Section 3

(1)

Other geometrical complexes & how efficient they are.

In the last section we have extensivelly studied cubical complexes. But, let us take a step back & remember the reason why we study them:

$$\text{Data} \rightsquigarrow \text{Combinatorial complex} \rightsquigarrow (\text{persistent homology}).$$

What are other ways of representing data, most notably point clouds in $\mathbb{R}^n$ or discrete metric spaces, as combinatorial complexes?

In this section we will focus on simplicial complexes

First, let us discuss shortly how to store simplicial complexes on a computer?

Firstly, we cannot store a geometric structure, as it is infinite. But, we can store combinatorial structure + coordinates of vertices.

Combinatorial structure on a set of vertices V is nothing more but an abstract simpliicial complex.

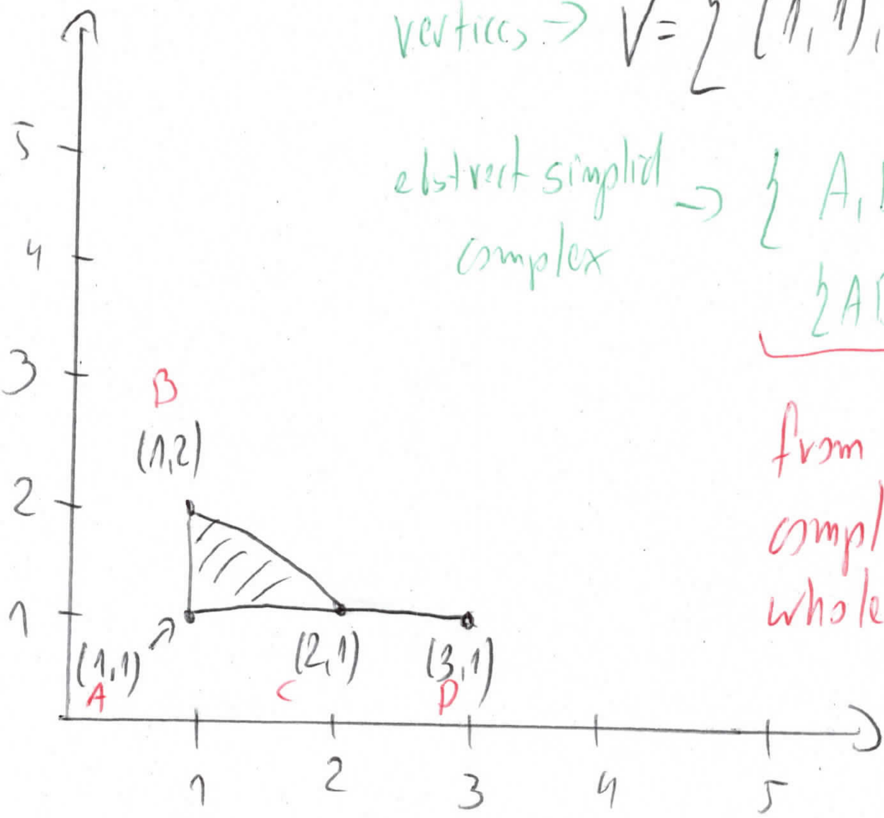
K is an abstract simplicial complex if ~~for~~

1. It is finite
2. For every $A \in K$ and for every $B \subseteq A, B \in K$

(in other words, K is a set of sets closed under operation of taking subset)

For simplicial complexes, coordinates of vertices V + the abstract simplicial complex, suffice to represent the input complex.

I.E.



vertices → $V = \{ \overset{A}{(1,1)}, \overset{B}{(1,2)}, \overset{C}{(2,1)}, \overset{D}{(3,1)} \}$

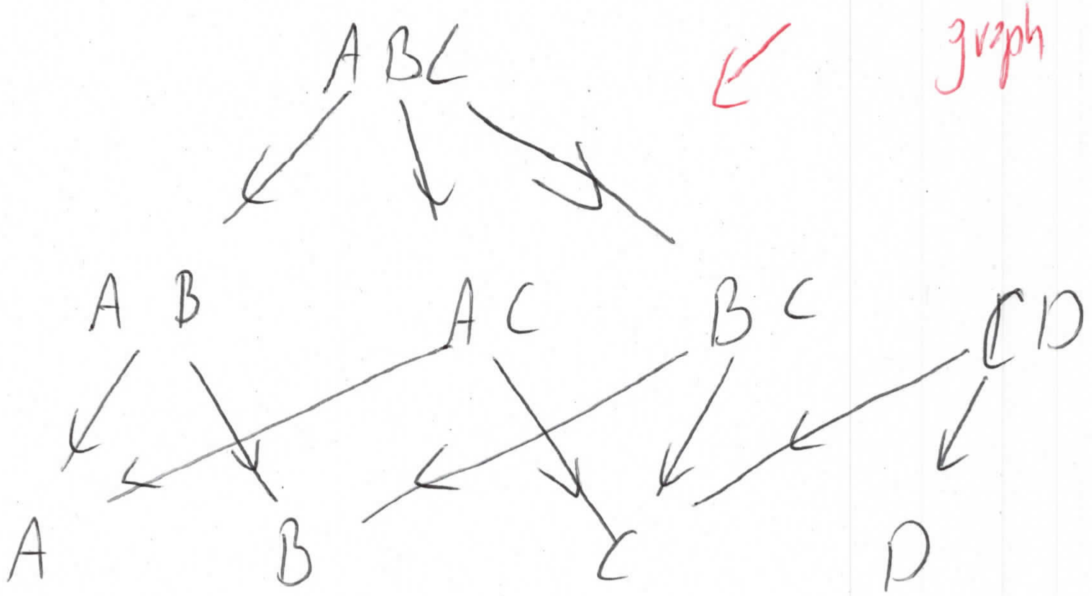
abstract simplicial complex → $\underbrace{\{ A, B, C, D, \{AB\}, \{AC\}, \{BC\}, \{ABC\}, \{C,D\} \}}$

from the abstract simplicial complex we can recover whole connectivity of the input simplicial complex.

Representations of Abstract Simplicial Complex:

1 Hasse Diagram:

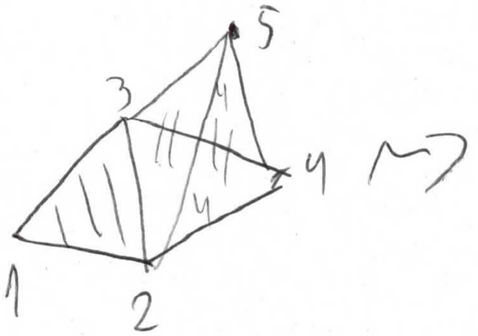
Directed Acyclic graph

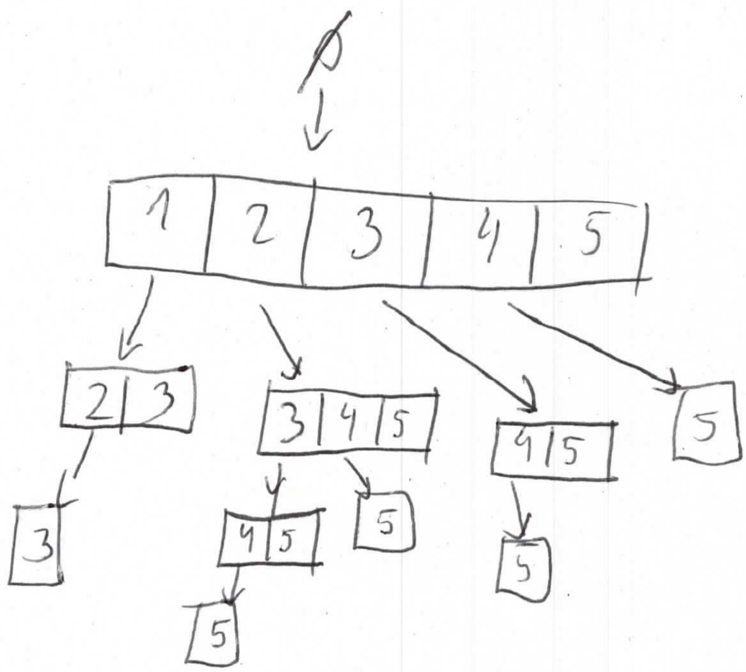


Clearly exponential complexity with respect to dimension

Explicit boundary relation encoded by edges of the graph.

2 Simplex Tree [Boissonat, Mariа]



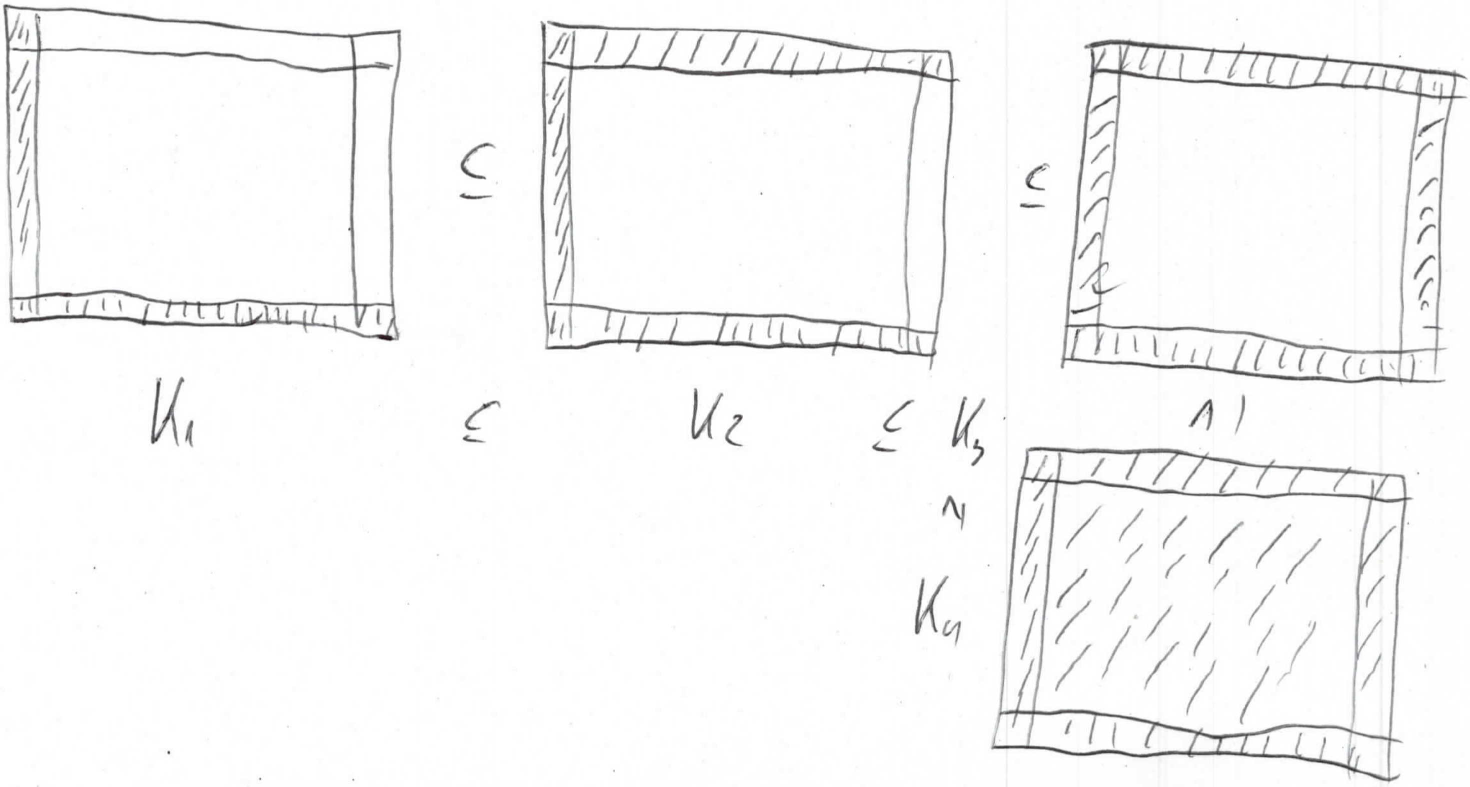


Efficient data structure, working horse in Gudhi

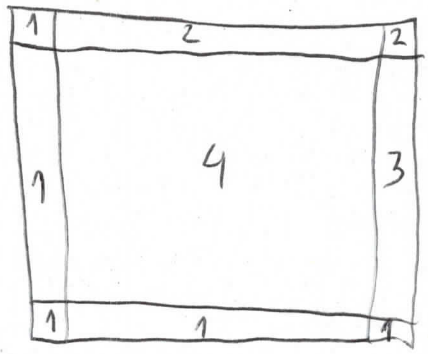
We now know how to store combinatorial structure of simplicial complex.
Let us remind a concept of filtration. Given a simplicial or a cubical complex (in this case, in this generality, we will be speaking about a cell complex) K a filtration is a nested sequence of subcomplexes

$$\emptyset \subseteq K_1 \subseteq K_2 \subseteq \ldots \subseteq K_{n-1} \subseteq K_n = K$$

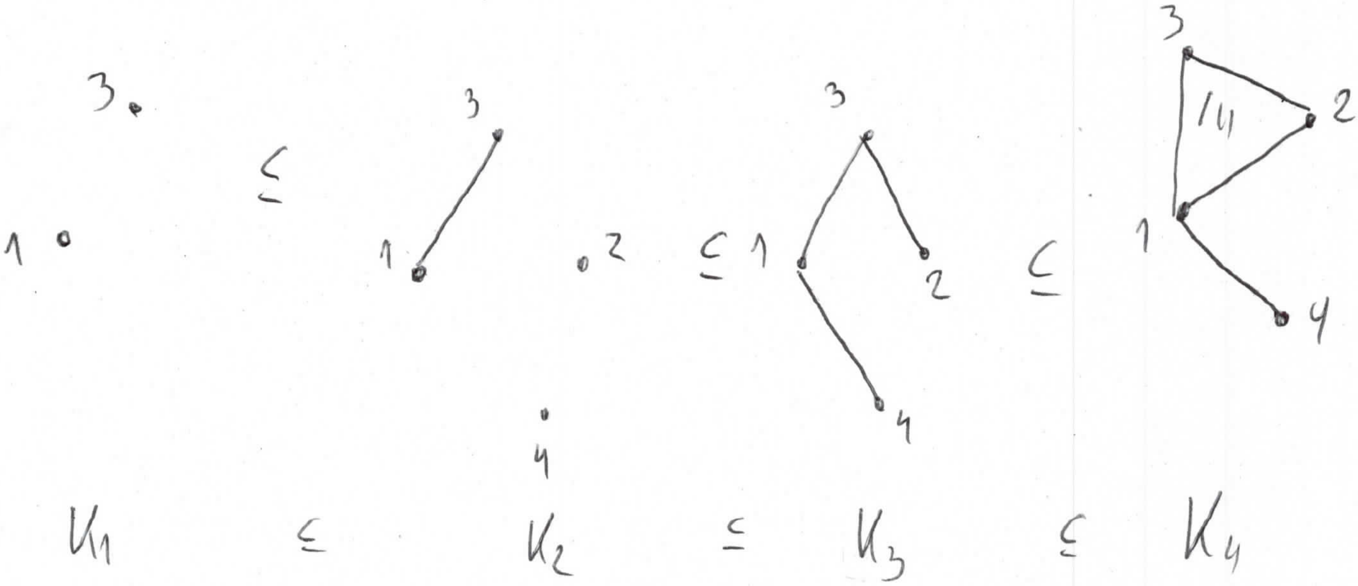
We assume that each intermediate element K_i is a complex, i.e. every element in K_i have all boundary elements in K_i.
This is an example of a cubical filtration:



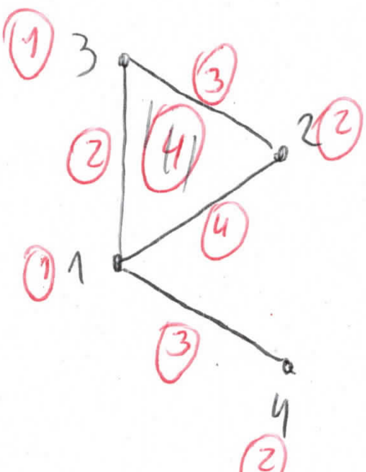
In this case we can mark each cube in this cubical complex by the time it enters the filtration:



Likewise in the case of simplicial complexes, a filtration is a nested sequence of subcomplexes:



Likewise, we can denote the filtration:



We will see, that filtration is the key concept in persistent homology.

But, before we get there, let us discuss & remind a few standard ways of building complexes from point cloud

a.k.a Victoris-Rips, or simply V-R complex

1 Rips complex - given a collection of points P, a metric d and $\varepsilon > 0$ a collection of points $x_0, \dots, x_n$ constitute a simplex in VR complex if $d(x_i, x_j) \leq \varepsilon \quad \forall\, i,j \in \{0, \dots, n\}$.

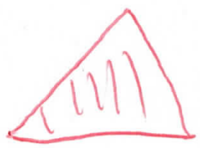
In general, maximal distance between two points gives a filtration value of a simplex.

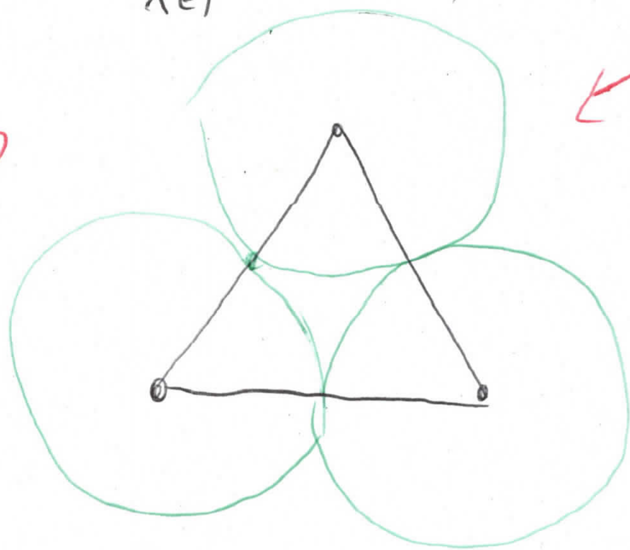
N.b. it is obvious that a simplex in VR complex enters filtration not before its boundary, because a diameter of a subcollection of points $\leq$ diameter of a collection of points.

V-R complex is conceptually simple & easy to implement. ⑥
It has however a couple of issues:

① Given a point cloud P, metric d and $\varepsilon > 0$ topology of V-R complex is not the same as topology of $\bigcup_{x \in P} B(x, \varepsilon)$. Here is an example

Rips complex is





equilateral triangle of a side length 2ε.
Balls of radius ε intersect.
Topologically equivalent to S^1

② If we choose large ε, then every point in P will be connected to every other point in P. That gives $\binom{n}{2}$ edges, $\binom{n}{3}$ 2-simplices, $\binom{n}{4}$ 3 simplices, ..., $\binom{n}{k}$ $k+1$ simplices

we call this a combinatorial explosion

Even for a planar point cloud V-R complex may have simplices of arbitrary high dimension.

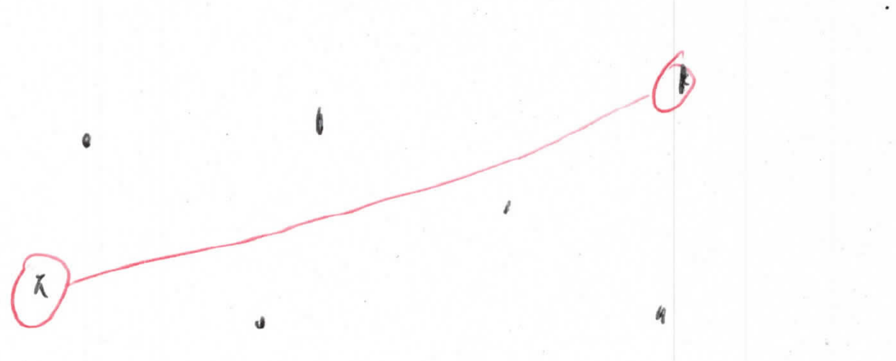
2) Čech / Nerve complex. Let P be a point cloud in $\mathbb{R}^n$ and $\varepsilon > 0$. Points $x_0, \ldots, x_m$ constitute m dimensional simplex iff

$$\bigcap_{x \in \{x_0, \ldots x_m\}} B(x, \varepsilon) \neq \emptyset$$

Čech complex, due to the Nerve theorem, have the same homotopy type as $\bigcup_{x \in P} B(x, \varepsilon)$. But it still suffer from a problem of combinatorial explosion.

3) A possible solution to the problem of combinatorial explosion is to use so called Alpha complexes. But let us first understand the nature of the problem.

If we have a point cloud P:



The problem is that for $\varepsilon \sim$ diameter of P points that are very far away will get connected. Clearly, at that scale, we do not expect much change in topology, as the contributions of points are localized. But, for those large values of ε, exponential number of simplices will enter a complex.

To proceed with alpha-complex construction we need two main ingredients:

1. Nerve Theorem: Let C be a collection of convex sets. Nerve complex have a simplex per every nonempty intersection of elements from C. Then, nerve complex

is homotopy equivalent to $\bigcup \mathcal{C}$.

N.b. this theorem can be stated in more general way, but this formulation suffice for our case.

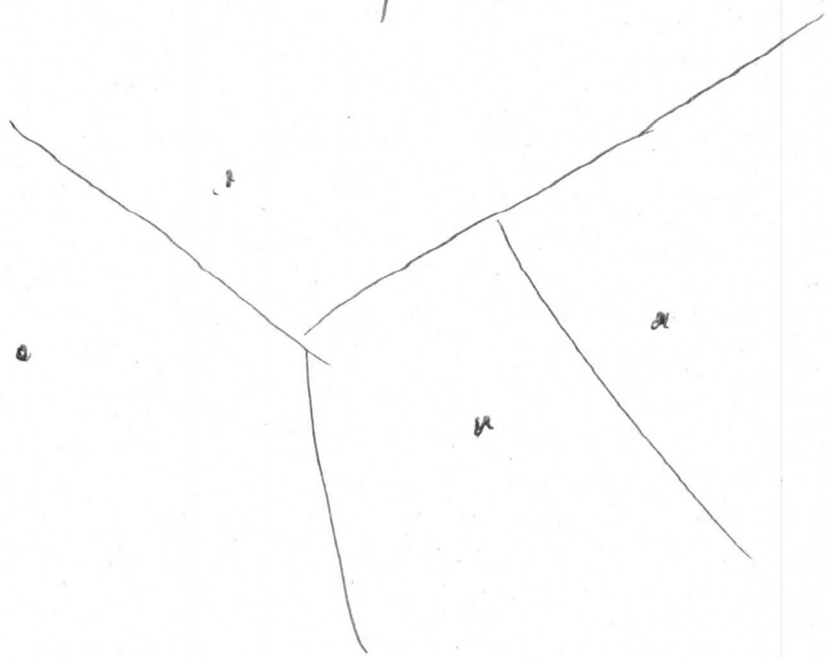
(2) Voronoi diagram of a point cloud P is a partition of the ambient space X into voronoi cells

$$R_k = \{ x \in X \mid d(p_k, x) \leq d(p_l, x) \},$$

where $P = \{ p_1, \dots, p_m \}$.

I.e. we partition the space X into regions that have unique closest point in P

Eg:



Fact: Voronoi cells are convex.

Fact: $B(x,r)$ is convex

Fact: Intersection of two convex sets is a convex set.

Therefore we can use the intersection of Balls from the Čech complex with the corresponding Voronoi cells and use this in Nerve theorem (to get a complex that is homotopically equivalent to union of balls - this is not trivial to show this!)

This is how alpha complex is constructed.

Cool fact from computational geometry → often the number of simplices in the alpha complex is a linear function of a number of points in P.
(function depends on the dimension of the ambient space).