Category Representations and Convergence

Reeb Space and Mapper

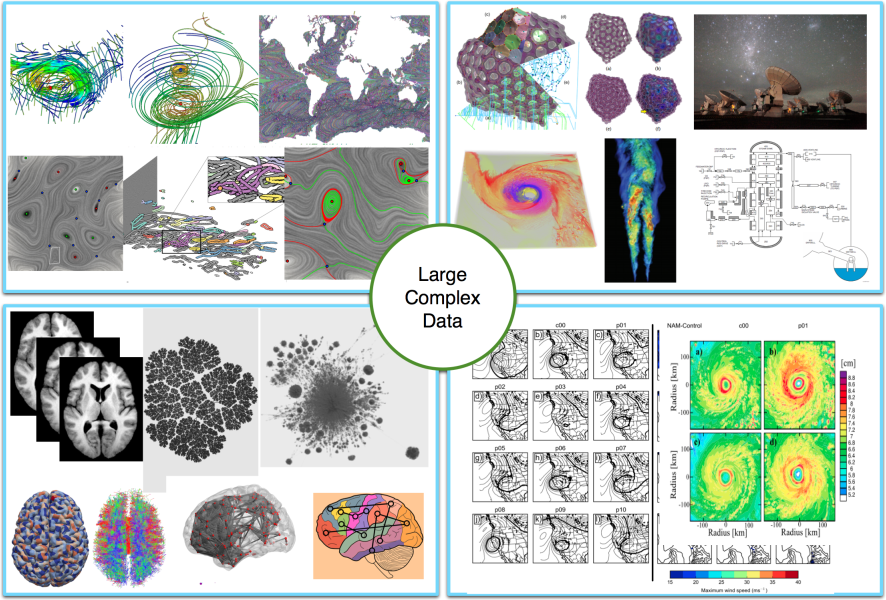
Bei Wang

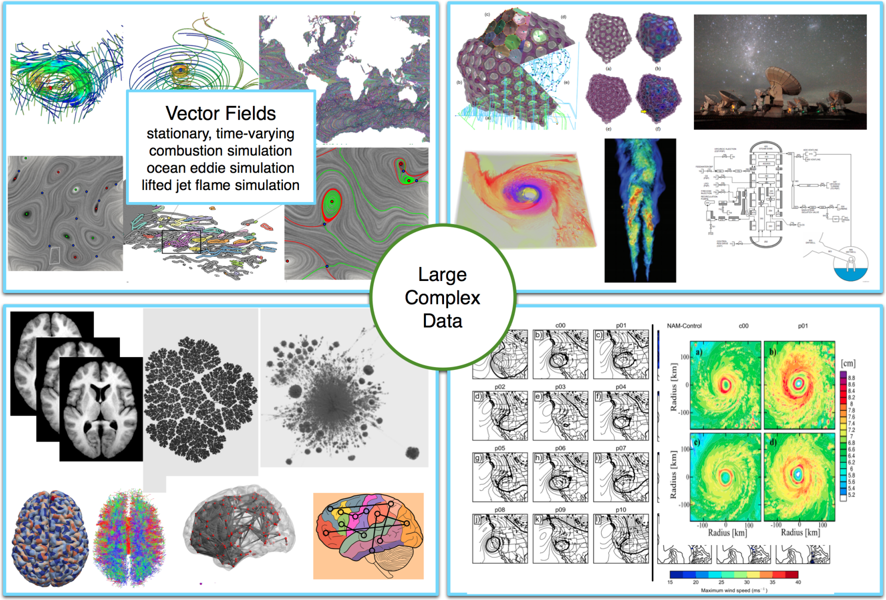
School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah

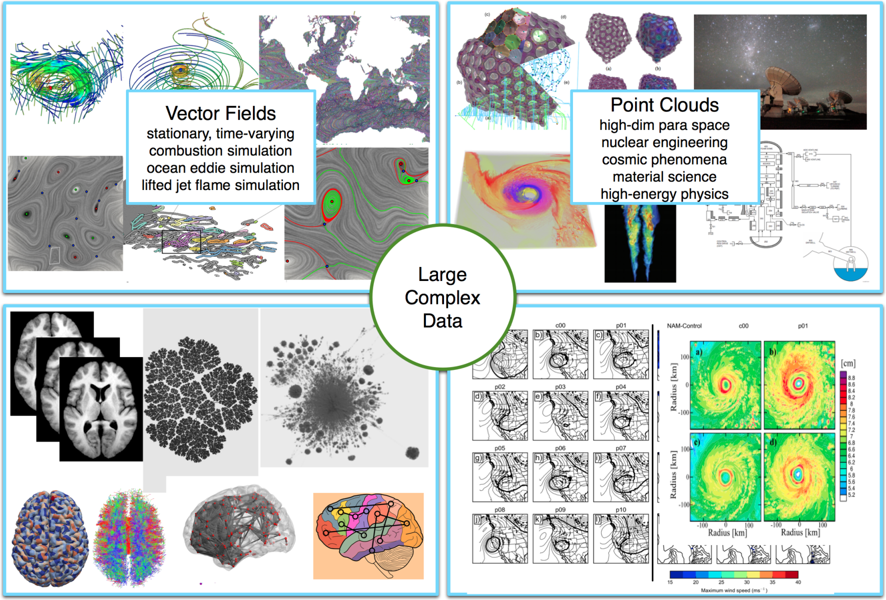
May 18, 2016

Joint work with Liz Munch (University at Albany - SUNY)

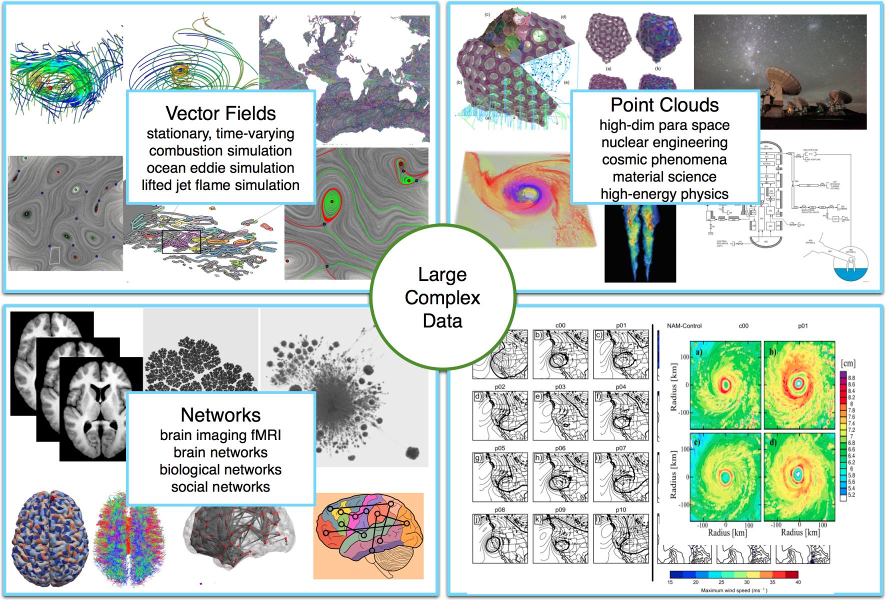
Motivation

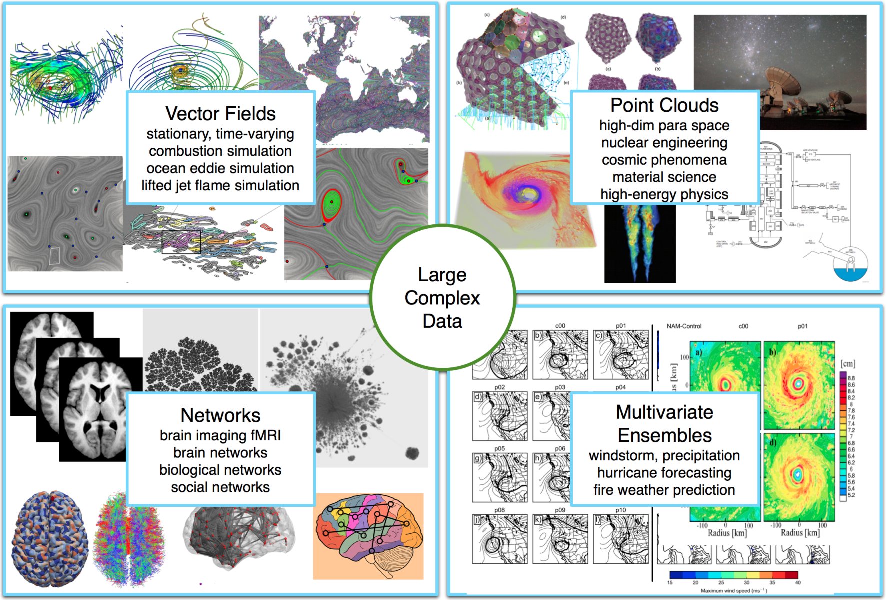

Large Complex Data
NAM-Control
c00
p01
p02
p03
p04
p05
p06
p07
p08
p09
p10
Radius [km]
[cm]
Maximum wind speed (ms⁻¹)

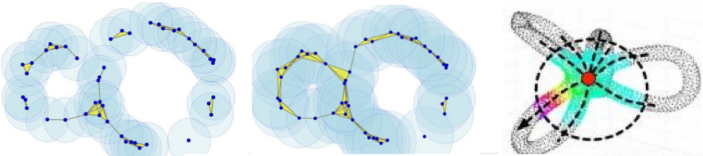

Vector Fields
stationary, time-varying
combustion simulation
ocean eddie simulation
lifted jet flame simulation
Large Complex Data
NAM-Control
c00
p01
Radius [km]
[cm]
Maximum wind speed (ms⁻¹)



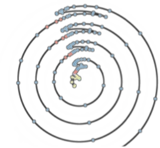
Vector Fields
stationary, time-varying
combustion simulation
ocean eddie simulation
lifted jet flame simulation
Point Clouds
high-dim para space
nuclear engineering
cosmic phenomena
material science
high-energy physics
Large Complex Data
NAM-Control
c00
p01
Radius [km]
[cm]
Maximum wind speed


Vector Fields
stationary, time-varying
combustion simulation
ocean eddie simulation
lifted jet flame simulation
Point Clouds
high-dim para space
nuclear engineering
cosmic phenomena
material science
high-energy physics
Large Complex Data
Networks
brain imaging fMRI
brain networks
biological networks
social networks
NAM-Control
c00
p01
Radius [km]
[cm]
Maximum wind speed


Vector Fields
stationary, time-varying
combustion simulation
ocean eddie simulation
lifted jet flame simulation
Point Clouds
high-dim para space
nuclear engineering
cosmic phenomena
material science
high-energy physics
Large
Complex
Data
Networks
brain imaging fMRI
brain networks
biological networks
social networks
Multivariate
Ensembles
windstorm, precipitation
hurricane forecasting
fire weather prediction

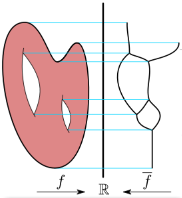


Persistent Homology, Cohomology, Local Homology

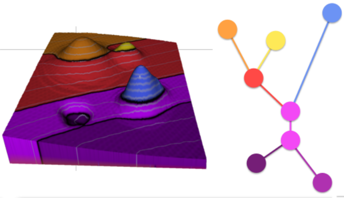


Cyclic Structures

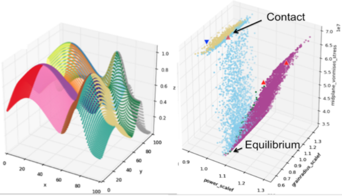
Topological data analysis and visualization capture the shape of complex data



Reeb Graph



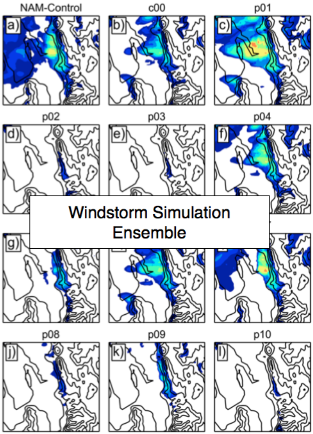
Contour and Contour Trees

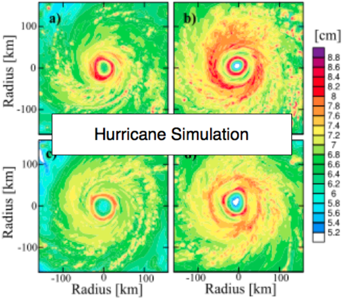


Morse-Smale Complexes

Multivariate data and atmospheric science

- Multivariate: data points with vector values
- Domain applications: weather ensembles from atmospheric science with large societal impacts, windstorms, smoke transport from wildfires, winter season precipitation, and hurricane forecasting

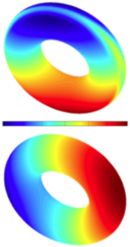


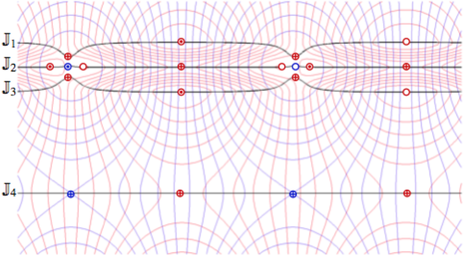


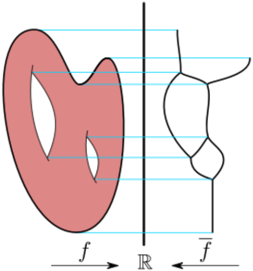
Credit: John Horel (Utah)
Kristen Corbosiero (Albany)

Explore multivariate topological data analysis tools

- Multivariate analysis and vis: Jacobi sets, Reeb spaces and mapper
- Study relations among level sets and critical pts of multiple functions



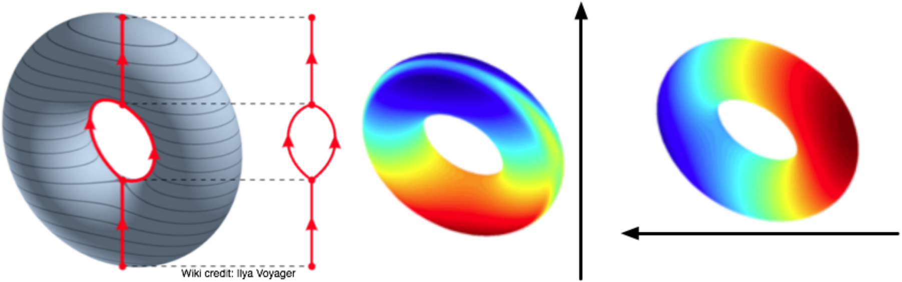




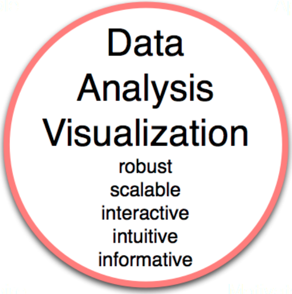
[Bhatia, **Wang**, Norgard, Pascucci, Bremer (CGTA) 2015]
[Munch, **Wang** (FWCG) 2015]
[Munch, **Wang** (SoCG) 2016]

Reeb space

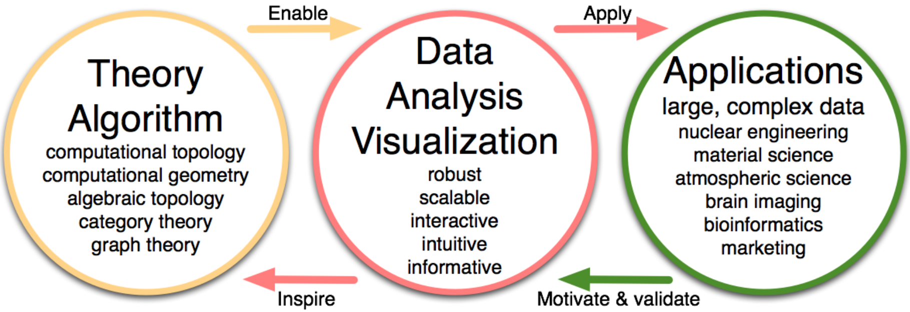
- Generalization of Reeb graph
- Compresses the contours of a multivariate mapping and obtains a summary representation of their relationships
- Fundamental to the study of multivariate scientific data



Multivariate analysis within in an end-to-end view

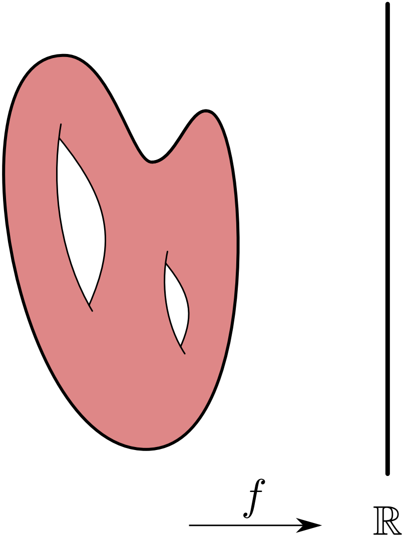


Multivariate analysis within in an end-to-end view

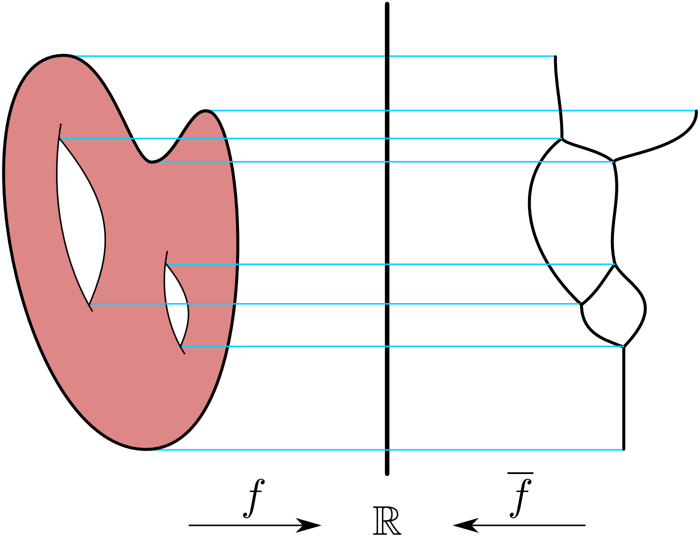


High-Level Question

Original Reeb graph construction



Original Reeb graph construction



Mapper

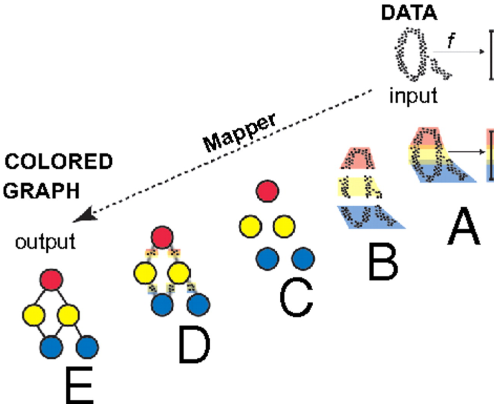


Image: Nicolau Levine Carlsson 2011

Mapper

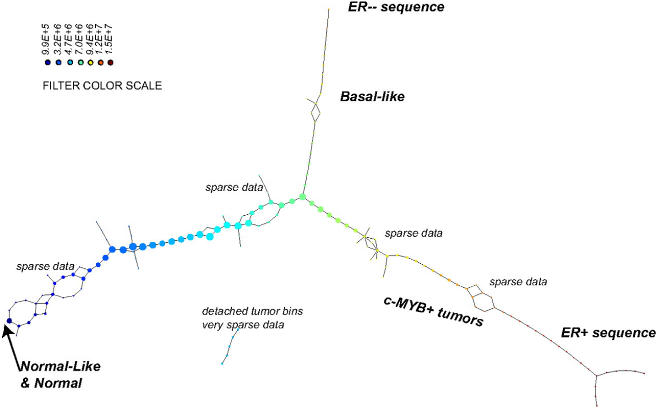


Image: Nicolau, Levine, Carlsson, PNAS 2011

Joint Contour Net

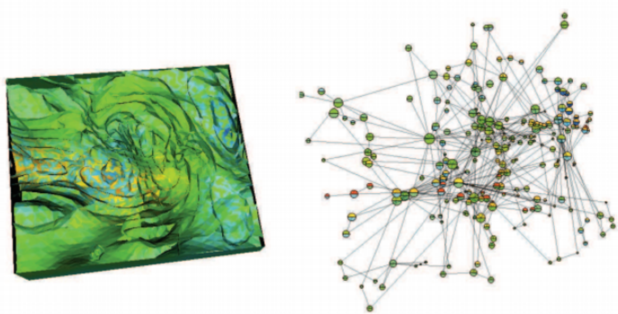


Image: Duke, Carr, 2013

Intuition

Mapper is an approximation of the Reeb space.

Intuition

Mapper is an approximation of the Reeb space.

Question

How do we formalize this?

Intuition

Mapper is an approximation of the Reeb space.

Question

How do we formalize this?

Goal

Develop some theoretical understanding of the relationship between the Reeb space and its discrete approximations to support its use in practical data analysis.

Intuition

Mapper is an approximation of the Reeb space.

Question

How do we formalize this?

Goal

Develop some theoretical understanding of the relationship between the Reeb space and its discrete approximations to support its use in practical data analysis.

Result

Prove the convergence between the Reeb space and mapper in terms of an interleaving distance between their categorical representations

(Partially) address some open questions

- [Carriére, Oudot, 2016]: is it possible to describe the mapper as a particular constructible cosheaf? Yes, $d = 1$; hopefully $d > 1$.
- [Dey, Mémoli, and Wang]: does the mapper construction converges to the Reeb space in the limit? Yes, in the categorical representations; hopefully geometrically on the space level.

Sheaves make baby cry.
During a talk given by: Amit Patel
Baby: Stanley Michael Phillips
Quote contributed to (possibly): Dmitriy Morozov
SIAM Conference on Applied Algebraic Geometry, August 2, 2013

Category Theory Basics

Category and opposite category

- Data of a category: the objects and the arrows
- A "generalization" of set theory: set and relationships between elements of a set
- Arrows: morphisms between the objects
- Arrows can be composed associatively; identity arrow for each object
- Intuitively, a big (probably infinite) directed multi-graph with extra underlying structures: objects – nodes, each possible arrow between the nodes – directed edge

Examples of categories

- $\mathbf{Top}$: the category of topological spaces with continuous functions between them
- $\mathbf{Set}$: the category of sets with set maps
- $\mathbf{Open}(\mathbb{R}^d)$: the category of open sets in $\mathbb{R}^d$ with inclusion maps
- $\mathbf{Vect}$: the category of vector spaces with linear maps
- $\mathbf{R}$: the category of real numbers with inequalities connecting them
- $\mathbf{Cell}(K)$: the category induced by any simplicial complex K, where the objects are the simplices of K, and there is a arrow $\sigma \to \tau$ if σ is a face of τ

Poset category

- A category $\mathbf{P}$ in which any pair of elements $x, y \in \mathbf{P}$ has at most one arrow $x \to y$.
- $\mathbf{Open}(\mathbb{R}^d)$: exactly one arrow $I \to J$ between open sets if $I \subseteq J$
- $\mathbf{R}$: exactly one arrow $a \to b$ between real numbers if $a \leq b$
- Intuitively, a poset category can be thought of as a directed graph which is not a multigraph

Opposite category

- The *opposite category* $\mathcal{C}^{op}$ of a given category $\mathcal{C}$ is formed by interchanging the source and target of each arrow

Functor

- A *functor* is a map between categories that maps objects to objects and arrows to arrows
- Functor $F : \mathcal{C} \to \mathcal{D}$ maps an object x in $\mathcal{C}$ to an object $F(x)$ in $\mathcal{D}$, and maps an arrow $f : x \to y$ of $\mathcal{C}$ to an arrow $F[f] : F(x) \to F(y)$ of $\mathcal{D}$ in a way that respects the identity and composition laws
- Intuitively, a functor is a map between graphs which sends nodes (objects) to nodes and edges (arrows) to edges in a way that is compatible with the structure of the graphs

Examples of functors

- $H_p : \mathbf{Top} \to \mathbf{Vect}$, sends a topological space $\mathbb{X}$ to its p-th singular homology group $H_p(\mathbb{X})$, and sends any continuous map $f : \mathbb{X} \to \mathbb{Y}$ to the linear map between homology groups, $H_p[f] := f_* : H_p(\mathbb{X}) \to H_p(\mathbb{Y})$
- $\pi_0 : \mathbf{Top} \to \mathbf{Set}$, sends a topological space $\mathbb{X}$ to a set $\pi_0(\mathbb{X})$ where each element represents a path connected component of $\mathbb{X}$, and sends a map $f : \mathbb{X} \to \mathbb{Y}$ to a set map $\pi_0[f] := f_* : \pi_0(\mathbb{X}) \to \pi_0(\mathbb{Y})$

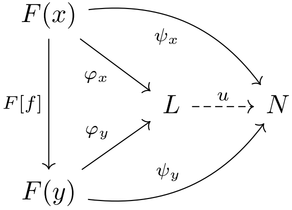
Natural Transformation

- A *natural transformation* $\varphi : F \Rightarrow G$ between functors $F, G : \mathcal{C} \to \mathcal{D}$ is a family of arrows φ in $\mathcal{D}$ such that (a) for each object x of $\mathcal{C}$, we have $\varphi_x : F(x) \to G(x)$, an arrow of $\mathcal{D}$; and (b) for any arrow $f : x \to y$ in $\mathcal{C}$, $G[f] \circ \varphi_x = \varphi_y \circ F[f]$
- Any collection of functors $F : \mathcal{C} \to \mathcal{D}$ can be turned into a category, with the functors themselves as objects and the natural transformations as arrows, notated as $\mathcal{D}^{\mathcal{C}}$
- Our case: $\mathcal{D} = \mathbf{Set}$

$$\begin{array}{ccc} F(x) & \xrightarrow{\varphi_x} & G(x) \\ {\scriptstyle F[f]}\downarrow & & \downarrow{\scriptstyle G[f]} \\ F(y) & \xrightarrow{\varphi_y} & G(y) \end{array}$$

Colimit

- The *cocone* (N, ψ) of a functor $F : \mathcal{C} \to \mathcal{D}$ is an object N of $\mathcal{D}$ along with a family of ψ of arrows $\psi_x : F(x) \to N$ for every object x of $\mathcal{C}$, such that for every arrow $f : x \to y$ in $\mathcal{C}$, we have $\psi_y \circ F[f] = \varphi_x$
- A cocone (N, ψ) *factors* through another cocone (L, φ) if there exists an arrow $u : L \to N$ such that $u \circ \varphi_x = \psi_x$ for every x in $\mathcal{C}$
- The *colimit* of $F : \mathcal{C} \to \mathcal{D}$, denoted as $\operatorname{colim} F$, is a cocone (L, φ) of F such that for any other cocone (N, ψ) of F, there exists a unique arrow $u : L \to N$ such that (N, ψ) factors through (L, φ).



Topological Notions

Reeb space

Reeb Space $\mathcal{R}(\mathbb{X}, f)$

- Given $f : \mathbb{X} \to \mathbb{R}^d$
- $x \sim y$ iff x and y in same (path) connected component of $f^{-1}(a)$
- Reeb space: quotient space obtained by identifying equivalent points with the quotient topology
- $\mathcal{R}(\mathbb{X}, f) := \mathbb{X}/\sim_f$

The Point

- The Reeb space of a (nice enough) $(\mathbb{X}, f)$ is a stratified space. [Edelsbrunner, Harer, Patel 2008]
- A Reeb space comes with a space *and* a function

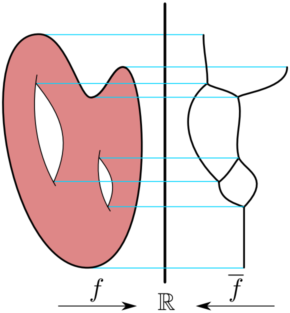
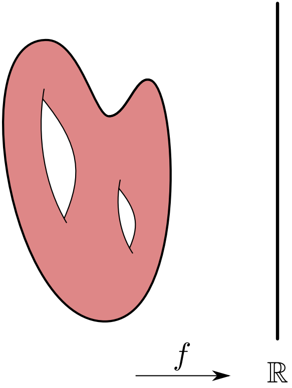


Image: [de Silva, Munch, Patel, 2015]

Data

Data $(\mathbb{X}, f)$

A compact topological space $\mathbb{X}$ with a function $f : \mathbb{X} \to \mathbb{R}^d$

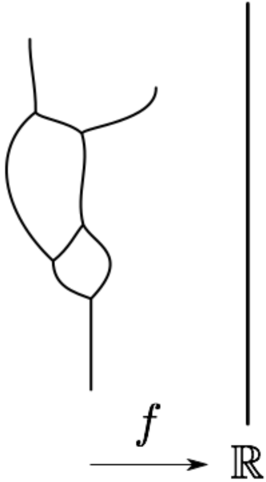


$\mathbb{X}$: manifold

Data of a Reeb space

Data $(\mathbb{X}, f)$

A compact topological space $\mathbb{X}$ with a function $f : \mathbb{X} \to \mathbb{R}^d$

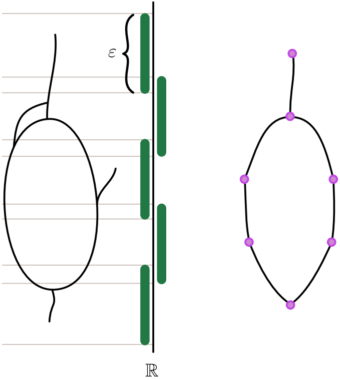


$\mathbb{X}$: Reeb graph / Reeb space

Mapper

Mapper $M(\mathcal{U}, f)$

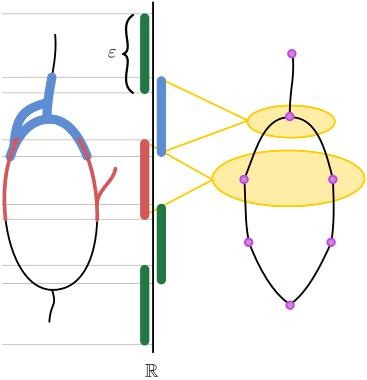
- Given $f : \mathbb{X} \to \mathbb{R}^d$
- Fix a *good* cover $\mathcal{U} = \{U_\alpha\}$ of $\mathbb{R}^d$
- $f^*(\mathcal{U})$: the cover of $\mathbb{X}$ obtained by considering the path connected components of $\{f^{-1}(U_\alpha)\}$
- Mapper is the nerve of this cover
- $M(\mathcal{U}, f) := \mathrm{Nrv}(f^*(\mathcal{U}))$
- [Singh, Mémoli, Carlsson 2007]



Mapper

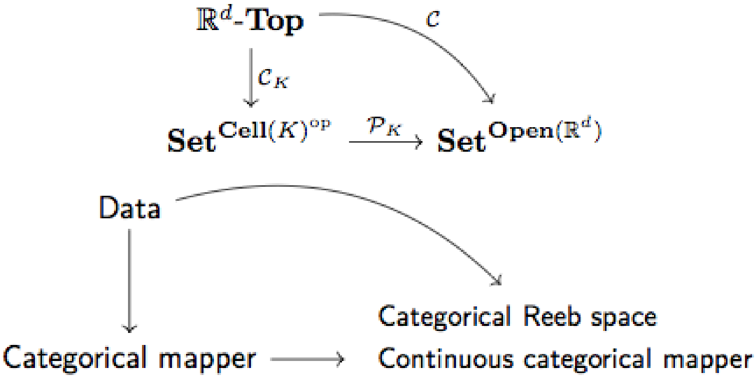
Mapper $M(\mathcal{U}, f)$

- Given $f : \mathbb{X} \to \mathbb{R}^d$
- Fix a *good* cover $\mathcal{U} = \{U_\alpha\}$ of $\mathbb{R}^d$
- $f^*(\mathcal{U})$: the cover of $\mathbb{X}$ obtained by considering the path connected components of $\{f^{-1}(U_\alpha)\}$
- Mapper is the nerve of this cover
- $M(\mathcal{U}, f) := \mathsf{Nrv}(f^*(\mathcal{U}))$
- [Singh, Mémoli, Carlsson 2007]



A Road Map

Connecting categorical representations



- Measures the amount the above diagram deviates from being commutative.

Results Overview

Convergence result

Theorem

Given a multivariate function $f : \mathbb{X} \to \mathbb{R}^d$ *defined on a compact topological space, the data is represented as an object* $(\mathbb{X}, f)$ *in* $\mathbb{R}^d$-$\mathbf{Top}$*. Let* $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ *be a good cover of* $f(\mathbb{X}) \subseteq \mathbb{R}^d$*,* K *be the nerve of the cover and* $\mathrm{res}(\mathcal{U})$ *be the resolution of the cover,* $\mathrm{res}(\mathcal{U}) = \sup\{\mathrm{diam}(U_\alpha) \mid U_\alpha \in \mathcal{U}\}$*. Then*

$$d_I(\mathcal{C}(\mathbb{X}, f), \mathcal{P}_K \mathcal{C}_K(\mathbb{X}, f)) \leq \mathrm{res}(\mathcal{U}).$$

- Convergence between continuous Reeb Space and discrete Mapper
- Distance between their categorical reps. requires only the knowledge of the cover
- Interleaving distance is an extended pseudometric.

Geometric convergence ($d = 1$)

Corollary

Given a constructible $\mathbb{R}$-space $(\mathbb{X}, f)$ with $f : \mathbb{X} \to \mathbb{R}$, let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be a good cover of $f(\mathbb{X}) \subseteq \mathbb{R}$, and let K be the nerve of the cover. Then

$$d_I(\mathcal{R}(\mathbb{X}, f), \mathcal{M}_K(\mathbb{X}, f)) \leq \mathrm{res}(\mathcal{U}).$$

In particular, a sequence of mappers for increasingly refined covers converges to the Reeb graph.

- Interleaving distance is an extended metric when $d = 1$ [de Silva, Munch, Patel, 2015].
- $d = 1$: convergence geometrically (i.e. on the space level).

Categorical Notions

Data

- Data is stored in the category $\mathbb{R}^d$-$\mathbf{Top}$
- Object: $\mathbb{R}^d$-space, a pair consisting of a topological space $\mathbb{X}$ with a continuous map $f : \mathbb{X} \to \mathbb{R}^d$, $(\mathbb{X}, f)$
- Arrow: $\nu : (\mathbb{X}, f) \to (\mathbb{Y}, g)$, a function-preserving map, i.e., a continuous map on the underlying spaces $\nu : \mathbb{X} \to \mathbb{Y}$ such that $g \circ \nu(x) = f(x)$ for all $x \in \mathbb{X}$
- Examples: PL functions on simplicial complexes or Morse functions on manifolds are objects in $\mathbb{R}^d$-$\mathbf{Top}$

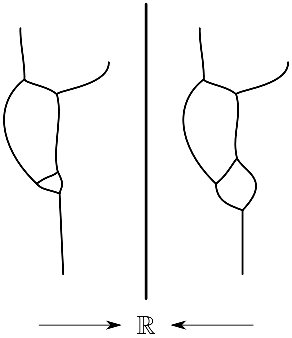
Function preserving maps

Definition

A *function preserving map* between two $\mathbb{R}^d$-spaces $(\mathbb{X}, f)$ and $(\mathbb{Y}, g)$ is a continuous map $\nu : \mathbb{X} \to \mathbb{Y}$ such that

$$\begin{array}{ccc} \mathbb{X} & \xrightarrow{\nu} & \mathbb{Y} \\ & {\scriptstyle f}\searrow \quad \swarrow{\scriptstyle g} & \\ & \mathbb{R}^d & \end{array}$$

commutes.



Categorical Reeb graph [de Silva, Munch, Patel, 2015]

Represented by functor $F = \pi_0 f^{-1} : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$

Category: $\mathbf{Set}^{\mathbf{Open}(\mathbb{R})}$

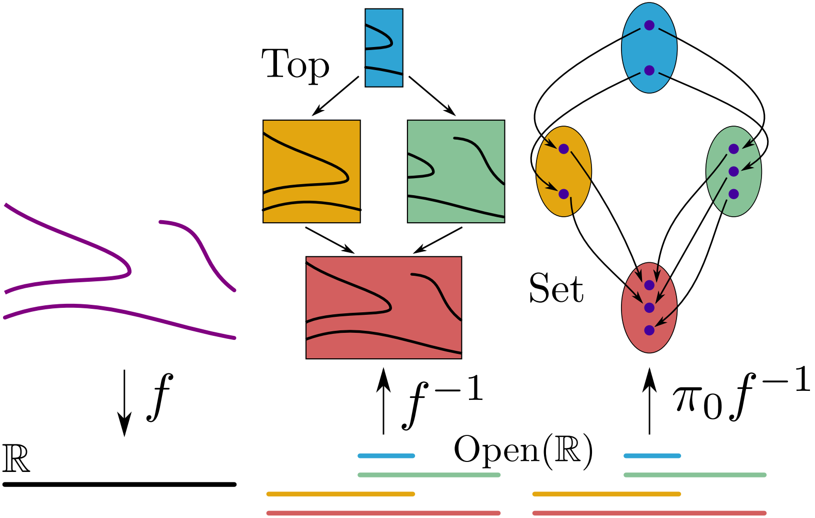
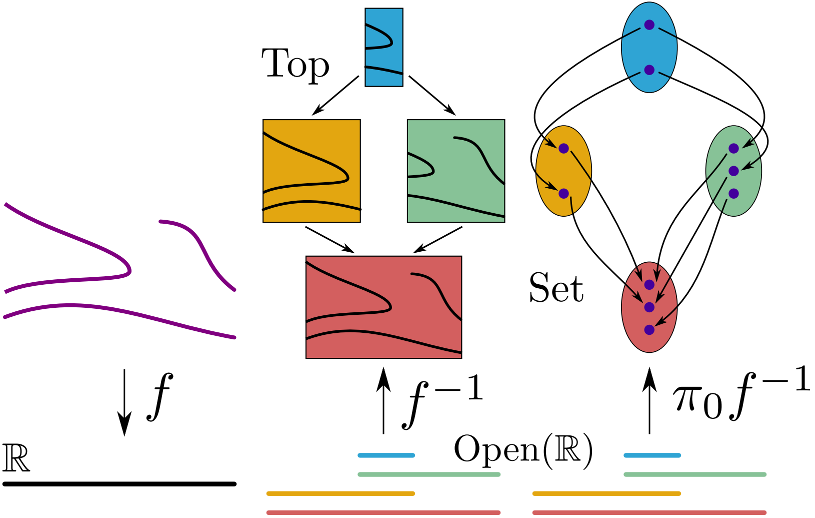


Image: [de Silva, Munch, Patel, 2015]

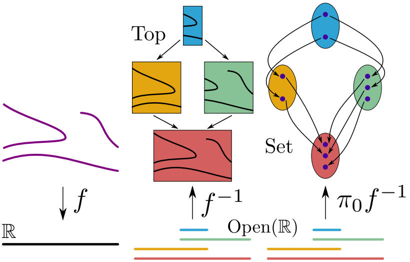
Categorical Reeb space

Represented by functor $F = \pi_0 f^{-1} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$

Category: $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$



Constructing a Reeb space from the data

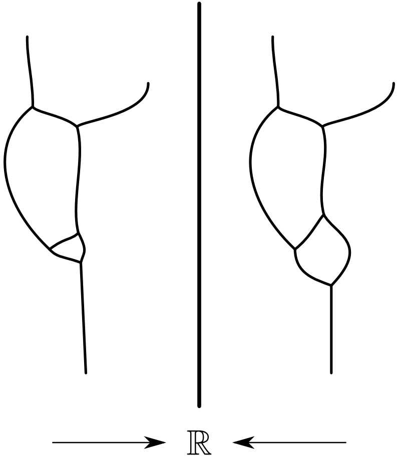


Represented by the functor $\mathcal{C} : \mathbb{R}^d\text{-}\mathbf{Top} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$

$\mathcal{C}$ maps:

- Data $(\mathbb{X}, f)$ to $\pi_0 f^{-1}$
- Function preserving map $\nu : (\mathbb{X}, f) \to (\mathbb{Y}, g)$ to natural transformation $\mathcal{C}[\nu]$ induced by the inclusion $\nu f^{-1}(I) \subseteq g^{-1}(I)$

Almost Isomorphisms



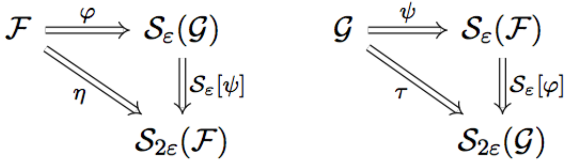
Interleaving Distance

- Perturb each Reeb graph/Reeb space by ε (Smoothing)
- Determine if there is an almost isomorphism (ε-interleaving)

Interleaving between categorical Reeb spaces

Definition (Interleaving distance between Categorical Reeb spaces)

An ε-interleaving between functors $\mathcal{F}, \mathcal{G} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ is a pair of natural transformations, $\varphi : \mathcal{F} \Rightarrow \mathcal{S}_\varepsilon(\mathcal{G})$ and $\psi : \mathcal{G} \Rightarrow \mathcal{S}_\varepsilon(\mathcal{F})$ such that the diagrams below commute.



Interleaving between categorical Reeb spaces

Definition

Given two functors $\mathcal{F}, \mathcal{G} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$, the interleaving distance is defined to be

$$d_I(\mathcal{F}, \mathcal{G}) = \inf\{\varepsilon \in \mathbb{R}_{\geq 0} \mid \mathcal{F}, \mathcal{G} \text{ are } \varepsilon\text{-interleaved}\}$$

Interleaving between categorical Reeb spaces

Definition

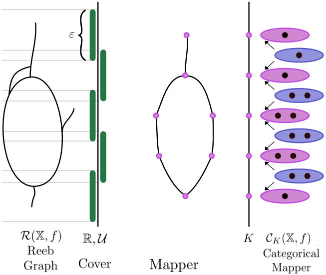
Given two functors $\mathcal{F}, \mathcal{G} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$, the interleaving distance is defined to be

$$d_I(\mathcal{F}, \mathcal{G}) = \inf\{\varepsilon \in \mathbb{R}_{\geq 0} \mid \mathcal{F}, \mathcal{G} \text{ are } \varepsilon\text{-interleaved}\}$$

Idea

The interleaving is a metric on Reeb spaces which takes into account both the space and the function

Categorical Mapper: data over nerve of cover

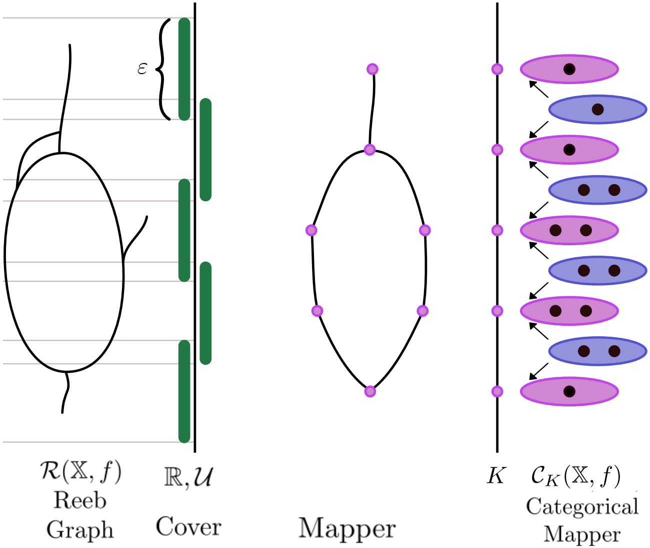


- Given a finite open cover for $\mathrm{im}\,(f) \subseteq \mathbb{R}^d$, $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$, $K = \mathrm{Nrv}(\mathcal{U})$.
- $\mathcal{U}_\sigma = \bigcap_{\alpha \in \sigma} U_\alpha$: open set in $\mathbb{R}^d$ associated to the simplex $\sigma \in K$
- Key: for $\sigma \leq \tau$ in K, "reversed arrow": $\mathcal{U}_\sigma \supseteq \mathcal{U}_\tau$.
- Face relation $\sigma \leq \tau$ induces a "backwards" mapping $\pi_0 f^{-1}(\mathcal{U}_\tau) \to \pi_0 f^{-1}(\mathcal{U}_\sigma)$.
- $\mathbf{Cell}(K)^{\mathrm{op}}$: simplices of K as objects and a unique arrow $\tau \to \sigma$ given by the face relation $\sigma \leq \tau$.
- Given a $(\mathbb{X}, f)$ in $\mathbb{R}^d$-$\mathbf{Top}$, functor $\mathcal{C}_K^f : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$ maps every σ to $\mathcal{C}_K^f(\sigma) := \pi_0 f^{-1}(\mathcal{U}_\sigma)$.

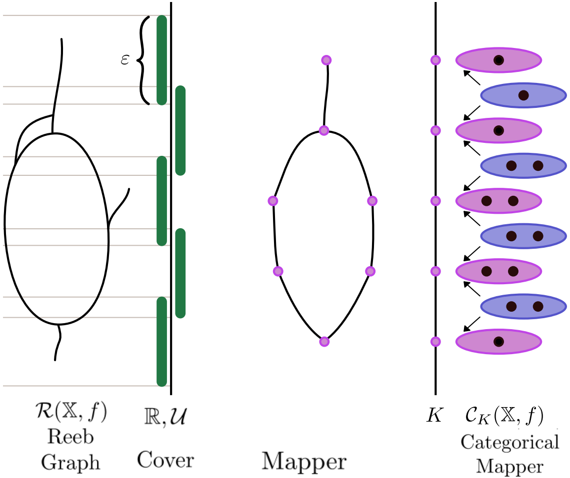
Categorical Mapper

Represented by functor $F : \mathbf{Cell}(K)^{op} \to \mathbf{Set}$

Category: $\mathbf{Set}^{\mathbf{Cell}(K)^{op}}$



Constructing mapper from the data



Represented by the functor $\mathcal{C}_K : \mathbb{R}^d\textbf{-Top} \to \textbf{Set}^{\textbf{Cell}(K)^{\text{op}}}$:

- Maps $(\mathbb{X}, f)$ to $\mathcal{C}_K^f$
- Maps $\nu : (\mathbb{X}, f) \to (\mathbb{Y}, g)$ to a natural transformation, $\mathcal{C}_K[\nu] : \mathcal{C}_K^f \to \mathcal{C}_K^g$.

Challenge

- Mapper doesn't have an obvious $\mathbb{R}^d$ function.
- Mapper and Reeb are not represented in the same category.

Compare Reeb space and mapper

$$\begin{array}{ccc} \mathbb{R}^d\text{-}\mathbf{Top} & \overset{\mathcal{C}}{\searrow} & \\ \downarrow \mathcal{C}_K & & \\ \mathbf{Set}^{\mathbf{Cell}(K)^{\mathrm{op}}} & \xrightarrow{\mathcal{P}_K} & \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)} \end{array}$$

Functor $\mathcal{P}_K : \mathbf{Set}^{\mathbf{Cell}(K)^{\mathrm{op}}} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$:

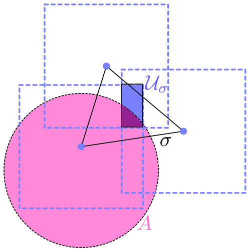
- Push mapper rep. to Reeb space rep.
- Prove convergence using interleaving distance between objects in $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$

Defining $\mathcal{P}_K$

- Simplicial complex K is the nerve of a good cover $\mathcal{U}$
- An open set $A \subseteq \mathbb{R}^d$
- $K_A = \{\sigma \in K \mid \bigcap_{\alpha \in \sigma} \mathcal{U}_\sigma \cap A \neq \emptyset\}$

$$\mathcal{P}_K : \mathbf{Set}^{\mathbf{Cell}(K)^{\mathrm{op}}} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}:$$

- Maps functor $F : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$ to functor $\mathcal{P}_K(F) : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$
- $\mathcal{P}_K(F)(I) = \mathrm{colim}_{\sigma \in K_I} F(\sigma)$ for every I in $\mathbf{Open}(\mathbb{R}^d)$



Highlight

$\mathcal{P}_K(F)(I) = \mathrm{colim}_{\sigma \in K_I} F(\sigma)$ for every I in $\mathbf{Open}(\mathbb{R}^d)$

- Colimit: "gluing", push discrete entities to continuous entities

Lemma

Let $\mathcal{F} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ be a functor which maps an open set I, to a set $\pi_0 f^{-1}(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma)$ with morphisms induced by π_0 on the inclusions. Then, the functor $\mathcal{P}_K \mathcal{C}_K(\mathbb{X}, f)$ is equivalent to $\mathcal{F}$.

Convergence result revisited

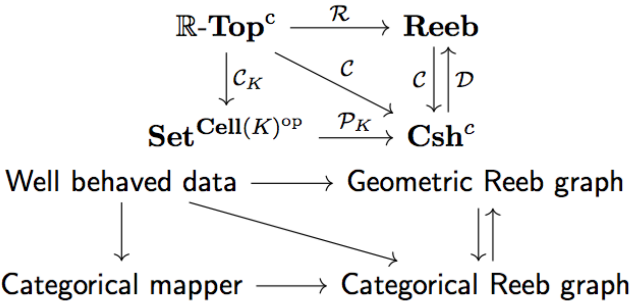
Theorem

Given a multivariate function $f : \mathbb{X} \to \mathbb{R}^d$ *defined on a compact topological space, the data is represented as an object* $(\mathbb{X}, f)$ *in* $\mathbb{R}^d$-$\mathbf{Top}$*. Let* $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ *be a good cover of* $f(\mathbb{X}) \subseteq \mathbb{R}^d$*,* K *be the nerve of the cover and* $\mathrm{res}(\mathcal{U})$ *be the resolution of the cover,* $\mathrm{res}(\mathcal{U}) = \sup\{\mathrm{diam}(U_\alpha) \mid U_\alpha \in \mathcal{U}\}$*. Then*

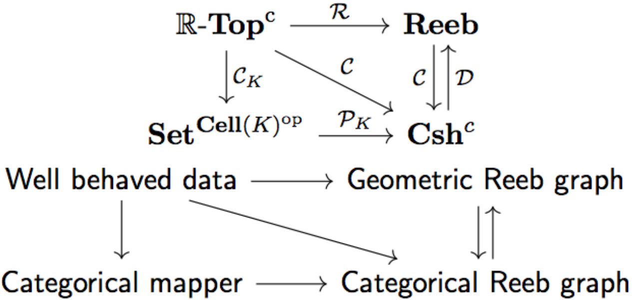
$$d_I(\mathcal{C}(\mathbb{X}, f), \mathcal{P}_K \mathcal{C}_K(\mathbb{X}, f)) \leq \mathrm{res}(\mathcal{U}).$$

- If we have a sequence of covers $\mathcal{U}_i$ such that $\mathrm{res}(\mathcal{U}_i) \to 0$, then the categorical representations of the mapper converge to the Reeb space in the interleaving distance

Connecting categorical rep. with geometric rep. $(d = 1)$



- Define a mapping that recovers the geometric rep. of mapper from its categorical rep.
- Convergence between mapper and Reeb graph geometrically (on the space level).



- Constructing the (geometric) Reeb graph from well behaved data is the same as creating its categorical representation, and then turning it back into a geometric object.

Well behaved data

$$\begin{array}{ccc} \mathbb{R}\text{-}\mathbf{Top}^{c} & \xrightarrow{\mathcal{R}} & \mathbf{Reeb} \\ \downarrow \mathcal{C}_K & \searrow^{\mathcal{C}} & \mathcal{C} \downarrow \uparrow \mathcal{D} \\ \mathbf{Set}^{\mathbf{Cell}(K)^{\mathrm{op}}} & \xrightarrow{\mathcal{P}_K} & \mathbf{Csh}^{c} \end{array}$$

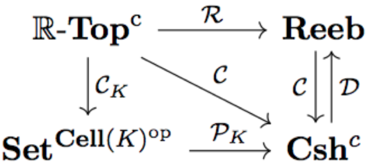
- Requiring the data to be constructible $\mathbb{R}$-spaces
- [de Silva, Munch, Patel, 2015], [Patel, Curry, 2016]
- $\mathbb{R}\text{-}\mathbf{Top}^{c}$: full subcategory of $\mathbb{R}\text{-}\mathbf{Top}$, objects are constructible $\mathbb{R}$-spaces
- $\mathbf{Reeb}$: full subcategory of $\mathbb{R}\text{-}\mathbf{Top}^{c}$, category of (geometric) Reeb graphs
- The construction of a (geometric) Reeb graph from well behaved data (a constructible $\mathbb{R}$-space) is captured by the functor $\mathcal{R} : \mathbb{R}\text{-}\mathbf{Top}^{c} \to \mathbf{Reeb}$.

Well behaved data

$$\begin{array}{ccc} \mathbb{R}\text{-}\mathbf{Top}^{c} & \xrightarrow{\ \mathcal{R}\ } & \mathbf{Reeb} \\ \downarrow \mathcal{C}_K & \searrow^{\mathcal{C}} & \mathcal{C}\downarrow\ \uparrow\mathcal{D} \\ \mathbf{Set}^{\mathbf{Cell}(K)^{\mathrm{op}}} & \xrightarrow{\ \mathcal{P}_K\ } & \mathbf{Csh}^{c} \end{array}$$

- Further restrict our objects of interest in $\mathbf{Set}^{\mathbf{Open}(\mathbb{R})}$ to be well behaved
- A *cosheaf* is a functor $F : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$ such that for any open cover $\mathcal{U}$ of a set U, the unique map $\mathrm{colim}_{U_\alpha \in \mathcal{U}} F(U_\alpha) \to F(U)$ is an isomorphism.
- A cosheaf is *constructible* if there is a finite set $S \subset \mathbb{R}$ such that if $A, B \in \mathbf{Open}(\mathbb{R})$ with $A \subseteq B$ and $S \cap A = S \cap B$, then $F(A) \to F(B)$ is an isomorphism. In addition, we require that if $A \cap S = \emptyset$ then $F(A) = \emptyset$.
- The category of constructible cosheaves with natural transformations is denoted $\mathbf{Csh}^{c}$.

Equivalence of categories [de Silva, Munch, Patel, 2015]



- $\mathbf{Reeb} \equiv \mathbf{Csh}^c$
- $\mathcal{C}$ has an "inverse" functor $\mathcal{D} : \mathbf{Csh}^c \to \mathbf{Reeb}$ which can turn a constructible cosheaf back into a geometric object
- Commutativity of the upper right triangle: $\mathcal{R} = \mathcal{D}\mathcal{C}$

Our geometric result

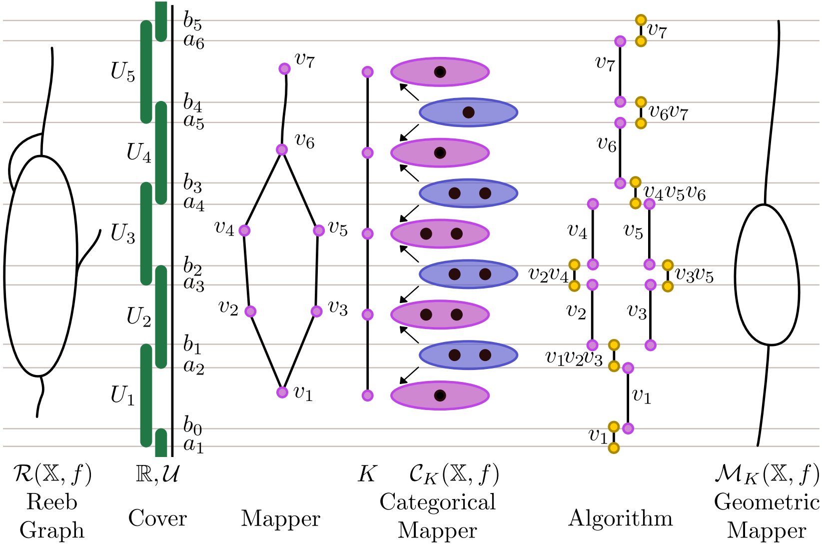
- Turn the categorical mapper back into a geometric one
- $\mathcal{M}_K(\mathbb{X}, f) := \mathcal{D}\mathcal{P}_K\mathcal{C}_K(\mathbb{X}, f)$ be the geometric rep.
- $\mathcal{R}(\mathbb{X}, f)$: geometric Reeb graph

Corollary

Given a constructible $\mathbb{R}$-space $(\mathbb{X}, f)$ with $f : \mathbb{X} \to \mathbb{R}$, let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be a good cover of $f(\mathbb{X}) \subseteq \mathbb{R}$, and let K be the nerve of the cover. Then

$$d_I(\mathcal{R}(\mathbb{X}, f), \mathcal{M}_K(\mathbb{X}, f)) \leq \mathrm{res}(\mathcal{U}).$$

Algorithm for geometric mapper



- Glue a collection of disjoint edges along equivalent vertices defined by the cover

Summary

- [Carriére, Oudot, 2016]: is it possible to describe the mapper as a particular constructible cosheaf? Yes, $d = 1$ with our geometric results: we described the mapper as a constructible cosheaf when it is passed to the continuous version.
- Suspect that our geometric results may hold in the case $d > 1$.
- Require proper notion of constructibility for $\mathbb{R}^d$-spaces and cosheaves: want an equivalence of categories, and a proof that the interleaving distance is an extended metric, not just a pseudometric; and therefore the mapper converges to the Reeb space on the space level.
- Algorithm strategy for building the associated geometric mapper may be generalized by considering k-dimensional cover elements and their intersections.
- First steps towards providing a theoretical justification for the use of discrete objects (mapper and JCN) as approximations to the Reeb space with guarantees.

On-going/future directions

- Categorical interpretations of Jacobi sets and their distances
- Categorical interpretations of multiscale mapper
- Geometric graphs

Take home message

- Category theory provides a simple, beautiful language that could potentially give us cleaner interpretation of some commonly used TDA constructs
- Simple language for convergence proofs that connect discrete with continuous entities (hard to prove otherwise)
- New interpretations for studying topological structures, and for multivariate data analysis

Thank you!

- Liz Munch
- Vin de Silva, Justin Curry, Amit Patel, Robert Ghrist...
- All the TDA researchers who make category theory less scary
- arXiv:1512.04108, SoCG 2016
- NSF-IIS-1513616