MATH 529 – Lecture 1 (01/09/2024)

Today: * syllabus, logistics
* a couple motivating applications

Call me Bala!

Introduction to **Computational** topology → focus for this course

This course will be offered completely electronically:
- scribes will be posted as "course notes"; videos will also be posted.
- assignments to be turned in electronically (you could submit scanned versions of handwritten assignments).
- web page has all the docs/info.

Topology

"Topo" → place or space
"logos" → study or talk
} in Greek

Topology talks about how space is connected.

topology
- point set topology (open/closed, connected, ...)
- algebraic topology (groups, addition, basis, ...)

We will concentrate on algebraic topology.

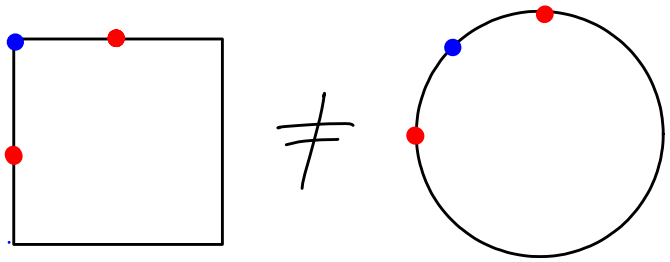
Computational topology

→ combine efficient algorithms and data structures with results from topology to analyze real-life data.

Let's start with an intuition for what we mean by connectivity of spaces.

An Example

According to geometry, the square and circle are not equal.

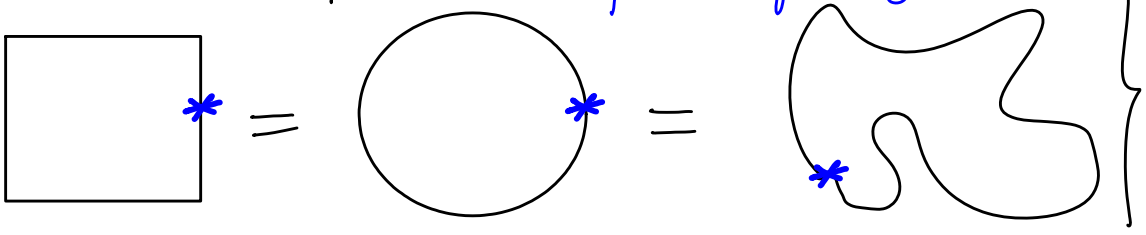


Size (length, here) is critical in geometry, but not so much in topology.

But topology says they are same as far as how they are connected!

For instance, take a string, and tie a knot to make a loop.

We want to study connectivity irrespective of size here!

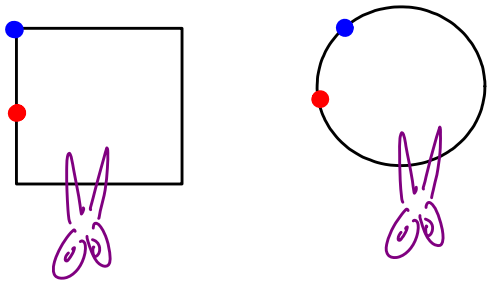


All these objects are same, i.e., they are connected the same way.

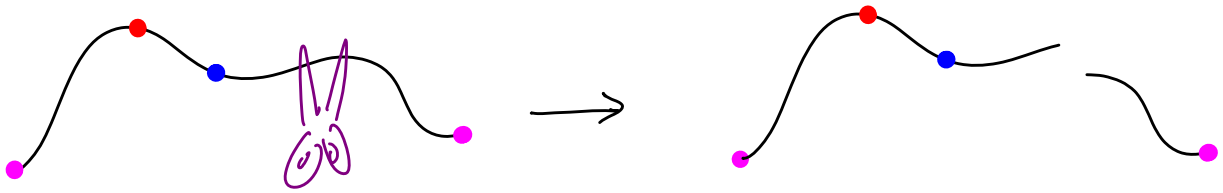
But, if you did not tie the knot, the loose string (open thread) differs from any of the above tied loops in connectivity.

Note that the two end points have different "neighborhoods."

Here is another way to understand connectivity. Consider cutting the string (tied into a loop) once. Such a cut leaves the string in one piece, i.e., connected.

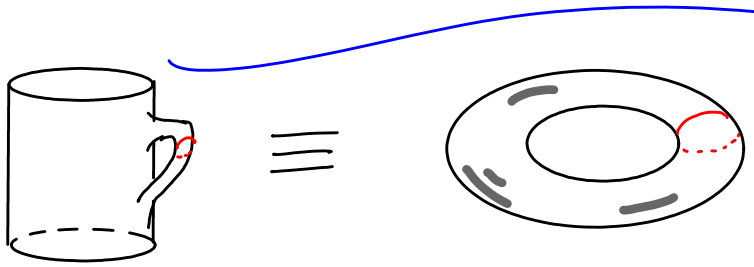


But cutting the open thread once leaves two pieces, i.e., it is disconnected.



two pieces

A popular quote: "A topologist cannot distinguish the coffee cup from a donut!"



the coffee cup, if made of clay, could be deformed to look like the donut, without changing its connectivity.

A more practical example:
how we are able to read (recognize) letters of the alphabet in different fonts.

A **A** $\mathcal{A}$ ≠ B **B** $\mathcal{B}$ ≠ C $\mathcal{C}$ **C**

Two Illustrations of Computational Topology

1. Patient antibiotic trajectories

https://doi.org/10.1145/3307339.3342143

(http://www.math.wsu.edu/faculty/bkrishna/Papers/PatientTrajectories.pdf → preprint)

? How do agents and doses affect length of stay?

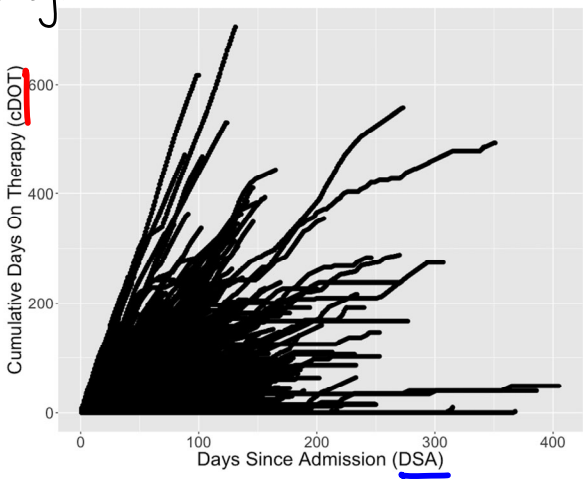
Number of hospitals	25
Number of hospital unit-categories	9
Number of distinct patient-admission records	349, 610
Number of adult patients	334, 207
Number of male patients	148, 540
Number of female patients	201, 052
Average LOS per admission	7 days
Longest LOS → length of stay	405 days
Number of antibacterials used	66
Most used antibacterial	Vancomycin
Average DOT per admission	6
Number of agent ranks	4
Most used agent rank	rank 3

→ Days On Therapy

If a patient gets one dose of an agent (antibiotic) that is counted as 1 Day On Therapy (DOT).

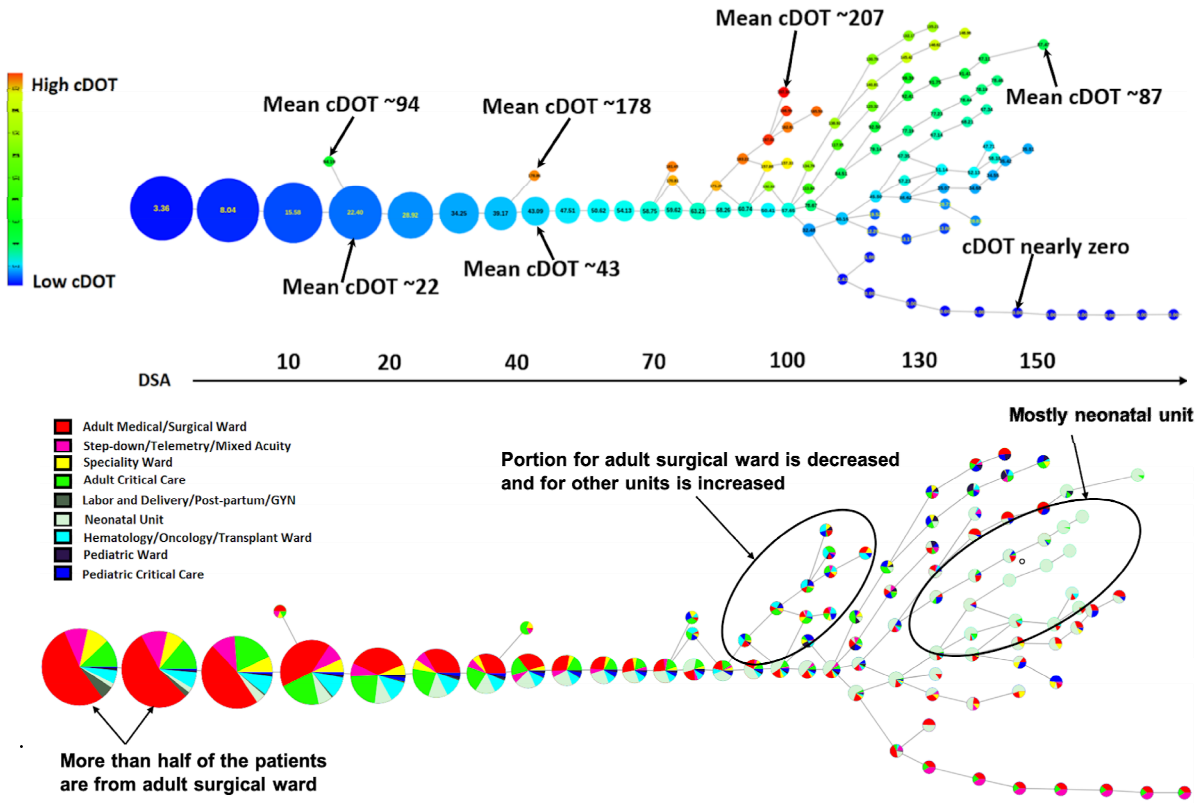
cDOT : cumulative Days On Therapy

DSA : Days Since Admission.



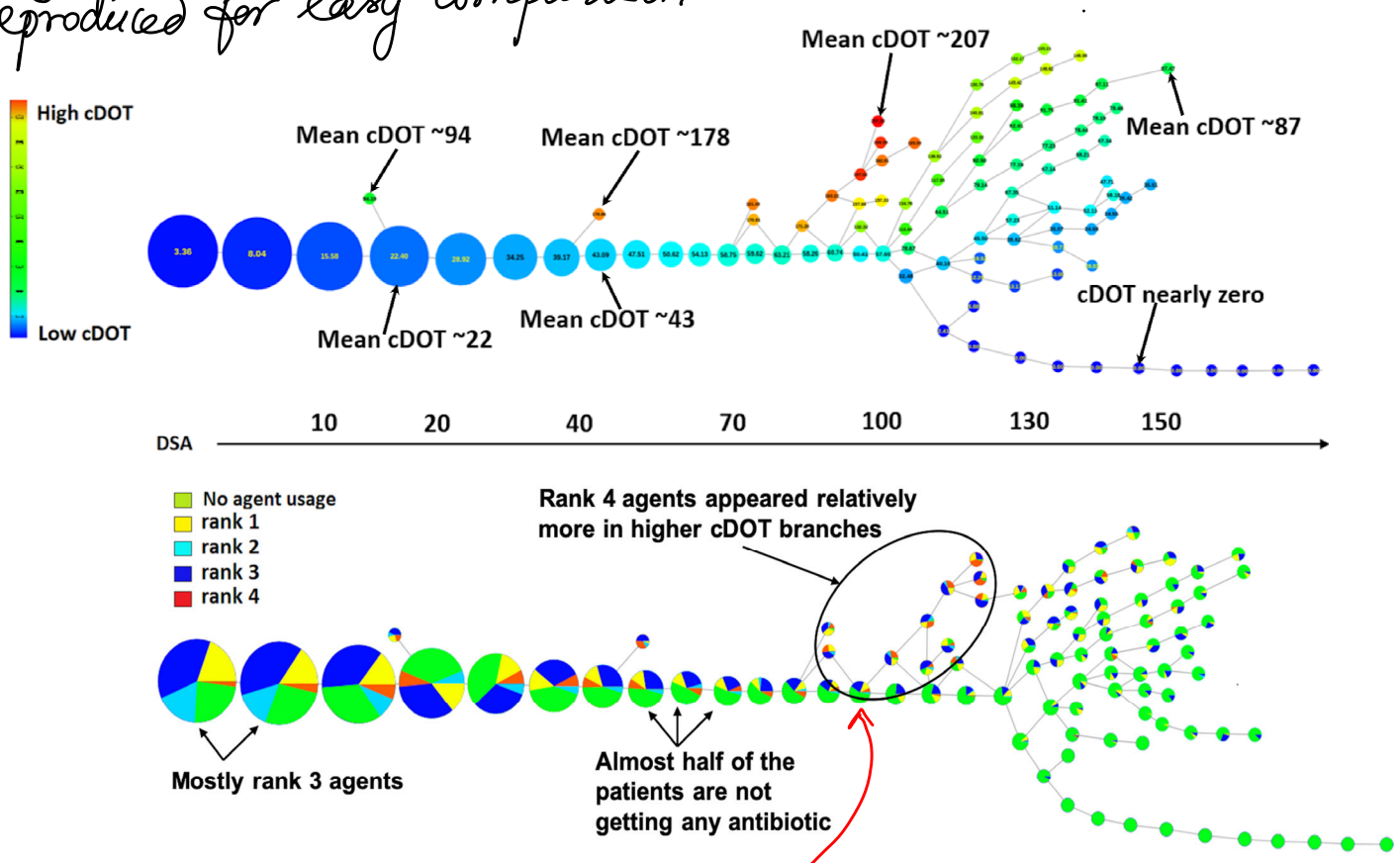
This plot by itself is not very informative or insightful (even if we were to use color...)

Here are two versions of a Mapper representation of the same data:



Each node represents a cluster of patient trajectories. A lot of the patients got low cDOTs and they had small(er) DSAs — as captured by the big clusters on the left. As the second Mapper shows, many of these patients were treated in the adult surgical ward — one of the most common types of admissions to hospitals.

Here is another version of the Mapper showing ranks (1–4, 4 is strongest) of the antibiotics. The first mapper (using cDOT) is reproduced for easy comparison.

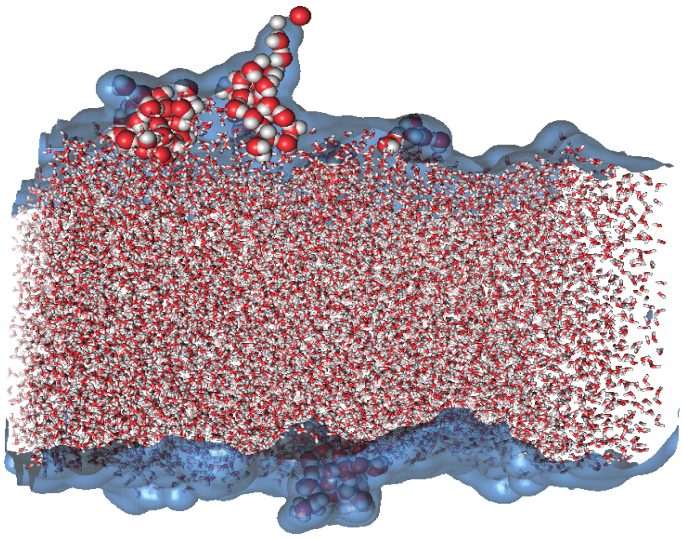


The high cDOT + high rank sub-branches had more patients in the other (higher risk) wards. Similarly, the much higher DSA group (120+) on the right-end with relatively smaller cDOT values turned out to be patients in neonatal ICUs.

Note that these nontrivial subgroups are identified in an unsupervised manner — no learning is involved!

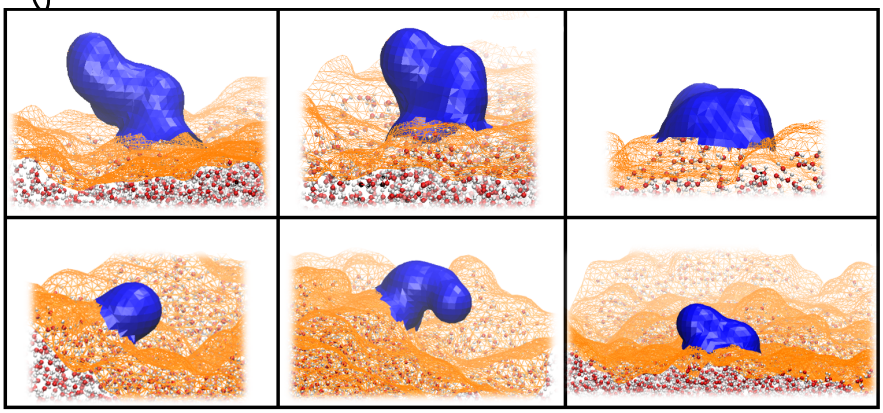
2. Interface features in Chemistry

https://doi.org/10.1021/acs.jctc.0c00260 (https://doi.org/10.26434/chemrxiv.11988048.v1) → preprint



An interface surface separates a water layer from a hexane (organic) layer. When a reagent is added, the reaction is initiated and the water molecules escape to the hexane layer through finger-like features in the interface called "protrusions". These features were identified manually (by observation!)

The goal was to identify and characterize protrusions using geometric measure theory and computational topology.



Which of these six features do you think are protrusions?

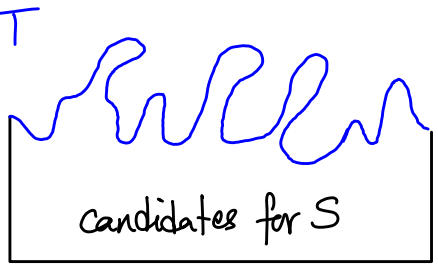
It is not easy to guess!

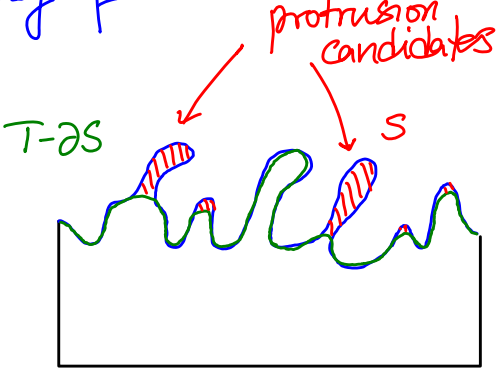
We use the notion of multiscale flat norm of surface T:

$$F_\lambda(T) = \min_S \{\text{Area}(T - \partial S) + \lambda \text{Volume}(S)\}, \quad \lambda \geq 0$$

S → 3D volume; ∂S → boundary of S; λ → scale parameter

Illustration in 2D:

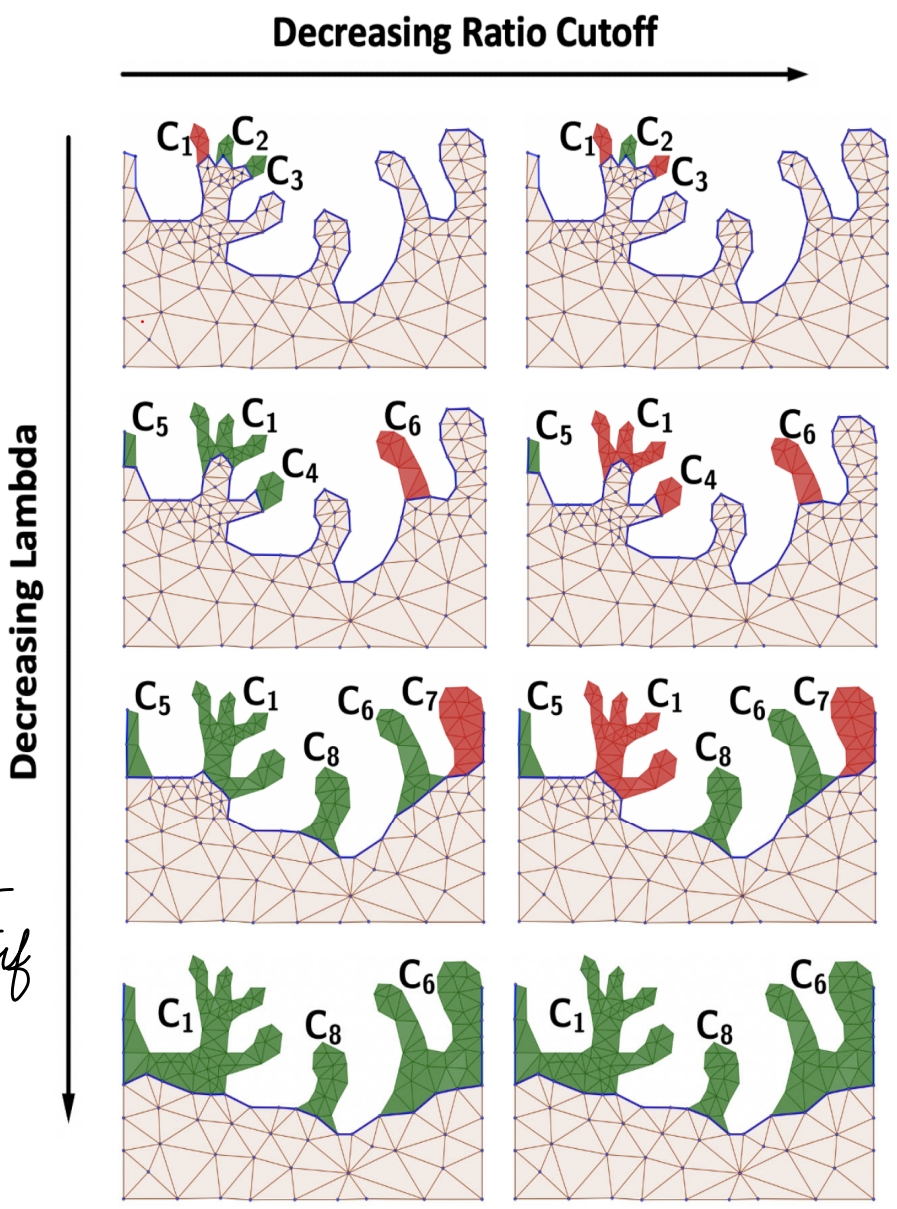




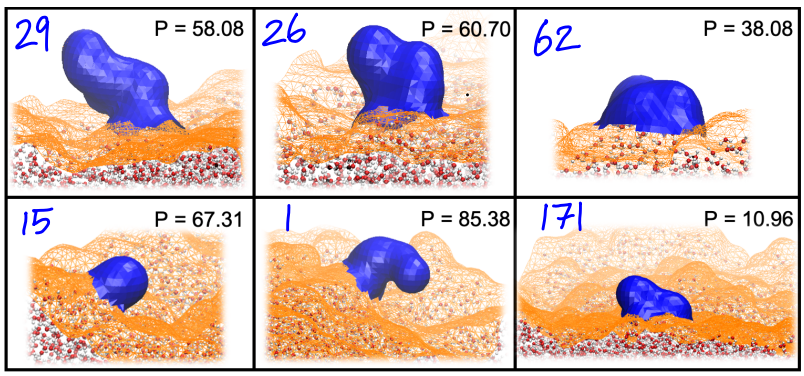
We keep track of connected components in S as $\lambda \downarrow$. We relabel them and also keep track of merging behavior.

We also track the ratio $\frac{\mathrm{Vol}(C)}{\mathrm{Vol}(\mathbb{B}(\frac{1}{\lambda}))}$ for each component C, where $\mathbb{B}(\frac{1}{\lambda})$ is the ball with radius $\frac{1}{\lambda}$ and $\mathrm{Vol}(C)$ is the volume of component C.

A component C is "alive" at ratio cutoff r and scale λ if $\frac{\mathrm{Vol}(C)}{\mathrm{Vol}(\mathbb{B}(\frac{1}{\lambda}))} > r$.



The longer a component is alive, the more likely it is to be a protrusion.



It turned out that all six of these features were protrusions! The probabilities (as %'s) along with their ranks among 195 candidate features (lower rank $\Rightarrow$ more likely to be a protrusion) are shown here.

While this example is described in 3D, the underlying concepts are more general, and in fact generate certain key fundamental questions in geometric measure theory (GMT).

In fact, when we talk about applied algebraic topology, the application could be to pure mathematics! We will talk about this aspect toward the end of the semester.

Note that we are showing a discrete version of the surface – in the form of a triangular mesh. Indeed, we need to discretize continuous spaces to perform computations!

Here is a notion of connectivity in the "discrete setting":

The neighborhood, i.e., the set of nearby points, of the two
- points are different – they each have only one neighbor, while the • points all have two neighbors each.

MATH 529 – Lecture 2 (01/11/2024)

Today: * topology, open/closed sets
* homeomorphism, examples

We define topology as a mathematical method to define and study how a space is connected.

Notation For a set X, we denote by 2^X the power set of X, which is the set of all subsets of X.

Def A **topology** on a set X is a subset T of 2^X such that the following conditions hold.

1. $A_1, A_2 \in T \Rightarrow A_1 \cap A_2 \in T$ (finite intersections)
2. $\{A_j \mid j \in J\} \in T \Rightarrow \bigcup_{j \in J} A_j \in T$ (infinite unions or finite)
   - J: index set, infinite or finite
3. $\emptyset, X \in T$ ($\emptyset$: empty set)

(X, T) is a topological space, denoted $\mathbb{X}$ (T is understood from context).

$A \in T$ is an **open set** of $\mathbb{X}$.

The complement of A, i.e., $X - A$ (or $X \setminus A$; "set minus") is a **closed set**.

Some sets can be both open and closed at the same time, e.g., $\emptyset, X$ are both open and closed in any topology.

We typically specify a topology by specifying its open sets.

interior $\mathring{A}$ of $A \subseteq \mathbb{X}$: $\mathring{A} = \bigcup$ (open sets contained in A) (union)

closure $\bar{A}$ of $A \subseteq \mathbb{X}$: $\bar{A} = \bigcap$ (closed sets containing A). (intersection)

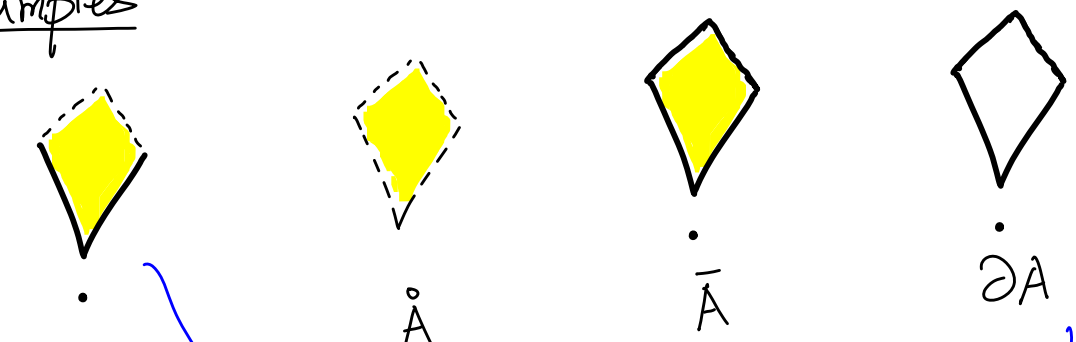
→ minimal closed set that contains A.

boundary ∂A of $A \subseteq \mathbb{X}$: $\partial A = \bar{A} - \mathring{A}$.

∂A = {points in A that intersect both $\bar{A}$ and $\overline{(X-A)}$}.

Examples

1.



→ A consists of the "half-open" rhombus and a separate point.

2. A discrete example. Let $X = \{a, b, c\}$.

We can define different topologies on X.

Let $T_1 = \{\emptyset, \{b\}, \{a,b\}, \{b,c\}, \{a,b,c\}\}$ and

$T_2 = \{\emptyset, \{a\}, \{b\}, \{a,b\}, \{a,b,c\}\}$.

Under T_1, $\{a,b\}$ is open, its complement $\{c\}$ is closed. With $A = \{a,b\}$, $\mathring{A} = \bigcup\{\emptyset, \{b\}, \{a,b\}\} = \{a,b\} = A$.

We can specify other topologies on X, e.g., $T_3 = 2^X$, where each set in T is both open and closed. But $T_4 = \{\emptyset, \{a\}, \{b\}, \{a,b,c\}\}$ is not a topology, as, e.g., $\{a\} \cup \{b\} = \{a,b\} \notin T_4$.

Neighborhood Let $\mathbb{X} = (X, T)$. A neighborhood of $x \in X$ is any $A \in T$ such that $x \in \mathring{A}$.

More generally, some books define a neighborhood as any set that includes, i.e., contains as a subset, an open set which contains X. Under this definition, the neighborhood could be a closed set (or neither open nor closed).

Now that we have defined topology, we consider the natural next question of comparing two spaces — how do we say two given spaces have the "same topology"? We introduce the notion of homeomorphism as a (strong) notion of topological similarity.

Homeomorphism

In geometry, we can study transformations that preserve "shape" of a rigid body, e.g., rotation and translation. These transformations "do not change" the geometry of the body.

In topology, we permit more types of transformations — e.g., stretch, shrink, expand, twist, etc., as long as you do not cut one piece into two or more, or join two pieces into one, or poke a hole in your object. All such permitted transformations "preserve topology".

A series of such permitted transformations that preserve topology constitute a homeomorphism. And two spaces are topologically "similar" if such "nice" functions exist from one space to the other and also back. We define what we mean by "nice" here.

We start with some background and definitions on functions.

Def Let A, B be sets. A function $f: A \to B$ is a rule that assigns exactly one $b \in B$ for every $a \in A$.

$\operatorname{dom} f$: domain of $f = A$, $\operatorname{cod} f$: codomain of $f = B$

$\operatorname{im} f$: image of $f = \{b \in B \mid f(a) = b \text{ for some } a \in A\} = \{f(a) \mid a \in A\}$.

$\operatorname{im} f$ is also called the range of f. Note that $\operatorname{im} f \subseteq \operatorname{cod} f$.

$f: A \to B$ is 1-to-1 or **injective** if $\forall b \in B$ ("for every"), there exists **at most** one $a \in A$ with $f(a) = b$: → can be none

$f: A \to B$ is onto or **surjective** if $\forall b \in B$, there exists **at least** one $a \in A$ with $f(a) = b$. → can be more

If f is both injective and surjective, we say that f is **bijective**, or that f is a bijection.

Def A function $f: X \to Y$ is **continuous** if for every open set $B \subseteq Y$, $f^{-1}(B)$ is open in X. "takes" open sets to open sets.

A continuous function is also called a **map**.

Def A **homeomorphism** $f: X \to Y$ is a bijective function such that both f and f^{-1} are continuous.

We say X is homeomorphic to Y, or $X \approx Y$.

We also say that X and Y have the same topological type.

Examples

It's often easier to argue why two spaces are not homeomorphic — we just identify one (or more) place(s) where things don't work.

1. $[0,1] \not\simeq S^1$ ("not homeomorphic")

$[0,1]$: 0 — 1

S^1: 1-sphere, or circle

We would need a map that assigns both end points of $[0,1]$ to a single point in S^1.

But the inverse of any map that takes both end points of $[0,1]$ to one point in S^1 is not bijective.

2. $S^1 \not\simeq$ Figure 8

The crossing point in ∞ (x) cannot be mapped to a corresponding point in S^1.

Also, we could map S^1 to one of the two circles in figure-8, but not both.

On the other hand, two show that two spaces are homeomorphic, we need to specify the maps f and f^{-1}.

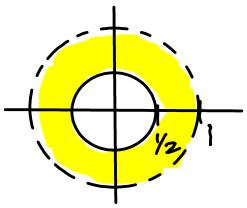
3. $D_1 \approx \mathbb{R}^2$

$$D_1 = \{\bar{x} \in \mathbb{R}^2 \mid \|\bar{x}\| < 1\} \rightarrow \text{open unit disc}$$

f: $D_1 \to \mathbb{R}^2$

$f^{-1} = g$ for ease of notation: $\mathbb{R}^2 \to D_1$

Intuitively, we can shrink all of $\mathbb{R}^2$ into D_1. Similarly, we can stretch D_1 to fill all of $\mathbb{R}^2$.



$$g(\bar{x}) = \frac{\bar{x}}{1+\|\bar{x}\|_2} \qquad g: \mathbb{R}^2 \to D$$

Euclidean norm

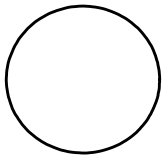
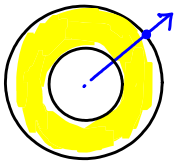
g maps all of D_1 (in $\mathbb{R}^2$) to fit within $D_{1/2} = \{\bar{x} \in \mathbb{R} \mid \|\bar{x}\| < 1/2\}$, and then fits all of $\mathbb{R}^2$ outside D_1 within the half open annulus with radii $1/2$ and 1.

The continuous function going from D_1 to $\mathbb{R}^2$ can be similarly defined.

$f: D \to \mathbb{R}^2$ where $f(\bar{x}) = \frac{\bar{x}}{1-\|\bar{x}\|}$. f is an "infinite stretch".

Note that points $\bar{x}$ in D that are close to the edge, i.e., have $\|\bar{x}\|$ close to 1, are mapped so as to fill up the entire $\mathbb{R}^2$ outside $\bar{D}$. We stretch the open disc so as to fill the entire plane, and hence it is called an infinite stretch.

Usually, we try to define the continuous maps f and f^{-1} to show that two spaces $\mathbb{X}$ and $\mathbb{Y}$ are homeomorphic. At the same time, the intuition (geometric when possible) is also important to grasp. On the other hand, to show that $\mathbb{X} \not\cong \mathbb{Y}$, it is often sufficient to identify subset(s) that create the obstructions, e.g., the $\times$ in figure-8 v/s S^1.

4.  $\not\simeq$ 

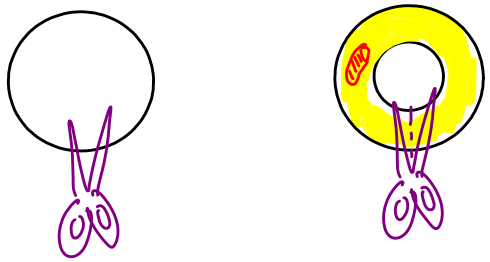
S^1 annulus

Both these spaces have the shape of a "hole".

We could shrink the annulus so that it reduces to the circle. The corresponding function maps every point on the annulus radially onto the outer circle, for instance. But we cannot uniquely map the circle back to the annulus – would need to "map" each point on the circle to (infinitely) many points on the thick strip of the annulus.

Another observation highlights the neighborhoods of points in the circle and the annulus. Every point on the circle has open neighborhoods that look like the number line ($\mathbb{R}^1$). On the other hand, points in the annulus have neighborhoods that look like the open disc ($\mathbb{R}^2$) or open half disc (the points on the boundary). Intuitively, the annulus is 2-dimensional, while S^1 is one-dimensional.

S^1 annulus



Notice that the two spaces behave the same way under a "cut" as we had been talking about earlier with the string.

In particular, a straight cut along one "edge" of either space would leave them both connected. At the same time, one could "carve out" a 2D disc ( ) from the annulus, but not from the circle.

If we "relax" our definition of topological similarity, the two spaces would be considered the same – they both look like a hole, after all. Indeed we will see that checking for homeomorphism is difficult (both theoretically and computationally). We'll work with looser concepts of topological similarity later on – homology!

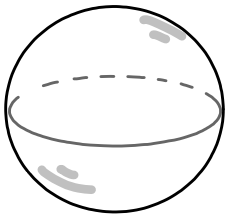
MATH 529 – Lecture 3 (01/16/2024)

Today: * 1 more example of homeomorphism
* manifolds

Examples of homeomorphism (continued...)

5. sphere $\not\cong \mathbb{R}^2$

$\mathbb{S}^2$ (2-sphere)



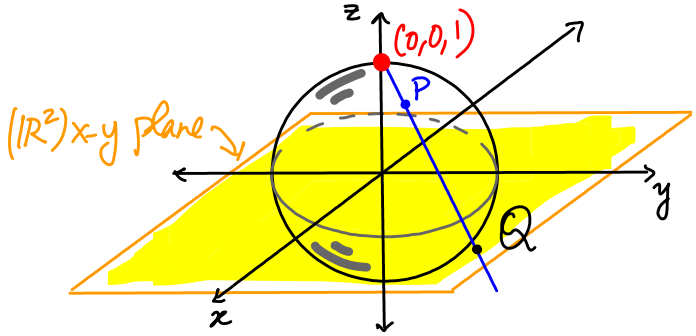
Here is an observation: the sphere encloses a 3D pocket (or void), while $\mathbb{R}^2$ does not do so.

Such enclosed voids are the 3D analogues of holes (which are 2D).

But $\mathbb{R}^2 \cup \{\infty\} \approx \mathbb{S}^2$ → By stereographic projection, which is used to map out the surface of the earth on to a (planar) map for instance.

↓ "point at infinity"

Recall the equation for $\mathbb{S}^2$: $x^2+y^2+z^2=1$.



If you poke a hole in the sphere, you can spread it out on the 2D plane, like a pierced balloon.

The equation of the line connecting $(0,0,1)$ and $P(x,y,z)$ is given by

$\bar{x} = (0,0,1) + t(x-0, y-0, z-1)$, $t \in \mathbb{R}$.

This line intersects the x-y plane at Q, which has $z=0$. Hence we get $t(z-1)+1=0 \Rightarrow t = \frac{1}{1-z}$.

Thus Q is $\left(\frac{x}{1-z}, \frac{y}{1-z}, 0\right)$.

$P(x,y,z)$ on $\mathbb{S}^2$ projected from $(0,0,1)$ (→ North pole) to $\mathbb{R}^2$ is

$Q\left(\frac{x}{1-z}, \frac{y}{1-z}, 0\right)$.

↙ This formula is valid for all points on $\mathbb{S}^2$, except the north pole $(0,0,1)$.

The lower hemisphere of $\mathbb{S}^2$ gets mapped to D (unit disc), and the upper hemisphere gets mapped to rest of $\mathbb{R}^2$.

According to a topologist, "a sphere is nothing but the plane with a point added at infinity"!

Note that every point on $\mathbb{S}^2$ has a neighborhood that looks like $\mathbb{R}^2$ i.e., "it feels locally Euclidean". Such objects are called **manifolds** and are among the most commonly studied spaces in computational topology.

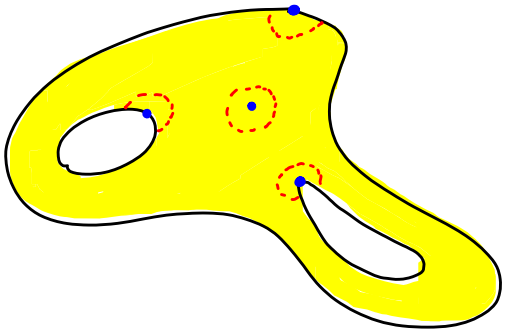
Def A topological space $\mathbb{M}$ is a **2-manifold** if all points in $\mathbb{M}$ lie in open discs, i.e., every point has a neighborhood $\approx D = \{\bar{x} \in \mathbb{R}^2 \mid \|x\| < 1\}$.

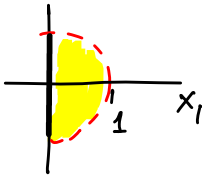
these are 2-manifolds without boundary.

e.g., $\mathbb{S}^2$, $\mathbb{R}^2$.

A **2-manifold with boundary** is a topological space $\mathbb{M}$ whose every point has a neighborhood homeomorphic to D or to $D_+ = \{\bar{x} \in \mathbb{R}^2 \mid \|\bar{x}\| < 1, x_1 \geq 0\}$ (but not both), and there exist some points of the latter type.

1st entry



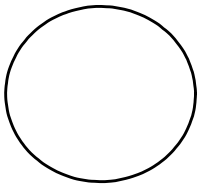
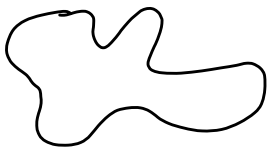
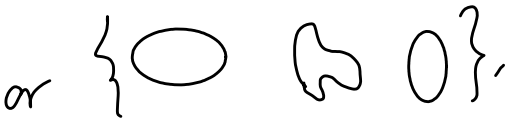


points on the boundary have neighborhoods that are half discs.

Def The **boundary** of a 2-manifold with boundary $\mathbb{M}$ is the set of points in $\mathbb{M}$ that have neighborhoods homeomorphic to the half disc.

(The definition of ∂A used for sets A is equivalent to this definition).

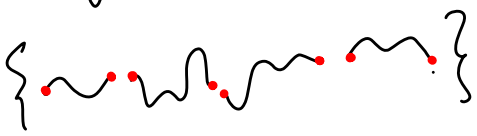
Notice that the 1-manifold is just the circle ($\mathbb{S}^1$),

 or  or 

a collection of disjoint circles.

A 1-manifold with boundary:  or

only one end point is included here!



boundary is indeed the set of end points.

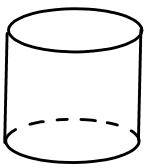
0-manifold: Any collection of distinct points (discrete set of points).

2-manifolds, also called surfaces, are a well studied class of spaces (objects), both from the theoretical as well as applied points of view. We will present several details of the properties of 2-manifolds first. To define and study d-manifolds for $d \geq 2$, we will need a few more definitions and concepts from analysis / point set topology.

By default, we assume a manifold (w/ or w/o boundary) is connected.

A 2-manifold (for that matter, d-manifold for $d \geq 2$) is orientable or non-orientable.

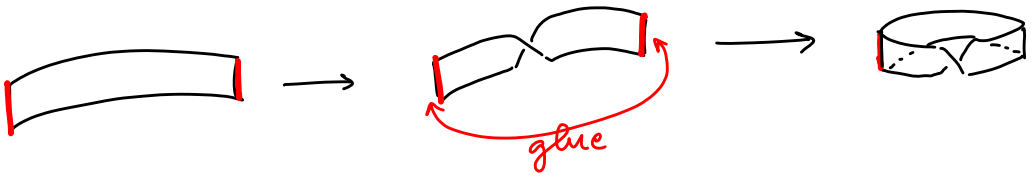
Möbius strip is non-orientable

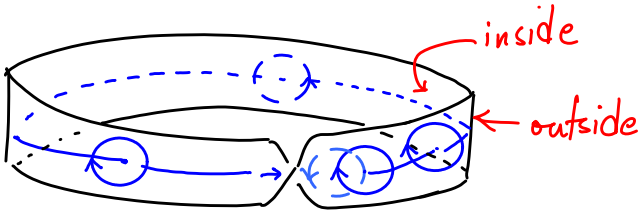


cylinder is orientable

The Möbius strip has only one "side", while the cylinder has two "sides" — inside and outside.

We obtain the Möbius strip by taking a rectangular strip of paper, and gluing its short edges together after twisting the strip once.





there are two "sides" at each point on the Möbius strip – front/back or inside/outside.

Consider sliding an oriented loop, or a clock ◯ along the surface of the Möbius strip. Look at the path followed by the center of the clock. Once the center comes back to where it started, its orientation is reversed (as it goes over the "twist").

The path traced by the center of the clock here is hence an **orientation reversing** closed curve. If the orientation is not reversed this way, the curve is said to be **orientation preserving**.

Def If all closed curves in a 2-manifold (with or without boundary) are orientation preserving, then the 2-manifold is **orientable**, else it is **nonorientable**.

↙ $\mathbb{R}^2$, $S^1 \times S^1$, torus, etc.

↳ Möbius strip, Klein bottle, etc.

Classification of Manifolds

Enumerate all possible manifolds up to homeomorphisms for each dimension.

We mention the classification for 0- and 1-dimensional manifolds, but will come back to give some more definitions to finish the discussion for 2-manifolds.

0-manifolds: a discrete space, e.g., $\mathbb{Z}^2$ – all points with integer coordinates in $\mathbb{R}^2$.

each point has to have a neighborhood $\approx \mathbb{R}^0$, i.e., a point.

Notice that $2\mathbb{Z}^2$, all points with even integer coordinates in $\mathbb{R}^2$ is homeomorphic to $\mathbb{Z}^2$.

1-manifolds

connected

- compact: $\approx S^1$, ◯ (We will define "compact" soon...)
- compact with boundary: $\approx [0,1]$ closed unit interval
- non-compact: $\approx \mathbb{R}^1$ or $(0,1)$
- non-compact with boundary $\approx \mathbb{R}^1_+$ ($\hookrightarrow x \geq 0$) or $[0,1)$ ↓ half-open unit interval

MATH 529 – Lecture 4 (01/18/2024)

Today: * d-manifolds in general
* Classification of 2-manifolds

We will now introduce some concepts which we will use to define d-manifolds in general (in particular, for $d \geq 2$).

Def A **cover** of $A \subseteq X$ is a family $\{C_j \mid j \in J\}$ in 2^X such that $A \subseteq \bigcup_{j \in J} C_j$. ($J$: index set)

An **open cover** is a cover made of open sets.

A **subcover** of A is a cover $\{C_k \mid k \in K\}$ such that $K \subseteq J$. (K: index set is a subset)

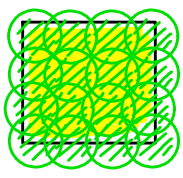
Def A set A is **compact** if every open cover of A has a finite subcover. Correspondingly, a topological space $A \subseteq X$ is compact if every open cover of A has a finite subcover.

Note: In $\mathbb{R}^d$, closed+bounded $\Leftrightarrow$ compact.

e.g., S^2 is compact, but $\mathbb{R}^2$ is not.

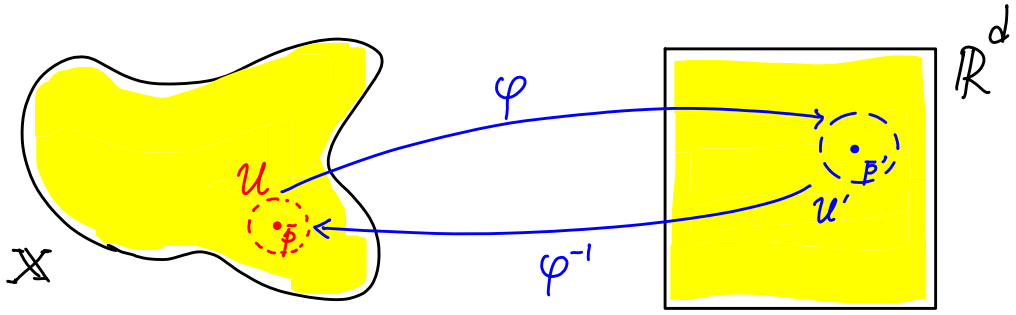
D_1 is not compact.

$\overline{D}_1$ (closed disc) is a compact 2-manifold with boundary



Example of finite subcover: consider a unit square, which is a subset of $\mathbb{R}^2$. Consider open discs of radius $\frac{1}{4}$ centered at each rational point within the square (denoted here as ◎). There are infinitely many such discs, which together cover the square. But a finite subset of those discs also covers the square.

Def A **chart** at $p \in X$ is a homeomorphism $\varphi : U \to \mathbb{R}^d$ for $U \in X$ an open set containing p. The **dimension** of the chart is d.



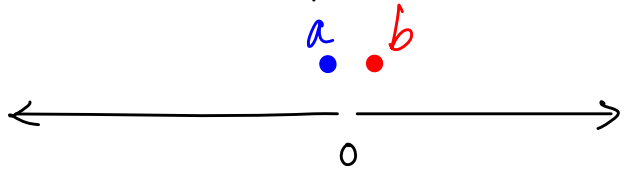
Def (Hausdorff) A topological space X is **Hausdorff** if $\forall x, y \in X$, $x \neq y$, there exist neighborhoods U, V of x, y, respectively, such that $U \cap V = \emptyset$.

→ an open set containing the point in question

e.g., $\mathbb{R}^2$.

Example of a non-Hausdorff space: $X = L \cup \{a, b\}$ where

a and b are both used together in place of the origin. Open sets are usual open intervals on the line.



$L = \mathbb{R} \setminus \{0\}$

and such that

$L \cup \{a\} \approx \mathbb{R}$ and $L \cup \{b\} \approx \mathbb{R}$.

But every pair of open sets U and V containing a and b, respectively, intersect!

Def A topological space is **completely separable** (or second countable) if it has a countable basis, i.e., it has a countable collection of open sets such that every open set can be written as a union of open sets from this collection (basis).

(think $\mathbb{Z}$, integers, as opposed to $\mathbb{R}$, which is uncountable)

e.g., $\mathbb{R}$ is completely separable – it can be shown that open intervals with rational lengths centered at only rational points works as a countable basis.

A space that is not completely separable: take uncountably many copies of $[0,1)$, e.g., with the 0 of $[0,1)$ anchored at all irrational points – called the long line or Alexandroff line.

Def (manifold) A completely separable, Hausdorff space $\mathbb{X}$ is a **d-manifold** (d-dimensional manifold) if there exists a d-dimensional chart at every $\bar{x} \in \mathbb{X}$, i.e., $\bar{x}$ has a neighborhood homeomorphic to $\mathbb{R}^d$.

$\mathbb{X}$ is a d-manifold with boundary if every $\bar{x} \in \mathbb{X}$ has a neighborhood homeomorphic to $\mathbb{R}^d$ or $\mathbb{H}^d = \{\bar{x} \in \mathbb{R}^d \mid x_1 \geq 0\}$ (d-dimensional half space).

The boundary of $\mathbb{X}$, denoted by $\partial\mathbb{X}$, is the set of $\bar{x} \in \mathbb{X}$ with a neighborhood homeomorphic to $\mathbb{H}^d$.

The **dimension** of the manifold is d here.

Notice the correspondence between the definition of 2-manifolds introduced previously, and the general definition here. The main condition is the existence of neighborhoods $\approx \mathbb{R}^d$ around each point.

Def (Embedding) An **embedding** of $\mathbb{X}$ in $\mathbb{Y}$ is a map $g : \mathbb{X} \to \mathbb{Y}$ whose restriction to $g(\mathbb{X})$ is a homeomorphism.

Manifolds are manifolds irrespective of their embedding! S^2 is a 2-manifold even if it is not sitting in $\mathbb{R}^3$.

We will introduce alternative representations of manifolds to highlight this point. In fact, in many cases, we can study the manifold easily using such representations.

Classification of Manifolds (continued)

Enumerate all possible manifolds of a given dimension upto homeomorphism. We already listed the classifications for 0- and 1-dimensional manifolds.

We now consider the case of compact, connected, closed 2-manifolds. We first list the "basic building blocks", so to speak, which include the 2-sphere, torus, Möbius strip, and the real projective plane. We can build larger 2-manifolds by gluing these building blocks together.

2-Manifolds (We consider compact 2-manifolds)

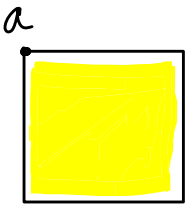
First, let us study several typical 2-manifolds, some of which we already saw previously.

1. S^2

2-sphere

$$\{\bar{x} \in \mathbb{R}^3 \mid \|\bar{x}\|_2 = 1\}$$

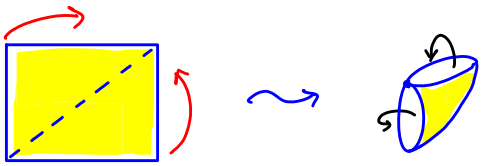
both are 2-spheres!



"identify" all points on boundary with the point a.

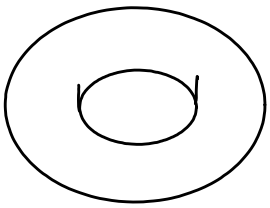
This is a "diagram" of S^2.

Start with a square sheet of paper and glue its all its edges together to make a sphere.



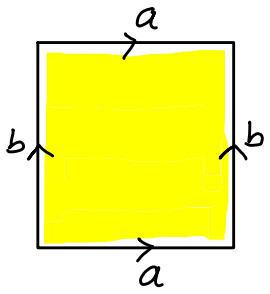
Arrows capture how edges are glued – with or without twist.

2.



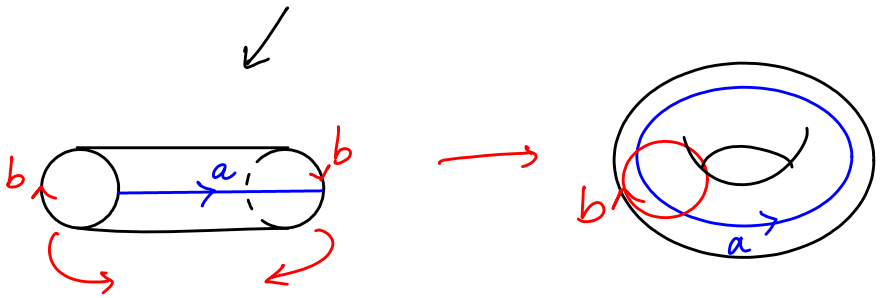
$\mathbb{T}^2$ torus

$\approx$

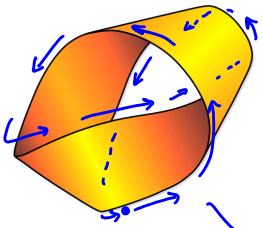


→ we could mathematically study this representation

Imagine folding a rectangular sheet of paper first into an open cylinder, and then gluing its end circles to form a torus.

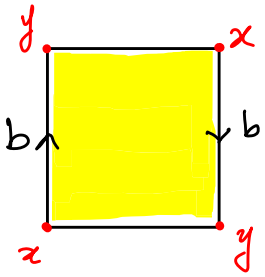


3. Möbius strip → 2-manifold with boundary



(image: www)

$\approx$

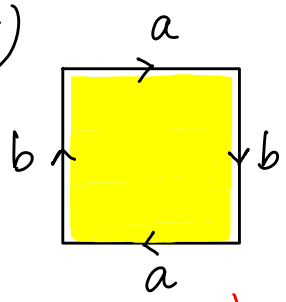


Notice that we can traverse the "edge" of the Möbius strip in one go — it's one big circle!

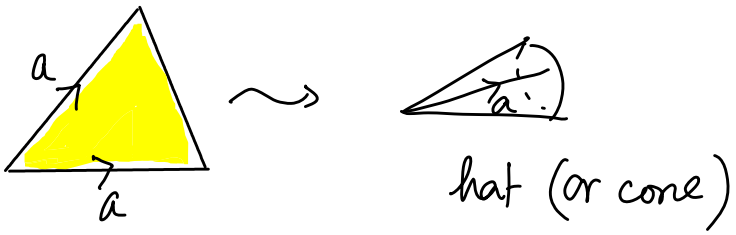
Notice that we are not identifying the horizontal edges. So they remain as boundaries. At the same time all four edges are identified pairwise in the case of the torus. Indeed, the Möbius strip is a manifold with boundary, while the torus is a manifold (without boundary).

→ cannot be embedded in $\mathbb{R}^3$!

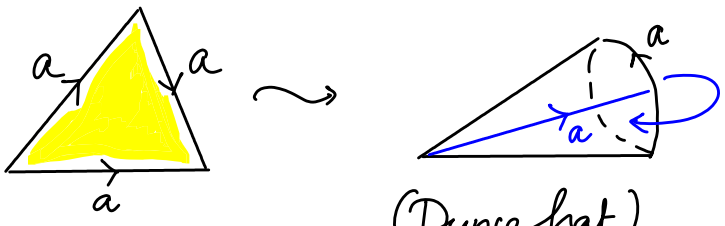
4. (Real) Projective plane ($\mathbb{RP}^2$)
(also, Dunce hat)



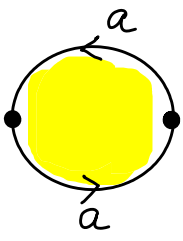
identify the free edges of Möbius strip in an opposing sense.



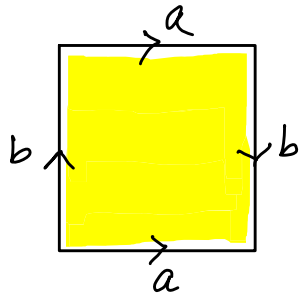
→ Same as gluing the boundary of a disc to the boundary of a Möbius strip.



Here's yet another representation →



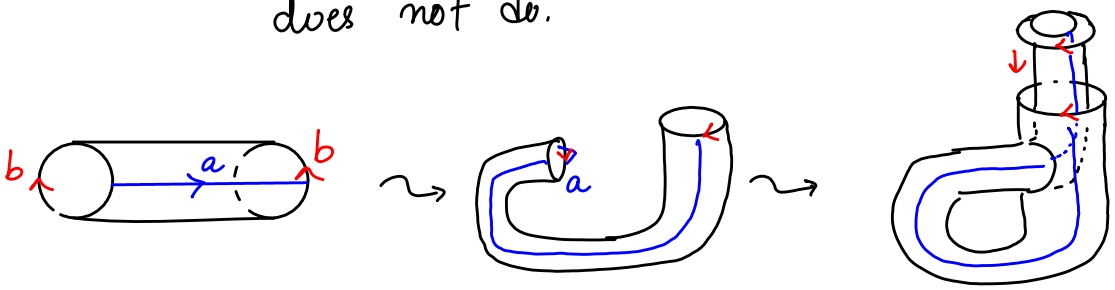
5. Klein bottle ($\mathbb{K}^2$)



Identify free edges of the Möbius strip in the same direction.

An "immersion" in 3D:

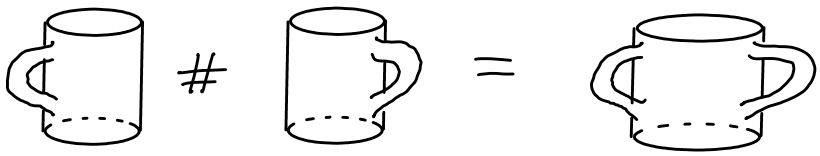
→ allows self intersection, which an embedding does not do.



We get $\mathbb{K}^2$ also by gluing together two Möbius strips along their boundary circles. Or, cut two discs out of $\mathbb{S}^2$ and glue a Möbius strip each along the edges of both cuts.

Note that $\mathbb{S}^2$ and $\mathbb{T}^2$ are orientable manifolds, while the Möbius strip, $\mathbb{RP}^2$, and $\mathbb{K}^2$ are non-orientable manifolds (with or without boundary).

We can obtain more general 2-manifolds by "gluing" these basic shapes together. For example, we can connect two coffee cups to get one coffee cup with two handles!



We can do this kind of "gluing" to join manifolds in any dimension (as long as the manifolds being joined have the same dimension). This gluing is formally termed connected sum.

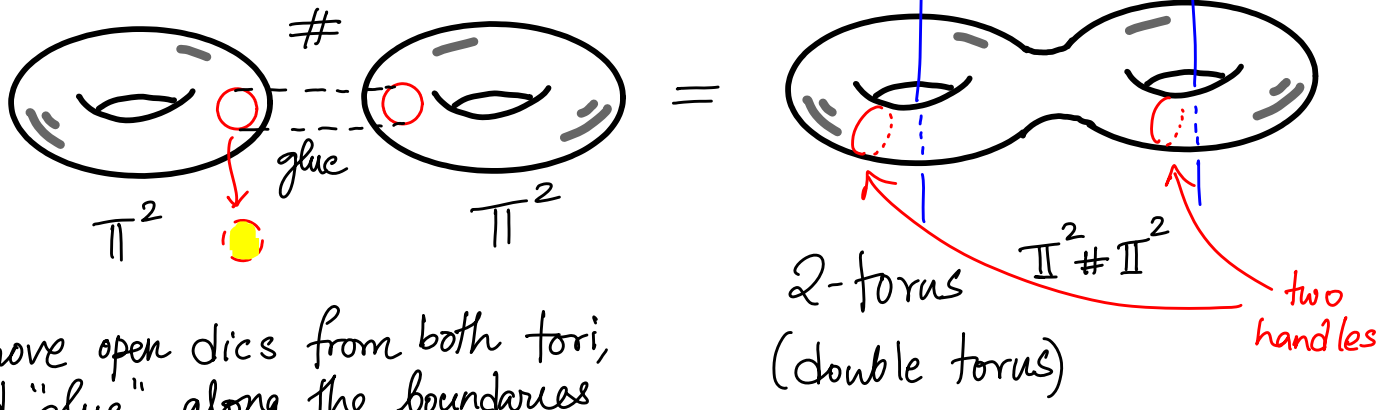
Def (connected sum). Let $\mathbb{M}_1, \mathbb{M}_2$ be d-manifolds. The **connected sum** of these d-manifolds is another d-manifold defined as follows.

$$\mathbb{M}_1 \# \mathbb{M}_2 = (\mathbb{M}_1 - D_1^d) \bigcup_{\partial\bar{D}_1^d \cong \partial\bar{D}_2^d} (\mathbb{M}_2 - D_2^d)$$

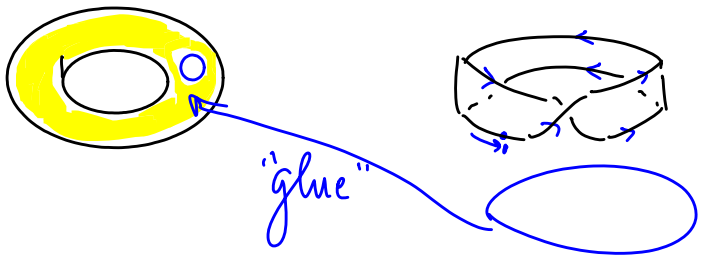
↳ identified by homeomorphism

D_1^d, D_2^d are d-dimensional open discs in $\mathbb{M}_1, \mathbb{M}_2$, respectively.

Here is an illustration:



Remove open dics from both tori, and "glue" along the boundaries of these circular holes.



Illustrating how to glue a Möbius strip to a hole in a torus.

MATH 529 – Lecture 5 (01/23/2024)

Today:
* classification of 2-manifolds
* simplices and simplicial complexes
* abstract simplicial complexes

Classification of compact, connected 2-manifolds

Result Every compact, connected 2-manifold is homeomorphic to S^2, or a connected sum of copies of $\mathbb{T}^2$, or a connected sum of copies of $\mathbb{RP}^2$.
If a 2-manifold is not connected, each component has this structure.

Examples

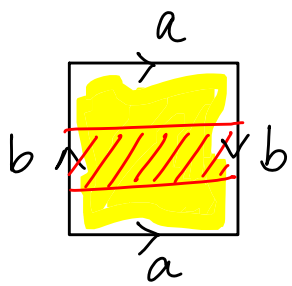
1. $S^2 \# \mathbb{RP}^2 \approx \mathbb{RP}^2$

(you just close back the open disc cut out from M!)

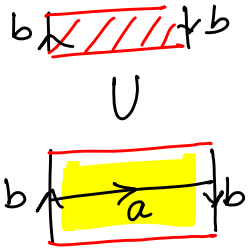
In fact, $S^2 \# M \approx M$ for any 2-manifold M.

2. $\mathbb{RP}^2 \# \mathbb{RP}^2 \approx \mathbb{K}^2$ Here's a "proof" by pictures:

$\mathbb{K}^2$:

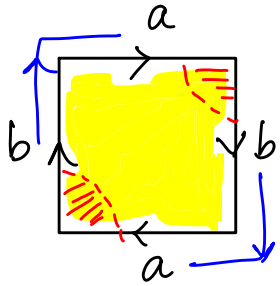


=

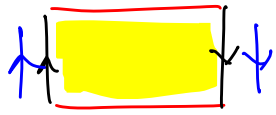


connected sum of 2 Möbius strips!

And $\mathbb{RP}^2 - D^2$:
↑ open 2-disc



→



Möbius strip

So, $\mathbb{K}^2$ is $\approx$ connected sum of 2 $\mathbb{RP}^2$'s.

Also, we get $\mathbb{RP}^2$ as $S^2 \# M$, for a Möbius strip M.

3. $\mathbb{T}^2 \# \mathbb{RP}^2 \approx \#(\mathbb{RP}^2)^3$

($\mathbb{T}^2$: orientable; $\mathbb{RP}^2$: nonorientable)

Once you join a non-orientable 2-manifold, the result stays non-orientable.

As one would expect, the corresponding result for 3-manifolds is much more complicated. Thurston's geometrization conjecture states that all compact 3-manifolds can be canonically decomposed into submanifolds that have geometric structure. This conjecture implies the famous Poincaré conjecture, which states that every compact, simply connected 3-manifold is homeomorphic to S^3, the 3-sphere.

(Informally, an object is simply connected if there are no "holes" passing through the object).

Perelman presented a proof of the Poincaré conjecture, almost 100 years after it was originally proposed (in 1904; Perelman's proof appeared in 2003). The corresponding result for n-manifolds with $n \geq 4$ turns out to be easier to prove, informally because of the "increased geometric freedom" one can afford in higher dimensions.

The main concepts used in Perelman's proof can be used to provide a proof for Thurston's geometrization conjecture (Perelman presented such a proof in 2003, along with his proof of the Poincaré conjecture).

Simplices

While we can study sample 2-manifolds as is, we cannot do computations on them. For this purpose, we need a discretized version of the spaces in question, which could be stored and handled naturally by a computer. Can we, for instance, use some sort of "counting arguments" to distinguish $\mathbb{S}^2$ from $\mathbb{T}^2$?

We introduce the concept of simplicial complexes in this context, and use concepts from combinatorial algebraic topology. The idea is that we can handle the combinatorics using efficient algorithms. Similarly, there are efficient data structures that can be used to work with simplicial complexes modeling the spaces. We also want to separate the topology from the geometry of the object. → we use the geometry in many applications.

We first introduce some definitions.

Def (combinations) Let $S = \{\bar{p}_0, \dots, \bar{p}_k\} \subseteq \mathbb{R}^d$. A

linear combination of $\bar{p}_i$ is $\bar{x} = \sum_{i=0}^{k} \lambda_i \bar{p}_i$, $\lambda_i \in \mathbb{R}\ \forall i$.

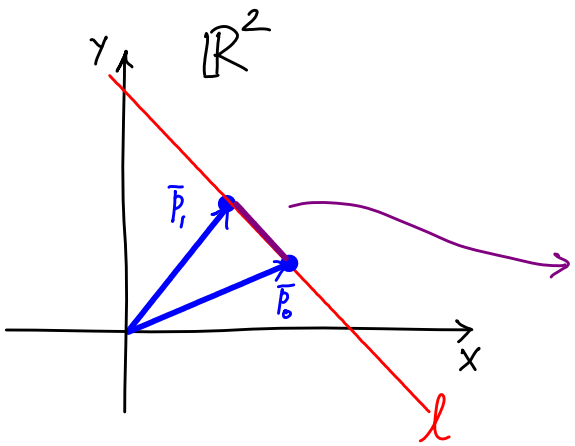
If $\sum_{i=0}^{k} \lambda_i = 1$, then $\bar{x}$ is an **affine combination** of $\bar{p}_i$'s.

In addition, if $\lambda_i \geq 0\ \forall i$, $\bar{x}$ is a **convex combination** of $\bar{p}_i$'s.

The set of all convex combinations of elements in S is the **convex hull** of S, denoted as

$$\text{conv}(S) = \left\{ \sum_{i=0}^{k} \lambda_i \bar{p}_i \;\middle|\; \lambda_i \geq 0\ \forall i,\ \sum_{i=0}^{k} \lambda_i = 1 \right\}.$$

Illustration in 2D



set of all linear combinations $= \mathbb{R}^2$

set of all affine combinations $= \ell$, the line through $\bar{P}_0, \bar{P}_1$

$\text{conv}(\{\bar{P}_0, \bar{P}_1\})$ = line segment connecting $\bar{P}_0, \bar{P}_1$.

$\bar{P}_0, \bar{P}_1$ are not parallel here.

Def (Independence) S with $|S| \geq 2$ is **linearly (affinely) independent** if no point in S is a linear (affine) combination of other points in S.

We denote linearly independent in short as LI, and affinely independent as AI.

$|S|=1$ case: $\{\bar{P}_0\}$ is LI if $\bar{P}_0 \neq \bar{0}$ (zero vector), but $\{\bar{P}_0\}$ is AI for all $\bar{P}_0$ (even if $\bar{P}_0 = \bar{0}$).

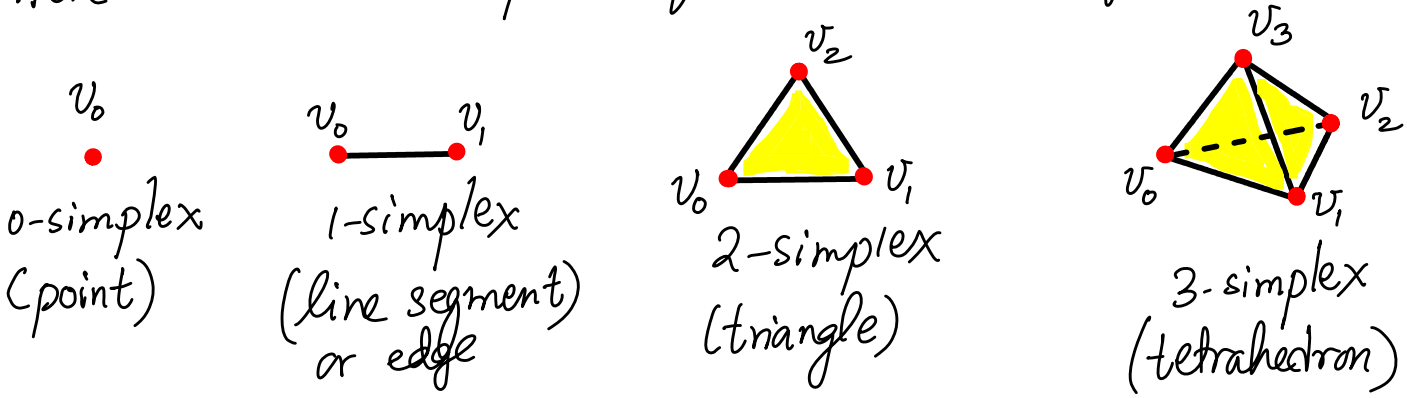
For example, 3 points in $\mathbb{R}^2$ are AI as long as they are not collinear.

We will use convex hulls of AI points as our building blocks — called simplices.

Def (simplex) The convex hull of $(k+1)$ affinely independent points $S = \{\bar{v}_0, \dots, \bar{v}_k\}$ is a **k-simplex**. The **dimension** of the simplex is k, and $\bar{v}_j$'s are the **vertices** of the k-simplex.

(vertices, and hence $\bar{v}_i$, rather than $\bar{P}_i$.)

Here are the k-simplices for small values of k: $0 \le k \le 3$.



Notice that the k-simplex is homeomorphic to the k-ball, i.e., $B_k = \{\bar{x} \in \mathbb{R}^k \mid \|\bar{x}\| \le 1\}$. "=" gives S^{k-1}, the $(k-1)$-sphere.

Indeed, the boundary of the k-ball is the $(k-1)$-sphere, e.g., 2-ball (or 2-disc) has the circle (1-sphere) as the boundary.

Each p-simplex is made of lower dimensional simplices, i.e., k-simplices with $k \le p$. Thus, $\triangle v_0 v_1 v_2$ contains vertices v_0, v_1, v_2, edges $\overline{v_0 v_1}$, $\overline{v_1 v_2}$, $\overline{v_0 v_2}$, and $\triangle v_0 v_1 v_2$ itself.

Def (face/coface). Let σ be the k-simplex defined on $S = \{\bar{v}_0, \bar{v}_1, \ldots, \bar{v}_k\}$. A simplex τ defined on a subset T of S, $|T| \neq \emptyset$, is a **face** of σ, and σ is a **coface** of τ. The notation is $\tau \leq \sigma$, $\sigma \geq \tau$ (some books use $\tau \preceq \sigma$, $\sigma \succeq \tau$ → \succeq, \preceq in LaTeX.

Thus, $\overline{v_0 v_1}$ is a face of $\Delta v_0 v_1 v_2$. So are $v_0, v_1, v_2, \overline{v_1 v_2}$, and $\overline{v_0 v_2}$.

A simplex is always a face of itself, i.e., $\sigma \leq \sigma$ and $\sigma \geq \sigma$.

We can attach simplices together "nicely" to form bigger objects called simplicial complexes.

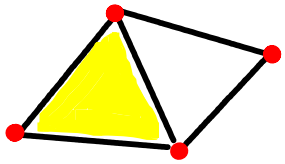
Def A **simplicial complex** K is a set of simplices such that

1. $\sigma \in K$, $\tau \leq \sigma \Rightarrow \tau \in K$; and — Every face of a simplex in K is also in K.
2. $\sigma, \sigma' \in K \Rightarrow \sigma \cap \sigma' \leq \sigma, \sigma'$ when $\sigma \cap \sigma' \neq \emptyset$. — In particular, the non-empty intersection of two simplices in K is a face of both of them, and hence in K as well.

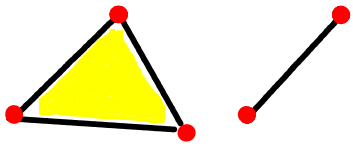
The above definition holds in the case of both finite and nonfinite K. In the latter case, K has infinitely many simplices satisfying the two conditions. But we will usually limit our attention in this course to finite simplicial complexes, unless mentioned otherwise.

Note that a simplex is the convex hull of a **finite** number of affinely independent points. So, we never talk about "infinite-dimensional" simplices. We could talk about each of these points sitting in infinite-dimensional space, but we restrict our attention to $\mathbb{R}^d$ for finite d.

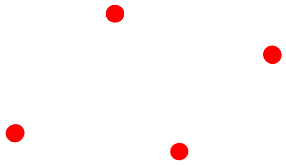
Examples Here are some simplicial complexes:



K_1



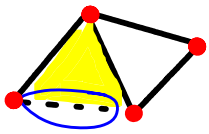
K_2



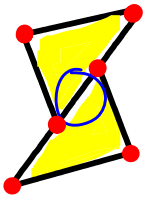
K_3

In particular, notice that a simplicial complex need not consist of just one connected component.

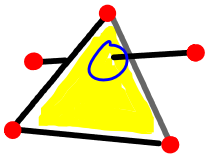
Here are some collections that are not simplicial complexes:



K_4



K_5



K_6

K_4 violates Condition 1, as one of the faces of the triangle in K_4 is not in the collection. K_5 and K_6 violate Condition 2, as the intersection of two simplices is not a face of either simplex in both cases.

Def The dimension of a simplicial complex is the same as that of the highest dimensional simplex in it, i.e.,

$$\dim K = \max \{\dim \sigma \mid \sigma \in K\}.$$

In the previous examples, $\dim(K_1) = 2$, $\dim(K_2) = 2$, and $\dim(K_3) = 0$.

Earlier we have been talking about continuous surfaces, e.g., the 2-sphere, torus, etc. And now we are talking about simplicial complexes as discrete objects. How do we "reconcile" the two notions? Indeed, we can formally define the "space" modeled by a simplicial complex. Later on, we will talk about a simplicial complex "triangulating", say, a torus, when this "space" is homeomorphic to $\mathbb{T}^2$.

Def The **underlying space** of a simplicial complex K the space made of all simplices in K together with the topology inherited from the ambient Euclidean space. We denote the underlying space of the simplicial complex K by $|K|$. Thus,

$$|K| = \bigcup_{\sigma \in K} \sigma.$$

$|K|$ is also called the polyhedron (or polytope) of K.

$A \subseteq |K|$ is closed in $|K|$ iff $A \cap \sigma$ is closed $\forall \sigma \in K$.

We can define simplicial complexes abstractly – simplices need not necessarily be sitting in $\mathbb{R}^d$ for any d. Indeed, this notion conveys the power and versatility of the concept!

Def An **abstract simplicial complex** (ASC) is a collection $\mathcal{S}$ of finite non-empty sets such that if $A \in \mathcal{S}$, and $B \subseteq A$ with $B \neq \emptyset$, then $B \in \mathcal{S}$.

Note that the condition specified in the above definition of an abstract simplicial complex is equivalent to the first condition in the definition of a (regular) simplicial complex, which says that every face of a simplex in the complex is also in the complex.

The second intersection condition is trivially satisfied in the case of abstract simplicial complexes. The intersection of two sets is indeed a subset of both sets. $\mathcal{S}$ itself can be finite or infinite, but each $A \subseteq \mathcal{S}$ is a finite set.

The sets in $\mathcal{S}$ are called the **simplices** of $\mathcal{S}$. The dimension of a simplex $A \in \mathcal{S}$ is $\dim(A) = |A| - 1$.

↳ cardinality (# entries) of A

Note the correspondence of the above definition to the definition of simplices in the usual sense. Recall that a k-simplex is the convex hull of $(k+1)$ affinely independent points, which are its vertices. We maintain this correspondence by defining $\dim A = |A| - 1$ for any set $A \in \mathcal{S}$.

The **dimension** of $\mathcal{S}$ is $\dim(\mathcal{S}) = \max\{\dim(A) \mid A \in \mathcal{S}\}$.

In the definition of $\mathcal{S}$, we do assume that all $A \in \mathcal{S}$ are finite sets. And there exists a maximum dimensional simplex in $\mathcal{S}$.

→ which is finite

The singleton sets in $\mathcal{S}$ are called its **vertices**, and is denoted by $\mathrm{Vert}(\mathcal{S})$.

Again note the correspondence of these singleton sets to the vertices (0-simplices) in (geometric) simplicial complexes.

Here is an example of an abstract simplicial complex.

$$\mathcal{S} = \{\{0\}, \{1\}, \{2\}, \{3\}, \{0,1\}, \{0,2\}, \{1,2\}, \{1,3\}, \{2,3\}, \{0,1,2\}\}.$$

$\{0\}, \{1\}, \{2\}, \{3\}$: vertices

We can indeed check that the condition on inclusion of subsets is satisfied. For instance, consider the set $\{0,1,2\}$.

$\{0,1,2\} \in \mathcal{S} \Rightarrow$ need $\{0,1,2\}$, $\{0,1\}, \{0,2\}, \{1,2\}, \{0\}, \{1\}, \{2\} \in \mathcal{S}$.

$\{0,1,2\}$: trivial

$\{0,1\}, \{0,2\}, \{1,2\}, \{0\}, \{1\}, \{2\} \in \mathcal{S}$: indeed!

$\dim \mathcal{S} = 2$.

MATH 529 – Lecture 6 (01/25/2024)

Today:
* geometric realization of ASCs
* Comparing ASCs
* topological invariants

Recall: ASC $\mathcal{S}$; $A \in \mathcal{S}$, $B \subset A$, $B \neq \emptyset \Rightarrow B \in \mathcal{S}$.

The singleton sets in $\mathcal{S}$ are called its **vertices**.

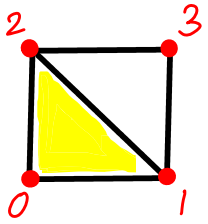
Given any (geometric) simplicial complex K, we can create an abstract simplicial complex $\mathcal{S}$ by taking just the sets of vertices in each simplex of K (and ignoring the geometry). $\mathcal{S}$ here is called the **vertex scheme** of K.

Symmetrically, K is a **geometric realization** of $\mathcal{S}$.

↳ there could be other geometric realizations

$$\mathcal{S} = \{ \underbrace{\{0\}, \{1\}, \{2\}, \{3\}}_{\text{vertices}}, \{0,1\}, \{0,2\}, \{1,2\}, \{1,3\}, \{2,3\}, \{0,1,2\} \}.$$

K



K can sit in $\mathbb{R}^2$

We have a vertex corresponding to each singleton set (i.e., vertex) in $\mathcal{S}$, an edge corresponding to each doublet, a triangle for each triplet, etc.

But this is just one geometric realization. In particular, if we were to specify the (x,y) coordinates of the vertices, we can imagine other realizations, e.g., by translating this one. We could also have a realization in $\mathbb{R}^3$, for instance.

Theorem (Geometric realization theorem) Every abstract simplicial complex $\mathcal{S}$ with $\dim \mathcal{S} = d$ has a geometric realization in $\mathbb{R}^{2d+1}$.

Idea of the Proof

We map vertices of $\mathcal{S}$ injectively to points in $\mathbb{R}^{2d+1}$, say, $f: \text{Vert}(\mathcal{S}) \to \mathbb{R}^{2d+1}$. ($\text{Vert}(\mathcal{S})$ → set of vertices in $\mathcal{S}$) Why $2d+1$? We use the fact that $2d+2$ or fewer points in $\mathbb{R}^{2d+1}$ that are in **general position** are affinely independent (AI).

Def $(d+1)$ points in $\mathbb{R}^d$ are in **general position** if no hyperplane contains more than d of those points.

The idea is that the points do not satisfy any more linear relationships than they must. For instance 3 points in $\mathbb{R}^2$ that are not collinear are in general position.

Recall that a d-simplex is the convex hull of $(d+1)$ AI points. We need to make sure that we will have "enough freedom", i.e., affine independence, among the mapped vertices so that we can map all the simplices in $\mathcal{S}$ to corresponding simplices in the geometric simplicial complex.

Consider $A, B \in \mathcal{S}$. Since $\dim \mathcal{S} = d$, $|A|, |B| \leq d+1$.

Hence, $|A \cup B| = |A| + |B| - |A \cap B| \leq d+1 + d+1 = 2d+2$.

Hence by going to $\mathbb{R}^{2d+1}$ and choosing points there in general position, we can ensure that (up to) $2d+2$ points are AI.

$\Rightarrow$ Any convex combination $\bar{x}$ in $A \cup B$ is unique.

$\Rightarrow$ $\bar{x} \in A$ and $\bar{x} \in B \iff \bar{x} \in A \cap B$.

↳ ensures the second requirement of nonempty intersections of two simplices being their faces.

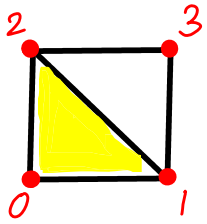
We could often find geometric realizations in $\mathbb{R}^{d'}$ for d' smaller than $2d+1$.

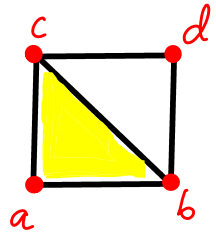
□

How do we compare abstract simplicial complexes? We had previously defined the concept of homeomorphism to study when two topological spaces are "similar". We now define corresponding notions for simplicial complexes. Recall that $\mathrm{Vert}(\mathcal{S})$ represents the vertex set of the ASC $\mathcal{S}$.

Def Two abstract simplicial complexes $\mathcal{S}_1$ and $\mathcal{S}_2$ are **isomorphic** if there exists a bijection $\varphi: \mathrm{Vert}(\mathcal{S}_1) \to \mathrm{Vert}(\mathcal{S}_2)$ such that $A \in \mathcal{S}_1$ iff $\varphi(A) \in \mathcal{S}_2$. φ is an isomorphism between $\mathcal{S}_1$ and $\mathcal{S}_2$. We write $\mathcal{S}_1 \approx \mathcal{S}_2$ here. (A: simplex)

In this setting, every simplex in $\mathcal{S}_1$ has a unique corresponding simplex in $\mathcal{S}_2$.

K_1 

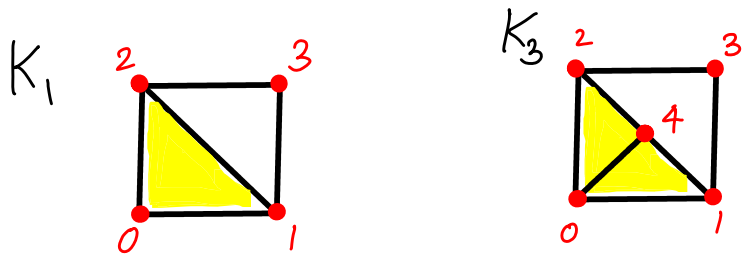
K_2 

Vertex schemes of K_1 and K_2 are isomorphic

Notice that the similarity here is defined for **abstract** simplicial complexes. Hence, K_1 and K_2 above need not be sitting in the same space. Still, they are isomorphic as ASCs. We make this notion precise in the following theorem.

Theorem Two (geometric) simplicial complexes K_1 and K_2 are isomorphic, or **simplicially homeomorphic**, iff their vertex schemes $\mathcal{S}_1$ and $\mathcal{S}_2$ are isomorphic as abstract simplicial complexes. We denote this fact by $K_1 \cong K_2$, which implies $|K_1| \approx |K_2|$, and $\mathcal{S}_1 \approx \mathcal{S}_2$.

The implication might not go the other way, though. For instance, $K_1 \cong K_2$ above. Now consider K_3 as shown below.



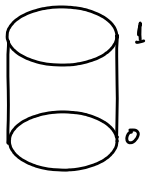
Notice that $K_1 \not\cong K_3$, even though $|K_1| \approx |K_3|$. In fact, their underlying spaces could very well be identical!

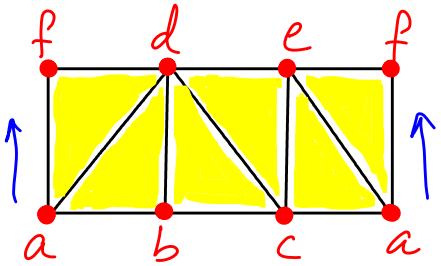
From the computational point of view, while simplicial complexes that are "smaller", i.e., have a smaller number of simplices while modeling the same topological space are usually preferred. At the same time, geometry might dictate that we need a large number of simplices to capture the complexity.

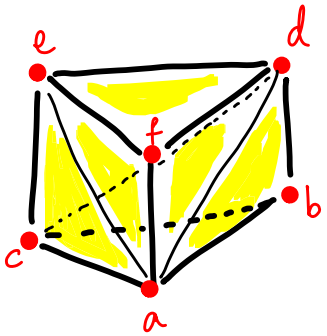
How do we use ASCs? We illustrate several examples

1. cylinder $\mathbb{S}^1 \times [0,1]$

geometric representation!







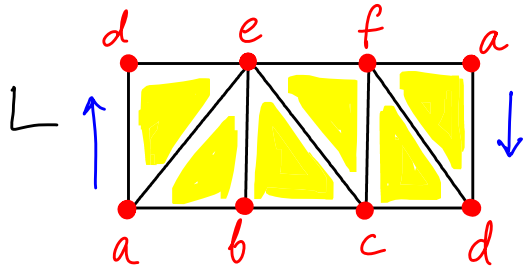
K

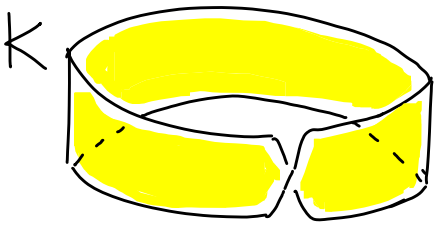
The underlying space of L appears to be rectangle, but notice how the vertex labels identify the left and right edges (both are $\overline{af}$).

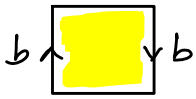
$L = \{\{a,b,d\}, \{a,d,f\}, \{b,c,d\}, \{c,d,e,\}, \{a,c,e\}, \{a,e,f\},$ and all nonempty subsets $\}$

K is one geometric realization of L here.

2. Möbius strip — We start with the ASC L here







The left and right vertical edges are identified, after a "twist". Like in Example 1, the underlying space is a rectangle, but vertex labels are different.

L represents the Möbius strip, i.e., K is a geometric realization of L.

We could also specify L abstractly:

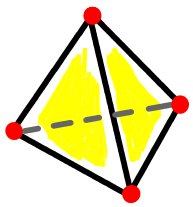
$L = \{ \{a,d,e\}, \{a,b,e\}, \{b,c,e\}, \{c,e,f\}, \{c,d,f\}, \{a,d,f\},$ and all nonempty subsets of these triplets $\}$.

The abstract simplicial complexes shown above consist of triangles - indeed they are triangulations. In general, triangulations consist of triangles (in 2D), and simplices in general as we formalize below.

Notice that each k-simplex $\approx$ k-ball (closed). 2-ball is the closed 2-disc.

Def (Triangulation). A **triangulation** of a topological space $\mathbb{X}$ is a simplicial complex K such that $|K| \approx \mathbb{X}$.

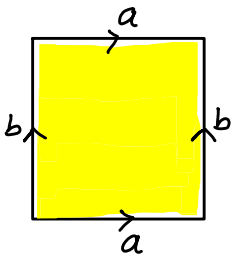
Example:



surface of a tetrahedron (triangles and their faces) is a triangulation of the 2-sphere S^2.

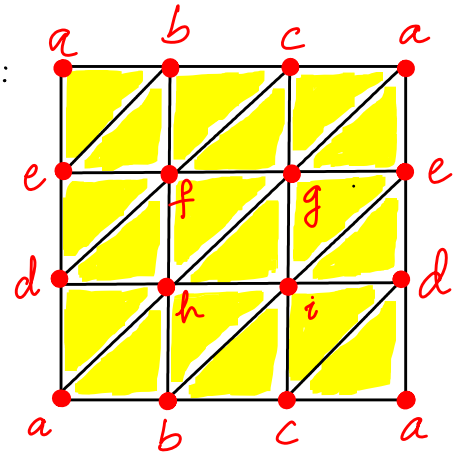
A triangulation is a piecewise linear representation of the topological space.

3. Torus ($\mathbb{T}^2$)



→ this diagram is not a triangulation!

K:

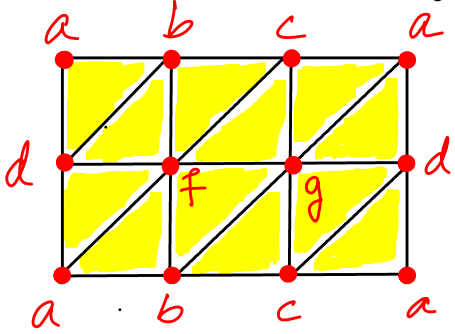


Why so many triangles?

This triangulation K of the torus has 18 (different) triangles. One might wonder if we could produce a triangulation using a much smaller number of triangles.

Consider the following candidate triangulation K':

K'



Is K' a triangulation of $\mathbb{T}^2$? No!

For instance, consider edge $\overline{ad}$. It is a face of four triangles: adb, adf, adc, adg.

Hence, points on $\overline{ad}$ do not have neighborhoods homeomorphic to $\mathbb{R}^2$. We're doing "too much gluing" here.

Q. What is the minimum number of triangles needed to produce a triangulation of $\mathbb{T}^2$?

A. We need at least 14 triangles.

Rule: In a triangulation of a 2-manifold (with boundary), each edge must be part of (one or) two triangles.

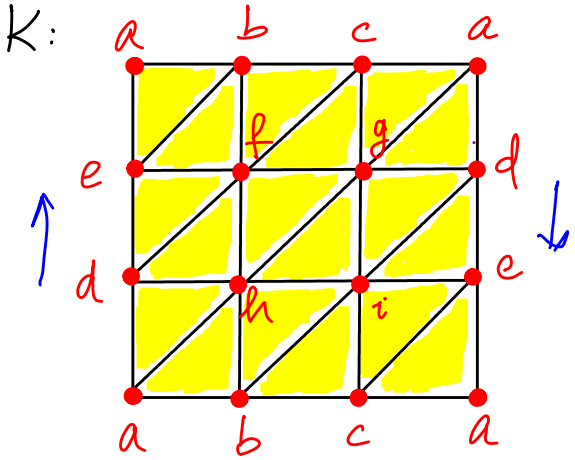
Edges that are part of only one triangle each form the boundary of the 2-manifold.

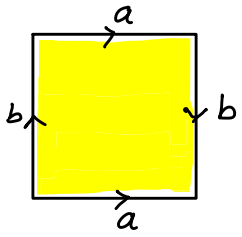
Recall that every point on a 2-manifold has a neighborhood homeomorphic to an open disc. The corresponding requirement for edges becomes that each edge has to be part of exactly two triangles. Similarly, points on the boundary of a 2-manifold have neighborhoods homeomorphic to half discs. Edges that are part of single triangles are indeed the boundary edges in the triangulation. The result extends to d-dimensions — every (d-1)-simplex is the face of one or two d-simplices.

The above rule could be used to check if a given simplicial complex is the triangulation of a manifold or not. But, be warned that satisfying this rule **alone** is not enough to identify the given simplicial complex as the triangulation of a specific manifold, e.g., the torus.

4. Klein bottle ($\mathbb{K}^2$)

K:





Same rule applies here — each edge is shared by two triangles exactly.

For example, edge $\overline{ae}$ is part of triangles $\triangle abe$ and $\triangle ace$.

We talked about distinguishing spaces that are not homeomorphic. In practice, spaces or objects are usually represented by triangulations. Checking for homeomorphisms between triangulations is not easy. How can we distinguish two topological spaces computationally? One option is to use an invariant.

Topological Invariants

Def A **topological invariant** is a map that assigns the same object to spaces of the same topological type.

↳ usually, a number; but we could also have a "barcode", for instance

Let $f(\cdot)$ be an invariant.

$$\mathbb{X} \approx \mathbb{Y} \Rightarrow f(\mathbb{X}) = f(\mathbb{Y}).$$

So, $f(\mathbb{X}) \neq f(\mathbb{Y}) \Rightarrow \mathbb{X} \not\approx \mathbb{Y}$ → could be used for contrapositive arguments.

But $f(\mathbb{X}) = f(\mathbb{Y})$ does not necessarily mean $\mathbb{X} \approx \mathbb{Y}$.

If $f(\mathbb{X}) = f(\mathbb{Y}) \Rightarrow \mathbb{X} \approx \mathbb{Y}$, then $f(\cdot)$ is called a **complete invariant**.

Notice that an invariant could assign the same object to spaces of different topological types. The main way to use the invariant is in the contrapositive, i.e., if the invariant is different for a pair of spaces, then the two spaces have different topological types.

MATH 529 – Lecture 7 (01/30/2024)

Today:
* Euler characteristic (χ)
* χ + orientability: complete invariant
* genus, cross cap

Recall: topological invariant $f(\cdot)$: $\mathbb{X} \approx \mathbb{Y} \Rightarrow f(\mathbb{X}) = f(\mathbb{Y})$ —— (*)

complete invariant: (*) & $f(\mathbb{X}) = f(\mathbb{Y}) \Rightarrow \mathbb{X} \approx \mathbb{Y}$.

The first invariant we will study is the Euler characteristic, which is not a complete invariant by itself. We will add orientability to get a complete invariant.

The Euler characteristic (χ) ("chi")

(originally defined for graphs)

Let K be a simplicial complex, and let s_i be the # i-simplices in K for $0 \le i \le \dim(K)$. Thus,

$$s_i = \left| \{ \sigma \in K \mid \dim \sigma = i \} \right|.$$

The **Euler characteristic** of K is defined as

$$\chi(K) = \sum_{i=0}^{\dim K} (-1)^i s_i.$$

Notice that $s_i = 0 \ \forall \ i > \dim K$.

Equivalently, $\chi(K) = v - e + f - t + \cdots$, where v = # vertices, e = # edges, f = # faces (→ triangles), t = # tetrahedra, ... and so on.

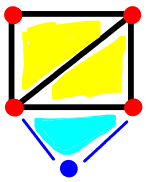
Let us find χ for a triangulation of the closed 2-disc.

χ () $= 3 - 3 + 1 = 1$

3: # vertices; 3: # edges; 1: # faces (or # triangles)

χ is an integer invariant, and it is an **invariant of the underlying space** $|K|$. So, χ is invariant over triangulations of a given space. Thus, any triangulation of a topological space $\mathbb{X}$ has the same $\chi(\mathbb{X})$ value.

Continuing with the example of the disc, we get the same χ using any other triangulation — see two examples below.

χ () $= 4 - 5 + 2 = 1.$

The triangulation is made of two triangles sharing an edge.

Now consider adding one more triangle to get another valid triangulation (as shown in blue).

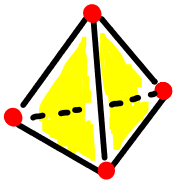
$\Delta(V) = 1, \Delta(E) = 2, \Delta(F) = 1$, so $\Delta(\chi) = 1 - 2 + 1 = 0!$

$\Delta(V)$: change in # vertices, edges, triangles.

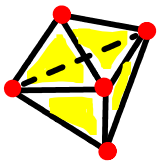
$\Delta(\chi)$: change in χ

Q: Can we use χ to distinguish compact 2-manifolds?

Let us find χ for $\mathbb{S}^2$, $\mathbb{T}^2$, $\mathbb{RP}^2$ and $\mathbb{K}^2$, 2-sphere, torus, projective plane, and the Klein bottle.

1. $\mathbb{S}^2$ K:  Surface of a tetrahedron is a triangulation of $\mathbb{S}^2$.

$\chi(K) = 4-6+4 = 2.$

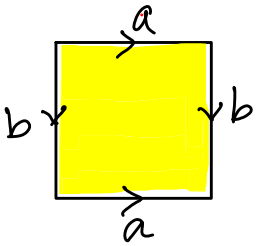
K': 

Let us consider another triangulation K' of $\mathbb{S}^2$, made of 3 triangles from top and 3 from bottom joined to form a "sphere".

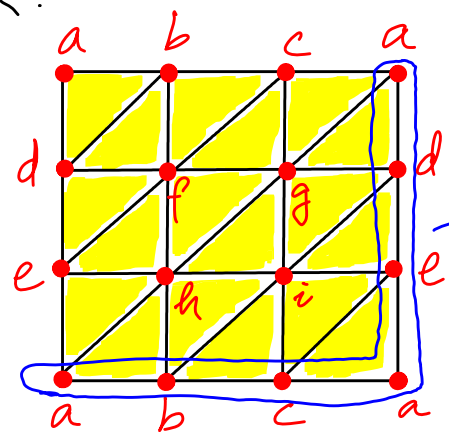
$\chi(K') = 5-9+6 = 2.$

Indeed, $\chi(\mathbb{S}^2) = 2.$ One could take any triangulation of $\mathbb{S}^2$, χ will be the same.

2. $\mathbb{T}^2$ (torus)



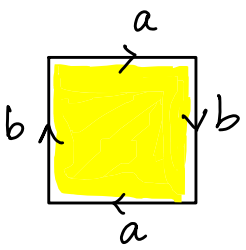
K:

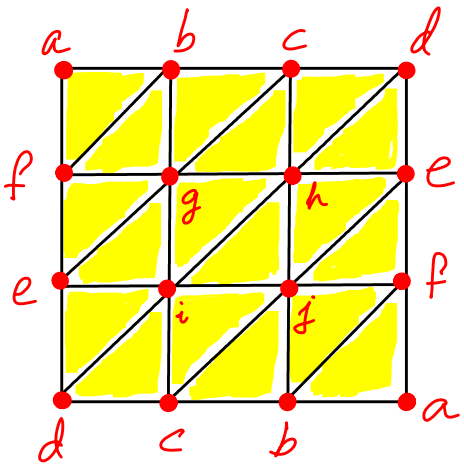


→ should not double count!

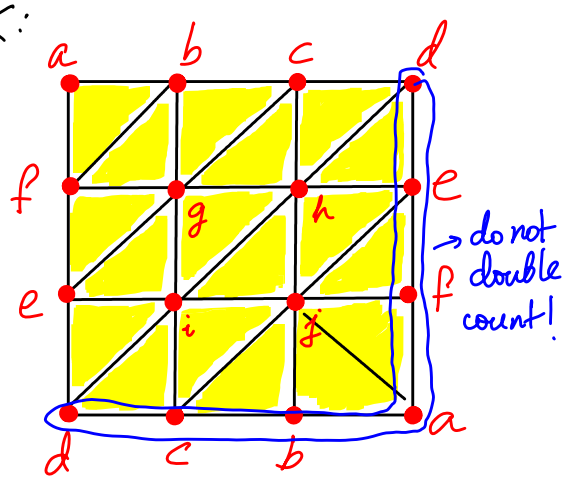
$\chi(\mathbb{T}^2) = 9-27+18 = 0.$

3. $\mathbb{RP}^2$





K:



$\overline{bf} \in abf, jbf, gbf$! ✗

$\overline{ab} \in abf$! ✗

$\overline{ab} \in abj, abf$ ✓

$\overline{bf} \in bfa, bfg$ ✓

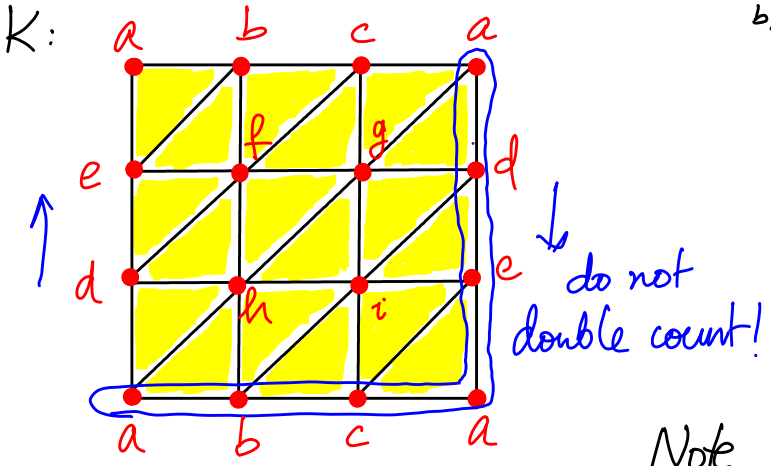
Our first attempt at finding a triangulation (left) of $\mathbb{RP}^2$ is not correct!
In particular, edge $\overline{bf}$ (bottom right square) is part of 3 $\triangle$es: abf, bfg, bfj!
Also, $\overline{ab}$ and $\overline{af}$ are part of only one triangle each. But $\mathbb{RP}^2$ has no boundary!
A correct triangulation is given on the right.

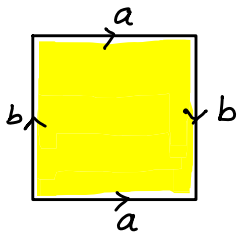
→ the left triangulation represents $\mathbb{RP}^2$ with a "flap" ($\triangle abf$)

$$\chi(K) = \underbrace{10}_{\text{vertices a-j}} - 27 + 18 = 1.$$

4. $\mathbb{K}^2$ (Klein bottle)

K:





$$\chi(K) = 9 - 27 + 18 = 0.$$

Note that $\chi(\mathbb{K}^2) = \chi(\mathbb{T}^2)$!

Here is the summary of the χ values we have seen so far.

	2-manifold	χ
orientable	S^2	2
orientable	$\mathbb{T}^2$	0
nonorientable	$\mathbb{RP}^2$	1
nonorientable	$\mathbb{K}^2$	0

So, χ alone is not sufficient to distinguish between all these surfaces!

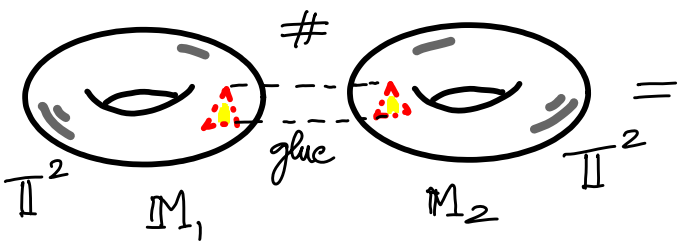
It turns out that if we add orientability to χ, we do get a complete invariant for all compact (connected) 2-manifolds (without boundary). Recall the original classification theorem, which states that every compact connected 2-manifold is homeomorphic to S^2, a connected sum of copies of $\mathbb{T}^2$, or a connected sum of copies of $\mathbb{RP}^2$. With this result in mind, let us first study how χ changes when we take the connected sum of two manifolds.

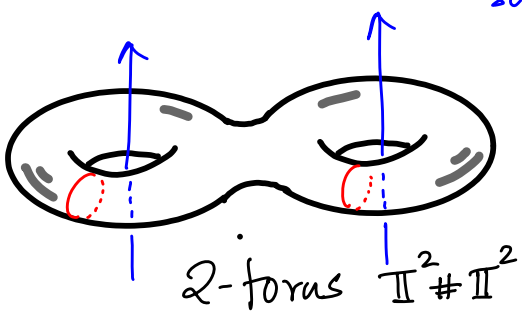
Theorem For compact, connected surfaces (2-manifolds) $\mathbb{M}_1$ and $\mathbb{M}_2$,

$$\chi(\mathbb{M}_1 \# \mathbb{M}_2) = \chi(\mathbb{M}_1) + \chi(\mathbb{M}_2) - 2.$$

← take one Δle out of each surface.

Illustration



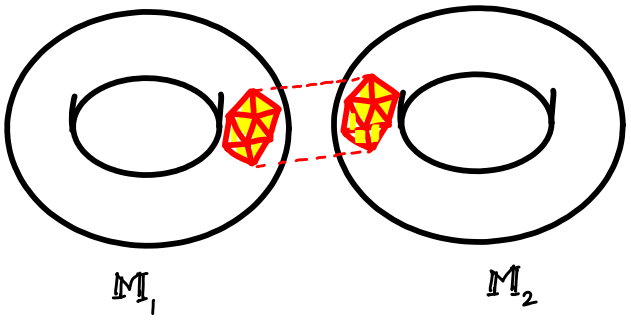


double torus
two tunnels (↑↑)
and two handles
C: C:

We remove a triangle each from both $\mathbb{M}_1$ and $\mathbb{M}_2$, and glue along the boundaries of these triangles.

$\Delta(V) = -3$, $\Delta(E) = -3$, $\Delta(F) = -2$. So, $\Delta(\chi) = -3 - (-3) + (-2) = -2$.

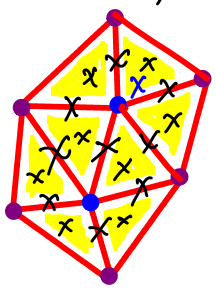
The result holds for the removal of a disc in general, and not just for the case of (the removal of) a triangle.



Here, we remove a patch (homeomorphic to the disc) from both $\mathbb{M}_1$ and $\mathbb{M}_2$, and identify the boundaries, which is composed of 6 edges and 6 vertices (to form a loop).

From the middle regions of each patch, we remove 2 vertices, 9 edges, and 8 triangles.

The change in $\chi(\mathbb{M}_1 \# \mathbb{M}_2)$ contributed by the simplices removed from $\mathbb{M}_1$ is



simplices marked with an 'x' are removed, and so are the two middle vertices.

$-(2-9+8) = -1$. A same change is contributed by the simplices removed from $\mathbb{M}_2$.

There is no change in χ from identifying the boundaries — as they are both cycles (so have same # of vertices & edges).

As such, $\chi(\mathbb{M}_1 \# \mathbb{M}_2) = \chi(\mathbb{M}_1) + \chi(\mathbb{M}_2) - 2$.

We get the same result even if we were to remove different "discs" from the two tori. Just that the homeomorphism defining the gluing would be more complicated there.

We could prove this result in general (for a removal of a general disk).

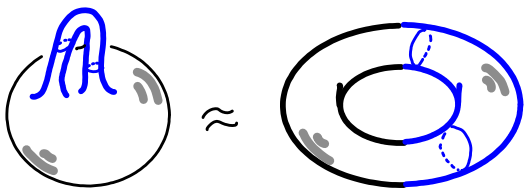
Theorem Two compact, closed, connected 2-manifolds $\mathbb{M}_1$ and $\mathbb{M}_2$ are homeomorphic **if and only if**

1. $\chi(\mathbb{M}_1) = \chi(\mathbb{M}_2)$ and
2. either $\mathbb{M}_1$ and $\mathbb{M}_2$ are both orientable, or are both nonorientable.

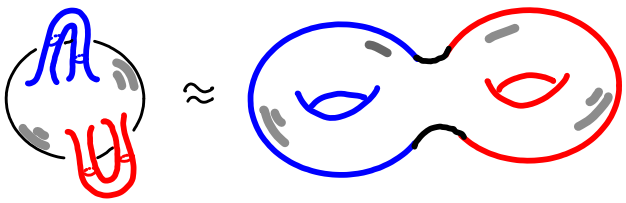
We could perform both checks 1 and 2 efficiently (→ in polynomial-time, to be precise) on a computer.

Genus and Cross-cap ⟶ Two more terms used in the context of 2-manifolds.

Def The connected sum of g tori is called a surface with **genus** g. Equivalently, a 2-sphere with 1 tube is a surface with genus $g = 1$.

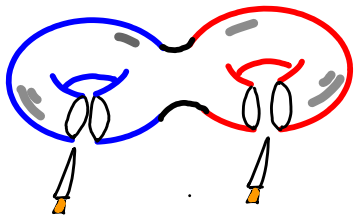


Sphere with one tube is homeomorphic to torus. (so, torus has genus = 1).



A sphere with two tubes is homeomorphic to the double torus.

$\mathbb{M}$ has genus $g \Rightarrow$ there are g disjoint closed curves on $\mathbb{M}$ along which you can cut without disconnecting $\mathbb{M}$.



$g = 2$ here. If we cut along one more closed curve now, we get two pieces that are disconnected.

Euler characteristic and genus

Recall $\chi(M_1 \# M_2) = \chi(M_1) + \chi(M_2) - 2$.

$$\chi(g\mathbb{T}^2) = 2 - 2g.$$

$g\mathbb{T}^2$: connected sum of g tori

We could easily prove this result using induction, using the above fact about $\chi(M_1 \# M_2)$, and $\chi(\mathbb{T}^2) = 0$.

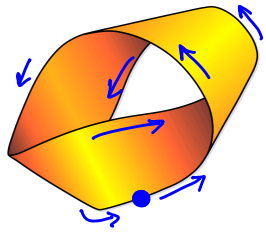
Recall, $\chi(\mathbb{S}^2) = 2$

Also, $\chi(\#(\mathbb{T}^2)^g)$

Cross cap

Recall that the Möbius strip has only one edge, i.e., its boundary is a single circle.

Starting from a point on the edge, we can traverse the entire boundary to come back to the same point (as shown by arrows).



(image: www)

If we remove an open disc from the 2-sphere, and glue a Möbius strip along its edge onto the boundary of this disc, we have added one **cross cap**.

A sphere with a single cross cap is homeomorphic to the real projective plane ($\mathbb{RP}^2$).
A sphere with two cross caps is homeomorphic to the Klein bottle ($\mathbb{K}^2$).

In general, a sphere with g cross caps is the connected sum of g projective planes, and we have

$$\chi(g\mathbb{RP}^2) = 2 - g.$$

also, sphere with g cross caps

MATH 529 – Lecture 8 (02/01/2024)

Today: * orientation of a simplex
* orienting surfaces
* ~~subdivision~~ didn't get to it :)!!

Before we discuss orientation, we present one last illustration of how we use the Euler characteristic and its properties.

We saw that $\chi(g\mathbb{RP}^2) = 2-g$ (g: # cross caps).

On the other hand, the classification theorem for compact connected 2-manifolds says that any non-orientable surface (2-manifold) is homeomorphic to the connected sum of copies of $\mathbb{RP}^2$, the projective plane. Recall that once we glue at least one cross cap, the surface becomes non-orientable. We could use the above result relating χ and the # cross caps and the result on how χ changes when we take the connected sum of two surfaces to identify the # copies of $\mathbb{RP}^2$ whose connected sum is homeomorphic to a given nonorientable surface.

Consider $\mathbb{T}^2 \# \mathbb{RP}^2$. We get

$$\chi(\mathbb{T}^2 \# \mathbb{RP}^2) = \chi(\mathbb{T}^2) + \chi(\mathbb{RP}^2) - 2 = 0 + 1 - 2 = -1.$$

To get the # cross caps in the homeomorphic surface, we set

$$\chi(g\mathbb{RP}^2) = 2 - g = -1 \implies g = 3.$$ Hence we should have

$$\mathbb{T}^2 \# \mathbb{RP}^2 \approx \#(\mathbb{RP}^2)^3,$$

as we stated in Lecture 5!

Moving on, we now consider orientations of simplices, and how to extend them to possibly orient entire simplicial complexes.

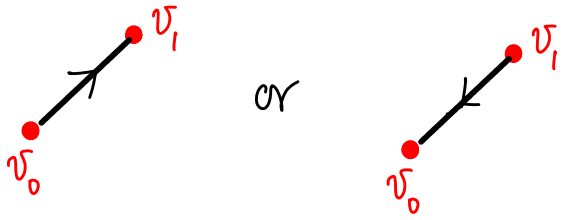
Def Let σ be a simplex (geometric or abstract). Two orderings of its vertices are equivalent if they differ by an even permutation. If $\dim(\sigma) > 0$, the orderings fall into **two** equivalence classes. Each equivalence class is an **orientation** of σ.
A 0-simplex has only one orientation.

An **oriented simplex** is a simplex σ with an orientation of σ.

An even permutation is obtained by doing an even number of pairwise swaps.

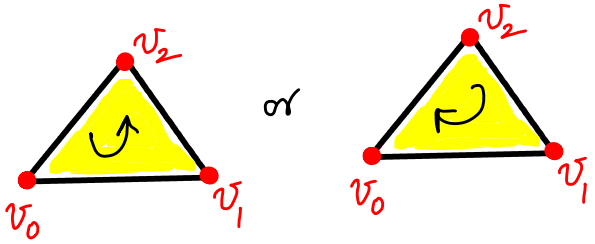
Notation For σ with vertices $\{v_0, \dots, v_k\}$, $\sigma = [v_0, \dots, v_k]$ denotes an oriented simplex. Note that we use σ to denote both the default and the oriented simplex.

Examples



1-simplex

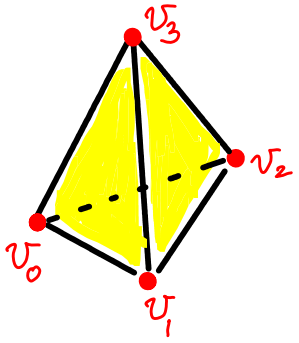
$[v_0 v_1]$ is opposite to $[v_1 v_0]$
(we can go forward or backward)



$[v_0 v_1 v_2]$ is the same orientation as $[v_2 v_0 v_1]$ — both go CCW (↺). But $[v_1 v_0 v_2]$ goes clockwise.

$[v_0 v_1 v_2]$ and $[v_2 v_0 v_1]$ are same orientations

$[v_0 v_1 v_2] \longrightarrow [v_0 v_2 v_1] \longrightarrow [v_2 v_0 v_1]$ two swaps, so they are even permutations of each other.



$[v_0 v_1 v_2 v_3]$ is the same orientation as $[v_2 v_0 v_1 v_3]$ while $[v_0 v_2 v_1 v_3]$ is the opposite orientation.

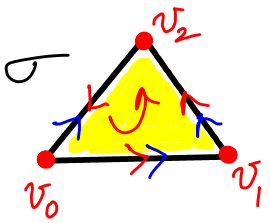
Induced orientation

Let σ have vertices $\{v_0, \dots, v_k\}$. When σ is oriented, it induces an "induced orientation" on all its $(k-1)$-faces. Each $(k-1)$-face of σ can be denoted as $\operatorname{conv}\{v_0, \dots, \hat{v}_i, \dots, v_k\}$, where the '^' (hat) above a vertex indicates that it is excluded.

"convex hull" here is used in general to cover both geometric and abstract cases.

Let σ be oriented as $[v_0, \dots, v_k]$. Then the orientation induced by σ on $\tau = \operatorname{conv}\{v_0, \dots, \hat{v}_i, \dots v_k\}$ is the same as that of $[v_0, \dots, \hat{v}_i, \dots, v_k]$ if i is even. Else, it is the opposite orientation.

For illustration, consider the triangle oriented as $\sigma = [v_0 v_1 v_2]$.



"$>$": induced orientations

"$>$": edges' own orientations, which are independent of the induced orientations

σ has three edges as faces, given by $\{v_0, v_1\}$, $\{v_0, v_2\}$, and $\{v_1, v_2\}$.

The induced orientations on these edges are $[v_0 v_1]$, $[v_2 v_0]$, and $[v_1 v_2]$, respectively.

We leave out v_1 to get the edge $\{v_0, v_2\}$, and hence its induced orientation is the reverse of $[v_0 v_2]$, i.e., it is $[v_2 v_0]$.

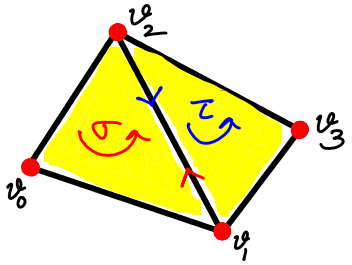
In the case of a triangle, the induced orientations indeed agree with the intuition of arrows "induced" on the edges by the CCW (or CW) arrow of the triangle.

CCW: counterclockwise

Comparing Orientations

Let σ, τ be simplices. If $\dim \sigma \neq \dim \tau$ we cannot compare their orientations. So let us consider the case when $\dim \sigma = \dim \tau = k$.

If σ and τ share a common $(k-1)$-face, they are **consistently oriented**, or oriented the same way, if they induce **opposite** orientations on the common $(k-1)$-face.



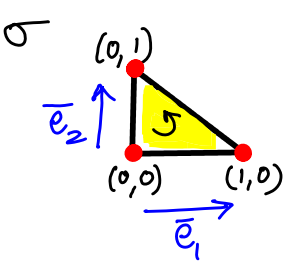
Cinduced orientations on $\overline{v_1 v_2}$ shown by $\wedge$ and $\vee$)

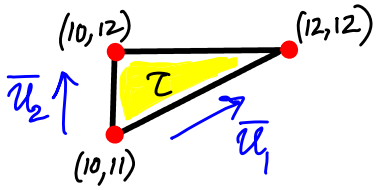
orientations induced by $\sigma = [v_0 v_1 v_2]$ and $\tau = [v_1 v_3 v_2]$ on $\overline{v_1 v_2}$ are opposite.

Hence the two triangles are consistently oriented — both CCW here.

Note that the induced orientations are separate from the edge's own orientation. Here, we could have $[v_1, v_2]$ or $[v_2, v_1]$ as the inherent orientation of $\overline{v_1 v_2}$. The induced orientations are still as shown in the figure.

Note: We could compare orientations of two k-simplices even if they do not share a common $(k-1)$-face, if they both are sitting in the same k-dimensional plane. For instance, consider two disjoint triangles in $\mathbb{R}^2$.





The 3 vertices of a triangle generate two vectors, whose cross-product can be used to calculate the **signed area** of the triangle.

For oriented triangle $[v_0 v_1 v_2]$, consider vectors $\bar{e}_1 = \bar{v}_1 - \bar{v}_0$ and $\bar{e}_2 = \bar{v}_2 - \bar{v}_0$.

For σ, we can take $\bar{e}_1 = \begin{bmatrix}1\\0\end{bmatrix} - \begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\bar{e}_2 = \begin{bmatrix}0\\1\end{bmatrix} - \begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$.

The signed area of σ is given by $\text{area}(\sigma) = \frac{1}{2}|\bar{e}_1 \times \bar{e}_2| = \frac{1}{2}\det([\bar{e}_1, \bar{e}_2])$.

($|\bar{e}_1 \times \bar{e}_2|$: area of the parallelogram generated by $\bar{e}_1$ & $\bar{e}_2$)

Thus, $\text{area}(\sigma) = \frac{1}{2}\left|\begin{bmatrix}1&0\\0&1\end{bmatrix}\right| = \frac{1}{2}$.

(determinant)

Similarly for τ, we choose $\bar{u}_1 = \begin{bmatrix}12\\12\end{bmatrix} - \begin{bmatrix}10\\11\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix}$ and $\bar{u}_2 = \begin{bmatrix}10\\12\end{bmatrix} - \begin{bmatrix}10\\11\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$,

giving $\text{area}(\tau) = \frac{1}{2}\left|[\bar{u}_1, \bar{u}_2]\right| = \frac{1}{2}\left|\begin{bmatrix}2&0\\1&1\end{bmatrix}\right| = 1$.

Since σ and τ have the same sign for their signed areas, they are consistently oriented. In this case, they are both oriented CCW. In fact, both sets of vectors $\{\bar{e}_1, \bar{e}_2\}$ and $\{\bar{u}_1, \bar{u}_2\}$ orient $\mathbb{R}^2$ in the same way here.

These computations are naturally extended to d-dimensions for $d \geq 3$. We can compute the signed d-volume in the same fashion.

We could compare orientations in the abstract setting as well.

Consistently oriented simplices

We consider an example in the abstract setting.

Let $\sigma_1 = [2\ 5\ 12\ 19]$ and $\sigma_2 = [12\ 19\ 7\ 2]$ be two oriented 3-simplices. Are they consistently oriented?

Notice that $\tau = \{2, 12, 19\}$ is the common 2-face.

(exclude to get τ)

$\sigma_1 = [\overset{0}{2}\ \overset{1}{5}\ \overset{2}{12}\ \overset{3}{19}]$ $\quad$ $\sigma_2 = [\overset{0}{12}\ \overset{1}{19}\ \overset{2}{7}\ \overset{3}{2}]$

We get τ by removing the first ($i=1$) vertex from σ_1. Similarly, we get τ by removing the second ($i=2$) vertex from σ_2.

Hence, the orientation induced on τ by σ_1 is $[12\ 2\ 19]$, which is the opposite orientation to $[2\ 12\ 19]$. (differ by 1 pairwise swap)

The orientation induced on τ by σ_2 is $[12\ 19\ 2]$.

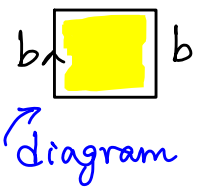
The two induced orientations are opposite here. Hence σ_1 and σ_2 are consistently oriented.

We extend the idea of when two d-simplices are consistently oriented to when the entire triangulation is consistently oriented.

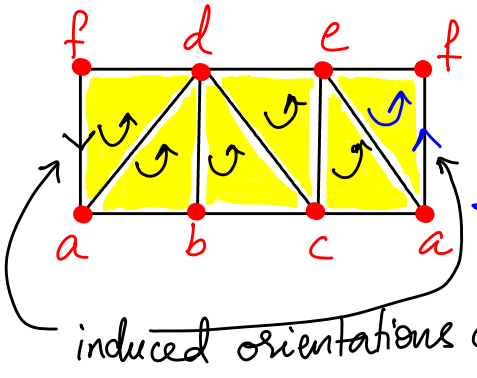
Def A triangulable d-manifold (with or without boundary) is called **orientable** if all the d-simplices in any triangulation of the manifold can be consistently oriented. Else, it is a **nonorientable** manifold.

Examples

1. Cylinder



diagram



all triangles oriented CCW.

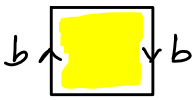
triangulation

induced orientations are opposite

It can be checked that the orientations induced by each pair of triangles on their common (shared) edge are indeed opposite. In particular, notice that this is indeed the case for $\overline{af}$ — induced orientations from $[adf]$ and $[afe]$ are $[fa]$ and $[af]$, respectively.

Thus, the cylinder is orientable.

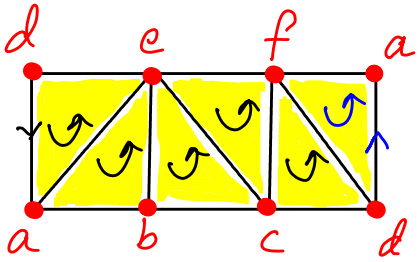
2. Möbius strip



diagram

We had noted previously that the Möbius strip is non-orientable.

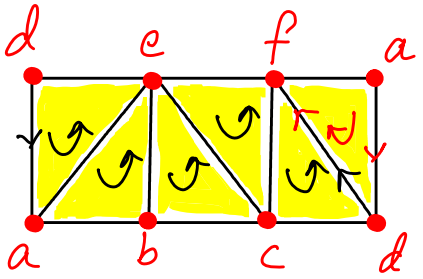
triangulation →



The orientations induced on edge $\overline{ad}$ by [aed] and [afd] are the same here — both being [da], or $d \to a$. Thus, $\triangle aeb$ and $\triangle adb$ are not consistently oriented.

Notice that the induced orientations on all remaining shared edges except $\overline{ad}$ are indeed opposite — check the induced orientations on $\overline{ae}$, $\overline{be}$, $\overline{ce}$, $\overline{cf}$ and $\overline{df}$.

If we fix the orientations such that induced orientations on $\overline{ad}$ are opposite, say, by orienting $\triangle adf$ clockwise, i.e., [adf], then the induced orientations on $\overline{df}$ are now identical — [df].



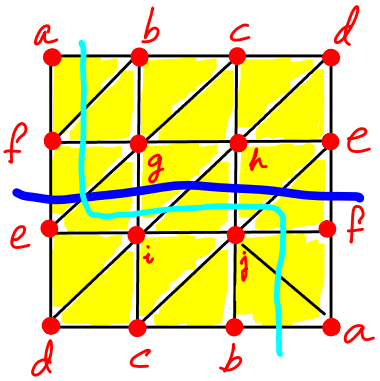
It turns out that however we orient the triangles, the induced orientations will be the same on one of the shared edges. As such, all triangles in the Möbius strip cannot be consistently oriented.

Hence the Möbius strip is non-orientable.

In fact, Möbius strips are the minimal non-orientable "objects" in 2D. For instance, you can identify Möbius strips in the triangulations of $\mathbb{RP}^2$ and $\mathbb{K}^2$ we introduced in Lecture 7!

Hence for surfaces, it is sufficient to identify a Möbius strip in the given triangulation to "certify" its nonorientability.

$\mathbb{RP}^2$:

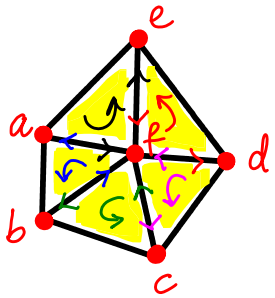


The middle strip of 6 triangles forms a Möbius strip! And there are several other instances of the Möbius strip here — another instance is identified by ~ curve.

In both the cylinder and the Möbius strip, the boundary edges can be oriented arbitrarily. These are the edges that are faces of only one triangle each. For example, in the Möbius strip above, edges $\overline{ab}$, $\overline{bc}$, $\overline{cd}$, $\overline{de}$, $\overline{ef}$ and $\overline{af}$ are boundary edges, and they can be assigned orientations arbitrarily without affecting the (non-)orientability of the manifold.

Checking orientability of a d-manifold Start by assigning an orientation to one d-simplex σ (pick one of the two possibilities). Then "propagate" this orientation to any other d-simplex σ' that shares a common $(d-1)$-simplex τ, say, with σ. In other words, orient σ' such that the orientations induced on τ by σ and σ' are opposite. Continue this process until all d-simplices are oriented. If we can consistently orient all d-simplices, the manifold is orientable. Else, it is non-orientable.

Example:



Start with [afe], and propagate this orientation in the order [abf], [bcf], [cdf], and [def]. Notice that the shared edges $\overline{af}$, $\overline{bf}$, $\overline{cf}$, $\overline{df}$, and $\overline{ef}$ have opposite induced orientations.

MATH 529 – Lecture 9 (02/06/2024)

Today:
* subdivision
* barycentric subdivision
* star, closed star, link

We now define several constructs on simplicial complexes that are simplicial analogues of standard constructs on (continuous) spaces, such as subspaces, open sets, neighborhoods of points, boundary of a neighborhood, etc.

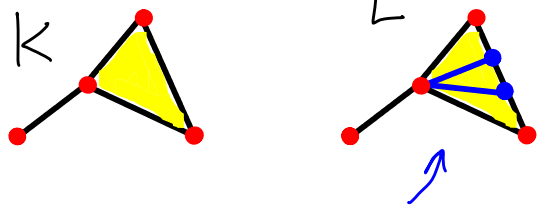
Subdivision

Def A simplicial complex L is a **subdivision** of another complex K if $|K| = |L|$, and $\forall \sigma \in L$, $\sigma \subseteq \tau \in K$.

($|K| = |L|$: not $\approx$!)

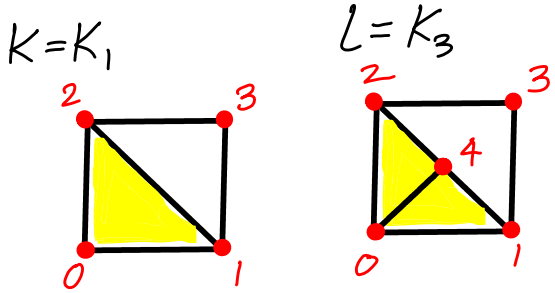
In words, every simplex in L is contained in a simplex in K.

Example:



the 3 triangles are contained in the one triangle in K, for instance.

Here is another example we saw in a different context (in Lecture 6):



Barycentric Subdivision

One way to create a subdivision of K is by forming the **barycentric subdivision**, denoted $\mathrm{Sd}\,K$ (this is a "standard" way to subdivide a complex).

The **barycenter** of a simplex is the centroid of its vertices.

We define (or construct) the barycentric subdivision inductively on the dimension of the simplices.

Inductive construction of $\mathrm{Sd}\,K$

Notation $K^{(j)} = \{\sigma \in K \mid \dim \le j\}$ is the j-skeleton of K. (all simplices in K with dimension $\le j$)

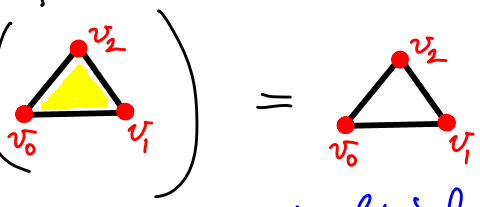
Thus, $K^{(0)}$ is the set of vertices of K.

We start by defining $\mathrm{Sd}\,K^{(0)} = K^{(0)}$. (barycenter of a vertex is the vertex itself).

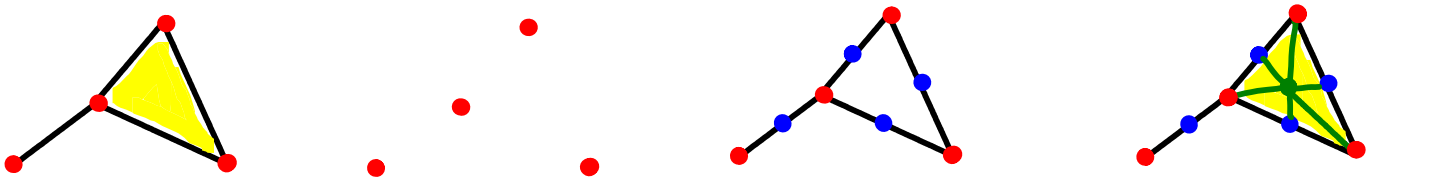
If we have $\mathrm{Sd}\,K^{(j-1)}$, we construct $\mathrm{Sd}\,K^{(j)}$ by adding the barycenter of every j-simplex as a new vertex, and connecting that vertex to each simplex that subdivides the boundary of that simplex.

$$\mathrm{Bd}\,\sigma = \bigcup_{\tau \prec \sigma} \tau$$

or Bd



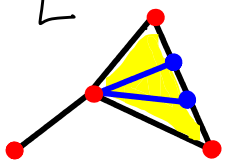
We illustrate this construction on the same simplicial complex we used to illustrate subdivisions.



K $\quad$ $\mathrm{Sd}\,K^{(0)} = K^{(0)}$ $\quad$ $\mathrm{Sd}\,K^{(1)}$ $\quad$ $\mathrm{Sd}\,K^{(2)} = \mathrm{Sd}\,K$

Note that $Sd\, K^{(0)} = K^{(0)}$ as the barycenter of a vertex is the vertex itself. Each edge in K is replaced by 2 edges, while each triangle is replaced by 6 triangles in $Sd\, K$.

L



In the previous example, L is a subdivision of of K, but is not a barycentric subdivision.

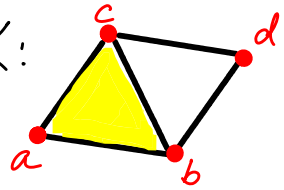
In certain applications, we may want not to subdivide certain simplices, e.g., keep a subset of original edges in fact.

We now present definitions on simplicial complexes corresponding to open neighborhoods around points in $\mathbb{R}^d$. We start with some preliminary definitions.

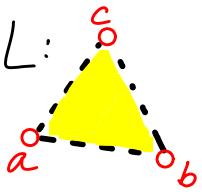
Def A **subcomplex** of K is a simplicial complex L, such that $L \subseteq K$.
A subset that is a simplicial complex by itself

Def The smallest subcomplex containing a subset $L \subseteq K$ is its **closure**, $Cl\, L = \{\tau \in K \mid \tau \le \sigma \in L\}$. → could also use $\overline{L}$

K:



L:



L is not a subcomplex of K here!

$Cl\,\{\overline{ac}, d\} = \{\overline{ac}, d, a, c\}$. $Cl\,\{\Delta abc\} = \{\Delta abc, \overline{ab}, \overline{ac}, \overline{bc}, a, b, c\}$.

Notice that $Cl\, L$ is a simplicial complex by itself.

Def For a simplex $\sigma \in K$, we define its **boundary** and **interior** as follows.

$$Bd\,\sigma = \bigcup_{\tau < \sigma} \tau \quad \text{and} \quad Int\,\sigma = \sigma \setminus Bd\,\sigma.$$

or $\partial\sigma$; or $\mathring{\sigma}$

Def For a vertex $\bar{v}$ of K, the **star** of $\bar{v}$, denoted $St\,\bar{v}$, is

$$St\,\bar{v} = \bigcup_{\sigma \geq \bar{v}} Int\,\sigma.$$

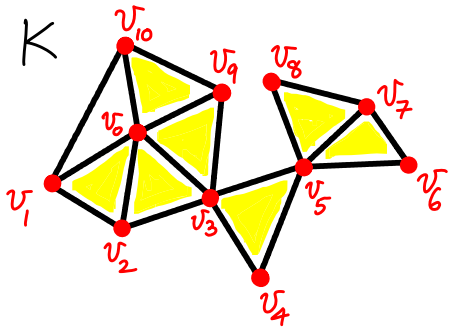
Union of interiors of all simplices that have $\bar{v}$ as a vertex.

$\sigma \geq \bar{v}$: coface

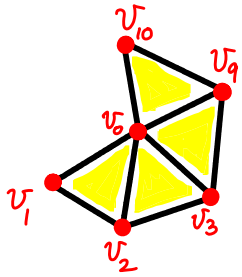
Also, $Cl\,St\,\bar{v}$ (or $\overline{St\,\bar{v}}$) is the **closed star** of $\bar{v}$ → take the closure of $St\,\bar{v}$, i.e., throw in all faces.

The set $Cl\,St\,\bar{v} \setminus St\,\bar{v}$ is the **link** of $\bar{v}$, denoted $Lk\,\bar{v}$.

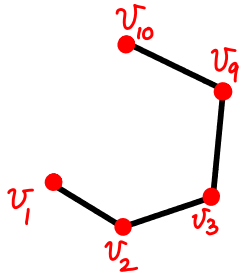
Here is an example – consider the complex K shown.



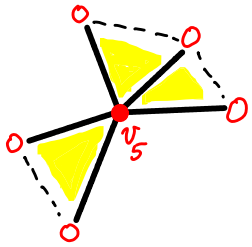
$St\,v_0$



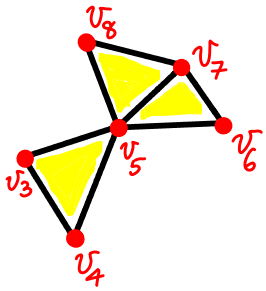
$Cl\,St\,v_0$



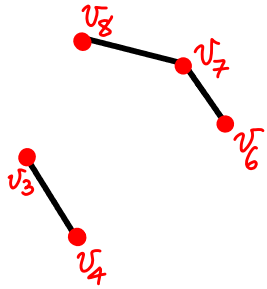
$Lk\,v_0$



$St\,v_5$



$Cl\,St\,v_5$



$Lk\,v_5$

Following these examples, we could make the following observations.

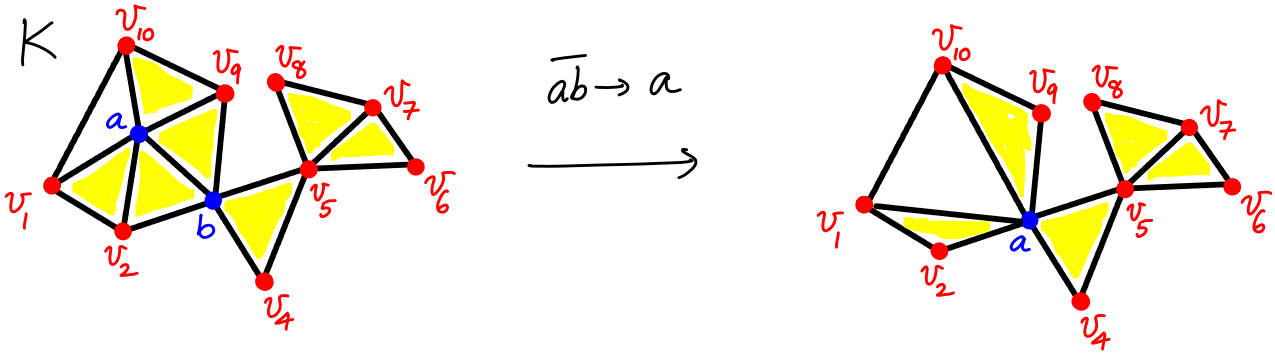
$St\,\bar{v}$ and $Cl\,St\,\bar{v}$ are both path connected, but $Lk\,\bar{v}$ need not be be connected.

↳ there is a path between any two points in the set

Also, $Cl\,St\,\bar{v}$ and $Lk\,\bar{v}$ are subcomplexes of K, but $St\,\bar{v}$ is typically not a subcomplex.

An aside on mesh simplification

Links of vertices are employed in mesh simplification — you want a simplicial complex with a smaller number of simplices while preserving topology. A standard operation used in this context is edge contraction, where we replace edge $\overline{ab}$, say, with vertex a. We also make associated changes to other simplices connected to edge $\overline{ab}$. Here is an illustration.



Here, contracting $\overline{ab}$ to a preserves the topology — at least, we do not close any holes in K. But if we were to contract edge $\overline{v_1 v_{10}}$, we will close a hole!

We could define a condition on how the links of a, b, and of $\overline{ab}$ (to be defined next) are related. This is a local condition, and can be checked quickly. If it is satisfied, we can contract $\overline{ab}$ without the fear of closing any holes.

Such operations are critical to efficient computations on simplicial complexes. We will talk about them later in the semester.

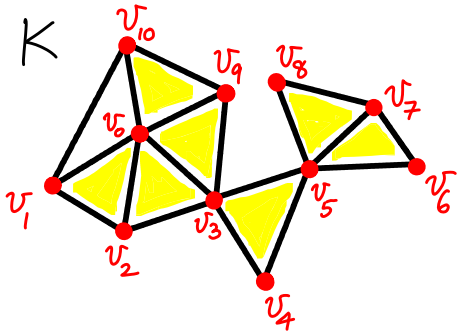
We now extend the definitions of star and link to simplices (of any dimension) and then to collections of simplices.

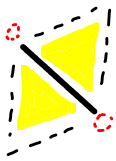
Def For a simplex $\sigma \in K$, we let

$$St\,\sigma = \bigcup_{\tau \geq \sigma} Int\,\tau, \quad Cl\,St\,\sigma \text{ is its closed star,}$$

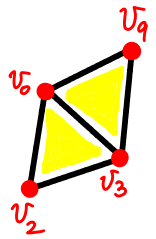
and $$Lk\,\sigma = \{ U\tau \mid \tau \in Cl\,St\,\sigma,\ \tau \cap \sigma = \emptyset \}.$$

In words, the link of σ is the set of simplices in its closed star which are disjoint from σ.





$St\,\overline{v_0v_3}$

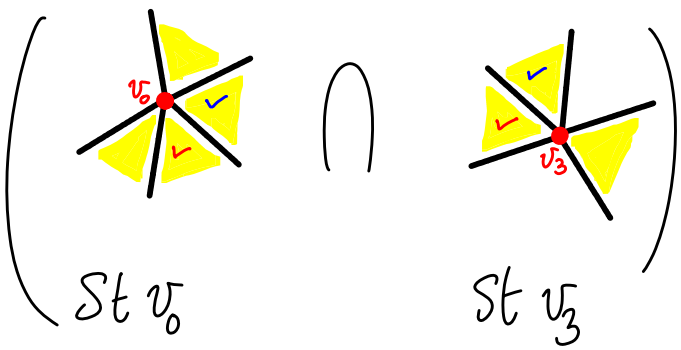


$Cl\,St\,\overline{v_0v_3}$

$Lk\,\overline{v_0v_3}$

For $\sigma = [v_0, \ldots, v_k]$, $\quad St\,\sigma = \left(\bigcap_{i=0}^{k} St\,v_i \right)$

For instance, with $\sigma = v_0v_3$, we get



$St\,v_0$ $\qquad$ $St\,v_3$ $\qquad$ $St\,\overline{v_0v_3}$

Def For $X \subseteq K$, we define

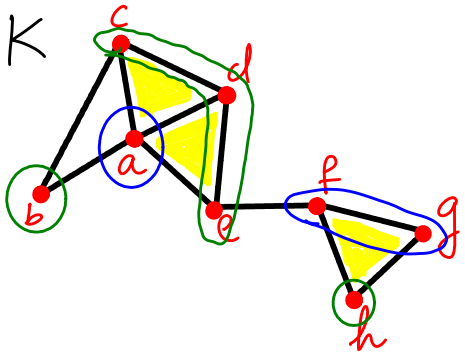
$$\mathrm{St}\, X = \{ \cup \tau \mid \tau \geq \sigma, \sigma \in X \}.$$

→ Set of cofaces of each simplex in the set X.

$\mathrm{Cl}\,\mathrm{St}\, X$ is the closure of $\mathrm{St}\, X$. We define

$\mathrm{Lk}\, X$ = Set of simplices in $\mathrm{Cl}\,\mathrm{St}\, X$ that do not belong to $\mathrm{St}(\bar{X})$.

The intuition of the star being the open neighborhood of X and link being the boundary of the closed neighborhood still holds.



Let $X = \{a, \overline{fg}\}$. Then we intuitively want

$\mathrm{Lk}\, X = \{b, c, d, e, \overline{cd}, \overline{de}, h\}$.

Following the definitions, we get

$$\mathrm{St}\, X = \{a, \overline{ab}, \overline{ac}, \overline{ad}, \overline{ae}, \triangle aed, \triangle ade, \overline{fg}, \triangle fgh\}.$$

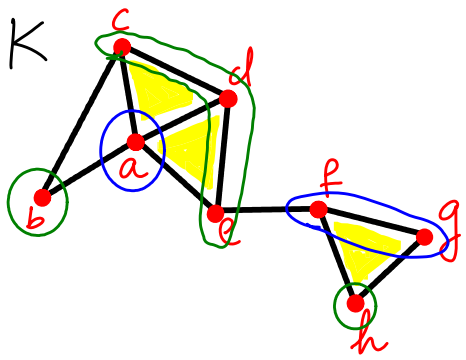
$$\mathrm{Cl}\,\mathrm{St}\, X = \{a, \overline{ab}, \overline{ac}, \overline{ad}, \overline{ae}, \triangle aed, \triangle ade, \overline{fg}, \triangle fgh, \ldots$$
$$b, c, d, e, \overline{cd}, \overline{de}, f, g, h, \overline{fh}, \overline{gh}, \}$$

We will finish this example in the next lecture...

MATH 529 – Lecture 10 (02/08/2024)

Today:
* poset representation
* retraction, homotopy equivalence
* Nerve

We first finish the example of $Lk\,X$ for a general set $X \subset K$...



Let $X = \{a, \overline{fg}\}$. Then we intuitively want

$Lk\,X = \{b, c, d, e, \overline{cd}, \overline{de}, h\}$.

Following the definitions, we get

$St\,X = \{a, \overline{ab}, \overline{ac}, \overline{ad}, \overline{ae}, \triangle acd, \triangle ade, \overline{fg}, \triangle fgh\}$.

$Cl\,St\,X = \{a, \overline{ab}, \overline{ac}, \overline{ad}, \overline{ae}, \triangle acd, \triangle ade, \overline{fg}, \triangle fgh, \ldots$
$b, c, d, e, \overline{cd}, \overline{de}, f, g, h, \overline{fh}, \overline{gh}\}$

(the simplices added as proper faces)

We also get

$Cl\,X = \{a, \overline{fg}, f, g\}$, and

(faces added to close X)

$St\,Cl\,X = \{a, \overline{ab}, \overline{ac}, \overline{ad}, \overline{ae}, \triangle acd, \triangle ade, \overline{fg}, \triangle fgh, \ldots$
$f, g, \overline{ef}, \overline{fh}, \overline{gh}\}$

(cofaces of the elements added to close X)

$\Rightarrow Lk\,X = Cl\,St\,X - St\,Cl\,X = \{b, c, d, e, \overline{cd}, \overline{de}, h\}$, as expected!

Note that $\overline{ef} \in St\,Cl\,X$, but is not in $Lk\,X$ (as per definition).

For a small example, we can easily eye-ball these sets. But how do you handle large simplicial complexes with, say, 10^4 simplices?

We describe a way to efficiently store simplicial complexes and to read off $\mathrm{St}X$, $\mathrm{Lk}X$, $\mathrm{Cl}X$, etc. from that representation.

Def (poset) Given a finite set S, a **partial order** is a binary relation $\leq$ on S that is reflexive, antisymmetric, and transitive, i.e., $\forall\, x, y, z \in S$,

(a) $x \leq x$;

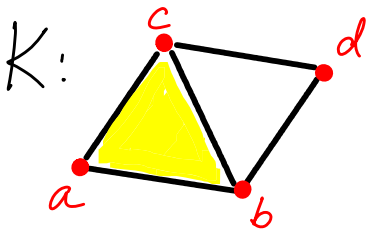
(b) $x \leq y$ and $y \leq x \Rightarrow x = y$; and

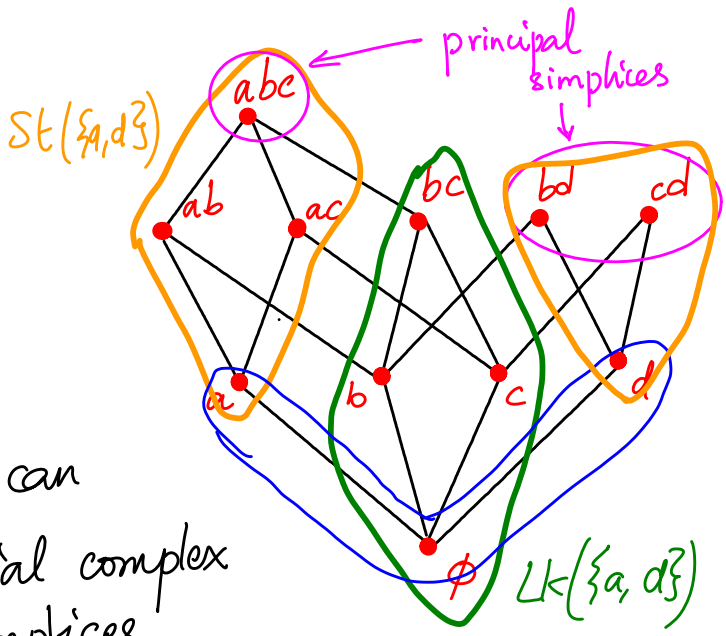
(c) $x \leq y, y \leq z \Rightarrow x \leq z$.

→ partial, as not every $x, y \in S$ are related by $\leq$.

A set S with a partial order is called a partially ordered set, or a poset. The face relationships of a simplicial complex is a partial order. So the vertex scheme of a simplicial complex with face relationships is a poset.

Illustration





The simplices "at the top" are called principal simplices. We can determine the entire simplicial complex if we know the principal simplices.

To find star ($St\,X$), take X and everything above. For instance,

$St(\{a,d\}) = \{a, ab, ac, abc, d, bd, cd\}$.

To find $Cl\,X$, take X and everything below; e.g., $Cl(\{a,d\}) = \{a, d, \emptyset\}$.

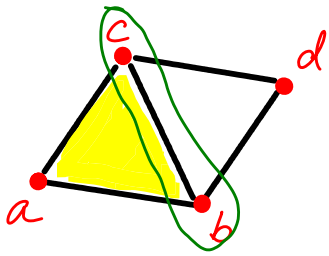
Notice that $Cl\,St(\{a,d\}) = K \cup \{\emptyset\}$ here.

As a convention, the empty simplex (or null set) is added at the bottom of this poset representation. It plays the role of the "root node" from which the poset representation "grows up".

Hence, we include the empty set $\emptyset$ in our definitions and discussions of closure, star, and link. In particular, we modify the definition of link slightly as follows:

$$Lk\,X = Cl\,St\,X - St\,(Cl\,X - \{\emptyset\}).$$

With $X = \{a,d\}$, we expect $Lk\,X$ to be $\{bc, b, c\}$.



Recall, $Cl\,X = \{a, d, \emptyset\}$. So,

$St\,(Cl\,X - \{\emptyset\}) = St(\{a,d\})$ here. Hence we indeed get

$$Lk\,X = \{bc, b, c, \emptyset\}.$$

We now define a notion of topological similarity that is weaker than homeomorphism. We then use this notion to define how to build simplicial complexes on data sets of points in $\mathbb{R}^d$.

Homotopy

Def Let $f, g : \mathbb{X} \to \mathbb{Y}$ be continuous maps from topological space $\mathbb{X}$ to space $\mathbb{Y}$. A **homotopy** between f and g is another continuous map

$$H : \mathbb{X} \times \underbrace{[0,1]}_{t} \longrightarrow \mathbb{Y}$$

such that H agrees with f at $t=0$, and with g at $t=1$. In other words,

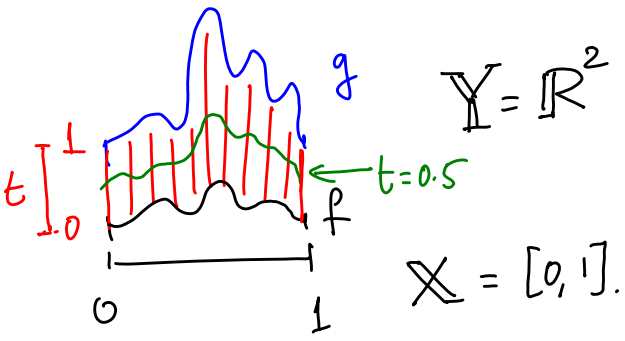
$$H(x,0) = f(x) \quad \forall x \in \mathbb{X}, \text{ and}$$
$$H(x,1) = g(x) \quad \forall x \in \mathbb{X}.$$

The index t can be thought of as time, varying from 0 to 1.

H could be thought of a time-series of functions $f_t(x) = H(x,t)$, where $f_t : \mathbb{X} \to \mathbb{Y}$ for $t \in [0,1]$, with $f_0 = f$ and $f_1 = g$.

We say that f is **homotopy equivalent** to g, or that f is homotopic to g. We denote this equivalence relation by $f \simeq g$. → reflexive, symmetric, and transitive.

Here is an illustration, with $\mathbb{X} = [0,1]$ and $\mathbb{Y} = \mathbb{R}^2$. The homotopy H is a 2D strip of functions going from f to g. All of f, g, and H are continuous.



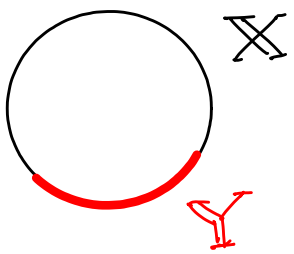
We extend the definition of homotopy to topological spaces. First we consider a special case.

Def $\mathbb{Y} \subseteq \mathbb{X}$ is a **retract** of $\mathbb{X}$ if there is a continuous map $r: \mathbb{X} \to \mathbb{Y}$ with $r(y) = y \ \forall y \in \mathbb{Y}$. r is called a **retraction**.

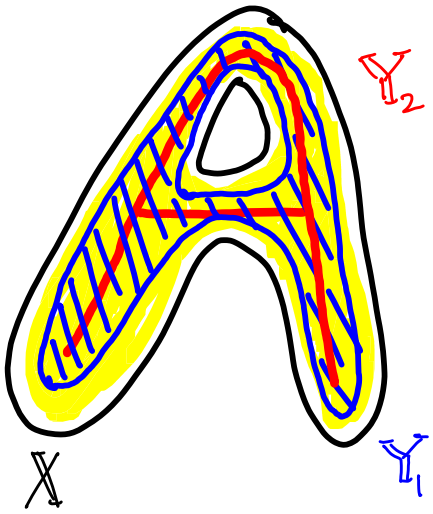
Def $\mathbb{Y}$ is a **deformation retract** of $\mathbb{X}$, and r is a **deformation retraction**, if there is a homotopy between the retract r and the identity map $id_{\mathbb{X}}$ on $\mathbb{X}$, i.e., $r \simeq id_{\mathbb{X}}$.

$\rightsquigarrow id_{\mathbb{X}}(x) = x \ \forall x \in \mathbb{X}$.

We also say that $\mathbb{X}$ deformation retracts to $\mathbb{Y}$.

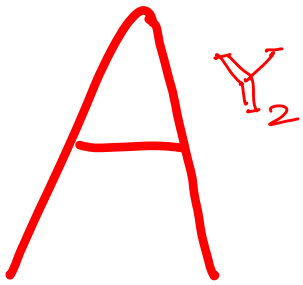


Here is an example of a retract that is not a deformation retract. Notice that $\mathbb{X}$ is S^1 (circle), while $\mathbb{Y}$ is just an open arc.



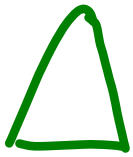
Y_1 is a deformation retract of X.

Continue to deform to obtain



$(Y_2 \subset X)$.
"skeleton" sitting inside the "fat A".

$Y_2 \not\subset X$, but Y_2 and X have the same homotopy type.
we'll define it formally soon!

Deforming ever further, we can get Y_3  $(Y_3 \subset Y_2)$.

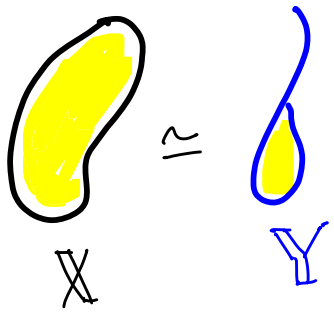
X, and Y_j, $j=1,2,3$ are all homotopy equivalent. Also, each Y_j is a deformation retract of Y_k for $k<j$, and also of X.

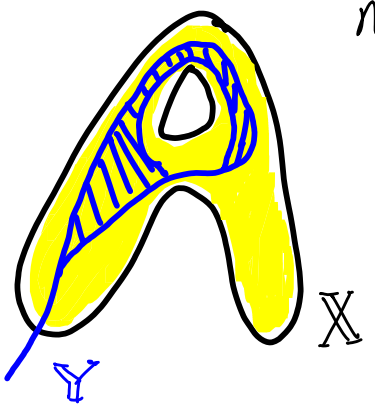
Notice that while X and Y_2, for instance, are not homeomorphic, they both are forms of the letter 'A'. Y_2 is, in some sense, the "skeleton" of X. These types of transformations are allowed in the less tight notion of topological similarity termed homotopy equivalence, which is not as strict as homeomorphism.

Def $\mathbb{X}$ and $\mathbb{Y}$ are homotopy equivalent, or have the **same homotopy type**, if there exists continuous maps $f: \mathbb{X} \to \mathbb{Y}$ and $g: \mathbb{Y} \to \mathbb{X}$ such that $f \circ g \simeq id_{\mathbb{Y}}$ and $g \circ f \simeq id_{\mathbb{X}}$.

We denote $\mathbb{X} \simeq \mathbb{Y}$.

Note we have $f \circ g \simeq id_{\mathbb{Y}}$ and $g \circ f \simeq id_{\mathbb{X}}$, and not **equal to** in each case.





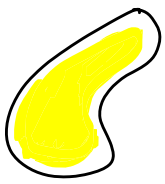
notice that $\mathbb{X} \simeq \mathbb{Y}$ here, but $\mathbb{Y}$ is not a retract of $\mathbb{X}$.

If two spaces are homeomorphic, they have the same homotopy type.

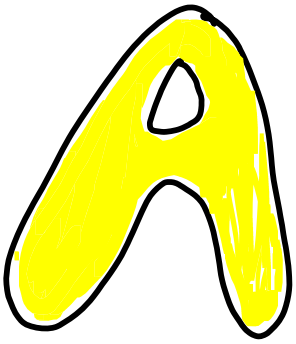
So, $\mathbb{X} \approx \mathbb{Y} \implies \mathbb{X} \simeq \mathbb{Y}$.

The implication does not go the other way, as many of the above examples show. For instance, $\mathbb{X}$ (fat 'A') is a 2-manifold with boundary, and $\mathbb{Y}_2$ (1-D 'A'), its 'skeleton', is a 1-manifold with boundary.

Def If $\mathbb{Y}$ is a single point, and $\mathbb{X} \simeq \mathbb{Y}$, then we say that $\mathbb{X}$ has the homotopy type of a point, and we say that $\mathbb{X}$ is **contractible**.



$\mathbb{X}_1$ is contractible



$\mathbb{X}_2$ is not contractible.

Our next goal is to study how to construct simplicial complexes from sets of points (in some space $\mathbb{R}^d$). Most applications analyze data in this format. We would like to construct the simplicial complex such that it captures the topology of the point set — if not up to homeomorphism, up to homotopy, or even up to a weaker level (to be defined later). We need one more concept to introduce such constructions.

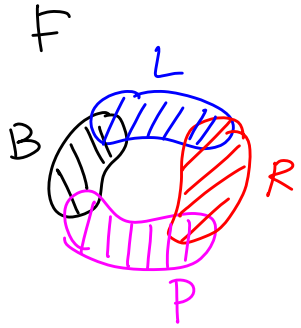
Def (Nerves) Let $\mathcal{F}$ be a finite collection of sets in $\mathbb{R}^d$. The **nerve** of $\mathcal{F}$ consists of all subcollections of $\mathcal{F}$ with nonempty intersections.

$$\text{Nrv}\,\mathcal{F} = \{\mathcal{X} \subseteq \mathcal{F} \mid \cap \mathcal{X} \neq \emptyset\}.$$

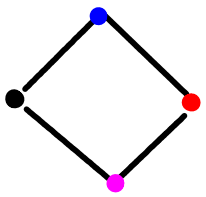
$\text{Nrv}\,\mathcal{F}$ is always an ASC, as $\cap \mathcal{X} \neq \emptyset$ and $\mathcal{Y} \subseteq \mathcal{X} \Rightarrow \cap \mathcal{Y} \neq \emptyset$.
↳ abstract simplicial complex

Example

Consider an instance of $\mathcal{F}$ consisting of four sets, shaded **B**lack, B**L**ue, **R**ed, and **P**ink. The four sets intersect in four pairs, as shown. Then $\text{Nrv}\,\mathcal{F}$ consists of the following intersecting subsets of $\{B, L, R, P\}$.



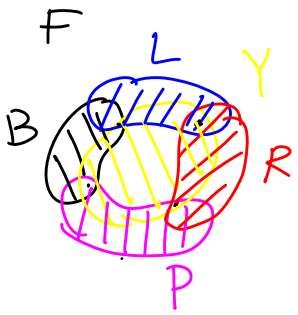
$$\text{Nrv}\,\mathcal{F} = \{\{B\}, \{L\}, \{R\}, \{P\}, \{B,L\}, \{L,R\}, \{R,P\}, \{B,P\}\}$$



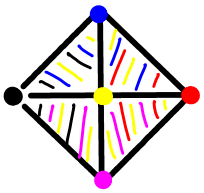
one geometric realization of $\text{Nrv}\,\mathcal{F}$

$\text{Nrv}\,\mathcal{F}$ has a geometric realization in the same space ($\mathbb{R}^2$) as $\mathcal{F}$ here.

Now consider adding another set to F, shaded Yellow, such that Y intersects each pair of intersections already present, as shown.



Now, Nrv F has a geometric realization as a disc made of four triangles, as shown here.

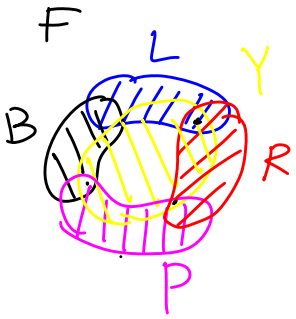


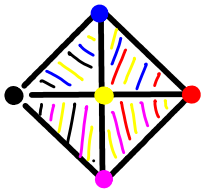
One geometric realization of Nrv F

MATH 529 – Lecture 11 (02/13/2024)

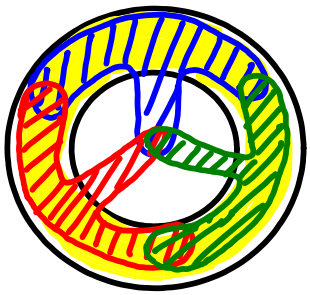
Today:
* Nerve theorem
* Čech complex
* Vietoris-Rips complex
* Delaunay complex

Recall $\mathrm{Nrv}\, F = \{X \subseteq F \mid \cap X \neq \emptyset\}$, and the examples ...





In this example, F and Nrv F are homotopy equivalent. But does this result hold in general? Let's consider another example...



$|F|$ is a closed disc with three holes

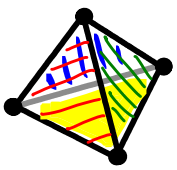
F is a collection of four regions, such that every subset of three regions has a common intersection.

With $F = \{R, B, G, Y\}$ for Red, Blue, Green, Yellow, we can write

$$\mathrm{Nrv}\, F = \{R, B, G, Y, \{R,B\}, \{R,G\}, \{R,Y\}, \{B,G\}, \{B,Y\}, \{G,Y\}, \{R,B,G\}, \{R,B,Y\}, \{R,G,Y\}, \{B,G,Y\}\}.$$

Indeed, Nrv F has a geometric realization as the surface of a tetrahedron as shown.

Nrv F



$\approx S^2$

So $\mathrm{Nrv}\, F \neq |F|$ here!

→ underlying space, disk with 3 holes.

But if the sets in F are "nice", we do get homotopy equivalence with Nrv F, as specified by the following theorem.

Nerve theorem Let $\mathcal{F}$ be a finite collection of closed **convex** sets in $\mathbb{R}^d$. Then $Nrv\,\mathcal{F}$ has the same homotopy type as the collection of sets in $\mathcal{F}$.

Our goal is to build simplicial complexes out of collections of points. We could consider a collection of convex sets, each containing one point from the set, and then form its nerve. A default convex set containing a point is a closed ball centered at that point. We will consider a few different ways of forming simplicial complexes out of points using balls centered on them.

Čech Complex (pronounced as "check") Let S be a finite set of points in $\mathbb{R}^d$.

We write $B_{\bar{x}}(r) = \bar{x} + rB^d = \{\bar{y} \in \mathbb{R}^d \mid \|\bar{y} - \bar{x}\| \le r\}$, for the closed ball of radius r and center $\bar{x}$.

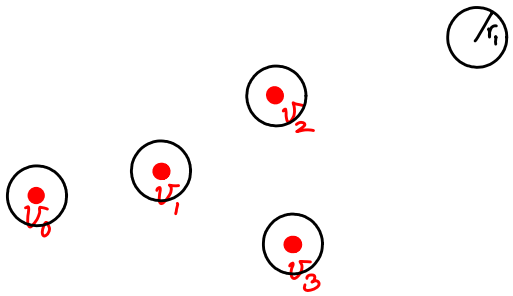
The **Čech complex** at radius r of the points in set S is the nerve of the collection of closed r-balls centered at the points.

$$\check{C}ech(r) = \left\{\sigma \subseteq S \;\middle|\; \bigcap_{\bar{x} \in \sigma} B_{\bar{x}}(r) \neq \emptyset\right\}.$$

one could write $\check{C}ech_S(r)$ to be complete, but S is understood, and hence omitted, typically.

to be exact, one should say $conv(\sigma)$ here. We do mean the simplex spanned by vertices in σ; $\sigma = conv\{\bar{v}_0, \ldots, \bar{v}_k\}$, $\bar{v}_i \in S$.

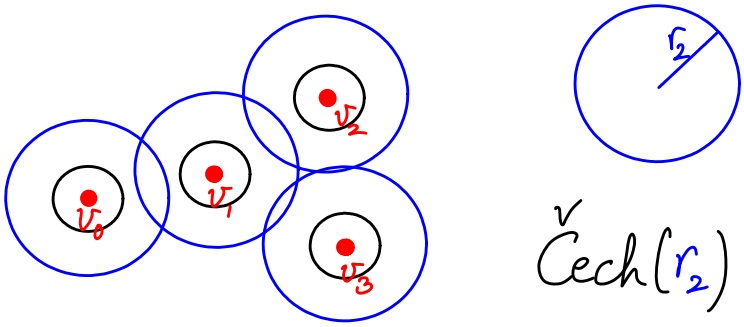
Consider an example with four points in $\mathbb{R}^2$ as shown.



$\check{C}ech_S(r_1) \simeq \{v_0, v_1, v_2, v_3\}$

$\check{C}ech$ complex is homotopic to the union of balls centered at v_i — at all radii (and not just for small values such as r_1 shown here)

r_1 is small enough that no two of the balls centered at v_i intersect. Hence, $\check{C}ech(r_1)$ has just the four points. Let's consider a bigger radius now.

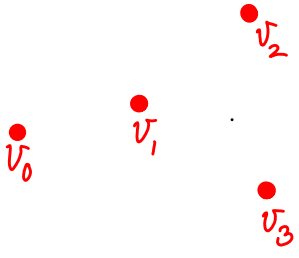


the balls at v_1 & v_2 intersect

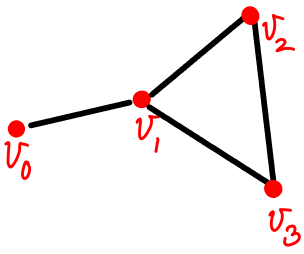
$\check{C}ech(r_2) = \{v_0, v_1, v_2, v_3, \overline{v_0v_1}, \overline{v_1v_2}, \overline{v_2v_3}, \overline{v_1v_3}\}$

Geometric realizations of $\check{C}ech(r_1)$ and $\check{C}ech(r_2)$:

$\check{C}ech(r_1)$:

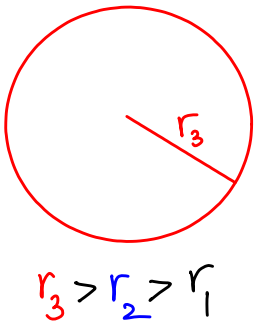


$\check{C}ech(r_2)$:

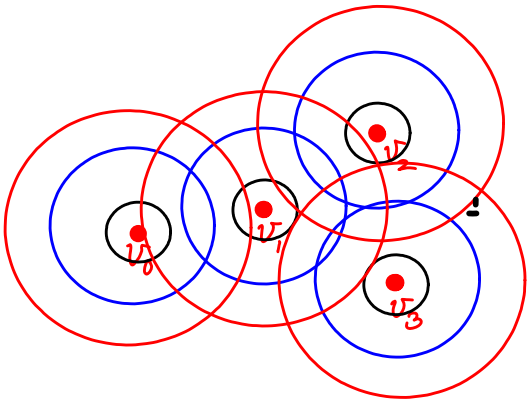


Notice that the balls centered at v_1, v_2, v_3 do not all interseet. Thus, there is a "hole" in between these three balls, which is represented by the empty triangle $v_1v_2v_3$ in $\check{C}ech(r_2)$.

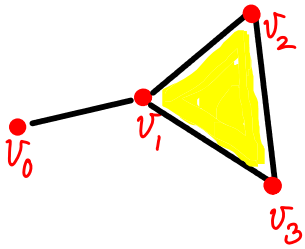
Increasing the radius a bit more brings in $\triangle v_1v_2v_3$:



$r_3 > r_2 > r_1$



$\check{C}ech(r_3)$:

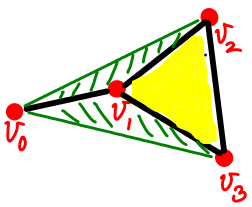


Note: The circles shown here are solid discs — shading is avoided for clarity.

Notice that $\check{C}ech(r_1)$ is a subcomplex of $\check{C}ech(r_2)$, which in turn is a subcomplex of $\check{C}ech(r_3)$.

In general, $\check{C}ech(r_i) \subseteq \check{C}ech(r_j)$ when $r_i \leq r_j$.

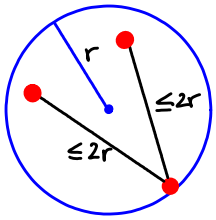
Also, $\check{C}ech(r)$ of a set S of points in $\mathbb{R}^d$ may not have a geometric realization in $\mathbb{R}^d$ itself. But you can always treat it as an abstract simplicial complex.



At larger radii (r_4), triangles $\triangle v_0v_1v_2$ and $\triangle v_0v_1v_3$ are included in $\check{C}ech(r_4)$, and at a still higher radius, tetrahedron $v_0v_1v_2v_3$ is included. But, of course, $\diamondsuit v_0v_1v_2v_3$ cannot be embedded in $\mathbb{R}^2$.

We will consider this aspect — the complex having a geometric realization in the input space itself — later on. First, we look at more properties of the $\check{C}ech$ complex.

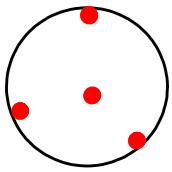
Another property central to the Čech complex is that balls of radius r have a common intersection iff their centers lie inside a ball of radius r.



So, $\sigma \subseteq S \in \check{C}ech(r) \iff$ smallest ball enclosing σ has radius $\leq r$.

Def The **miniball** of a set $\sigma \subseteq S$ is the smallest closed ball containing σ.

similar to circumsphere/circumcircle



Hence, radius of miniball of $\sigma \leq r \iff \sigma \in \check{C}ech(r)$.

To build (or define) $\check{C}ech(r)$, we need to check intersections of multiple (≥ 3) balls. This step could be computationally expensive, especially in large sets of points, since we have to go upto checking all points together in the data set!
But, here is a better option.

Vietoris-Rips Complexes Instead of checking the intersection of all balls, if we check just pairwise intersections, and add 2- or higher dimensional simplices whenever all edges are in, we get the **Vietoris-Rips** or VR complex.

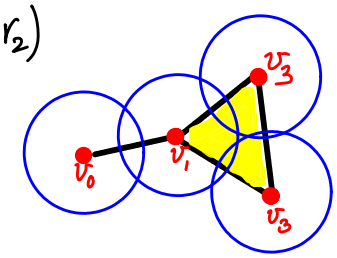
We write $VR_S(r) = \{\sigma \subseteq S \mid \operatorname{diam} \sigma \leq 2r\}$.

or Vietoris-Rips$_S(r)$

diameter of σ

Def The **diameter** of σ is the supremum of all pairwise distances between points in σ.

$VR(r_2)$



Compared to $\check{C}ech(r_2)$, we add $\Delta v_1 v_2 v_3$ to the Vietoris-Rips complex at $r = r_2$.

How do $\check{C}ech(r)$ and $VR(r)$ compare?

Naturally, $\check{C}ech_S(r) \subseteq VR_S(r)$. But notice that $VR_S(r_2)$ does not have a hole, as $\Delta v_1 v_2 v_3$ is included. At the same time, $\left| \bigcup_{i=0}^{q} B_{v_i}(r_2) \right|$ does have a hole, and so does $\check{C}ech(r_2)$.

So, homotopy is not preserved in $VR_S(r_2)$. Nonetheless, we get an inclusion going the other way, i.e., $VR(r) \subseteq \check{C}ech(r')$, at a larger radius r'.

<u>Vietoris-Rips Lemma</u> Let S be a finite set of points in $\mathbb{R}^d$, and let $r \geq 0$. Then

$$VR_S(r) \subseteq \check{C}ech(\sqrt{2}r).$$

The inclusions going both ways mean that VR and $\check{C}ech$ complexes are "quite comparable" when we consider all possible radii $(-\infty < r < \infty)$. While we may not get the same series of complexes, either family would be sufficient for most topological computations of interest. Hence, VR complexes are almost always preferred for computations, while $\check{C}ech$ complexes are sometimes preferred when used in proofs.

Proof (IDEA)

Consider Δ^d, the regular d-simplex in $\mathbb{R}^{d+1}$. Each vertex is a unit vector in this space. Thus,

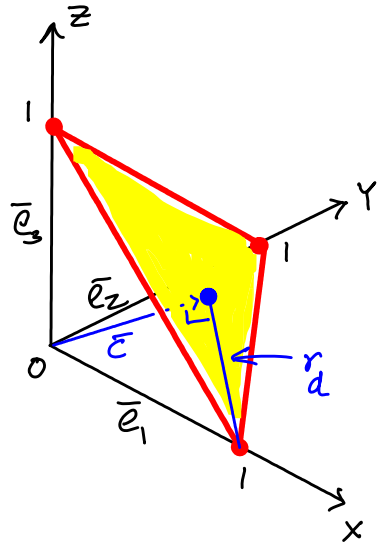
$\Delta^d = \text{conv}(\bar{e}_1, \ldots, \bar{e}_{d+1})$, where $\bar{e}_j$ is the j^{th} unit vector in $\mathbb{R}^{d+1}$.

Regular simplices are the "limiting" cases to consider here, due to their symmetry.

Let $\bar{c}$ be the barycenter of Δ^d.

$$\bar{c} = \begin{bmatrix} \frac{1}{d+1} \\ \vdots \\ \frac{1}{d+1} \end{bmatrix} \qquad \|\bar{c}\| = \frac{1}{\sqrt{d+1}}$$ is the length from origin of Δ^d.

↳ perpendicular distance



We compute $r_d = \sqrt{\frac{d}{d+1}} \left(= \sqrt{1 - \|\bar{c}\|^2}\right)$.

Note: $r_d \longrightarrow 1$ as $d \to \infty$.

The pairwise distance between $\bar{e}_i$ and $\bar{e}_j$ in σ is $\sqrt{2}$.

Also, the miniball of Δ^d has radius r_d.

Hence, simplex Δ^d of diameter $\sqrt{2}$ also belongs to $\check{C}ech(r_d)$. Multiplying by $\sqrt{2}r$, we get that,

$VR(r) \subseteq \check{C}ech(\sqrt{2}r\, r_d)$. But $r_d \leq 1$,

and hence $VR(r) \subseteq \check{C}ech(\sqrt{2}r)$. □

We saw $\check{C}ech_S(r)$ and $VR_S(r)$ of S (points) in $\mathbb{R}^d$.

Can we limit the dimension of the simplices we get from Nrv S? Yes!

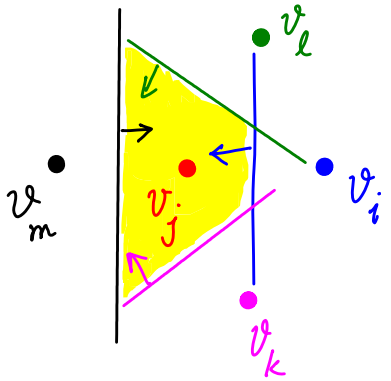
We can build the Delaunay complex. We first describe its dual construction.

Voronoi Diagram

Recall: $S = \{\bar{v}_1, \ldots, \bar{v}_n\}$ is a finite set of points in $\mathbb{R}^d$.

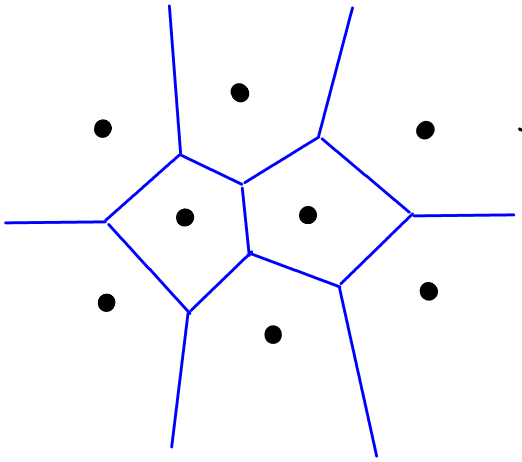
The Voronoi cell of $\bar{v}_j \in S$ is the set of points in $\mathbb{R}^d$ closest to $\bar{v}_j$:

$$V_{\bar{v}_j} = \{\bar{x} \in \mathbb{R}^d \mid \|\bar{x} - \bar{v}_j\| \leq \|\bar{x} - \bar{v}_i\| \; \forall \, \bar{v}_i \in S\}.$$



When we have just two points, say, v_i and v_j, the perpendicular bisector between them is the set of points equidistant from both of them. The half plane on the side of v_j then is V_{v_j}, its Voronoi cell.

$V_{\bar{v}_j}$ is a convex polyhedron, as it is the intersection of a set of half spaces, each being convex. $V_{\bar{v}_j}$ for all $\bar{v}_j \in S$ together tile or cover all of $\mathbb{R}^d$.



The collection of $V_{\bar{v}_j}$ for all $\bar{v}_j \in S$ is called the **Voronoi diagram** of S.

$V_{\bar{v}_i}$ and $V_{\bar{v}_j}$ meet at most in a common boundary. In $\mathbb{R}^2$, Voronoi cells meet at points or edges.

Notice that $V_{\bar{v}_j}$ can be open or closed. Intuitively, the boundary of $V_{\bar{v}_j}$ can be thought of as the "fence" around $\bar{v}_j$'s "house" – everything within the fence "belongs" to $\bar{v}_j$.

Delaunay Triangulation

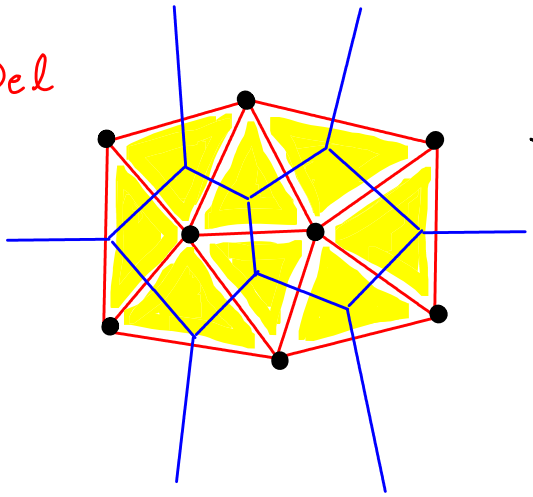
The **Delaunay complex** of S is (isomorphic to) the nerve of its Voronoi diagram.

$$\mathrm{Del}_S = \left\{ \sigma \subseteq S \;\middle|\; \bigcap_{\bar{v}_j \in \sigma} V_{\bar{v}_j} \neq \emptyset \right\}.$$

or $\mathrm{Delaunay}_S$

Similar to the Čech complex, we start with a convex set or cell associated with each point in S, and then take the nerve. But instead of balls, we use the Voronoi cells for each vertex.

Del_S



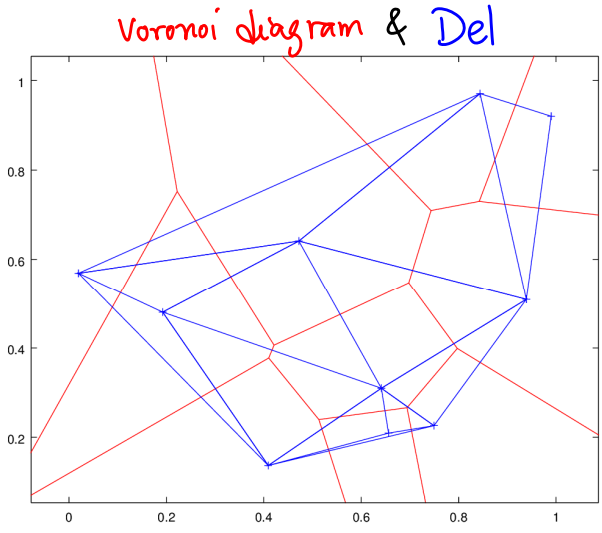
Shown here is one geometric realization of $\mathrm{Delaunay}_S$. This is the "natural" realization, as well.

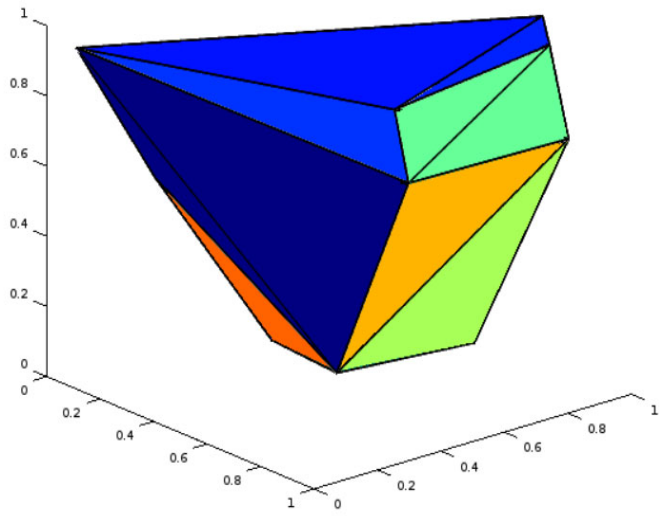
MATH 529 – Lecture 12 (02/15/2024)

Today:
* filtration
* alpha complexes
* weighted alpha complexes

Check out the commands voronoi, delaunay, and delaunayn, as well as related commands in Matlab. Similar commands are available in Python as well.

Here are a couple sample pictures of 2D and 3D Delaunay complexes produced in Matlab. The 2D picture shows the Voronoi diagram as well.

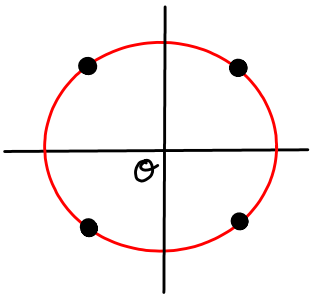




Looks like we only get (upto) triangles here. Recall one of the main motivations we stated for introducing Delaunay complexes – that we wanted to get only upto d-simplices for point sets S in $\mathbb{R}^d$.

Q. Do we always get only triangles in Del_S is 2D?

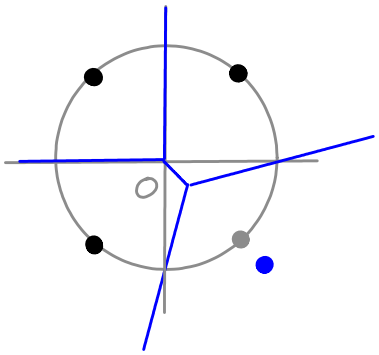
No!



The Voronoi cells V_{v_j} of the four points meet at the central point o (origin) here.

Hence, Del_S contains the tetrahedron!

But, if we move even one of the four points just ever so slightly away from the circle, we can avoid the 4-way intersection of their Voronoi cells.



Mathematically, we need to move only one (out of the four) points by $\varepsilon > 0$ in one of the coordinate directions; ε could be really small, as long as it is > 0.

Def The set of points S in $\mathbb{R}^d$ is in **general position** if no $(d+2)$ points in S lie on a common $(d-1)$-sphere.

e.g., $d=2 \Rightarrow$ no 4 points lie on a circle.

No $(d+2)$ V_{v_j}'s have a common intersection $\Rightarrow$ ($\sigma \in Del_S$ means $\dim \sigma \leq d$)

We assume general position usually. If this condition is not satisfied, we could perturb a single coordinate of a single point from any $(d+2)$ such points by a small $\epsilon > 0$.

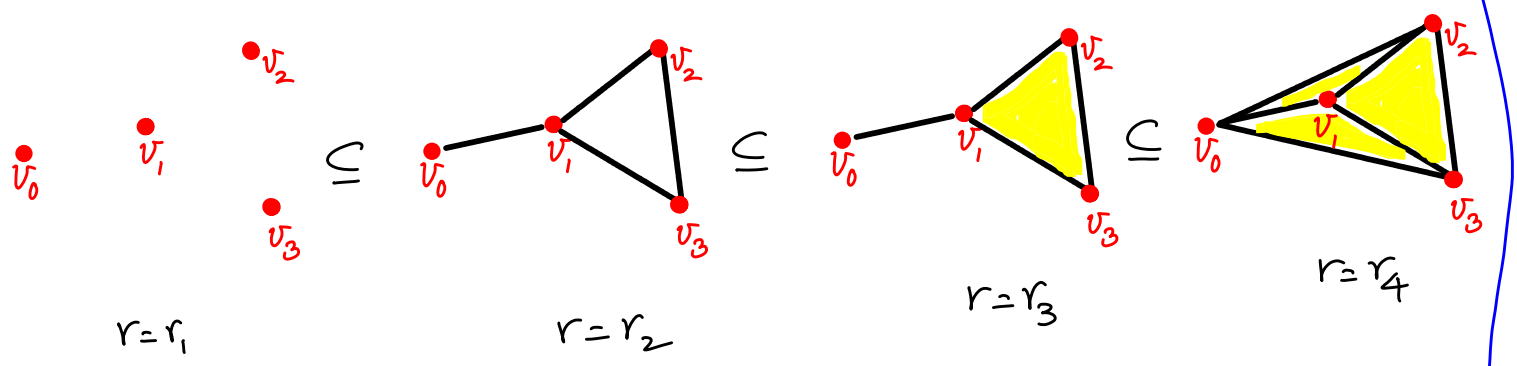
There are very efficient (polynomial time) algorithms for constructing Delaunay tessellations, at least in 2D and 3D.

We have seen 3 families of simplicial complexes: $\check{C}ech_S(r)$, $VR_S(r)$, and Del_S. Note that the Delaunay complex is independent of any distance cutoffs (or radii of balls).

But, what do we gain by varying r?

Given S, we could study the family of $\check{C}ech_S(r)$ or $VR_S(r)$ as r varies from 0 to ∞ (or, even from $-\infty$ to $+\infty$).

Since S is finite, we will only a finite number of such complexes.



$\subseteq$ tetrahedron $v_0v_1v_2v_3$ (at $r = r_5$)

→ Since there are only four points here, their tetrahedron is the largest dimensional simplex we get in $\check{C}ech(r)$, even if we keep increasing r beyond r_5.

We capture the fact that v_1, v_2, v_3 are closer to each other than, say, $\{v_2, v_3\}$ are to v_0, since $\Delta v_1v_2v_3$ comes in to $\check{C}ech(r)$ before the other triangles.

We will study such families of complexes in detail.

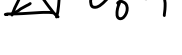
Def A **filtration** of a simplicial complex K is a nested sequence of subcomplexes of K

$$\phi = K^0 \subseteq K^1 \subseteq \dots \subseteq K^m = K.$$

The simplicial complex K along with a filtration is a **filtered simplicial complex**.

A **filtration ordering** is a full ordering of all simplices in K such that each prefix of the ordering is a subcomplex.

Filtration ordering: Example

$K =$ 

a subcomplex

$$v_0 < v_1 < v_2 < v_3 < v_0v_1 < v_0v_2 < v_0v_3 < v_1v_2 < v_1v_3 < v_2v_3$$
$$< v_0v_1v_2 < \dots < v_0v_1v_2v_3.$$

More generally, '$<$' could assign the same rank for several simplices, e.g., all vertices $<$ all edges $<$ all triangles $< \dots$.
In this case, $<$ is not a full ordering but we can convert it to one by breaking ties arbitrarily.

Q. What do we use filtrations for?

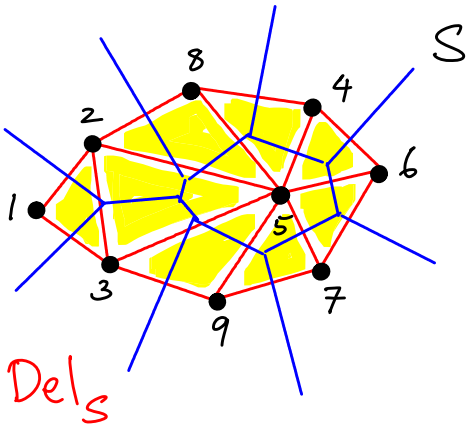
We could study signature functions: $\lambda : \{0, 1, \dots, m\} \longrightarrow \mathbb{R}^d$, $d \geq 1$. λ assigns a value in $\mathbb{R}^d$ for each $k \in \{0, \dots, m\}$. We could compare the signatures for two point sets S_1 and S_2 to distinguish them — by comparing $\lambda(S_1)$ and $\lambda(S_2)$.

For instance, χ (Euler characteristic). We could compute $\chi(K^i)$ for each $i \in \{0, \dots, m\}$, and study the Euler characteristic curve (or vector).

But we will study other more involved signatures soon.

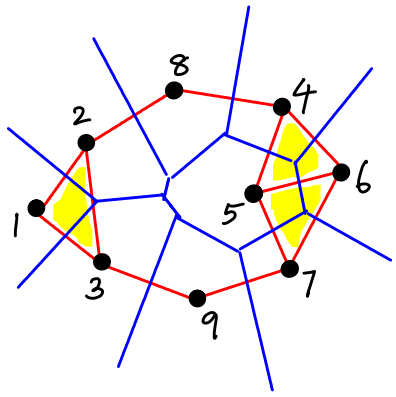
Could we vary r to create a family of nested Delaunay complexes? Equivalently, could we create a filtration for Del_S?

Consider S with 9 points shown here:



Del_S

Observe: $\{1,2,3\}$ and $\{4,5,6,7\}$ form two clusters of nearby points, which are further away from each other.



A subcomplex of Del_S as shown here would capture the topology of S "better".

For the given set of 9 points, how do we define the complex shown here?

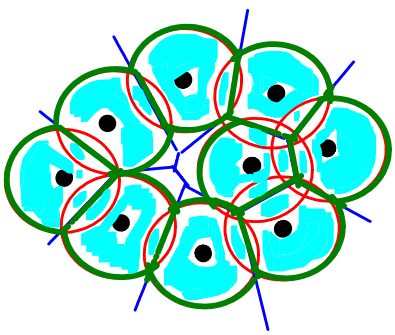
Alpha Complexes

Let $S = \{\bar{v}_1, \dots, \bar{v}_n\}$, $\bar{v}_j \in \mathbb{R}^d$, $r \geq 0$. Recall that $B_{\bar{v}_j}(r) = \bar{v}_j + rB^d$

$= \{\bar{x} \in \mathbb{R}^d \mid \|\bar{x} - \bar{v}_j\| \leq r\}$ is the r-ball (d-dimensional) centered at $\bar{v}_j$.

We "combine" balls around the points, and their Voronoi cells. We consider $B_{\bar{v}_j}(r) \cap V_{\bar{v}_j}$.

$\bigcup_{\bar{v}_j \in S} R_{\bar{v}_j}(r)$



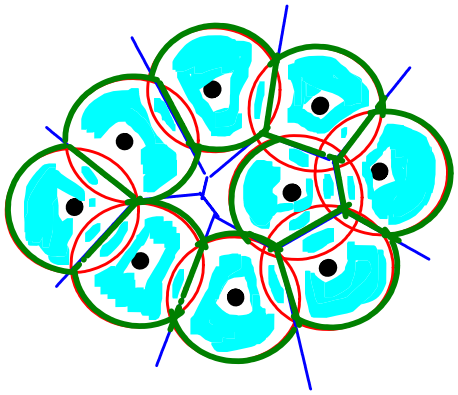
Note that there is a "hole" in the union of regions (in the middle).

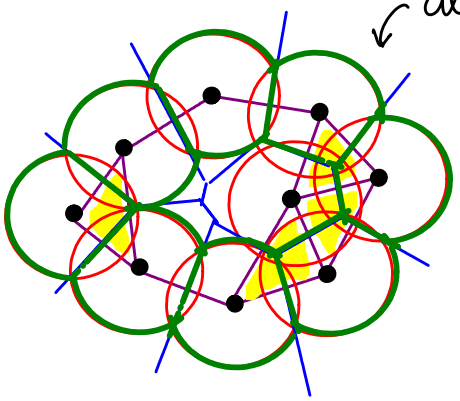
The idea is to combine the good properties of balls and Voronoi cells – get at most d-simplices, but still get a hierarchy.

Let $R_{\bar{v}_j}(r) = B_{\bar{v}_j}(r) \cap V_{\bar{v}_j}$. (region of $\bar{v}_j$) Since both $B_{\bar{v}_j}(r)$ and $V_{\bar{v}_j}$ are convex, so is $R_{\bar{v}_j}(r)$. The regions $R_{\bar{v}_j}(r)$ intersect, if at all, along common boundaries, and together they tile $\bigcup_{\bar{v}_j \in S} B_{\bar{v}_j}(r)$. (or, cover)

The **alpha complex** is the nerve of this union.

$$\text{Alpha}_S(r) = \left\{ \sigma \subseteq S \;\middle|\; \bigcap_{\bar{v}_j \in \sigma} R_{\bar{v}_j}(r) \neq \emptyset \right\}.$$





$R_{\bar{v}_j}(r) \subseteq V_{\bar{v}_j} \Rightarrow \text{Alpha}(r) \subseteq \text{Del}_S$. Also,

$R_{\bar{v}_j}(r) \subseteq B_{\bar{v}_j}(r) \Rightarrow \text{Alpha}(r) \subseteq \check{\text{C}}\text{ech}_S(r)$.

By the nerve lemma, $\bigcup_{\bar{v}_j \in S} B_{\bar{v}_j}(r) \simeq |\text{Alpha}(r)|$.

By varying r and considering Alpha(r) $\forall r$, we get a filtration of the Delaunay complex.

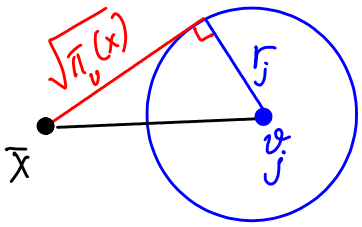
Q: Can we use balls of different radii?

Motivation: modeling proteins – made up of atoms, which could be modeled using balls of different radii.

Weighted Alpha Complexes

Def The weighted squared distance, or the **power distance**, of $\bar{x} \in \mathbb{R}^d$ from $\bar{v}_j$ with weight w_j is

$$\pi_{\bar{v}_j}(\bar{x}) = \|\bar{x} - \bar{v}_j\|^2 - w_j.$$



When $w_j = r_j^2$, we get $\pi_{\bar{v}_j}(\bar{x})$ is the squared length of the tangent from $\bar{x}$ to $B_{\bar{v}_j}(r_j)$.

The **weighted** or **power Voronoi** cell of $\bar{v}_j$ is

$$W_{\bar{v}_j} = \{\bar{x} \in \mathbb{R}^d \mid \pi_{\bar{v}_j}(\bar{x}) \leq \pi_{v_i}(\bar{x}) \ \forall \bar{v}_i \in S\}.$$

weighted Voronoi diagram

The **power Voronoi complex** is the collection of $W_{\bar{v}_j}$'s. And the **weighted Delaunay complex** is the nerve of the **weighted Voronoi diagram**.

We could apply the concept of weighted Voronoi diagram also to sets of points where different points have different weights (and not only to cases where the balls are differently sized).

The definition of power distance appears somewhat involved – we will describe the motivation soon!

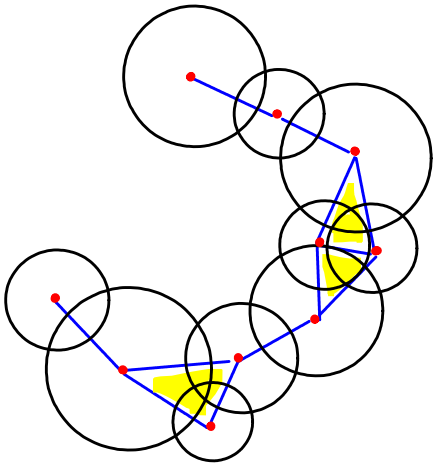
Weighted Alpha Complex

We set $w_j = r_j^2$ and define

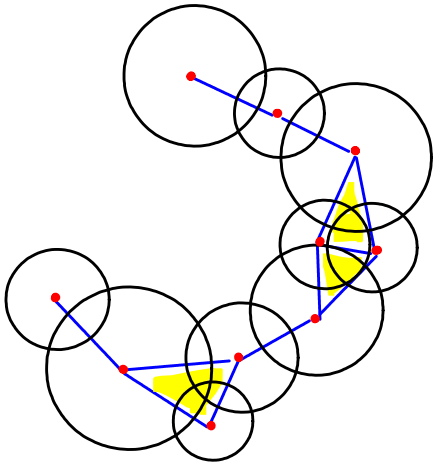
$$R^w_{\bar{v}_j}(r) = B_{\bar{v}_j}(r) \cap W_{\bar{v}_j}$$

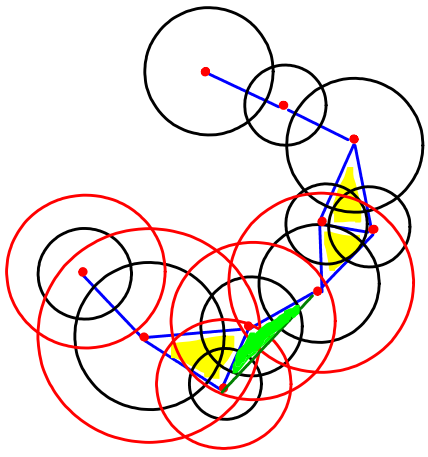
The **weighted alpha complex** is the nerve of the collection of $R^w_{\bar{v}_j}(r)$. This complex is a subcomplex of the weighted Delaunay complex.

Here is an illustration. The discs could model different atoms, and the weighted alpha complex shown here is one "skeleton" of the protein, which is the collection of atoms.



Here is an illustration of how the weighted alpha complex grows as we increase r_j^2 linearly, i.e., we set $w_j = r_j^2 \leftarrow r_j^2 + r$ and let $r \to \infty$.





A subset of the bigger balls (2D-discs in this case) are shown here for illustration. The extra triangle added to the original nerve is shown in green.

MATH 529 – Lecture 13 (02/20/2024)

Today:
* Varying r_j in weighted alpha complexes
* Empty circle property of Delaunay triangulation
* Witness complexes

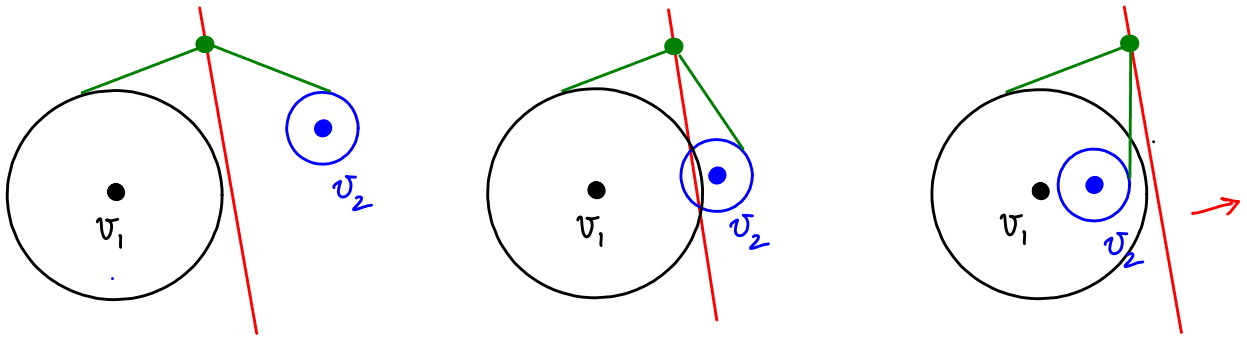
Recall: power distance $\pi_{\bar{v}_j}(\bar{x}) = \|\bar{x}-\bar{v}_j\|^2 - w_j$, $w_j = r_j^2$, $W_{\bar{v}_j}$, $R_{\bar{v}_j}(r) = W_{\bar{v}_j} \cap B_{\bar{v}_j}(r)$...

We could "vary" the different radii to get a weighted alpha complex filtration $\emptyset = K^0 \subseteq \cdots \subseteq K^m$ = weighted Delaunay complex.

Q: How to "vary" the radii? r_j are not same to start with.

1. Set $w_j = r_j^2$, and then increase all radii r_j at the same linear rate, i.e., $r_j \leftarrow r_j + r$ for $r > 0$. Then let $r \to \infty$ (uniformly increase all r_j).
But, $W_{\bar{v}_j}$ for different r may not be the same. Hence, it could happen that $K^j \not\subseteq K^{j+1}$ for some j. This situation is best avoided!

Here is an observation: Bisector of two weighted points



Note that the bisector stays a straight line!

Similar to the default alpha complex construction, where the Voronoi cells stay the same while the balls grow, it is desirable to have the weighted Voronoi cells stay same as well.

2. We set $w_j = r_j^2$, and grow the square of the radii uniformly, i.e., set $r_j^2 \leftarrow r_j^2 + r$, as $r \to \infty$.

Since we are using the power distance,

$$\pi_{\bar{v}_j}(\bar{x}) = \pi_{\bar{v}_i}(\bar{x}) \Rightarrow \|\bar{x} - \bar{v}_j\|^2 - (r_j^2 + r) = \|\bar{x} - \bar{v}_i\|^2 - (r_i^2 + r)$$

Hence, the bisectors using $\pi_{\bar{v}}(\bar{x})$ stay the same as $r \to \infty$. So, the power Voronoi cells remain the same, just like $V_{\bar{v}_j}$.

So, as r increases, $W_{\bar{v}_j}$ remains same. We do get the nesting of simplicial complexes as r increases.

In fact, $W_{\bar{v}_j}$'s here have most of the nice properties that $V_{\bar{v}_j}$, the default Voronoi cells, have. As such, the alpha complex filtration also has most nice properties that the default (same r for each r_j) alpha complexes.

Originally, Edelsbrunner and Mücke (1983) defined the weighted alpha complexes to study structure of biomolecules. The notation, used was $r_j^2 \leftarrow r_j^2 + \alpha$ for the growth parameter α $(-\infty < \alpha < \infty)$. Hence the name alpha complex.

There are efficient algorithms to construct the weighted alpha complexes in 2D and 3D. We will discuss a version in a future lecture.

For large sets of points, all of the complexes we introduced — Čech, VR, alpha, etc., become intensive to compute. We would be better off sampling a subset of points!

Čech and VR complexes grow too large even for moderately large point sets S. For instance, the VR complex of a set with ~2000 points in $\mathbb{R}^3$ could have more than a million triangles! Further, computing Delaunay and alpha complexes are also computationally expensive in high dimensions. We look at a possible alternative now.

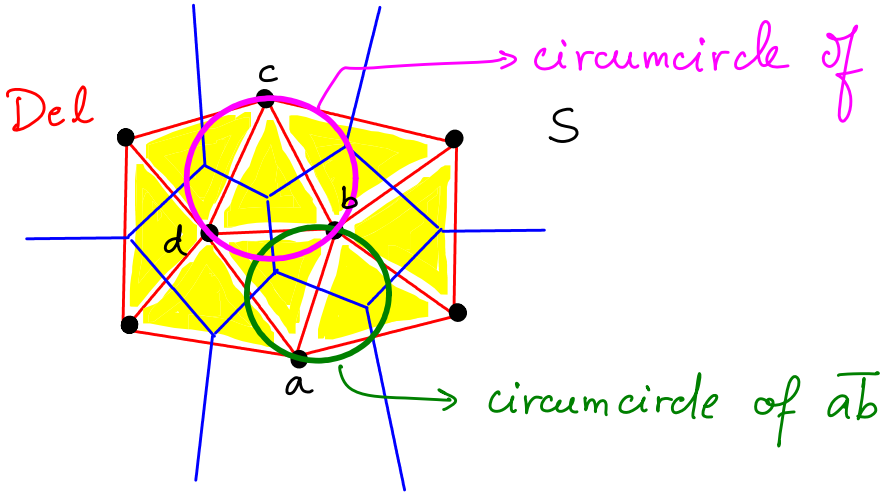
We will now consider two families of complexes that are designed to be much more sparse — they sample from the input point set, and build the complex by generalizing or relaxing some of the conditions used to define the complexes we have already seen.

A crucial property of the Delaunay complex

The **empty circumsphere** property: (boundary of miniball)

$\sigma \in \mathrm{Del}_S \iff$ circumsphere(σ) has no points of S in its interior.

Of course, with $\sigma = \mathrm{conv}\{\bar{v}_0, \dots, \bar{v}_k\}$, $\bar{v}_0, \dots, \bar{v}_k$ lie on the circumsphere, and the center of its circumsphere is in the intersection of the Voronoi cells of $\bar{v}_0, \dots, \bar{v}_k$.



Witness Complex

Idea: Choose $L \subseteq S$ (L is typically small), and build your complex on L. Use the remaining points in $S \setminus L$ as possible "witnesses" for the simplices in the complex.

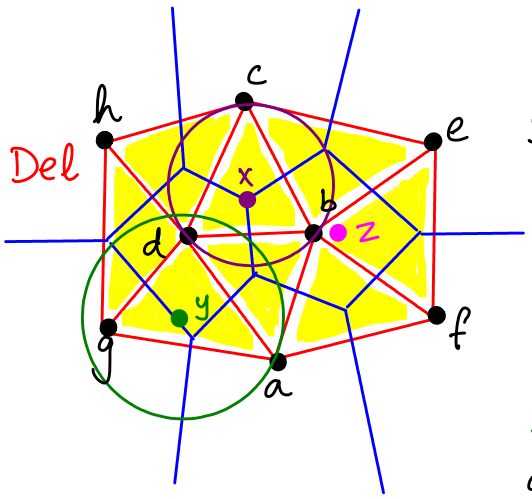
Def Let $\sigma = [\bar{v}_0 \dots \bar{v}_k]$, $\bar{v}_j \in S$, be a k-simplex, where $\bar{v}_j \in \mathbb{R}^d$. $\bar{x} \in \mathbb{R}^d$ is a **weak witness** for σ with respect to S if $\|\bar{x} - \bar{v}\| \leq \|\bar{x} - \bar{u}\|$ $\forall$ $\bar{v} \in \{\bar{v}_0, \dots, \bar{v}_k\}$ and $\bar{u} \in S \setminus \{\bar{v}_0, \dots, \bar{v}_k\}$.

$\bar{x} \in \mathbb{R}^d$ is a **strong witness** for σ w.r.t S if it is a weak witness, and in addition, $\|\bar{x} - \bar{v}_0\| = \|\bar{x} - \bar{v}_1\| = \dots = \|\bar{x} - \bar{v}_k\|$.

Equivalently, we say that σ is weakly (or strongly) witnessed by $\bar{x}$.

We could define the Delaunay complex in terms of existence of strong and weak witnesses for the simplices. Subsequently, we could relax (some of) the requirements to build complexes that are more manageable in size.

We first illustrate weak and strong witness points on the Delaunay complex we have seen previously.



x is a strong witness for $\triangle bcd$. Notice that x is the center of the circumcircle of $\triangle bcd$, which is also the (point of) intersection of the Voronoi cells of the vertices b, c, and d.

y is a weak witness for $\overline{dg}$, and also for $\triangle adg$. Notice that as drawn, $\|y-g\| < \|y-d\| < \|y-a\| < \|y-v\|$ for $v = b, c, e, f, h$.

z is a weak witness for $\triangle abe$. Also, $\triangle abe$ as shown does not have a strong witness. Intuitively, the center of the circumcircle of $\triangle abe$ lies closer to f than to a, b, and e.

Indeed, every simplex in Del_S will have a strong witness, which is the center of the empty circumsphere of that simplex. Ideally, we want to sample from S, and look for witnesses for simplices in the points not included in the sample. At same time, looking for a strong witness among a given set of vertices ($\subseteq S$) is futile, as such a point occurs with zero probability. But the following result bails us out.

Result (de Silva, 2003) $\sigma = [v_0 \cdots v_k] \subseteq S$ has a strong witness iff every $\tau \leq \sigma$ (face) has a weak witness.

Notice that one direction is obvious – if σ has a strong witness, the same point is a strong witness for all its faces too, and hence every face has a weak witness. The other direction is more technical – see the paper posted on the course web page for details.

We define witness complexes in the more general setting of a metric space with pairwise distances between points provided.

Let $D = [d_{ij}]$ be the $l \times n$ distance matrix between $L \subseteq S$ of landmark points and all points in S. Here, $|L| = l$, and $|S| = n$.

Def The (strict) **witness complex** $W_\infty(L,S)$ is the collection of all simplices $\sigma \subseteq L$ whose all subsimplices have weak witnesses in S.

This restriction of all faces having weak witnesses is required to insure that $W_\infty(L,S)$ is a simplicial complex. Notice that we are building simplices on points just from L, and not from all of S.

In particular, if $\sigma = [v_0 \dots v_k] \in W_\infty(L,S)$, then there exists a j with $1 \le j \le n$ such that d_{ij} for $i = v_0, \dots, v_k$ are the $(k+1)$ smallest entries in the j^{th} column of D in some order.

We also say that $\bar{u}_j \in S$ (corresponding to the j^{th} column) is a witness to the existence of σ in $W_\infty(L,S)$.

Relationship to Del_L

Result $\sigma \in Del_L$ iff σ is strongly witnessed.
(just follows from the empty circumsphere property).

But $\sigma \in W_\infty(L,S)$ implies that σ is strongly witnessed, as all its faces have weak witnesses. Hence, we get that

$$W_\infty(L,S) \subseteq Del_L.$$

So, we are first choosing a possibly much smaller set L of points from S as landmarks. We then build a complex on L which also has a bound on the dimension of the simplices being included.

Similar to how we defined the Vietoris-Rips complex by relaxing the definition of the Čech complex, we now define an easier to construct version of the strict witness complex by requiring that only edges need to be present for a higher dimensional simplex to be included.

Def The **lazy witness complex** $W_1(L,S) \supseteq W_\infty(L,S)$ has the same 1-skeleton as $W_\infty(L,S)$. After that, $\sigma = [v_0 \dots v_k] \subseteq L$ is in $W_1(L,S)$ iff all edges of σ are in $W_1(L,S)$.

In practice, we almost always work with the lazy witness complex $W_1(L,S)$, and write $W(L,S)$ to mean $W_1(L,S)$ (and not $W_\infty(L,S)$).

Q: Is $W_1(L,S) \subseteq \mathrm{Del}_L$? If not, can you give a counterexample?

Think about it!

How to choose the landmarks L

First decide how many landmarks you want ($|L| = \ell$). Then,

two methods:
- random selection (select ℓ points randomly)
- maxmin selection.

Maxmin selection of ℓ landmarks

Choose the first landmark ℓ_1 randomly. After that, inductively, with $\{\ell_1, \dots, \ell_{i-1}\}$ chosen, pick ℓ_i that maximizes the following function:

$$z \mapsto \min\{D(z, \ell_1), D(z, \ell_2) \dots, D(z, \ell_{i-1})\}, \text{ where}$$

$D(z, \ell_j)$ is the distance from z to ℓ_j.

Maxmin provides widespread coverage of S, but could also end up picking outliers.

Guidelines for choosing $\ell = |L|$

(de Silva, Carlsson, 2004) For data sampled from surfaces (in 3D), $\ell \leq \frac{n}{20}$ works reasonably well.

Javaplex and Gudhi provide functions to build witness complexes.

https://github.com/appliedtopology and https://gudhi.inria.fr/.

Now that we have seen how to build several simplicial complexes on point sets, we will talk about how to infer the topology of the built complex. We first review basic results from algebra, and use them to define and study groups on the simplicial complex.

Review of Algebra (Groups)

A binary operation $*$ on a set S is a rule that assigns to each ordered pair $(a,b) \in S$ some element in S.

e.g., $a * b = c$, for $c \in S$.

If $a*b = b*a \ \forall a,b \in S$, $*$ is **commutative**.

$(a*b)*c = a*(b*c) \ \forall a,b,c \in S \Rightarrow$ $*$ is **associative**.

A **group** $\langle G, * \rangle$ is a set G with a binary operation $*$ defined on elements of G such that the following conditions hold.

(a) $*$ is associative.

(b) $\exists\, e \in G$ such that $e*a = a*e = a \ \forall a \in G$.
e is an identity element for $*$ on G.

(c) $\forall\, a \in G$, $\exists\, a' \in G$ such that $a*a' = a'*a = e$.
a' is the inverse of a with respect to $*$.

We assume that G is **closed under** $*$ to begin with, i.e., $\forall a,b \in G$, $a*b = c$ for some $c \in G$.

If G is finite, the **order** of the group is $|G|$. Oftentimes, G itself is used to denote the group, with operation $*$ understood.

e.g., $\langle \mathbb{Z}, + \rangle$, $\langle \mathbb{Z}_4, +_4 \rangle$

$\mathbb{Z}$: integers

$\mathbb{Z}_4$: $\{0,1,2,3\}$, $+_4$: add modulo 4

$\mathbb{Z}_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

If $*$ is commutative, G is an **Abelian group**. We also say that G is abelian.

$\mathbb{Z}$ and $\mathbb{Z}_4$ are both abelian. Also, notice that the order of $\mathbb{Z}_4$ is 4.

MATH 529 – Lecture 14 (02/22/2024)

Today: * more on groups, homomorphisms
* chains and chain groups on simplicial complexes

Recall group G with operation $*$...

Induced operation Let $\langle G, * \rangle$ be a group, and $S \subseteq G$. If S is closed under $*$, then $*$ is the **induced operation** on S from G.

Subgroup $H \subseteq G$ (subset) is a **subgroup** of $\langle G, * \rangle$ if H is a group and is closed under $*$.

$\{e\}$ is the trivial subgroup of G. All other subgroups are nontrivial.

Theorem $H \subseteq G$ of a group $\langle G, * \rangle$ is a subgroup of G if and only if
(a) H is closed under $*$;
(b) identity e of G is in H; and
(c) $\forall a \in H$, $a^{-1} \in H$.

This theorem could also be used as the definition of a subgroup.

e.g., Consider $\mathbb{Z}_4$

$H = \{0, 2\}$ is a subgroup.

Notice that $2+2 = 0 \pmod 4$, and hence H is indeed closed under the operation in question (addition modulo 4).

$\mathbb{Z}_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$H = \{0, 2\}$ is the only nontrivial subgroup of $\mathbb{Z}_4$ (distinct from $\mathbb{Z}_4$ itself).

Cosets Let H be a subgroup of G. Let the relation $\sim_L$ be defined on G by $a \sim_L b$ iff $a^{-1}b \in H$. Similarly, $\sim_R$ is defined on G by $a \sim_R b$ iff $ab^{-1} \in H$. Note that $\sim_L$ and $\sim_R$ are equivalence relations on G. Also ($a^{-1}b \to a^{-1} * b$)

$$a^{-1}b \in H \text{ or } a^{-1}b = h \in H \implies b = ah.$$

For $a \in G$, the subset $aH = \{ah \mid h \in H\}$ of G is the **left coset** of H containing a, and $Ha = \{ha \mid h \in H\}$ is the **right coset** of H containing a.

If G is abelian, then $ah = ha \ \forall a \in G, h \in H$. Then the left and right cosets match, i.e., $aH = Ha$.

e.g., $\mathbb{Z}_4$, $H = \{0,2\}$ is a subgroup. The coset of 1 is $1 +_4 \{0,2\} = \{1,3\}$.

Our goal is to use groups to characterize topological spaces. Hence, we need to be able to characterize the "structure" of groups. We could use maps between groups for this purpose.

To simplify notation, we write ab for $a * b$, with $*$ understood.

Homomorphisms A map $\varphi : G \to G'$ is a **homomorphism** if $\varphi(ab) = \varphi(a)\varphi(b) \ \forall a, b \in G$.

$\varphi(a *_G b) = \varphi(a) *_{G'} \varphi(b)$

We can always define a trivial homomorphism by setting $\varphi(g) = e' \ \forall\, g \in G$, where e' is the identity of G'.

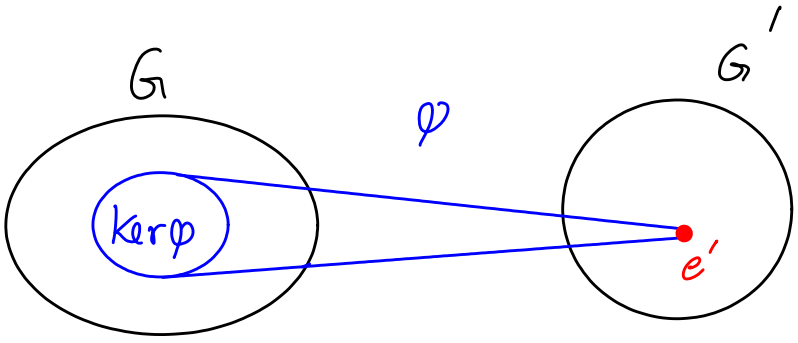
Homomorphisms preserve identity, inverse, and subgroups.

Theorem Let $\varphi: G \to G'$ be a homomorphism. Then

1. $\varphi(e) = e'$, where e, e' are identities of G, G', respectively.
2. $\varphi(a^{-1}) = (\varphi(a))^{-1} \ \forall\, a \in G$.
3. $H \subseteq G$ is a subgroup of $G \Rightarrow \varphi(H)$ is a subgroup of G'.
4. $K' \subseteq G'$ is a subgroup of $G' \Rightarrow \varphi^{-1}(K')$ is a subgroup of G.

This theorem could be used, alternatively, as a definition of homomorphism.

Kernel Let $\varphi: G \to G'$ be a homo'm. The subgroup $\varphi^{-1}(\{e'\}) \subseteq G$ is the **kernel** of φ.



Notice that $\{e'\}$ is the trivial subgroup of G'. Hence by (4) of the Theorem above, $\ker\varphi$ is a subgroup of G.

Since $\ker\varphi$ is a subgroup of G_1, we can define kernel cosets. Let $H = \ker\varphi$, $a \in G_1$, then

$$aH = \varphi^{-1}\{\varphi(a)\} = \{x \in G_1 \mid \varphi(x) = \varphi(a)\} = Ha.$$

Intuitively, any $h \in \ker\varphi$ gets mapped to the identity (e'). So, we could add h to a to get x, and x also gets mapped to the image of a.

If $gH = Hg$ $\forall g \in G_1$ for a subgroup H of G_1, then H is a **normal** subgroup of G_1. All subgroups of Abelian groups are normal, and so is $\ker\varphi$ for a homomorphism φ.

The properties of a function being injective, surjective, or both can be studied for homomorphisms as well. But for homomorphisms, we refer to these properties using terms specific to groups.

Maps in general	homomorphisms between groups
1-1	monomorphism
onto	epimorphism
1-1 and onto (bijection)	isomorphism ($\cong$) → notation

Isomorphism between groups is like homeomorphism between topological spaces. Recall previous discussion about ASCs being isomorphic.

Finitely generated Let $a_i \in G_1$ for $i \in I$, an index set. The smallest subgroup of G_1 containing $\{a_i \mid i \in I\}$ is the subgroup generated by $\{a_i \mid i \in I\}$. If this subgroup is all of G_1, then $\{a_i \mid i \in I\}$ **generates** G_1, and a_i are the generators. If I is finite, then G_1 is **finitely generated**.

For instance, both $\langle \mathbb{Z}, + \rangle$ and $\langle \mathbb{Z}_4, +_4 \rangle$ are finitely generated.

Homology Groups

We will talk about homology for simplicial complexes, i.e., simplicial homology. Homology studies "holes" in spaces, i.e., holes in 1D, enclosed voids in 2D, etc. Interestingly, holes are characterized by what surrounds them!

Let K be a simplicial complex. A **p-chain** (for $p \le \dim K$) of K is a formal sum (→ "linear combination") of the p-simplices of K.

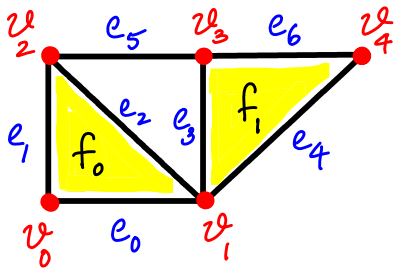
$$\bar{c} = \sum_{i=1}^{m} a_i \sigma_i$$

, K has m p-simplices, $\sigma_1, \dots, \sigma_m$, and a_i's are their coefficients.

To define groups using addition modulo 2, i.e., in $\mathbb{Z}_2$, $a_i \in \{0,1\}$. When using addition over $\mathbb{Z}$, $a_i \in \mathbb{Z}$. We could define homology groups over $\mathbb{Z}_2, \mathbb{Z}, \mathbb{R}, \mathbb{Q}$, (→ rational numbers) or over any ring.

Examples

K:



0-chain (over $\mathbb{Z}_2$)

$\bar{c}_0$:

$a_i = 1$ $i = 0,2,3$
$a_i = 0$, $i = 1,4$

$$\bar{c}_0 = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

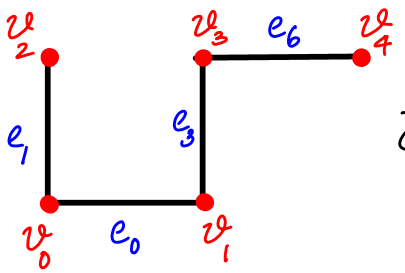
0-chain (over $\mathbb{Z}$)

$\bar{c}_0'$:

$$\bar{c}_0' = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}$$

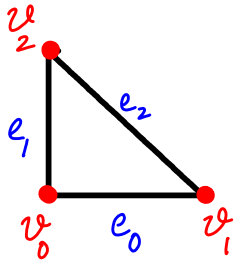
1-chains over $\mathbb{Z}_2$:

$\bar{c}_1$:

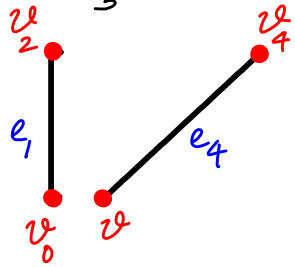


$$\bar{c}_1 = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$\bar{c}_2$:



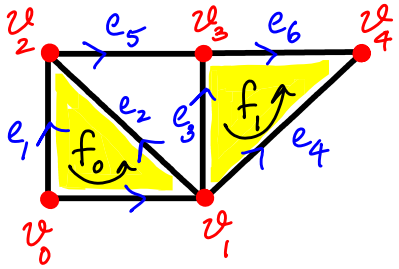
$\bar{c}_3$:



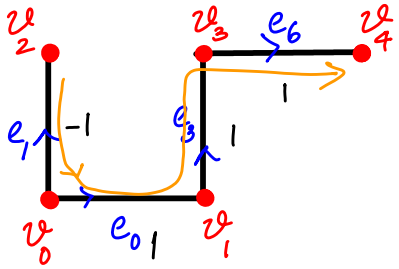
Note that the chain need not be connected.

We consider orientations over $\mathbb{Z}$.

K:



$\bar{c}_1'$:

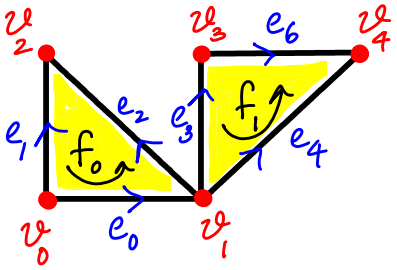


$$\bar{c}_1' = \begin{matrix}0\\1\\2\\3\\4\\5\\6\end{matrix}\begin{bmatrix}1\\-1\\0\\1\\0\\0\\1\end{bmatrix}$$

We could have an "overall orientation" for the 1-chain, but this is not needed always.

2-chains

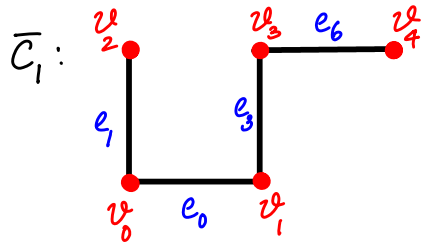
$\bar{d}_1 = \begin{matrix}0\\1\end{matrix}\begin{bmatrix}1\\1\end{bmatrix}$ is a 2-chain (over $\mathbb{Z}_2$ or $\mathbb{Z}$)



$\bar{d}_2 = \begin{bmatrix}2\\-3\end{bmatrix}$ is a 2-chain over $\mathbb{Z}$.

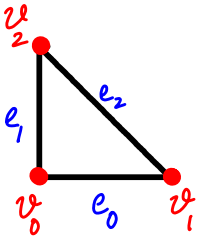
We can add two p-chains component- or p-simplex-wise, i.e., by adding their vectors. If $\bar{c} = \sum_{i=1}^{m} a_i\sigma_i$, $\bar{c}' = \sum_{i=1}^{m} b_i\sigma_i$, then $\bar{c} +_2 \bar{c}' = \sum_{i=1}^{m} (a_i +_2 b_i)\sigma_i$

$\bar{c}_1$:



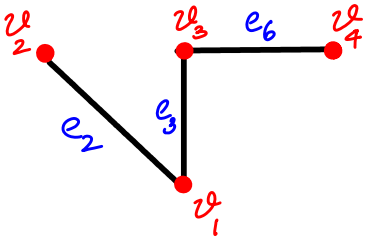
$+_2$

$\bar{c}_2$:



$=$

$\bar{c}_4$:



Equivalently, we add the corresponding vectors:

$$\begin{bmatrix}1\\1\\0\\1\\0\\0\\1\end{bmatrix}\begin{matrix}e_0\\e_1\\e_2\\e_3\\e_4\\e_5\\e_6\end{matrix} \quad \bar{c}_1 \qquad +_2 \qquad \begin{bmatrix}1\\1\\1\\0\\0\\0\\0\end{bmatrix}\begin{matrix}e_0\\e_1\\e_2\\e_3\\e_4\\e_5\\e_6\end{matrix} \quad \bar{c}_2 \qquad = \qquad \begin{bmatrix}0\\0\\1\\1\\0\\0\\1\end{bmatrix}\begin{matrix}e_0\\e_1\\e_2\\e_3\\e_4\\e_5\\e_6\end{matrix} \quad \bar{c}_4$$

p-chains with addition ($+_2$ over $\mathbb{Z}_2$ or $+$ over $\mathbb{Z}$) form the group of p-chains of K, denoted $\langle C_p(K), +_2 \rangle$ or $\langle C_p(K), \mathbb{Z} \rangle$, or simply $C_p(K)$ (or just C_p).

$C_p(K)$ is indeed a group, and is an Abelian group.

* identity: $\bar{0} = \sum_{i=1}^{m} 0\sigma_i$
* inverse: inverse of $\bar{c}$ is $-\bar{c}$ over $\mathbb{Z}$, and $\bar{c}$ over $\mathbb{Z}_2$ (as $\bar{c} +_2 \bar{c} = \bar{0}$).
* $+_2$ and $+$ are associative
* $+_2$ and $+$ are commutative.

For K, there is $C_p(K)$ for $0 \le p \le \dim(K)$.

If K has m p-simplices, each p-chain can be represented by an m-vector. The p-chains corresponding to the m unit vectors are the **elementary p-chains** of K, i.e., they correspond to each σ_i.

$$\bar{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \sigma_i$$

When K is finite, these m elementary chains generate $C_p(K)$, making it finitely generated.

How are $C_p(K)$ related for various p?

We use "boundary", defined as homomorphisms between $C_p(K)$ and $C_{p-1}(K)$ to connect these chain groups.

We first define the boundary of a single simplex, i.e., an elementary chain, and then extend it naturally to chains. We provide the definition over both $\mathbb{Z}_2$ and $\mathbb{Z}$.

Boundary The **boundary** of a p-simplex is the sum of all its $(p-1)$-faces.

→ same notation used for orientation, which we do not use over $\mathbb{Z}_2$.

If $\sigma = \text{conv}\{v_0, \dots, v_p\}$, or $\sigma = [v_0 v_1 \dots v_p]$, then

p-boundary → $\partial_p \sigma = \sum_{j=0}^{p} [v_0 \dots \hat{v}_j \dots v_p]$, where $\hat{v}_j$ means v_j is omitted.

Over $\mathbb{Z}$, we have $\partial_p \sigma = \sum_{j=0}^{p} (-1)^j [v_0, \dots, \hat{v}_j, \dots v_p]$, which is the sum of all its $(p-1)$-faces with their induced orientations.

Notice that $\partial_p \sigma$ is a $(p-1)$-chain (in both cases).

MATH 529 – Lecture 15 (02/27/2024)

Today: * boundary of chains
* cycles, boundaries, homology groups

Recall boundary of p-simplex σ (over $\mathbb{Z}_2$ or $\mathbb{Z}$).

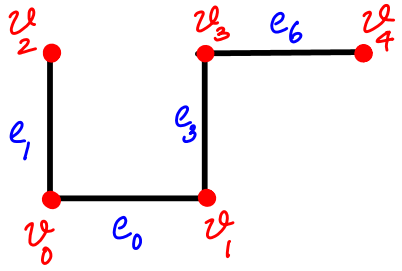
Note that $\partial_p \sigma$ is a $(p-1)$-chain (in both cases).

The boundary of a p-chain is the sum of the boundaries of its p-simplices.

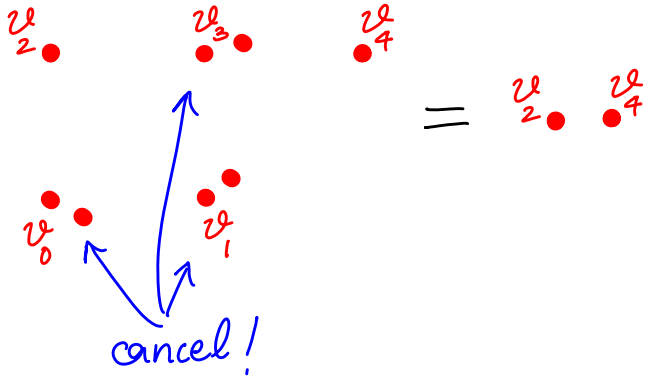
$$\bar{c} = \sum_{i=1}^{m} a_i \sigma_i \Rightarrow \partial_p \bar{c} = \sum_{i=1}^{m} a_i (\partial_p \sigma_i),$$ which is also a $(p-1)$-chain.

Examples

$\bar{c}_1$:



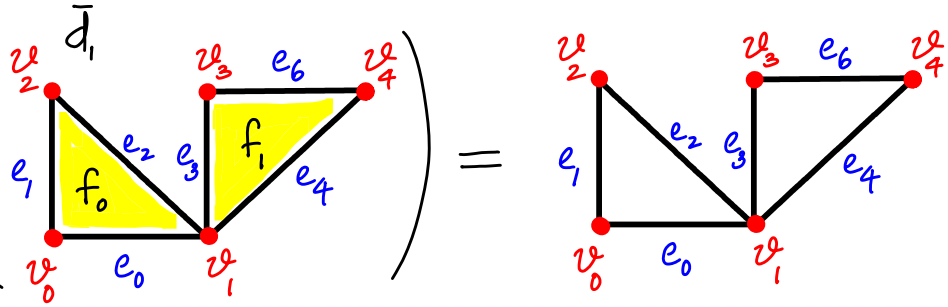
$\Rightarrow \partial_1(\bar{c}_1) =$



Note that vertices shared by two edges in the chain do not appear in its boundary.

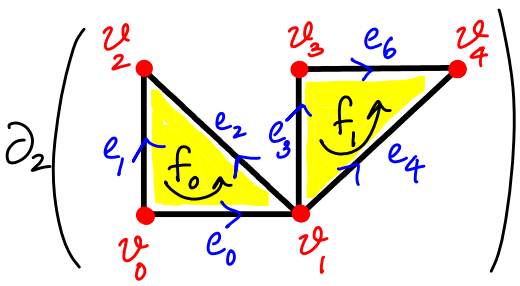
$\bar{d}_1$:

$\partial_2 \left(\right) =$

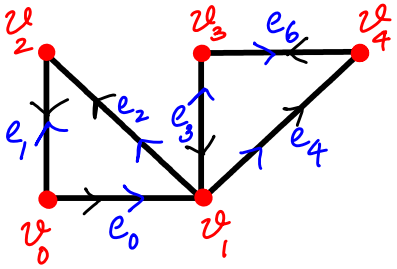


The boundary of a triangle is made of its edges.

Consider $\partial_2 \bar{d}_1$ over $\mathbb{Z}$ now:



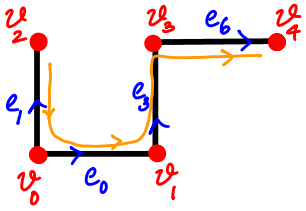
$=$



$$\partial_2 \bar{d}_1 = \begin{matrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{matrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

Notice that the induced orientation on $\bar{e}_1$ is opposite to its own orientation.

$\bar{c}_1$:



$$\partial \bar{c}_1 = -(v_2 - v_0) + (v_1 - v_0) + (v_3 - v_1) + (v_4 - v_3) = v_4 - v_2.$$

The dimension is often omitted, and we just talk about $\partial \bar{c}$ of a p-chain $\bar{c}$, with the dimension understood.

Taking the boundary maps a p-chain to a $(p-1)$-chain. Equivalently, we can talk about the map $\partial_p : C_p \to C_{p-1}$. Notice that such a map is defined for each p in the range $1 \le p \le \dim K$.

Also, $\partial_p(\bar{c} +_2 \bar{c}') = \partial_p \bar{c} +_2 \partial_p \bar{c}'$ for 2 p-chains $\bar{c}, \bar{c}'$. Hence ∂_p is a homomorphism, referred to as the **p-th boundary map** or homomorphism.
→ commutes with $+_2$, which is the definition of a homomorphism

Similarly, we have ∂_{p-1}, which is the $(p-1)$-st boundary homomorphism, and ∂_{p-2}, and so on. $\partial_{p-1} : C_{p-1} \to C_{p-2}$.

∂_p is a homomorphism over $\mathbb{Z}$ (or $\mathbb{Q}$, $\mathbb{R}$) as well, not just over $\mathbb{Z}_2$.

A **chain complex** is a sequence of chain groups connected by boundary homomorphisms

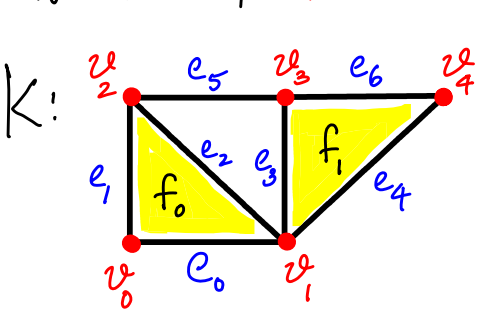
$$\cdots \xrightarrow{\partial_{p+2}} C_{p+1} \xrightarrow{\partial_{p+1}} C_p \xrightarrow{\partial_p} C_{p-1} \longrightarrow \cdots$$

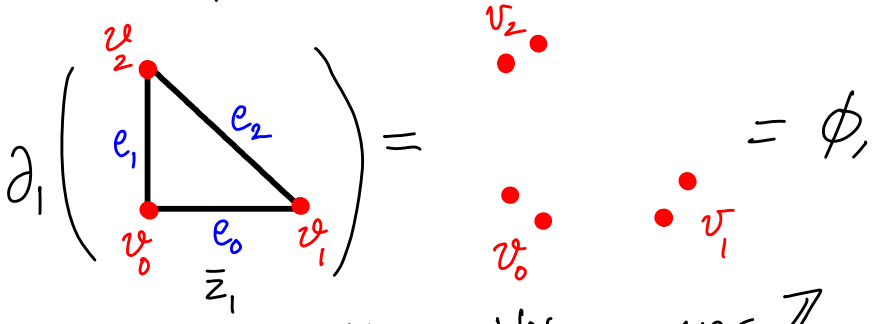
The word complex as used here is different from a simplicial complex. At the same time, the chain complex is indeed an abstract simplicial complex, with elements connected by boundary homomorphisms.

Cycles

A **p-cycle** is a p-chain with empty boundary

For instance, consider the 1-chain $\bar{z}_1$ of K:

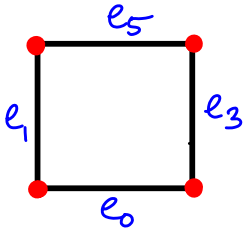




Hence $\bar{z}_1$ is a 1-cycle.

Alternatively, a p-chain $\bar{c}$ is a p-cycle if $\partial \bar{c} = 0$.

Here is another 1-cycle: $\bar{z}_2$:



Since ∂ commutes with $+$, the p-cycles of K form a group, denoted by Z_p or $Z_p(K)$. Z_p is a subgroup of C_p. Also, $Z_p = \ker \partial_p$, i.e., Z_p is the kernel of the p^{th} boundary homomorphism. Notice that $\partial_p \bar{z} = 0$ for $\bar{z} \in Z_p$. Also, just as C_p is, Z_p is an Abelian group.

The surface of a tetrahedron (made of four triangles) is a 2-cycle. Notice that its boundary is empty.

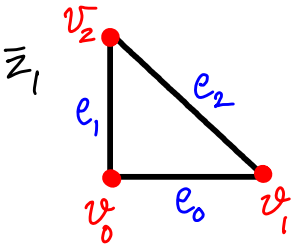
So all cycles are also chains, but not the other way usually. But for $p=0$, $\partial_0 v_j = 0$, i.e., the boundary of a vertex is empty (by definition). Hence every 0-chain is also a 0-cycle, i.e., $Z_0 = C_0$.

But typically, $Z_p \subset C_p$ for $p \geq 1$.

Boundaries A **p-boundary** $\bar{b}$ is a p-chain that is the boundary of some $(p+1)$-chain $\bar{d}$, i.e., $\bar{b} = \partial_{p+1} \bar{d}$ for $\bar{d} \in C_{p+1}$.

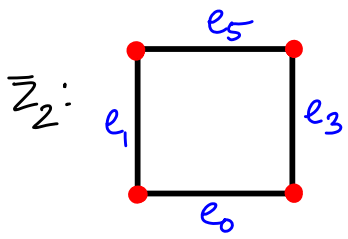
Again, since ∂ commutes with $+_2$ (or $+$) we get a group of p-boundaries B_p (or $B_p(K)$). B_p is a subgroup of Z_p, and of C_p.

B_p is the image of $\partial_{p+1} : C_{p+1} \to C_p$. $B_p = \text{im}\, \partial_{p+1}$. B_p is abelian.



$\bar{z}_1 = \partial \bar{f}_0$, and hence is a 1-boundary.

→ 2-chain made of triangle f_0



But $\bar{z}_2$ is not a 1-boundary.

We have $B_p \subset Z_p \subset C_p$ (in general).

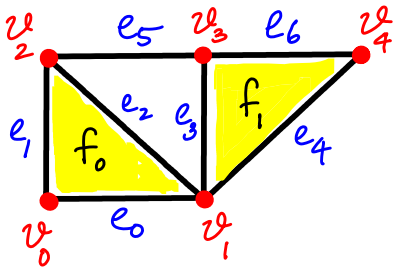
So, in summary, p-cycle $\bar{z}$: $\partial \bar{z} = 0$; $Z_p = \ker \partial_p$.

p-boundary $\bar{b}$: $\bar{b} = \partial_{p+1} \bar{d}$; $B_p = \operatorname{im} \partial_{p+1}$.

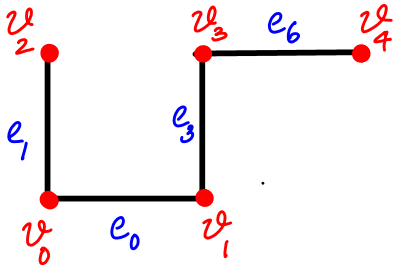
Examples

is a 0-boundary, as it is $\partial_1 \bar{c}_1$.

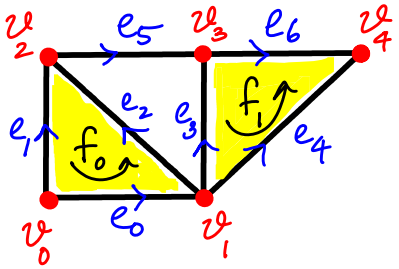
K_1:



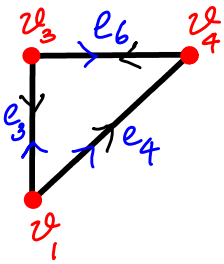
$\bar{c}_1$:



Now consider oriented K_1 (over $\mathbb{Z}$):



$\bar{b}_1$:



$$\bar{b} = \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

$\bar{b}_1$ is a 1-boundary, as $\bar{b}_1 = \partial_2 \bar{f}_1$. → elementary 2-chain of f_1

Let's try to enumerate how many p-cycles and p-boundaries are there for general p. We want to study cycles that are not boundaries.

$p=0$ case (over $\mathbb{Z}_2$)

$\bar{c}$, a 1-chain, is a collection of edges. $\partial_1 \bar{c}$ gives endpoints of edges with duplicate endpoints canceled in pairs, leaving an even number of distinct v_j's.

If K is connected, for every set of even number of vertices, we can find paths made of edges that connect the vertices such that the 1-chain made of these edges has as its boundary the collection of vertices.

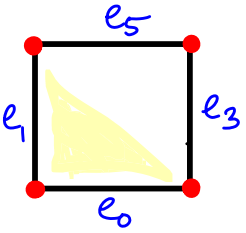
e.g., $\partial_1(e_0+e_2) = \{v_0, v_1, v_2, v_3\}$.

Thus, every set of even number of vertices (v_j's) is a 0-boundary, while every odd set is not. Hence, if K is connected, exactly half of the 0-cycles are 0-boundaries.

But we typically cannot make similar statements about counts of p-cycles and p-boundaries for $p \geq 1$.

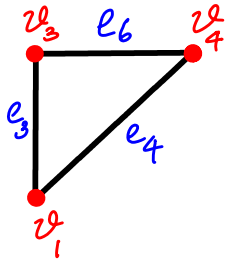
We want to characterize cycles that are not boundaries, as they capture holes, e.g., as $\bar{z}_2$ here.

$\bar{z}_2$:



An observation: Consider $\bar{b}_1 = \partial_2 \bar{f}_1$.

$\bar{b}_1$:



$\partial \bar{b}_1 = 0$ (as each v_j cancels in pairs).

In other words, $\partial_1 \partial_2 \bar{f}_1 = 0$. This result holds in general!

Fundamental Lemma of Homology $\partial_p \partial_{p+1} \bar{d} = 0 \quad \forall p \in \mathbb{Z}$.

In words, boundary of a boundary is empty.

Proof (over $\mathbb{Z}_2$) For each $(p+1)$-simplex τ, $\partial_p \partial_{p+1} \tau = 0$, as $\partial_{p+1}\tau$ consists of all p-faces of τ. Each $(p-1)$-face of τ belongs to exactly two p-faces.

$$\partial_2\left(\partial_3\left(\text{tetrahedron } v_0 v_1 v_2 v_3\right)\right) = \partial_2\left(\text{union of 4 triangles}\right) = 0,$$

as each edge is shared by exactly two triangles.

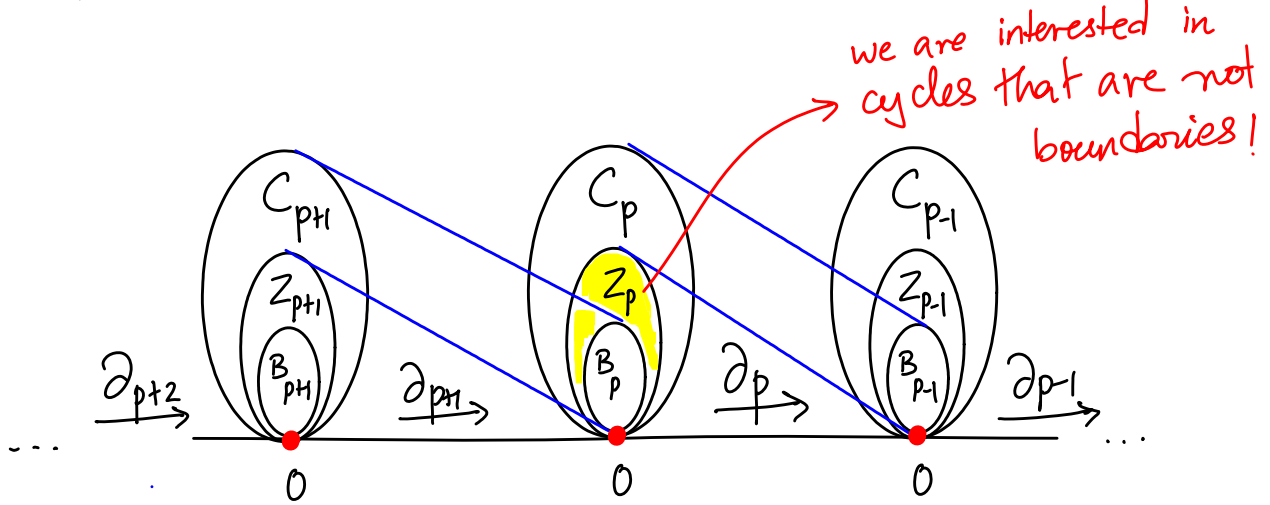
$$\partial_p\left(\partial_{p+1}\tau = \partial_{p+1}\left([v_0 \dots v_{p+1}]\right)\right) = \partial_p\left(\sum_{j=0}^{p+1} [v_0 \dots \hat{v}_j \dots v_{p+1}]\right)$$

WLOG, the $(p-1)$-simplex $[v_0 \dots v_{p-1}]$ is a face of $[v_0 \dots v_{p-1} v_p]$ and $[v_0 \dots v_{p-1} v_{p+1}]$.

The result holds over $\mathbb{Z}$ as well (over any ring, in fact). We must consider induced orientations when taking ∂_{p+1} and ∂_p.

A p-boundary is also a p-cycle. Hence B_p is a subgroup of Z_p.

The groups C_p, Z_p, B_p for various p are related as follows:



Homology Groups

Since B_p is a subgroup of Z_p, we can take quotients.

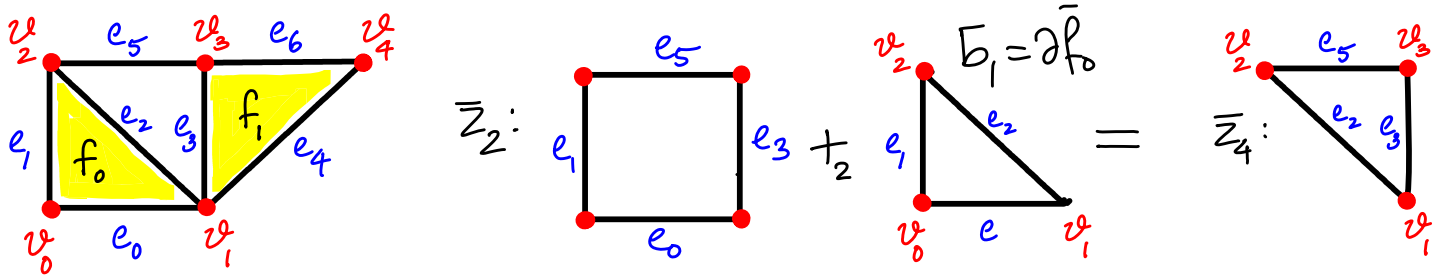
Def The **p-th homology group** is the p-th cycle group modulo the p-boundary group, $H_p = Z_p / B_p$.

H_p has the classes of cycles that are not boundaries.

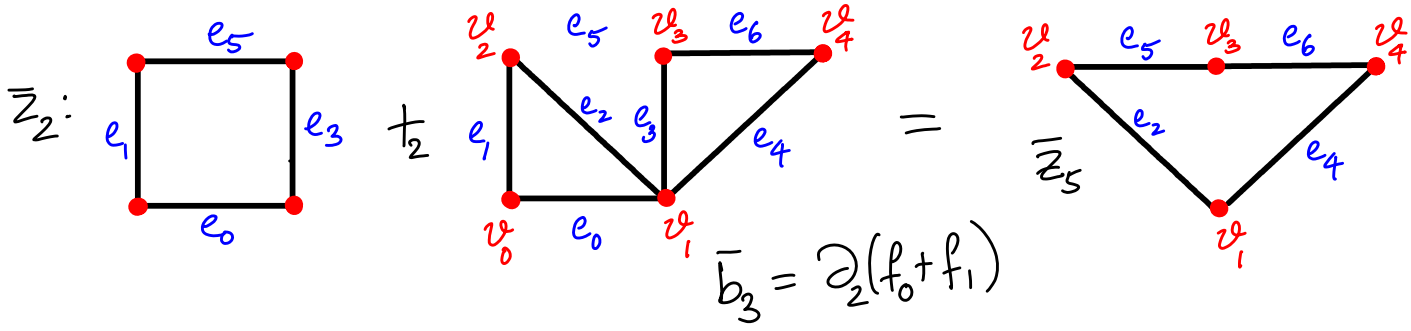
Each element of $H_p(K)$ is obtained by adding p-boundaries to given p-cycles (in their class), i.e., $\bar{z} + B_p$, where $\bar{z} \in Z_p$.

$\bar{z} + B_p$ is a coset of B_p in Z_p.

Example



$\bar{z}_2$ and $\bar{z}_4$ are cycles going around the same hole.



$\bar{z}_5$ also goes around the same hole as $\bar{z}_2$ (and $\bar{z}_4$).

We could use any one cycle going around the hole ($\bar{z}_2, \bar{z}_4, \bar{z}_5$) as a representative of $H_1(K)$.

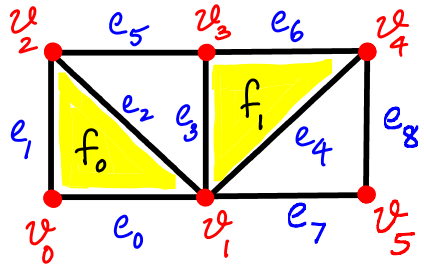
In general, another cycle $\bar{z}' = \bar{z} + \bar{b}$ for $\bar{z}$ in $H_p(K)$ and $\bar{b} \in B_p$ is in the same class as $\bar{z}$, i.e., $\bar{z}' + B_p = \bar{z} + B_p$ (as $\bar{b} + B_p = B_p$ itself). This is a class in $H_p(K)$, and any two cycles in this class are said to be **homologous**, written as $\bar{z} \sim \bar{z}'$.

In this setting, addition of classes is well-defined:
$(\bar{z} + B_p) + (\bar{z}' + B_p) = (\bar{z} + \bar{z}' + B_p)$, independent of the particular representatives $\bar{z}$ and $\bar{z}'$.

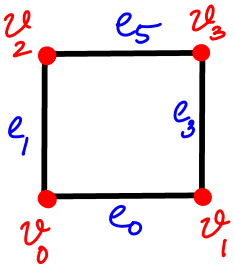
Thus H_p is indeed a group, and is abelian.

Another example

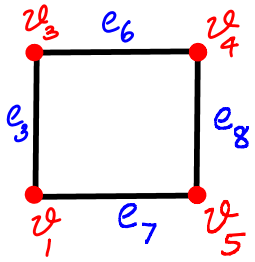
K_2:



$\bar{z}_2$:



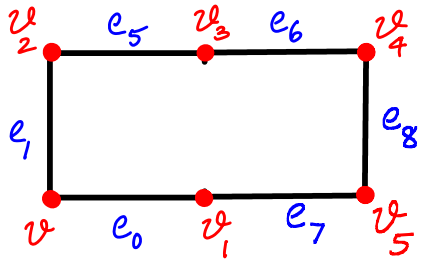
$\bar{z}_6$:



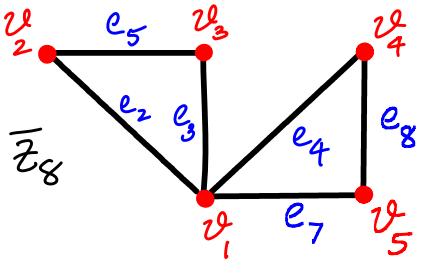
Notice $\bar{z}_6 \not\sim \bar{z}_2$, as we cannot get $\bar{z}_6 = \bar{z}_2 + \bar{b}$ for any $\bar{b} \in B_1$.

Now consider two more cycles $\bar{z}_7$ and $\bar{z}_8$ as shown.

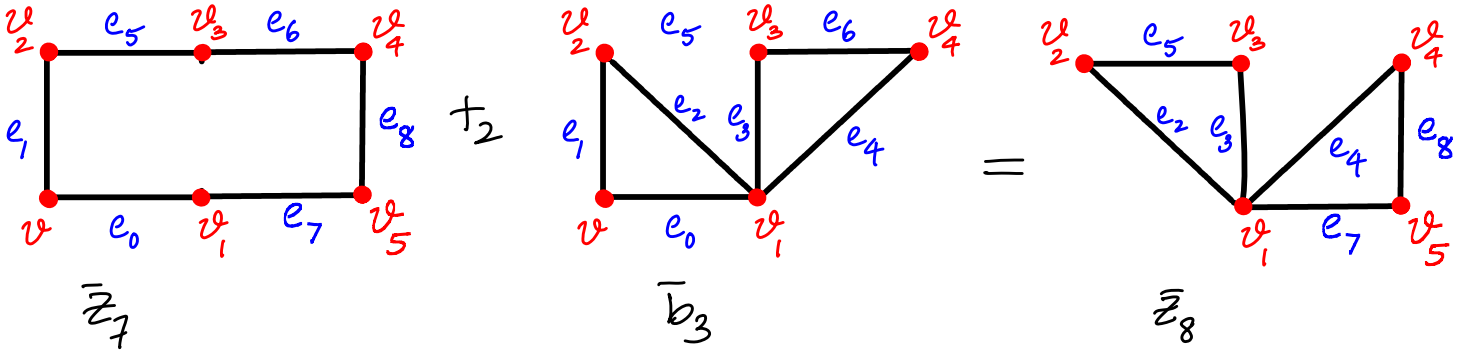
$\bar{z}_7$:



$\bar{z}_8$

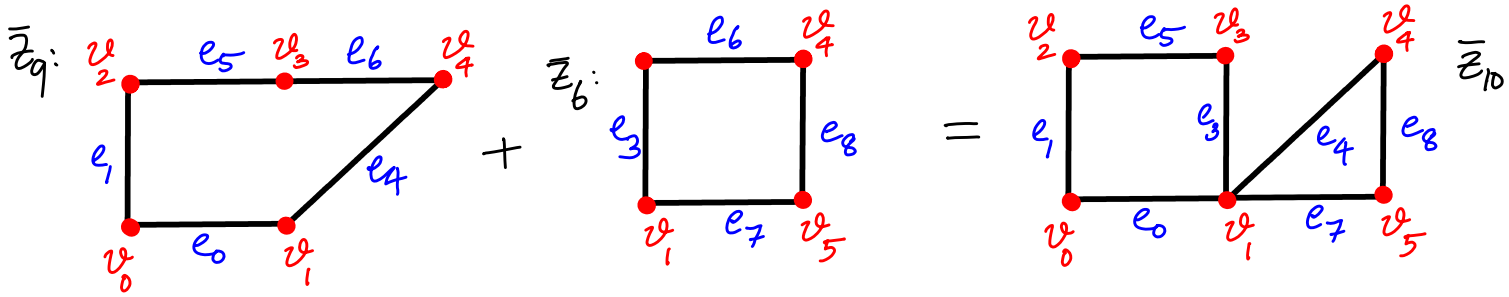


$\bar{z}_7 \sim \bar{z}_8$ as



$\bar{z}_7$ $\bar{b}_3$ $\bar{z}_8$

Now, consider $\bar{z}_9$ and $\bar{z}_6$. Notice that $\bar{z}_9 \not\sim \bar{z}_6$. But



$\bar{z}_{10} = \bar{z}_9 + \bar{z}_6$. $\bar{z}_{10} \not\sim \bar{z}_9$, but $\bar{z}_{10} \sim \bar{z}_8$

homology class of $\bar{z}_2$

To describe $H_1(K_2)$ completely, we could present $[\bar{z}_2]$ and $[\bar{z}_6]$, or equivalently, $[\bar{z}_2]$ and $[\bar{z}_7]$, or $[\bar{z}_6]$ and $[\bar{z}_7]$.

Intuitively, since K_2 has two holes, we expect $H_1(K_2)$ to have two classes.

MATH 529 – Lecture 16 (02/29/2024)

Today:
* Betti numbers
* Euler-Poincaré theorem
* boundary matrix

Recall: $B_p \subseteq Z_p \subseteq C_p$, $H_p = Z_p / B_p$ ...

Order and Rank

Recall that the cardinality of a group is called its **order**.

$\underset{\text{order}}{\text{ord}}\, C_p = |C_p|$. Since we are working over $\mathbb{Z}_2$, if there are n p-simplices in K, there are 2^n p-chains.

So, $\text{ord}\, C_p = 2^n$. Each p-simplex is either present or absent (coefficient of 1 or 0, resp.) in a p-chain.

Also, the **rank** of the chain group is $\text{rank}\, C_p = n$ here.

C_p is isomorphic to $\mathbb{Z}_2^n$, group of binary n-vectors, with addition satisfying $1+1=0$ (componentwise).

$\mathbb{Z}_2^n$ is an n-dimensional vector space. For instance, think of the n unit vectors, and the space generated by them (with $\mathbb{Z}_2$ addition).

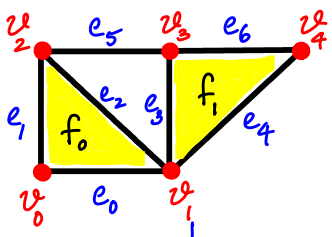
$\text{rank}\, C_p$ is the dimension of this vector space.

Z_p and B_p also show the same vector space structure, but typically have smaller orders and ranks.

e.g., K has 7 edges $(e_0, \dots, e_6)$.

$|C_1| = 2^7$ and $\text{rank}\, C_1 = 7$.

K:



All 1-chains, including cycles and boundaries, can be represented by 7-vectors.

$\text{ord } H_p = \text{ord } Z_p / \text{ord } B_p$

can add some $\bar{b} \in B_p$ to $\bar{z} \in H_p$ to get a single cycle in the same homology class.

Equivalently, $\text{rank } H_p = \text{rank } Z_p - \text{rank } B_p$.

We study these ranks in particular. We define

$\text{rank } H_p = \beta_p$, the **p-th Betti number**.

β_p's have intuitive interpretations for $p = 0, 1, 2$.

β_0 = # connected components.

β_1 = # holes

β_2 = # enclosed spaces or voids

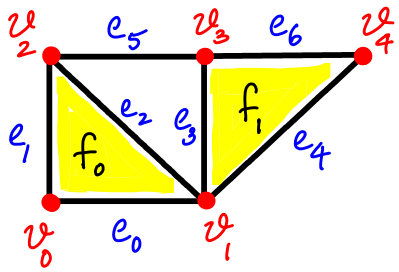
e.g.,

$\beta_0(K) = 1$ there is a single connected component

$\beta_1(K) = 1$ there is a single hole

$\beta_2(K) = 0$. There are no enclosed voids.

K:



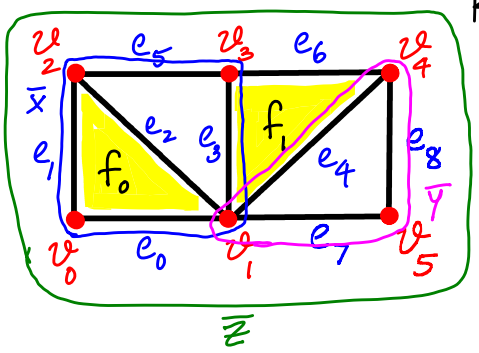
Intuitively, β_p captures the topology of K at dimension p upto homology.

Examples of Homology groups

1. To describe of $H_1(K_2)$, we could first choose one cycle around each hole.

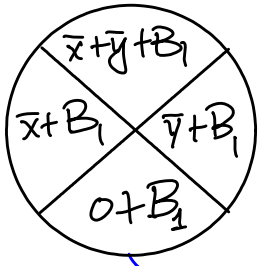
$\bar{x} = \{e_0, e_3, e_5, e_1\}$

$\bar{y} = \{e_4, e_7, e_8\}$



Here is a coset decomposition of $H_1(K_2)$:

H_1:



→ empty cycle $+B_1$

Thus, $\bar{z} = \{e_0, e_1, e_5, e_6, e_8, e_7\} \in \bar{x}+\bar{y}+B_1$, as

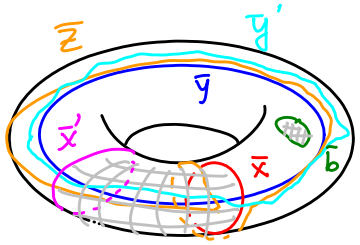
$\bar{z} = \bar{x}+\bar{y}+\{\bar{e}_3, \bar{e}_4, \bar{e}_6\}$ → $\partial f_1 \in B_1$

$H_1(K_2) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$ (isomorphic), and $\beta_1(K_2) = 2$. Also,

$\beta_0(K_2) = 1$ and $\beta_p(K_2) = 0 \ \forall p \geq 2$.

Intuitively, K_2 has 1 component, 2 holes, and no voids (or other higher dimensional "holes").

2. Torus ($\mathbb{T}^2$)



$\bar{y}$: tunnel loop
$\bar{x}$: handle loop

$\bar{x} \sim \bar{x}'$, $\{\bar{x}, \bar{x}'\}$ form the boundary of the solid 2D patch in between.

$\bar{b}$ is also a cycle, but is in fact a boundary, and hence is not a member of H_1.

H_1 is generated by $\bar{x}$ and $\bar{y}$. Here $H_1 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$, and
($\bar{x}$ $\quad$ $\bar{y}$)

$\beta_0 = 1$, $\beta_1 = 2$, and $\beta_2 = 1$ ⟶ one enclosed void

↳ one component

$\bar{z} \in \bar{x} + \bar{y} + B_1$ here.

Similarly, $\bar{y}' \sim \bar{y}$ here (both are tunnel loops).

To be exact, we should consider a triangulation of the torus, and $\bar{x}, \bar{y}, \bar{z}, \bar{b}$ are 1-cycles in the triangulation. Alternatively, we could talk about general curves instead of simplicial chains — this setting leads to singular homology, which is equivalent to simplicial homology. We will concentrate on simplicial homology, as it is more amenable to computation.

But as we will see soon, homology groups are invariants of the underlying space (up to ranks), and do not depend on the specific complex used.

3. p-ball $B_p = \{\bar{x} \in \mathbb{R}^d \mid \|\bar{x}\| \leq 1\}$ $(d \geq p)$.

For instance, when $p=3$, $B_3 \simeq |K|$ where K has a single solid tetrahedron and all its faces.

For general p, we get that $H_0 \cong \mathbb{Z}_2$ (recall, all 0-chains are also 0-cycles, and half of them are 0-boundaries), and hence $\beta_0 = 1$. one connected component

Further, H_p is trivial for $p \geq 1$, and $\beta_p = 0$ for $p \geq 1$.

Intuitively, there are no holes, no pockets, etc.

How to compute homology?

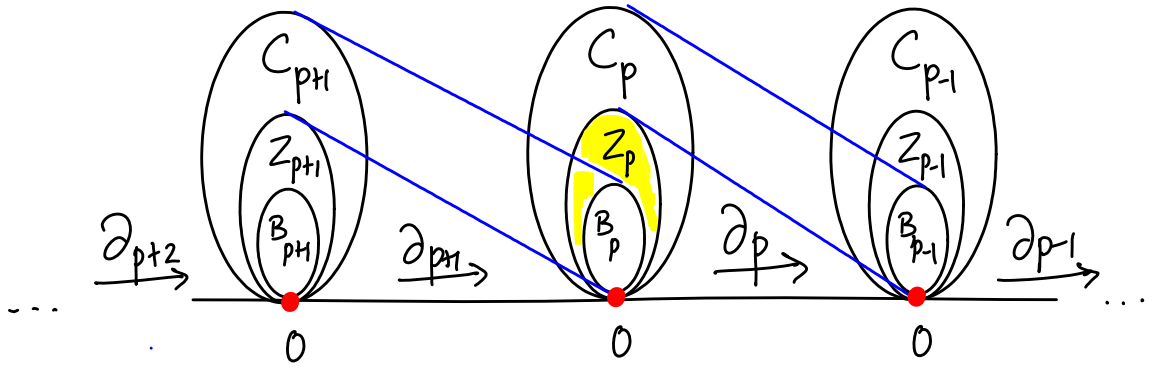
→ calculate β_p's, and find representative cycles for each homology class

We first give a result which shows that β_p's are invariants of the underlying space. Recall the Euler characteristic:

$$\chi(K) = \sum_{i=0}^{\dim K} (-1)^i s_i, \quad \text{where } s_i = \#\ i\text{-simplices in } K.$$

Let us write $z_p = \operatorname{rank} Z_p$, $b_p = \operatorname{rank} B_p$, and write $s_p = \operatorname{rank} C_p$ (# p-simplices). We get that

$$s_p = z_p + b_{p-1}.$$



This result could be seen as the implication of a general result on linear transformations (LTs). Let $f: U \to V$ be an LT from vector space U to vector space V. Then

$$\dim(U) = \dim(\ker f) + \dim(\text{im } f).$$

Applying this result to $\partial_p : C_p \to C_{p-1}$ gives $s_p = z_p + b_{p-1}$.

Alternatively, we could think about $Z_p = C_p / B_{p-1}$, and hence $\text{rank } Z_p = \text{rank } C_p - \text{rank } B_{p-1}$, i.e., $z_p = s_p - b_{p-1}$. Intuitively, if a p-chain is not a p-cycle, it will generate a $(p-1)$-boundary.

By definition, $\beta_p = z_p - b_p$ $(\text{rank } Z_p - \text{rank } B_p)$.

$$\Rightarrow \quad \chi = \sum_{p \geq 0} (-1)^p s_p$$

$$= \sum_{p \geq 0} (-1)^p (z_p + b_{p-1})$$

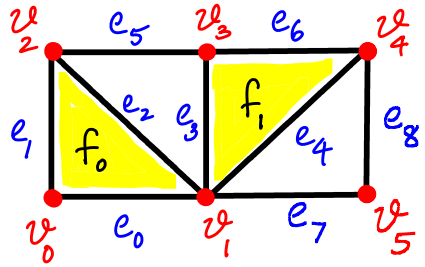
$$= \sum_{p \geq 0} (-1)^p (z_p - b_p) = \sum (-1)^p \beta_p.$$

Euler-Poincaré theorem: $\chi(K) = \sum_{p \geq 0} (-1)^p \beta_p$.

Significance Homology groups do not depend on the particular triangulation chosen. χ is an invariant of $|K|$, and so are β_p's.

Illustration

K_2:



$\chi(K_2) = 6 - 9 + 2 = -1.$

$\beta_p = 0 \quad \forall p \geq 2.$

$\beta_0 = 1$ (one connected component).

$\Rightarrow \chi = -1 = \beta_0 - \beta_1 = 1 - \beta_1 \Rightarrow \beta_1 = 2$, which agrees with the intuition that K_2 has two holes.

Boundary Matrices

To compute, i.e., find z_p, b_p, β_p and bases for Z_p, B_p, H_p, we combine the information about cycles and boundaries into a single matrix, the **boundary matrix**.

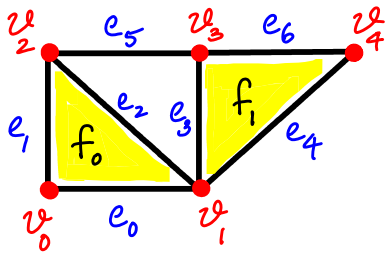
$[\partial_p(K)]$ is the p-th boundary matrix of K, $0 \leq p \leq \dim K$. It is an $m \times n$ matrix when K has m $(p-1)$-simplices $\tau_1, \ldots, \tau_m$, and n p-simplices $\sigma_1, \ldots, \sigma_n$.

$$[\partial_p]_{ij} = \begin{cases} 1, & \text{if the } i^{th} \text{ } (p-1)\text{-simplex is a face of the } j^{th} \text{ } p\text{-simplex, i.e., } \tau_i \leq \sigma_j. \\ 0, & \text{otherwise} \end{cases}$$

This is the definition over $\mathbb{Z}_2$. We will introduce the same matrices over $\mathbb{Z}$, where the nonzero entries are ± 1 depending on orientations.

Example

K:



$$[\partial_0] = \begin{matrix} \\ 1 \end{matrix}\begin{matrix} \begin{matrix} v_0 & v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

↳ default, as there are no (-1)-dimensional simplices

$$[\partial_1] = \begin{matrix} \\ v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{matrix}\begin{matrix} \begin{matrix} e_0 & e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{matrix} \\ \begin{bmatrix} 1 & 1 & & & & & \\ 1 & & 1 & 1 & 1 & & \\ & 1 & 1 & & & 1 & \\ & & & 1 & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{matrix}$$

(entries not listed are zeros).

$$[\partial_2] = \begin{matrix} \\ e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{matrix}\begin{matrix} \begin{matrix} f_0 & f_1 \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \end{matrix}$$

A collection of columns represents a p-chain, and a collection of rows represents a $(p-1)$-chain. Given a p-chain $\bar{c} = \sum_{j=1}^{n} a_j \sigma_j$, $a_j \in \{0,1\}$, its p-boundary $\partial_p \bar{c}$ is given by the matrix-vector product $[\partial_p]\bar{c}$, i.e., the sum of the corresponding columns with weights a_j gives its p-boundary.

Similarly, a collection of rows represents a $(p-1)$-chain, and their sum gives the $(p-1)$-coboundary. Coboundary and cohomology is a dual concept to boundary and homology.

Since every p-chain can be written as $\bar{c} = \sum_{j=1}^{n} a_j \sigma_j$, the columns of $[\partial_p]$ generate C_p. Similarly, rows of $[\partial_p]$ generate C_{p-1}.

To compute z_p, b_p, β_p for all p, we will use operations similar to Gaussian elimination in linear algebra. Here we perform both elementary row operations (EROs) as well as elementary column operations (ECOs).

MATH 529 – Lecture 17 (03/05/2024)

Today: * Smith normal form of $[\partial_p]$
* read off β_p, b_p, and maintain bases

Recall $[\partial_p]$: $m \times n$ $\{0,1\}$-matrix (over $\mathbb{Z}_2$) when K has m $(p-1)$-simplices and n p-simplices.

$[\partial_p]_{ij} = \{1 \text{ if } \tau_i \le \sigma_j, 0, \text{otherwise}\}$

EROs on $A_{m\times n}$: $R_i \rightleftharpoons R_j$ (swap rows i and j)

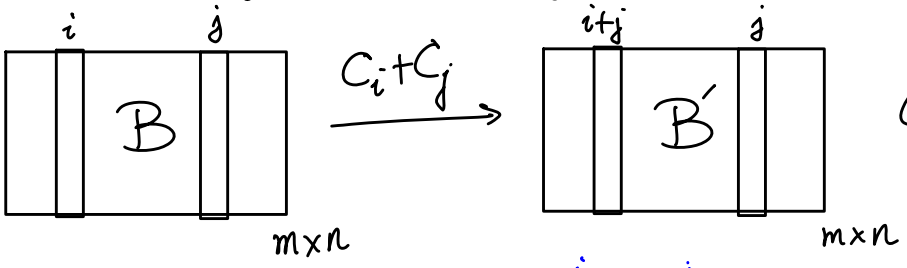
over $\mathbb{R}$, $c \in \mathbb{R}$, and $R_i \leftarrow R_i + cR_j$ ($R_i + cR_j$, replacement)

over $\mathbb{Z}$, $c \in \mathbb{Z}$. $R_i \leftarrow cR_i$, $c \neq 0$ (cR_i, scaling)

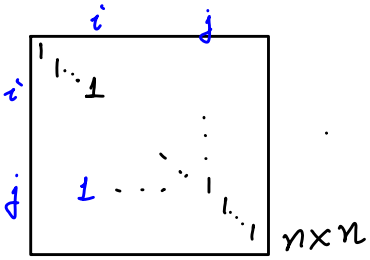
Over $\mathbb{Z}_2$, $c \in \{0,1\}$, and only swap and replacement need to be considered. We write $R_i \rightleftharpoons R_j$, $R_i + R_j$, in short.

ECOs are defined similarly: $C_i \rightleftharpoons C_j$, $C_i + C_j$ (over $\mathbb{Z}_2$).

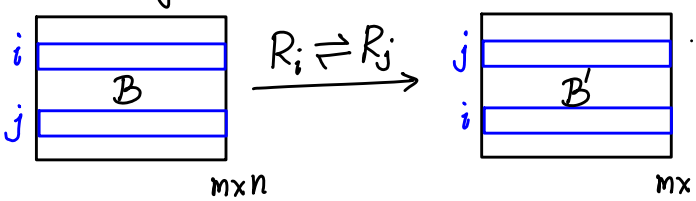
ECOs can be represented by multiplication on the right by the corresponding **elementary matrix**.

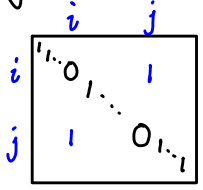
e.g., 

can be written as

$B' = BV$, where $V =$ 

Similarly, EROs can be represented by multiplication the left by an elementary matrix.



can be written as $B' = UB$ for $U =$ 

.

Smith Normal Form (SNF)

This is the analogue of reduced row echelon form (RREF) of a matrix, from linear algebra. SNF is defined over $\mathbb{Z}$, in general, and accounts for both EROs and ECOs.

Def A matrix $B \in \mathbb{Z}^{m \times n}$ is in **Smith normal form** if

$$B = \begin{bmatrix} D & O \\ O & O \end{bmatrix} = \left[\begin{array}{ccc|c} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_\ell \\ \hline & O & & O \end{array}\right] \quad \text{where}$$

Each 'O' is a submatrix of all zeros.

$d_i \in \mathbb{Z}$, $d_i \geq 1$ and $d_1 | d_2 | d_3 | \cdots | d_\ell$.

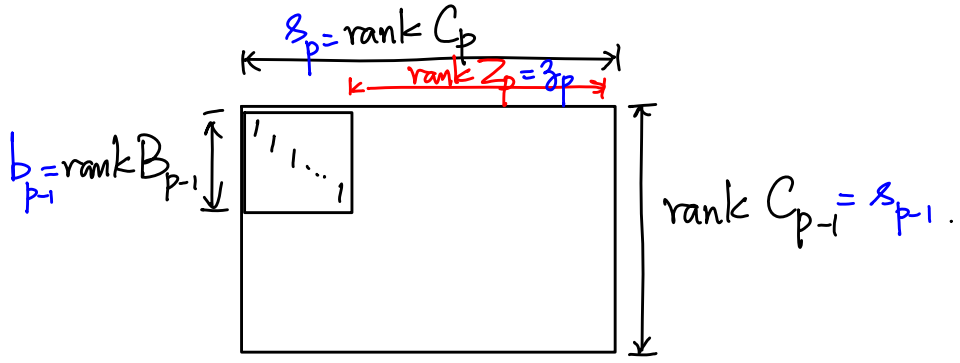
↳ "divides": $a|b$ means a divides b.

When working over $\mathbb{Z}_2$, we will get

$$\text{SNF}([\partial_p]) = \left[\begin{array}{ccc|c} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ \hline & O & & O \end{array}\right].$$

In more detail, we will have the following structure for $\text{SNF}([\partial_p])$.

$$\text{SNF}([\partial_p]) =$$



We get z_p and b_{p-1} from $\text{SNF}([\partial_p])$. To compute β_p, we need to know also b_p, which is obtained from $\text{SNF}([\partial_{p+1}])$.

Overall, we can write $SNF([\partial_p]) = U_{p-1}[\partial_p]V_p$, where U_{p-1} captures all EROs performed, and V_p captures all ECOs performed. We also get information about bases of Z_p and B_{p-1} from U_{p-1} and V_p.

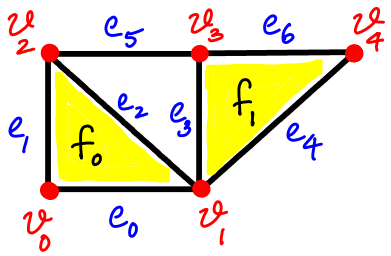
More specifically, a basis for the cycle group Z_p is encoded in the last β_p columns of V_p. Similarly, a basis for the boundary group B_{p-1} is encoded in U_{p-1} — in the first b_{p-1} columns of U_{p-1}^{-1}, to be exact.

Thus we can reduce each $[\partial_p]$ to SNF, and in that process, compute all β_p, and also identify bases for Z_p and B_p. In the following illustration, we keep track of the bases as we perform the corresponding EROs or ECOs on $[\partial_p]$. Later on, we describe an algorithm that does all the operations in a unified manner.

Example

We consider our favorite example and reduce each $[\partial_p]$ to SNF for $p = 0, 1, 2$.

K:



$$[\partial_0] = \begin{matrix} \\ 1 \end{matrix} \begin{matrix} v_0 & v_1 & v_2 & v_3 & v_4 \\ \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

↳ default, as there are no (−1)-dimensional simplices

$$[\partial_1] = \begin{matrix} \\ v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{matrix} \begin{matrix} e_0 & e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \\ \end{matrix} \begin{bmatrix} 1 & 1 & & & & & \\ 1 & & 1 & 1 & 1 & & \\ & 1 & 1 & & & 1 & \\ & & & 1 & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix}$$

(entries not listed are zeros).

$$[\partial_2] = \begin{matrix} \\ e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{matrix} \begin{matrix} f_0 & f_1 \\ \end{matrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

We reduce each $[\partial_p]$ to SNF.

We consider $[\partial_0]$ and $[\partial_2]$ first, since they are simpler compared to $[\partial_1]$ (we consider that the last).

$$[\partial_0] = \begin{matrix} & v_0 & v_1 & v_2 & v_3 & v_4 \end{matrix} \; 1\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \xrightarrow[\text{for } j=2,\dots,5]{C_j + C_1} \; 1\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\begin{matrix} v_0 & v_1+v_0 & v_2+v_0 & v_3+v_0 & v_4+v_0 \end{matrix}$

$\underbrace{\qquad\qquad}_{z_0 = 4}$

But $s_0 = 5 = z_0$!

We had noticed previously that every 0-chain is also a 0-cycle, as vertices have empty boundaries. So, we should have got $s_0 = z_0 = 5$ here! We will address this discrepancy in $p=0$ soon. For now, work with $z_0 = s_0 = 5$.

A basis for $B_0(K) = \{ v_j + v_0 \mid j = 1,2,3,4 \}$.

A collection of even pairs of vertices, such that any even set of vertices can be written as a combination of these pairs.

Recall that every 0-chain is also a 0-cycle since it has no boundary (when we work over $\mathbb{Z}_2$). And every 0-chain with an even number of vertices is a 0-boundary (assuming K is connected).

Adding any one vertex (by itself) to the above basis will give another basis for Z_0 (and hence C_0), e.g., $\{ v_j + v_0 \}_{j=1}^{4}, v_0 \}$. Recall that $\{ v_j \}_{j=0}^{4}$ gives the default (elementary chain) basis for $Z_0(C_0)$.

Notation: $e_{ijk} \equiv e_i + e_j + e_k$

$$[\partial_2] = \begin{array}{c} \\ e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{array}\!\!\begin{array}{c} \begin{array}{cc} f_0 & f_1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \end{array} \xrightarrow[\substack{R_3+R_1 \\ R_5+R_4 \\ R_7+R_4}]{R_2+R_1} \begin{array}{c} \\ e_0 \\ e_{01} \\ e_{02} \\ e_3 \\ e_{34} \\ e_5 \\ e_{36} \end{array}\!\!\begin{array}{c} \begin{array}{cc} f_0 & f_1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \end{array} \xrightarrow{R_2 \rightleftharpoons R_4} \begin{array}{c} \\ e_0 \\ e_3 \\ e_{02} \\ e_{01} \\ e_{34} \\ e_5 \\ e_{36} \end{array}\!\!\begin{array}{c} \begin{array}{cc} f_0 & f_1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \end{array}$$

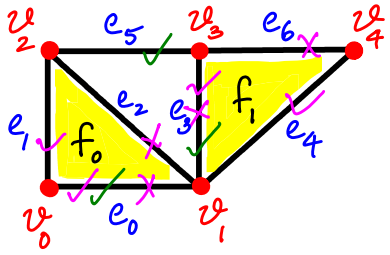
there are no 2-cycles here!

no 0-matrix to the right!

$z_2 = 0, \; b_1 = 2$

Basis for $C_2(K) = \{f_0, f_1\}$.

K:



A basis for $C_1(K) = \{e_0, e_3, e_5, e_0+e_1, e_0+e_2, e_3+e_4, e_3+e_6\}$.

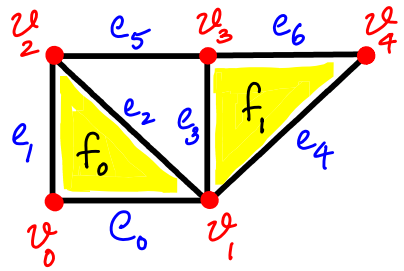
Notice that we have found also a basis for $B_1(K)$ — consisting of the boundaries of f_0 and f_1.

We are numbering the rows and columns starting from 1. As we proceed further into the reduction (to SNF), the labels for the rows/columns could become more complicated.

Similar notation for column labels: $C_{ijk} \equiv C_i + C_j + C_k$.

Let's look at $[\partial_1]$ now.

K:



$$[\partial_1] = \begin{array}{c} \\ v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & 1 & & & & & \\ 1 & & 1 & 1 & 1 & & \\ & 1 & 1 & & & 1 & \\ & & & 1 & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{array}$$

$R_2 + R_1$ and then $C_2 + C_1$

$$\begin{array}{c} \\ v_0 \\ v_{01} \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_{01} & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & 0 & & & & & \\ 0 & 1 & 1 & 1 & 1 & & \\ & 1 & 1 & & & 1 & \\ & & & 1 & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{array}$$

$R_3 + R_2$ then $C_j + C_2$, $j = 3,4,5$

$$\begin{array}{c} \\ v_0 \\ v_{01} \\ v_{012} \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_{01} & e_{012} & e_{013} & e_{014} & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & & & & & & \\ & 1 & 0 & 0 & 0 & & \\ & 0 & 0 & 1 & 1 & 1 & \\ & & & 1 & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{array}$$

$C_3 \rightleftharpoons C_4$

$$\begin{array}{c} \\ v_0 \\ v_{01} \\ v_{012} \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_{01} & e_{013} & e_{012} & e_{014} & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & & & & & & \\ & 1 & 0 & 0 & 0 & & \\ & 0 & 1 & 0 & 1 & 1 & \\ & & 1 & & & 1 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{array}$$

$R_4 + R_3$ then $C_5 + C_3$, $C_6 + C_3$

$$\begin{array}{c} \\ v_0 \\ v_{01} \\ v_{012} \\ v_{0123} \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_{01} & e_{013} & e_{012} & e_{34} & e_{0135} & e_6 \end{array} \\ \begin{bmatrix} 1 & & & & & & \\ & 1 & 0 & 0 & 0 & & \\ & 0 & 1 & 0 & 0 & 0 & \\ & & 0 & & 1 & 0 & 1 \\ & & & & 1 & & 1 \end{bmatrix} \end{array}$$

$C_5 \rightleftharpoons C_4$ then $R_5 + R_4$ then $C_7 + C_4$

$$\begin{array}{c} \\ v_0 \\ v_{01} \\ v_{012} \\ v_{0123} \\ v_{01234} \end{array} \begin{array}{c} \begin{array}{ccccccc} e_0 & e_{01} & e_{013} & e_{34} & e_{012} & e_{0135} & e_{346} \end{array} \\ \begin{bmatrix} 1 & & & & & & \\ & 1 & 0 & 0 & 0 & & \\ & 0 & 1 & 0 & 0 & 0 & \\ & & 0 & 1 & & 0 & 0 \\ & & & 0 & & & 0 \end{bmatrix} \end{array}$$

$= SNF([\partial_1])$.

$z_1 = 3$, $b_0 = 4$

Thus we get $\beta_1 = z_1 - b_1 = 3 - 2 = 1$, and $\beta_0 = z_0 - b_0 = 5 - 4 = 1$, and both of these numbers agree with intuition (1 hole and 1 component).

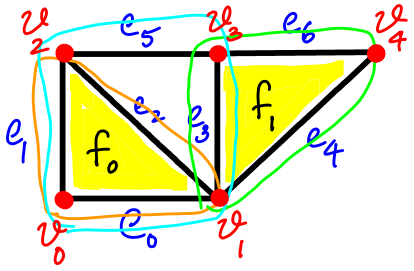
$\beta_p = 0 \quad \forall p \geq 2$.

MATH 529 – Lecture 18 (03/07/2024)

Today: * Reduction algorithm for $[\partial_p]$
* reduced homology
* relative homology

A final observation on the matrix reduction example...

K:



$SNF([\partial_1]) =$

	e_0	e_{01}	e_{013}	e_{34}	e_{012}	e_{0135}	e_{346}
v_0	1						
v_{01}		1	0	0	0		
v_{012}		0	1	0	0	0	
v_{0123}			0	1		0	0
v_{01234}				0			0

The labels for the Z_1 columns in $SNF([\partial_1])$: e_{012} e_{0135} e_{346}

The column labels of the last three (zero) columns of $SNF([\partial_1])$ gives a basis for Z_1. They are e_{012}, e_{0135}, and e_{346}, i.e., $\{\{e_0, e_1, e_2\}, \{e_0, e_1, e_3, e_5\}, \{e_3, e_4, e_6\}\}$. These three loops consist of ∂f_0, ∂f_1, and $\{e_0, e_1, e_3, e_5\}$, which captures the hole. $\{e_0, e_1, e_3, e_5\}$ can be used to represent H_1 here.

Notice that we may not obtain the actual "boundary of the hole" – we're guaranteed to get one loop around the hole, but it may not be the shortest or tightest loop!

Algorithm for SNF over $\mathbb{Z}_2$ of $B \in \{0,1\}^{m \times n}$

Here is the main block:

```
void REDUCE(i)
→ if ∃ j≥i, k≥i with B_jk = 1 then
      R_j ⇌ R_i ; C_k ⇌ C_i
      for h = i+1 to m
          if B_hi = 1 then
              R_h + R_i ;
          endif
      endfor
      for l = i+1 to n
          if B_il = 1 then
              C_l + C_i ;
          endif
      endfor
      REDUCE(i+1);
→ end if
```

$B_{ii} = 1$ after this step (this step allows the possibility of $B_{ii} = 1$ to start with)

Zero out 1's below B_{ii} using replacement EROs

Zero out 1's to the right of B_{ii} using replacement ECOs

→ proceed to next smaller block to right and below

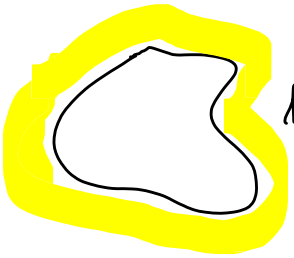
We could initiate (as identity matrices) the elementary matrices $U_{m \times m}$ and $V_{n \times n}$, as well as U^{-1}, and update them along the way to identify bases for Z_p, B_{p-1}, etc.

You will get the chance to implement this reduction algorithm in HW5...

Reduced Homology groups

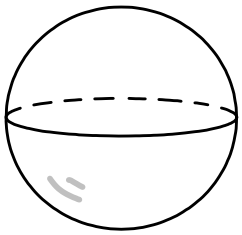
To take care of the discrepancy at $p=0$ ($z_0 = s_0$ or not?)

IDEA: "A prison cell in p-dimensions has $(p-1)$-dimensional walls."



$p=2, \beta_1 = 1$

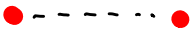
2D prison cell with 1D wall



$p=3, \beta_2 = 1$

3D prison cell with 2D wall

We count $\beta_1 = 1$ for the hole when the 2D solid patch is missing in the middle. Similarly, when the solid 3D object/space is missing from the middle of a sphere, we set $\beta_2 = 1$.



$\beta_0 = 1$?

Note that we're removing a p-ball in each case to make the prison cell
(2-ball $\approx$ disc, 3-ball $\approx$ solid 3D patch,
1-ball $\approx$ edge).

Under this idea, we would count $\beta_0 = 1$ when we have two vertices (with the "edge in between missing", capturing a "hole"). But if β_0 is to count the number of connected components, $\beta_0 = 2$ here.

To resolve this confusion, we add the augmentation map $\epsilon : C_0 \to \mathbb{Z}_2$, where $\epsilon(v_j) = 1$ for each $v_j \in K$.

We push off the "boundary effect" one step more, so that the general structure holds for $p=0$ as well!

$$\cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\epsilon} \mathbb{Z}_2 \xrightarrow{0} 0$$

ϵ: new, as compared to the original chain complex

Notice that $\epsilon \circ \partial_1 = 0$, similar to $\partial_p \partial_{p+1} = 0 \; \forall p$, as each edges has two vertices, and both vertices have value 1 under ϵ and $1+_2 1 = 0$.

Under ϵ, a 0-cycle must have an even # vertices. As the 1-boundary of an edge (1-simplex) will have two vertices, and each of them has a value of 1 under ϵ, giving $1+_2 1 = 0$.

The new chain complex with homomorphisms ∂_p along with ϵ gives rise to **reduced homology groups** $\tilde{H}_p(K)$. We also get the **reduced Betti numbers**, $\tilde{\beta}_p$, as the corresponding ranks: $\tilde{\beta}_p = \text{rank } \tilde{H}_p$.

Assuming $K \neq \emptyset$, $\tilde{\beta}_p = \beta_p$ for $p \geq 1$, and $\tilde{\beta}_0 = \beta_0 - 1$.

A similar augmentation map is defined for homology over $\mathbb{Z}$. $\epsilon(v_j) = 1$ for vertex v_j.

$$\cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\epsilon} \mathbb{Z} \xrightarrow{0} 0 \to \cdots$$

$\epsilon \partial_1 = 0$ here as well.

$$\epsilon\left(\partial_1 \left[\; \overset{e}{\overrightarrow{v_0 \; \; v_1}} \; \right]\right) = \epsilon(v_1 - v_0) = 1 - 1 = 0.$$

Relative Homology

One way to extend the concepts of homology from finite simplicial complexes, to pairs of spaces.

Let K be a simplicial complex and $K_0 \subseteq K$ be a subcomplex. We can talk about relative chain groups as

$$C_p(K, K_0) = C_p(K) / C_p(K_0),$$

i.e., as the quotient of the chain groups in K and K_0.

IDEA: Divide $C_p(K)$ into cosets of p-chains that differ in p-simplices in K_0, but not over $K - K_0$.

We can consider ∂_p from $C_p(K, K_0)$ to $C_{p-1}(K, K_0)$ as the homomorphism induced by ∂_p originally defined on K.

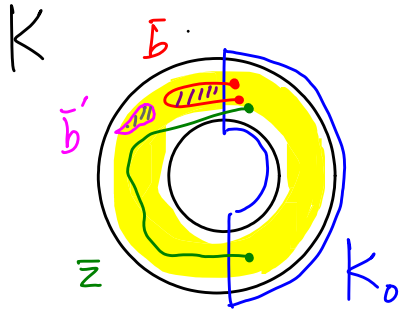
$$\partial_p : C_p(K, K_0) \longrightarrow C_{p-1}(K, K_0).$$

These boundary maps also satisfy $\partial_p \circ \partial_{p+1} = 0$. Hence we can define cycles, boundaries, and homology groups in the relative setting.

$$Z_p(K, K_0) = \ker \partial_p, \quad B_p(K, K_0) = \operatorname{im} \partial_{p+1}, \text{ and}$$

$$H_p(K, K_0) = Z_p(K, K_0) / B_p(K, K_0).$$

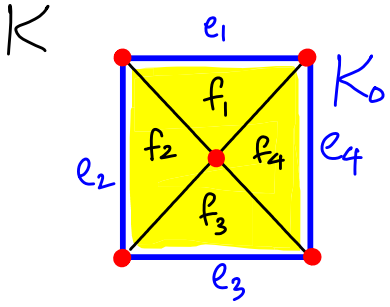
Examples Let K be an annulus, and K_0 be half of it.



$\bar{z}$ is a 1-cycle in K/K_0, i.e., $\bar{z} \in Z_1(K, K_0)$. So is $\bar{b}$. But $\bar{b}$ is also a 1-boundary in K/K_0, i.e., $\bar{b} \in B_1(K, K_0)$. Similarly, $\bar{b}' \in B_1(K, K_0)$, but is also in $B_1(K)$.

Intuitively we consider everything in K_0 as "trivial", or as "reduced to empty". For instance, $\partial\bar{z}$ above is the set of its two end points, which are both in K_0. Hence, $\partial\bar{z}$ is considered empty in K/K_0, and hence $\bar{z}$ is a relative cycle.

Here is another example. Let K_0 be the four boundary edges of K, which is made of the four triangles.



$K_0 = \{e_1, e_2, e_3, e_4, \text{ and faces thereof}\}$.

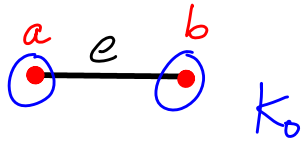
$\bar{c} = \sum_1^4 a_j f_j$ is a general 2-chain of K

When $a_j = a \in \mathbb{Z}\ \forall j$, $\partial\bar{c} \in K_0$ (when $a = 1$, we get all of K_0 as $\partial\bar{c}$). In this case, $\partial\bar{c}$ is not empty in K. Hence, $\bar{c}$ is not a 2-cycle in K.

But $\bar{c} \in Z_2(K, K_0)$ as $\partial\bar{c}$ is contained in K_0.

Intuitively, we "shrink" all of the boundary of $\bar{c}$ to empty, thus "creating" a 2-sphere out of it!

Notice that K_0 need not be connected. Consider the example where K consists of the single edge $e = \bar{ab}$ (and the vertices a and b). Let $K_0 = \{a, b\}$, the two vertices. Here, e is a relative 1-cycle in (K, K_0).



We can compute ranks of, and bases for, relative homology groups using the same concept of SNF of $[\partial_p]$. We start off by defining the relative boundary matrix for (K, K_0) as the submatrix of $[\partial_p]$ obtained by removing the rows and columns corresponding to simplices in K_0. The rest of the procedure is identical to the one used in the absolute, i.e., default, case.

We will use the idea of relative homology in defining some of the concepts in the upcoming lectures.

MATH 529 – Lecture 19 (03/19/2024)

Today:
* persistence
* incremental algorithm for betti numbers

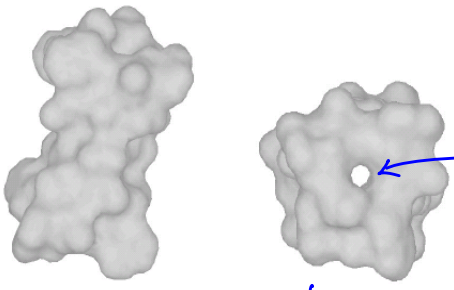
Persistent Homology

We will consider a filtered complex K $\phi = K^0 \subseteq K^1 \subseteq \dots \subseteq K^m = K$, $|K| = \mathbb{X}$, some space.

Let $\beta_k^l = \text{rank}(H_k^l)$ where $H_k^l = Z_k^l / B_k^l$ with $Z_k^l = Z_k(K^l)$, etc.

We could track β_k^l for fixed k as l varies, getting more information than obtained from any single l.

Example (from Edelsbrunner, Letscher, Zomorodian, 2002).



Gramicidin A, a protein with an open tunnel running through it.

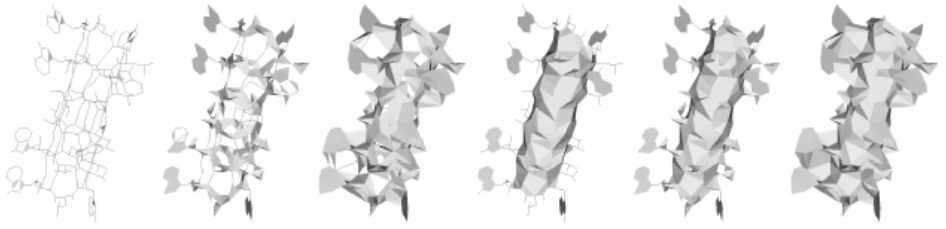
tunnel going through

side view

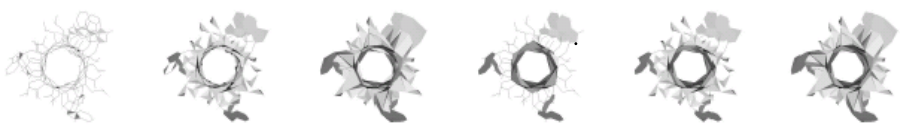
top view

Shown below are certain subcomplexes from the α-complex filtration of gramicidin.

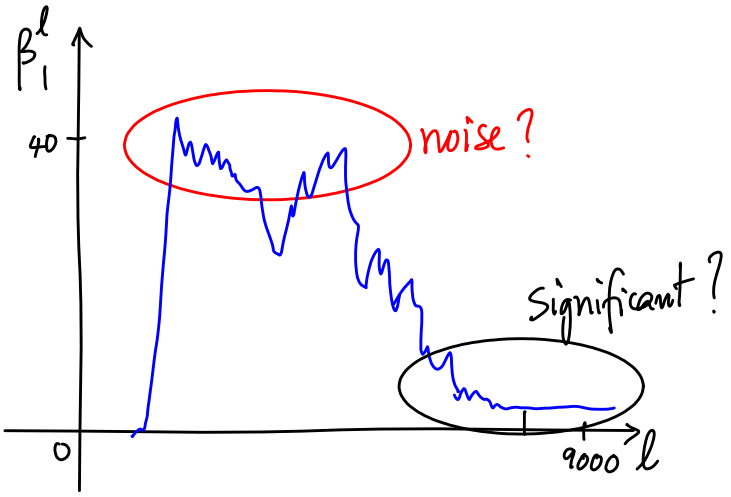
→ with atoms as points, having different starting radii.



→ side views



→ top views



The appears to be lots of holes showing up early on, which all get "closed" later on. We need to separate topological noise from features.

A "significant" feature should have a long "lifetime" in the filtration. We look for non-bounding cycles that do **not** turn into boundaries "in the near future," i.e., for the next p complexes in the filtration. Thus, we look for cycles in K^l that stay non-bounding till K^{l+p}.

Def The p-persistent k^{th} homology group of K^l is

$$H_k^{l,p} = Z_k^l / \left(B_k^{l+p} \cap Z_k^l\right).$$

And $\beta_k^{l,p} = \text{rank}\left(H_k^{l,p}\right)$ is the p-persistent k^{th} betti number of K^l.

Let the non-bounding k-cycle $\bar{z}$ be created at time i (for index l, $l+p$, etc.) with the arrival of simplex σ^i. So $[\bar{z}]$, the homology class represented by $\bar{z}$ is in H_k^i. Let σ^j arrive at time j, and turn $\bar{z}' \in [\bar{z}]$ into a boundary. The "k-dimensional hole" captured by $[\bar{z}]$ is "closed" by σ^j.

So $\bar{z}' \in B_k^j$, and $[\bar{z}]$ is now merged with some older class of cycles, i.e., it no longer exists independently. But it did exist independently for all times $i \le t < j$, i.e., for $j-i-1$ steps.

The **persistence** of $[\bar{z}]$ (or of $\bar{z}$ itself) is $j-i-1$, and $[i, j)$ is its **life-time** in the filtration.

σ^i is the **creator** of $[\bar{z}]$, and σ^j is the **destroyer**.

A simplex that is a creator is a positive simplex, and one that destroys is a negative simplex.

If a class has no destroyer, its persistence is $+\infty$.

Time-based persistence

$$H_k^{\lambda,\mu} = Z_k^\lambda / \left(B_k^{\lambda+\mu} \cap Z_k^\lambda\right) \quad \text{for times } \lambda, \mu \ge 0.$$

We could have the filtration growing with time (instead of adding one simplex in each step). By default, though, we will consider index-based persistence. The two notions are equivalent in some sense (under appropriate assumptions).

Incremental Algorithm for Betti numbers (Delfinado & Edelsbrunner, 1995)

The algorithm assumes K is in $\mathbb{R}^3$ and is torsion-free. Equivalently, K is a subcomplex of (a triangulation) of $S^3 \simeq \mathbb{R}^3 \cup \{\infty\}$.
We add one simplex at each step l to grow K. Equivalently, in the filtration of K, $K^j = K^{j-1} \cup \{\sigma^j\}$ $\forall j$.

We want to classify k-simplices for $k=0,1,2,3$ as creators or destroyers. Later on we will talk about pairing a destroyer with a creator, and use these pairings to compute persistence.

To simplify notation for this discussion, we write $L = K \cup \{\sigma\}$, instead of $K^j = K^{j-1} \cup \{\sigma^j\}$. We consider four cases, corresponding to σ being a vertex, edge, triangle, or a tetrahedron.

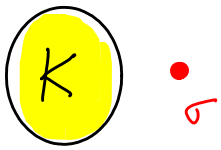
In fact, this classification is valid in general when we add a simplex σ to a simplicial complex K (i.e., not necessarily in a filtration framework).

Recall that $K^{j-1} \subseteq K^j$ in the filtration. Hence when we add a simplex σ in step j, we are guaranteed that all proper faces of σ are already present in the previous complex.

For instance, when we add $\sigma = \triangle abc$, the 0-simplices a, b, c, and the 1-simplices ab, ac, and bc are already present.

A simple way to create such a filtration is to order all simplices by their dimensions, with lower dimensional simplices appearing before higher dimensional ones (we can break ties arbitrarily). Standard filtrations such as alpha or Vietoris-Rips complexes also provide such orderings once we use dimension to break ties among simplices (so, if a triangle and its faces are present at a given radius, the edges come before the face).

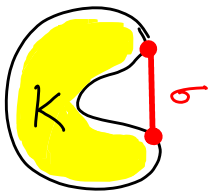
Case 1. σ is a 0-simplex



We get $\beta_0(L) = \beta_0(K) + 1$.

σ here adds a new class of 0-cycles. Other homology classes are not changed. Thus, 0-simplices are always creators, i.e., they are always positive.

Case 2 σ is a 1-simplex

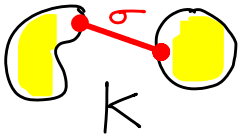


If σ belongs to a 1-cycle in L,

$$\beta_1(L) = \beta_1(K) + 1,$$

else

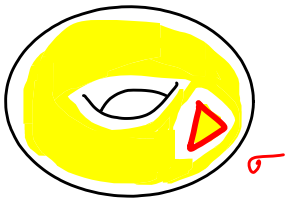
$$\beta_0(L) = \beta_0(K) - 1.$$



Other β_k's are not affected.

So, 1-simplices can be creators or destroyers.

Case 3 σ is a 2-simplex



If σ belongs to a 2-cycle in L

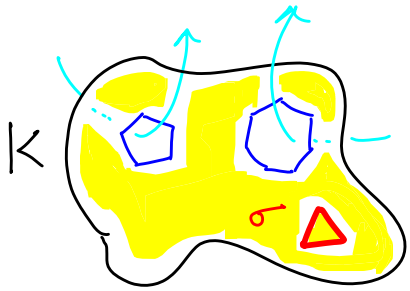
$$\beta_2(L) = \beta_2(K) + 1;$$

else

$$\beta_1(L) = \beta_1(K) - 1.$$

(other β_k's remain unchanged)

all other triangles in the surface of the torus are already present, and σ just closes the final hole, thus trapping the 3D space inside.



σ closes one of the holes.

Thus, 2-simplices can be creators or destroyers, just like 1-simplices.

Case 4 σ is a 3-simplex.

$$\beta_2(L) = \beta_2(K) - 1.$$

The tetrahedron coming in closes the void formed by its four faces.

So, 3-simplices are always destroyers (or, are negative) in our setting.

Since we assumed that K is a subcomplex of a triangulation of S^3 in $\mathbb{R}^3$, we do not get any 3-cycles. So, $\beta_3(K) = 0$.

IDEA: To understand the above statement about $\beta_3 = 0$, we will go down in dimension by 1. Think about K being a subcomplex of S^2 and consider triangles. The only case when a triangle is positive is when it comes in to complete the surface of S^2. But, if we further require K to be embedded in $\mathbb{R}^2$, then we cannot have a 2-cycle and all triangles are negative ($\beta_2(K) = 0$). The case of tetrahedra in K that is a subcomplex of a triangulation of S^3, sitting in $\mathbb{R}^3$, is similar (we get $\beta_3(K) = 0$).

The idea of classifying simplices as positive and negative can be extended to arbitrary dimensions, though.

Algorithm for $(\beta_0, \beta_1, \beta_2)$ when $K \subset$ triangulation of S^3

integer^3 BETTI

$\beta_0 = \beta_1 = \beta_2 = 0$;

for $j = 1$ to m do → filtration is $K^0 \subseteq \cdots \subseteq K^m = K$, $K^0 = \emptyset$ (m simplices in K)

$k = \dim \sigma^j$;

if σ^j belongs to a k-cycle in K^j then

β_k++; ($\beta_k \leftarrow \beta_k + 1$) ↳ non-trivial in general!

else

β_{k-1}--; ($\beta_{k-1} \leftarrow \beta_{k-1} - 1$)

end if

endfor

return $(\beta_0, \beta_1, \beta_2)$;

We can understand the logic behind this algorithm using the formulae for Betti numbers. Recall,

$\beta_k = z_k - b_k$ —(1)

$\beta_{k-1} = z_{k-1} - b_{k-1}$ —(2)

$s_k = z_k + b_{k-1}$ —(3)

When we add a k-simplex, s_k goes up by 1, which we indicate by $s_k \uparrow$.

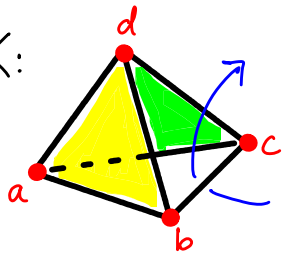
If $s_k \uparrow$, then (3) $\Rightarrow$ either $z_k \uparrow$ or $b_{k-1} \uparrow$.

If $z_k \uparrow$, then (1) $\Rightarrow$ $\beta_k \uparrow$; and

if $b_{k-1} \uparrow$, then (2) $\Rightarrow$ $\beta_{k-1} \downarrow$.

Example

K:



filtration: a, b, c, d, ab, ac, ad, bc, bd, cd, abd, acd ($m=12$)

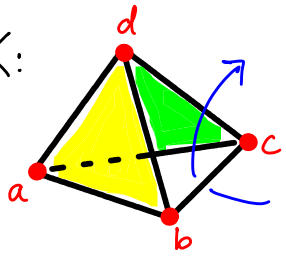
There is one connected component, and one hole.

MATH 529 – Lecture 20 (03/21/2024)

Today:
* Example for incremental algorithm
* UNION-FIND data structure
* persistence algorithm

Example

K:



filtration: $a, b, c, d, ab, ac, ad, bc, bd, cd, abd, acd$ $(m=12)$

There is one connected component, and one hole.
↳ There are two triangles missing from the tetrahedron, but there is only one hole, as indicated by the arrow 'C'. Also, adding one more triangle gives a disc!

So we expect $\beta_0=1, \beta_1=1$.

Steps of the algorithm ($\beta_0=\beta_1=0$, $\beta_2=0$ to start)

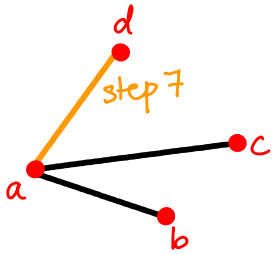
↳ we have no positive triangles. So β_2 remains at zero.

1-4: a, b, c, d: β_0 goes $1\to2\to3\to4$ (all are positive simplices)

5: ab: β_0--; $(\beta_0=3)$ (ab is a negative simplex)

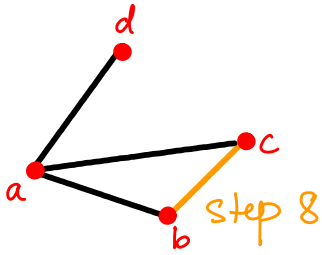
6: ac: β_0--; $(\beta_0=2)$ (ac is a negative simplex)

7: ad: β_0--; $(\beta_0=1)$ (ad is a negative simplex)

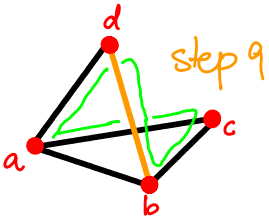


8: bc: β_1++; $(\beta_1=1)$ (bc is a positive 1-simplex)

bc is part of exactly one 1-cycle: $abca$



9. bd: β_1++; $(\beta_1=2)$ (bd is positive)
bd is part of at least one 1-cycle: $abda$, $acbda$.
We will study this aspect later. For now, note that bd is positive.

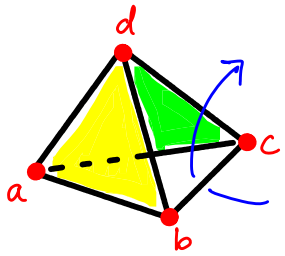


10. cd: β_1++; $(\beta_1=3)$ (cd is positive) → as with bd, cd is part of several 1-cycles

We now have $\beta_0=1, \beta_1=3$.

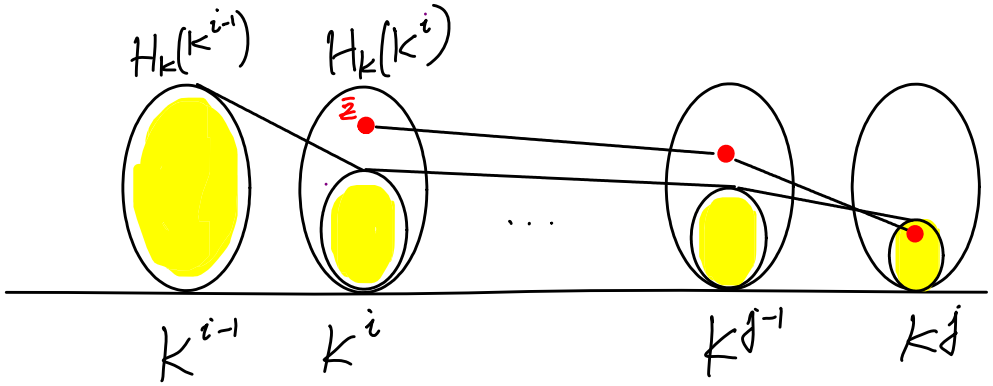
11. abd; β_1 -- ; ($\beta_1 = 2$) (abd is negative)

12. acd; β_1 -- ; ($\beta_1 = 1$) (acd is negative)



Finally, we get $\beta_0 = 1$, $\beta_1 = 1$, as expected.

Here is a visualization of how a homology class evolves:



The class $[\bar{z}]$ is created at K^i, as σ^i comes in, and is destroyed as it enters K^j, as σ^j comes in.

We can maintain connected components and classify (0- and) 1-simplices efficiently as positive or negative using the UNION-FIND data structure. This data structure maintains (and updates) a collection of connected components — think of each component as a set, with a label or ID. There are two operations one can perform on the collection, which are UNION and FIND.

Illustration of UNION-FIND Data Structure

The data structure maintains a set $\mathcal{S}$ of connected components.

Two operations (on members of $\mathcal{S}$)

$$\text{FIND}(v) = \begin{cases} \text{returns (the identity of) the set that } v \\ \text{belongs to, if it already belongs to one in } \mathcal{S} \\ \phi, \text{ otherwise.} \end{cases}$$

UNION(U,V): replaces U with $U \cup V$ (and deletes V).

Steps in the algorithm for the example ($\beta_0 = 0, \beta_1 = 0$ at start)

1-4: For $v = a,b,c,d$, FIND$(v) = \phi$, so we add $\{v\}$ as a new set to $\mathcal{S}$ each time, and do β_0++. ($\beta_0 = \beta_0 + 1$)

Now $\mathcal{S} = \{\{a\},\{b\},\{c\},\{d\}\}$, and $\beta_0 = 4$.

5-7: For $e = uv = ab, ac, ad$, do FIND(u), FIND(v)

If FIND$(u) \neq$ FIND(v)

UNION(FIND(u), FIND(v)); β_0--;

$\mathcal{S} = \{\{\underline{a},b\},\{c\},\{d\}\} \longrightarrow \{\{\underline{a},b,c\},\{d\}\} \longrightarrow \{\{\underline{a},b,c,d\}\}$

a represents its set in each case

We have $\beta_0 = 1$ now.

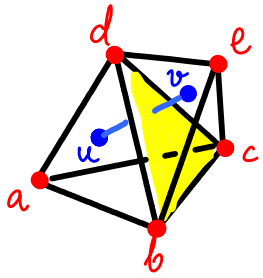
8-10: $uv = bc, bd, cd$.

In each case, FIND(u) = FIN(v) = a → set represented by a

so, β_1++;

$\mathcal{S} = \{\{\underline{a},b,c,d\}\}$ $\beta_0 = 1, \beta_1 = 3$.

For triangles, we use duality: consider a vertex for each tetrahedron and an edge for each triangle:



u is dual to $\triangle abcd$
v is dual to $\triangle bcde$
uv is dual to $\triangle bcd$

We can traverse the filtration in the reverse order, and classify triangles as positive or negative. Recall that all tetrahedra are negative.

The notions of positive/negative will be flipped when one considers the corresponding dual edge or vertex.

The incremental algorithm gives

$$\beta_k^l = pos_k^l - neg_{k+1}^l \quad \text{where}$$

pos_k^l = # positive k-simplices in K^l and

neg_{k+1}^l = # negative $(k+1)$-simplices in K^l.

This result also holds for each subcomplex K^l in the filtration. We indicate this fact by adding the superscript l in each term.

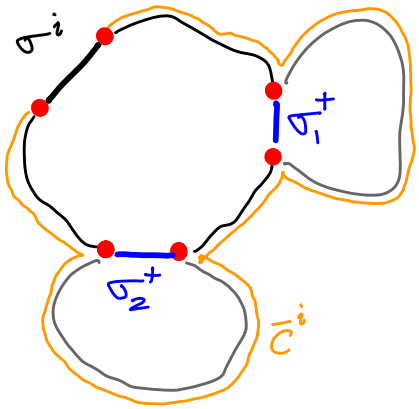
Persistence Algorithm (over $\mathbb{Z}_2$) (Edelsbrunner, Letscher, Zomorodian, 2002)

We can pair positive and negative simplices so that each pair represents a member in the homology class. The longer its "life", the more significant it is as a feature.

This original algorithm can be described as building on top of the incremental algorithm. But the matrix algorithm (SNF) can also be expanded to compute persistence, and the latter option generalizes to arbitrary dimensions.

We maintain a basis for H_k, which is empty to start with. For each positive k-simplex σ^i, we find a non-bounding k-cycle $\bar{c}^i$ that contains σ^i but no other positive k-simplex. $\bar{c}^i$ is the **canonical cycle** of σ^i.

We can show the canonical cycle always exists. Recall in the example, $\overline{bd}$ comes in at step 9, and is part of two cycles.



If we start with a cycle that has, say, two other positive k-simplices σ_1 and σ_2 apart from σ^i. Then we could add the canonical cycles corresponding to σ_1 and σ_2 to this cycle (addition is $\pm_2$), to get $\bar{c}^i$.

Now add $[\bar{c}^i]$ to the basis for H_k, i.e., the homology class of $\bar{c}^i$ is added as a new element of H_k. So, we represent $\bar{c}^i + B_k$ using $\bar{c}^i$, which in turn is represented using σ^i.

For each negative $(k+1)$-simplex σ^j, we find the corresponding positive k-simplex σ^i, and remove its class from H_k.

A general class is represented as $\bar{d} + B_k = \sum_g' (\bar{c}^g + B_k)$.

But $\sum_g' (\bar{c}^g + B_k) = \sum_g' \bar{c}^g + B_k$, so $\bar{d} \sim \sum_g' \bar{c}^g$.

Each $\bar{c}^g$ is represented by the positive k-simplex σ^g for $g < j$, that are not yet paired. This collection of all positive k-simplices in $[\bar{d}]$, written as $\Gamma' = \Gamma'(\bar{d})$, is uniquely determined by $\bar{d}$.

Note that $\bar{d}$ here is $\partial \sigma^j$ when considering the negative simplex σ_j.

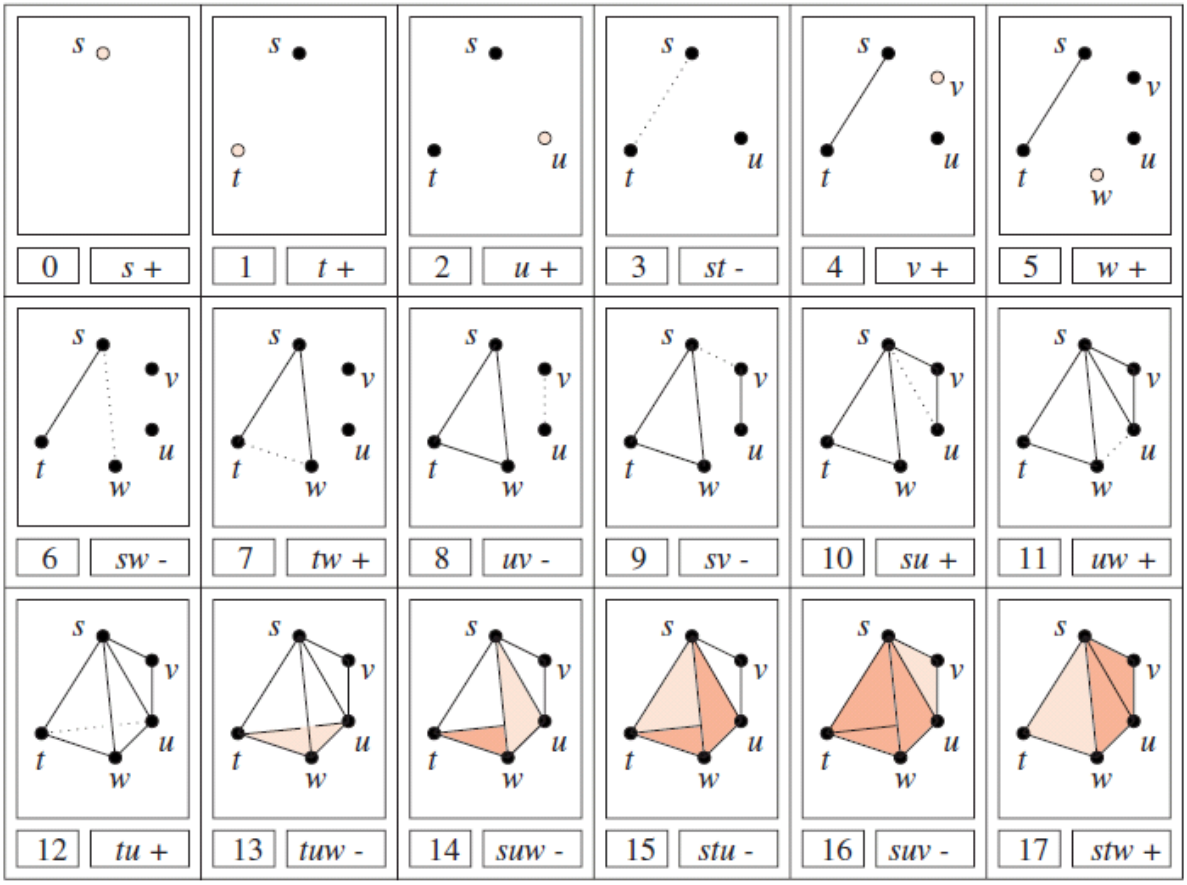
youngest : comes in the latest.

Let σ^i = the youngest positive k-simplex in $\Gamma'(\bar{d})$.

We record the pair (σ^i, σ^j), and its persistence as $j-i-1$.

We consider the example from the paper by Edelsbrunner, Letzcher, and Zomorodian (Topological Persistence and Simplification, 2002).

The final complex K is a hollow tetrahedron with a flap (i.e., it is a 2-complex). We get a filtration with 18 subcomplexes, corresponding to the 18-simplices, as shown in the next page.



We will go through all pairings in the next lecture...

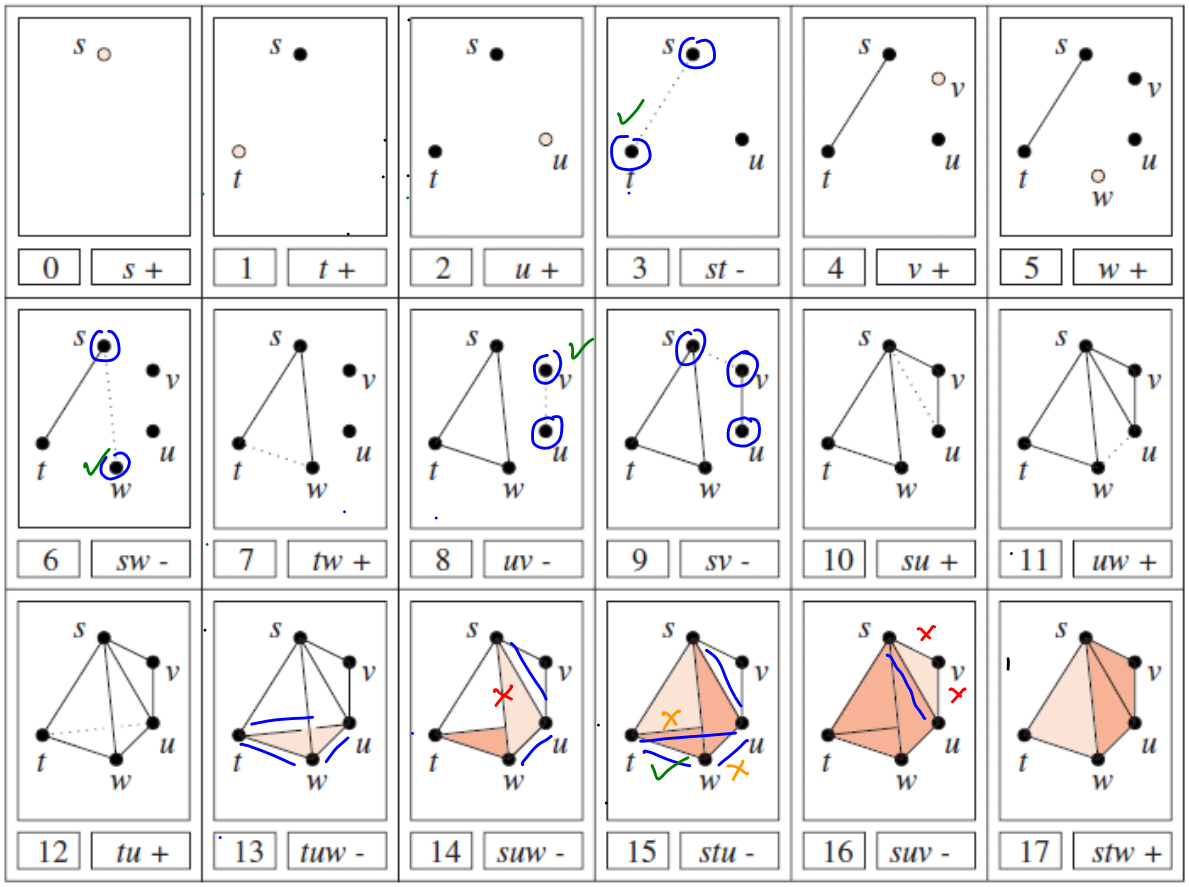
MATH 529 – Lecture 21 (03/26/2024)

Today: * example of persistence algorithm
* implementation of pairing

Example

The convention: the latest simplex to come in is shaded lightly (for vertices and triangles) or is shown dashed (edges).

In this filtration, all k-simplices do **not** come in before the $(k+1)$-simplices. Indeed, edge st comes in before vertices v and w. All we need to insure is that $K^i \subseteq K^{i+1}$ for the filtration, by insuring that all proper faces of σ^i come in before σ^i itself.



Pairings
(t, st)
(w, sw)
(v, uv)
(u, sv)
(tu, tuw)
(uw, suw)
(tw, stw)
(su, suv)

$\Gamma(tuw) = \{tu, uw, tw\}$

$\Gamma(suw) = \{uw, su\}$

$\Gamma(stu) = \{tu, su\}$ → uw, tw

$\Gamma(suv) = \{su\}$

Details of the pairing

For the negative 1-simplex sw, we consider $\Gamma = \{s, w\}$, and we choose the younger 0-simplex between s and w, which came in at time (or index) 0 and 5, respectively, pairing (w, sw).

Consider the negative 1-simplex sv. We start with $\Gamma = \{s, v\}$, and v is younger. But v is already paired, so we add u to Γ, as u is homologous to v (since uvs is present). u is positive, and is younger than s. So we pair (u, sv).

This calculation illustrates that an edge could (often) be paired with a simplex that is not its own face.

Pairing for stu: Γ has ~~st~~ $\{tu, su\}$. Here, tu is younger but is already paired.
↳ st is $-ve$

Due to Δtuw, $\{tw, uw\}$ is homologous to tu. tw and uw are both positive, and hence are candidates. But uw is already paired. Hence we pair tw with stu.

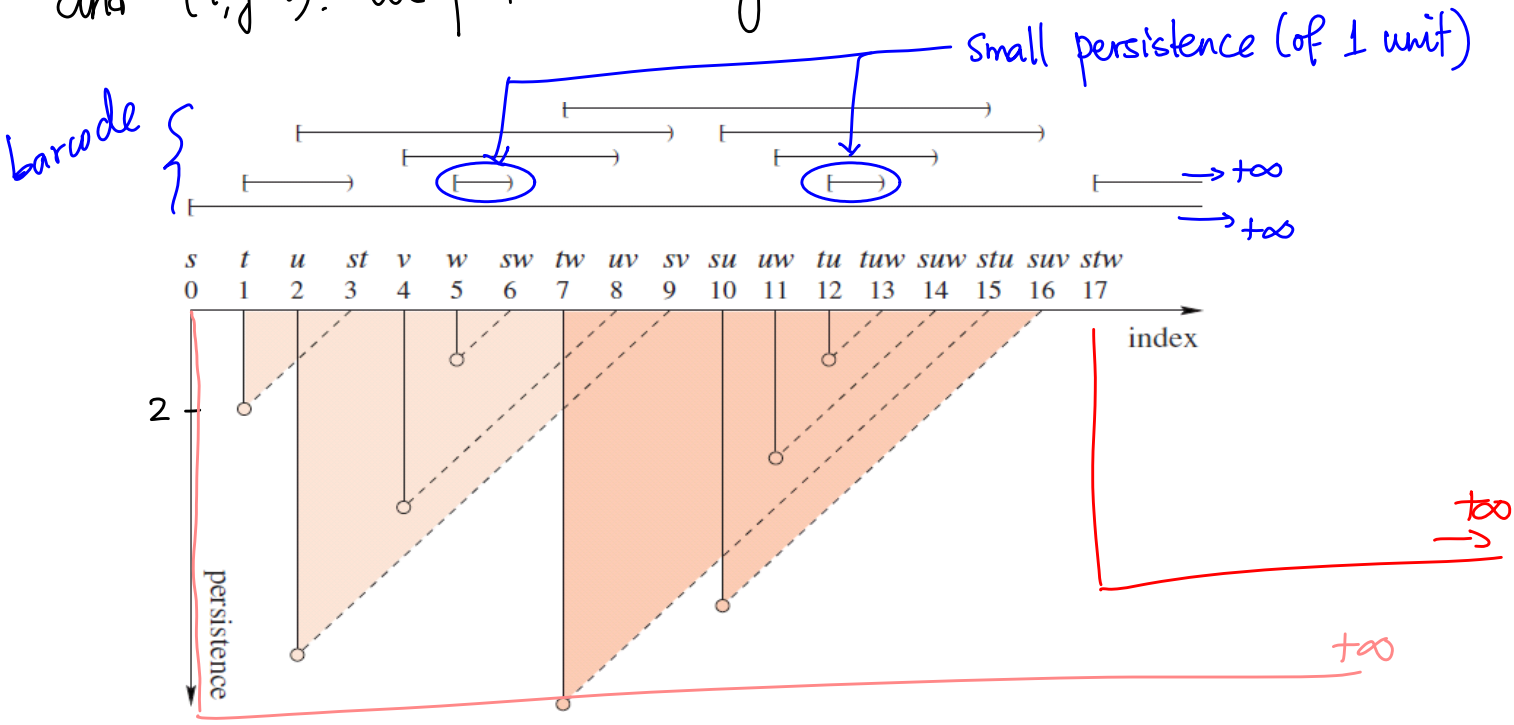
Finally, Γ for suv has just su, since sv and uv are both negative. Since su is still unpaired, we pair su with suv.

Notice that the 2-simplex stw is positive, and is left unpaired as there are no (negative) 3-simplices. Similarly, the first 0-simplex s, which is positive, is also left unpaired, representing the single connected component that is the final simplicial complex.

We visualize the pairings by converting the intervals to triangles that are open on one side.

Index-Persistence Diagram

For a pair (σ^i, σ^j), consider the triangle with vertices $(i,0)$, $(j,0)$ and $(i, j-i)$. We plot the triangles on the index-persistence axes.



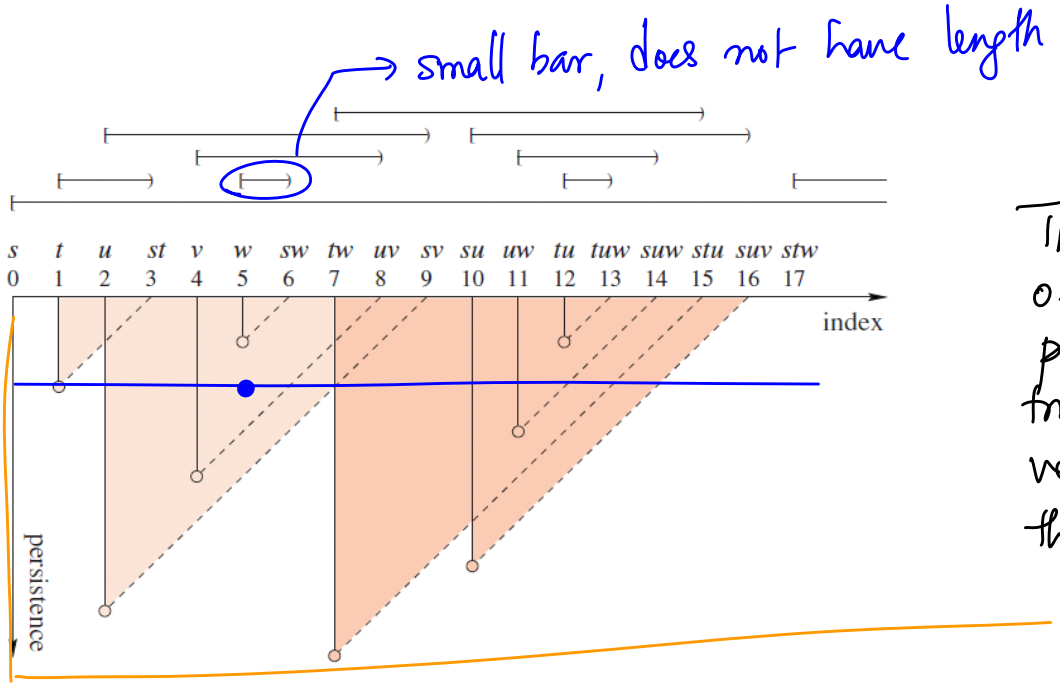
The positive 0-simplex s and positive 2-simplex stw are not paired, and hence have infinite persistence. These two infinite triangles are not shown.

Here, the lighter shade triangles correspond to (vertex, edge) pairs and the darker shade ones correspond to (edge, triangle) pairs.

We can compute the persistent Betti numbers by simply counting the numbers of triangles, as described below.

Theorem $\beta_k^{l,p}$ = the number of k-triangles containing the point (l,p) in the index-persistence plane.

For example, consider $(l,p)=(5,2)$, the point marked on the index-persistence diagram. There are 3 0-triangles containing the point $(5,2)$, and no 1-triangles. Thus, there are 3 connected components that have persistence at least 2 in K^5. Notice that the subcomplex K^5 has 4 connected components (refer to the last box in Row 1 of the filtration figure). But one of these 4 components — represented by vertex w, merges with the bigger component represented by vertex s in the next step, and hence is not 2-persistent.



There are indeed 3 0-triangles containing the point $(5,2)$. The infinite triangle corresponding to vertex s (at index 0) is the non-obvious one!

Implementation of Pairing

How do we find the youngest k-simplex in $\Gamma(\bar{d})$?

Store the index of pairings in a linear array $T[0, \ldots, m-1]$. The pair (σ^i, σ^j) is stored by setting $T[i] = j$. We also store Λ^i = list of positive simplices representing the cycle created by σ^i and destroyed by σ^j. These simplices in Λ^i are not necessarily only the ones in $\bar{d} = \partial_{k+1}(\sigma^j)$, but could include simplices in chains homologous to these simplices (in $\bar{d}$), and contains the youngest positive simplex.

For example, consider the negative edge sv coming in at step 9. $\partial_1(\overline{sv}) = \{s, v\}$. Here, v is younger than s, but it has already been paired (with edge uv in the previous step). But $u \sim v$, because of edge uv being already present. Hence $\Gamma(\overline{sv}) = \{s, u\}$, and u is indeed unpaired (and younger than s). Hence we pair u with sv.

We now describe the function to find the youngest positive k-simplex for pairing with a negative $(k+1)$-simplex.

```
integer YOUNGEST(σ j)
  Λ = {σ ∈ ∂_{k+1}(σ j) | σ is positive};
  while (true)
  {
    i = max(Λ);          largest index
    if T[i] is empty     → associate this Λ with T[i]
    {
      T[i] = j;  "store" Λ also in T[i]
      break;
    }
    else
      Λ = Λ +_2 Λ[i];    → the list associated with T[i]
  }
  return i;
```

Illustration on the Example filtration

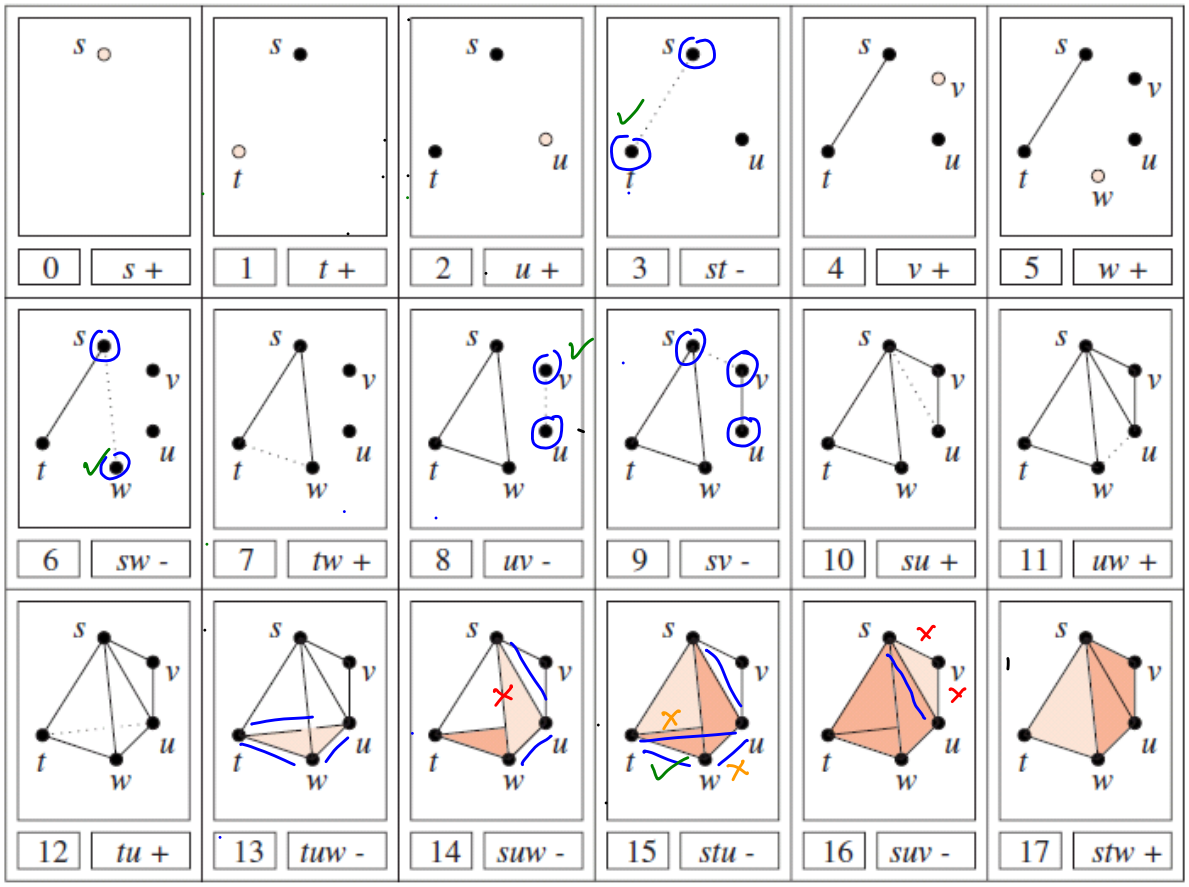
Negative simplices are indicated in blue

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

latest addition shown in blue (to T[4]).

For st, sw, and uv, the function does not go into the "else" part of if-statement (indices 3, 6, 8, respectively).

Here are the filtration and pairings again:



Pairings

(t, st)

(w, sw)

(v, uv)

(u, sv)

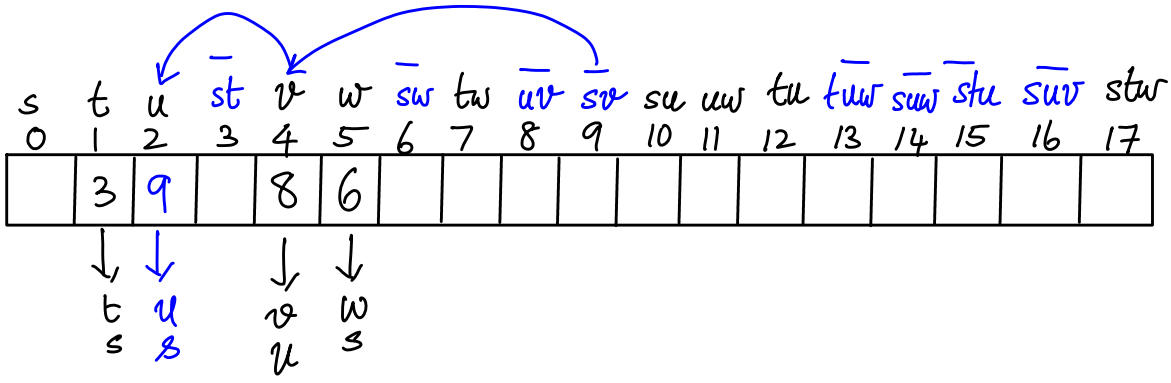
(tu, tuw)

(uw, suw)

(tw, stu)

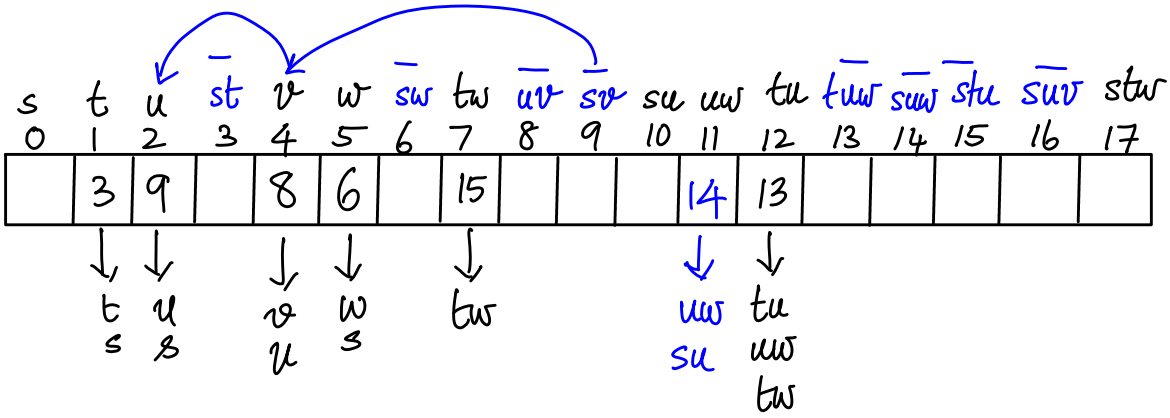
(su, suv)

For sv at index 9: $\Lambda = \{s, v\}$ v is younger, at 4. But $T[4]$ is not empty, and has $\Lambda[4] = \{v, u\}$. So, we set $\Lambda = \Lambda +_2 \Lambda[4] = \{s, v\} +_2 \{v, u\} = \{s, u\}$.



A step where we update Λ by taking sum modulo 2 as above is termed a **collision**.

We keep proceeding with the function till step/index 15, when Δstu comes in. The array T has the following form before Step 15.



stu (at step 15): $\Lambda = \{su, tu\}$.

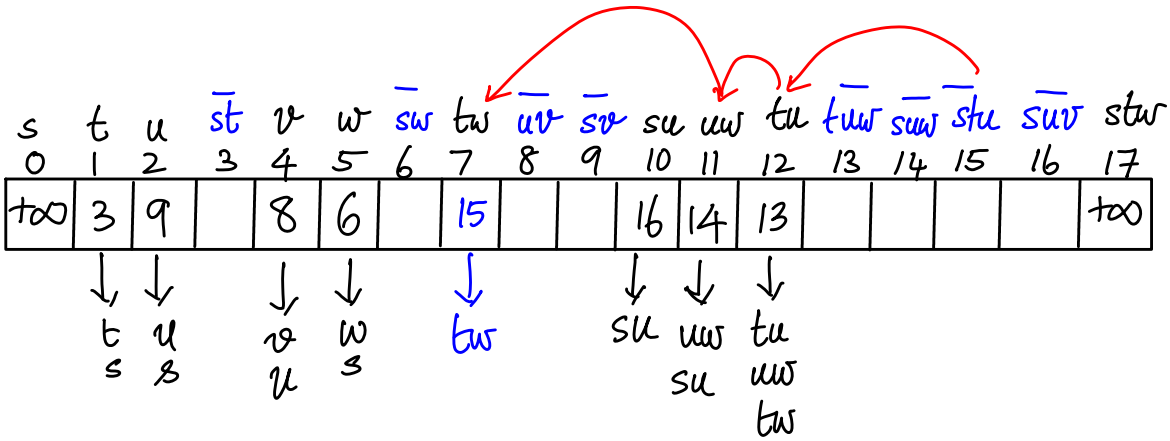
tu is positive, and is youngest, at 12, but $T[12]$ is not empty.

Notice that $st \in \partial_2(stu)$ is negative, and hence not included in Λ here.

$$\Lambda + \Lambda[12] = \{su, tu\} +_2 \{tu, uw, tw\} = \{su, uw, tw\}$$

Now, uw is youngest at 11. But, $T[11]$ is occupied, so we take another sum mod 2 (i.e., there is a second collision here).

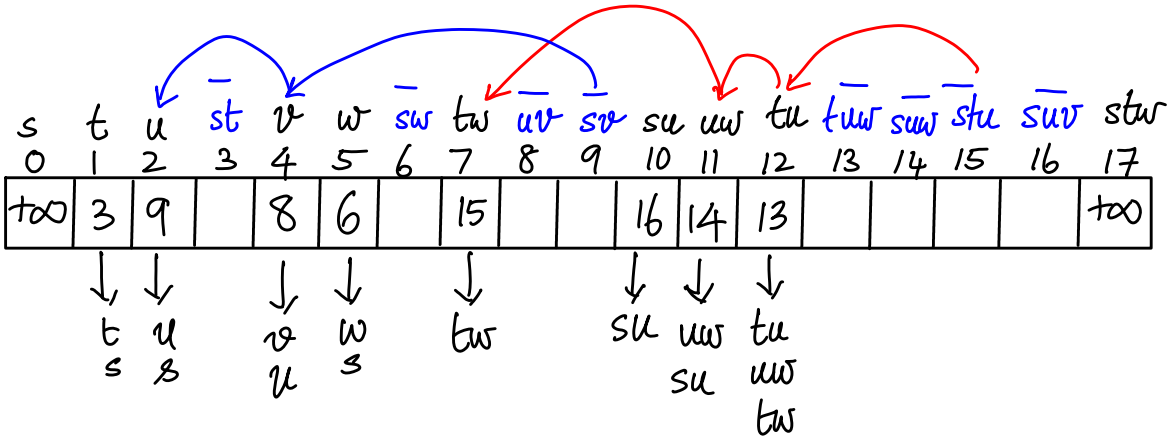
$$\Lambda + \Lambda[11] = \{su, uw, tw\} +_2 \{uw, su\} = \{tw\}.$$



We then pair suv (step 16) with edge su – there are no collisions in this case.

Finally, slots corresponding to the two remaining positive simplices s and stw are filled with ∞ (some large number in practice).

The final filled in array T is given below:



MATH 529 – Lecture 22 (03/28/2024)

Today: * more details of pairing/persistence
* matrix algorithm for persistence

Recall: $\emptyset = K^0 \subseteq \dots \subseteq K^m = K$ a filtered simplicial complex.

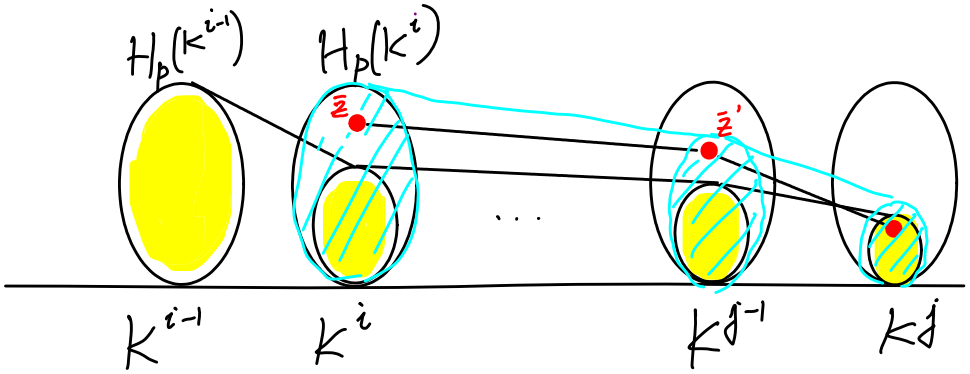
→ change of notation: we had $H_k^{l,p}$ previously; we'll need i,j,k,l,p soon!

$$H_p^{i,j} = Z_p(K^i) \Big/ \left(B_p(K^j) \cap Z_p(K^i)\right), \quad \beta_p^{i,j} = \operatorname{rank} H_p^{i,j}$$

$j > i$, usually.

$H_p^{i,j}$ captures non-trivial p-classes in K^i that stay active in K^j.
→ cycles that are not boundaries

We want to count exactly the # classes that are born in K^i and die in K^j. How do we do that? Recall the following picture (from Lecture 20):



We have $\beta_p^{i,j-1}$ = # classes in $H_p(K^i)$ that are still active in K^{j-1}

$\beta_p^{i,j}$ = # { subset of these classes that remain active in K^j }

$\beta_p^{i,j-1} - \beta_p^{i,j}$ = # classes active in K^i that die entering K^j.

Similarly, $\beta_p^{i-1,j-1} - \beta_p^{i-1,j}$ = # classes alive in K^{i-1} that die entering K^j.

Hence we get

$$\mu_p^{i,j} = \left(\beta_p^{i,j-1} - \beta_p^{i,j}\right) - \left(\beta_p^{i-1,j-1} - \beta_p^{i-1,j}\right)$$

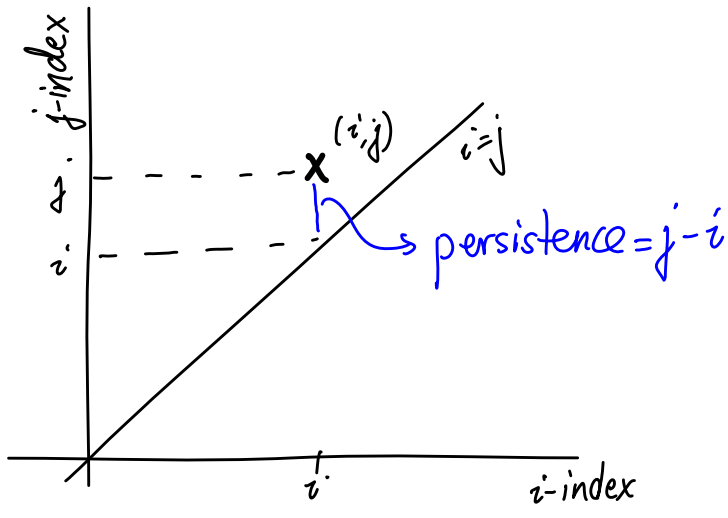
as the number of p-dimensional classes born at K^i and die entering K^j.

$\mu_p^{i,j}$ counts multiplicities (there could be more than one class born at K^i that dies entering K^j in general; but when we insist that $K^j = K^{j-1} \cup \{\sigma^j\}$, only one class will die).

p^{th}-persistence diagram

We pair σ^i a positive p-simplex with σ^j a negative $(p+1)$-simplex to capture a particular p-dimensional feature (or class). We plot the point (i,j) for each such pair on the index-index plane.

Since we assume a monotonic filtration, i.e., all proper faces are already present when simplex σ^j enters, we have that $i<j$ for all such pairs (i,j).



Hence all points are plotted above the $i=j$ ($45°$) diagonal.

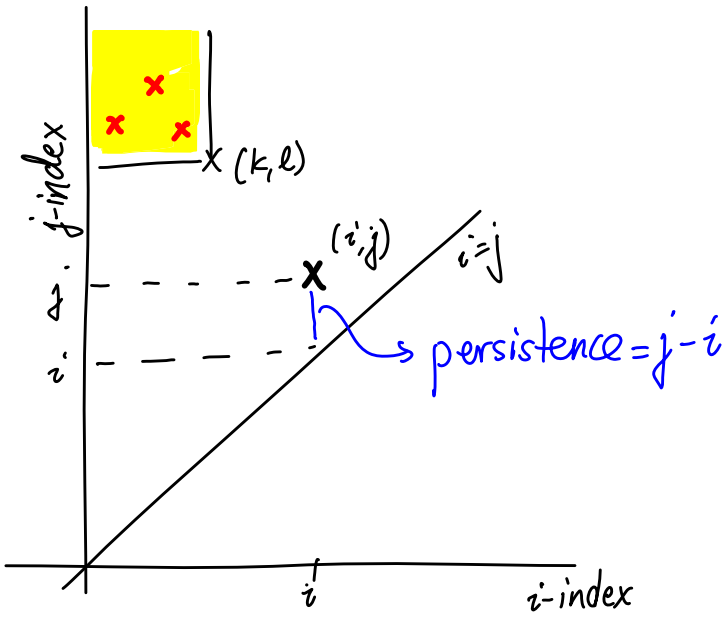
The persistence of (i,j) is then the vertical distance from the $i=j$ line ($45°$ line). In other words, the higher a point is above the diagonal line, the more persistent is the corresponding feature.

Technically, we also include $\{+\infty\}$ in both axes. Recall that unpaired positive simplices are assigned persistences of $+\infty$, and hence we plot the points $(i,+\infty)$ corresponding to such a simplex σ^i.

We can count the points in the persistence diagram to compute $\beta_p^{k,\ell}$:

$\beta_p^{k,\ell}$ = # points (with multiplicities) in the upper left quadrant with corner point (k,ℓ)

$\beta_p^{k,\ell} = 3$, as shown here using ×'s.



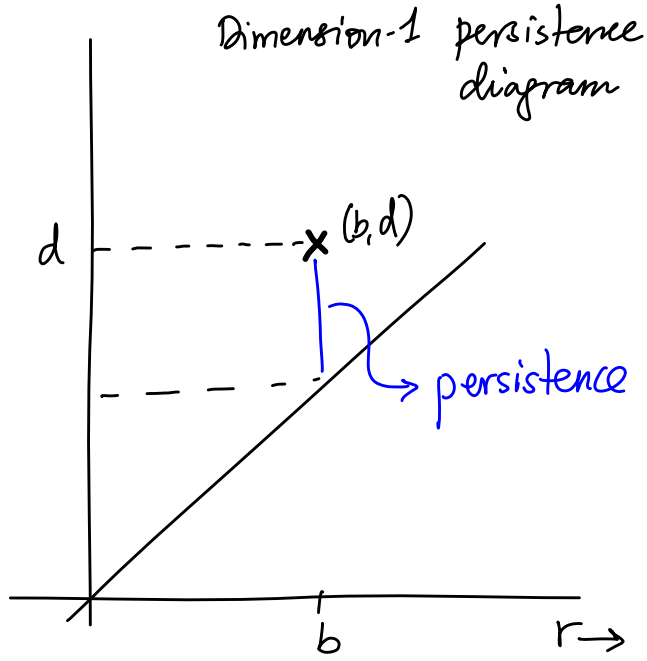
Fundamental lemma of persistent homology

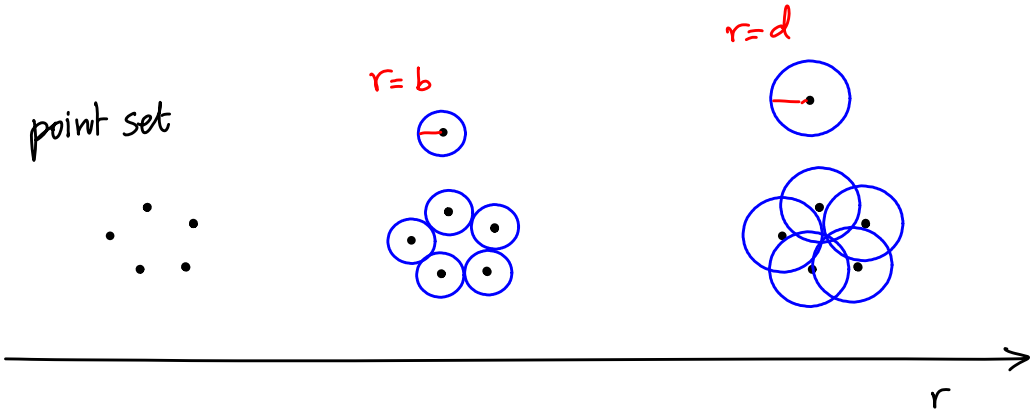
$$\beta_p^{k,\ell} = \sum_{i \le k} \sum_{j > \ell} \mu_p^{i,j} \qquad \text{for } k \le \ell.$$

$\mu_p^{i,j}$ → plotted (with multiplicities) in the persistence diagram.

Radius-based persistence diagram

While we've presented the entire pipeline of persistence for index-based filtrations, similar diagrams are obtained for points sets using, e.g., VR complexes, in terms of the radius values. We record pairs as (b, d), for instance, where the feature is born at $r = b$, and dies at $r = d$.

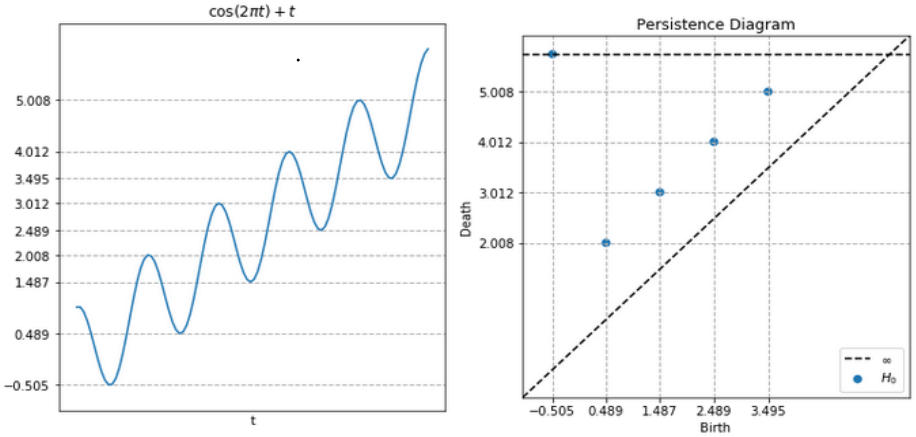




Sublevel-set persistence

Similar persistence diagrams are obtained when tracking the topology of evolving sublevel sets of functions.

(image from scikit-tda.org)



Matrix Reduction

We devise a matrix reduction algorithm to compute all pairings, i.e., compute the persistence diagram. The algorithm is motivated by the SNF algorithm to compute ranks of Z_p, B_p, H_p from $[\partial_p]$ $\forall p$. But we can also get the (σ^i, σ^j) pairings in the process!

We combine all boundary matrices with simplices ordered as in the monotonic filtration (σ^i enters in K^i, and all its faces are already present).

$$\partial[i,j] = \begin{cases} 1 & \text{if } \sigma^i \prec \sigma^j \text{ and } \dim(\sigma^i) = \dim(\sigma^j) - 1; \\ 0 & \text{otherwise} \end{cases}$$

→ in the final complex

$[\partial]$ is an $m \times m$ $\{0,1\}$-matrix (m = # simplices).

The filtration info is captured in the order of rows & column, i.e., the rows and columns are ordered the same way as the simplices come in.

We use replacement elementary column operations (over $\mathbb{Z}_2$) to "reduce" $[\partial]$ matrix to a matrix R.

→ $C_j + C_j$

new definition of reduction; not SNF!

Def Let $low(j)$ be the row index of the lowest 1 in column j. We call R **reduced** if $low(j) \neq low(j')$ whenever $j \neq j'$ specify two nonzero columns of R.

The algorithm reduces ∂ to R by adding columns from left to right ($C_j + C_i$ for $j > i$).

Here is the main block/function in the algorithm.

```
void REDUCE
    R = [∂];
    for j = 1 to m do
        while ∃ j' < j with low(j') = low(j) do
            add column j' to column j
        end
    end
```

This algorithm can be shown to run in $O(m^3)$ time.

$[\partial]$ is upper triangular to start with, and by the nature of these ECOs, we get that R is also upper triangular.

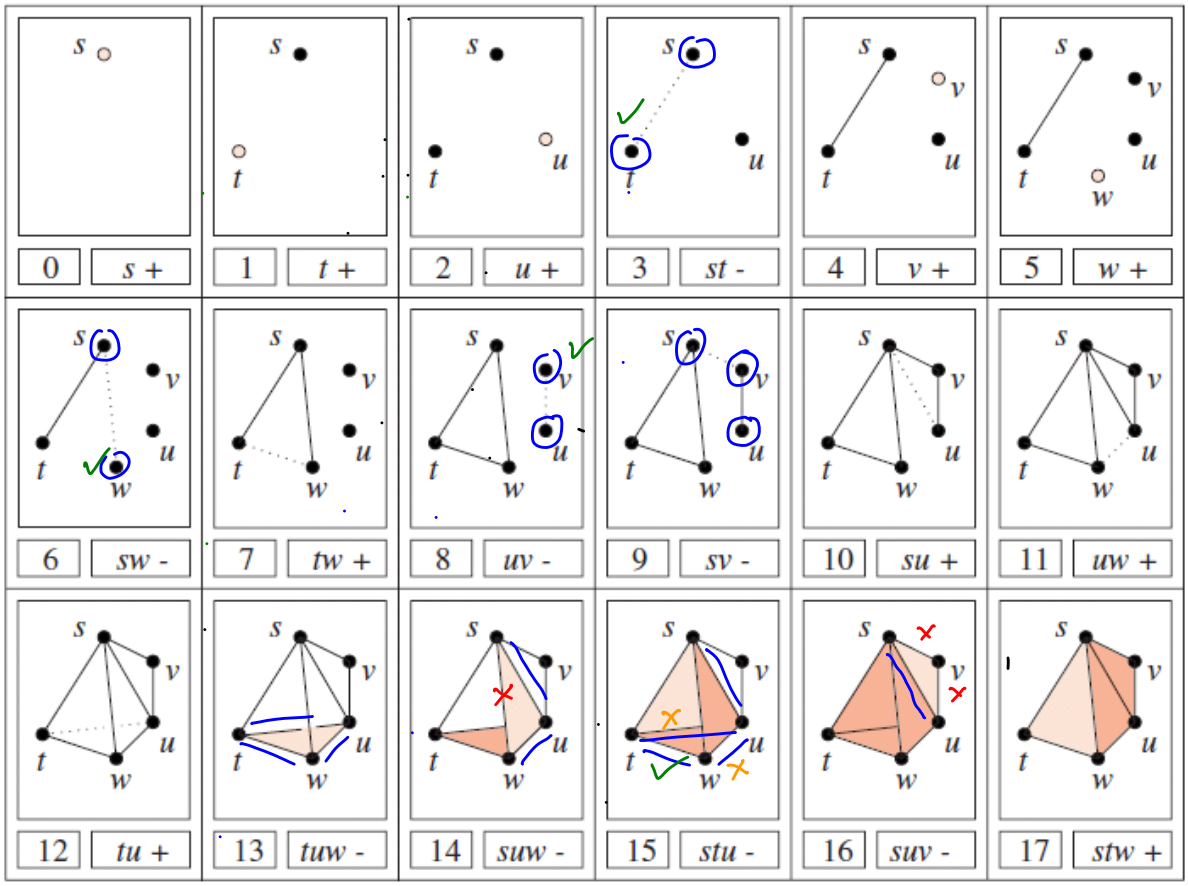
$$R = [\partial] V$$ → elementary matrix

In fact, R, $[\partial]$, and V are upper-triangular.

Note that we may not get block structure for $[\partial]$ in general. But if the filtration has the property that all p-simplices come in before any of the $(p+1)$-simplices do, then we will get a block structure — in fact, $[\partial]$ will be the block assembly of individual $[\partial_p]$ in that case.

Example

We apply matrix reduction to the complex we saw earlier. Recall the filtration and the pairings we obtained.



Pairings

(t, st)

(w, sw)

(v, uv)

(u, sv)

(tu, tuw)

(uw, suw)

(tw, stu)

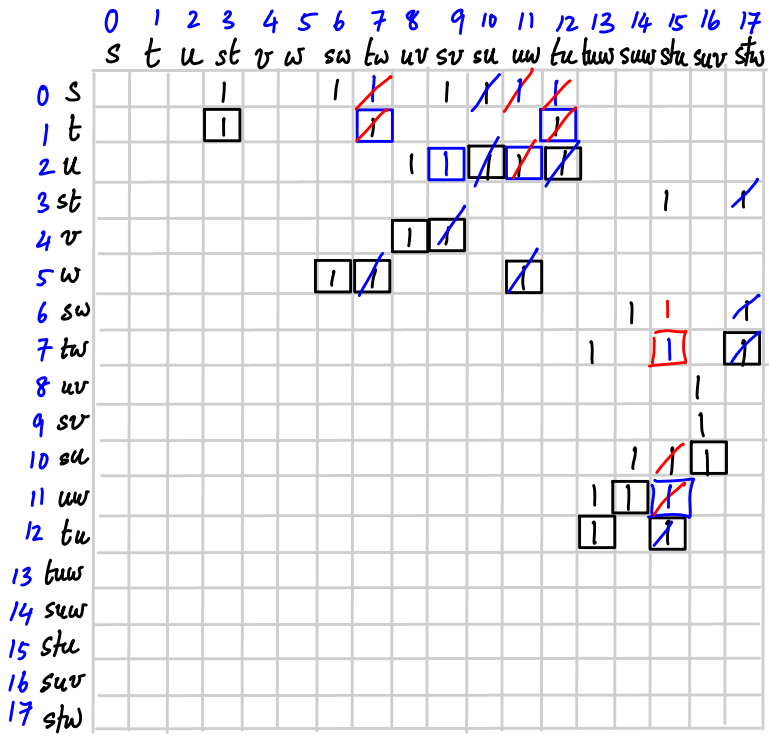
(su, suv)

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

Here is the combined boundary matrix. This [∂] is not necessarily in blocks above the diagonal — depending on the filtration, columns could be mixed up (while still maintaining upper triangularity).

We use the following conventions to show steps in the reduction algorithm.

1. $\text{low}(j)$ is indicated by $\boxed{1}$ (a box on the (i,j)-th entry).
2. Changes in the first ECO (for each column) are indicated in blue.
3. Changes in the second ECO (for each column) are shown in red.



ECOs

$j=7$ $\overset{7}{tw} + \overset{6}{sw} + \overset{3}{st}$

$j=9$ $\overset{9}{sv} + \overset{8}{uv}$

$j=10$ $\overset{10}{su} + \overset{9}{sv} + uv$

$j=11$ $\overset{11}{uw} + \overset{6}{sw} + \overset{9}{sv} + uv$

$j=12$ $\overset{12}{tu} + \overset{9}{sv} + uv + \overset{3}{st}$

$j=15$ $\overset{15}{stu} + \overset{13}{tuw} + \overset{4}{suw}$

$j=17$ $\overset{17}{stw} + \overset{15}{stu} + tuw + suw$

pairings:

$\overset{1}{(t,} \overset{3}{st)}$, $\overset{5}{(w,} \overset{6}{sw)}$, $\overset{4}{(v,} \overset{8}{uv)}$, $\overset{2}{(u,} \overset{9}{sv)}$

$\overset{12}{(tu,} \overset{13}{tuw)}$, $\overset{11}{(uw,} \overset{14}{suw)}$, $\overset{7}{(tw,} \overset{15}{stu)}$, $\overset{10}{(su,} \overset{16}{suv)}$

These are the same pairings obtained by the YOUNGEST algorithm!

Details of reduction steps: For $j=7$, we first do $C_7 \overset{+}{\leftarrow}_2 C_6$, as $\text{low}(7) = \text{low}(6) = 5$ at that moment. The $(5,7)$-entry is zeroed out, $(7,0)$-entry is turned to $\underline{1}$, and $\text{low}(7)=1$ after this ECO. Now, $\text{low}(7) = \text{low}(3)$, so we do $C_7 \overset{+}{\leftarrow}_2 C_3$, which zeros out both 1's in the $j=7$ column.

MATH 529 – Lecture 23 (04/02/2024)

Reminder: Start thinking about options for the project (explore tutorials on scikit-tda.org)

Today: * details of pairing
* mapper algorithm

Details of Pairing

We show that the lowest 1's are unique, i.e., they do not depend on the final reduced matrix R.

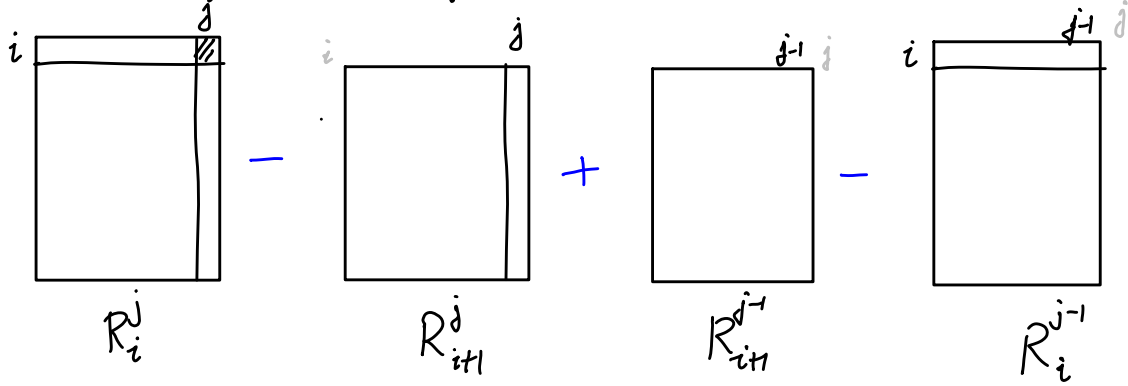
Let's look at R_i^j, the lower left submatrix whose corner element is $R[i,j]$, i.e., remove rows 1 to $i-1$, and columns $j+1$ to m (last $m-j$ columns).

Since we do ECOs left to right, the ranks of R_i^j are preserved.

In particular, $\operatorname{rank} R_i^j = \operatorname{rank} [\partial]_i^j \quad \forall i,j$.

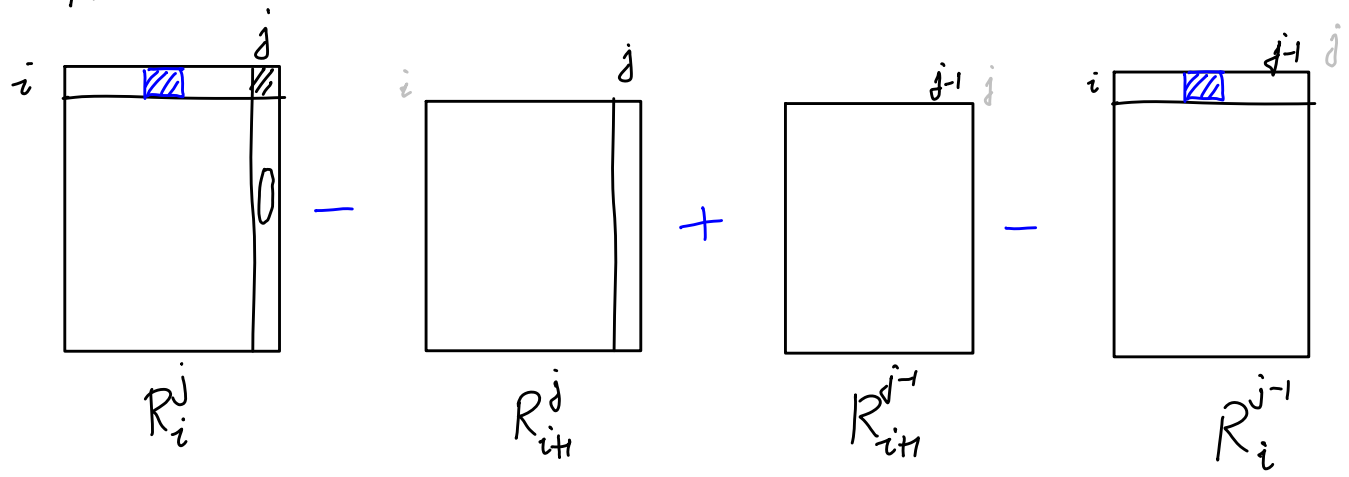
Also $\operatorname{rank} R_i^j$ = # nonzero columns in R_i^j. You can see this result by observing that any combinations of nonzero columns of R remain non-zero, following how we apply ECOs to obtain R in the first place.

Consider $r_R(i,j) = \operatorname{rank} R_i^j - \operatorname{rank} R_{i+1}^j + \operatorname{rank} R_{i+1}^{j-1} - \operatorname{rank} R_i^{j-1}$.



Note that $r_R(i,j) = r_{[\partial]}(i,j) \quad \forall i,j$, the similar submatrix obtained from $[\partial]$ instead of R.

$$r_R(i,j) = \text{rank } R_i^j - \text{rank } R_{i+1}^j + \text{rank } R_{i+1}^{j-1} - \text{rank } R_i^{j-1}.$$



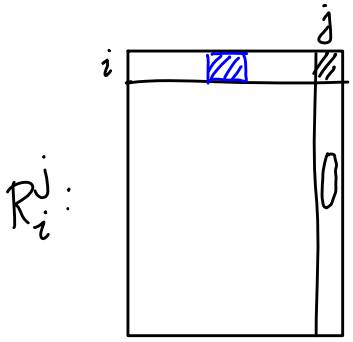
We consider two possibilities for $R[i,j]$:

$R[i,j]$ is a lowest 1 $\Rightarrow$ R_i^j has one extra nonzero row than other 3 matrices

$\Rightarrow$ $r_R(i,j) = 1$.

$R[i,j]$ is not a lowest 1: We consider two subcases here.

- There is no lowest 1 in Row i in columns 1 to $j-1$.
  - $\Rightarrow$ # nz columns in R_i^j = # nz columns in R_{i+1}^j
  - Similarly for R_{i+1}^{j-1} and R_i^{j-1}
  - $\Rightarrow$ $r_R(i,j) = 0$.
- There is a lowest 1 in row i in columns 1 to $j-1$:



$r_R(i,j) = 0$ here as well, since R_i^j will have an extra nonzero column than R_{i+1}^j, and similarly, R_i^{j-1} has an extra nonzero column than R_{i+1}^{j-1}.

Hence we get the result that (i, j) are paired independent of the exact form of R, i.e., it depends only on $[\partial]$.

Pairing Lemma $i = low(j)$ iff $r_j(i,j) = 1$. The pairings are independent of R, and depend only on $[\partial]$.

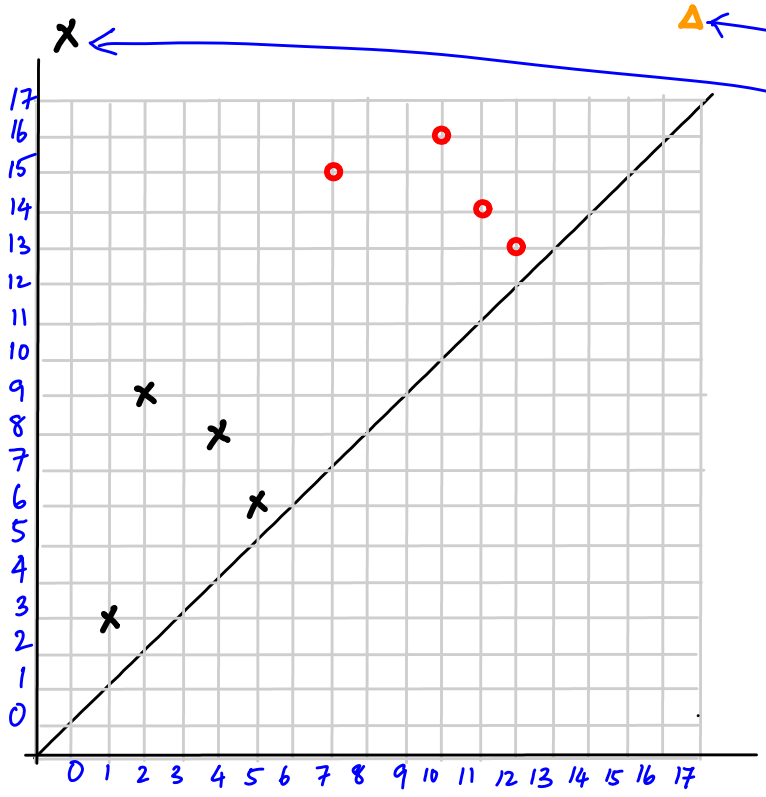
Classification of Simplices

* If column j of R is zero, then σ^j is positive.
* If column j of R is non-zero, σ^j is negative.

column j of R stores the boundary of the chain accumulated in the j^{th} column of V, where $R = [\partial]V$.

Here are the 0-th and 1^{st} persistence diagrams drawn together — typically, they are drawn separately.

→ along with the single unpaired triangle triangle that gives a point in the H_2 persistence diagram.



s $(0,\infty)$ stw $(17,\infty)$ unpaired positive simplices

pairings:

$(t, st)^{1,3}$, $(w, sw)^{5,6}$, $(v, uv)^{4,8}$, $(u, su)^{2,9}$

$(tu, tuw)^{12,13}$, $(uw, suw)^{11,14}$, $(tw, stu)^{7,15}$, $(sv, suv)^{10,16}$

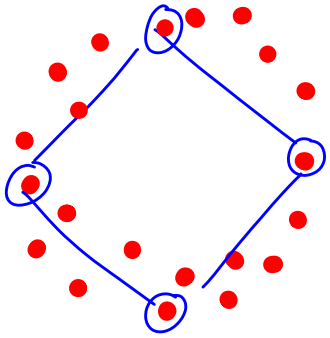
× : (vertex, edge)

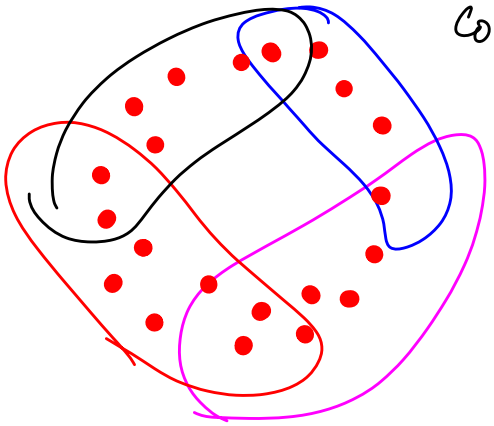
o : (edge, triangle)

Check out Ripser (part of Scikit-tda), GUDHI, Dionysis, etc.

The Mapper Algorithm

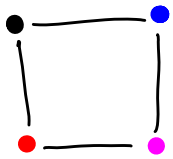
The idea is to create "highly compact" summaries of point cloud data. To this end we "cover" the range of values using some open sets, and then represent the points within each such set using a small number of nodes representing clusters. We could then capture intersections between two such sets by edges connecting their nodes. Here is a simple example for a set of points sampled from a circle.





Cover

"Nrv(Cover)"

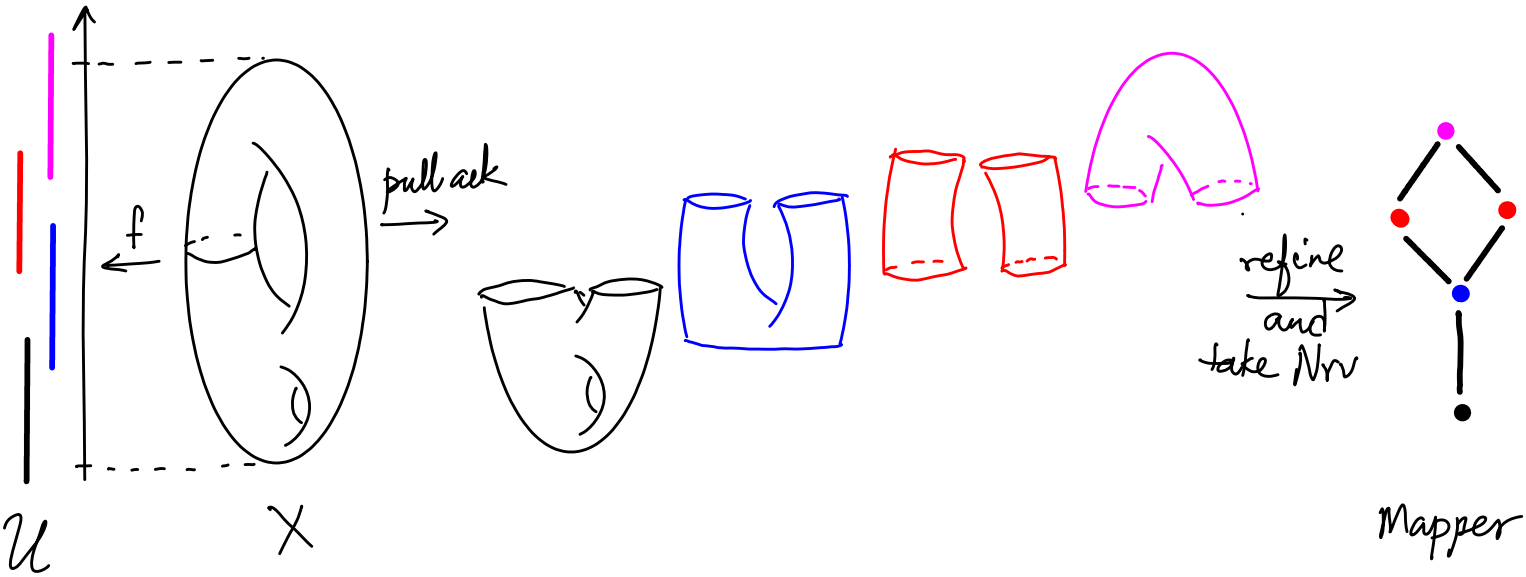


More formally, we take the nerve of the pull back of the cover — in other words, we get a simplicial complex that is much smaller in size than the input point cloud (for appropriate choices of the cover elements). Here, we get a circle with 4 nodes and 4 edges.

A (more) formal definition of mapper

Def Let $f: X \to \mathbb{R}^d$, $d \geq 1$, be a continuous real-valued function, and $\mathcal{U} = (U_i)_{i \in I}$ be an open cover of $f(X)$. The **pullback cover** of X induced by $(f, \mathcal{U})$ is a collection of open sets $(f^{-1}(U_i))_{i \in I}$. The **refined pullback** is the collection of connected components of the open sets $f^{-1}(U_i)$, $i \in I$. The **mapper** complex $M(f, \mathcal{U})$ is defined as the nerve of the refined pullback.

Here is another pictorial illustration on a connected object. We cover the range of the height of points on the object. → as opposed to a point cloud



Notice that we did "lose" the hole in the bottom portion of the object here!

MATH 529 – Lecture 24 (04/04/2024)

Today: * mapper algorithm
* Reeb graph

The Mapper Algorithm

INPUT: A set of points X with a metric

Function(s) $f_j : X \to \mathbb{R}$ (or $\mathbb{R}^d$), a cover $\mathcal{U}$ of $f(X)$ $\quad \mathcal{U} = (U_i)_{i \in I}$

$\hookrightarrow$ all f_j together taken as $f : X \to \mathbb{R}^d$

Steps: For each $U_i \in \mathcal{U}$, decompose $f^{-1}(U_i)$ into clusters $C_{i,1}, \dots, C_{i,k_i}$. $\quad k_i \geq 1 \ \forall i$

Compute Nrv of cover of X defined by $\{C_{i,1}, \dots, C_{i,k_i}\}, i \in I$.

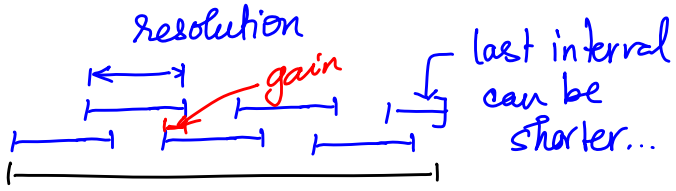
OUTPUT: The simplicial complex, which is the Nrv.

Def The functions f_i are called as **filter** functions. The metric (on X) used for clustering is called the **distance** function.

There are a lot of choices to be made by the user — which filters to use, what covers to use, etc. There are some theoretical results on when such choices are stable. But practitioners have been trying various values heuristically, and mapper has been getting used widely.

Default choice for filter functions:

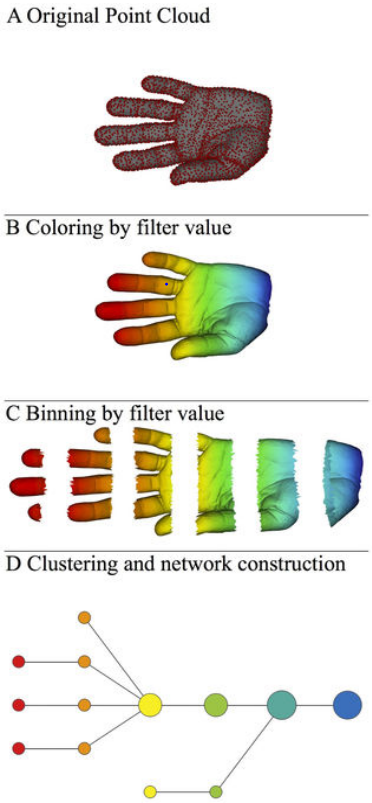
$f : X \to \mathbb{R}$



Length of (uniform) intervals: resolution

% overlap: gain

Here is another (more realistic) illustration from Lum et al. (2013). The input is ~1000 points sampled from the surface of a hand. Distance of a point from the wrist (base of hand) is used as the filter function.



The mapper representation takes us from ~1000 pts in $\mathbb{R}^3$ to a mapper (graph) with 13 nodes and 12 edges.

We now introduce some motivating mathematical concepts.

Reeb Graph

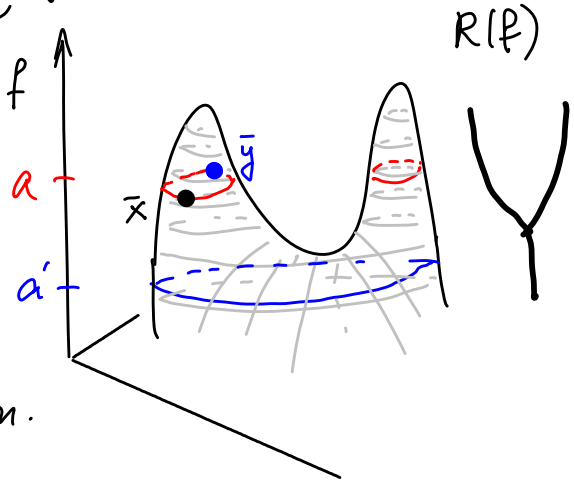
We study $\mathbb{X}$ and $f(\mathbb{X})$ and level sets of f on $\mathbb{X}$.

super / sub Level set: $\mathbb{X}_a = \{\bar{x} \in \mathbb{X} \mid f(\bar{x}) \overset{\geq}{\leq} a\}$.

The collection of all level sets forms a partition of $\mathbb{X}$.

Def We say $\bar{x}, \bar{y} \in \mathbb{X}$ are equivalent if they belong to the same connected component of a level set (or contour) of f.

Def The **Reeb graph** of f is the set of contours $R(f)$ together with the quotient topology induced by this equivalence relation.



We consider $\psi : \mathbb{X} \to R(f)$ as the contour map, where $\psi(\bar{x})$ is the contour that contains $\bar{x}$. We can show that ψ maps components to components, but could merge some cycles. Hence we get

$$\beta_0(R(f)) = \beta_0(\mathbb{X}), \text{ and}$$
$$\beta_1(R(f)) \leq \beta_1(\mathbb{X}).$$

Hence if $\mathbb{X}$ has no "holes", $R(f)$ is a truthful representation for sure. But even when $\mathbb{X}$ has holes, Reeb graphs find many applications, e.g., medical image analysis.

Morse theory studies the case where $\mathbb{X} = \mathbb{M}$, a compact manifold, and f is a Morse function (continuous function whose critical points are non-degenerate).

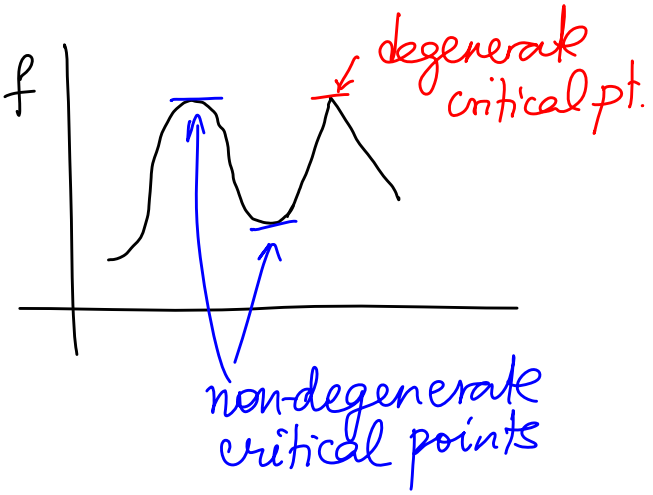
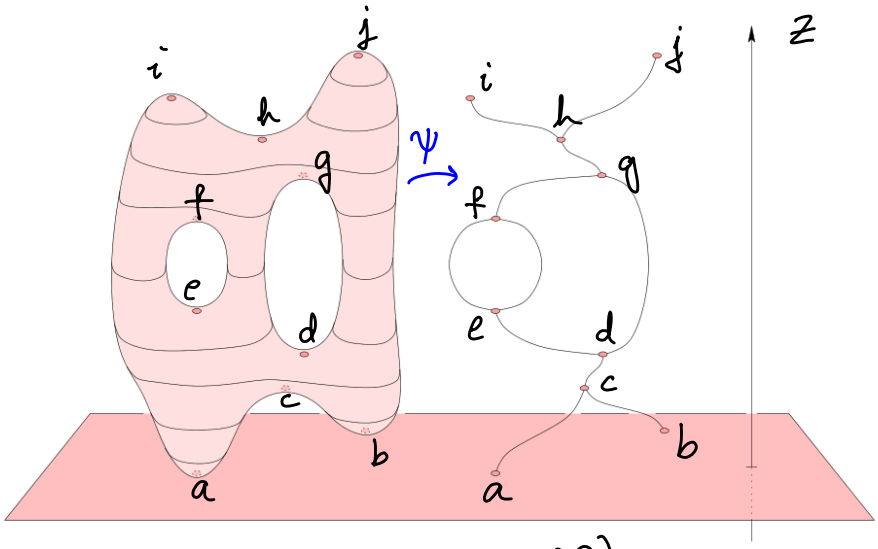


Illustration of Reeb graph

From Edelsbrunner and Harer



$$(\mathbb{X}, f) \longrightarrow R(f)$$

u is a node in $R(f)$ if it is the image of a critical point (under ψ).

Mapper is motivated by the Reeb graph construction, but aspires to be much more general. We will highlight some aspects of the cover choices and their (refined) pullbacks here.

Typical setting in Mapper: $f: \mathbb{X} \to Z$ a parameter space. Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be an open cover of Z. We study cover of $\mathbb{X}$ by $f^{-1}(U_\alpha)$.

If $\mathcal{V} = \{V_\beta\}_{\beta \in B}$ is another cover of Z, a **map of coverings** is function $g: A \to B$ s.t. $\forall \alpha \in A$, $U_\alpha \subseteq V_{f(\alpha)}$.

Given such a map of coverings g, there is an induced map of the simplicial complexes

$$N(g): Nrv(\mathcal{U}) \longrightarrow Nrv(\mathcal{V}).$$

Example 1 Let $\mathbb{X} = [0,n] \subseteq \mathbb{R}$, and $\epsilon > 0$. Then $I_l^\epsilon = (l-\epsilon, l+1+\epsilon) \cap \mathbb{X}$ for $l = 0,1,\dots,n-1$ forms an open cover $\mathcal{I}_\epsilon$ of $\mathbb{X}$.

For different choices of ϵ, we get different covers. The map $g: A \to B$ is identity when $I_l^\epsilon \subseteq I_l^{\epsilon'}$ for $\epsilon \le \epsilon'$.

Example 2 Let $\mathbb{X} = [0,2n] \subseteq \mathbb{R}$, and $\epsilon > 0$. We consider

$$I_l^\epsilon = (l-\epsilon, l+1+\epsilon) \cap \mathbb{X} \quad \text{for } l = 0,1,\dots,2n-1, \text{ and}$$

$$J_m^\epsilon = (2m-\epsilon, 2m+2+\epsilon) \cap \mathbb{X} \quad \text{for } m = 0,\dots,n-1.$$

$\mathcal{I}_\epsilon$: covering of $\mathbb{X}$ by I_l^ϵ, $l = 0,\dots,2n-1$.

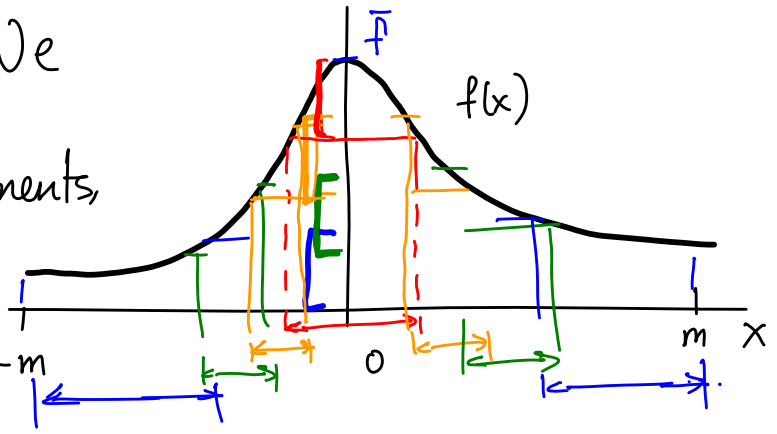
$\mathcal{J}_{\epsilon'}$: covering of $\mathbb{X}$ by $J_m^{\epsilon'}$, $m = 0,\dots,n-1$.

Then $g : \{0,\dots,2n-1\} \longrightarrow \{0,\dots,n-1\}$ is $g(l) = \left\lfloor \frac{l}{2} \right\rfloor$. This function induces the map between $\mathcal{I}_\epsilon$ and $\mathcal{J}_{\epsilon'}$ for $\epsilon \le \epsilon'$.

Example 3 We can extend to 2D easily. Consider $\mathbb{X} = [0,n] \times [0,n]$ $(\mathbb{X} \subset \mathbb{R}^2)$. A cover similar to the one in Example 1 can use

$$I_{i,j}^\epsilon = (l-\epsilon, l+1+\epsilon) \times (j-\epsilon, j+1+\epsilon).$$

Example 4 Let $\mathbb{X} = (-m, m) \subseteq \mathbb{R}$ and $f : \mathbb{X} \to \mathbb{R}$ be given as $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$. Then we have $Z = f(x)$ given by $\left[0, \frac{1}{\sigma\sqrt{2\pi}}\right] = [0, \bar{f}]$. We consider the pullback cover elements, i.e., $f^{-1}(U_\alpha)$ for $\alpha = 1, 2, 3, 4$.



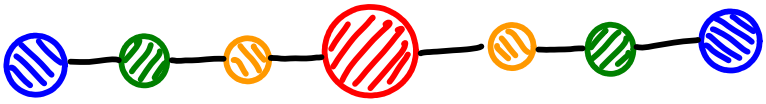
$f^{-1}(U_1)$ is 1 component.

$f^{-1}(U_2)$ is 2 components.

$f^{-1}(U_3)$ is 2 components.

$f^{-1}(U_4)$ is 2 components.

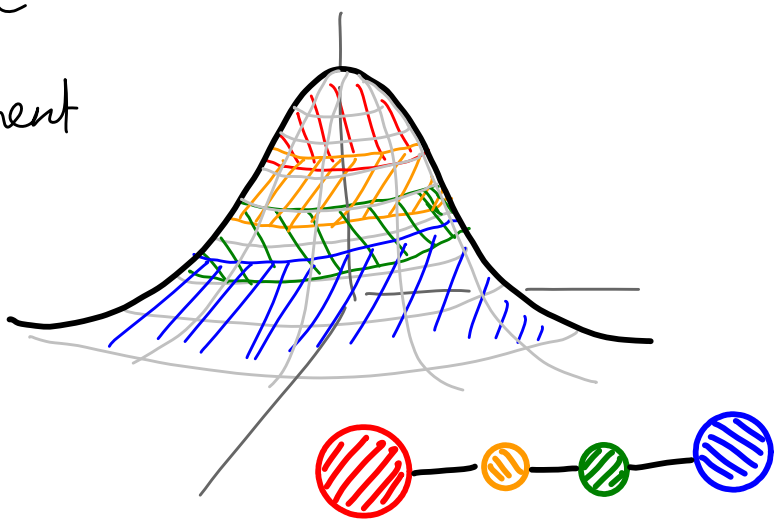
Here is the mapper. Sizes of nodes are proportional to relative sizes of the $f^{-1}(U_\alpha)$ sets.



Example 5 Let $\mathbb{X} = \mathbb{R}^2$, and we apply the same covering of Z as in Example 4 with

$$f(x,y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x^2+y^2)}{2\sigma^2}}.$$

Here, $f^{-1}(U_\alpha)$ is a single component for all $\alpha = 1, 2, 3, 4$. The corresponding mapper is shown, and has fewer nodes and edges than in Example 4.



Implementation

Find the range of values of a function of interest, Z.
Cover Z by choosing two parameters:

* interval length ℓ, and
* percentage overlap, p.

} We get $\mathcal{C}$, the (set of intervals in the) cover.

Example 6: Let $Z=[0,3]$, $\ell=1$, $p=20\%$. We get

$$\mathcal{C} = \{[0,1], [0.8,1.8], [1.6,2.6], [2.4,3]\}.$$

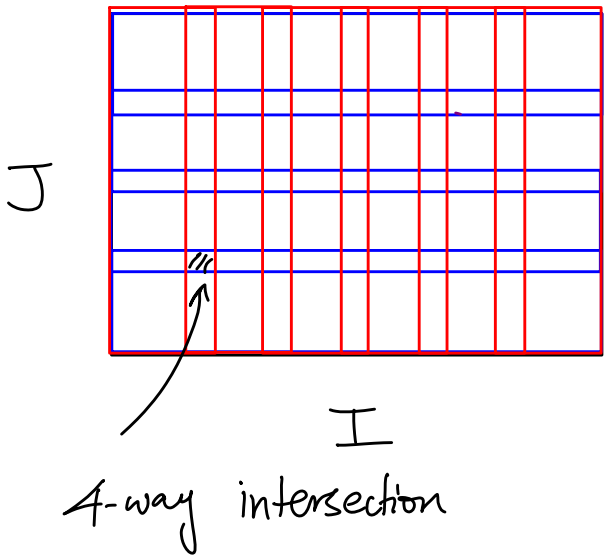
→ could go to 3.4;
→ $p=20\%$ is honored.

If the last interval is set as $[2.4, 3.4]$, then the length choice is also honored. But it is not too critical for the last interval.

For each interval $I_j \in \mathcal{C}$, find $X_j = \{\bar{x} \mid f(\bar{x}) \in I_j\}$, the set of points that form the domain of I_j. The X_j's form a cover of $\mathbb{X}$. For each X_j, find clusters $\{X_{jk}\}$. We could use a subset of original dimensions (even a single one), or all dimensions to do this clustering.

We represent each cluster by a node, and draw an edge between X_{jk} and X_{lm} when $X_{jk} \cap X_{lm} \neq \emptyset$, a triangle when three clusters have non-empty intersection, and so on.

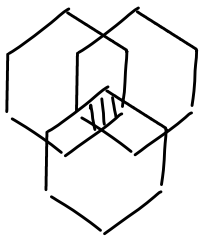
With two functions, your intervals are now rectangles (by default). But other shapes, e.g., hexagons, could be used for these intervals.



4-way intersection

With rectangular intervals, one would get 4-way intersections, and even higher order intersections if the overlap is larger. Hence we could get tetrahedra, or even higher dimensional simplices.

But with hexagonal cover elements, and appropriate overlap percentage(s), we could ensure we get at most 3-way intersections. Hence we get at most triangles in the mapper.



Check out the mapper implementations available in scikit-tda and giotto-tda.

MATH 529 – Lecture 25 (04/09/2024)

Today: * optimality in homology
* homology over $\mathbb{Z}$

Optimality in Homology

We consider homology groups over $\mathbb{Z}$ to define optimization problems. While $\mathbb{Z}_2$ might be the simpler ring to consider, it turns out $\mathbb{Z}$ opens up other possible approaches not possible with $\mathbb{Z}_2$ in this context.

Homology over $\mathbb{Z}$

$\mathbb{Z}$ has both $+z$ and $-z$ for any integer z. $\mathbb{Z}_2$ has only 1 (no $\pm$ signs)

We can use orientation of simplices/chains to capture the $\pm$ signs.

Recall induced orientation

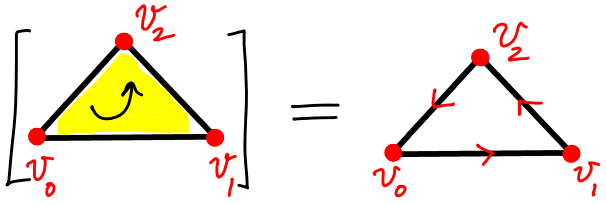
Let $\sigma = [v_0 v_1 \dots v_p]$ be a p-simplex and let $\tau = \text{conv}\{v_0, \dots \hat{v}_i, \dots v_p\}$ ($\hat{v}_i$ is omitted) be a $(p-1)$-face. The orientation induced on τ by σ is the same as $[v_0 \dots \hat{v}_i \dots v_p]$ if i is even.

The boundary of a simplex is defined as follows.

$$\partial_p \sigma = \sum_{i=0}^{p} (-1)^i [v_0 \dots \hat{v}_i \dots v_p].$$

For example,

$$\partial_2 \left[\begin{array}{c} \\ \end{array} \right] = \quad [v_1 v_2] - [v_0 v_2] + [v_0 v_1]$$

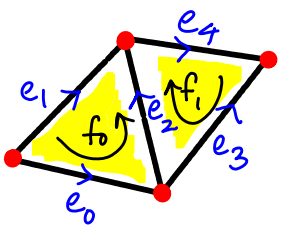


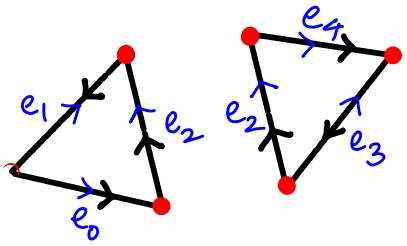
For a p-chain $\bar{c} = \sum_{j=1}^{n} a_j \sigma_j$, where $a_j \in \mathbb{Z}$ and σ_j's are p-simplices,

$$\partial_p \bar{c} = \partial \left(\sum_{j=1}^{n} a_j \sigma_j \right) = \sum_{j=1}^{n} a_j (\partial \sigma_j).$$

Choices of orientations for p-simplices become critical in boundary computations when working over $\mathbb{Z}$. Consider the 2-chain consisting of two triangles shown below.

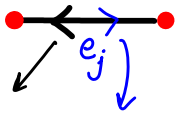
$$\partial_2 \left[\right] =$$





$$\bar{c}_1 = \begin{matrix} f_0 \\ f_1 \end{matrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

notation:



induced orientation from face

orientation of e_j

$$\partial_2 \bar{c}_1 = \bar{b}_1 = \begin{matrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{matrix} \begin{bmatrix} 1 \\ -1 \\ 2 \\ -1 \\ 1 \end{bmatrix}$$

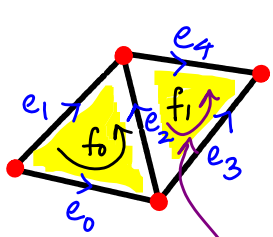
Notice that f_0 and f_1 are not consistently oriented. Induced orientations on e_2 are same, and hence e_2 has weight 2 in $\partial_2 \bar{c}_1$.

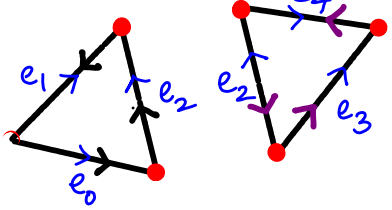
When working over $\mathbb{Z}_2$, e_2 will be counted twice, and hence will cancel when taking $\partial_2 \bar{c}_1$.

But, consider $\partial \bar{c}_2$ where $\bar{c}_2 = \begin{matrix} f_0 \\ f_1 \end{matrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

We get $\partial \bar{c}_2 = \bar{b}_2 = \begin{matrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{matrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \\ -1 \end{bmatrix}$

$$\partial_2 \left[\right] =$$





$$\bar{c}_2 = \begin{matrix} f_0 \\ f_1 \end{matrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

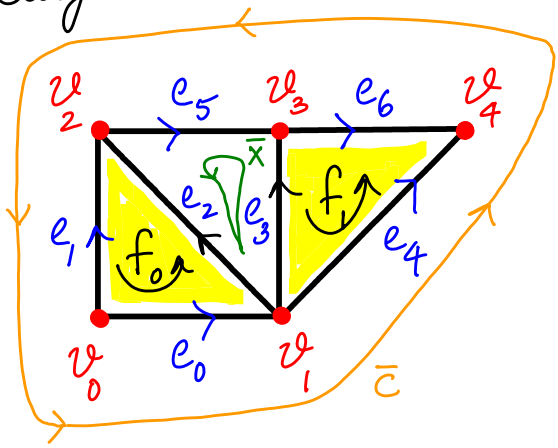
f_0 and f_1 are consistently oriented in $\bar{c}_2$, and hence e_2 cancels when taking boundary. e_2 will cancel as long as the weights for f_0 and f_1 are same in absolute value, and f_0 and f_1 are consistently oriented.

The definitions of C_p, Z_p, B_p, and H_p are similar to what we have seen before (when we were working over $\mathbb{Z}_2$). So, $\partial_p : C_p \to C_{p-1}$ are the boundary homomorphisms with addition over $\mathbb{Z}$ now, and $Z_p = \ker \partial_p$, $B_p = \operatorname{im} \partial_{p+1}$. Similarly, $H_p = Z_p / B_p$.

Let us consider our popular example now. We orient the triangles CCW, and edges lexicographically.

We now consider 1-cycles in K. In particular, we consider representative cycles around the hole. Each cycle is also oriented here.

K:



Consider the two cycles $\bar{c}$ and $\bar{x}$ shown here, which both go CCW around the hole in K. As vectors, we can write them as follows.

$$\bar{c} = \begin{matrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{matrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

$\to$ e_1 is directed opposite to the orientation of $\bar{c}$

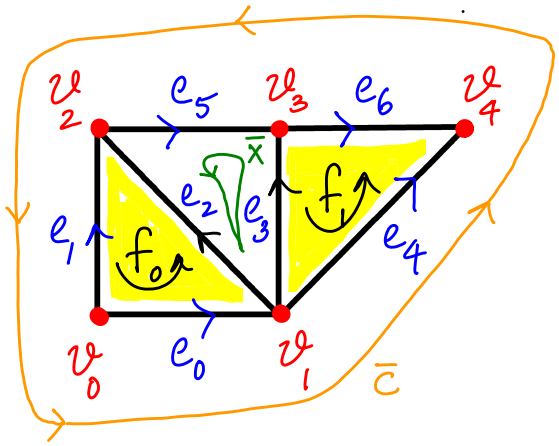
$$\bar{x} = \begin{matrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{matrix} \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

$\bar{c}$ and $\bar{x}$ are both cycles around the same hole, and hence are homologous. Similar to how we wrote such homology relations over $\mathbb{Z}_2$ using the boundary matrix, we form the boundary matrix over $\mathbb{Z}$ here. Now, the boundary matrices have entries in $\{-1, 0, 1\}$.

Compare with $[\partial_p]$ over $\mathbb{Z}_2$, which had entries in $\{0, 1\}$.

$$[\partial_2] = \begin{array}{c} \\ e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{array} \begin{array}{c} \begin{array}{cc} f_0 & f_1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} \end{array}$$

K:



In general, $[\partial_p]$ is an $m \times n$ matrix when K has m $(p-1)$-simplices τ_i and n p-simplices σ_j. The (i,j)-entry is nonzero if $\tau_i \leq \sigma_j$, and 0 otherwise. This nonzero entry is $+1$ if orientations of τ_i and σ_j agree, and -1 if they are opposite.

Homology between $\bar{c}$ and $\bar{x}$ can be written in terms of $[\partial_2]$, as follows.

$$\bar{x}: \begin{array}{c} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{array} \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \bar{c}: \begin{array}{c} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{array} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 1 \\ -1 \\ -1 \end{bmatrix} + \underbrace{\begin{array}{c} \\ e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{array} \begin{array}{c} \begin{array}{cc} f_0 & f_1 \end{array} \\ \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} \end{array} \begin{bmatrix} -1 \\ -1 \end{bmatrix}}_{[\partial_2]\bar{y} \text{ where } \bar{y} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}}$$

So, $\bar{x} \sim \bar{c}$ here.

Add boundaries of two triangles f_0, f_1, after reversing their orientations, to $\bar{c}$ to get $\bar{x}$.

Note that $\bar{c}$ has 5 edges and $\bar{x}$ has 3. Even if we were to include (Euclidean) lengths of the edges, and define

$$\|\bar{c}\| = \sum_{i=0}^{6} w_i |c_i| \quad \text{where} \quad w_i = 1 \text{ for horizontal/vertical edges } \{0,1,3,5,6\}$$
$$= \sqrt{2} \text{ for diagonal edges } \{2,4\}$$

We see that $\|\bar{x}\| = 2+\sqrt{2}$ and $\|\bar{c}\| = 4+\sqrt{2}$.

More generally, we could take w_i as the length of an edge, or the area of a triangle. Irrespective of the direction of traversal of an edge, we want to count its length — hence the absolute value $|x_i|$. We can seek a minimal weight chain $\bar{x}$ that is homologous to $\bar{c}$.

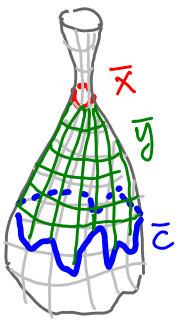
→ could generalize to chain when seeking $\bar{x} \sim \bar{c}$ for a chain $\bar{c}$.

Optimal Homologous Cycle Problem (OHCP)

Given $\bar{c}$ in $[\bar{c}] \in H_p(K)$, find $\bar{x} \in [\bar{c}]$ with $\|\bar{x}\| = \sum_{i=1}^{m} w_i |x_i|$, where $w_i \geq 0$ is the weight for σ_i the i^{th} p-simplex, is smallest.

We can prove an optimal homologous cycle always exists assuming K (simplicial complex) is finite, and $w_i \geq 0 \ \forall i$.

We can cast OHCP as the following optimization problem:



$$\text{minimize} \quad \sum_{i=1}^{m} w_i |x_i| + \lambda \left(\sum_{j=1}^{n} v_j |y_j| \right)$$

→ generalization where we also minimize the area of triangles used to define $\bar{x} \sim \bar{c}$.

$$\text{subject to} \quad \bar{x} = \bar{c} + [\partial_{p+1}] \bar{y}$$
$$\bar{x} \in \mathbb{Z}^m, \ \bar{y} = \mathbb{Z}^n$$

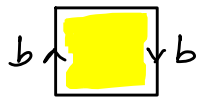
$\lambda \geq 0$ is a scale parameter that controls the relative importance of the two terms in the objective function.

Working over $\mathbb{Z}$ instead of $\mathbb{Z}_2$, gives us more "freedom", and actually makes the OHCP easier to solve under certain settings!

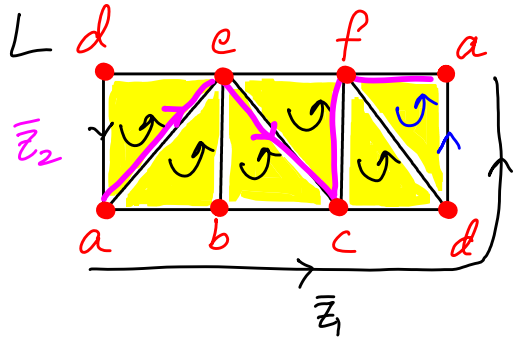
Homology over $\mathbb{Z}$ for some familiar objects

1. Möbius strip

M



L



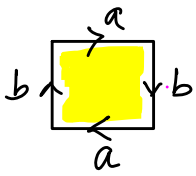
$\bar{z}_1 = a\text{-}b\text{-}c\text{-}d\text{-}a$

$\bar{z}_2 = a\text{-}e\text{-}c\text{-}f\text{-}a$

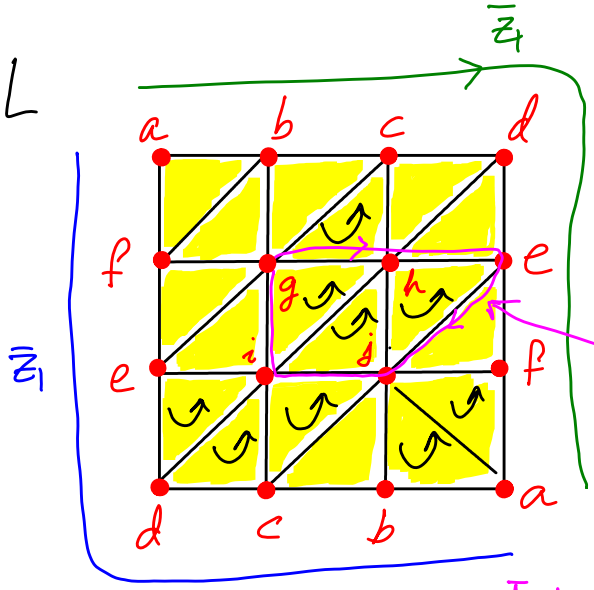
$\bar{z}_1 - \bar{z}_2 = \partial_2 \bar{d}$, where $\bar{d}$ is the 2-chain of triangles between $\bar{z}_1$ and $\bar{z}_2$.

We could go around $\bar{z}_1$ any integer # times. Hence $H_1 \cong \mathbb{Z}$, $\{\bar{z}_1\}$ is a basis, and $\beta_1(L;\mathbb{Z}) = \text{rank}(H_1;\mathbb{Z}) = 1$ here. $H_0(L;\mathbb{Z}) \simeq \mathbb{Z}$ as well (could use any vertex as a basis. Here, $\{\beta_0, \beta_1\}$ over $\mathbb{Z}$ coincide with those over $\mathbb{Z}_2$.

2. Projective Plane



L



$\bar{b}_1$ is a 1-boundary, bounding the 3 triangles inside.

$H_0(L;\mathbb{Z})$ has same rank as $H_0(L;\mathbb{Z}_2)$ ($\beta_0 = 1$ over both $\mathbb{Z}$ and $\mathbb{Z}_2$)

There is 1 connected component, so $H_0(K;\mathbb{Z})$ could be expected to be similar to $H_0(K;\mathbb{Z}_2)$.

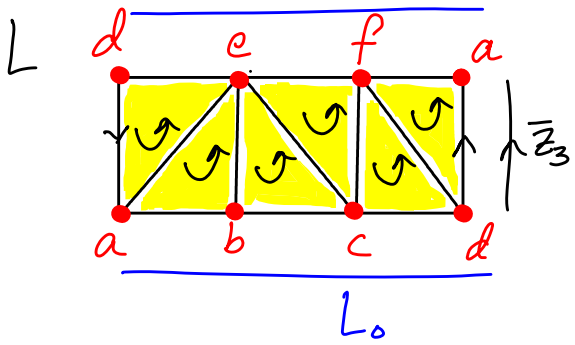
Let's consider H_1.

We can start by assigning CCW orientations for all triangles. Then the induced orientations for all edges "in the middle" are indeed opposite.

Consider the cycle $\bar{z}_1 = abcdefa$. It turns out that $2\bar{z}_1$ is a boundary — it bounds the 2-chain that has all triangles with coefficient 1

So, we could repeat $\bar{z}_1$ an odd # times to get another cycle in H_1. But an even # times gives a boundary. So, here $H_1 \cong \mathbb{Z}_2$.

Back to Example 1: Möbius strip



Let $L_0 \subset L$ consist of the edges $\bar{ab}, \bar{bc}, \bar{cd}, \bar{de}, \bar{ef}, \bar{fa}$, and the vertices. Let $\bar{z}_3 = \bar{ad}$.

The induced orientations on $\bar{ad}$ from $\triangle ade$ and $\triangle adf$ do not cancel.

What is $H_1(L, L_0; \mathbb{Z})$?

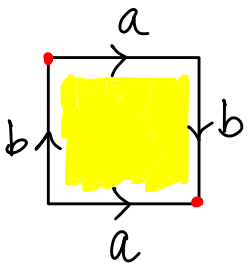
Note that all edges going "across" ($\bar{ad}, \bar{ae}, \bar{be}, \bar{ce}$, etc.) are relative 1-cycles. Also, if $\bar{d}$ is the 2-chain with coefficients 1 for each triangle, then $\partial_2 \bar{d} = 2\bar{z}_3$. Thus, $\bar{z}_3$ is not a relative 1-boundary, but $2\bar{z}_3$ is.

$$\Rightarrow \quad H_1((L, L_0); \mathbb{Z}) \cong \mathbb{Z}_2.$$

We will see later that Möbius strips are a canonical way in which homology groups that are not $\cong \mathbb{Z}$ show up in surfaces, and in 2-complexes more generally.

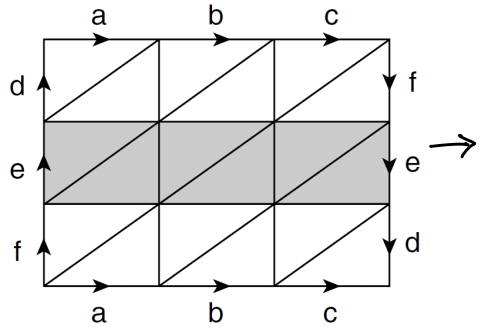
In fact, the "canonical" way in which a 2-manifold (w/ or w/o boundary) is non-orientable is by having a Möbius strip as a subcomplex.

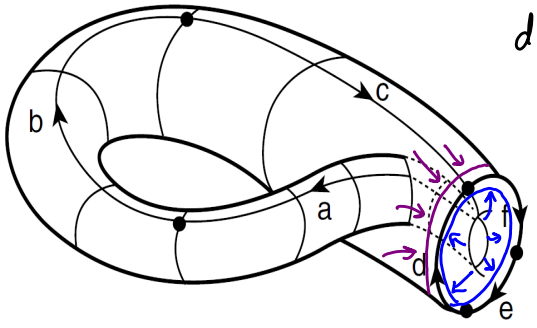
3. Klein Bottle ($\mathbb{K}^2$)



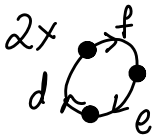
Notice that gluing along b produces an "intermediate" Möbius strip (before we also glue along a). Going multiple times around a shows the same behavior — a essentially behaves like the "skeleton" of this Möbius strip, thus capturing the main "hole" of the strip.

But, going around b twice after gluing along a gives a boundary — the boundary of the entire piece of paper. See the alternative triangulation and gluing below.





(images from P.J. Giblin – "Graphs, Surfaces, & Homology")



forms the boundary of the entire surface, when taken twice — once for the part "outside" and the second for the part "coming through".

Notice that going just once around gives a cycle that is not a boundary.

Hence $H_1(\mathbb{K}^2) \cong \mathbb{Z} \otimes \mathbb{Z}_2$ (also written as $\mathbb{Z}/_2$)

"free part" of H_1 — $\mathbb{Z}$

"torsion part" of H_1 — $\mathbb{Z}_2$

MATH 529 – Lecture 26 (04/11/2024)

Today: * torsion in homology groups
* more on OHCP: LP/IP, total unimodularity ...

Recall: $H_1(\mathbb{K}^2;\mathbb{Z}) \cong \mathbb{Z}\oplus\mathbb{Z}_2$, $H_1(\mathbb{RP}^2;\mathbb{Z}) \cong \mathbb{Z}_2$, $H_1(\mathbb{T}^2;\mathbb{Z}) \cong \mathbb{Z}\oplus\mathbb{Z}$.

Torsion subgroup: If G is an Abelian group, an element $g\in G$ has **finite order** if $ng = 0$ for some $n\in\mathbb{Z}$, $n>0$. ($ng \to g*g*\cdots*g$ (n times))

The set of all elements of finite order in G is a subgroup T of G, called the **torsion** subgroup.
If T does not contain any element other than the identity of G, we say G is **torsion-free**.

Fundamental Theorem of Finitely Generated Abelian Groups

Every finitely generated Abelian group G has a decomposition $G \cong F\oplus T$, where F is free abelian and T is the torsion subgroup of G, such that

$F \cong \underbrace{\mathbb{Z}\oplus\cdots\oplus\mathbb{Z}}_{\beta \text{ copies}}$ and $T \cong \mathbb{Z}_{t_1}\oplus\mathbb{Z}_{t_2}\oplus\cdots\oplus\mathbb{Z}_{t_k}$

where $t_i\in\mathbb{Z}$, $t_i>1$, and $t_1 | t_2 | \cdots | t_k$. ($t_1 | t_2$ means t_1 divides t_2)

This form $F\oplus T$, with the structure of F and T as shown, is a canonical form of G.

β is the **Betti number**, and t_i's are the **torsion coefficients** of G.

e.g., $H_1(\mathbb{K}^2;\mathbb{Z}) \cong \mathbb{Z}\oplus\mathbb{Z}_2 \Rightarrow \beta = 1,\ t_1 = 2.$
$H_1(\mathbb{RP}^2;\mathbb{Z}) \cong \mathbb{Z}_2 \Rightarrow \beta = 0,\ t_1 = 2.$

Back to OHCP → cycle

Given $\bar{c}$ in $[\bar{c}] \in H_p(K)$, find $\bar{x} \in [\bar{c}]$ with $\|\bar{x}\| = \sum_{i=1}^{m} w_i |x_i|$, where $w_i \geq 0$ is the weight for σ_i the i^{th} p-simplex, is smallest.

Q. Does an optimal homologous cycle always exist?

Yes, as long as K is finite and all $w_i \geq 0$.

We can prove that an optimal chain does indeed exist, mainly because of the finiteness of the complex. The challenge is in finding it efficiently.

Let's consider the optimization model for OHCP:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} w_i |x_i| \\ \text{s.t.} \quad & \bar{x} = \bar{c} + [\partial_{p+1}]\bar{y} \\ & \bar{x} \in \mathbb{Z}^m, \ \bar{y} \in \mathbb{Z}^n \end{aligned}$$

minimize ← min; s.t. → "subject to"

→ piecewise linear → can linearize by replacing x_i with $x_i^+ - x_i^-$ in constraints, and $x_i^+ + x_i^-$ in the objective function as shown below, along with $x_i^+, x_i^- \geq 0$.

K has m p-simplices and n $(p+1)$-simplices

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} w_i (x_i^+ + x_i^-) \\ \text{s.t.} \quad & \bar{x}^+ - \bar{x}^- = \bar{c} + [\partial_{p+1}]\bar{y} \\ & \bar{x}^+, \bar{x}^- \in \mathbb{Z}^m_{\geq 0}, \ \bar{y} \in \mathbb{Z}^n \end{aligned}$$

linear optimization problem with integrality constraints.

nonnegative

Intuition:

$|-3| = 3$. We write

$-3 = 2-5$, or

$-3 = 2024-2027$, or

$-3 = 0-3$.

In general, we want $-3 = x^+ - x^-$ (x^+: positive part; x^-: negative part)

$x^+, x^- \geq 0$,

and choose x^+, x^- such that $x^+ + x^-$ is minimum.

Similarly, $5 = 5-0$ (rather than $11-6$, for instance).

We can show that only one of x^+ and x^- will be >0 in each such representation.

$\Rightarrow$ The OHCP can be modeled as an integer program (IP) (an integer linear program, or ILP, to be exact).

But IPs are hard to solve in general (NP-hard), while LPs can be solved in polynomial time.

$$\min\{\bar{c}^T\bar{x} \mid A\bar{x}=\bar{b}, \bar{x}\geq\bar{0}, \bar{x}\in\mathbb{Z}\} \quad \text{——— (IP)}$$

$$\min\{\bar{c}^T\bar{x} \mid A\bar{x}=\bar{b}, \bar{x}\geq\bar{0}\} \quad \text{——— (LP)}$$

drop the integrality restriction to get the (LP) relaxation

The OHCP IP can be written in the above standard form of (IP).

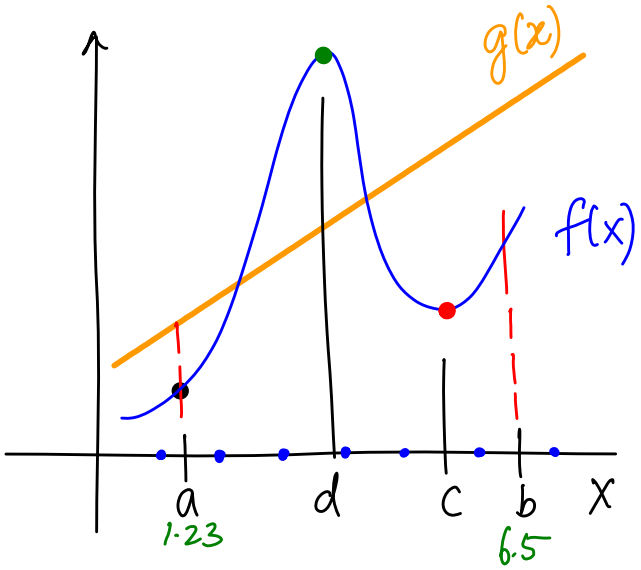
Let's review LP and IP from a 10,000-ft level...

Optimization for dummies

In calculus, we learned how to min $f(x)$ over $x \in [a,b]$.

We need:

$f'(x)=0$, and $f''(x)>0$ for x to be a minimum



$x=c$ is a (local) minimum,

but $x=a$ is the minimum of $f(x)$ over $x \in [a,b]$.

If $g(x)$ is linear, we need to check only the end points $x=a$ and $x=b$ to find the minimum.

Linear programming (LP) extends this linear case (of $g(x)$) to higher dimensions. The feasible set is the intersection of half spaces, and we need to still look only at vertices (or corner points). We can solve LPs "efficiently", i.e., in polynomial time.

Integer (linear) programming (IP) aims to find solutions to (LP) that have integer values. (IP) cannot be solved in polynomial time in the worst case.

Recall OHCP LP/IP

$$\min \sum_{i=1}^{m} w_i (x_i^+ + x_i^-)$$

$$\text{s.t.} \quad \bar{x}^+ - \bar{x}^- = \bar{c} + [\partial_{pn}]\bar{y}$$

$$\bar{x}^+, \bar{x}^- \in \mathbb{Z}^m_{\geq 0}, \quad \bar{y} \in \mathbb{Z}^n$$

} (IP)

— (LP) with $\bar{x}^+, \bar{x}^- \geq \bar{0}$

Let's look at a 2D example that highlights the differences between (IP) and (LP).

A 2D LP/IP Example

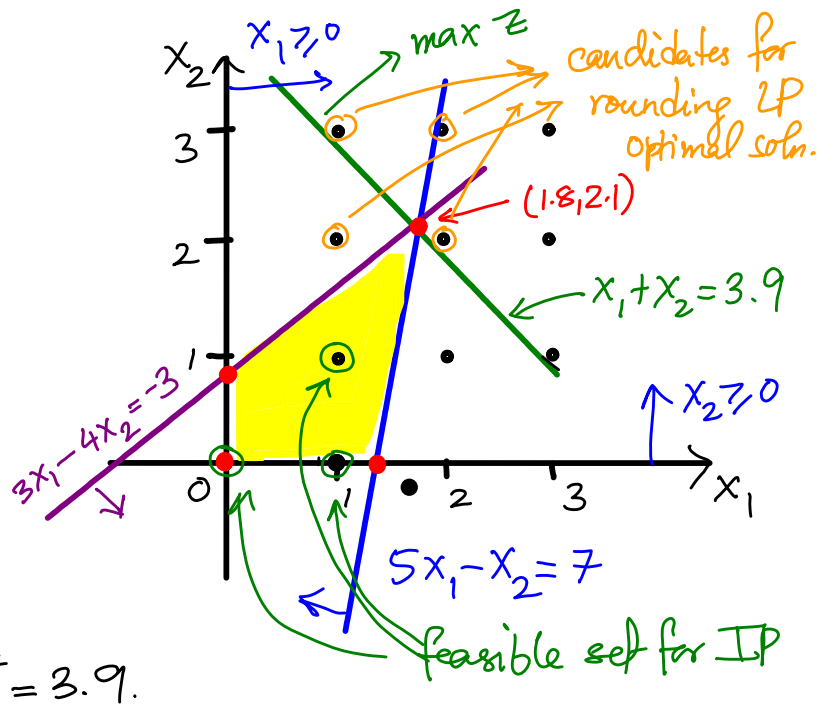
$$\max \quad z = x_1 + x_2$$

$$\text{s.t.} \quad 5x_1 - x_2 \leq 7$$

$$3x_1 - 4x_2 \geq -3$$

$$x_1, x_2 \geq 0$$

$$x_1, x_2 \in \mathbb{Z}$$



We slide the z-line in the direction of increase of z, while staying feasible.

Optimal solution is at $(1.8, 2.9)$, $z^* = 3.9$.

For solving the IP, rounding $(1.8, 2.9)$ will not work, as none of the options are even feasible here! Hence ignoring the integrality restriction could be quite bad.

But- there are some special cases where we solve just the LP, and get integrality "for free"! One such special case involves the total unimodularity of the constraint matrix, and we will explore this concept.

Q. Why not solve OHCP over $\mathbb{Z}_2$?

Result: Chen and Freedman (2010)

OHCP over $\mathbb{Z}_2$ is NP-hard (even to approximate).

We could simulate addition over $\mathbb{Z}_2$ in the OHCP LP by setting $\bar{x}^+ - \bar{x}^- = \bar{c} + [\partial_{PH}]\bar{y} + 2\bar{u}$, with $\bar{u} \in \mathbb{Z}^m$.
But the $2\bar{u}$ term destroys "nice structure"!

OHCP LP in standard form:

$$\left.\begin{aligned} \min \sum_{i=1}^{m} w_i (x_i^+ + x_i^-) \\ \text{s.t.} \quad \bar{x}^+ - \bar{x}^- = \bar{c} + [\partial_{PH}](\bar{y}^+ - \bar{y}^-) \\ \bar{x}^+, \bar{x}^- \geq \bar{0}, \ \bar{y}^+, \bar{y}^- \geq \bar{0} \end{aligned}\right\} \text{(LP)} \quad \text{(IP)}$$
$$\bar{x}^+, \bar{x}^- \in \mathbb{Z}^m, \ \bar{y}^+, \bar{y}^- \in \mathbb{Z}^n$$

Writing it in standard form, we get

$$\min \ \underbrace{[\bar{w}^T \ \bar{w}^T \ \bar{0} \ \bar{0}]}_{\bar{c}^T} \bar{x}$$

$$\text{s.t.} \ \underbrace{[I \ -I \ -B \ B]}_{A} \begin{bmatrix} \bar{x}^+ \\ \bar{x}^- \\ \bar{y}^+ \\ \bar{y}^- \end{bmatrix}}_{\bar{x}} = \underbrace{\bar{c}}_{\bar{b}}$$

$$\bar{x} \geq \bar{0}$$

$$\bar{x} \in \mathbb{Z}^{2m+2n}$$

standard form LP:
$\min \bar{c}^T \bar{x}$
s.t. $A\bar{x} = \bar{b}$
$\bar{x} \geq \bar{0}$

There is a well-studied special case where IP can be solved as an LP.

Total Unimodularity (TU)

Result Let $A \in \mathbb{Z}^{m \times n}$, $\bar{b} \in \mathbb{Z}^m$

$\min \{\bar{c}^T \bar{x} \mid A\bar{x} = \bar{b}, \bar{x} \geq \bar{0}, \bar{x} \in \mathbb{Z}^n\}$ ——— (IP)

and $\min \{\bar{c}^T \bar{x} \mid A\bar{x} = \bar{b}, \bar{x} \geq \bar{0}\}$. ——— (LP)

IP can **always** ($\forall \bar{c} \in \mathbb{R}^n$) be solved in polynomial time by solving (LP) iff A is totally unimodular (TU).

Equivalently, the vertices of (LP) have integer coordinates.

We consider whether we could use this result for the OHCP. We ask

When is $A = [I \; -I \; -B \; B]$ TU?

Def A matrix $B \in \mathbb{Z}^{m \times n}$ is **totally unimodular** (TU) if every subdeterminant of B is in $\{-1, 0, 1\}$.

B is **unimodular** if every nonsingular $m \times m$ submatrix of B has determinant ± 1.

→ In particular, $B_{ij} \in \{0, 1, -1\}$.

MATH 529 – Lecture 27 (04/16/2024)

Today: * total unimodularity (TU)
* OHCP and TU

Recall $B \in \mathbb{Z}^{m \times n}$ is TU if every subdeterminant is $-1, 0$, or 1.

Examples

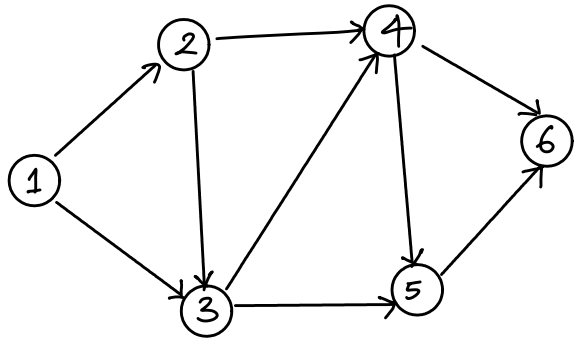
$B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is not TU, as $\det B = 2$. But smaller subdeterminants are $0, \pm 1$ in both cases

$B' = \begin{bmatrix} 1 & -1 & \\ & 1 & \\ & 1 & 1 \end{bmatrix}$ is TU (note that $\det B' = 0$).

We will study these types of matrices in detail soon!

The node-arc incidence matrix of a directed network is TU:

$$B = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccccccc} (1,2) & (1,3) & (2,3) & (2,4) & (3,4) & (3,5) & (4,5) & (4,6) & (5,6) \end{array} \\ \begin{bmatrix} -1 & -1 & & & & & & & \\ 1 & & -1 & -1 & & & & & \\ & 1 & 1 & & -1 & -1 & & & \\ & & & 1 & 1 & & -1 & -1 & \\ & & & & & 1 & 1 & & -1 \\ & & & & & & & 1 & 1 \end{bmatrix} \end{array}$$



row ≡ node, column ≡ arc

Many network flow problems including min-cost flow, max flow, shortest path, etc, are easy because of this network matrix property. At the same time, they are easier than general LP — there are efficient algorithms to solve them that exploit the network structure.

We consider whether we could use the TU result for OHCP.

When is $A = [I \; -I \; -B \; B]$ TU? → A in the OHCP LP written as $A\bar{x} = \bar{b}, \bar{x} \geq 0$.

Theorem 1 A is TU iff B is TU.

Proof There are several elementary operations that preserve TU.

* taking transpose
* multiply a column/row by -1
* add copy of a row/column
* swap two rows (or two columns)
* add a new singleton row/column with the single non-zero entry being ± 1.
* ...

We could prove these results using arguments that show preservation of determinant (absolute) values under each operation.

We get the constraint matrix A from B using a series of these TU-preserving operations:

$$A = [I \; -I \; -B \; B]$$

* duplicate columns of B
* scale columns of (one copy of) B by -1
* add $2m$ columns of unit vectors

□

Q. When is $B = [\partial_{ph}]$ TU?

This is the big question now. If B is TU, then we could solve all OHCP instances easily on that K.

Before that, let's revisit OHCP over $\mathbb{Z}_2$.

We could implement homology over $\mathbb{Z}_2$ by modifying the constraints of the OHCP IP as follows:

$$\bar{x}^+ - \bar{x}^- = \bar{c} + B(\bar{y}^+ - \bar{y}^-) + 2\bar{u}, \quad u \in \mathbb{Z}^m.$$

Here, $A = [I \; -I \; -B \; B \; -2I]$, and A is not TU even when B is, because of the $2I$ term.

But we could "simulate" working over $\mathbb{Z}_2$ differently. We could add the constraints $\bar{x}^+, \bar{x}^- \leq \bar{1}_m$, the vector of m ones, and also $\bar{y}^+, \bar{y}^- \leq \bar{1}_n$, the vector of n ones. The modified OHCP LP constraints now become the following.

$$\left.\begin{array}{rl} x^+ - \bar{x}^- - B\bar{y}^+ + B\bar{y}^- = \bar{c} & \\ \bar{x}^+ \leq \bar{1} & \\ \bar{x}^- \leq \bar{1} & \\ \bar{y}^+ \leq \bar{1} & \\ \bar{y}^- \leq \bar{1} & \end{array}\right\} \Rightarrow A' = \begin{bmatrix} I & -I & -B & B \\ I_m & & & \\ & I_m & & \\ & & I_n & \\ & & & I_n \end{bmatrix}.$$

Using arguments similar to ones used in Theorem 1, we get that A' is TU iff B is TU. In the optimal solution, we are now guaranteed to get $x_i, y_j \in \{\pm 1, 0\}$.

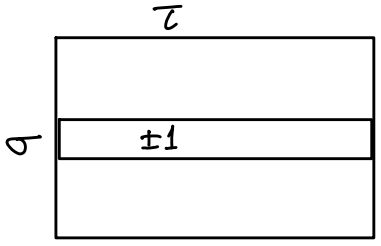
We now present the first result characterizing when B is TU.

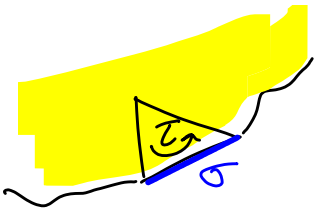
Theorem 2 Let K be a finite simplicial complex triangulating a compact orientable $(p+1)$-manifold. Then $[\partial_{p+1}]$ is TU.

($(p+1)$-manifold: with or without boundary)

Proof idea

Case 1: σ is a boundary p-simplex



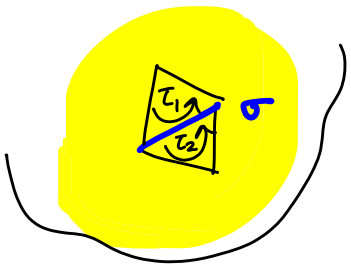


τ : $(p+1)$-simplex
σ : p-simplex, $\sigma \leq \tau$

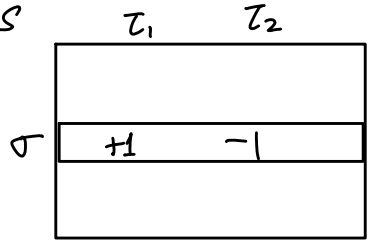
Row in B corresponding to σ has exactly one nonzero, which is a ± 1 (at column corresponding to τ).

Case 2: σ is a "manifold" p-simplex.

Assume K is consistently oriented.



Here, the row corresponding to σ has exactly two nonzeros, (at columns corresponding to τ_1 and τ_2, and these entries are a $+1$ and a -1.



$\sigma \leq \tau_1$ and $\sigma \leq \tau_2$ here.

To obtain a consistent orientation, we might have to scale some columns (τ_j) by -1, but those operations preserve TU.

So, every row of $[\partial_{p+1}]$ has 1 or 2 non-zeros. If it has two nonzeros, they are $+1$ and -1.

Now, consider any $r \times r$ submatrix S of $[\partial_{p+1}]$.

Rows of S:
* could be all zero
* could have a single ± 1
* have one $+1$ and one -1.

If S has a zero row, then $\det(S) = 0$. If there is a singleton row, we could expand along that row, and look at an $(r-1)\times(r-1)$ subdeterminant instead. In the nontrivial case, every row has one $+1$, one -1.

So assume every row of S has two nonzeros: $+1, -1$.

$\Rightarrow S\mathbb{1} = \bar{0}$ (adding all columns gives zero vector!)

$\Rightarrow \det S = 0$ (columns are linearly dependent)

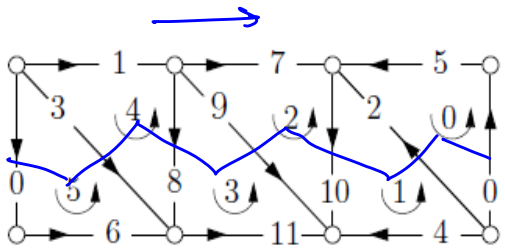
$\Rightarrow B$ is TU.

If K is not consistently oriented to start with, then we multiply a subset of columns of B by -1 to orient K consistently. These scaling operations preserve TU.

$\Rightarrow [\partial_{p+1}(K)]$ is TU for orientable manifold K (with or without boundary). □

What about $[\partial_{p+1}]$ of arbitrary simplicial complexes, that are not necessarily orientable manifolds? We consider perhaps the quintessential nonorientable manifold first – the Möbius strip.

Illustration on Möbius strip



Notice that we have a minimal Möbius strip here — if we remove one triangle, we get a disc, and the Möbius strip disappears.

Hence, to possibly find an obstruction to TU, we look at a submatrix that uses the entire Möbius strip, i.e., all triangles, and hence all columns.

$$S = \begin{array}{c} \begin{matrix} 5 & 4 & 3 & 2 & 1 & 0 \end{matrix} \\ \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \end{array} \begin{matrix} 0 \\ 3 \\ 8 \\ 9 \\ 10 \\ 2 \end{matrix}$$

Möbius cycle matrix (MCM)

$\det S = -2$

$[\partial_2]$ for Möbius strip :

	0:	1:	2:	3:	4:	5:
→ 0:	1	0	0	0	0	1
1:	0	0	0	0	−1	0
→ 2:	−1	1	0	0	0	0
→ 3:	0	0	0	0	1	−1
4:	0	−1	0	0	0	0
5:	1	0	0	0	0	0
6:	0	0	0	0	0	1
7:	0	0	−1	0	0	0
→ 8:	0	0	0	1	−1	0
→ 9:	0	0	1	−1	0	0
→ 10:	0	1	−1	0	0	0
11:	0	0	0	1	0	0

Similarly, the boundary edges, i.e., the edges that are faces of only one triangle each, cannot contribute in a nontrivial manner to any determinants. So, we take all the "manifold" edges shared by the 6 triangles to consider the 6×6 submatrix using rows 0, 2, 3, 8, 9, 10; and all the columns 0–5. Indeed, this submatrix has determinant −2. Furthermore, if we rearrange the rows and columns in the order in which we see the edges and triangles from left to right, we see a canonical matrix, which we call **Möbius cycle matrix** (MCM).

$$MCM(n) = \begin{bmatrix} 1 & & & & \alpha \\ 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \end{bmatrix}, \quad \alpha = (-1)^{n+1} \qquad \det(MCM(n)) = 2.$$

We can prove that in 2D (i.e., for the edge-triangle case), the only obstructions to $[\partial_2]$ being TU are these Möbius strips. The corresponding subcomplex is called a Möbius subcomplex.

We can show that the only way in which a 2-complex has a non-TU boundary matrix is by it having a Möbius subcomplex.

Theorem 3 $[\partial_2(K)]$ is TU iff K has no Möbius subcomplex.

Proof uses result on minimal violators of TU of matrices. These are matrices that are not TU, but every proper submatrix is TU. These minimal violators belong to two classes. One is the MCMs. In the second class, every row/column has at least 4 non-zeros. For a 2-complex, every column has exactly 3 nonzeros, and hence cannot have a submatrix of the second class. □

MATH 529 – Lecture 28 (04/18/2024)

Today:
* OHCP and TU
* torsion
* other optimal homology problems

Recall: **Theorem 3** $[\partial_2(K)]$ is TU iff K has no Möbius strips.

$$MCM(n) = \begin{bmatrix} 1 & & & \alpha \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix}, \quad \alpha = (-1)^{n+1}$$

A similar (sub)matrix could be defined for a cylinder subcomplex, called cylinder cycle matrix (CCM).

$$CCM(n) = \begin{bmatrix} 1 & & & \alpha \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix} \qquad \alpha = (-1)^{n}.$$

$$\det(CCM(n)) = 0.$$

Notice that the corresponding cylinder subcomplex is orientable, while the Möbius subcomplex is not.

Recall examples B, B' from Lecture 27 when we introduced TU: $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ and $B' = \begin{bmatrix} 1 & & -1 \\ 1 & 1 & \\ & 1 & 1 \end{bmatrix}$.

Note that $B = MCM(3)$, $B' = CCM(3)$

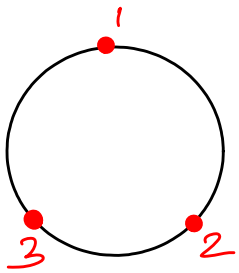
Thus, Möbius strips are canonical shapes that correspond to non-total unimodularity. How about the result in general?

Theorem 4 $[\partial_{p+1}(K)]$ is TU iff K has no relative torsion in $(p, p+1)$-dimensions.

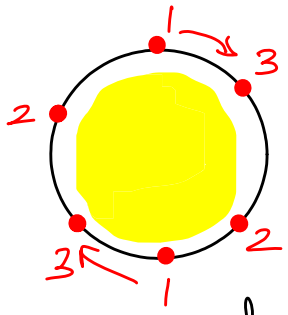
↪ torsion in relative homology groups

Torsion

Here is how Poincaré visualized torsion, a concept he originally introduced in late 1890s. Consider gluing the boundary of a disc to a circle. If we do it in the simple manner, we get a hat. But if we do the gluing such that the boundary of the disc goes twice around the circle, we get the 2-fold dunce hat, which is the projective plane $\mathbb{RP}^2$. If we go k-times around, we get the k-fold dunce hat. And k-fold dunce hats have torsion in its H_1 (first homology group). → $H_1 \cong \mathbb{Z}_k$ here.



circle

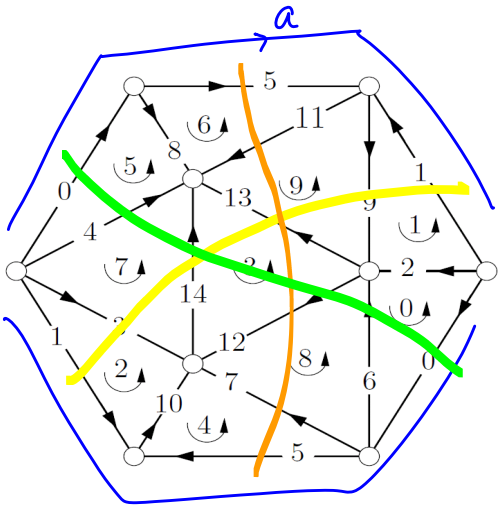


disc

2-fold Dunce hat $\approx \mathbb{RP}^2$

k-fold dunce hat in general, for $k \geq 2$.

Here is a triangulation of the 2-fold dunce hat ($\mathbb{RP}^2$). Notice the boundary edges going 0-5-(-1) twice around. Indeed, we could spot at least 3 Möbius strips, as indicated here. We could spot other Möbius strips that wander left and right more...



Theorem 4 gives an algebraic topological characterization of TU. We saw the algebraic definition (all subdeterminants are $0, \pm 1$) as well as a geometric characterization (all vertices of LP polytope have integer coordinates) previously.

How useful is this characterization? Could we check this condition efficiently?

YES. Can check if $[\partial_{p+1}]$ is TU or not in polynomial time.

Seymour's matroid decomposition (1980) $\left(O(n^3)\right)$.

Truemper – practical algo (2012) $\left(O(n^5)\right)$.

→ check out https://discopt.github.io/cmr/ .

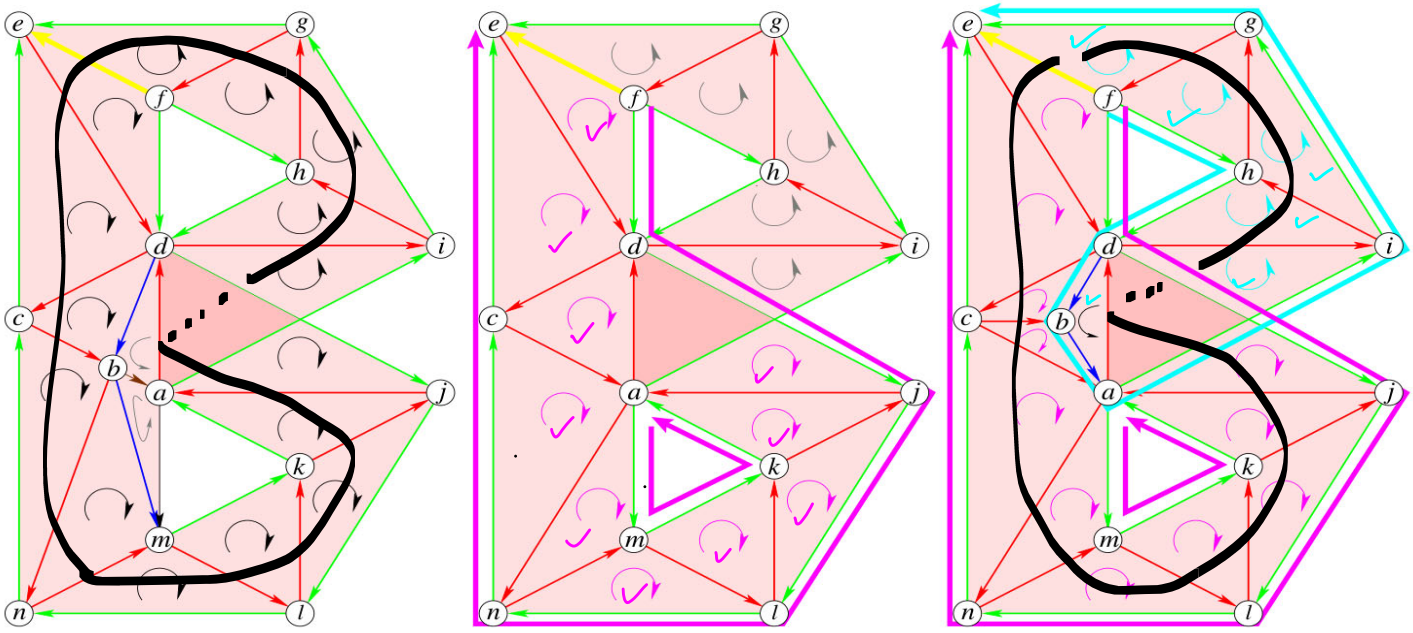
Here is a summary of the special cases of OTEP that are easy.

1. $[\partial_{p+1}]$ is TU if K is an orientable $(p+1)$-manifold.
2. $[\partial_2]$ is TU if K has "no Möbius strip".
3. $[\partial_d]$ is TU for K in $\mathbb{R}^d$. → tetrahedra-triangle case in $\mathbb{R}^3$ is easy!

Result 3 is similar to Result 1: we cannot realize/embed a Möbius strip in $\mathbb{R}^2$ for instance. But the proof presented used different techniques.

More recently, people have studied optimal homology problems over a filtration (rather than on a single simplicial complex).

If K is not TU, is all hope lost? No!
Non-Total Unimodularity Neutralized (NTUN) complexes!



not TU TU not TU

There are Möbius strips (as subcomplexes) in both the first and third complexes, while the middle (second) complex has a TU $[\partial_2]$ matrix. But we are guaranteed to find an integer solution for every OTEP instance of K_3, as it is NTU-neutralized.

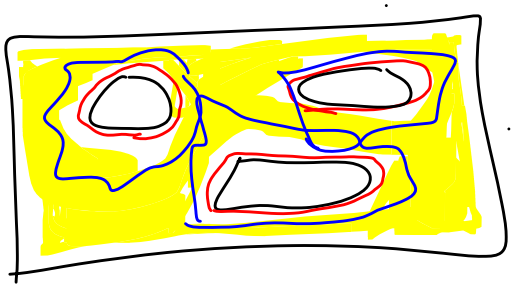
See the paper for details, at http://arxiv.org/abs/1304.4985

Being NTU neutralized is a property of K, and not of |K|, as the above example illustrates.

There are several interesting open problems in this context. For instance, **all** instances of OTEP for 1D inputs in a convex 3D domain, e.g., 2-skeleton of a solid unit cube in computations were solved by their corresponding LPs – there ought to be some deeper theory that's waiting to be unearthed!

Related optimal homology problems

The Optimal Homology Basis Problem (OBAS)



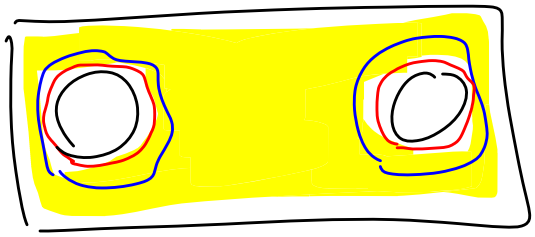
input

OBAS solution.

Given a loop around each hole, find a set of tightest loops around each hole, or a minimal basis in general.

Q: Could we solve OBAS as an OHCP?

What complexes exist for which the OHCP solution with the collection of cycles in a homology basis as input is actually an optimal homology basis?



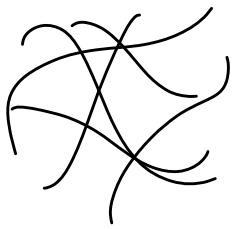
May be we could characterize the case where the holes are "far apart" from each other, and the loops around each hole are also disjoint. In this case, solving the OHCP with the collection of input cycles as the single input cycle should solve the OBAS problem.

MATH 529 – Lecture 29 (04/23/2024)

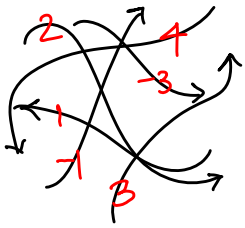
Today: * GMT : currents and flat norm
* flat norm as an OHCP

Connection to Geometric Measure Theory (GMT)

Currents: generalized surfaces used in GMT



A "rectifiable" 1-set



1-current

Orient each piece, and assign an integer multiplicity

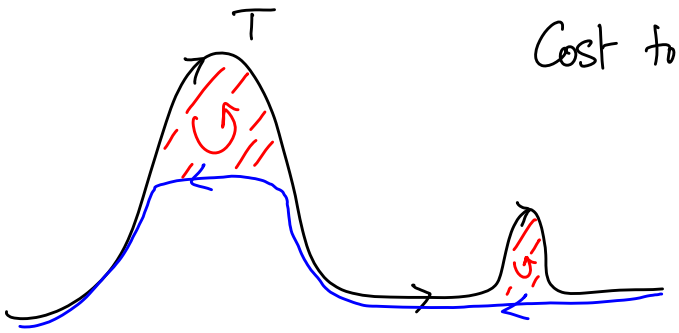
In simplicial setting, we study 1-chains (over $\mathbb{Z}$).

Flat norm of d-current T with scale ($\lambda \geq 0$)

$$\mathbb{F}(T) = \min_S \left\{ M_d(T - \partial S) + \lambda M_{d+1}(S) \right\}$$

T: d-current
S: (d+1)-current

M_d: Mass (d-dim volume ≡ length, area, volume etc.)



Cost to erase T using two operations:

Operation 1: draw 1-curves of opposite orientation
cost: length of curve

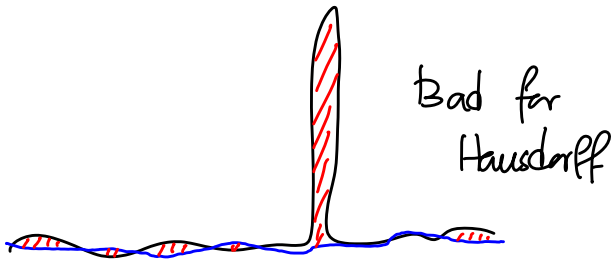
Operation 2: trace boundaries of 2D patches
cost: area of patch

Total cost = length + area

Flat norm (T) = min total cost over all such decompositions.

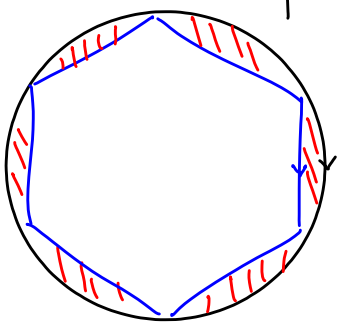
Distance between two currents P and Q: $\text{dist}(P,Q) = \mathbb{F}(P-Q)$.

The flat norm distance is the "natural" distance between spaces modeled as currents. Hausdorff distance is sensitive to outliers. Consider the two curves (one blue and another black), which are close to each other except for a small narrow spike in the black curve. Hausdorff distance between the curves will be offset by this spike. But the flat norm distance will be small – the area of the spike in essence.



Another option is the Frechét distance, or "dog leash" distance. Imagine you're walking on the blue curve and your dog on the black curve. Both could stop or move at different speeds, but cannot go backwards. The Frechét distance is the length of the smallest (non-stretching) leash that'll work. But this distance is not defined for more general inputs, e.g., the spaghetti goop!

What about the mass difference $M(P-Q)$? Even this distance might not make sense in all cases, though. Consider the following example.



T: unit circle (oriented CW) ↻ clockwise

T_n: regular n-gon inscribed in T

Notice that $M(T-T_n) \to 4\pi$ as $n \to \infty$.

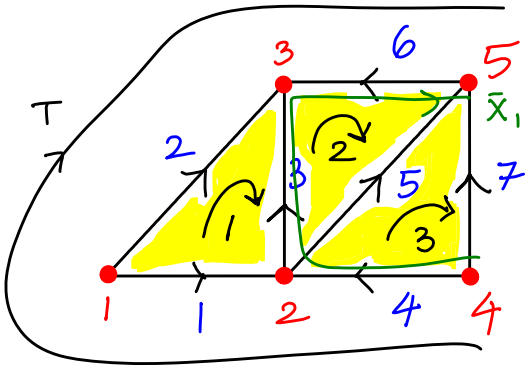
But $\mathbb{F}(T-T_n) \to 0$, as the flat norm distance between T and T_n is measured by the area between T and T_n, which goes to zero as $n \to \infty$.

Simplicial Flat Norm

We can represent currents as chains on simplicial complexes, and can compute the simplicial flat norm as an optimal homology problem.

Example

$$\bar{t} = \begin{matrix}1\\2\\3\\4\\5\\6\\7\end{matrix}\begin{bmatrix}-1\\-1\\0\\-1\\0\\-1\\0\end{bmatrix}$$



Note that
length$(T) = 3+\sqrt{2}$
$= 4.41$

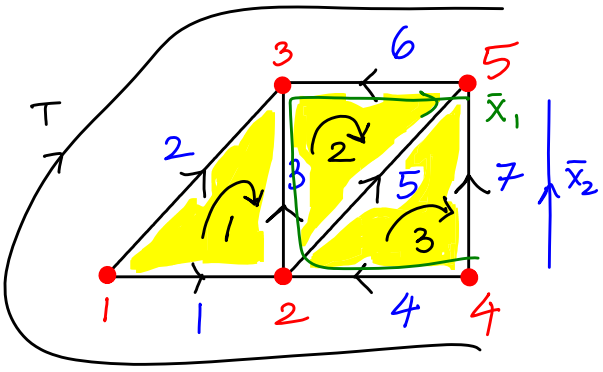
$$w_i = \begin{cases} 1, & i = 1,3,4,6,7 \\ \sqrt{2}, & i = 2,5 \end{cases}$$

Let $v_j = \frac{1}{2} \forall j$. (area of each triangle)

For $\lambda = 2.5$ we get

$$F_\lambda(\bar{t}) = 3 + \left(\frac{5}{2}\right)\left(\frac{1}{2}\right) = 4.25$$

$<$ length $(\bar{t}) = 4.41$



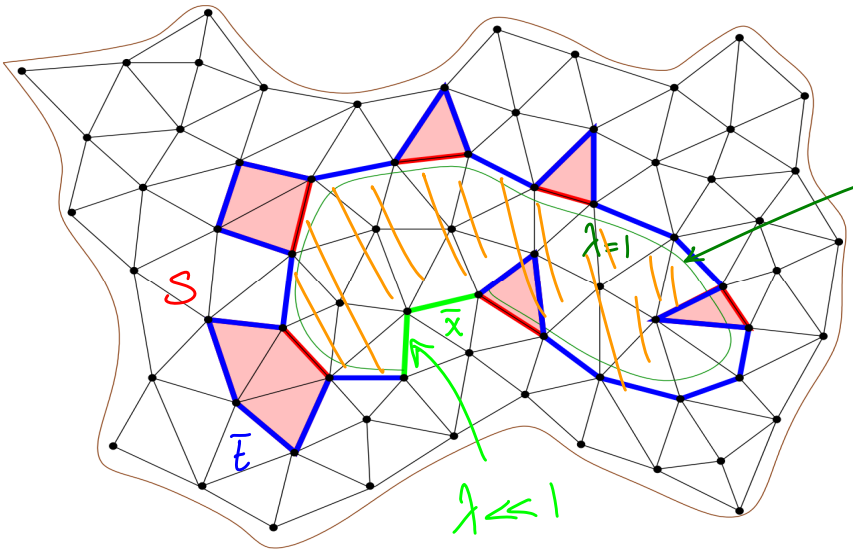
For $\lambda = 1$, we get $\bar{x}_2$ as the result;

$$F_\lambda(\bar{t}) = 1 + (1)^{\lambda}\left(3\left(\frac{1}{2}\right)\right) = 2.5$$

area of 3 triangles

$\bar{x}_2$ remains the flat norm decomposition for any $\lambda \leq 1$.

λ is a scale parameter. At smaller values of λ, we smooth out larger scale features.



flat norm decomposition for $\lambda = 1$ (using S)

For $\lambda \ll 1$, we use almost all the 2D space "enclosed" by $\bar{t}$.

Simplicial Flat Norm as an LP

We could write down an integer program for the simplicial flat norm problem that is quite similar to the OTEP IP. We start with the optimization model with absolute value terms.

$$\min \sum_{i=1}^{m} w_i |x_i| + \lambda \left(\sum_{j=1}^{n} v_j |s_j| \right)$$

piecewise linear

$$\bar{x} = \bar{t} - [\partial_{p+1}]\bar{s}$$

$$\bar{x} \in \mathbb{Z}^m, \quad \bar{s} \in \mathbb{Z}^n$$

Using the standard transformation to handle absolute values, we get the following IP:

$$\min \sum_{i=1}^{m} w_i (x_i^+ + x_i^-) + \lambda \sum_{j=1}^{n} v_j (s_j^+ + s_j^-)$$

recall that we had used $\bar{y}^+, \bar{y}^-$ in place of $\bar{s}^+, \bar{s}^-$ previously

$$\text{s.t.} \quad \bar{x}^+ - \bar{x}^- = \bar{t} - [\partial_{p+1}](\bar{s}^+ - \bar{s}^-)$$

$$\bar{x}^+, \bar{x}^- \geq \bar{0}, \quad \bar{s}^+, \bar{s}^- \geq \bar{0}$$

$$\bar{x}^+, \bar{x}^- \in \mathbb{Z}^m, \quad \bar{s}^+, \bar{s}^- \in \mathbb{Z}^n$$

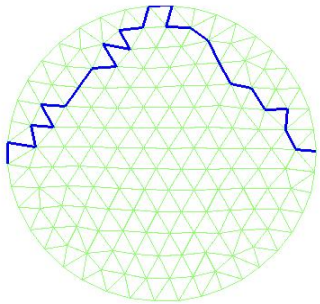
→ relax to get MSFN LP

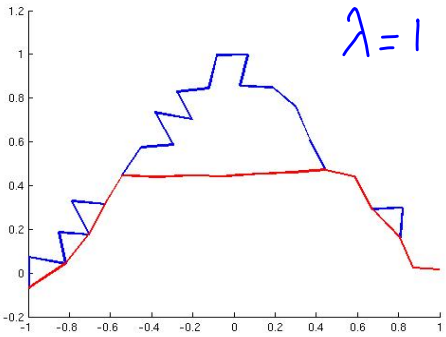
This is the Multiscale simplicial flat norm (MSFN) LP. Note that we can use results on TU of $[\partial_{p+1}]$ here, just as we did for the OHCP.

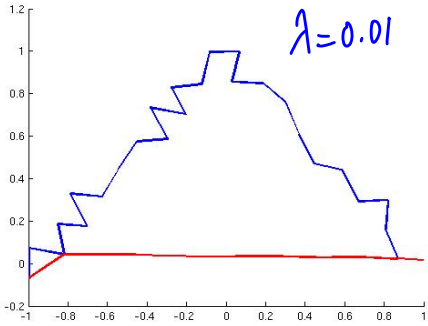
When $\lambda = 0$, the MSFN LP $\equiv$ OHCP LP.

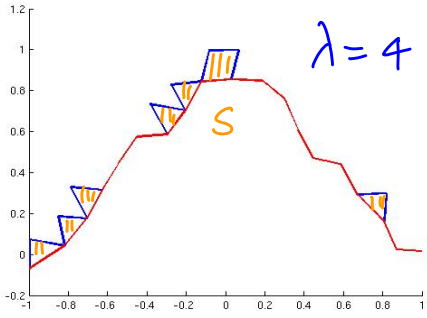
When $\lambda \to \infty$, the MSFN LP solves the area minimizing surface problem for a cycle $\bar{E}$.

Here is an illustration in 2D (for 1-chain $\bar{E}$). While the definition allows one to consider all of $\mathbb{R}^2$ as candidates for S, we could restrict to the "convex hull" of $\bar{E}$.









At larger values of $\lambda = 4$, we "smooth out" only the small scale bumps in $\bar{E}$. At $\lambda = 0.01$, we flatten out the entire input curve.

It appears the S chains seem to only grow as λ decreases.

Open problem: When can you guarantee that as $\lambda \to 0$, the $(d+1)$-chains S (defining $\bar{E} \sim \bar{x}$ homology) form a filtration?

Could we devise an "incremental algorithm" for computing the flat norm?

MATH 529 – Lecture 30 (04/25/2024)

Today: * integral decompositions of currents
* median shapes

How do we reconcile continuous currents with simplicial ones?

There are results on how well the simplicial chains approximate the (continuous) currents – **simplicial deformation theorem.**

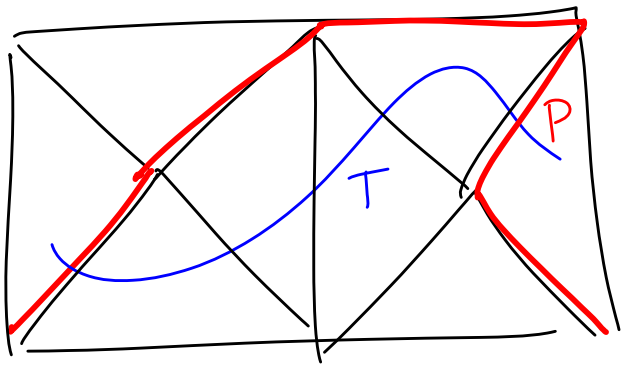
Any current T can be deformed to a simplicial chain P in a simplicial complex K such that

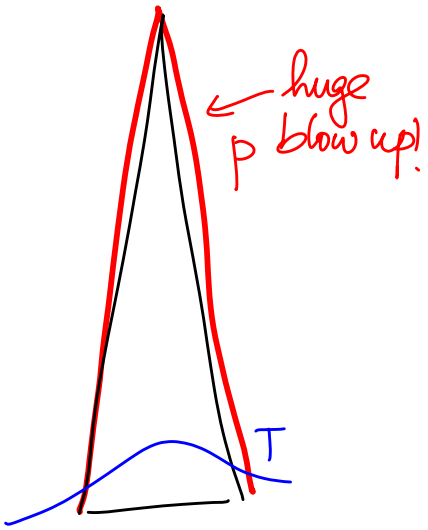
$$M(P) \leq C_1 M(T) + \Delta C_2 M(\partial T)$$

$$M(\partial P) \leq C_2 M(\partial T), \quad \text{and}$$

$$\mathbf{F}_\lambda(T, P) \leq \Delta C_1 \left[M(T) + C_3 M(\partial T) \right]$$

$C_1, C_2, C_3 > 1$ constants, and $\Delta \to 0$ as we take finer and finer simplicial complexes.





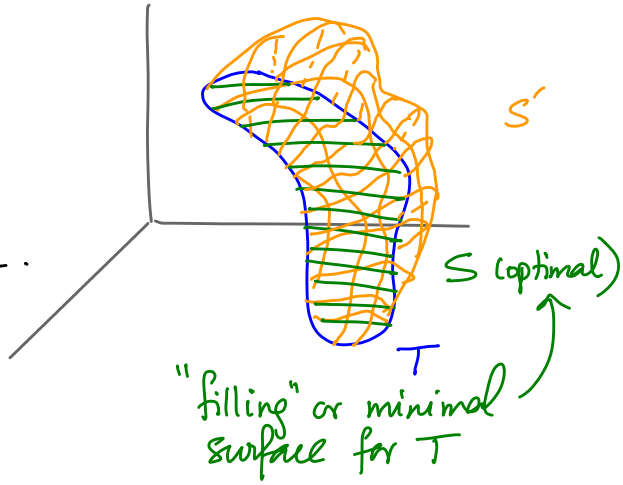
In the trivial case, T could be already simplicial, and we don't have to do any pushing at all. But in the general setting, the regularity of the complex will affect our deformation. Hence C_1, C_2, C_3 are > 1, and could be large, in fact.

IU connection to Prove Fundamental Results

Apart from efficient computability of optimal homologous chains, the total unimodularity result could also be used to prove (more) fundamental results.

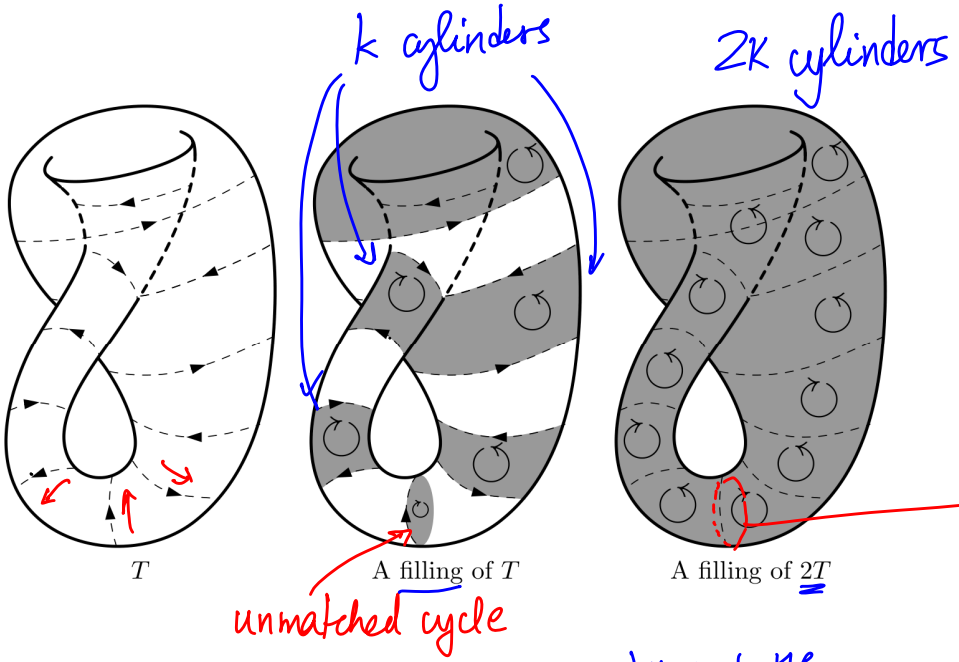
Here is a motivating problem:

Q. When is the minimal surface for $2\bar{t}$ the same as 2(minimal surface for $\bar{t}$)?



The answer is YES in codimension 1 and 2.

But No in codimension 3!
Here is a counterexample.



R. Young (2013)

Klein bottle

T: $2k+1$ cycles around $\mathbb{K}^2$, with adjacent cycles oriented oppositely.

Area of min surface for $2T$ < 2(area of min surface for T).

A similar question on flat norm decomposition:

$$\mathbb{F}_\lambda(T) = \min_S \{ M_d(\underbrace{T - \partial S}_{X}) + \lambda M_{d+1}(S) \}$$

Q: For integral T (with integer multiplicities), when is its flat norm decomposition (X, S) also integral?

The answer is YES in the finite simplicial setting in codimension 1, as $[\partial_{d+1}]$ is TU. But the answer is not known in general.

F. Almgren considered this related problem:

"If a sequence of currents $\{2T_i\}$ converge in the integral flat topology, must the sequence $\{T_i\}$ also converge?"

B. White: "He used a lot of weapons, including his enormous $(m-2)$-regularity paper..."

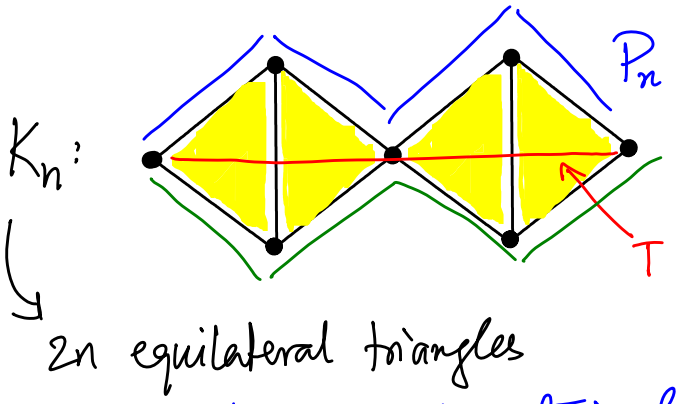
→ single paper that is 1,500+ pages long!

Could we use the TU result, and then somehow take finer and finer simplicial complexes and somehow the result holds in the limit for the continuous case?

Challenges The simplicial flat norm may not converge to the (continuous) flat norm.

$$\mathbb{F}_{K_n}(P_n) = \frac{2}{\sqrt{3}}\,\mathbb{F}(P_n)$$

→ gives T as the optimal decomposition



$2n$ equilateral triangles

The simplicial flat norm is always a multiple of flat norm even as $n \to \infty$!

We could use all the triangles but end up with another chain of the same length as the "upper roof" input! While the continuous flat norm decomposition is T, the flat line.

We cannot apply simplicial deformation directly!

we want mass expansion of T and ∂T to $\to 0$:

$$M(P) \leq M(T) + \rho(*)$$
$$M(\partial P) \leq M(\partial T) + \rho(*)$$

and $\rho \to 0$ as we take finer and finer simplicial complexes.

Compare these statements to the bounds specified by the simplicial deformation theorem.

See the paper for details: http://arxiv.org/abs/1411.0882

Median Shapes

Notions of average of $x_1, \dots, x_k \in \mathbb{R}$?

mean: $\hat{x} = \underset{x}{\operatorname{argmin}} \sum_{i=1}^{k} (x - x_i)^2$

median: $\hat{x} = \underset{x}{\operatorname{argmin}} \sum_{i=1}^{k} |x - x_i|$ → less sensitive to outliers

Generalizing to vectors: geometric median of $\bar{x}_1, \dots, \bar{x}_k \in \mathbb{R}^d$:

$$\hat{\bar{x}} = \underset{\bar{x}}{\operatorname{argmin}} \sum_{i=1}^{k} \|\bar{x} - \bar{x}_i\|$$ → we get mean if minimizing $\sum \|\bar{x} - \bar{x}_i\|^2$

We defined the flat norm distance between currents. It's natural to use that distance to define an average shape of input shapes. We concentrate on the median shape.

Median of shapes represented as currents $T_1, \dots, T_k$?

$$\hat{T} = \underset{T}{\operatorname{argmin}} \sum_{i=1}^{k} \mathbb{F}_\lambda(T, T_i)$$ → we get the mean shape if we minimize $\sum_{i=1}^{k} \mathbb{F}_\lambda(T, T_i)^2$

We can write down the median shape computation as an optimization problem similar to the OTCP and flat norm in the finite simplicial setting. We specify the flat norm decomposition of each T_i w.r.t. T as a homology equation.

$$\bar{t} - \bar{t}_h = \bar{r}_h + [\partial_{p+1}]\bar{s}_h, \quad h = 1, \dots, k.$$

→ flat norm decomposition of each t_h

(we'll need i, j as other indices soon. Hence the switch to h ☺!)

Consider the simplicial complex with m and n p- and $(p+1)$-simplices, respectively. We use the standard approach to handle absolute value terms in the minimization objective function.

$$\min \sum_{h=1}^{k} \left(\sum_{i=1}^{m} w_i |r_{hi}| + \lambda \sum_{j=1}^{n} v_j |s_{hj}| \right)$$

$$\text{s.t.} \quad \bar{t} - \bar{t}_h = \bar{r}_h + B\bar{s}_h, \quad h = 1, \ldots, k$$

$$\bar{t} \in \mathbb{Z}^m, \ \bar{r}_h \in \mathbb{Z}^m, \ s_h \in \mathbb{Z}^n$$

In a modified version, we consider median shape with mass regularization — $|t_i| \to t_i^+ + t_i^-$ is used in the objective function in this case.

$$\min \sum_{h=1}^{k} \left(\sum_{i=1}^{m} w_i (r_{hi}^+ + r_{hi}^-) + \lambda \sum_{j=1}^{n} v_j (s_{hj}^+ + s_{hj}^-) \right)$$

$$\text{s.t.} \quad \bar{t}^+ - \bar{t}^- - (\bar{r}_h^+ - \bar{r}_h^-) - B(\bar{s}_h^+ - \bar{s}_h^-) = \bar{t}_h, \quad h = 1, \ldots, k$$

$$\bar{t}^+, \bar{t}^-, \bar{r}_h^+, \bar{r}_h^- \geq \bar{0}_m, \ \bar{s}_h^+, \bar{s}_h^- \geq \bar{0}_n$$

$$\bar{t}^+, \bar{t}^-, \bar{r}_h^+, \bar{r}_h^- \in \mathbb{Z}^m, \ \bar{s}_h^+, \bar{s}_h^- \in \mathbb{Z}^n$$

The constraint matrix A has the following form:

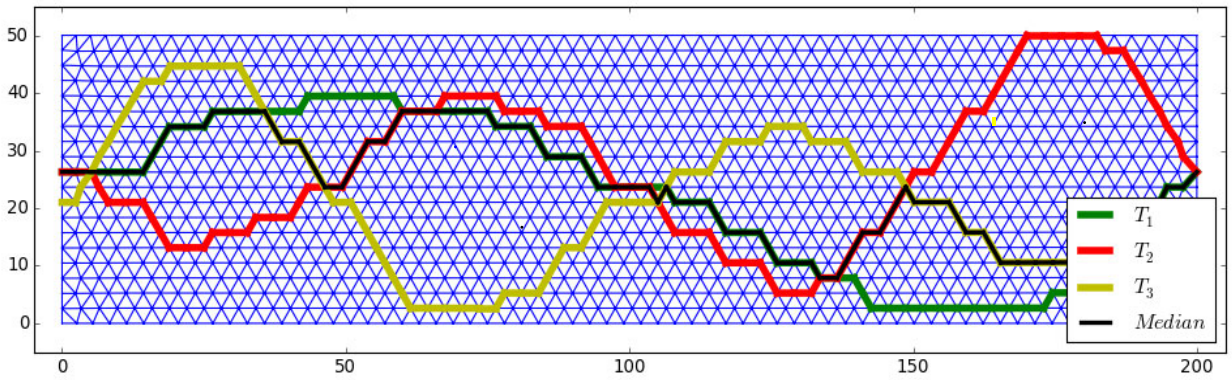
$$A = \begin{array}{c} \begin{matrix} \bar{t}^+ & \bar{t}^- & \bar{r}_1^+ & \bar{r}_1^- & \bar{s}_1^+ & \bar{s}_1^- & \bar{r}_2^+ & \bar{r}_2^- & \bar{s}_2^+ & \bar{s}_2^- & \cdots & \bar{r}_k^+ & \bar{r}_k^- & \bar{s}_k^+ & \bar{s}_k^- \end{matrix} \\ \begin{bmatrix} [I & -I] & [-I & I & -B & B] & & & & & & & & & \\ [I & -I] & & & & & [-I & I & -B & B] & & & & & \\ \vdots & & & & & & & & & & \ddots & & & & \\ [I & -I] & & & & & & & & & & [-I & I & -B & B] \end{bmatrix} \end{array} \quad \left. \vphantom{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}} \right\} k \text{ blocks}$$

Because of the common $[I \ -I]$ columns $(\bar{t}^+, \bar{t}^-)$, A may not be TU even when B is! Still, every instance of this median shape LP yields integer solutions in practice!

See the paper for details: https://arxiv.org/abs/1802.04968.
(Theoretical and computational results).

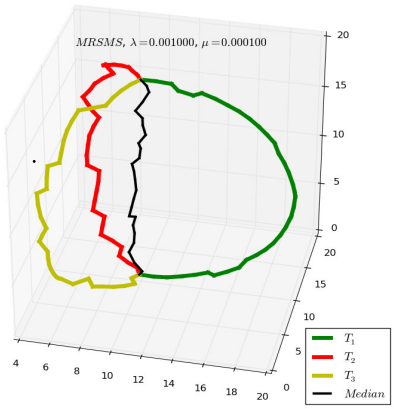
Here are some computational experiments.

1. Input is 3 (simplicial) curves in $\mathbb{R}^2$:



2. Similar to OTEP and flat norm LP producing integer solutions (for fee) in codimension 2 (1-d input in $\mathbb{R}^3$), the median shape LP works as well!

The input is 3 semi-circular curves connecting the North and South poles with some noise on the surface of S^2. They are separated by $\sim 120°$ angles @ the poles. The domain is the 3-ball (solid) tetrahedralized. We get the "axle" from North to South pole as the median shape (with mass regularization).

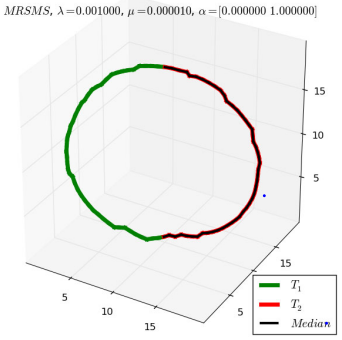


↳ The simplicial complex (2-skeleton) is not shown for clarity reasons.

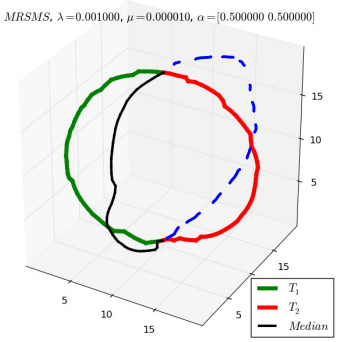
The LP framework for median shape is quite versatile, and can model several variants of the default median shape problem. One variant considers weighting the flat distances from input T_h by α_h $(0 \le \alpha_h \le 1)$ such that $\sum_{h=1}^{k} \alpha_h = 1$. This approach could be used to traverse the entire shape space of all T_h's in a convex combination fashion. We could also restrict the candidates to lie on a subspace, e.g., $\mathbb{S}^{12}$ instead of $\mathbb{B}^3$.

3. Two curves that form great semi-circles from North to South poles on $\mathbb{S}^{12}$, median shape sought on $\mathbb{S}^{12}$ as well.

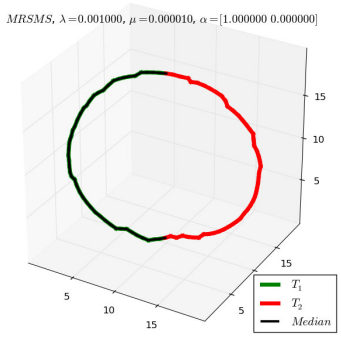
$\alpha_1 = 0 \; \alpha_2 = 1$



$\alpha_1 = \alpha_2 = \frac{1}{2}$



$\alpha_1 = 1, \alpha_2 = 0$



there is non-uniqueness here. One could take the great semicircle going on the "other" side as the median shape as well!!