

METRICS STRATIFICATION THEORY IN MODERN APPLIED TOPOLOGY

A TALK BY
Asst. Prof. JUSTIN CURRY
OUF DYNAMICS SEMINAR 1/18/2022

HISTORY

- TOPOLOGY IS THE YOUNGEST BRANCH OF MATHEMATICS → Riemann 1850s

cf. ALGEBRA, GEOMETRY
ANALYSIS, LOGIC

Betti	1870s
Poincaré	~1900

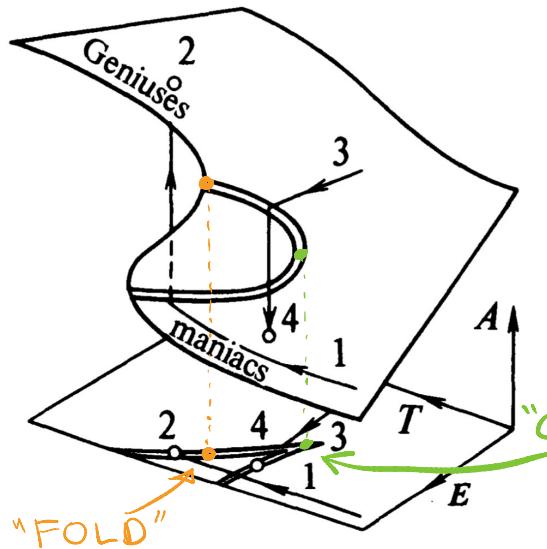
- SINGULARITY THEORY WAS ONE OF THE FIRST APPLICATIONS OF DIFFERENTIAL TOPOLOGY → Maxwell, Morse, Whitney, Thom, Mather

1850s 1920s 1950s 1960s 1970s } Arnold
 $\hookrightarrow f: M \rightarrow N$ smooth mapping

study where $Df: TM \rightarrow TN$
 is not full rank as well as
 "normal forms" of $Df(p)$

CATASTROPHE THEORY

- FIRST REVOLUTION OF APPLIED TOPOLOGY
- ENDED IN DISGRACE by late 70s
 - ~ Sussman \ Zahler 1977 , Smale 1978



cf. ARNOLD's parody of an application of WHITNEY's CUSP \ FOLD CLASSIFICATION by ZEEMAN.

A = Achievement

T = Technical Proficiency

E = Enthusiasm

LESSONS LEARNED

BE HUMBLE

USE NUMBERS / COLLECT DATA

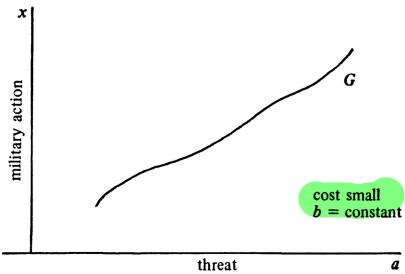


FIGURE 6, 316, ZCT

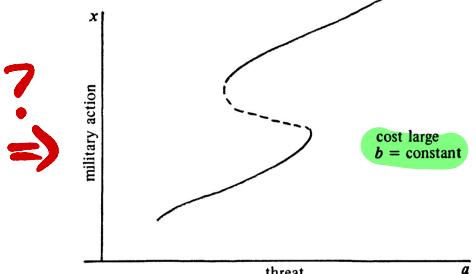
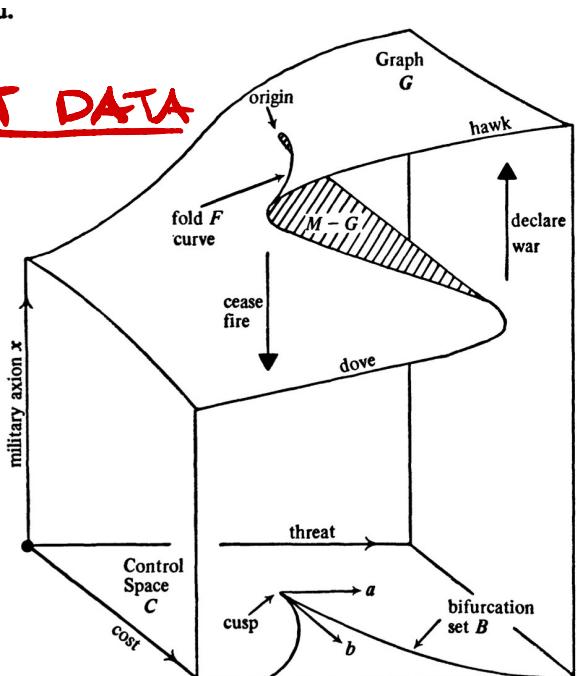


FIGURE 7, 317, ZCT

One trouble is the definition for “small” cost and “large” cost that Zeeman and Isard use. They mean there exists b_0 such that “small b ” means $b < b_0$ and “large b ” means $b > b_0$, the same b_0 for “small” and “large”. Thus Figures 6 and 7 (the sociological hypotheses) already describe the model for all b save b_0 ; Figure 6 applies if $b < b_0$, Figure 7 if $b > b_0$ [p. 331, ZCT]. No evidence or justification is given for this sociological hypothesis, that such a b_0 exists (there are arguments given earlier however to justify a “delay rule” vs. “Maxwell’s rule”).

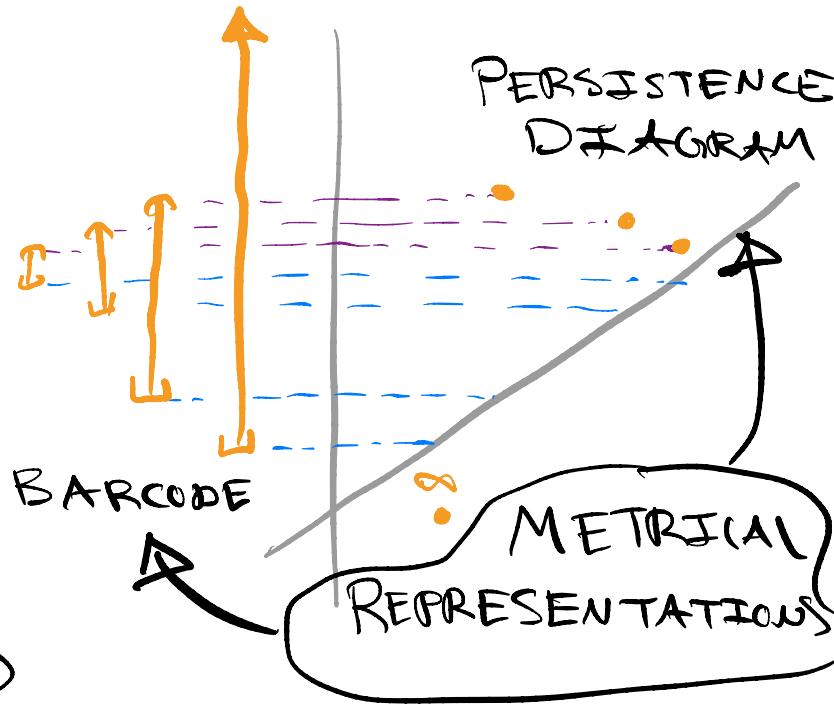
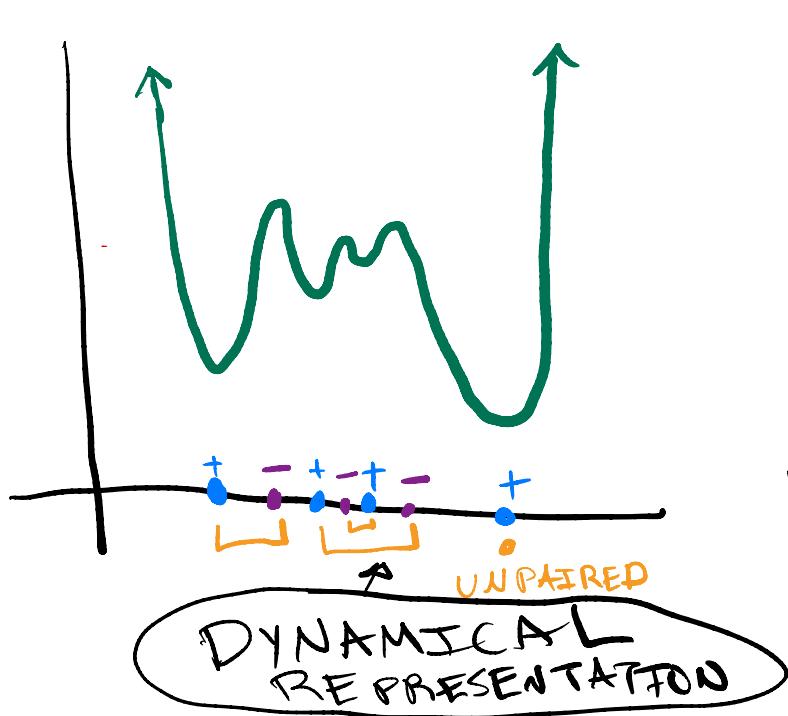
On the other hand around CT, not only does mathematics come first but one sees a sort of mathematical egocentricity; understanding the world is a mathematical (even geometrical) problem. Thom's position on this is clear; e.g. “... Eliminate the “obvious” meaning and replace it by the purely abstract geometrical manipulation of forms. The only possible theoretisation is Mathematical.” (638, ZCT) or “... I agree with P. Antonelli, when he states that theoretical biology should be done in Mathematical Departments; we have to let biologists busy themselves with their very concrete—but almost meaningless—experiments; in developmental Biology, how could they hope to solve a problem they cannot even formulate?” (636, ZCT).



DO STATISTICS!

TOPOLOGICAL DATA ANALYSIS : CATASTROPHE THEORY REDUX?

- Uses COMPUTERS and tries to empower users
- Incorporates METRIC information, e.g. SCALE



INTERLEAVING DISTANCE

- GENERALIZATION OF THE HAUSDORFF DISTANCE

$$X = \square \rightsquigarrow X^\varepsilon = \text{shaded ball}$$

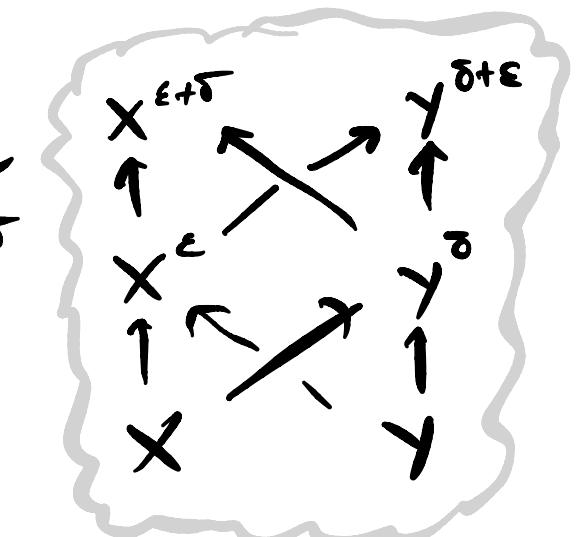
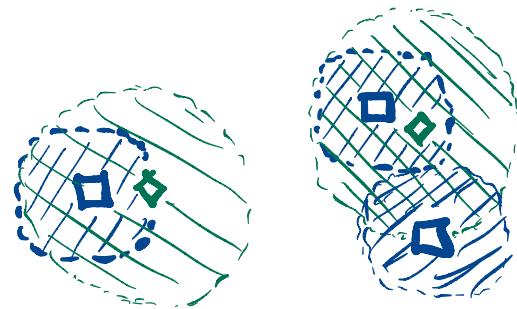
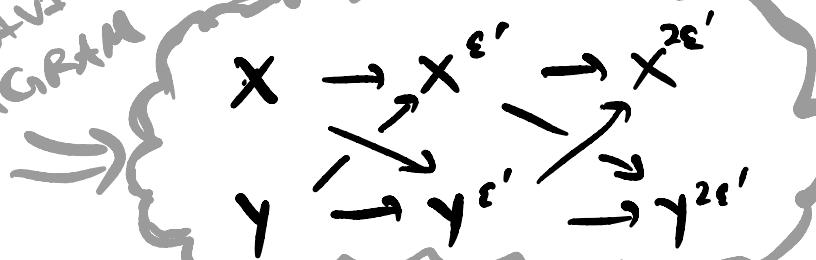
$$Y = \square \rightsquigarrow Y^\delta = \text{shaded ball}$$

- FIND ε s.t.

- FIND δ s.t.

- CHOOSE $\varepsilon' := \max(\varepsilon, \delta)$

ε' -INTER
LEAVING
DIAGRAM



Inclusion
of sets & their thickenings

HAUSDORFF STABILITY ~2005 Cohen-Steiner, et al

- $d_H(x, y) = \inf \{ \varepsilon' \mid \exists \varepsilon' \text{ interleaving} \}$

Here we view x, y as functors from

$\overline{\mathbb{R}_{\geq 0}}$ to $\frac{\text{subsets of } \mathbb{R}^2 \text{ w/ inclusion}}{\mathcal{C}}$ or $\frac{\text{TOP}}{\text{top. spaces}} \xrightarrow{\text{inclusion}}$

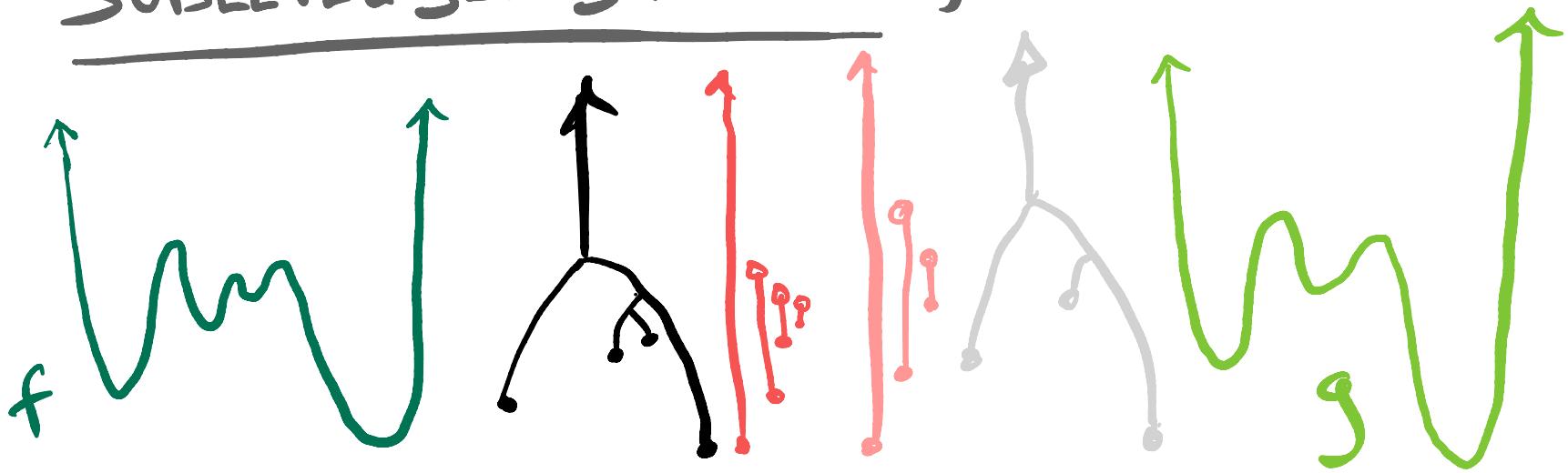
- Of course, applying Homology, provides a completely analogous definition

$$d_I(H_n X, H_n Y) = \inf_{\varepsilon} \text{s.t. } \left\{ \begin{array}{ccc} H_n X & \xrightarrow{\sigma} & H_n X^\varepsilon \xrightarrow{\pi} H_n X^{2\varepsilon} \\ & \searrow & \downarrow \\ H_n Y & \xrightarrow{\sigma} & H_n Y^\varepsilon \xrightarrow{\pi} H_n Y^{2\varepsilon} \end{array} \right.$$

THM for $n > 0$

$$d_I(H_n X, H_n Y) \leq d_H(X, Y)$$

SUBLEVEL SET STABILITY



$$F : \mathbb{R} \rightarrow \text{Top}$$

$$t \mapsto f^{-1}(-\infty, t]$$

Def: ε -shift $F^\varepsilon(t) = F(t + \varepsilon)$

$$d_I(F, G) \leq \|f - g\|_\infty = f^{-1}(-\infty, t + \varepsilon]$$

$$G : \mathbb{R} \rightarrow \text{Top}$$

$$t \mapsto g^{-1}(-\infty, t]$$

and $H_n : \text{Top} \rightarrow \text{Vert}$ } $\pi_0 : \text{Top} \rightarrow \text{Set}$
are LIPSCHITZ.

THMS

$$d_I(H_n F, H_n G) \leq d_I(\pi_0 F, \pi_0 G) \leq \|f - g\|_\infty$$

COMPUTABILITY

- IN GENERAL, d_I is NP HARD to compute, but

LESNICK'S ISOMETRY THM: $d_I(H_n F, H_n G) = d_B(H_n F, H_n G)$

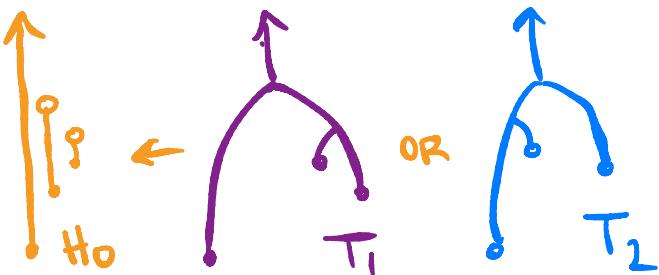
where d_B = BOTTLENECK DIST., which is computable using the HUNGARIAN ALGORITHM

- EVEN MERGE TREE INTERLEAVING DISTANCE IS HARD.

- WOULD BE NICE TO LEVERAGE

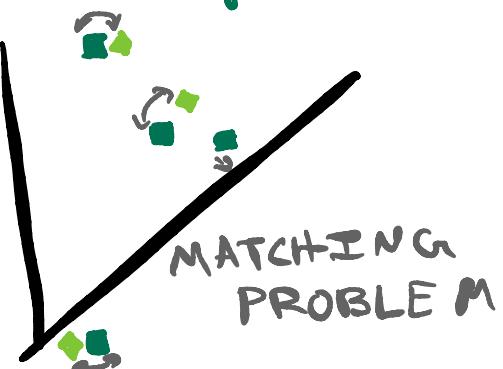
$$\pi_0 \rightarrow h_0 \text{ MAP, BUT}$$

EXponential INVERSE PROBLEMS



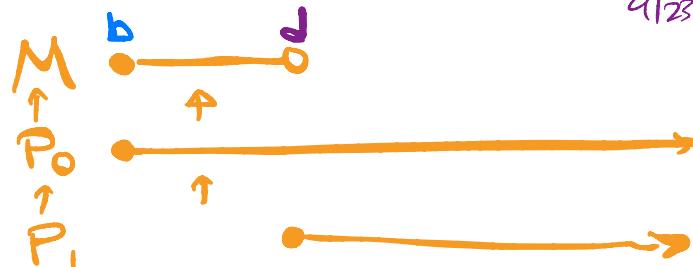
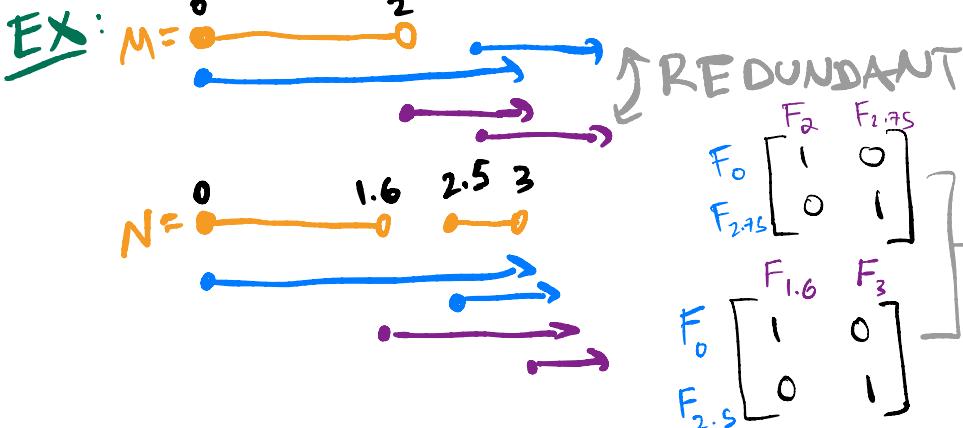
$T_1 \neq T_2$ BUT $h_0 T_1 \cong h_0 T_2$

\Rightarrow NONTRIVIAL FIBERS
(cf. CURRY "FIBER")



PRESENTATIONS

- Consider projective resolution
- Elementary projectives are "born" in a single graded degree.
- By inserting redundant generators & relations we can find compatible presentations for $M \setminus N$ assuming $\varprojlim M \rightarrow \varprojlim M \cong \varprojlim N \rightarrow \varprojlim N$



Use to define $\delta P dist D$

LABEL VECTORS

$$L_M = [0, 2.75, 2, 2.75]$$

$$\begin{matrix} F_0 & F_{2.75} \\ F_{2.75} & \end{matrix} \left[\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right]$$

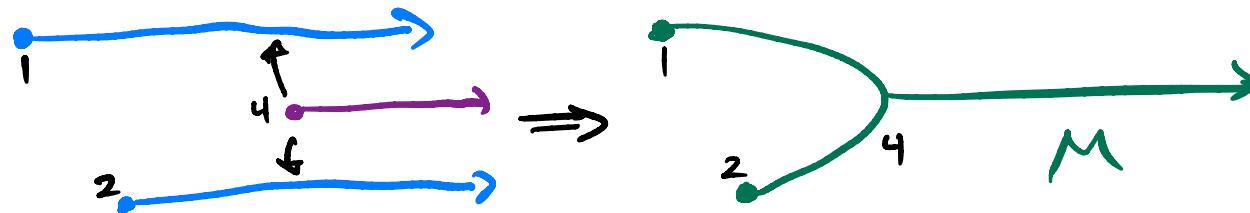
$$\begin{matrix} F_{1.6} & F_3 \\ F_0 & \end{matrix} \left[\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right]$$

$$L_N = [0, 2.5; 1.6, 3] \Rightarrow d_I^2 = \sqrt{(1/4)^2 + (1/4)^2 + (1/4)^2}$$

PRESERNTATIONS FOR MERGE TREES

10/23

- GENERAL PHILOSOPHY: REPLACE COKERS w/ COEQUALIZERS



- PRESERNTATION DISTANCE

$$\hat{d}_I^P(M, N) := \inf \left\{ \|L(P_M) - L(P_N)\|_P \text{ s.t. } P_M \setminus P_N \text{ compatible} \right\}$$

FAILS Δ -Ineq BUT "UNIVERSAL" REPLACEMENT

$$d_I^P(M, N) := \inf \left\{ \sum_{i=0}^{n-1} \hat{d}_I^P(M_i, M_{i+1}) \mid M_0 = M \quad M_n = N \right\}$$

- WASSERSTEIN STABILITY (for MTs w/ Caronat/Lam/Lesnick)

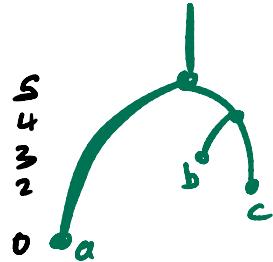
$$d_{WP}(H_0M, H_0N) \leq d_I^0(M, N) \quad \begin{matrix} \text{also} \\ \text{for} \\ \text{pers. modules} \end{matrix} \quad d_{WP}(H_nF, H_nG) \leq d_I^P(F, G)$$

BJERKVIK & LESNICK

INCLUDE ℓ^p PENALTIES

COUNTEREXAMPLE TO COPHENETIC STABILITY

11/23



$$\Rightarrow \begin{matrix} & a & b & c \\ a & 0 & s & s \\ b & s & 3 & 4 \\ c & s & 4 & 2 \end{matrix} = C_M \quad \text{LEAST COMMON ANCESTOR MATRIX}$$

• MUNCH & STEFANOU '17: For ordered MTs w/ same # leaf nodes S

$$d_I(M, N) = \|C_M - C_N\|_\infty$$

• GASPAROVIC, MUNCH, OUPOT, TURNER, B.WANG, V.WANG '19

$$d_I(M, N) = \inf_{\pi: A \rightarrow L} \|C_M - C_N\|_\infty \quad \xrightarrow{\text{This is the NP part!}}$$

• ALTHOUGH $\|C_M - C_N\|_p$ makes sense, it is unstable

BOTH FILTRATIONS OF

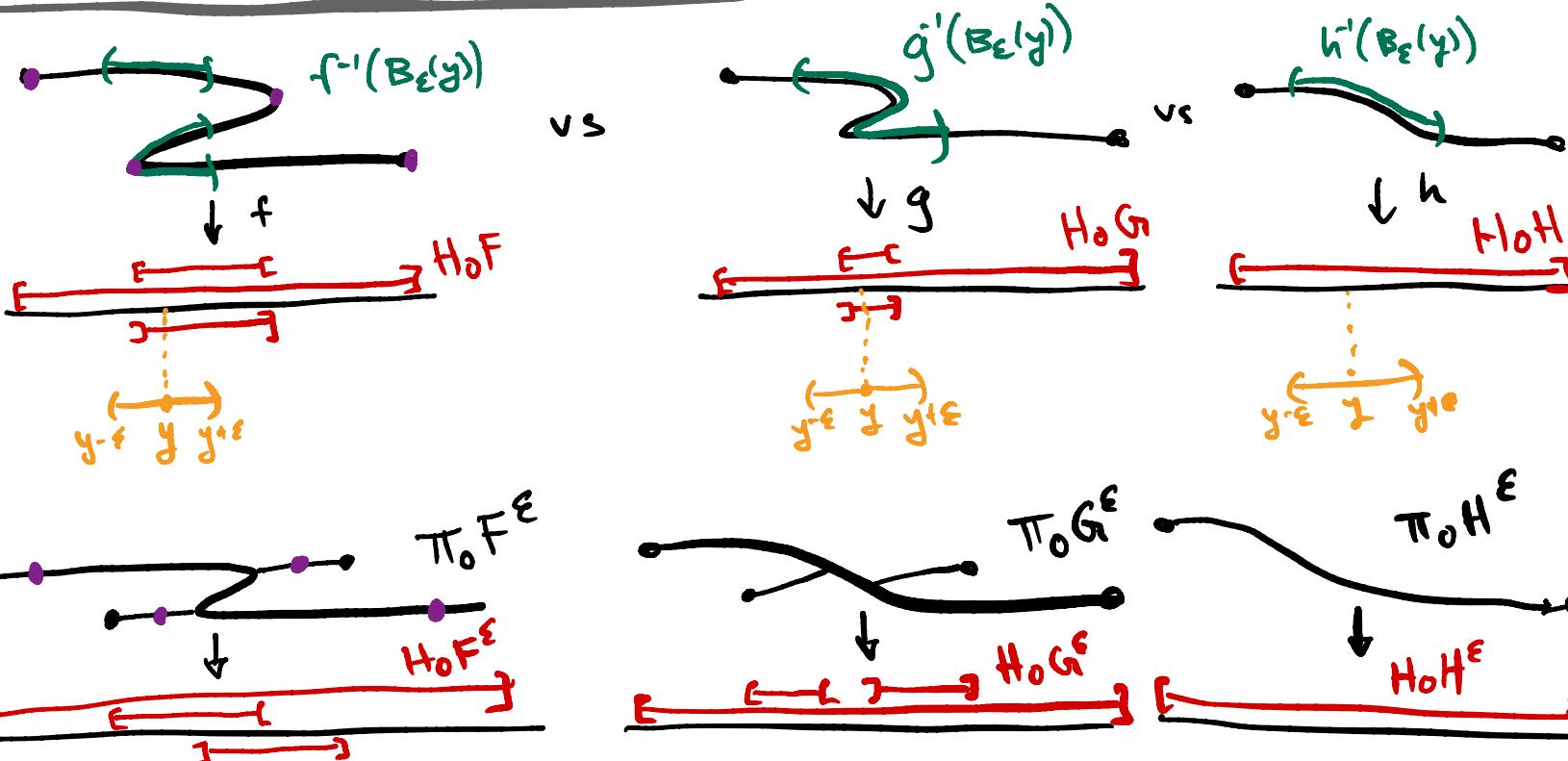
\leadsto LP distance $\sqrt{2}$

$\| \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \|_2 = \sqrt{3}$ BUT $\| \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \|_2 = \sqrt{2}$

QUADRATIC vs. LINEAR!

LEVEL SET STABILITY (Revist our catastrophe)

12/23



Def $F: \text{Open}(\mathbb{R}) \rightarrow \text{Top}$ $F(u) = f^{-1}(u)$ and $F^\varepsilon(u) = f^{-1}(u^\varepsilon)$

THM: $d_I(H_0 F, H_0 G) \leq d_I(\pi_0 F, \pi_0 G) \leq \|f-g\|_\infty$

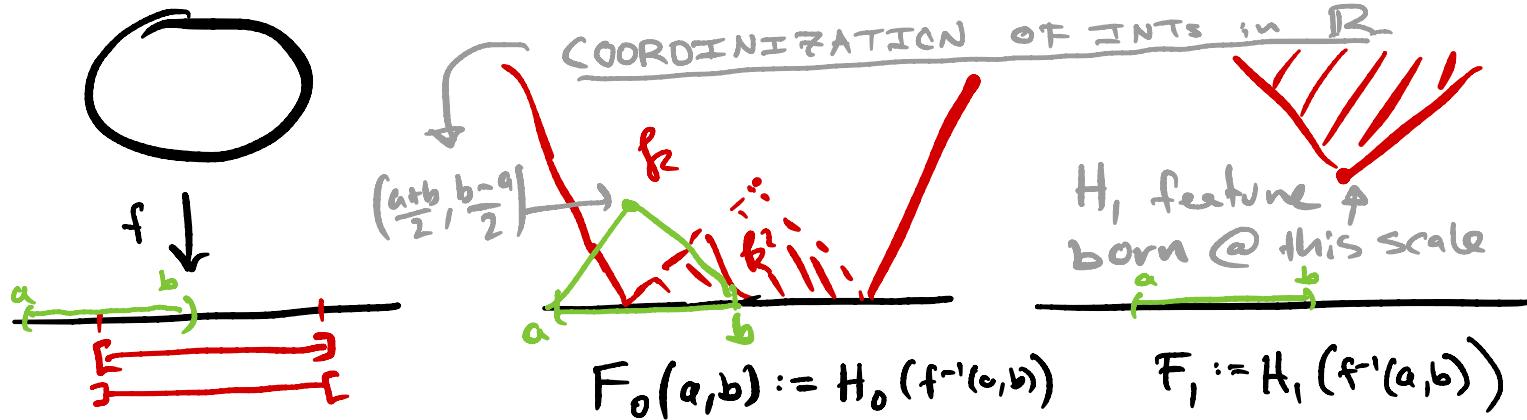
DYNAMICS ON BARCODES \nmid THE MOVE FROM PRE-SHEAVES \rightarrow SHEAVES

- FOUR TYPES OF INDECOMPOSABLES/BARCODES:

CLOSED , LEFT CLOSED , RIGHT CLOSED , OPEN



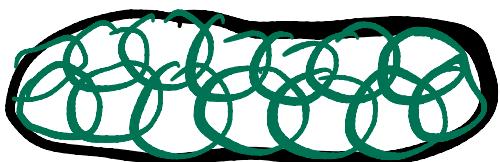
- BUT THIS REPRESENTATION REQUIRES SHEAVES OTHERWISE YOU USE 2D BLOCKS!



SHEAVES & COSHEAVES

14/23

- GENERAL FUNCTORS $F : \text{Open}(X)^{(op)} \rightarrow \mathcal{D}$ = DATA CATEGORY
 ARE CALLED PRE-SHEAVES ($\circ p$)
 OR PRE-COSHEAVES ($\circ \circ op$)
 SET or VECT
 or $\mathbf{Ch}(\text{VECT})$
- IN ORDER TO ENSURE LOCAL DETERMINATION
 i.e. $F(U)$ CAN BE ASSEMBLED FROM ANY COVER OF U



$U = \{U_i\}$
 closed under
 intersection

- THERE IS A PROCESS FOR FORCING THIS
 ☆ (co) SHEAFIFICATION ☆



SHEAF AXIOM

$$F(U) \cong \varinjlim F(U_i)$$

$$\downarrow \quad \downarrow$$

$$F(U_i) \rightarrow F(U_{ij}) \leftarrow F(U_j)$$

COSHEAF AXIOM

$$F(U) \cong \varprojlim F(U_i)$$

$$\uparrow \quad \uparrow$$

$$F(U_i) \wedge F(U_{ij}) \rightarrow F(U_j)$$

STABILITY FOR ((0) SHEAVES)

- INSTEAD OF INCREASING DIMENSION, USE SHEAVES

Given $f: X \rightarrow \mathbb{R}^k$ the n^{th} Leray sheaf is $\mathcal{F}_x^n = H^n(f^{-1}(x))$ if x cpt, then

$\mathcal{F}^n = Rf_* \mathbb{R}_X$ where \mathcal{F}^n is the sheafification ^{STALK} _{\mathbb{R}_X}

For our example
const.
sheaf on X

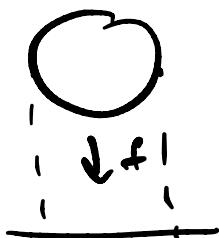


$$\begin{aligned} F' &= H^1(f^{-1}(U)) \\ &\neq 0 \end{aligned}$$

BUT $\mathcal{F}' = 0$!

$$\begin{aligned} \text{if } F'(U) &= H^1(f^{-1}(U)) \\ \text{Leray Pre-Sheaf} \end{aligned}$$

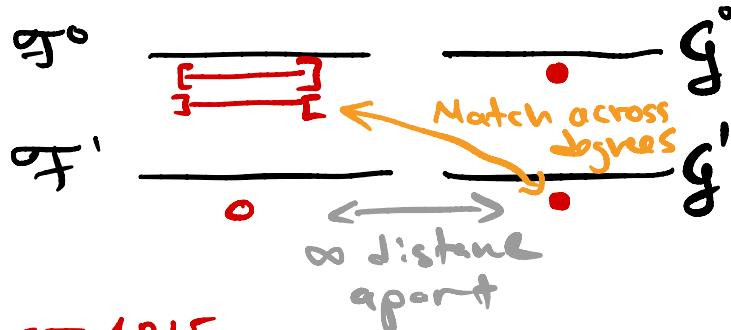
- SHEAVES ARE UNSTABLE DEGREE BY DEGREE



vs



(and GRADED!)



- BUT DERIVED SHEAVES ARE STABLE

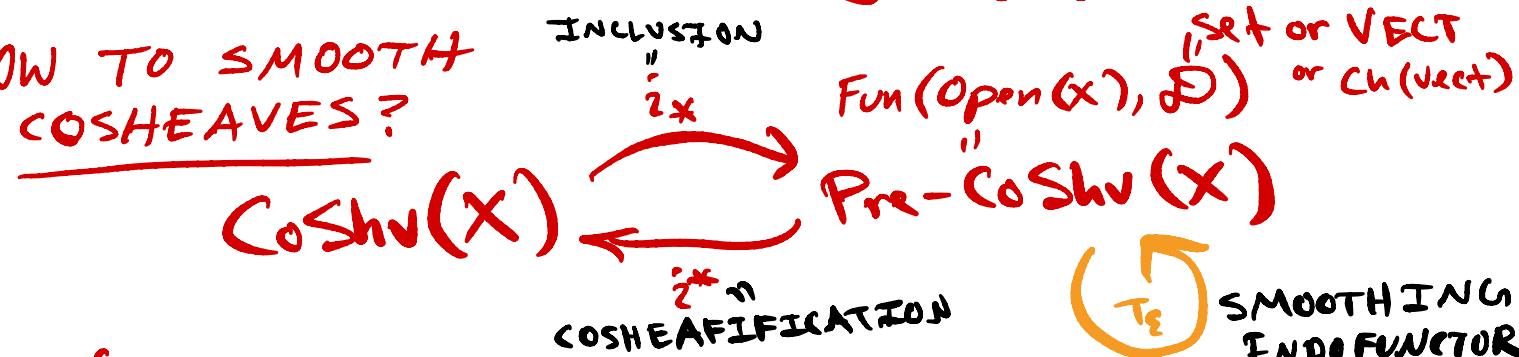
$Rf_* \mathbb{R}_X := f_* S^\bullet$ where S^\bullet = sheet of singular cohomology

PRESERVING LOCALITY

16/23

- EASY TO SHIFT OR "SMOOTH" PRE-COSHEAVES $F_n^{\varepsilon(x)}$
 $T_{\varepsilon}(F) := F_n^{\varepsilon}(U) = H_n(f^{-1}(U^{\varepsilon}))$ but $\mathcal{F}_n^{\varepsilon}(U) = \varinjlim \overline{h_n(f^{-1}(B_{\varepsilon}(x)))}$
 Uses "Small" cover, but ε is getting bigger!

- HOW TO SMOOTH COSHEAVES?



$$\mathcal{F}^{\varepsilon} := i^* T_{\varepsilon} i_* \mathcal{F}$$

$$(\mathcal{F}^{\varepsilon})^{\varepsilon} \rightarrow \mathcal{F}^{2\varepsilon} \text{ uses } i^* T_{\varepsilon} i_* i^* T_{\varepsilon} i_* \rightarrow i^* T_{\varepsilon} T_{\varepsilon} i_* \xrightarrow{i^* T_{2\varepsilon} i_*}$$

ADJOINT PAIR $i_* \dashv i^*$



CATEGORIES w/ A FLOW (THESIS WORK OF A. STEFANOU)

17/23

- IN GENERAL METRIC SPACES, THICKENING HAS POOR PROPERTIES: GIVEN $A \subseteq (X, d)$ $A^\epsilon = \{y \mid d(A, y) \leq \epsilon\}$
THIS DEFINES AN ACTION $T: P(X) \times [0, \infty) \rightarrow P(X)$ BUT
 $T_0 A \neq A \quad \downarrow T_\epsilon T_\delta A = (A^\epsilon)^\delta \leq A^{2\epsilon} = T_{2\epsilon} A$
- desILVA, MUNCH, and STEFANOU abstracted this into a general framework using category theory.

Definition 1.2. A category with a flow (or translation)¹ (\mathbf{C}, T) consists of a category \mathbf{C} , together with

- a functor $T: \mathbb{R}_{\geq 0} \rightarrow \text{End}(\mathbf{C})$, $\epsilon \mapsto T_\epsilon$, called the **flow** or **translation**,
- a natural transformation $u: I_{\mathbf{C}} \Rightarrow T_0$, where $I_{\mathbf{C}}$ is the identity endofunctor of \mathbf{C} , and
- a collection of natural transformations $\mu_{\epsilon, \zeta}: T_\epsilon T_\zeta \Rightarrow T_{\epsilon+\zeta}$, $\epsilon, \zeta \geq 0$,

such that the diagrams

$$\begin{array}{ccc}
 \begin{array}{c}
 \begin{array}{ccccc}
 & u_{I_{\mathbf{C}}} & & & \\
 & \swarrow & \searrow & & \\
 T_0 & T_\epsilon & & T_\epsilon & \\
 & \swarrow & \searrow & & \\
 & \mu_{0,\epsilon} & & &
 \end{array}
 & \quad &
 \begin{array}{ccccc}
 & I_{T_\epsilon} u & & & \\
 & \swarrow & \searrow & & \\
 T_\epsilon & T_0 & & T_\epsilon & \\
 & \swarrow & \searrow & & \\
 & \mu_{\epsilon,0} & & &
 \end{array}
 \end{array}
 & \quad &
 \begin{array}{ccc}
 & & \\
 T_\epsilon T_\zeta T_\delta & \xrightarrow{I_{T_\epsilon} \mu_{\zeta, \delta}} & T_\epsilon T_{\zeta+\delta} \\
 \downarrow \mu_{\epsilon, \zeta} I_{T_\delta} & & \downarrow \mu_{\epsilon, \zeta+\delta} \\
 T_\epsilon T_\delta & \xrightarrow{\mu_{\epsilon+\zeta, \delta}} & T_{\epsilon+\zeta+\delta} \\
 & &
 \end{array}
 \quad
 \begin{array}{ccc}
 & & \\
 T_\epsilon T_\zeta & \xrightarrow{\mu_{\epsilon, \zeta}} & T_{\epsilon+\zeta} \\
 \downarrow T_{(\epsilon \leq \delta)} T_{(\zeta \leq \kappa)} & & \downarrow T_{(\epsilon+\zeta \leq \delta+\kappa)} \\
 T_\delta T_\kappa & \xrightarrow{\mu_{\delta, \kappa}} & T_{\delta+\kappa} \\
 & &
 \end{array}$$

commute for every $\epsilon, \zeta, \delta, \kappa \geq 0$.

Definition 1.3. Let M, N be two objects in \mathbf{C} . A weak ϵ -interleaving of M and N , denoted (φ, ψ) , consists of a pair of morphisms $\varphi: M \rightarrow T_\epsilon N$ and $\psi: N \rightarrow T_\epsilon M$ in \mathbf{C} such that the following pentagons

$$\begin{array}{ccccc}
 T_0 M & \xleftarrow{u_M} & M & \xrightarrow{u_N} & T_0 N \\
 \downarrow T_{(0 \leq 2\epsilon), M} & & \varphi \swarrow & \searrow \psi & \downarrow T_{(0 \leq 2\epsilon), N} \\
 T_\epsilon M & & T_\epsilon N & & \\
 \downarrow T_\epsilon \varphi & \swarrow & \searrow T_\epsilon \psi & & \downarrow T_\epsilon \psi \\
 T_{2\epsilon} M & \xleftarrow{\mu_{\epsilon, \epsilon, M}} & T_\epsilon T_\epsilon M & \xrightarrow{\mu_{\epsilon, \epsilon, N}} & T_{2\epsilon} N
 \end{array}$$

commute. We say that M, N are weakly ϵ -interleaved if there exists a weak ϵ -interleaving (φ, ψ) of M and N . The (weak) interleaving distance with respect to T for a pair of objects M, N in \mathbf{C} is defined to be

$$d_{(\mathbf{C}, T)}(M, N) = \inf\{\epsilon \geq 0 \mid M, N \text{ are weakly } \epsilon\text{-interleaved}\}.$$

If M and N are not weakly interleaved for any ϵ , we set $d_{(\mathbf{C}, T)}(M, N) = \infty$.

PULLING BACK A FLOW

18/23

(WORK W/ MAGNUS BOTNAN & LIZ MUNAO)

cf. "A RELATIVE

THEORY

- GIVEN A CATEGORY WI A FLOW (\mathcal{D}, τ) AND AN ADJOINT PAIR $F \dashv G$ INTERLEAVINGS

$$\tilde{\tau}_{\mathcal{E}} = G\tau_{\mathcal{E}}F \quad \begin{array}{c} \text{---} \\ \mathcal{E} \end{array} \xrightarrow{F} \mathcal{D} \xleftarrow{G} \mathcal{D} \xrightarrow{\tau_{\mathcal{E}}} \mathcal{T}_{\mathcal{E}}$$

(BUT THIS IS
A MAJOR REWRITE)

WE HAVE AN INDUCED PULLBACK FLOW $\tilde{\tau}$

- MOREOVER, THIS EQUIPS $\mathcal{E} \setminus \mathcal{D}$ WI THE STRUCTURE OF LAWVERE METRIC SPACES WHERE

$$d_{(\mathcal{E}, \tilde{\tau})}(M, N) = d_{(\mathcal{D}, \tau)}(f_{\#}M, f_{\#}N)$$

IS AN ISOMETRY ONTO ITS IMAGE.

APPLICATION TO DISCRETIZATIONS

14/23

- Suppose $f: P \rightarrow Q$ is a map of posets

e.g. $x_1 \xrightarrow{[x_1, x_2]} x_2 \xrightarrow{[x_2, x_3]} x_3 \xrightarrow{[x_3, x_4]} x_4 = P$

$$Q = \text{CLOSED}(B) \quad ; \quad Q \xrightarrow{\text{TE}} S \mapsto S^\varepsilon$$

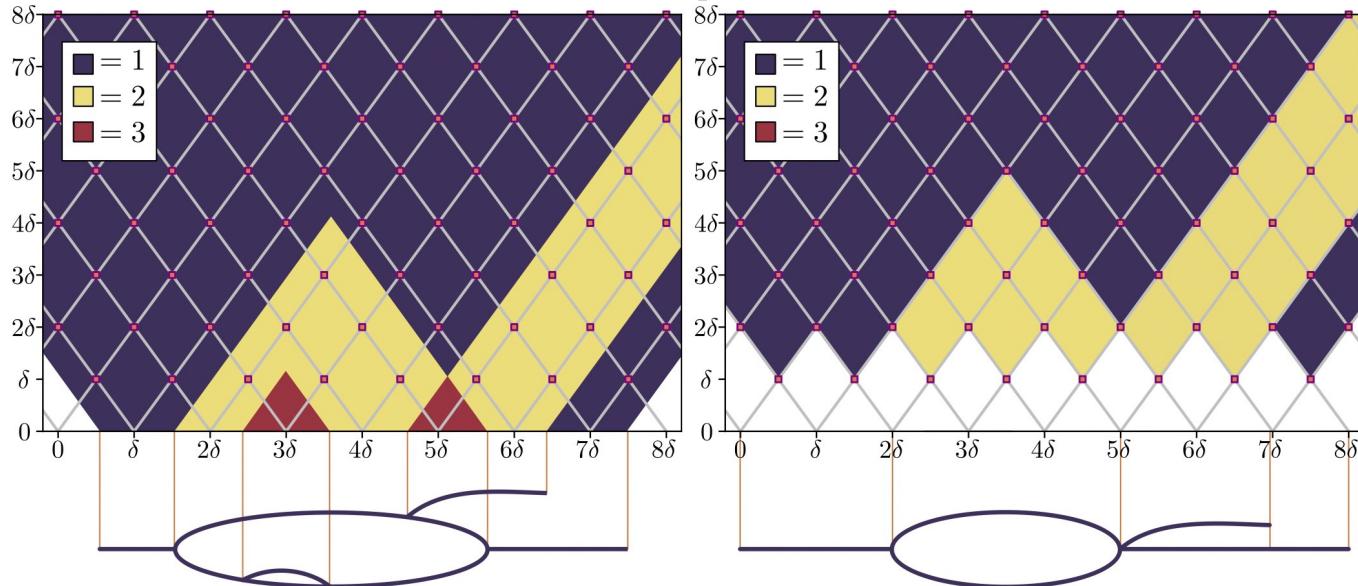
AND $Q \times B_{\geq 0} \rightarrow Q$ HAS A TRANSLATION

THEN (f_*, f^*) IS AN ADJOINT PAIR
THAT ALLOWS US TO COMPUTE INTERLEAVINGS
OF $M, N \in \text{Fun}(Q, \text{VECT})$ WRT $f: P \rightarrow Q$

- WE THINK OF $f_* f^* M = \text{"PIXELIZATION OF } M\text{"}$

PIXELIZATION OF REEB GRAPHS

(Figure courtesy of Liz Munch)



- PIXELIZATION ACTUALLY BREAKS THE COSHEAF AXIOM
- cf. MUNCH/WANG or KASHIWARA/SCHAPIRA

INFERENCE RESULT

f^*M
"

- IMAGINE $f: P \rightarrow Q$ is a "SAMPLING" OF A POSET Q e.g. $P = \uparrow, \uparrow \wedge \uparrow, \uparrow \vee \uparrow, \uparrow$ $Q = \text{OPEN}(\mathbb{R})$
- THE Δ -INEQUALITY PROVIDES "INFERENCE"

$$d_Q(M, N) \leq d_Q(M, f_* f^* M) + \underbrace{d_Q(f_* f^* M, f_* f^* N)}_{d_P(f^* M, f^* N)} + d_Q(f_* f^* N, N)$$

INTRINSIC DISTANCE B/W SAMPLINGS

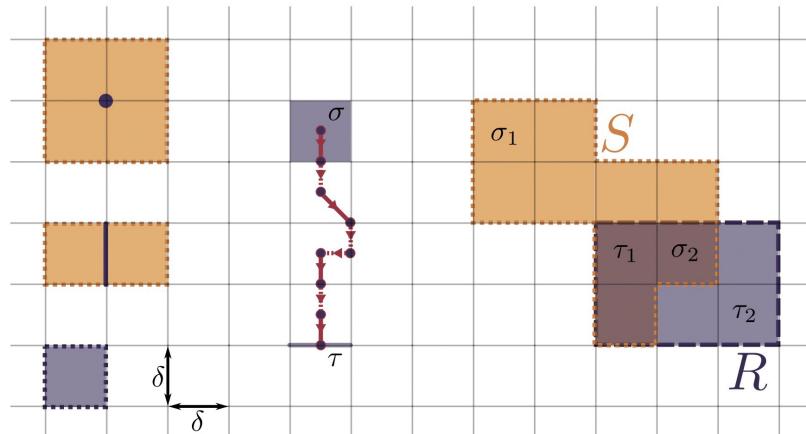
LEMMA: $|d_Q(M, N) - d_P(f^* M, f^* N)| \leq d_Q(M, f_* f^* M) + d_Q(N, f_* f^* N)$

DEF: $f: P \rightarrow Q$ is a δ -APPROX IF $\forall q \in Q \exists p \in P \cup q \leq f(p) \leq T_\delta(q)$

THM: If $P \not\cong Q$ are LATTICES $\nexists f: P \rightarrow Q$ is a δ -approx that preserves JOINS THEN WE CAN INFERENCE INTERLEAVINGS UP TO A 2δ TOLERANCE!

WORKING LOCALLY w/ WEIGHTED POSETS

- PRACTICALLY, WE CAN REDUCE INTERLEAVINGS TO A GRID ON \mathbb{R}^k
- \tilde{T}_ϵ THEN ONLY "JUMPS" AT DISCRETE VALUES OF ϵ ADAPTED TO THE GRID SIZE



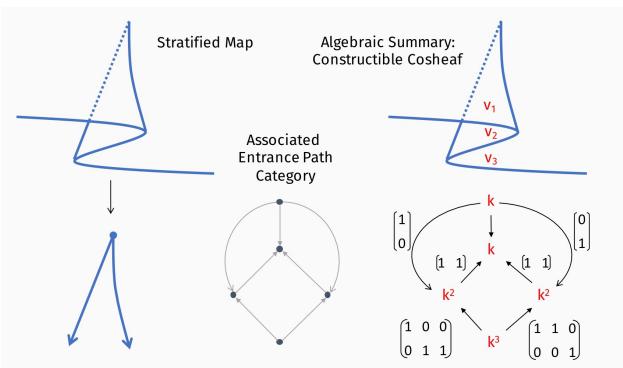
- THIS INDUCES A "WEIGHTING" ON THE FACE RELATION POSET (MONOIDAL ENRICHMENT) THAT ALLOWS US TO METRIZE CELLULAR SHEAVES & COSHEAVES

FUTURE DIRECTIONS

- THIS PROVIDES A METRIC VERSION OF MACPHERSON'S OBSERVATION

CONSTRUCTIBLE
(CO)SHEAVES \cong REPRESENTATIONS
OF THE

(ENTRANCE) EXIT
PATH CATEGORY



← PREVIEW OF FRIDAY'S
TALK @
UF TDA 2022

AND MAKES IT EVEN MORE PERTINENT
TO MODERN APPLIED TOPOLOGY!



THANK
YOU!