The Reeb Graph Edit Distance is Universal

Ulrich Bauer TUM

TGDA @ OSU TRIPODS Workshop
Theory & Foundations of TGDA

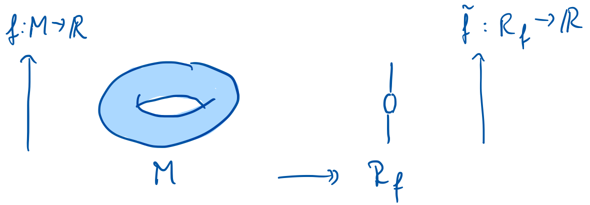
May 21, 2018

Joint work with Claudia Landi (U Modena)
and Facundo Mémoli (Ohio State U)

"Here are two things that are reasonably close to each other, and I want to compare them."

S. WEINBERGER

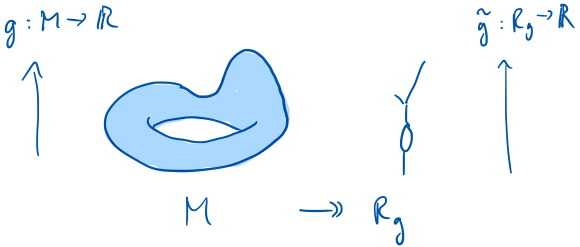
REEB GRAPHS



identify components of level sets $f^{-1}(t)$:

$$R_f = M/\sim_f ,$$

where $x \sim_f y \iff x, y$ in same component of some $f^{-1}(t)$, $t \in \mathbb{R}$.



Formal Setting

We consider

- locally compact Hausdorff spaces (Reeb domains)
- proper quotient maps with connected fibers (Reeb quotient maps)

These maps are closed under composition, and stable under pullbacks.

Define a Reeb graph as

- a Reeb domain R_f with
- a function $\tilde{f} : R_f \to \mathbb{R}$ with discrete fibers (Reeb function)

A Reeb graph R_f is the Reeb graph of a function $f : X \to \mathbb{R}$ if

- $f = \tilde{f} \circ p$ for some Reeb quotient map $p : X \twoheadrightarrow R_f$.

In this case, $R_f \cong X/{\sim_f}$.

Moreover: let $q : Y \twoheadrightarrow X$ be a Reeb quotient map.
Then R_f is also the Reeb graph of $g = f \circ q$.

$$\begin{tikzcd} Y \arrow[r, two heads, "q"] \arrow[drr, bend right, "g"'] & X \arrow[r, two heads, "p"] \arrow[dr, "f"'] & R_f \arrow[d, "\tilde{f}"] \\ & & \mathbb{R} \end{tikzcd}$$

- Reeb quotient maps preserve Reeb graphs.

GOALS

How to compare two Reeb graphs R_f, R_g? ($f, g : M \to \mathbb{R}$ are unknown)

Assign distance (extended pseudo-metric) $d(R_f, R_g)$.

Desirable properties:

Stability: For any space X and $f, g : X \to \mathbb{R}$ yielding Reeb graphs R_f, R_g,

$$d(R_f, R_g) \leq \|f - g\|_\infty .$$

Universality: For any other stable distance d_S,

$$d_S(R_f, R_g) \leq d(R_f, R_g) .$$

A CANONICAL UNIVERSAL DISTANCE

Given Reeb graphs R_f, R_g with functions $\tilde{f}, \tilde{g}$, define

$$d_u(R_f, R_g) = \inf \quad \|f-g\|_\infty$$

$$\begin{tikzcd}
& X \arrow[dl, "p_f"'] \arrow[dr, "p_g"] \arrow[ddl, "f"] \arrow[ddr, "g"'] & \\
R_f \arrow[d, "\tilde{f}"'] & & R_g \arrow[d, "\tilde{g}"] \\
\mathbb{R} & & \mathbb{R}
\end{tikzcd}$$

taken over all Reeb domains X and Reeb quotient maps p_f, p_g.

- stability: infimum is lower bound on all $\|f-g\|_\infty$
- universality: infimum is greatest lower bound

TRIANGLE INEQUALITY FOR THE CANONICAL UNIVERSAL DISTANCE

Triangle inequality: for all

$$\begin{tikzcd}
 & X \arrow[dl, two heads] \arrow[dr, two heads] & & Y \arrow[dl, two heads] \arrow[dr, two heads] & \\
R_f & & R_g & & R_h
\end{tikzcd},$$

consider pullback

$$\begin{tikzcd}
 & & P \arrow[dl] \arrow[dr] \arrow[d, phantom, "\vee"] \arrow[dddll, bend right, "f_P"'] \arrow[ddd, bend left, "g_P"] \arrow[dddrr, bend left, "h_P"] & & \\
 & X \arrow[dl, two heads] \arrow[dr, two heads] \arrow[ddl, "f_X"] \arrow[ddr, "g_X"] & & Y \arrow[dl, two heads] \arrow[dr, two heads] \arrow[ddl, "g_Y"] \arrow[ddr, "h_Y"] & \\
R_f \arrow[d, "\tilde f"'] & & R_g \arrow[d, "\tilde g"'] & & R_h \arrow[d, "\tilde h"'] \\
\mathbb{R} & & \mathbb{R} & & \mathbb{R}
\end{tikzcd}$$

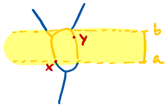
Then
$$\|f_P - h_P\| \leq \|f_P - g_P\| + \|g_P - h_P\|$$
$$= \|f_X - g_X\| + \|g_Y - h_Y\|.$$

- working with arbitrary spaces X is unfeasible

PREVIOUS WORK: FUNCTIONAL DISTORTION DISTANCE [B., Ge, Wang 2014]

- On a Reeb graph R_f with $\tilde{f}: R_f \to \mathbb{R}$, consider the metric

$$d_f : (x,y) \mapsto \inf \{ b-a \mid x,y \text{ in same component of } \tilde{f}^{-1}[a,b] \}.$$



- Given maps $\phi : R_f \to R_g$, $\psi : R_g \to R_f$, consider

$$G(\phi,\psi) = \{ (x, \phi(x)) \mid x \in R_f \} \cup \{ (\psi(y), y) \mid y \in R_g \}.$$

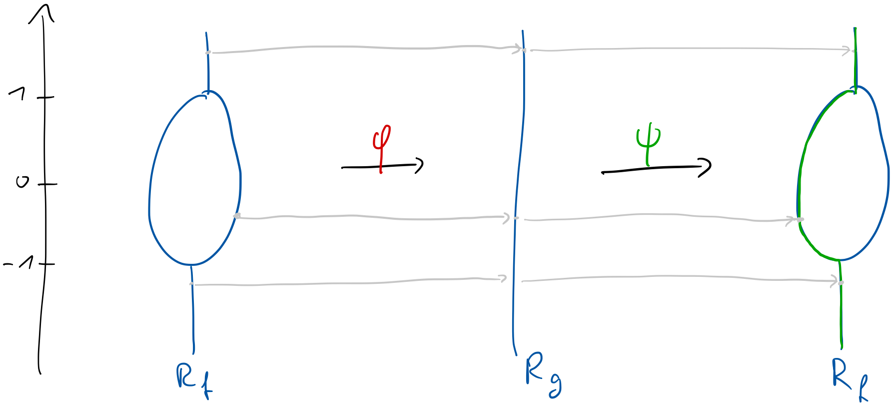
- Define the distortion of (ϕ,ψ) as

$$D(\phi,\psi) = \sup_{(x,y),(\tilde{x},\tilde{y}) \in G(\phi,\psi)} \frac{1}{2} \left| d_f(x,\tilde{x}) - d_g(y,\tilde{y}) \right|.$$

- Define the functional distortion distance as

$$d_{FD}(R_f, R_g) = \inf_{\phi,\psi} \left(\max \{ D(\phi,\psi), \| f - g\circ\phi \|_\infty, \| g - f\circ\psi \|_\infty \} \right).$$

EXAMPLE: FUNCTIONAL DISTORTION DISTANCE



$$D(\varphi, \psi) = \sup \frac{1}{2}\left(d(x, \tilde{x}) - d(y, \tilde{y})\right) = \frac{1}{2}$$

where $x, \tilde{x} \in R_f$, $y, \tilde{y} \in R_g$

with $\varphi(x) = y$ or $x = \psi(y)$,

$\varphi(\tilde{x}) = \tilde{y}$ or $\tilde{x} = \psi(\tilde{y})$

PREVIOUS WORK: INTERLEAVING DISTANCE [Bubenik & al. 2015; de Silva & al. 2016]

- Interpret Reeb graph R_f as a functor (in fact, a cosheaf)

$$F : \mathrm{Int}_{\mathbb{R}} \to \mathbf{Set}, \quad I \mapsto \pi_0(\tilde{f}^{-1}(I))$$

($\mathrm{Int}_{\mathbb{R}}$: open intervals, as a poset wrt. $\subseteq$)

- A δ-interleaving between F and G is a pair of natural transformations φ, ψ (with components $\varphi_I : F(I) \to G(B_\delta(I)), \dots$) such that

$$\begin{tikzcd}
F(I) \arrow[r] \arrow[dr, "\varphi_I"'] & F(B_\delta(I)) \arrow[r] \arrow[dr, "\varphi_{B_\delta(I)}"'] & F(B_{2\delta}(I)) \\
G(I) \arrow[r] \arrow[ur, "\psi_I"] & G(B_\delta(I)) \arrow[r] \arrow[ur, "\psi_{B_\delta(I)}"] & G(B_{2\delta}(I))
\end{tikzcd}$$

commutes for all $I \in \mathrm{Int}_{\mathbb{R}}$.

- $B_\delta(I)$: thickening I by δ
- Unlabeled maps: induced by inclusion

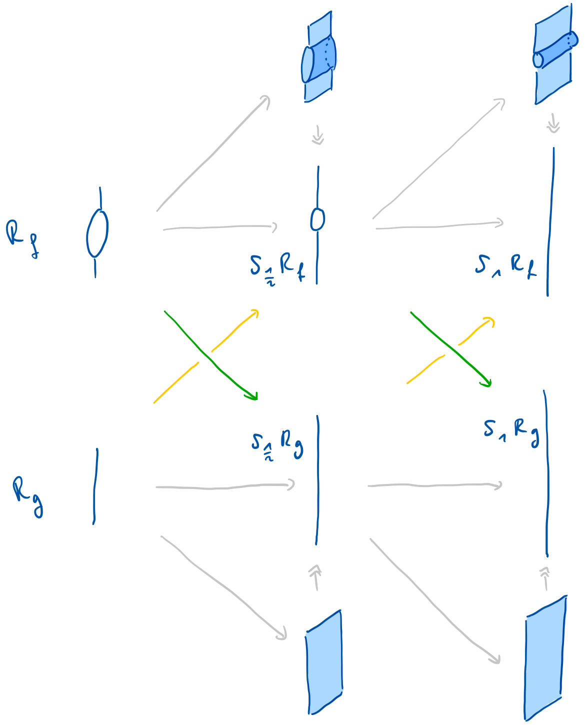
- The interleaving distance is

$$d_I(R_f, R_g) := \inf \{ \delta \mid \exists\ \delta\text{-interleaving between } F \text{ and } G \}$$

Thm [B., Munch, Wang 2015] $\frac{1}{3} d_{FD} \le d_I \le d_{FD}$.

Open problem:
is the lower bound tight?

ABSTRACT AND TOPOLOGICAL INTERLEAVINGS [de Silva & al. 2016]



Reeb graph $\rightsquigarrow$ cosheaf

$$R_f \rightsquigarrow F = \pi_0 \circ f^{-1}$$

$$S_\delta R_f \rightsquigarrow F \circ B_\delta$$

"smoothing" thickening

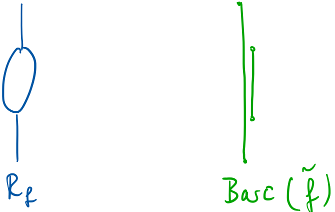
LEVEL SET PERSISTENT HOMOLOGY

Thm [Carlsson, de Silva, Morozov 2009]

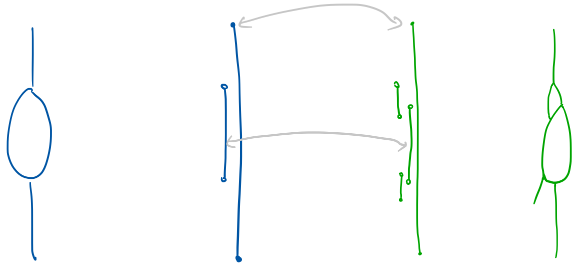
Given $f : X \to \mathbb{R}$ (PL, with X compact):

Homology of level sets $H_*(f^{-1}(t); \mathbb{F})$ (and more generally, of inclusions $f^{-1}(I) \hookrightarrow f^{-1}(J)$ for intervals $I \subseteq J$) is encoded (up to isomorphism) by a unique collection of intervals (level set persistence barcode).

Example for a Reeb graph:



THE BOTTLENECK DISTANCE BETWEEN PERSISTENCE BARCODES



A **δ-matching** between two barcodes $\mathrm{Barc}(f)$, $\mathrm{Barc}(g)$ satisfies:

- matched intervals (I, J) have distance $d_H(I, J) \leq \delta$
- unmatched intervals have length $\leq 2\delta$

The **bottleneck distance** $d_B(f, g)$ is

$$\inf \delta : \exists \ \delta\text{-matching between } \mathrm{Barc}(f), \mathrm{Barc}(g)$$

A Zoo of Distances and Inequalities

[Carlsson, de Silva, Morozov 2009]

$$d_B(R_f, R_g) \leq \|f-g\|_\infty$$

[B., Ye, Wang 2014]

$$\frac{1}{3} d_B(R_f, R_g) \leq d_{FD}(R_f, R_g) \leq \|f-g\|_\infty$$

[B., Munch, Wang 2015]

$$\frac{1}{3} d_{FD}(R_f, R_g) \leq d_I(R_f, R_g) \leq d_{FD}(R_f, R_g)$$

[Botnan, Lesnick 2016]

$$\frac{1}{5} d_B(R_f, R_g) \leq d_I(R_f, R_g)$$

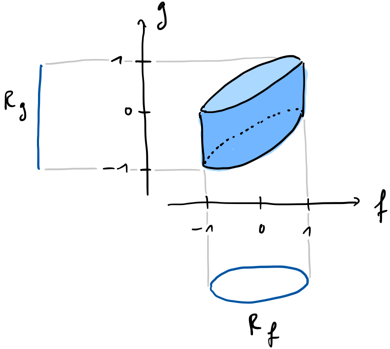
[Bjerkevik 2016]

$$\frac{1}{2} d_B(R_f, R_g) \leq d_I(R_f, R_g)$$

- d_B : bottleneck distance (of level set H_0 barcode)
- d_{FD} : functional distortion distance
- d_I : interleaving distance

FUNCTIONAL DISTORTION & INTERLEAVING DISTANCES ARE NOT UNIVERSAL

Consider a cylinder with two functions f, g:



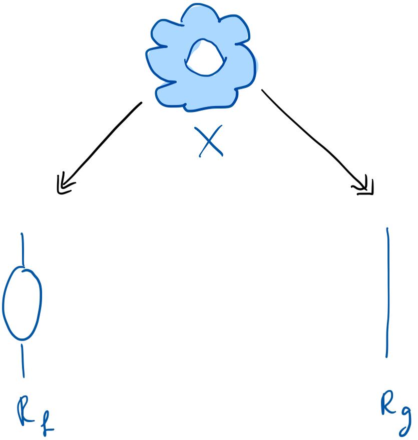
- $d_u(R_f, R_g) \leq \|f-g\|_\infty = 1$

- $d_I(R_f, R_g) \leq d_{FD}(R_f, R_g) \leq \frac{1}{2} < d_u(R_f, R_g)$:

FROM CLOSE REEB GRAPHS TO CLOSE FUNCTIONS

Open problem

Given two Reeb graphs R_f, R_g with $d_I(R_f, R_g) = \delta$.
Is there a space X with $f, g : X \to \mathbb{R}$, $\|f - g\| \leq C \cdot \delta$,
yielding Reeb graphs R_f, R_g, for some fixed constant C?



By the previous example:
if yes, then $C \geq 2$.

The Topological Edit Distance

- Consider zig-zag diagrams Z of Reeb quotient maps

$$\begin{tikzcd}
\mathbb{R} & & \mathbb{R} & & \mathbb{R} & & & \mathbb{R} \\
R_f = R_1 \arrow[u, "\tilde f_1"] & & R_2 \arrow[u, "\tilde f_2"] & & R_3 \arrow[u, "\tilde f_3"] & & & R_n = R_g \arrow[u, "\tilde f_n"] \\
& X_1 \arrow[ul, two heads] \arrow[ur, two heads] & & X_2 \arrow[ul, two heads] \arrow[ur, two heads] & & \cdots & X_{n-1} \arrow[ur, two heads] & \\
& & & L_Z \arrow[ull, two heads] \arrow[u, two heads] \arrow[urrr, two heads] & \cdots & & &
\end{tikzcd}$$

and take the limit L_Z (note: all maps are Reeb quotient maps).

- Define the *spread* of the functions $f_i : L_Z \twoheadrightarrow R_i \to \mathbb{R}$ as

$$s_Z : L_Z \to \mathbb{R}, \qquad x \mapsto \max_i f_i(x) - \min_j f_j(x).$$

(Note: $\|f-g\|_\infty \leq \|s_Z\|_\infty$, with equality for $n=2$)

- Define the (topological) *edit distance* as

$$d_{eTop}(R_f, R_g) = \inf_Z \|s_Z\|_\infty .$$

Prop. d_{eTop} is stable and universal.

THE REEB GRAPH EDIT DISTANCE

- Consider zig-zag diagrams Z of Reeb quotient maps

$$R_f = R_1 \twoheadleftarrow G_1 \twoheadrightarrow R_2 \twoheadleftarrow G_2 \twoheadrightarrow R_3 \quad \cdots \quad G_{n-1} \twoheadrightarrow R_n = R_g$$

as before, but restrict R_i, G_j in Z to be finite graphs.

Interpretation:

$$\begin{tikzcd}[column sep=small] & \mathbb{R} & \\ & R_{i+1} \arrow[u, "\tilde{f}_i"'] & \\ G_i \arrow[ur, two heads] & & G_{i+1} \arrow[ul, two heads] \end{tikzcd} \qquad \text{edit} \qquad \begin{tikzcd}[column sep=small] \mathbb{R} & & \mathbb{R} \\ R_i \arrow[u, "\tilde{f}_i"] & & R_{i+1} \arrow[u, "\tilde{f}_{i+1}"'] \\ & G_i \arrow[ul, two heads] \arrow[ur, two heads] & \end{tikzcd} \qquad \text{relabel}$$

modify G_i to G_{i+1}, maintaining the Reeb graph R_{i+1}

modify f_i to $f_{i+1} : G_i \to \mathbb{R}$, maintaining the domain G_i

- Define the Reeb graph edit distance analogously as

$$d_{eGraph}(R_f, R_g) = \inf_Z \|s_Z\|_\infty .$$

MAIN RESULT

Thm [B., Landi, Mémoli] The Reeb graph edit distance is stable & universal.

- We restrict to the (compact) PL category here.
- The hard part is stability:

  given $f, g : X \to \mathbb{R}$ (PL, for triangulation $X = |K|$),

  construct edit zigzag between R_f and R_g with spread $\leq \|f-g\|_\infty$?

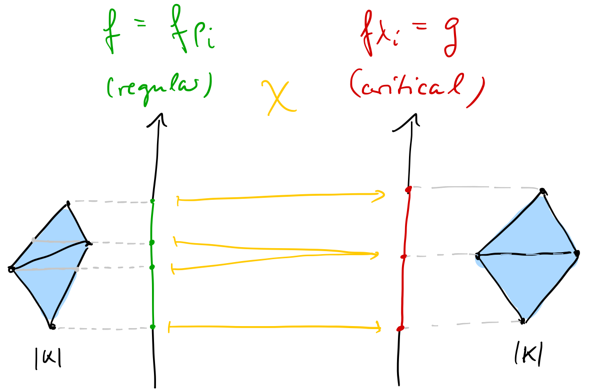
- Idea:
  - Consider straight-line homotopy $f_\lambda = \lambda f + (1-\lambda) g$
  - The structure of $R_\lambda = R_{f_\lambda}$ changes only finitely often (say, at parameters $0 = \lambda_0 < \dots < \lambda_n = 1$). Choose $p_i \in (\lambda_i, \lambda_{i+1})$.
  - Construct zigzag

$$\begin{tikzcd}[column sep=small]
R_f = R_{\lambda_0} & & R_{\lambda_i} & & R_{\lambda_{i+1}} & & R_{\lambda_n} = R_g \\
& \cdots \arrow[ul] \arrow[ur, dotted] & & R_{p_i} \arrow[ul] \arrow[ur] & & \cdots \arrow[ul, dotted] \arrow[ur] &
\end{tikzcd}$$

  - How to get the Reeb quotient maps in this zigzag?

CRITICAL INSTANTS OF A PL STRAIGHT-LINE HOMOTOPY

W.l.o.g.:



order-preserving surjection

$\chi : \operatorname{im} f \twoheadrightarrow \operatorname{im} g$ (PL)

- We have $\chi \circ f(v) = g(v) \quad \forall \; v \in \operatorname{Vert}(K)$
- But $\chi \circ f \neq g$!
- However: $\chi \circ f$ and g have the same Reeb graph...

LIFTING REPARAMETRIZATIONS

Lemma Let $h = \chi \circ f$. Then χ lifts to a Reeb quotient map $\zeta : R_f \to R_h$.

$$\begin{tikzcd}
\operatorname{im} f \arrow[r, two heads, "\chi"] & \operatorname{im} h \\
R_f \arrow[u, "\tilde{f}"] \arrow[r, two heads, "\zeta"] & R_h \arrow[u, "\tilde{h}"'] \\
{|K|} \arrow[u, two heads, "q_f"] \arrow[ur, two heads, "q_h"] \arrow[uu, bend left, "f"] \arrow[uur, bend right, "h"']
\end{tikzcd}$$

THE WORKHORSE: REEB QUOTIENT MAPS FROM INTERPOLATION

$\chi : \operatorname{im} f \to \operatorname{im} g$ (order-preserving PL surjection),

$$g(v) = \underbrace{\chi \circ f}_{h}(v) \qquad \forall\, v \in \operatorname{Vert} K .$$

$$\begin{tikzcd}
\operatorname{im} f \arrow[r, two heads, "\chi"] & \operatorname{im} h \arrow[r, phantom, "="] & \operatorname{im} g \\
R_f \arrow[u, "\tilde f"] \arrow[r, two heads] & R_h \arrow[u, "\tilde h"'] \arrow[r, "\cong"'] & R_g \arrow[u, "\tilde g"] \\
{|K|} \arrow[u, two heads, "q_f"] \arrow[uu, bend left, "f"] \arrow[ur, two heads, "q_h"'] \arrow[uur, bend right, "h"] & & {|K|} \arrow[u, two heads, "q_g"'] \arrow[uu, bend right, "g"'] \arrow[ul, dashed, "k"]
\end{tikzcd}$$

Lemma The relation

$$k = q_h \circ \big((h^{-1} \circ g) \cap st_K \big)$$

is a Reeb quotient map.

Corollary $R_h \cong R_g$, and χ lifts to a Reeb quotient map $R_f \twoheadrightarrow R_g$.

This provides the maps $R_{\lambda_i} \twoheadleftarrow R_{p_i} \twoheadrightarrow R_{\lambda_{i+1}}$ for our interpolation zigzag.

THE WORKHORSE: REEB QUOTIENT MAPS FROM INTERPOLATION

$\chi : \operatorname{im} f \to \operatorname{im} g$ (order-preserving PL surjection),

$$g(v) = \underbrace{\chi \circ f}_{h}(v) \qquad \forall\, v \in \operatorname{Vert} K.$$

$$\begin{tikzcd}
\operatorname{im} f \arrow[r, two heads, "\chi"] & \operatorname{im} h \arrow[r, phantom, "="] & \operatorname{im} g \\
R_f \arrow[u, "\tilde f"] \arrow[r, two heads] & R_h \arrow[u, "\tilde h"'] \arrow[r, "\cong"] & R_g \arrow[u, "\tilde g"'] \\
{|K|} \arrow[u, two heads, "q_f"] \arrow[uu, bend left, "f"] \arrow[uur, "h"] \arrow[ur, two heads, "q_h"'] & & {|K|} \arrow[u, two heads, "q_g"'] \arrow[uu, bend right, "g"'] \arrow[ul, dashed, two heads, "k"]
\end{tikzcd}$$

Lemma The relation

$$k = q_h \circ \left((h^{-1} \circ g) \cap st_K \right)$$

is a Reeb quotient map.

Corollary $R_h \cong R_g$, and χ lifts to a Reeb quotient map

$$R_f \twoheadrightarrow R_g.$$

This provides the maps $R_{\lambda_i} \twoheadleftarrow R_{\rho_i} \twoheadrightarrow R_{\lambda_{i+1}}$ for our interpolation zigzag.

CONCLUSION

- A universal distance is the most discriminative stable distance between Reeb graphs
- There is a simple construction of a universal distance
- Interleaving and functional distortion distances are not universal
- A universal distance in PL can be constructed using graph edit zigzags

Questions:

- What is the complexity of computing the distance?
- Is $d_u \leq C \cdot d_I$ for some constant C?