TDA - fall 2025

Homology

First: an announcement for a talk next week.
Please come if you can!

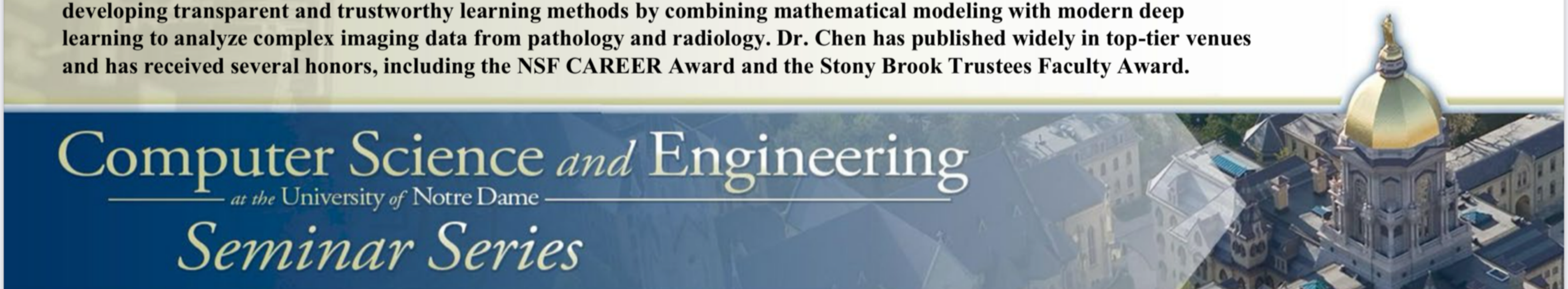
Topology-Driven Learning for Biomedical Images – from Uncertainty to Generation

Thursday, September 11, 2025
3:30 – 4:45 p.m.
303 Cushing Hall

Dr. Chao Chen
Stony Brook University

With advanced imaging techniques, we are collecting images of various complex structures such as neurons, vessels, tissues and cells. These structures encode important information about underlying biological mechanisms. To fully exploit these structures, we propose to enhance learning pipelines with topology, the branch of abstract mathematics that deals with structures such as connections, loops and branches. Under-the-hood is a formulation of the topological computation as a robust and differentiable operator. This inspires a series of novel methods for segmentation, uncertainty estimation, generation, and analysis of these topology-rich biomedical structures. These methods are applied to various problems in cancer research and neuroscience.

Chao Chen, Ph.D., is an Associate Professor in the Department of Biomedical Informatics at Stony Brook University, with affiliated appointments in the Departments of Computer Science and Applied Mathematics and Statistics. His research integrates biomedical imaging informatics, robust machine learning, and topological data analysis. He focuses on developing transparent and trustworthy learning methods by combining mathematical modeling with modern deep learning to analyze complex imaging data from pathology and radiology. Dr. Chen has published widely in top-tier venues and has received several honors, including the NSF CAREER Award and the Stony Brook Trustees Faculty Award.



Vector space

A vector space over a field k is a set V with vector addition: $\forall v, w \in V$, $v+w \in V$

& scalar multiplication: $\forall a \in k$, $a\vec{v} \in V$

s.t. it is

- associative $(+)$: $(v+w)+x = v+(w+x)$
- commutative $(+)$: $v+w = w+v$
- identity $(+ \,\&\, \cdot)$: $\exists 0_v \in V$ & $1_k \in k$ s.t. $\forall v \in V$, $0_v + v = v$ & $1_k \cdot v = v$
- inverse $(+)$: $\forall v \in V$ $\exists w \in V$ s.t. $v+w = 0_v$
- scalar mult: $a(b\vec{v}) = (ab)\vec{v}$
- 2 kinds of distributivity:

  in k $\leftarrow$ $a(v+w) = av + aw$

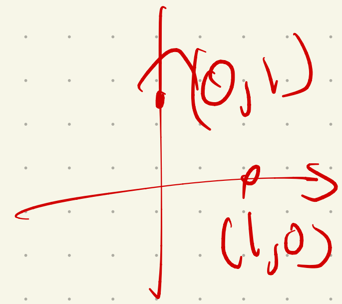
  $(a+b)w = aw + bw$

Bases

A **basis** for a vector space V is a collection of vectors $\{b_\alpha\}_{\alpha \in A}$ s.t.

- They are **linearly independent**:
  if $\sum_{\alpha \in A} \underbrace{c_\alpha}_{\text{coefficient}} b_\alpha = 0$

  then $c_\alpha = 0$.



- They **span** V:
  $\forall v \in V, \ \exists c_\alpha \in K$ s.t. $\sum c_\alpha b_\alpha = v$

Note: All bases have the same cardinality, called the **dimension** of V.

Goal: Build a vector space from a simplicial complex

Let K be a simplical complex, + fix a dimension p

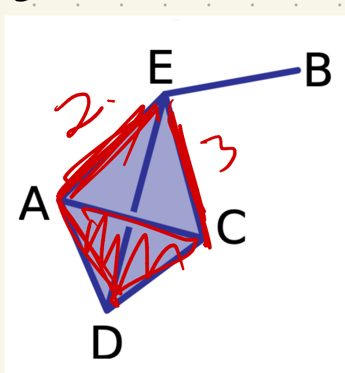
A p-chain is a formal sum of p-simlices, written

$$\alpha = \sum a_i \sigma_i$$

where $\sigma_i \in K$

Usually, each $a_i \in$ some field (or ring).

Example:



1 chain: $2\{a,e\} + 3\{ce\} + 0\{ad\}$

2 chain: $1\{acd\}$

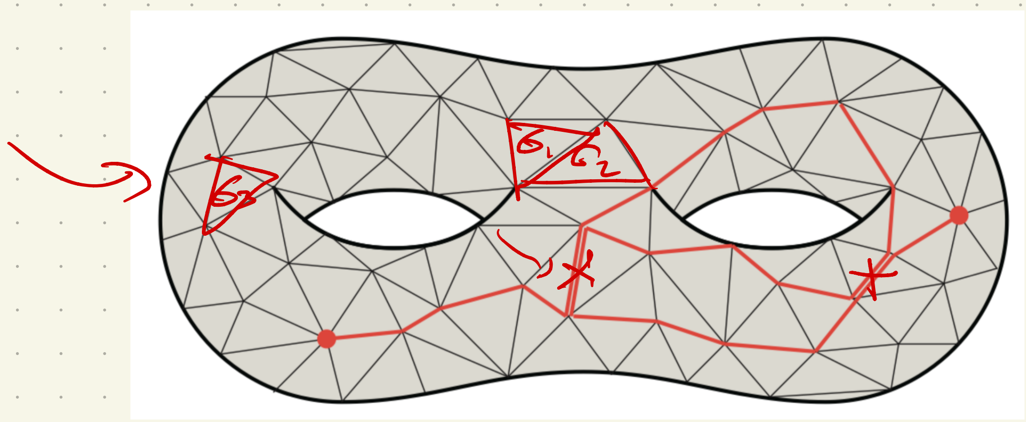
Adding Chains

If $\alpha = \sum a_i \sigma_i$ and $\beta = \sum b_i \sigma_i$

then

$$\alpha + \beta = \sum a_i \sigma_i + \sum b_i \sigma_i$$
$$= \sum (a_i + b_i) \sigma_i$$

Example: 2-dim complex with coefficients in $\mathbb{Z}_2 = \{0, 1\}$.



2-chain:
set of cycles
+ paths

$$\gamma = 1 \cdot \sigma_1 + 1 \sigma_2 + \sigma_3 + 0 \cdot \sigma_i$$

Chain group

The collection of p-chains with addition is called the p^{th}-chain group $C_p(K)$.

It is a vector space:

- associative $+$: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
- commutative $+$: $\alpha + \beta = \beta + \alpha$
- zero: $\vec{0} + \alpha = \alpha \quad + \quad 0 = \sum 0_F \cdot \sigma_i$
- inverses: How to build $-\alpha$?

$\alpha = \sum a_i \sigma_i \quad -\alpha = \sum (-a_i) \sigma_i$

Linear Transformations

A linear transformation between 2 vector spaces V + W is a map $T: V \to W$
V: dim n, W: dim m

such that:

1) $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w})$

2) $T(a\vec{v}) = aT(\vec{v})$

Representation: A matrix! Fix basis $v_1 - v_n$ + $w_1 - w_m$

$v = \sum_i a_i v_i$

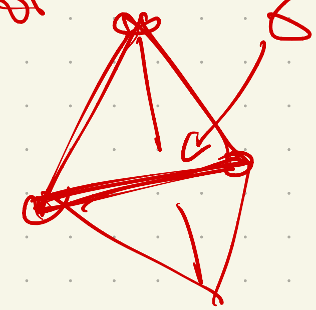
$\hookrightarrow v = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ If $\mathbb{Z}_2$ coeff., 0 or 1

then

$$\begin{pmatrix} | & & | \\ T(v_1) & \cdots & T(v_n) \\ | & & | \end{pmatrix} \underset{n \times 1}{\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}} = \underset{m \times 1}{\begin{pmatrix} \circ \\ \circ \\ \vdots \\ \circ \end{pmatrix}} \begin{matrix} w_1 \\ w_2 \\ \\ w_m \end{matrix}$$

Maps on Chain complexes

fix $\mathbb{Z}_2$ for now for field



The boundary map

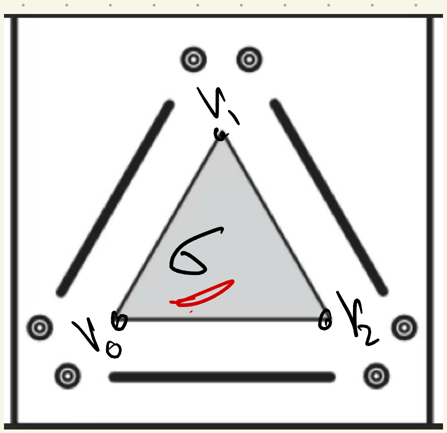
$$\partial_p : C_p(K) \longrightarrow C_{p-1}(K)$$

takes $\sigma = [v_0, \dots, v_p]$

$$\longmapsto \sum_{j=0}^{p} [v_0, \dots, \hat{v}_j, \dots, v_p]$$

Here, $\hat{v}_j$ means removing simplex j.

Example:



1) $\sigma = [v_0 v_1 v_2]$

$$\partial_2(\sigma) = [v_1 v_2] + [v_0 v_2] + [v_0 v_1]$$

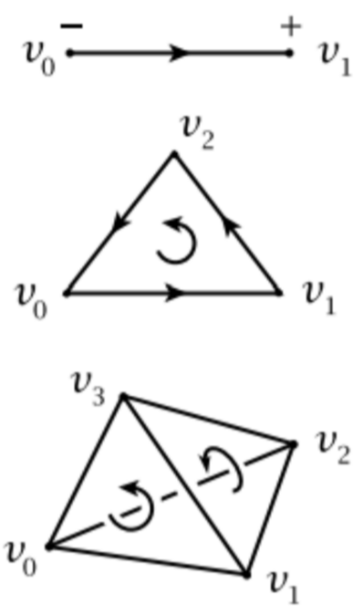
2) $\partial_1\left([v_0 v_1] + [v_1 v_2]\right) = v_0 + \cancel{v_1} + \cancel{v_1} + v_2 = v_0 + v_2$

Choices of K

Generally speaking, can study any field

↳ or even rings!

Algebraic Topology (Hatcher)



$$\partial[v_0, v_1] = [v_1] - [v_0]$$

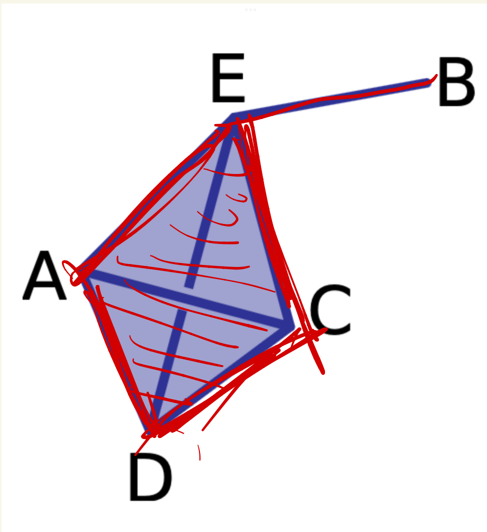
$$\partial[v_0, v_1, v_2] = [v_1, v_2] - [v_0, v_2] + [v_0, v_1]$$

$$\partial[v_0, v_1, v_2, v_3] = [v_1, v_2, v_3] - [v_0, v_2, v_3] + [v_0, v_1, v_3] - [v_0, v_1, v_2]$$

But (following book), we'll focus on $\mathbb{Z}_2$.

Why? Computation + rich geometric interpretation

Let's try:

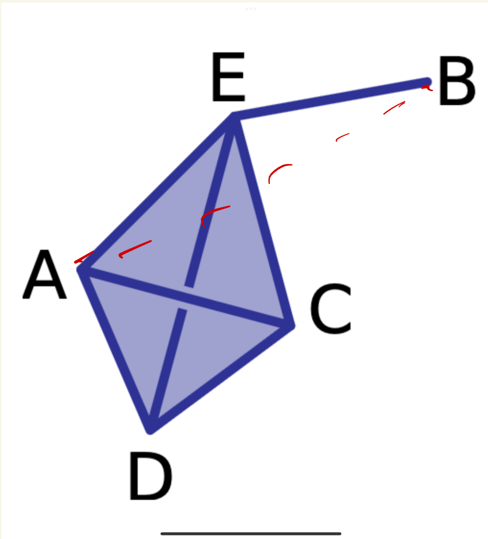


$$\partial_1\left([a,e]+[b,e]\right) = a+e+b+e$$
$$= a+b$$

$$\partial_1\left[[a,e]+[c,e]+[c,d]+[a,d]\right)$$
$$= \cancel{a}+\cancel{e}+\cancel{c}+\cancel{e}+\cancel{c}+\cancel{d}+\cancel{a}+\cancel{d}$$
$$= 0$$

$$\partial_2\left([a\,c\,e]+[a\,c\,d]\right) =$$
$$\cancel{ac}+ae+ce$$
$$+\cancel{ac}+ad+cd$$
$$= ae+ce+ad+cd$$

Matrix representation



$$\partial_1 : C_1(K) \to C_0(K)$$

take $\alpha = \sum a_i \sigma_i$

edges

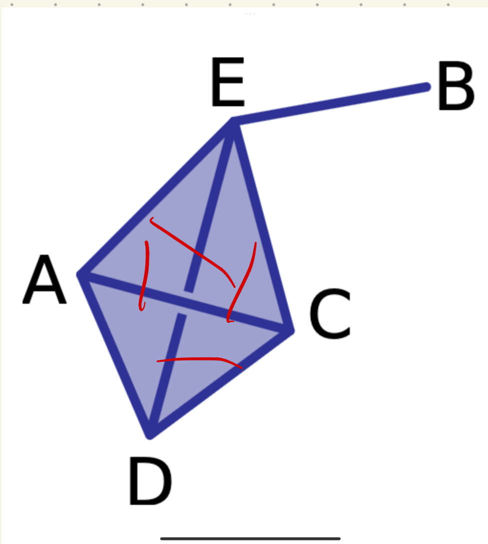
basis?

vertices

6

edges

$$\partial_1 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccccc} AC & AD & AE & BE & CD & CE & DE \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

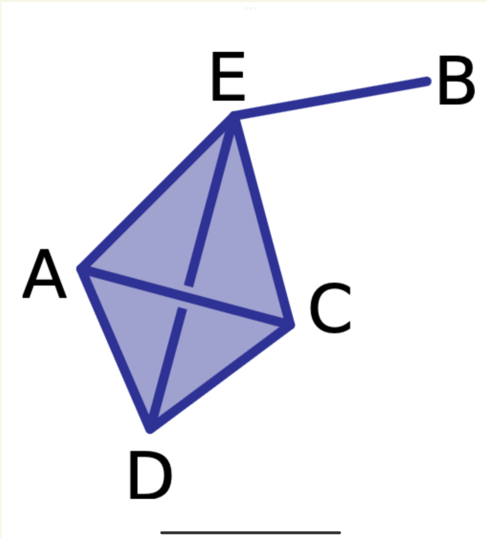


∂_2

$\triangle$'s

$$
\begin{array}{c|cccc}
 & ACE & ACD & ADE & CDE \\
\hline
AC & 1 & & & \\
AD & 0 & & & \\
AE & 1 & & & \\
BE & 0 & & & \\
CD & 0 & & & \\
CE & 1 & & & \\
DE & 0 & & &
\end{array}
$$

edges



$$\text{take } \overset{\text{some}}{\sigma} = \underline{[a,e]} + [c,e] + [c,d] + \underline{[a,d]}$$

non-zeros

$$\text{so } \sigma = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

We know $\partial_1(\sigma) = 0$:

$$\begin{array}{c|c} & \begin{array}{ccccccc} AC & AD & AE & BE & CD & CE & DE \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix} \end{array} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{array}{c} 1+1 \\ \\ 1+1 \\ \\ \\ \end{array} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \checkmark$$

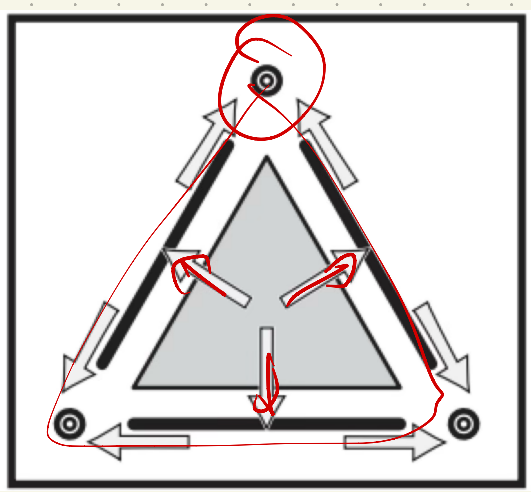
Chain Complex:

matrix matrix vertex

$$\cdots \longrightarrow C_{p+1}(K) \xrightarrow{\partial_{p+1}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K) \to \cdots C_0 \to C_{-1}$$

$C_{-1} = \emptyset$

vectors of simplices



Note: $\forall \alpha \in C_p(K)$,

$\alpha = \sum a_i \sigma_i$

Lemma:

$$\partial_{p-1} \circ \partial_p (\alpha) = 0.$$

Proof: For any p-simplex σ:

$C_p \to C_{p-1} \to C_{p-2}$

every p-2 dim simplex will appear twice

Cycles

Any chain in the kernal of ∂_p is called a p-cycle.

Reminder: an element x is in $\ker(f)$ if $f(x) = 0$

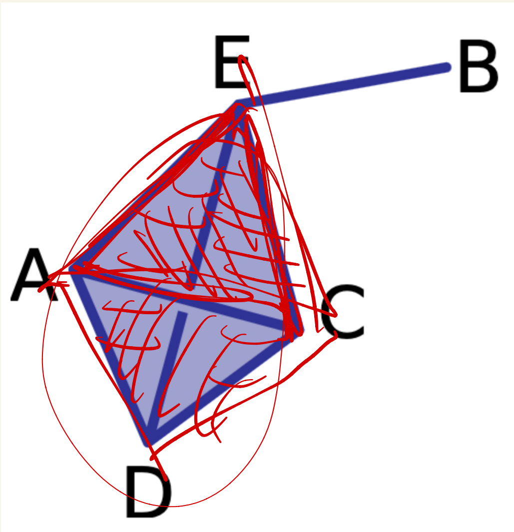
Here: $C_{p+1}(K) \xrightarrow{\partial_{p+1}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K)$

0

So: a set of simplices that, after ∂_p, cancel each other out. $\rightarrow 0$

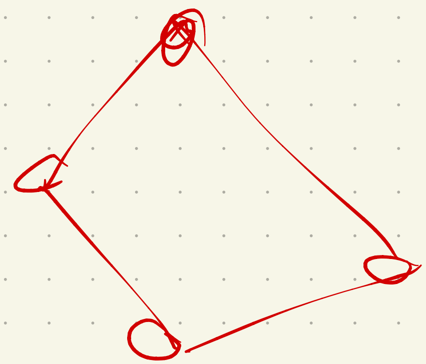
The set of p-cycles forms a subspace $Z_p(K) \subseteq C_p(K)$

What is a 1-cycle or 2-cycle?



Here:

path that ends at its start point



2cycles: enclosed 2-ball

Boundaries

A chain which is in the image of ∂_{p+1} is a p-boundary.

Reminder: $x \in \mathrm{Im}(f)$, $f: A \to B$, if $\exists y \in A$ s.t. $f(y) = x$

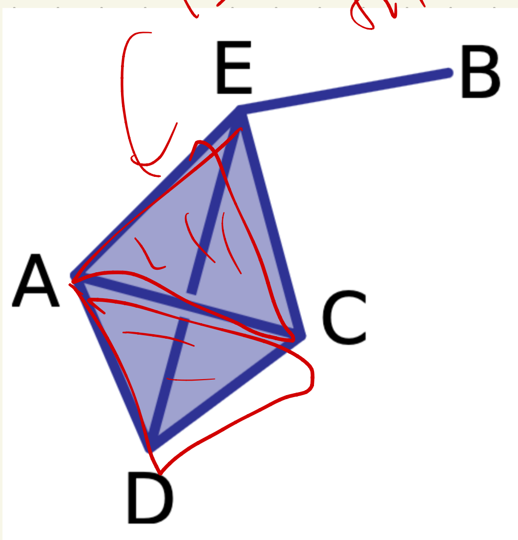
Here: $C_{p+1}(K) \xrightarrow{\partial_{p+1}} C_p(K) \xrightarrow{\partial_p} C_{p-1}(K)$

& the set of p-boundaries forms a subspace $B_p(K) \subseteq C_p(K)$.

What types of things are boundaries?

Example:

is this filled



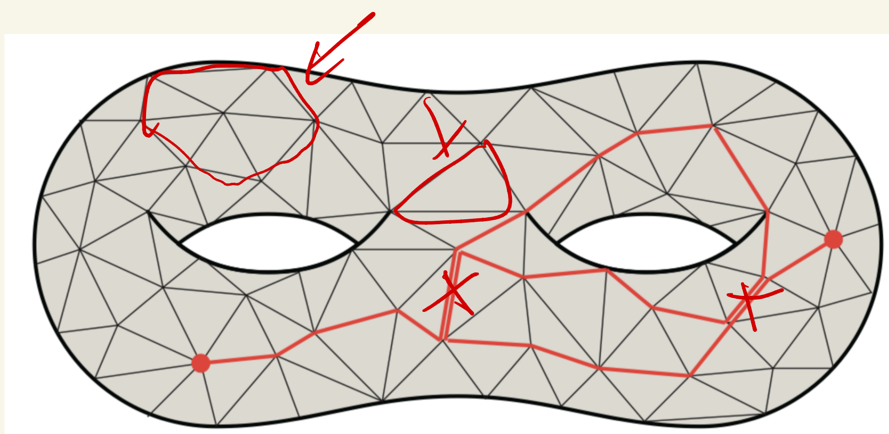
2-boundary:

$C_3 \neq \emptyset$, then {acde}

$\partial_2 (acde) = acd + ade + ace + cde$

1 boundary: bound some set of △s

Another:

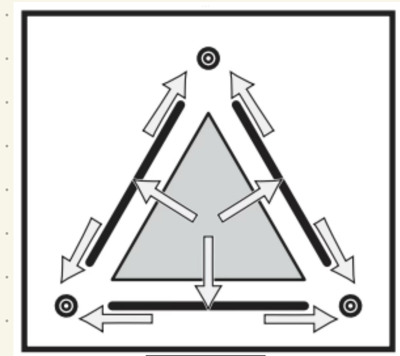


0 chain: endpoints of paths

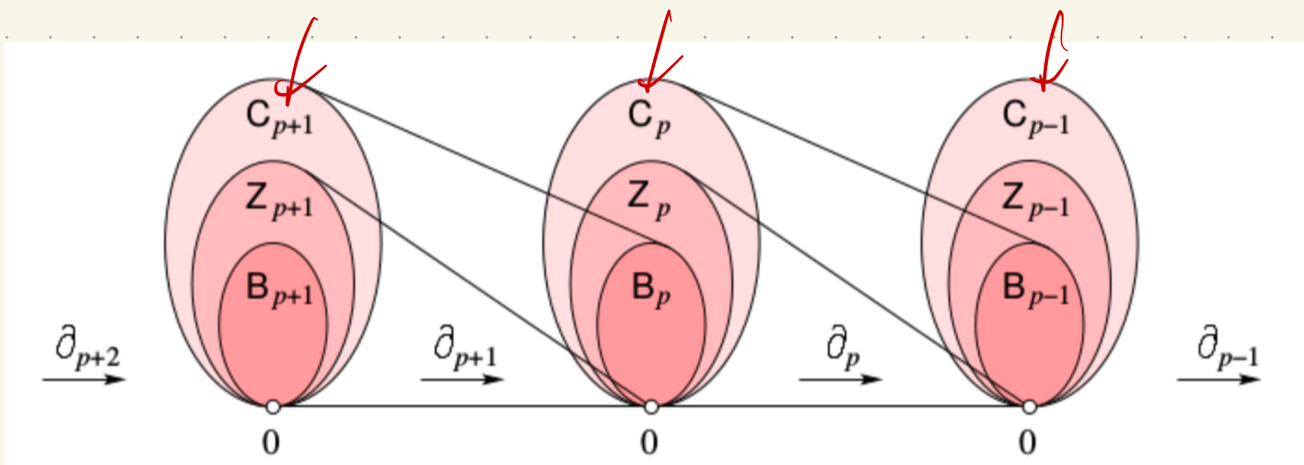
1 chain: even subgraphs

Note: Since $\partial_p \partial_{p+1}(\alpha) = 0 \ \forall \alpha \in C_{p+1}(K)$

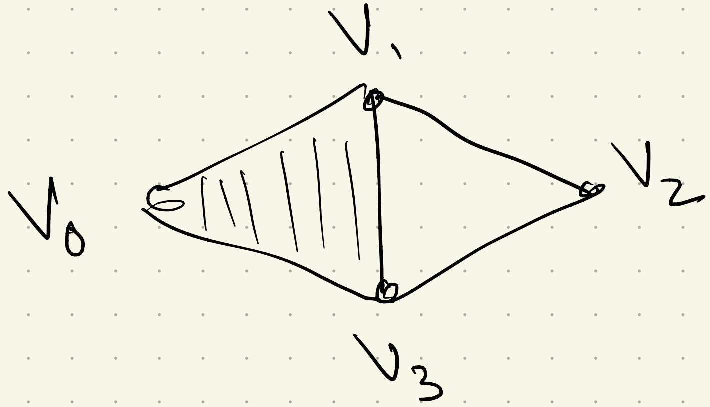
$\Rightarrow$ every p-boundary is also a p-cycle



So we get:



Example:

$K =$ 

Generators of $B_1(K)$?

Generators of $Z_1(K)$?

Quotient space

Take a vector space V over field F, and $W \subset V$ a subspace.

Define $\sim$ on V by $x \sim y$ iff $x - y \in W$.

Equivalence class of x:

$$[x] = x + W = \{x + w \mid w \in W\}$$

$$y \in [x] \Rightarrow x - y \in W$$

Then, quotient space V/W is $\{[x] \mid x \in V\}$.

Fact: V/W is a vector space with

- Scalar multiplication

$$a[x] =$$

if $y \in [x]$,
(& $y \neq x$)

- Addition:

Homology

The p^{th} homology group is the quotient space:

$$H_p(K) := Z_p(K) / B_p(K)$$

Recall: $C_{p+1} \xrightarrow{\partial_{p+1}} C_p \xrightarrow{\partial_p} C_{p-1}$

$[\alpha] \in H_p(K)$ $\alpha \in Z_p$

$$\Rightarrow \{\alpha + \beta \mid \beta \in B_p\}$$

$$= \{\alpha + \partial_{p+1}\gamma \mid \gamma \in C_{p+1}\}$$

We say $\alpha, \beta \in C_p(K)$ are homologous

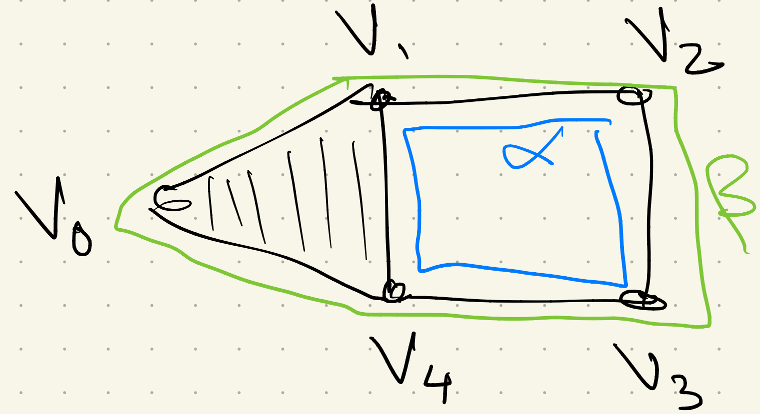
if $[\alpha] = [\beta]$ in $H_p(K)$

so: $\alpha = \beta + \partial\gamma$ for $\gamma \in C_{p+1}(K)$

cycle   cycle   boundary of higher dim chain

Time for an example ...

Can we find homologous 1-cycles?

$K =$ 

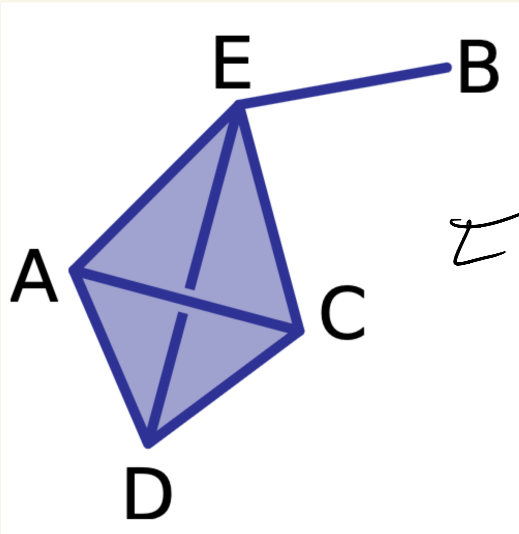
Consider: $\alpha =$

$\beta =$

If homologous, need a 2-chain γ s.t.
$\alpha = \beta + \partial_2 \gamma$.

Here, $H_1(K) = \langle \quad \rangle$

Another: What is $H_2(K)$?



no tetrahedron inside this time!

Well: $C_3(K) \xrightarrow{\partial_3} C_2(K) \xrightarrow{\partial_2} C_1(K)$

& $H_2(K) = \ker(\partial_2) / \operatorname{im}(\partial_3)$

$= Z_2 / B_2$

What is in $\operatorname{im}(\partial_3)$?

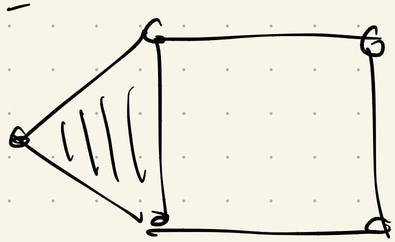
What about $\ker(\partial_2)$?

So: $H_1(K) =$

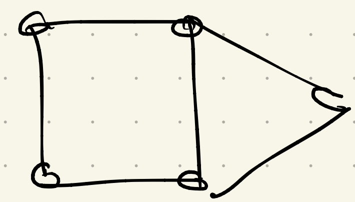
Betti numbers

The p^{th} Betti number is the rank of the p-dim homology: $\beta_p = \text{rank}(H_p)$

$K_1 =$



$K_2 =$



$\beta_1(K_1) =$

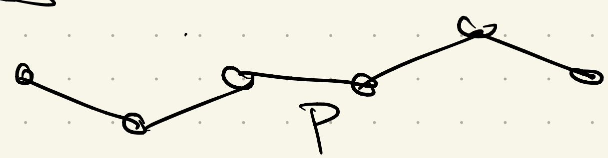
$\beta_1(K_2) =$

Some common spaces

① **Graphs**: 1d simplicial spaces

$$C_2(G)=0 \xrightarrow{\partial_2} C_1(G) \xrightarrow{\partial_1} C_0(G) \xrightarrow{\partial_2} \phi$$

- $\partial_0 = 0$, so every vertex is a 0-cycle
- B_0: boundaries of 1-chains (=paths)



$\partial(p) =$

- So $H_0(G) =$
- For H_1: no 2-cells! $\Rightarrow B_1 =$

  What is Z_1?

  basis for H_1:

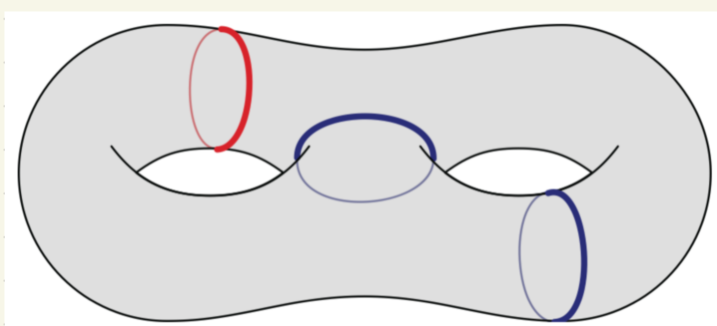
② Surfaces:

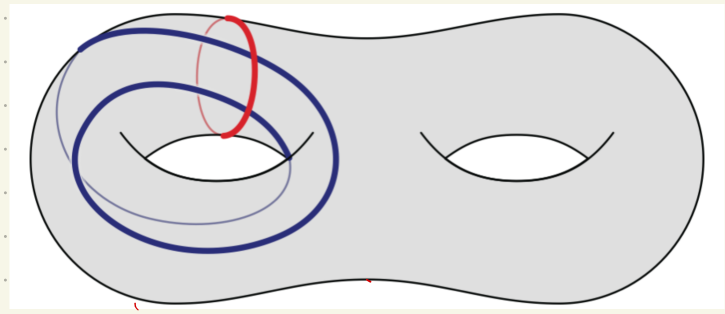
$$C_3 = \emptyset \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} \emptyset$$

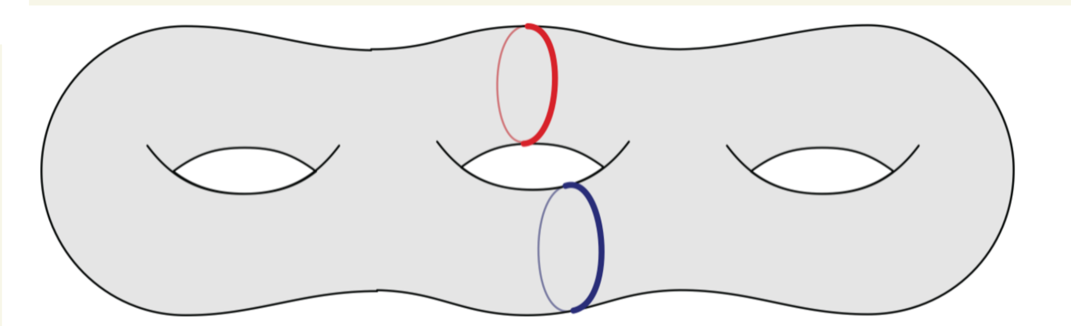
- H_0: same as graphs
- $H_1 = Z_1 / B_1 = \ker(\partial_1) / \mathrm{im}(\partial_2)$

Z_1: Still unions of cycles

B_1: differ by ∂(some $\cup$ of Δ's)



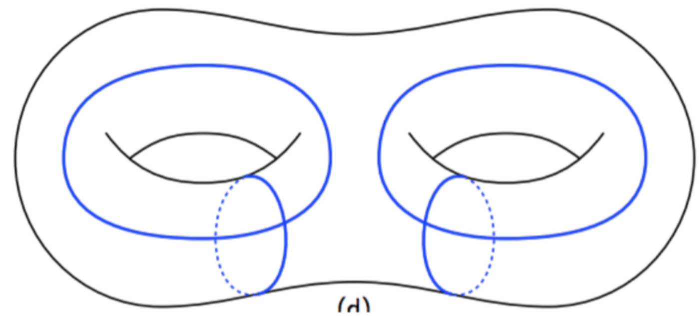




Surfaces (cont):

In the end: non-zero for H_0, H_1, & H_2:

$$\dim H_k(S_g) = \begin{cases} 1 & : \quad k = 0 \\ 2g & : \quad k = 1 \\ 1 & : \quad k = 2 \\ 0 & : \quad k > 2 \end{cases} .$$



Erickson-Whittlesey 2005

H_2: the only 2-cycle is the union of all Δ's

H_1: 2g cycles per handle