

Introduction & Topology Basics

Dr. Elena Xinyi Wang

14 September, 2026

Who am I?

- Mathematician by training, works at the intersection of math, CS, and life sciences
- I received my PhD from Michigan State University in 2025
- I was a postdoc at TU Graz in 2025 summer
- I am now a postdoctoral researcher at the AIDOS (AI for Data Oriented Science) Lab at the University of Fribourg led by Prof. Bastian Rieck
- I am also a classically trained pianist, flutist, and vocalist
- I have a cat (Shosty, after Shostakovich), and a dog (Wolfie, after Wolfgang Mozart)
- I am actively trying to improve my French and German :) , I also speak Mandarin and Korean

Syllabus & Course Policies

- Available on course website (<https://elenawang93.github.io/TDA/>)
- Prerequisites
 - Linear Algebra (we will review some today)
 - Some programming (the material I will provide will be in Python)

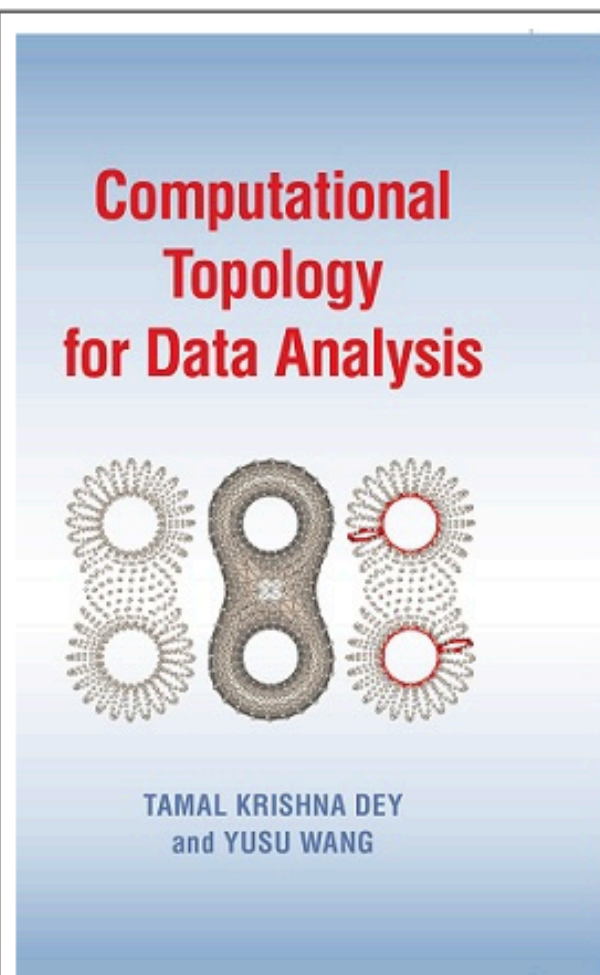
Course Schedule

Subject to change

Date	Lec.	Topic (Planned)	Reading	Notes	Due
16/9	1	Overview of course; basic topology definitions	DW 1.1–1.3	HW 1 posted	
23/9	2	Manifolds and Morse theory	DW 1.4, 1.5		
30/9	3	Simplicial complexes	DW 2.1–2.3		HW 1
7/10	4	Reeb graphs and merge trees	DW 7.1, 7.2	HW 2 posted	
14/10	5	Mapper	DW 9.1, 9.3		
21/10	6	Homology	DW 2.4, 2.5	Final project introduced; paper list posted.	HW 2
28/10	7	Computing homology and filtrations	DW 2.5.1, 3.1; EH IV.2	HW 3 posted	
4/11	8	Persistent homology	DW 3.2, 3.3; EH VII.1		Project proposal
11/11	9	Persistence in practice	DW 3.5, 6.1	Proposal feedback returned	HW 3
18/11	10	Distances and stability	DW 3.4; EH VIII.2	HW 4 posted	
25/11	11	Persistence and machine learning	DW 13.1, 13.3		
2/12	12	Directional transforms	ECT survey ; PHT	Presentation schedule posted	HW 4
9/12	13	Discrete Morse theory, multiparameter persistence	DW 10.1, 10.2, 11.1; Multiparameter survey		
16/12		Final project presentations		Send slides by Tue 15/12	Project presentation

Textbook

See https://elenawang93.github.io/TDA/Course_Info/Textbook.html for the full list



Copyright: Tamal Dey and Yusu Wang 2016-2021

Get an electronic copy.

[MAA review](#) and [zbMath review](#)

Errata detected in the print are corrected in electronic version marked with red text

This material has been / will be published by Cambridge University Press as Computational Topology for Data Analysis by Tamal Dey and Yusu Wang. This pre-publication version is free to view and download for personal use only. Not for re-distribution, re-sale, or use in derivative works.

Book : Computational Topology for Data Analysis

(published by Cambridge University Press)

Tamal K Dey

Department of Computer Science, Purdue University, IN

Yusu Wang

Halicioglu Data Science Institute, UC San Diego, CA

Computational topology has played a synergistic role in bringing together research work from computational geometry, algebraic topology, data analysis, and many other related scientific areas. In recent years, the field has undergone an explosive growth in the area of data analysis. The application of topological techniques to traditional data analysis, which has earlier developed mostly on a statistical setting, has opened up new opportunities. This book is intended to cover this aspect of computational topology along with the developments of generic techniques for various topology-centered problems. Since the development of persistent homology, the area has grown both in its methodology and applicability. One consequence of this growth has been the development of various algorithms which intertwine with the discoveries of various mathematical structures in the context of processing data. The purpose of this book is to capture these algorithmic developments with the associated mathematical guarantees.

- **Contents**

Chapter 1: Basic Topology

- a. Topological spaces, metric space topology
- b. Maps: homeomorphisms, homotopy equivalence, isotopy
- c. Manifolds
- d. Functions on smooth manifolds
- e. Notes and Exercises

Chapter 2 (i) . Complexes

- a. Simplicial complexes
- b. Nerves, Čech and Vietoris-Rips complexes
- c. Sparse complexes (Delaunay, Alpha, Witness)
- d. Graph induced complexes
- e. Notes and Exercises

Grading

- **In-class presentations (20%)**

- I will assign approximately 2-3 problems related to the lecture
- Students will be assigned to present problem solutions in the following week (starting next week)

- **Homework assignments (20%)**

- 4 written or programming assignments
- See course website for detailed late policies

LaTeX

- **Final project (60%)**

Final Project

- The project can range from applied to theoretical, from the known to completely original. I will provide examples/papers/ideas/datasets, but you have almost limitless freedom
- **Proposal (10%)**: schedule a meeting with me to talk through the idea, then submit a short (1-2 pages) written proposal. Proposals are only accepted upon completion of the meeting
- **Presentation (25%)**: this will take place in the last stage of the course period, you will be able to explain your whole project, and answer questions intentionally designed to probe your understanding of both the source material and your own work. Passing is required for the rest of the project to be assessed.
- **Report (20%)**: a written report of the project, with the appropriate mathematical, computational, and bibliographic details.
- **Oral exam (5%)**: this will take place during the exam session, aimed to examine your understanding of the whole project.

Why is presentation 45% of the grade

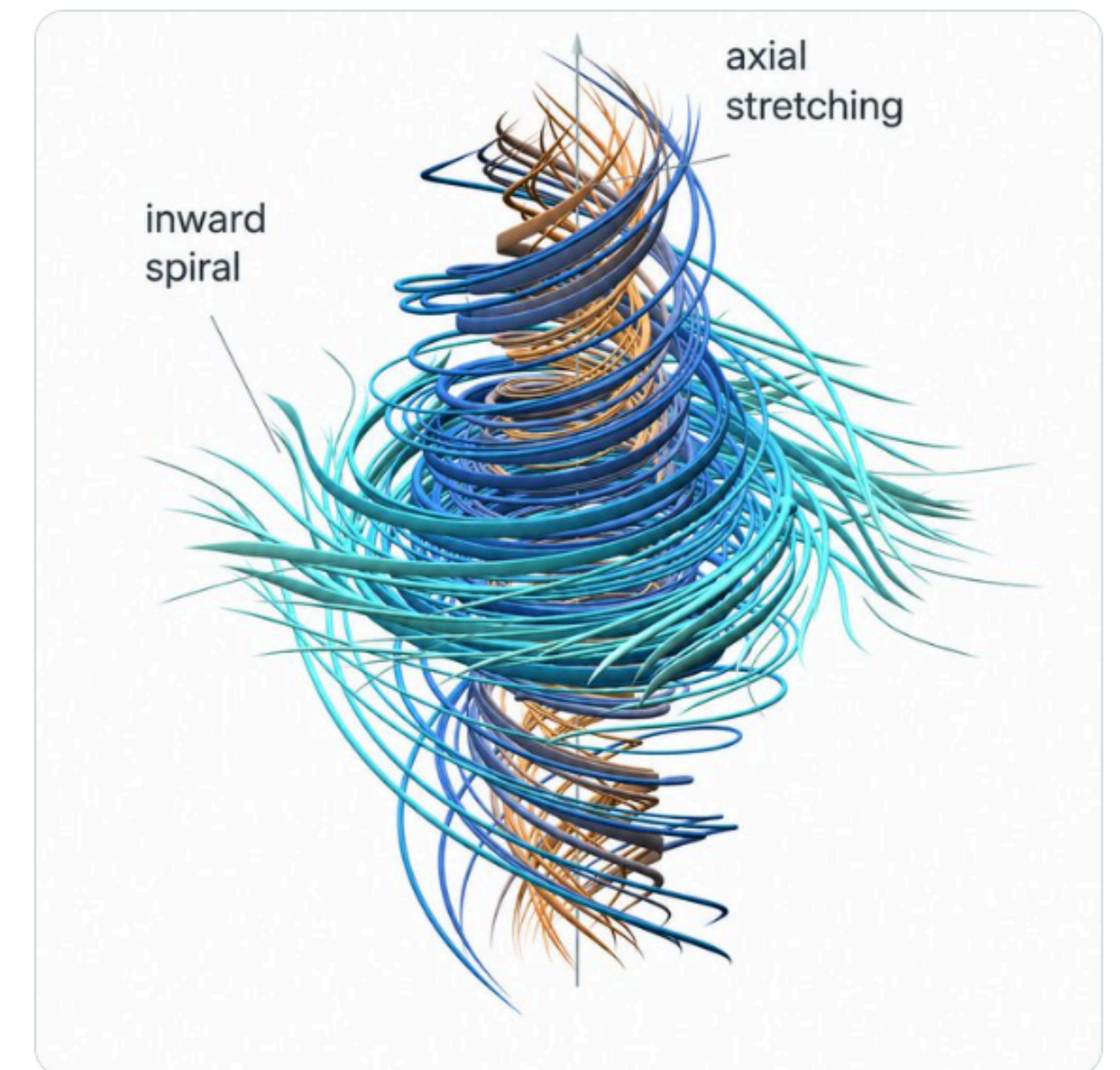
- AI can solve the “most difficult” problems in math now, what does that mean for us?
- What was math/CS research like before?
- What was math/CS education like?
- And what now?
- Learning is about the process, not the result
- Let’s slow down, ask basic questions, make mistakes, and enjoy understanding things together



We’re sharing a solution to the Navier-Stokes Millennium Prize Problem, one of the deepest problems at the frontier of mathematics.

The proof was produced by a group of agents, using an OpenAI next-generation model significantly more capable than GPT-6 Astra.

The problem concerns whether the description of smooth three-dimensional fluid motion modeled by the Navier-Stokes equations can break down. It has remained unresolved for roughly 90 years.



AI Policy

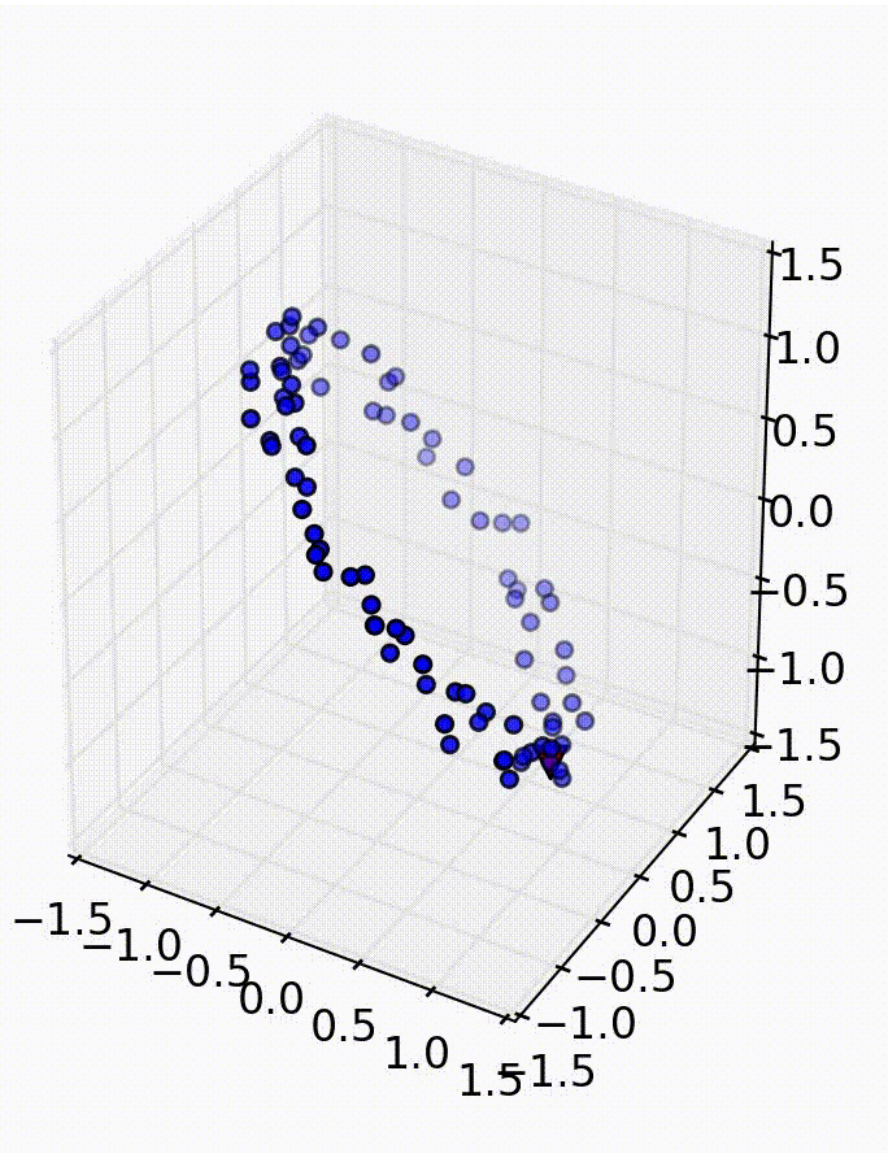
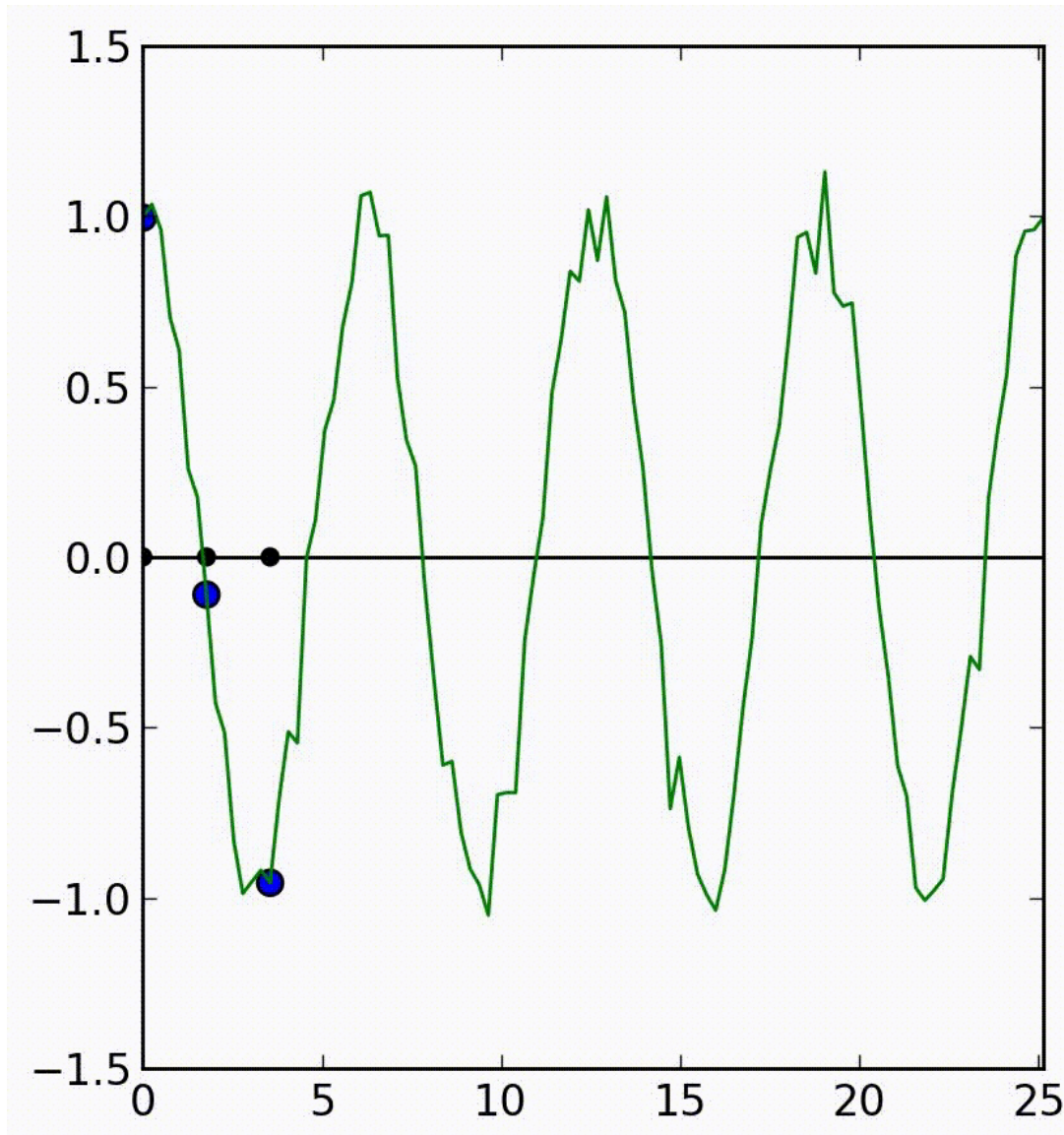
See <https://elenawang93.github.io/TDA/Policies/Integrity.html> for details

- AI is allowed, but you have to let me know
- Presentations will show me your level of understanding
- Be prepared for questions
- Passing the course is only possible if you pass the presentaion of the final project

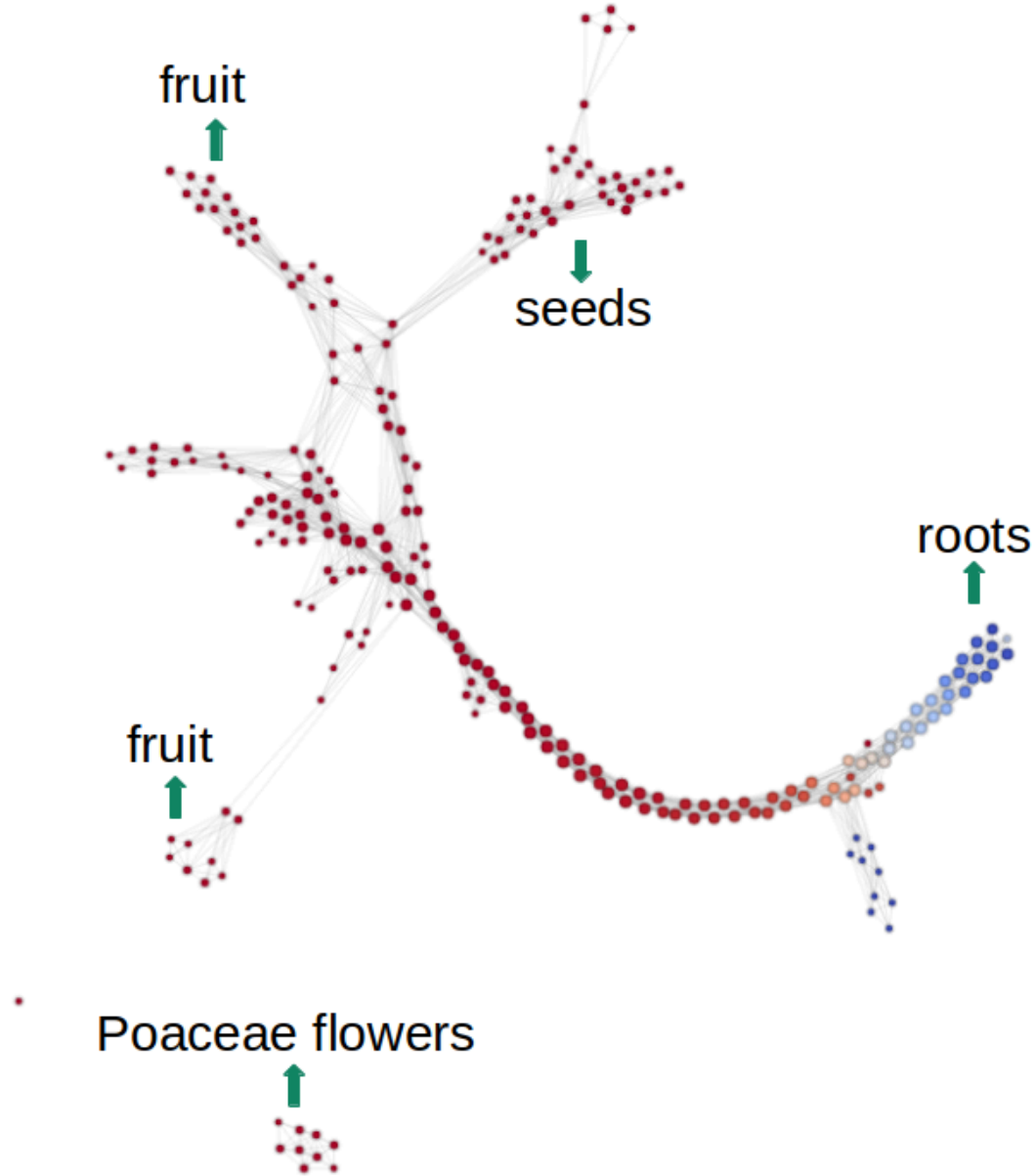
Data has shape, shape has meaning — Gunnar Carlsson



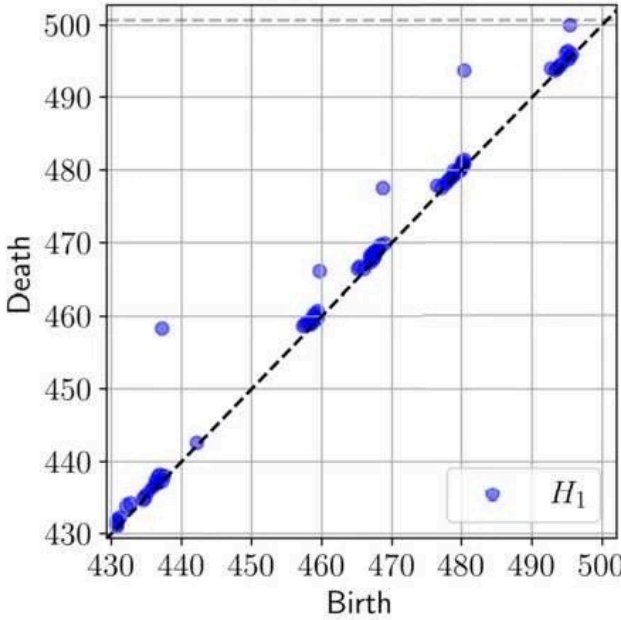
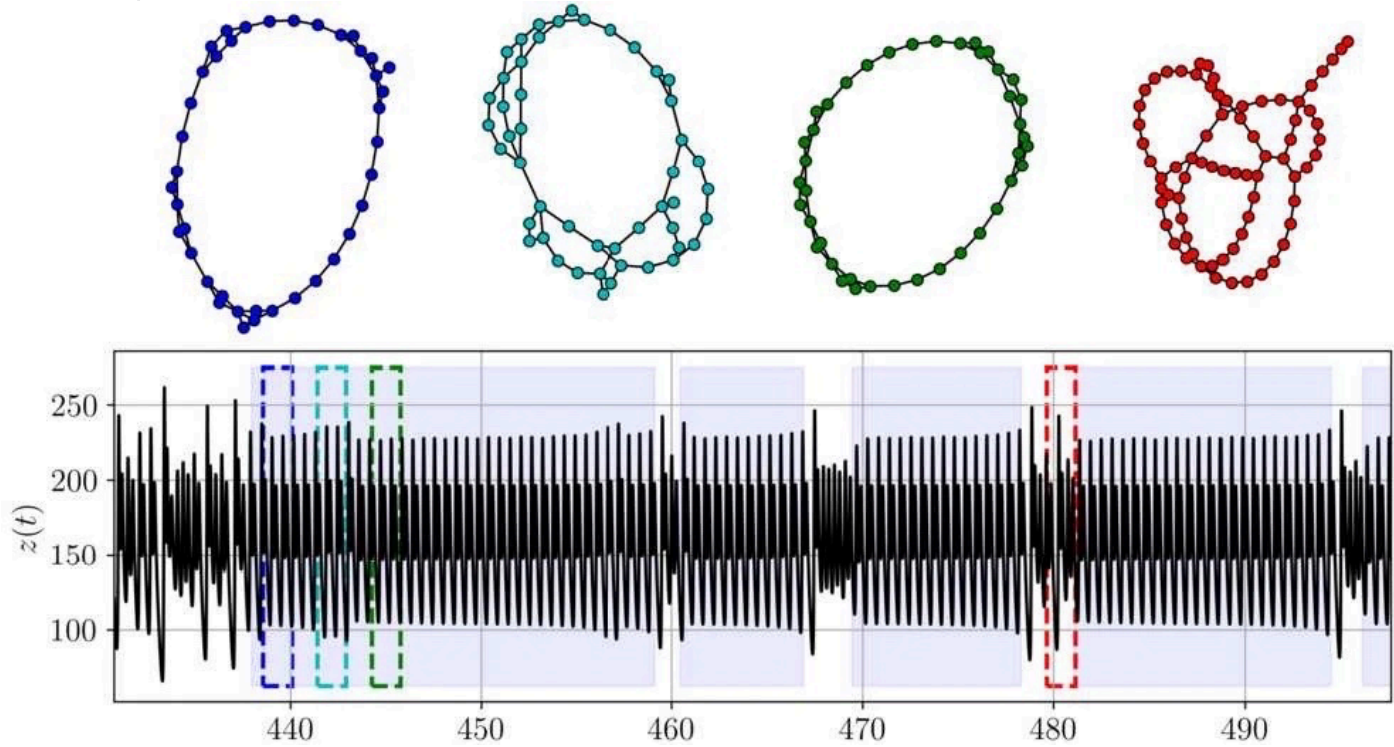
Amézquita et al., 2020



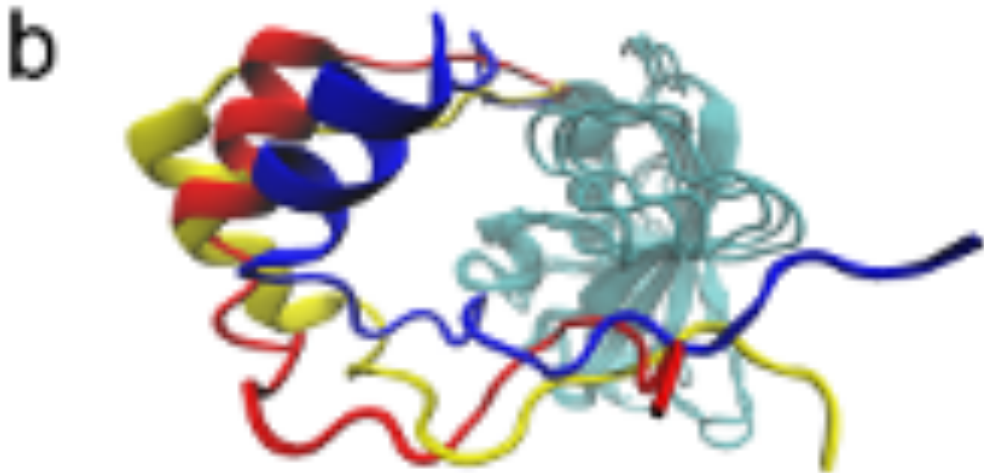
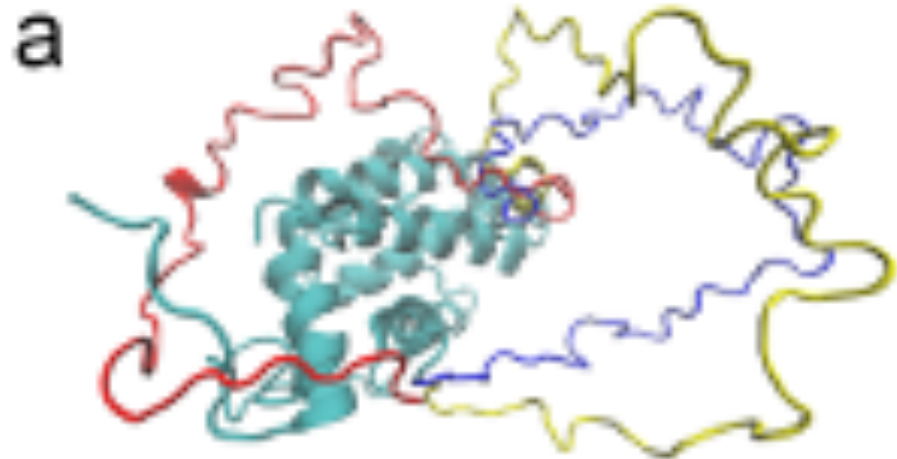
Munch et al., 2017



Chitwood et al., 2022



Myers et al., 2023



Qin et al., 2019

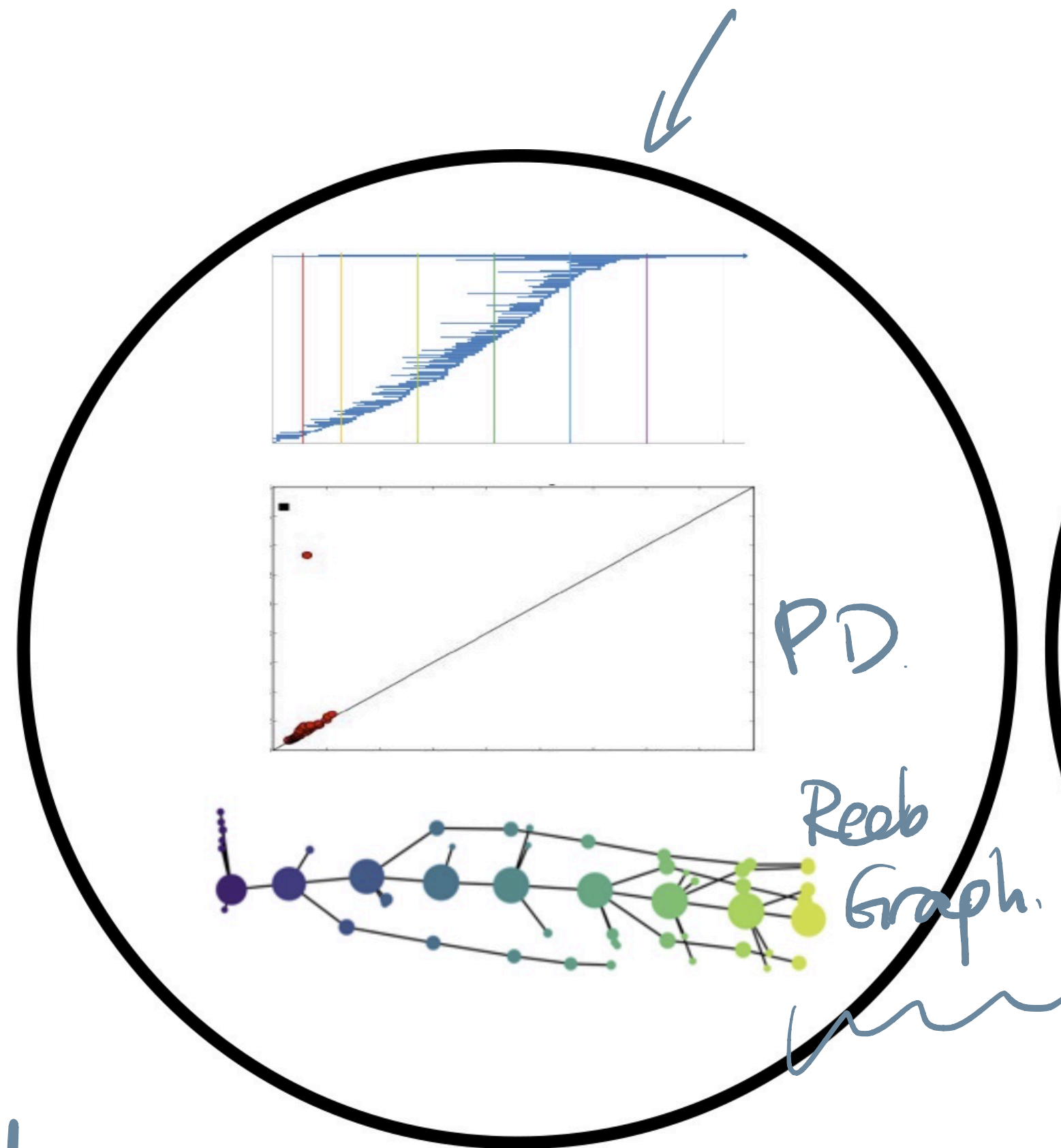
The Topological Data Analysis (TDA) Pipeline



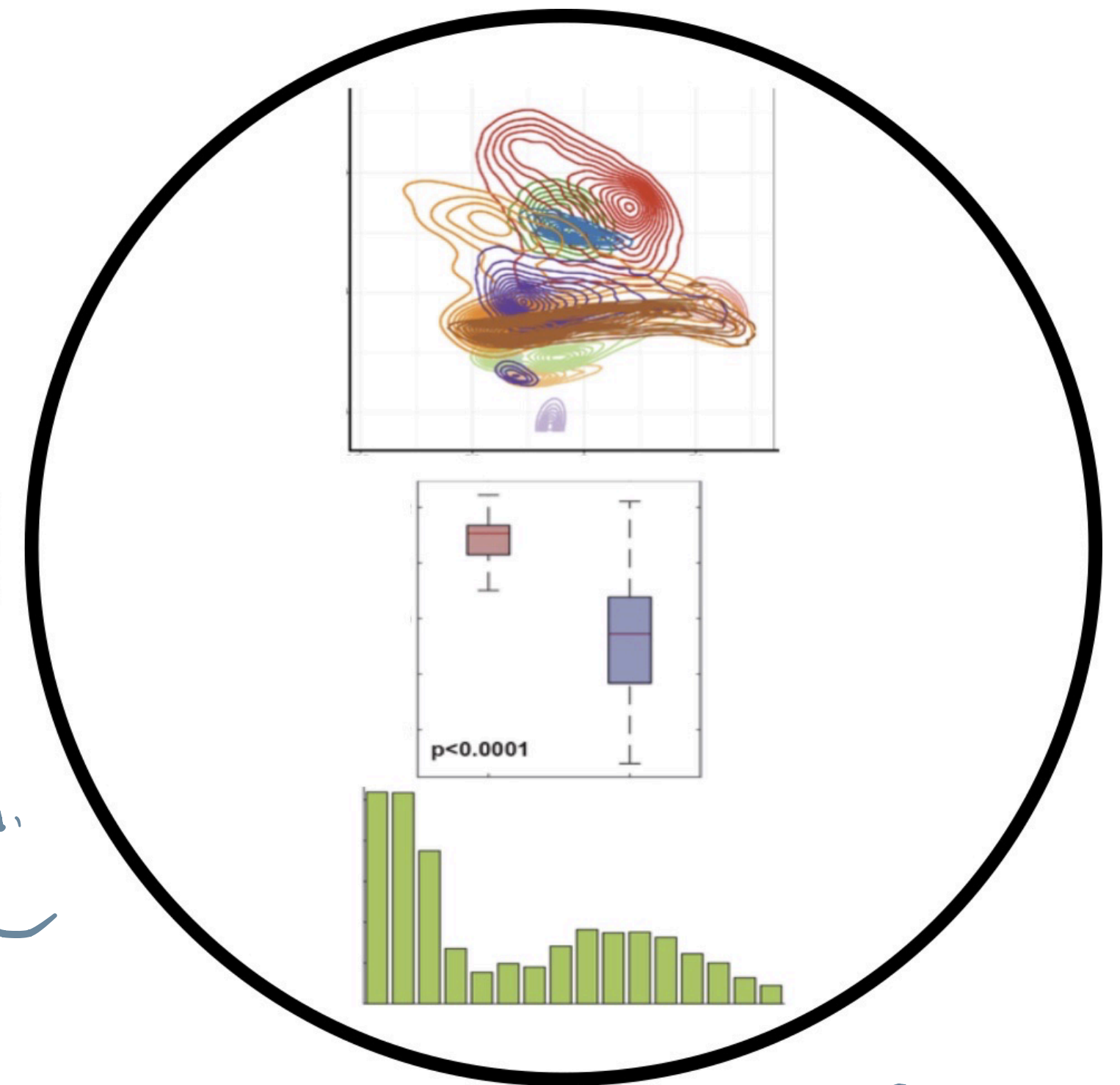
Input Data

X-ray
Point Cloud
Networks

Raw data



Topological Signatures



Analysis

Stats,
ML/DL

Main goal of TDA

We want to provide:

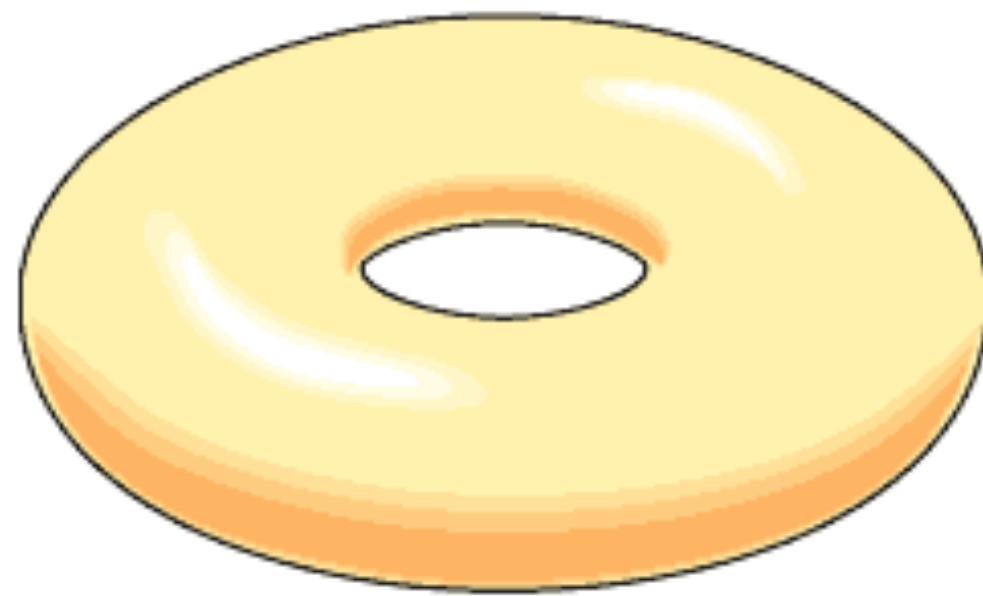
- 
1. Quantifiable, $\rightarrow$ *computable*.
 2. Comparable, $\rightarrow$ *metric*.
 3. Robust, $\rightarrow$ *noise*
 4. Consise $\rightarrow$ *smaller than input data*.

summaries of the shape of data.

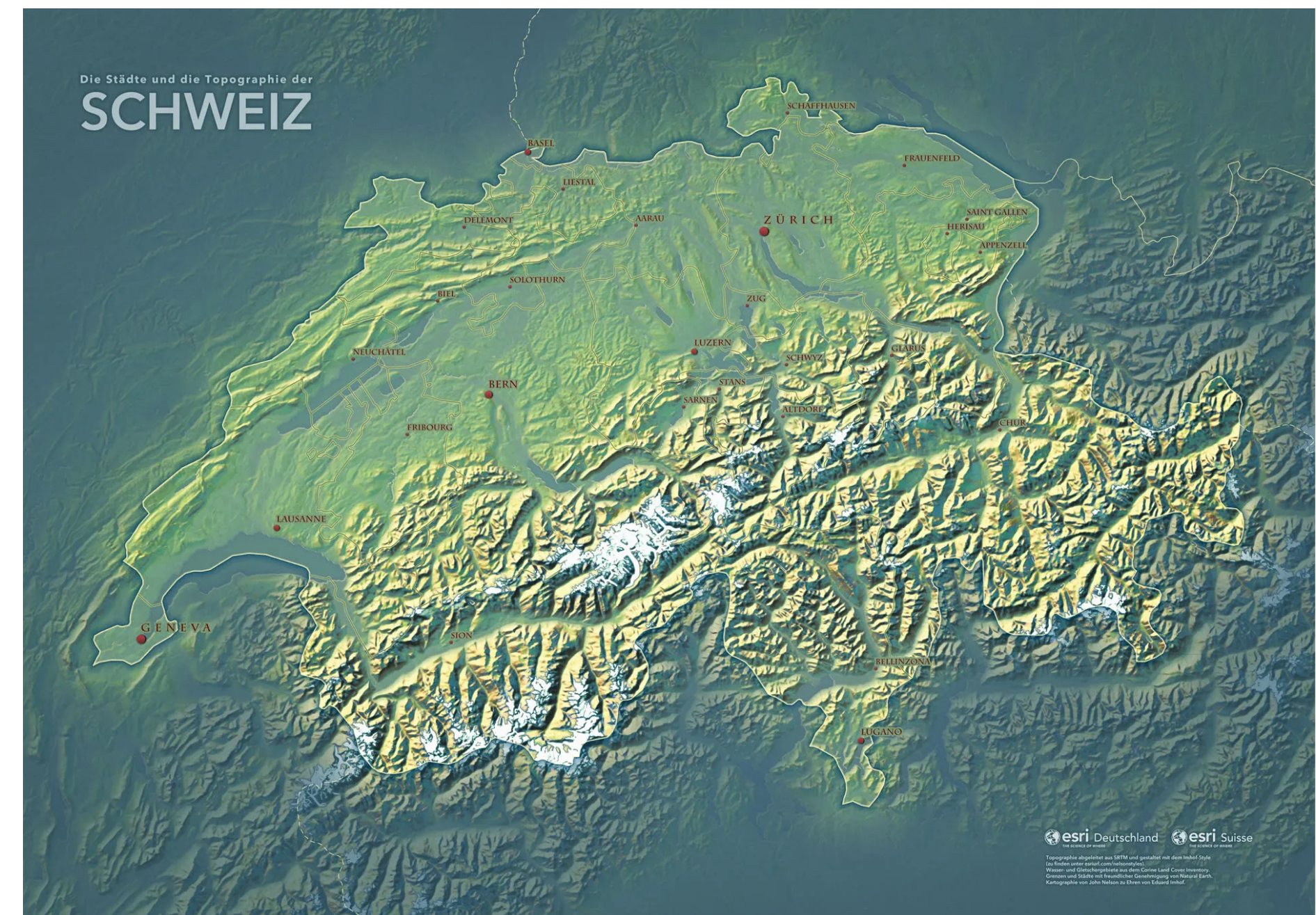
What is Topology?

Topology $\neq$ Topography

Mathematical study of spaces preserved under continuous deformation



Study of the shape and features of the surface of the earth

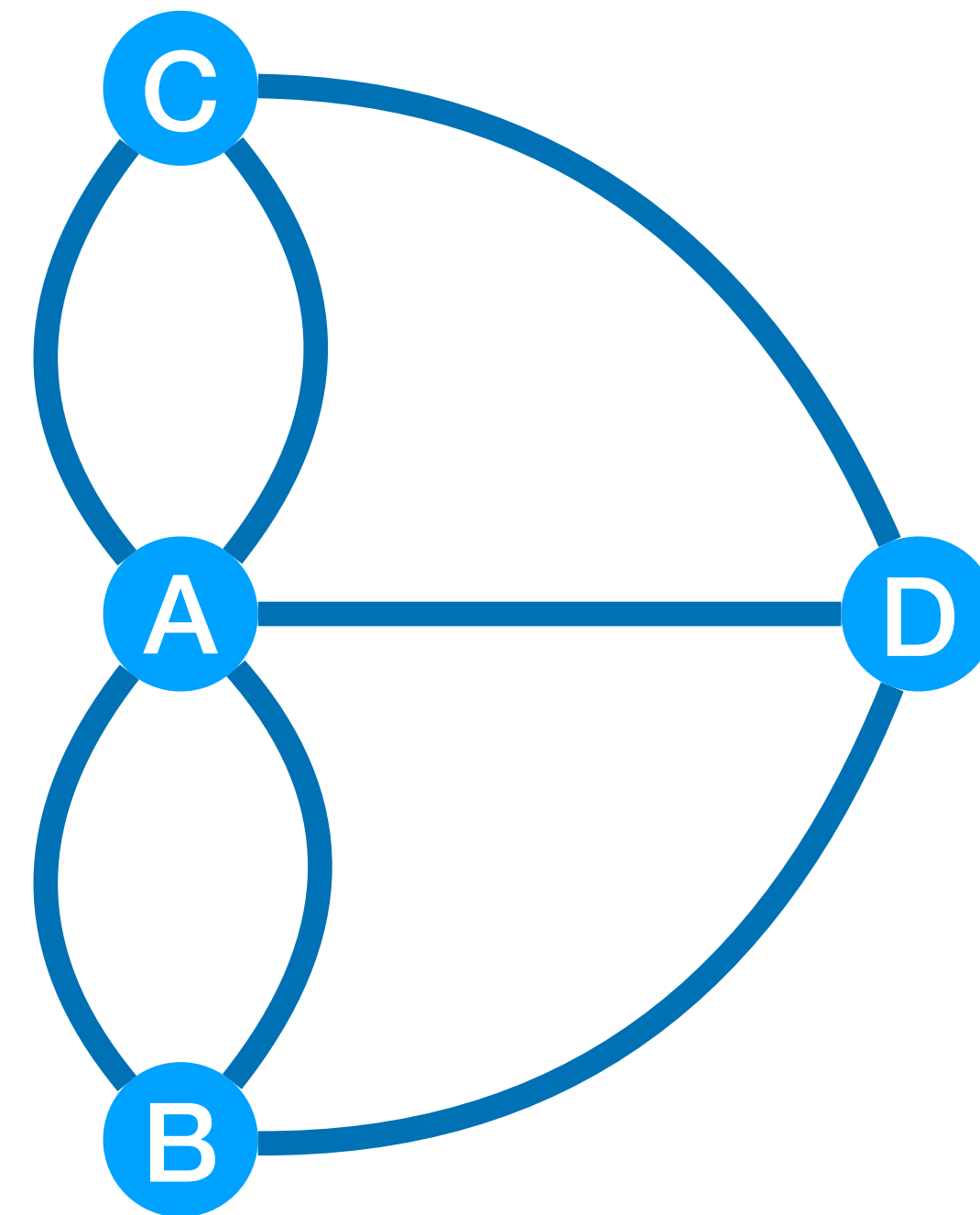
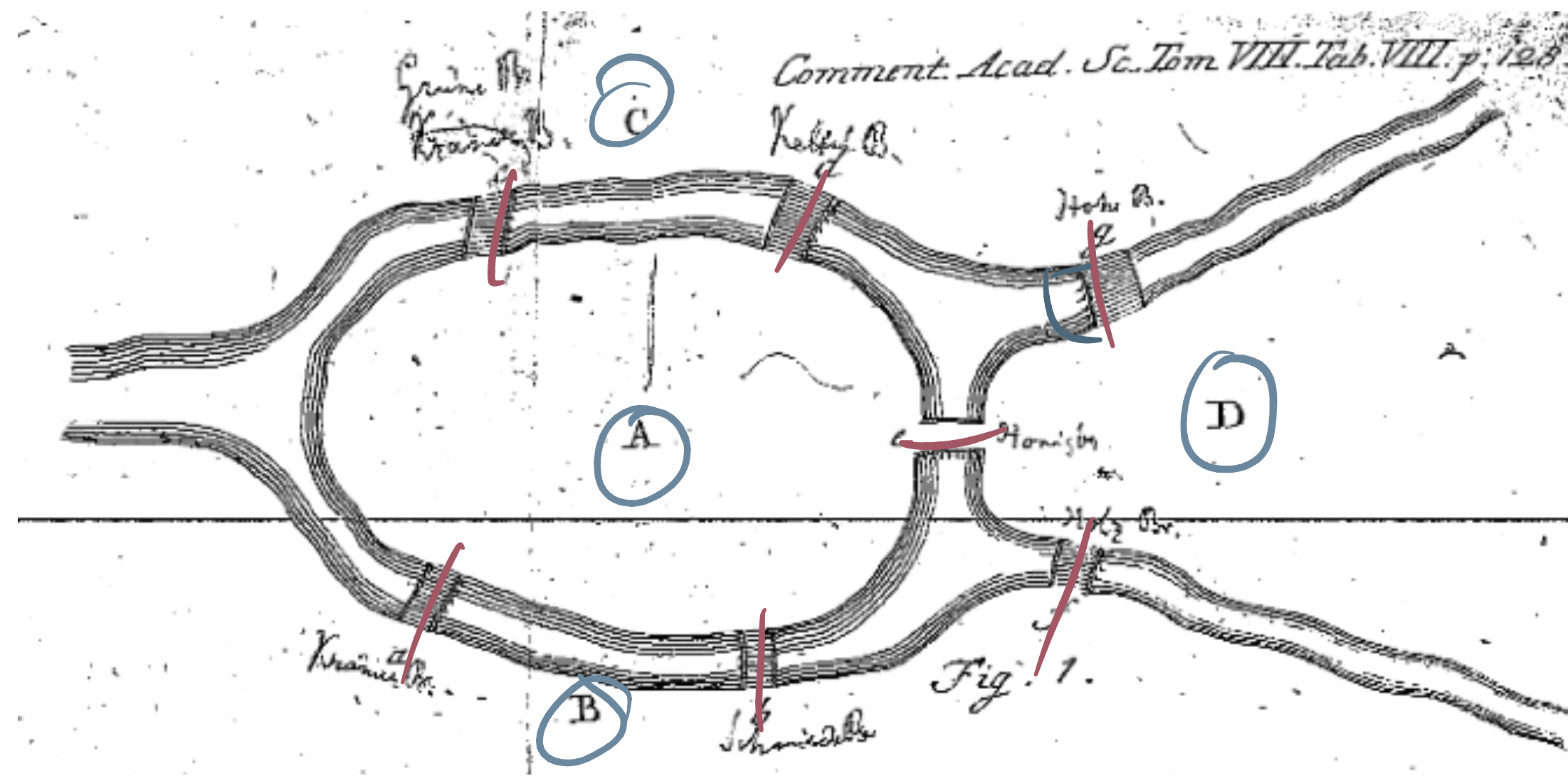


History



Leonhard Euler,
1707 - 1783

Königsberg.



Graph.
Eulerian Path.

History

Topology, the stratosphere of human thoughts!

In the ~~XXIV~~ century it might just possibly be useful

4 decades.

— Alexander Solzhenitzin, “*The First Circle*,” 1968

Things changed ca. 2000: introduction of Persistent Homology

Ca. 2005: Algebraic and computational development

2007: stability of persistence diagrams (mathematical foundation); Mapper as another data summary

Different ways of quantifying things

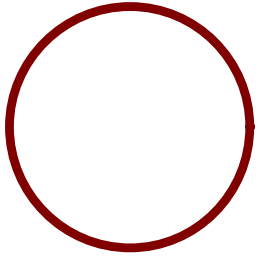
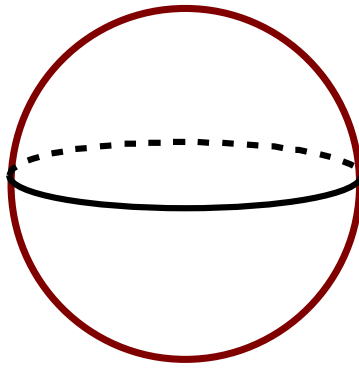
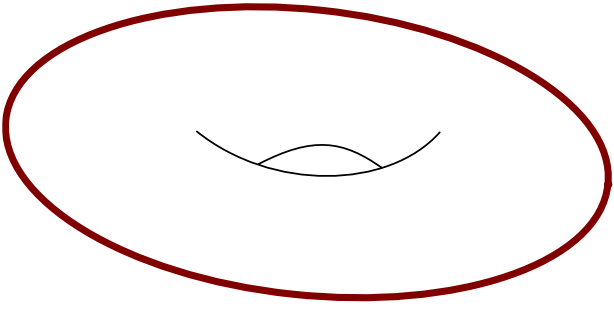
Euler Characteristic

$$\chi(K) = V - E + F$$



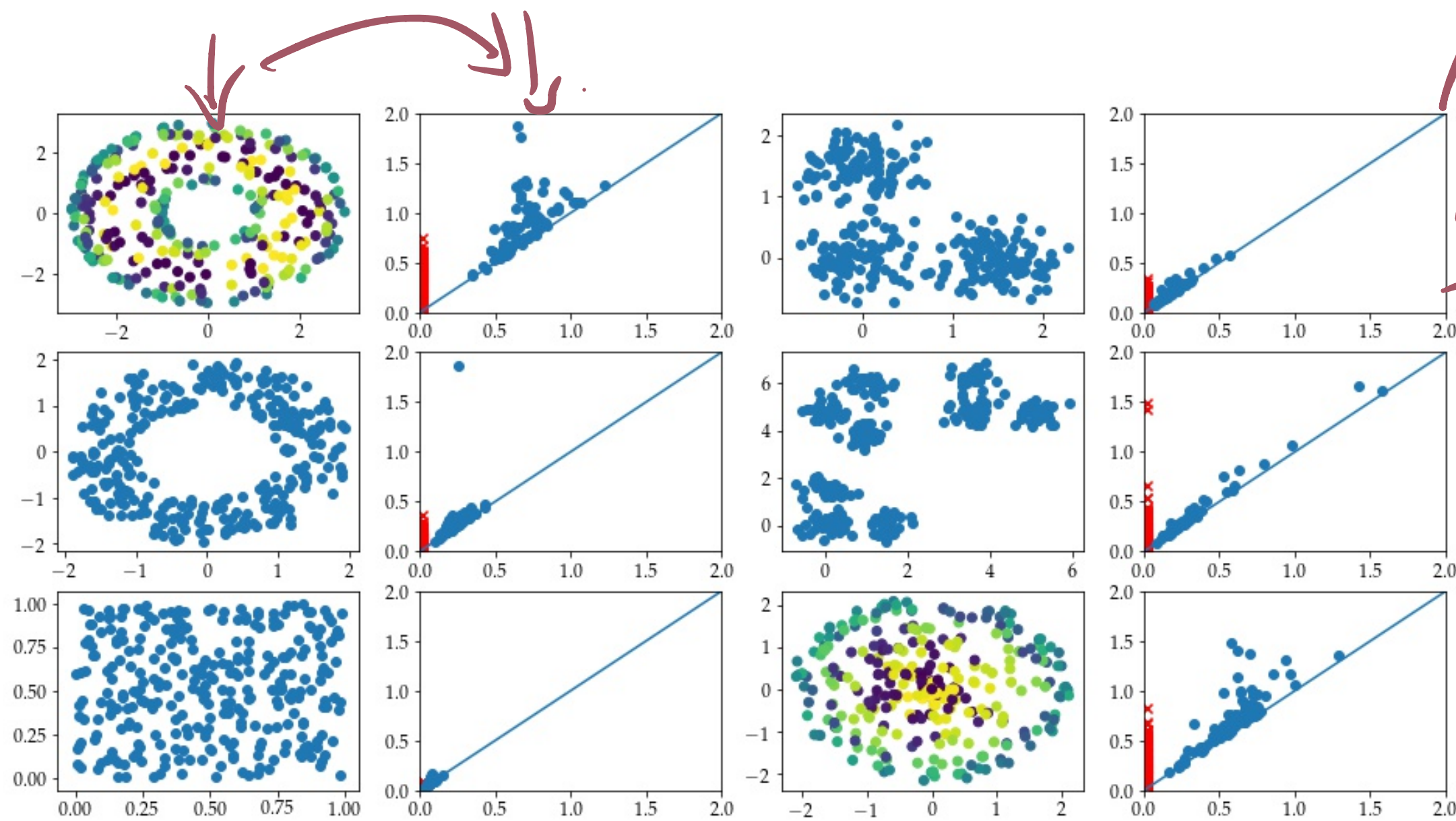
	Tetrahedron	Cube	Octahedron	Dodecahedron
V	4	8	6	20
$-E$	-6	-12	-12	-30
$+F$	4	6	8	12
χ	2	2	2	2

Topological Signature

K				
$\chi(K)$	1	0	2	0

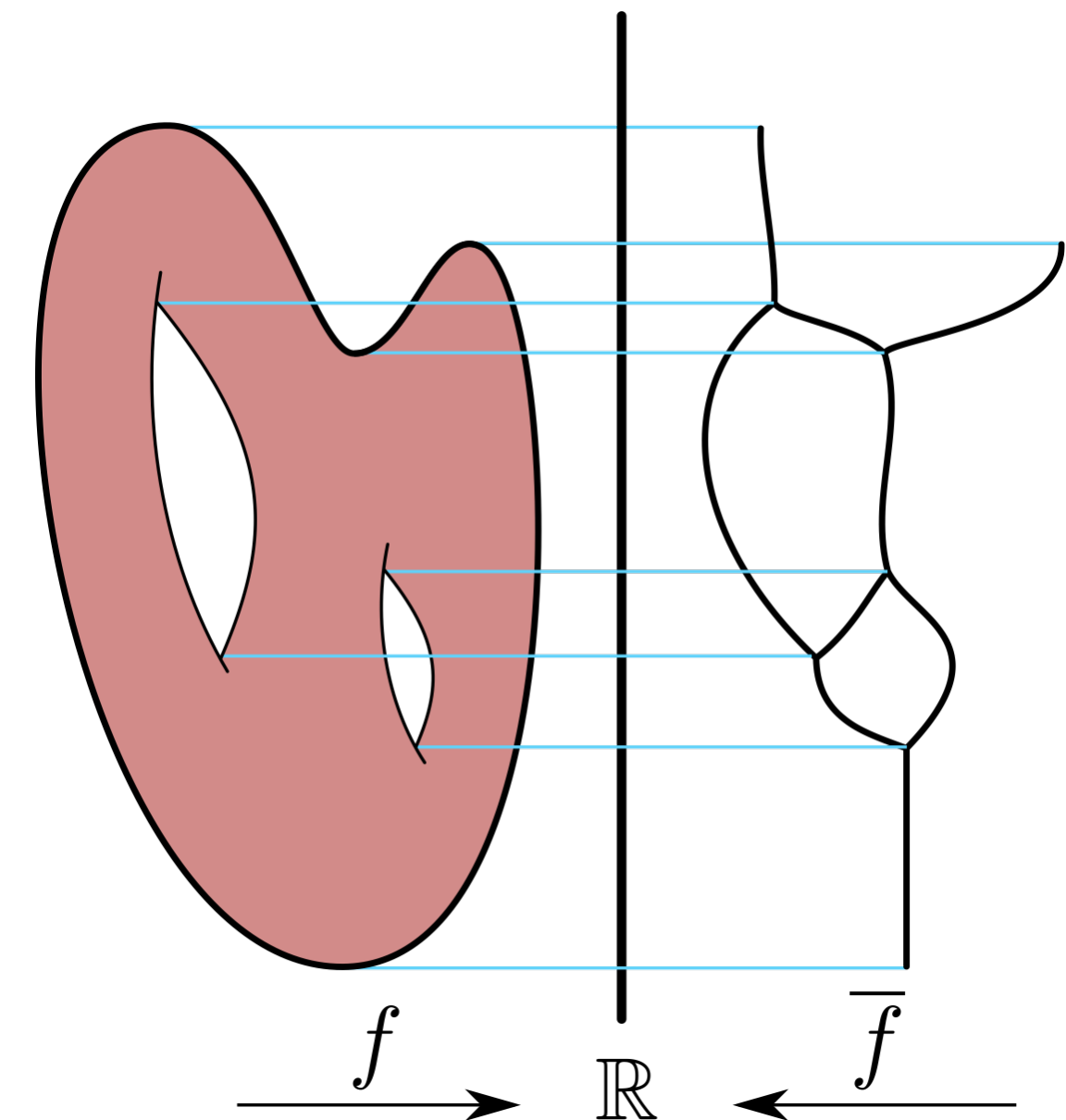
- If the objects are the **same**, the ECs are the **same**
- If the ECs are **different**, the objects are **different**

Quantification vs Representation of Shape



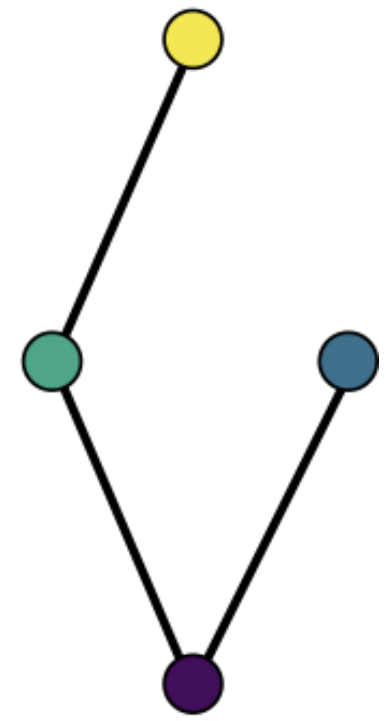
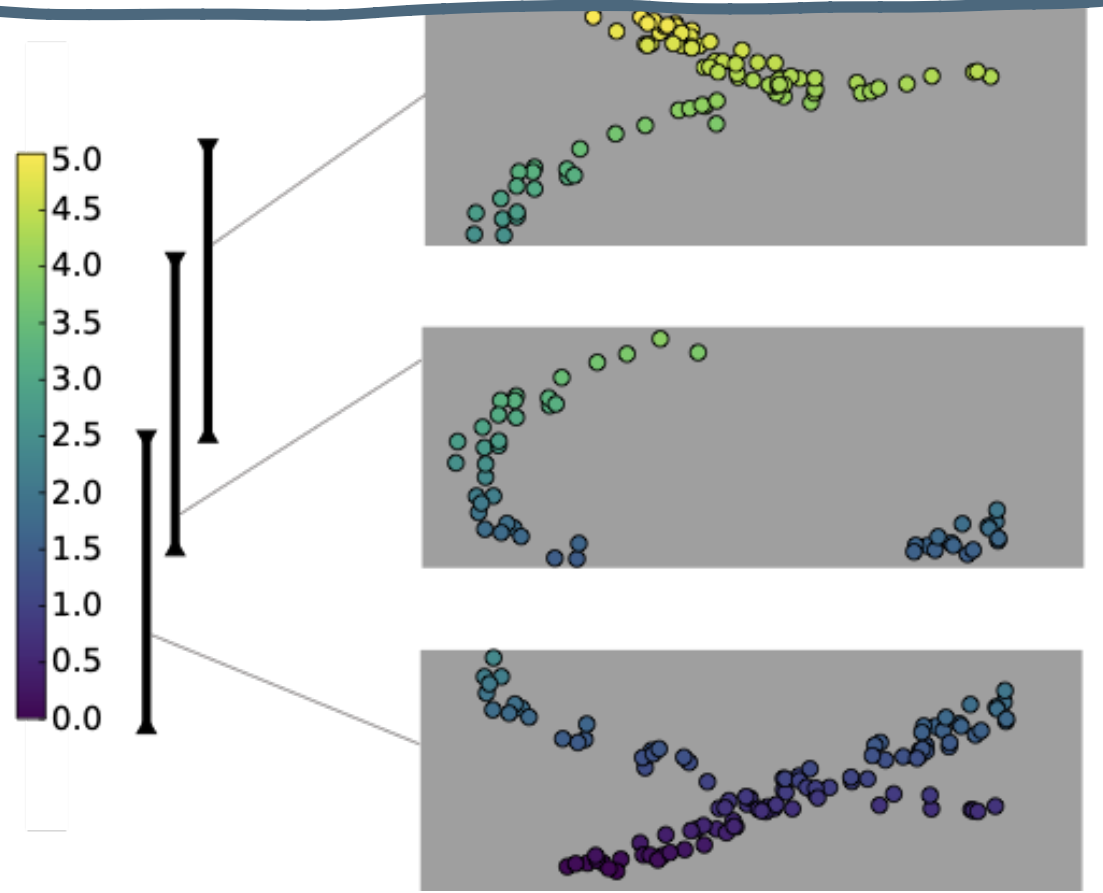
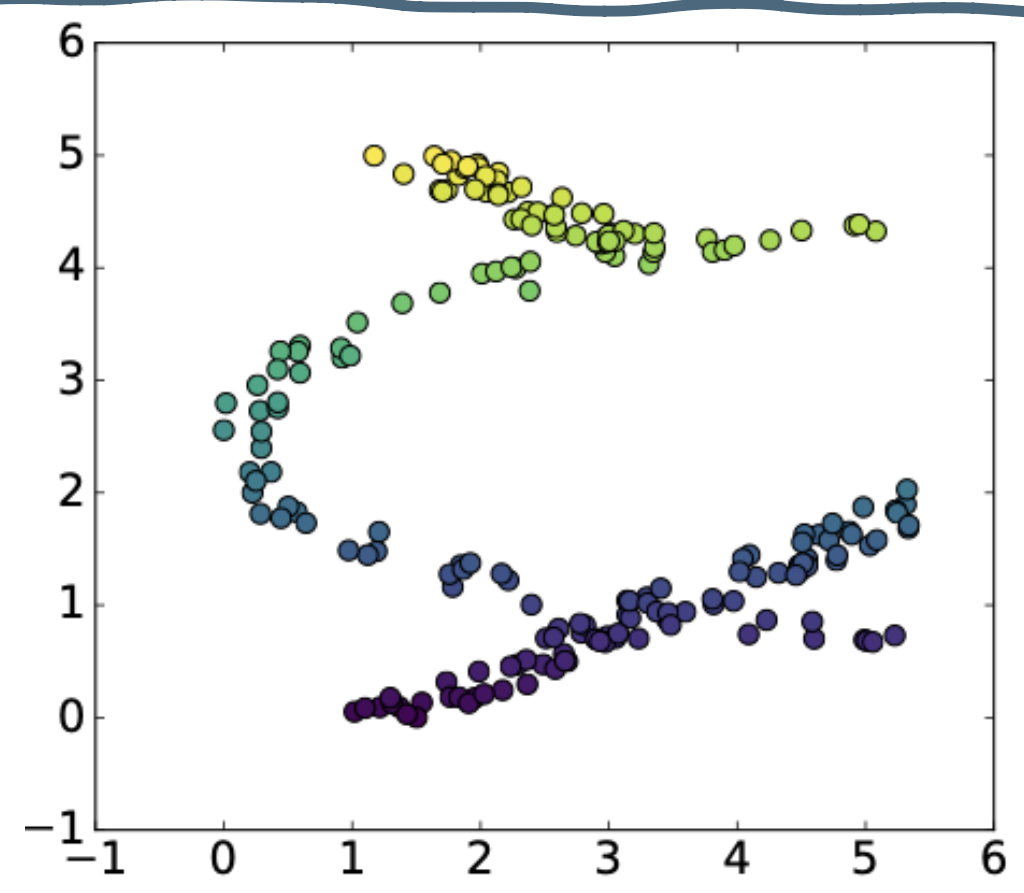
Persistent Homology

*Algebraic
Constructions*
- Easy to compute.
- Lossy.



Reeb graph

Graphical constructions
- Encode more info.
- Harder to compute.



Mapper

Current active research directions

- Multidimensional persistence
 - Machine learning, deep learning
 - Time series analysis
 - Dynamical systems
 - Visualization
- Application areas:
 - Neuroscience
 - Gene expression
 - Image processing
 - Sensor networks
 - Climate science

Goal for this course

- Understand the computation and interpretation of several commonly used tools in TDA
 - **Spaces and functions**
 - Complexes
 - Reeb graphs/Mapper *Merge Tree.*
 - **Homology**
 - Computation
 - Persistence and Stability
 - Maching/Deep learning, directional transforms, and further topics
- Know what types of data and/or are amenable to TDA methods
- Have experience working with current research (understanding new results, code with open-source code banks)

Intro to Topology Vocabulary

- Point set topology
- Cover some basic terms from Ch 1.1-1.3
- Depending on your math background, this might be obvious or this might seem ***IMPOSSIBLE***. If the latter, spend some time tonight trying some examples and read the textbook! :)

Topology

A topological space is a point set $\mathbb{T}$ with a set of subsets T (called open sets) such that

- $\emptyset, \mathbb{T} \in T$ ✓
- For every $\underline{U} \subseteq T$, the union of the sets in U is in T ✓
- For every finite $U \subseteq T$, the common intersection of the subsets in U is in T ✓

Ex. $\mathbb{T} = \underline{\{a, b, c\}}$, $T = \{\emptyset, \underline{T_1 = \{a\}}, \underline{T_2 = \{b\}}, T_3 = \{a, b\}, \underline{\mathbb{T} = \{a, b, c\}}\}$

$$\emptyset \cup T_1 = \emptyset \cup \{a\} = \{a\}$$

$$T_2 \cup T_3 = \{b\} \cup \{a, b\} = \{a, b\}$$

$$T_1 \cap T_2 = \{a\} \cap \{b\} = \underline{\emptyset}$$

$$T_3 \cap T_1 = \{a, b\} \cap \{a\} = \underline{\{a\}}$$

$$\mathbb{R} \setminus \{ \underline{(n-1, n+1)} \mid \underline{n \in \mathbb{Z}} \}$$

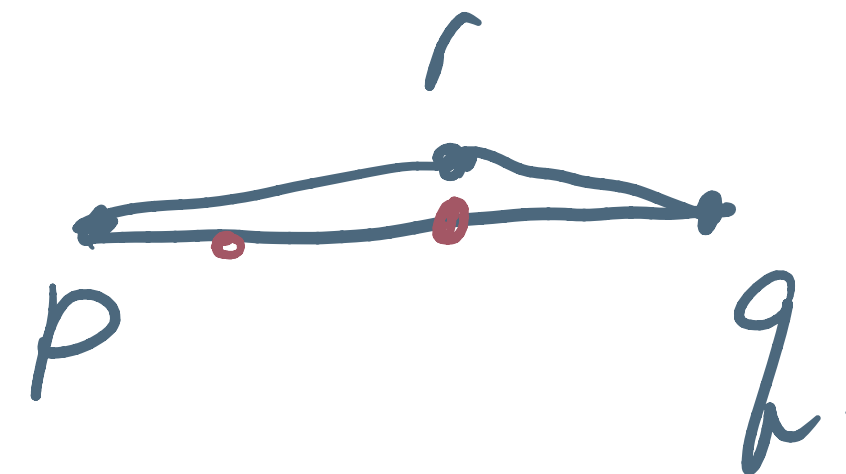
Metric

A metric is a pair $(\mathbb{T}, d)$ where $\mathbb{T}$ is a set, and $d : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{R}$ satisfies

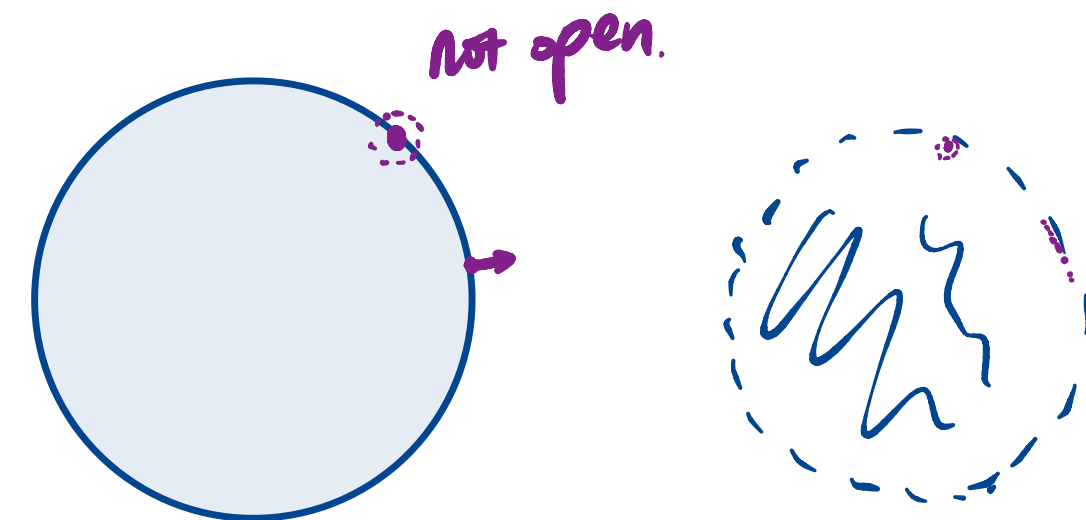
- $d(p, q) = 0$ if and only if $p = q$ ✓
- $d(p, q) = d(q, p)$ for all $p, q \in \mathbb{T}$
- $d(p, q) \leq d(p, r) + d(r, q)$ for all $p, q, r \in \mathbb{T}$

Ex. $\mathbb{T} = \mathbb{R}^2$, $d((a, b), (c, d)) = \sqrt{(c - a)^2 + (b - d)^2}$

$$\begin{aligned}
 d((a, b), (c, d)) &= \sqrt{(c-a)^2 + (b-d)^2} \\
 &= \underline{0} \geq 0 \quad \begin{matrix} c-a=0 \\ c=a \end{matrix} \quad \begin{matrix} b-d=0 \\ b=d \end{matrix}
 \end{aligned}$$



Open and Closed Sets



A set is open if it is in the topology T . A set is closed if its complement is open.

Ex. $\mathbb{T} = \{a, b, c\}, T = \{\emptyset, T_1 = \{a\}, T_2 = \{b\}, T_3 = \{a, b\}, \mathbb{T} = \{a, b, c\}\}$ *not open $\neq$ closed*
not closed $\neq$ open.

Ex. $\mathbb{T} = \mathbb{R}$ with metric topology. Intervals: $[1, 2], (3, 5), (2, \infty), [0, 3)$

	open?	closed?
$[1, 2]$	X	closed
$(3, 5)$	✓	not closed.
$(2, \infty)$	✓	closed.
$[0, 3)$	X	X

$$\mathbb{R} \setminus [1, 2] = (-\infty, 1) \cup (2, \infty)$$

$$\mathbb{R} \setminus (3, 5) = (-\infty, 3] \cup [5, \infty)$$

$$(-\infty, 2]$$

$$\mathbb{R} \setminus [0, 3) = (-\infty, 0) \cup [3, \infty)$$

Metric Topology

Given a metric space $(\mathbb{T}, d)$, an open metric ball is $B_0(c, r) = \{p \in \mathbb{T} \mid d(p, c) < r\}$

The metric topology consists of all unions of open metric balls

$$\{B_0(c, r) \mid c \in \mathbb{T}, 0 < r \leq \infty\}$$

Ex. $\mathbb{R}, \mathbb{R}^2$ $(\mathbb{T}, d)$

$\mathbb{R}^2$ Euclidean distance.

$\mathbb{R}$ $d(x, y) = |x - y|$

$$B_0(c, r) = \{x : |x - c| < r\} = (c - r, c + r)$$

$$B_0(2, 1) = (1, 3)$$

$$d((a, b), (c, d)) = \sqrt{(c - a)^2 + (d - b)^2}$$



Open Cover and Compactness

An open (closed) cover of a topological space $(\mathbb{T}, T)$ is a collection $\mathcal{C}$ of open (closed) sets so that $T \subseteq \bigcup_{U \in \mathcal{C}} U$.

Ex. $\mathbb{R}, \mathcal{C} = \{(n-1, n+1) \mid n \in \mathbb{Z}\}$

$$\mathbb{R} = \bigcup_{n \in \mathbb{Z}} (n-1, n+1)$$

$x \in \mathbb{R}$, let $n = \lfloor x \rfloor$

$$n \leq x < n+1$$

$$n-1 < x < n+1$$

$$x \in (n-1, n+1)$$

$x = 1.5$
 $\lfloor x \rfloor = 1$



If every open cover of $\mathbb{T}$ has a finite subcover, then it is compact. ←

Connectedness

A topological space $(\mathbb{T}, T)$ is disconnected if there are two disjoint non-empty open sets $U, V \in T$ so that $T = U \cup V$. A topological space is connected if it is not disconnected

Ex A $\neq (1,2) \cup (5,7) \subset \mathbb{R}$

$\nearrow U = (1,2), V = (5,7). \neq \emptyset \checkmark$

$$U \cap V = \emptyset \checkmark$$

$$A = U \cup V$$

disconnected.

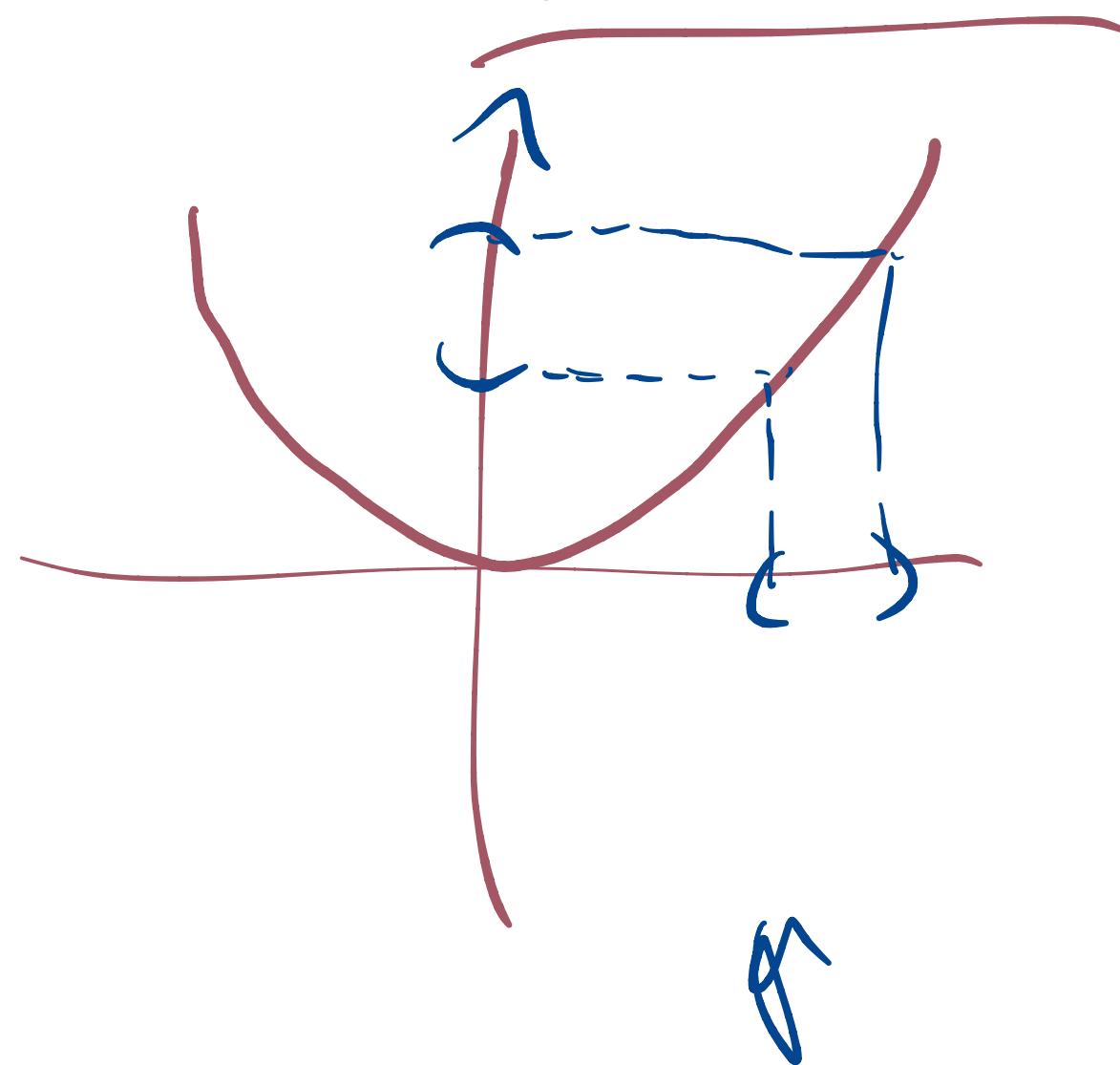
Maps

$$f^{-1}(Q) = \{t \in T \mid f(t) \in Q\}$$

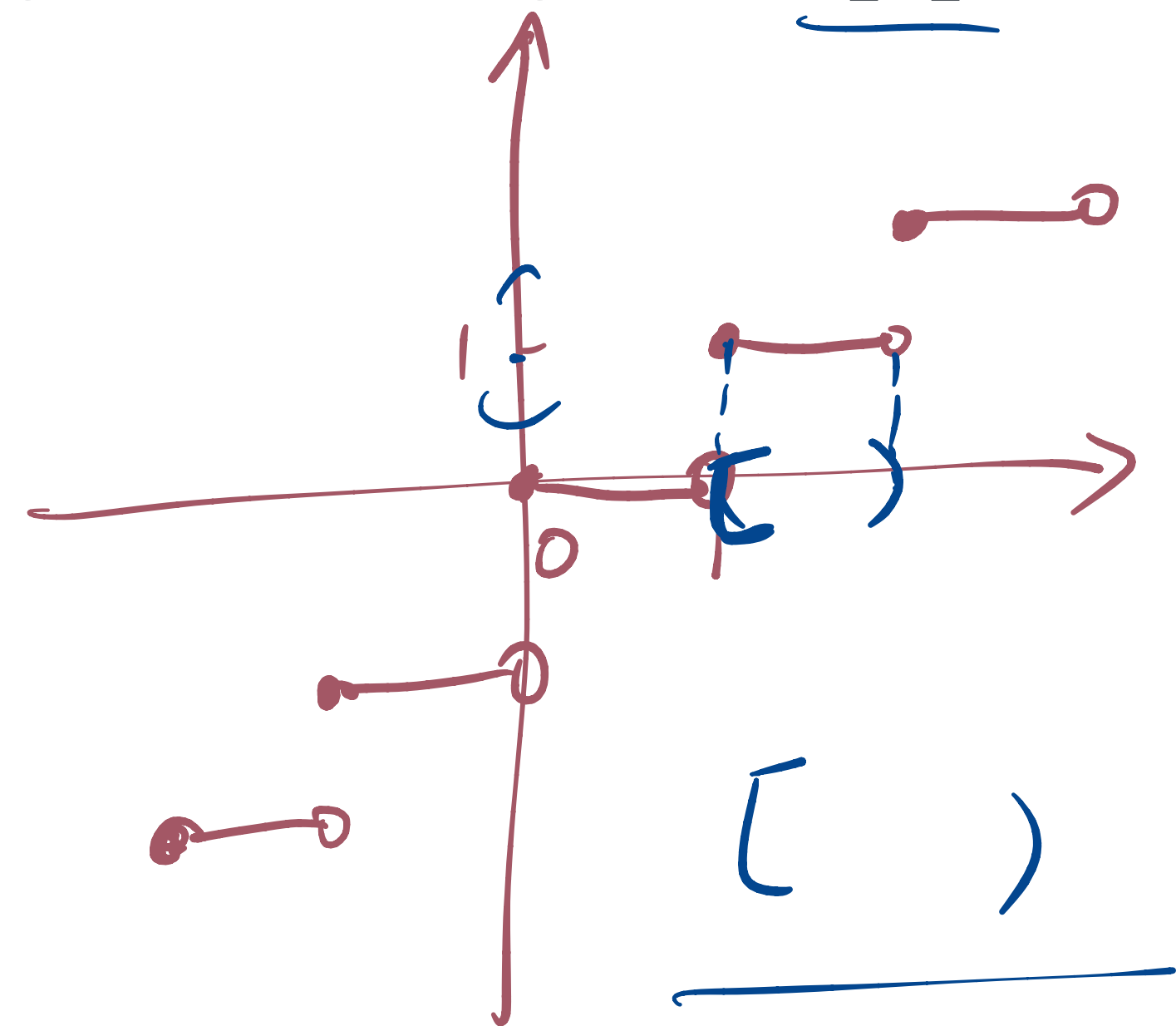
A function $f: T \rightarrow U$ is continuous if for every open set $Q \subseteq U$, $f^{-1}(Q)$ is open.
Continuous functions are also called maps.

Ex. $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$

$$f^{-1}(x) = \sqrt{x}$$



Ex. $g: \mathbb{R} \rightarrow \mathbb{R}, g(x) = \lfloor x \rfloor$ x



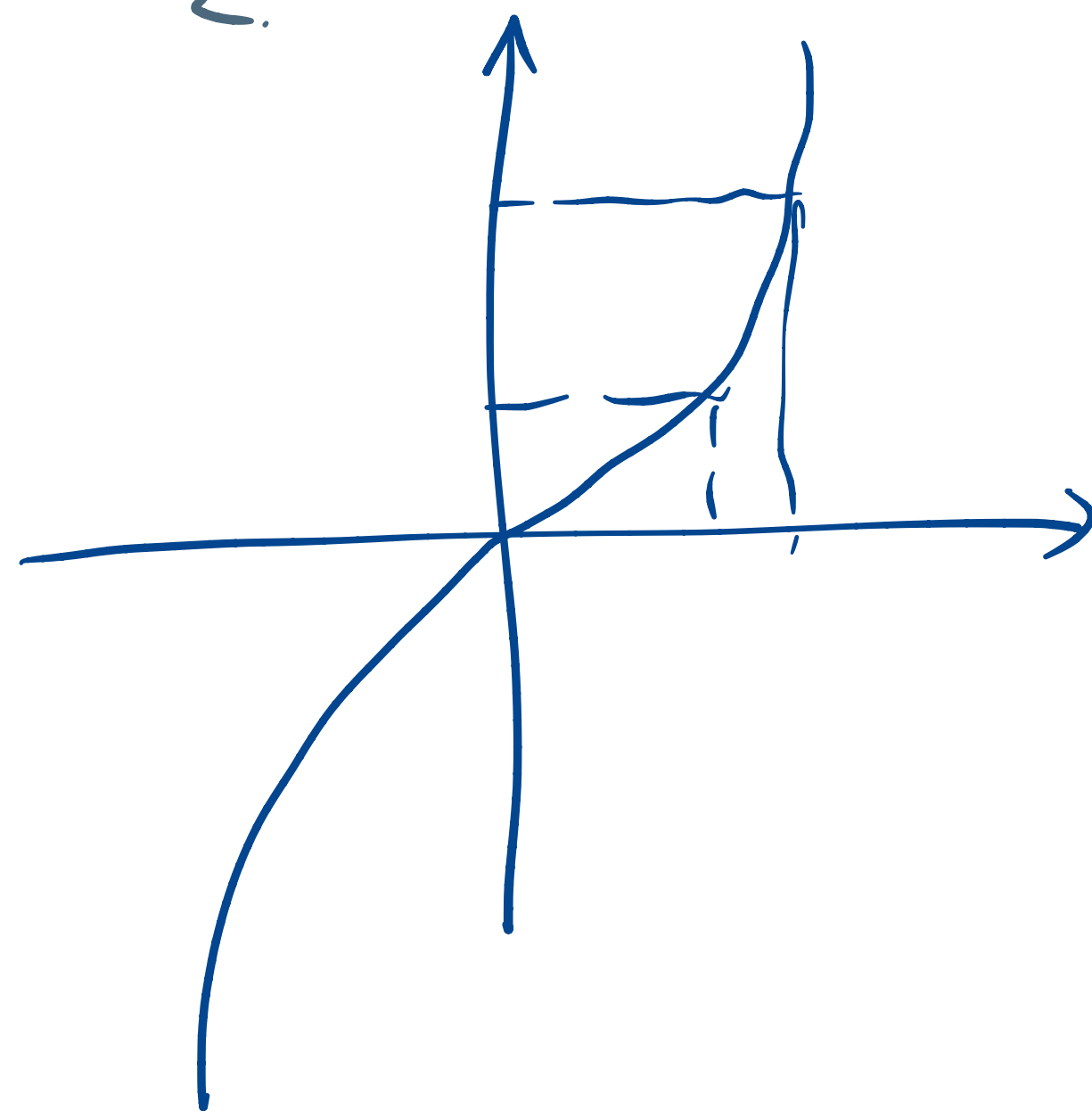
Embedding

continuous, continuous inverse

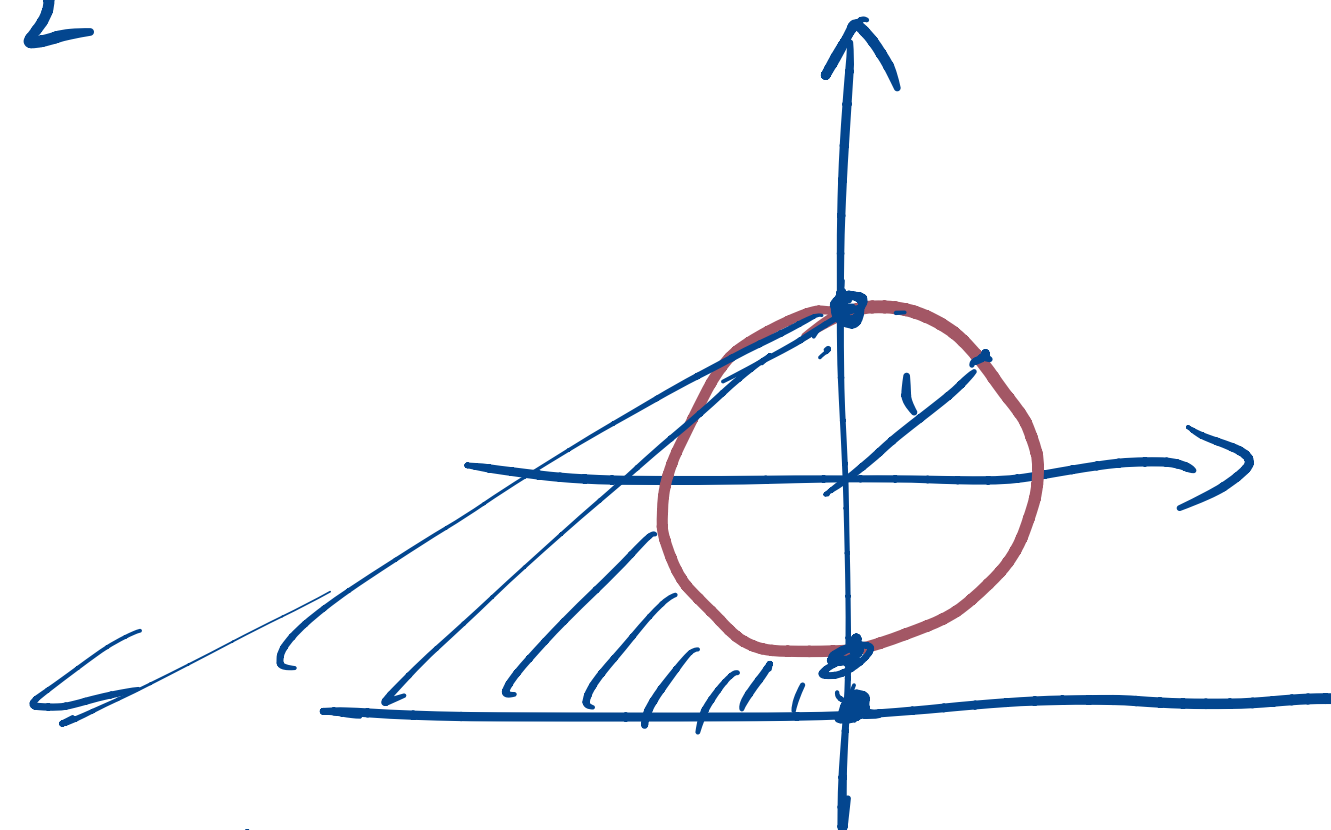
A map $g : T \rightarrow U$ is an embedding of T into U if g is injective

Injective (1-1): $f(x) = f(y)$ if and only if $x = y$

Ex. $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$



Ex. $g : S^1 \rightarrow \mathbb{R}^2$



$h : S^1 \rightarrow \mathbb{R}$
x embedding

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

$$g(x, y) = (x, y)$$

$$g(x_1, y_1) = g(x_2, y_2)$$

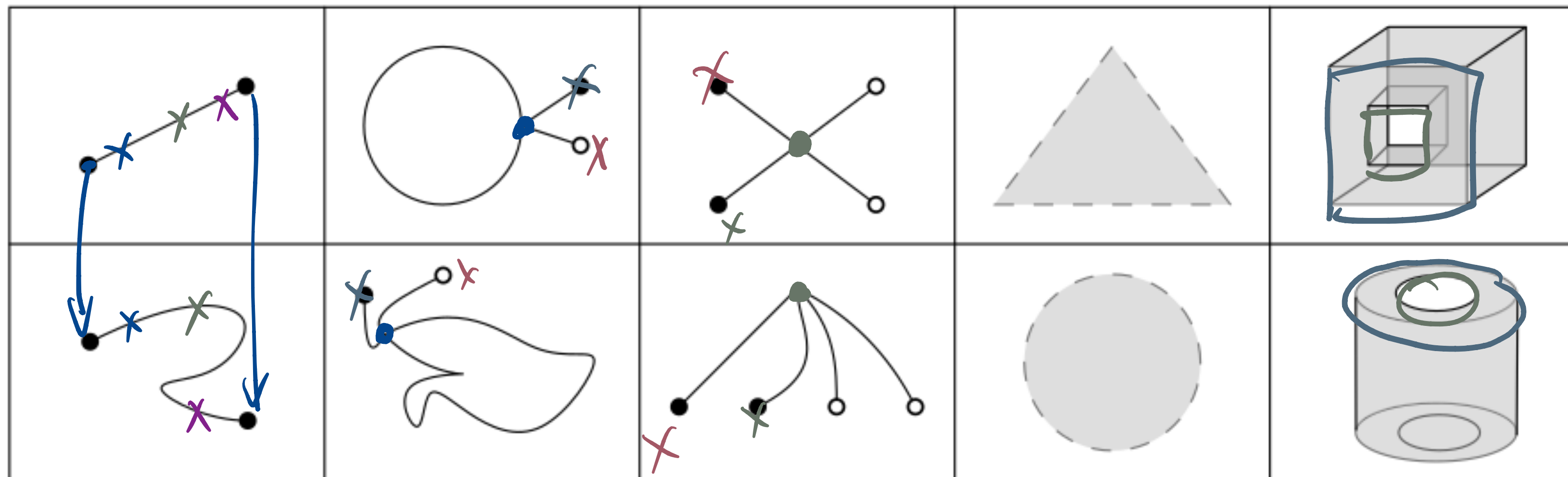
$$g^{-1}(x, y) = (x, y)$$

Homeomorphism

Let T and U be topological spaces.

A homeomorphism is a bijjective map $h : T \rightarrow U$ whose inverse is also continuous.

Two topological spaces are homeomorphic if there exists a homeomorphism between them.



Homeomorphism

Cheap trick:

If T and U are compact metric spaces, every bijective map from T to U has a continuous inverse.

if every open cover of T has a finite subcover.

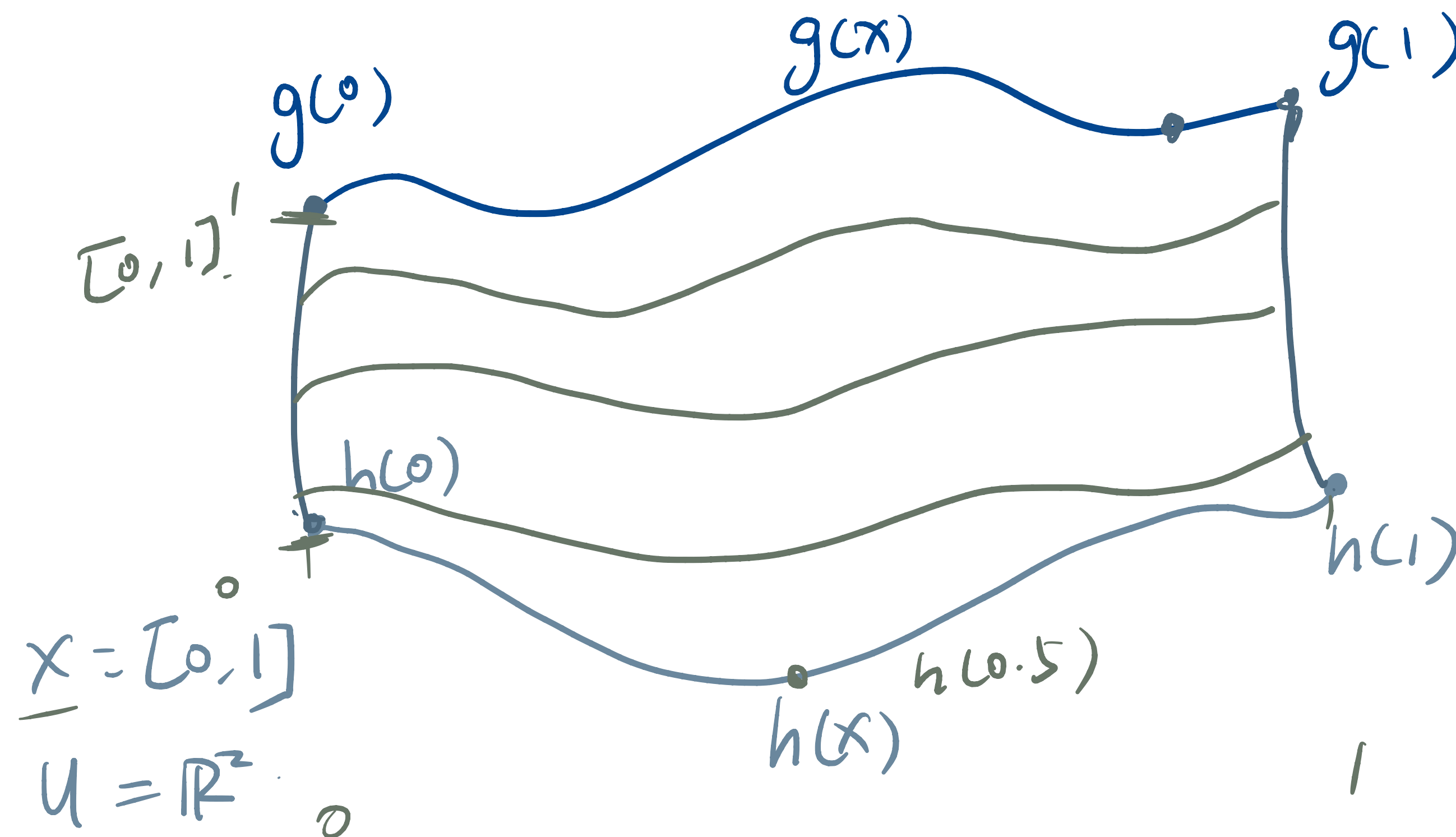
$$\mathbb{R} \quad \{(n-1, n+1) \mid n \in \mathbb{Z}\}$$

Homotopy

Let $\underline{g} : \underline{X} \rightarrow \underline{U}$ and $\underline{h} : \underline{X} \rightarrow \underline{U}$ be maps.

A homotopy is a map $\underline{H} : \underline{X} \times [0,1] \rightarrow \underline{U}$ such that $\underline{H}(\cdot, 0) = \underline{g}$ and $\underline{H}(\cdot, 1) = \underline{f}$

Two maps are homotopic if there is a homotopy connecting them.



$$H : [0, 1] \times [0, 1] \rightarrow \mathbb{R}^2$$

where you are on the curve.

how far along the deformation you are.

Definitions from Today.

Topology $\mathcal{T}$

Open / closed sets.

Clopen Set

Metric Topology

Continuous map

Connected

Compact.

Homeomorphism.

Embedding

Homotopy

Homotopy equivalent.

Retraction.

Deformation retract.

Problems to work on

- Presentation assignment starts next week (unless you volunteer!)
- DW 1.6.1. Be sure to explain why the constructions you have created are/are not Hausdorff.
- DW 1.6.2
- DW 1.6.3
- DW 1.6.4
- ~~DW 1.6.5~~
- DW 1.6.6

Python Set Up

- Install Miniconda if you do not already have Conda (<https://www.anaconda.com/docs/getting-started/installation>)
- Download tda-environment.yml to your Downloads folder
- Open Miniconda Prompt/Anaconda Prompt (Windows), or Terminal (Mac)
- Run the commands below.

- Mac User:

```
cd ~/Downloads
```

```
conda env create -f tda-environment.yml
```

```
conda activate tda-course
```

```
python -m ipykernel install --user \  
--name tda-course \  
--display-name "Python (tda-course)"
```

```
jupyter lab
```

- Windows User:

```
cd %USERPROFILE%\Downloads
```

```
conda env create -f tda-environment.yml
```

```
conda activate tda-course
```

```
python -m ipykernel install --user \  
--name tda-course \  
--display-name "Python (tda-course)"
```

```
jupyter lab
```


Try to run the Jupyter Notebook

TDA Lecture 1 — Python Environment Check

What to do

1. Make sure the kernel is **Python (tda-course)**.
2. Choose **Run → Run All Cells** (or run each cell with **Shift+Enter**).
3. If all cells run without errors and the plots appear, you are ready for the course.

```
import sys

print("Python version:", sys.version.split()[0])
print("Python executable:")
print(sys.executable)

print("\nIf the path above contains 'tda-course', you are using the intended environment.")
```

[]

Python

1. Check the course packages

If this cell runs without an error, the environment is installed correctly.

▶

```
import numpy as np
import matplotlib.pyplot as plt
```

Exercise 1

$\mathbb{T} = \{a, b, c\}, T = \{\emptyset, \underline{a}, \underline{b, c}, \mathbb{T}\}$ a set is closed if its complement is open.

- List all clopen subsets of $\mathbb{T}$

→ Find two disjoint nonempty open sets U, V , such that $\mathbb{T} = U \cup V$

- Prove that $\mathbb{T}$ is connected if and only if $\emptyset$ and $\mathbb{T}$ are the only clopen subsets of T

$\Rightarrow$ Suppose T is connected.
For contradiction, assume $\exists$ clopen set
 $U \subset T, U \neq \emptyset, U \neq T$.

$\hookrightarrow U$ is closed, its complement $V = T \setminus U$
is open, $\hookrightarrow U \neq T, V \neq \emptyset$.

Since U is open, $U \cap V = \emptyset$.

$$U \cup V = T$$

T is disconnected $\Rightarrow$ contradiction

$\Leftarrow$ Suppose $\emptyset$ and T are the only
clopen sets.

For contradiction, assume T is
disconnected.

there exist U, V , s.t. $U \cup V = T$
 $V = T \setminus U$

V is open, so U is closed.

U is clopen. $\Rightarrow$ contradiction.

Exercise 2

Show that any metric on a finite set induces a discrete topology