



AIDOS LAB
AI FOR DATA-ORIENTED SCIENCE



Lecture 6: Persistent Homology

Geometry and Topology in Machine Learning

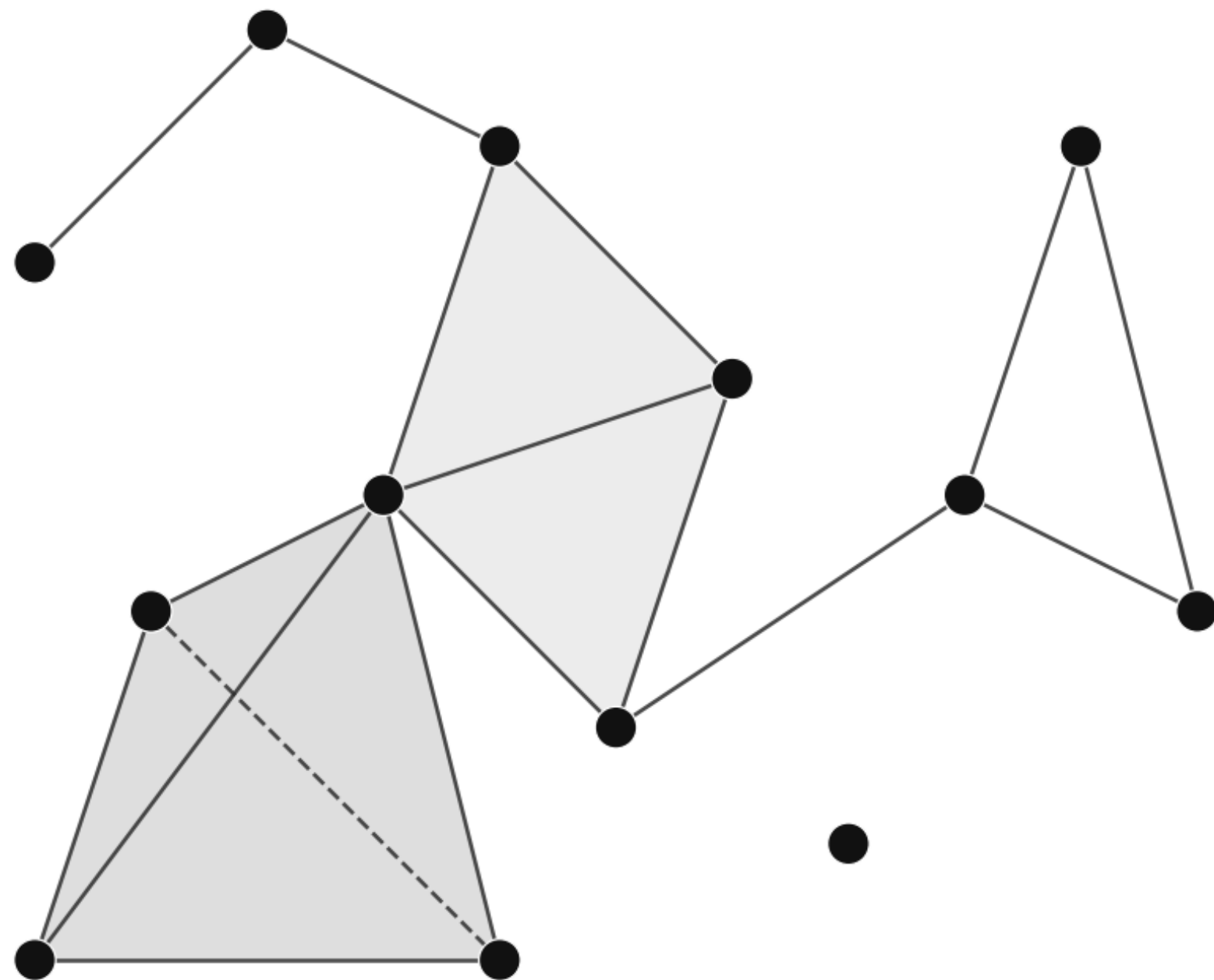
Inés García Redondo — October 20, 2026

Recap: simplicial complexes

Definition

Def. (Abstract) Simplicial Complex

Given a discrete set $S = \{s_1, \dots, s_n\}$, a simplicial complex X is a set of subsets $\sigma \subset S$ closed under inclusion (if $\sigma \in X$, then for all $\tau \subseteq \sigma$ we must have $\tau \in X$)



We can *stratify* a simplicial complex by dimension:

Vertices: 0-dimensional simplices

Edges: 1-dimensional simplices

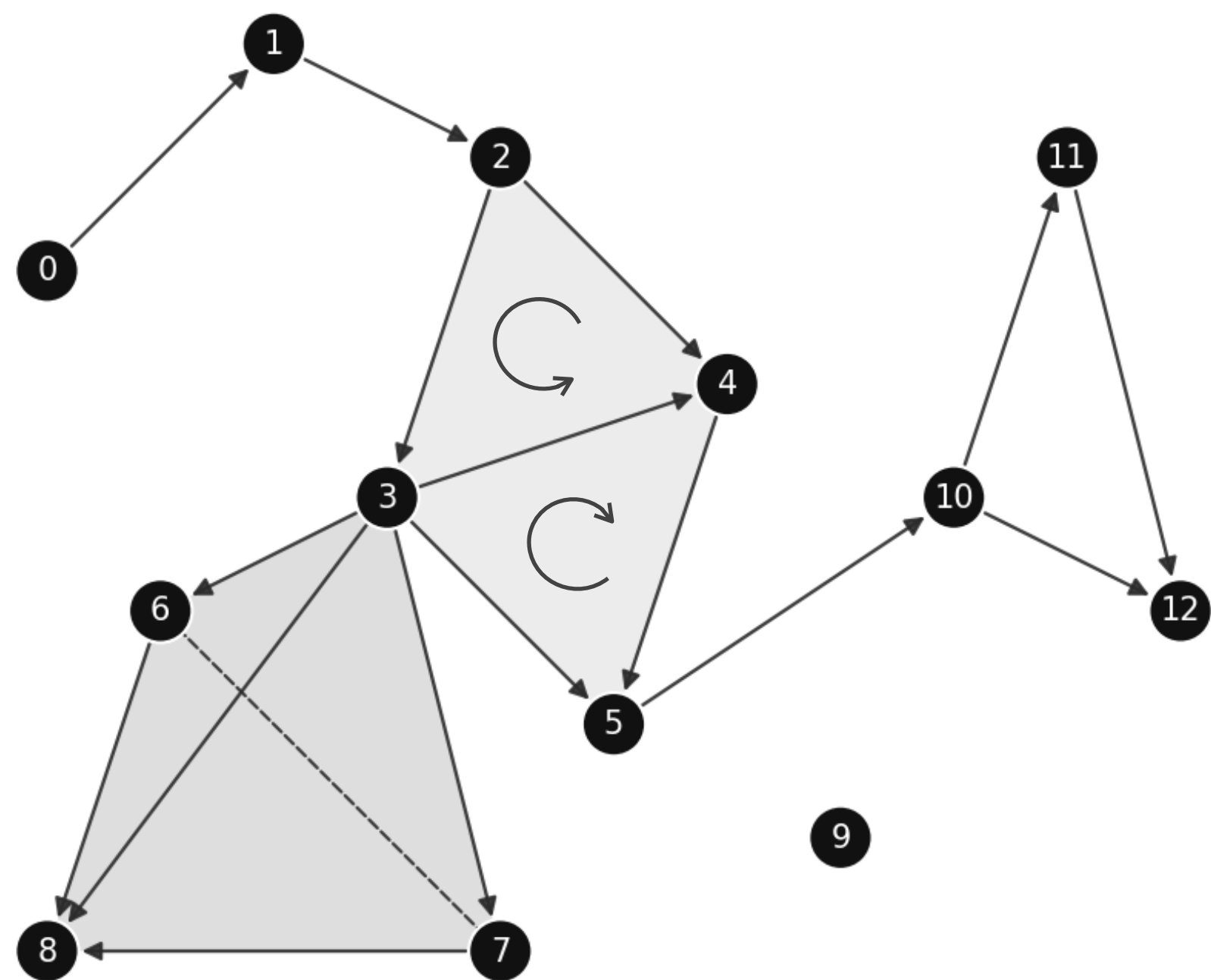
Triangles: 2-dimensional simplices

In general:

- **k -dimensional simplex:** $\sigma = \{s_{i_0}, \dots, s_{i_k}\}$ $k + 1$ elements!
- X_k : set of k -dimensional simplices
- **k -skeleton** ($X_{\leq k}$): all simplices with dimension $\leq k$

Recap: simplicial homology

Orientations



The first thing we do is choosing an **orientation** for the simplices, that is, a unique way of representing a set $\sigma = \{s_{i_0}, \dots, s_{i_k}\}$

Note that we use the same notation that we had in graphs to distinguish oriented and un-oriented simplices!

How to choose an orientation?

- Choose an ordering of the vertices $S = \{1, \dots, n\}$
- Orient higher order simplices accordingly: $(s_{i_0}, \dots, s_{i_k}), i_0 < \dots < i_k$
- We will simply denote the simplices as: $(s_{i_0}, \dots, s_{i_k}) \equiv (i_0, \dots, i_k)$

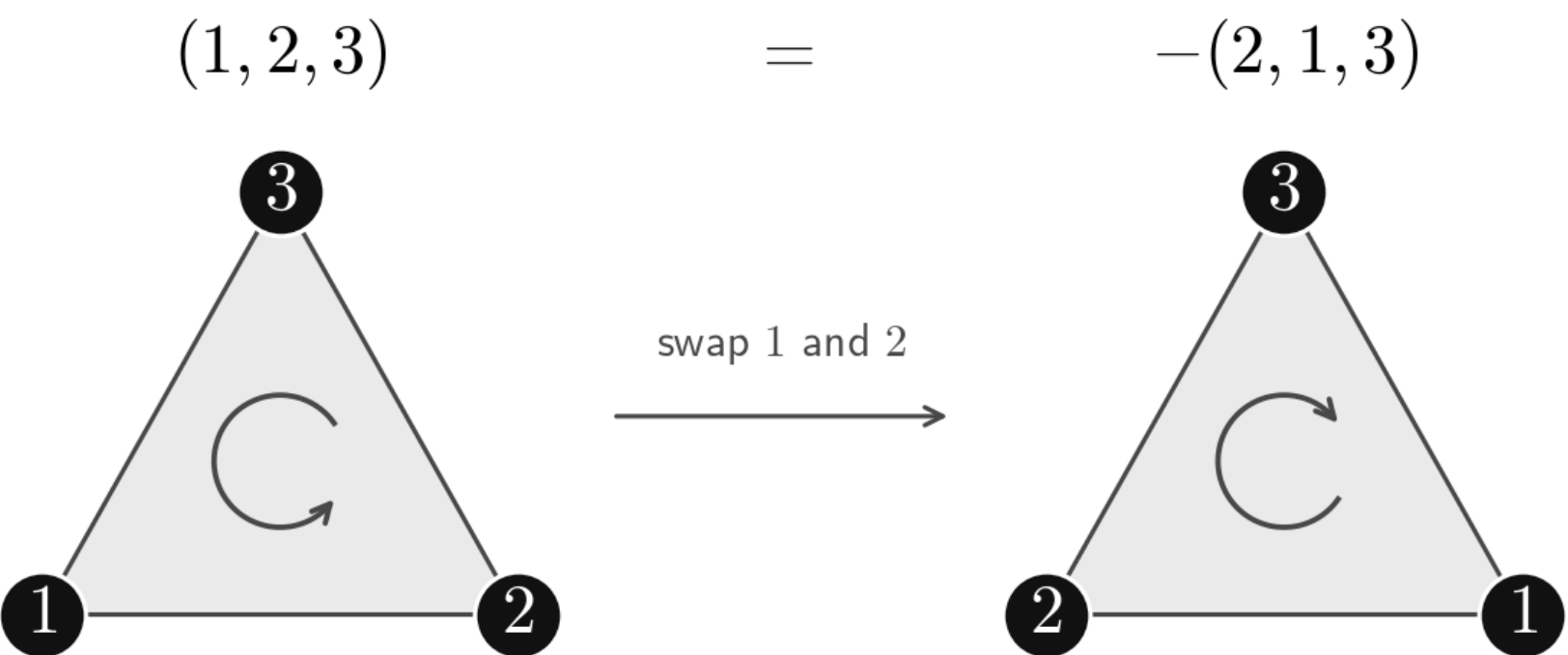
■ Any other ordering of the vertices is related to this one by:

$$(i_0, \dots, i_k) = (-1)^{\mathfrak{p}}(j_0, \dots, j_k)$$

Where $\{s_{i_0}, \dots, s_{i_k}\} = \{s_{j_0}, \dots, s_{j_k}\}$ (the simplices contain the same vertices) and $\mathfrak{p}$ is the number of pairwise permutations to go from $(s_{j_0}, \dots, s_{j_k})$ to $(s_{i_0}, \dots, s_{i_k})$

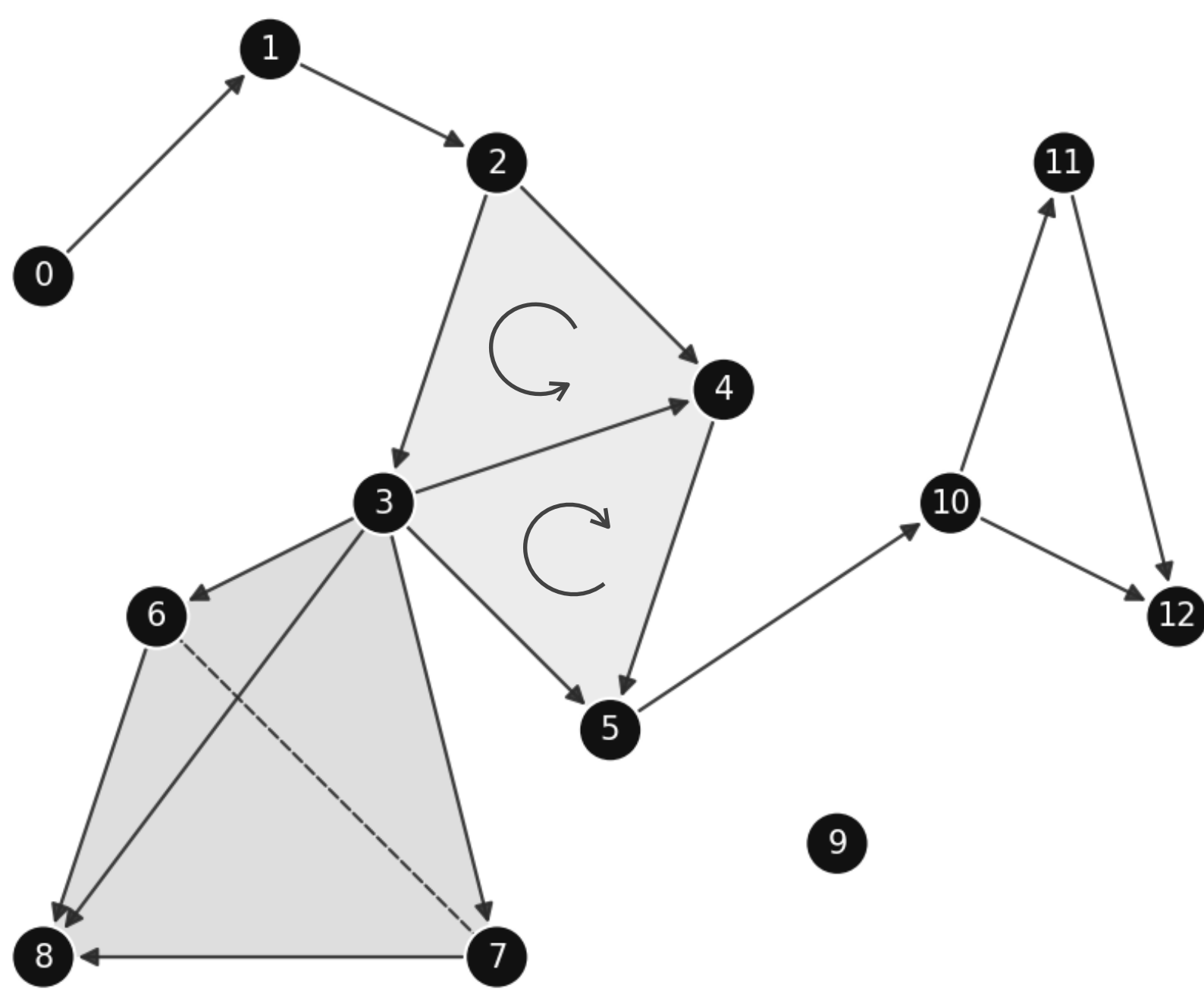
Examples:

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



Recap: simplicial homology

Chain spaces



Chain space: let $k \geq 0$, the k -chain space is the *vector space* spanned by the *oriented simplices*

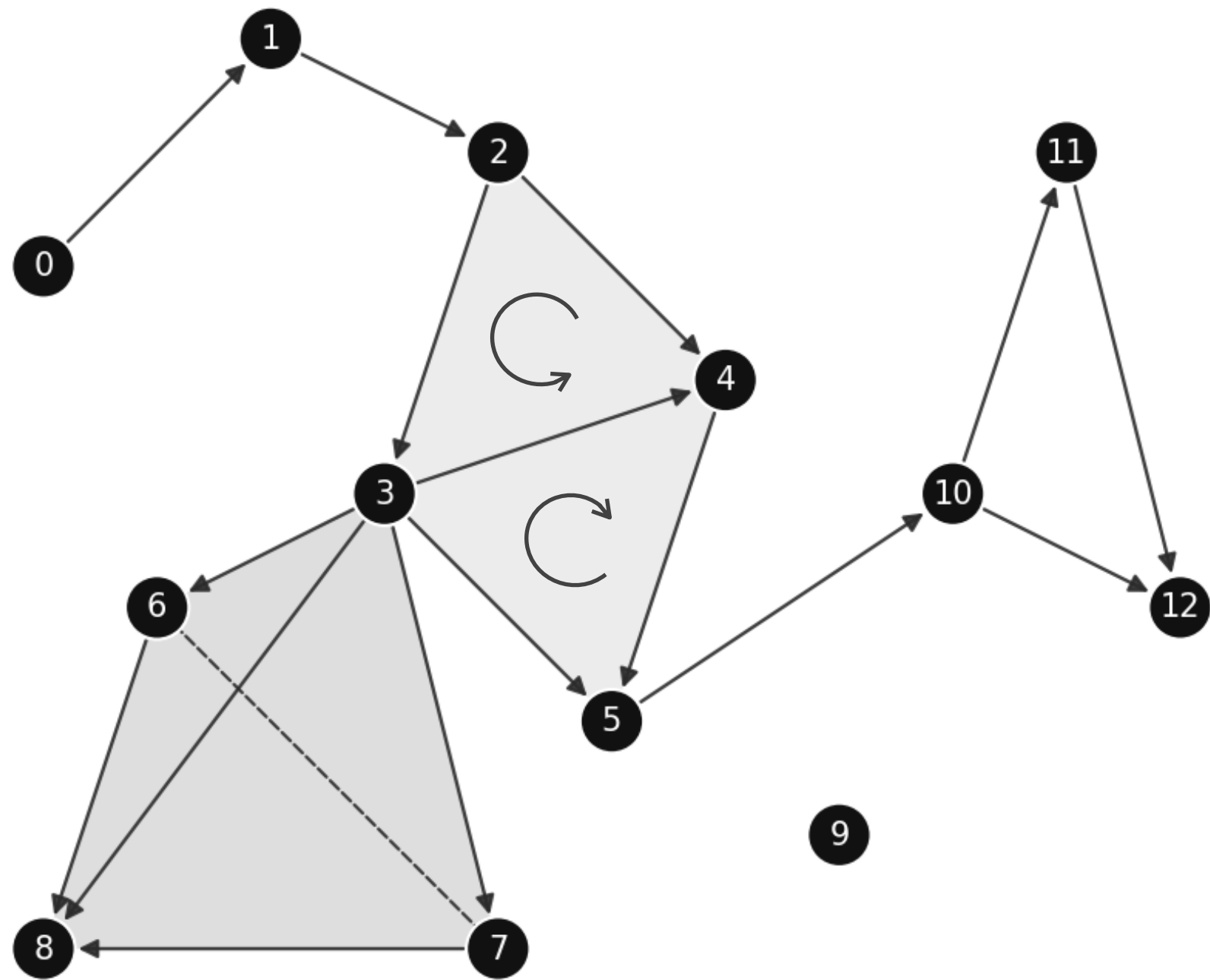
$$C_k(X) = \text{span}\{(i_0, \dots, i_k) : i_0 < \dots < i_k\} \cong \mathbb{R}^{|X_k|}$$

$$c = \sum_{\sigma \in X_k} a_\sigma \cdot \sigma, \quad a_\sigma \in \mathbb{R}$$

Try it: What are the chain groups of the complex on the left?

Recap: simplicial homology

Boundary operators



Boundary operator: *linear map* connecting chain spaces of consecutive dimension $\partial_k : C_k(X) \rightarrow C_{k-1}(X)$ defined on the basis elements as

$$\partial_k(i_0, \dots, i_k) = \sum_{j=0}^k (-1)^j (i_0, \dots, \widehat{i_j}, \dots, i_k)$$

And extended linearly:

$$c = \sum_{\sigma \in X_k} a_{\sigma} \cdot \sigma, \quad a_{\sigma} \in \mathbb{R}$$

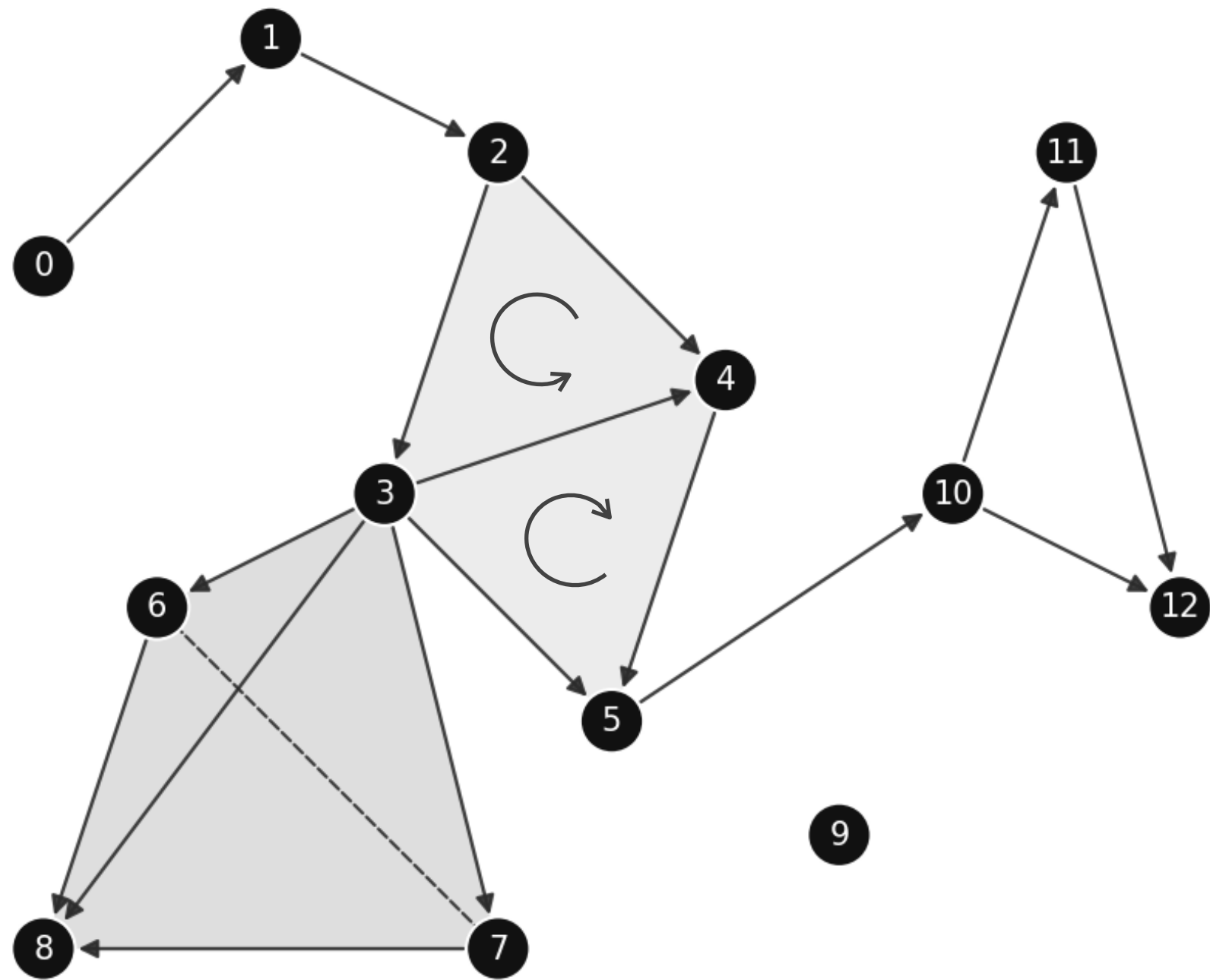
$$\partial_k c = \sum_{\sigma \in X_k} a_{\sigma} \cdot \partial_k \sigma, \quad a_{\sigma} \in \mathbb{R}$$

Try it:

$$\begin{aligned} \partial_1(0, 1) &= \\ \partial_2(2, 3, 4) &= \\ \partial_3(3, 6, 7, 8) &= \end{aligned}$$

Recap: simplicial homology

Boundary operators



Boundary operator: *linear map* connecting chain spaces of consecutive dimension $\partial_k : C_k(X) \rightarrow C_{k-1}(X)$ defined on the basis elements as

$$\partial_k(i_0, \dots, i_k) = \sum_{j=0}^k (-1)^j (i_0, \dots, \widehat{i_j}, \dots, i_k)$$

And extended linearly:

$$c = \sum_{\sigma \in X_k} a_{\sigma} \cdot \sigma, \quad a_{\sigma} \in \mathbb{R}$$

$$\partial_k c = \sum_{\sigma \in X_k} a_{\sigma} \cdot \partial_k \sigma, \quad a_{\sigma} \in \mathbb{R}$$

Try it:

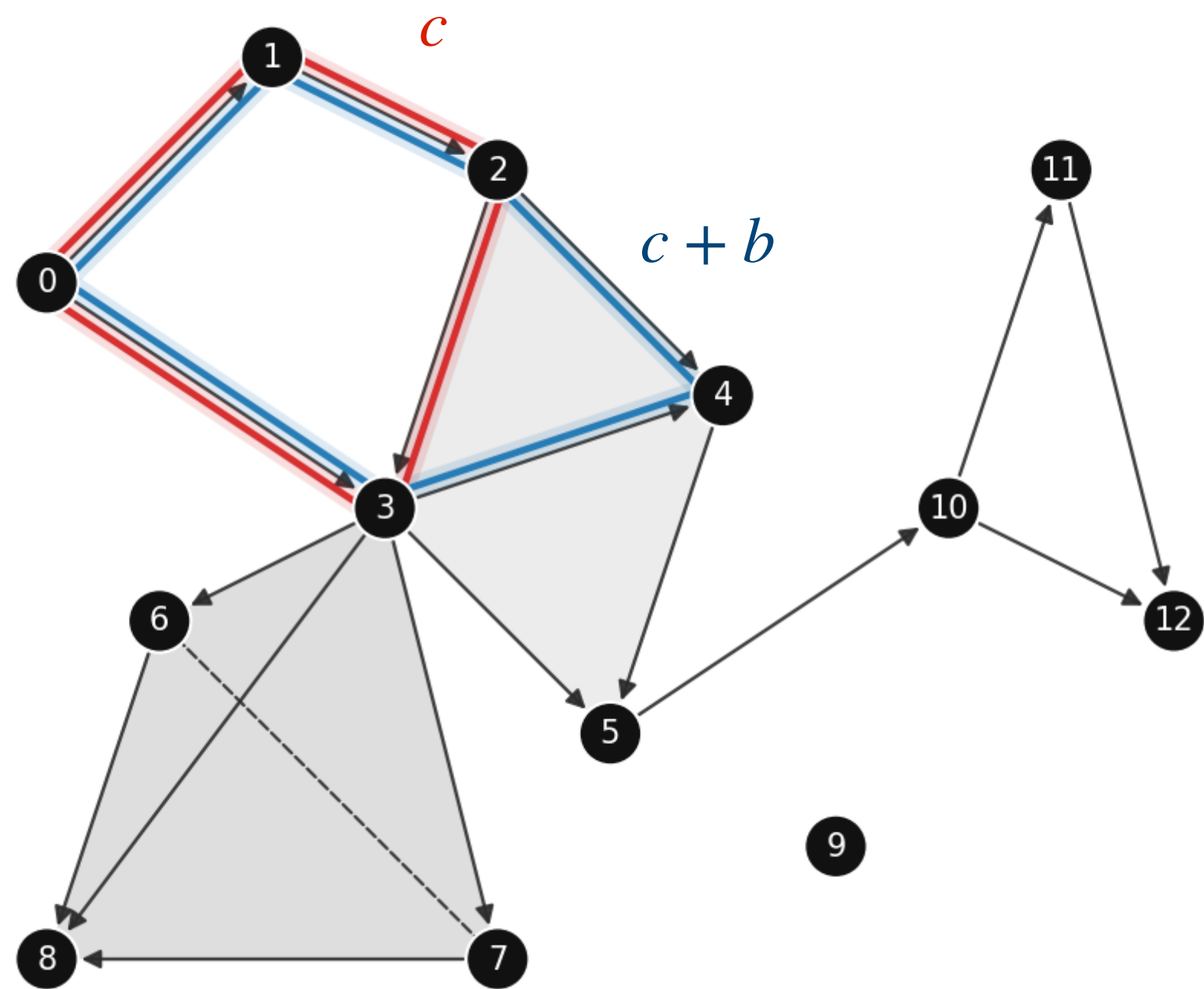
$$\begin{aligned} \partial_1(0, 1) &= \\ \partial_2(2, 3, 4) &= \\ \partial_3(3, 6, 7, 8) &= \end{aligned}$$

Important property:

$$\partial_k \partial_{k+1} = 0 \implies \ker \partial_k \subseteq \operatorname{im} \partial_{k+1}$$

Recap: simplicial homology

Homology groups



Def. Homology group

$$H_k(X) = \frac{\ker \partial_k}{\text{im } \partial_{k+1}}$$

Cycles

Boundaries

Example: What are the elements of the homology group?

First, we **find a cycle**:

$$c = (0,1) + (1,2) + (2,3) - (0,3) \in \ker \partial_1$$

$$\partial_1(c) = (1) - (0) + (2) - (1) + (3) - (2) - (3) + (0) = 0$$

An **element of** $H_1(X)$ is given by the whole subset:

$$[c] := c + \text{im } \partial_2 = \{c + \partial_2 \beta : \beta \in C_2(X)\} \quad \text{Homology class}$$

For instance, an **element of this subset** is:

$$c + b = (0,1) + (1,2) + (2,4) - (3,4) - (0,3)$$

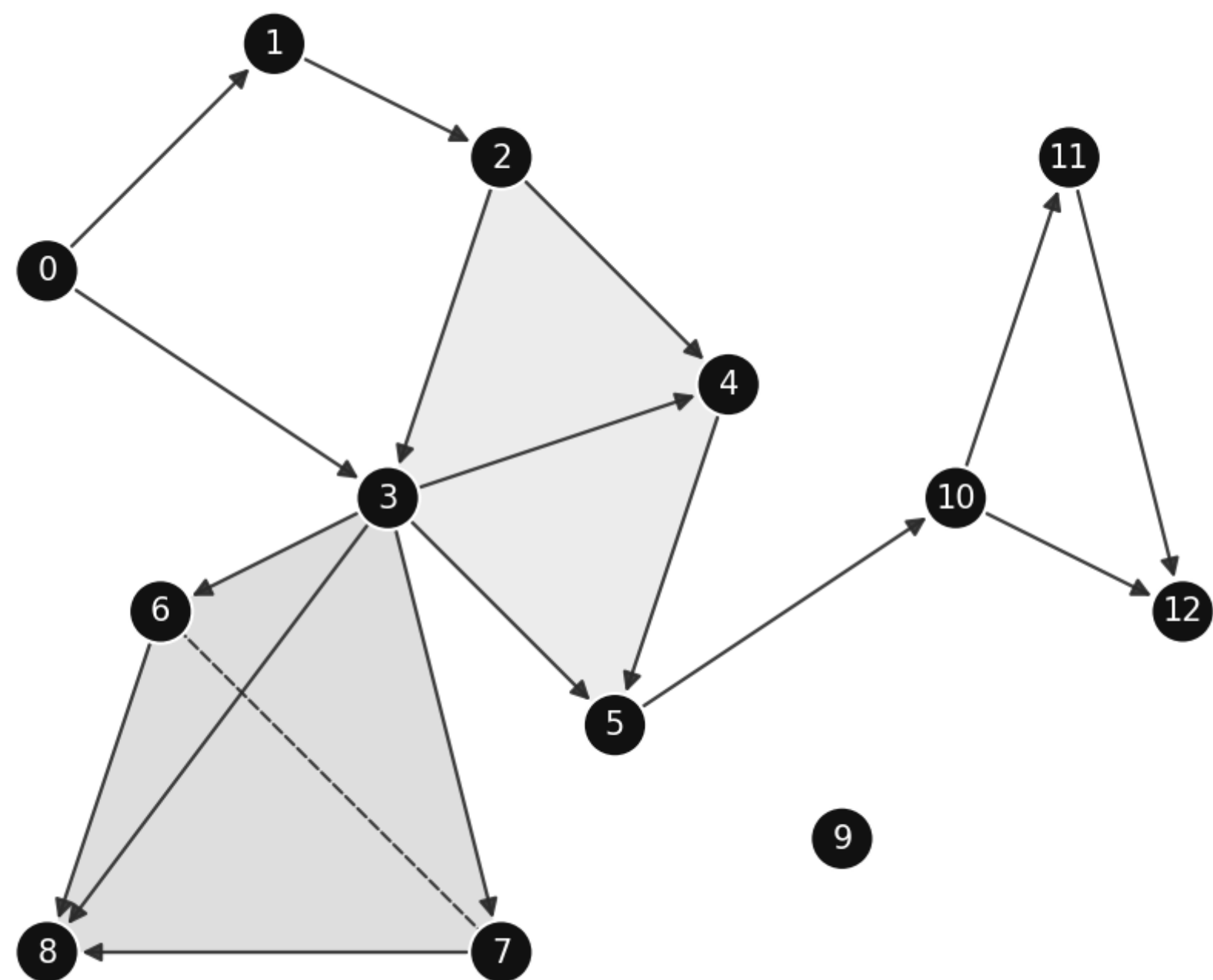
$$b = - (2,3) - (3,4) + (2,4) = \partial_2[- (2,3,4)] \in \text{im } \partial_2$$

We call homology representatives to the elements of the homology class

If I can move through filled in triangles, I stay inside of the homology class

Recap: simplicial homology

Betti numbers



Def. Betti number

$$\beta_k(X) = \dim H_k(X)$$

“Number of k -dimensional holes in X ”

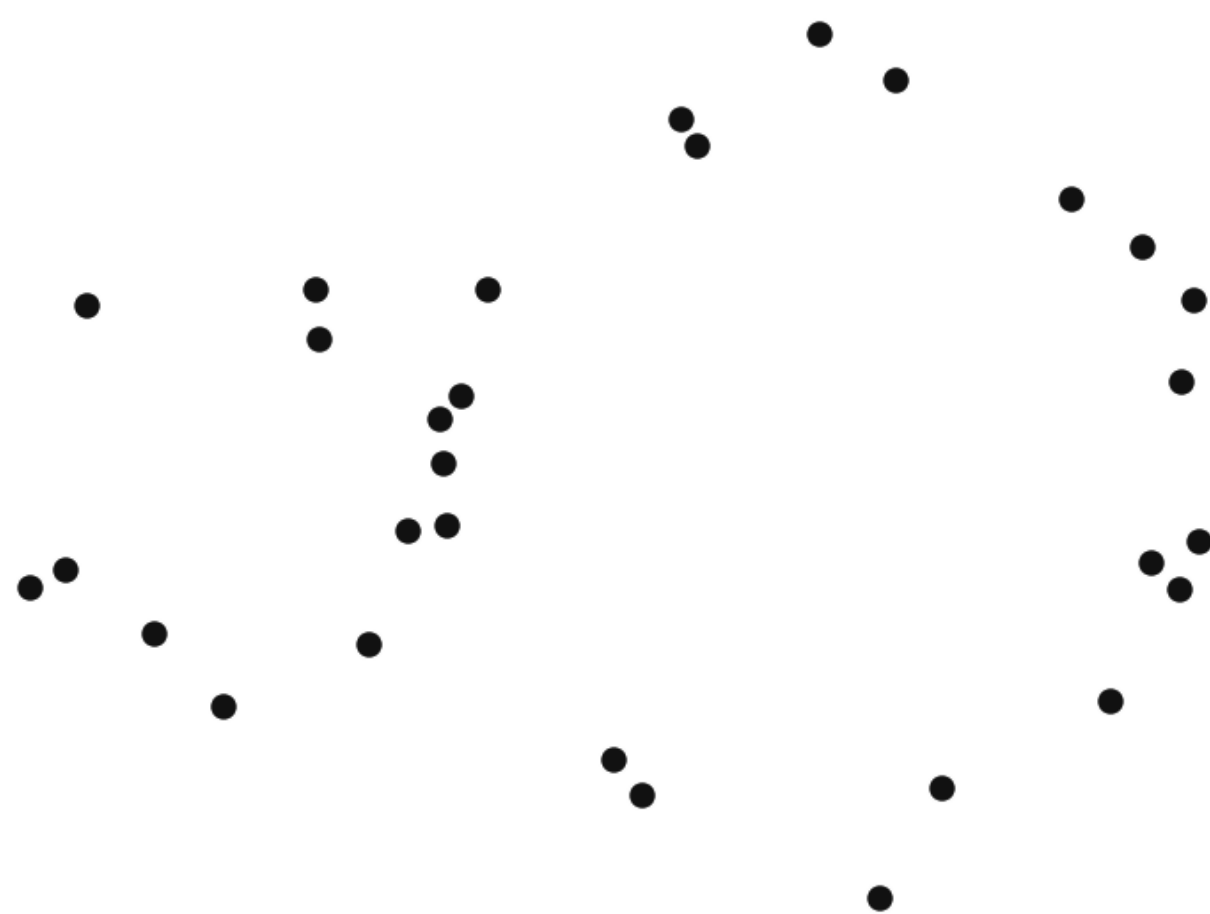
- $k = 0$: number of connected components
- $k = 1$: number of loops
- $k = 2$: number of cavities or bubbles

Try it: What are the Betti numbers of the complex on the left?

Motivation: what we want to solve today

Topological inference problem

Suppose you are given the following point cloud, how can you design a test to conclude that this sample comes from a distribution over two circles?



Session 6 — Persistent Homology

Outline

- Induced maps in homology
- How to build simplicial complexes from point clouds
- Persistent Homology
 - Filtrations
 - Persistence modules
 - Barcodes

Induced maps in homology

Session 6 — Persistent Homology

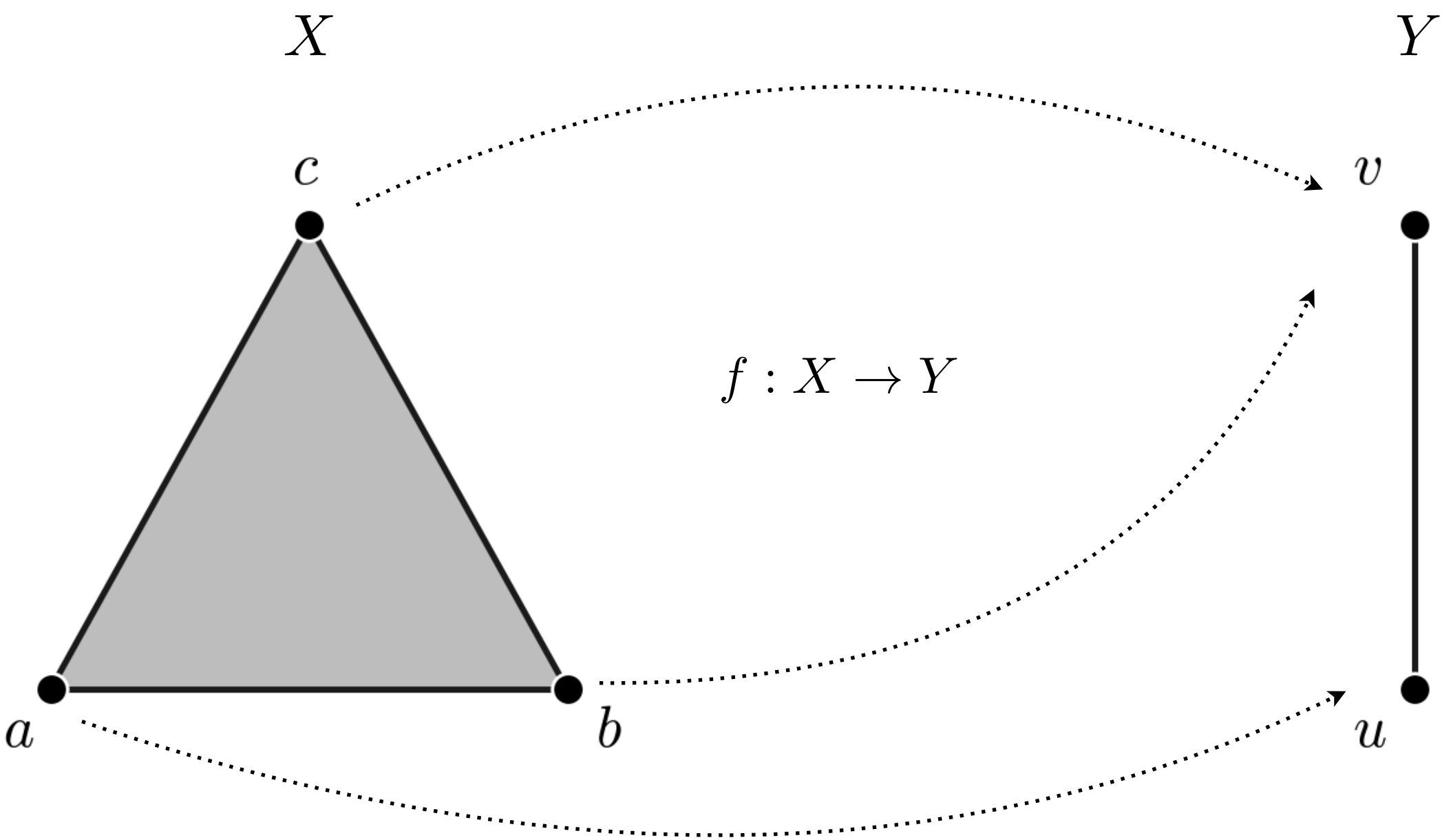
Simplicial maps

Maps between simplicial complexes

Def. Simplicial map

A simplicial map $f : X \rightarrow Y$ is a *map between the sets of vertices* of the simplicial complexes such that for every simplex $\sigma = \{x_0, \dots, x_k\} \in X$, its image $f(\sigma) = \{f(x_0), \dots, f(x_k)\} \in Y$ is also a simplex

Example:



$$f(a) = u, \quad f(b) = v, \quad f(c) = v$$

$$f(\{a, b\}) = \{u, v\}, \quad f(\{b, c\}) = \{v\}, \quad f(\{a, c\}) = \{u, v\}$$

$$f(\{a, b, c\}) = \{u, v\}$$

Induced maps in the chain space

Definition of induced maps at the chain level

Def. Induced map at the chain level

A simplicial map $f: X \rightarrow Y$ induces a linear map $f_{\#}: C_k(X) \rightarrow C_k(Y)$ for all $k \geq 0$ defined in the following way:

- We can write $\alpha \in C_k(X)$ in the basis of k -simplices as:

$$\alpha = \sum_{\sigma \in X_k} a_{\sigma} \cdot \sigma, \quad a_{\sigma} \in \mathbb{R}$$

- Applying the simplicial map at each simplex, we can define the following chain in Y :

$$f_{\#}(\alpha) = \sum_{\sigma \in X_k : f(\sigma) \in Y_k} a_{\sigma} \cdot f(\sigma) \in C_k(Y)$$

Notice that we require the image to be a k -simplex too!

Induced maps in the chain space

Induced maps commute with boundary operators

Important property

The induced map $f_{\#}$ commutes with the boundary maps $\partial_k^X : C_k(X) \rightarrow C_{k-1}(X)$ and $\partial_k^Y : C_k(Y) \rightarrow C_{k-1}(Y)$, that is:

$$f_{\#}(\partial_k^X(\alpha)) = \partial_k^Y(f_{\#}(\alpha))$$

Proof.

For a k -simplex $\sigma = \{x_0, \dots, x_k\} \in X$ we have two options:

■ $f(\sigma) = (y_0, \dots, y_k) \in Y_k$, that is, all vertices in σ are mapped to different vertices $f(x_i) = y_i \neq y_j = f(x_j)$, for $i \neq j$, and $f(\sigma)$ is a k -simplex of Y .

In that case:

$$f_{\#}(\partial_k^X(\sigma)) = f_{\#} \left(\sum_{i=0}^k (-1)^i (x_0, \dots, \widehat{x_i}, \dots, x_k) \right) = \sum_{i=0}^k (-1)^i f(x_0, \dots, \widehat{x_i}, \dots, x_k) = \sum_{i=0}^k (-1)^i (y_0, \dots, \widehat{y_i}, \dots, y_k) = \partial_k^Y(f_{\#}(\sigma))$$

Induced maps in the chain space

Induced maps commute with boundary operators

Important property

The induced map $f_{\#}$ commutes with the boundary maps $\partial_k^X : C_k(X) \rightarrow C_{k-1}(X)$ and $\partial_k^Y : C_k(Y) \rightarrow C_{k-1}(Y)$, that is:

$$f_{\#}(\partial_k^X(\alpha)) = \partial_k^Y(f_{\#}(\alpha))$$

- Two vertices in σ are mapped to the same vertex of Y , $f(x_i) = y_i = y_j = f(x_j)$ for $i < j$. In this case $f_{\#}(\sigma) = 0$ because $f(\sigma) \notin Y_k$, and $\partial_k^Y(f_{\#}(\sigma)) = 0$.

We now check $f_{\#}(\partial_k^X(\sigma)) = 0$ too:

$$f_{\#}(\partial_k^X(\sigma)) = \sum_{l=0}^k (-1)^l (y_0, \dots, \widehat{y_l}, \dots, y_k)$$

Where the sum on the right only includes $(y_0, \dots, \widehat{y_l}, \dots, y_k) \in Y_{k-1}$

- If $l \notin \{i, j\}$, then $(y_0, \dots, \widehat{y_l}, \dots, y_k)$ has at most $k - 2$ distinct vertices, and it is not in Y_{k-1} , so all these terms do not appear in the sum
- We only have left $(-1)^i (y_0, \dots, \widehat{y_i}, \dots, y_k) + (-1)^j (y_0, \dots, \widehat{y_j}, \dots, y_k) = (-1)^i (y_0, \dots, \widehat{y_i}, \dots, y_k) + (-1)^j (-1)^{j-i-1} (y_0, \dots, \widehat{y_i}, \dots, y_k) = 0$, where in the second equality we have moved y_i to the j -th position, changing the orientation accordingly.
- If more vertices in σ are mapped to the same vertex in Y , we apply a similar argument to conclude that both sides of the equality are 0 too.

Induced maps in the homology space

Induced maps

Def. Induced map at the homology level

A simplicial map $f : X \rightarrow Y$ induces a linear map $f_* : H_k(X) \rightarrow H_k(Y)$ for all $k \geq 0$ defined in the following way:

- An element $[\alpha] \in H_k(X)$ is a whole set $[\alpha] = \{\alpha + \partial_{k+1}\beta : \beta \in C_{k+1}(X)\}$ where, also, $\alpha \in \ker \partial_k^X \iff \partial_k^X \alpha = 0$
- We define $f_*([\alpha]) = [f_{\#}(\alpha)] = \{f_{\#}(\alpha) + \partial_k^Y \beta : \beta \in C_{k+1}(Y)\}$

We need to check that this map is well-defined:

- First, we check that $f_{\#}(\alpha)$ is a cycle:

$$\partial_k^Y(f_{\#}(\alpha)) = f_{\#}(\partial_k^X(\alpha)) = f_{\#}(0) = 0$$

- Now we check that two elements of the same class $[\alpha_1] = [\alpha_2] \in H_k(X)$ go to the same class in $H_k(Y)$. Two elements in the same class are related as: $\alpha_1 = \alpha_2 + \partial_k^X \beta$, $\beta \in C_{k+1}(X)$.

Then: $f_{\#}(\alpha_1) = f_{\#}(\alpha_2) + f_{\#}(\partial_k^X \beta) = f_{\#}(\alpha_2) + \partial_k^Y(f_{\#}(\beta))$ and thus

$$f_*([\alpha_1]) = [f_{\#}(\alpha_1)] = [f_{\#}(\alpha_2)] = f_*([\alpha_2])$$

Induced maps in the homology space

Two important properties

Properties of the induced maps

The induced homology map $f_* : H_k(X) \rightarrow H_k(Y)$ for all $k \geq 0$ satisfies the following two properties:

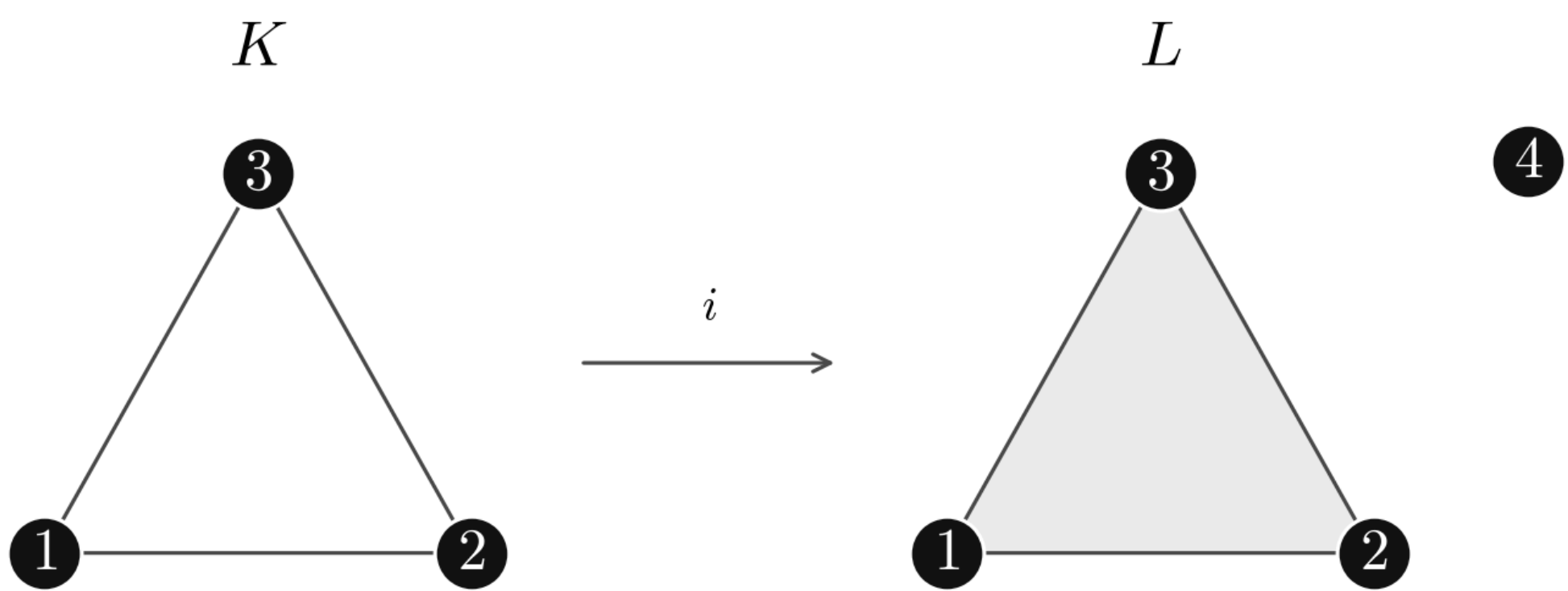
- If $\text{id} : X \rightarrow X$ is the identity map, sending each vertex in X to itself, then $\text{id}_* : H_k(X) \rightarrow H_k(X)$ is the identity map of vector spaces, which leaves every vector unchanged
- If $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ then

$$(g \circ f)_* = g_* \circ f_*$$

An example for you to work on

Induced maps

Describe the induced maps for the 0, 1 and 2-dimensional homology groups of the following inclusion of simplicial complexes

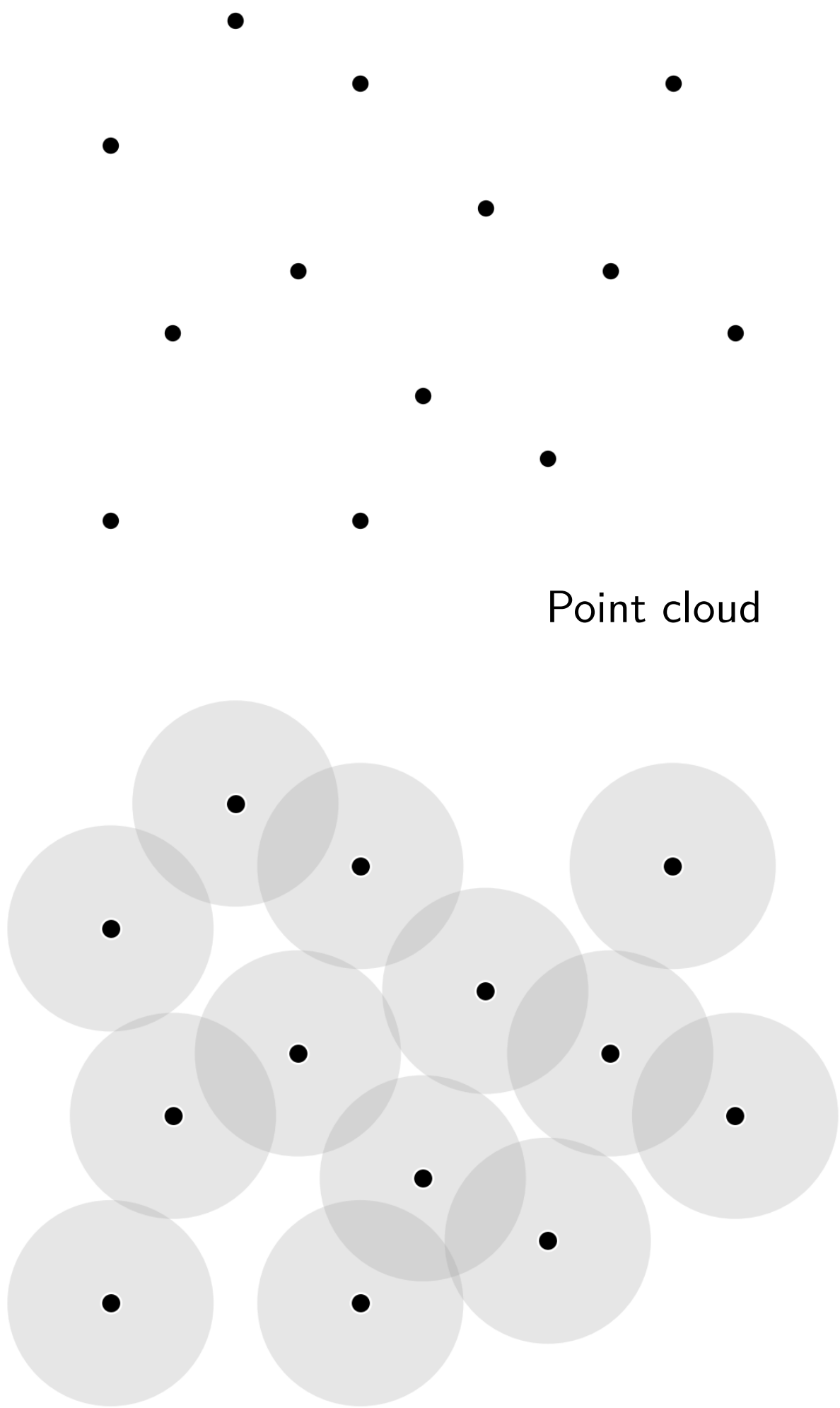


How to build simplicial complexes from point clouds

Session 6 — Persistent Homology

Point clouds

Our starting point



Point cloud

Point cloud + closed balls

A point cloud is a finite metric space (S, d) where:

- $S = \{s_1, \dots, s_n\}$ is a **discrete set of points** (the vertices of the abstract simplicial complex we will build)
- $d : S \times S \rightarrow \mathbb{R}$ is a **distance function on the set**, which we can represent as the matrix of pairwise distances

$$\begin{pmatrix} d(s_1, s_1) & d(s_1, s_2) & \dots & d(s_1, s_n) \\ d(s_2, s_1) & d(s_2, s_2) & \dots & d(s_2, s_n) \\ \vdots & \vdots & \ddots & \vdots \\ d(s_n, s_1) & d(s_n, s_2) & \dots & d(s_n, s_n) \end{pmatrix}$$

If we consider these points as living in a Euclidean space: $S \subset \mathbb{R}^D$, we can consider the **closed balls**:

$$B(x, r) := \{y \in \mathbb{R}^D : d(x, y) \leq r\}$$

Ball centered at x with radius r

Simplicial complexes

Building simplicial complexes from point clouds

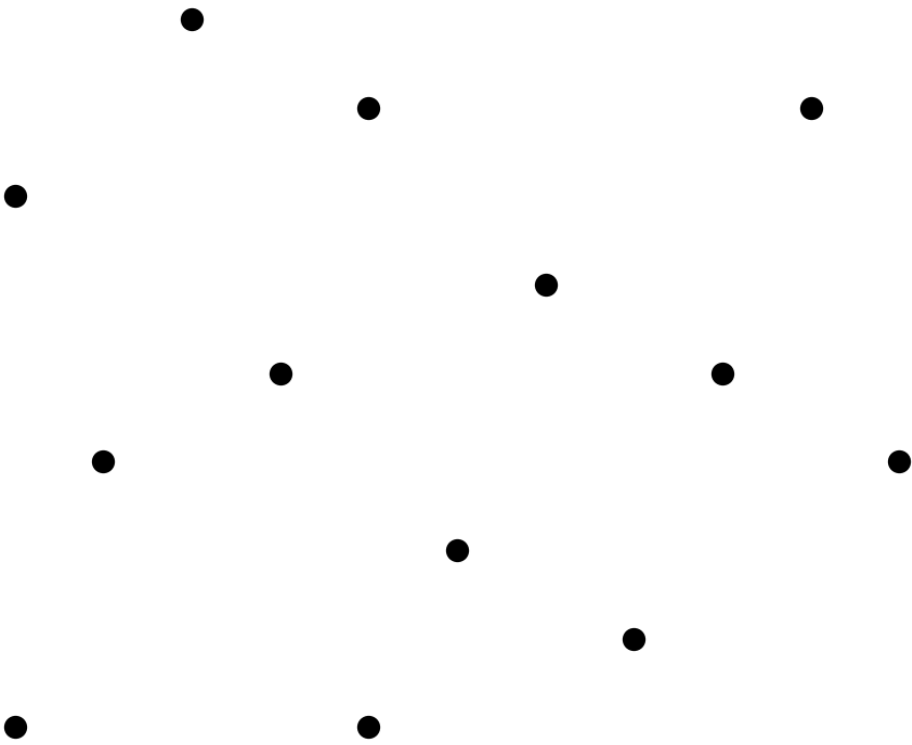
Finite metric space

Given (S, d) and $r \geq 0$, the **Čech complex** (or Nerve complex) at scale r is:

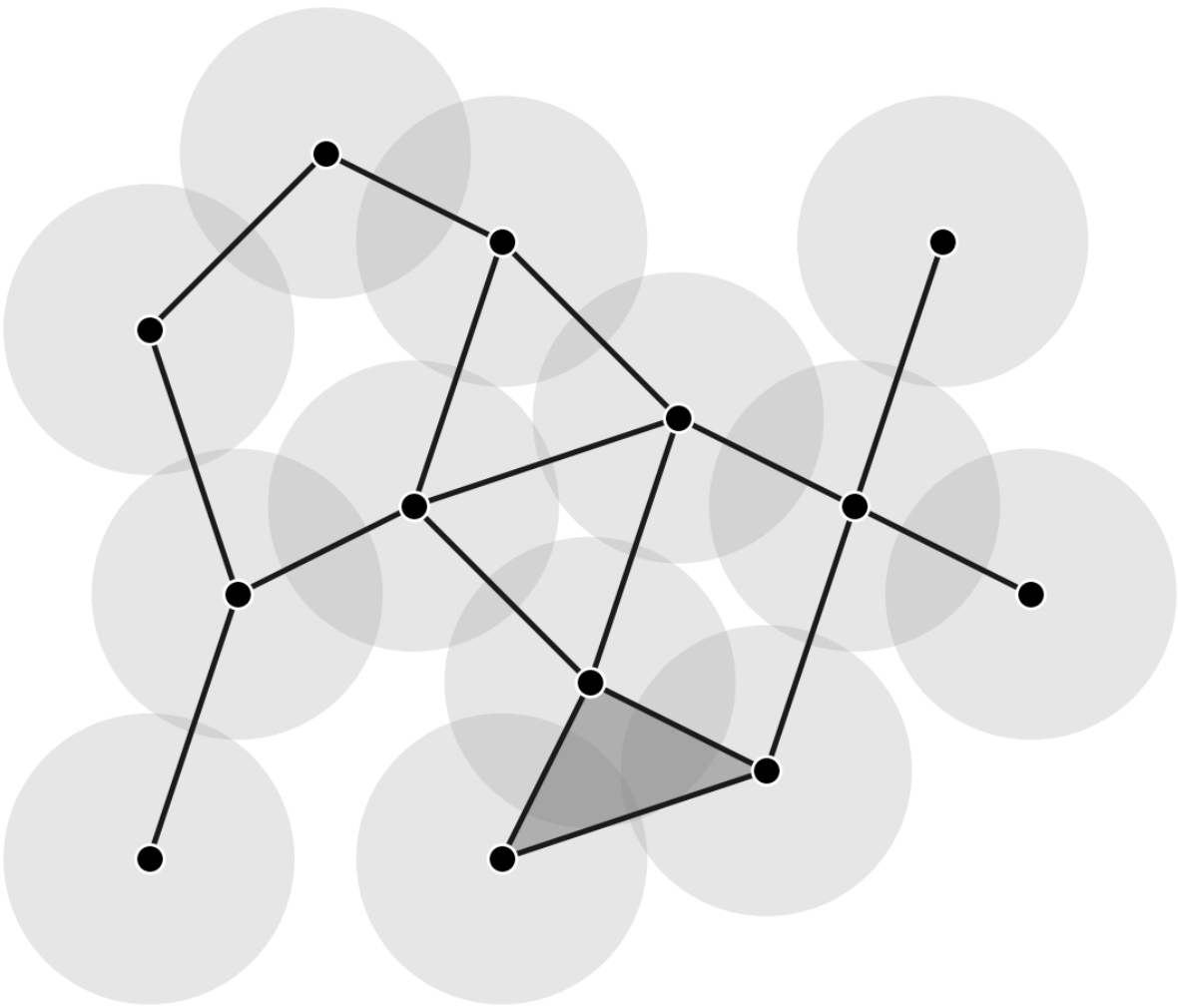
$$C_r(S, d) = \left\{ \{s_{i_0}, \dots, s_{i_k}\} : \bigcap_{j=0}^k B(s_{i_j}, r) \neq \emptyset \right\}$$

Monotonicity

For $r \leq s$: $C_r(S) \subseteq C_s(S)$



Point cloud



Čech complex

Simplicial complexes

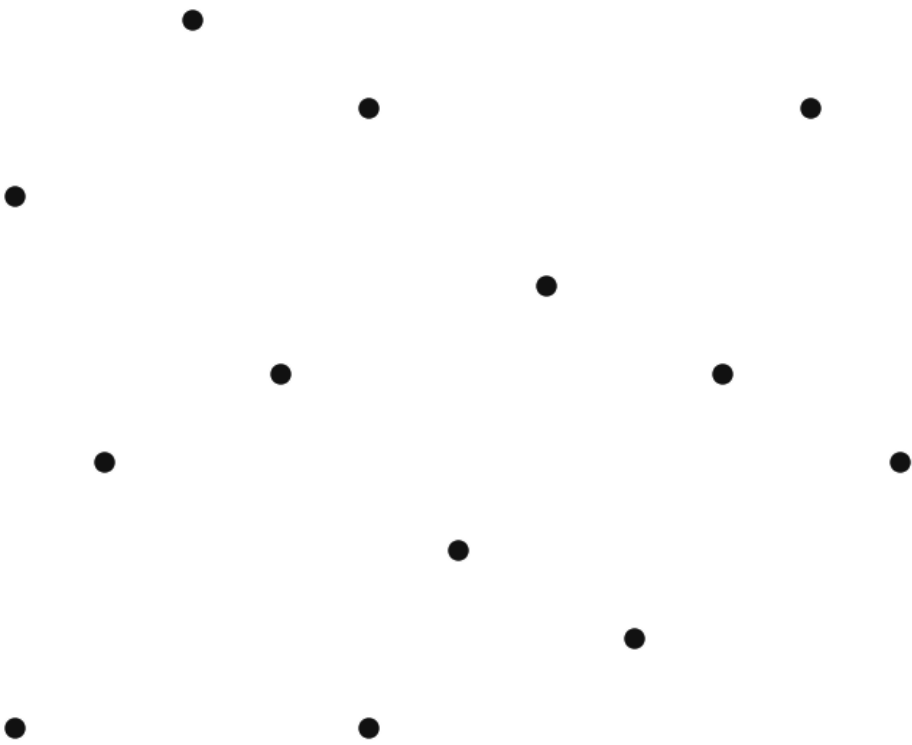
Building simplicial complexes from point clouds

Given (S, d) and $r \geq 0$, the **Vietoris-Rips complex** at scale r is:

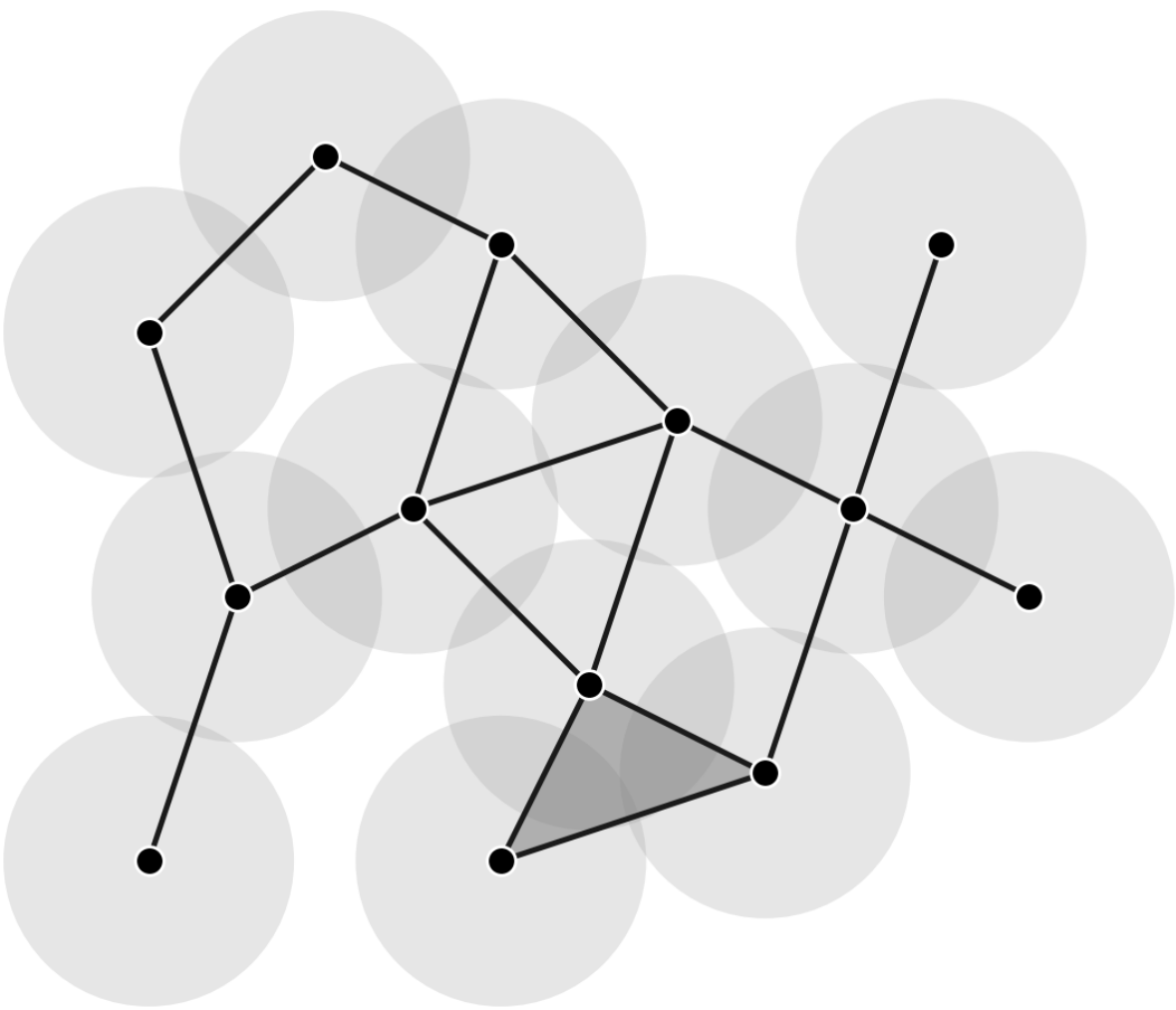
$$VR_r(S, d) = \left\{ \{s_{i_0}, \dots, s_{i_k}\} : d(s_i, s_j) \leq 2r, \forall i_0 \leq i < j \leq i_k \right\}$$

Monotonicity

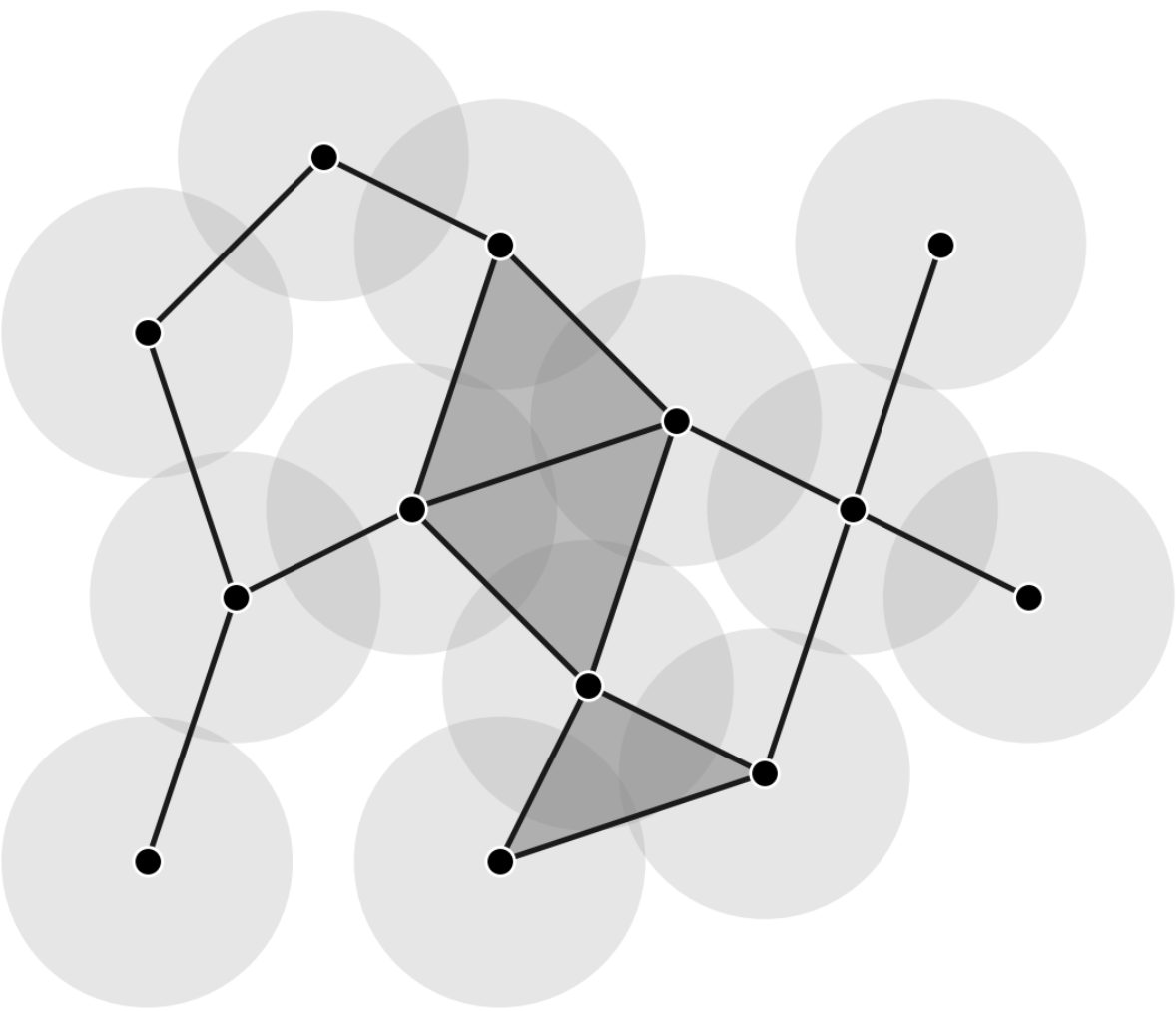
For $r \leq s$: $VR_r(S) \subseteq VR_s(S)$



Point cloud



Čech complex



VR complex

Simplicial complexes

Relating Čech and VR complexes

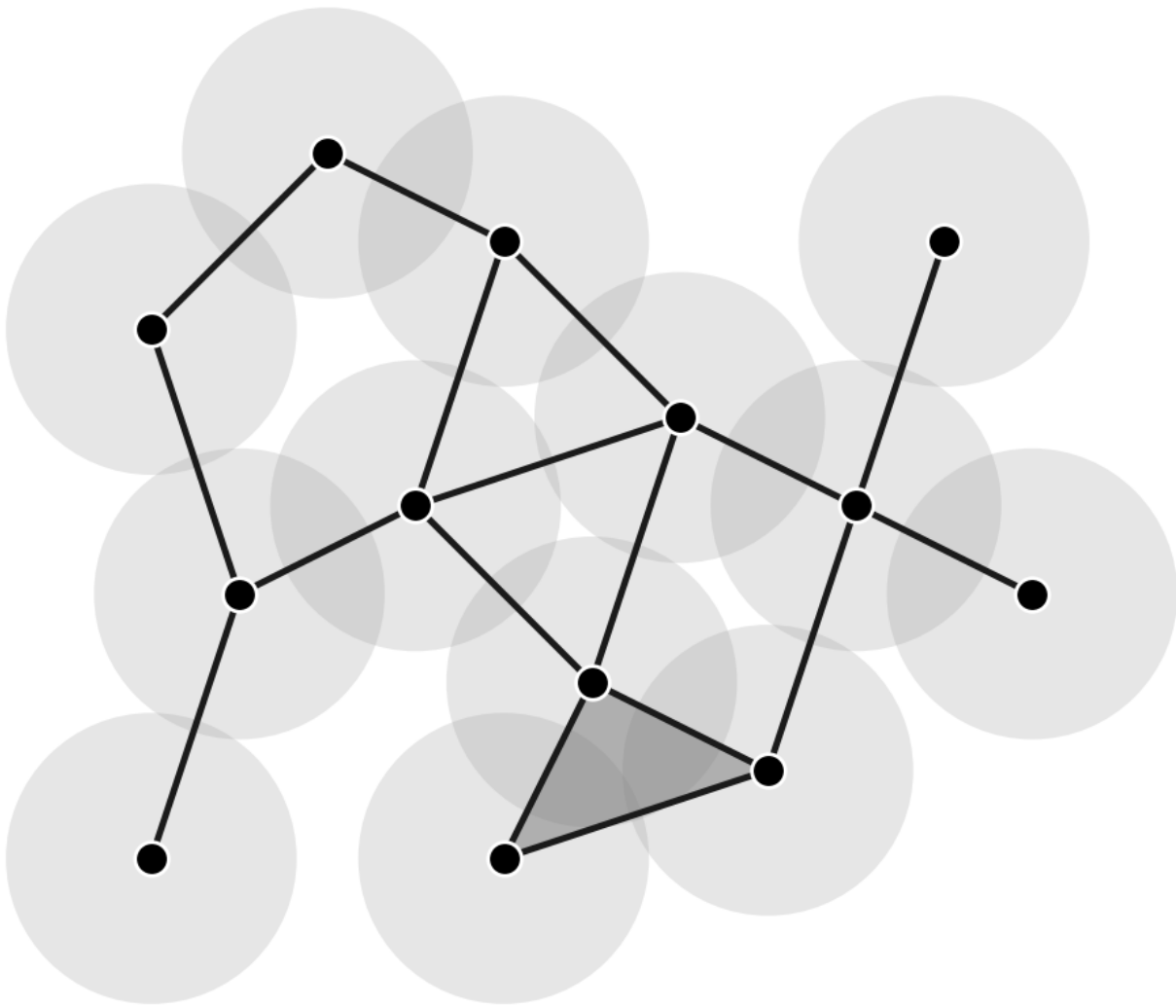
This containment is easy to prove

This one requires a bit more work

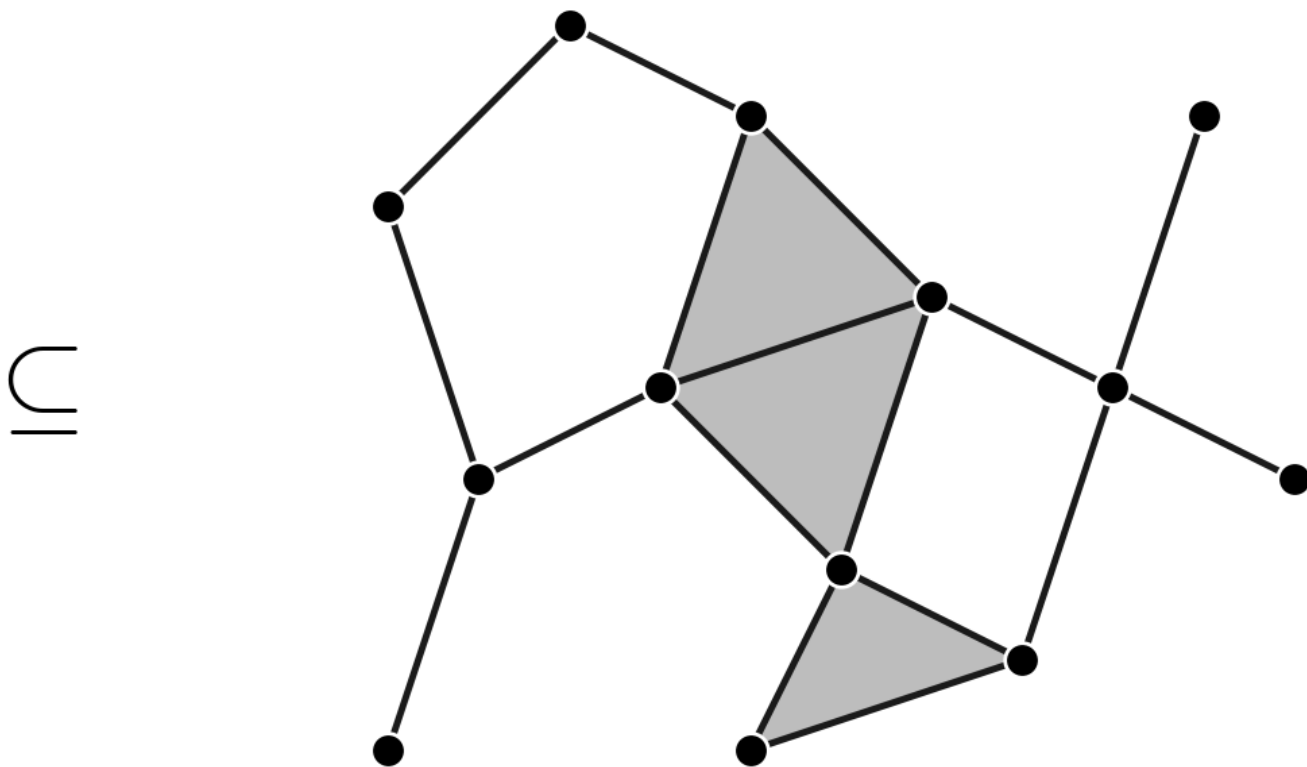
Lemma:

For $r \geq 0$: $C_r(S) \subseteq VR_r(S) \subseteq C_{\sqrt{2}r}(S)$

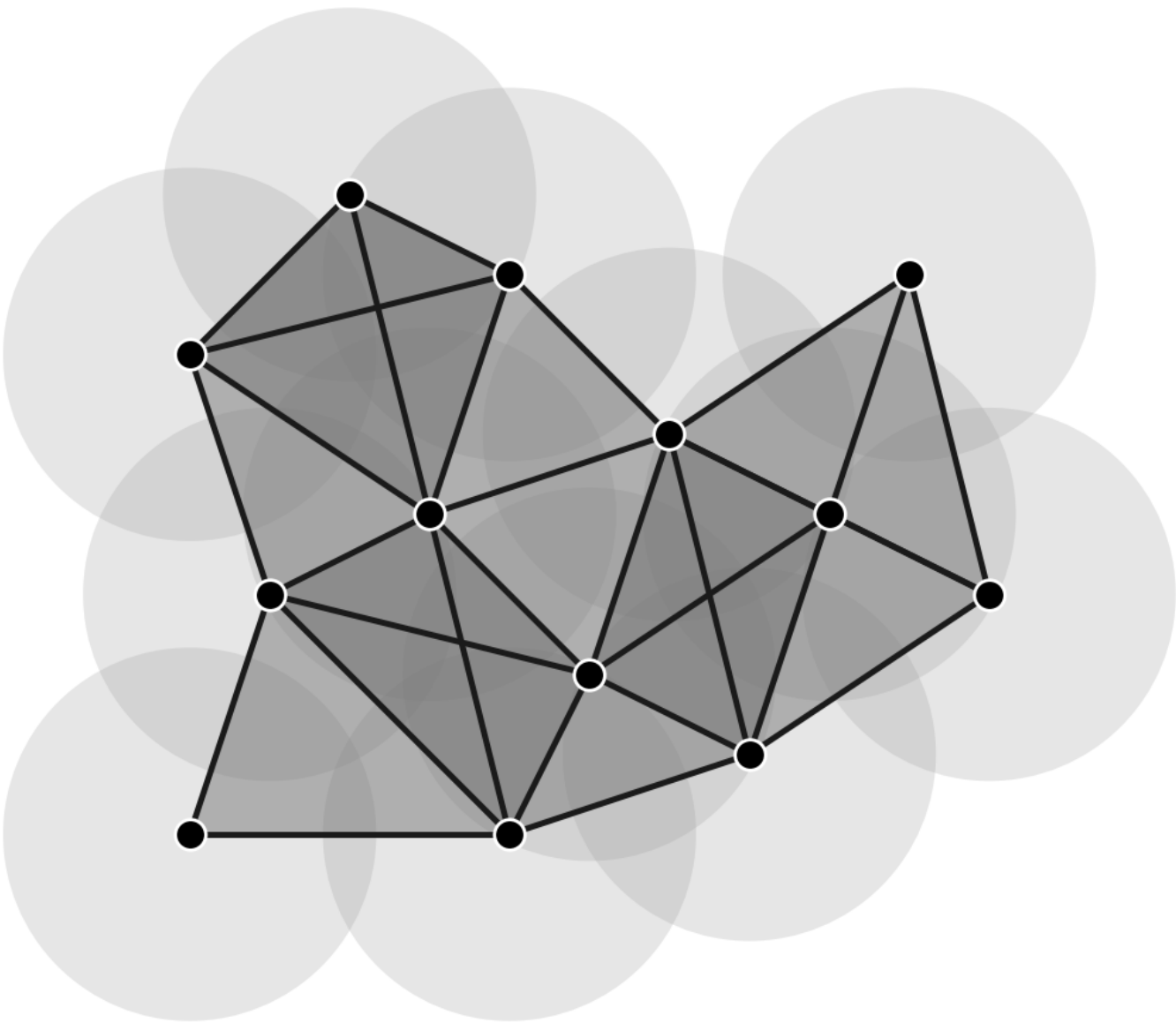
$C_r(S)$



$VR_r(S)$



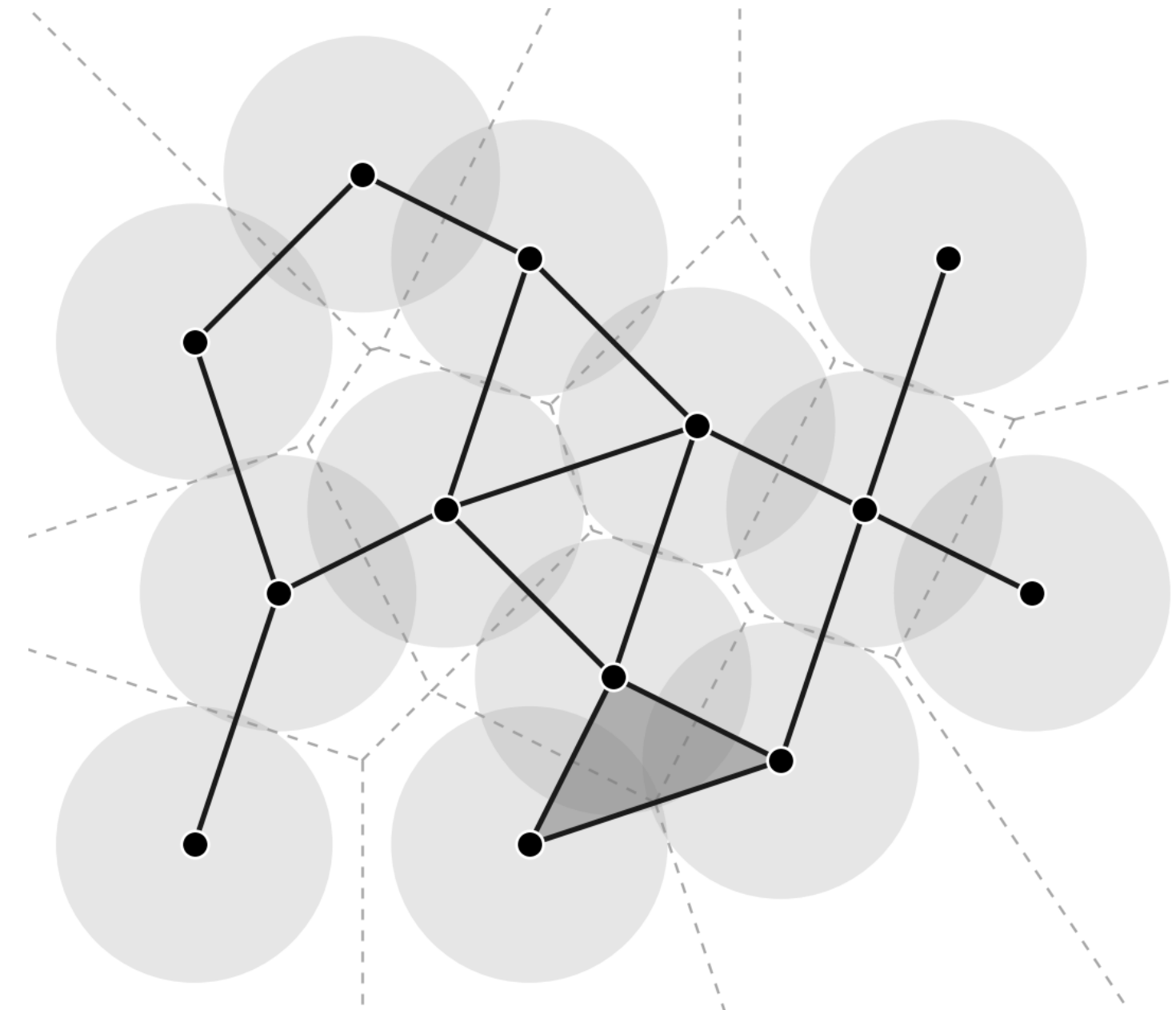
$C_{\sqrt{2}r}(S)$



Simplicial complexes

Other simplicial complexes

- Alpha shapes and Alpha complexes
- Dowker complexes
- k -NN complexes



H. Edelsbrunner, D. Kirkpatrick and R. Seidel, “On the shape of a set of points in the plane”, *IEEE Transactions on Information Theory* 29.4, 1983, pp. 551–559

C. H. Dowker, “Homology Groups of Relations”, *Annals of Mathematics* 56.1, 1952, pp. 84–95

M. Q. Le and D. Taylor, “Persistent Homology with k-nearest-neighbor Filtrations reveals Topological Convergence of PageRank”, 2022, arXiv: 2206.04725

The nerve theorem

Session 6 — Persistent Homology

Simplicial homology

Nerve theorem

Def + thm. Nerve theorem (informal)

Let X be a topological space and $\mathcal{U} = \{B(x, r) : x \in S \subset X, r > 0\}$ a cover of this space, that is

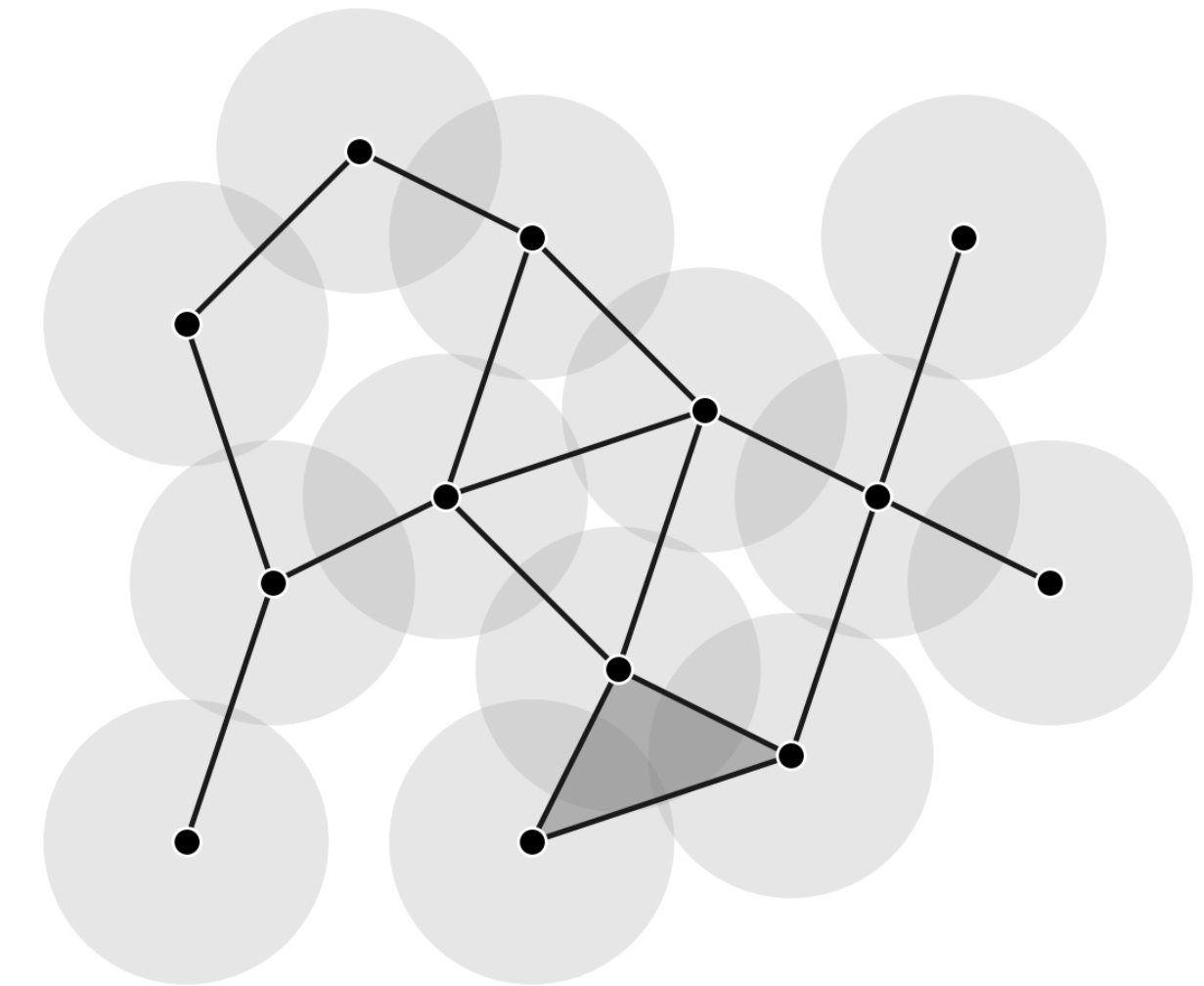
$$X \subset \bigcup_{x \in S} B(x, r)$$

The nerve of this cover is the following simplicial complex:

$$\text{Nerve}(\mathcal{U}) = \left\{ \sigma \subset S : \bigcap_{x \in \sigma} B(x, r) \neq \emptyset \right\}$$

With some *nice ness assumptions* on X and $\mathcal{U}$ we have:

$$\text{Nerve}(\mathcal{U}) \cong X \quad \textbf{(homotopy equivalent, same topological properties)}$$



The Čech complex is a nerve

$$C_r(S, d) = \left\{ \{s_{i_0}, \dots, s_{i_k}\} : \bigcap_{j=0}^k B(s_{i_j}, r) \neq \emptyset \right\}$$

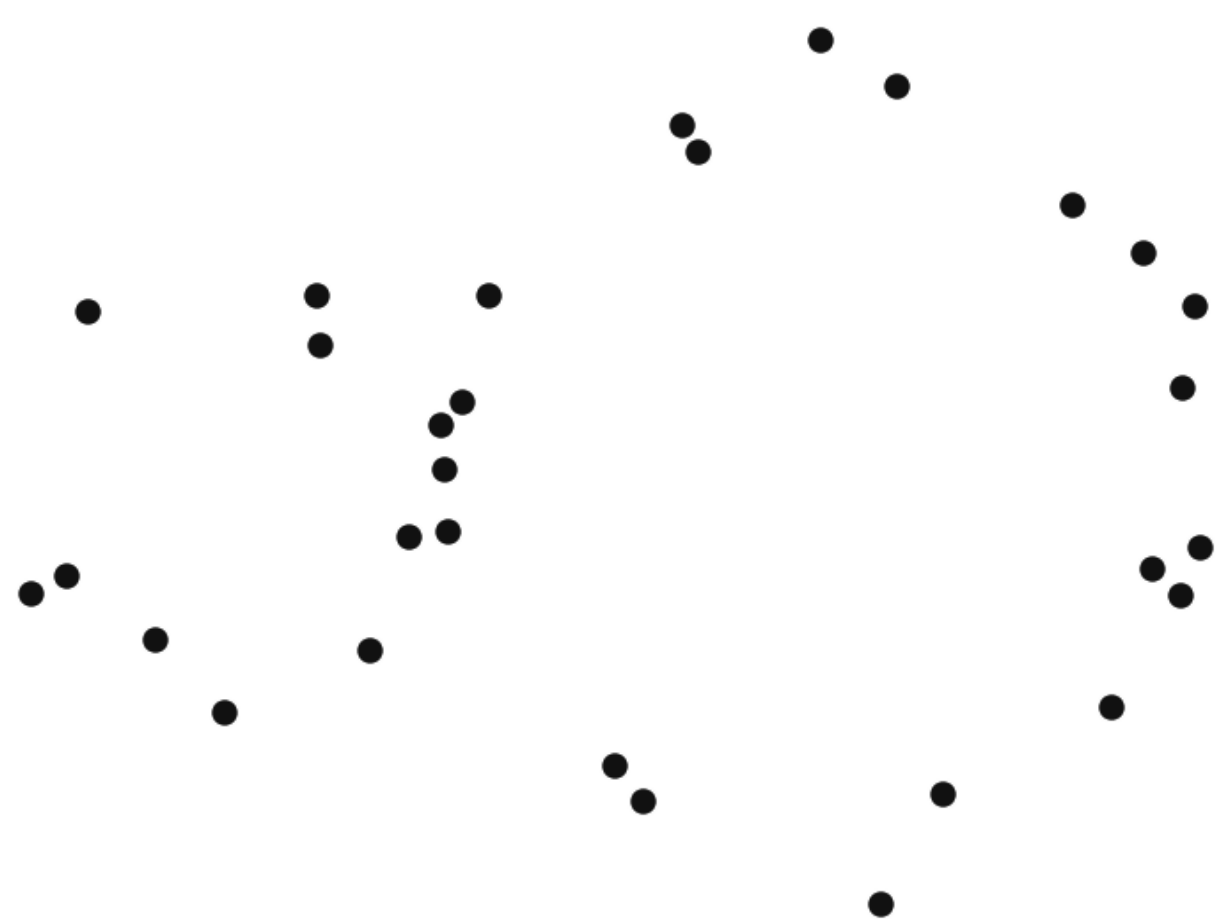
¹Ghrist, R., and Muhammad, A. “Coverage and Hole-Detection in Sensor Networks via Homology.” *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks*, 2005., April 2005, 254–60.

¹Bauer, U., Kerber, M., Roll, F., and Rolle, A. “A Unified View on the Functorial Nerve Theorem and Its Variations.” *Expositiones Mathematicae* 41, no. 4 (2023): 125503.

Simplicial homology

Nerve theorem — the practical consequences

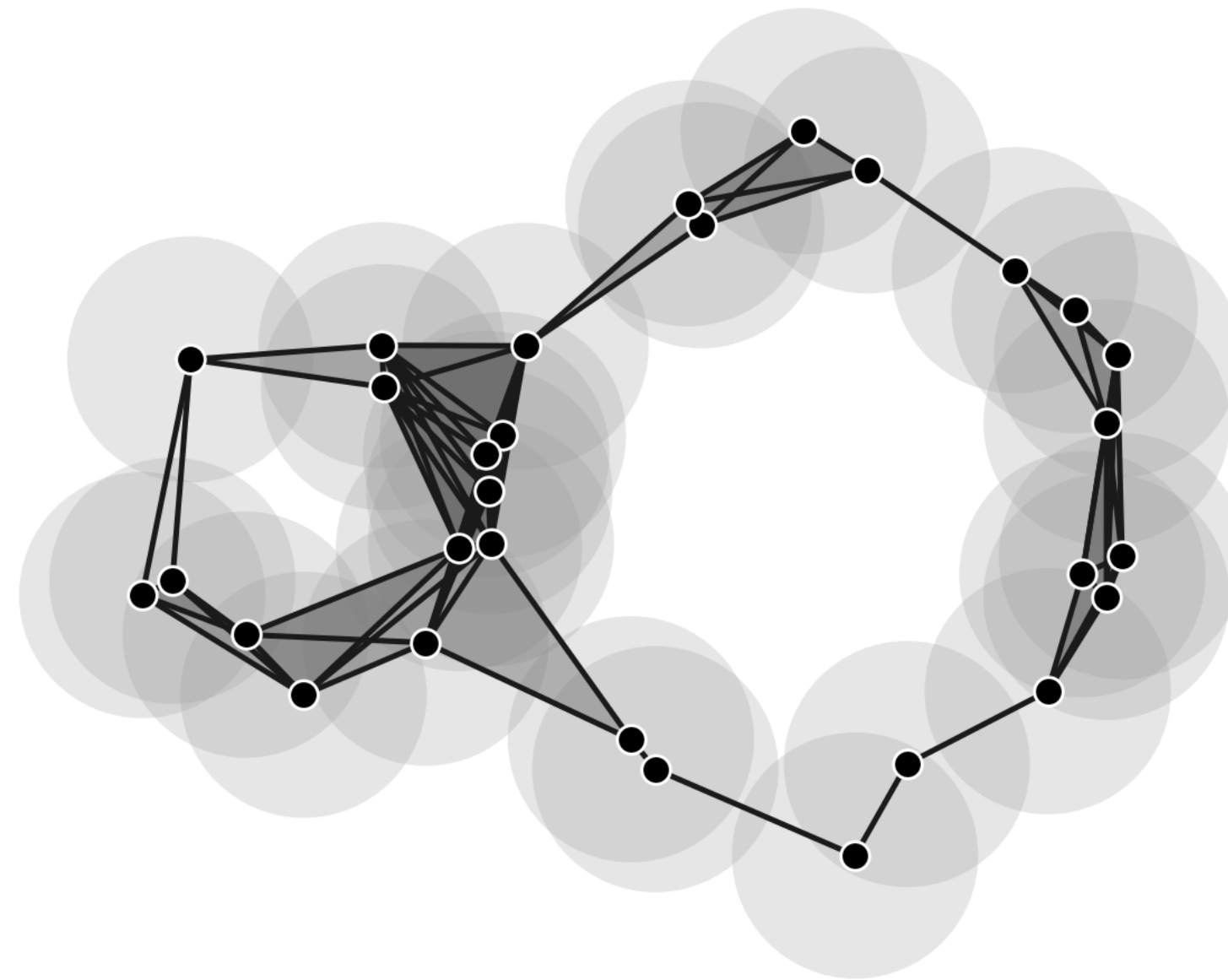
Suppose you are given the following point cloud, how can you design a test to conclude that this sample comes from a distribution over two circles?



Simplicial homology

Nerve theorem — the practical consequences

1. Approximate S with a Čech complex for a some radius that captures the holes: $C_r(S)$
2. Compute homology in degree 1: $H_1(C_r(S))$
3. If $\beta_1(C_r(S)) = 2$, then you are observing two holes



How do you find the radius r ?

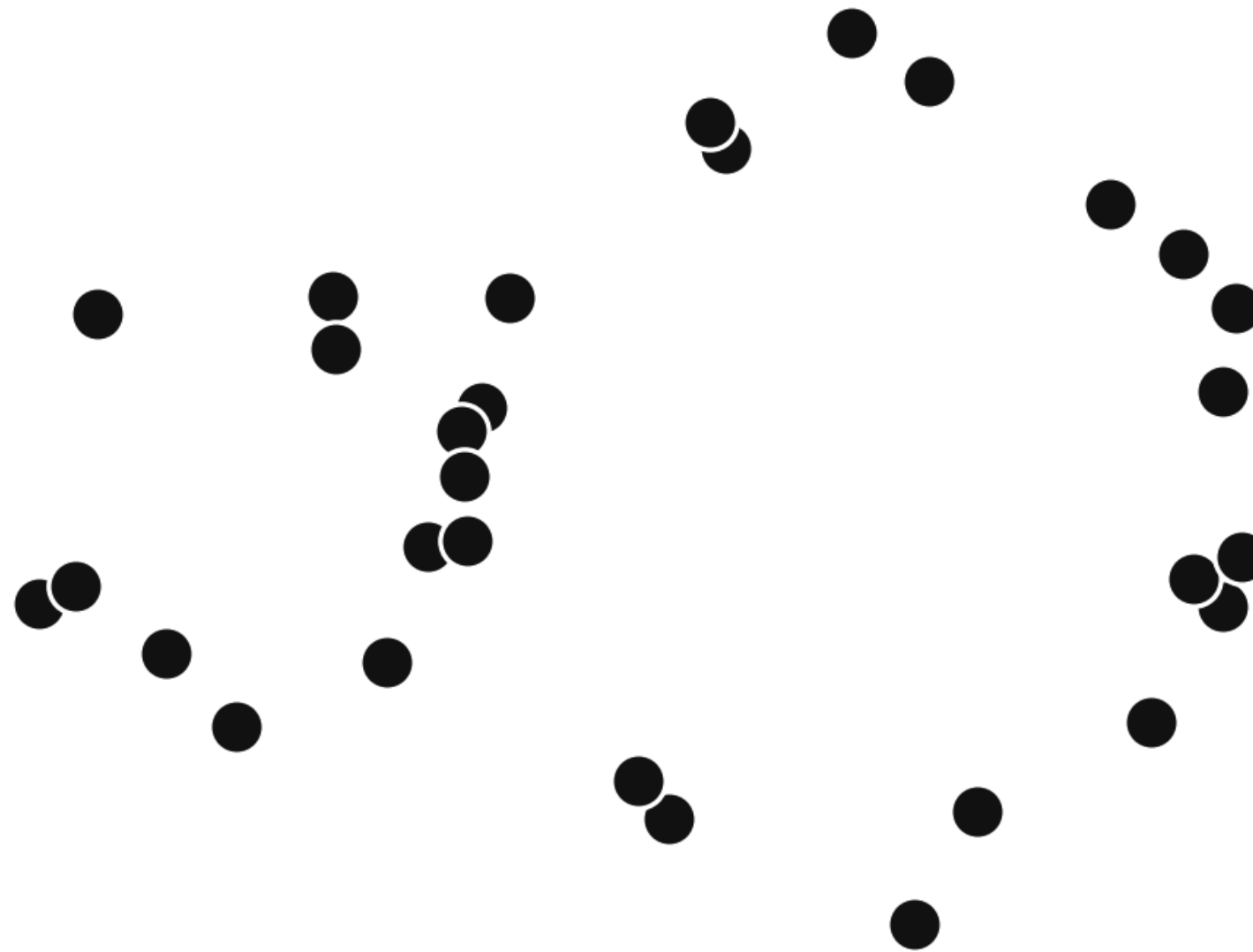
Persistent homology

Lecture 6 — Persistent homology

Persistent homology

Filtrations

Consider all possible radii!



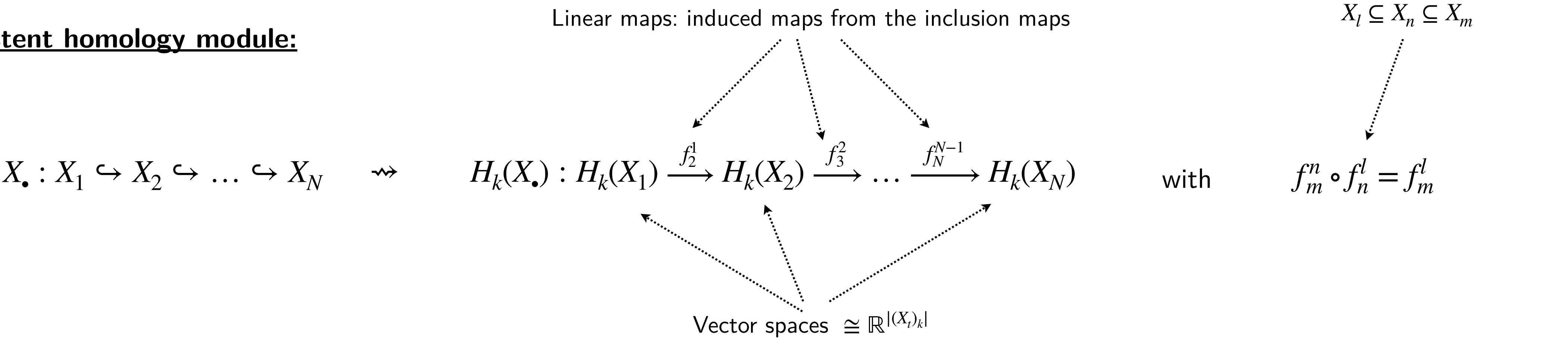
Def. Filtration

$\{X_t : t \in \mathbb{R}\}$ such that for $s \leq t$, $X_s \subseteq X_t$

Persistent homology

Persistence modules

Persistent homology module:



Def. Persistence module (in full generality)

- A family $\{V_n\}_{1 \leq n \leq N}$ of vector spaces and a family $\{\phi_m^n : V_n \rightarrow V_m\}_{n \leq m}$ of linear maps satisfying:
- $\phi_n^n = \text{id}_{V_n}$
 - $\phi_n^m \circ \phi_m^l = \phi_n^l$ for all $l \leq m \leq n$

Persistent homology

Interval modules

An example of persistence modules are interval modules:

$$I[a, b]_n = \begin{cases} 0 & \text{if } n < a, \text{ or } n > b \\ \mathbb{R} & \text{if } a \leq n \leq b \end{cases}$$

$$I[a, b]_m^n = \begin{cases} 0 & \text{if } n < a, \text{ or } m > b \\ \text{id}_{\mathbb{R}} & \text{if } a \leq n \leq m \leq b \end{cases}$$

Module:

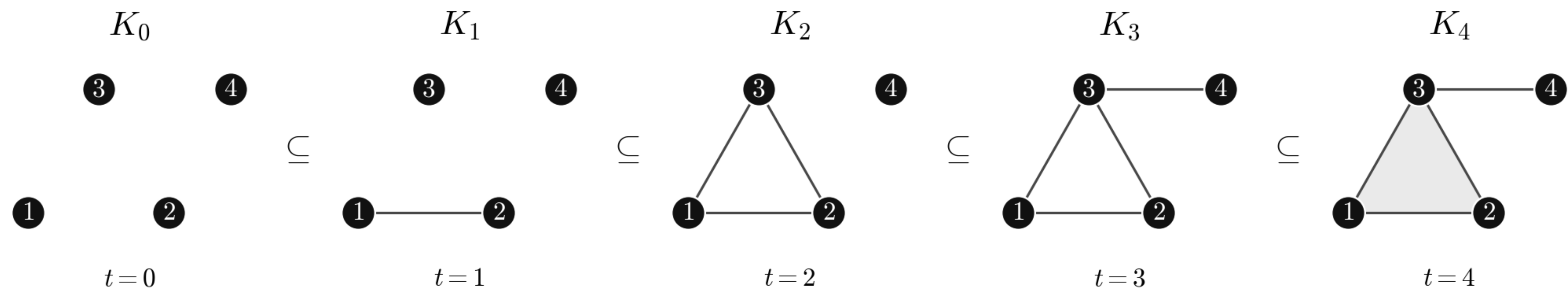
$$0 \rightarrow 0 \rightarrow \cdots \rightarrow 0 \longrightarrow \mathbb{R} \xrightarrow{\text{id}_{\mathbb{R}}} \mathbb{R} \longrightarrow \cdots \longrightarrow \mathbb{R} \xrightarrow{\text{id}_{\mathbb{R}}} \mathbb{R} \longrightarrow 0 \rightarrow \cdots \rightarrow 0$$

Parameters:

$$1 \qquad 2 \qquad \qquad a-1 \qquad a \qquad a+1 \qquad \qquad b-1 \qquad b \qquad b+1 \qquad N$$

Persistent homology

An example for you to try



For 0 and 1-dimensional homology:

- Choose generators for the homology groups at each step in the filtration
- Connect the generators in subsequent steps in the filtration whenever the map induced in homology by the inclusion connects them
- No line needs to be drawn when the generator is directly mapped to 0 by the induced map

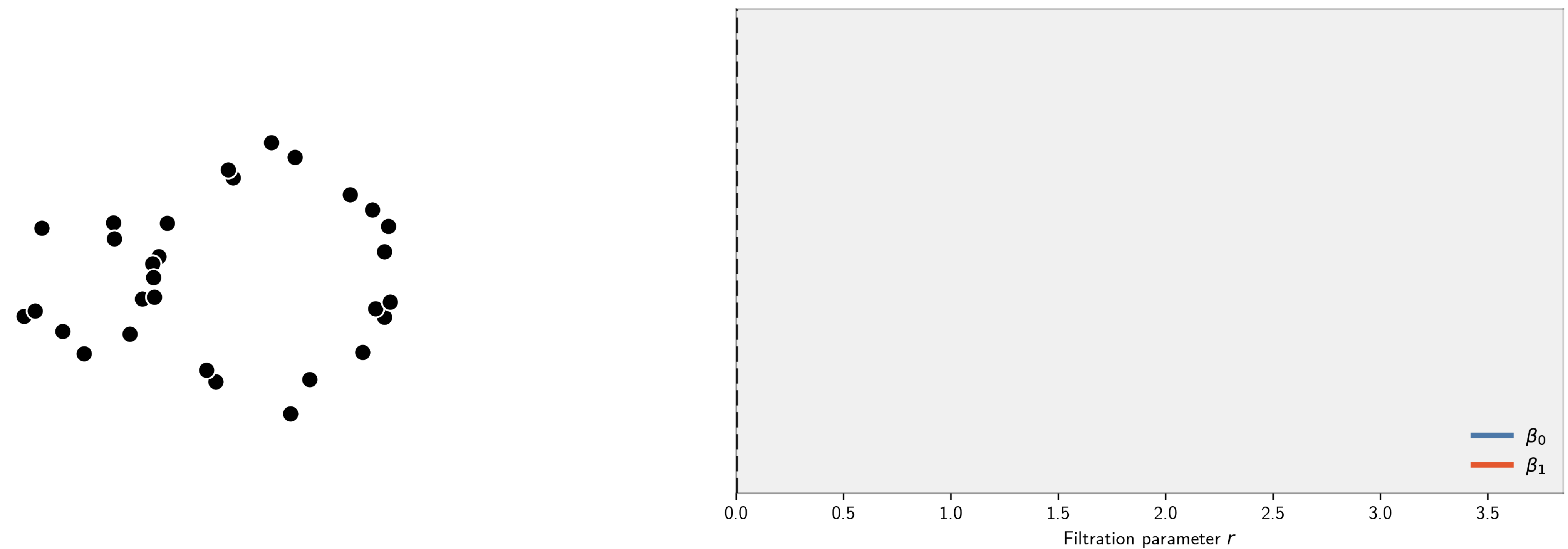
Persistent homology

Structure theorem and barcodes

This is true in general by the **structure theorem**, whose statement and proof go beyond the scope of the course

It is enough that you know that for each filtration (and corresponding persistent homology module) we can compute a **barcode** that tracks how homology classes are created and destroyed as the filtration grows

This is the same as tracking the evolution of the Betti numbers of the complexes in the filtration



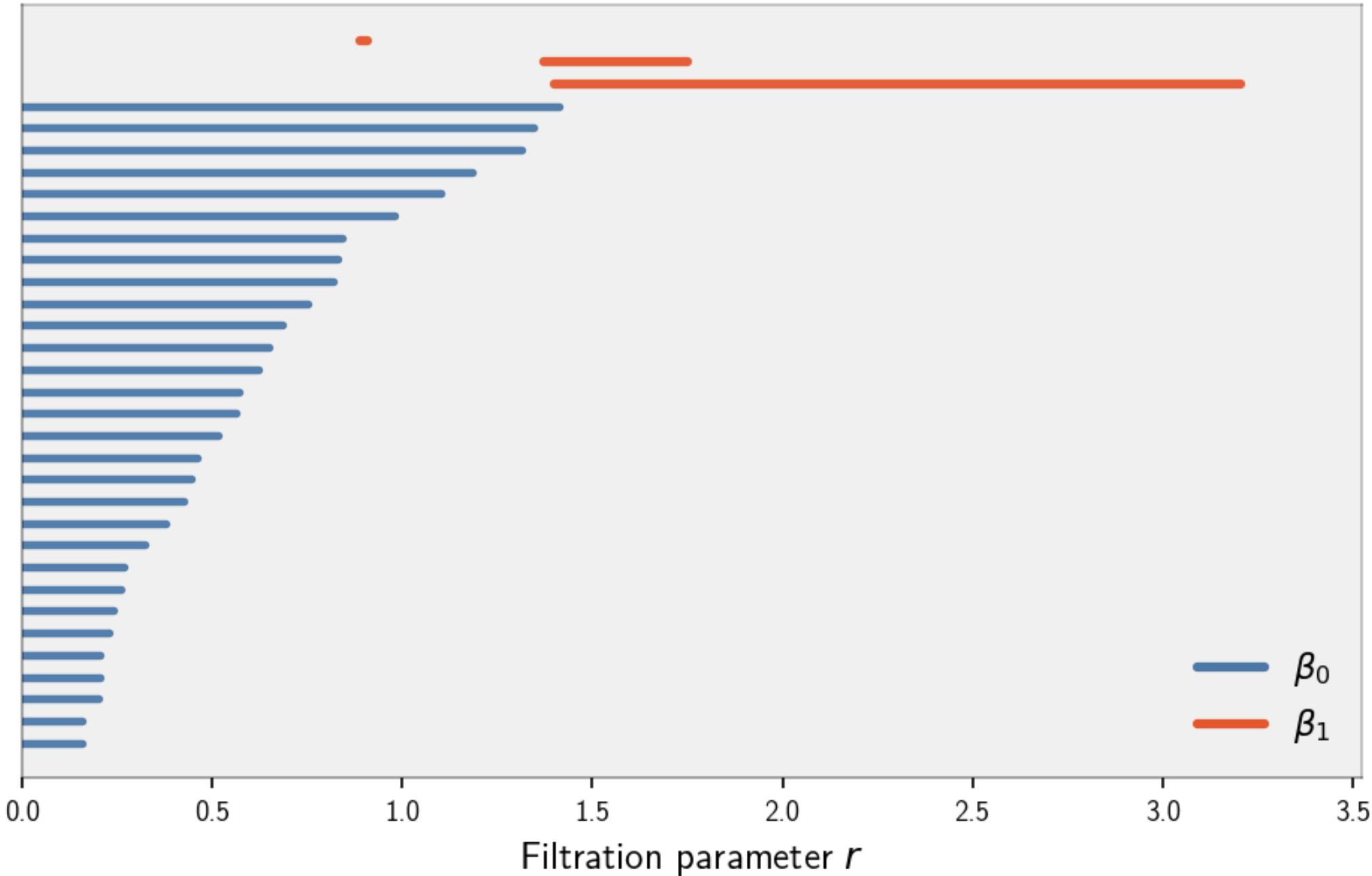
Azumaya, Gorô. “Corrections and Supplementaries to My Paper Concerning Krull-Remak-Schmidt’s Theorem.” *Nagoya Mathematical Journal* 1 (June 1950): 117–24.

Crawley-Boevey, William. “Decomposition of Pointwise Finite-Dimensional Persistence Modules.” *Journal of Algebra and Its Applications* 14, no. 05 (2015): 1550066.

Persistent homology

Persistence barcode

As many bars as points,
bars *die* as points get
connected



Two long bars: we have two
persistent loops in the point
cloud

The point cloud had the shape of
two circles!

Persistence barcode

To sum up

Today we have learnt how to...

- Obtain maps between homology groups from simplicial maps between simplicial complexes
- Build simplicial complexes from point clouds, approximating their shape at different scales
- Combine these ideas to define the persistent homology module
- Summarize the topological information of the persistent module in a barcode

Thank you for your attention!



AIDOS LAB
AI FOR DATA-ORIENTED SCIENCE

