TDA - fall 2025

Reeb
Metrics

Recap

- This week: Reeb + Mapper graphs
- Next week: no class
- Paper choose: due Friday
- Next assignment: after break
  final project proposals

Question: Given 2 Reeb graphs, how to compare them?

Goal: At least an extended pseudo metric:

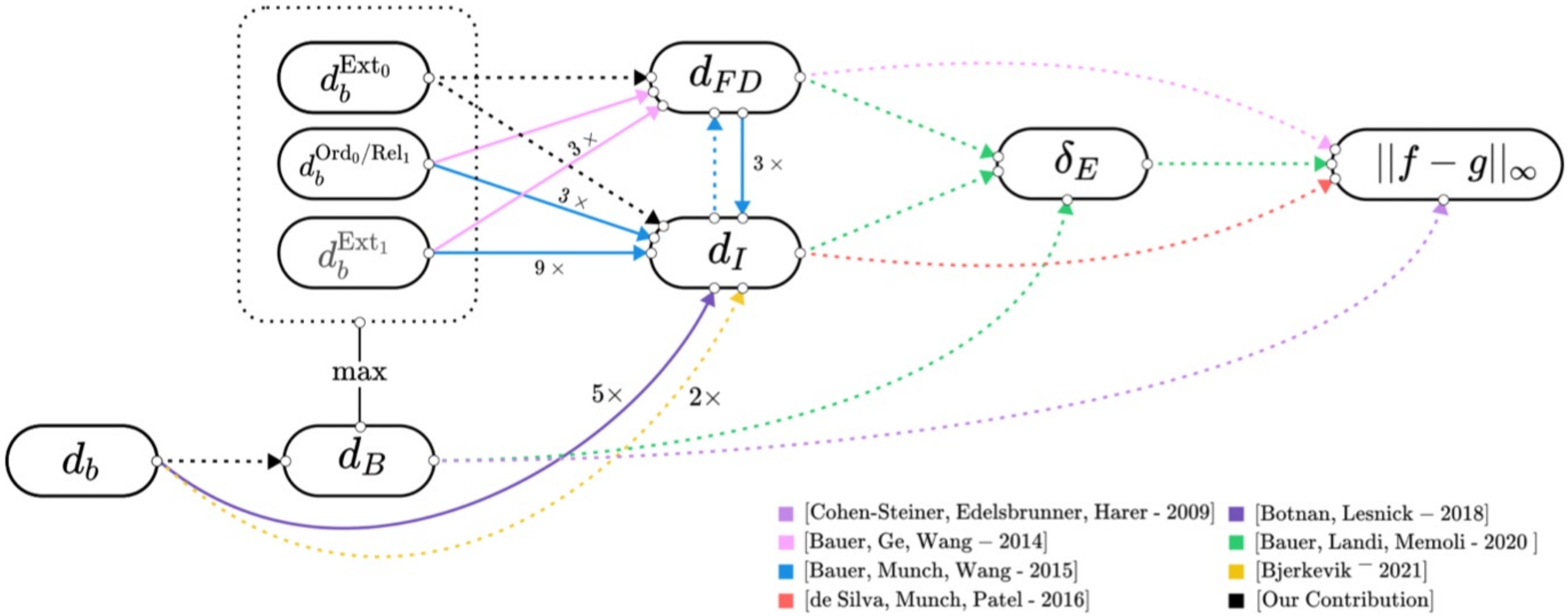
1) $d(X, X) = 0$
2) $d(X, Y) = d(Y, X)$
3) $d(X, Z) \leq d(X, Y) + d(Y, Z)$

Weaker than metric:

- extended: could be ∞
- pseudo: $d(x, y) = 0 \not\Rightarrow x = y$

Lots of well-studied metrics!

Bollen - C. - Levine - Munch 2023



d_B : bottleneck (last time)

d_I : interleaving (cont.)

d_{FD} : functional distortion

d_E : Edit distance

Interleaving distance: Some Category theory

A category C consists of objects (X, Y, Z, etc) with morphisms $(\to)$ between them, satisfying some "niceness" properties (associativity, identity).

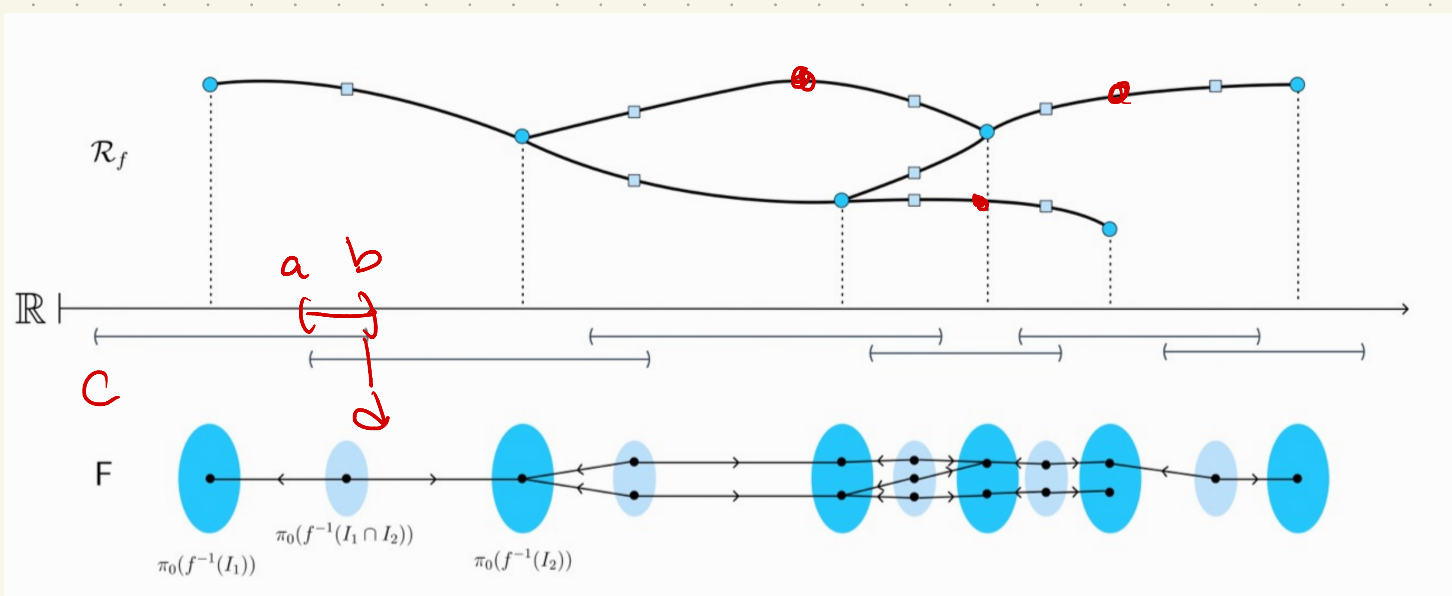
Examples	Objects	Morphisms
"Sets"	Sets	functions
"Vect"	Vector spaces	Linear Transformations
"Top"	Topological spaces	Continuous Maps
"Int"	Intervals $(a,b) \subset \mathbb{R}$	$\subseteq$

Interleaving distance:

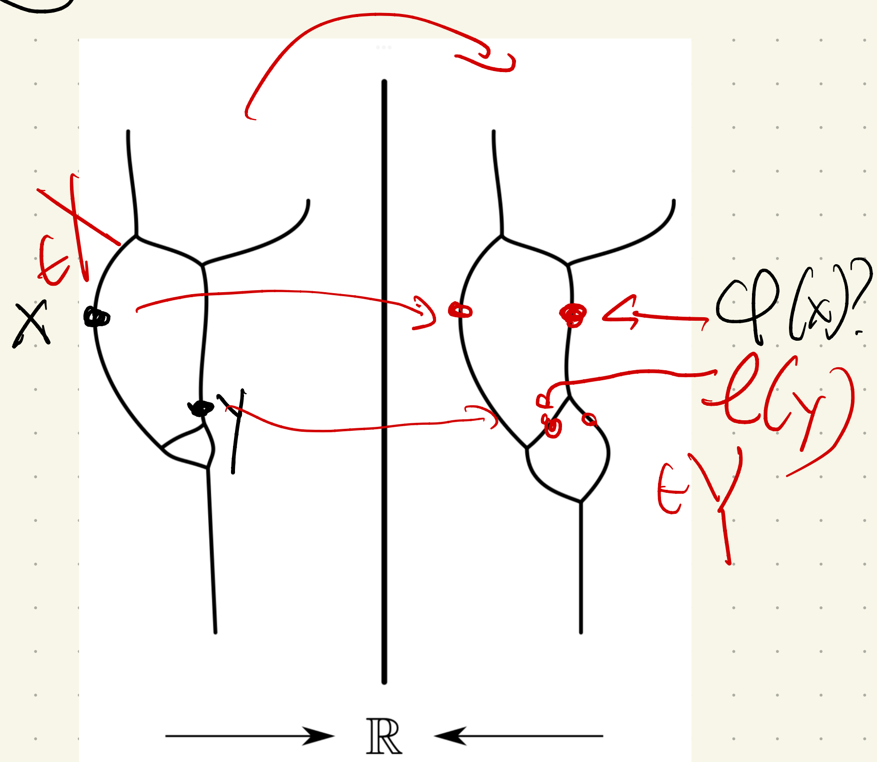
Takes a categorical point of view:
Reeb graph is a set valued cosheaf,

$$F : \mathrm{Int} \longrightarrow \mathrm{Set}$$

$$\begin{tikzcd}
I=(a,b) \arrow[r] \arrow[d, hook] & \pi_0\left(f^{-1}((a,b))\right) \arrow[d, "{F[I \subseteq J]}"] \\
J=(c,d) \arrow[r] & \pi_0\left(f^{-1}((c,d))\right)
\end{tikzcd}$$



Definition: "Reeb" is the category with Reeb graphs as objects, & morphisms as function preserving maps φ



function preserving map:

$\varphi : (X,f) \to (Y,g)$

is a continuous map s.t.

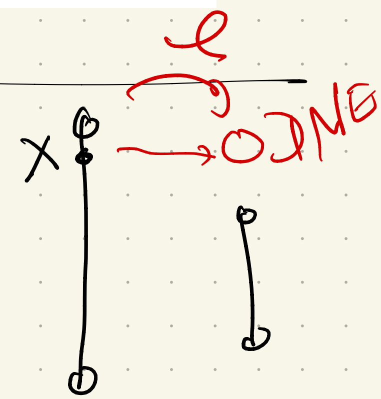
Reeb graph

Reeb graphs

$$\begin{tikzcd} X \arrow[rr, "\varphi"] \arrow[dr, "f"'] & & Y \arrow[dl, "g"] \\ & \mathbb{R} & \end{tikzcd}$$

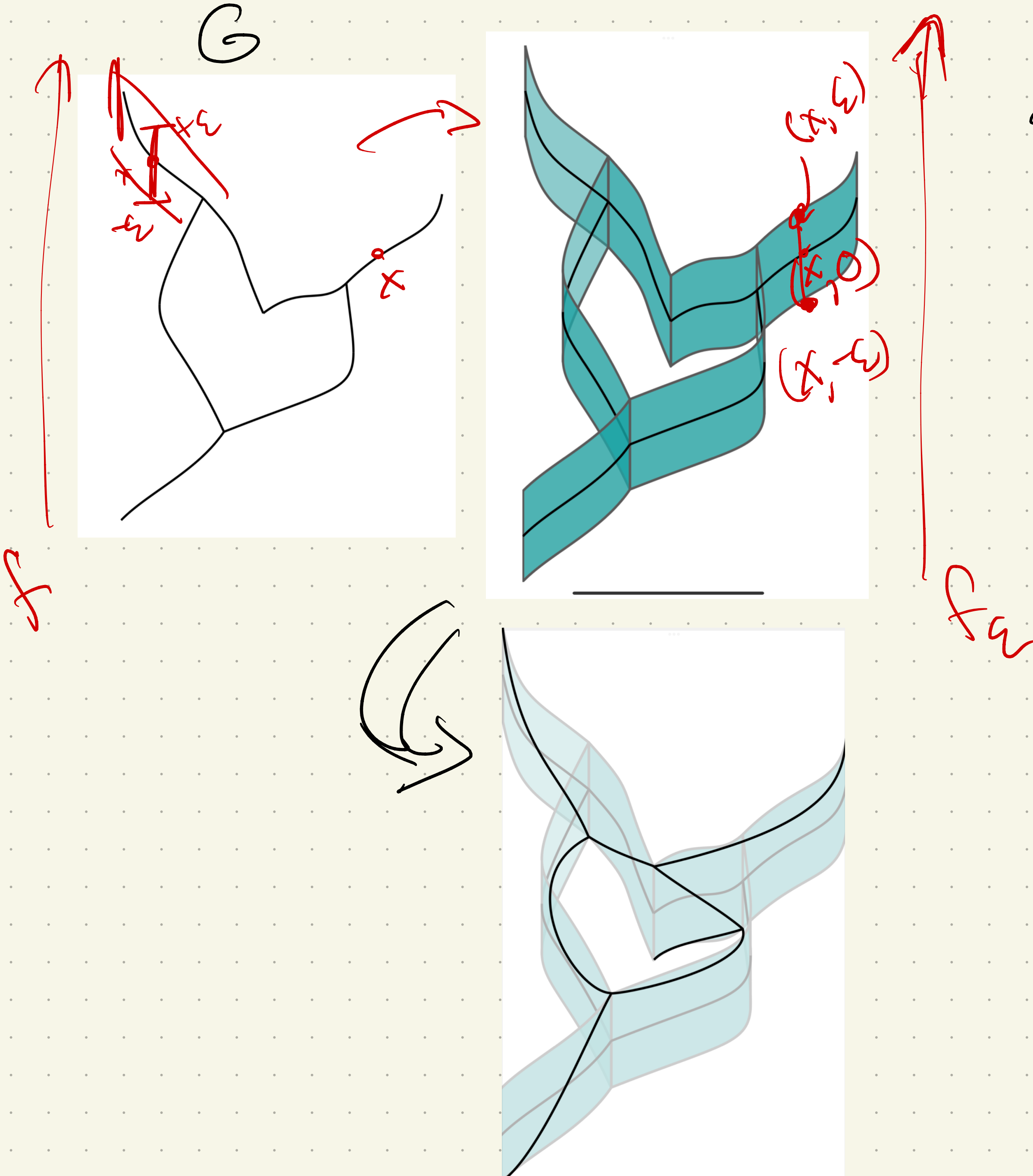
commutes.

Note: Don't always exist! $\boxed{\text{Ex!}}$



Smoothing a Reeb co-sheaf: (Definition by example)

$G\times[-\varepsilon,\varepsilon]$



- Given (G,f)
- Thicken!

  $f_\varepsilon : G\times[-\varepsilon,\varepsilon]\to\mathbb{R}$

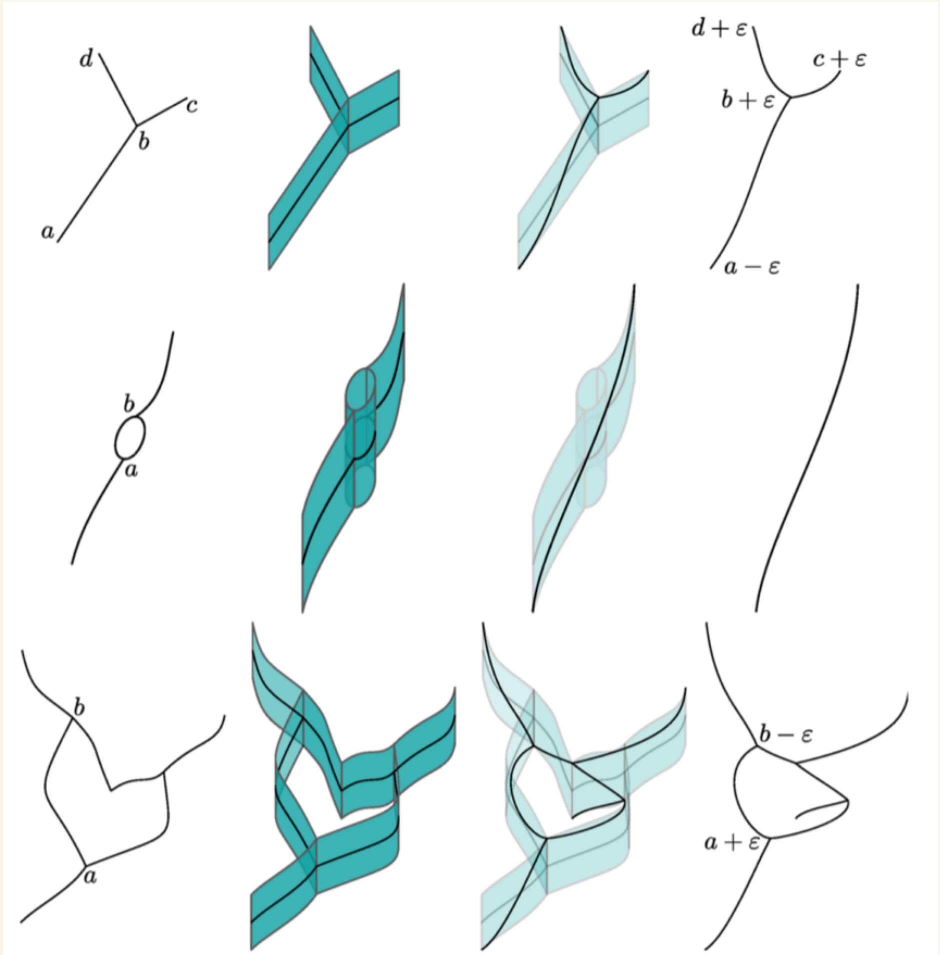
  $(x,t)\mapsto f(x)+t$

- Take Reeb graph of $(G\times[-\varepsilon,\varepsilon], f_\varepsilon)$

Result of smoothing: $S_\varepsilon(G,f)$:

- loops get 2ε smaller
- Max + mins "stretch"

Munch, de Silva + Patel 2016



Interleaving distance: Why do we care??

Given 2 Reeb cosheaves

$$F, G : \mathrm{Int} \to \mathrm{Set}$$

an ε-interleaving is φ and ψ such that the diagram commutes:

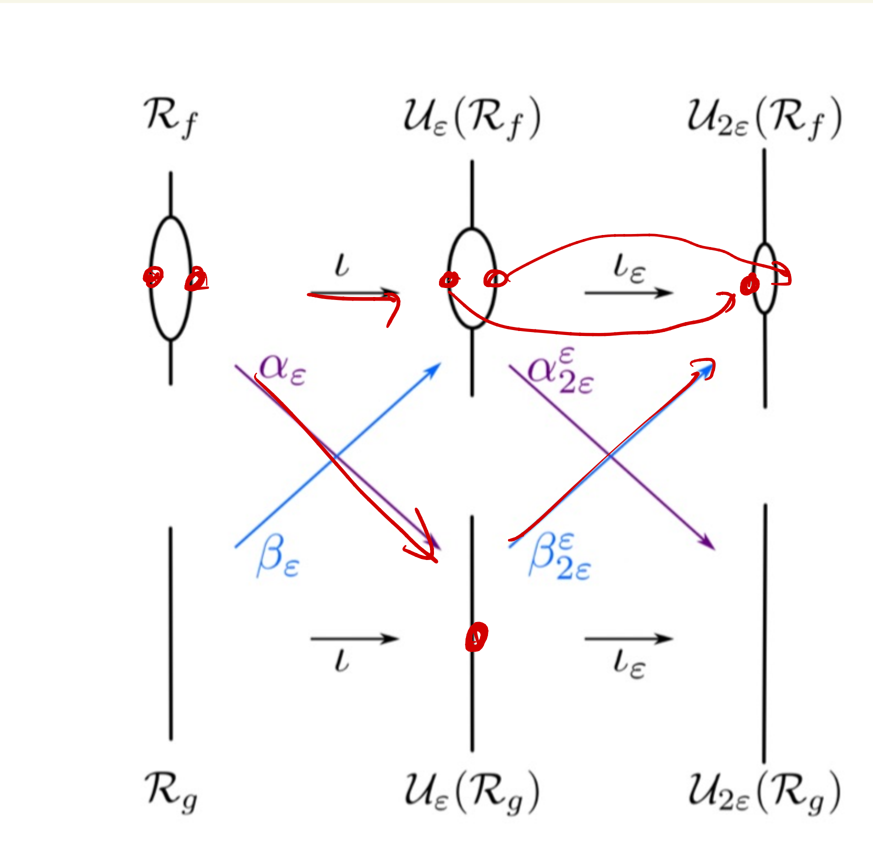
$$\begin{tikzcd}
F \arrow[r] \arrow[dr] & S_\varepsilon F \arrow[r] \arrow[dr] & S_{2\varepsilon} F \\
G \arrow[r] \arrow[ur] & S_\varepsilon G \arrow[r] \arrow[ur] & S_{2\varepsilon} G
\end{tikzcd}$$

Reeb Interleaving distance is

$$d_I(F,G) = \inf\{\varepsilon \geq 0 \mid \exists\ \varepsilon\text{-interleaving}\}$$

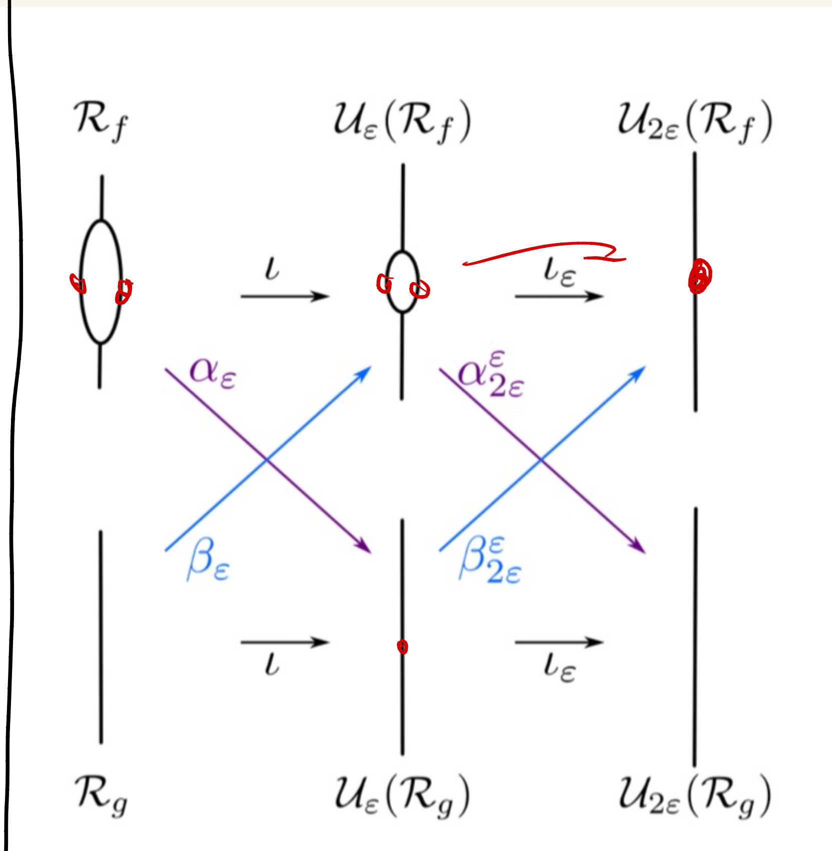
Example:

Bad ε:



no ε-interleaving

Good ε: (larger)



found φ + Φ showing interleaving is $\leq$ this one

Interleaving distance pros + cons:

It is: - discriminative [Bauer et al 2015]
- stable
- path component sensitive
- isomorphism invariant

[de Silva et al, 2016]

Unfortunately, it's Graph-Isomorphism hard to compute. (Not quite NP-Hard)

↳ Essentially, can use 0-interleaved to check for graph isomorphism.

Functional distortion

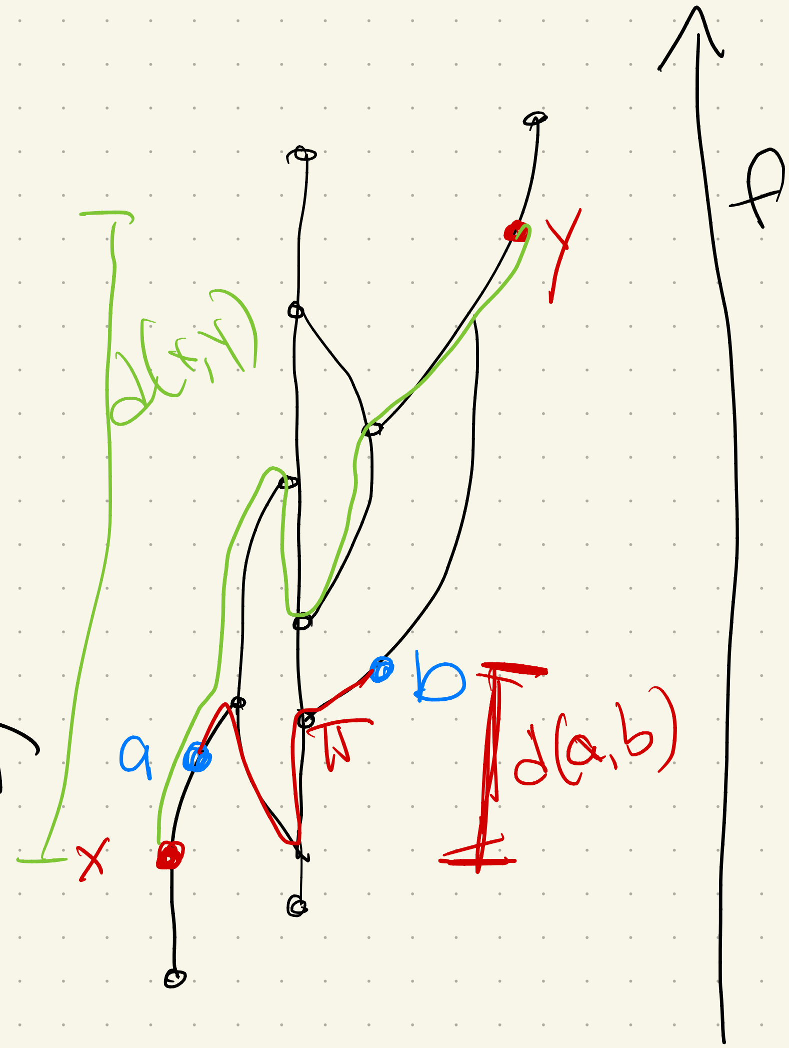
Based on Gromov-Hausdorff idea:

Height of a path:

given path π,

$$\max_{x \in \pi} f(x) - \min_{x \in \pi} f(x)$$

$d(x,y)$ = height of smallest such π

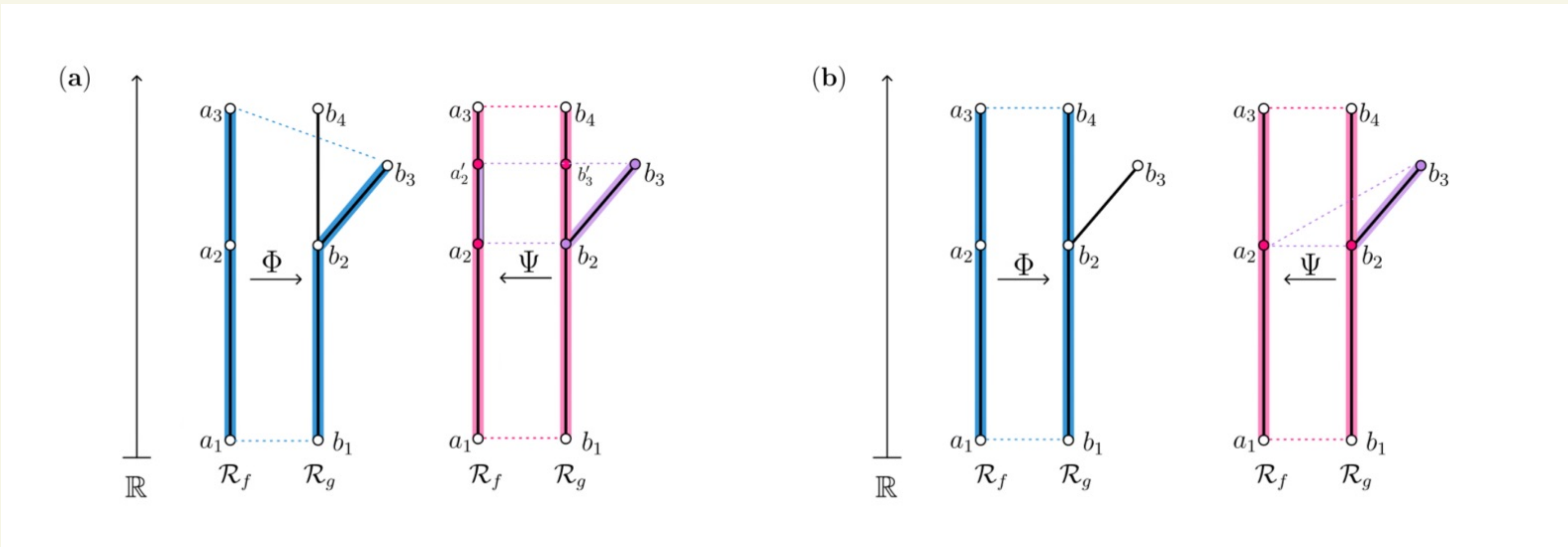


Point distortion:

Fix maps $\Psi : R_f \to R_g$

$\Phi : R_g \to R_f$

(continuous, but not function preserving)

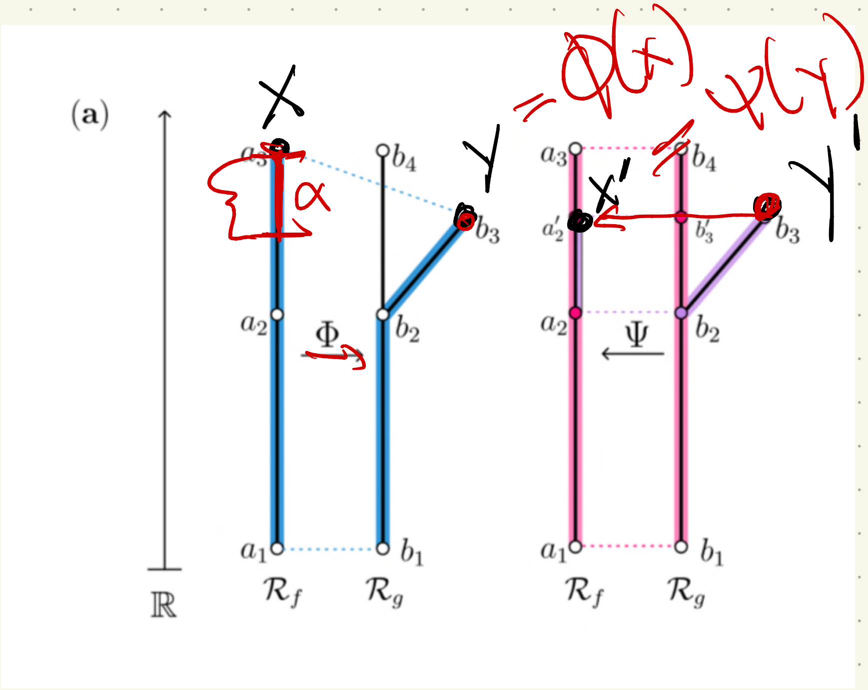


Note: Many choices here!

$$G(\Psi, \Phi) = \{(x, \Psi(x)) \mid x \in R_f\} \cup \{(\Phi(y), y) \mid y \in R_g\}$$

Point distortion For $(x,y), (x',y') \in G(\Phi, \Psi)$,

$$\Lambda((x,y),(x',y')) = \frac{1}{2}\left| d_f(x,x') - d_g(y,y') \right|$$

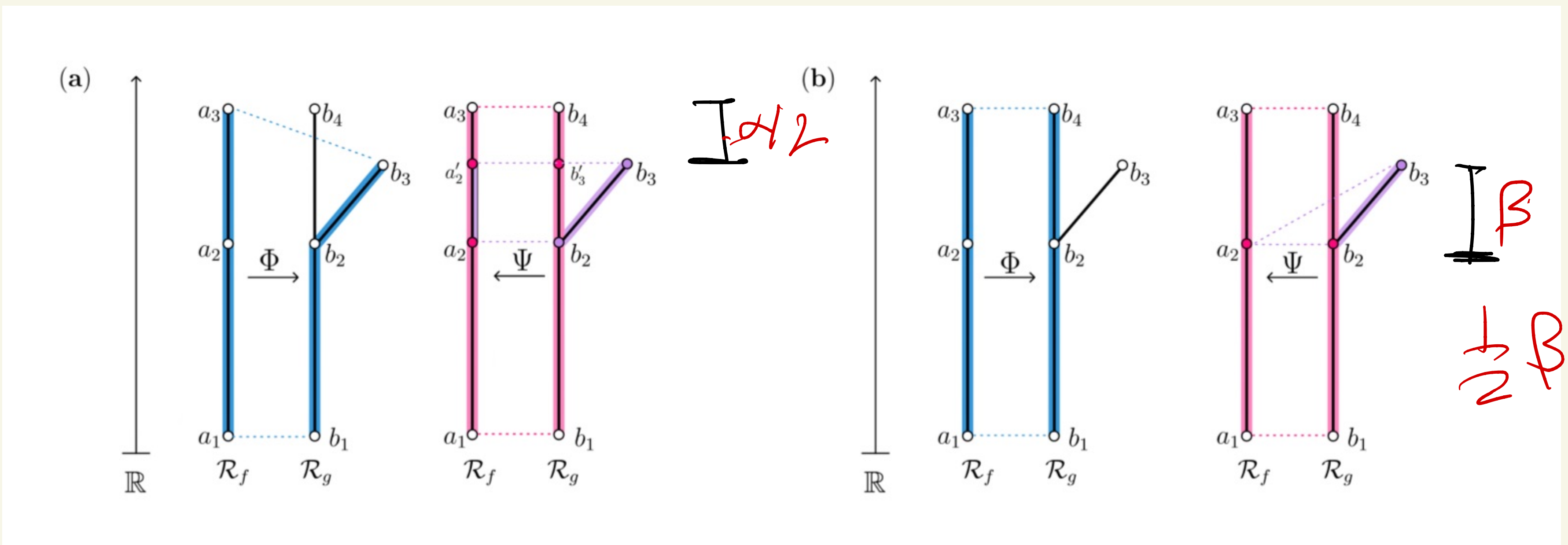


$$d = \frac{1}{2}\alpha$$

Map distortion:

supremum over all pairs of points of the point distortion

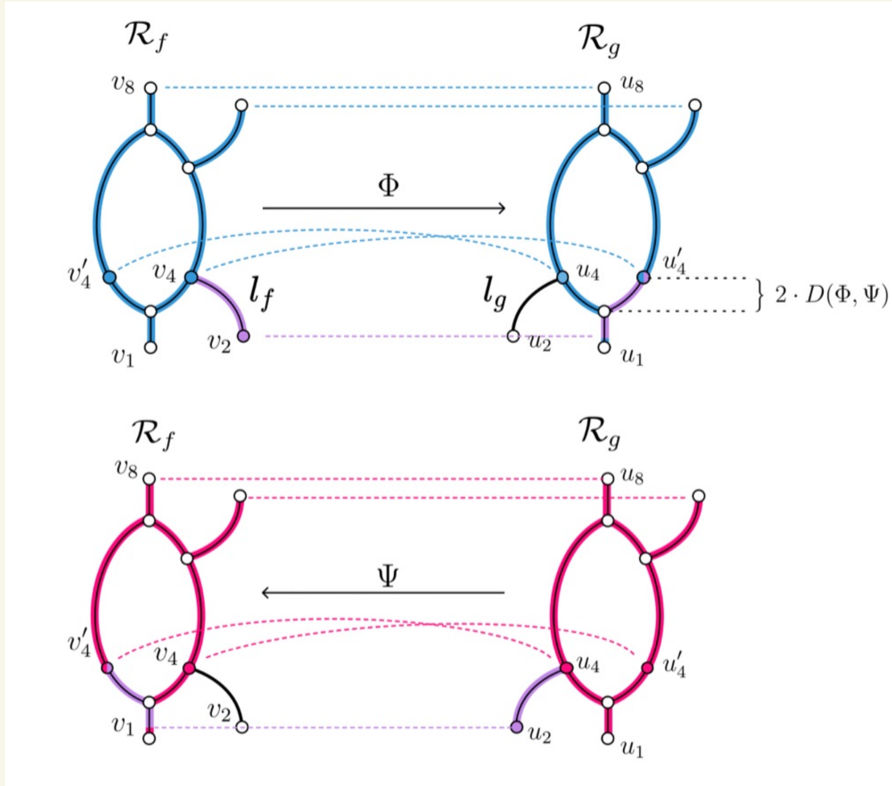
Note: all depends upon maps Ψ & Φ !



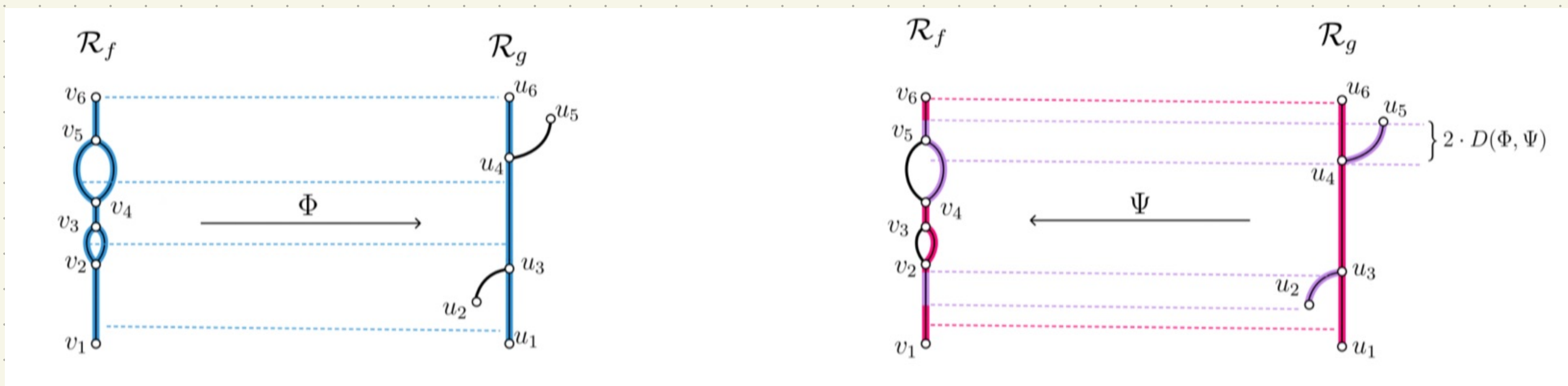
Functional distortion

$$d_{FD}(R_f, R_g) = \inf_{\Phi, \Psi} \max \{ \underbrace{D(\Phi, \Psi)}_{\substack{\text{Map distortion} \\ \text{(intrinsic to graph)}}}, \underbrace{\|f - g \circ \Phi\|_\infty, \|f \circ \Psi - g\|_\infty}_{\text{global rescaling}} \}$$

Examples:



$\mathcal{R}_f$
$\mathcal{R}_g$
v_8
u_8
Φ
v_4'
v_4
l_f
l_g
u_4
u_4'
$2 \cdot D(\Phi, \Psi)$
v_2
u_2
v_1
u_1
$\mathcal{R}_f$
$\mathcal{R}_g$
v_8
u_8
Ψ
v_4'
v_4
u_4
u_4'
v_2
v_1
u_2
u_1



$\mathcal{R}_f$
$\mathcal{R}_g$
v_6
u_6
v_5
u_5
u_4
v_4
Φ
v_3
v_2
u_3
u_2
v_1
u_1
$\mathcal{R}_f$
$\mathcal{R}_g$
v_6
u_6
v_5
u_5
$2 \cdot D(\Phi, \Psi)$
u_4
v_4
Ψ
v_3
v_2
u_3
u_2
v_1
u_1

Functional distortion pros + cons

It is:
- stable
- discriminative

[Bauer et al 2014]

- isomorphism invariant
- path component sensitive

[Bauer et al 2016]

No idea if it's computable in any efficent way.
(or even at all, except on trees!)

Edit distances:

Edit distances are well studied for strings and abstract graphs.
↳ see [Bille 2005] for the many variants on graphs.

In a graph, usually have:

- vertex insertion / deletion
- edge insertion / deletion

and some cost associated with each operation.

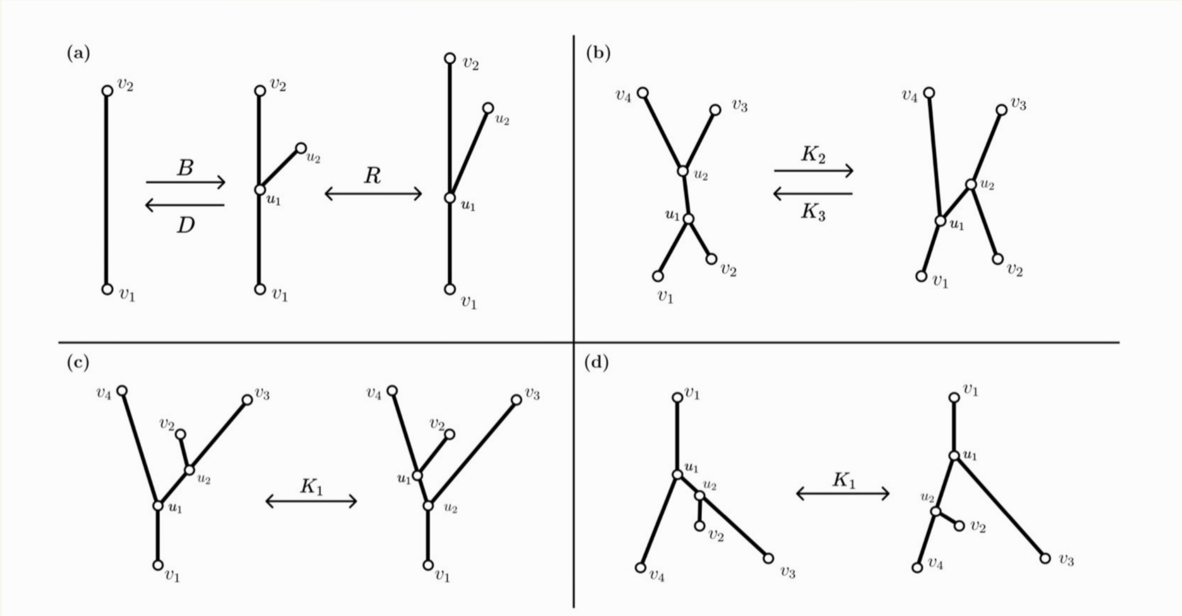
$$\text{Total Edit distance} = \min\left\{\sum(\text{edit costs})\right\}$$

Edits:

In Reeb graphs, only certain edits will correspond to topological deformations on original space.

[Di Fabio & Landi, 2012 and 2016]

- insert
- delete
- relabel
- K-type

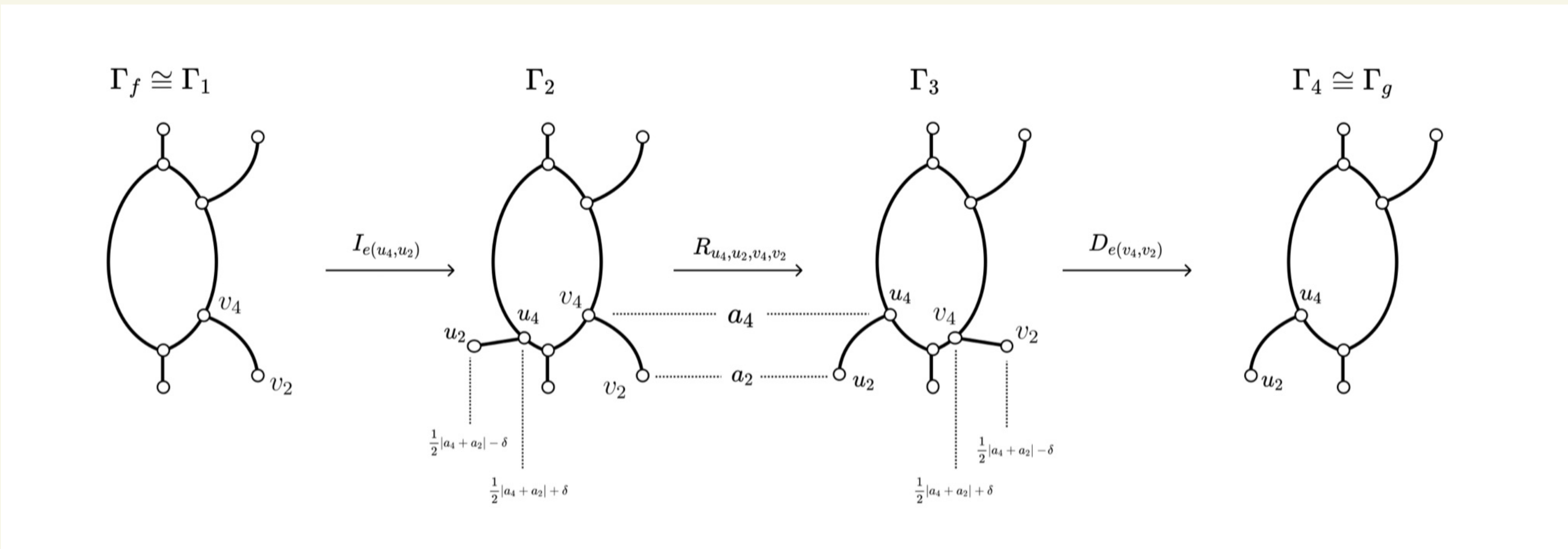


(Reason: Morse theory!)

Downside

Could only compare homeomorphic spaces, originally.

Also a more subtle flaw: can pay less by infinitesmally shifting vertices, since relabels cost max of operations.



Categorical edit distance

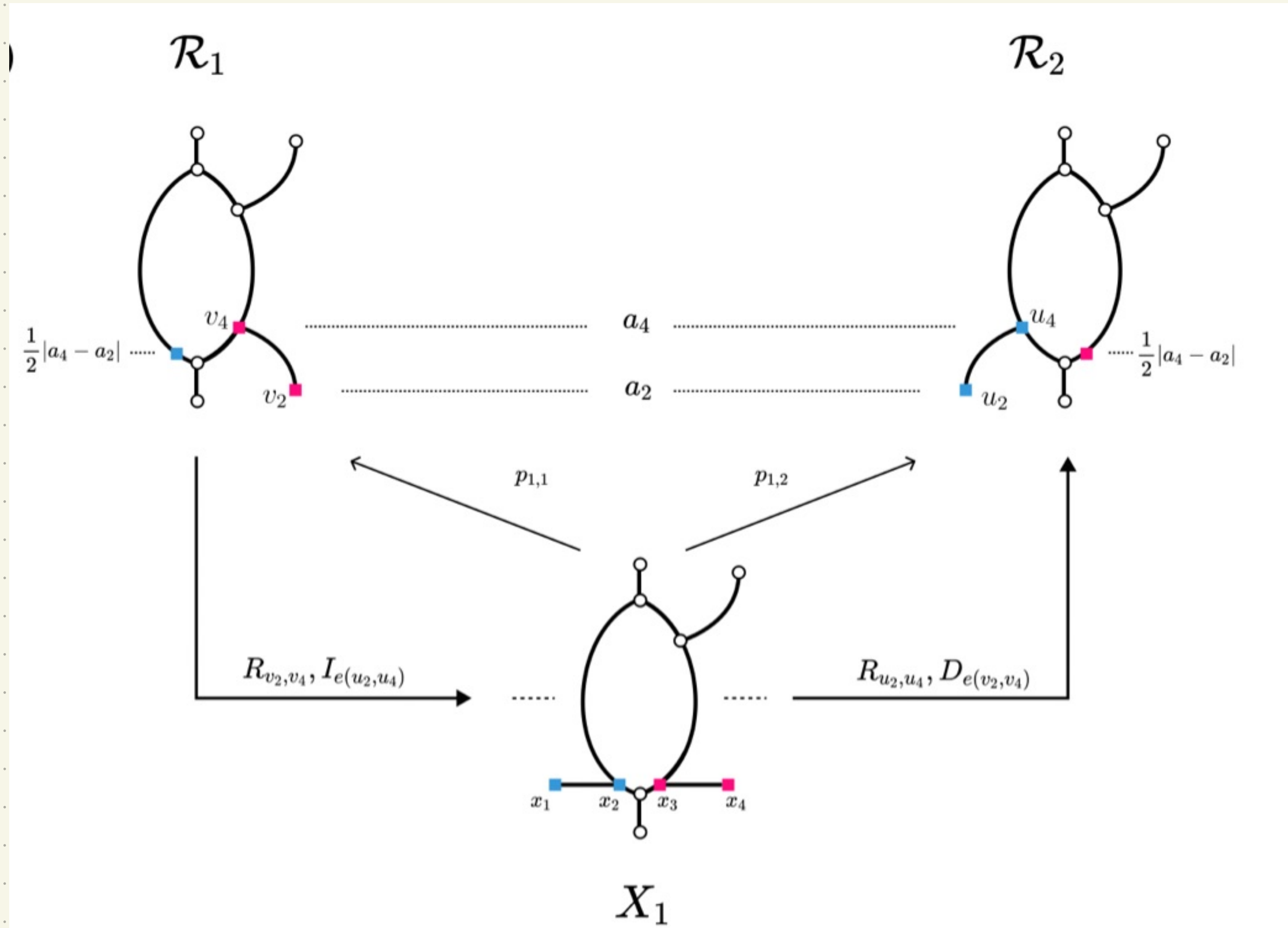
[Bauer et al 2021]

Instead of "direct" edits to Reeb graphs, goal it to find a sequence of topological spaces making this diagram commute.

$$
\begin{tikzcd}
\mathbb{R} & & \mathbb{R} & & & & \mathbb{R} & & \mathbb{R} \\
\mathcal{R}_f = \mathcal{R}_1 \arrow[u, "\tilde{f}_1"] & & \mathcal{R}_2 \arrow[u, "\tilde{f}_2"] & & \cdots & & \mathcal{R}_{n-1} \arrow[u, "\tilde{f}_{n-1}"] & & \mathcal{R}_n = \mathcal{R}_g \arrow[u, "\tilde{f}_n"] \\
& X_1 \arrow[ul, "p_{1,1}"] \arrow[ur, "p_{1,2}"] & & X_2 \arrow[ul, "p_{2,2}"] \arrow[ur, dotted] & & X_{n-2} \arrow[ul, dotted] \arrow[ur, "p_{n-2,n-1}"] & & X_{n-1} \arrow[ul, "p_{n-1,n-1}"] \arrow[ur, "p_{n-1,n}"] &
\end{tikzcd}
$$

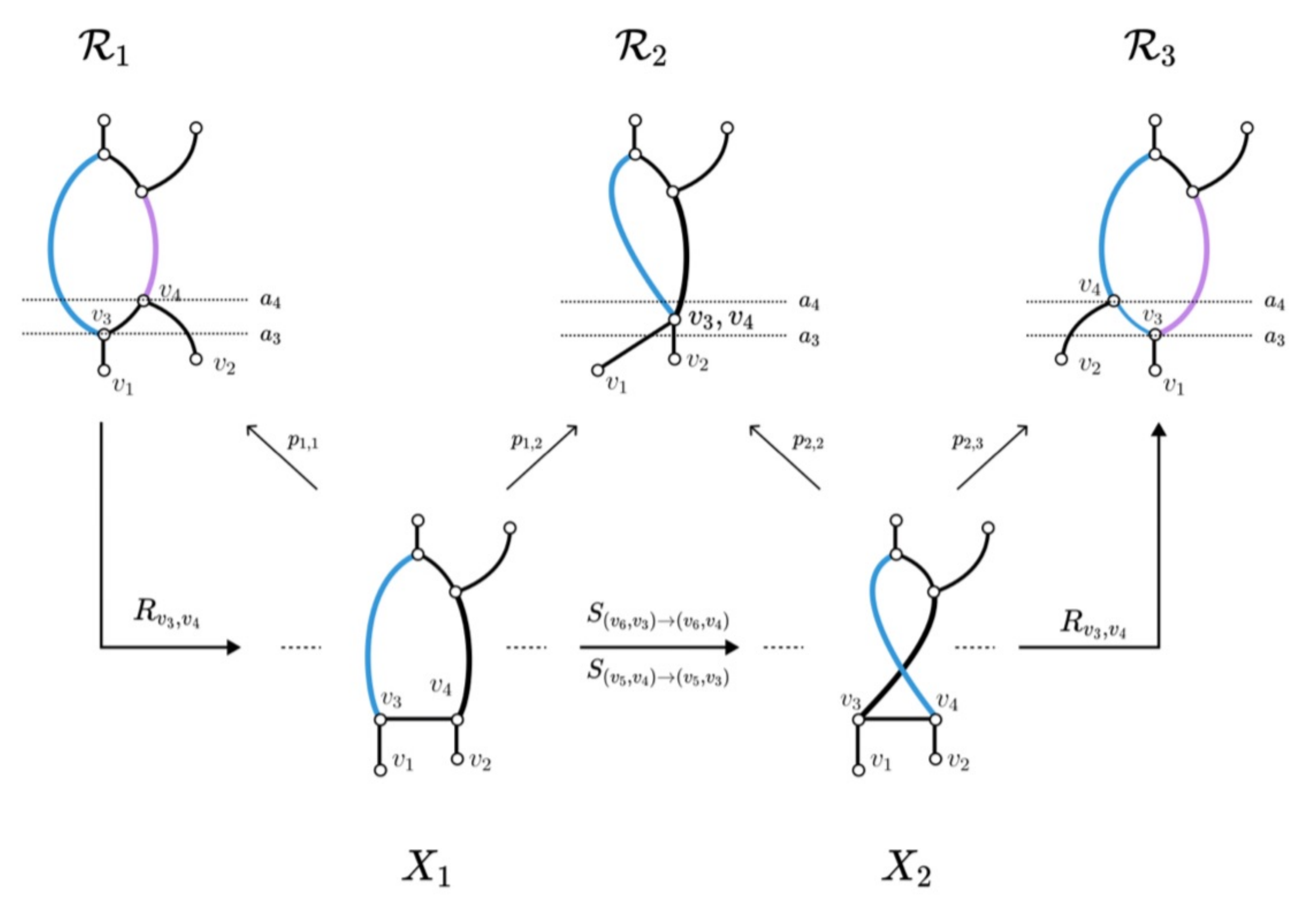
Then, pullback (or cone) of this diagram must exist, + can measure how much things more across the maps.

Example:



$\mathcal{R}_1$
$\mathcal{R}_2$
v_4
v_2
$\frac{1}{2}|a_4 - a_2|$
a_4
a_2
u_4
u_2
$\frac{1}{2}|a_4 - a_2|$
$p_{1,1}$
$p_{1,2}$
$R_{v_2,v_4}, I_{e(u_2,u_4)}$
$R_{u_2,u_4}, D_{e(v_2,v_4)}$
x_1
x_2
x_3
x_4
X_1

Alternative:



Categorical edit distance pros + cons:

It is:
- stable
- discriminative
- isomorphism invariant
- path component sensitive

and universal:
for any stable distance S,

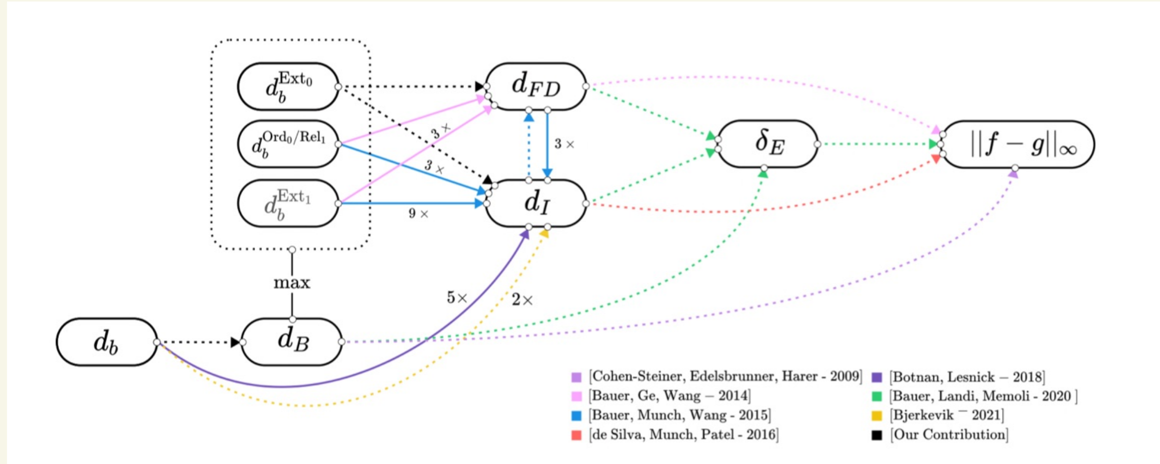
$$d_S(R_f, R_g) \leq d_E(R_f, R_g)$$

[Bauer et al 2020]

Unfortunately, no idea how to compute it!

Take-aways:

① This are very connected:



In fact, we conclude our survey with a conjecture:

Conjecture 9.3. *The functional distortion distance, interleaving distance, and universal edit distance are equivalent on the space of PL Reeb graphs where the domain $\mathbb{X}$ is simply connected. That is*

$$d_B \leq d_I = d_{FD} = \delta_E.$$

(up to constant factors)

② They do capture different types of features, even though there are connections!

	d_B	d_I	d_I^m	d_{FD}	d_E	δ_E
Stable	[27]	[62]	-	[4]	[35]	[5]
Discriminative	-	[6], 7.24	[22], C.3	[4], 7.20	[35]	[5]
Isomorphism Invariant	-	[62]	[22]	[6], 7.10	[35]	[5]
Path Component Sensitive	-	[62]	[22]	[6]	-	[5], 7.39
Universal	-	-	-	-	-	[5]

Table 1: Table of distance properties. Entry corresponds to a citation where the distance was proved. We supply additional references to statements we contributed in this work which solidify these properties. We denote disproven properties or properties that are not applicable to the given distance with "-".

We consider 4 simple graph classes, to find differences:

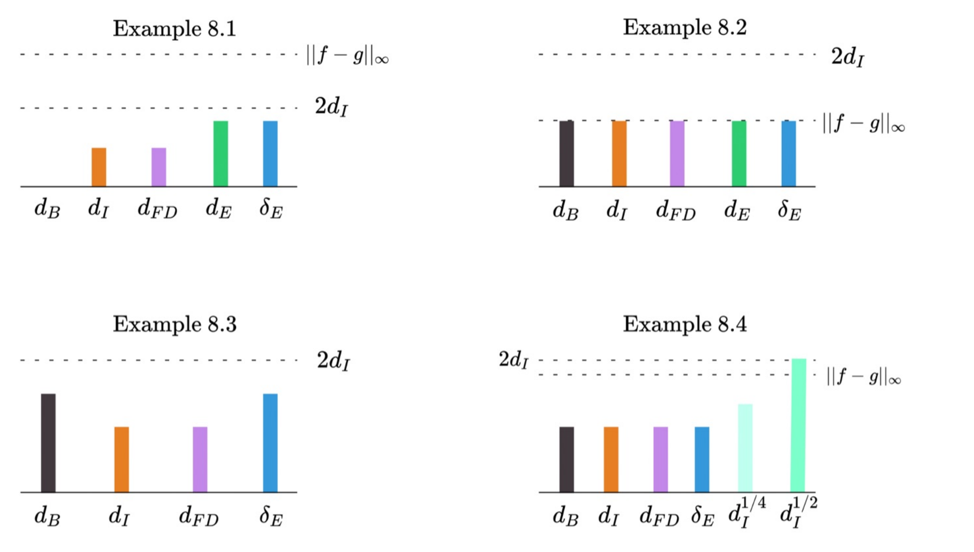


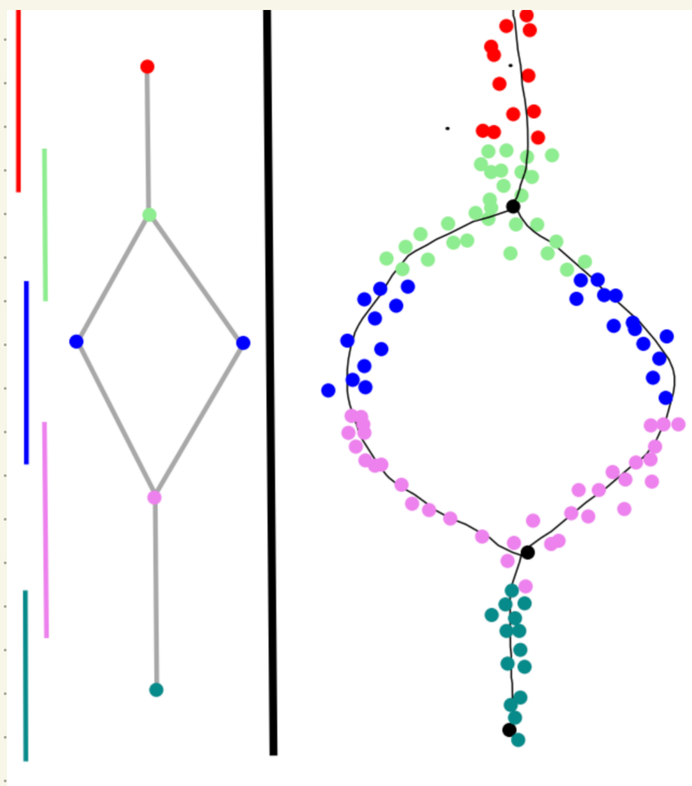
Figure 16: Visual summary of the distance values attained for each example. For Example 4, we show two different choices of m for the truncated interleaving distance to illustrate the affect that m plays on the distance values.

Mapper graphs

Idea: Approximate Reeb graphs

Won't always have a PL-space!

What about point clouds?



Idea:

- Give $\mathbb{R}$-values to dots
- Use a cover of $\mathbb{R}$
- Cluster into components + build a graph

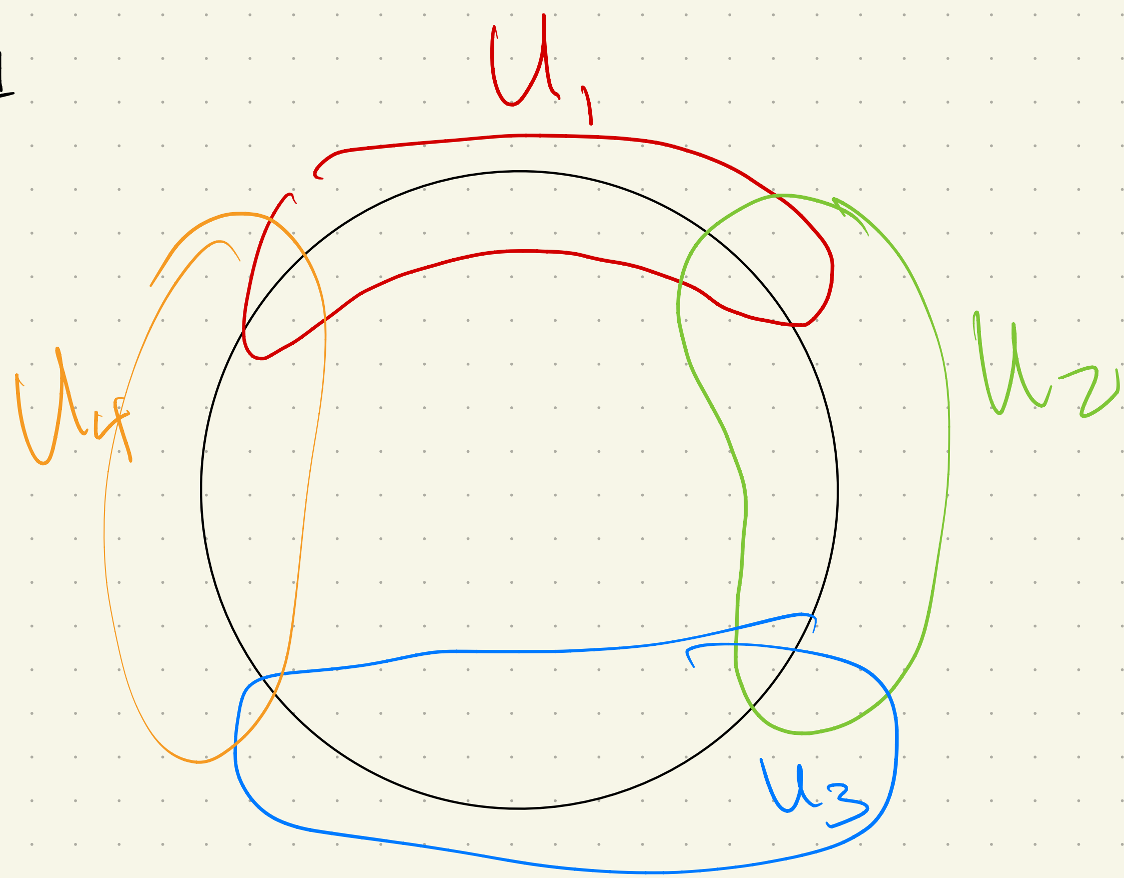
More details:

A cover of a set X is a collection of sets $\mathcal{U} = \{U_1, \dots, U_k\}$ s.t. $X \subseteq \bigcup_i U_i$

open cover $\rightarrow$ each U_i open

Ex: S^1

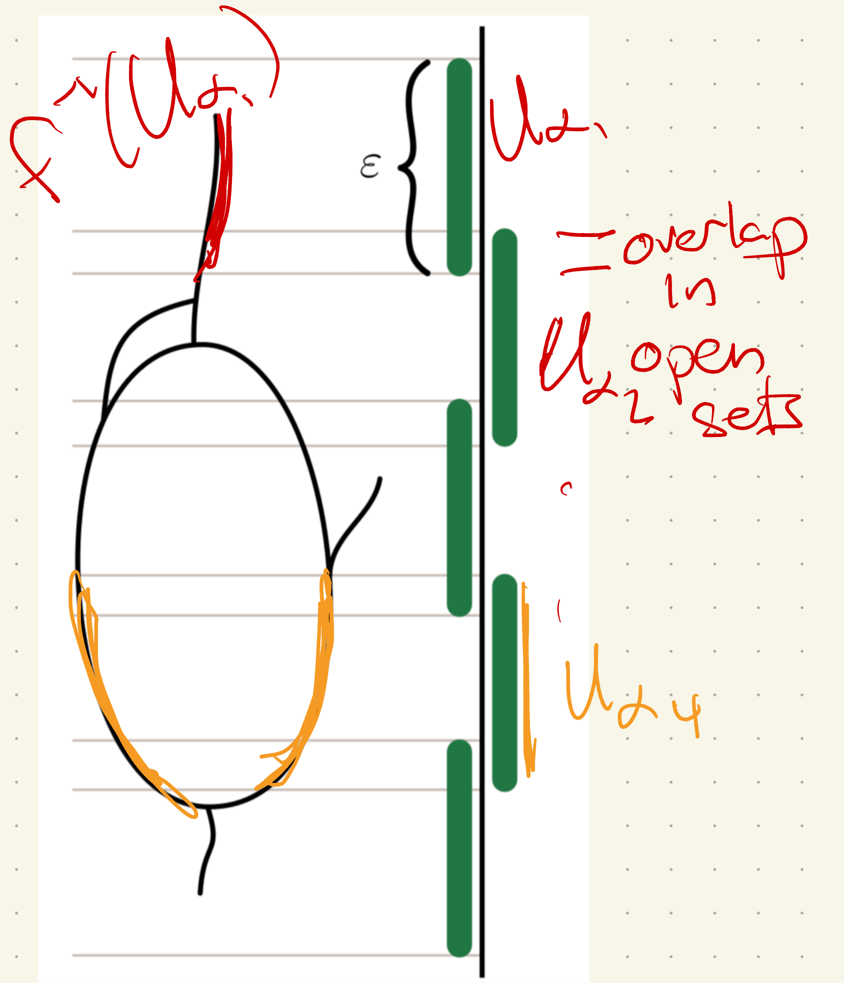
Cover:



Let's start on a simplicial complex:

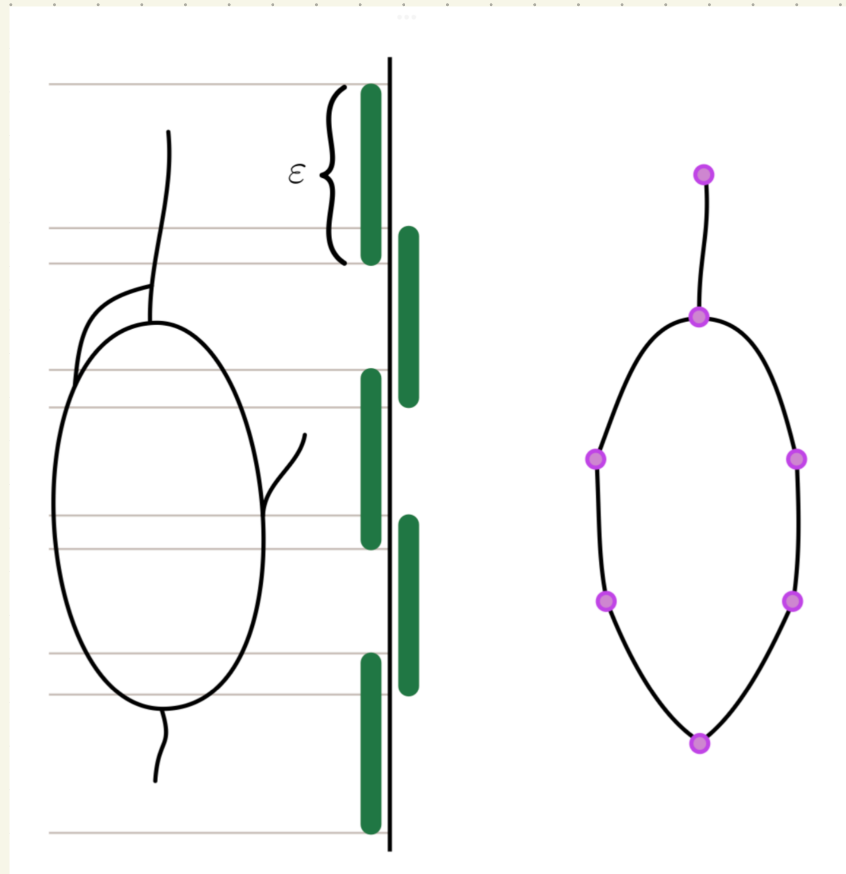
- Given $f : |K| \to \mathbb{R}$
- Fix a cover $\mathcal{U} = \{U_\alpha\}$ of $\mathbb{R}$
- The collection $f^{-1}(\mathcal{U}) = \{f^{-1}(U_\alpha)\}$ is a cover of X.

$\mathbb{R}$ contained in $\bigcup U_\alpha$



Mapper for simplicial complex (part 2)

- Let $f^{-1}(U)$ be the cover which splits sets into connected components



- Then, Mapper is the nerve of this cover

Recall: Given a finite collection of sets $\mathcal{F}$ in $\mathbb{R}^n$, the nerve is

$$\mathrm{Nrv}(\mathcal{F}) = \{X \subseteq \mathcal{F} \mid \bigcap_{U \in X} U \neq \emptyset\}$$

Let's try: