From Algebraic Topology to Data Analysis

Part II/III: Homological inference

https://raphaeltinarrage.github.io

Last update: January 26, 2023

Part I/III: Topological invariants

Thursday 26, 9∼11am

Part II/III: Homological inference

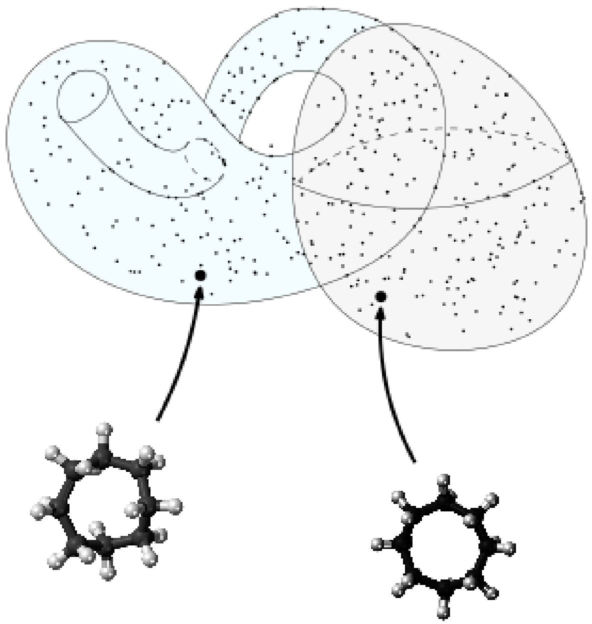
Friday 27, 9∼11am

Part III/III: Persistent Homology

Friday 27, 3∼5pm

Lembrete de ontem

Some datasets contain topology

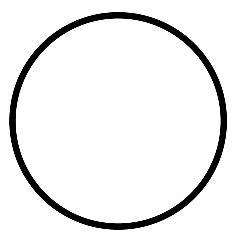
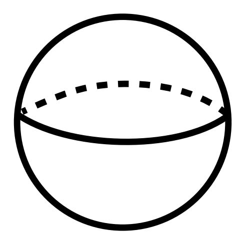
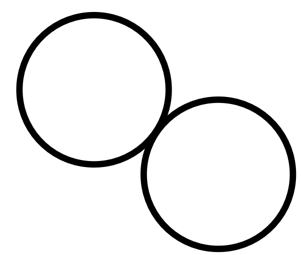
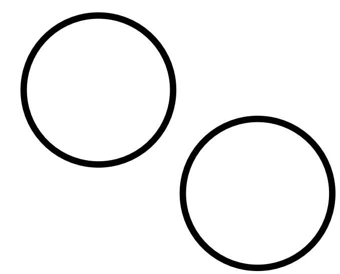


Invariants of homotopy classes allow to describe and understand topological spaces

Number of connected components

Euler characteristic χ

Betti numbers $\beta_0, \beta_1, \beta_2, \ldots$

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

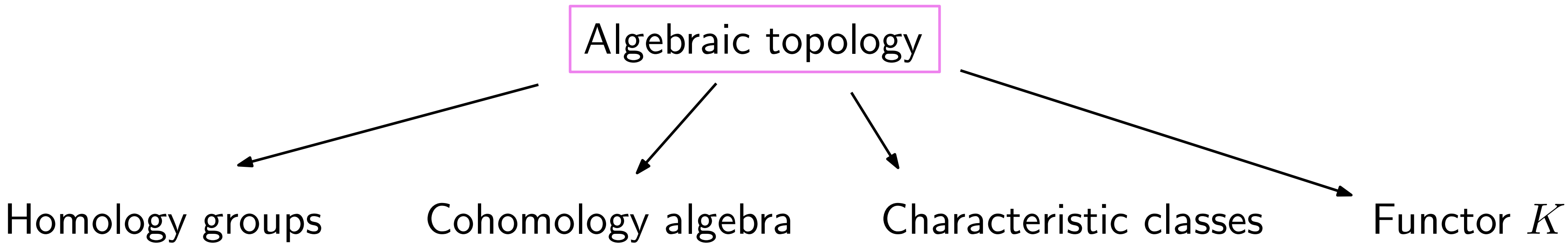
Cardápio

Today we will define a powerful invariant, **homology groups**, that already contains the number of connected components, and the Euler characteristic.

Algebraic topology

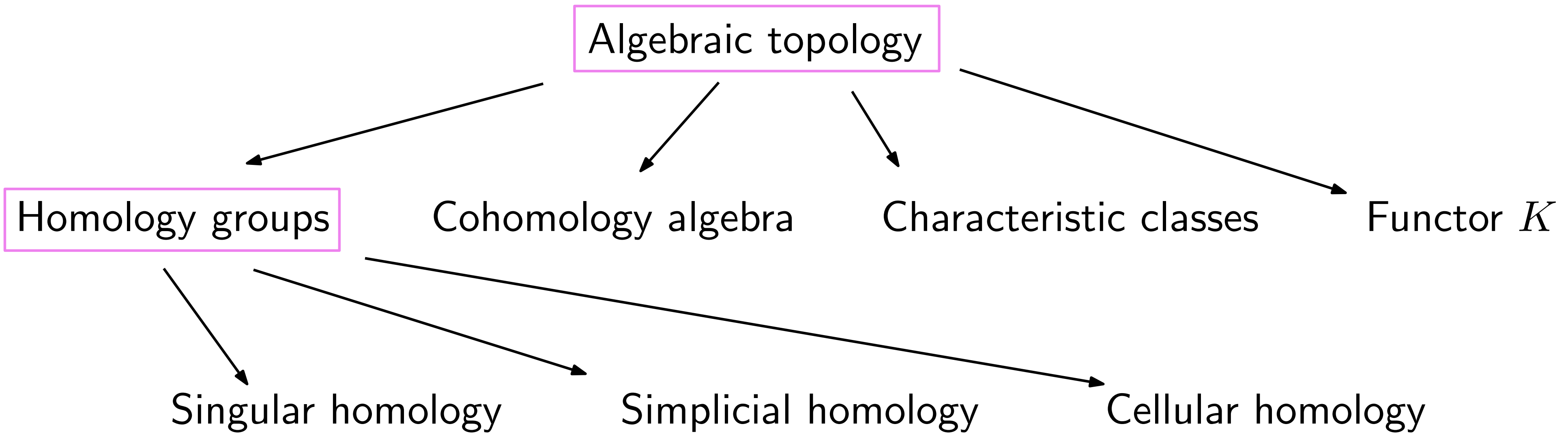
Cardápio

Today we will define a powerful invariant, **homology groups**, that already contains the number of connected components, and the Euler characteristic.



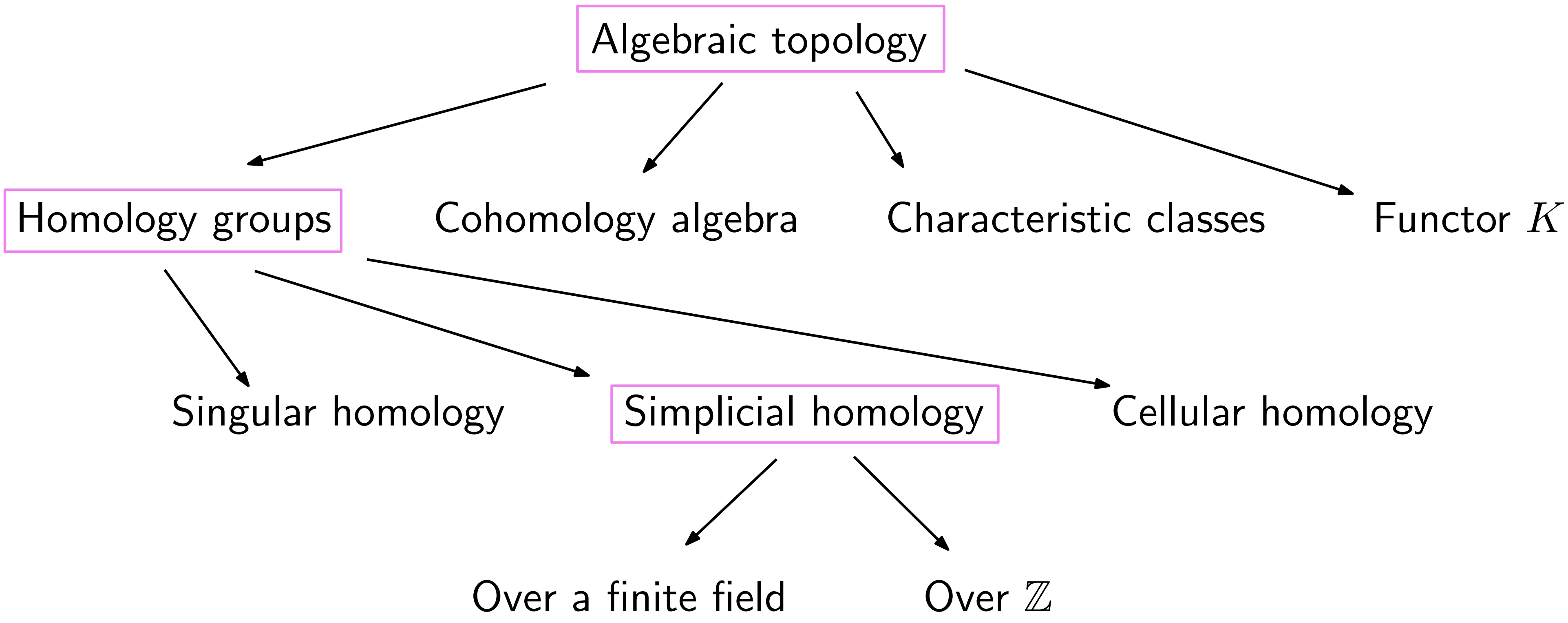
Cardápio

Today we will define a powerful invariant, **homology groups**, that already contains the number of connected components, and the Euler characteristic.



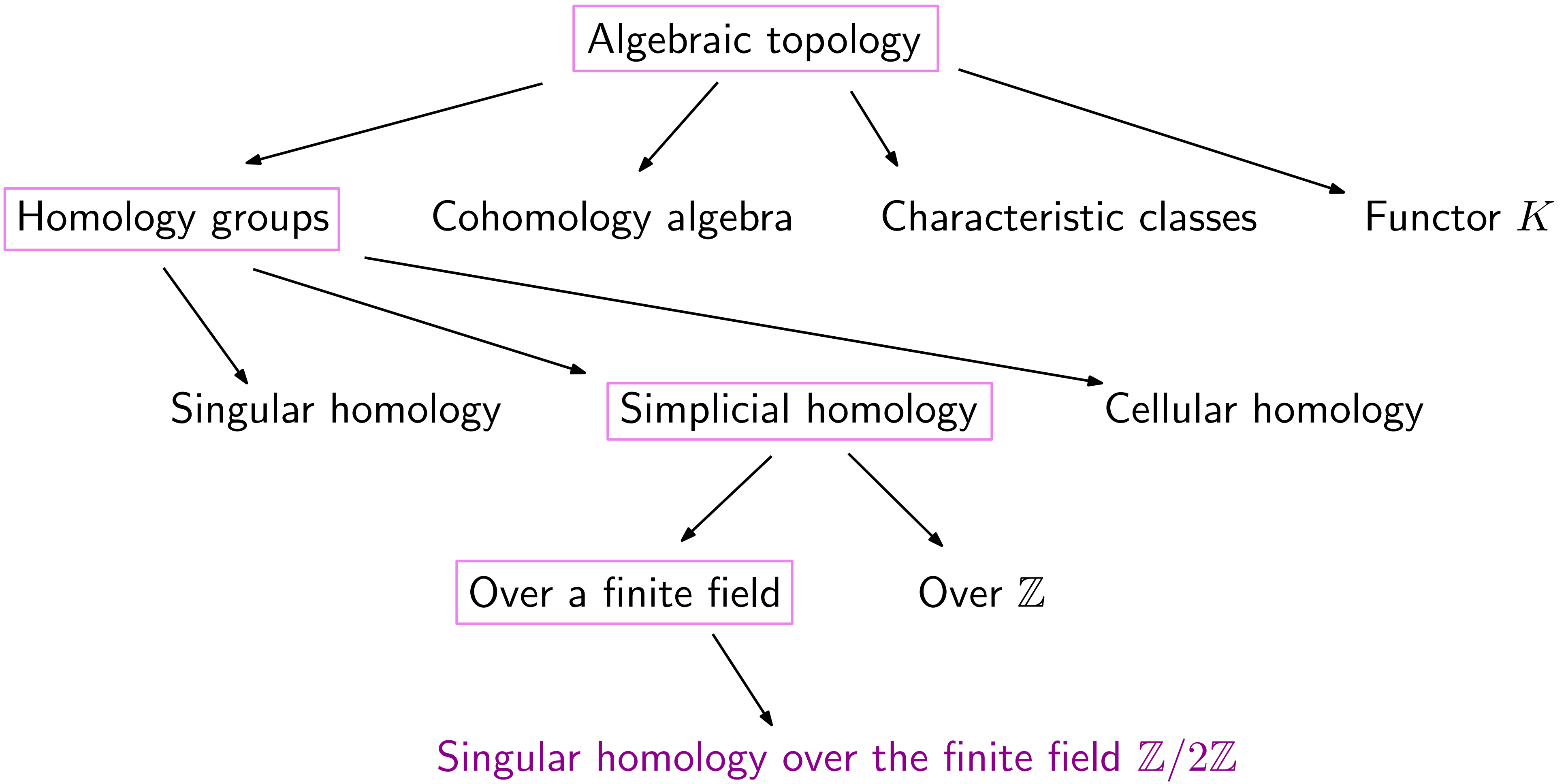
Cardápio

Today we will define a powerful invariant, **homology groups**, that already contains the number of connected components, and the Euler characteristic.



Cardápio

Today we will define a powerful invariant, **homology groups**, that already contains the number of connected components, and the Euler characteristic.



I - Simplicial homology

1 - Reminder of algebra

2 - Homological algebra

3 - Incremental algorithm

II - More about homology

1 - Topology of simplicial complexes

2 - Singular homology

3 - Functoriality

III - Homological inference

1 - Thickening parameter selection

2 - Čech complex

3 - Rips complex

O grupo $\mathbb{Z}/2\mathbb{Z}$

The **group** $\mathbb{Z}/2\mathbb{Z}$ can be seen as the set $\{0, 1\}$ with the operation

$$\begin{aligned} 0+0&=0\\ 0+1&=1\\ 1+0&=1\\ 1+1&=0 \end{aligned}$$

For any $n \geq 1$, the **product group** $(\mathbb{Z}/2\mathbb{Z})^n$ is the group whose underlying set is

$$(\mathbb{Z}/2\mathbb{Z})^n = \{(\epsilon_1, ..., \epsilon_n),\ \epsilon_1, ..., \epsilon_n \in \mathbb{Z}/2\mathbb{Z}\}$$

and whose operation is defined as

$$(\epsilon_1, ..., \epsilon_n) + (\epsilon'_1, ..., \epsilon'_n) = (\epsilon_1 + \epsilon'_1, ..., \epsilon_n + \epsilon'_n).$$

The group $\mathbb{Z}/2\mathbb{Z}$ can be given a **field** structure

$$\begin{aligned} 0\times 0&=0\\ 0\times 1&=0\\ 1\times 0&=0\\ 1\times 1&=1 \end{aligned}$$

and $(\mathbb{Z}/2\mathbb{Z})^n$ can be seen as a $\mathbb{Z}/2\mathbb{Z}$-**vector space** over the field $\mathbb{Z}/2\mathbb{Z}$.

Espaços vetoriais sobre $\mathbb{Z}/2\mathbb{Z}$

Definition: A **vector space** over $\mathbb{Z}/2\mathbb{Z}$ is a set V endowed with two operations

$$V \times V \longrightarrow V \qquad \mathbb{Z}/2\mathbb{Z} \times V \longrightarrow V$$
$$(u, v) \longmapsto u + v \qquad (\lambda, v) \longmapsto \lambda \cdot v$$

such that

(associativity) $\forall u, v, w \in V,\ (u + v) + w = u + (v + w)$,

(identity) $\exists 0 \in V, \forall v \in V,\ v + 0 = 0 + v = v$,

(inverse) $\forall v \in V, \exists w \in V,\ u + v = v + u = 0$,

(commutativity) $\forall u, v \in V,\ u + v = v + u$,

(compatibility of multiplication) $\forall \lambda, \mu \in \mathbb{Z}/2\mathbb{Z}, \forall v \in V, \lambda \cdot (\mu \cdot v) = (\lambda \times \mu) \cdot v$,

(scalar identity) $\forall v \in V, 1 \cdot v = v$,

(scalar distributivity) $\forall \mu, \nu \in \mathbb{Z}/2\mathbb{Z}, \forall v \in V,\ (\lambda + \nu) \cdot v = \lambda \cdot v + \nu \cdot v$,

(vector distributivity) $\forall \mu \in \mathbb{Z}/2\mathbb{Z}, \forall v, w \in V,\ \lambda \cdot (u + v) = \lambda \cdot v + \nu \cdot v$.

Espaços vetoriais sobre $\mathbb{Z}/2\mathbb{Z}$

Definition: A **vector space** over $\mathbb{Z}/2\mathbb{Z}$ is a set V endowed with two operations

$$V \times V \longrightarrow V \qquad \mathbb{Z}/2\mathbb{Z} \times V \longrightarrow V$$
$$(u,v) \longmapsto u+v \qquad (\lambda, v) \longmapsto \lambda \cdot v$$

such that

$$\begin{aligned}
\text{(associativity)}\ & \forall u,v,w \in V,\ \ (u+v)+w = u+(v+w),\\
\text{(identity)}\ & \exists 0 \in V, \forall v \in V,\ \ v+0=0+v=v,\\
\text{(inverse)}\ & \forall v \in V, \exists w \in V,\ \ u+v=v+u=0,\\
\text{(commutativity)}\ & \forall u,v \in V,\ \ u+v=v+u,\\
\text{(compatibility of multiplication)}\ & \forall \lambda,\mu \in \mathbb{Z}/2\mathbb{Z}, \forall v \in V, \lambda \cdot (\mu \cdot v) = (\lambda \times \mu)\cdot v,\\
\text{(scalar identity)}\ & \forall v \in V, 1 \cdot v = v,\\
\text{(scalar distributivity)}\ & \forall \mu,\nu \in \mathbb{Z}/2\mathbb{Z}, \forall v \in V,\ (\lambda+\nu)\cdot v = \lambda \cdot v + \nu \cdot v,\\
\text{(vector distributivity)}\ & \forall \mu \in \mathbb{Z}/2\mathbb{Z}, \forall v,w \in V,\ \lambda \cdot (u+v) = \lambda \cdot v + \nu \cdot v.
\end{aligned}$$

Proposition: Le $(V,+)$ be a commutative group.
It can be given a $\mathbb{Z}/2\mathbb{Z}$-vector space structure iff $\forall v \in V, v+v=0$.

Proposition: Let $(V,+,\cdot)$ be a finite $\mathbb{Z}/2\mathbb{Z}$-vector space. Then there exists $n \geq 0$ such that V has cardinal 2^n, and $(V,+,\cdot)$ is isomorphic to the vector space $(\mathbb{Z}/2\mathbb{Z})^n$.

Proof: Consequence of the theory of vector spaces.

Espaços vetoriais sobre $\mathbb{Z}/2\mathbb{Z}$

A **linear subspace** of $(V, +, \cdot)$ is a subset $W \subset V$ such that

$$\forall u, v \in W, u + v \in W \qquad \text{and} \qquad \forall v \in W, \forall \lambda \in \mathbb{Z}/2\mathbb{Z}, \lambda v \in W.$$

We define the following equivalence relation on V: for all $u, v \in V$,

$$u \sim v \iff u - v \in W.$$

Denote by V/W the quotient set of V under this relation. For any $v \in V$, one shows that the equivalence class of v is equal to $v + W = \{v + w \mid w \in W\}$.

One defines a group structure $\oplus$ on V/W as follows:

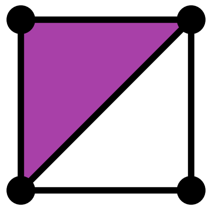
$$(u + W) \oplus (u' + W) = (u + u') + W.$$

Definition: The vector space $(V/W, \oplus, \cdot)$ is called the **quotient vector space**.

Proposition: We have $\dim V/W = \dim V - \dim W$.

Cadeias

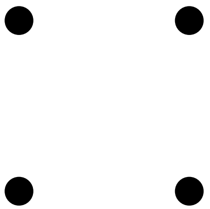
Definition (reminder): Let V be a set (called the set of *vertices*). A **simplicial complex** over V is a set K of subsets of V (called the *simplices*) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$.
The dimension of a simplex $\sigma \in K$ is $\dim(\sigma) = |\sigma| - 1$.



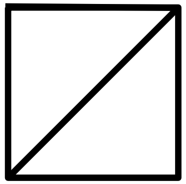
K

Let K be a simplicial complex. For any $n \geq 0$, define

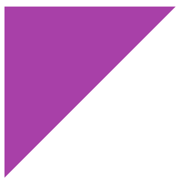
$$K_{(n)} = \{\sigma \in K \mid \dim(\sigma) = n\}.$$



$K_{(0)}$



$K_{(1)}$

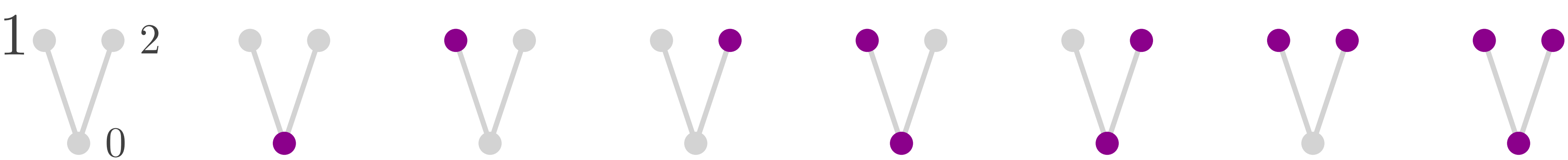


$K_{(2)}$

Let $n \geq 0$. The n-**chains** of K is the set $C_n(K)$ whose elements are the formal sums

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \quad \text{where} \quad \forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}.$$

Example: The 0-chains of $K = \{[0], [1], [2], [0,1], [0,2]\}$ are:



0 $\quad [0]$ $\quad [1]$ $\quad [2]$ $\quad [0]+[1]$ $\quad [0]+[2]$ $\quad [1]+[2]$ $\quad [0]+[1]+[2]$

and the 1-chains

0 $\quad [0,1]$ $\quad [0,2]$ $\quad [0,1]+[0,2]$

Cadeias

Let $n \geq 0$. The n-**chains** of K is the set $C_n(K)$ whose elements are the formal sums

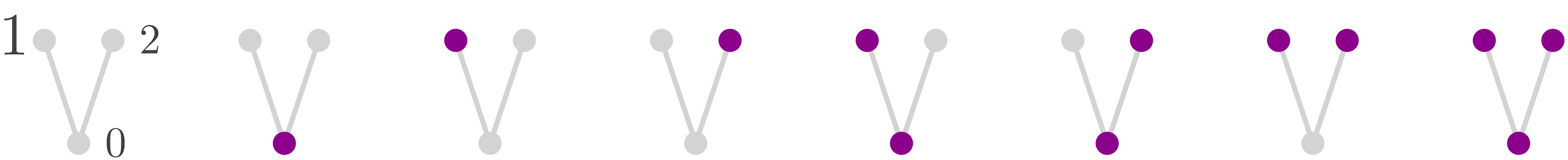
$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \quad \text{where} \quad \forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}.$$

We can give $C_n(K)$ a **group structure** via

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma + \sum_{\sigma \in K_{(n)}} \eta_\sigma \cdot \sigma = \sum_{\sigma \in K_{(n)}} (\epsilon_\sigma + \eta_\sigma) \cdot \sigma.$$

Moreover, $C_n(K)$ can be given a $\mathbb{Z}/2\mathbb{Z}$-vector space structure.

Example: The 0-chains of $K = \{[0], [1], [2], [0,1], [0,2]\}$ are:



0 $\quad [0]$ $\quad [1]$ $\quad [2]$ $\quad [0]+[1]$ $\quad [0]+[2]$ $\quad [1]+[2]$ $\quad [0]+[1]+[2]$

and the 1-chains

0 $\quad [0,1]$ $\quad [0,2]$ $\quad [0,1]+[0,2]$

Cadeias

Let $n \geq 0$. The n-**chains** of K is the set $C_n(K)$ whose elements are the formal sums

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \quad \text{where} \quad \forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}.$$

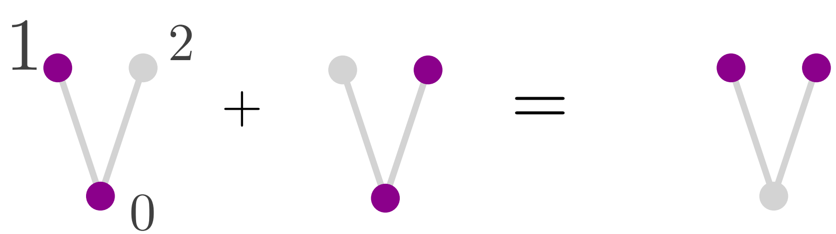
We can give $C_n(K)$ a **group structure** via

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma + \sum_{\sigma \in K_{(n)}} \eta_\sigma \cdot \sigma = \sum_{\sigma \in K_{(n)}} (\epsilon_\sigma + \eta_\sigma) \cdot \sigma.$$

Moreover, $C_n(K)$ can be given a $\mathbb{Z}/2\mathbb{Z}$-vector space structure.

Example: In the simplicial complex $K = \{[0], [1], [2], [0,1], [0,2]\}$, the sum of the 0-chains $[0] + [1]$ and $[0] + [2]$ is $[1] + [2]$:

$$([0] + [1]) + ([0] + [2]) = [0] + [0] + [1] + [2] = [1] + [2].$$



Operador bordo

Let $n \geq 1$, and $\sigma = [x_0, ..., x_n] \in K_{(n)}$ a simplex of dimension n. We define its **boundary** as the following element of $C_{n-1}(K)$:

$$\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$$

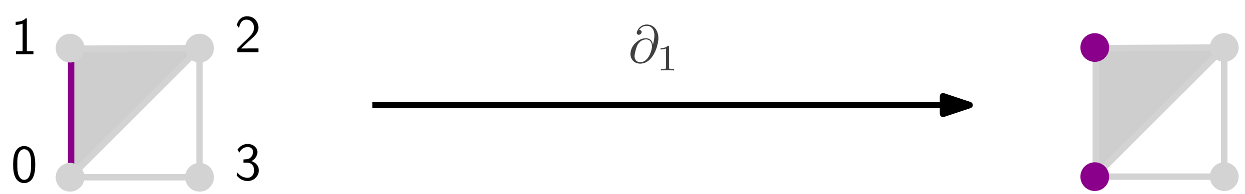
We can extend the operator ∂_n as a linear map $\partial_n \colon C_n(K) \to C_{n-1}(K)$.

Example: Consider the simplicial complex

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

The simplex $[0,1]$ has the faces $[0]$ and $[1]$. Hence

$$\partial_1 [0,1] = [0] + [1].$$



Operador bordo

Let $n \geq 1$, and $\sigma = [x_0, ..., x_n] \in K_{(n)}$ a simplex of dimension n. We define its **boundary** as the following element of $C_{n-1}(K)$:

$$\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$$

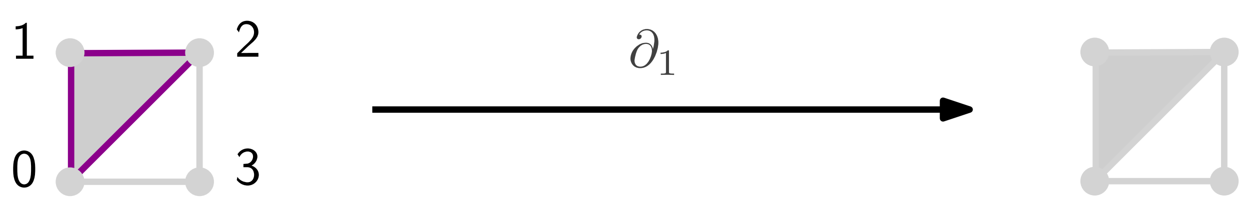
We can extend the operator ∂_n as a linear map $\partial_n \colon C_n(K) \to C_{n-1}(K)$.

Example: Consider the simplicial complex

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

The boundary of the 1-chain $[0,1] + [1,2] + [2,0]$ is

$$\begin{aligned} \partial_1\big([0,1] + [1,2] + [2,0]\big) &= \partial_1[0,1] + \partial_1[1,2] + \partial_1[2,0] \\ &= [0] + [1] + [1] + [2] + [2] + [0] = 0 \end{aligned}$$



Operador bordo

Let $n \geq 1$, and $\sigma = [x_0, ..., x_n] \in K_{(n)}$ a simplex of dimension n. We define its **boundary** as the following element of $C_{n-1}(K)$:

$$\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$$

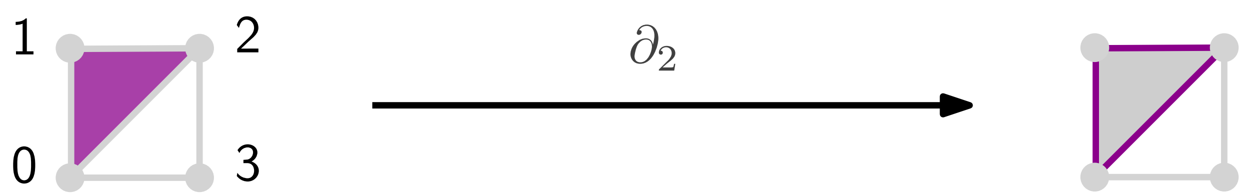
We can extend the operator ∂_n as a linear map $\partial_n \colon C_n(K) \to C_{n-1}(K)$.

Example: Consider the simplicial complex

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

The simplex $[0,1,2]$ has the faces $[0,1]$ and $[1,2]$ and $[2,0]$. Hence

$$\partial_2 [0,1,2] = [0,1] + [1,2] + [2,0].$$



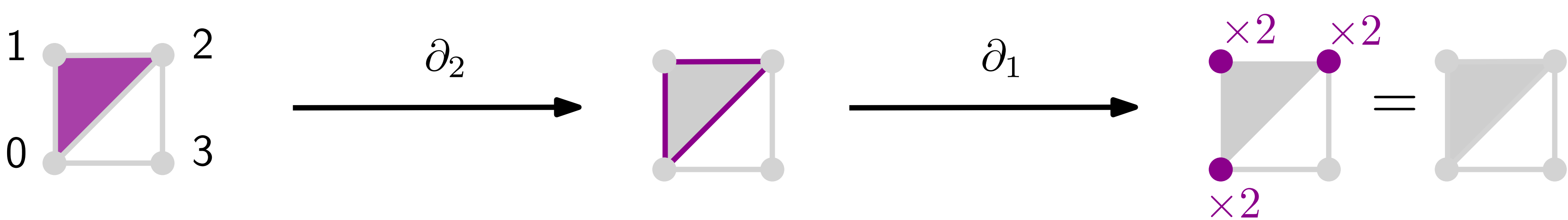
Operador bordo

Let $n \geq 1$, and $\sigma = [x_0, ..., x_n] \in K_{(n)}$ a simplex of dimension n. We define its **boundary** as the following element of $C_{n-1}(K)$:

$$\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$$

We can extend the operator ∂_n as a linear map $\partial_n : C_n(K) \to C_{n-1}(K)$.

Proposition: For any $n \geq 1$, for any $c \in C_n(K)$, we have $\partial_{n-1} \circ \partial_n(c) = 0$.



Operador bordo

Proposition: For any $n \geq 1$, for any $c \in C_n(K)$, we have $\partial_{n-1} \circ \partial_n(c) = 0$.

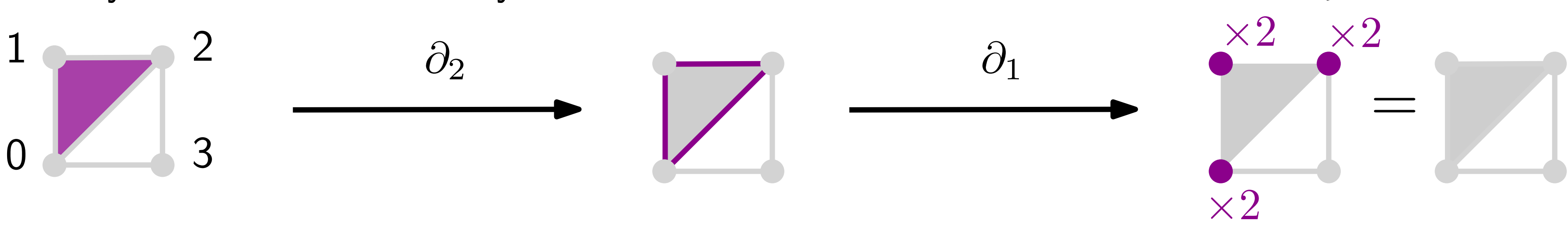
Proof: Suppose that $n \geq 2$, the result being trivial otherwise.
Since the boundary operators are linear, it is enough to prove that $\partial_{n-1} \circ \partial_n(\sigma) = 0$ for all simplex $\sigma \in K_{(n)}$.
By definition, $\partial_n(\sigma) = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$, and

$$\partial_{n-1} \circ \partial_n(\sigma) = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \partial_{n-1}(\tau) = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \sum_{\substack{\nu \subset \tau \\ |\nu| = |\tau| - 1}} \nu$$

We can write this last sum as

$$\sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \sum_{\substack{\nu \subset \tau \\ |\nu| = |\tau| - 1}} \nu = \sum_{\substack{\nu \subset \sigma \\ |\nu| = |\sigma| - 2}} \alpha_\nu \nu$$

where $\alpha_\nu = \{\tau \subset \sigma \mid |\tau| = |\sigma| - 1, \nu \subset \tau\}$.
It is easy to see that for every ν such that $\dim \nu = \dim \tau - 2$, we have $\alpha_\nu = 2 = 0$.



Ciclos e bordos

Let $n \geq 0$. We have a sequence of vector spaces

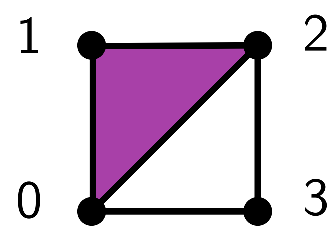
$$\cdots \longrightarrow C_{n+1}(K) \xrightarrow{\partial n+1} C_n(K) \xrightarrow{\partial n} C_{n-1}(K) \longrightarrow \dots$$

The maps ∂_{n+1} and ∂_n are linear maps, and we can consider their kernel and image.

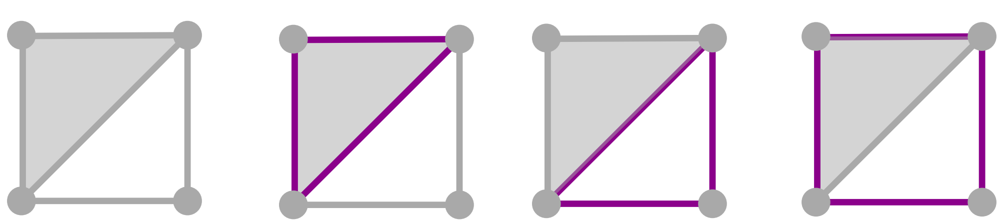
Definition: We define:

- The n-**cycles**: $Z_n(K) = \mathrm{Ker}(\partial_n) = \{c \in C_n(K) \mid \partial_n(c) = 0\}$,
- The n-**boundaries**: $B_n(K) = \mathrm{Im}(\partial_{n+1}) = \{\partial_{n+1}(c) \mid c \in C_{n+1}(K)\}$.

Example: Consider the simplicial complex



The 1-cycles are:



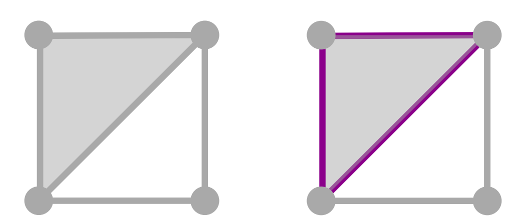
0

$[0,1]+[1,2]+[0,2]$

$[0,2]+[2,3]+[0,3]$

$[0,1]+[1,2]+[2,3]+[0,3]$.

The 1-boundaries are:



0

$[0,1]+[1,2]+[0,2]$

Ciclos e bordos

Let $n \geq 0$. We have a sequence of vector spaces

$$\cdots \longrightarrow C_{n+1}(K) \xrightarrow{\partial n+1} C_n(K) \xrightarrow{\partial n} C_{n-1}(K) \longrightarrow \ldots$$

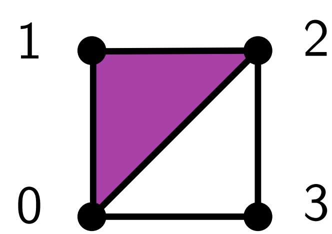
The maps ∂_{n+1} and ∂_n are linear maps, and we can consider their kernel and image.

Definition: We define:

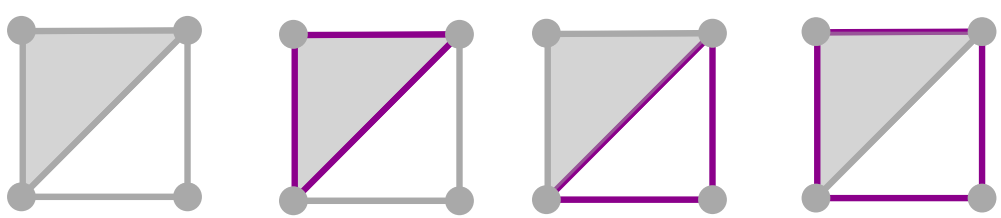
- The n-**cycles**: $Z_n(K) = \mathrm{Ker}(\partial_n) = \{c \in C_n(K) \mid \partial_n(c) = 0\}$,
- The n-**boundaries**: $B_n(K) = \mathrm{Im}(\partial_{n+1}) = \{\partial_{n+1}(c) \mid c \in C_{n+1}(K)\}$.

Proposition: We have $B_n(K) \subset Z_n(K)$.

Example: Consider the simplicial complex

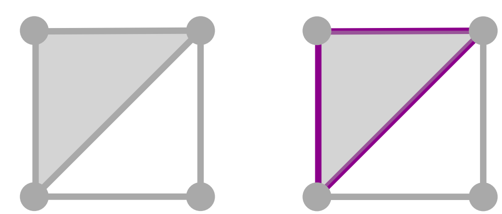


The 1-cycles are:



0 $\quad [0,1]+[1,2]+[0,2]$ $\quad [0,2]+[2,3]+[0,3]$ $\quad [0,1]+[1,2]+[2,3]+[0,3]$.

The 1-boundaries are:



0 $\quad [0,1]+[1,2]+[0,2]$

Ciclos e bordos

Let $n \geq 0$. We have a sequence of vector spaces

$$\cdots \longrightarrow C_{n+1}(K) \xrightarrow{\partial n+1} C_n(K) \xrightarrow{\partial n} C_{n-1}(K) \longrightarrow \ldots$$

The maps ∂_{n+1} and ∂_n are linear maps, and we can consider their kernel and image.

Definition: We define:

- The n-**cycles**: $Z_n(K) = \text{Ker}(\partial_n) = \{c \in C_n(K) \mid \partial_n(c) = 0\}$,
- The n-**boundaries**: $B_n(K) = \text{Im}(\partial_{n+1}) = \{\partial_{n+1}(c) \mid c \in C_{n+1}(K)\}$.

Proposition: We have $B_n(K) \subset Z_n(K)$.

Definition: We say that two chains $c, c' \in C_n(K)$ are **homologous** if there exists $b \in B_n(K)$ such that $c = c' + b$.

⟶ interpretation: two cycles are homologous if they represent the same 'hole'

Example:

$$[0,2] + [2,3] + [0,3] = [0,1] + [1,2] + [2,3] + [0,3] + [0,1] + [0,2] + [1,2].$$

Ciclos e bordos

Let $n \geq 0$. We have a sequence of vector spaces

$$\cdots \longrightarrow C_{n+1}(K) \xrightarrow{\partial n+1} C_n(K) \xrightarrow{\partial n} C_{n-1}(K) \longrightarrow \dots$$

The maps ∂_{n+1} and ∂_n are linear maps, and we can consider their kernel and image.

Definition: We define:

- The n-**cycles**: $Z_n(K) = \text{Ker}(\partial_n) = \{c \in C_n(K) \mid \partial_n(c) = 0\}$,
- The n-**boundaries**: $B_n(K) = \text{Im}(\partial_{n+1}) = \{\partial_{n+1}(c) \mid c \in C_{n+1}(K)\}$.

Proposition: We have $B_n(K) \subset Z_n(K)$.

Proof: Let $b \in B_n(K)$ be a boundary. By definition, there exists $c \in C_{n+1}(K)$ such that $b = \partial_{n+1}(c)$. Using $\partial_n \partial_{n+1} = 0$, we get

$$\partial_n(b) = \partial_n \partial_{n+1}(c) = 0,$$

hence $b \in Z_n(K)$.

We have defined a sequence of vector spaces, connected by linear maps

$$\cdots \longrightarrow C_{n+1}(K) \longrightarrow C_n(K) \longrightarrow C_{n-1}(K) \longrightarrow \cdots$$

and for every $n \geq 0$, we have defined the cycles and the boundaries $Z_n(K)$ and $B_n(K)$.

Since $B_n(K) \subset Z_n(K)$, we can see $B_n(K)$ as a linear subspace of $Z_n(K)$.

Definition: The $n^{\textbf{th}}$ **(simplicial) homology group** of K is the quotient vector space

$$H_n(K) = Z_n(K)/B_n(K).$$

Grupos de homologia

We have defined a sequence of vector spaces, connected by linear maps

$$\cdots \longrightarrow C_{n+1}(K) \longrightarrow C_n(K) \longrightarrow C_{n-1}(K) \longrightarrow \cdots$$

and for every $n \geq 0$, we have defined the cycles and the boundaries $Z_n(K)$ and $B_n(K)$.

Since $B_n(K) \subset Z_n(K)$, we can see $B_n(K)$ as a linear subspace of $Z_n(K)$.

Definition: The n^{th} **(simplicial) homology group** of K is the quotient vector space

$$H_n(K) = Z_n(K)/B_n(K).$$

Remark: A finite $\mathbb{Z}/2\mathbb{Z}$-vector space must be isomorphic to $(\mathbb{Z}/2\mathbb{Z})^k$ for some k.

Definition: Let K be a simplicial complex and $n \geq 0$. Its n^{th} **Betti number** is the integer $\beta_n(K) = \dim H_n(K)$.

$$H_n(K) = (\mathbb{Z}/2\mathbb{Z})^k \longrightarrow \beta_n(K) = k$$

Grupos de homologia

We have defined a sequence of vector spaces, connected by linear maps

$$\cdots \longrightarrow C_{n+1}(K) \longrightarrow C_n(K) \longrightarrow C_{n-1}(K) \longrightarrow \cdots$$

and for every $n \geq 0$, we have defined the cycles and the boundaries $Z_n(K)$ and $B_n(K)$.

Since $B_n(K) \subset Z_n(K)$, we can see $B_n(K)$ as a linear subspace of $Z_n(K)$.

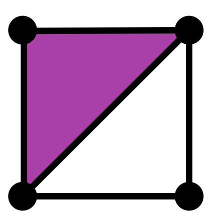
Definition: The n^{th} **(simplicial) homology group** of K is the quotient vector space

$$H_n(K) = Z_n(K)/B_n(K).$$

Remark: A finite $\mathbb{Z}/2\mathbb{Z}$-vector space must be isomorphic to $(\mathbb{Z}/2\mathbb{Z})^k$ for some k.

Definition: Let K be a simplicial complex and $n \geq 0$. Its n^{th} **Betti number** is the integer $\beta_n(K) = \dim H_n(K)$.

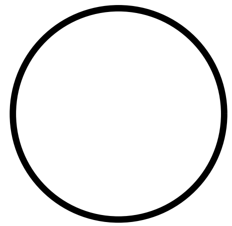
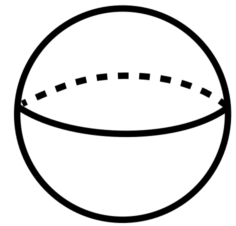
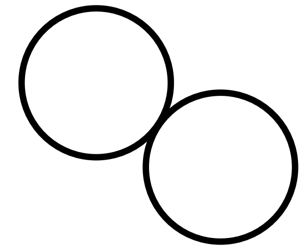
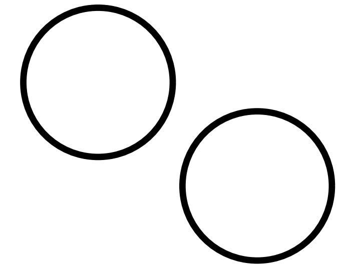
Example:



$H_0(K) = \mathbb{Z}/2\mathbb{Z} \longrightarrow \beta_0(K) = 1$

$H_1(K) = \mathbb{Z}/2\mathbb{Z} \longrightarrow \beta_1(K) = 1$

$H_2(K) = 0 \longrightarrow \beta_2(K) = 0$

X					
$H_0(X)$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\beta_0(X)$	1	1	1	1	2
$H_1(X)$	0	$\mathbb{Z}/2\mathbb{Z}$	0	$(\mathbb{Z}/2\mathbb{Z})^2$	$(\mathbb{Z}/2\mathbb{Z})^2$
$\beta_1(X)$	0	1	0	2	2
$H_2(X)$	0	0	$\mathbb{Z}/2\mathbb{Z}$	0	0
$\beta_2(X)$	0	0	1	0	0

Ordenar o complexo simplicial

Let K be a simplicial complex with n simplices. Choose a total order of the simplices

$$\sigma^1 < \sigma^2 < ... < \sigma^n$$

such that

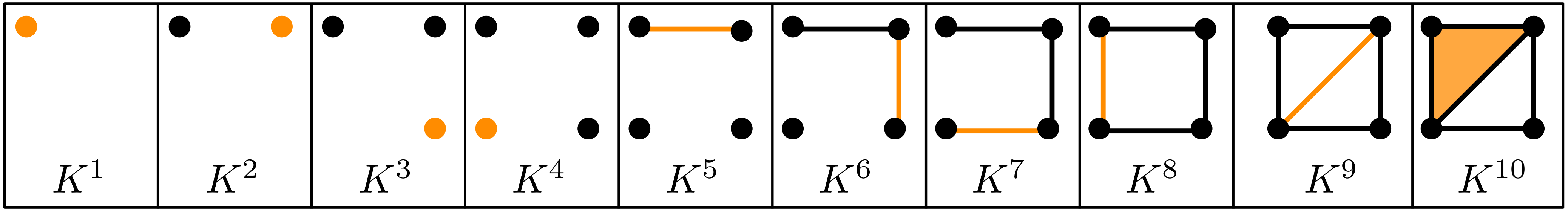
$$\forall \sigma, \tau \in K,\ \tau \subsetneq \sigma \implies \tau < \sigma.$$

In other words, a face of a simplex is lower than the simplex itself.
For every $i \leq n$, consider the simplicial complex

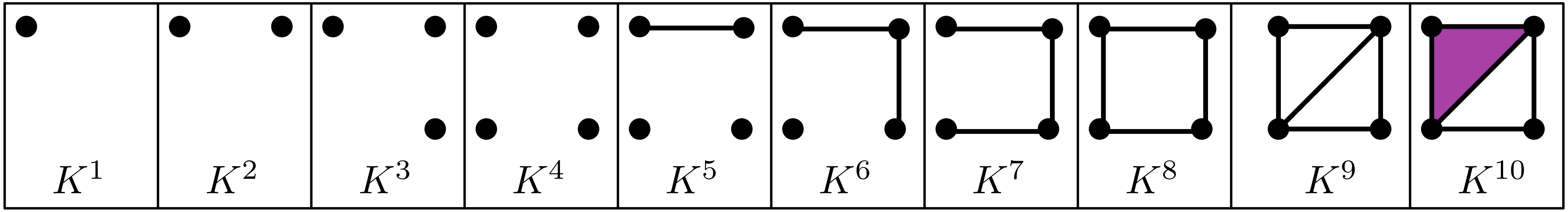
$$K^i = \{\sigma^1, ..., \sigma^i\}.$$

We have $\forall i \leq n, K^{i+1} = K^i \cup \left\{\sigma^{i+1}\right\}$, and $K^n = K$. They form an inscreasing sequence of simplicial complexes

$$K^1 \ \subset \ K^2 \ \subset \ ... \ \subset \ K^n.$$

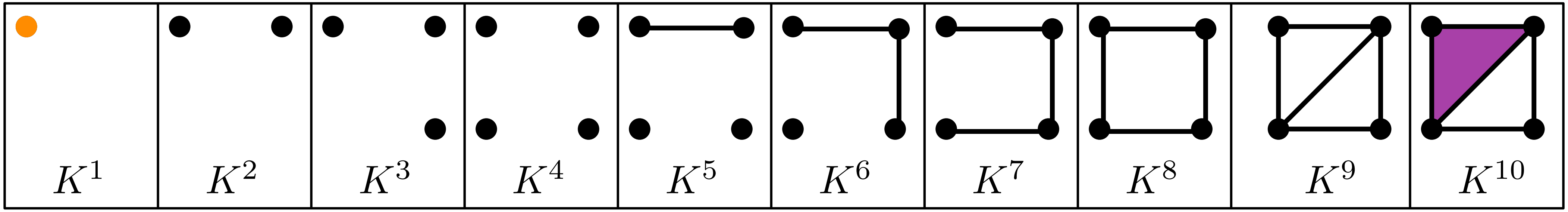


Positivity dos simplexos



Let $k \geq 0$. We will compute the homology groups of K^i incrementally:

$$H_k(K^1),\ H_k(K^2),\ H_k(K^3),\ H_k(K^4),\ H_k(K^5),\ H_k(K^6),\ H_k(K^7),\ H_k(K^8),\ H_k(K^9),\ H_k(K^{10})$$



Let $k \geq 0$. We will compute the homology groups of K^i incrementally:

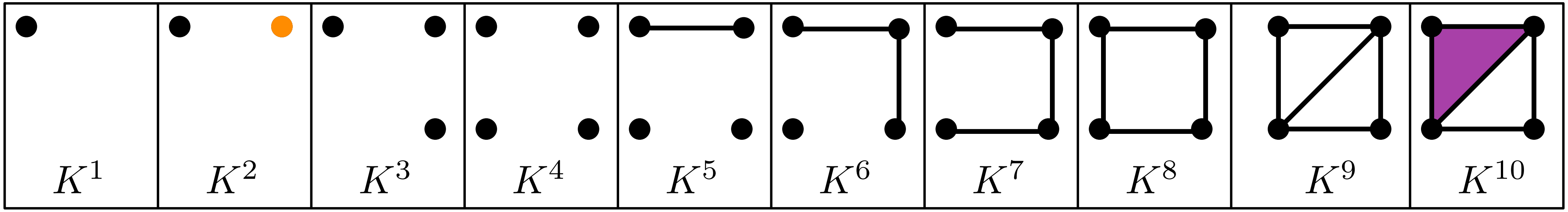
$$H_k(K^1),\ H_k(K^2),\ H_k(K^3),\ H_k(K^4),\ H_k(K^5),\ H_k(K^6),\ H_k(K^7),\ H_k(K^8),\ H_k(K^9),\ H_k(K^{10})$$

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma^i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ^i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ^i.
In other words, there exist $c = \sum_{\sigma \in K^i_{(n)}} \epsilon_\sigma \cdot \sigma \in C_n(K^i)$ such that $\epsilon_{\sigma^i} = 1$ and $\partial_n(c) = 0$. Otherwise, σ^i is **negative**.

Example:

- $\sigma^1 \in K^1$ is **positive** because it is included in the cycle $c = \sigma^1$ (indeed, $\partial_0(\sigma^1) = 0$).

Positividade dos simplexos



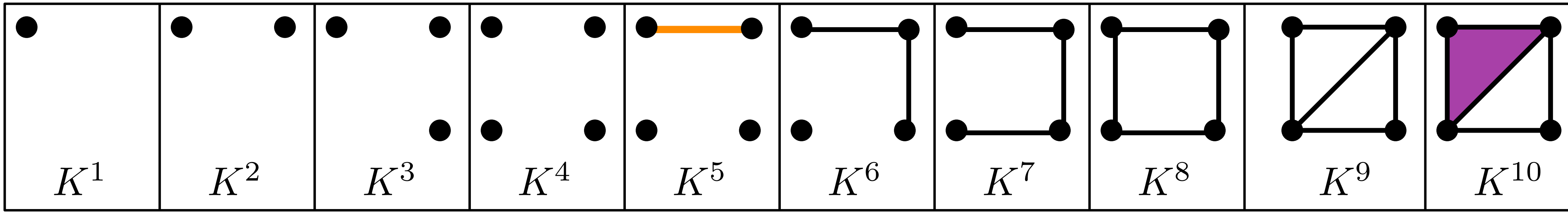
Let $k \geq 0$. We will compute the homology groups of K^i incrementally:

$$H_k(K^1),\ H_k(K^2),\ H_k(K^3),\ H_k(K^4),\ H_k(K^5),\ H_k(K^6),\ H_k(K^7),\ H_k(K^8),\ H_k(K^9),\ H_k(K^{10})$$

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma^i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ^i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ^i.
In other words, there exist $c = \sum_{\sigma \in K^i_{(n)}} \epsilon_\sigma \cdot \sigma \in C_n(K^i)$ such that $\epsilon_{\sigma^i} = 1$ and $\partial_n(c) = 0$. Otherwise, σ^i is **negative**.

Example:

- $\sigma^1 \in K^1$ is **positive** because it is included in the cycle $c = \sigma^1$ (indeed, $\partial_0(\sigma^1) = 0$).
- $\sigma^2 \in K^2$ is **positive** because it is included in the cycle $c = \sigma^2$ (indeed, $\partial_0(\sigma^2) = 0$).



Let $k \geq 0$. We will compute the homology groups of K^i incrementally:

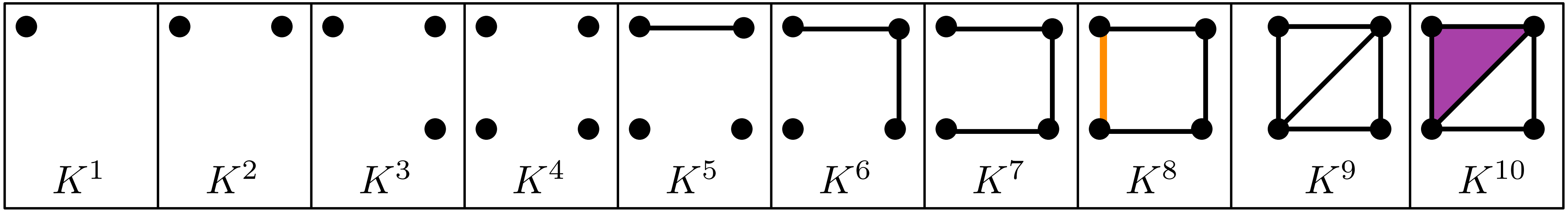
$$H_k(K^1),\ H_k(K^2),\ H_k(K^3),\ H_k(K^4),\ H_k(K^5),\ H_k(K^6),\ H_k(K^7),\ H_k(K^8),\ H_k(K^9),\ H_k(K^{10})$$

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma^i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ^i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ^i.
In other words, there exist $c = \sum_{\sigma \in K^i_{(n)}} \epsilon_\sigma \cdot \sigma \in C_n(K^i)$ such that $\epsilon_{\sigma^i} = 1$ and $\partial_n(c) = 0$. Otherwise, σ^i is **negative**.

Example:

- $\sigma^1 \in K^1$ is **positive** because it is included in the cycle $c = \sigma^1$ (indeed, $\partial_0(\sigma^1) = 0$).
- $\sigma^2 \in K^2$ is **positive** because it is included in the cycle $c = \sigma^2$ (indeed, $\partial_0(\sigma^2) = 0$).
- $\sigma^6 \in K^5$ is **negative** because it is not included in a cycle $Z_1(K^5)$. Indeed, $C_1(K^5)$ only contains 0 and σ_5, and $\partial_1(\sigma^5) = \sigma^1 + \sigma^2 \neq 0$.

Positividade dos simplexos



Let $k \geq 0$. We will compute the homology groups of K^i incrementally:

$$H_k(K^1),\ H_k(K^2),\ H_k(K^3),\ H_k(K^4),\ H_k(K^5),\ H_k(K^6),\ H_k(K^7),\ H_k(K^8),\ H_k(K^9),\ H_k(K^{10})$$

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma^i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ^i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ^i.
In other words, there exist $c = \sum_{\sigma \in K^i_{(n)}} \epsilon_\sigma \cdot \sigma \in C_n(K^i)$ such that $\epsilon_{\sigma^i} = 1$ and $\partial_n(c) = 0$. Otherwise, σ^i is **negative**.

Example:

- $\sigma^1 \in K^1$ is **positive** because it is included in the cycle $c = \sigma^1$ (indeed, $\partial_0(\sigma^1) = 0$).
- $\sigma^2 \in K^2$ is **positive** because it is included in the cycle $c = \sigma^2$ (indeed, $\partial_0(\sigma^2) = 0$).
- $\sigma^6 \in K^5$ is **negative** because it is not included in a cycle $Z_1(K^5)$. Indeed, $C_1(K^5)$ only contains 0 and σ_5, and $\partial_1(\sigma^5) = \sigma^1 + \sigma^2 \neq 0$.
- $\sigma^8 \in K^8$ is **positive** because it is included in the cycle $c = \sigma^5 + \sigma^6 + \sigma^7 + \sigma^8$ (indeed, $\partial_1(c) = 2\sigma^1 + 2\sigma^2 + 2\sigma^3 + 2\sigma^4 = 0$).

Positividade dos simplexos

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma_i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ_i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ_i.
Otherwise, σ_i is **negative**.

Remark: By adding σ^i in the simplicial complex, the only groups that may change are $Z_d(K^i)$ and $B_{d-1}(K^i)$.

Positividade dos simplexos

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma_i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ_i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ_i.
Otherwise, σ_i is **negative**.

Remark: By adding σ^i in the simplicial complex, the only groups that may change are $Z_d(K^i)$ and $B_{d-1}(K^i)$.

Lemma: If σ^i is positive, then $\beta_d(K^i) = \beta_d(K^{i-1}) + 1$,
and for all $d' \neq d$, $\beta_{d'}(K^i) = \beta_{d'}(K^{i-1})$.

Proof: We start by proving the following fact: if $c \in Z_d(K^i)$ is a cycle that contains σ_i, then c is not homologous (in K^i) to a cycle of $c' \in Z_d(K^{i-1})$.

By contradiction: if $c = c' + b$ with $c' \in Z_d(K^{i-1})$ and $b \in B_d(K^i)$, then $c - c' = b \in B_d(K^i)$. This is absurd because we just added σ_i: it cannot appear in a boundary of K^i.

As a consequence, $\dim Z_d(K^i) = \dim Z_d(K^{i-1}) + 1$.

We conclude by using the relation $\beta_d(K^i) = \dim Z_d(K^i) - \dim B_d(K^i)$.

Positividade dos simplexos

Definition: Let $i \in [\![1, n]\!]$, and $d = \dim(\sigma_i)$. Recall that $K^i = K^{i-1} \cup \{\sigma_i\}$.
The simplex σ_i is **positive** if there exists a cycle $c \in Z_d(K^i)$ that contains σ_i.
Otherwise, σ_i is **negative**.

Remark: By adding σ^i in the simplicial complex, the only groups that may change are $Z_d(K^i)$ and $B_{d-1}(K^i)$.

Lemma: If σ^i is positive, then $\beta_d(K^i) = \beta_d(K^{i-1}) + 1$,
and for all $d' \neq d$, $\beta_{d'}(K^i) = \beta_{d'}(K^{i-1})$.

Lemma: If σ^i is negative, then $\beta_{d-1}(K^i) = \beta_{d-1}(K^{i-1}) - 1$,
and for all $d' \neq d-1$, $\beta_{d'}(K^i) = \beta_{d'}(K^{i-1})$.

Proof: We start by proving the following fact: $\partial_d(\sigma^i)$ is not a boundary of K^{i-1}.

Otherwise, we would have $\partial_d(\sigma^i) = \partial_d(c)$ with $c \in C_d(K^{i-1})$, i.e. $\partial_d(\sigma^i + c) = 0$.
Hence $\sigma^i + c$ would be a cycle of K^i that contains c, contradicting the negativity of σ^i.

As a consequence, $\dim B_{d-1}(K^i) = \dim B_{d-1}(K^{i-1}) + 1$.

We conclude by using the relation $\beta_{d-1}(K^i) = \dim Z_{d-1}(K^i) - \dim B_{d-1}(K^i)$.

Algoritmo incremental

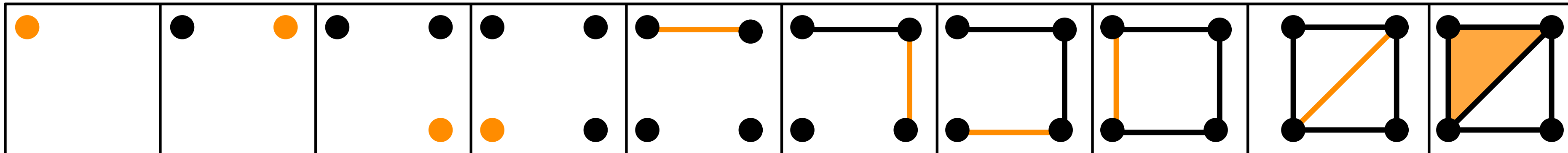
Lemma: If σ^i is positive, then $\beta_d(K^i) = \beta_d(K^{i-1}) + 1$, and for all $d' \neq d$, $\beta_{d'}(K^i) = \beta_{d'}(K^{i-1})$.

Lemma: If σ^i is negative, then $\beta_{d-1}(K^i) = \beta_{d-1}(K^{i-1}) - 1$, and for all $d' \neq d-1$, $\beta_{d'}(K^i) = \beta_{d'}(K^{i-1})$.

We deduce the following algorithm:

Input: an increasing sequence of simplicial complexes $K^1 \subset \cdots \subset K^n = K$
Output: the Betti numbers $\beta_0(K), ...\beta_d(K)$
$\beta_0 \leftarrow 0, ..., \beta_d \leftarrow 0$;
for $i \leftarrow 1$ **to** n **do**
  $d = \dim(\sigma^i)$;
  if σ^i *is positive* **then**
    $\beta_k(K^i) \leftarrow \beta_k(K^i) + 1$;
  else if $d > 0$ **then**
    $\beta_{k-1}(K^i) \leftarrow \beta_{k-1}(K^{i-1}) - 1$;

Algoritmo incremental



	K^1	K^2	K^3	K^4	K^5	K^6	K^7	K^8	K^9	K^{10}
Dimension	0	0	0	0	1	1	1	1	1	2
Positivity	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$+$	$+$	$-$
$\beta_0(K^i)$	1	2	3	4	3	2	1	1	1	1
$\beta_1(K^i)$	0	0	0	0	0	0	0	1	2	1

We deduce the following algorithm:

Input: an increasing sequence of simplicial complexes $K^1 \subset \cdots \subset K^n = K$
Output: the Betti numbers $\beta_0(K), ...\beta_d(K)$
$\beta_0 \leftarrow 0, ..., \beta_d \leftarrow 0$;
for $i \leftarrow 1$ **to** n **do**
  $d = \dim(\sigma^i)$;
  if σ^i *is positive* **then**
    $\beta_k(K^i) \leftarrow \beta_k(K^i) + 1$;
  else if $d > 0$ **then**
    $\beta_{k-1}(K^i) \leftarrow \beta_{k-1}(K^{i-1}) - 1$;

Característica de Euler

Reminder: the Euler characteristic of a simplicial complex K is

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

Proposition: The Euler characteristic is also equal to

$$\chi(K) = \sum_{0 \leq i \leq n} (-1)^i \cdot \beta_i(K).$$

Característica de Euler

Proposition: The Euler characteristic of K is equal to

$$\chi(K) = \sum_{0 \leq i \leq n} (-1)^i \cdot \beta_i(K).$$

Proof: Pick an ordering $K^1 \subset \cdots \subset K^n = K$ of K, with $K^i = K^{i-1} \cup \{\sigma^i\}$ for all $2 \leq i \leq n$.

By induction, let us show that, for all $1 \leq m \leq n$,

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

For $m = 1$, σ^m is a 0-simplex, and the equality reads $1 = 1$.

Now, suppose that the equality is true for $1 \leq m < n$, and consider the simplex σ^{m+1}. Let $d = \dim \sigma^{m+1}$. The right-hand side of the Equation is increased by $(-1)^d$.

If σ^{m+1} is positive, then $\beta_d(K^{m+1}) = \beta_d(K^m) + 1$, hence the left-hand side of the Equation is increased by $(-1)^d$.

Otherwise, it is negative, and $\beta_{d-1}(K^{m+1}) = \beta_{d-1}(K^m) - 1$, hence the left-hand side of the Equation is increased by $-(-1)^{d-1} = (-1)^d$.

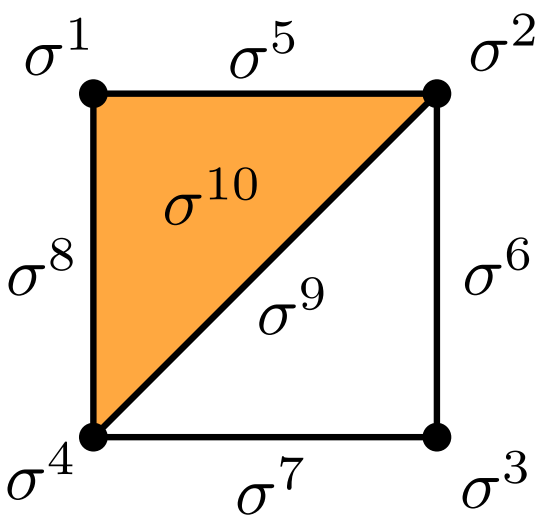
Matriz de bordo

The only thing missing to apply the incremental algorithm is to determine whether a simplex is positive or negative.

Let K be a simplicial complex, and $\sigma^1 < \sigma^2 < \cdots < \sigma^n$ and ordering of its simplices.

Define the **boundary matrix** of K, denoted Δ, as follows: Δ is a $n \times n$ matrix, whose (i, j)-entry (i^{th} row, j^{th} column is)

$$\Delta_{i,j} = \begin{cases} 1 & \text{if } \sigma^i \text{ is a face of } \sigma^j \text{ and } |\sigma^i| = |\sigma^j| - 1 \\ 0 & \text{else.} \end{cases}$$



	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	1	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

By adding columns one to the others, we create chains.
If we were able to reduce a column to zero, then we found a cycle.

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	1	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

$$\partial_1(\sigma^6) = \sigma^2 + \sigma^3$$

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

$$\partial_1(\sigma^5 + \sigma^6 + \sigma^7 + \sigma^8) = 0$$

Matriz de bordo

The process of reducing columns to zero is called **Gauss reduction**.
For any $j \in [\![1, n]\!]$, define
$$\delta(j) = \max\{i \in [\![1, n]\!] \mid \Delta_{i,j} \neq 0\}.$$

If $\Delta_{i,j} = 0$ for all j, then $\delta(j)$ is *undefined*.

We say that the boundary matrix Δ is *reduced* if the map δ is injective on its domain of definition.

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

Algorithm 2: Reduction of the boundary matrix

Input: a boundary matrix Δ
Output: a reduced matrix $\widetilde{\Delta}$
for $j \leftarrow 1$ **to** n **do**
  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**
    add column i to column j;

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	1	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8 + \sigma^7$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	1	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

Algorithm 2: Reduction of the boundary matrix

Input: a boundary matrix Δ
Output: a reduced matrix $\widetilde{\Delta}$
for $j \leftarrow 1$ **to** n **do**
  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**
    add column i to column j;

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8+\sigma^7$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	1	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8+\sigma^7+\sigma^6$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	1	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

Algorithm 2: Reduction of the boundary matrix

Input: a boundary matrix Δ

Output: a reduced matrix $\widetilde{\Delta}$

for $j \leftarrow 1$ **to** n **do**

  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**

    add column i to column j;

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8+\sigma^7+\sigma^6$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	1	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8+\sigma^7+\sigma^6+\sigma^5$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

Algorithm 2: Reduction of the boundary matrix

Input: a boundary matrix Δ

Output: a reduced matrix $\widetilde{\Delta}$

for $j \leftarrow 1$ **to** n **do**

  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**

    add column i to column j;

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^8 + \sigma^7 + \sigma^6 + \sigma^5$	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5 + \sigma^6 + \sigma^7 + \sigma^8$	$\sigma^9 + \sigma^6 + \sigma^7$	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Matriz de bordo

Algorithm 2: Reduction of the boundary matrix

Input: a boundary matrix Δ
Output: a reduced matrix $\widetilde{\Delta}$
for $j \leftarrow 1$ **to** n **do**
  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**
    add column i to column j;

Lemma: Suppose that the boundary matrix is reduced. Let $j \in [\![1, n]\!]$.
If $\delta(j)$ is defined, then the simplex σ^j is negative.
Otherwise, it is positive.

$$
\begin{array}{c}
 \\
\sigma^1\\ \sigma^2\\ \sigma^3\\ \sigma^4\\ \sigma^5\\ \sigma^6\\ \sigma^7\\ \sigma^8\\ \sigma^9\\ \sigma^{10}\\
 \\

\end{array}
\begin{array}{cccccccccc}
\sigma^1 & \sigma^2 & \sigma^3 & \sigma^4 & \sigma^5 & \sigma^6 & \sigma^7 & \sigma^5+\sigma^6+\sigma^7+\sigma^8 & \sigma^9+\sigma^6+\sigma^7 & \sigma^{10}\\
0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&\textcircled{1}&1&0&0&0&0\\
0&0&0&0&0&\textcircled{1}&1&0&0&0\\
0&0&0&0&0&0&\textcircled{1}&0&0&0\\
0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&\textcircled{1}\\
0&0&0&0&0&0&0&0&0&0\\
\sigma^1 & \sigma^2 & \sigma^3 & \sigma^4 & \textcircled{σ^5} & \textcircled{σ^6} & \textcircled{σ^7} & \sigma^8 & \sigma^9 & \textcircled{σ^{10}}\\
+&+&+&+&-&-&-&+&+&-
\end{array}
$$

Algoritmo final

Incremental computation of the homology

Input: an increasing sequence of simplicial complexes $K^1 \subset \cdots \subset K^n = K$
Output: the Betti numbers $\beta_0(K), ...\beta_d(K)$
$\beta_0 \leftarrow 0, ..., \beta_d \leftarrow 0;$
for $i \leftarrow 1$ **to** n **do**
  $d = \dim(\sigma^i);$
  if σ^i *is positive* **then**
    $\beta_k(K^i) \leftarrow \beta_k(K^i) + 1;$
  else if $d > 0$ **then**
    $\beta_{k-1}(K^i) \leftarrow \beta_{k-1}(K^{i-1}) - 1;$

Gauss reduction of the boundary matrix

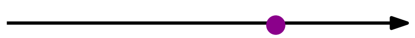
Input: a boundary matrix Δ
Output: a reduced matrix $\widetilde{\Delta}$
for $i \leftarrow 1$ jo n **do**
  while *there exists* $i < j$ *with* $\delta(i) = \delta(j)$ **do**
    add column i to column j;

Simplexo padrão

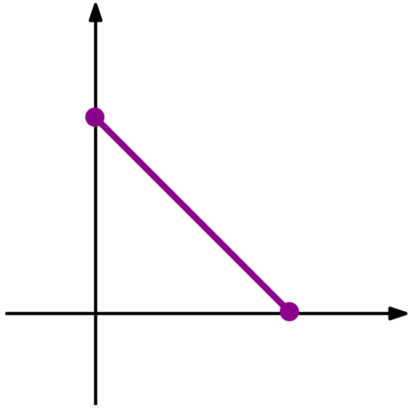
In order to describe topological spaces, we will decompose them into simpler pieces. The pieces we shall consider are the standard simplices.

The **standard simplex of dimension** n is the following subset of $\mathbb{R}^{n+1}$

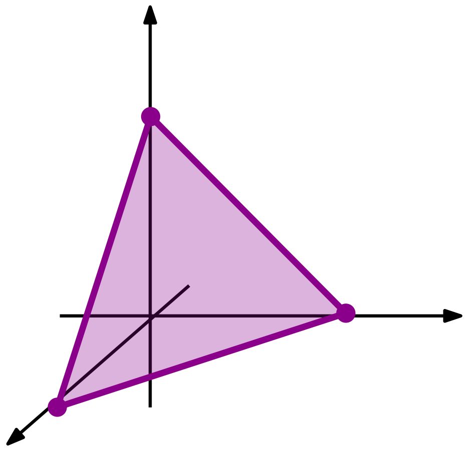
$$\Delta_n = \{x = (x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1, ..., x_{n+1} \geq 0 \text{ and } x_1 + ... + x_{n+1} = 1\}$$



Δ_0



Δ_1



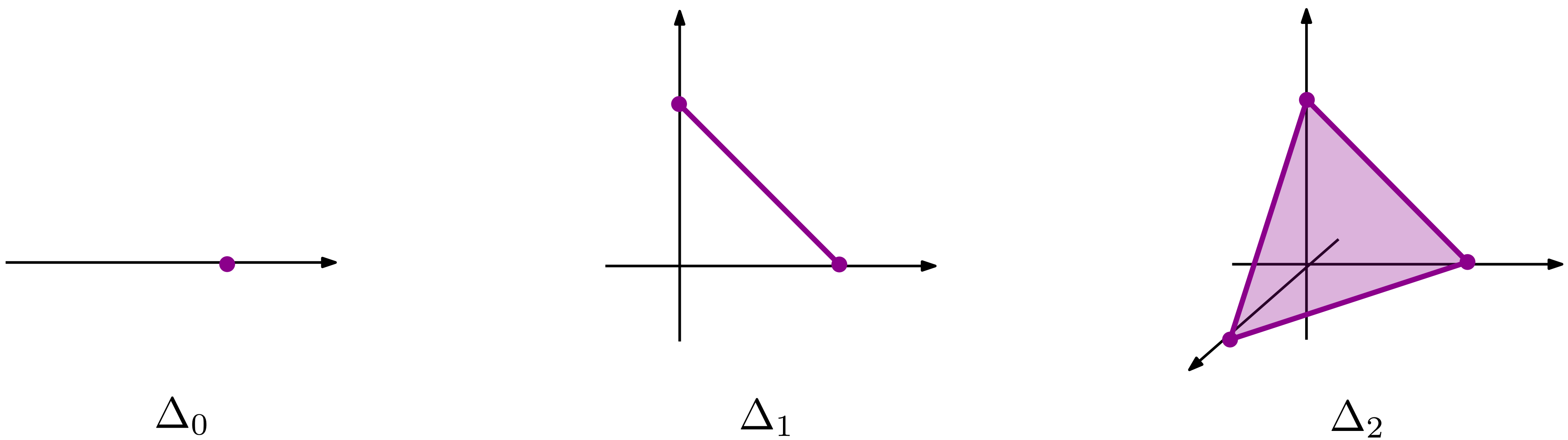
Δ_2

Simplexo padrão

In order to describe topological spaces, we will decompose them into simpler pieces. The pieces we shall consider are the standard simplices.

The **standard simplex of dimension** n is the following subset of $\mathbb{R}^{n+1}$

$$\Delta_n = \{x = (x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1, ..., x_{n+1} \geq 0 \text{ and } x_1 + ... + x_{n+1} = 1\}$$



Remark: For any collection of points $a_1, ..., a_k \in \mathbb{R}^n$, their convex hull is defined as:

$$\text{conv}(\{a_1 ... a_k\}) = \left\{\sum_{1 \leq i \leq k} t_i a_i \mid t_1 + ... + t_k = 1, \quad t_1, ..., t_k \geq 0\right\}.$$

We can say that Δ_n is the convex hull of the vectors $e_1, ..., e_{n+1}$ of $\mathbb{R}^{n+1}$, where

$$e_i = (0, ..., 1, 0, ..., 0) \qquad (i^{\text{th}} \text{ coordinate } 1, \text{ the other ones } 0).$$

Realização topológica

Let us give simplicial complexes a topology.

Definition: Let K be a simplicial complex, with vertex $V = \{1, ..., n\}$.
In $\mathbb{R}^n$, consider, for every $i \in [\![1, n]\!]$, the vector $e_i = (0, ..., 1, 0, ..., 0)$ (i^{th} coordinate 1, the other ones 0).
Let $|K|$ be the subset of $\mathbb{R}^n$ defined as:

$$|K| = \bigcup_{\sigma \in K} \text{conv}\left(\{e_j, j \in \sigma\}\right)$$

where conv represent the convex hull of points.
Endowed with the subspace topology, $(|K|, \mathcal{T}_{||K|})$ is a topological space, that we call the *topological realization of* K.

If $a_1, ..., a_k \in \mathbb{R}^n$, the convex hull is defined as:

$$\text{conv}(\{a_1 ... a_k\}) = \left\{\sum_{1 \leq i \leq k} t_i a_i \mid t_1 + ... + t_k = 1, \;\; t_1, ..., t_k \geq 0\right\}.$$

Realização topológica

Let us give simplicial complexes a topology.

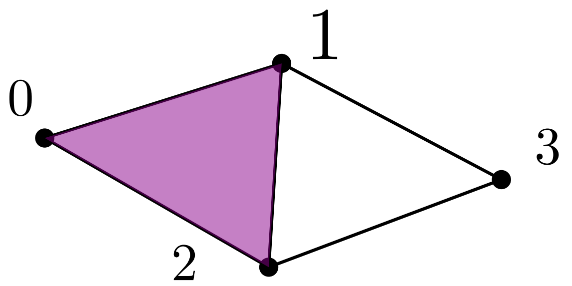
Definition: Let K be a simplicial complex, with vertex $V = \{1, ..., n\}$.
In $\mathbb{R}^n$, consider, for every $i \in [\![1, n]\!]$, the vector $e_i = (0, ..., 1, 0, ..., 0)$ (i^{th} coordinate 1, the other ones 0).
Let $|K|$ be the subset of $\mathbb{R}^n$ defined as:

$$|K| = \bigcup_{\sigma \in K} \text{conv}\left(\{e_j, j \in \sigma\}\right)$$

where conv represent the convex hull of points.
Endowed with the subspace topology, $(|K|, \mathcal{T}_{||K|})$ is a topological space, that we call the *topological realization of K*.

Remark: If the simplicial complex can be drawn in the plane (or space) without crossing itself, then its topological realization simply is the subspace topology.

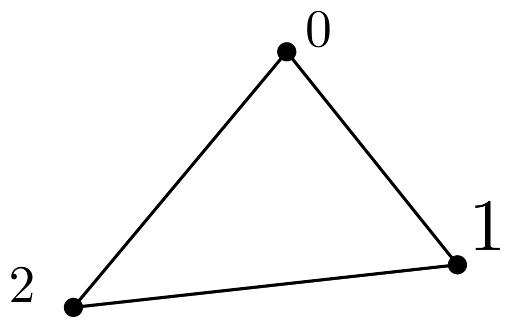
Example: $K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 0], [1, 3], [2, 3], [0, 1, 2]\}$.



Definition: Let X be a topological space. A **triangulation** of X is a simplicial complex K such that its topological realization $|K|$ is homeomorphic to X.

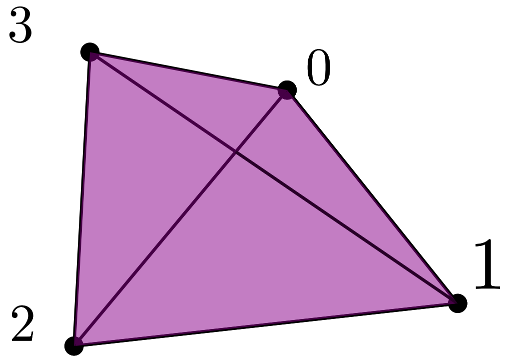
Example: The following simplicial complex is a triangulation of the circle:

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



Example: The following simplicial complex is a triangulation of the sphere:

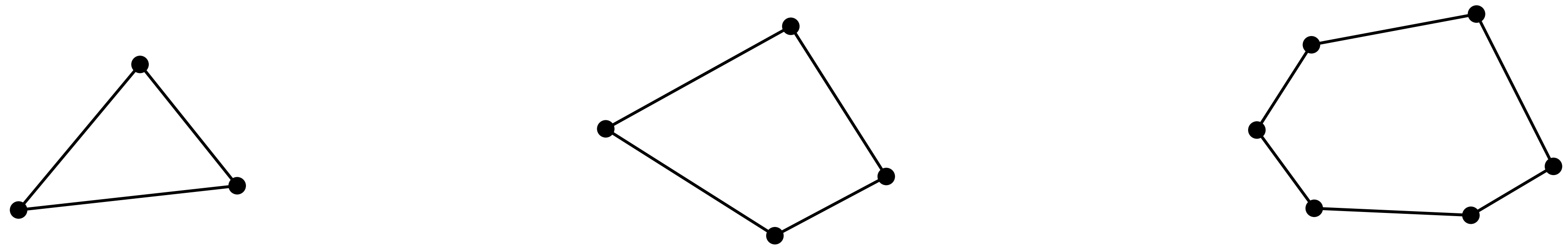
$$K = \{[0], [1], [2], [3], [0,1], [1,2], [2,3], [3,0], [0,2], [1,3], [0,1,2], [0,1,3], [0,2,3], [1,2,3]\}.$$



Triangulações

Definition: Let X be a topological space. A **triangulation** of X is a simplicial complex K such that its topological realization $|K|$ is homeomorphic to X.

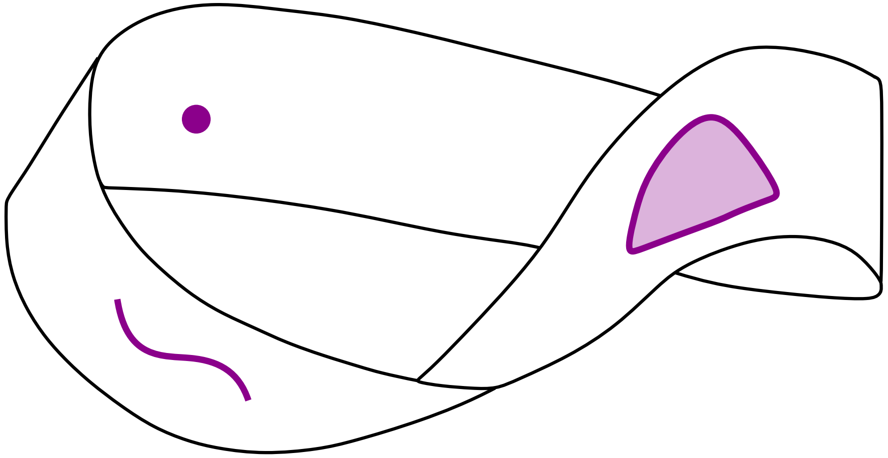
Given a topological space, it is not always possible to triangulate it.
However, when it is, there exists many different triangulations.



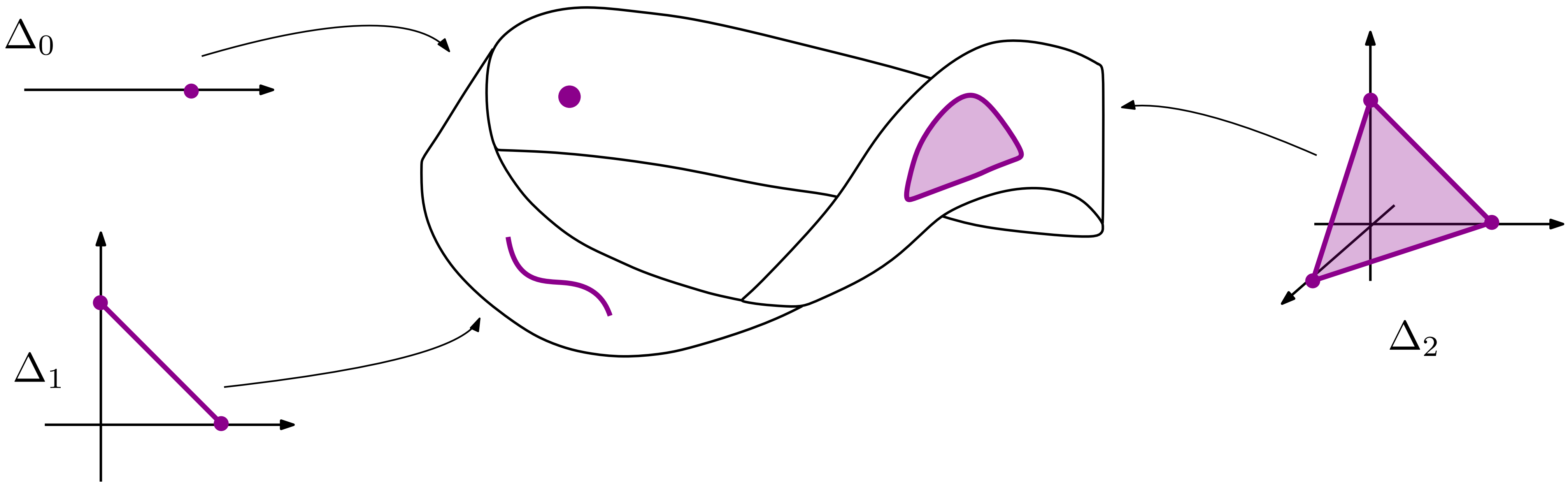
Theorem (Manolescu, 2016): For any dimension $n \geq 5$ there is a compact topological manifold which does not admit a triangulation.

Simplexo singular

Let us consider a **topological space** X. We want a notion of **simplices**.



Let us consider a **topological space** X. We want a notion of **simplices**.

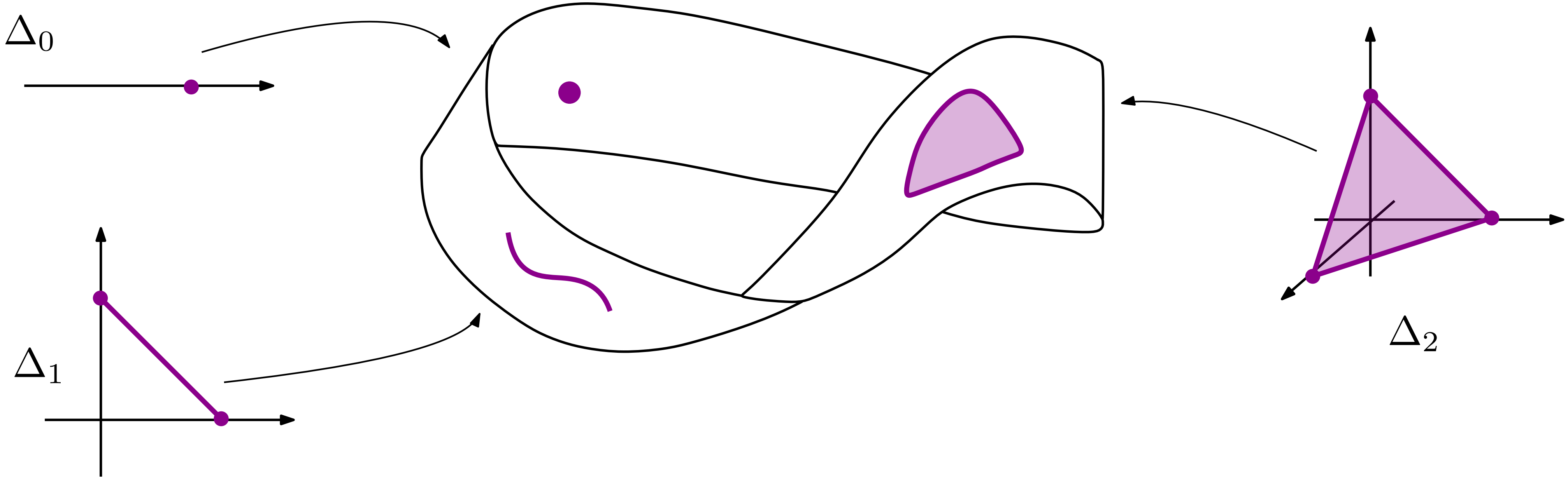


Definition: A **singular** n-**simplex** is a continuous map $\Delta_n \to X$, where Δ_n is the standard n-simplex. We denote S_n their set.

We now want a notion of **boundary**.

Simplexo singular

Let us consider a **topological space** X. We want a notion of **simplices**.



Definition: A **singular** n-**simplex** is a continuous map $\Delta_n \to X$, where Δ_n is the standard n-simplex. We denote S_n their set.

We now want a notion of **boundary**.

The boundary of Δ_n consists in $n+1$ copies of Δ_{n-1}.

We can restrict a singular n-simplex $\Delta_n \to X$ to the boundaries, giving $n+1$ singular $(n-1)$-simplices $\Delta_{n-1} \to X$.

Definition: The **boundary** of a singular n-simplex $\Delta_n \to X$ is the formal sum of the $n+1$ singular $(n-1)$-simplices $\Delta_{n-1} \to X$

Homologia singular

For a **simplicial complex** K, we have defined

n-chains $\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma$ where $\forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}$

boundary operator $\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$

chain complex $\dots \xrightarrow{\partial_{n+2}} C_{n+1}(K) \xrightarrow{\partial_{n+1}} C_n(K) \xrightarrow{\partial_n} C_{n-1}(K) \xrightarrow{\partial_{n-1}} \dots$

n-cycles and n-boundaries $Z_n(K) = \text{Ker}(\partial_n)$ $B_n(K) = \text{Im}(\partial_{n+1})$

n^{th}**simplicial** homology group $H_n(K) = Z_n(K)/B_n(K)$

Homologia singular

For a **simplicial complex** K, we have defined

n-chains $\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma$ where $\forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}$

boundary operator $\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$

chain complex $\dots \xrightarrow{\partial_{n+2}} C_{n+1}(K) \xrightarrow{\partial_{n+1}} C_n(K) \xrightarrow{\partial_n} C_{n-1}(K) \xrightarrow{\partial_{n-1}} \dots$

n-cycles and n-boundaries $Z_n(K) = \mathrm{Ker}(\partial_n)$ $B_n(K) = \mathrm{Im}(\partial_{n+1})$

n^{th}**simplicial** homology group $H_n(K) = Z_n(K)/B_n(K)$

For a **topological space** X, we can define

n-chains $\sum_{\sigma \in S_n} \epsilon_\sigma \cdot \sigma$ where $\forall \sigma \in S_n,\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}$

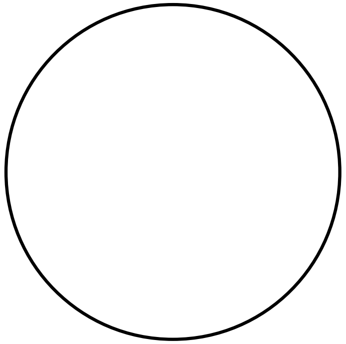
boundary operator $\partial_n \sigma = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} \tau$

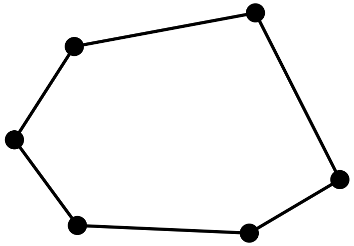
chain complex $\dots \xrightarrow{\partial_{n+2}} C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \dots$

n-cycles and n-boundaries $Z_n(X) = \mathrm{Ker}(\partial_n)$ $B_n(X) = \mathrm{Im}(\partial_{n+1})$

n^{th}**singular** homology group $H_n(X) = Z_n(X)/B_n(X)$

Theorem: If X is a topological space and K a triangulation of it, then for all $n \geq 0$, $H_n(X) = H_n(K)$.





$H_0(X) = \mathbb{Z}/2\mathbb{Z}$

$H_1(X) = \mathbb{Z}/2\mathbb{Z}$

$H_2(X) = 0$

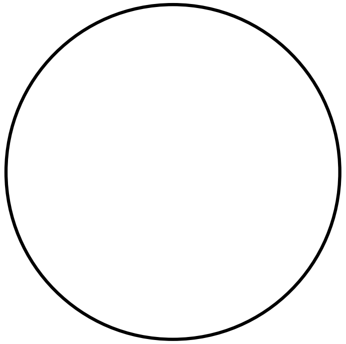
$H_0(K) = \mathbb{Z}/2\mathbb{Z}$

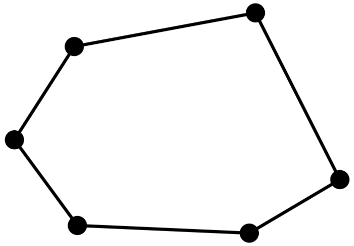
$H_1(K) = \mathbb{Z}/2\mathbb{Z}$

$H_2(X) = 0$

Homologia singular

Theorem: If X is a topological space and K a triangulation of it, then for all $n \geq 0$, $H_n(X) = H_n(K)$.





$$H_0(X) = \mathbb{Z}/2\mathbb{Z} \qquad H_0(K) = \mathbb{Z}/2\mathbb{Z}$$
$$H_1(X) = \mathbb{Z}/2\mathbb{Z} \qquad H_1(K) = \mathbb{Z}/2\mathbb{Z}$$
$$H_2(X) = 0 \qquad H_2(X) = 0$$

Theorem: If X and Y are homotopy equivalent topological spaces, then for all $n \geq 0$, $H_n(X) = H_n(Y)$.

Corollary: If K and L are homotopy equivalent simplicial complexes, then for all $n \geq 0$, $H_n(K) = H_n(L)$.

the homology groups are **invariants** of homotopy classes

Homologia é um functor

We have seen that homology transforms topological spaces into vector spaces

$$\begin{aligned} H_i \colon \mathrm{Top} &\longrightarrow \mathrm{Vect} \\ X &\longmapsto H_i(X) \end{aligned}$$

Actually, it also transforms **continous maps** into **linear maps**

$$X \xrightarrow{\ f\ } Y \qquad\qquad H_n(X) \xrightarrow{H_n(f)} H_n(Y)$$

This operation preserves **commutative diagrams**:

$$X \xrightarrow[f]{} Y \xrightarrow[g]{} Z, \quad (X \xrightarrow{g\circ f} Z) \qquad\qquad H_n(X) \xrightarrow[H_n(f)]{} H_n(Y) \xrightarrow[H_n(g)]{} H_n(Z). \quad (H_n(X) \xrightarrow{H_n(g\circ f)} H_n(Z))$$

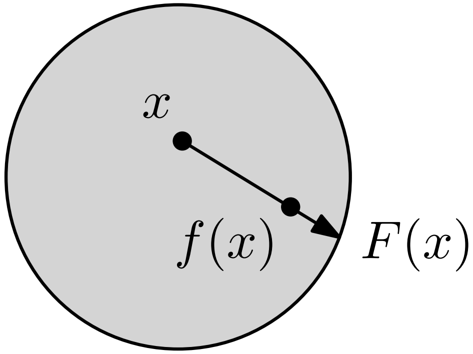
$$\boxed{H_n(g \circ f) = H_n(g) \circ H_n(f)}$$

Aplicação - na teoria

Application (Brouwer's fixed point theorem):

Let $f\colon \mathcal{B} \to \mathcal{B}$ be a continous map, where $\mathcal{B}$ is the unit closed ball of $\mathbb{R}^n$. Let us show that f has a fixed point ($f(x) = x$).

If not, we can define a map $F\colon \mathcal{B} \to \partial\mathcal{B}$ such that F restricted to $\partial\mathcal{B}$ is the identity. To do so, define $F(x)$ as the first intersection between the half-line $[x, f(x))$ and $\partial\mathcal{B}$.



Denote the inclusion $i\colon \partial\mathcal{B} \to \mathcal{B}$. Then $F \circ i\colon \partial\mathcal{B} \to \partial\mathcal{B}$ is the identity.
By functoriality, we have commutative diagrams

$$\overbrace{\partial\mathcal{B} \xrightarrow[i]{} \mathcal{B} \xrightarrow[F]{} \partial\mathcal{B}}^{\mathrm{id}}, \qquad \overbrace{H_i(\partial\mathcal{B}) \xrightarrow[H_i(i)]{} H_i(\mathcal{B}) \xrightarrow[H_i(F)]{} H_i(\partial\mathcal{B})}^{H_i(\mathrm{id})}.$$

But for $i = n - 1$, we have an absurdity:

$$\overbrace{\mathbb{Z}/2\mathbb{Z} \longrightarrow 0 \longrightarrow \mathbb{Z}/2\mathbb{Z}}^{\mathrm{id}}.$$

I - Simplicial homology

1 - Reminder of algebra

2 - Homological algebra

3 - Incremental algorithm

II - More about homology

1 - Topology of simplicial complexes

2 - Singular homology

3 - Functoriality

III - Homological inference

1 - Thickening parameter selection

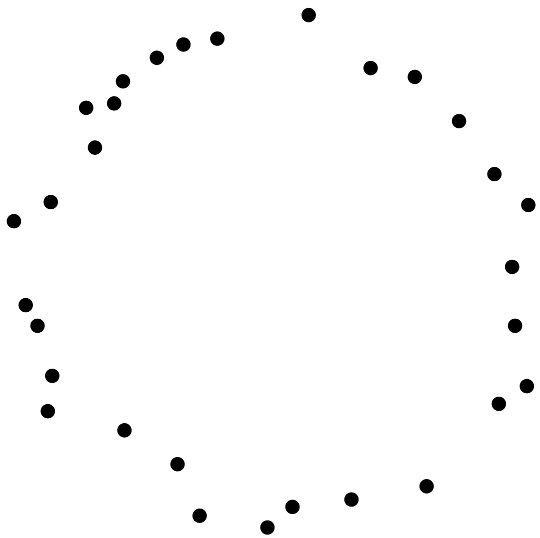
2 - Čech complex

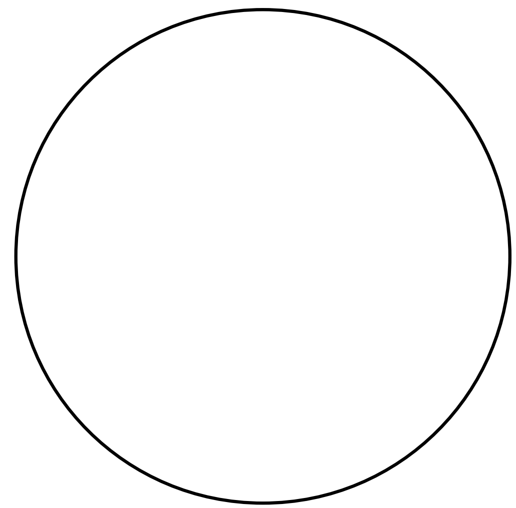
3 - Rips complex

O problema da inferência homológica

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.

X

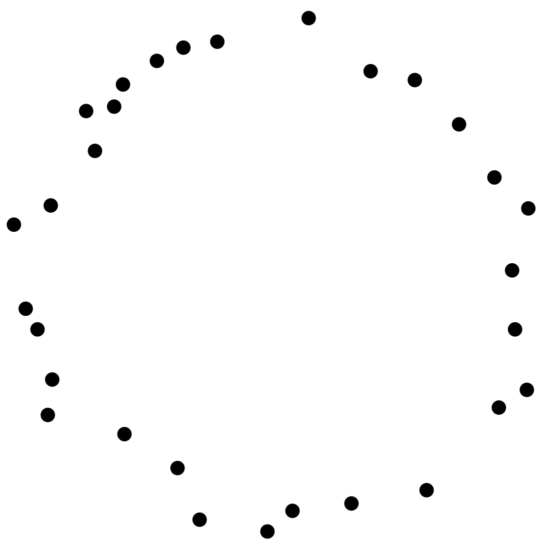




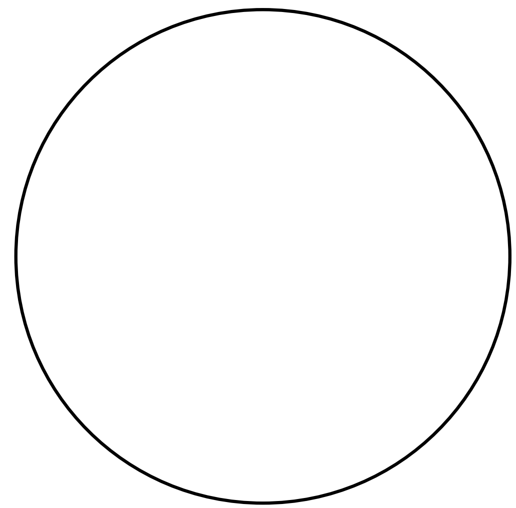
$\mathcal{M}$

O problema da inferência homológica

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.



X



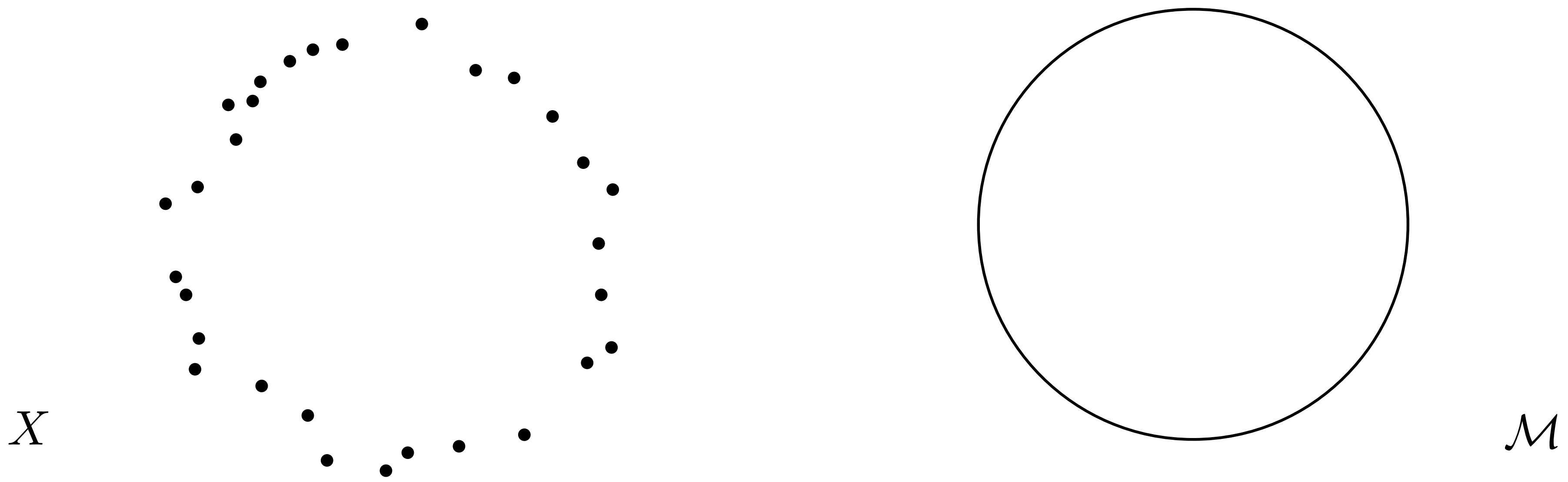
$\mathcal{M}$

We cannot use X directly. Its homology is disapointing:

$$\beta_0(X) = 30 \quad \text{and} \quad \beta_i(X) \text{ for } i \geq 1$$

number of connected components
$=$ number of points of X

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.



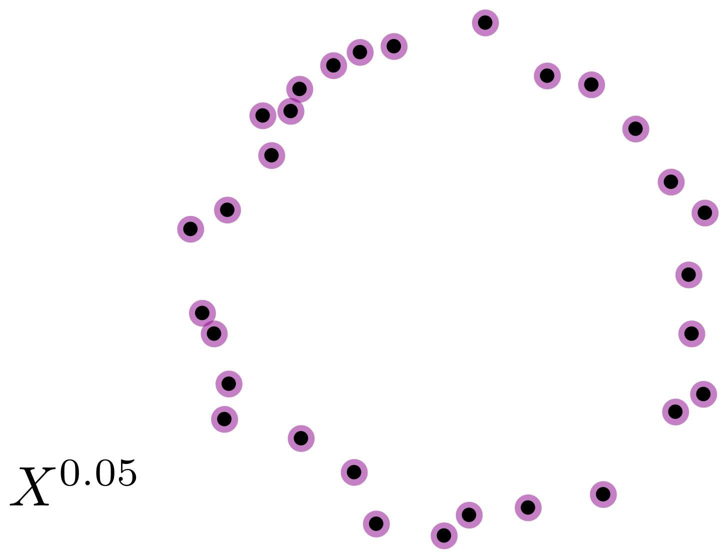
We cannot use X directly.

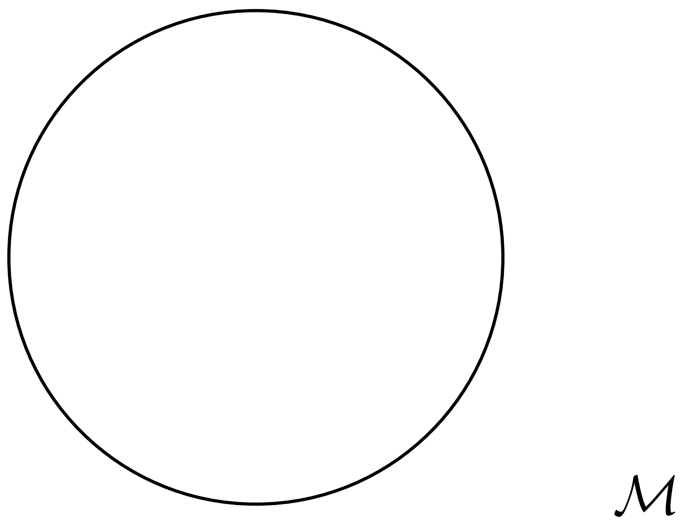
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.





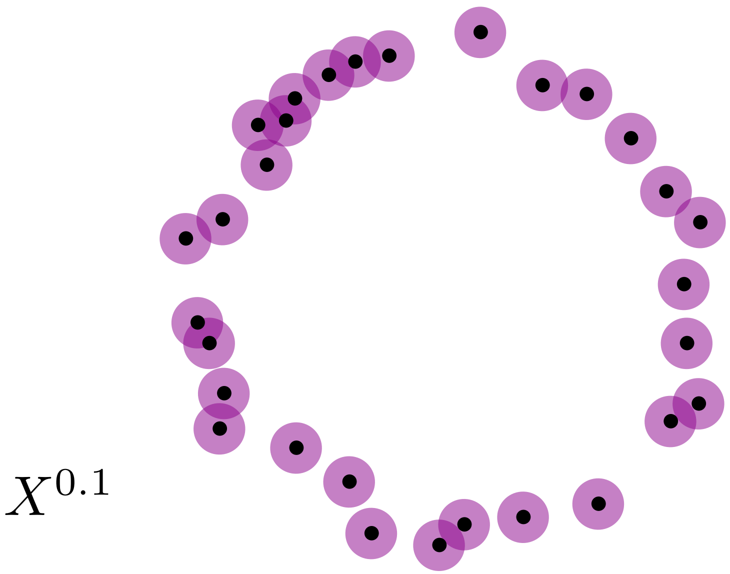
We cannot use X directly.

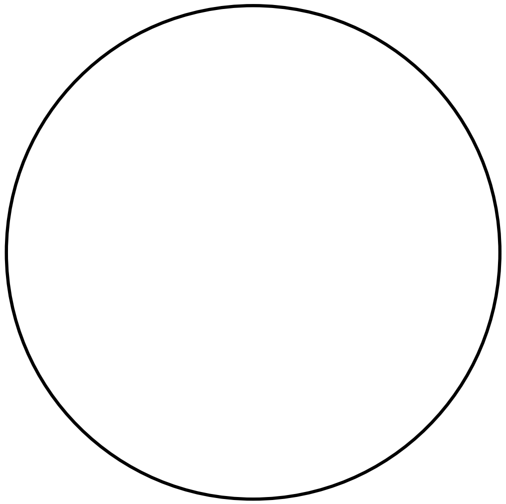
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.





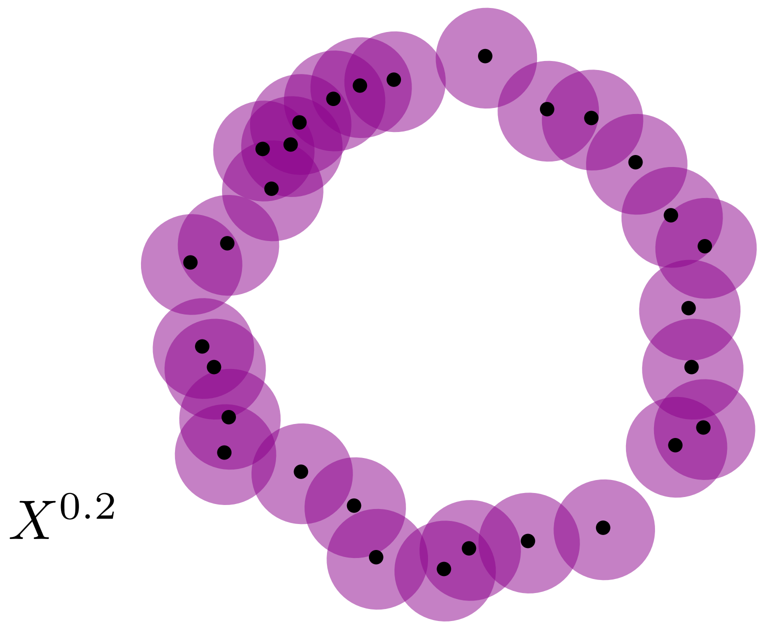
We cannot use X directly.

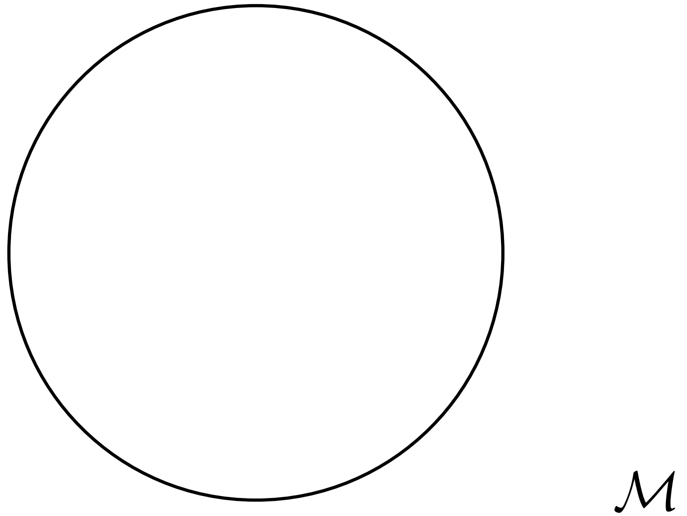
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.





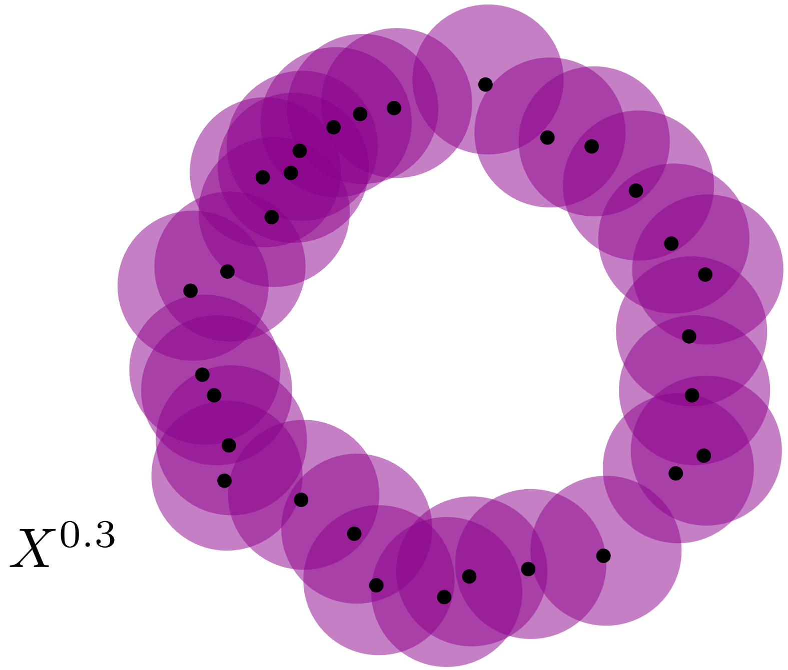
We cannot use X directly.

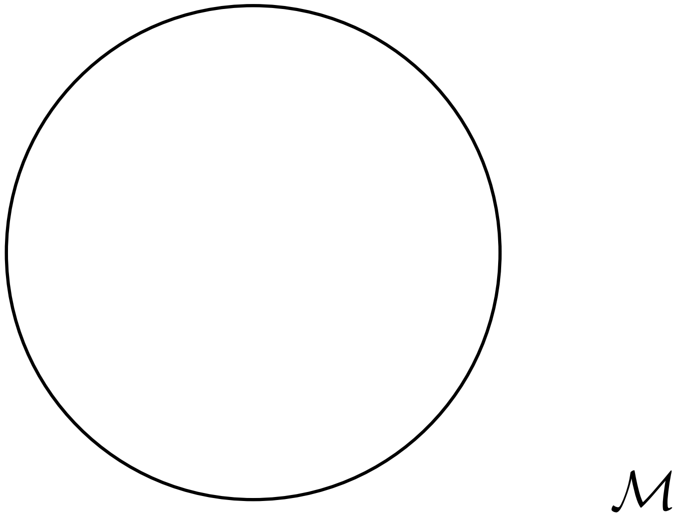
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.





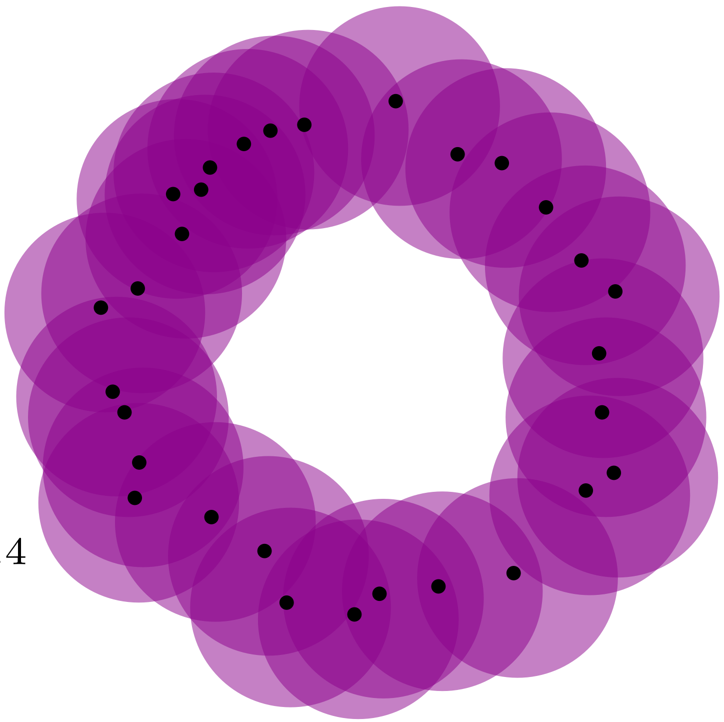
We cannot use X directly.

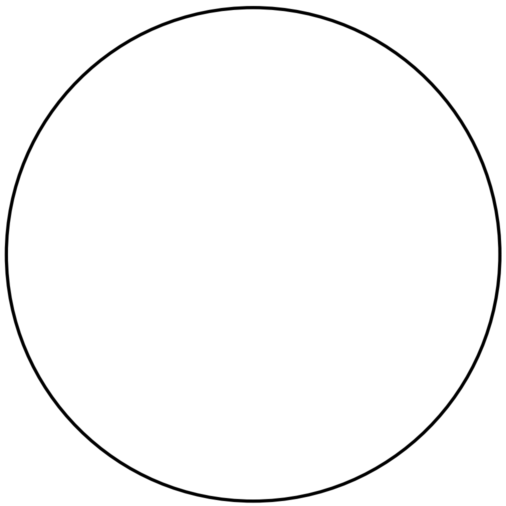
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.


$X^{0.4}$


$\mathcal{M}$

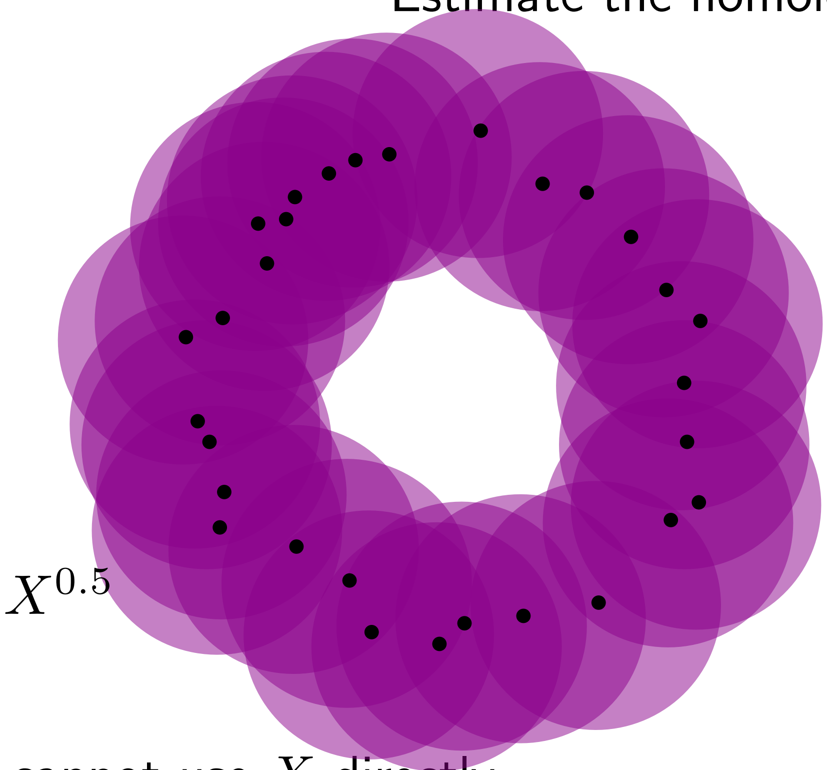
We cannot use X directly.

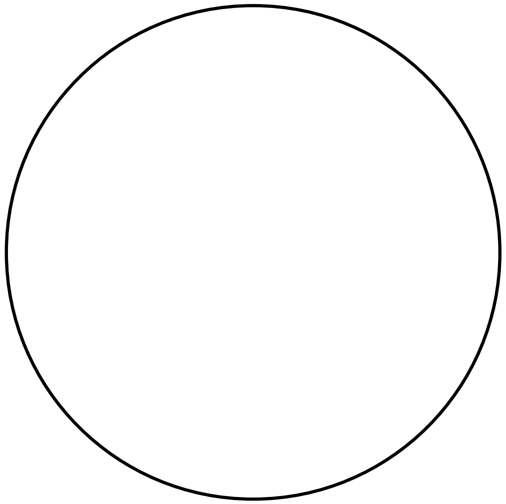
Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.





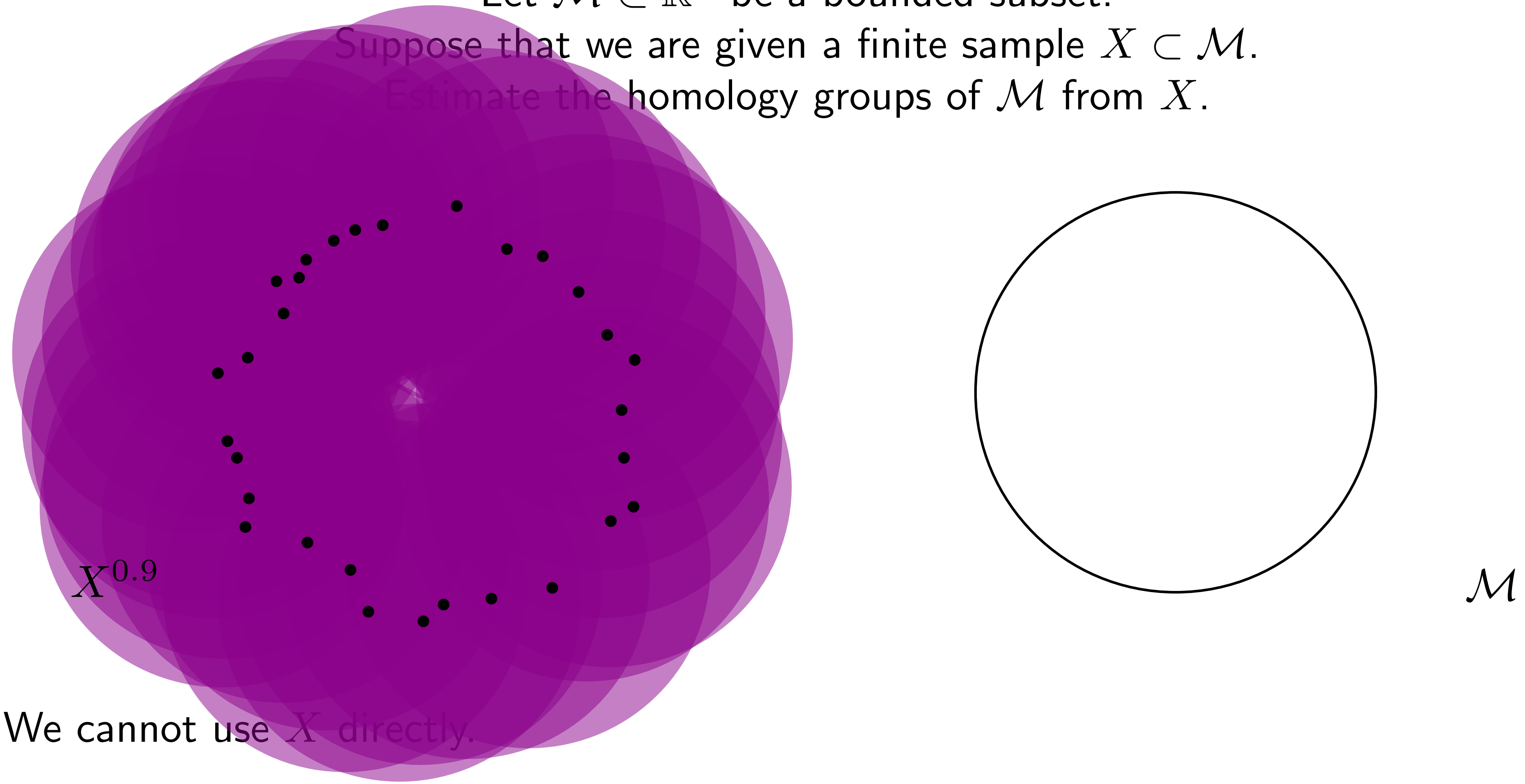
We cannot use X directly.

Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

Let $\mathcal{M} \subset \mathbb{R}^n$ be a bounded subset.
Suppose that we are given a finite sample $X \subset \mathcal{M}$.
Estimate the homology groups of $\mathcal{M}$ from X.



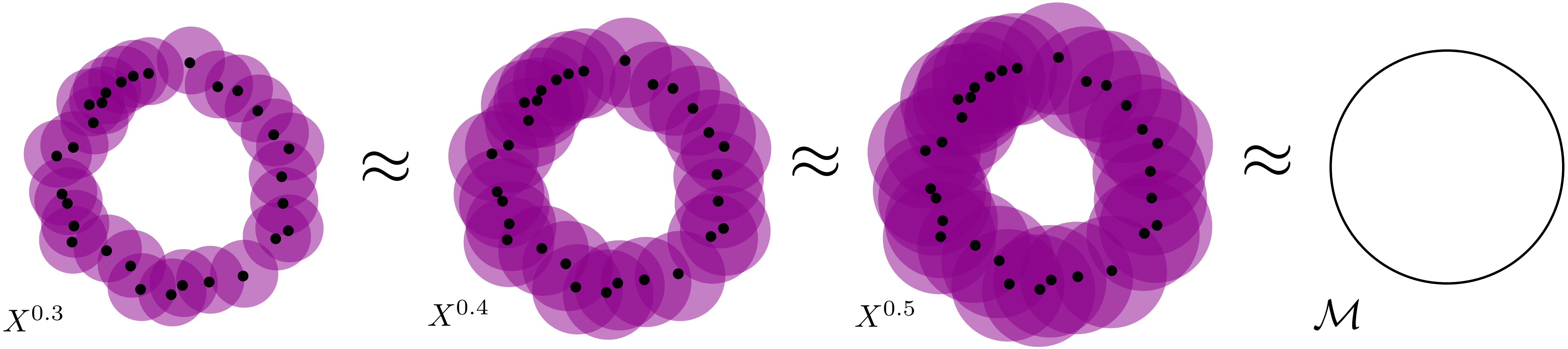
We cannot use X directly.

Idea: Thicken X.

Definition: For every $t \geq 0$, the t-**thickening** of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:

$$X^t = \{y \in \mathbb{R}^n \mid \exists x \in X, \|x - y\| \leq t\}.$$

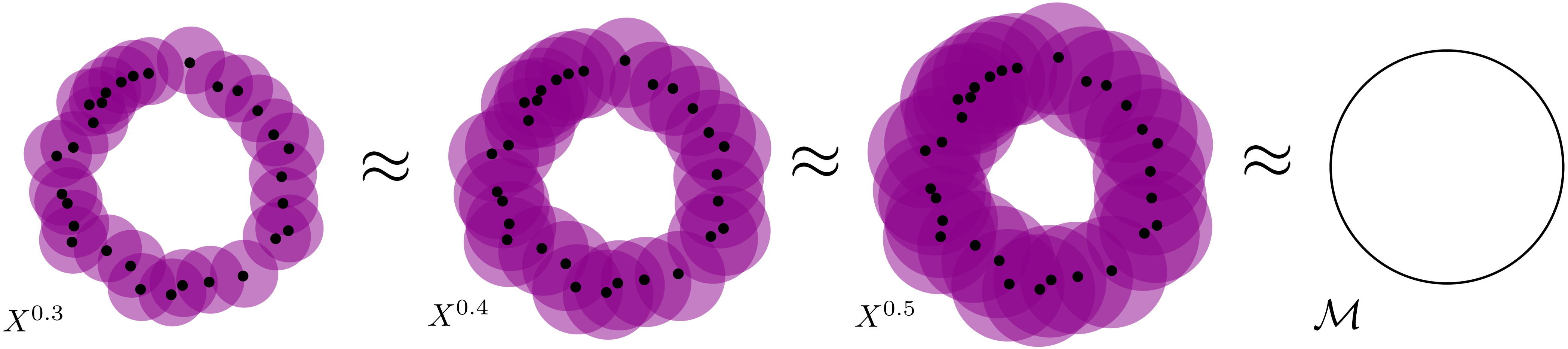
Some thickenings are homotopy equivalent to $\mathcal{M}$.



Hence we can recover the homology of $\mathcal{M}$:

$$\beta_0(\mathcal{M}) = \beta_0(X^{0.3})$$
$$\beta_1(\mathcal{M}) = \beta_1(X^{0.3})$$
$$\beta_2(\mathcal{M}) = \beta_2(X^{0.3})$$
$$\ldots$$

Some thickenings are homotopy equivalent to $\mathcal{M}$.



Hence we can recover the homology of $\mathcal{M}$:

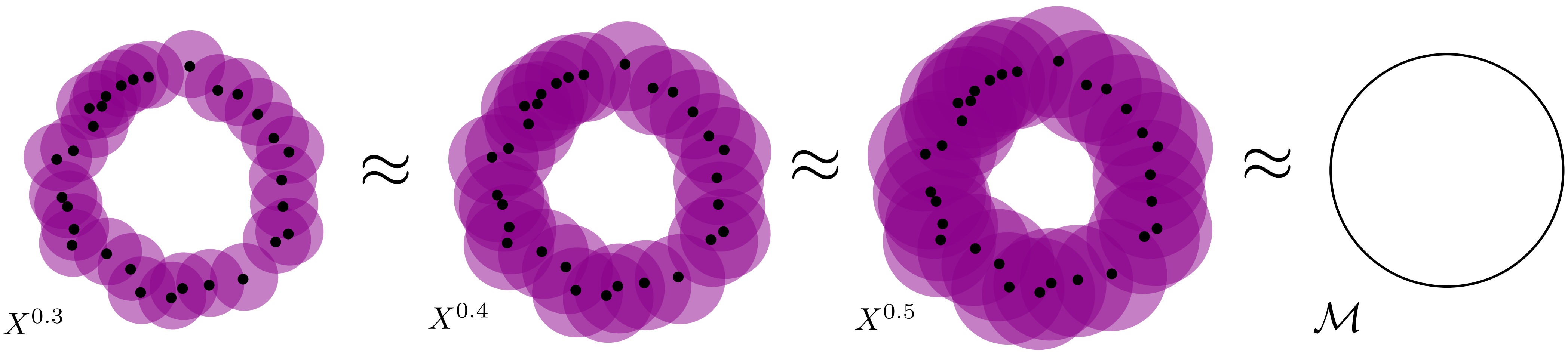
$$\beta_0(\mathcal{M}) = \beta_0(X^{0.3})$$
$$\beta_1(\mathcal{M}) = \beta_1(X^{0.3})$$
$$\beta_2(\mathcal{M}) = \beta_2(X^{0.3})$$
$$\cdots$$

Question 1: How to select a t such that $X^t \approx \mathcal{M}$?

Question 2: How to compute the homology groups of X^t?

O problema da inferência homológica

Some thickenings are homotopy equivalent to $\mathcal{M}$.



Hence we can recover the homology of $\mathcal{M}$:

$$\beta_0(\mathcal{M}) = \beta_0(X^{0.3})$$
$$\beta_1(\mathcal{M}) = \beta_1(X^{0.3})$$
$$\beta_2(\mathcal{M}) = \beta_2(X^{0.3})$$
$$\dots$$

Question 1: How to select a t such that $X^t \approx \mathcal{M}$? → Hausdorff distance, Reach

Question 2: How to compute the homology groups of X^t?

Distância de Hausdorff

Let X be any subset of $\mathbb{R}^n$. The function **distance to** X is the map

$$\begin{aligned} \operatorname{dist}(\cdot, X) : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ y &\longmapsto \operatorname{dist}(y, X) = \inf\{\|y - x\|, x \in X\} \end{aligned}$$

A **projection** of $y \in \mathbb{R}^n$ on X is a point $x \in X$ which attains this infimum.

Distância de Hausdorff

Let X be any subset of $\mathbb{R}^n$. The function **distance to** X is the map

$$\begin{aligned} \operatorname{dist}(\cdot, X) : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ y &\longmapsto \operatorname{dist}(y, X) = \inf\{\|y - x\|, x \in X\} \end{aligned}$$

A **projection** of $y \in \mathbb{R}^n$ on X is a point $x \in X$ which attains this infimum.

Definition: Let $Y \subset \mathbb{R}^n$ be another subset. The **Hausdorff distance** between X and Y is

$$\begin{aligned} \mathrm{d_H}(X, Y) &= \max\left\{ \sup_{y \in Y} \operatorname{dist}(y, X), \ \sup_{x \in X} \operatorname{dist}(x, Y) \right\} \\ &= \max\left\{ \sup_{y \in Y} \inf_{x \in X} \|x - y\|, \ \sup_{x \in X} \inf_{y \in Y} \|x - y\| \right\}. \end{aligned}$$

Distância de Hausdorff

Let X be any subset of $\mathbb{R}^n$. The function **distance to** X is the map

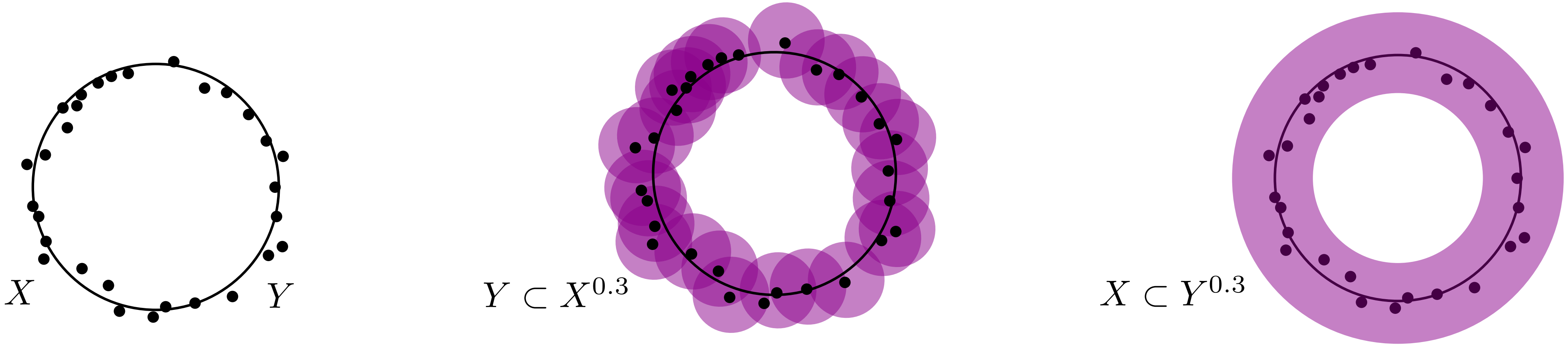
$$\begin{aligned} \mathrm{dist}\,(\cdot, X) : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ y &\longmapsto \mathrm{dist}\,(y, X) = \inf\{\|y - x\|\,, x \in X\} \end{aligned}$$

A **projection** of $y \in \mathbb{R}^n$ on X is a point $x \in X$ which attains this infimum.

Definition: Let $Y \subset \mathbb{R}^n$ be another subset. The **Hausdorff distance** between X and Y is

$$\begin{aligned} \mathrm{d_H}\,(X, Y) &= \max\left\{\sup_{y\in Y} \mathrm{dist}\,(y, X)\,,\ \sup_{x \in X} \mathrm{dist}\,(x, Y)\right\} \\ &= \max\left\{\sup_{y\in Y}\inf_{x\in X} \|x - y\|\,,\ \sup_{x\in X}\inf_{y\in Y} \|x - y\|\right\}. \end{aligned}$$

Proposition: The Hausdorff distance is equal to $\inf\{t \geq 0 \mid X \subset Y^t \text{ and } Y \subset X^t\}$.



Medial axis e reach

The **medial axis** of X is the subset $\operatorname{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\operatorname{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \operatorname{dist}(y, X)\}.$$

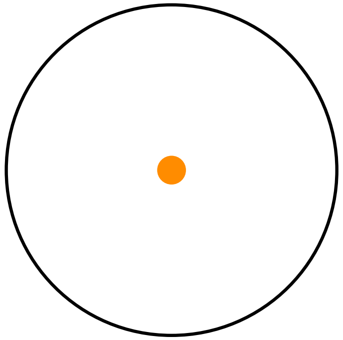
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}\,(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\mathrm{med}\,(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}\,(y, X)\}\,.$$

Examples:

The medial axis of the unit circle is the origin



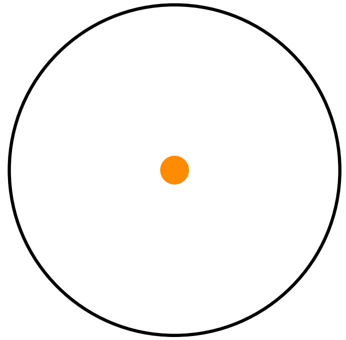
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}\,(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

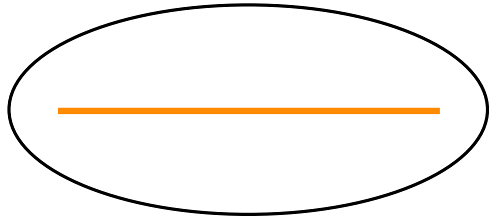
$$\mathrm{med}\,(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}\,(y, X)\}\,.$$

Examples:

The medial axis of the unit circle is the origin



The medial axis of an ellipse is a segment



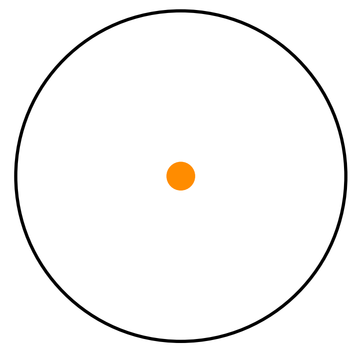
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}\,(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

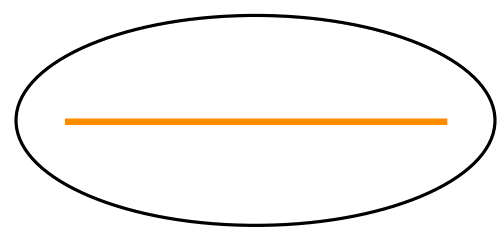
$$\mathrm{med}\,(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}\,(y, X)\}\,.$$

Examples:

The medial axis of the unit circle is the origin



The medial axis of an ellipse is a segment



The medial axis of a point is the empty set

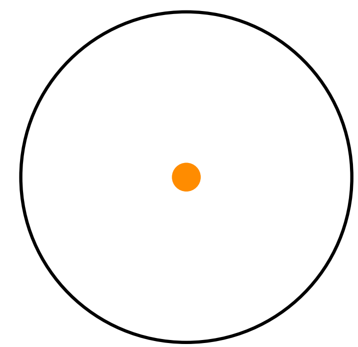
Medial axis e reach

The **medial axis** of X is the subset $\operatorname{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

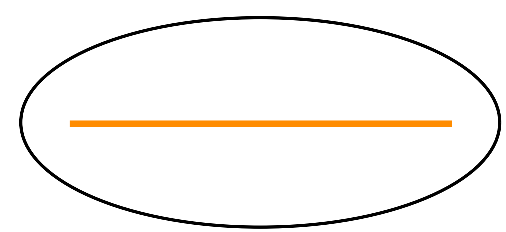
$$\operatorname{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \operatorname{dist}(y, X)\}.$$

Examples:

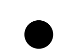
The medial axis of the unit circle is the origin



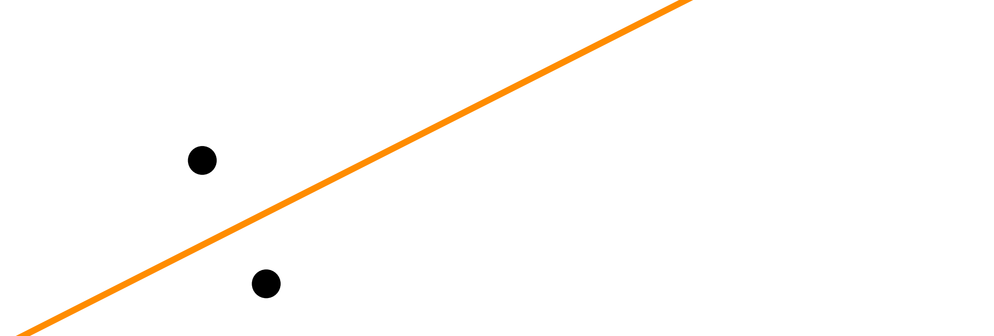
The medial axis of an ellipse is a segment



The medial axis of a point is the empty set



The medial axis of two points is their bisector



Medial axis e reach

The **medial axis** of X is the subset $\operatorname{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\operatorname{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \operatorname{dist}(y, X)\}.$$

The **reach** of X is

$$\begin{aligned} \operatorname{reach}(X) &= \inf \{\operatorname{dist}(y, X) \mid y \in \operatorname{med}(X)\} \\ &= \inf \{\|x - y\| \mid x \in X, y \in \operatorname{med}(X)\}. \end{aligned}$$

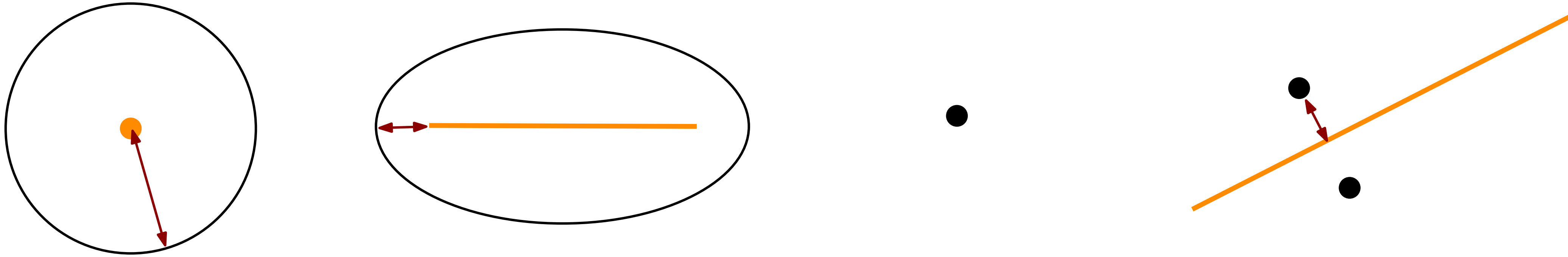
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}\,(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\mathrm{med}\,(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}\,(y, X)\}\,.$$

The **reach** of X is

$$\begin{aligned}\mathrm{reach}\,(X) &= \inf\,\{\mathrm{dist}\,(y, X) \mid y \in \mathrm{med}\,(X)\} \\ &= \inf\,\{\|x - y\| \mid x \in X, y \in \mathrm{med}\,(X)\}\,.\end{aligned}$$



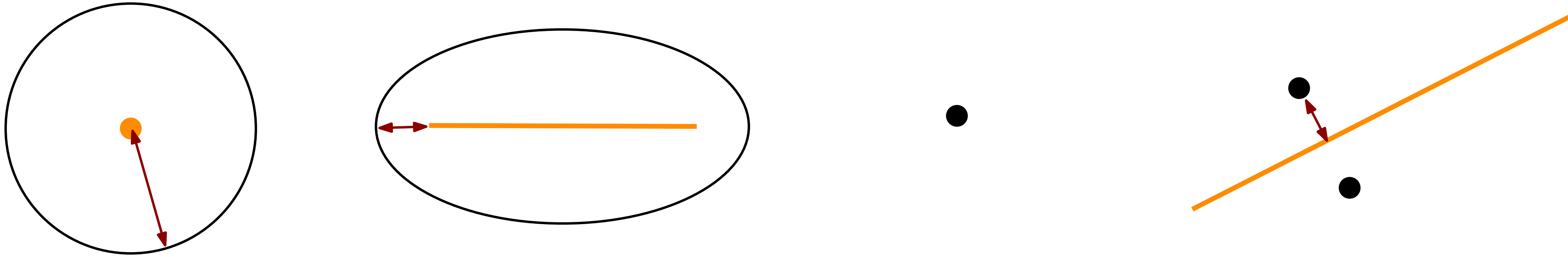
Medial axis e reach

The **medial axis** of X is the subset $\operatorname{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\operatorname{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y-x\| = \|y-x'\| = \operatorname{dist}(y, X)\}.$$

The **reach** of X is

$$\begin{aligned} \operatorname{reach}(X) &= \inf\{\operatorname{dist}(y, X) \mid y \in \operatorname{med}(X)\} \\ &= \inf\{\|x-y\| \mid x \in X, y \in \operatorname{med}(X)\}. \end{aligned}$$



Proposition: For every $t \in [0, \operatorname{reach}(X))$, the spaces X and X^t are homotopy equivalent.

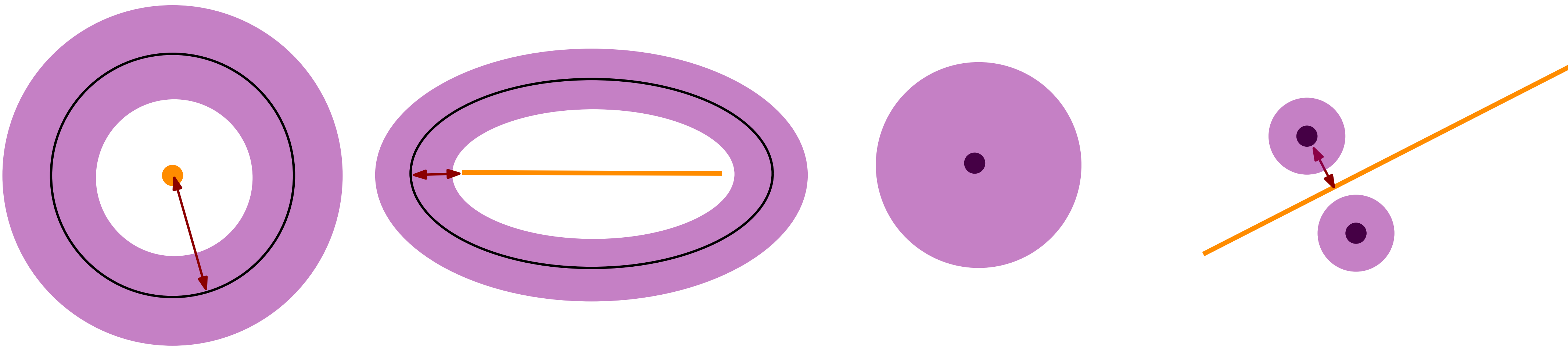
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\mathrm{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}(y, X)\}.$$

The **reach** of X is

$$\begin{aligned}\mathrm{reach}(X) &= \inf \{\mathrm{dist}(y, X) \mid y \in \mathrm{med}(X)\} \\ &= \inf \{\|x - y\| \mid x \in X, y \in \mathrm{med}(X)\}.\end{aligned}$$



Proposition: For every $t \in [0, \mathrm{reach}(X))$, the spaces X and X^t are homotopy equivalent.

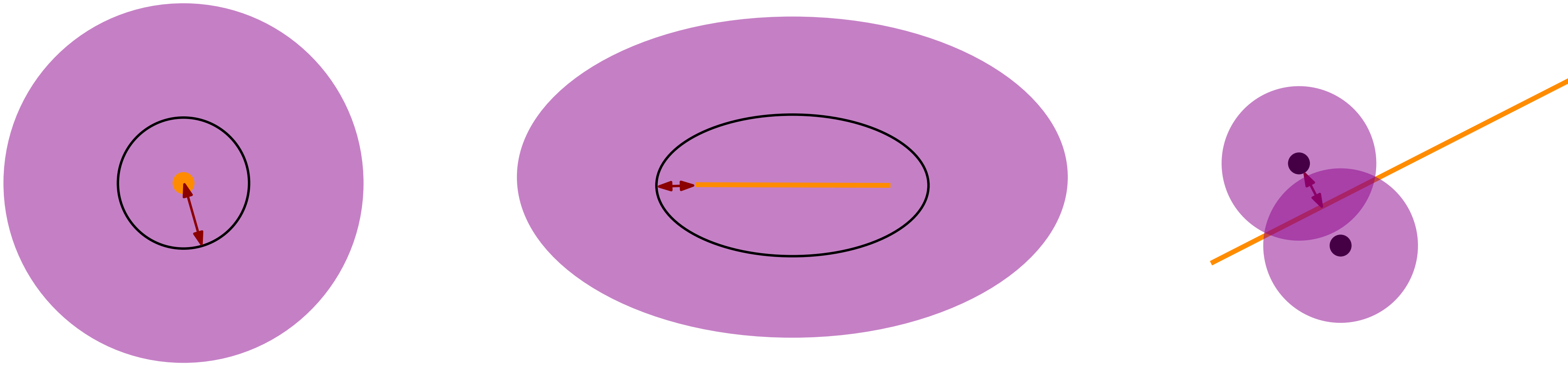
Medial axis e reach

The **medial axis** of X is the subset $\mathrm{med}(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:

$$\mathrm{med}(X) = \{y \in \mathbb{R}^n \mid \exists x, x' \in X, x \neq x', \|y - x\| = \|y - x'\| = \mathrm{dist}(y, X)\}.$$

The **reach** of X is

$$\begin{aligned}\mathrm{reach}(X) &= \inf\{\mathrm{dist}(y, X) \mid y \in \mathrm{med}(X)\} \\ &= \inf\{\|x - y\| \mid x \in X, y \in \mathrm{med}(X)\}.\end{aligned}$$



Proposition: For every $t \in [0, \mathrm{reach}(X))$, the spaces X and X^t are homotopy equivalent.

If $t \geq \mathrm{reach}(X)$, the sets X and X^t may not be homotopy equivalent.

Medial axis e reach

Proposition: For every $t \in [0, \mathrm{reach}\,(X))$, the spaces X and X^t are homotopy equivalent.

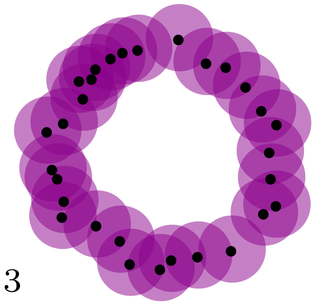
Proof: For every $t \in [0, \mathrm{reach}\,(X))$, the thickening X^t deform retracts onto X. A homotopy is given by the following map:

$$\begin{aligned} X^t \times [0,1] &\longrightarrow X^t \\ (x,t) &\longmapsto (1-t)x + t \cdot \mathrm{proj}\,(x, X)\,. \end{aligned}$$

Indeed, the projection $\mathrm{proj}\,(x, X)$ is well defined (it is unique).

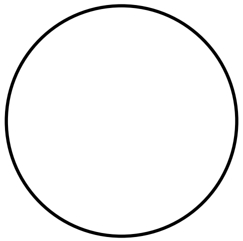
Seleção do parâmetro t

Remember Question 1: How to select a t such that $X^t \approx \mathcal{M}$?



$X^{0.3}$

$\approx$

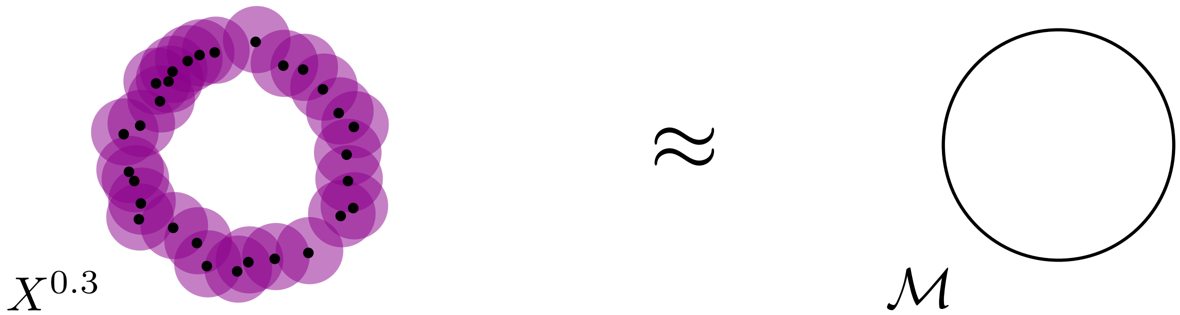


$\mathcal{M}$

Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}\left(X, \mathcal{M}\right) \leq \frac{1}{17}\mathrm{reach}\left(\mathcal{M}\right)$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \left[4\mathrm{d_H}\left(X, \mathcal{M}\right), \mathrm{reach}\left(\mathcal{M}\right) - 3\mathrm{d_H}\left(X, \mathcal{M}\right)\right).$$

Remember Question 1: How to select a t such that $X^t \approx \mathcal{M}$?



Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in [4\mathrm{d_H}(X, \mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M})).$$

Theorem (Partha Niyogi, Stephen Smale, and Shmuel Weinberger, 2008):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$.
Suppose that $\mathcal{M}$ has positive reach.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \left[2\mathrm{d_H}(X, \mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right).$$

Triangulações (fracas)

Let us consider Question 2: How to compute the homology groups of X^t?

We must a triangulation of X^t, that is: a simplicial complex K homeomorphic to X.

Actually, we will define something weaker: a simplicial complex K that is homotopy equivalent to X.

Triangulações (fracas)

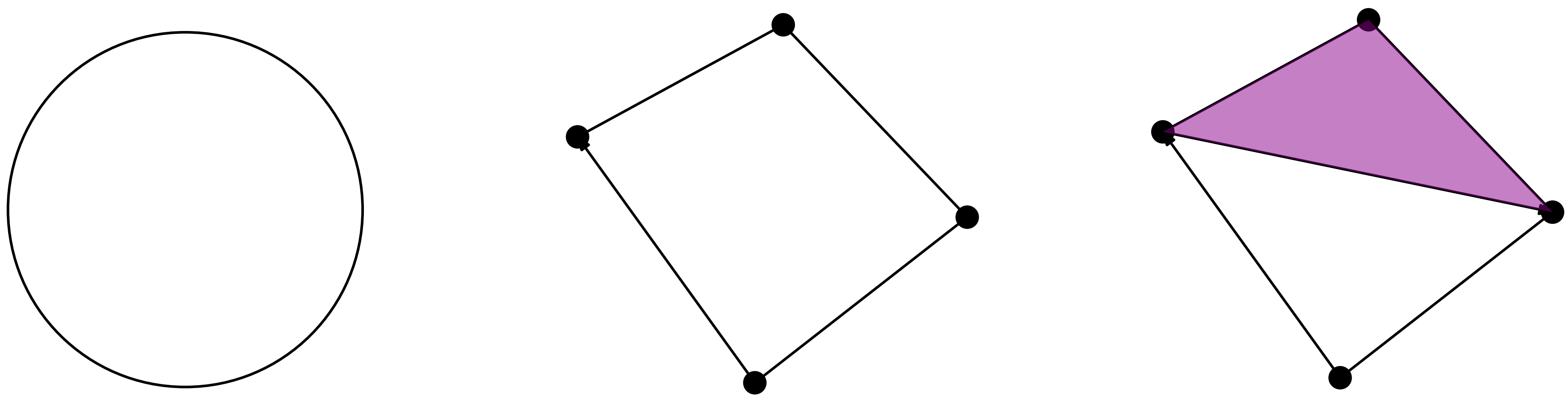
Let us consider Question 2: How to compute the homology groups of X^t?

We must a triangulation of X^t, that is: a simplicial complex K homeomorphic to X.

Actually, we will define something weaker: a simplicial complex K that is homotopy equivalent to X.

weak triangulation

Either case, we will have $\beta_i(X) = \beta_i(K)$ for all $i \geq 0$.

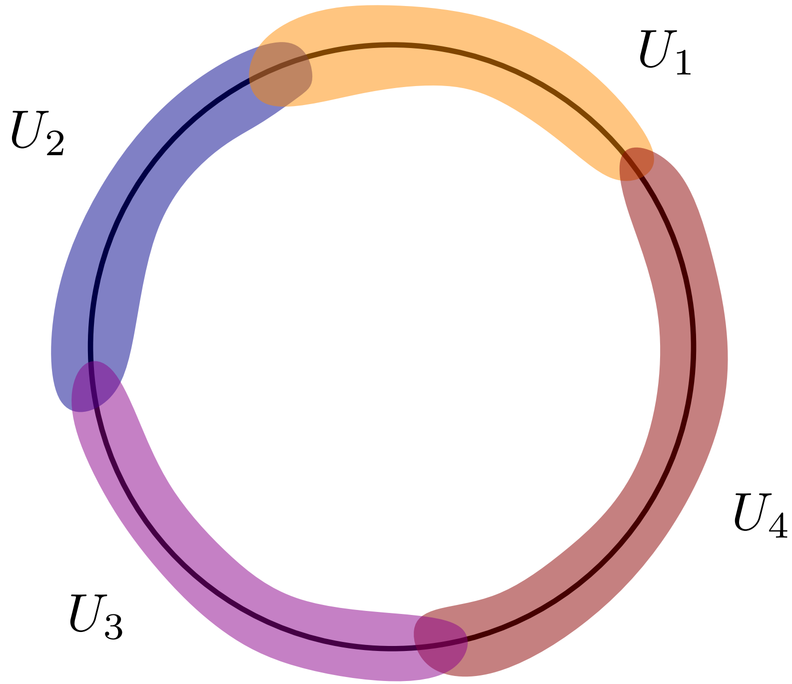


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

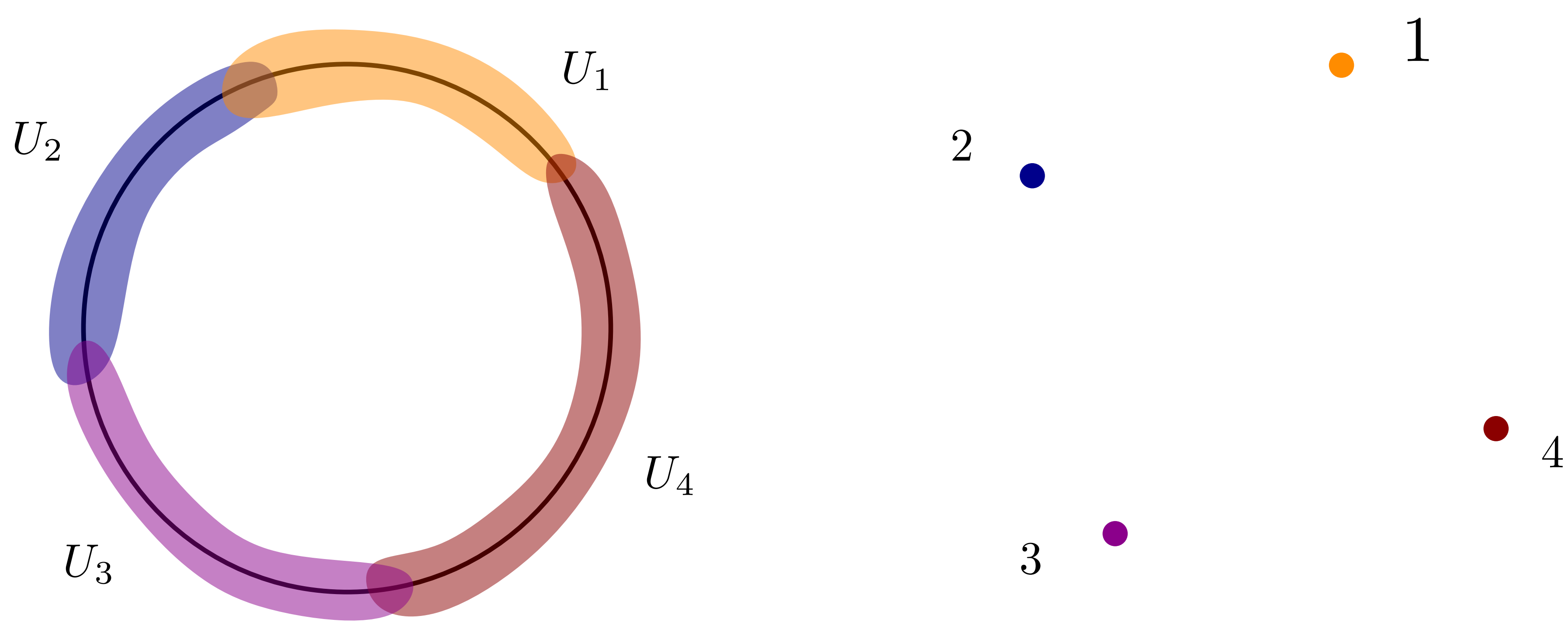


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \le i \le N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \le i \le N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, \ldots, N\}$ and whose m-simplices are the subsets $\{i_1, \ldots, i_m\} \subset \{1, \ldots, N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

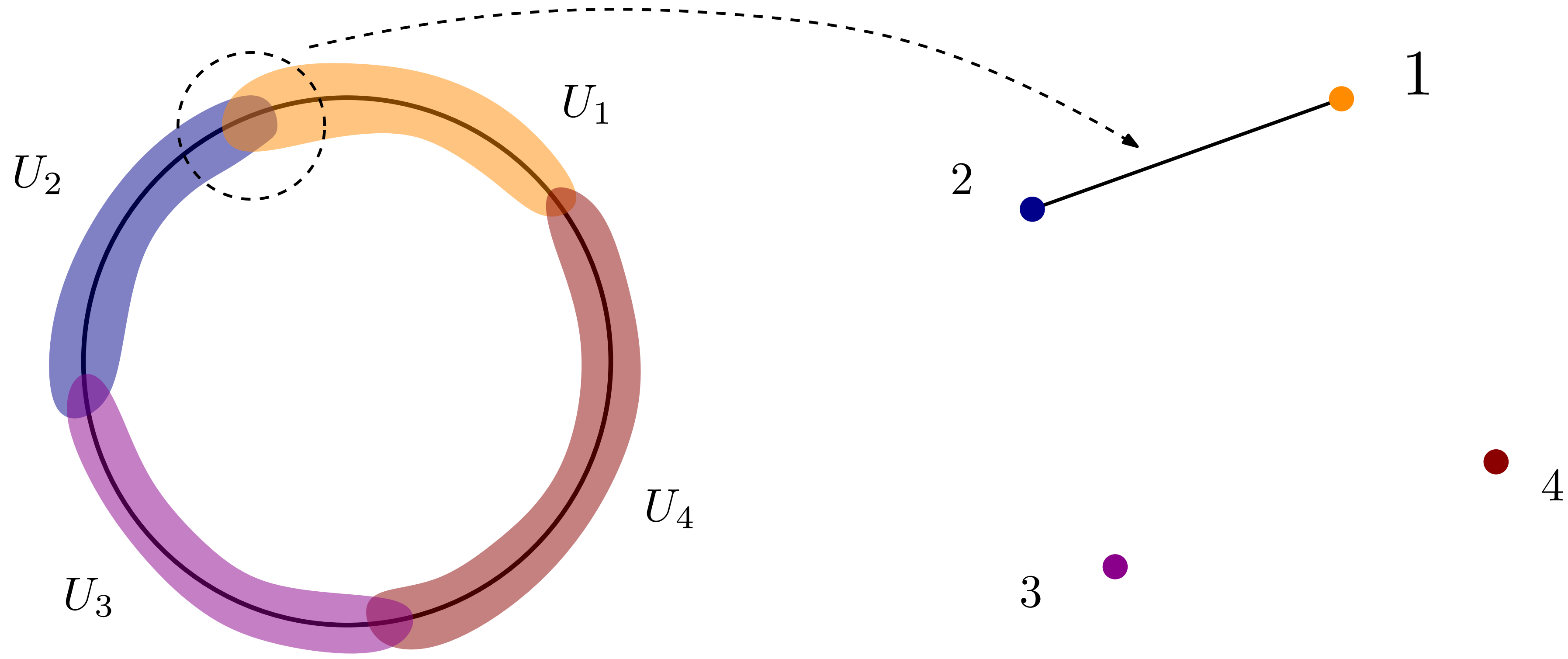


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

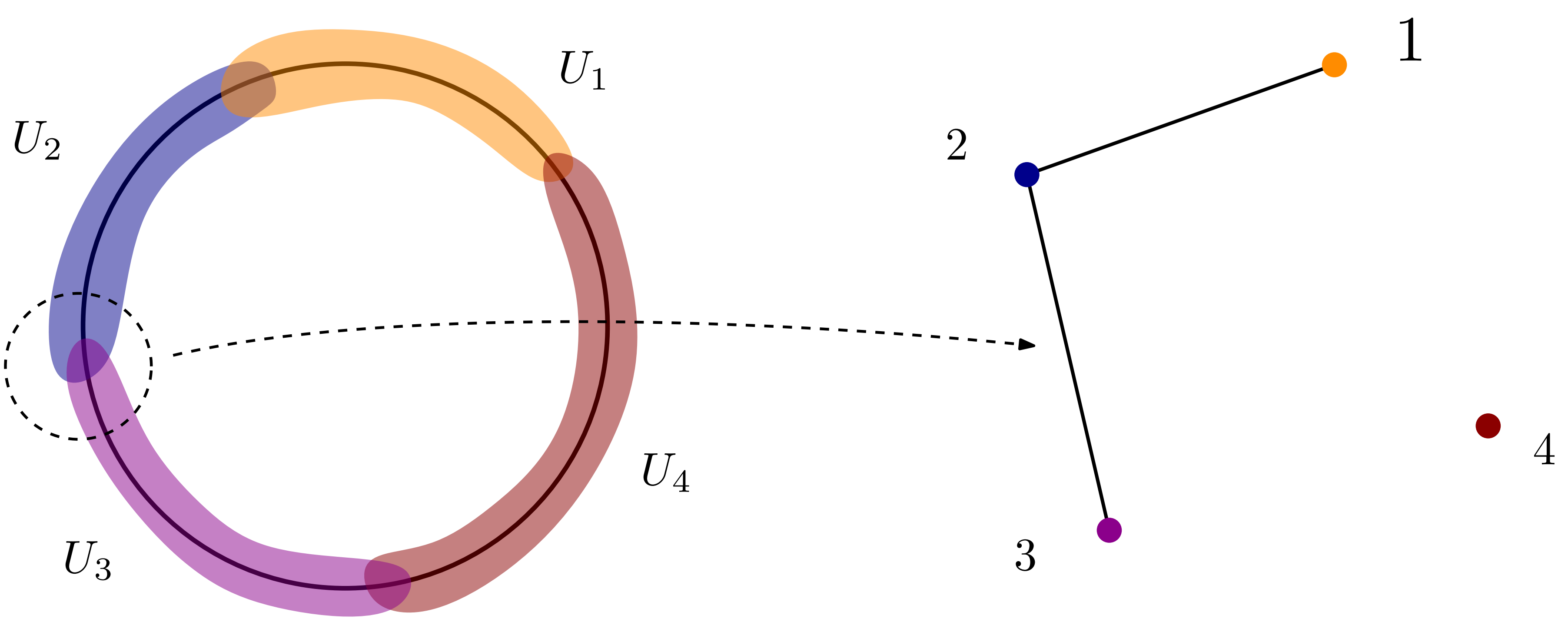


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

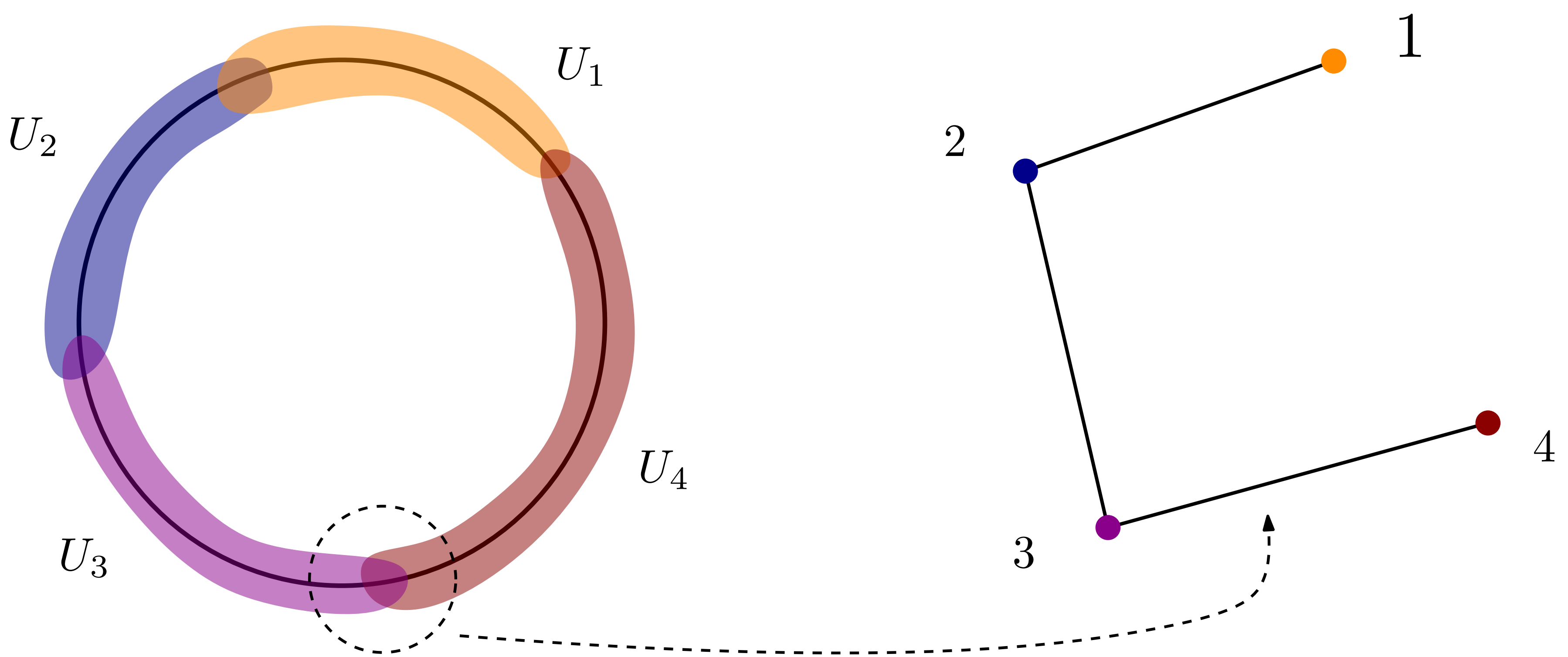


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

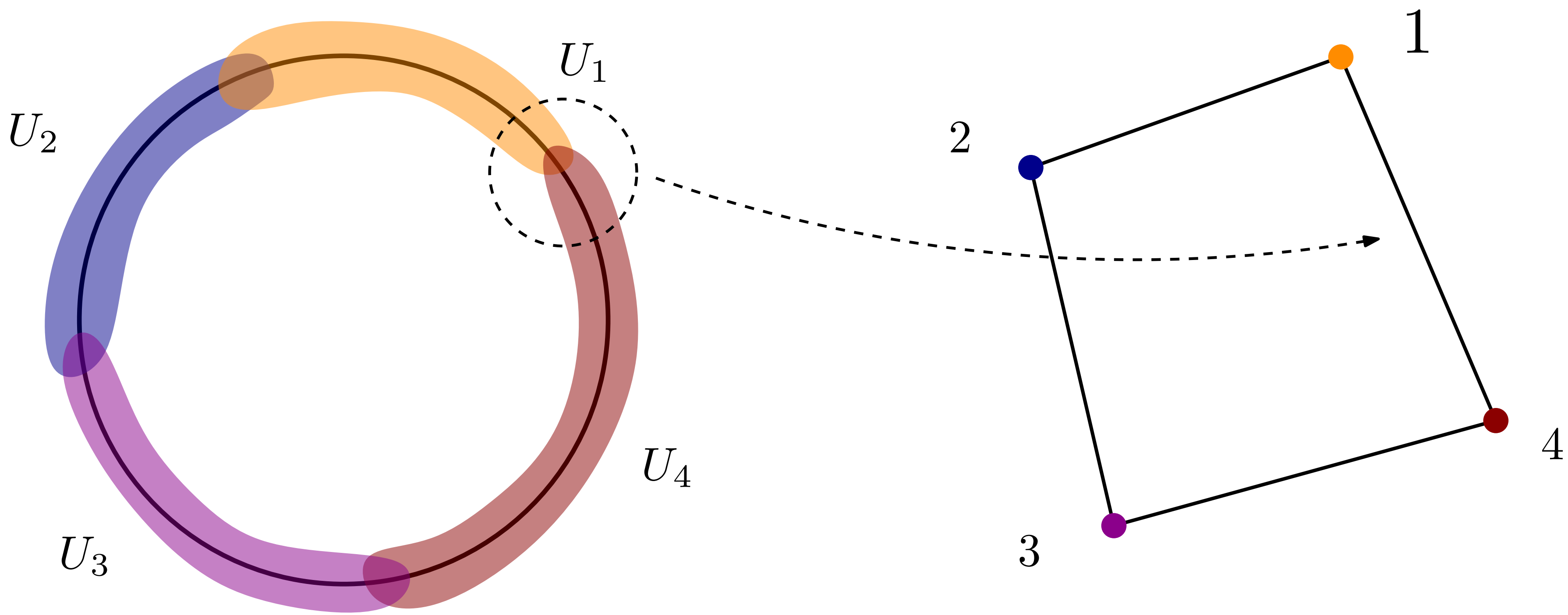


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \le i \le N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \le i \le N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

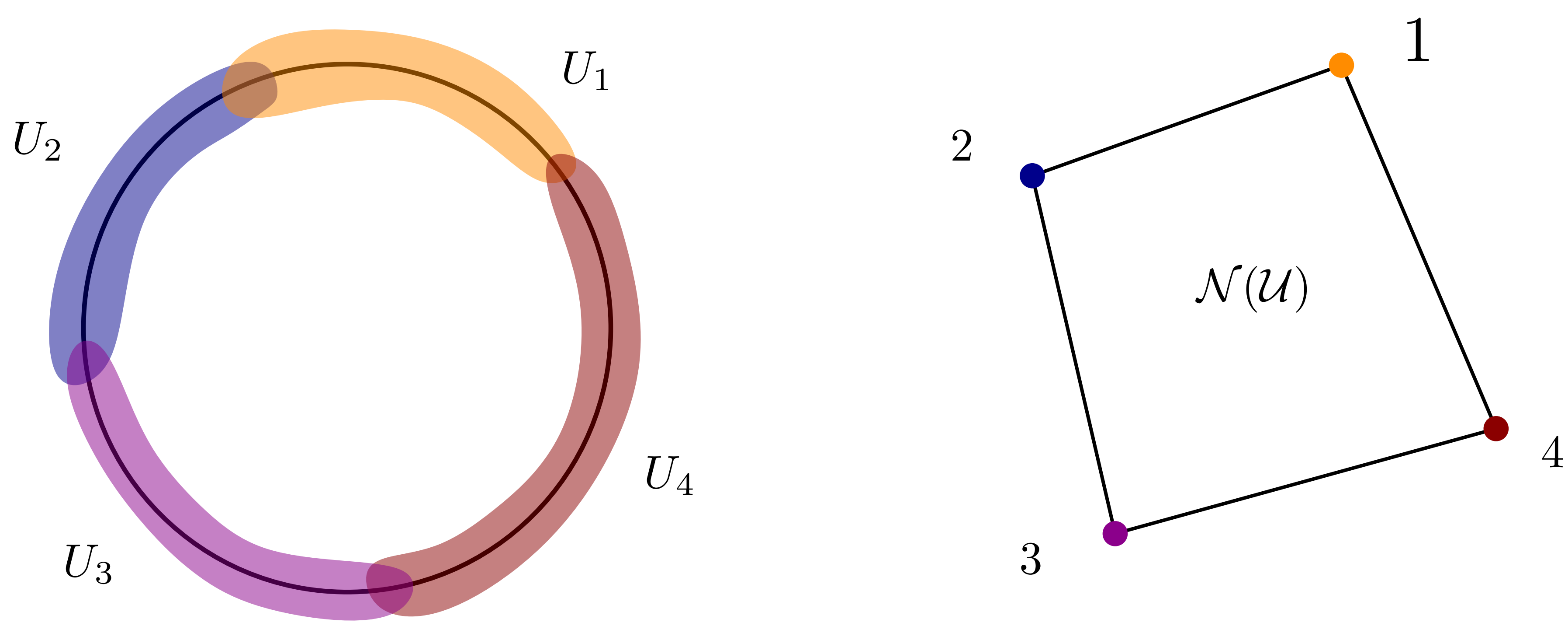


Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1\leq i\leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1\leq i\leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

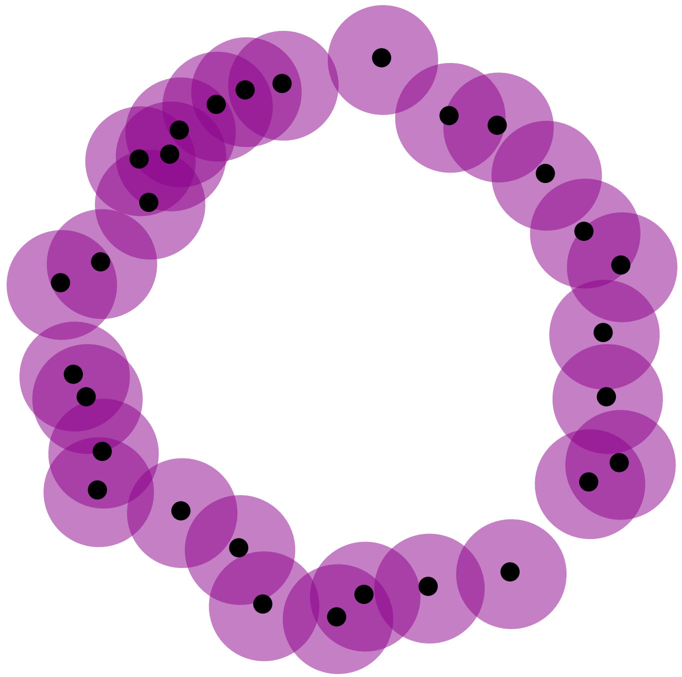


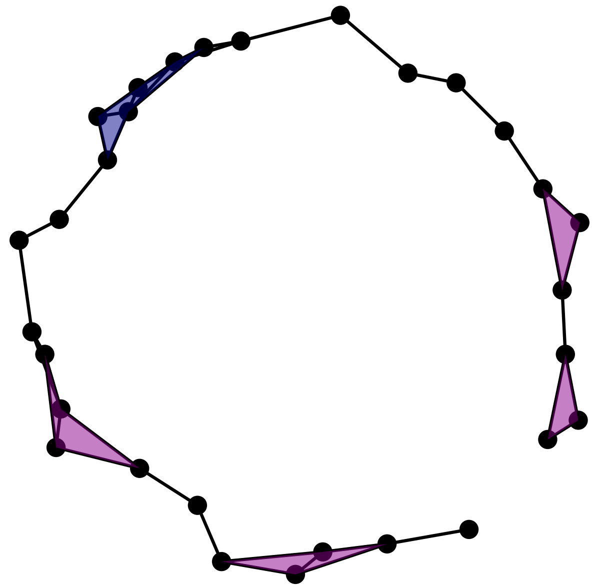
Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \le i \le N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \le i \le N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.





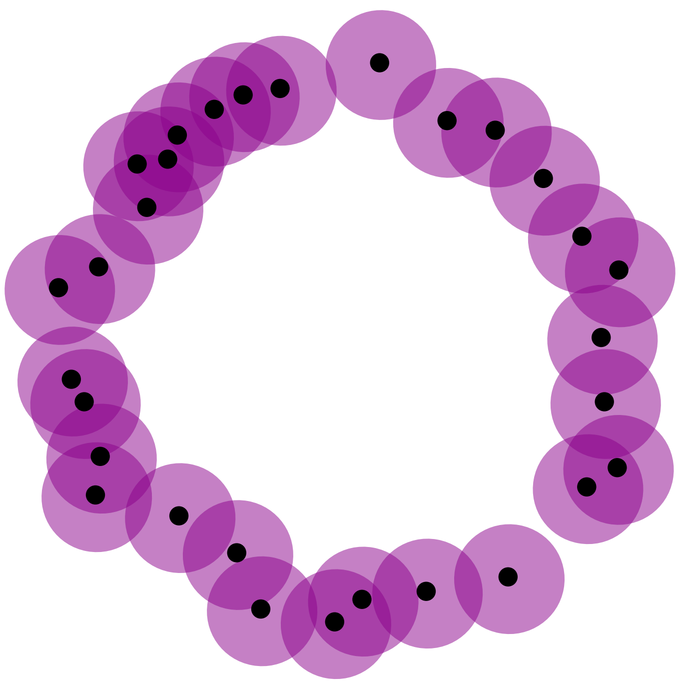
$X^{0.2} = \bigcup_{x \in X} \overline{\mathcal{B}}(x, 0.2)$ is covered by $\mathcal{U} = \{\overline{\mathcal{B}}(x, 0.2) \mid x \in X\}$

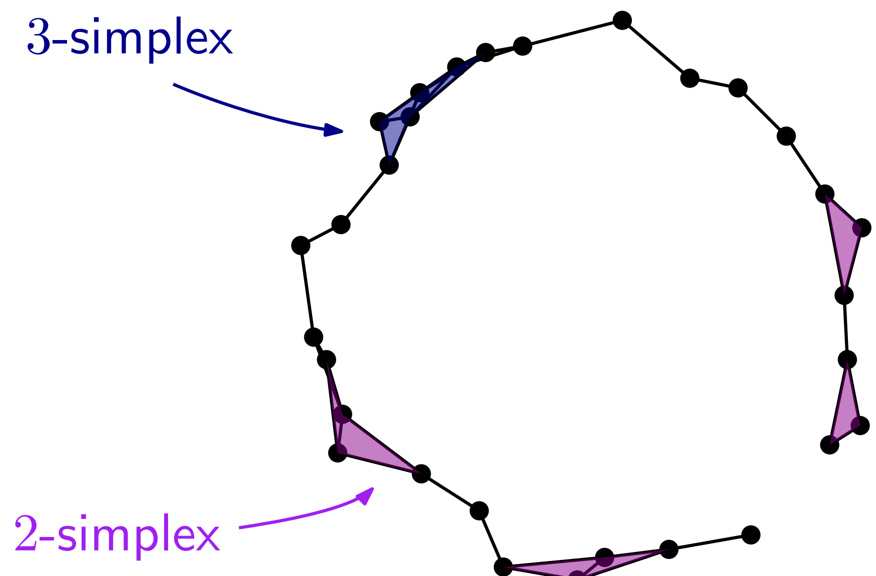
Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.



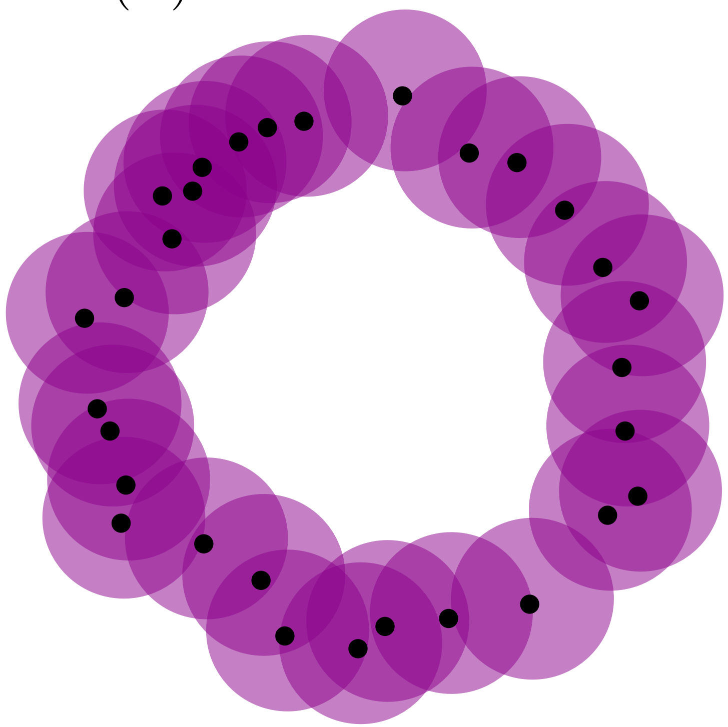


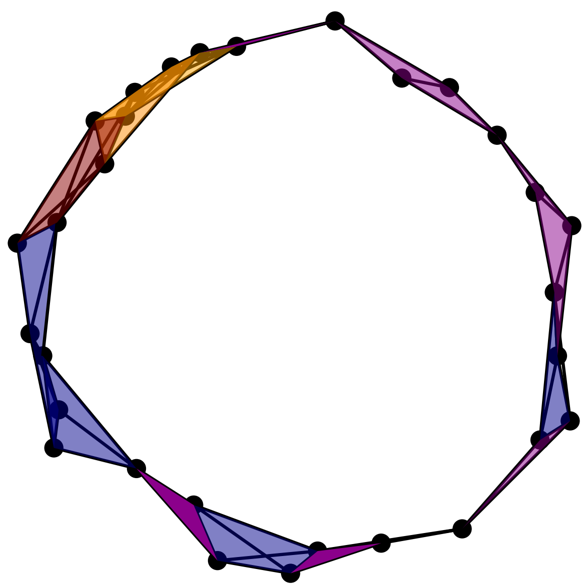
$X^{0.2} = \bigcup_{x \in X} \overline{\mathcal{B}}(x, 0.2)$ is covered by $\mathcal{U} = \{\overline{\mathcal{B}}(x, 0.2) \mid x \in X\}$

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.





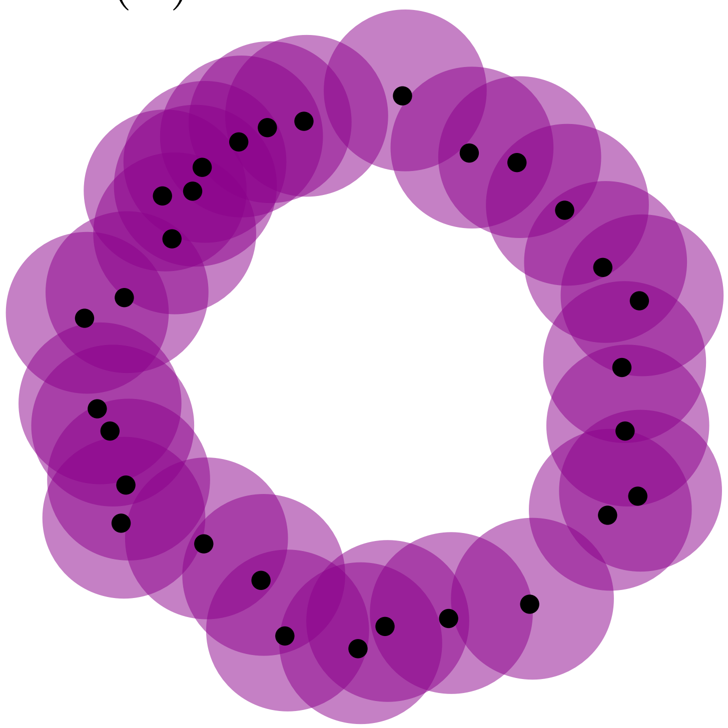
$X^{0.3} = \bigcup_{x \in X} \overline{\mathcal{B}}(x, 0.3)$ is covered by $\mathcal{U} = \{\overline{\mathcal{B}}(x, 0.3) \mid x \in X\}$

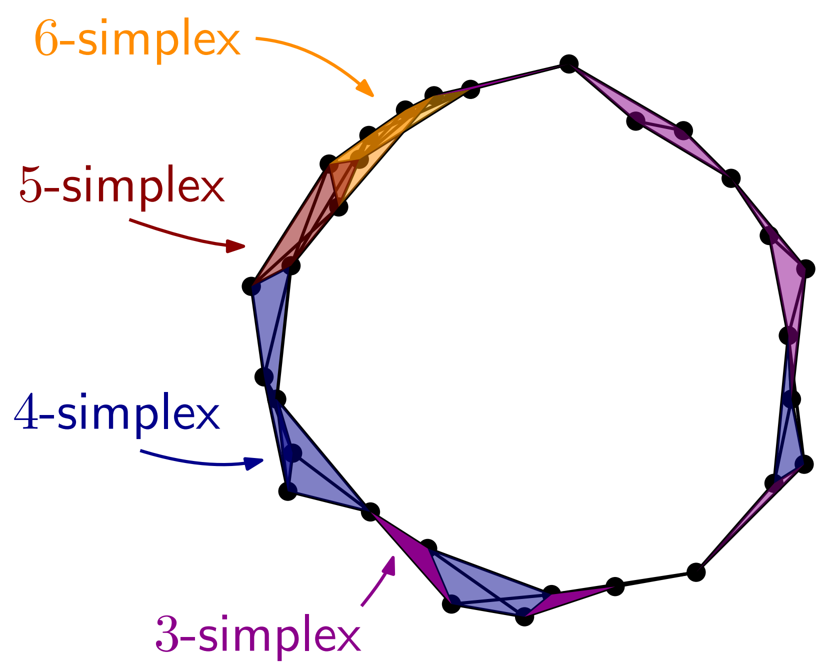
Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \le i \le N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \le i \le N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.





$X^{0.3} = \bigcup_{x \in X} \overline{\mathcal{B}}(x, 0.3)$ is covered by $\mathcal{U} = \{\overline{\mathcal{B}}(x, 0.3) \mid x \in X\}$

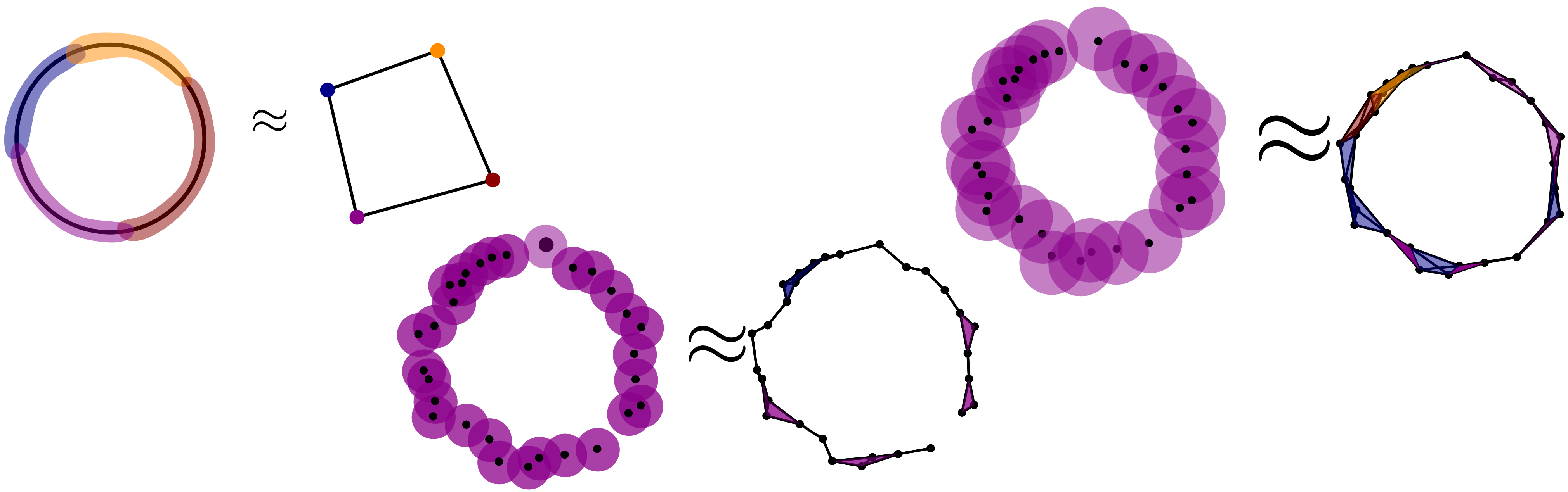
Nervos

Definition: Let X be a topological space, and $\mathcal{U} = \{U_i\}_{1 \leq i \leq N}$ a cover of X, that is, a collection of subsets $U_i \subset X$ such that

$$\bigcup_{1 \leq i \leq N} U_i = X.$$

The **nerve** of $\mathcal{U}$ is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\bigcap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(\mathcal{U})$.

Nerve theorem: Consider $X \subset \mathbb{R}^n$. Suppose that each U_i are balls (or more generally, closed and convex). Then $\mathcal{N}(\mathcal{U})$ is homotopy equivalent to X.



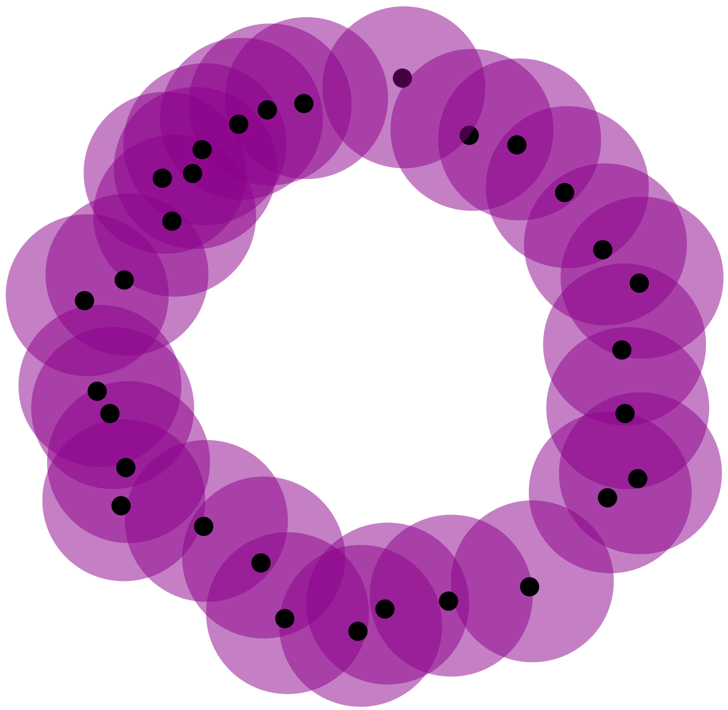
Complexo de Čech

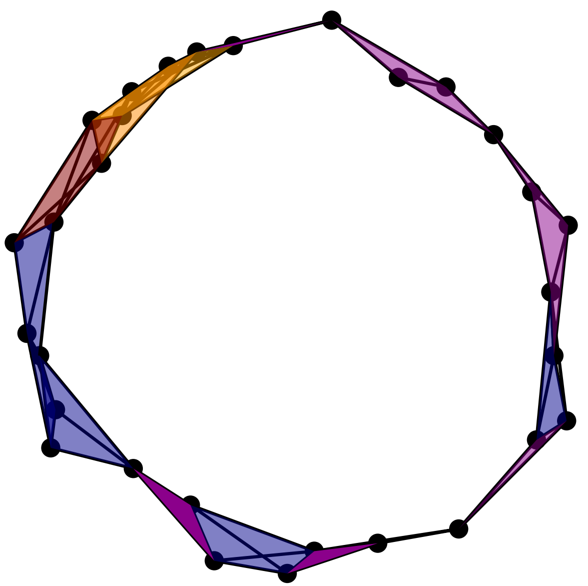
Let X be a finite subset of $\mathbb{R}^n$, and $t \geq 0$. Consider the collection

$$\mathcal{V}^t = \left\{ \overline{\mathcal{B}}\left(x,t\right), x \in X \right\}.$$

This is a cover of the thickening X^t, and each components are closed balls.
By Nerve Theorem, its nerve $\mathcal{N}(\mathcal{V}^t)$ has the homotopy type of X^t.

Definition: This nerve is denoted $\check{\mathrm{C}}\mathrm{ech}^t(X)$ and is called the **Čech complex** of X at time t.



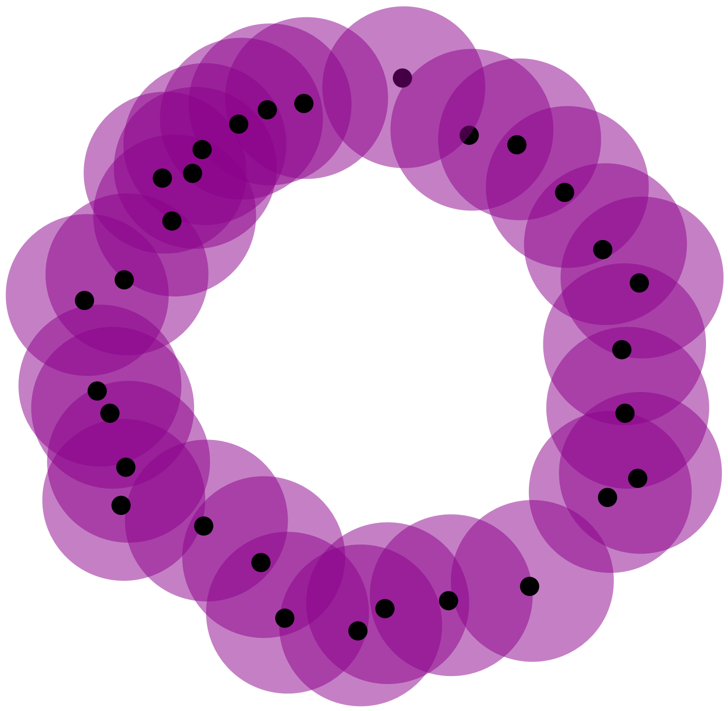


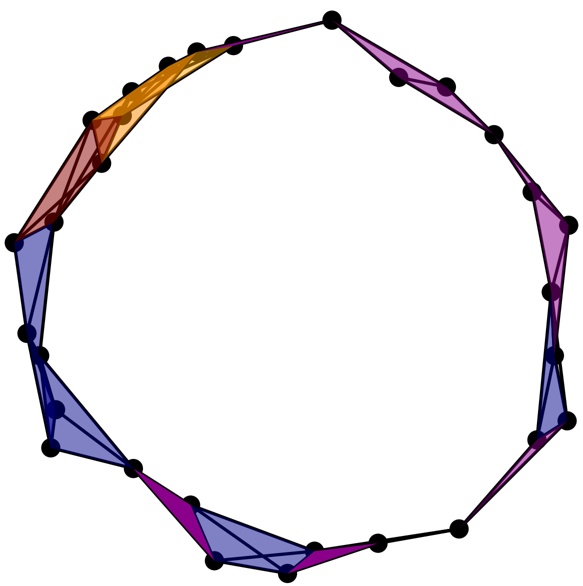
Let X be a finite subset of $\mathbb{R}^n$, and $t \geq 0$. Consider the collection

$$\mathcal{V}^t = \left\{ \overline{\mathcal{B}}\left(x, t\right), x \in X \right\}.$$

This is a cover of the thickening X^t, and each components are closed balls.
By Nerve Theorem, its nerve $\mathcal{N}(\mathcal{V}^t)$ has the homotopy type of X^t.

Definition: This nerve is denoted $\check{\mathrm{C}}\mathrm{ech}^t(X)$ and is called the **Čech complex** of X at time t.





→ The Question 2 (How to compute the homology groups of X^t?) is solved.

Computação do complexo de Čech

Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ be finite, let $t \geq 0$ and consider the t-thickening

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}\left(x, t\right).$$

By definition, its nerve, $\check{\mathrm{C}}\mathrm{ech}^t(X)$, the Čech complex at time t, is a simplicial complex on the vertices $\{1, \ldots, N\}$ whose simplices are the subsets $\{i_1, ..., i_m\}$ such that

$$\bigcap_{1 \leq k \leq m} \overline{\mathcal{B}}\left(x_{i_k}, t\right) \neq \emptyset.$$

Computação do complexo de Čech

Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ be finite, let $t \geq 0$ and consider the t-thickening

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}\,(x, t)\,.$$

By definition, its nerve, $\check{\mathrm{C}}\mathrm{ech}^t(X)$, the Čech complex at time t, is a simplicial complex on the vertices $\{1, \ldots, N\}$ whose simplices are the subsets $\{i_1, ..., i_m\}$ such that

$$\bigcap_{1 \leq k \leq m} \overline{\mathcal{B}}\,(x_{i_k}, t) \neq \emptyset.$$

Therefore, computing the Čech complex relies on the following geometric predicate:

Given m closed balls of $\mathbb{R}^n$, do they intersect?

This problem is known as the *smallest circle problem*.
It can can be solved in $O(m)$ time, where m is the number of points.

Computação do complexo de Čech

Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ be finite, let $t \geq 0$ and consider the t-thickening

$$X^t = \bigcup_{x \in X} \overline{\mathcal{B}}(x, t)\,.$$

By definition, its nerve, $\check{\mathrm{C}}\mathrm{ech}^t(X)$, the Čech complex at time t, is a simplicial complex on the vertices $\{1, \ldots, N\}$ whose simplices are the subsets $\{i_1, \ldots, i_m\}$ such that

$$\bigcap_{1 \leq k \leq m} \overline{\mathcal{B}}(x_{i_k}, t) \neq \emptyset.$$

Therefore, computing the Čech complex relies on the following geometric predicate:

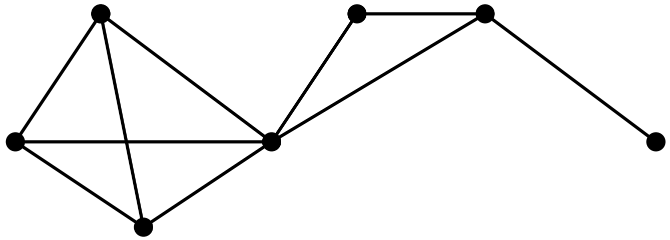
Given m closed balls of $\mathbb{R}^n$, do they intersect?

This problem is known as the *smallest circle problem*.
It can can be solved in $O(m)$ time, where m is the number of points.

⟶ in practice, we prefer a more simple version

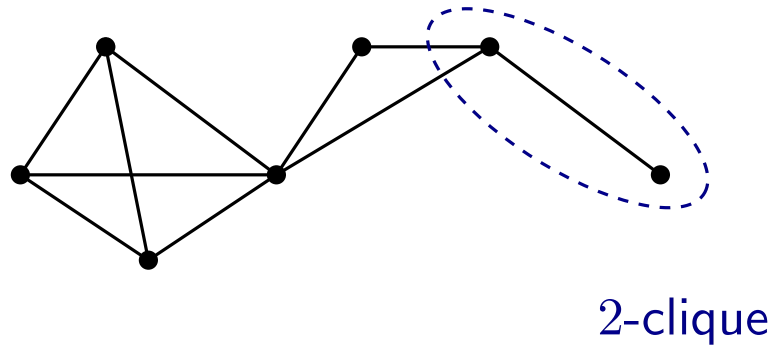
Complexo de clique

Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.



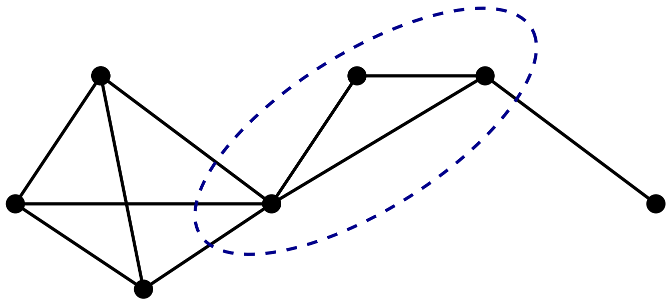
Complexo de clique

Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.



Complexo de clique

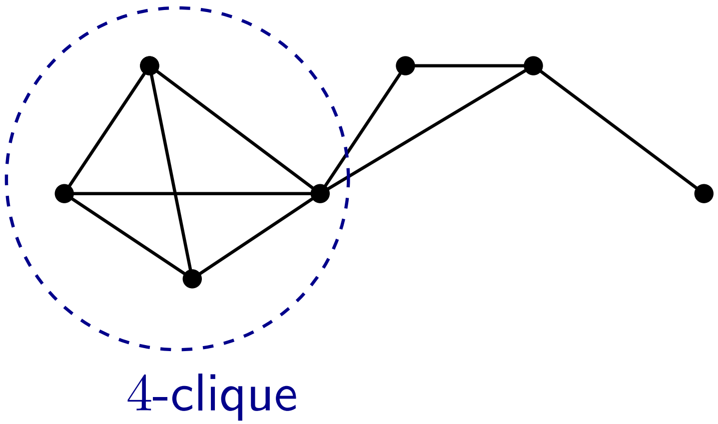
Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.



3-clique

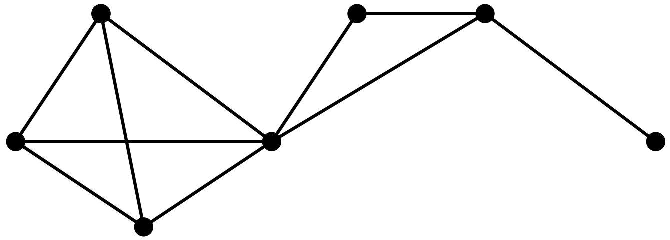
Complexo de clique

Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.



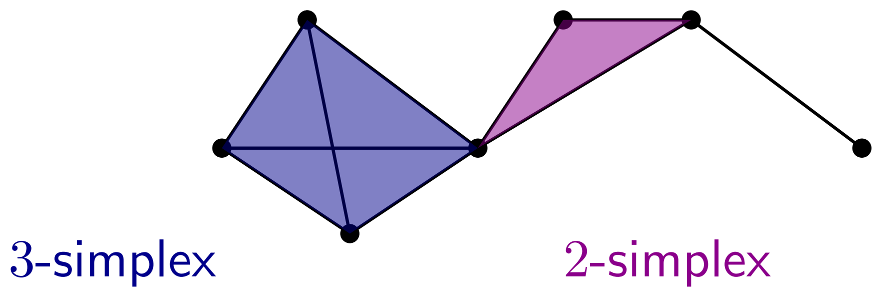
Complexo de clique

Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.

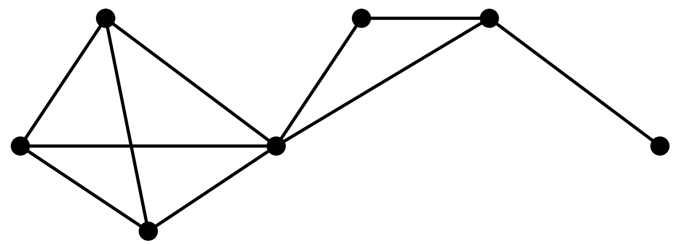


Definition: Given a graph G, the corresponding **clique complex** is the simplicial complex whose

- vertices are the vertices of G,
- simplices are the sets of vertices of the cliques of G.

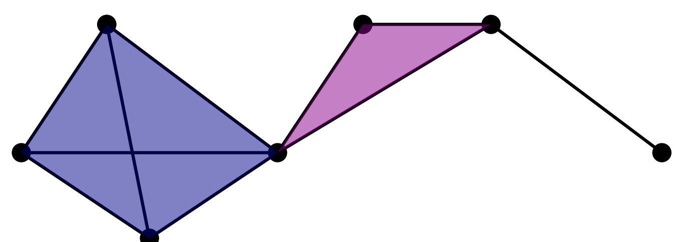


Let G be a graph.
We call a **clique** of G a set of vertices $v_1, ..., v_m$ such that for every $i, j \in [\![1, m]\!]$ with $i \neq j$, the edge $[v_i, v_j]$ belongs to G.



Definition: Given a graph G, the corresponding **clique complex** is the simplicial complex whose

- vertices are the vertices of G,
- simplices are the sets of vertices of the cliques of G.



3-simplex

2-simplex

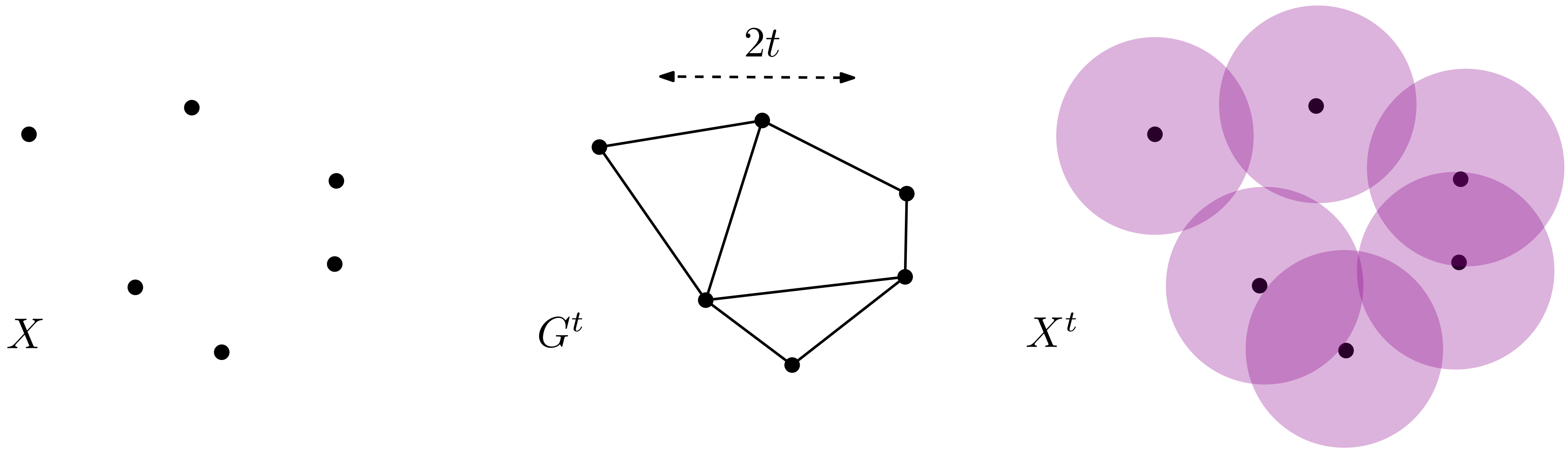
Observation: The clique complex of a graph is a simplicial complex.

Complexo de Rips

Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ and $t \geq 0$.

Consider the graph G^t whose vertex set is $\{1, \ldots, N\}$, and whose edges are the pairs (i, j) such that $\|x_i - x_j\| \leq 2t$.

Alternatively, G^t can be seen as the 1-skeleton of the Čech complex $\check{\mathrm{C}}\mathrm{ech}^t(X)$.

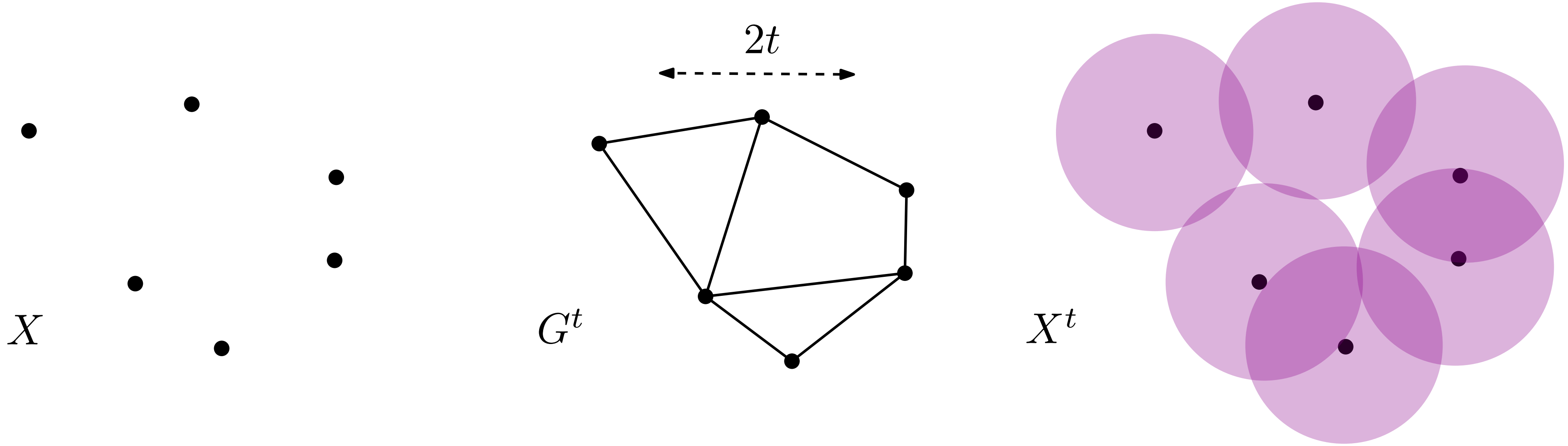


Complexo de Rips

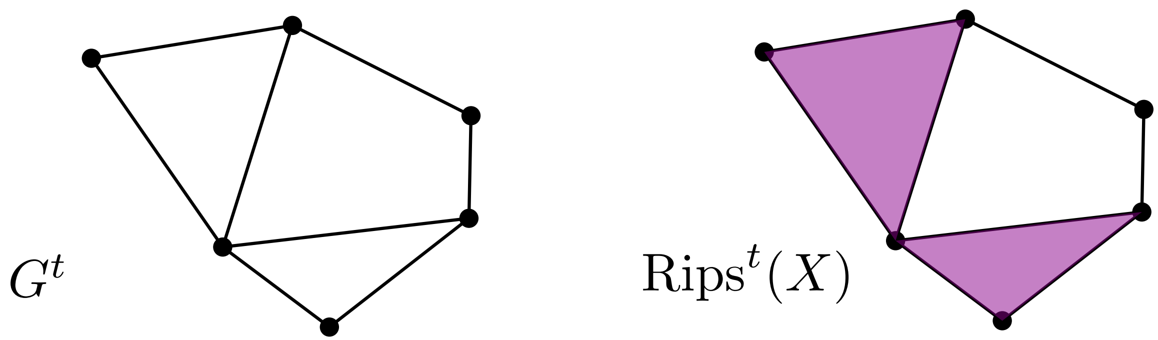
Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ and $t \geq 0$.

Consider the graph G^t whose vertex set is $\{1, \ldots, N\}$, and whose edges are the pairs (i, j) such that $\|x_i - x_j\| \leq 2t$.

Alternatively, G^t can be seen as the 1-skeleton of the Čech complex $\check{\mathrm{C}}\mathrm{ech}^t(X)$.



Definition: The **Rips complex** of X at time t is the clique complex of the graph G^t. We denote it $\mathrm{Rips}^t(X)$.

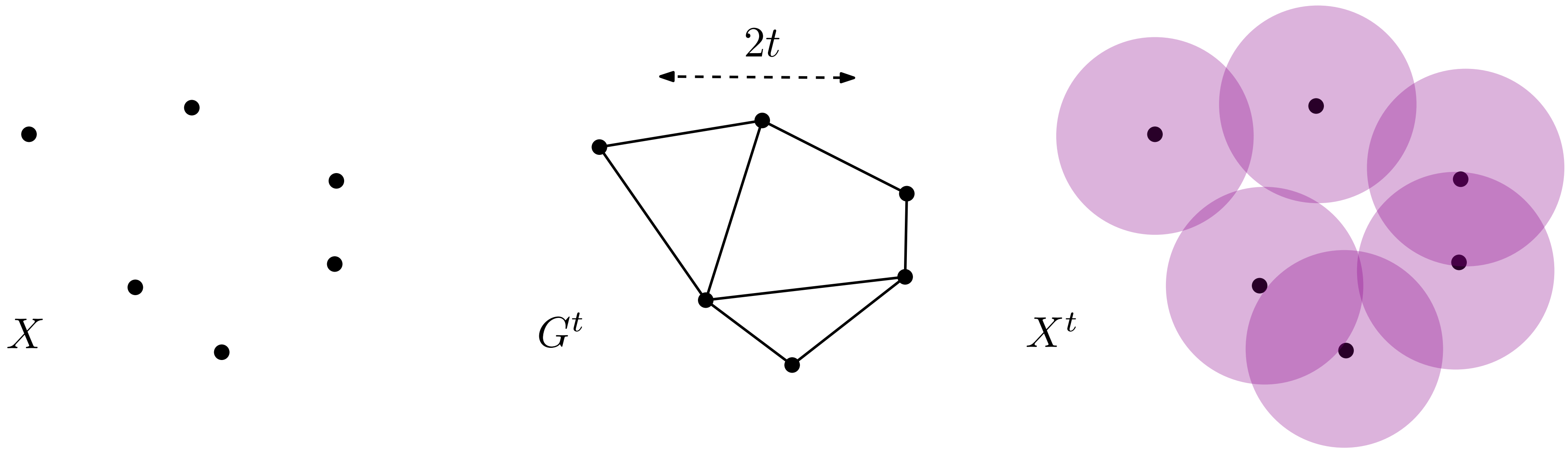


Complexo de Rips

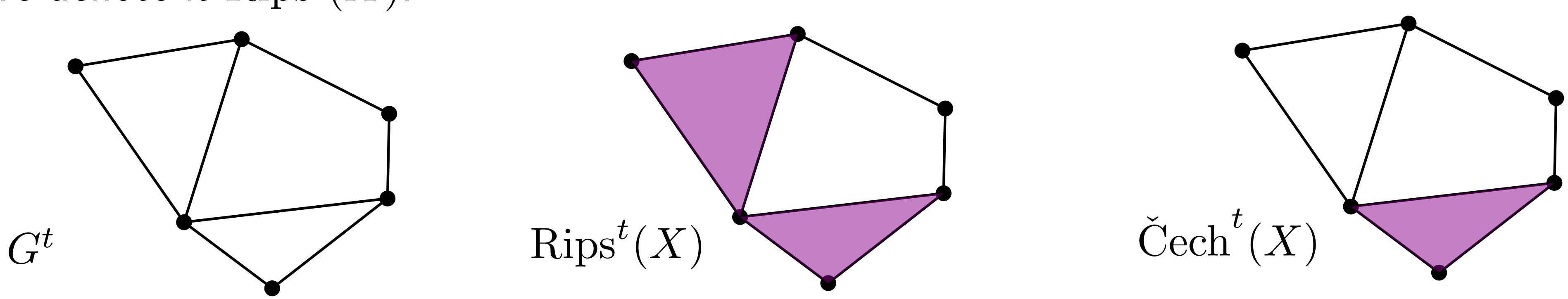
Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n$ and $t \geq 0$.

Consider the graph G^t whose vertex set is $\{1, \ldots, N\}$, and whose edges are the pairs (i, j) such that $\|x_i - x_j\| \leq 2t$.

Alternatively, G^t can be seen as the 1-skeleton of the Čech complex $\check{\mathrm{C}}\mathrm{ech}^t(X)$.



Definition: The **Rips complex** of X at time t is the clique complex of the graph G^t. We denote it $\mathrm{Rips}^t(X)$.

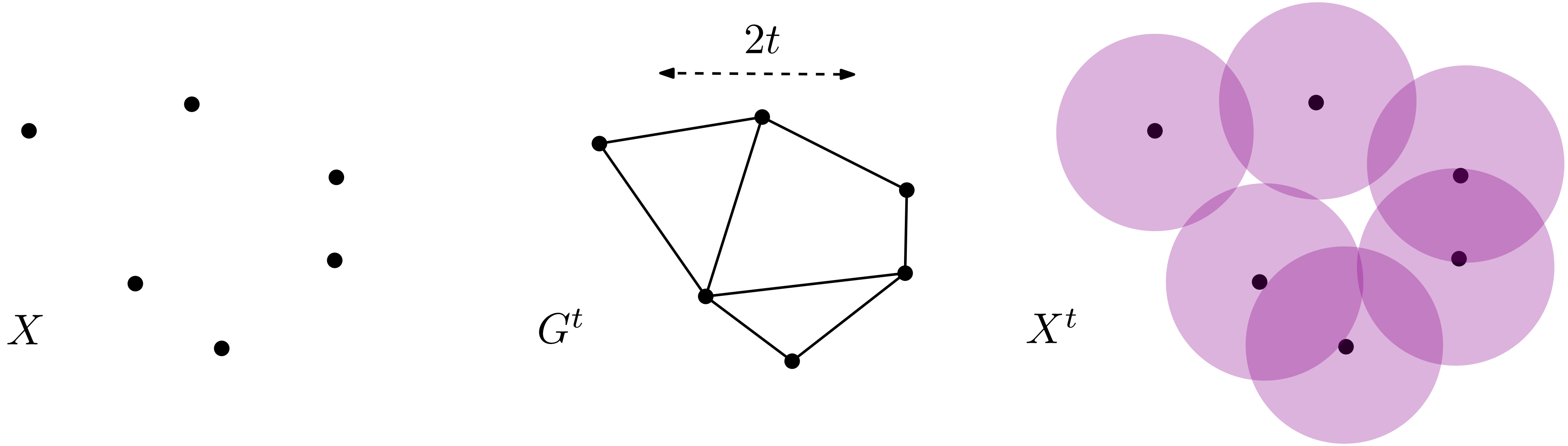


Complexo de Rips

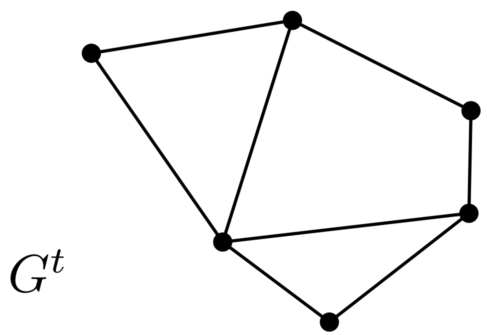
Let $X = \{x_1, \dots, x_N\} \subset \mathbb{R}^n$ and $t \geq 0$.

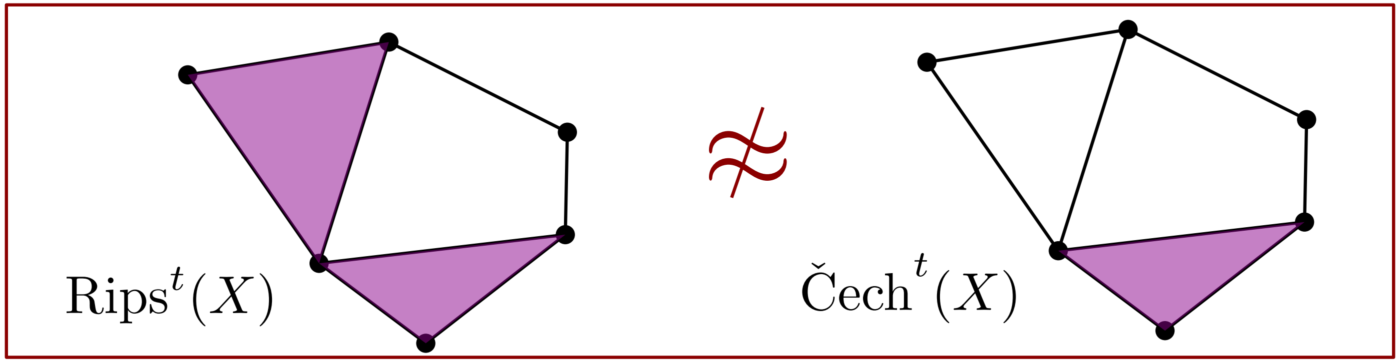
Consider the graph G^t whose vertex set is $\{1, \dots, N\}$, and whose edges are the pairs (i, j) such that $\|x_i - x_j\| \leq 2t$.

Alternatively, G^t can be seen as the 1-skeleton of the Čech complex $\check{\mathrm{C}}\mathrm{ech}^t(X)$.



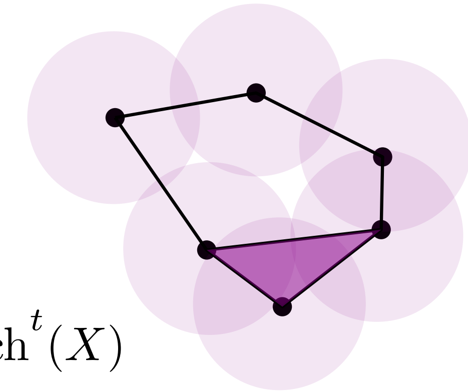
Definition: The **Rips complex** of X at time t is the clique complex of the graph G^t. We denote it $\mathrm{Rips}^t(X)$.



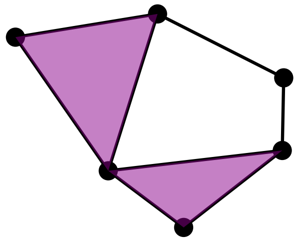


Proposition: For every $t \geq 0$, we have

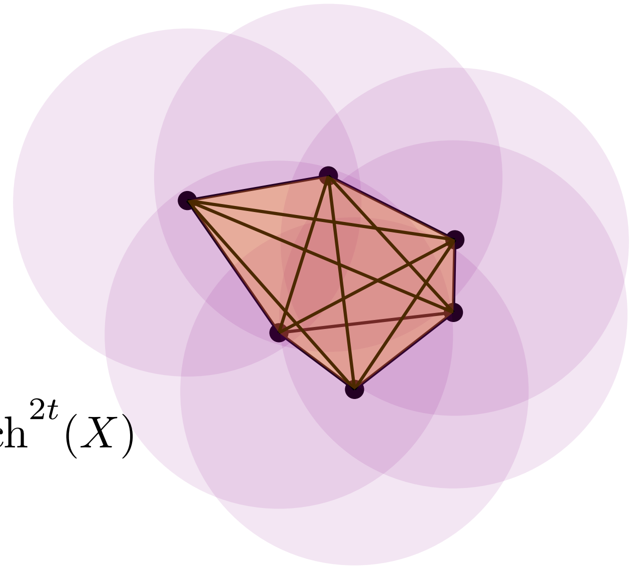
$$\check{\mathrm{C}}\mathrm{ech}^{t}(X) \subset \mathrm{Rips}^{t}(X) \subset \check{\mathrm{C}}\mathrm{ech}^{2t}(X).$$



$\check{\mathrm{C}}\mathrm{ech}^{t}(X)$



$\mathrm{Rips}^{t}(X)$

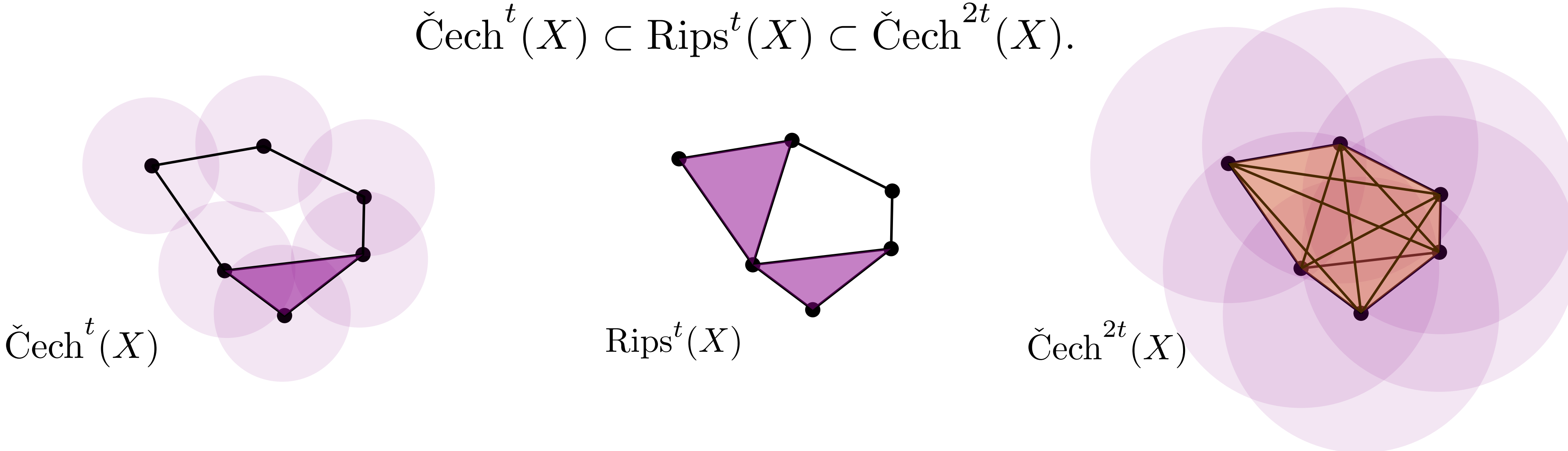


$\check{\mathrm{C}}\mathrm{ech}^{2t}(X)$

Complexo de Rips

Proposition: For every $t \geq 0$, we have

$$\check{\mathrm{C}}\mathrm{ech}^t(X) \subset \mathrm{Rips}^t(X) \subset \check{\mathrm{C}}\mathrm{ech}^{2t}(X).$$



Proof: Let $t \geq 0$. The first inclusion follows from the fact that $\mathrm{Rips}^t(X)$ is the clique complex of $\check{\mathrm{C}}\mathrm{ech}^t(X)$.

To prove the second one, choose a simplex $\sigma \in \mathrm{Rips}^t(X)$. Let us prove that $\omega \in \check{\mathrm{C}}\mathrm{ech}^{2t}(X)$.

Let $x \in \sigma$ be any vertex. Note that $\forall y \in \sigma$, we have $\|x - y\| \leq 2t$ by definition of the Rips complex. Hence

$$x \in \bigcap_{y \in \sigma} \overline{\mathcal{B}}\left(y, 2t\right).$$

The intersection being non-empty, we deduce $\sigma \in \check{\mathrm{C}}\mathrm{ech}^{2t}(X)$.

Conclusão

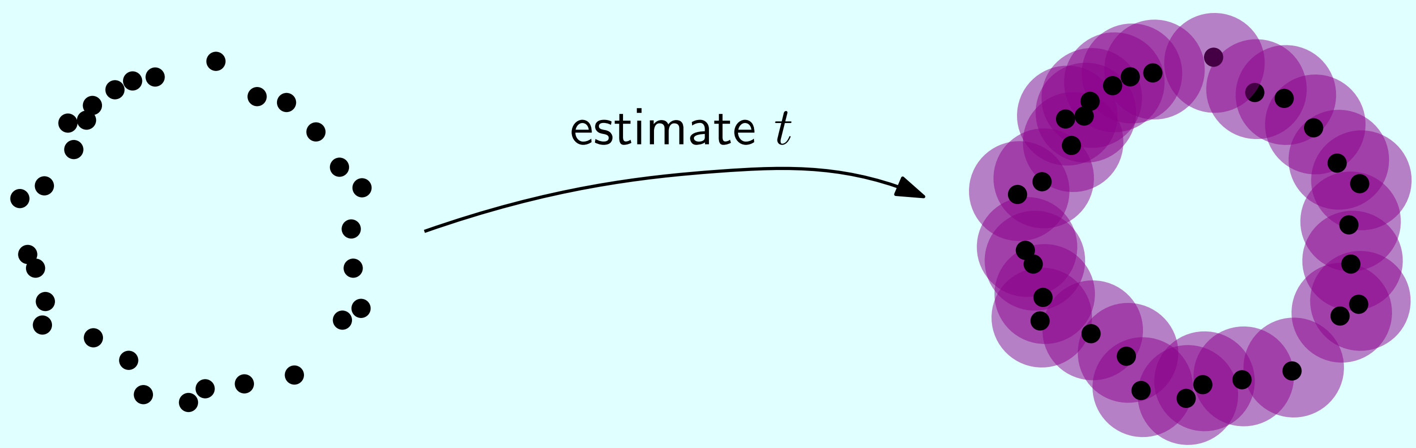
Question 1: How to select a t such that $X^t \approx \mathcal{M}$?

Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

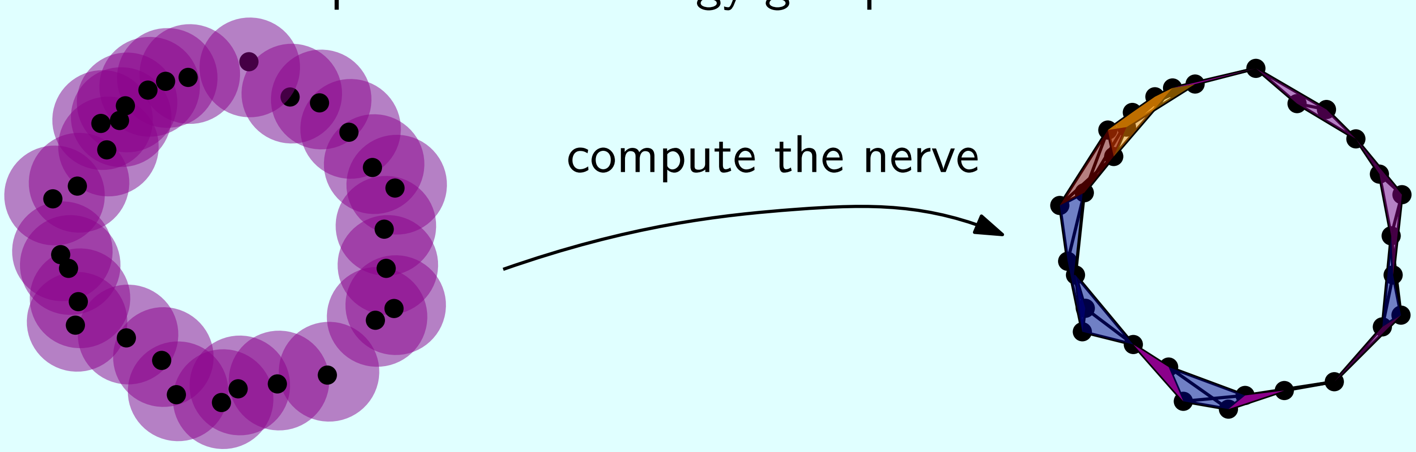
$$t \in [4\mathrm{d_H}(X, \mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M}))\,.$$

Theorem (Partha Niyogi, Stephen Smale, and Shmuel Weinberger, 2008):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$. Suppose that $\mathcal{M}$ has positive reach.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \left[2\mathrm{d_H}(X, \mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right)\,.$$



Question 2: How to compute the homology groups of X^t?



Conclusão

Question 1: How to select a t such that $X^t \approx \mathcal{M}$?

Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

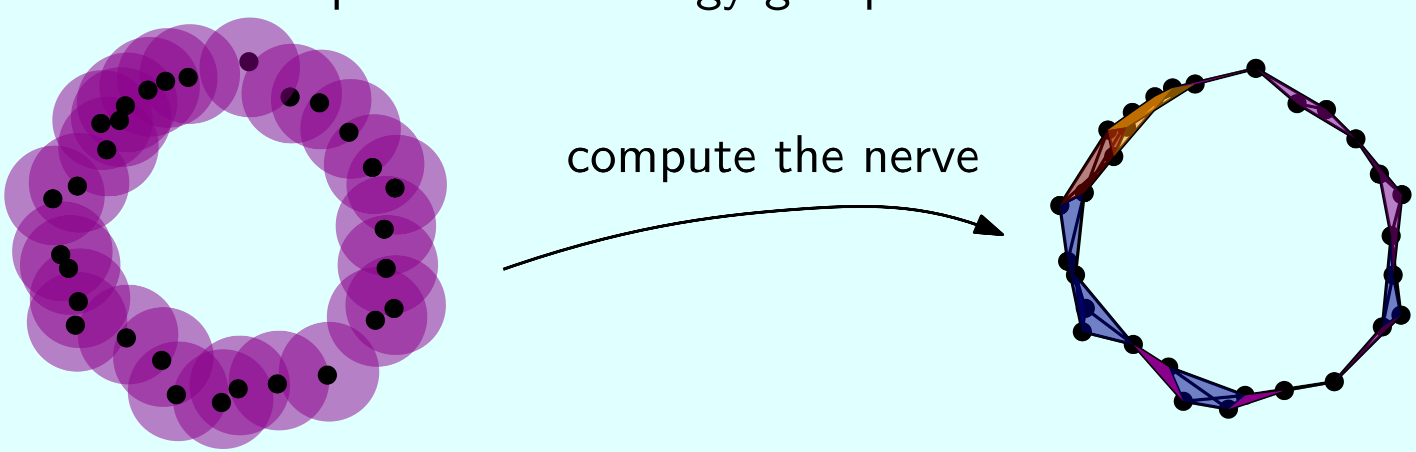
$$t \in \boxed{[4\mathrm{d_H}(X, \mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M}))\,.}$$

Theorem (Partha Niyogi, Stephen Smale, and Shmuel Weinberger, 2008):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$. Suppose that $\mathcal{M}$ has positive reach.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \boxed{\left[2\mathrm{d_H}(X, \mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right)\,.}$$

these quantities are not known!

Question 2: How to compute the homology groups of X^t?



Conclusão

Question 1: How to select a t such that $X^t \approx \mathcal{M}$?

Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

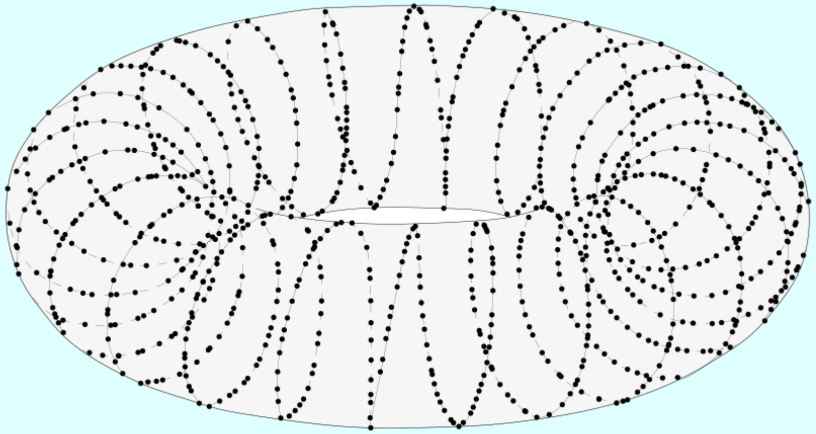
$$t \in [4\mathrm{d_H}(X, \mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X, \mathcal{M}))\,.$$

Theorem (Partha Niyogi, Stephen Smale, and Shmuel Weinberger, 2008):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$. Suppose that $\mathcal{M}$ has positive reach.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \left[2\mathrm{d_H}(X, \mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right).$$

these quantities are not known!

Is this object 1- or 2-dimensional?



Conclusão

Question 1: How to select a t such that $X^t \approx \mathcal{M}$?

Theorem (Frédéric Chazal, David Cohen-Steiner, and André Lieutier, 2009):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$. Suppose that $\mathcal{M}$ has positive reach, and that $\mathrm{d_H}(X, \mathcal{M}) \leq \frac{1}{17}\mathrm{reach}(\mathcal{M})$.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

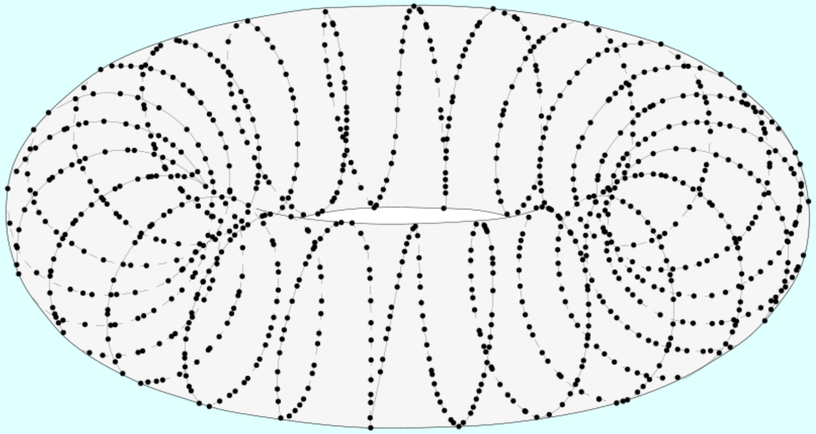
$$t \in \boxed{[4\mathrm{d_H}(X,\mathcal{M}), \mathrm{reach}(\mathcal{M}) - 3\mathrm{d_H}(X,\mathcal{M})).}$$

Theorem (Partha Niyogi, Stephen Smale, and Shmuel Weinberger, 2008):
Let X and $\mathcal{M}$ be subsets of $\mathbb{R}^n$, with $\mathcal{M}$ a submanifold, and X a finite subset of $\mathcal{M}$. Suppose that $\mathcal{M}$ has positive reach.
Then X^t and $\mathcal{M}$ are homotopic equivalent, provided that

$$t \in \boxed{\left[2\mathrm{d_H}(X,\mathcal{M}), \sqrt{\frac{3}{5}}\mathrm{reach}(\mathcal{M})\right).}$$

these quantities are not known!

Is this object 1- or 2-dimensional?



Idea (multiscale analysis): Instead of estimating a value of t, we will choose all of them.