TOPOLOGICAL PERSISTENCE

$$
\begin{array}{ccccccc}
\partial_3 \big\downarrow & & \partial_3 \big\downarrow & & \partial_3 \big\downarrow & & \\
\mathsf{C}_2^0 & \xrightarrow{f^0} & \mathsf{C}_2^1 & \xrightarrow{f^1} & \mathsf{C}_2^2 & \xrightarrow{f^2} & \cdots \\
\partial_2 \big\downarrow & & \partial_2 \big\downarrow & & \partial_2 \big\downarrow & & \\
\mathsf{C}_1^0 & \xrightarrow{f^0} & \mathsf{C}_1^1 & \xrightarrow{f^1} & \mathsf{C}_1^2 & \xrightarrow{f^2} & \cdots \\
\partial_1 \big\downarrow & & \partial_1 \big\downarrow & & \partial_1 \big\downarrow & & \\
\mathsf{C}_0^0 & \xrightarrow{f^0} & \mathsf{C}_0^1 & \xrightarrow{f^1} & \mathsf{C}_0^2 & \xrightarrow{f^2} & \cdots
\end{array}
$$

Afra Zomorodian

Stanford University

April 28, 2004

PLAN

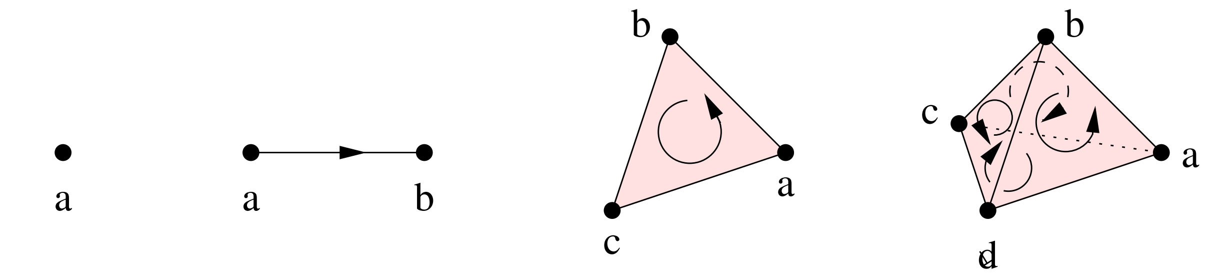
- Main approach: “Go with the flow” - Mehrdad
- Today: Topological Persistence
- Tomorrow: Shape Description via Persistent Homology
  - Theory
  - Practice

OVERVIEW

- Simplicial Complexes
- Homology
- Computing Homology
- Filtrations
- Persistent Homology
- Computing Persistence

SIMPLEX

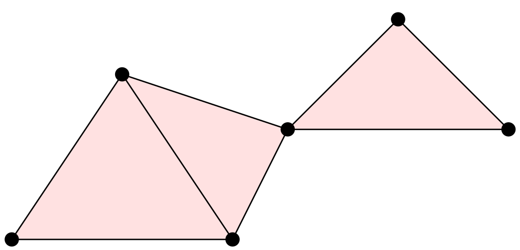
- S, a set of points
- A k-simplex is a subset of S of size $k+1$.
- A simplex may be realized geometrically as the convex hull of $k+1$ affinely independent points in $\mathbb{R}^d$, $d \geq k$.
- An orientation is an equivalence class of orderings $[\sigma]$.

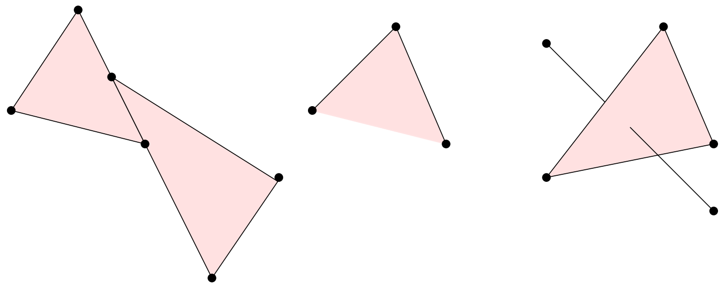


vertex a | edge [a, b] | triangle [a, b, c] | tetrahedron [a, b, c, d]

Simplicial Complex

- A simplex τ defined by $T \subseteq S$ is a face of σ and has σ as a coface.
- A simplicial complex is a set K of simplices on S such that if $\sigma \in K$, then all of σ's faces are in K.
- Realized simplicial complexes have simplices that match along faces.



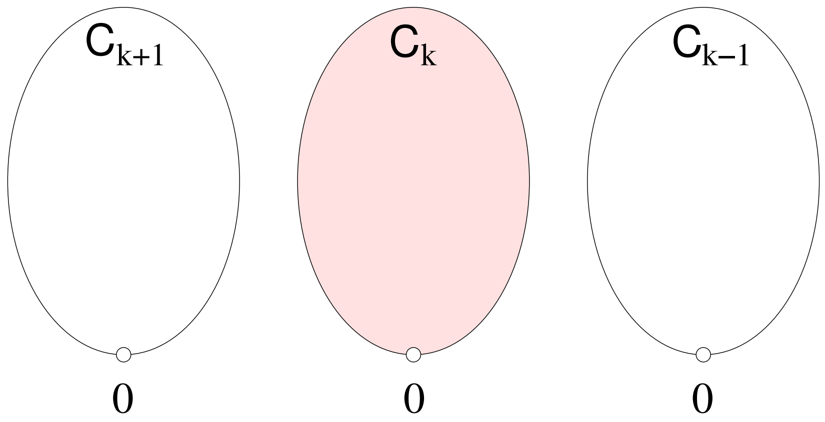


HOMOLOGY

- Algebraization of first layer of geometry in structures
- How cells of dimension n attach to cells of dimension $n - 1$
- This lecture: cells are simplices
- Con:
  - Coarse
  - Less transparent
  - More machinery
- Pro:
  - Combinatorial
  - Finite description
  - Computable

CHAIN GROUP

- Simplicial complex K
- k-chain: $c = \sum_i n_i[\sigma_i], n_i \in \mathbb{Z}, \sigma_i \in K$ (like a path)
- $[\sigma] = -[\tau]$ if $\sigma = \tau$ and σ and τ have different orientations.
- The kth chain group C_k of K is the free abelian group on its set of oriented k-simplices



BOUNDARY OPERATOR

- The boundary operator $\partial_k : \mathsf{C}_k \rightarrow \mathsf{C}_{k-1}$ is a homomorphism defined linearly on a chain c by its action on any simplex $\sigma = [v_0, v_1, \ldots, v_k] \in c$,

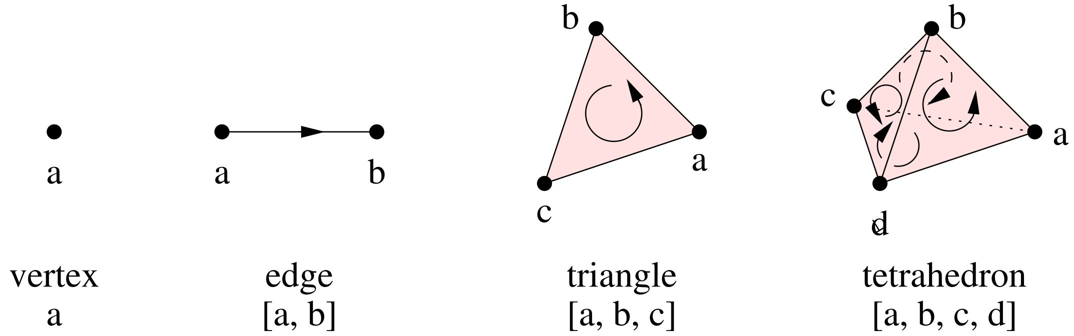
$$\partial_k \sigma = \sum_i (-1)^i [v_0, v_1, \ldots, \hat{v}_i, \ldots, v_k],$$

  where $\hat{v}_i$ indicates that v_i is deleted from the sequence.

- (Theorem) $\partial_{k-1}\partial_k = 0$, for all k.

ORIENTED BOUNDARIES

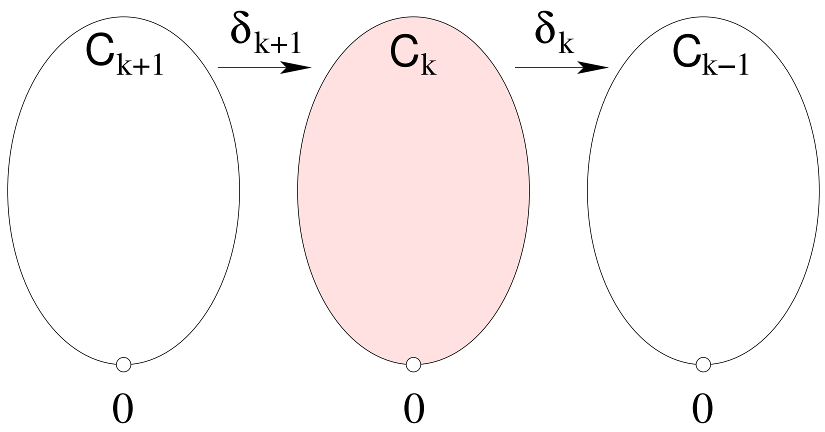
- $\partial_1[a, b] = b - a.$
- $\partial_2[a, b, c] = [b, c] - [a, c] + [a, b] = [b, c] + [c, a] + [a, b].$
- $\partial_3[a, b, c, d] = [b, c, d] - [a, c, d] + [a, b, d] - [a, b, c].$
- $\partial_1\partial_2[a, b, c] = [c] - [b] - [c] + [a] + [b] - [a] = 0.$



CHAIN COMPLEX

- The boundary operator connects the chain groups into a chain complex C_*:

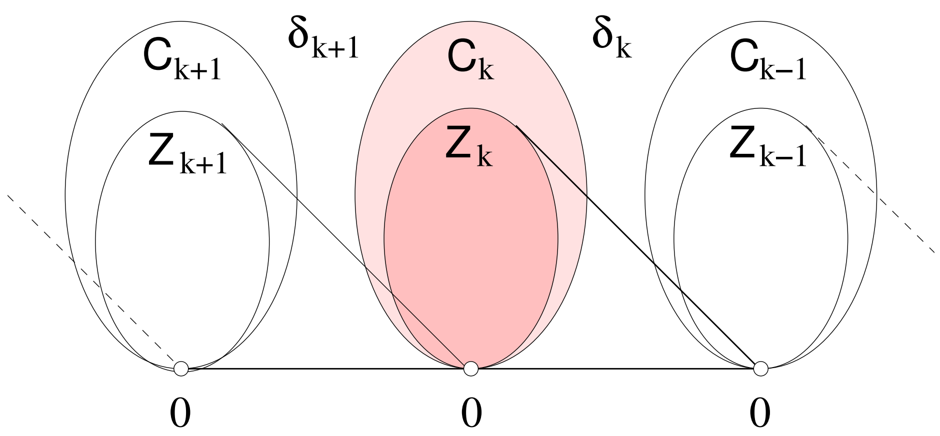
$$\ldots \rightarrow \mathsf{C}_{k+1} \xrightarrow{\partial_{k+1}} \mathsf{C}_k \xrightarrow{\partial_k} \mathsf{C}_{k-1} \rightarrow \ldots.$$



CYCLE GROUP

- Let c be a k-chain
- If it has no boundary, it is a k-cycle
- $\partial_k c = \emptyset$, so $c \in \ker \partial_k$
- The kth cycle group is

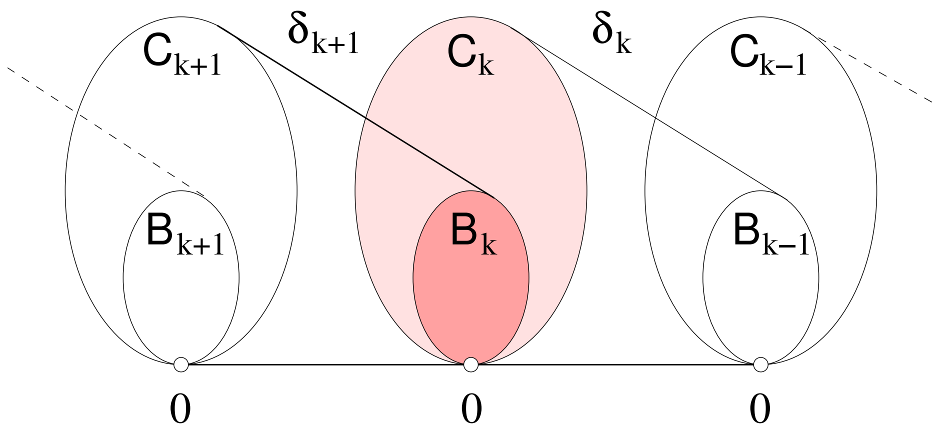
$$\mathsf{Z}_k = \ker \partial_k = \{c \in \mathsf{C}_k \mid \partial_k c = \emptyset\}.$$



BOUNDARY GROUP

- Let b be a k-chain
- If b is a boundary of something, it is a k-boundary.
- The kth boundary group is

$$\mathsf{B}_k = \operatorname{im} \partial_{k+1} = \{c \in \mathsf{C}_k \mid \exists d \in \mathsf{C}_{k+1} : c = \partial_{k+1} d\}.$$

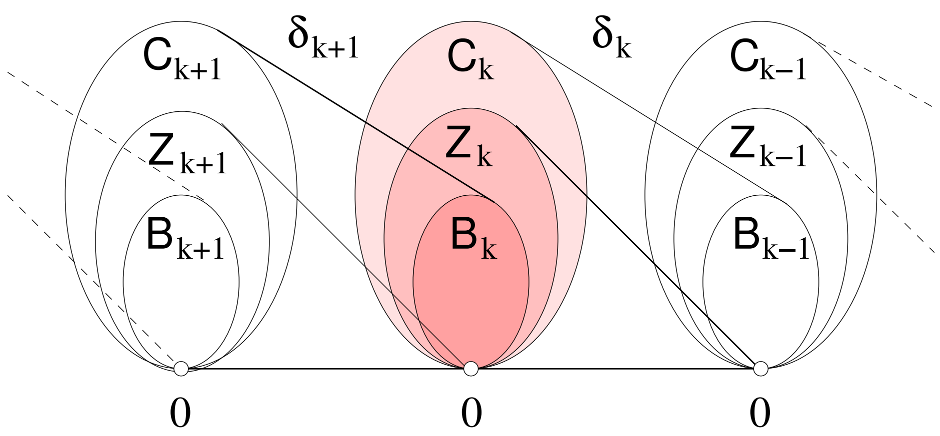


SIMPLICIAL HOMOLOGY

- The kth homology group is

$$\mathsf{H}_k = \mathsf{Z}_k / \mathsf{B}_k = \ker \partial_k / \mathrm{im}\, \partial_{k+1}.$$

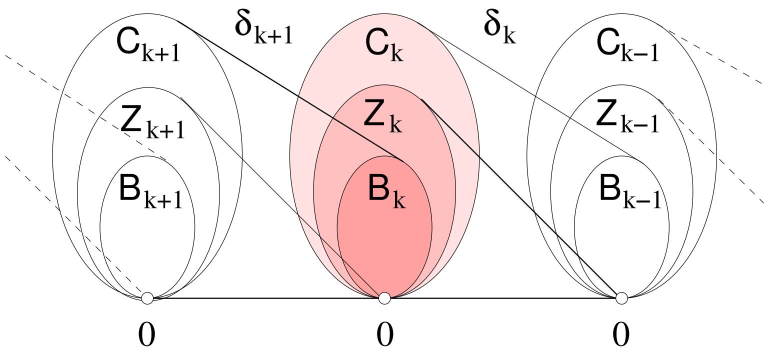
- Betti Numbers: $\beta_k = \operatorname{rank} \mathsf{H}_k = \operatorname{rank} \mathsf{Z}_k - \operatorname{rank} \mathsf{B}_k$



Interpretation

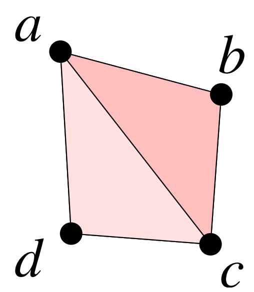
- Subcomplexes of $\mathbb{S}^3$ are torsion-free.
- Alexander Duality:
  - β_0 measures the number of components of the complex.
  - β_1 is the rank of a basis for the tunnels.
  - β_2 counts the number of voids in the complex.

COMPUTING HOMOLOGY



- Compute a basis for $\ker \partial_k$ to get $\operatorname{rank} \mathsf{Z}_k$
- Compute a basis for $\operatorname{im} \partial_{k+1}$ to get $\operatorname{rank} \mathsf{B}_k$
- $\partial_k : \mathsf{C}_k \to \mathsf{C}_{k-1}$ is linear, so it has a matrix
- Use oriented simplices as bases for domain and codomain, so matrix is $m_{k-1} \times m_k$
- M_k is the standard matrix representation for ∂_k

Example



$$M_1 = \left[\begin{array}{c|ccccc} \partial_1 & ab & bc & cd & ad & ac \\ \hline a & -1 & 0 & 0 & -1 & -1 \\ b & 1 & -1 & 0 & 0 & 0 \\ c & 0 & 1 & -1 & 0 & 1 \\ d & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

ELEMENTARY OPERATIONS

- The elementary row operations on M_k are
  1. exchange row i and row j,
  2. multiply row i by -1,
  3. replace row i by (row i) $+ q$(row j), where q is in the coefficient ring and $j \neq i$.
- Similar elementary column operations on columns
- Effect: change of basis, but no change in rank
  - Column operation (3): replaces basis element e_i with $e_i + qe_j$
  - Row operation (3): replaces basis element $\hat{e}_j$ with $\hat{e}_j - q\hat{e}_i$.

REDUCTION ALGORITHM

- Like Gaussian elimination, we keep changing the basis to get to the (Smith) normal form:

$$\tilde{M}_k = \left[\begin{array}{ccc|c} b_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & b_{l_k} & \\ \hline & 0 & & 0 \end{array}\right]$$

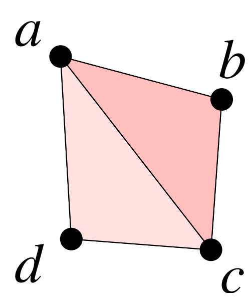
- $l_k = \operatorname{rank} M_k = \operatorname{rank} \tilde{M}_k$, $b^i \geq 1$
- $b_i | b_{i+1}$ for all $1 \leq i < l_k$

Normal Form

$$\tilde{M}_k = \left[\begin{array}{c|ccc|ccc} \partial_k & e_1 & \cdots & e_{l_k} & e_{l_k+1} & \ldots & e_{m_k} \\ \hline \hat{e}_1 & b_1 & & 0 & & & \\ \vdots & & \ddots & & & 0 & \\ \hat{e}_{l_k} & 0 & & b_{l_k} & & & \\ \hline \hat{e}_{l_k+1} & & & & & & \\ \vdots & & 0 & & & 0 & \\ \hat{e}_{m_{k-1}} & & & & & & \end{array}\right]$$

1. the torsion coefficients of H_{k-1} are $b_i \geq 1$.
2. $\{e_i \mid l_k + 1 \leq i \leq m_k\}$ is a basis for Z_k. $\Rightarrow \operatorname{rank} \mathsf{Z}_k = m_k - l_k$.
3. $\{b_i \hat{e}_i \mid 1 \leq i \leq l_k\}$ is a basis for B_{k-1}. $\Rightarrow \operatorname{rank} \mathsf{B}_k = l_{k+1}$.
4. $\beta_k = \operatorname{rank} \mathsf{Z}_k - \operatorname{rank} \mathsf{B}_k = m_k - l_k - l_{k+1}$

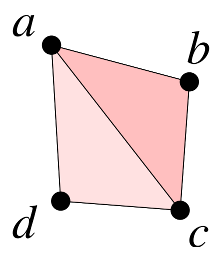
REDUCED EXAMPLE



$$\tilde{M}_1 = \left[\begin{array}{c|ccccc} & cd & bc & ab & z_1 & z_2 \\ \hline d-c & 1 & 0 & 0 & 0 & 0 \\ c-b & 0 & 1 & 0 & 0 & 0 \\ b-a & 0 & 0 & 1 & 0 & 0 \\ a & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

- $z_1 = ad - bc - cd - ab$, $z_2 = ac - bc - ab$, and rank $\mathsf{Z}_1 = 2$

Reduced Example



$$\tilde{M}_2 = \left[\begin{array}{c|cc} & -abc & -acd+abc \\ \hline ac-bc-ab & 1 & 0 \\ ad-cd-bc-ab & 0 & 1 \\ cd & 0 & 0 \\ bc & 0 & 0 \\ ab & 0 & 0 \end{array}\right]$$

- $\operatorname{rank} \mathsf{B}_1 = 2$, so $\beta_1 = \operatorname{rank} \mathsf{H}_1 = 2 - 2 = 0$

FILTRATIONS

- A subcomplex of K is a simplicial complex $L \subseteq K$.
- A filtration of a complex K is a nested sequence of complexes $\emptyset = K^0 \subseteq K^1 \subseteq \ldots \subseteq K^m = K$.
- K is a filtered complex.
- Natural:
  - Cěch-like complexes
  - Density measure
  - Manifold equipped with Morse function
  - Demo
- Looking for features in this growing space

PERSISTENCE

- Persistent Homology groups

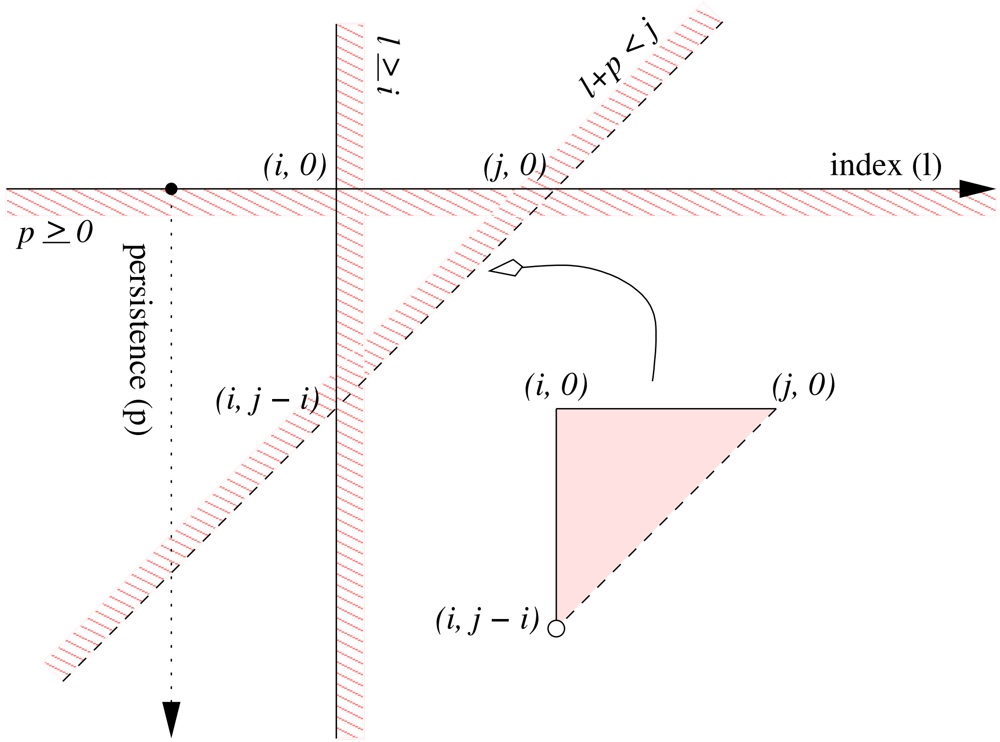
$$\mathsf{H}_k^{\ell,p} = \mathsf{Z}_k^\ell / (\mathsf{B}_k^{\ell+p} \cap \mathsf{Z}_k^\ell)$$

- $\beta_k^{\ell,p} = \mathrm{rank}\, \mathsf{H}_k^{\ell,p}$.
- The *persistence* of a cycle is its lifetime: death $-$ birth $-$ 1.
- As we increase p, cycles are killed earlier, so their persistence for p is lower.
- Pairing:
  - positive: creates k-cycle
  - negative: destroys $(k-1)$-cycle

LIFETIME REGIONS

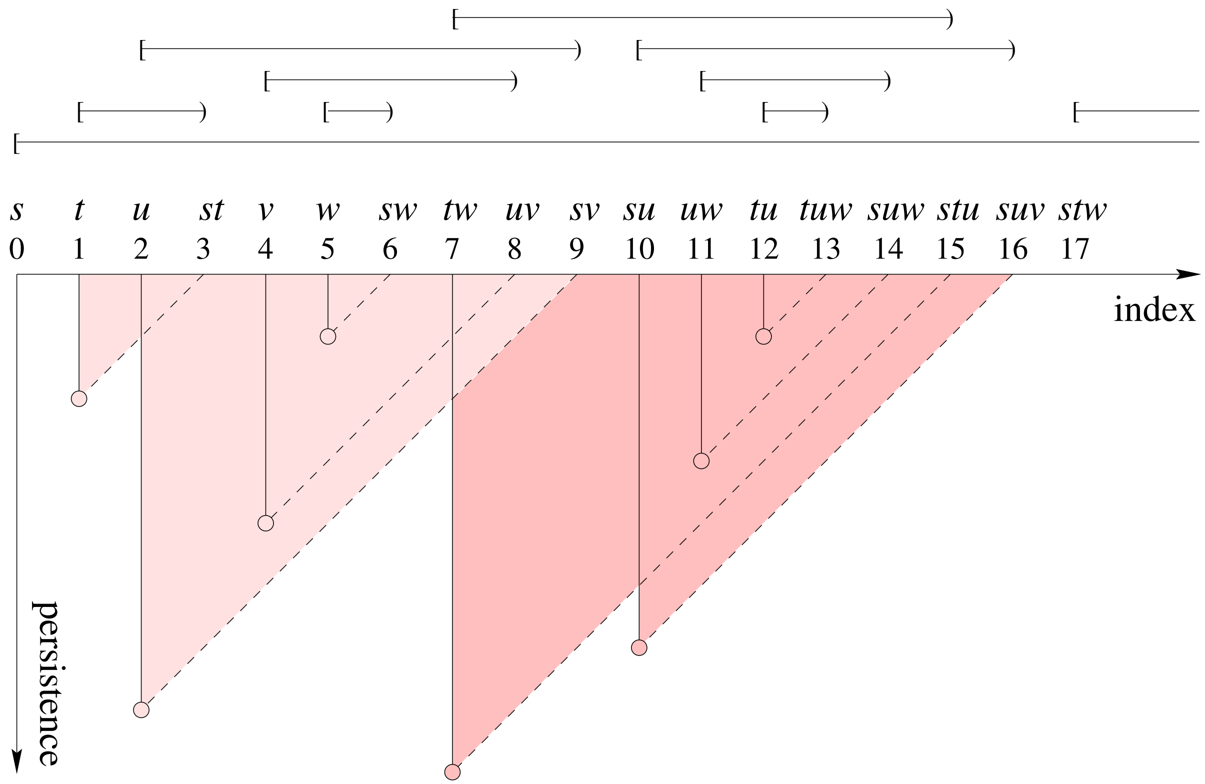
- $\mathsf{H}_k^{l,p} = \mathsf{Z}_k^l \,/\, (\mathsf{B}_k^{l+p} \cap \mathsf{Z}_k^l)$
- Basis element $z + \mathsf{B}_k^l$ lives during $l \in [i, j)$
- $z \notin \mathsf{B}_k^l$ for $l \leq j$
- Therefore, $z \notin \mathsf{B}_k^{l+p}$ for $l + p < j$.
- $p \geq 0$
- $l \geq i$

TRIANGLE



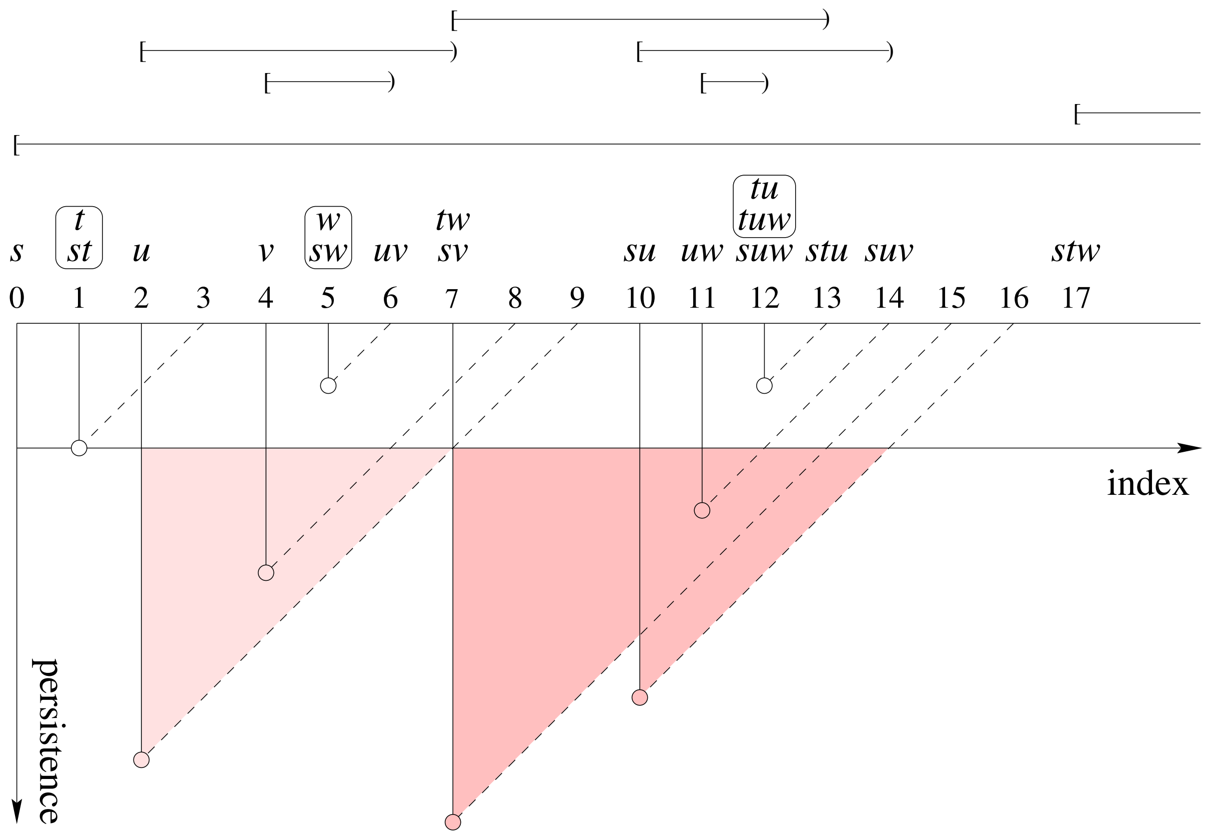
- $p \geq 0$
- $l \geq i$
- $l < j$

VISUALIZATION



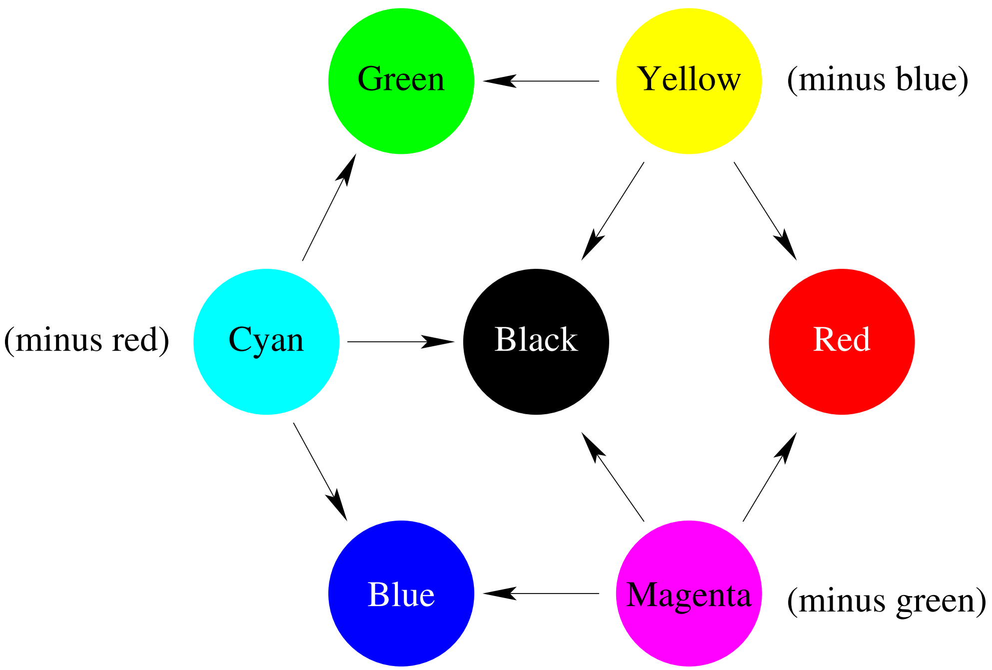
- Original algroithm for $\mathbb{Z}_2$-homology, subcomplexes of $\mathbb{S}^3$

REORDERING

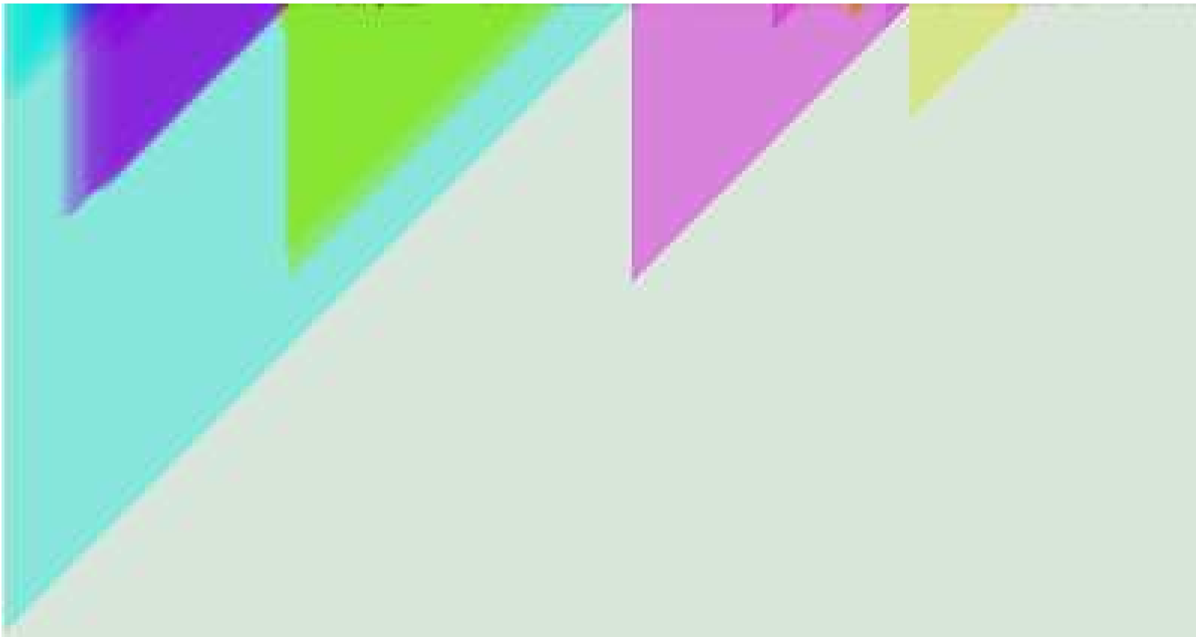


t
st
w
sw
tw
sv
tu
tuw
s
u
v
uv
su
uw
suw
stu
suv
stw
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
index
persistence

CMY Color Space



TOPOLOGY MAPS



- Demo

PERSISTENCE COMPLEX

- $\mathsf{C}_*^0 \xrightarrow{f^0} \mathsf{C}_*^1 \xrightarrow{f^1} \mathsf{C}_*^2 \xrightarrow{f^2} \cdots$.
- Expanding:

$$
\begin{array}{ccccccc}
\partial_3 \downarrow & & \partial_3 \downarrow & & \partial_3 \downarrow & & \\
\mathsf{C}_2^0 & \xrightarrow{f^0} & \mathsf{C}_2^1 & \xrightarrow{f^1} & \mathsf{C}_2^2 & \xrightarrow{f^2} & \cdots \\
\partial_2 \downarrow & & \partial_2 \downarrow & & \partial_2 \downarrow & & \\
\mathsf{C}_1^0 & \xrightarrow{f^0} & \mathsf{C}_1^1 & \xrightarrow{f^1} & \mathsf{C}_1^2 & \xrightarrow{f^2} & \cdots \\
\partial_1 \downarrow & & \partial_1 \downarrow & & \partial_1 \downarrow & & \\
\mathsf{C}_0^0 & \xrightarrow{f^0} & \mathsf{C}_0^1 & \xrightarrow{f^1} & \mathsf{C}_0^2 & \xrightarrow{f^2} & \cdots
\end{array}
$$

ARTIN-REES CONSTRUCTION

- $\mathcal{M} = \{M^i, \varphi^i\}_{i \geq 0}$ defined over R
- Define a graded $R[t]$-module over by

$$\alpha(\mathcal{M}) = \bigoplus_{i=0}^{\infty} M^i,$$

- R-module structure is the the sum on the individual components
- Action of t is

$$t \cdot (m^0, m^1, m^2, \ldots) = (0, \varphi^0(m^0), \varphi^1(m^1), \varphi^2(m^2), \ldots).$$

- t simply shifts elements of the module up in the gradation.

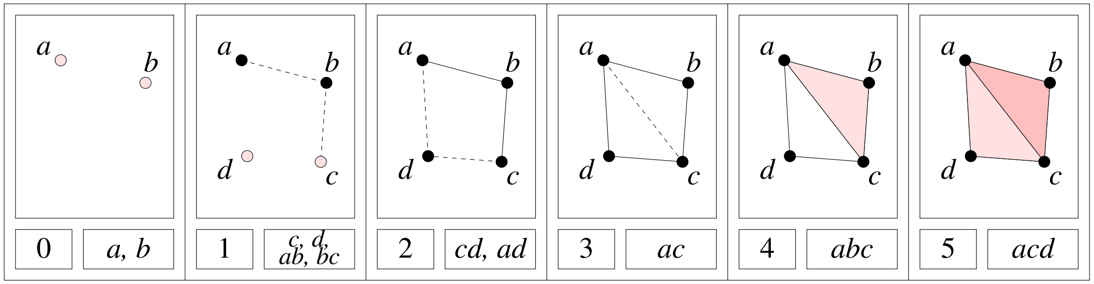
STRUCTURE

- Equivalent categories
- R, a field
- $R[t]$ is a PID
- Structure Theorem for graded $R[t]$-modules:

$$\left(\bigoplus_{i=1}^{n} \Sigma^{\alpha_i} F[t] \right) \oplus \left(\bigoplus_{j=1}^{m} \Sigma^{\gamma_j} F[t]/(t^{n_j}) \right).$$

- Intervals:
  - $\Sigma^{\alpha_i} F[t] \mapsto (\alpha^i, \infty)$
  - $\Sigma^{\gamma_j} F[t]/(t^{n_j}) \mapsto (\gamma_j, \gamma_j + n_j)$

Computing Persistence



- Degrees of homogeneous elements:

a	b	c	d	ab	bc	cd	ad	ac	abc	acd
0	0	1	1	1	1	2	2	3	4	5

$$M_1 = \left[\begin{array}{c|ccccc} \partial_1 & ab & bc & cd & ad & ac \\ \hline d & 0 & 0 & t & t & 0 \\ c & 0 & 1 & t & 0 & t^2 \\ b & t & t & 0 & 0 & 0 \\ a & t & 0 & 0 & t^2 & t^3 \end{array}\right]$$

- $\deg \hat{e}_i + \deg M_k(i,j) = \deg e_j$

Column Echelon Form

$$\tilde{M}_1 = \left[\begin{array}{c|ccccc} & cd & bc & ab & z_1 & z_2 \\ \hline d & \boxed{t} & 0 & 0 & 0 & 0 \\ c & t & \boxed{1} & 0 & 0 & 0 \\ b & 0 & t & \boxed{t} & 0 & 0 \\ a & 0 & 0 & t & 0 & 0 \end{array}\right]$$

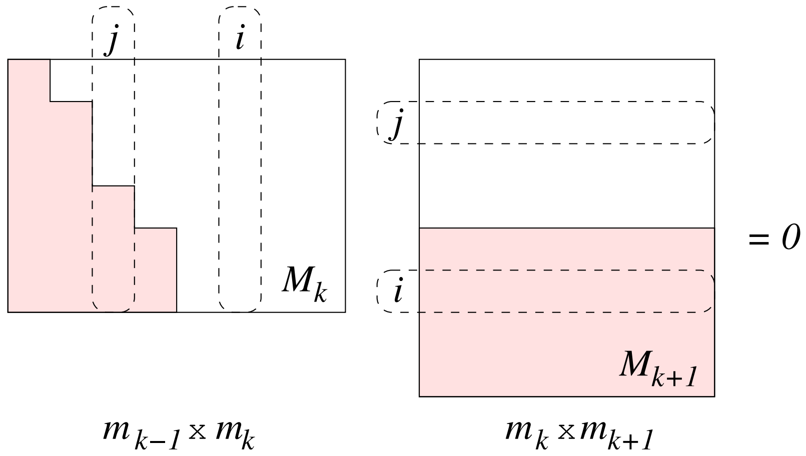
- Only column operations of type (1, 3)
- $z_1 = ad - cd - t \cdot bc - t \cdot ab$
- $z_2 = ac - t^2 \cdot bc - t^2 \cdot ab$
- $\{z_1, z_2\}$ form homogeneous basis for Z_1
- $\text{rank}\, M_k = \text{rank}\, \mathsf{B}_{k-1}$ is number of pivots

ECHELON FORM LEMMA

$$\tilde{M}_1 = \left[\begin{array}{c|ccccc} & cd & bc & ab & z_1 & z_2 \\ \hline d & \boxed{\text{t}} & 0 & 0 & 0 & 0 \\ c & t & \boxed{1} & 0 & 0 & 0 \\ b & 0 & t & \boxed{\text{t}} & 0 & 0 \\ a & 0 & 0 & t & 0 & 0 \end{array}\right]$$

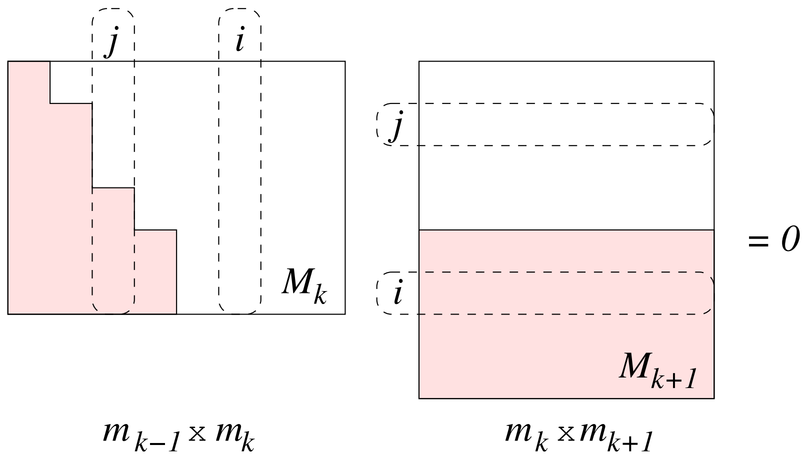
- (Lemma) The pivots in column-echelon form are the same as the diagonal elements in normal form. Moreover, the degree of the basis elements on pivot rows is the same in both forms.
- If only interested in degree of basis elements, read them off the column echelon form.

BASIS CHANGE LEMMA



- Represent ∂_{k+1} in terms of the basis computed for Z_k
- $\partial_k \partial_{k+1} = \emptyset$, $M_k M_{k+1} = 0$
- (Lemma) To represent ∂_{k+1} relative to the standard basis for C_{k+1} and the basis computed for Z_k, simply delete rows in M_{k+1} that correspond to pivot columns in $\tilde{M}_k$.

PROOF



- Replace column i by (column i) $+ q$(column j) to eliminate element in pivot row j
- $\equiv$ replacing column basis element e_i by $e_i + qe_j$ in M_k
- $\equiv$ replacing row j with (row j) $- q$(row i) in M_{k+1}
- But row j is eventually zero and row i is not changed. QED

ALGORITHM

- No need for row operations
- Free columns correspond to positive simplices
- Pivot columns correspond to negative simplices
- No need for matrix representation
- Sparse matrix computation of Betti numbers based on persistence

CONCLUSION

- Over fields, persistent homology of a filtration has a compact description
- The algorithm is $O(n^3)$, but fast in practice
- Papers available off graphics.stanford.edu/˜afra
  - Topological Persistence and Simplification – FOCS ’00, DCG ’02
  - Computing Persistent Homology – SoCG ’04
- `C` code available
- CGAL package soon