Extracting Geometrical Features From Data

Topological Data Analysis

Ulderico Fugacci

CNR - IMATI

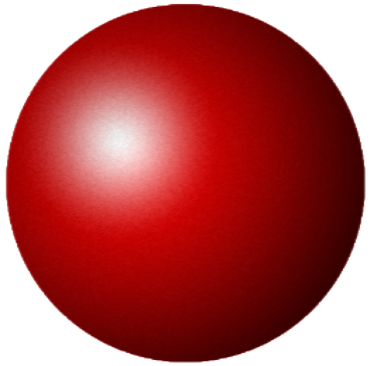
Topological Data Analysis

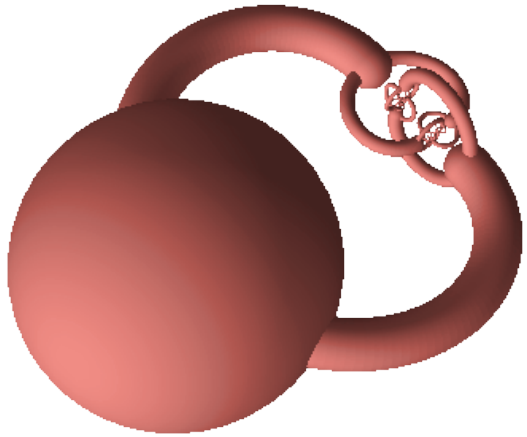
Outline:

Persistent Homology

Persistent Homology

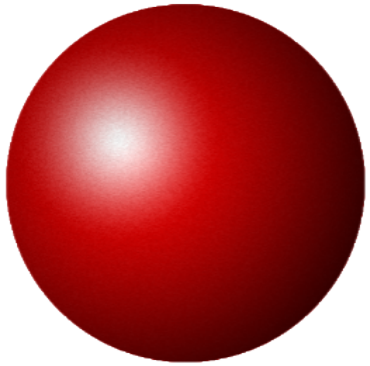
- *Do they have the same shape?*

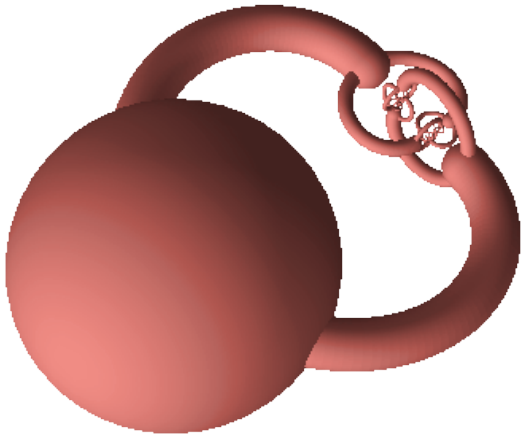




Persistent Homology

- ***Do they have the same shape?***

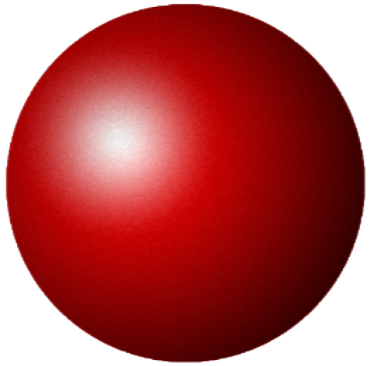


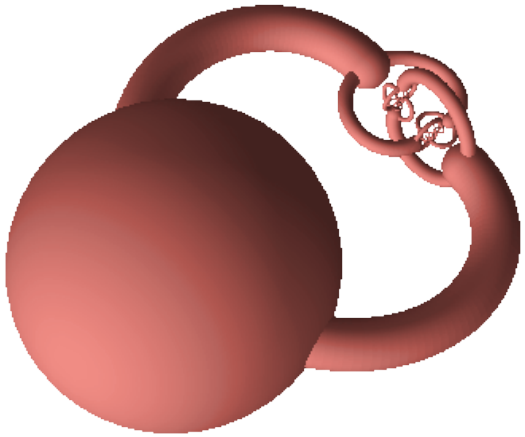


In Practice?	In Theory?

Persistent Homology

- ***Do they have the same shape?***





In Practice?

In Theory?

*They are **homeomorphic***

Persistent Homology

- *Do they have the same shape?*

Persistent Homology

✦ *Do they have the same shape?*

In Practice?	In Theory?

Persistent Homology

✦ ***Do they have the same shape?***

In Practice?

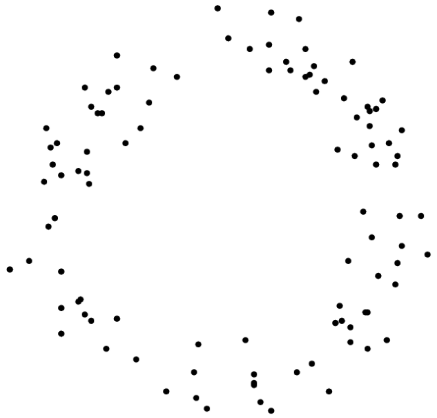
In Theory?

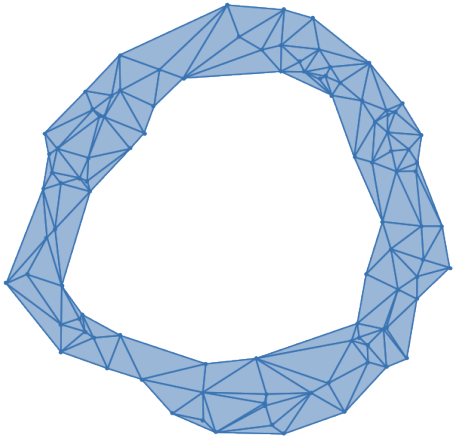
*They are **not homeomorphic***

Persistent Homology

✦ ***Which is the shape of a given data?***

Persistent homology allows for the retrieval of the ***"actual"*** homological information of a data





Point Cloud Dataset

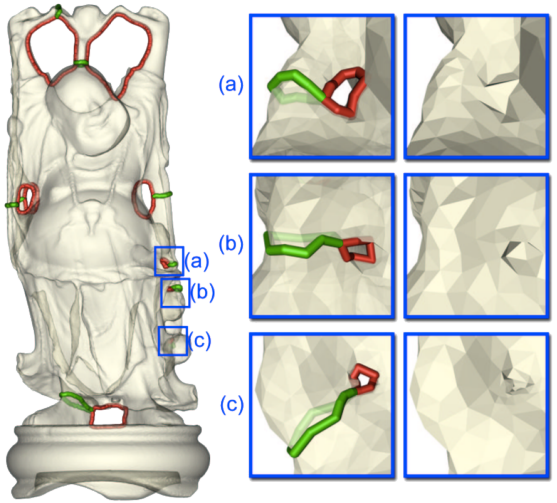
Topological Nature of the "Underlying" Shape

Image from [Bauer 2015]

Persistent Homology

✦ ***Which is the shape of a given data?***

Persistent homology allows for the retrieval of the ***"actual"*** homological information of a data



Noisy Dataset → ***Relevant Homological Information***

Image from [Dey et al. 2008]

Persistent Homology

In a Nutshell:

Persistent homology allows for
describing the changes in the shape of an evolving object

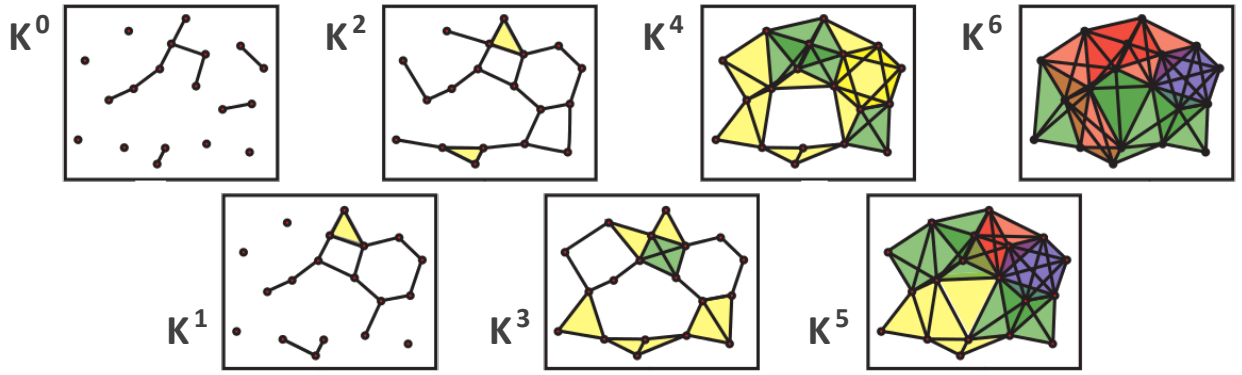


Image from [Ghrist 2008]

Persistent Homology

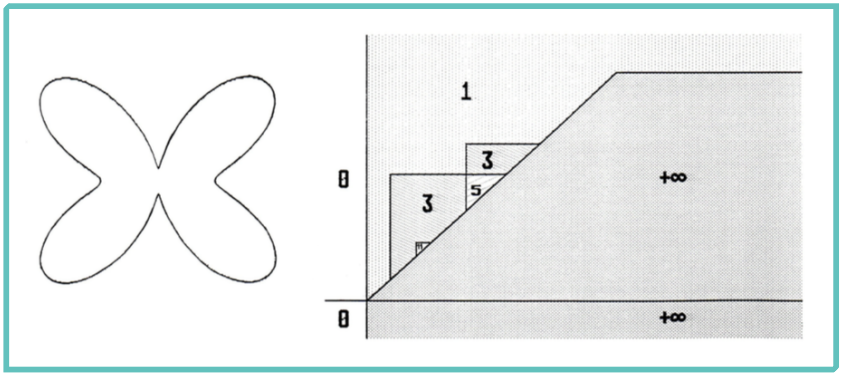
An Evolving Notion:

1990

Frosini

Size Functions:

- ***Estimation of natural pseudo-distance*** between shapes endowed with a function *f*
- Tracking of the ***connected components*** of a shape along its evolution induced by *f*



Actually, this coincides with ***persistent homology in degree 0***

Image from [Frosini 1992]

Persistent Homology

An Evolving Notion:

Incremental Algorithm for Betti Numbers:

- Introduction of the notion of ***filtration***
- De facto computation of ***persistence pairs***

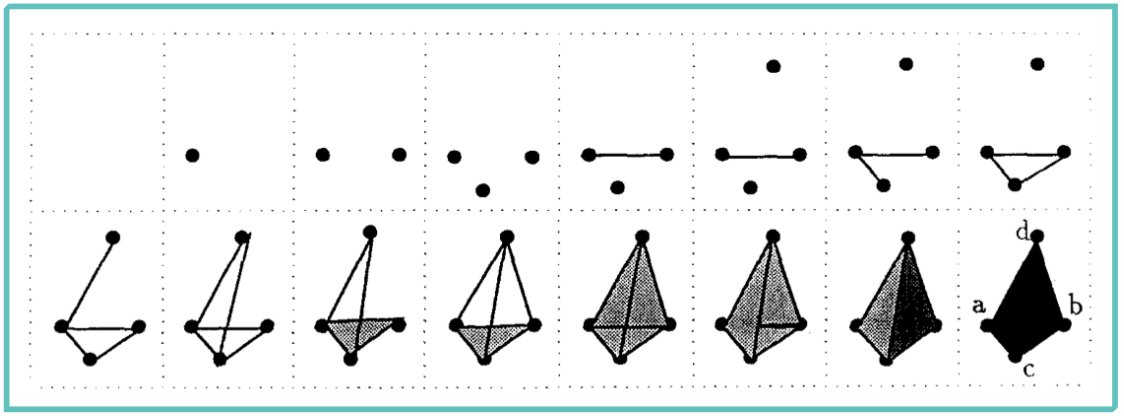


Image from [Delfinado, Edelsbrunner 1995]

Persistent Homology

An Evolving Notion:

Homology from Finite Approximations:

- ***Extrapolation of the homology*** of a metric space from a ***finite point-set approximation***
- Introduction of ***persistent Betti numbers***

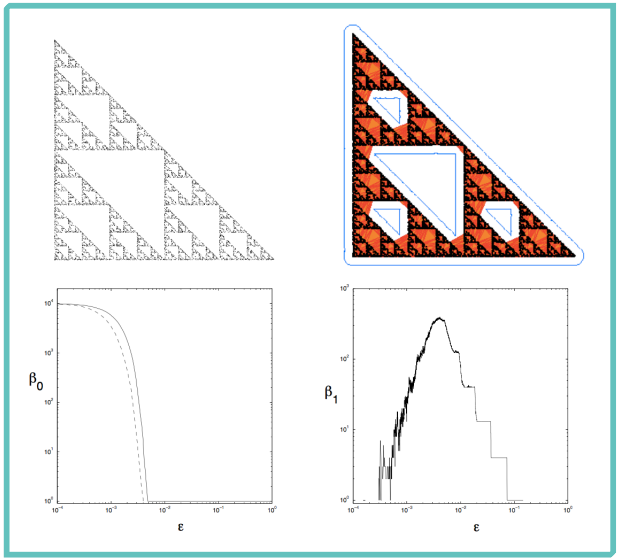


Image from [Robins 1999]

Persistent Homology

An Evolving Notion:

Topological Persistence:

- Introduction and algebraic formulation of the notion of ***persistent homology***
- ***Description of an algorithm*** for computing persistent homology

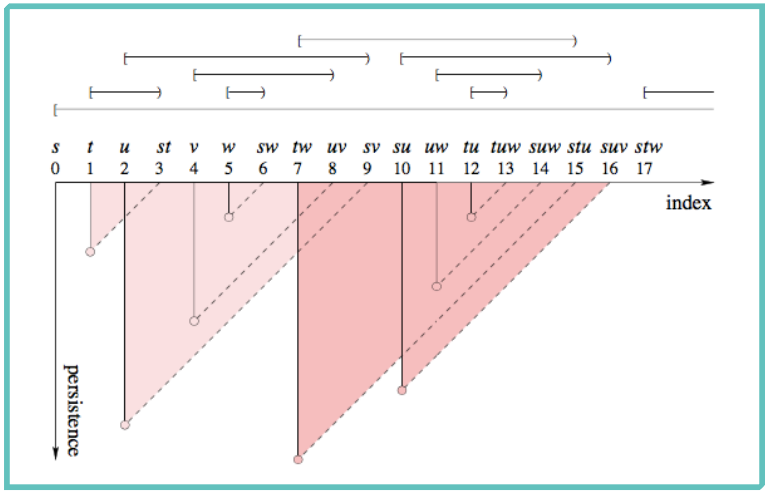
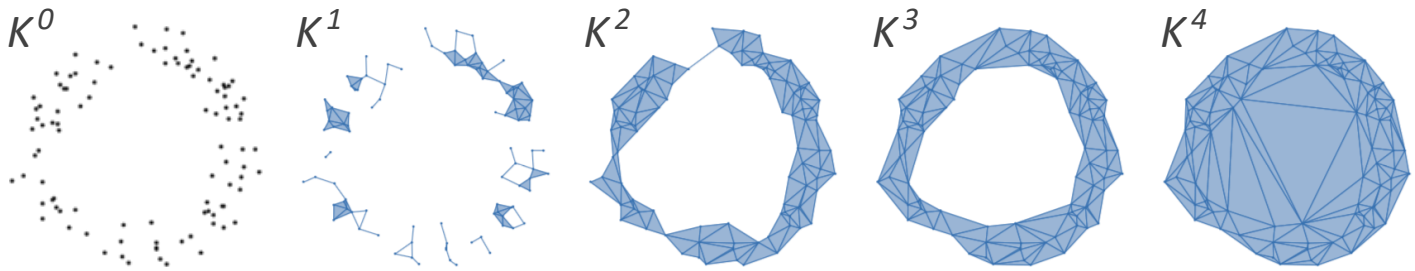


Image from [Edelsbrunner et al. 2002]

Persistent Homology

Definition:

Intuitively, a ***filtration $\mathcal{F}$*** is a finite ***"growing" sequence*** of simplicial complexes

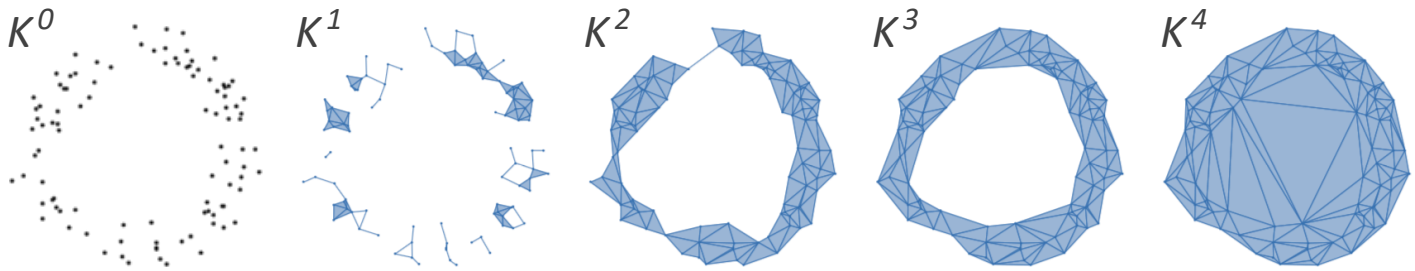


*Formally, a **filtration $\mathcal{F}$** of a simplicial complex K is a collection of subcomplexes $\{ K^p \}_{p \in \mathbb{R}}$ of K for which, **given any p, q $\in \mathbb{R}$ such that p ≤ q,***

$$K^p \subseteq K^q$$

Persistent Homology

Most of the techniques ***transforming a dataset into a simplicial complex depending on the choice of a parameter*** *actually produce a filtration* $\{ K^p \}_{p \in \mathbb{R}}$



Working Assumption:

We can always pretend that parameter *p varies over* $\mathbb{N}$

Persistent Homology

Definition:

Given a filtration $\mathcal{F} := \{ K^p \}_{p \in \mathbb{N}}$, a value $i \in \mathbb{N}$, and a field $\mathbb{F}$, the ***i^{th} persistence module M of $\mathcal{F}$ over $\mathbb{F}$*** *is defined as the* ***finitely generated graded $\mathbb{F}[x]$-module***

$$M := \bigoplus_{p \in \mathbb{N}} M_p$$

where:

- ***$M_p := H_i(K^p; \mathbb{F})$*** *, the set of* ***homogeneous elements of grade p***
- *The* ***action $x^{q-p} \cdot h$ over an element h of grade p*** *is defined as* ***$\mu_i^{p,q}(h)$****, where:*
  - ***$\mu_i^{p,q}(h): H_i(K^p; \mathbb{F}) \longrightarrow H_i(K^q; \mathbb{F})$*** *is the linear map induced by the inclusion $K^p \subseteq K^q$*

Persistent Homology

Theorem (structure for finitely generated graded modules over a PID):

Any persistence module M can be expressed as

$$M \cong \bigoplus_{k=1}^{n} \mathbb{F}[x](-r_k) \oplus \bigoplus_{j=1}^{m} \left(\mathbb{F}[x]/(x^{q_j - p_j}) \right)(-p_j)$$

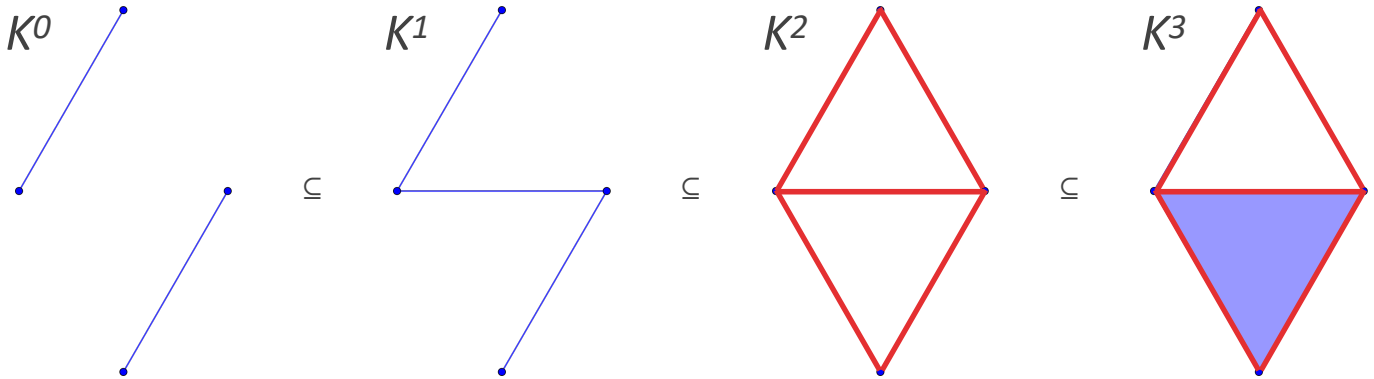
So, ***M is completely determined by the collection of values r_k and of pairs (p_j, q_j)***

Such descriptors are typically expressed as pairs, called ***persistence pairs*** of M, of the kind ***(r_k, ∞)*** and ***(p_j, q_j)***

Persistent Homology

Intuitively:

Given a filtration $\mathcal{F} := \{ K^p \}_{p \in \mathbb{N}}$, *a* ***persistence pair*** $(p,q) \in \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$ *with* $p < q$ *represents a* ***homological class*** *that is* ***born at step p*** *and* ***dies at step q***



$(2, \infty)$ essential pair **$(2, 3)$**

Persistent Homology

Differently from homology, persistent homology provides
a notion of "shape" closer to our everyday perception

It is possible to ***compare two shapes*** by comparing their ***homology groups***

Persistent Homology

Differently from homology, persistent homology provides
a notion of "shape" closer to our everyday perception

It is possible to ***compare two shapes*** by comparing their ***hom… …ups***

PERSISTENCE PAIRS

Persistent Homology

Differently from homology, persistent homology provides
a notion of "shape" closer to our everyday perception

It is possible to ***compare two shapes*** by comparing their ***hom... ...ups***

PERSISTENCE PAIRS

In order to better perform the above task, we need:

- ***Visual*** and ***descriptive representations*** for persistence pairs
- Notions of ***distance*** between sets of persistence pairs and ***stability results***

Bibliography

Some References:

- ***Persistent Homology:***
  - U. Fugacci, S. Scaramuccia, F. Iuricich, L. De Floriani. ***Persistent homology: a step-by-step introduction for newcomers***. Eurographics Italian Chapter Conference, pages 1-10, 2016.