

# THE REEB GRAPH EDIT DISTANCE IS UNIVERSAL

ULRICH BAUER 

TRENDS IN PERSISTENT HOMOLOGY

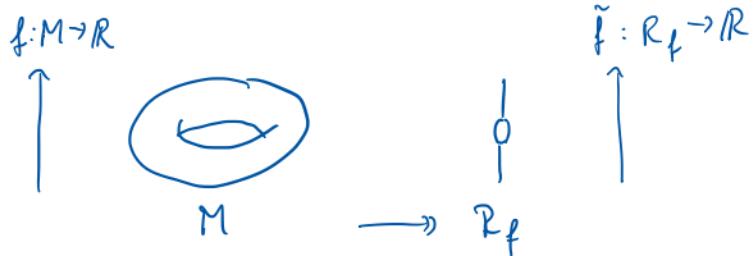
Ö}M{G  
D}V MEETING 2017, SALZBURG

JOINT WORK WITH CLAUDIA LANDI (UNIVERSITY OF MODENA)  
AND FACUNDO MÉMOLI (OHIO STATE U)

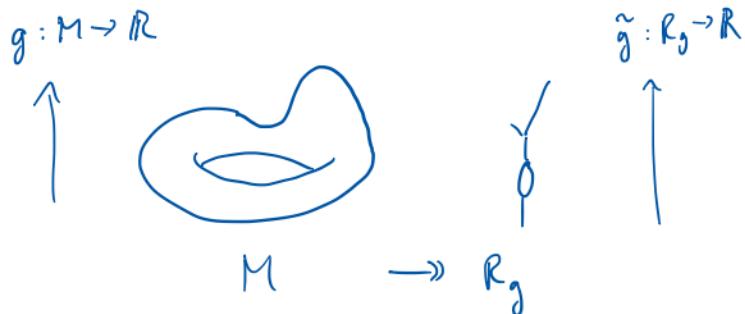
"Here are two things that are reasonably close to each other,  
and I want to compare them." 66

S. WEINBERGER, FoCM 2017

# REEB GRAPHS



identify components of level sets  $f^{-1}(t)$ :  $R_f = M / \sim_f$ , where  
 $x \sim_f y \Leftrightarrow x, y$  in same component of  $f^{-1}(t)$ ,  $t \in \mathbb{R}$ .



## Goals

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

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

Desirable properties:

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

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

Universality: For any other stable distance  $d_0$ ,

$$d_0(R_f, R_g) \leq d(R_f, R_g).$$

## FORMAL SETTING

We consider

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

This forms a category (which has all finite limits).

Define a Reeb graph as

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

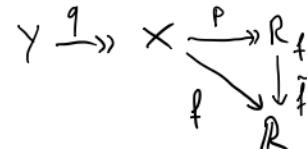
A Reeb graph  $R_f$  is the Reeb graph of a function  $f: X \rightarrow \mathbb{R}$  if

- $f = \tilde{f} \circ p$  for some Reeb quotient map  $p: X \rightarrow R_f$ .
- In this case,  $R_f \cong X / \sim_f$ .

Moreover: let  $q: Y \rightarrow X$  be a Reeb quotient map.

Then  $R_f$  is also the Reeb graph of  $f \circ q$ .

- Reeb quotient maps preserve Reeb graphs.



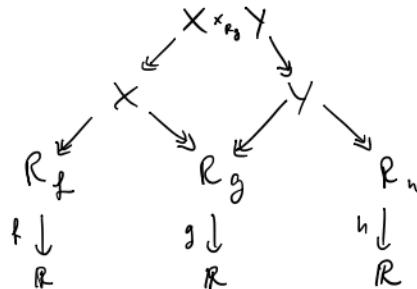
# A CANONICAL UNIVERSAL DISTANCE

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

$$d_u(R_f, R_g) = \inf_{\substack{f = \tilde{f} \circ p_f, \\ g = \tilde{g} \circ p_g : X \rightarrow \mathbb{R}}} \|f - g\|_\infty$$

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

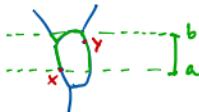
This is a distance (triangle inequality): consider pullbacks



- stability and universality immediate from definition
- working with arbitrary spaces  $X$  is unfeasible

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

- On a Reeb graph  $R_f$  with  $\tilde{f}: R_f \rightarrow \mathbb{R}$ , consider the metric  $d_f: (x, y) \mapsto \inf \{b-a | x, y \text{ are in same component of } \tilde{f}^{-1}[a, b]\}$ .



- Given maps  $\phi: R_f \rightarrow R_g$ ,  $\psi: R_g \rightarrow R_f$ , consider  $G(\phi, \psi) = \{(x, \phi(x)) \mid x \in R_f\} \cup \{\psi(\psi(y), y) \mid y \in R_g\}$ .

- Define the distortion of  $(\phi, \psi)$  as

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

- Define the functional distortion distance as

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

PREVIOUS WORK : INTERLEAVING DISTANCE [Bubenik, deSilva, Scott 2015;  
deSilva, Munch, Patel 2016]

- Interpret Reeb graph  $R_f$  as a functor  $F : \text{Int}_{\mathbb{R}} \rightarrow \text{Set}$ ,  $I \mapsto \pi_0(\tilde{f}^{-1}(I))$  ( $\text{Int}_{\mathbb{R}}$  are the open intervals, as a poset wrt.  $\subseteq$ )
- A  $\delta$ -interleaving between  $F$  and  $G$  is a pair of natural transformations  $\varphi, \psi$  (with components  $\varphi_I : F(I) \rightarrow G(B_\delta(I)), \dots$ ) such that

$$\begin{array}{ccccc}
 F(I) & \longrightarrow & F(B_\delta(I)) & \longrightarrow & F(B_{2\delta}(I)) \\
 & \searrow \varphi_I & \nearrow \psi_{F(I)} & & \nearrow \psi_{F(B_\delta(I))} \\
 G(I) & \longrightarrow & G(B_\delta(I)) & \longrightarrow & G(B_{2\delta}(I))
 \end{array}$$

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

(unlabelled maps: induced by inclusion)

- The interleaving distance is

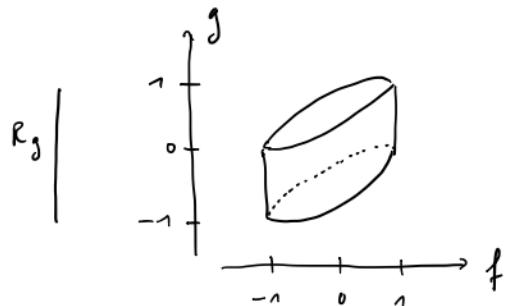
$$d_I(F, G) = \inf \{ \delta \mid \exists \text{ } \delta\text{-interleaving between } F \text{ and } G \}$$

$$d_{I, \pi_0}(R_f, R_g) := d_I(\pi_0 \circ \tilde{f}^{-1}, \pi_0 \circ \tilde{g}^{-1})$$

Thm [B., Munch, Wang 2015]  $\frac{1}{3}d_{FD} \leq d_{I, \pi_0} \leq d_{FD}$ .

# NON-UNIVERSALITY OF FUNCTIONAL DISTORTION & INTERLEAVING DISTANCES

Consider a cylinder with two functions  $f, g$ :



$R_f$

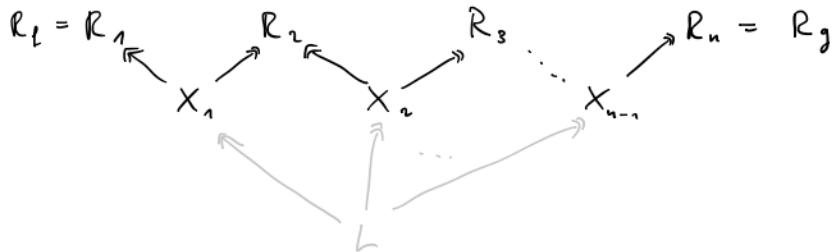
- $d_n(R_f, R_g) \leq \|f-g\|_\infty = 1$
- $d_B(Dgm_{\infty}(\mathbb{R}_f), Dgm_{\infty}(\mathbb{R}_g)) = 1 \leq d_n(R_f, R_g)$  (level set persistence is stable)
- $d_{I, \pi_0}(R_f, R_g) \leq d_{FD}(R_f, R_g) \leq \frac{1}{2} < d_n(R_f, R_g)$

$$R_f \circlearrowleft \xrightarrow{\phi} |R_g|$$

$$R_g | \xrightarrow{\psi} {}^{inv}\circlearrowright R_f$$

# THE TOPOLOGICAL EDIT DISTANCE

- Consider zig-zag diagrams  $\mathcal{Z}$  of Reeb quotient maps



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

Let  $q_i$  be the unique map  $L \rightarrow R_i$  in this diagram, and  $f_i = f_i \circ q_i$ .

- Define the spread of the functions  $f_1, \dots, f_n : L \rightarrow R_n \rightarrow R$  as

$$S_2 : L \rightarrow \mathbb{R}, \quad x \mapsto \max_i f_i(x) - \min_j f_j(x).$$

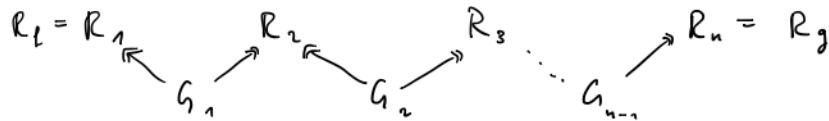
- Define the topological edit distance as

$$d_{\text{Top}}(R_f, R_g) = \inf_z \|s_z\|_\infty.$$

Prop.  $d_{top}$  is stable and universal.

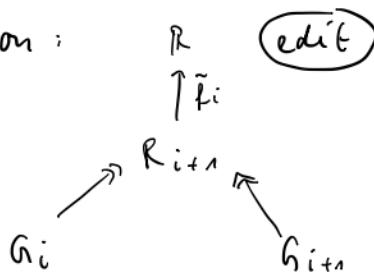
# THE REEB GRAPH EDIT DISTANCE

- Consider zig-zag diagrams  $\mathcal{Z}$  of Reeb quotient maps

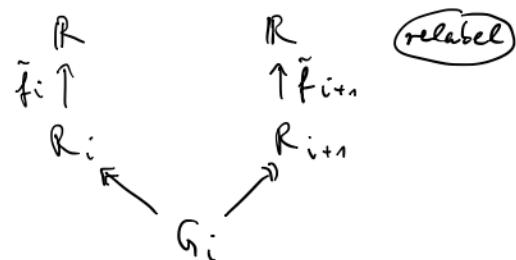


as before, but restrict  $R_i, G_j$  to finite graphs.

Interpretation:



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



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

- Define the Reeb graph edit distance analogously (using the limit) as

$$d_{\text{graph}}(R_f, R_g) = \inf_{\mathcal{Z}} \|s_{\mathcal{Z}}\|_{\infty}.$$

## MAIN RESULT

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

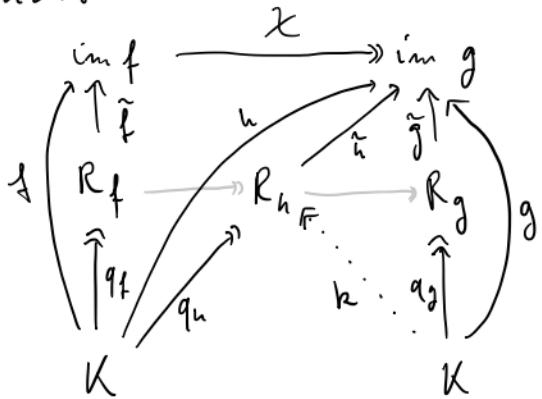
- We restrict to the PL category here.
- The hard part is stability:  
given  $f, g : K \rightarrow \mathbb{R}$  (PL), how to construct an edit zigzag  
between  $R_f$  and  $R_g$  with cost  $\leq \|f - g\|_\infty$ ?
- Idea:
  - Consider straight-line homotopy  $f_\lambda = \lambda f + (1-\lambda)g$
  - The structure of  $R_{f_\lambda} = R_{f_\lambda}$  changes only finitely often  
(say, at parameters  $0 = \lambda_0 < \dots < \lambda_n = 1$ . Choose  $\rho_i \in (\lambda_i, \lambda_{i+1})$ .
  - Construct zigzag  $R_f = R_0 \xrightarrow{\quad} R_{\lambda_1} \xrightarrow{\quad} R_{\lambda_2} \xrightarrow{\quad} R_{\lambda_{i+1}} \xrightarrow{\quad} \dots \xrightarrow{\quad} R_n = R_g$
  - But what are the maps here?

## THE WORKHORSE

Assume that there is a monotonic PL surjection  $\chi : \text{im } f \rightarrow \text{im } g$  such that  $g(v) = \chi \circ f(v)$  for every vertex  $v$  of  $K$ .

Then  $g$  is not the same as  $h = \chi \circ f$ , but has the same Reeb graph!

Construction:



Lemma The relation

$k = q_h \circ ((h^{-1} \circ g) \cap \text{st}_K)$   
is a Reeb quotient map.

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

This provides the maps

$$R_{\lambda_i} \xrightarrow{R_{\pi_i}} R_{\lambda_{i+1}} \quad \text{in our zigzag.}$$

## SUMMARY

- Interleaving and functional distortion distances are not universal
- There is a simple construction of a universal distance between Reeb graphs
- A universal distance in PL can be constructed using graph edit zigzags
- What is the complexity of computing the distance?
- Is there a constant  $C$  such that  $d_u \leq C \cdot d_I$ ?
- Further questions?