{\mathsf{F}}_1) = k$$

$$H_0(X;\hat{\mathsf{F}}_0) = k \qquad H_1(X;\hat{\mathsf{F}}_0) = k$$

$$H_0(T) = k \;\; H_1(T) = k^2 \;\; H_2(T) = k$$

Cosheaves on the Line

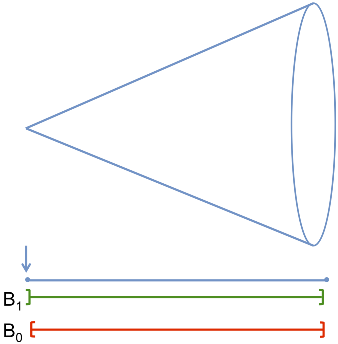


$H_0(X; \hat{\mathrm{F}}_1) = 0$ $\quad$ $H_1(X; \hat{\mathrm{F}}_1) = k$

$H_0(X; \hat{\mathrm{F}}_0) = k$ $\quad$ $H_1(X; \hat{\mathrm{F}}_0) = 0$

$$H_0(S^2) = k \ H_1(S^2) = 0 \ H_2(S^2) = k$$

Cosheaves on the Line



$$H_0(X;\hat{\mathsf{F}}_1) = 0 \qquad H_1(X;\hat{\mathsf{F}}_1) = 0$$

$$H_0(X;\hat{\mathsf{F}}_0) = k \qquad H_1(X;\hat{\mathsf{F}}_0) = 0$$

$$H_0(C) = k \;\; H_1(C) = 0 \;\; H_2(C) = 0$$

Total Homology from Level-Set Homology

Theorem (C. '12, Dey & Burghalea '11, Leray 1943)

Suppose

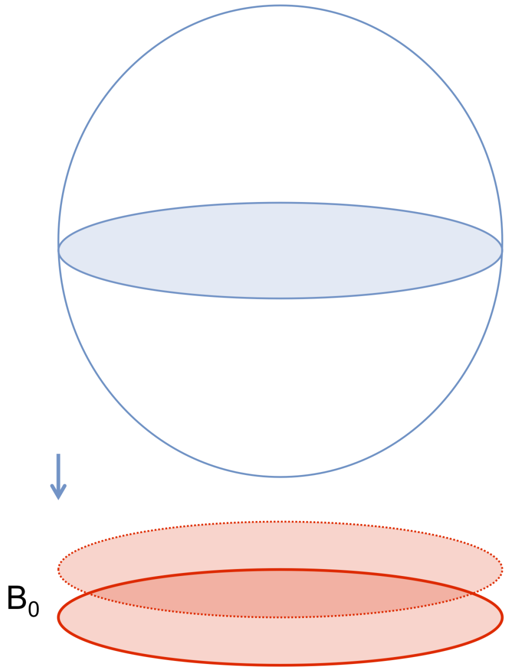
$$\begin{tikzcd} Y \arrow[d, "f"'] \\ X \subset \mathbb{R} \end{tikzcd}$$

is a stratified map with Y compact, then for each i consider the barcodes B_i associated to the cosheaf $\hat{\mathsf{F}}_i$, then

$$H_i(Y) \cong H_0(\mathbb{R}; \hat{\mathsf{F}}_i) \oplus H_1(\mathbb{R}; \hat{\mathsf{F}}_{i-1}) = H_0^{BM}(B_i) \oplus H_1^{BM}(B_{i-1}).$$

If X is 1D cell complex (graph or circle) then

$$H_i(Y) \cong H_0(\mathbb{R}; \hat{\mathsf{F}}_i) \oplus H_1(\mathbb{R}; \hat{\mathsf{F}}_{i-1}).$$


B_0

Spectral Sequences

Theorem (C.'12, Leray 1943)

Suppose

$$\begin{tikzcd} Y \arrow[d, "f"'] \\ X \end{tikzcd}$$

is a stratified map with Y compact and X stratified as a cell complex, then there is a spectral sequence converging to $H_*(Y)$.

$$\begin{array}{ccccc}
\oplus \hat{\mathsf{F}}_i(v) & \longleftarrow & \oplus \hat{\mathsf{F}}_i(e) & \longleftarrow & \oplus \hat{\mathsf{F}}_i(\sigma) \\
\oplus \hat{\mathsf{F}}_{i-1}(v) & \longleftarrow & \oplus \hat{\mathsf{F}}_{i-1}(e) & \longleftarrow & \oplus \hat{\mathsf{F}}_{i-1}(\sigma) \\
\vdots & & \vdots & & \vdots \\
\oplus \hat{\mathsf{F}}_0(v) & \longleftarrow & \oplus \hat{\mathsf{F}}_0(e) & \longleftarrow & \oplus \hat{\mathsf{F}}_0(\sigma)
\end{array}$$

Spectral Sequences

$$\begin{tikzcd}
H_0(X;\hat{\mathsf{F}}_i) & H_1(X;\hat{\mathsf{F}}_i) & H_2(X;\hat{\mathsf{F}}_i) & \cdots \\
H_0(X;\hat{\mathsf{F}}_{i-1}) & H_1(X;\hat{\mathsf{F}}_{i-1}) & H_2(X;\hat{\mathsf{F}}_{i-1}) \arrow[ull] & \cdots \\
\vdots & \vdots & \vdots & \\
H_0(X;\hat{\mathsf{F}}_1) & H_1(X;\hat{\mathsf{F}}_1) & H_2(X;\hat{\mathsf{F}}_1) & \cdots \\
H_0(X;\hat{\mathsf{F}}_0) & H_1(X;\hat{\mathsf{F}}_0) & H_2(X;\hat{\mathsf{F}}_0) \arrow[ull] & \cdots
\end{tikzcd}$$

Total Homology?

- So What?

Total Homology?

- So What?
  - Persistence is not supposed to compute total homology. It is supposed to give statistical signals for the topology of uncertain spaces.

Total Homology?

- So What?
  - Persistence is not supposed to compute total homology. It is supposed to give statistical signals for the topology of uncertain spaces.

Theorem (C. '13, Carlsson, de Silva & Morozov '09, Leray 1943)

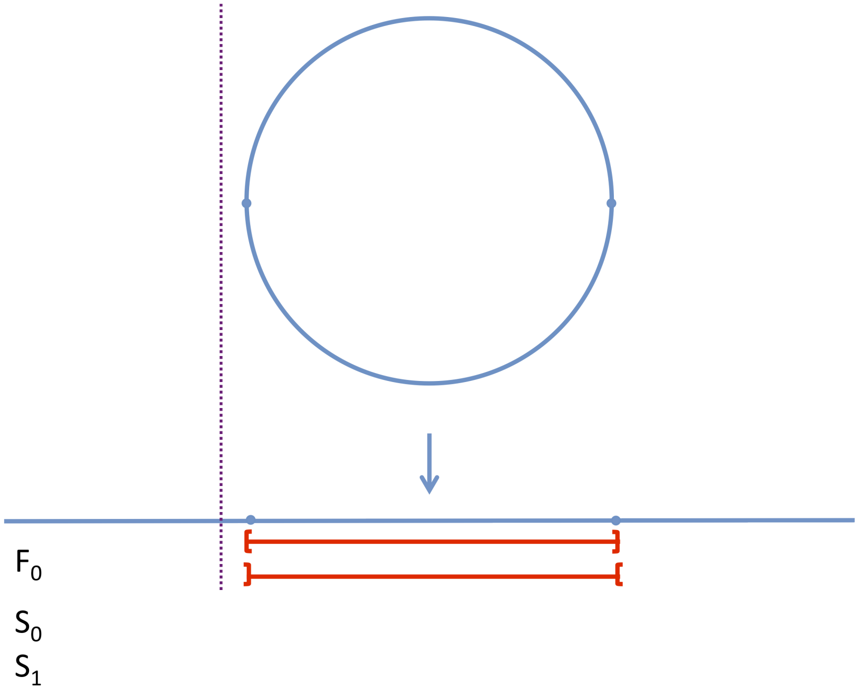
Level-Set Persistence determines Sublevel-set Persistence By defining

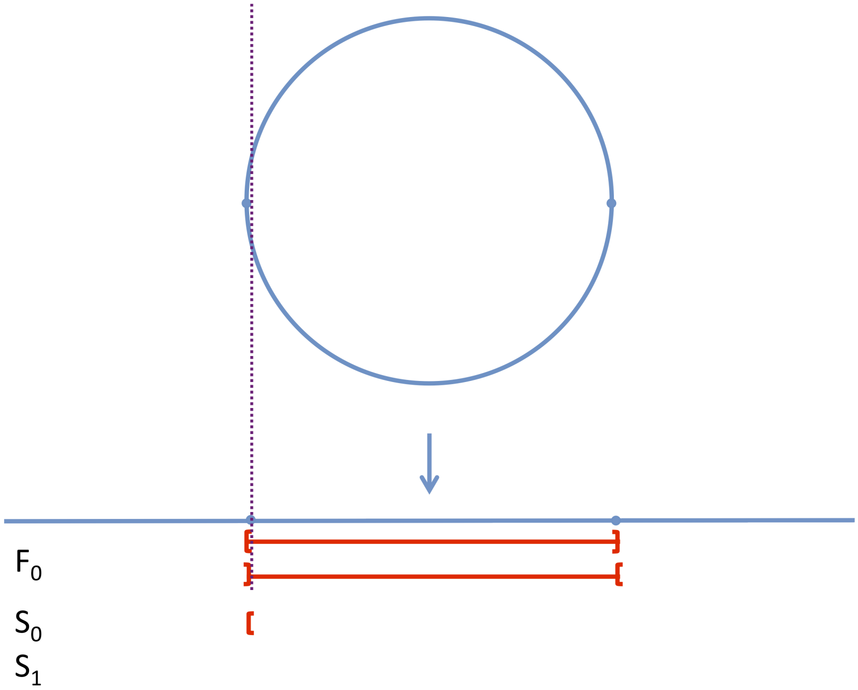
$$\begin{aligned} S_i(t) &:= H_0((-\infty, t]; \hat{\mathsf{F}}_i) \oplus H_1((-\infty, t]; \hat{\mathsf{F}}_{i-1}) \\ &\cong H_0^{BM}(B_i \cap (-\infty, t]) \oplus H_1^{BM}(B_{i-1} \cap (-\infty, t]) \end{aligned}$$

we get the homology of the entire sublevel set. Using covariance we get for $t \leqslant s$ the associated maps

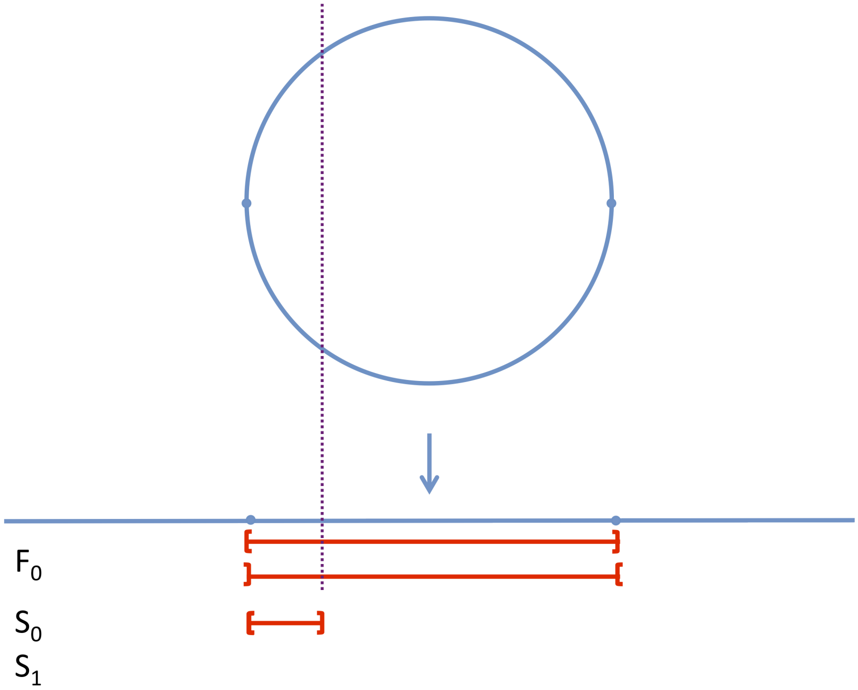
$$S_i(t) \to S_i(s)$$

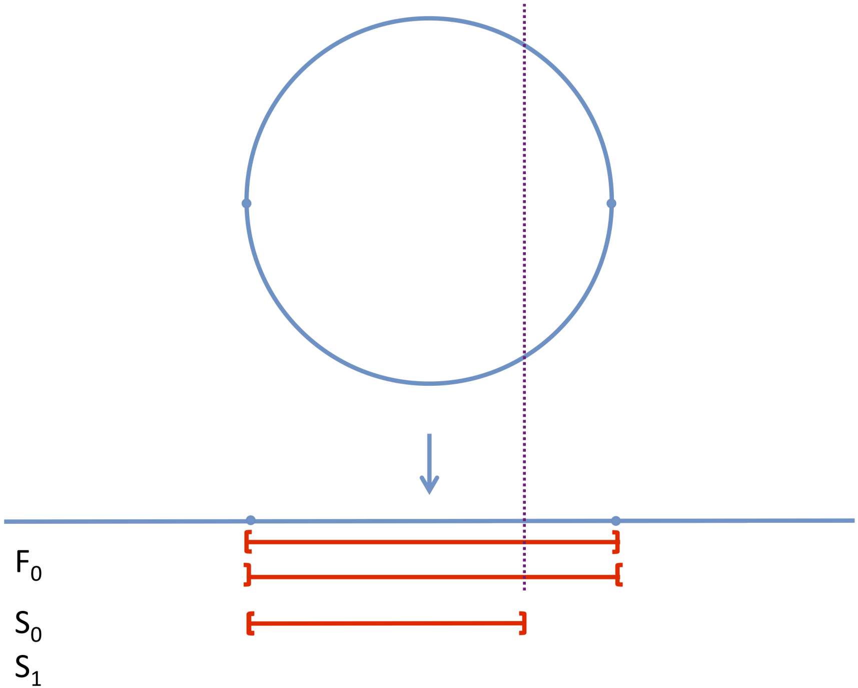
as before.

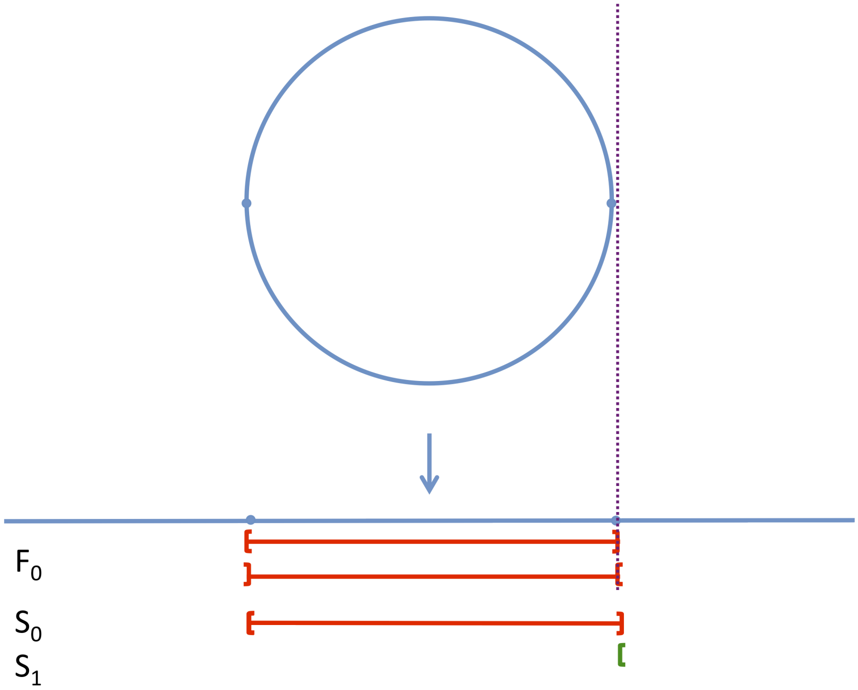

F_0
S_0
S_1

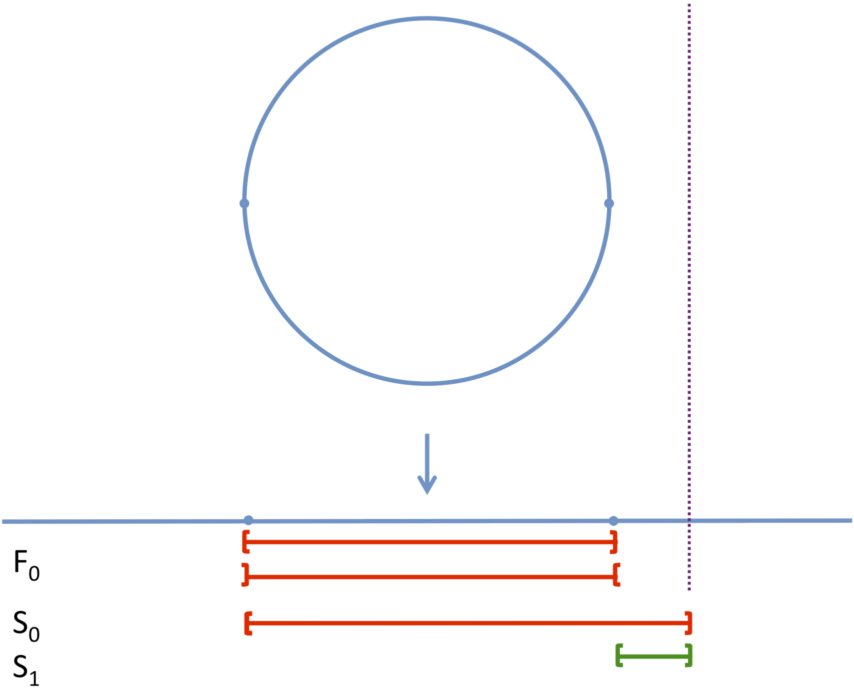


F_0
S_0
S_1

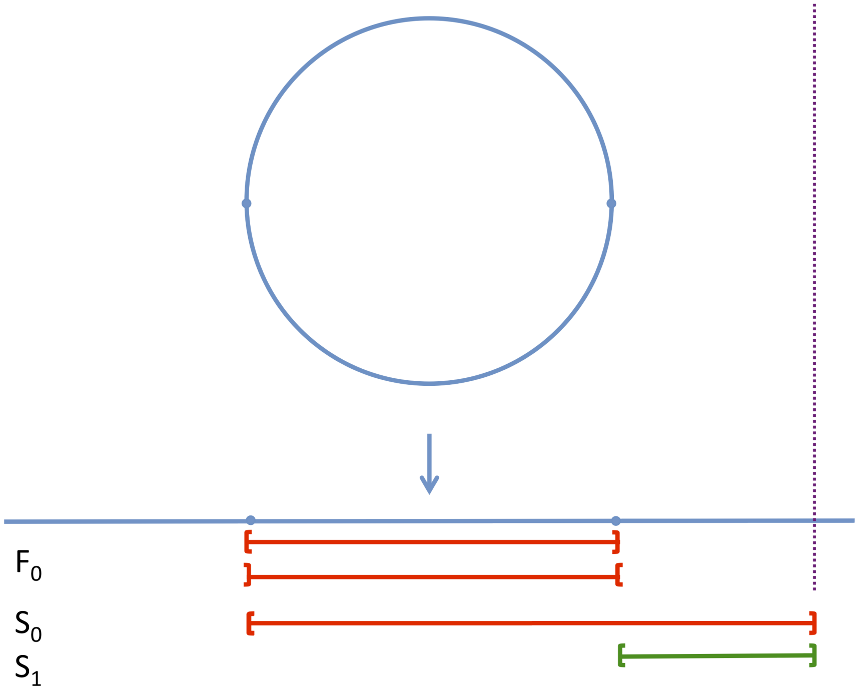

F_0
S_0
S_1


F_0
S_0
S_1


F_0
S_0
S_1



F_0
S_0
S_1



F_0
S_0
S_1

Remarks

- This is a purely (co)sheaf-theoretic treatment of Carlsson, de Silva & Morozov's **Pyramid Theorem**.

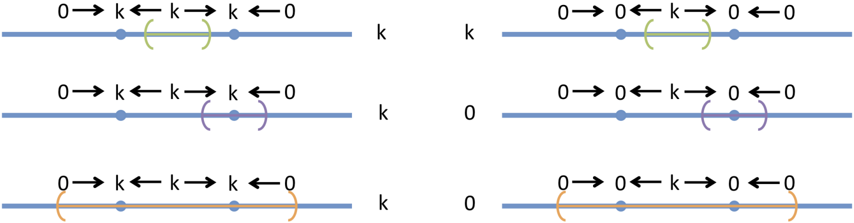
Remarks

- This is a purely (co)sheaf-theoretic treatment of Carlsson, de Silva & Morozov's **Pyramid Theorem**.
- It "generalizes" to a higher-dimensional pyramid theorem, with the only caveat that the associated spectral sequences are harder to compute.

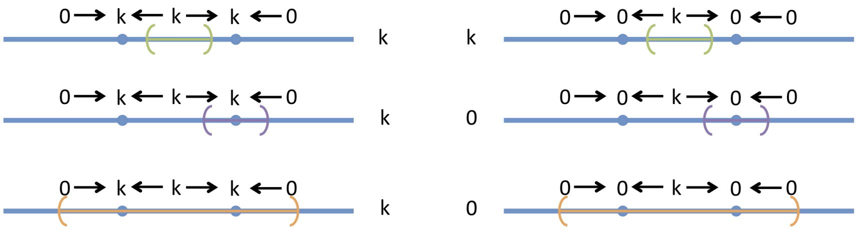
Remarks

- This is a purely (co)sheaf-theoretic treatment of Carlsson, de Silva & Morozov's **Pyramid Theorem**.
- It "generalizes" to a higher-dimensional pyramid theorem, with the only caveat that the associated spectral sequences are harder to compute.
- Any higher-dimensional attempt at persistence must interact with spectral sequences.

Remarks - Interleavings

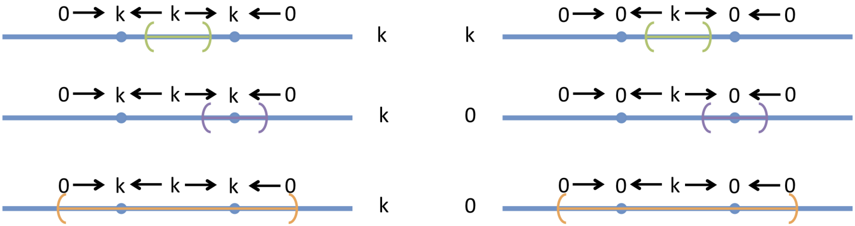


Remarks - Interleavings



- These diagrams define functors $\hat{\mathsf{F}}_i : \mathbf{Open}(X) \to \mathbf{Vect}$ by taking colimits of the diagram over an open set, called pre-cosheaves.

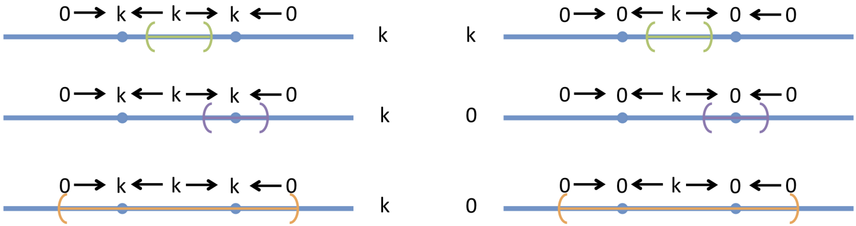
Remarks - Interleavings



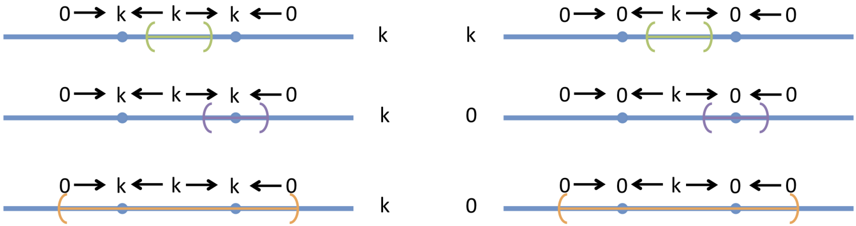
- These diagrams define functors $\hat{\mathsf{F}}_i : \mathbf{Open}(X) \to \mathbf{Vect}$ by taking colimits of the diagram over an open set, called pre-cosheaves.
- For X a metric space, define a functor $\epsilon : \mathbf{Open}(X) \to \mathbf{Open}(X)$ by

$$U \rightsquigarrow U^{\epsilon} := \{y \in X \mid \exists x \in U \,,\ d(x, y) \leqslant \epsilon\}$$

Remarks - Interleavings

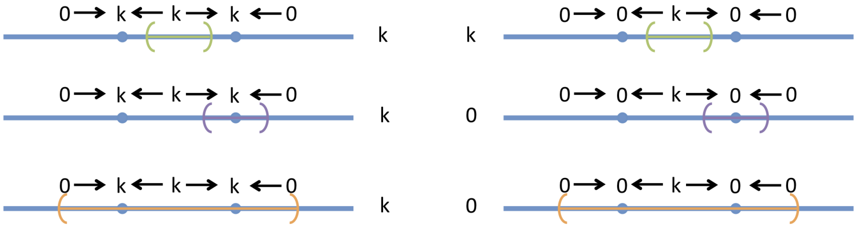


Remarks - Interleavings



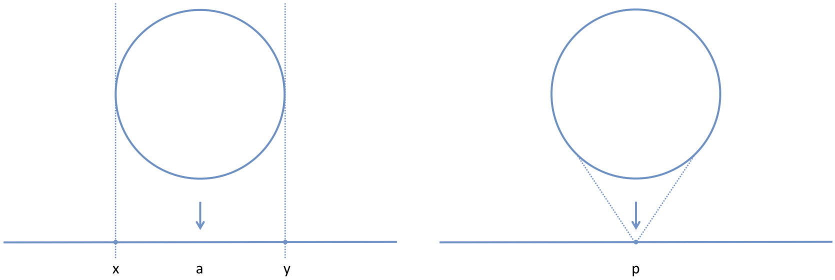
- Define $\hat{F}^{\epsilon} := \hat{F} \circ \epsilon$ for a thickened pre-cosheaf.

Remarks - Interleavings

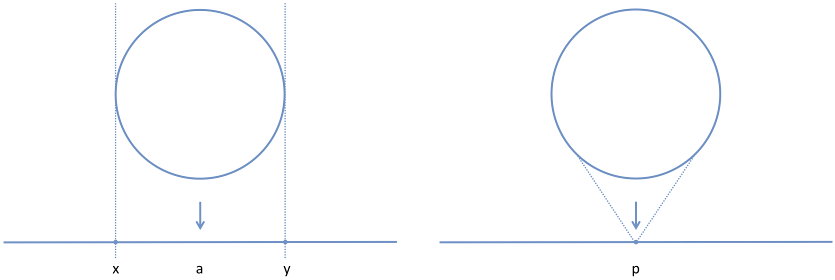


- Define $\hat{\mathsf{F}}^{\epsilon} := \hat{\mathsf{F}} \circ \epsilon$ for a thickened pre-cosheaf.
- Have **interleavings** of pre-(co)sheaves. Since $X^{\epsilon} = X$, $\hat{\mathsf{F}}(X)$ gives an obstruction to interleavings. So a point barcode is not interleaved with an empty barcode. An open barcode is not interleaved with a closed one.

Remarks - Interleavings



Remarks - Interleavings

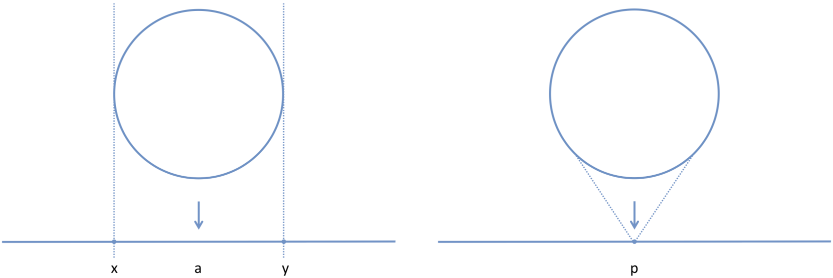


- To $f : Y \to X$, the assignment

$$U \rightsquigarrow H_i(f^{-1}(U); k)$$

is a stable pre-cosheaf.

Remarks - Interleavings



- To $f : Y \to X$, the assignment

$$U \rightsquigarrow H_i(f^{-1}(U); k)$$

  is a stable pre-cosheaf.
- The functors I described which locally agree with this one, are **stable**. But, these are the homology cosheaves of the derived pushforward, which is stable.

Acknowledgements

- This work benefited from conversations with Vin de Silva, Michael Lesnick, Bob MacPherson, and Amit Patel.

Special Thanks to the Organizers!