Persistent homology of spaces and groups

Graham Ellis

Outline

1. Introduction to persistence bar codes
   (three motivating examples)

2. Computation of persistence bar codes
   (discrete vector fields & contracting homotopies)

3. Potential application to group cohomology

Mathematics Algorithms Proofs, 8-12 November 2010

Motivating Example I: statistical data analysis

Given a set S of points randomly sampled from an unknown manifold M, what can we infer about the topology of M?

Motivating Example I: statistical data analysis

Given a set S of points randomly sampled from an unknown manifold M, what can we infer about the topology of M?

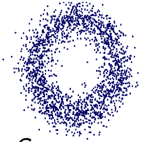
For instance, S sampled from $M \subset \mathbb{R}^2$.

One approach to data analysis

Repeatedly "thicken" the set S to produce a sequence of inclusions

$$S = S_1 \subset S_2 \subset S_3 \subset \cdots .$$

Then search for "persistent" topological features in the sequence.



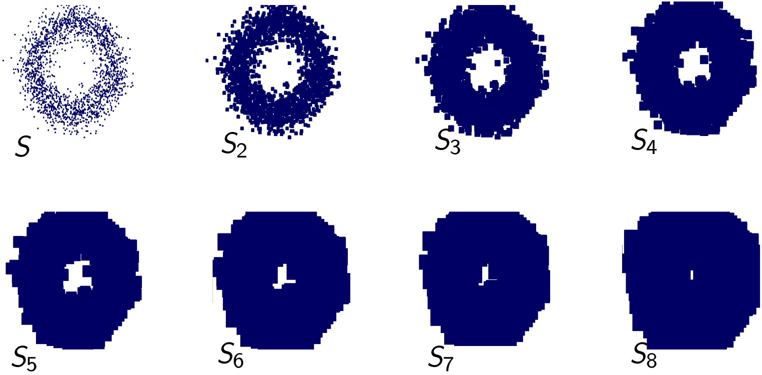
S

One approach to data analysis

Repeatedly "thicken" the set S to produce a sequence of inclusions

$$S = S_1 \subset S_2 \subset S_3 \subset \cdots .$$

Then search for "persistent" topological features in the sequence.



Betti numbers

$\beta_0(X) =$ number of path components of X

Betti numbers

$\beta_0(X) =$ number of path components of X

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
β_0	478	32	9	2	1	1	1	1

Betti numbers

$\beta_0(X)$ = number of path components of X

$\beta_1(X) = \dim(H_1(X, \mathbb{Q}))$ = number of 1-dimensional holes in X

	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
β_0	478	32	9	2	1	1	1	1
β_1	0	115	18	4	1	1	1	1

Betti numbers

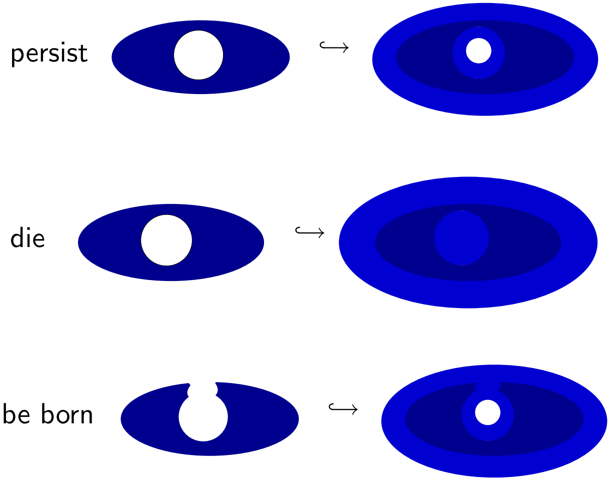
$\beta_0(X)$ = number of path components of X

$\beta_1(X) = \dim(H_1(X, \mathbb{Q}))$ = number of 1-dimensional holes in X

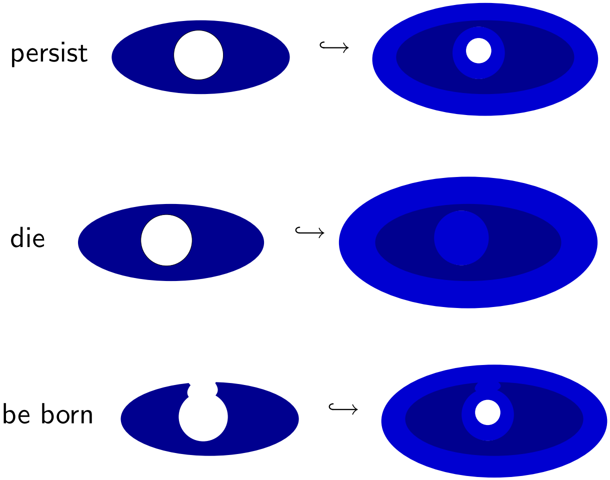
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
β_0	478	32	9	2	1	1	1	1
β_1	0	115	18	4	1	1	1	1

These numbers are consistent with the sample coming from some region with the homotopy type of a circle.

During an inclusion $S_i \hookrightarrow S_j$ holes can

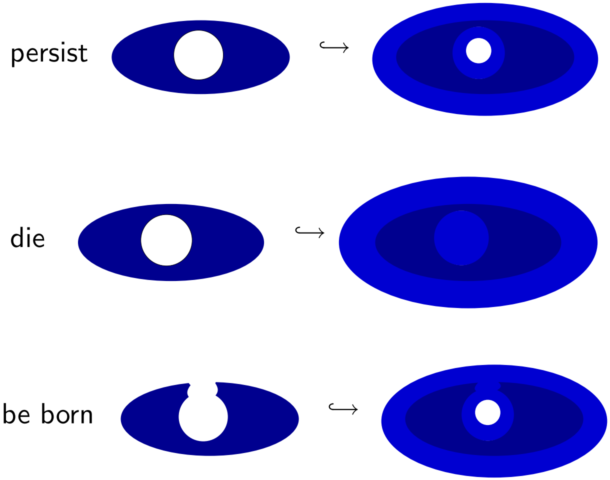


During an inclusion $S_i \hookrightarrow S_j$ holes can



β_n^{ij} = number of n-dimensional holes in S_i that persist to S_j

During an inclusion $S_i \hookrightarrow S_j$ holes can



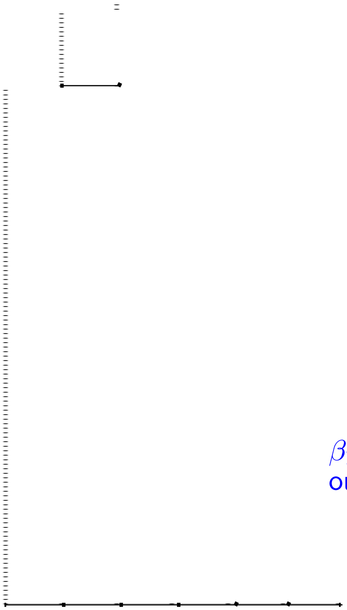
$$\beta_n^{ij} = \text{number of } n\text{-dimensional holes in } S_i \text{ that persist to } S_j$$

$$= \text{rank}\ (H_n(S_i, \mathbb{Q}) \longrightarrow H_n(S_j, \mathbb{Q}))$$

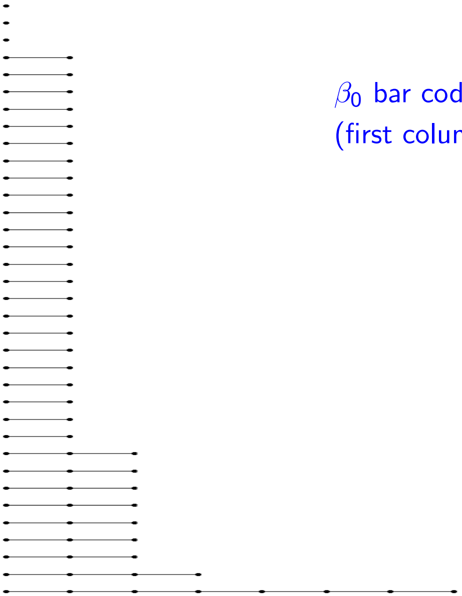
Bar codes

Matrix (β_n^{ij}) represented by graph with:

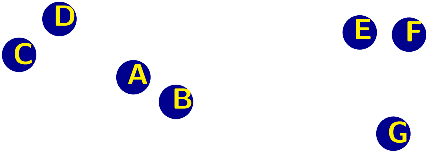
- vertices arranged in columns
- only horizontal edges
- ith column has $\beta_n^{ii} = \beta_n(S_i)$ vertices
- β_n^{ij} paths from ith column to jth column



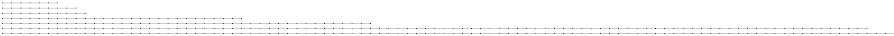
β_1 bar code for our example



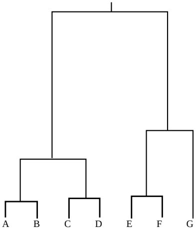
β_0 bar code for our example (first column cropped)



β_0 bar codes



contain less information than dendrograms (or phylogenetic trees)



How to thicken data?

Low-dimensional data in $\mathbb{R}^n$:
(digital images, dynamical systems, ...)

Construct a filtered cubical subcomplex of $\mathbb{R}^n$.

High-dimensional data:
(statistical data sample of N points, ...)

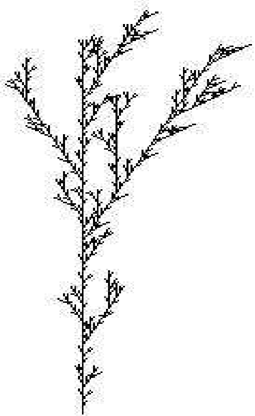
Construct a filtration on the simplex Δ^N.

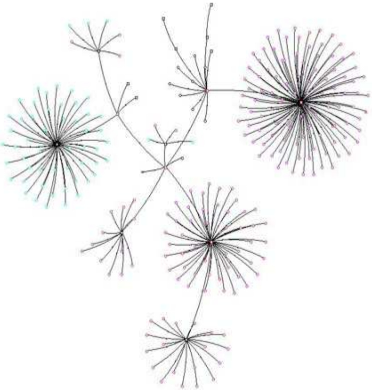
Group-theoretic data:

Construct a filtration on a (regular CW) Eilenberg-MacLane space.

Motivating Example II: polymer growth

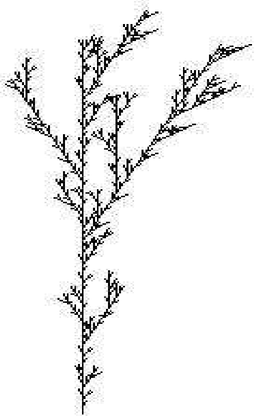
How do the shapes of the following planar graphs differ?

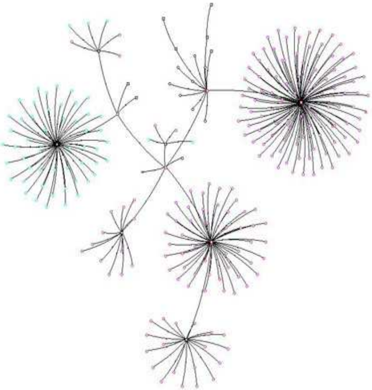




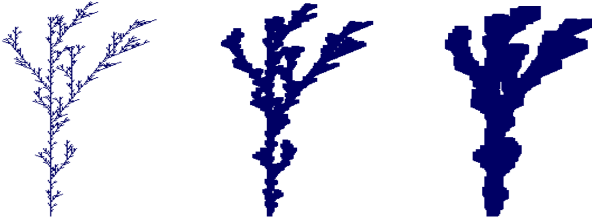
Motivating Example II: polymer growth

How do the shapes of the following planar graphs differ?

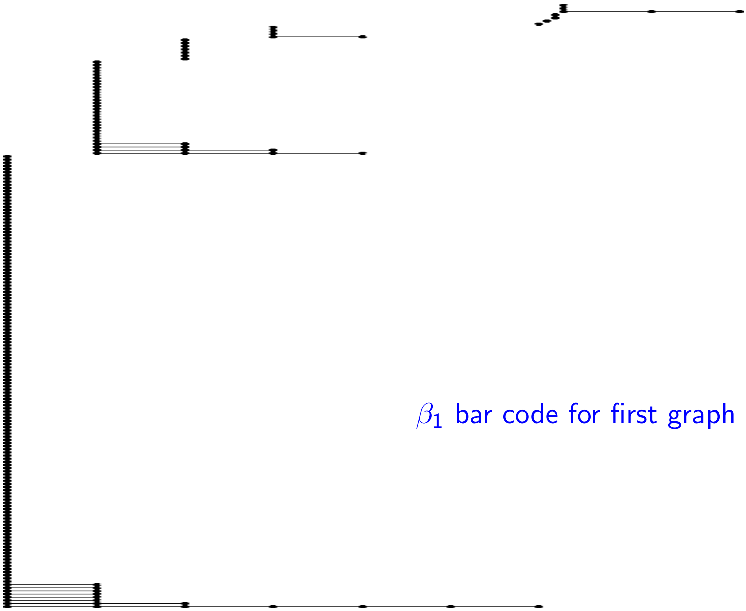




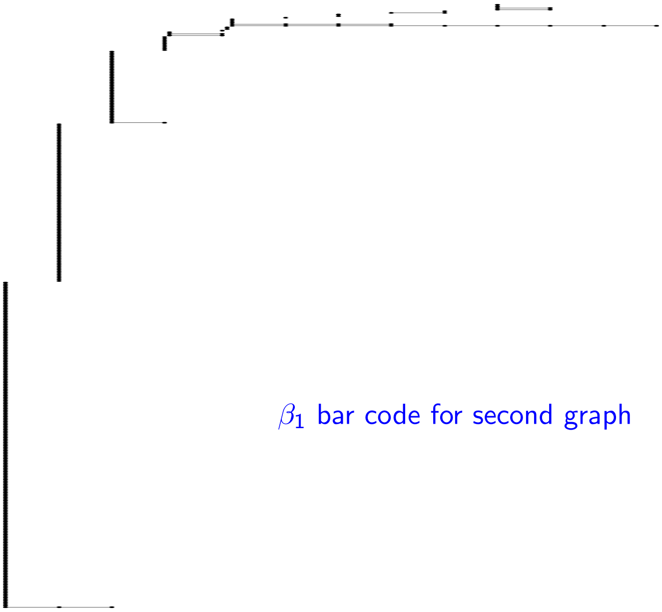
MacPherson & Srolovitz: Persistent homology can capture shape.



Various thickenings of the first graph



β_1 bar code for first graph



β_1 bar code for second graph

MacPherson & Srolovitz define the "homological dimension" of a graph in terms of:

- The number of bars in its bar code
- the length of these bars
- the centres of these bars

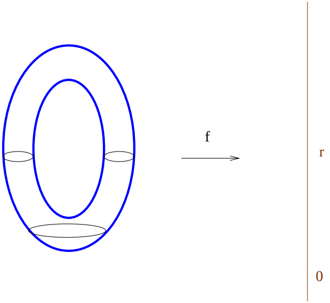
Motivating Example III: medical images

Digital images $f : \mathbb{M} \to \mathbb{R}$ could be analyzed using bar codes.

Motivating Example III: medical images

Digital images $f : \mathbb{M} \to \mathbb{R}$ could be analyzed using bar codes.

Consider a torus $\mathbb{M}$, height function f



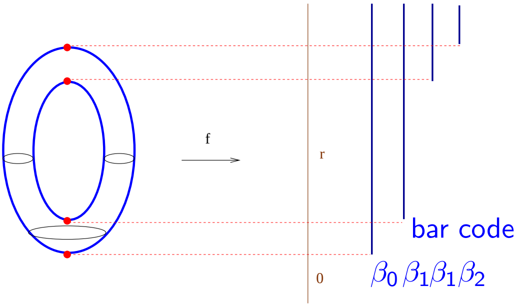
and filtration

$$\mathbb{M}_r = f^{-1}([0, r]).$$

Motivating Example III: medical images

Digital images $f : \mathbb{M} \to \mathbb{R}$ could be analyzed using bar codes.

Consider a torus $\mathbb{M}$, height function f



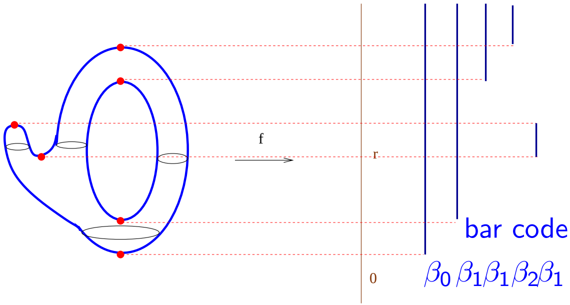
and filtration

$$\mathbb{M}_r = f^{-1}([0, r]).$$

Motivating Example III: medical images

Digital images $f : \mathbb{M} \to \mathbb{R}$ could be analyzed using bar codes.

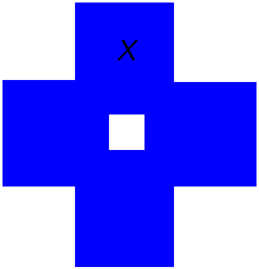
Consider a torus $\mathbb{M}$, height function f



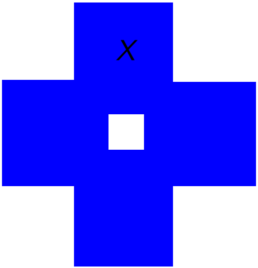
and filtration

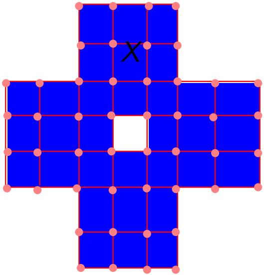
$$\mathbb{M}_r = f^{-1}([0, r]).$$

To compute the homlogy of a space



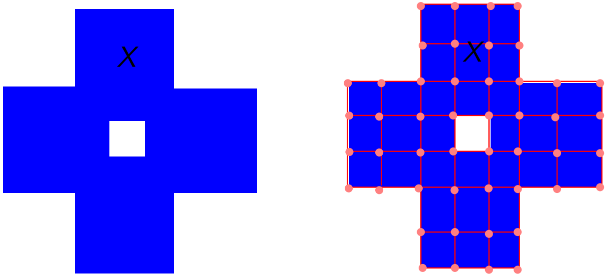
To compute the homlogy of a space





we impose some cell structure,

To compute the homlogy of a space

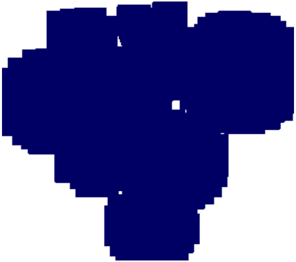


we impose some cell structure, and consider

$$\cdots \to C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \to 0$$

- $C_n(X)$ = vector space, basis $\leftrightarrow$ n-cells
- ∂_n induced by cell boundaries
- $H_n(X) = \ker(\partial_n)/\mathrm{image}(\partial_{n+1})$

Our cubical representation of the thickened planar graph $X =$



has 45467 2-cells, 91531 edges and 46060 vertices. A naive computation of

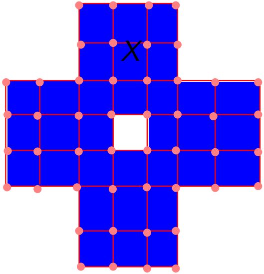
$$H_1(X, \mathbb{F}) = \mathbb{F}^5$$

is slow.

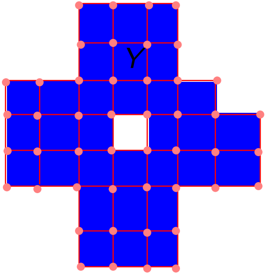
Simple homotopy collapses can yield homotopy retracts $Y \subset X$.

Simple homotopy collapses can yield homotopy retracts $Y \subset X$.

If $X = Y \cup \overline{e^n}$ and $Y \cap \overline{e^n} \simeq *$ then $X \simeq Y$.



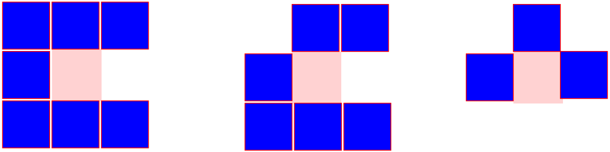
$\simeq$



For cubical subspaces of low-dimensional $\mathbb{R}^n$ the test $Y \cap \overline{e^n} \simeq *$ can be performed quickly.

For cubical subspaces of low-dimensional $\mathbb{R}^n$ the test $Y \cap \overline{e^n} \simeq *$ can be performed quickly.

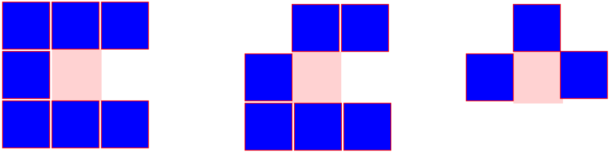
For cubcial $X \subset \mathbb{R}^2$ a cell $\overline{e^2}$ can be deleted without changing homotopy type iff its neighbourhood is one of a storable list:



etc.

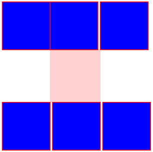
For cubical subspaces of low-dimensional $\mathbb{R}^n$ the test $Y \cap \overline{e^n} \simeq *$ can be performed quickly.

For cubcial $X \subset \mathbb{R}^2$ a cell $\overline{e^2}$ can be deleted without changing homotopy type iff its neighbourhood is one of a storable list:



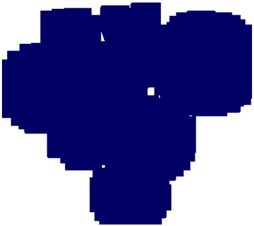
etc.

Many neighbourhoods not in list:

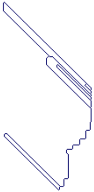


etc.

The retract

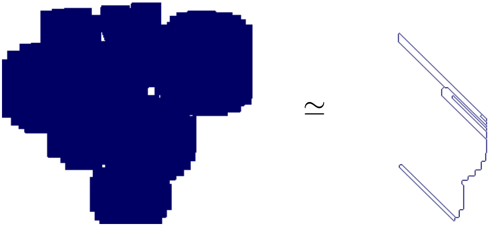


$\simeq$



has only 1717 vertices, 2342 edges and 621 faces.

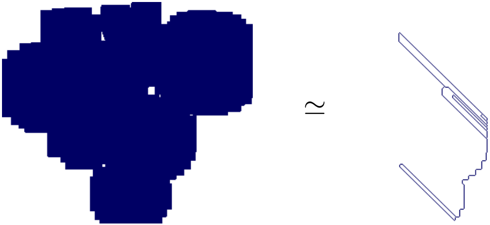
The retract



has only 1717 vertices, 2342 edges and 621 faces.

The retract Y has contractible subspace $Z \subset Y$ with 1713 vertices, 2329 edges and 617 faces.

The retract



has only 1717 vertices, 2342 edges and 621 faces.

The retract Y has contractible subspace $Z \subset Y$ with 1713 vertices, 2329 edges and 617 faces.

The computation

$$H_1(X,\mathbb{Z}) \cong H_1(C_*(Y)/C_*(Z)) = \mathbb{Z}^5$$

takes a fraction of a second.

Contracting homotopies

From a homotopy retract $Y \subset X$ we often need

- the chain inclusion $\iota_* \colon C_*(Y) \hookrightarrow C_*(X)$
- its quasi-inverse $\phi_* \colon C_*(X) \to C_*(Y)$
- and a family of homomorphisms

$$h_n \colon C_n(X) \to C_{n+1}(X) \quad (n \geq 0)$$

satisfying

$$\iota_n \phi_n - 1 = \partial_{n+1} h_n + h_{n-1} \partial_n \qquad (h_{-1} = 0).$$

Contracting homotopies

From a homotopy retract $Y \subset X$ we often need

- the chain inclusion $\iota_* \colon C_*(Y) \hookrightarrow C_*(X)$
- its quasi-inverse $\phi_* \colon C_*(X) \to C_*(Y)$
- and a family of homomorphisms

$$h_n \colon C_n(X) \to C_{n+1}(X) \quad (n \geq 0)$$

satisfying

$$\iota_n \phi_n - 1 = \partial_{n+1} h_n + h_{n-1} \partial_n \qquad (h_{-1} = 0).$$

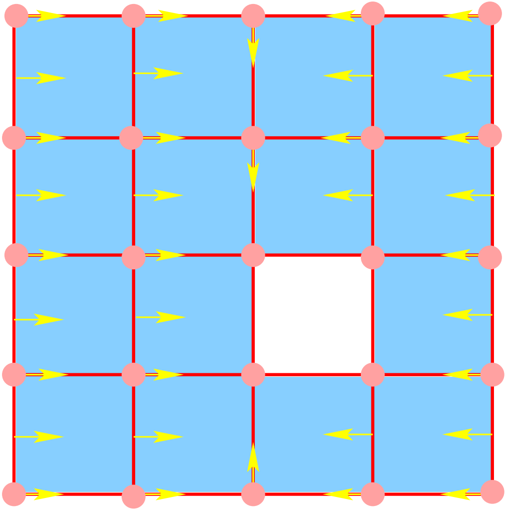
Discrete Morse Theory is handy for computing h_n, ϕ_n.

A discrete vector field on a regular CW-space X is a collection of arrows $s \longrightarrow t$ where

- s, t are cells and any cell is involved in at most one arrow
- $\dim(t) = \dim(s) + 1$
- s lies in the boundary of t

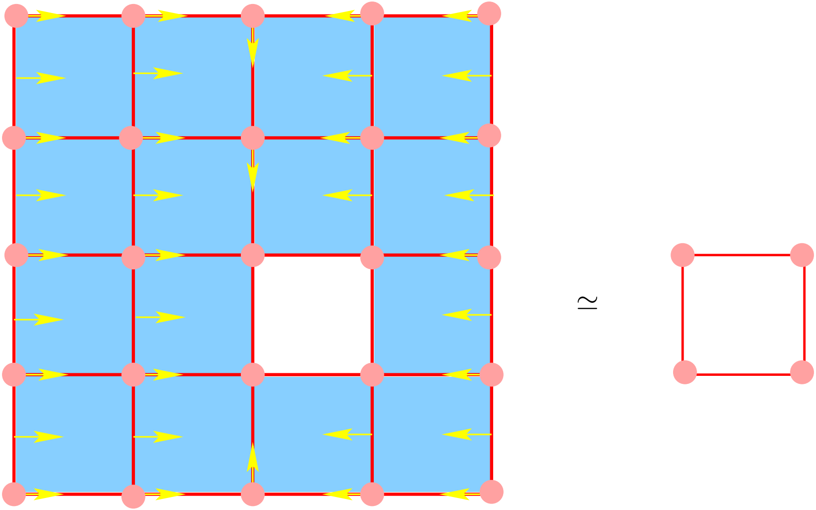
A discrete vector field on a regular CW-space X is a collection of arrows $s \longrightarrow t$ where

- s, t are cells and any cell is involved in at most one arrow
- $\dim(t) = \dim(s) + 1$
- s lies in the boundary of t

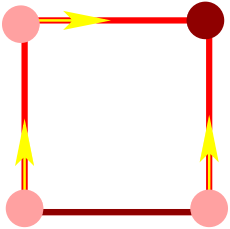


A discrete vector field on a cellular space X is a collection of arrows $s \longrightarrow t$ where

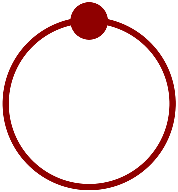
- s, t are cells and any cell is involved in at most one arrow
- $\dim(t) = \dim(s) + 1$
- s lies in the boundary of t



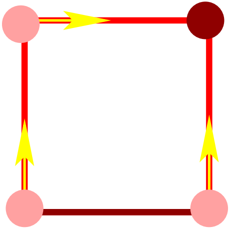
Continued example



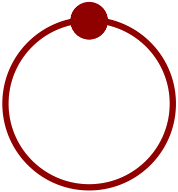
$\simeq$



Continued example



$$\simeq$$



Theorem:

If X is a regular CW-space with discrete vector field then there is a homotopy equivalence

$$X \simeq Y$$

where Y is a CW-space whose cells correspond to those of X not involved in any arrow.

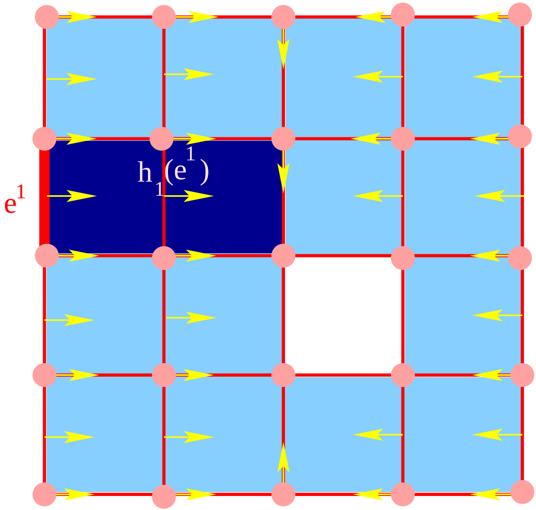
Contracting homotopy

Given a discrete vector field we define the contracting homotopy

$$h_n\colon C_n(X) \longrightarrow C_{n+1}(X)$$

on generators e^n by

$$h_n(e^n) = \begin{cases} 0 & \text{if } e^n \text{ is not a source} \\ \sum e_i^{n+1} & \partial_{n+1}(\sum e_i^{n+1}) \text{ contains just the one} \\ & \text{source } e^n \text{ of dimension } n \end{cases}$$



e^1
$h_1(e^1)$

Group (co)homology

Definition: The (co)homology of a group G is the (co)homology of X/G where X is any contractible space admitting a free G-action.

Group (co)homology

Definition: The (co)homology of a group G is the (co)homology of X/G where X is any contractible space admitting a free G-action.

Theorem: A CW-space X is contractible if $\pi_n(X^{n+1}) = 0$ for $n \geq 0$.

Group (co)homology

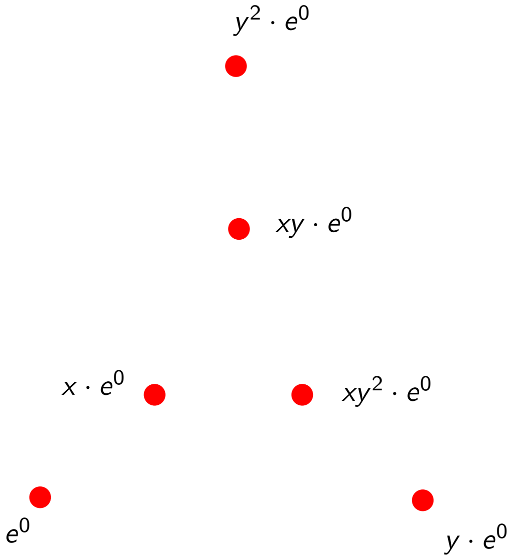
Definition: The (co)homology of a group G is the (co)homology of X/G where X is any contractible space admitting a free G-action.

Theorem: A CW-space X is contractible if $\pi_n(X^{n+1}) = 0$ for $n \geq 0$.

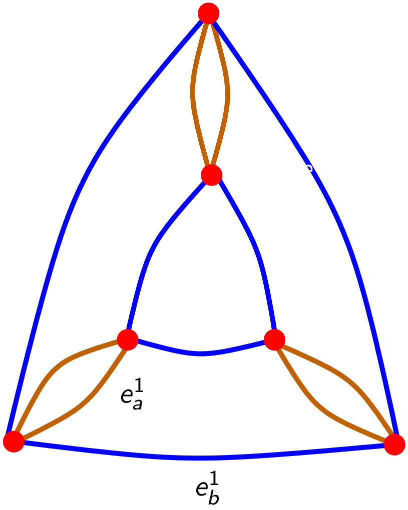
Let's illustrate for $G = S_3$.

$$x = (1,2), y = (1,2,3)$$
$$G = \langle x, y \rangle$$

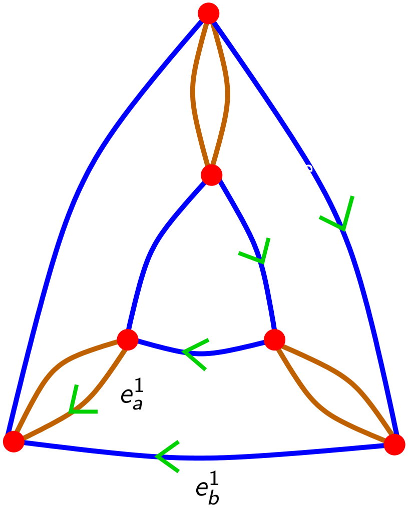
X^0 = one free orbit of vertices



$X^1 = X^0 \cup$ enough free orbits of edges to ensure $\pi_0(X^1) = 0$



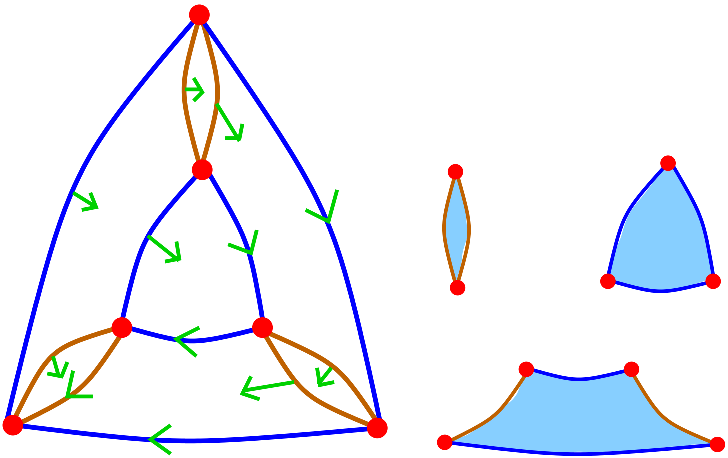
Discrete vector field on X^1 ensures $\pi_0(X^1) = 0$.



$X^2 = X^1 \cup$ enough free orbits of 2-cells to ensure $\pi_1(X^2) = 0$

$X^2 = X^1 \cup$ enough free orbits of 2-cells to ensure $\pi_1(X^2) = 0$

Discrete vector field on X^2

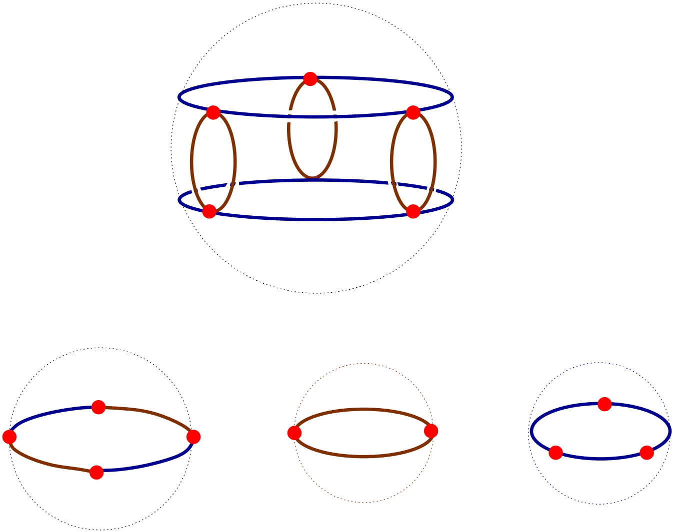


ensures that three orbits suffice.

$X^3 = X^2 \cup$ enough free orbits of 3-cells to ensure $\pi_2(X^3) = 0$

$X^3 = X^2 \cup$ enough free orbits of 3-cells to ensure $\pi_2(X^3) = 0$

Discrete vector field on X^3 ensures that four orbits suffice.



Algorithm produces a small regular CW-space X with free G-action and homotopy retraction $X \simeq *$.

Algorithm produces a small regular CW-space X with free G-action and homotopy retraction $X \simeq *$.

$$C_*(X): \quad \cdots \to C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \to 0$$

is a complex of free $\mathbb{Z}G$-modules with contracting homotopy

$$h_n\colon C_n(X) \to C_{n+1}(X) \quad (n \geq 0).$$

Algorithm produces a small regular CW-space X with free G-action and homotopy retraction $X \simeq *$.

$$C_*(X): \quad \cdots \to C_2(X) \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \to 0$$

is a complex of free $\mathbb{Z}G$-modules with contracting homotopy

$$h_n\colon C_n(X) \to C_{n+1}(X) \qquad (n \geq 0).$$

```
gap> R:=ResolutionFiniteGroup(SymmetricGroup(4),16);;

gap> C:=TensorWithIntegers(R);;

gap> Homology(C,15);
[ 2, 2, 2, 2, 2, 12 ]
```

An element of choice in homological algebra

Let X' be contractible. Often need to choose a homomorphism f_{n+1} so that the following diagram commutes.

$$
\begin{array}{ccc}
C_{n+1}(X) & \overset{f_{n+1}}{\dashrightarrow} & C_{n+1}(X') \\
\Big\downarrow{\scriptstyle \partial_{n+1}} & & \Big\downarrow \\
C_n(X) & \xrightarrow{\;f_n\;} & C_n(X')
\end{array}
$$

An element of choice in homological algebra

Let X' be contractible. Often need to choose a homomorphism f_{n+1} so that the following diagram commutes.

$$\begin{array}{ccc} C_{n+1}(X) & \overset{f_{n+1}}{\dashrightarrow} & C_{n+1}(X') \\ \downarrow{\scriptstyle \partial_{n+1}} & & \downarrow \\ C_n(X) & \xrightarrow{f_n} & C_n(X') \end{array}$$

Choice is algorithmic if some contracting homotopy $h_n \colon C_n(X') \to C_{n+1}(X')$ has already been specified for X'.

$$f_{n+1}(x) = h_n(\ f_n(\ \partial_{n+1}(x)\ )\ )$$

Homology of bigger groups?

Homology of bigger groups?

Theorem (Dutour, E, Shürmann)

$$H_n(PSL_4(\mathbb{Z}), \mathbb{Z}) = \begin{cases} 0 & n = 1 \\ (\mathbb{Z}_2)^3 & n = 2 \\ \mathbb{Z} \oplus (\mathbb{Z}_4)^2 \oplus (\mathbb{Z}_3)^2 \oplus \mathbb{Z}_5, & n = 3 \\ (\mathbb{Z}_2)^4 \oplus \mathbb{Z}_5, & n = 4 \\ (\mathbb{Z}_2)^{13}, & n = 5. \end{cases}$$

Homology of bigger groups?

Theorem (Dutour, E, Shürmann)

$$H_n(PSL_4(\mathbb{Z}),\mathbb{Z}) = \begin{cases} 0 & n=1 \\ (\mathbb{Z}_2)^3 & n=2 \\ \mathbb{Z}\oplus(\mathbb{Z}_4)^2\oplus(\mathbb{Z}_3)^2\oplus\mathbb{Z}_5, & n=3 \\ (\mathbb{Z}_2)^4\oplus\mathbb{Z}_5, & n=4 \\ (\mathbb{Z}_2)^{13}, & n=5. \end{cases}$$

Proof

$PSL_4(\mathbb{Z})$ acts on the contractible space $\mathbb{H}^4$ of positive definite symmetric 4×4 matrices. The action has finite stabilizers.

Homology of bigger groups?

Theorem (Dutour, E, Shürmann)

$$H_n(PSL_4(\mathbb{Z}),\mathbb{Z}) = \begin{cases} 0 & n=1 \\ (\mathbb{Z}_2)^3 & n=2 \\ \mathbb{Z}\oplus(\mathbb{Z}_4)^2\oplus(\mathbb{Z}_3)^2\oplus\mathbb{Z}_5, & n=3 \\ (\mathbb{Z}_2)^4\oplus\mathbb{Z}_5, & n=4 \\ (\mathbb{Z}_2)^{13}, & n=5. \end{cases}$$

Proof

$PSL_4(\mathbb{Z})$ acts on the contractible space $\mathbb{H}^4$ of positive definite symmetric 4×4 matrices. The action has finite stabilizers.

Homological Perturbation Lemma inputs:

1) $PSL_4(\mathbb{Z})$-equivariant CW-structure on $\mathbb{H}^4$.

2) CW-spaces X with free G-action and explicit contracting homotopy for each conjugacy class of cell stabilizer group G.

Homological Perturbation Lemma outputs:

Contractible CW-space with free $PSL_4(\mathbb{Z})$-action.

Potential application of bar codes : the shape of p-groups?

Group surjections

$$G \twoheadrightarrow G'$$

correspond to classifying space inclusions

$$B(G) = X/G \hookrightarrow B(G) = X'/G' \ .$$

The lower central series

$$\gamma_1 G = G, \quad \gamma_2 G = [G, G], \quad \ldots, \quad \gamma_{i+1} G = [G, \gamma_i G]$$

corresponds to a series of inclusions

$$\cdots \hookrightarrow B(\frac{G}{\gamma_4 G}) \hookrightarrow B(\frac{G}{\gamma_3 G}) \hookrightarrow B(\frac{G}{\gamma_2 G}) \hookrightarrow B(\frac{G}{\gamma_1 G}) \simeq *$$

The lower central series

$$\gamma_1 G = G, \quad \gamma_2 G = [G, G], \quad \dots, \quad \gamma_{i+1} G = [G, \gamma_i G]$$

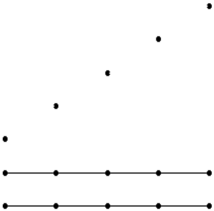
corresponds to a series of inclusions

$$\cdots \hookrightarrow B(\frac{G}{\gamma_4 G}) \hookrightarrow B(\frac{G}{\gamma_3 G}) \hookrightarrow B(\frac{G}{\gamma_2 G}) \hookrightarrow B(\frac{G}{\gamma_1 G}) \simeq *$$

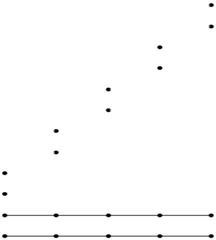
Definition:
We denote the persistent homology module of these inclusions by

$$H_n^{**}(G) = \{H_n^{ij}(G, \mathbb{F}_p)\}_{i<j}.$$

$H_1^{**}(D_{32})$

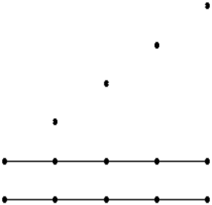


$H_2^{**}(D_{32})$

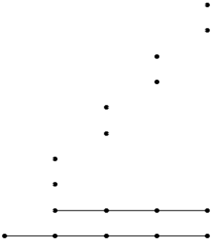


$H_3^{**}(D_{32})$

$H_1^{**}(Q_{32})$



$H_2^{**}(Q_{32})$



$H_3^{**}(Q_{32})$

Proposition : (E & King)

The invariant $H_*^{**}(G)$ partitions the 366 prime-power groups of order ≤ 81 into 227 classes with maximum class size equal to 7.

Proposition : (E & King)

The invariant $H_*^{**}(G)$ partitions the 366 prime-power groups of order ≤ 81 into 227 classes with maximum class size equal to 7.

Proposition: For a p-group G all β_2 bars start in the first column.

A group G of order p^n with $\gamma_{c+1} G = 1$, $\gamma_c G \neq 1$ is said to have coclass

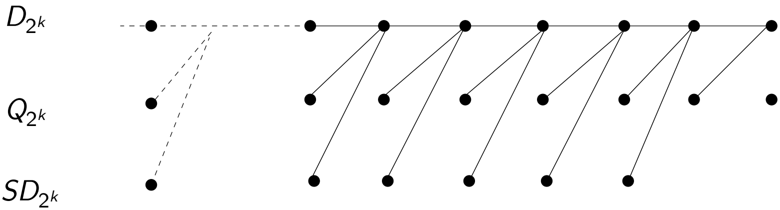
$$r = n - c \ .$$

A group G of order p^n with $\gamma_{c+1} G = 1$, $\gamma_c G \neq 1$ is said to have coclass

$$r = n - c\ .$$

The coclass graph $\mathbb{G}(p, r)$ has as vertices the p-groups of coclass r. Two vertices G, Q are connected by an edge if

$$Q \cong G/L_c(G) \quad \text{with } |L_c(G)| = p\ .$$



$\mathbb{G}(2, 1)$

Theorem (J. Carlson):

The groups $G \in \mathbb{G}(2, r)$ give rise to just finitely many non-isomorphic cohomology rings $H^*(G, \mathbb{F}_2)$.

Theorem (J. Carlson):

The groups $G \in \mathbb{G}(2, r)$ give rise to just finitely many non-isomorphic cohomology rings $H^*(G, \mathbb{F}_2)$.

Question:

Does (persistent) homology reflect the structure of coclass trees in a way that would allow us to compute the homology of large p-groups by determining their coclass tree and calculating homology of the initial period of the tree?

Theorem (J. Carlson):

The groups $G \in \mathbb{G}(2, r)$ give rise to just finitely many non-isomorphic cohomology rings $H^*(G, \mathbb{F}_2)$.

Question:

Does (persistent) homology reflect the structure of coclass trees in a way that would allow us to compute the homology of large p-groups by determining their coclass tree and calculating homology of the initial period of the tree?

Observation (E & King):

Almost all non-leaves G in a coclass tree have the same $H_2(G, \mathbb{F}_p)$.

A coclass 2 tree and its mainline β_3 bar code