Dynamical Systems & Applications — 26/09/23

An introduction to Topological Data Analysis

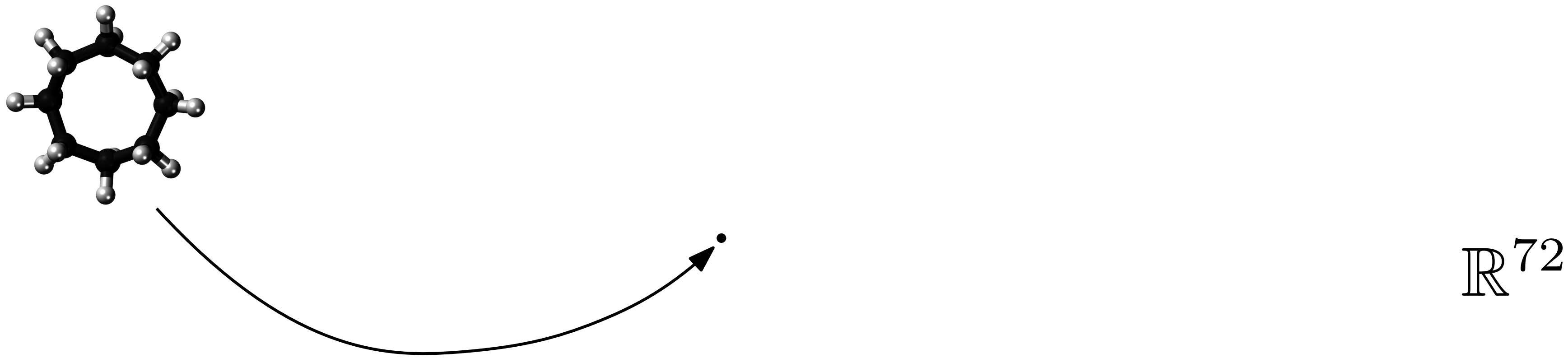
Part I/IV: Topological invariants

https://raphaeltinarrage.github.io

Last update: September 26, 2023

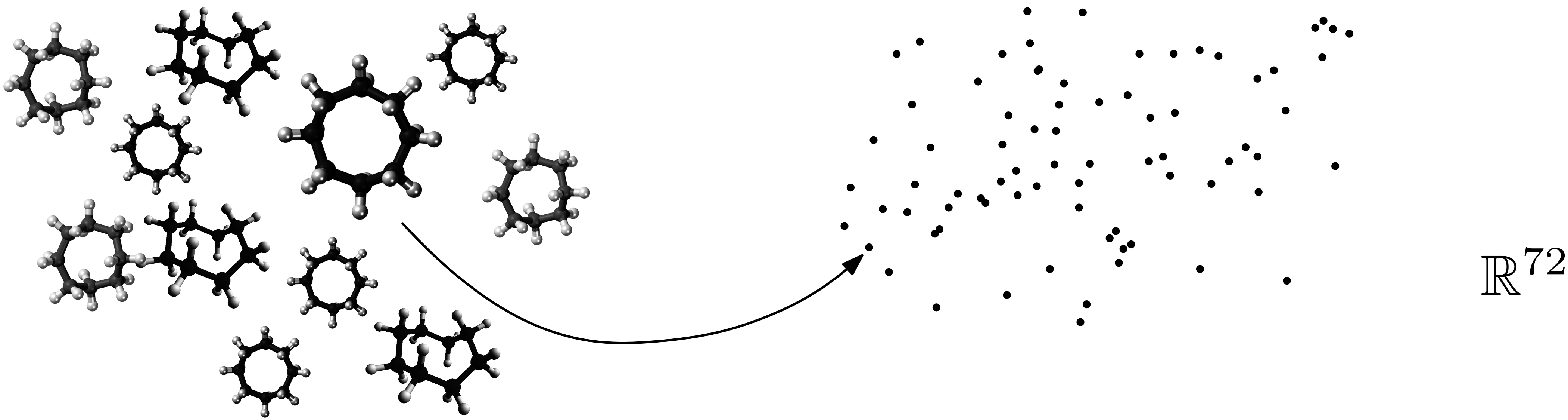
[Martin, Thompson, Coutsias and Watson, Topology of cyclo-octane energy landscape, 2010]

The cyclo-octane molecule $\mathrm{C_8H_{16}}$ contains 24 atoms.
Each atom has 3 spatial coordinates.
Hence a conformation of a molecule can be summarized by **a point** in $\mathbb{R}^{72}$ $(3 \times 24 = 72)$.



[Martin, Thompson, Coutsias and Watson, Topology of cyclo-octane energy landscape, 2010]

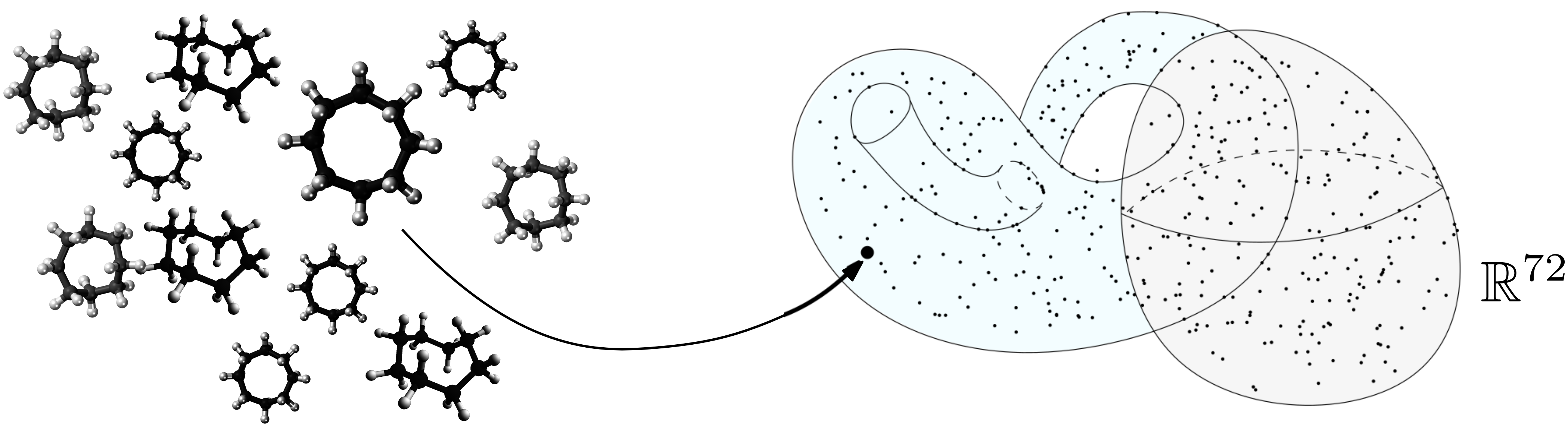
The cyclo-octane molecule $\mathrm{C_8H_{16}}$ contains 24 atoms.
Each atom has 3 spatial coordinates.
Hence a conformation of a molecule can be summarized by **a point** in $\mathbb{R}^{72}$ $(3 \times 24 = 72)$.



By considering a lot of such molecules, we obtain a **point cloud** in $\mathbb{R}^{72}$.

[Martin, Thompson, Coutsias and Watson, Topology of cyclo-octane energy landscape, 2010]

The cyclo-octane molecule $\mathrm{C_8H_{16}}$ contains 24 atoms.
Each atom has 3 spatial coordinates.
Hence a conformation of a molecule can be summarized by **a point** in $\mathbb{R}^{72}$ $(3 \times 24 = 72)$.

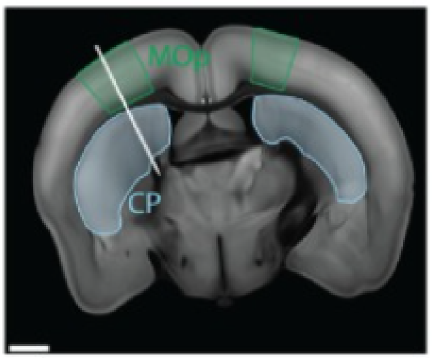


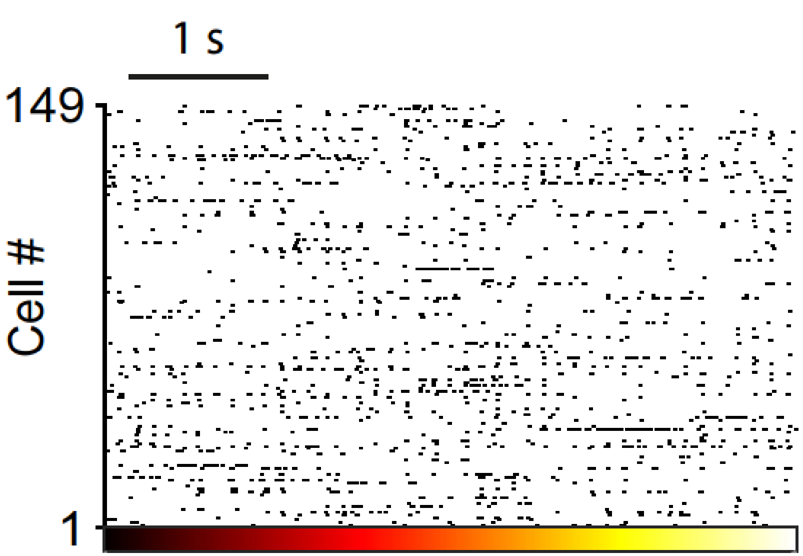
By considering a lot of such molecules, we obtain a **point cloud** in $\mathbb{R}^{72}$.

The authors show that this point cloud lies close to a small dimensional object: **the union of a sphere and a Klein bottle**.

[Richard J. Gardner et al, Toroidal topology of population activity in grid cells, 2022]

The authors recorded spikes of grid cells from rat brains. Then, they applied dimensionality reduction to the firing matrix.

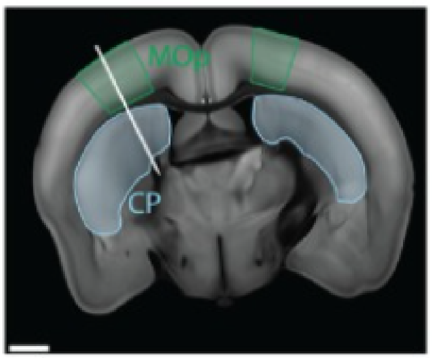


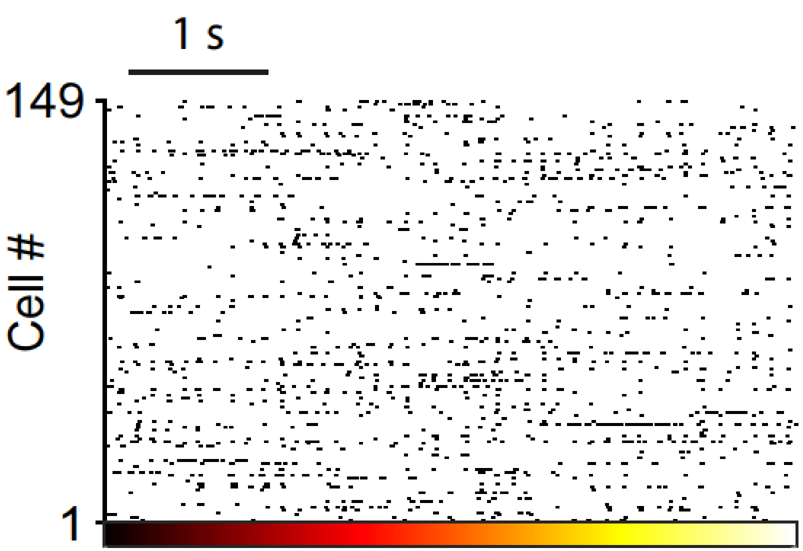


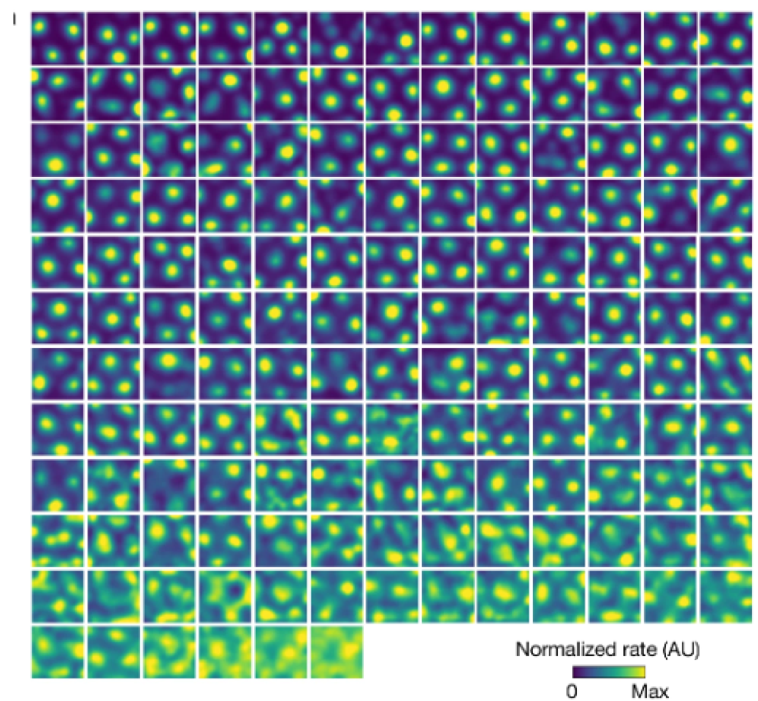
Aperitivo topológico — Neurofisiologia

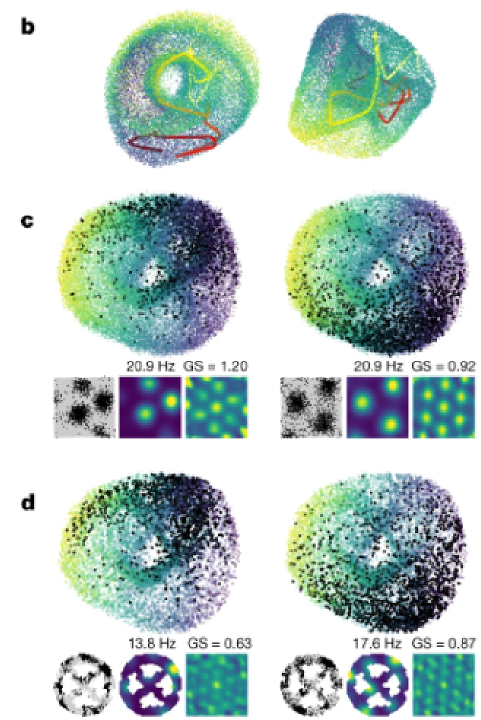
[Richard J. Gardner et al, Toroidal topology of population activity in grid cells, 2022]

The authors recorded spikes of grid cells from rat brains. Then, they applied dimensionality reduction to the firing matrix.
By applying persistent homology, they observed the homology of a **torus**.







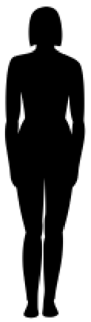


Aperitivo topológico — Biologia

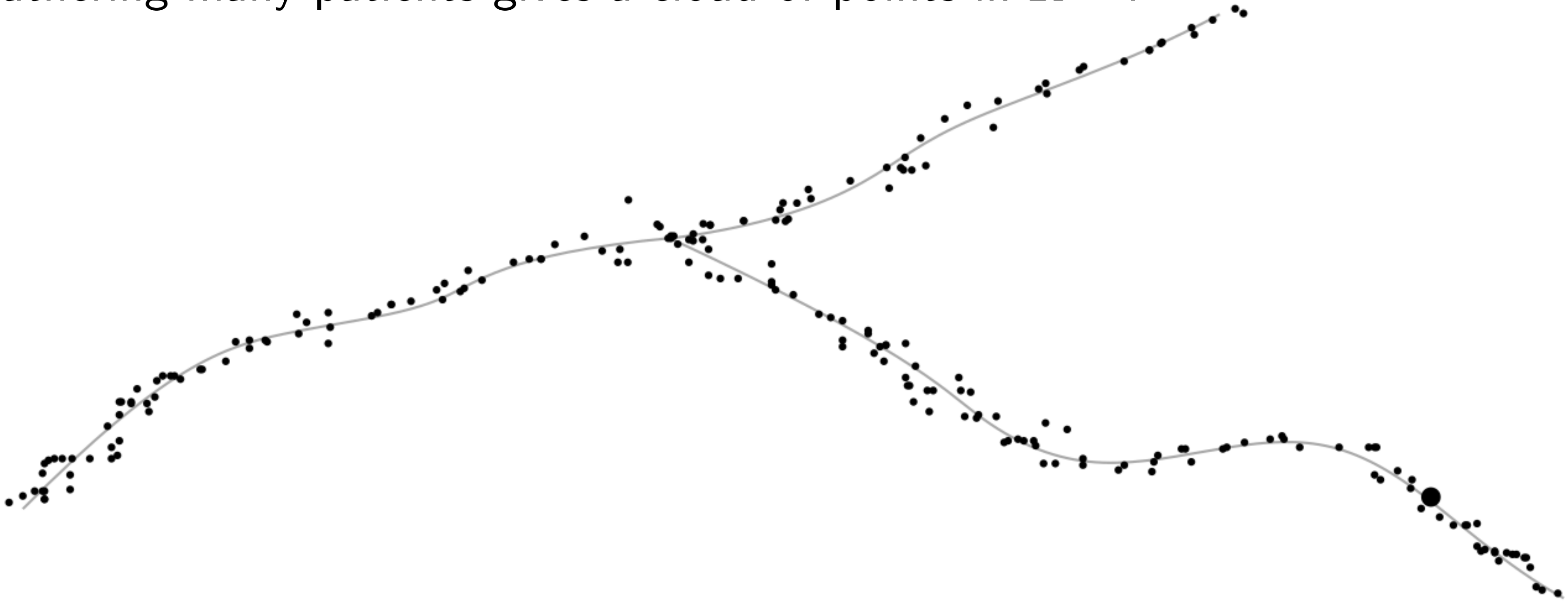
[Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival, Monica Nicolau, Arnold J Levine, and Gunnar Carlsson, *Proceedings of the National Academy of Sciences*, 2011]

The authors study tissues from patients infected by breast cancer. They obtain 262 genomic variables per patient.

$$(x_1, x_2, ..., x_{262}) \longleftarrow$$



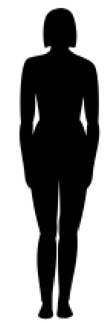
Gathering many patients gives a cloud of points in $\mathbb{R}^{262}$.



[Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival, Monica Nicolau, Arnold J Levine, and Gunnar Carlsson, *Proceedings of the National Academy of Sciences*, 2011]

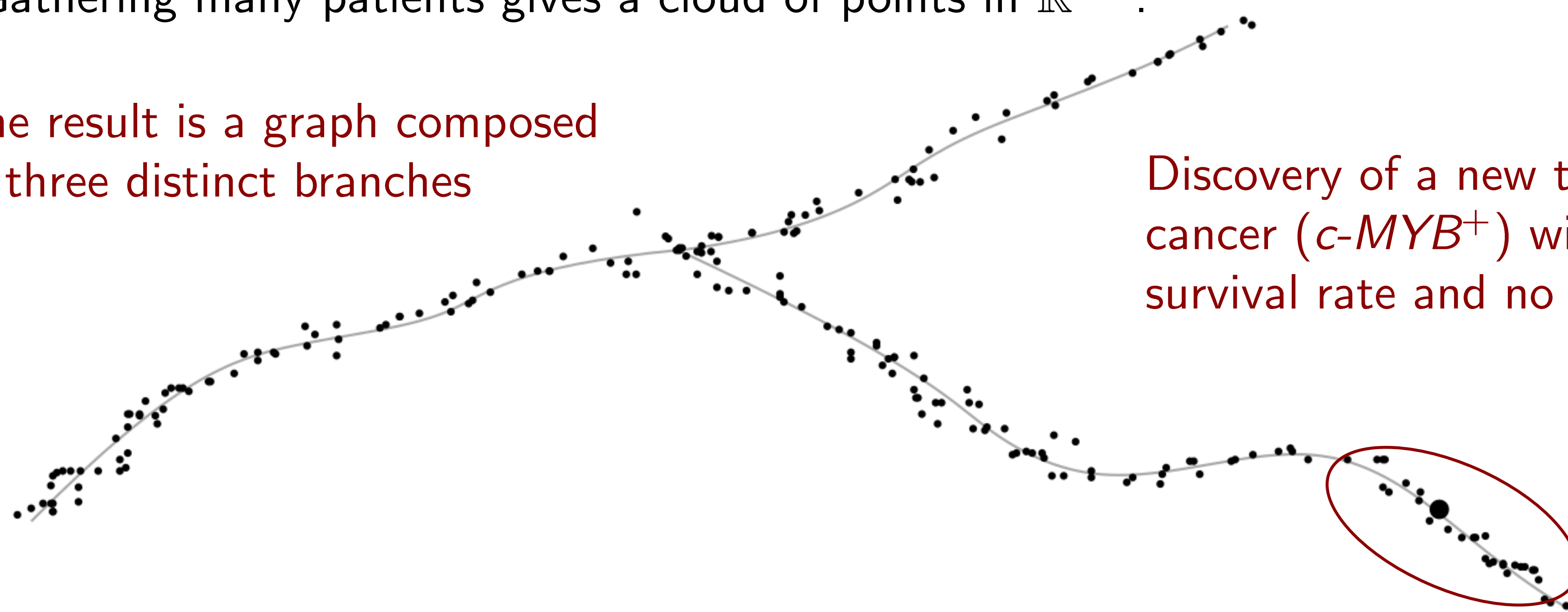
The authors study tissues from patients infected by breast cancer. They obtain 262 genomic variables per patient.

$$(x_1, x_2, ..., x_{262}) \longleftarrow$$



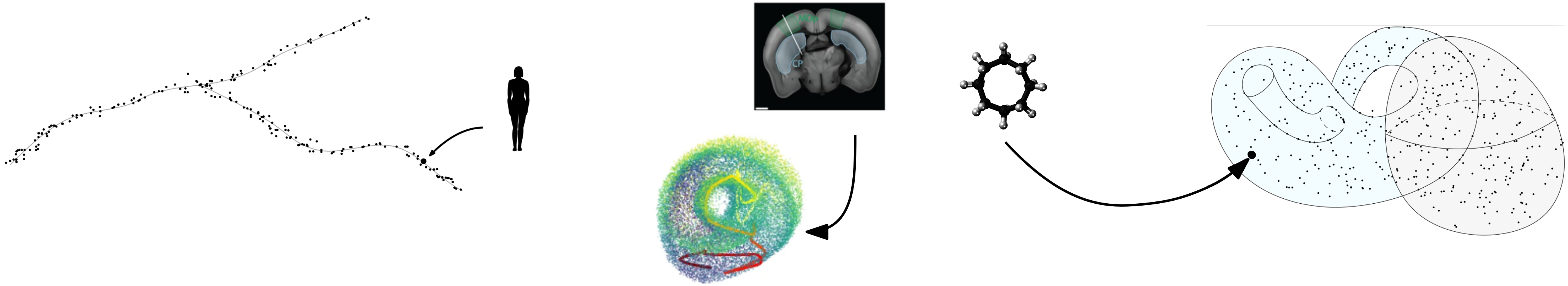
Gathering many patients gives a cloud of points in $\mathbb{R}^{262}$.

The result is a graph composed of three distinct branches

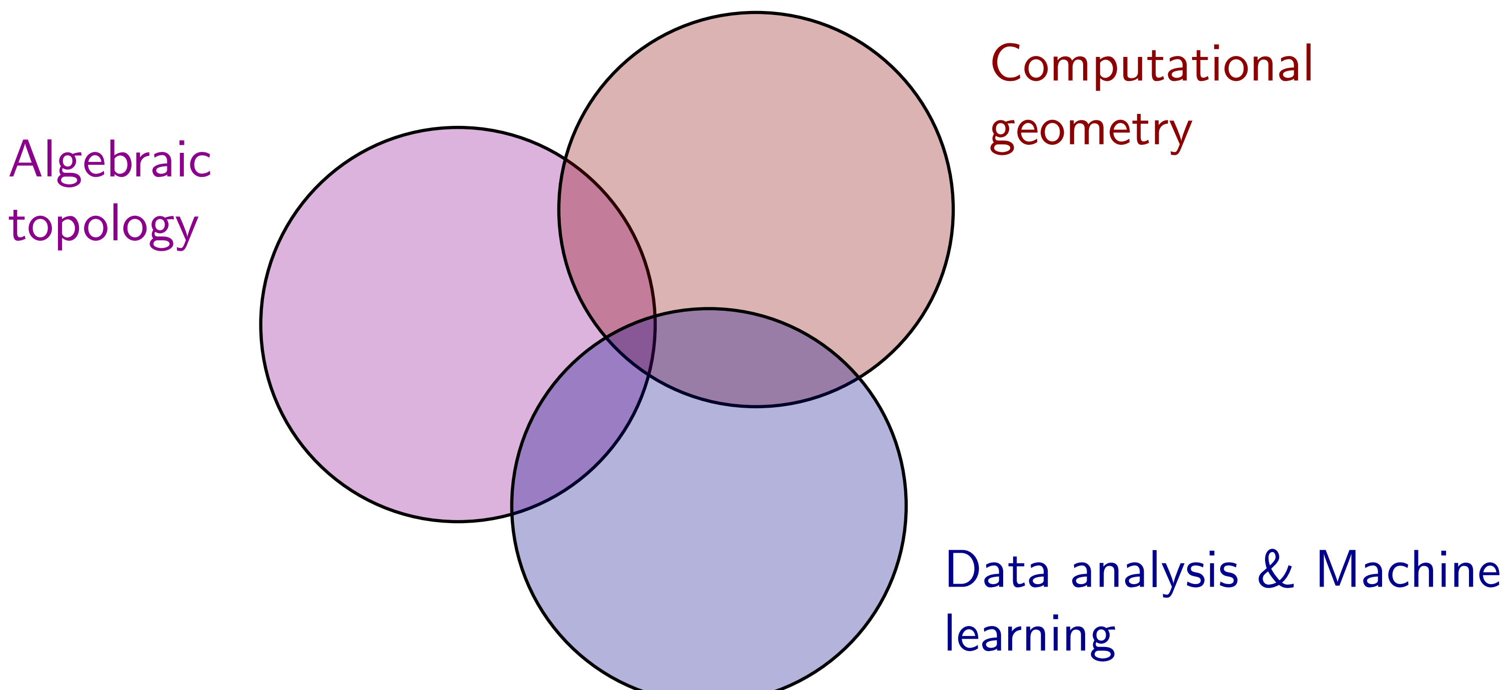


Discovery of a new type of breast cancer (*c-MYB*$^+$) with a 100% survival rate and no metastases

Qual é a forma dos dados?



Topological Data Analysis (TDA) allows to **explore** and **understand** the topology of datasets.



Part I/IV: Topological invariants

Tuesday 26th

Part II/IV: Homology

Thursday 28th

Part III/IV: Persistent Homology

Tuesday 3rd

Part IV/IV: Python tutorial

Thursday 5th

Algumas figuras históricas

Euler

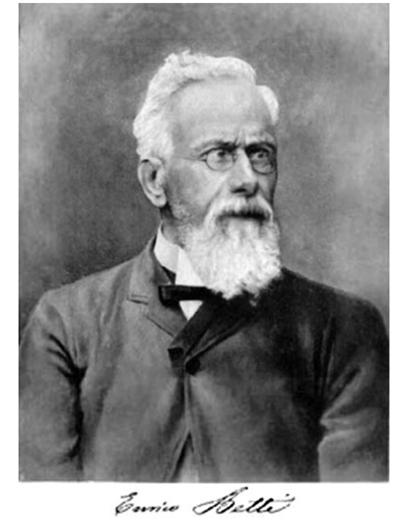
1736: Solutio problematis ad geometriarn situs pertinentis

Möbius

1865: Bestimmung des Inhaltes eines Polyëders

Riemann

1857: Theorie der Abel'schen Functionen



Betti

1871: Sopra gli spazi di un numero qualunque di dimensioni

Algumas figuras históricas

Euler

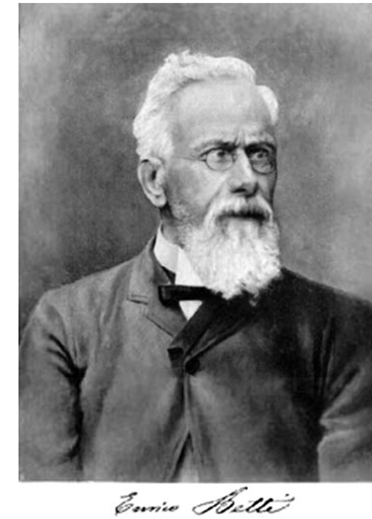
1736: Solutio problematis ad geometriarn situs pertinentis

Möbius

1865: Bestimmung des Inhaltes eines Polyëders

Riemann

1857: Theorie der Abel'schen Functionen



Betti

1871: Sopra gli spazi di un numero qualunque di dimensioni

Poincaré

1895: Analysis Situs

Toute 3-variété compacte sans bord et simplement connexe est-elle homéomorphe à la 3-sphère ?

[Cinquième complément à l'analysis situs, 1904]

Algumas figuras históricas

Euler

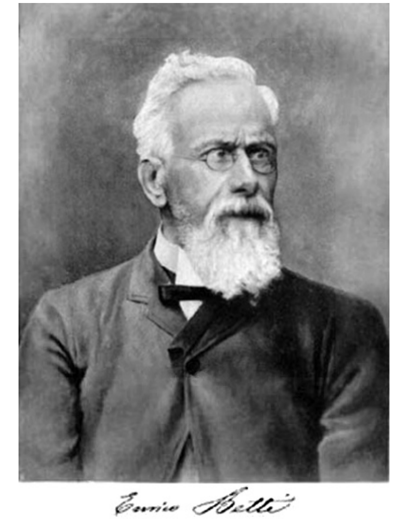
1736: Solutio problematis ad geometriarn situs pertinentis

Möbius

1865: Bestimmung des Inhaltes eines Polyëders

Riemann

1857: Theorie der Abel'schen Functionen



Betti

1871: Sopra gli spazi di un numero qualunque di dimensioni

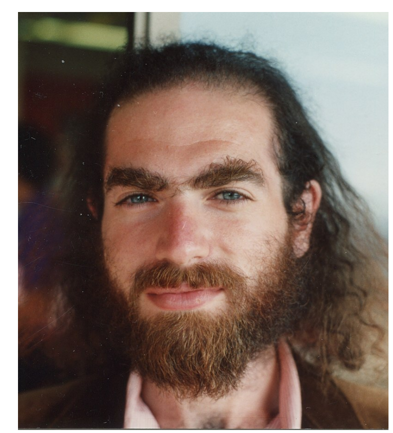
Poincaré

1895: Analysis Situs

Toute 3-variété compacte sans bord et simplement connexe est-elle homéomorphe à la 3-sphère ?

[Cinquième complément à l'analysis situs, 1904]

уи



Perelman

2002: The entropy formula for the Ricci flow and its geometric applications

Protagonistas da topologia

In topology we study **topological spaces**.

Definition: a topological space is a set X endowed with a collection of **open sets** $\{O_\alpha \mid \alpha \in A\}$, with $O_\alpha \subset X$, such that

- $\emptyset$ and X are open sets,
- an infinite union of open sets is an open set,
- a finite intersection of open sets is an open set.

Definition: Given two topological spaces X and Y, a map $f\colon X \to Y$ is **continuous** if for every open set $O \subset Y$, the preimage $f^{-1}(O)$ is an open set of X.

$$X \longrightarrow Y$$

Protagonistas da topologia

In topology we study **topological spaces**.

Definition: a topological space is a set X endowed with a collection of **open sets** $\{O_\alpha \mid \alpha \in A\}$, with $O_\alpha \subset X$, such that

- $\emptyset$ and X are open sets,
- an infinite union of open sets is an open set,
- a finite intersection of open sets is an open set.

Definition: Given two topological spaces X and Y, a map $f\colon X \to Y$ is **continuous** if for every open set $O \subset Y$, the preimage $f^{-1}(O)$ is an open set of X.

$$X \longrightarrow Y$$

translation in
ϵ-δ calculus

One can think of **subsets** $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$,

and maps $f\colon X \to Y$ **continuous** in the following sense:

$$\forall x \in X, \forall \epsilon > 0, \exists \eta > 0, \forall y \in X, \|x - y\| < \eta \implies \|f(x) - f(y)\| < \epsilon.$$

1 - We can build a topological space by seeing it as a **subspace** of another one.

In $\mathbb{R}^n$, we can define:

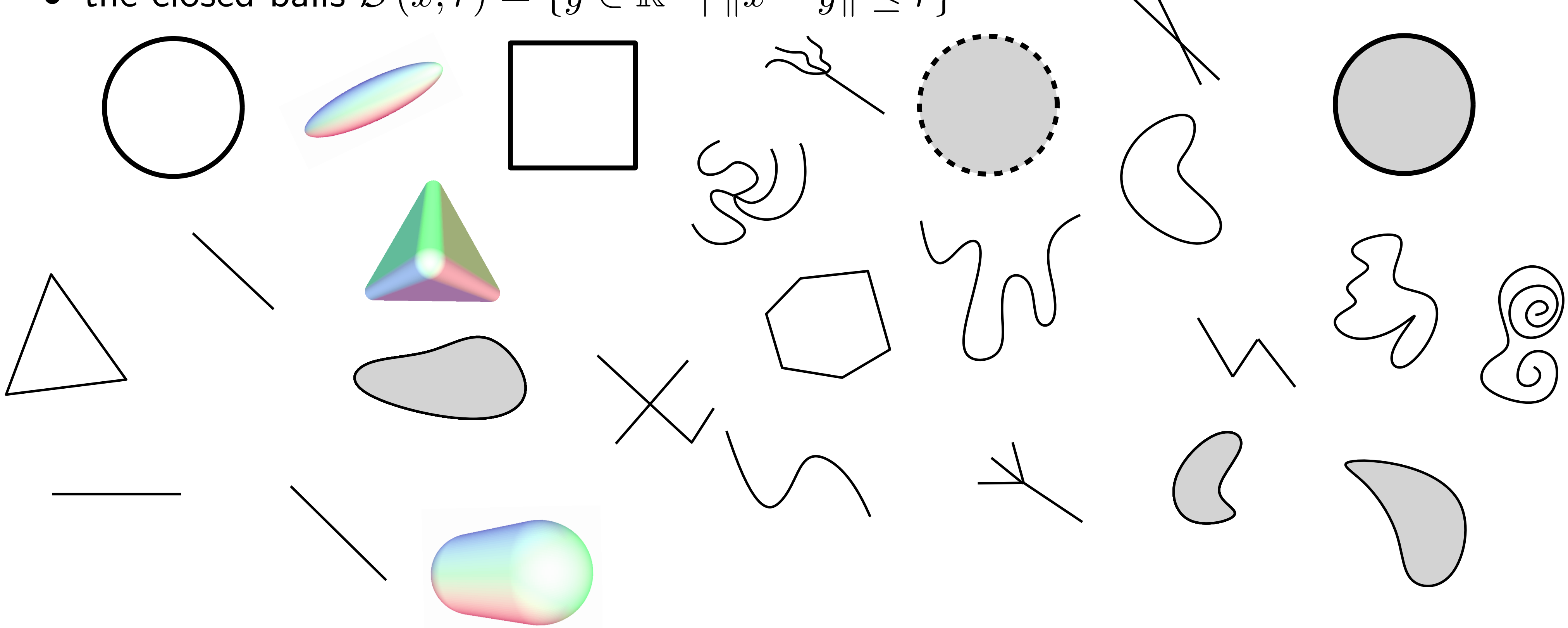
- the unit sphere $\mathbb{S}_{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$
- the unit cube $\mathcal{C}_{n-1} = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid \max(|x_1|, ..., |x_n|) = 1\}$
- the open balls $\mathcal{B}(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$
- the closed balls $\overline{\mathcal{B}}(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| \leq r\}$

Construção de espaços topológicos

1 - We can build a topological space by seeing it as a **subspace** of another one.

In $\mathbb{R}^n$, we can define:

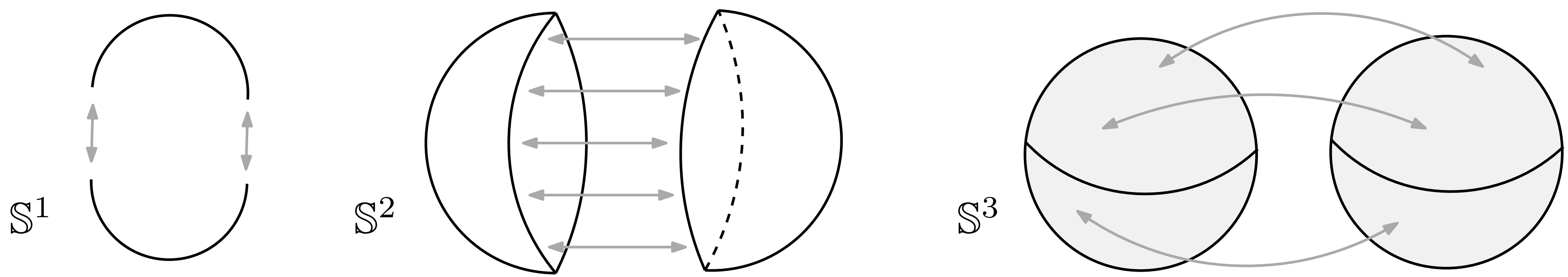
- the unit sphere $\mathbb{S}_{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$
- the unit cube $\mathcal{C}_{n-1} = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid \max(|x_1|, ..., |x_n|) = 1\}$
- the open balls $\mathcal{B}(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$
- the closed balls $\overline{\mathcal{B}}(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| \leq r\}$



Most of the time, we do not have a nice algebraic definition...

1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

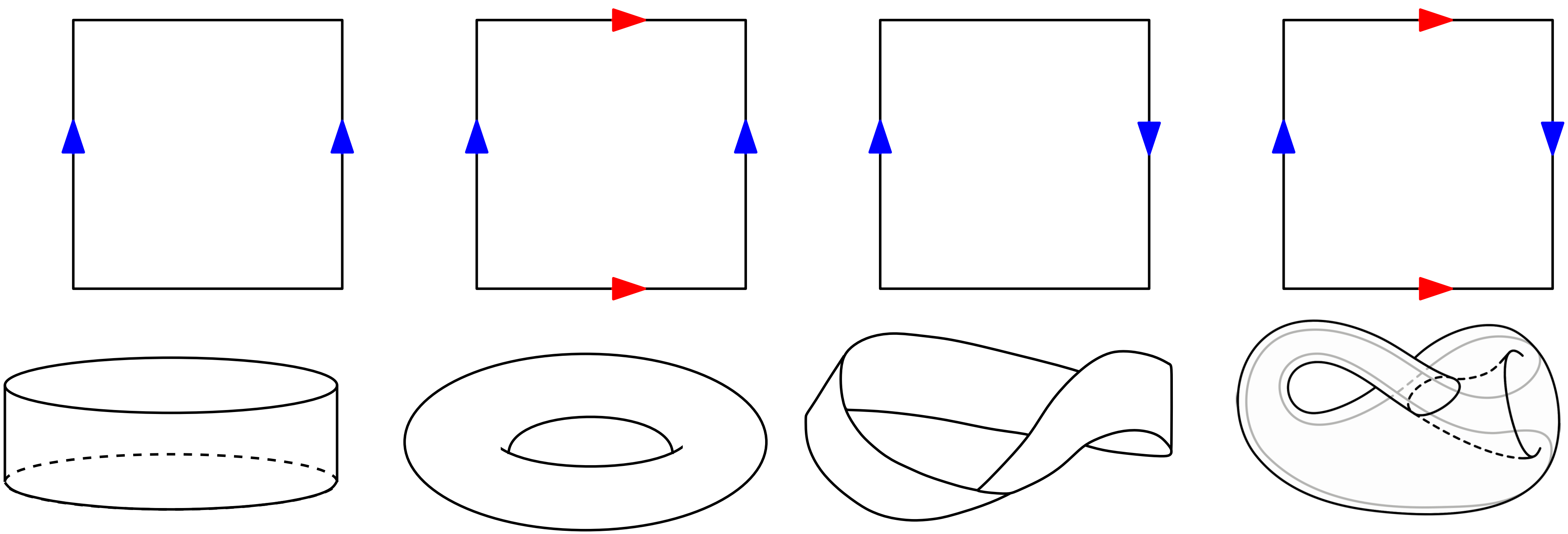


Construção de espaços topológicos

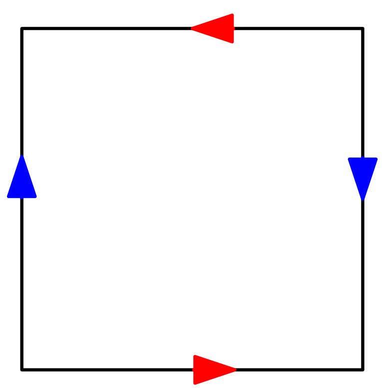
1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

3 - We can build a topological space by **quotienting** another one.



Bonus:

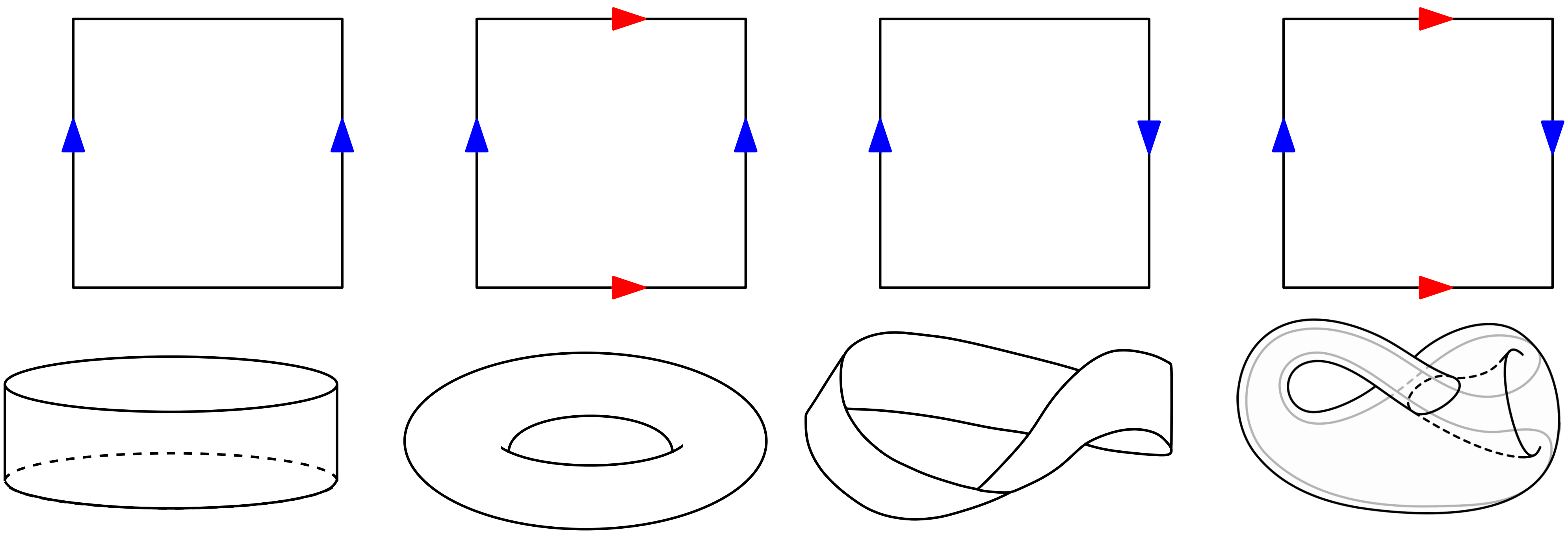


Construção de espaços topológicos

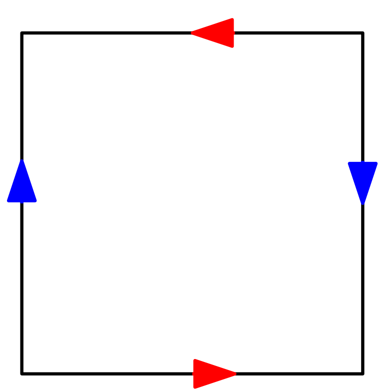
1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

3 - We can build a topological space by **quotienting** another one.



Bonus:



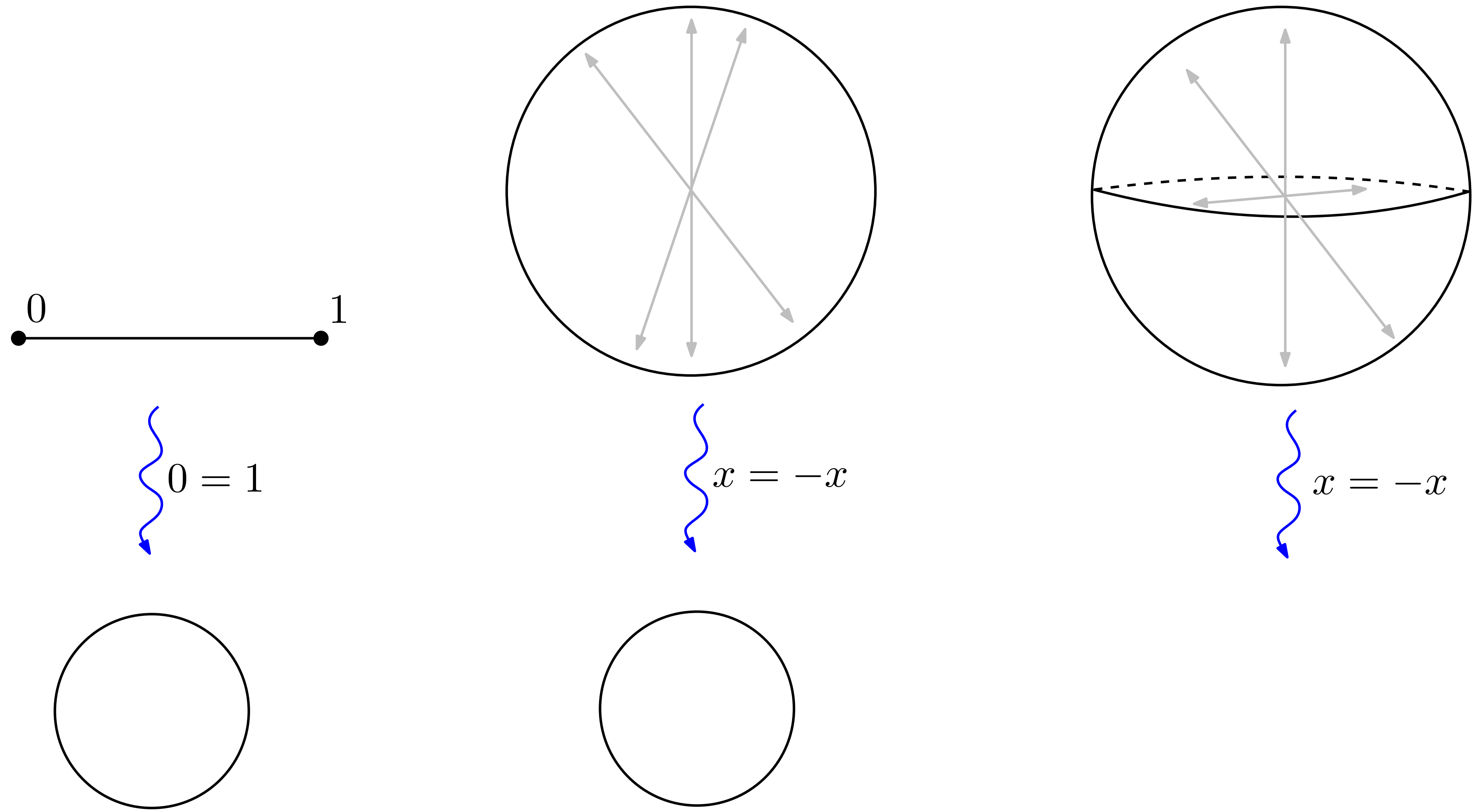
Projective plane $\mathbb{R}P^2$

Construção de espaços topológicos

1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

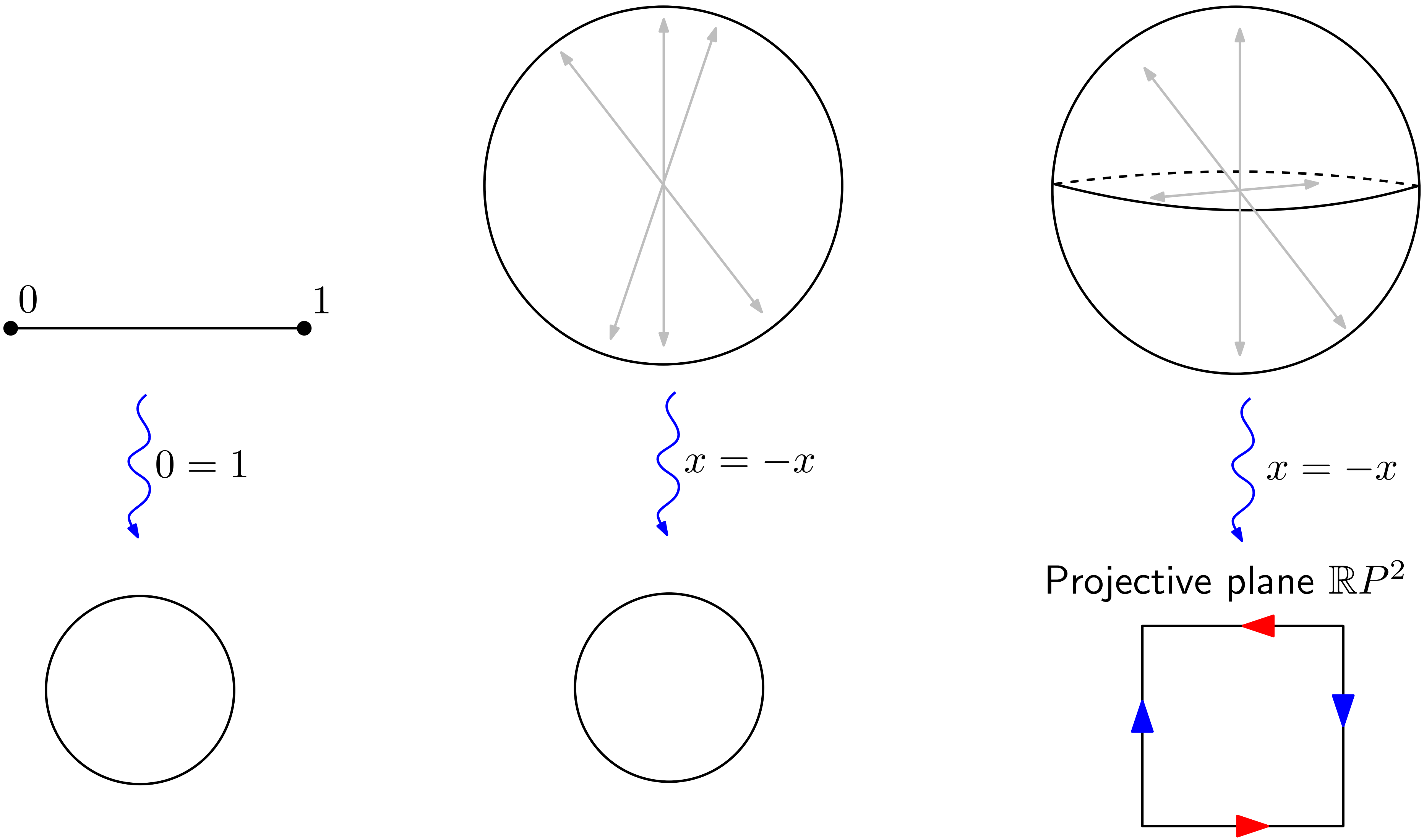
3 - We can build a topological space by **quotienting** another one.



1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

3 - We can build a topological space by **quotienting** another one.

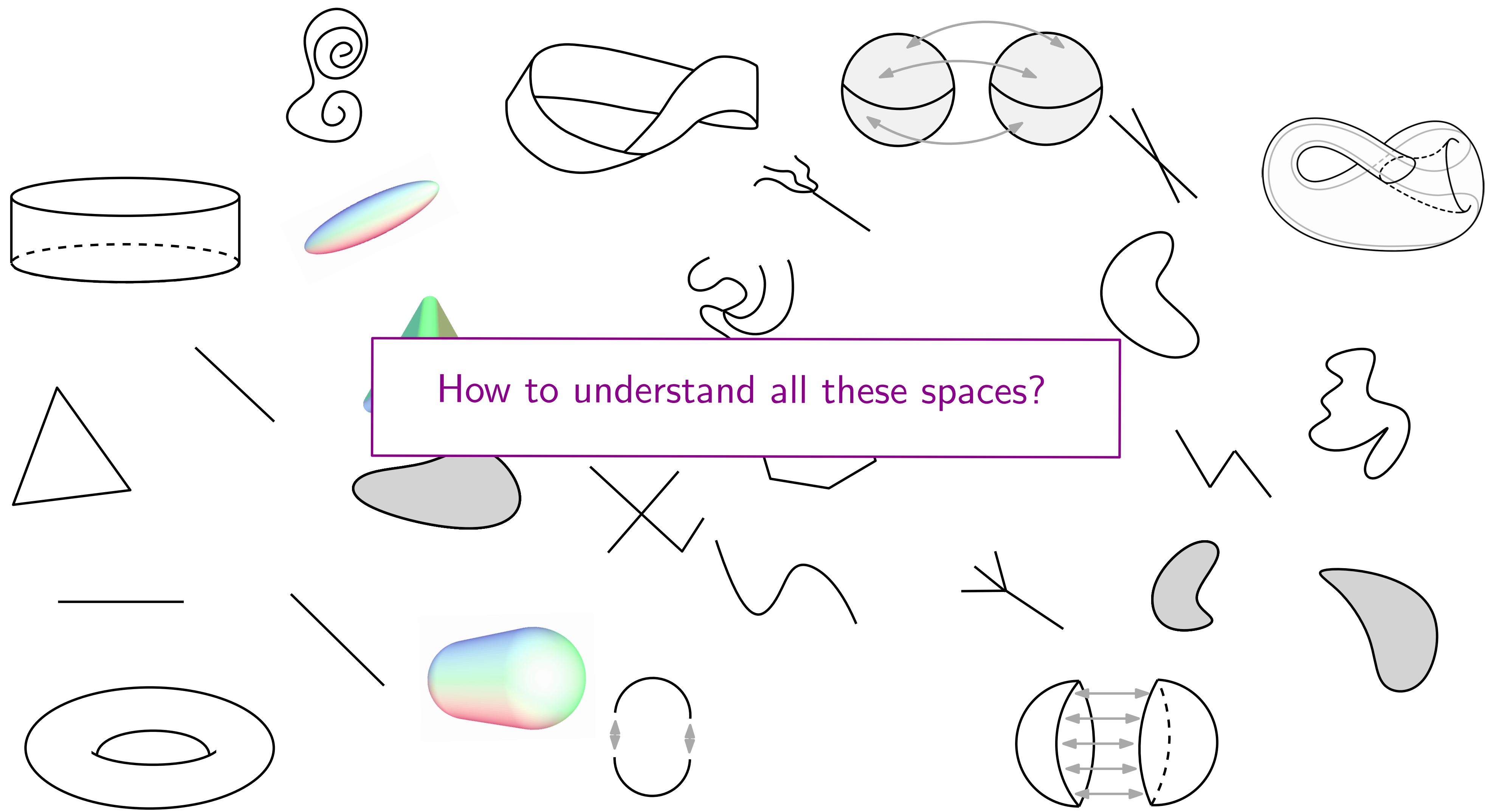


Construção de espaços topológicos

1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

3 - We can build a topological space by **quotienting** another one.

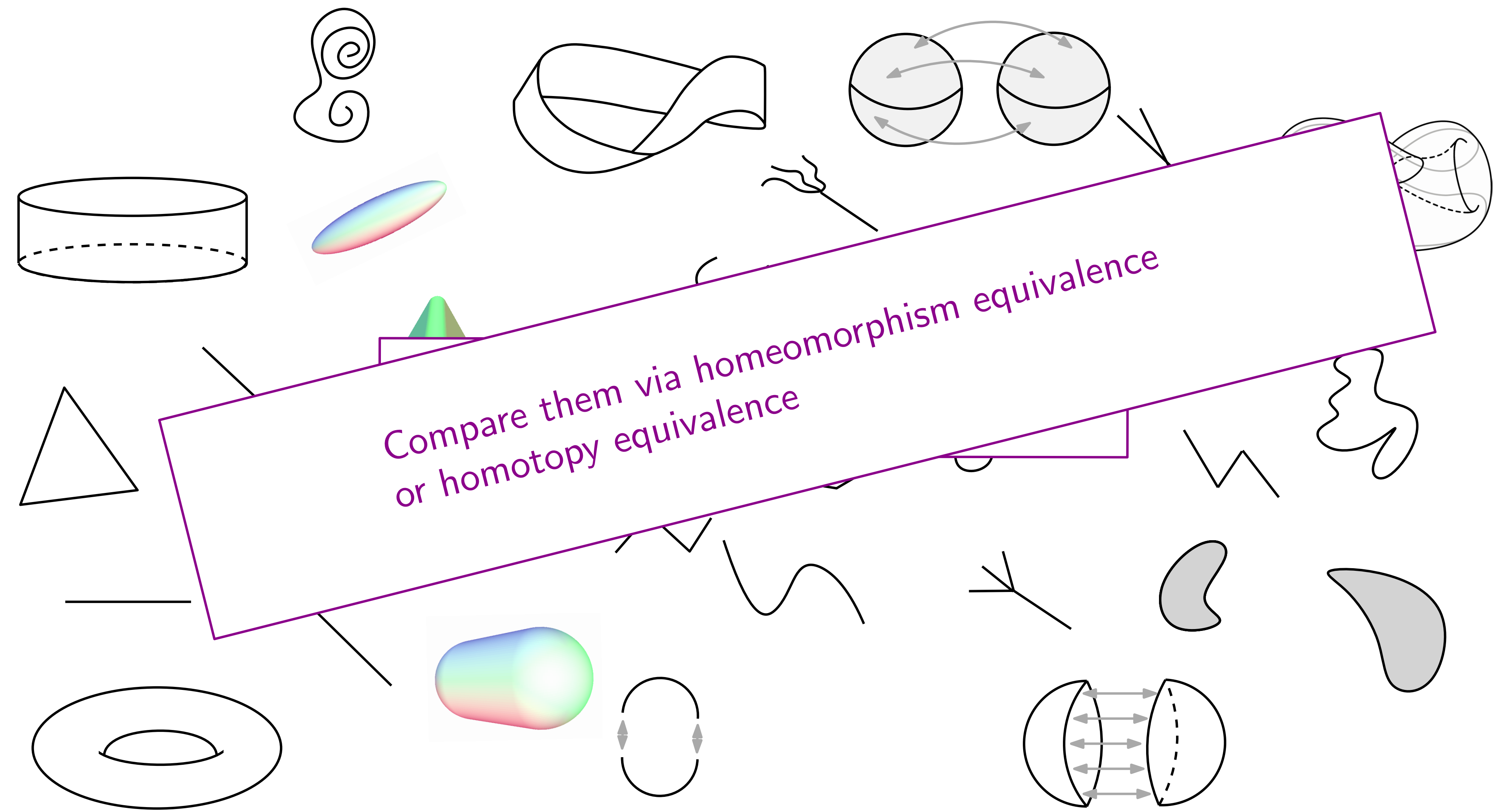


How to understand all these spaces?

1 - We can build a topological space by seeing it as a **subspace** of another one.

2 - We can build a topological space by **gluing** the boundaries of another one.

3 - We can build a topological space by **quotienting** another one.



Homeomorfismos

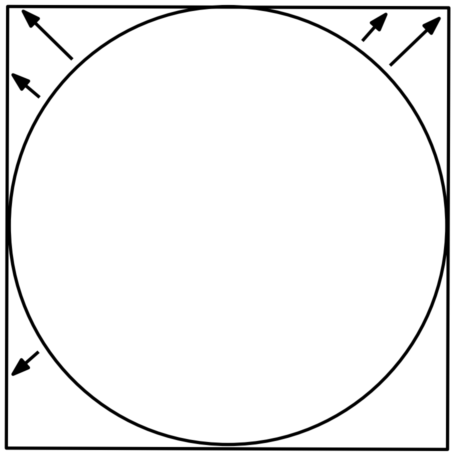
Definition: Let X and Y be two topological spaces, and $f\colon X \to Y$ a map.
We say that f is a **homeomorphism** if

- f is a bijection,
- $f\colon X \to Y$ is continuous,
- $f^{-1}\colon Y \to X$ is continuous.

If there exist such a homeomorphism, we say that the two topological spaces are **homeomorphic**.

Example: The unit circle and the unit square are homeomorphic via

$$\begin{aligned} f\colon \mathbb{S}_1 &\longrightarrow \mathcal{C} \\ (x_1, x_2) &\longmapsto \frac{1}{\max(|x_1|, |x_2|)}(x_1, x_2) \end{aligned}$$



Interpretation: Homeomorphisms allow 'continuous deformations'.

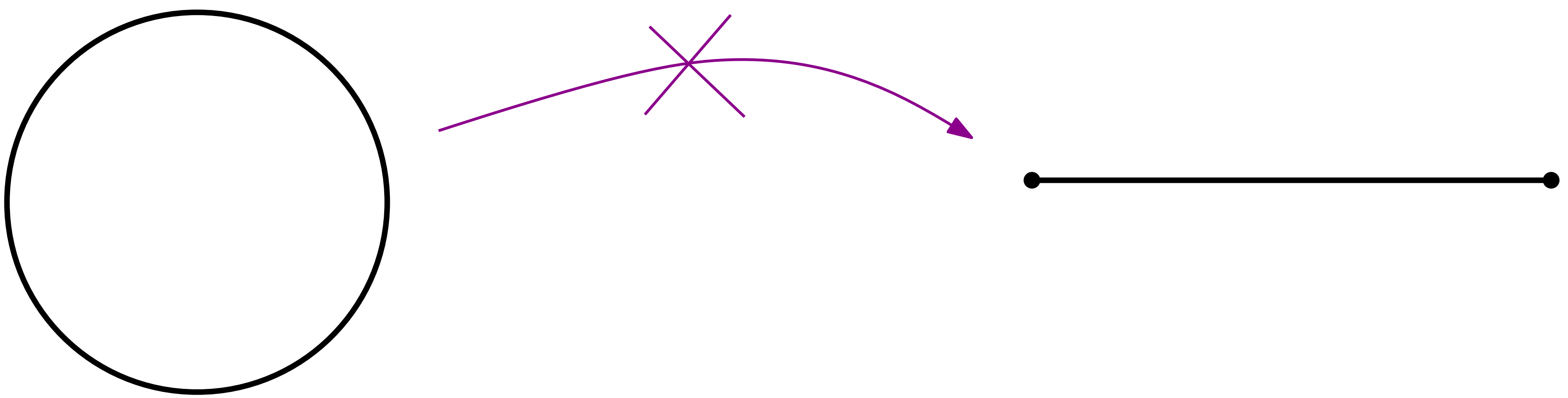
Homeomorfismos

Definition: Let X and Y be two topological spaces, and $f\colon X \to Y$ a map.
We say that f is a **homeomorphism** if

- f is a bijection,
- $f\colon X \to Y$ is continuous,
- $f^{-1}\colon Y \to X$ is continuous.

If there exist such a homeomorphism, we say that the two topological spaces are **homeomorphic**.

Example: The unit circle and the interval $[0, 1]$ are not homeomorphic.



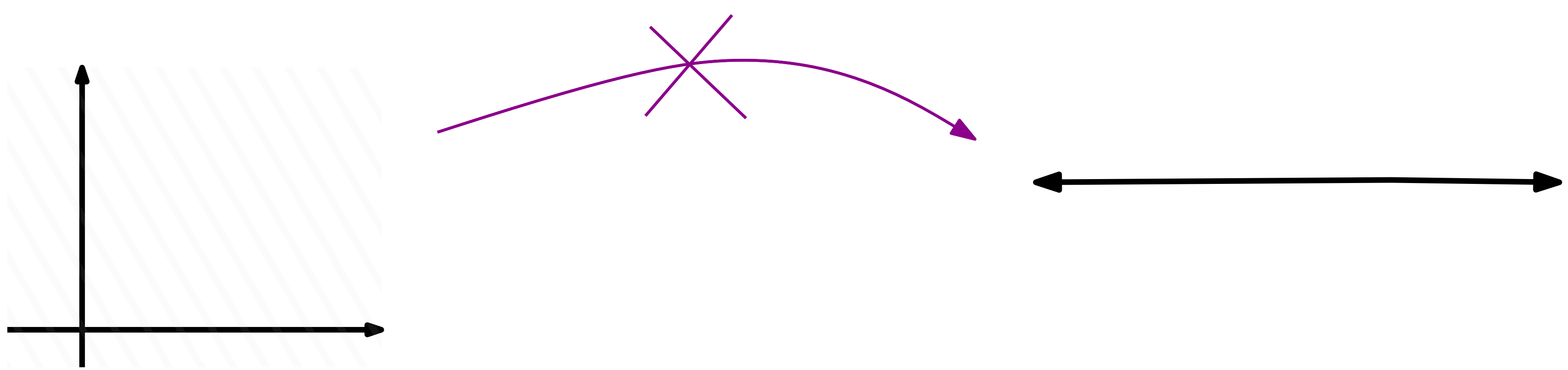
Interpretation: Homeomorphisms allow 'continuous deformations'.

Definition: Let X and Y be two topological spaces, and $f\colon X \to Y$ a map.
We say that f is a **homeomorphism** if

- f is a bijection,
- $f\colon X \to Y$ is continuous,
- $f^{-1}\colon Y \to X$ is continuous.

If there exist such a homeomorphism, we say that the two topological spaces are **homeomorphic**.

Example (Invariance of domain): [Brouwer, 1912] If $n \neq m$, the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ are not homeomorphic.



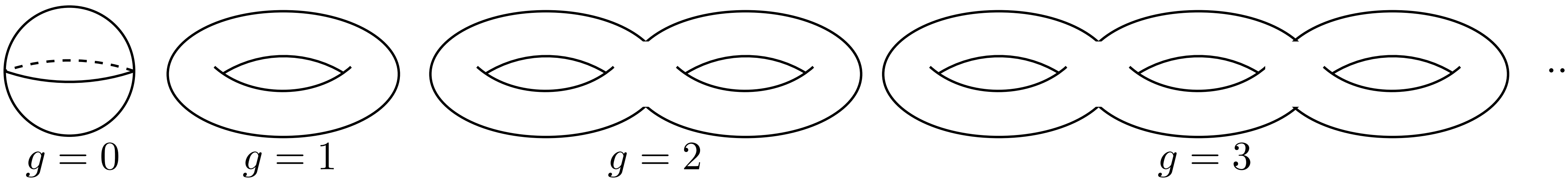
Interpretation: Homeomorphisms allow 'continuous deformations'.

Definition: Let X and Y be two topological spaces, and $f\colon X \to Y$ a map.
We say that f is a **homeomorphism** if

- f is a bijection,
- $f\colon X \to Y$ is continuous,
- $f^{-1}\colon Y \to X$ is continuous.

If there exist such a homeomorphism, we say that the two topological spaces are **homeomorphic**.

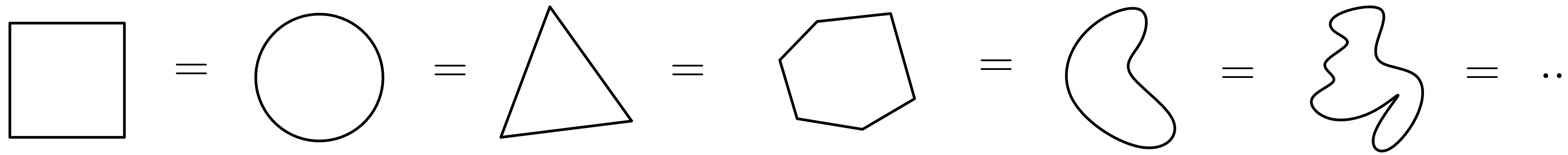
Example (Classification of surfaces): [Möbius, Jordan, von Dyck, Dehn and Heegaard, Alexander, Brahana, 1863-1921] If $g \neq g'$, the surfaces of genus g and g' are not homeomorphic.



Interpretation: Homeomorphisms allow 'continuous deformations'.

Classes de homeomorfismo

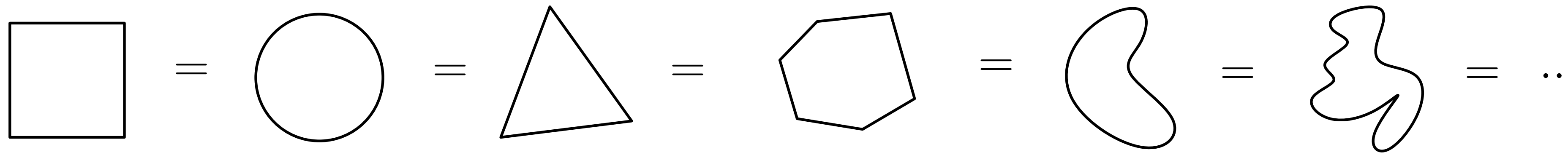
We can gather topological spaces that are homeomorphic



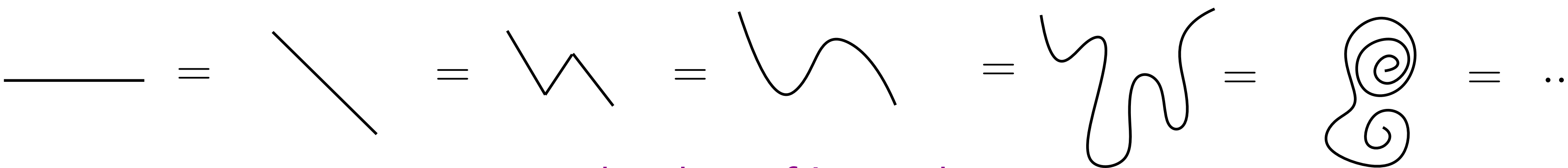
the class of circles

Classes de homeomorfismo

We can gather topological spaces that are homeomorphic



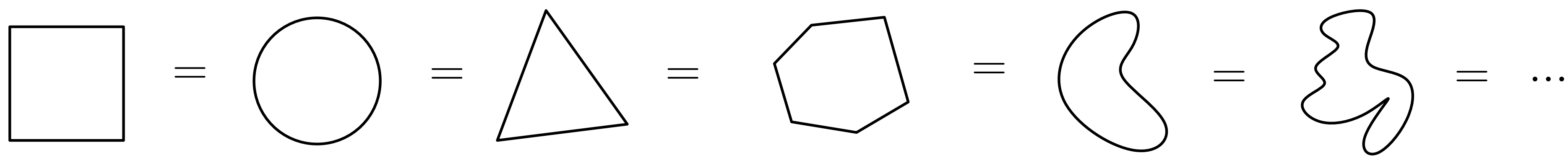
the class of circles



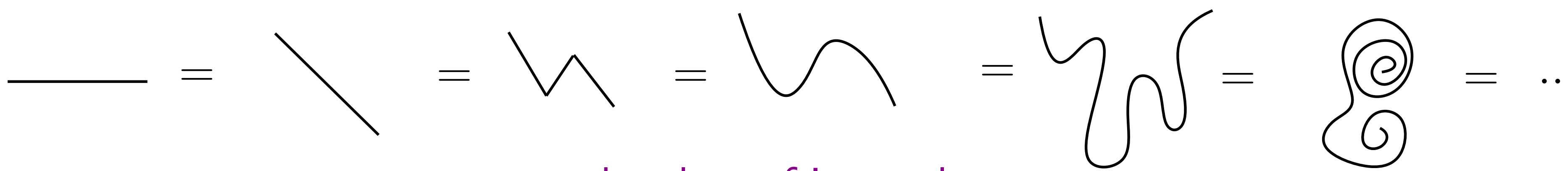
the class of intervals

Classes de homeomorfismo

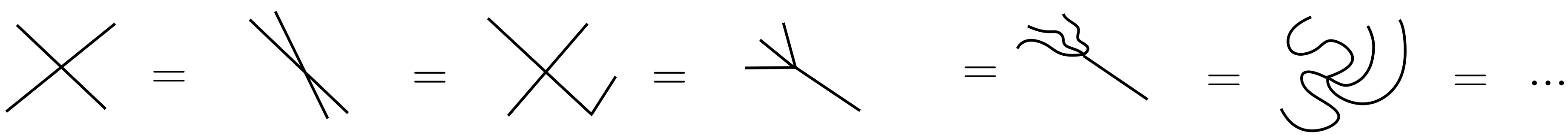
We can gather topological spaces that are homeomorphic



the class of circles



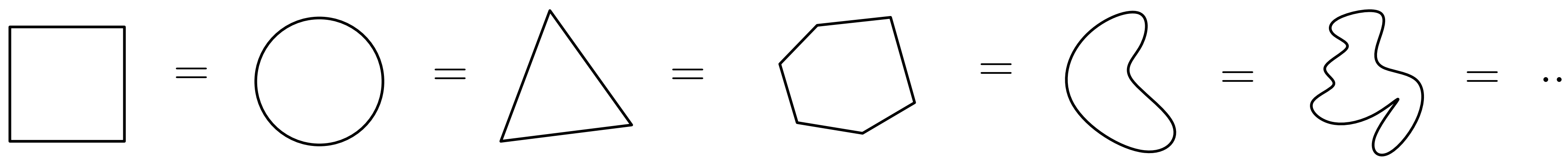
the class of intervals



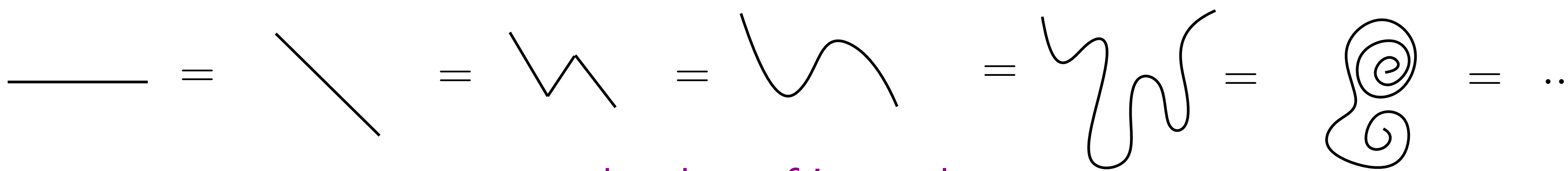
the class of crosses

Classes de homeomorfismo

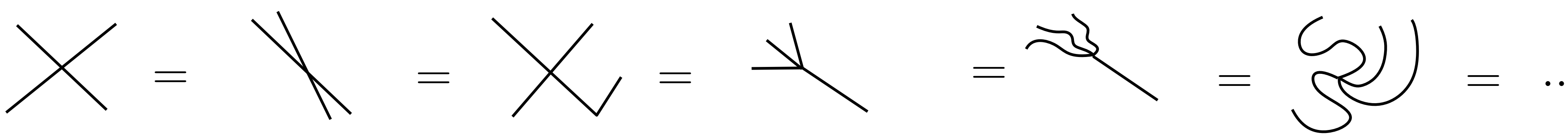
We can gather topological spaces that are homeomorphic



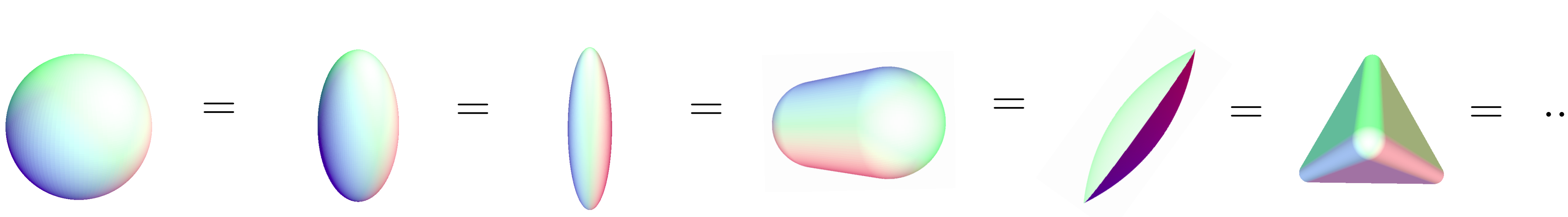
the class of circles



the class of intervals



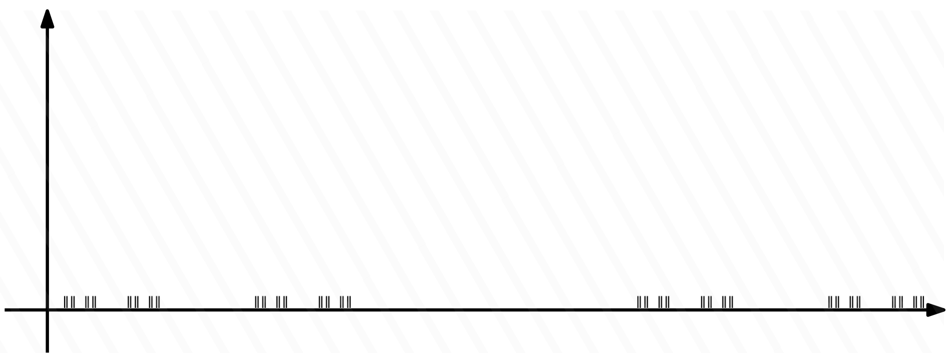
the class of crosses



the class of spheres

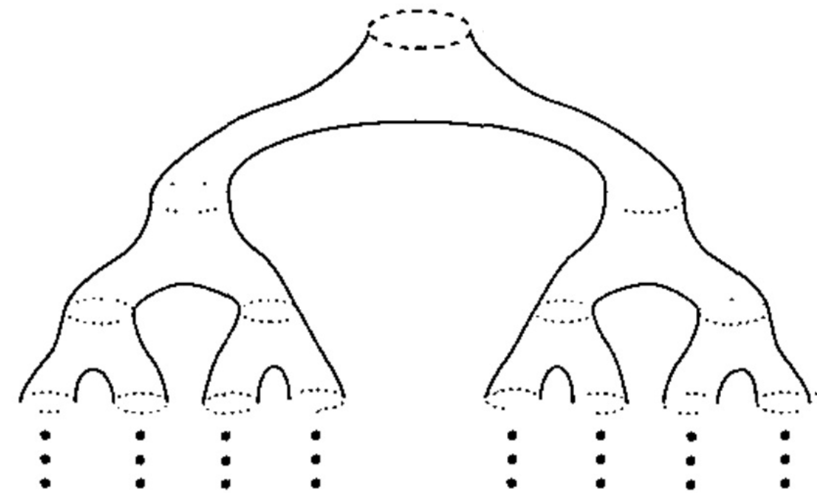
Classes de homeomorfismo

In general, it may be complicated to determine whether two spaces are homeomorphic.



$\mathbb{R}^2\setminus$ Cantor set

$$\stackrel{?}{=}$$



To answer this problem, we will use the notion of **invariant**.

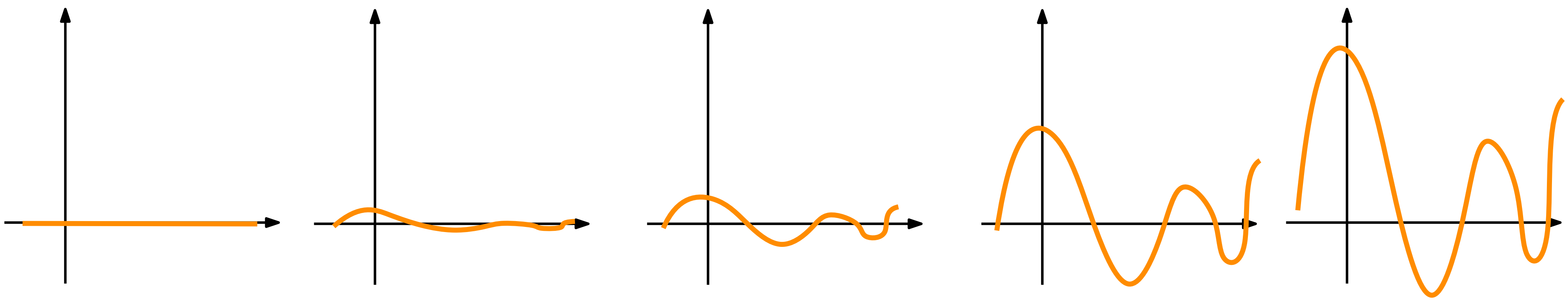
Homotopias

Definition: Let X, Y be two topological spaces, and $f, g\colon X \to Y$ two continuous maps. A **homotopy** between f and g is a map $F\colon X \times [0,1] \to Y$ such that:

- $x \mapsto F(x,0)$ is equal to f,
- $x \mapsto F(x,1)$ is equal to g,
- $F\colon X \times [0,1] \to Y$ is continuous.

If such a homotopy exists, we say that the maps f and g are **homotopic**.

Example: Homotopy between $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$.



$F(\cdot,0) = f$ $\quad$ $F(\cdot,0.2)$ $\quad$ $F(\cdot,0.5)$ $\quad$ $F(\cdot,0.6)$ $\quad$ $F(\cdot,1) = g$

Homotopias

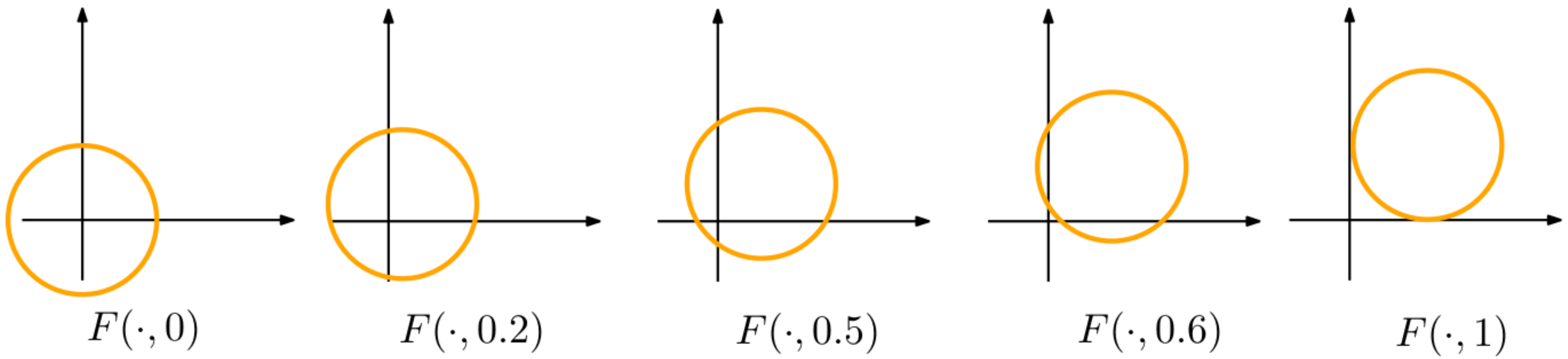
Definition: Let X, Y be two topological spaces, and $f, g\colon X \to Y$ two continuous maps. A **homotopy** between f and g is a map $F\colon X \times [0,1] \to Y$ such that:

- $x \mapsto F(x,0)$ is equal to f,
- $x \mapsto F(x,1)$ is equal to g,
- $F\colon X \times [0,1] \to Y$ is continuous.

If such a homotopy exists, we say that the maps f and g are **homotopic**.

Example: The map $F\colon (x,t) \in \mathbb{S}^1 \times [0,1] \mapsto (\cos(\theta) + 2t, \sin(\theta) + 2t)$ is a homotopy between

$$\begin{aligned} f\colon \mathbb{S}^1 &\to \mathbb{R}^2 \\ \theta &\mapsto (\cos(\theta), \sin(\theta)) \end{aligned} \quad \text{and} \quad \begin{aligned} g\colon \mathbb{S}^1 &\to \mathbb{R}^2 \\ \theta &\mapsto (\cos(\theta)+2, \sin(\theta)+2) \end{aligned}$$



Homotopias

Definition: Let X, Y be two topological spaces, and $f, g\colon X \to Y$ two continuous maps. A **homotopy** between f and g is a map $F\colon X \times [0,1] \to Y$ such that:

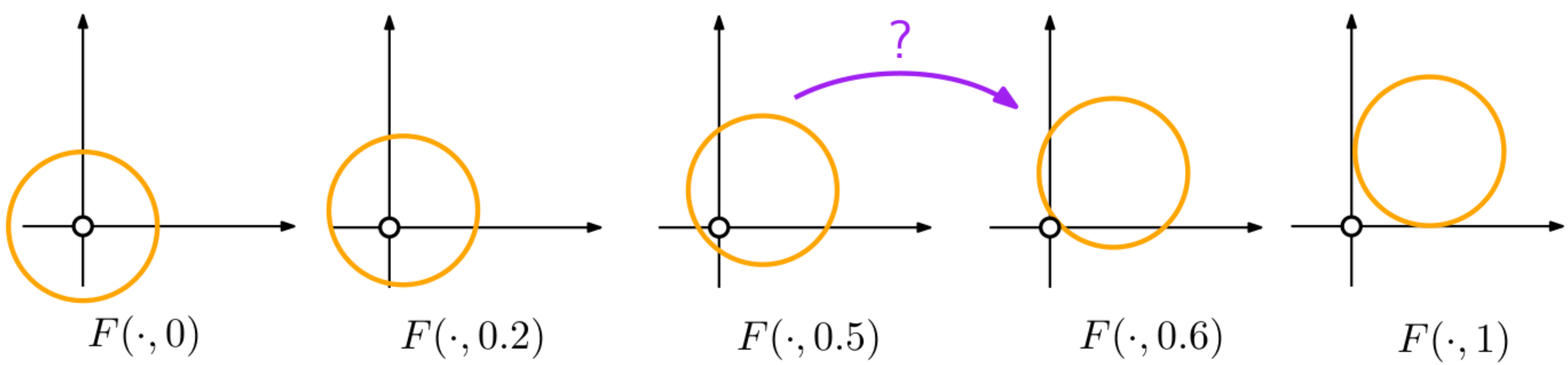
- $x \mapsto F(x,0)$ is equal to f,
- $x \mapsto F(x,1)$ is equal to g,
- $F\colon X \times [0,1] \to Y$ is continuous.

If such a homotopy exists, we say that the maps f and g are **homotopic**.

not well-defined

Example: The map $F\colon (x,t) \in \mathbb{S}^1 \times [0,1] \mapsto (\cos(\theta)+2t, \sin(\theta)+2t)$ ~~is a homotopy~~ between

$$\begin{aligned} f\colon \mathbb{S}^1 &\to \mathbb{R}^2 \setminus \{0\} \\ \theta &\mapsto (\cos(\theta), \sin(\theta)) \end{aligned} \quad \text{and} \quad \begin{aligned} g\colon \mathbb{S}^1 &\to \mathbb{R}^2 \setminus \{0\} \\ \theta &\mapsto (\cos(\theta)+2, \sin(\theta)+2) \end{aligned}$$

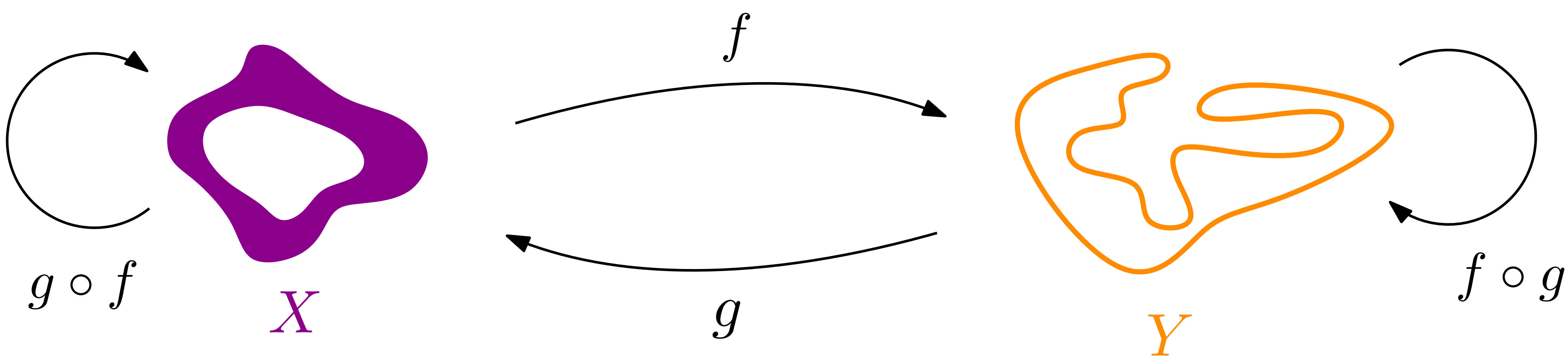


This is not true anymore if we remove the origin from the plane.

Defintion: Let X and Y be two topological spaces. A **homotopy equivalence** between X and Y is a pair of continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that:

- $g \circ f\colon X \to X$ is homotopic to the identity map $\mathrm{id}\colon X \to X$,
- $f \circ g\colon Y \to Y$ is homotopic to the identity map $\mathrm{id}\colon Y \to Y$.

If such a homotopy equivalence exists, we say that X and Y are **homotopy equivalent**.

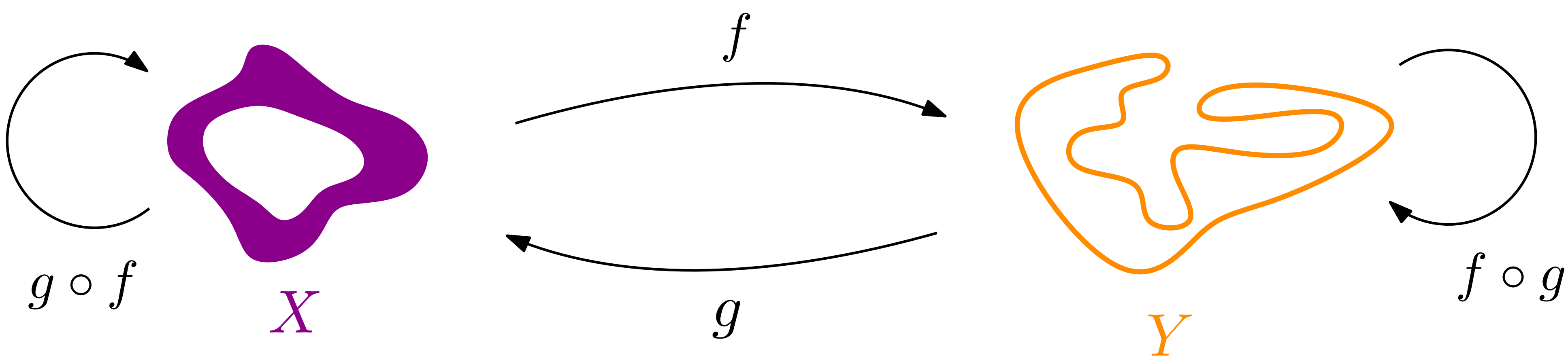


Equivalência de homotopia

Defintion: Let X and Y be two topological spaces. A **homotopy equivalence** between X and Y is a pair of continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that:

- $g \circ f\colon X \to X$ is homotopic to the identity map $\mathrm{id}\colon X \to X$,
- $f \circ g\colon Y \to Y$ is homotopic to the identity map $\mathrm{id}\colon Y \to Y$.

If such a homotopy equivalence exists, we say that X and Y are **homotopy equivalent**.



Observation: A homotopy equivalence is a weaker formulation of homeomorphism.

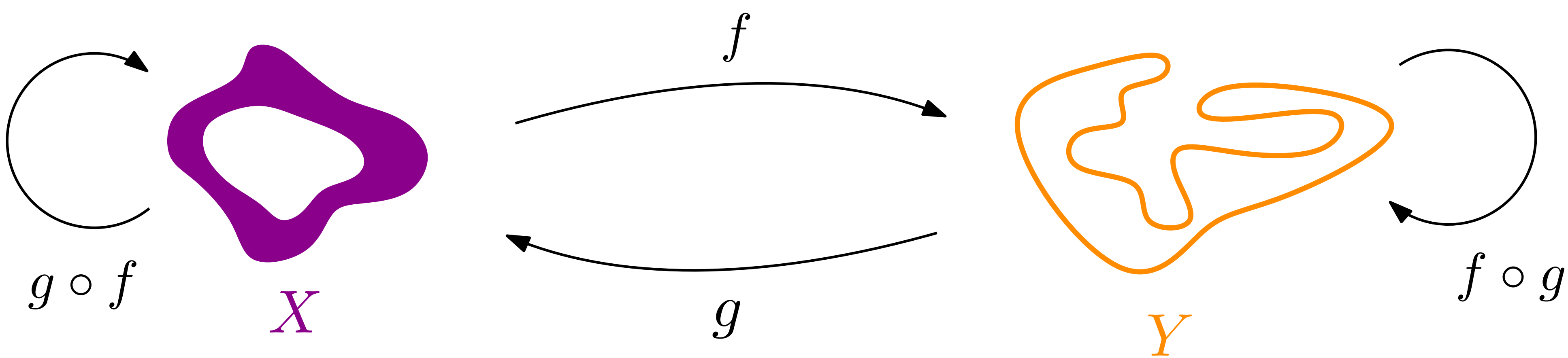
$$g \circ f = \mathrm{id}_X \qquad X \underset{g = f^{-1}}{\overset{f}{\rightleftarrows}} Y \qquad f \circ g = \mathrm{id}_Y$$

Equivalência de homotopia

Defintion: Let X and Y be two topological spaces. A **homotopy equivalence** between X and Y is a pair of continuous maps $f\colon X \to Y$ and $g\colon Y \to X$ such that:

- $g \circ f\colon X \to X$ is homotopic to the identity map $\mathrm{id}\colon X \to X$,
- $f \circ g\colon Y \to Y$ is homotopic to the identity map $\mathrm{id}\colon Y \to Y$.

If such a homotopy equivalence exists, we say that X and Y are **homotopy equivalent**.

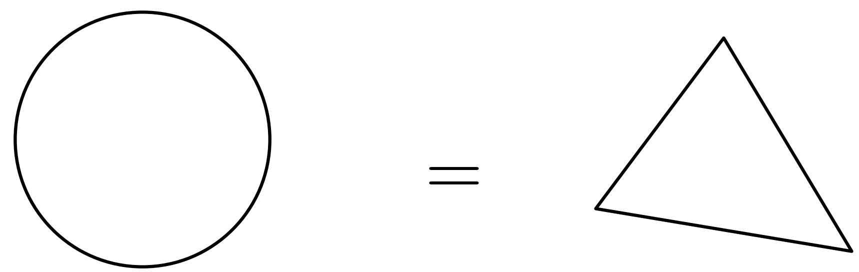


Observation: A homotopy equivalence is a weaker formulation of homeomorphism.

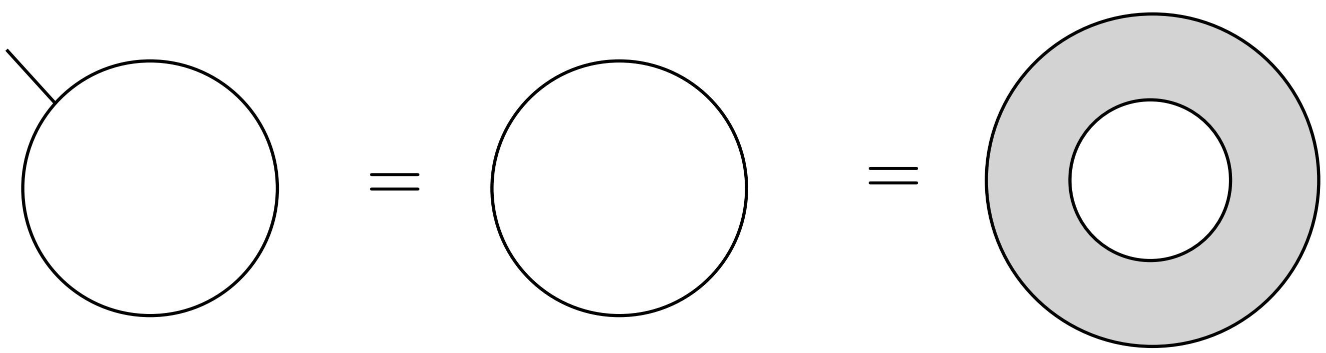
Proposition: If two topological spaces are homeomorphic, then they are homotopy equivalent.

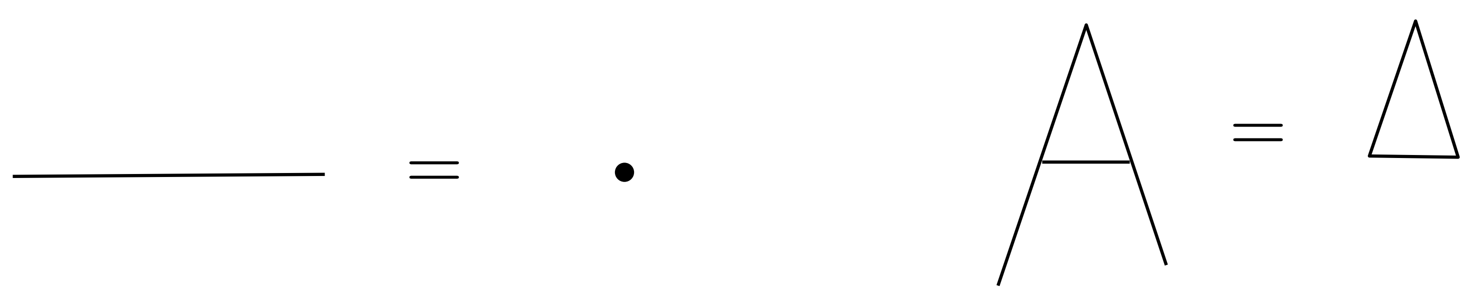
Equivalência de homotopia

Homotopy equivalence allows to continuously **deform** the space



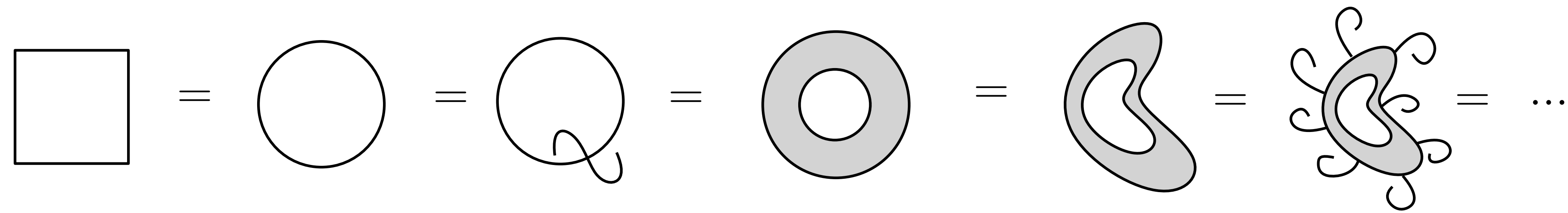
and to **retract** it.



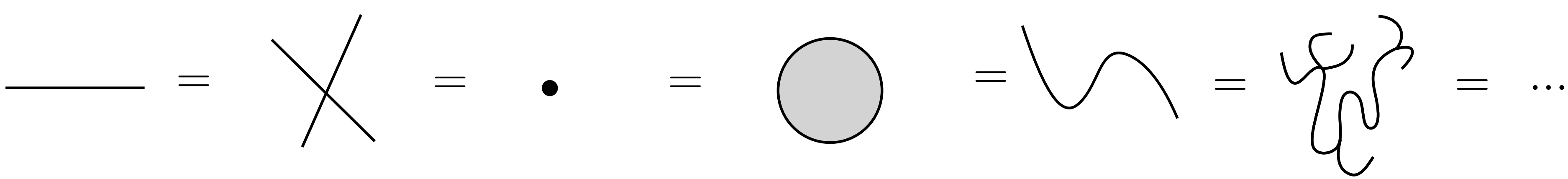


Classes de homotopia

Just as before, we can classify topological spaces according to this relation, and obtain **classes of homotopy equivalence**:



the class of circles



the class of points

the class of spheres, the class of torii, the class of Klein bottles, ...

Example: Classification, up to homotopy equivalence, of the alphabet.

A B C D E F

G H I J K L

M N O P Q R

S T U V W X

Y Z

Classes de homotopia

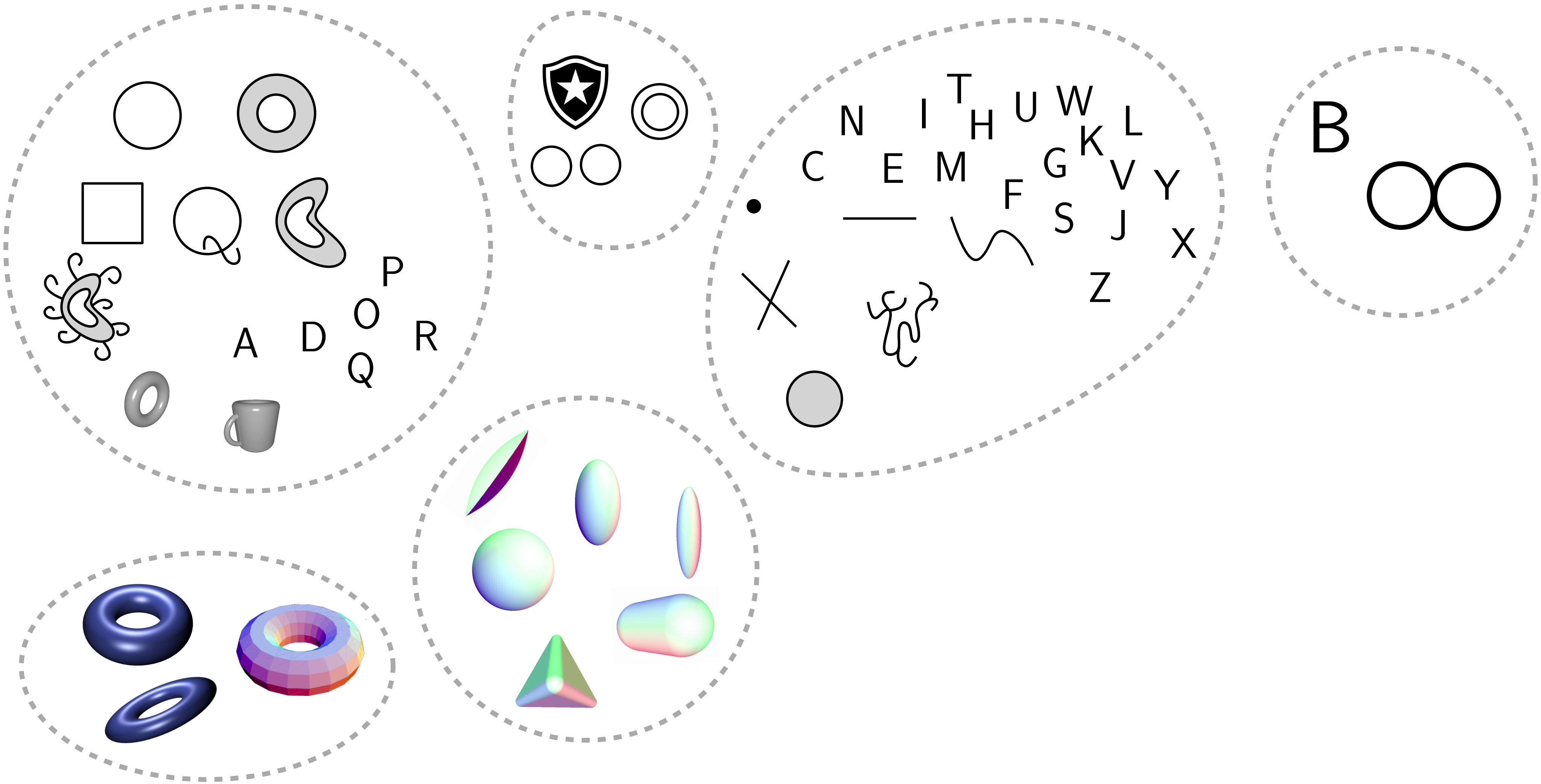
Example: Classification, up to homotopy equivalence, of the alphabet.

A D O P Q R $\approx \bigcirc$

B $\approx \bigcirc\!\!\bigcirc$

C E F G H I J K L
M N S T U V W X Y Z $\approx \bullet$

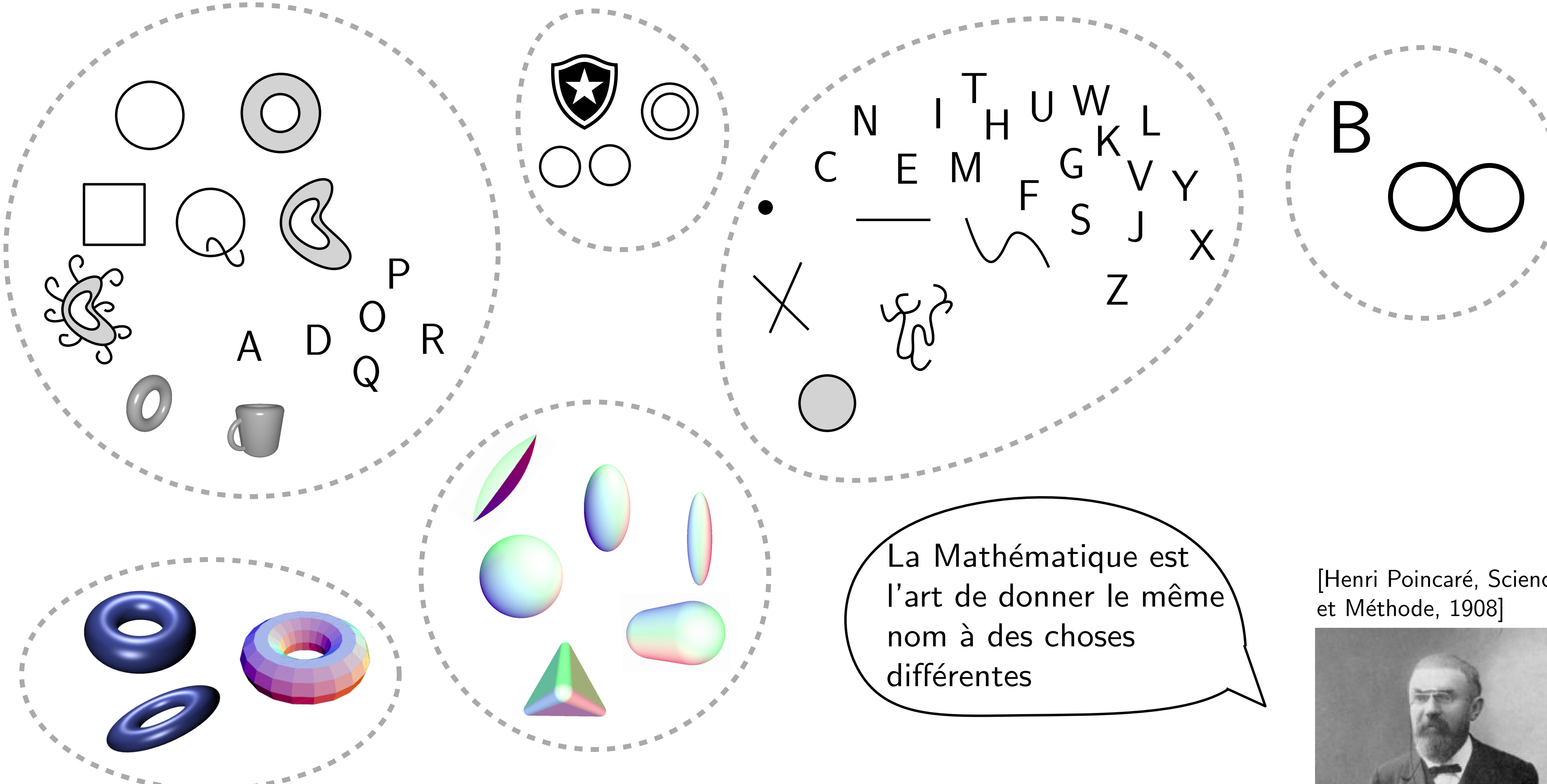
We gathered topological spaces into homotopy classes.



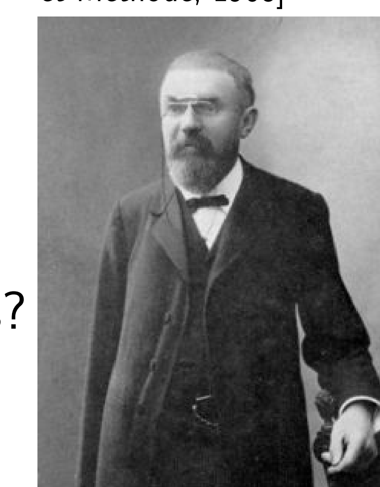
- Given a topological space X, how to recognize in which class it belongs?
- What are the common **features** of spaces in a same class?

Características comuns

We gathered topological spaces into homotopy classes.



[Henri Poincaré, Science et Méthode, 1908]



- Given a topological space X, how to recognize in which class it belongs?
- What are the common **features** of spaces in a same class?

I - Topology

- 1 - History
- 2 - Topological spaces

II - Comparing topological spaces

- 1 - Homeomorphism equivalence
- 2 - Homotopy equivalence

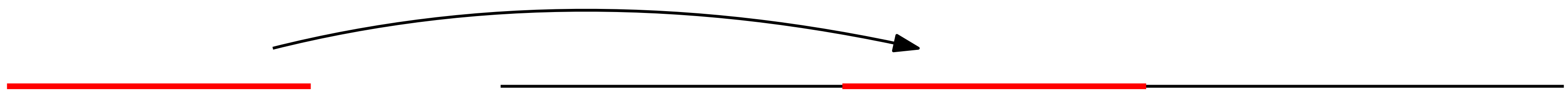
III - Topological invariants

- 1 - Embeddability
- 2 - Number of connected components
- 3 - Euler characteristic
- 4 - Betti numbers

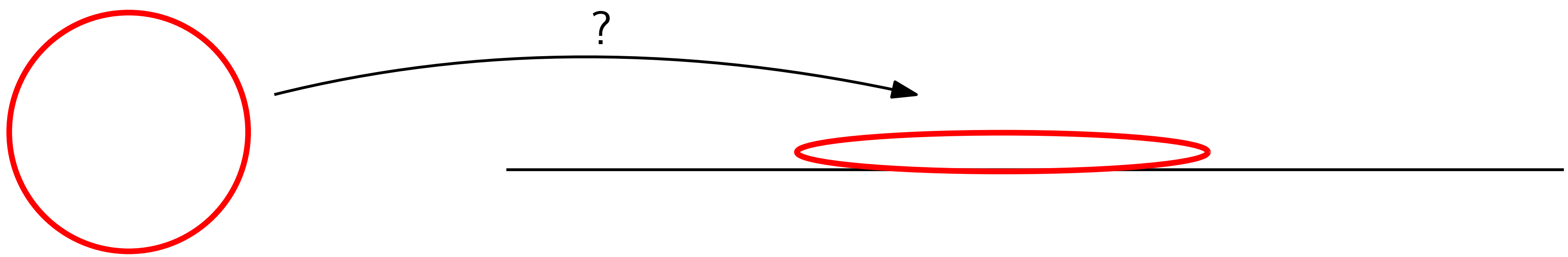
Imersibilidade

Definition: Let $n \in \mathbb{N}$. A topological space X is **embeddable** in $\mathbb{R}^n$ if there exists a continuous injective map $X \to \mathbb{R}^n$.

Example: The interval $(0,1)$ is embeddable in $\mathbb{R}$.

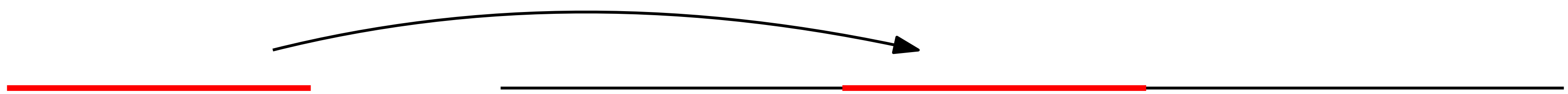


The circle $\mathbb{S}^1$ is not.

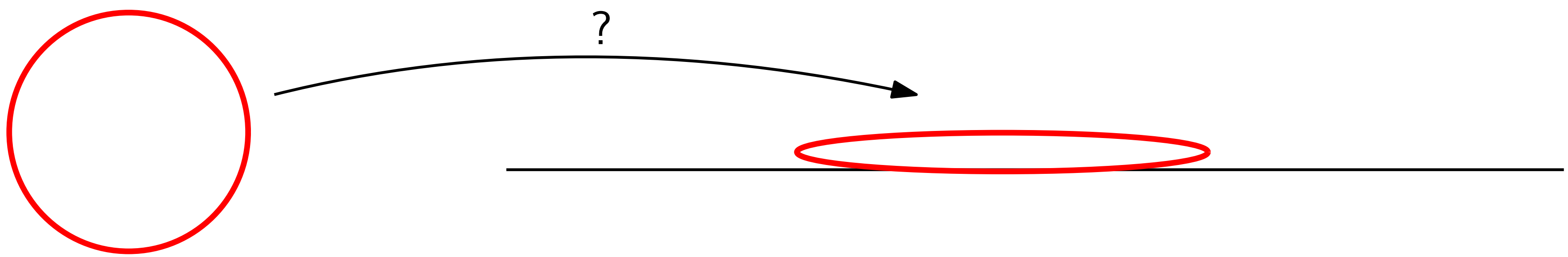


Definition: Let $n \in \mathbb{N}$. A topological space X is **embeddable** in $\mathbb{R}^n$ if there exists a continuous injective map $X \to \mathbb{R}^n$.

Example: The interval $(0,1)$ is embeddable in $\mathbb{R}$.



The circle $\mathbb{S}^1$ is not.

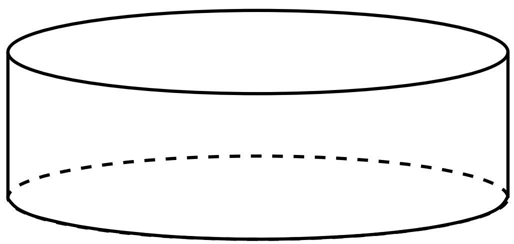


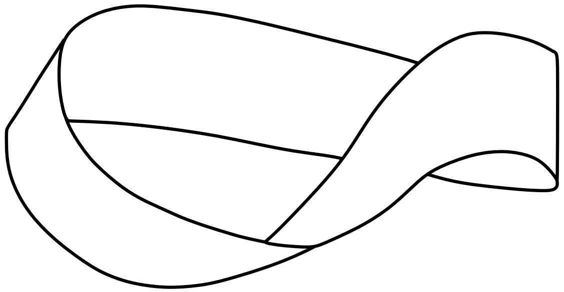
Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

We say that 'being embeddable in $\mathbb{R}^n$' is an **invariant of homeomorphism classes**. It can be used to show that two spaces are not homeomorphic.

Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

Example: The cylinder and the Möbius strip are not homeomorphic.



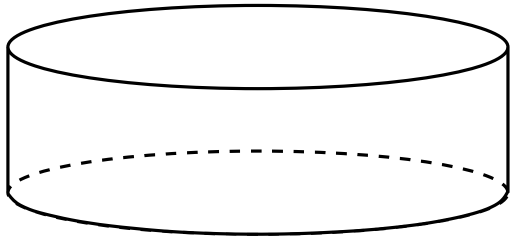


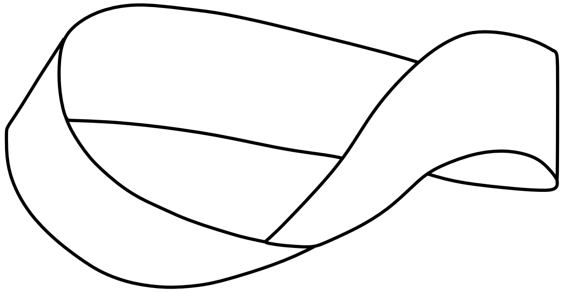
Indeed, the cylinder can be embedded in $\mathbb{R}^2$ (as an annulus).

If the strip was homeomorphic to the cylinder, then it would be also embeddable in $\mathbb{R}^2$.

Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

Example: The cylinder and the Möbius strip are not homeomorphic.

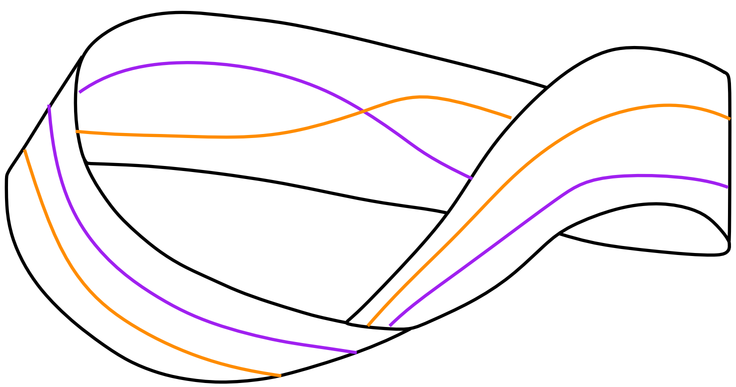




Indeed, the cylinder can be embedded in $\mathbb{R}^2$ (as an annulus).

If the strip was homeomorphic to the cylinder, then it would be also embeddable in $\mathbb{R}^2$.

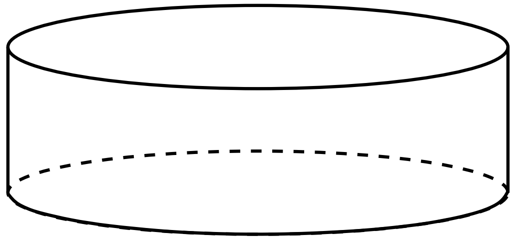
We draw two circles on the strip, C_1 and C_2, that only intersect once.

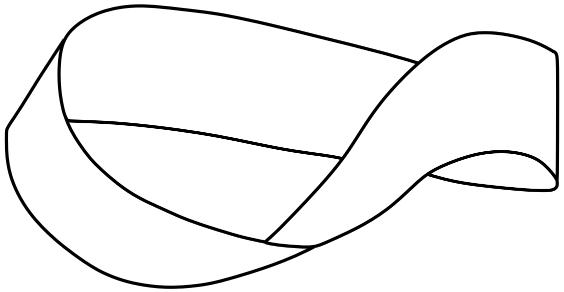


Propriedade de invariância - na teoria

Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

Example: The cylinder and the Möbius strip are not homeomorphic.

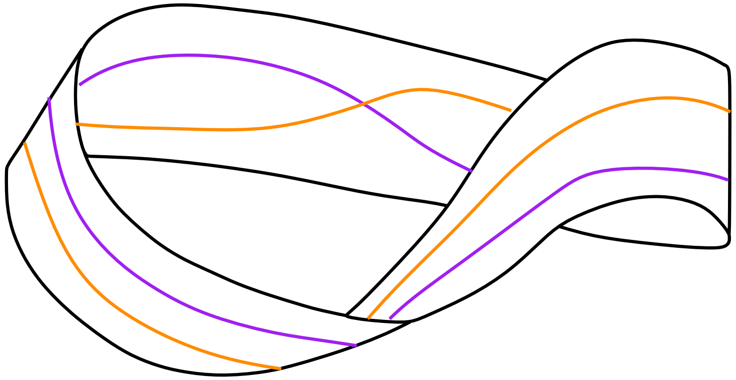


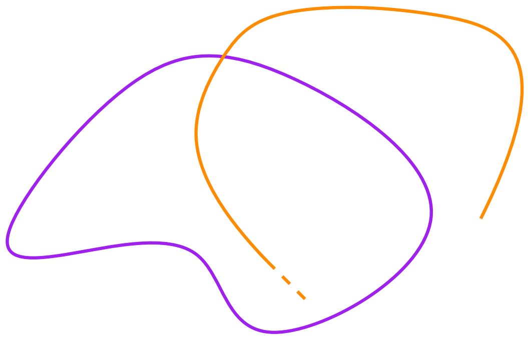


Indeed, the cylinder can be embedded in $\mathbb{R}^2$ (as an annulus).

If the strip was homeomorphic to the cylinder, then it would be also embeddable in $\mathbb{R}^2$.

We draw two circles on the strip, C_1 and C_2, that only intersect once.





Embedded in $\mathbb{R}^2$, the circles C^1 and C^2 only intersect once.
This is impossible by Jordan's theorem.

Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

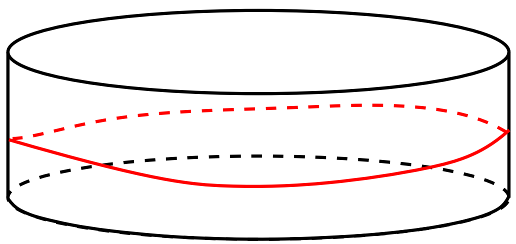
Example: The cylinder and the Möbius strip are not homeomorphic.

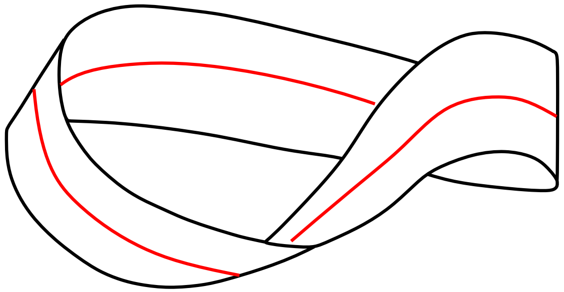
Remark: The property 'being embeddable in $\mathbb{R}^n$' is **not** an invariant of *homotopy* classes.

Indeed, the space and the cylinder are homotopy equivalent, but only one of them is embeddable in $\mathbb{R}^2$.

Proposition: Let $n \in \mathbb{N}$. If two spaces X and Y are homeomorphic, then either both are embeddable in $\mathbb{R}^n$, or neither.

Example: The cylinder and the Möbius strip are not homeomorphic.



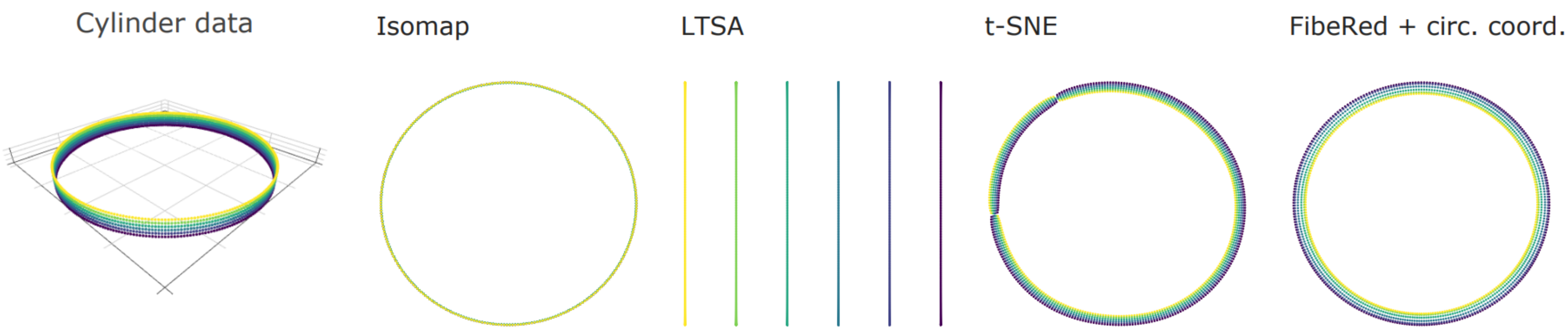


Remark: The property 'being embeddable in $\mathbb{R}^n$' is **not** an invariant of *homotopy* classes.

Indeed, the space and the cylinder are homotopy equivalent, but only one of them is embeddable in $\mathbb{R}^2$.

They can both be retracted onto their inner circle.

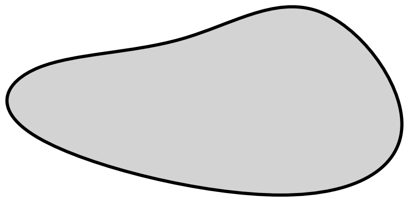
In applications, finding an embedding corresponds to the problem of **dimensionality reduction**.



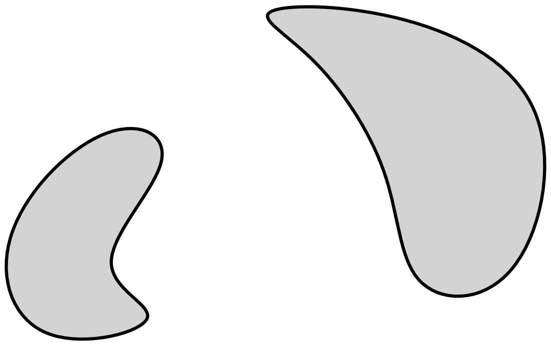
Illustrations from [Luis Scoccola, Jose A. Perea, Fiberwise dimensionality reduction of topologically complex data with vector bundles, 2022]

Componentes conexas

Definition: A subset $X \subset \mathbb{R}^n$ is (path-) **connected** if for every $x, y \in X$, there exists a continuous map $f\colon [0,1] \to X$ such that $f(0) = x$ and $f(1) = y$.

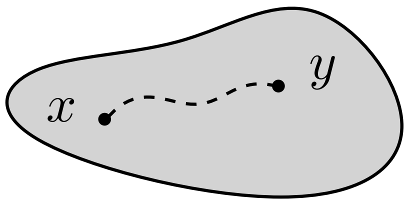


connected space

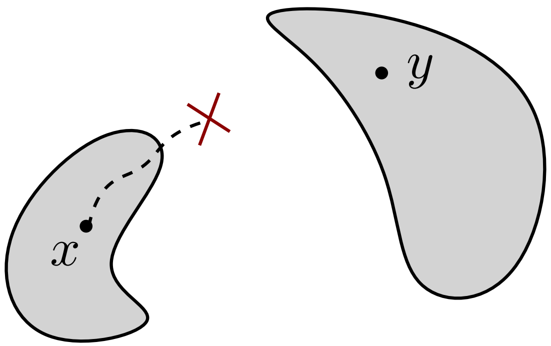


non-connected space

Definition: A subset $X \subset \mathbb{R}^n$ is (path-) **connected** if for every $x, y \in X$, there exists a continuous map $f\colon [0,1] \to X$ such that $f(0) = x$ and $f(1) = y$.



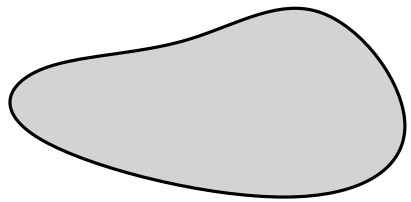
connected space



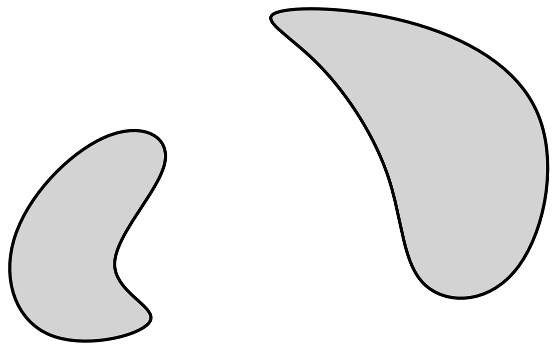
non-connected space

Componentes conexas

Definition: A subset $X \subset \mathbb{R}^n$ is (path-) **connected** if for every $x, y \in X$, there exists a continuous map $f\colon [0,1] \to X$ such that $f(0) = x$ and $f(1) = y$.

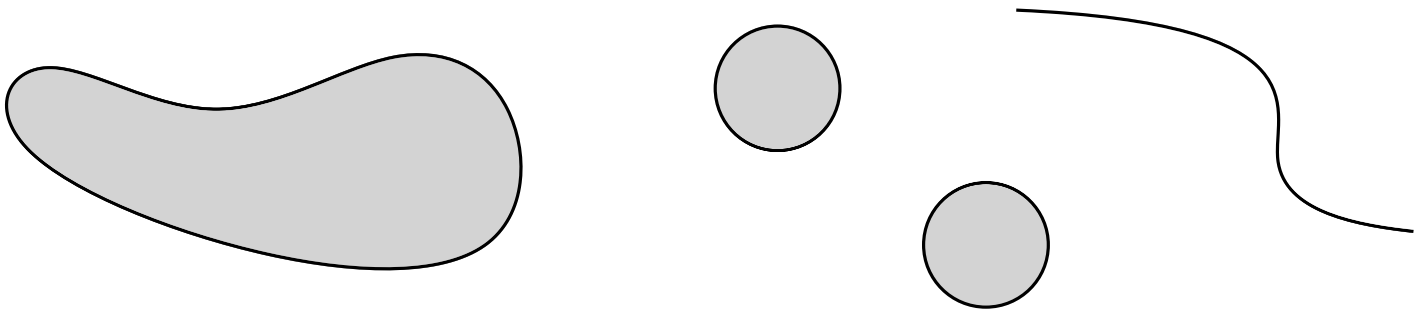


connected space



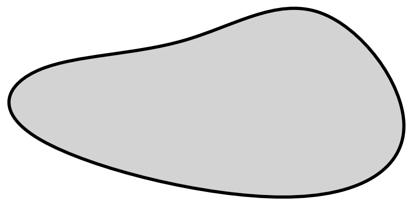
non-connected space

More generally, any topological space X can be partitioned into **connected components**.

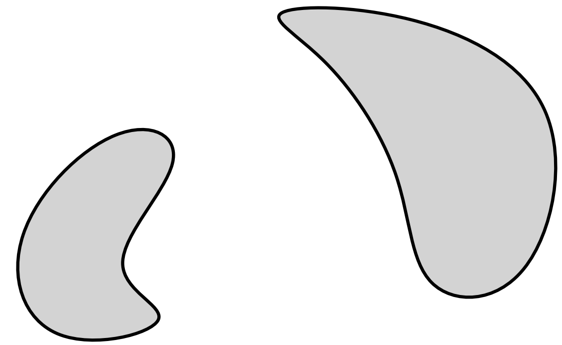


Componentes conexas

Definition: A subset $X \subset \mathbb{R}^n$ is (path-) **connected** if for every $x, y \in X$, there exists a continuous map $f\colon [0,1] \to X$ such that $f(0) = x$ and $f(1) = y$.

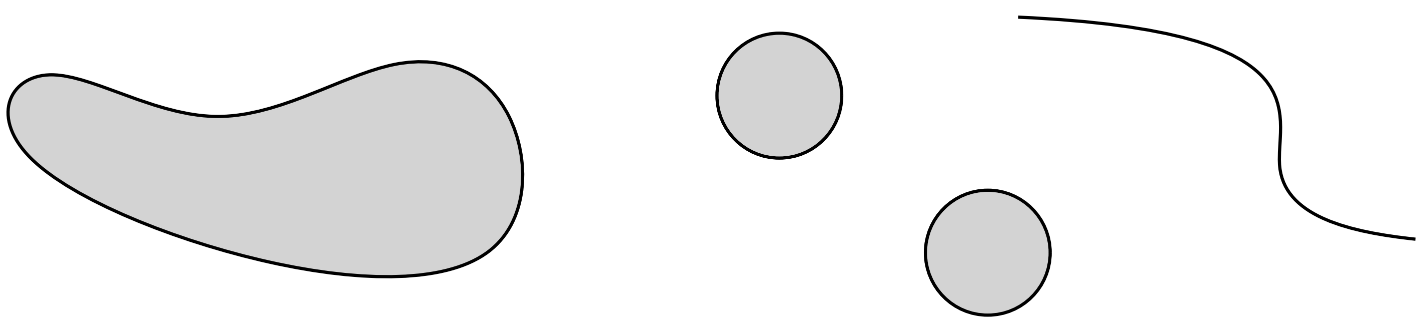


connected space



non-connected space

More generally, any topological space X can be partitioned into **connected components**.



Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Propriedade de invariância - na teoria

Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

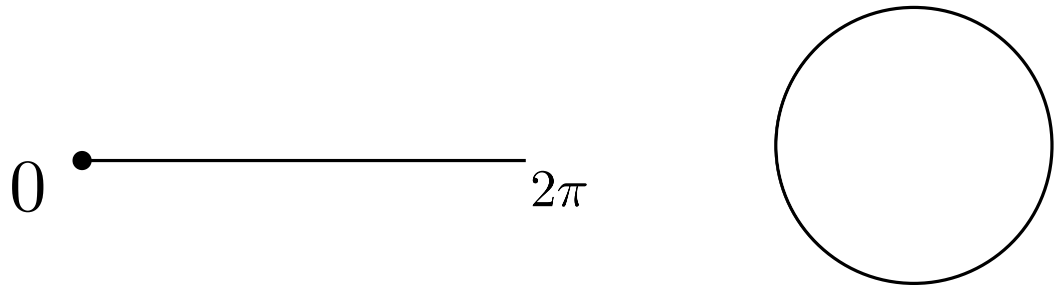
Consequence: If two spaces X and Y are homeomorphic, then they have the same number of connected components.

Example: The subsets $[0, 1]$ and $[0, 1] \cup [2, 3]$ of $\mathbb{R}$ are not homeomorphic, neither homotopy equivalent.
Indeed, the first one has one connected component, and the second one two.

Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Example: The interval $[0, 2\pi)$ and the circle $\mathbb{S}_1 \subset \mathbb{R}^2$ are not homeomorphic.

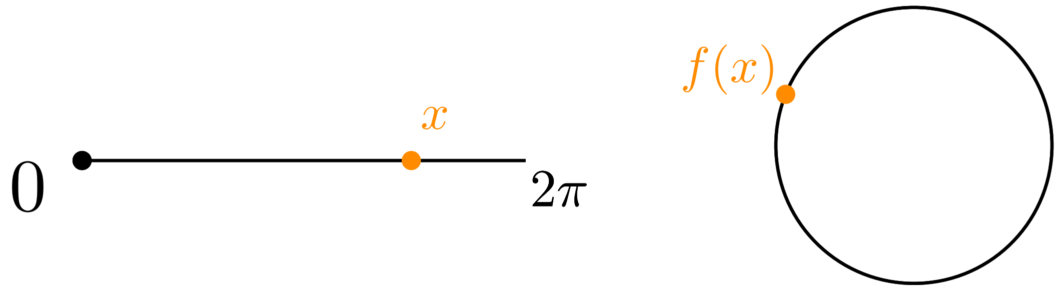
We will prove this by contradiction. Suppose that they are homeomorphic. By definition, this means that there exists a map $f\colon [0, 2\pi) \to \mathbb{S}_1$ which is continuous, invertible, and with continuous inverse.



Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Example: The interval $[0, 2\pi)$ and the circle $\mathbb{S}_1 \subset \mathbb{R}^2$ are not homeomorphic.

We will prove this by contradiction. Suppose that they are homeomorphic. By definition, this means that there exists a map $f\colon [0, 2\pi) \to \mathbb{S}_1$ which is continuous, invertible, and with continuous inverse.



Let $x \in [0, 2\pi)$ such that $x \neq 0$. Consider the subsets $[0, 2\pi) \setminus \{x\} \subset [0, 2\pi)$ and $\mathbb{S}_1 \setminus \{f(x)\} \subset \mathbb{S}_1$, and the induced map

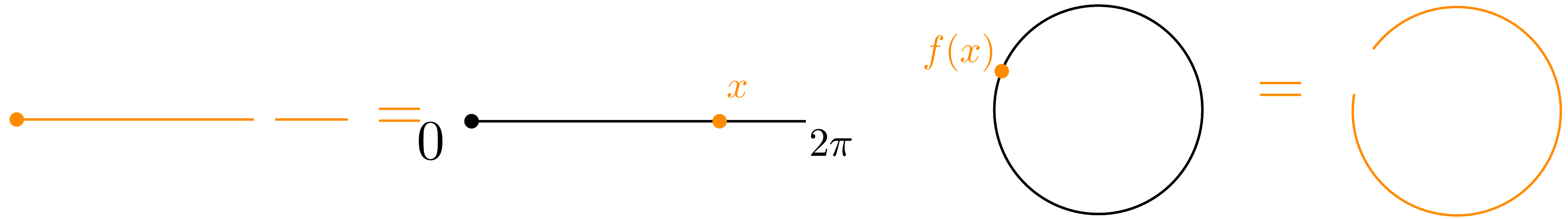
$$g\colon [0, 2\pi) \setminus \{x\} \to \mathbb{S}_1 \setminus \{f(x)\}.$$

The map g is a homeomorphism.

Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Example: The interval $[0, 2\pi)$ and the circle $\mathbb{S}_1 \subset \mathbb{R}^2$ are not homeomorphic.

We will prove this by contradiction. Suppose that they are homeomorphic. By definition, this means that there exists a map $f\colon [0, 2\pi) \to \mathbb{S}_1$ which is continuous, invertible, and with continuous inverse.



Let $x \in [0, 2\pi)$ such that $x \neq 0$. Consider the subsets $[0, 2\pi) \setminus \{x\} \subset [0, 2\pi)$ and $\mathbb{S}_1 \setminus \{f(x)\} \subset \mathbb{S}_1$, and the induced map

$$g\colon [0, 2\pi) \setminus \{x\} \to \mathbb{S}_1 \setminus \{f(x)\}.$$

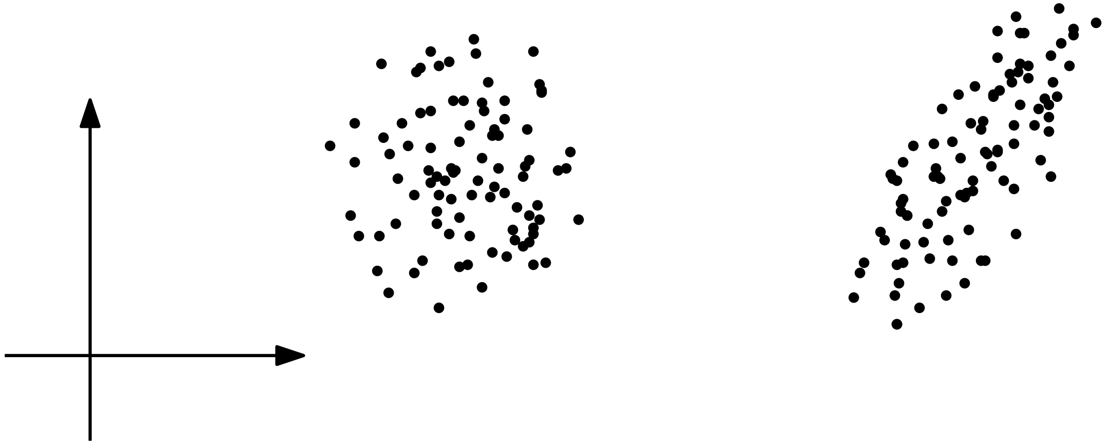
The map g is a homeomorphism.

Moreover, $[0, 2\pi) \setminus \{x\}$ has two connected components, and $\mathbb{S}_1 \setminus \{f(x)\}$ only one. This is absurd.

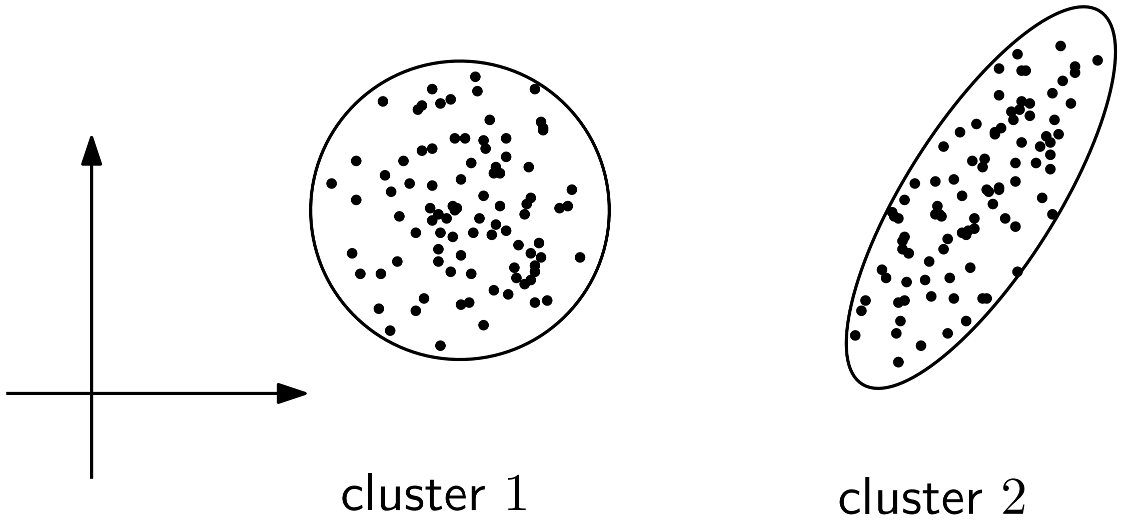
Proposition: If two spaces X and Y are homotopy equivalent, then they have the same number of connected components.

Homework: The intervals $[0, 1)$ and $(0, 1)$ are not homeomorphic.

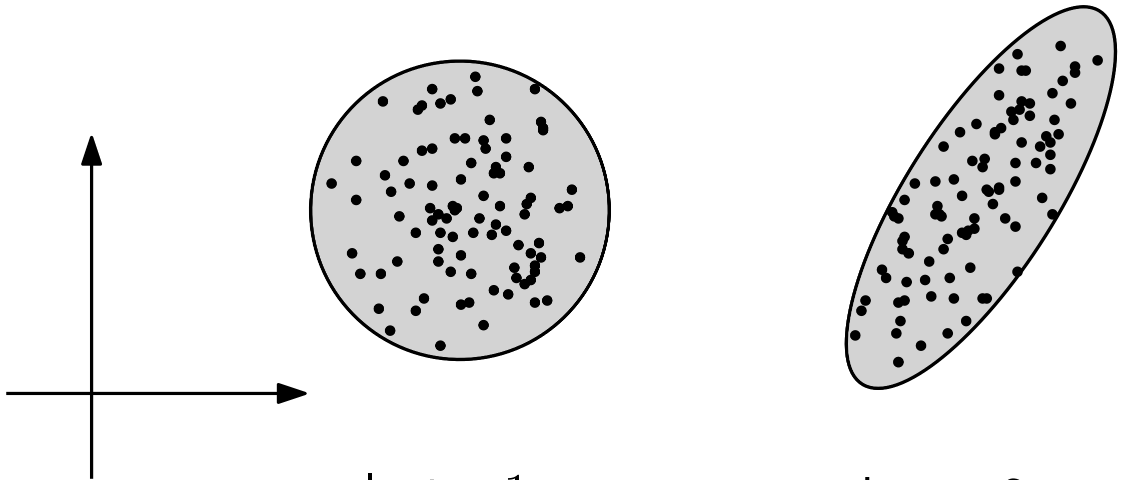
In applications, finding connected components corresponds to a **classification** task.



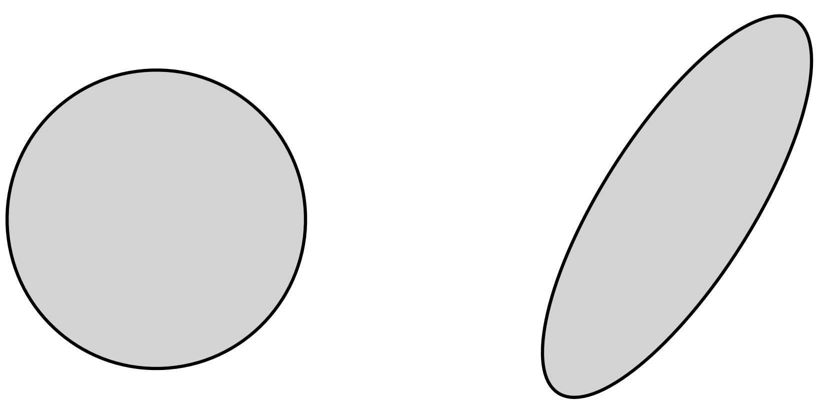
In applications, finding connected components corresponds to a **classification** task.



In applications, finding connected components corresponds to a **classification** task.



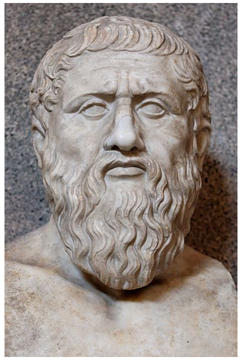
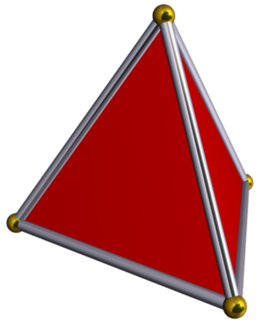
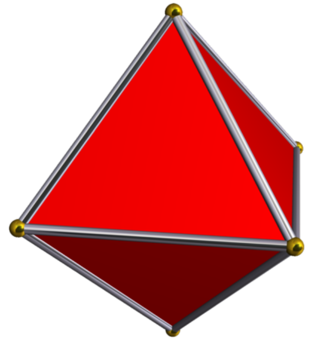
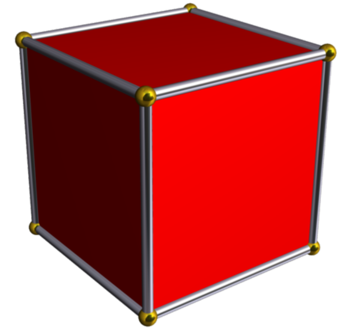
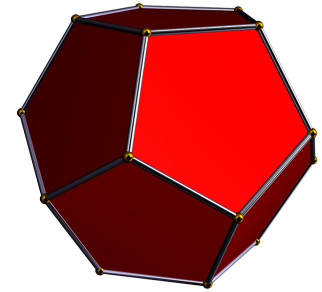
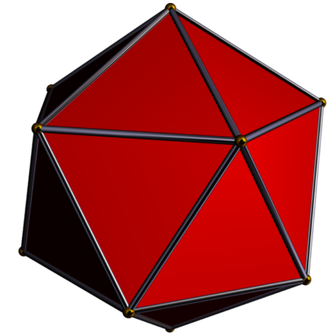
cluster 1 cluster 2



connected component 1 connected component 2

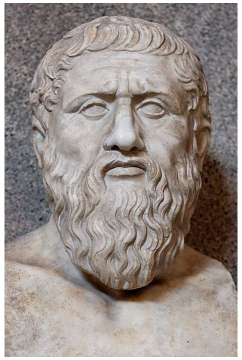
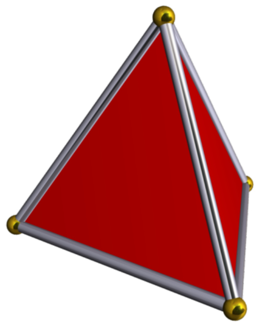
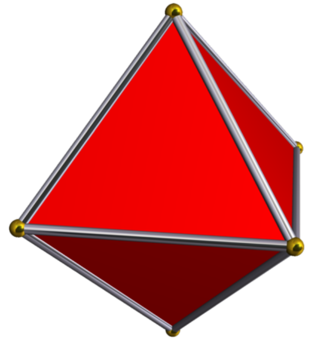
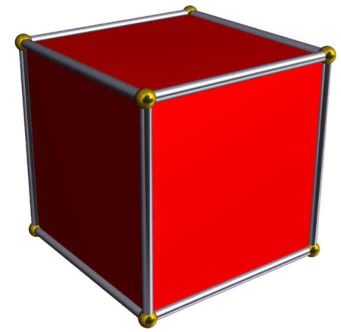
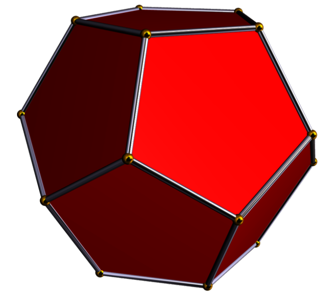
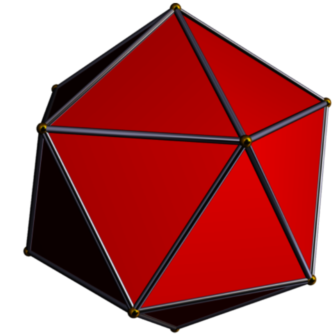
We can think of these sets as an **underlying topological space**, on which the points are sampled.

Característica de Euler

number of faces	4	8	6	12	20
number of edges	6	12	12	30	30
number of vertices	4	6	8	20	12
χ	2	2	2	2	2

Característica de Euler

					
number of faces	4	8	6	12	20
number of edges	6	12	12	30	30
number of vertices	4	6	8	20	12
χ	2	2	2	2	2

Proposition [Euler, 1758]: In any convex polyhedron, we have
number of faces $-$ number of edges $+$ number of vertices $= 2$

Definition: Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.

The **dimension** of a simplex $\sigma \in K$ is defined as $|\sigma| - 1$.

Característica de Euler

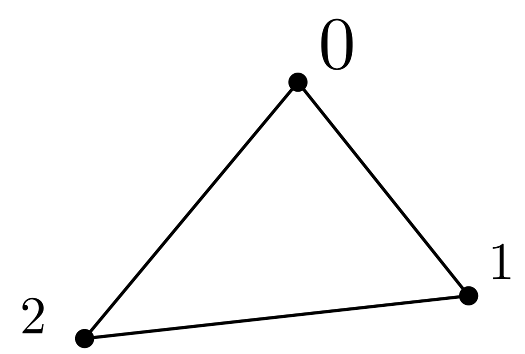
Definition: Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.

The **dimension** of a simplex $\sigma \in K$ is defined as $|\sigma| - 1$.

Example: Let $V = \{0, 1, 2\}$ and

$$K = \{[0], [1], [2], [0, 1], [1, 2], [0, 2]\}.$$

This is a simplicial complex.



It contains three simplices of dimension 0 ($[0]$, $[1]$ and $[2]$) and three simplices of dimension 1 ($[0, 1]$, $[1, 2]$ and $[0, 2]$).

Característica de Euler

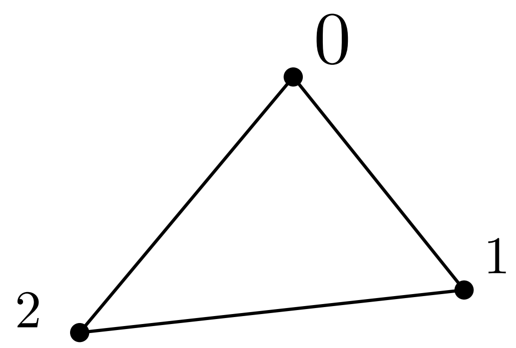
Definition: Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.

The **dimension** of a simplex $\sigma \in K$ is defined as $|\sigma| - 1$.

Example: Let $V = \{0, 1, 2\}$ and

$$K = \{[0], [1], [2], [0,1], [1,2], [0,2]\}.$$

This is a simplicial complex.



(this is a circle)

It contains three simplices of dimension 0 ($[0]$, $[1]$ and $[2]$) and three simplices of dimension 1 ($[0,1]$, $[1,2]$ and $[0,2]$).

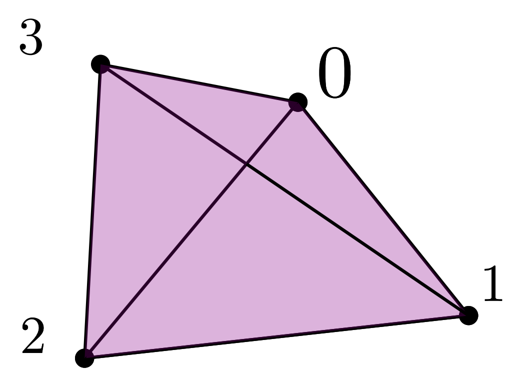
Definition: Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.

The **dimension** of a simplex $\sigma \in K$ is defined as $|\sigma| - 1$.

Example: Let $V = \{0, 1, 2, 3\}$ and

$$K = \{[0], [1], [2], [3], [0,1], [1,2], [2,3], [3,0], [0,2], [1,3], [0,1,2], [0,1,3], [0,2,3], [1,2,3]\}$$

It a simplicial complex.



It contains four simplices of dimension 0 ($[0]$, $[1]$, $[2]$ and $[3]$), six simplices of dimension 1 ($[0,1]$, $[1,2]$, $[2,3]$, $[3,0]$,$[0,2]$ and $[1,3]$) and four simplices of dimension 2 ($[0,1,2]$, $[0,1,3]$, $[0,2,3]$ and $[1,2,3]$).

Característica de Euler

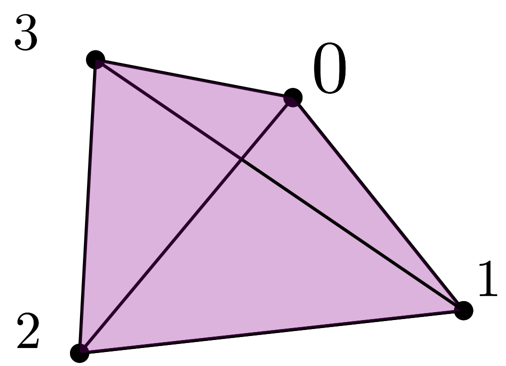
Definition: Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.

The **dimension** of a simplex $\sigma \in K$ is defined as $|\sigma| - 1$.

Example: Let $V = \{0, 1, 2, 3\}$ and

$$K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 3], [3, 0], [0, 2], [1, 3], [0, 1, 2], [0, 1, 3], [0, 2, 3], [1, 2, 3]\}$$

It a simplicial complex.



(this is a sphere)

It contains four simplices of dimension 0 ($[0]$, $[1]$, $[2]$ and $[3]$), six simplices of dimension 1 ($[0, 1]$, $[1, 2]$, $[2, 3]$, $[3, 0]$,$[0, 2]$ and $[1, 3]$) and four simplices of dimension 2 ($[0, 1, 2]$, $[0, 1, 3]$, $[0, 2, 3]$ and $[1, 2, 3]$).

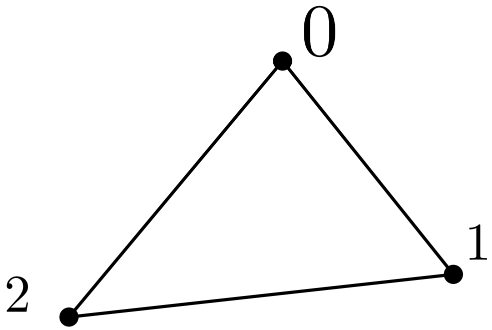
Característica de Euler

Definition: Let K be a simplicial complex of dimension n. Its **Euler characteristic** is the integer

$$\chi(K) = \sum_{0 \le i \le n} (-1)^i \cdot (\text{number of simplices of dimension } i).$$

Example: The simplicial complex $K = \{[0], [1], [2], [0,1], [1,2], [2,0]\}$ has Euler characteristic

$$\chi(K) = 3 - 3 = 0$$



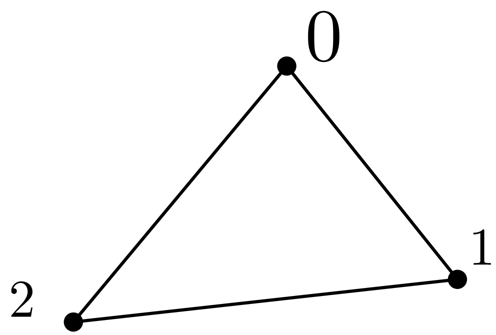
Característica de Euler

Definition: Let K be a simplicial complex of dimension n. Its **Euler characteristic** is the integer

$$\chi(K) = \sum_{0 \leq i \leq n} (-1)^i \cdot (\text{number of simplices of dimension } i).$$

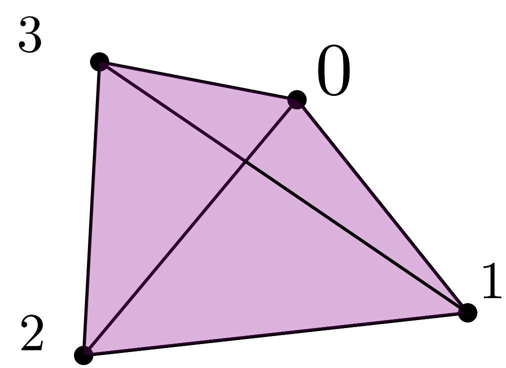
Example: The simplicial complex $K = \{[0],[1],[2],[0,1],[1,2],[2,0]\}$ has Euler characteristic

$$\chi(K) = 3 - 3 = 0$$



Example: The simplicial complex
$K = \{[0],[1],[2],[3],[0,1],[1,2],[2,3],[3,0],[0,2],[1,3],[0,1,2],[0,1,3],[0,2,3],[1,2,3]\}$
has Euler characteristic

$$\chi(K) = 4 - 6 + 4 = 2$$

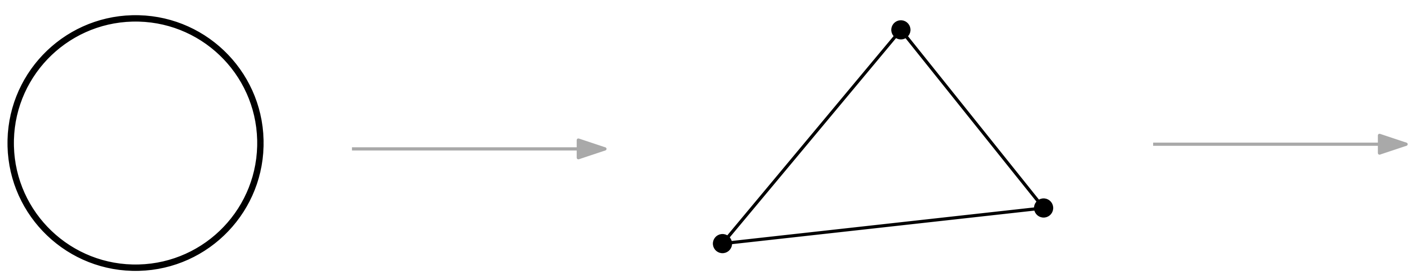


Característica de Euler

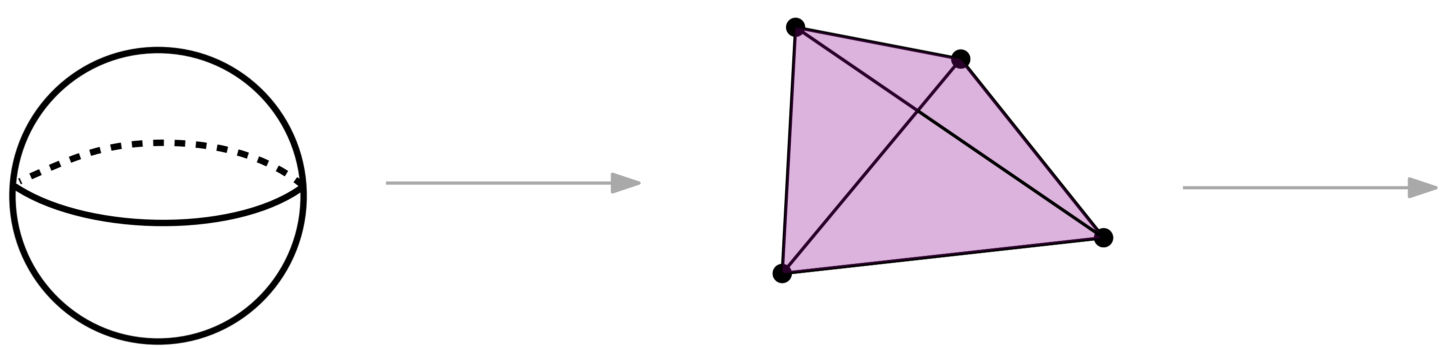
Definition: Let K be a simplicial complex of dimension n. Its **Euler characteristic** is the integer

$$\chi(K) = \sum_{0 \leq i \leq n} (-1)^i \cdot (\text{number of simplices of dimension } i).$$

Definition: Let X be a topological space. Its **Euler characteristic** is defined as the Euler caracteristic of a *triangulation* of X.



$\chi(X) = 0$



$\chi(X) = 2$

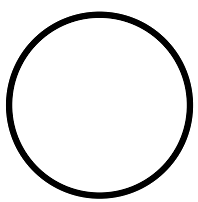
Propriedade de invariância - na teoria

Proposition: If X and Y are two homotopy equivalent topological spaces, then $\chi(X) = \chi(Y)$.

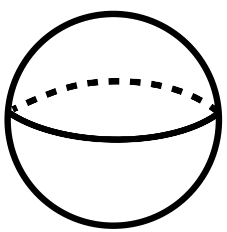
Therefore, the Euler characteristic is an **invariant** of homotopy equivalence classes.

We can use this information to prove that two spaces are not homotopy equivalent.

Example: The circle has Euler characteristic 0, and the sphere Euler characteristic 2. Therefore, they are not homotopy equivalent.



$\chi(\mathbb{S}_1) = 0$



$\chi(\mathbb{S}_2) = 2$

Propriedade de invariância - na teoria

Proposition: If X and Y are two homotopy equivalent topological spaces, then $\chi(X) = \chi(Y)$.

Therefore, the Euler characteristic is an **invariant** of homotopy equivalence classes.

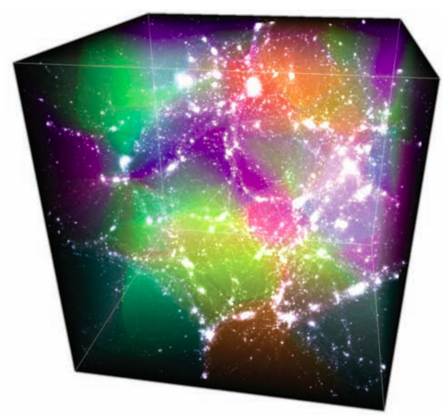
We can use this information to prove that two spaces are not homotopy equivalent.

Example (Classification of surfaces): The homeomorphism classes of *connected and compact surfaces* are classified by their Euler characteristic.

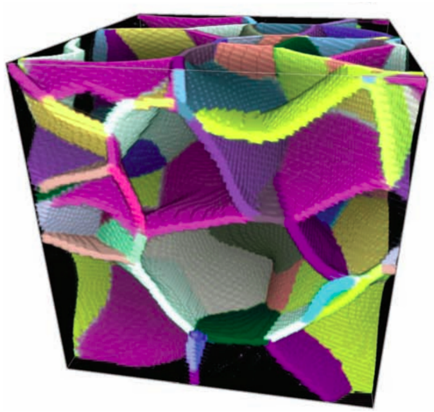
$\chi \quad 2 \quad 0 \quad -2 \quad -4 \quad 2 - 2 \times \text{genus}$

The Euler characteristic contains information about the homeomorphism class (and homotopy class) of the space.

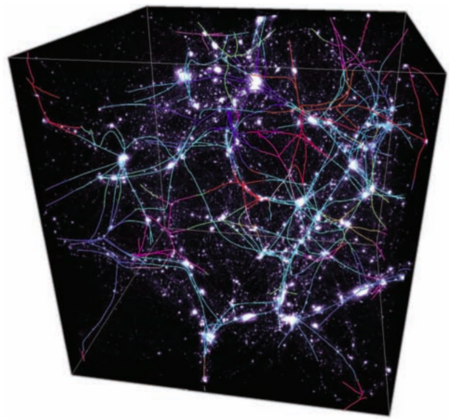
[T. Sousbie, The persistent cosmic web and its filamentary structure, 2011]



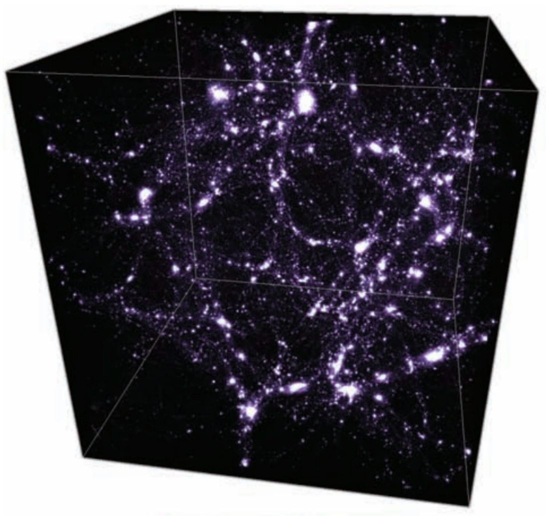
seen as an object of dimension 3



of dimension 2

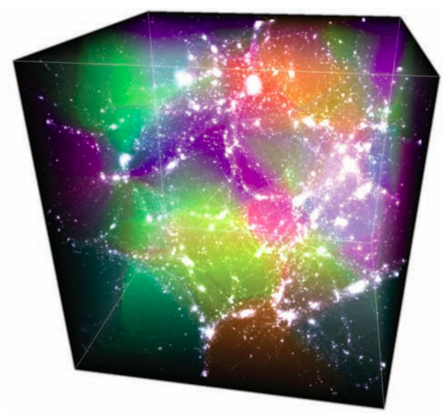


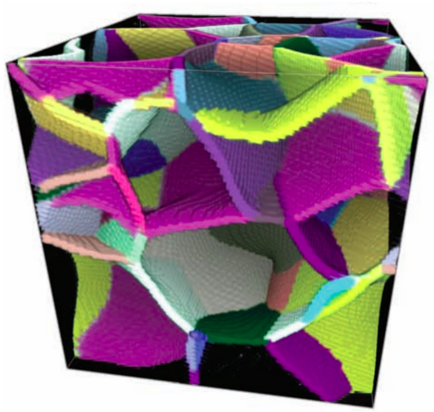
of dimension 1

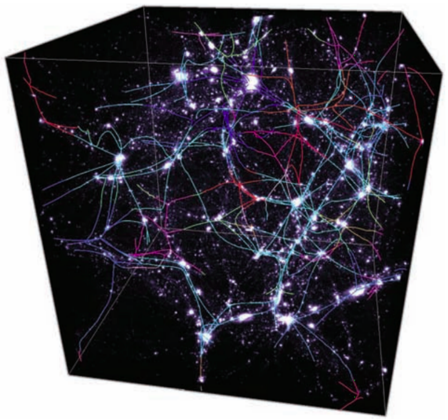


The Euler characteristic contains information about the homeomorphism class (and homotopy class) of the space.

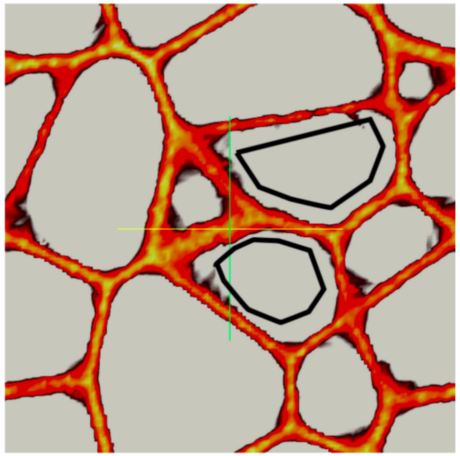
[T. Sousbie, The persistent cosmic web and its filamentary structure, 2011]







[P. Pranav, H. Edelsbrunner, R. de Weygaert, G. Vegter, M. Kerber, B. Jones and M. Wintraecken, The topology of the cosmic web in terms of persistent Betti numbers, 2016]



The Euler characteristic 'counts' the number of holes

Números de Betti

For any topological space X, one defines a sequence of integers

$$\beta_0(X), \quad \beta_1(X), \quad \beta_2(X), \quad \beta_3(X), \quad \ldots$$

called the **Betti numbers**.

Construction of Betti numbers:

Números de Betti

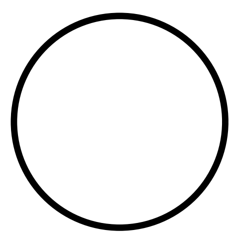
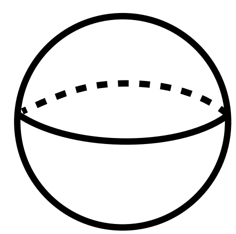
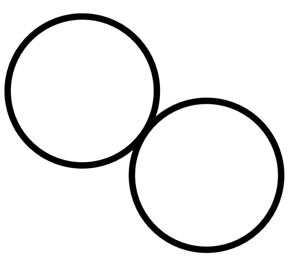
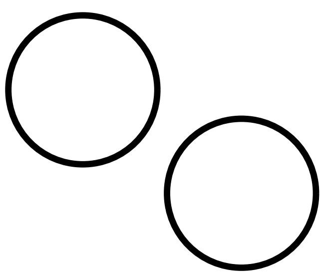
For any topological space X, one defines a sequence of integers

$$\beta_0(X), \quad \beta_1(X), \quad \beta_2(X), \quad \beta_3(X), \quad \ldots$$

called the **Betti numbers**.

Construction of Betti numbers: rendez-vous on Thursday! (based on homology theory)

Example: Let us give some examples instead.

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Números de Betti

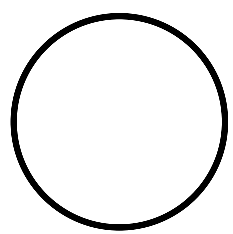
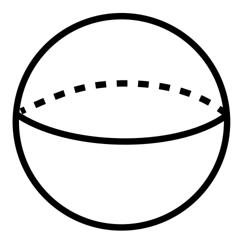
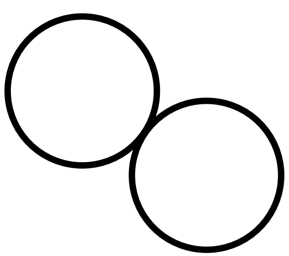
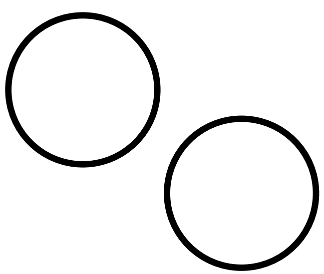
For any topological space X, one defines a sequence of integers

$$\beta_0(X), \quad \beta_1(X), \quad \beta_2(X), \quad \beta_3(X), \quad \ldots$$

called the **Betti numbers**.

Construction of Betti numbers: rendez-vous on Thursday! (based on homology theory)

Example: Let us give some examples instead.

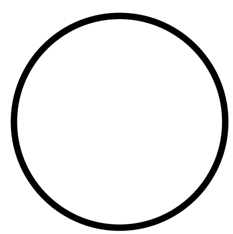
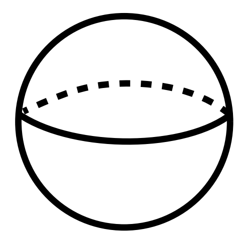
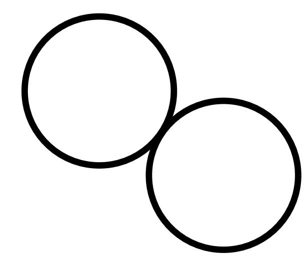
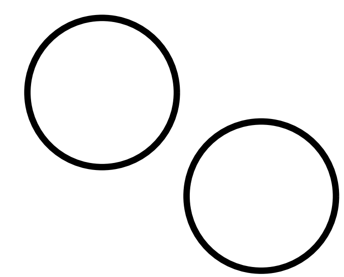
X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\dots$

Example: Let us give some examples instead.

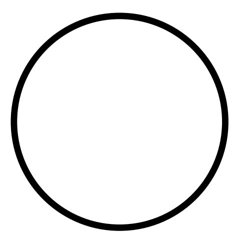
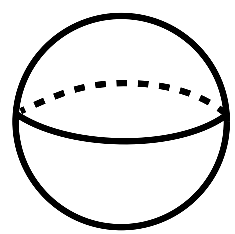
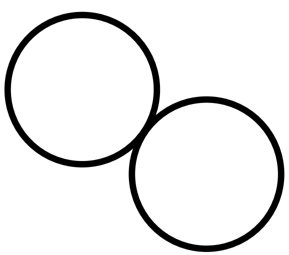
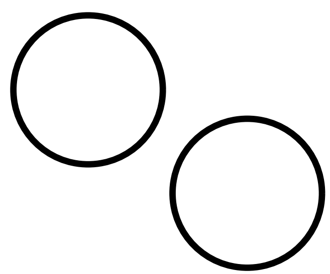
X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Números de Betti

Interpretation: We have:

- $\boxed{\beta_0(X) \text{ is the number of connected components of } X}$
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\ldots$

Example: Let us give some examples instead.

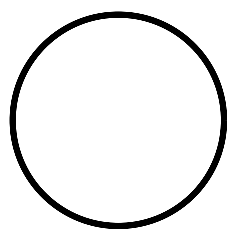
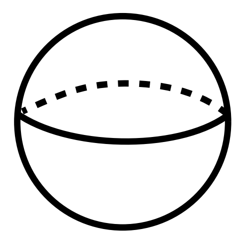
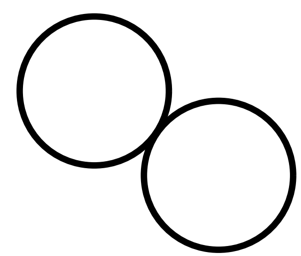
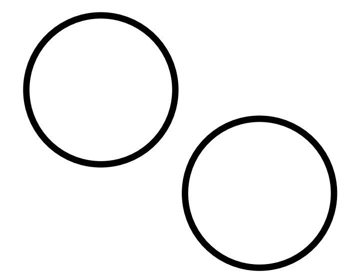
X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\boxed{\beta_1(X) \text{ is the number of 'holes' in } X}$
- $\beta_2(X)$ is the number of 'voids' in X
- ...

Example: Let us give some examples instead.

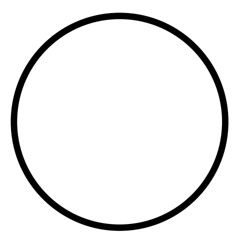
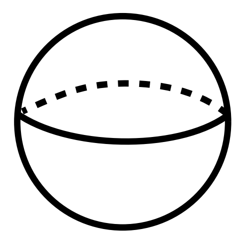
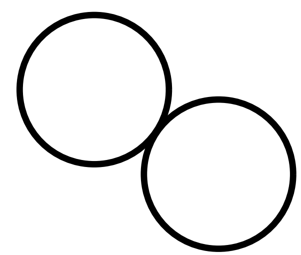
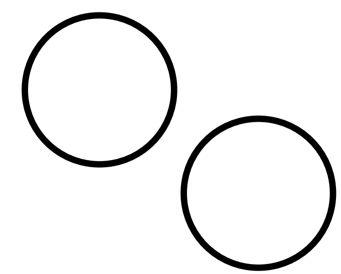
X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\boxed{\beta_2(X) \text{ is the number of 'voids' in } X}$
- $\dots$

Example: Let us give some examples instead.

X					
$\beta_0(X)$	1	1	1	1	2
$\beta_1(X)$	0	1	0	2	2
$\beta_2(X)$	0	0	1	0	0

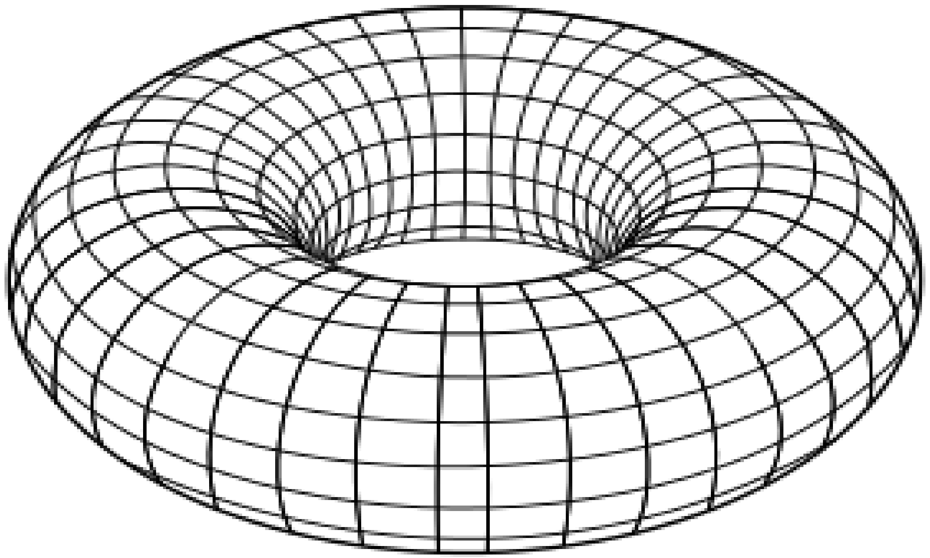
Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\ldots$

Example: Betti numbers of the torus:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \ldots$$



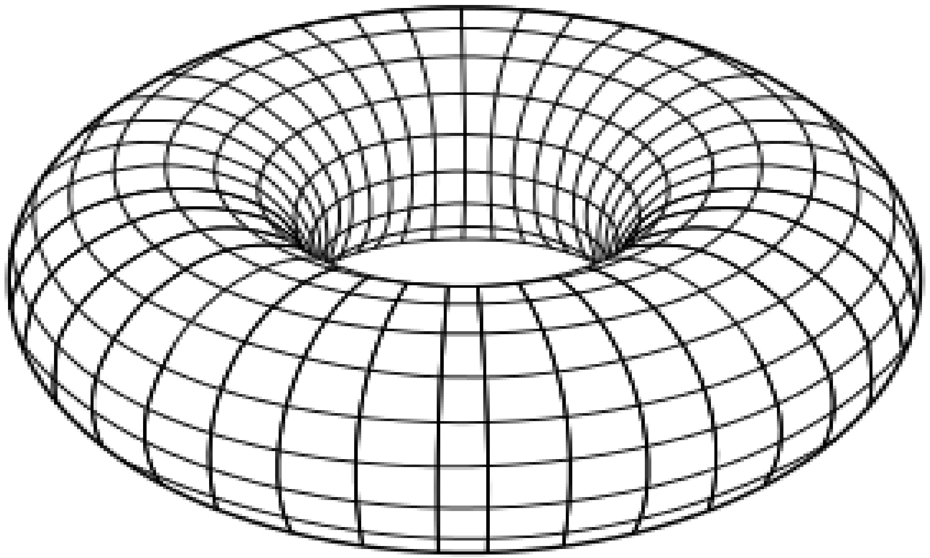
Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\ldots$

Example: Betti numbers of the torus:

$$\boxed{\beta_0(X) = 1,} \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \ldots$$

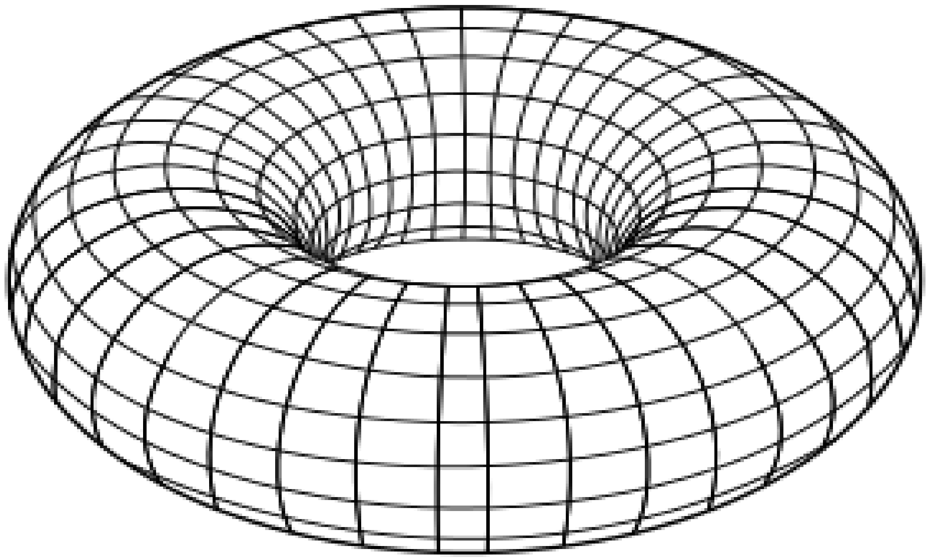


Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\ldots$

Example: Betti numbers of the torus:

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \boxed{\beta_2(X) = 1,} \quad \beta_3(X) = 0, \quad \ldots$$



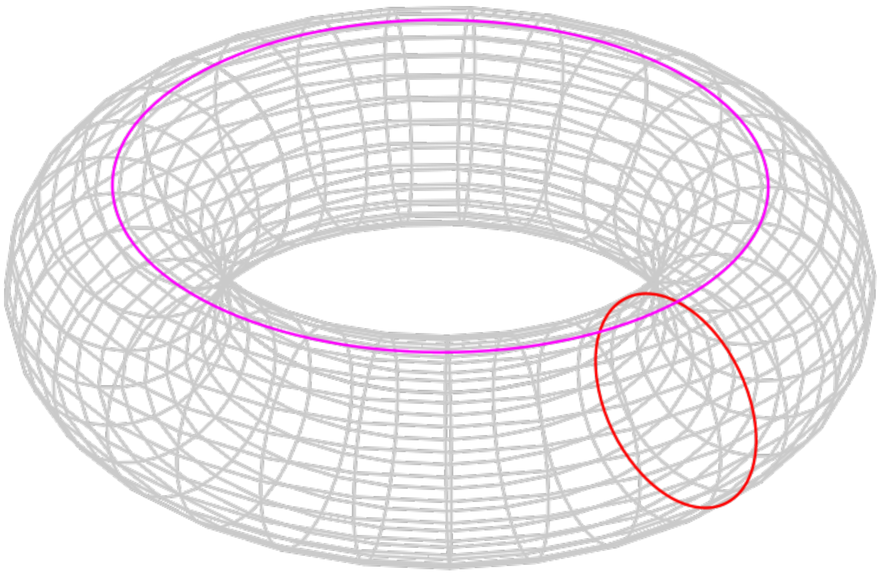
Números de Betti

Interpretation: We have:

- $\beta_0(X)$ is the number of connected components of X
- $\beta_1(X)$ is the number of 'holes' in X
- $\beta_2(X)$ is the number of 'voids' in X
- $\ldots$

Example: Betti numbers of the torus:

$$\beta_0(X) = 1, \quad \boxed{\beta_1(X) = 2,} \quad \beta_2(X) = 1, \quad \beta_3(X) = 0, \quad \ldots$$



Proposition: If two spaces X and Y are homotopy equivalent, then they have the same Betti numbers.

As a consequence, two spaces with different Betti numbers cannot be homotopy equivalent.

Example: The n-dimensional sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$ has Betti numbers

$$\begin{aligned} \beta_i(X) &= 1 \quad \text{if } \ i = 0 \text{ or } n, \\ \beta_i(X) &= 0 \quad \text{else.} \end{aligned}$$

Hence, if $n \neq m$, then $\mathbb{S}^n$ and $\mathbb{S}^n$ are not homotopy equivalent.

Proposition: If two spaces X and Y are homotopy equivalent, then they have the same Betti numbers.

As a consequence, two spaces with different Betti numbers cannot be homotopy equivalent.

Example (Brouwer's invariance of domain):

Let us show that $\mathbb{R}^n$ and $\mathbb{R}^m$, with $n \neq m$, are not homeomorphic.

Let $h\colon \mathbb{R}^n \to \mathbb{R}^m$ be a homeomorphism.
Choose any $x \in \mathbb{R}^n$and consider the restricted map

$$h\colon \mathbb{R}^n \setminus \{x\} \longrightarrow \mathbb{R}^m \setminus \{h(x)\}$$

It is still a homemorphism.

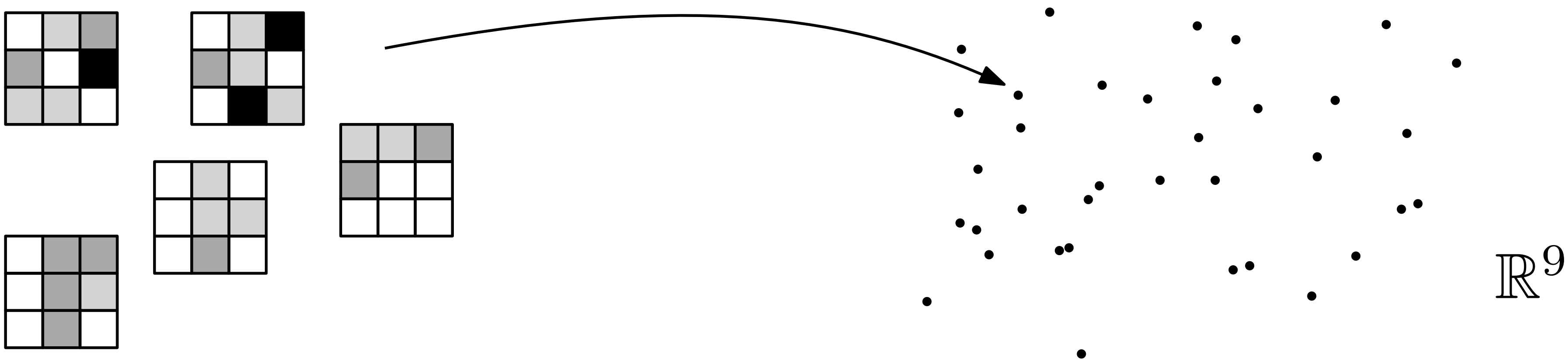
But $\mathbb{R}^n \setminus \{x\}$ is homotopic to the sphere $\mathbb{S}^{n-1}$, and $\mathbb{R}^m \setminus \{x\}$ is homotopic to the sphere $\mathbb{S}^{m-1}$

We have seen before that $\mathbb{S}^{n-1}$ and $\mathbb{S}^{m-1}$ are homotopic if and only if $m = n$. This is a contradiction.

The Betti numbers contain information about the space we study.

[G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, On the Local Behavior of Spaces of Natural Images, 2008.]

From a large collection of natural images, the authors extract 3×3 patches.
Since it consists of 9 pixels, each of these patches can be seen as a 9-dimensional vector, and the whole set as a point cloud in $\mathbb{R}^9$.



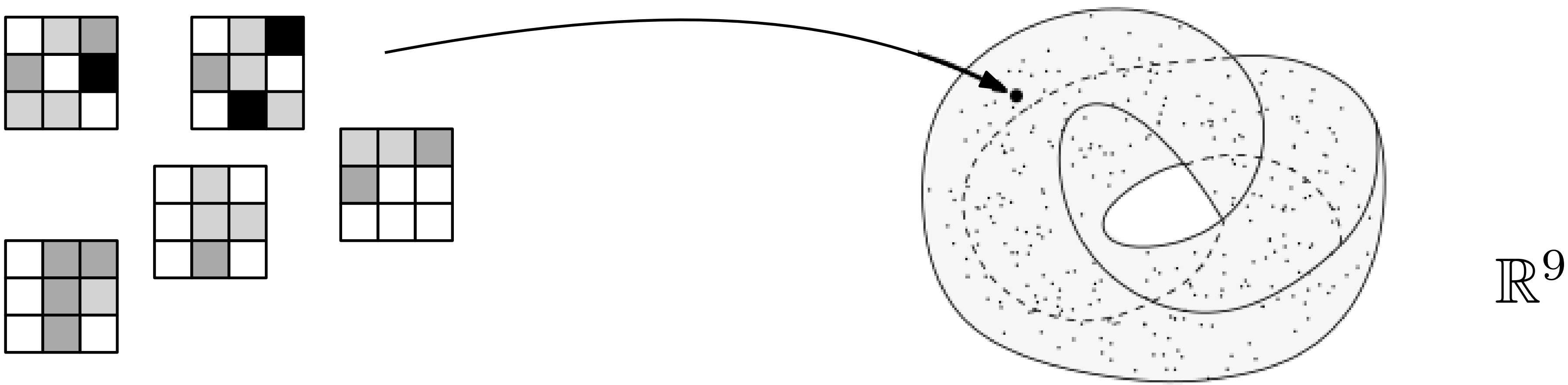
They observe that the point cloud lies close to a shape whose Betti numbers (over $\mathbb{Z}/2\mathbb{Z}$) are

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0$$

The Betti numbers contain information about the space we study.

[G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, On the Local Behavior of Spaces of Natural Images, 2008.]

From a large collection of natural images, the authors extract 3×3 patches. Since it consists of 9 pixels, each of these patches can be seen as a 9-dimensional vector, and the whole set as a point cloud in $\mathbb{R}^9$.



They observe that the point cloud lies close to a shape whose Betti numbers (over $\mathbb{Z}/2\mathbb{Z}$) are

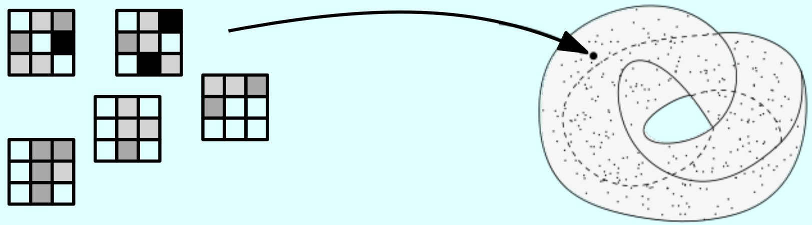
$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1, \quad \beta_3(X) = 0$$

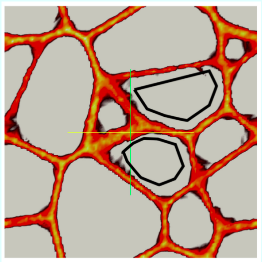
These are the Betti numbers of a Klein bottle!

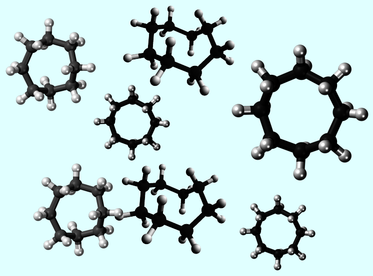
(and the authors actually show that the dataset concentrates near a Klein bottle embedded in $\mathbb{R}^9$.)

Conclusão

We can find interesting topology in datasets.







Invariants of homotopy classes allow to describe and understand them.

$$\beta_0(X) = 1, \quad \beta_1(X) = 2, \quad \beta_2(X) = 1$$

Thursday 28th: a stronger invariant, **homology**.

Tuesday 3rd: how to compute these invariants in practice? **persistent homology**.

Thursday 5th: **python tutorial** with the gudhi library.

A course about TDA: `https://raphaeltinarrage.github.io/EMAp.html`