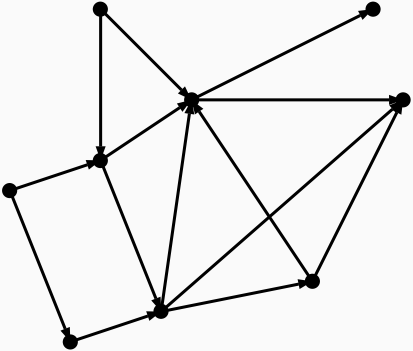
Product-simplicial complexes in a word graph

A different homology computation for directed graphs

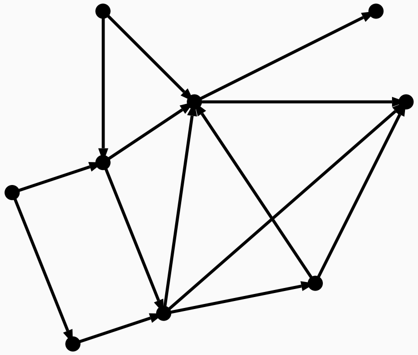
Lina Fajardo Gómez

University of South Florida

Homology on directed graphs

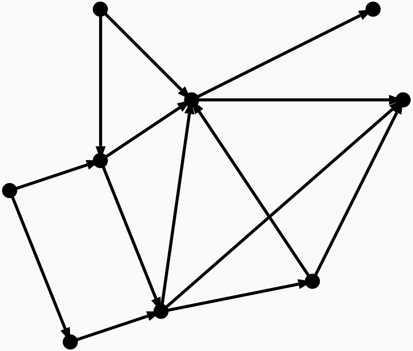


Homology on directed graphs



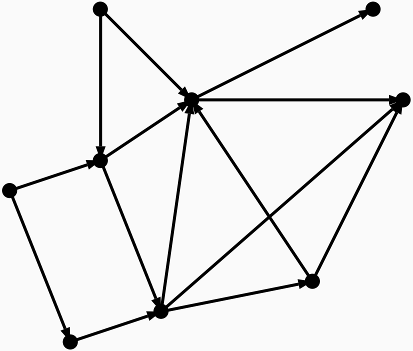
plane figures
loops

Homology on directed graphs



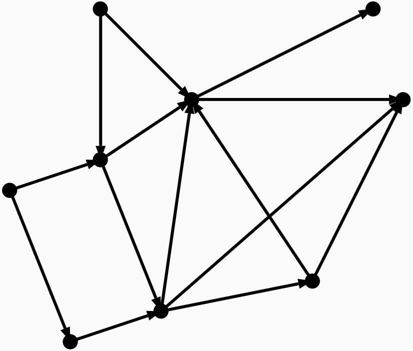
plane figures	solids
loops	cavities

Homology on directed graphs



plane figures	solids	cells
loops	cavities	holes

Homology on directed graphs

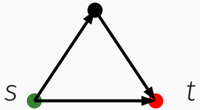


plane figures	solids	cells
loops	cavities	holes

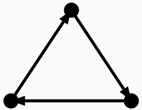
What should cells look like?

Consistently oriented graphs (CO)

Single source, single target graphs are consistently oriented:



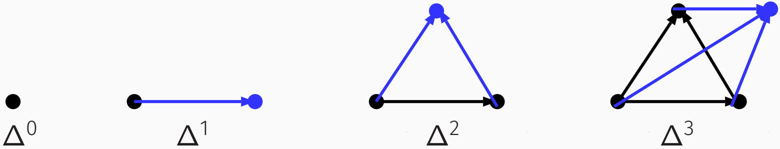
Consistently oriented



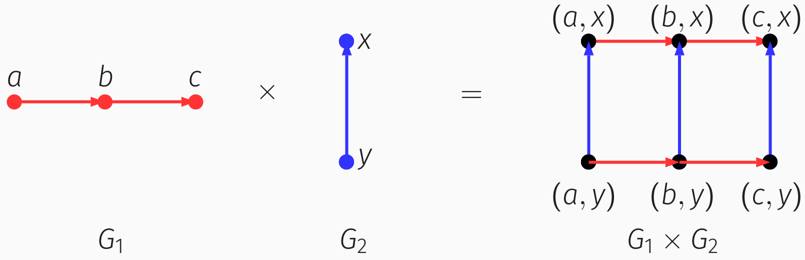
Not consistently oriented

Simplices

Graphs that are consistently oriented:



Cartesian product of graphs



$((a, x), (b, x)) \in E$, as $(a, b) \in E(G_1)$. $((a, y), (a, x)) \in E$, as $(y, x) \in E(G_2)$.

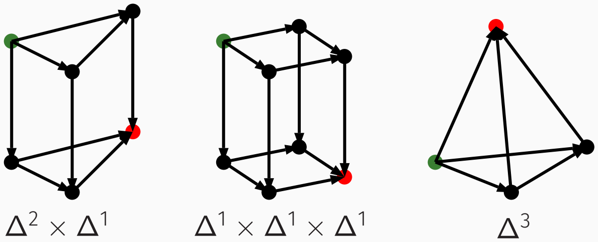
Up to relabeling of vertices, this product is **associative**.

Prodsimplicial cells

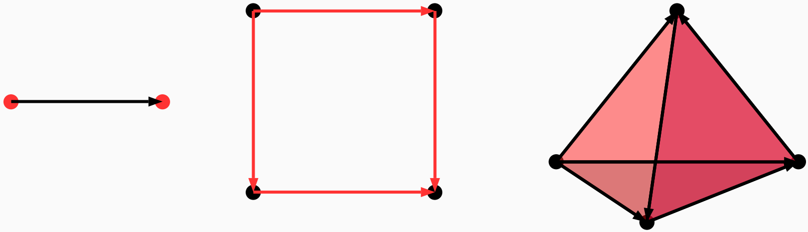
Lemma

Cartesian products of CO graphs are CO graphs.

$$G = \Delta^{n_1} \times \cdots \times \Delta^{n_k}$$



Faces and boundaries

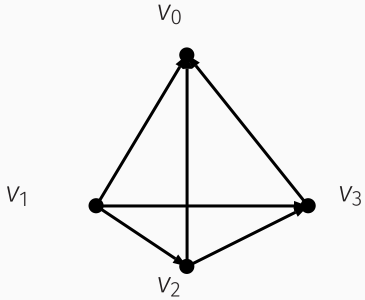


Boundary operator - simplices

$$\partial_n : \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

$[\hat{v}_i] = [v_0, v_1, \dots, v_{i-1}, \hat{v}_i, v_{i+1}, \dots, v_n]$

= the simplex generated by the vertices left after removing $\hat{v}_i$

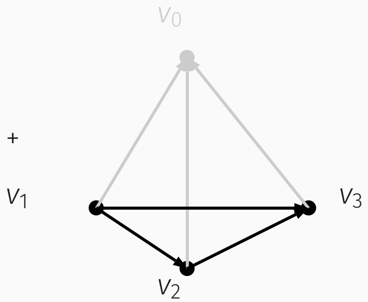


Boundary operator - simplices

$$\partial_n : \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

$[\hat{v}_i]$ = $[v_0, v_1, \ldots, v_{i-1}, \hat{v}_i, v_{i+1}, \ldots, v_n]$

= the simplex generated by the vertices left after removing $\hat{v}_i$

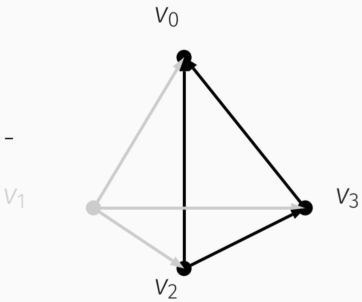


Boundary operator - simplices

$$\partial_n : \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

$[\hat{v}_i] = [v_0, v_1, \ldots, v_{i-1}, \hat{v}_i, v_{i+1}, \ldots, v_n]$

= the simplex generated by the vertices left after removing $\hat{v}_i$



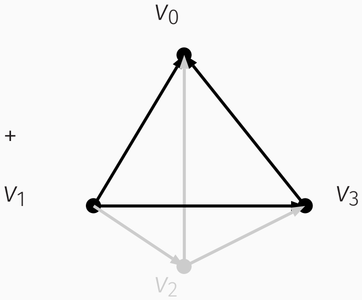
Boundary operator - simplices

$$\partial_n : \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

$[\hat{v}_i] = [v_0, v_1, \ldots, v_{i-1}, \hat{v}_i, v_{i+1}, \ldots, v_n]$

= the simplex generated by the vertices left after removing $\hat{v}_i$

+

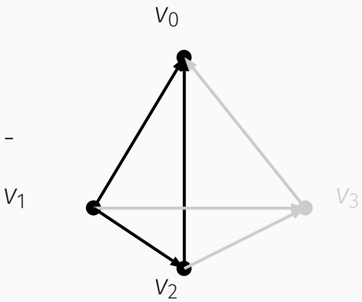


Boundary operator - simplices

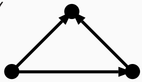
$$\partial_n : \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

$[\hat{v}_i]$ = $[v_0, v_1, \ldots, v_{i-1}, \hat{v}_i, v_{i+1}, \ldots, v_n]$

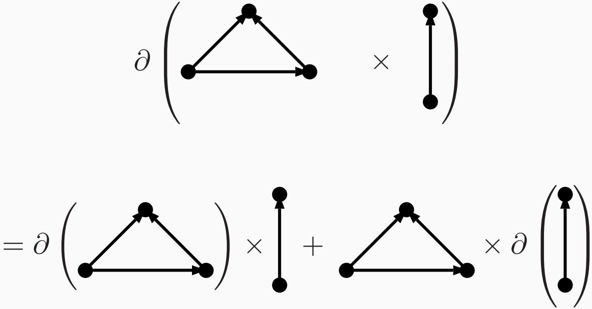
= the simplex generated by the vertices left after removing $\hat{v}_i$



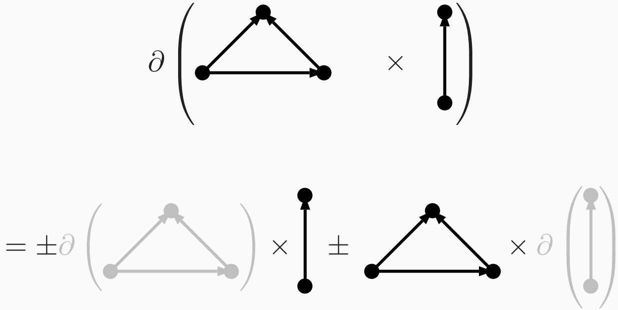
Boundary operator - cells

$\partial \Big($  $\times$  $\Big) = ?$

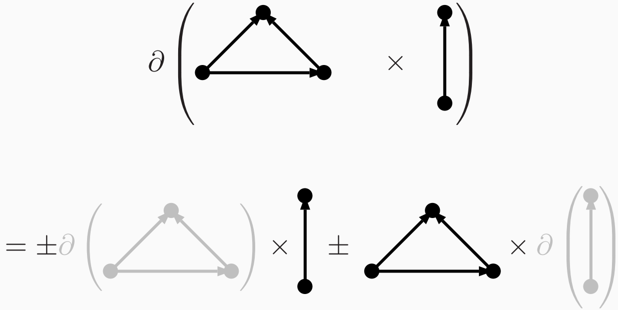
Boundary operator - cells



Boundary operator - cells



Boundary operator - cells



$$\Delta^{n_1} \times \Delta^{n_2} \times \cdots \times \Delta^{n_{i-1}} \times \partial_{n_i}(\Delta^{n_i}) \times \Delta^{n_{i+1}} \times \cdots \times \Delta^{n_k}$$

Boundary operator - cells

$$\partial_n : \; \Delta^n \mapsto \sum_{i=0}^{n} (-1)^i [\hat{v}_i]$$

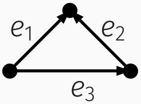
$$\partial_N : \prod_{j=1}^{k} \Delta^{n_j} \mapsto \sum_{j=1}^{k} (-1)^{\alpha(j)} [\overline{\partial_{n_j} \Delta^{n_j}}]$$

$N = \sum_i n_i$, $\alpha(j) = \sum_{l=1}^{j-1} n_l + j$

$[\overline{\partial_{n_i} \Delta^{n_i}}]$ denotes the product
$\Delta^{n_1} \times \Delta^{n_2} \times \cdots \times \Delta^{n_{i-1}} \times \partial_{n_i}(\Delta^{n_i}) \times \Delta^{n_{i+1}} \times \cdots \times \Delta^{n_k}$

Boundary operator - cells

An N-chain is a linear combination of prodsimplicial N-cells.



An example of a 2-chain is $e_1 - e_2 - e_3$.

The set of all N-chains is denoted $\mathcal{C}_N$.

∂_N induces a linear map on chains.

Boundary operator - cells

Proposition

The boundary operator ∂ applied to prodsimplicial cells, as defined, satisfies $\partial^2 = 0$.

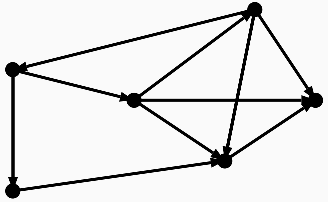
Chains with boundary zero enclose holes.

Cell complex on a graph

A prodsimplicial cell complex is a collection Γ of prodsimplicial cells such that the nonempty intersection of any two prodsimplicial cells in Γ is in the union of their face sets.

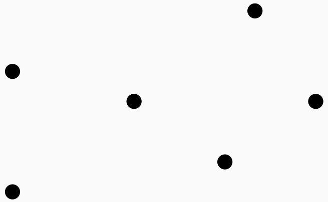
Cell complex on a graph

Given a directed graph, we associate to it a complex by identifying cells.



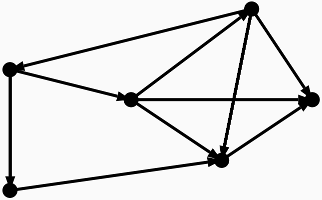
Filtrations on a graph - $n = 0$

1. vertices



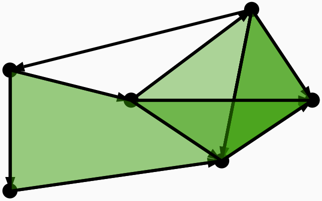
Filtrations on a graph - $n = 1$

2. edges



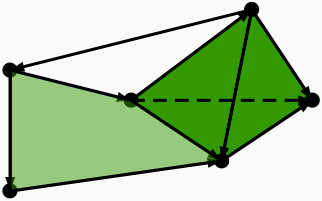
Filtrations on a graph - $n = 2$

3. 2-dimensional faces



Filtrations on a graph - $n = 3$

4. 3-dimensional faces



Homology groups

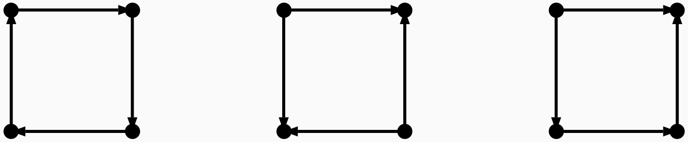
Recall, ∂_N can be seen as a linear map.

$$H_N = \frac{\ker \partial_N}{\operatorname{im} \partial_{N+1}}$$

$$H_N \cong \underbrace{\mathbb{Z} \oplus \mathbb{Z} \cdots \oplus \mathbb{Z}}_{n=\operatorname{rank}(H_N)} + \underbrace{\mathbb{Z}_{{p_1}^{s_1}} \oplus \mathbb{Z}_{{p_2}^{s_2}} \cdots \oplus \mathbb{Z}_{{p_m}^{s_m}}}_{\text{torsion group}}$$

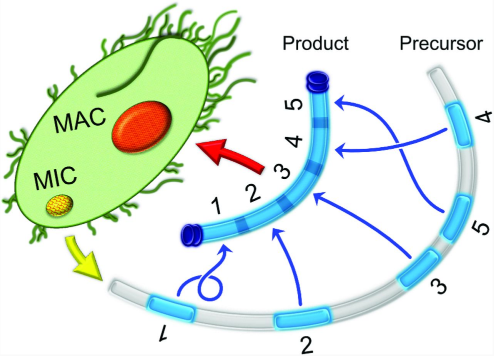
The Nth homology group (rank) counts N dimensional holes:

- non-CO
- non-product



DNA rearrangement

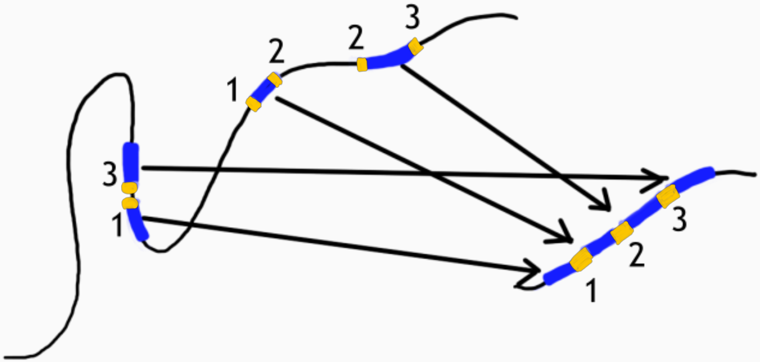
Oxytricha trifallax has somatic nuclei (MAC) and germline nuclei (MIC).



J. Burns, D. Kukushkin, K. Lindblad, X. Chen, N. Jonoska, L. Landweber. A Database of Ciliate Genome Rearrangements. Nucleic Acids Res. 44:D1 (2015) pp. D703-D709.

Pointers

Pointers mark where segments will be joined.



The pointer sequence formed is 311223.

Double occurrence words

Double ocurrence words (DOWs) are sequences of symbols where each symbol appears exactly twice.

Pointer sequences form DOWs.

Ascending order

A DOW is in ascending order if its symbols are labeled by order of appearance. Two DOWs are equivalent if they have the same ascending order representation.

Example

$311223 \sim 122331$

$$\begin{array}{cccccc} 3 & 1 & 1 & 2 & 2 & 3 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 2 & 2 & 3 & 3 & 1 \end{array}$$

Repeat and return words in DOWs

uu is a repeat word in w if $w = z_1 u z_2 u z_3$.

uu^R is a return word in w if $w = z_1 u z_2 u^R z_3$. In both of these cases, we say that w reduces to $z_1 z_2 z_3$.

Repeat and return words describe over 90% of DNA rearrangement in *Oxytricha trifallax*.

J. Burns, D. Kukushkin, X. Chen, L. Landweber, M. Saito, N. Jonoska. Recurring Patterns Among Scrambled Genes in the Encrypted Genome of the Ciliate Oxytricha trifallax. Journal of Theoretical Biology 410 (2016) pp. 171-180.

DOW reduction

Reduction of repeat and return words

1234523541 has one nontrivial repeat word and one nontrivial return word:

2323 is a repeat word: 1**23**45**23**541;

4554 is a return word:123**45**23**54**1

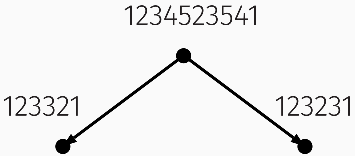
DOW reduction

Reduction of repeat and return words

1234523541 can be reduced two different ways:

1234523541 reduces to 145541=123321: 1234523541;

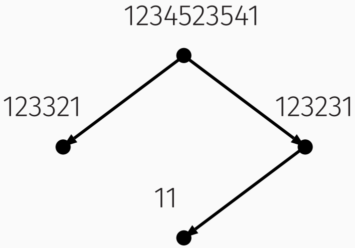
and also reduces to 123231: 1234523541



DOW reduction

Reduction of repeat and return words

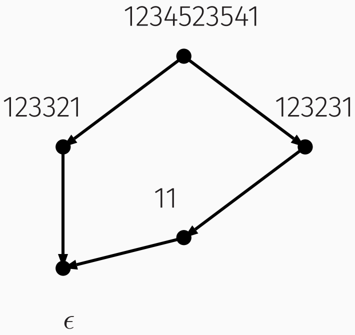
123231 reduces to 11



DOW reduction

Reduction of repeat and return words

123321 and 11 reduce to the empty word, ϵ



Future work

1. Create word reduction graphs based on DNA rearrangement
2. Compute homology groups
3. Compare homology groups corresponding to different species

Research group

This work is supported by the NSF grants CCF-1526485 and DMS-1800443 and the NIH grant R01GM109459-01.

Questions?

Thank you