Persistent Homology

the basics

kelly davis (founder forty.to)

Image:Jean-Marie Hullot (CC BY 3.0)

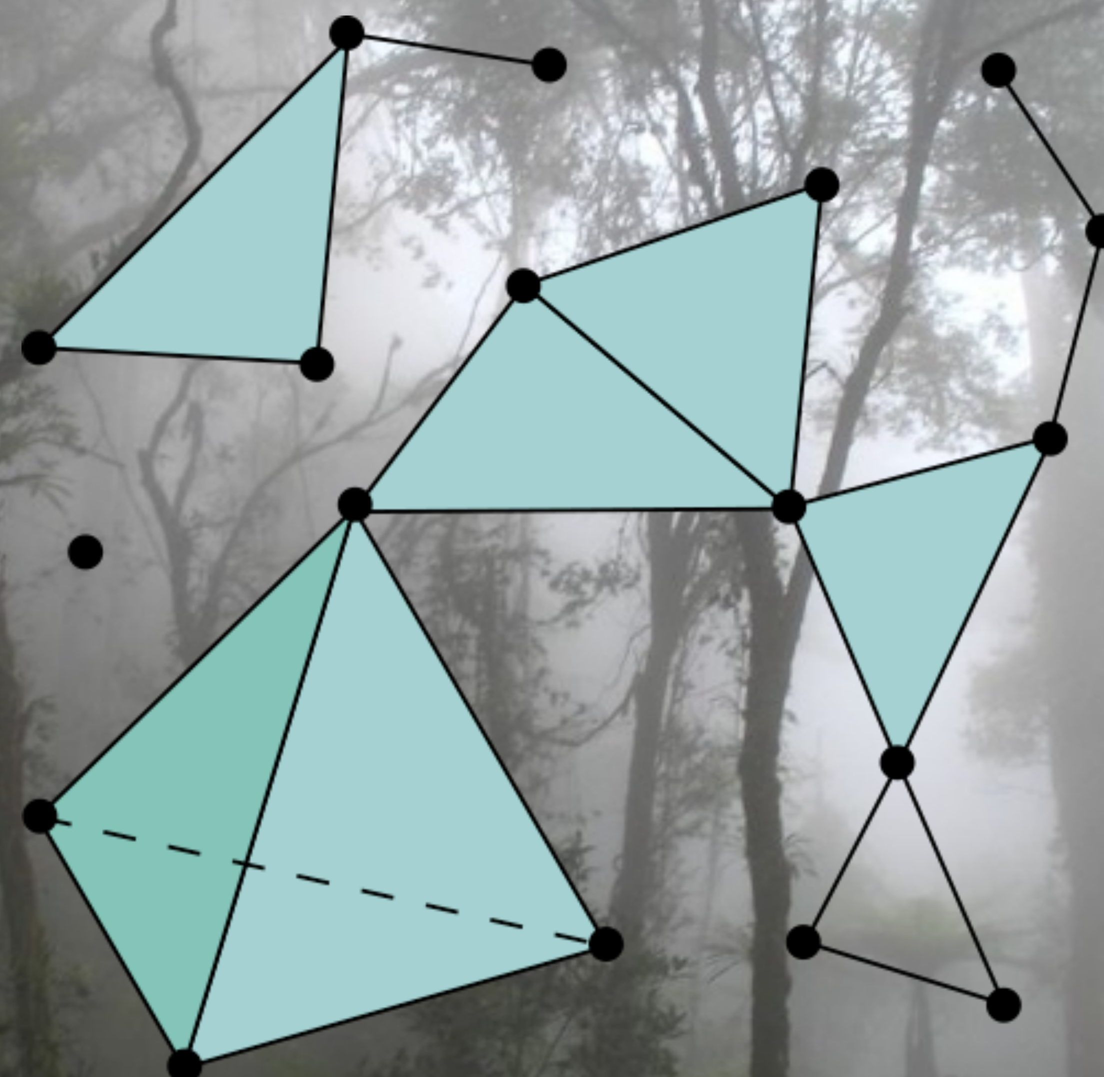
problem: what's the topology of a point cloud?

Image: Jean-Marie Hullot (CC BY 3.0)

solution: persistent homology

Image:Jean-Marie Hullot (CC BY 2.0)

simplicial complex

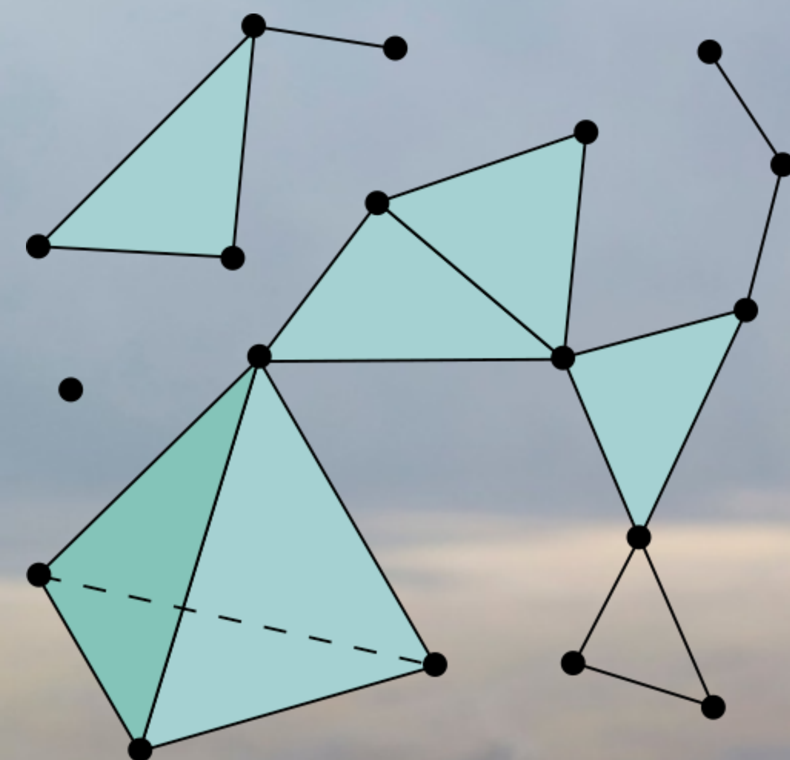


Simplicial Complex - A set K of simplicies such that:

- Any face of a simplex in K is also in K.
- The intersection of distinct simplicies in K is a face of both

Image:Antoine Hubert (CC BY 2.0)

simplicial homology



p-chain - A formal sum of p-simplicies $c = \sum_i \gamma_i \sigma_i$ with $\gamma_i \in \mathbb{F}$ a field.

$C_p(K;\mathbb{F})$ - Vector space of all p-simplicies in K over $\mathbb{F}$

p-cycles - Elements in $C_p(K;\mathbb{F})$ in the kernel of ∂, denoted $Z_p(K;\mathbb{F})$

p-boundaries - Elements in $C_p(K;\mathbb{F})$ in the image of ∂, denoted $B_p(K;\mathbb{F})$

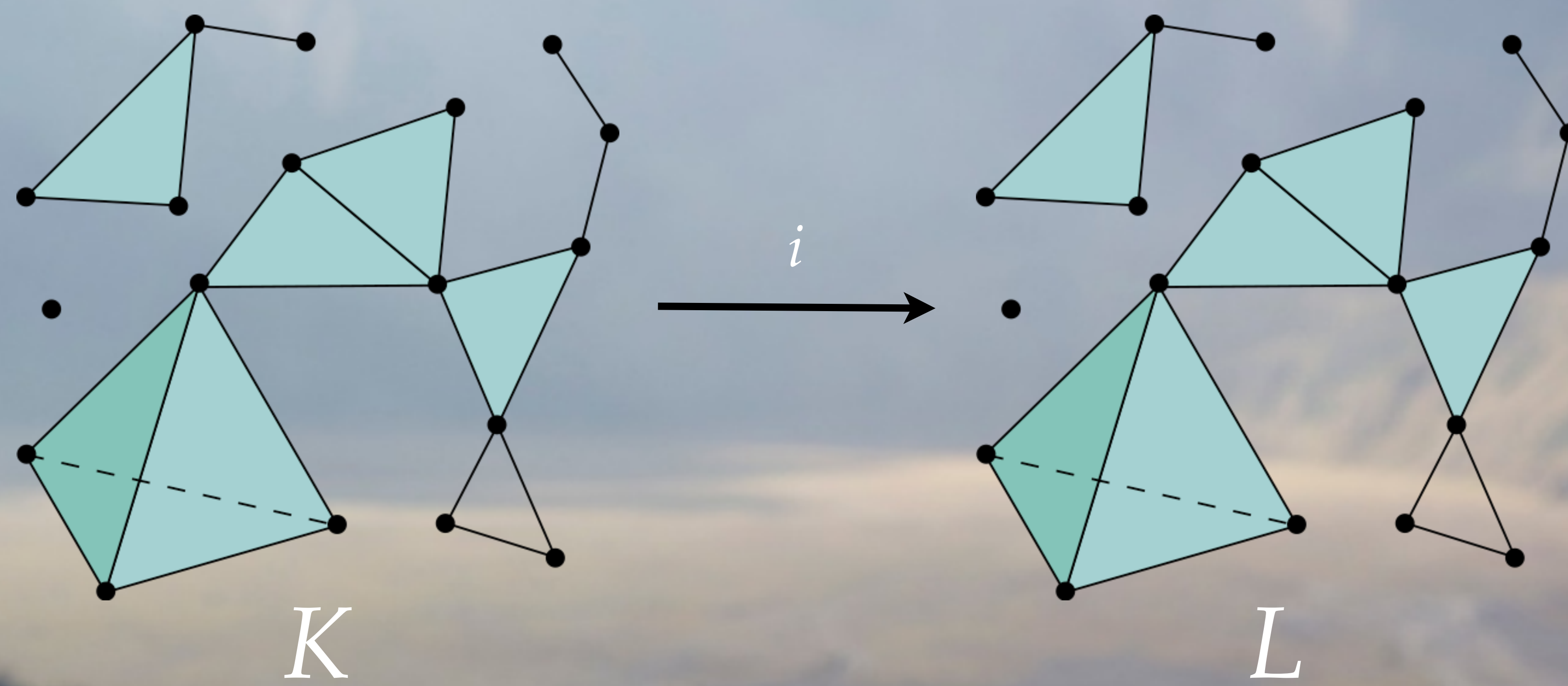
Simplicial Homology Group - $H_p(K;\mathbb{F}) := Z_p(K;\mathbb{F})/B_p(K;\mathbb{F})$

Betti Numbers - The numbers $\beta_p(K;\mathbb{F}) := dim(H_p(K;\mathbb{F}))$

Note: We will always work over $\mathbb{Z}_2$

Image:Jean-Marie Hullot (CC BY 2.0)

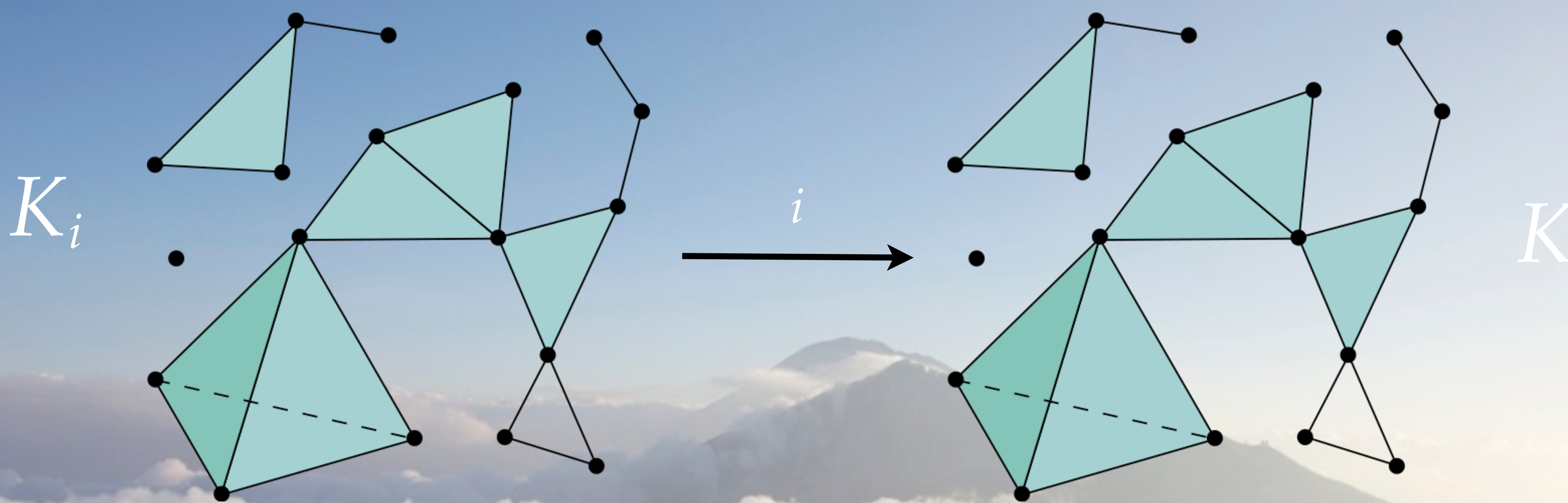
simplicial homology



Theorem. *If K and L are simplicial complexes and there exists an inclusion $i : K \hookrightarrow L$, the inclusion generates a homomorphism $i^* : H_p(K, \mathbb{F}) \to H_p(L, \mathbb{F})$.*

Image: Jean-Marie Hullot (CC BY 2.0)

simplicial filtrations



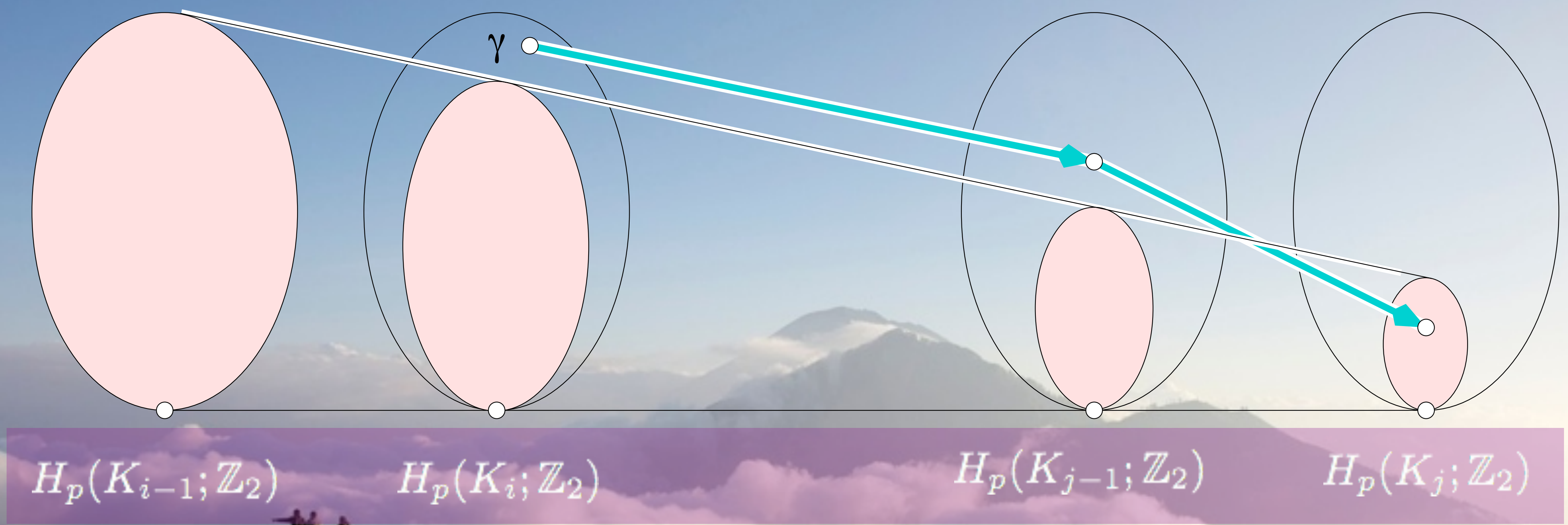
Definition. *Subcomplex - A subcomplex K_i of a simplicial complex K is a simplicial complex such that $K_i \subset K$*

Definition. *Filtration - A nested sequence of subcomplexes $\{K_i\}$ that starts with $\emptyset$ and ends with K*

$$\emptyset = K_0 \hookrightarrow K_1 \hookrightarrow K_2 \cdots \hookrightarrow K_{m-1} \hookrightarrow K_m = K$$

Image:Jean-Marie Hullot (CC BY 3.0)

birth, death, and taxes



Remark. *A filtration* $\{K_i\}$ *generates a sequence of homomorphisms*

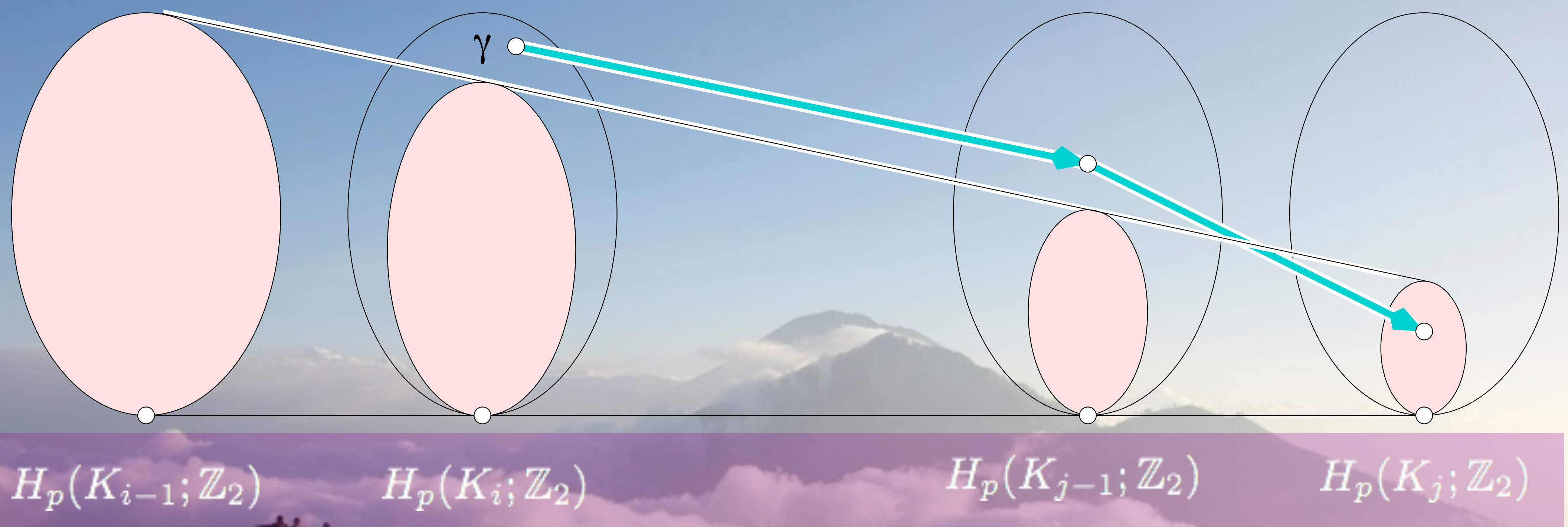
$$H_p(K_0;\mathbb{Z}_2) \to H_p(K_1;\mathbb{Z}_2) \to H_p(K_2;\mathbb{Z}_2) \cdots \to H_p(K_m;\mathbb{Z}_2)$$

Definition. *Born - A homology class* γ *is born at* K_i *if it is not in the image of the homomorphism induced by* $K_{i-1} \hookrightarrow K_i$

Definition. *Dies - A homology class* γ *dies at* K_j *if the image of the homomorphism induced by* $K_{i-1} \hookrightarrow K_{j-1}$ *does not contain the image of* γ*, but the image of the homomorphism induced by* $K_{i-1} \hookrightarrow K_j$ *does*

Image:Jean-Marie Hullot (CC BY 3.0)

persistence



Definition. *Persistence - The persistence of a homology element γ born at K_i and dying at K_j is $j - i$*

Remark. *What does this have to do with our original question?*

Remark. *How do we compute the Betti numbers $\beta_p(K;\mathbb{Z}_2)$*

Remark. *How do we compute persistence?*

Image:Jean-Marie Hullot (CC BY 3.0)

so what!

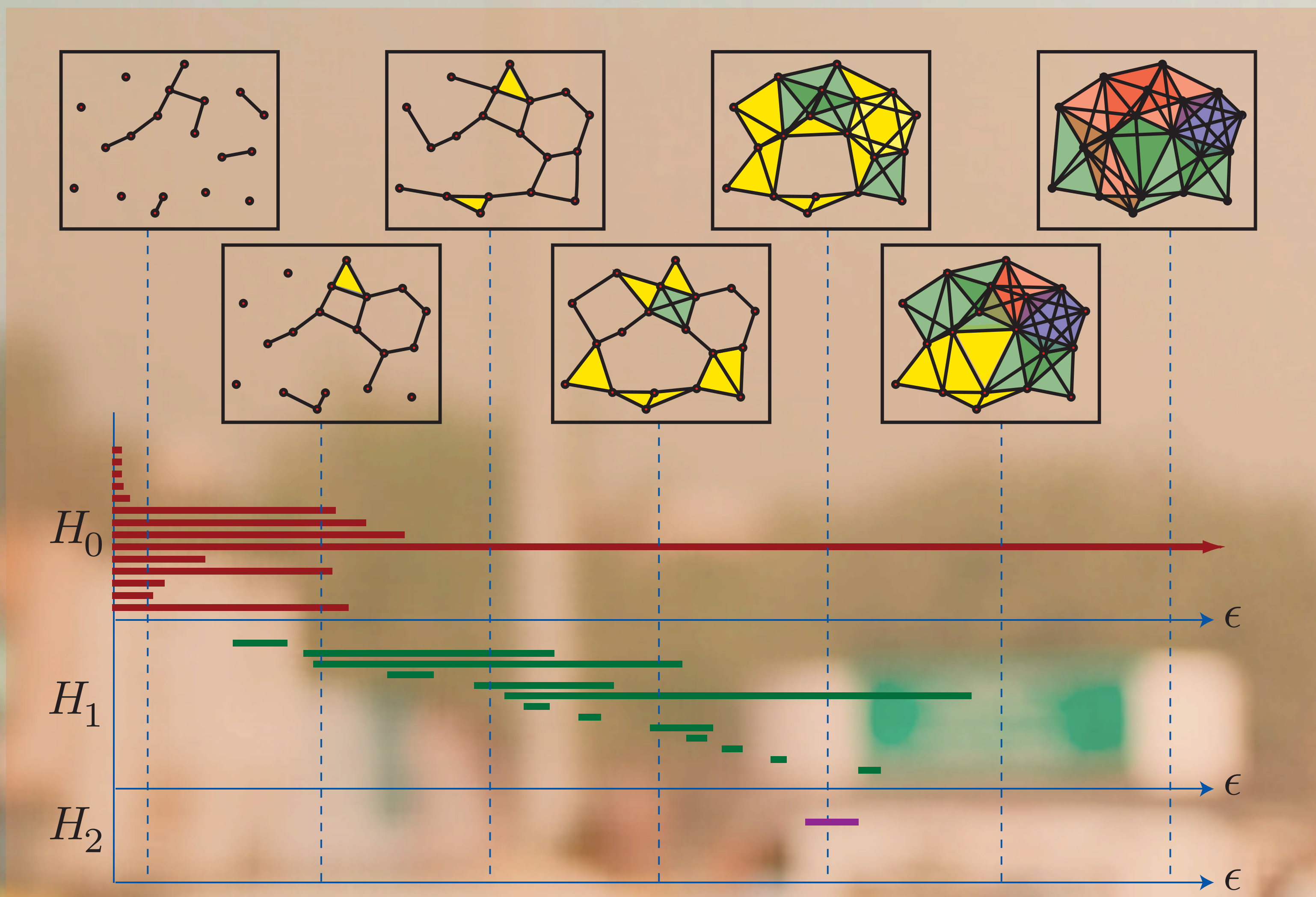


Image:Padmanaba01 (CC BY 2.0)

calm before the algorithm

Definition. *Compatible Total Ordering - A compatible total ordering on a simplicial complex K with filtration $\{K_i\}$ is a total ordering on the simplicies $\sigma_i \in K$ such that:*

- Simplicies in each K_i precede those in $K - K_i$
- Faces of a simplex precede the simplex

Definition. *Boundary Matrix - Matrix D_{ij} with rows and columns corresponding to simplicies of K in a compatible total ordering such that $D_{ij} = 1$ if σ_i is a codimension 1 face of σ_j and $D_{ij} = 0$ otherwise.*

Definition. *$low(j)$ - The map $low(j)$ is defined to be the row number in D_{ij} of the lowest non-zero entry in column j. If column j contains only 0's, then $low(j) = 0$*

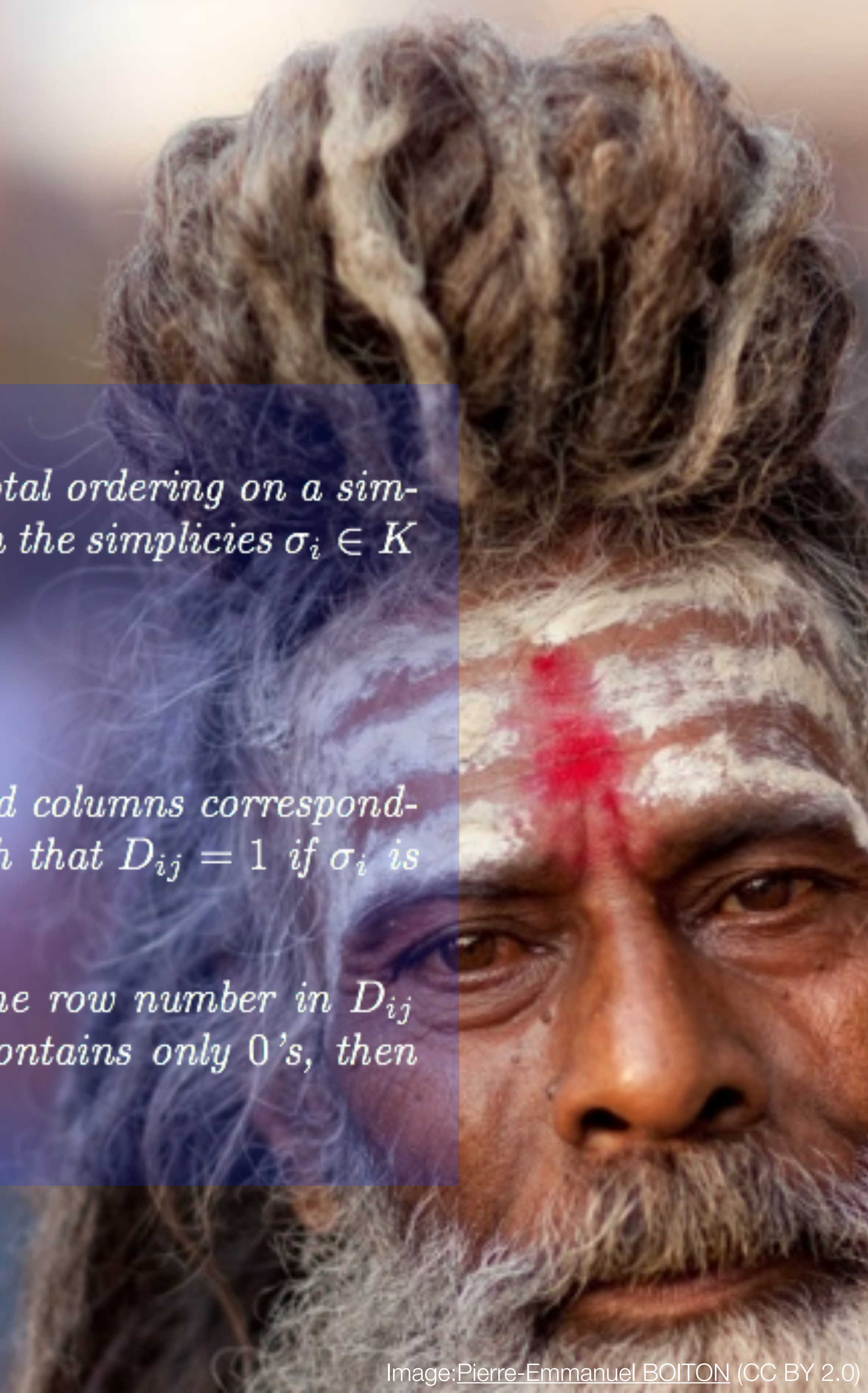


Image:Pierre-Emmanuel BOITON (CC BY 2.0)

the algorithm: betti numbers

for $j = 1$ *to* $|K|$ **do**
    while $\exists\, j' < j$ with $low(j') = low(j) \neq 0$ **do**
        add column j' of D_{ik} to column j of D_{ik}
    end while
end for

Definition. $\#Zero_p(R)$ - *Number of zero columns in R that correspond to p-simplicies.*

Definition. $\#Low_p(R)$ - *Number of lowest ones in rows that correspond to p-simplicies*

Lemma. $B_p(K;\mathbb{Z}_2) = \#Zero_p(R) - \#Low_p(R)$

Image:Pierre-Emmanuel BOITON (CC BY 2.0)

the algorithm: persistence

for $j = 1$ *to* $|K|$ **do**
  while $\exists\ j' < j$ with $low(j') = low(j) \neq 0$ **do**
    add column j' of D_{ik} to column j of D_{ik}
  end while
end for

Remark. *The persistence of the simplicies can be calculated as follows:*

- *If* $low(j) = i > 0$, *then the persistence of* σ_i *is* $j - i$.
- *If* $low(j) = 0$ *and there exists no* k *with* $low(k) = j$, *then* σ_j *essential, i.e. it doesn't die.*

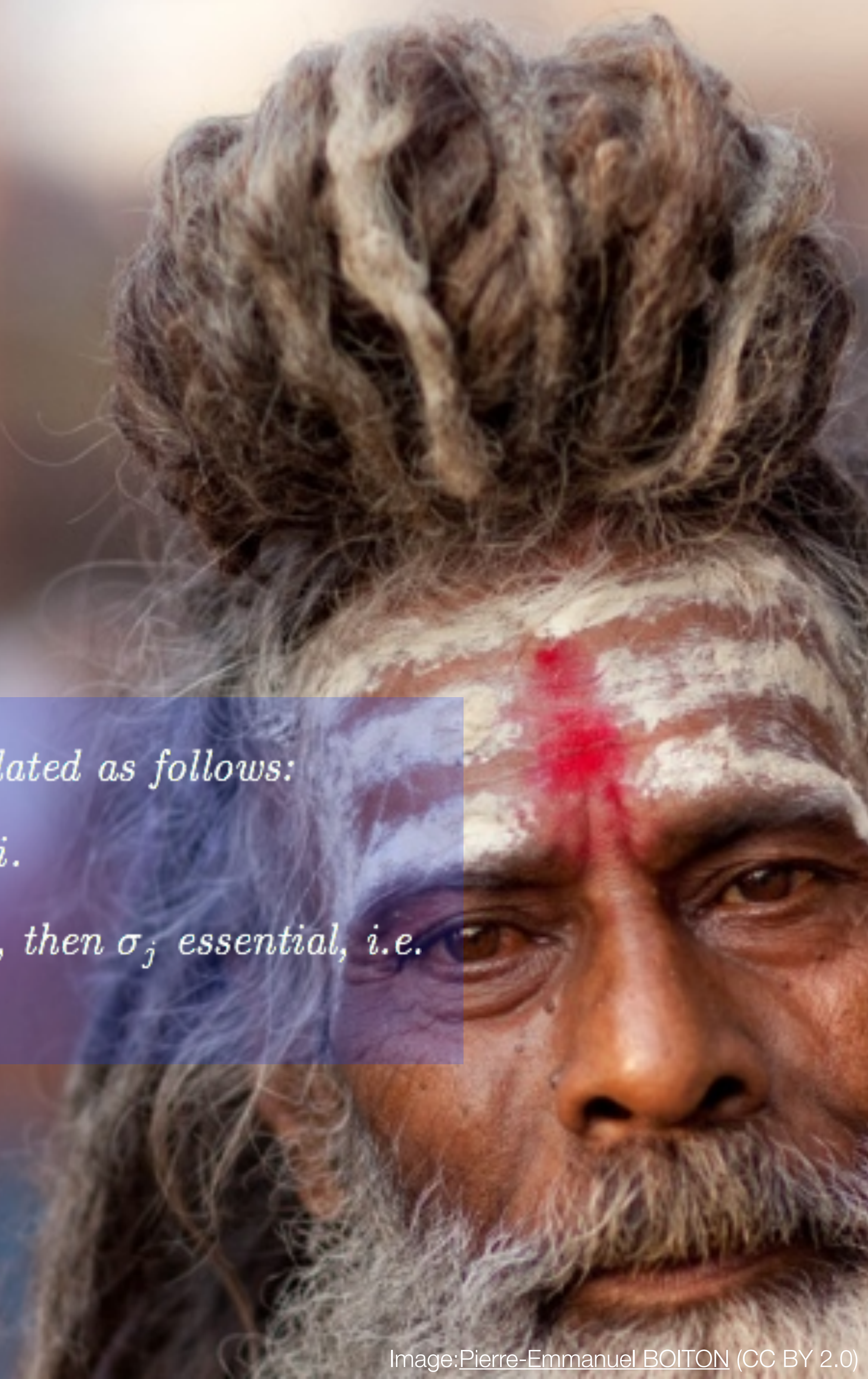


Image:Pierre-Emmanuel BOITON (CC BY 2.0)

Persistent Homology

the basics

kelly davis (founder forty.to)

implementations:

phat - c++ with c++ api
dionysus - c++ with python api
plex - java with java api

Image:Jean-Marie Hullot (CC BY 3.0)