Section 4

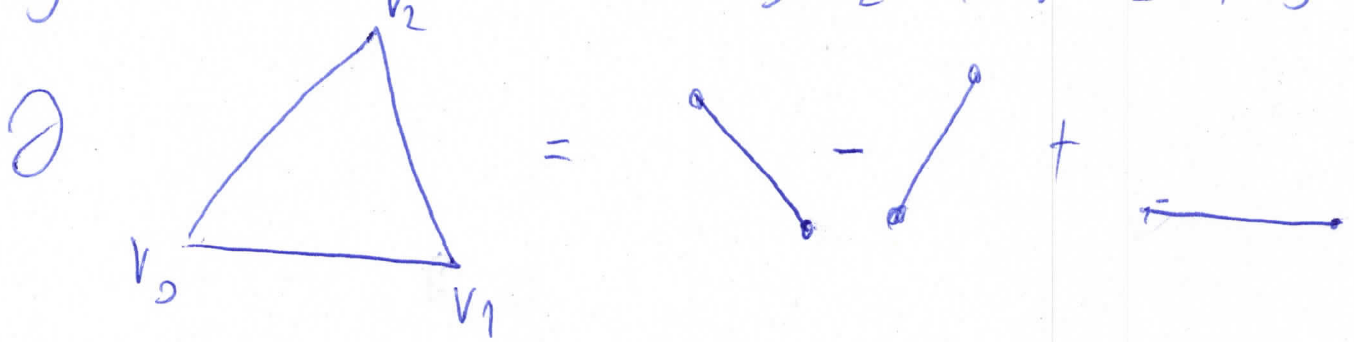
Short reminder on what is Persistent Homology.

You have already covered the standard concept of homology and persistent homology on your lectures so far.
In this section we will quickly revisit those ideas and talk about their computability.
Let us revisit first the boundary operator.
For simplicial complexes:

it means that this vertex is missing

$$\partial [v_0, v_1, \dots, v_n] = \sum_{i=0}^{n} (-1)^i [v_0, v_1, \dots, \bar{v}_i, v_{i+1}, \dots v_n]$$

Eg $\partial [v_0 \, v_1 \, v_2] = [v_1, v_2] - [v_0, v_2] + [v_0, v_1]$



(note, we swap under the rug the orientation of simplices - this is not required for persistent homology calculation).

We will skip the definition for cubical complexes, as it is more technical. Consult Kaczyński - Mischaikow - Mrozek book for details.

It is a classical exercise to show that $\partial\partial = 0$

Given a simplicial complex K and a field F, let us define

Typically we take a finite field here.

$$C(K, F) = \left\{ \sum_{S \in K} \alpha_S S, \text{ where } \alpha_S \in F \right\}$$

Those are chains of K.

space of all possible linear combinations of elements of S with coefs from F

If we restrict to simplices of a certain dimension, we will get chains in this dimension:

$$C_i(K, F) = \left\{ \sum_{S \in K_i} \alpha_S S, \text{ where } \alpha_S \in F \right\}$$

i dimensional simplices in the complex K

Given i-dimensional chains $C_i(K,F)$, there are two subgroups of them (N.B. $C_i(K,F)$ is a group):

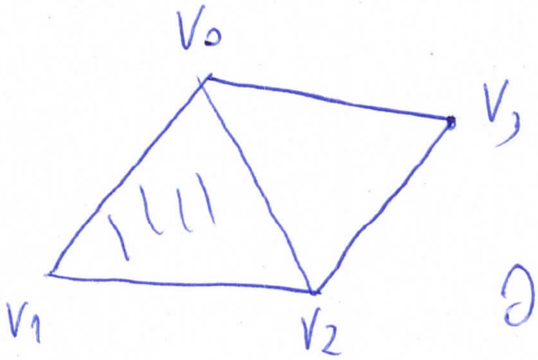
① **Cycles**, $z \in C_i(K,F)$ is a cycle, if $\partial z = 0$, where boundary of a chain $\partial \sum_{s \in K_i} \alpha_s s = \sum_{s \in K_i} \alpha_s \partial s$.

∂ is a linear operator.

$Z_n(K,F)$ is a classical symbol to denote cycles

② **Boundaries**, $z \in C_i(K,F)$ is a boundary if there exist $c \in C_{i+1}(K,F)$ such that $\partial c = z$.

Eg.



$c = [V_0 V_2] + [V_0 V_1] + [V_1 V_2]$ is a Z_2 cycle, yes:

$\partial c = V_0 + V_2 + V_0 + V_1 + V_1 + V_2 = 0$

$1+1=0 \mod 2$

$d = [V_1 V_0] + [V_1 V_2] + [V_0 V_2]$ is a boundary over Z_2, yes

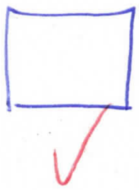
$\partial [V_0 V_1 V_2] = d$ (over Z_2)

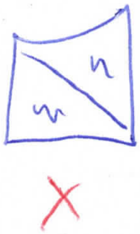
$B_n(K,F)$ is a classical symbol to denote boundaries

Observation $\partial\partial = 0 \Rightarrow B_n(K,F)$ is a subgroup of $Z_n(K,F)$.

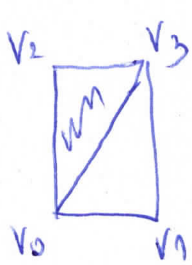
Now, homology $H_n(K,F) = Z_n(K,F) / B_n(K,F)$

quotient group.

Informally, we detect cycles  that are not bounded:



Note another consequence of taking the quotient group. Suppose we have two cycles surrounding the same hole: ← psst, it is a common intuition to understand homology groups as descriptors of holes & cycles that are not boundaries as their "approximations".



$[v_0 v_1] + [v_1 v_3] + [v_0 v_3]$ surround the same hole.

as well as

$[v_0 v_1] + [v_1 v_3] + [v_2 v_3] + [v_0 v_2]$

Their difference is the boundary $[v_0 v_3] + [v_2 v_3] + [v_0 v_2]$.

From the perspective of homology they are the same.

Now... persistent homology.

Let us take a filtration

$$K_1 \subseteq K_2 \subseteq \dots \subseteq K_{n-1} \subseteq K_n$$

If we apply the homology functor to this sequence of inclusions we get:

maps induced by inclusion

$$0 \to H(K_1) \to H(K_2) \to \dots \to H(K_{n-1}) \to H(K_n) \to 0$$

If there is a class $c \in H(K_i)$ that is not present in an image of $H(K_{i-1})$ we say that a class is born.

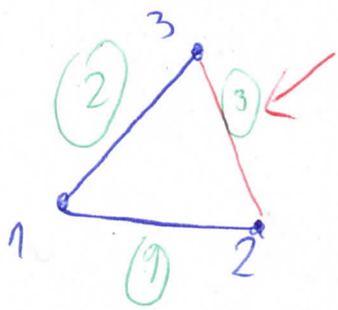
If a class $c \in H(K_{i-1})$, $c \neq 0$, after mapping to $H(K_i)$ become trivial, or become identical (homologous) to a class that is born earlier, then we say that c dies in $H(K_i)$

The birth-death events constitute persistence intervals that we have played with so far.

Please consider the jupyter notebook for a number of examples.

How to compute persistent homology? It is really the question if new cell that come to a filtration creates or destroy a nontrivial cycle (it have to do either of those)

Eg



Suppose vertices appear first in the filtration followed by edges [1,2], [1,3]. At this point the boundary of an edge [2,3] can be generated with what already exist in the filtration, namely a chain

[1,2] + [1,3]. Since

$$\partial([1,2]+[1,3]) = [2]+[3] = \partial[2,3]$$

$$\Downarrow$$

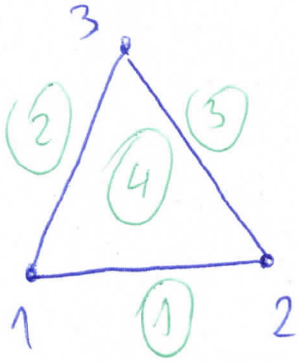
$$\partial(\overbrace{[1,2]+[1,3]+[2,3]}) = 0$$

this is a cycle

At the moment the edge [2,3] enters the filtration it cannot be a boundary, as there is no simplex having [2,3] in its boundary at this level of filtration.

In this case we have a cycle that is not a boundary and therefore a new homology class

Let us now extend the filtration from a previous example by adding a 2-simplex [1, 2, 3] at time 4:

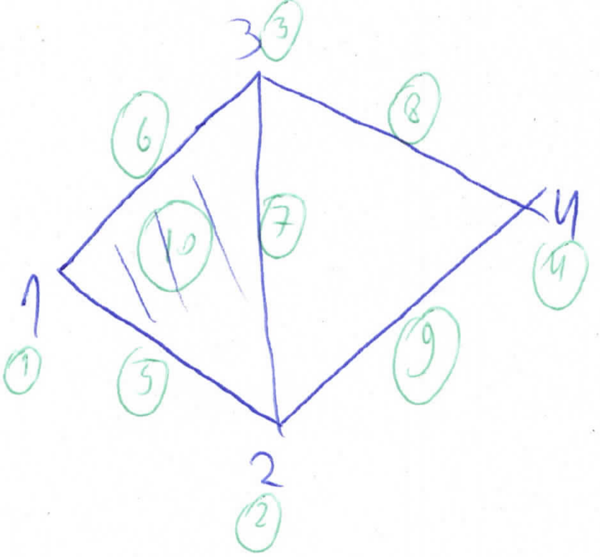


In this case $\partial[1,2,3] \overset{\text{over } \mathbb{Z}_2}{=} [1,2]+[1,3]+[2,3]$ and therefore the homologically nontrivial cycle that we have created in the last step / that was born in the last step became trivial / dies.

It turns out that the question if a new simplex/cell closes a new homology classes, or terminates the existing one, can be answered using a version of a Gaussian elimination procedure.

Let us take a look at one example:

(8)



Let us consider a simplicial complex on the left. It has:

① Four vertices: 1, 2, 3, 4 ← appearing in this order

② Five edges:

[1,2], [1,3], [2,3], [3,4], [2,4]

↑ appearing in this order in the filtration.

③ One 2-simplex: [1,2,3]

Filtration provides a total ordering on simplices. When constructing a boundary matrix they will appear in this order (both for rows and columns). ← filtered boundary matrix

In the matrix ~~abo~~ below we skip the columns corresponding to vertices, as they have empty boundary.

We also mark only nonzero entries ← matrices that stores only nonzero entries can be represented by SPARSE MATRIX data structures.

	12	13	23 13 + 12	34	+13+12 24+34	123	
1	1	1	✗	conflict	~~1~~		
2	1		~~1~~	conflict	~~1~~		
3		1	~~1~~	1	~~1~~		
4				1	~~1~~		
12		conflict				1	
13				conflict		1	
23	subsequent conflict					1	
	low(12) = 2	low(13) = 3					

The matrix reduction algorithm has the following form:

Input : M - filtered boundary matrix

Output : Reduced boundary matrix, i.e. $\forall$ column of the reduced matrix the lowest nonzero entry is unique.

```
for i = 1 to number of rows in M
    while low(i) = low(j) for j < i
        row(i) = row(i) + row(j)
    end
end
```

Here is the reduced matrix (this time with all the columns)

	1	2	3	4	12	13	23 13+12 +	34	24 +13+12 +34	123
1	zero column, homology class born	zero column, homology class born	zero column, homology class born	zero column, homology class born	1	1	zero column, homology class born		zero column, homology class born	
2					1					
3					↑	1		1		
4					class created	↑		1		
12					at 2 killed	class created		↑ class created		1
13					by 12	at 3 killed		of 4 killed by		1
23						by 13		34		1

class created at 23 killed at 123

Lowest nonzero entries for reduced columns are unique.

Observation 1: The process of addition as in columns 23 & 24 correspond to finding a cycle that is closed by a given simplex. Once the column become zero, we know that the cycle was found and that it generate a new homology class

Lowest nonzero entry correspond to a cell that cannot close a cycle in the complex up to its level of filtration.

Algebraically it is because the lowest nonzero entry cannot be reduced by anything to the left from it.

Geometrically, because lowest nonzero entry is defined, that means that simplex under consideration S have nonzero dimension. That implies that there used to be a nontrivial cycle, represented for instance by ∂S, that has been killed by adding S. This cycle was created by the ~~oldest simplex~~ simplex in ∂S that appear last in the filtration

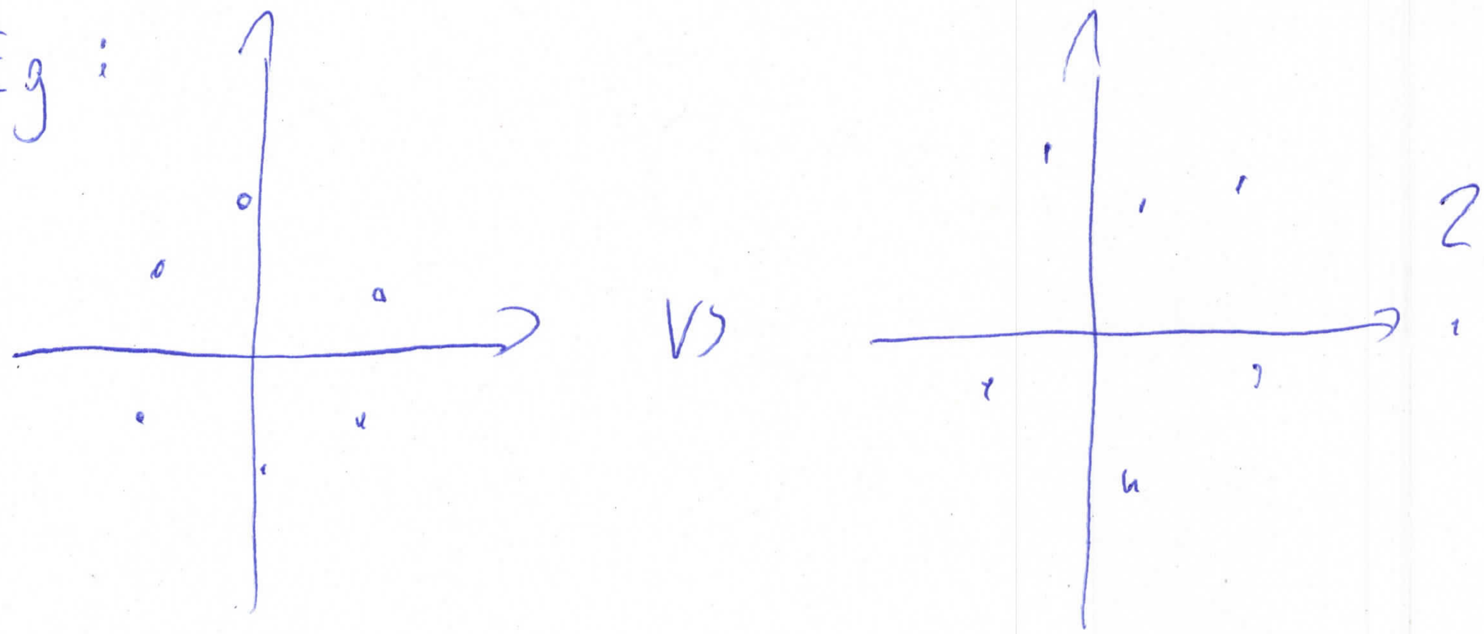
→ the filtration value of this simplex & filtration value of S give a persistence pair.

Supplementary: Geometrical stability of persistence.

We know & encounter examples of stability of persistent homology in case where we change the filtering function.

But what happen if we have a point cloud P & perturb it by adding a <u>bounded</u> <u>noise</u> to obtain P'?

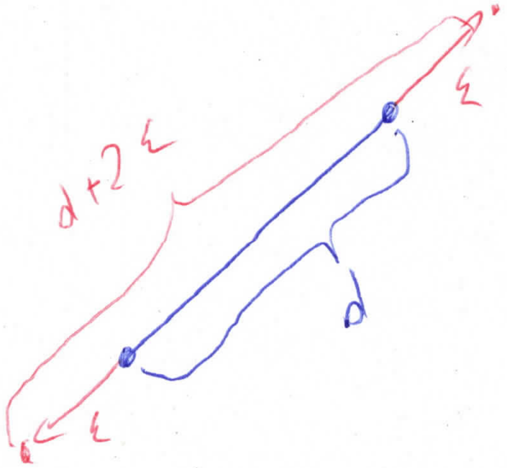
Eg:



Suppose we pick ε in VR large or use α-complex so that all homology classes but the unique infinite interval in dimension 0 die.

Let us then pick any simplex S from non perturbed point cloud complex & its perturbed version S'.

Then the edge lenghts of each simplex will move/change by no more than ε:



where ε is the bound ofth the noise.

Therefore the filtration of those simplices will change by at most 2ε.

Assuming that the final complex in both cases is topologically trivial (what implies that all nontrivial homology classes die) it implies, from the stability theorem, that the persistence of each of the class will not change by more than 2ε.

Let us consult the Jupyter notebook for an example.