

# MATH 529: Lecture 1 (01/13/2026)

Today:  $\begin{cases} \text{* syllabus, logistics} \\ \text{* two motivating applications} \end{cases}$

Call me Bala!

Introduction to Computational topology  $\xrightarrow{\text{focus for this course}}$

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"  $\rightarrow$  place or space  
"logos"  $\rightarrow$  study or talk } in Greek

Topology talks about how space is connected.

topology

```

graph LR
    A[Topology] --> B[point set topology]
    A --> C[algebraic topology]
    
```

point set topology (open/closed, connected, ...)

algebraic topology (groups, addition, basis, ...)

We will concentrate on algebraic topology.

## Computational topology

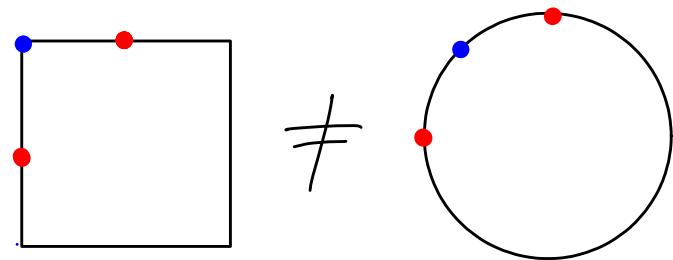
combine efficient algorithms and data structures with results from topology to analyze real-life data.

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.

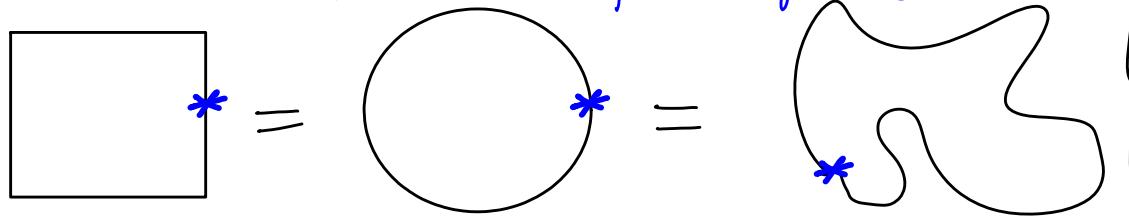
Size (length, here) is critical in geometry, but not so much in topology.



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

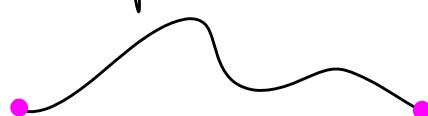
For instance, take a string, and tie a knot to make a loop.

We want to study connectivity irrespective of size here!



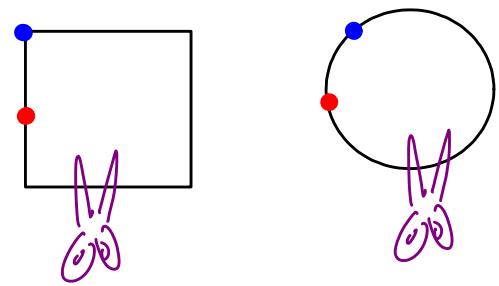
All these objects are same, i.e., they are connected the same way.

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

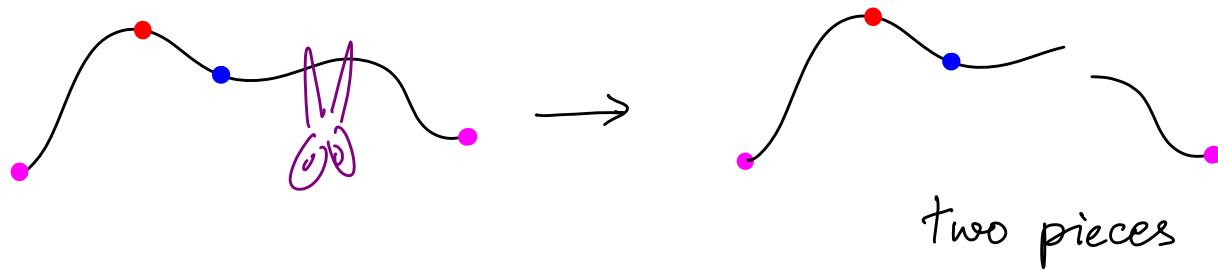


Note that the two end points have different "neighborhoods."

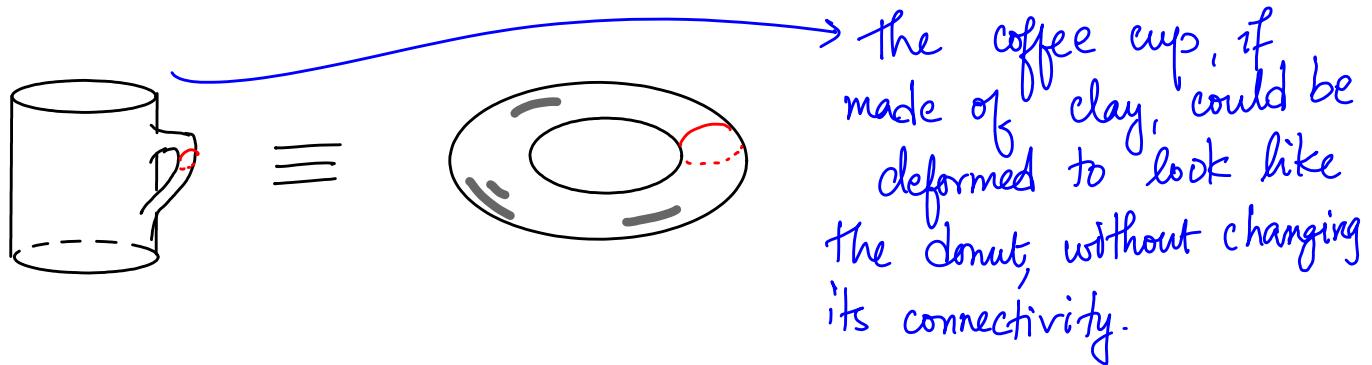
Here is another way to understand connectivity. Consider cutting the string (tied into a loop) once. Such a cut leaves the string in one piece, i.e., connected.



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



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



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

A **A**  $\neq$  B **B**  $\neq$  C **C**

# Two Illustrations of Computational Topology

## 1. Patient antibiotic trajectories

<https://doi.org/10.1145/3307339.3342143>

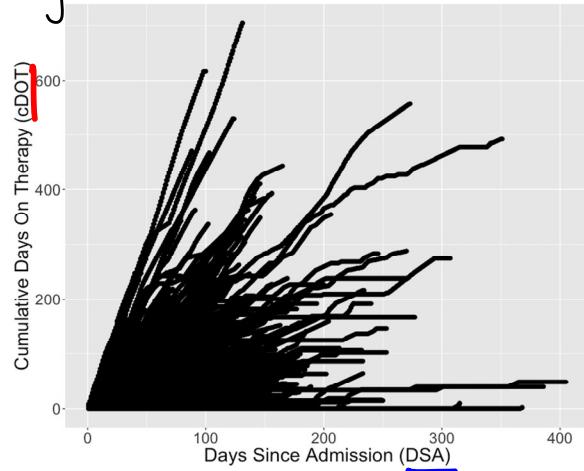
? How do agents and doses affect length of stay?

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

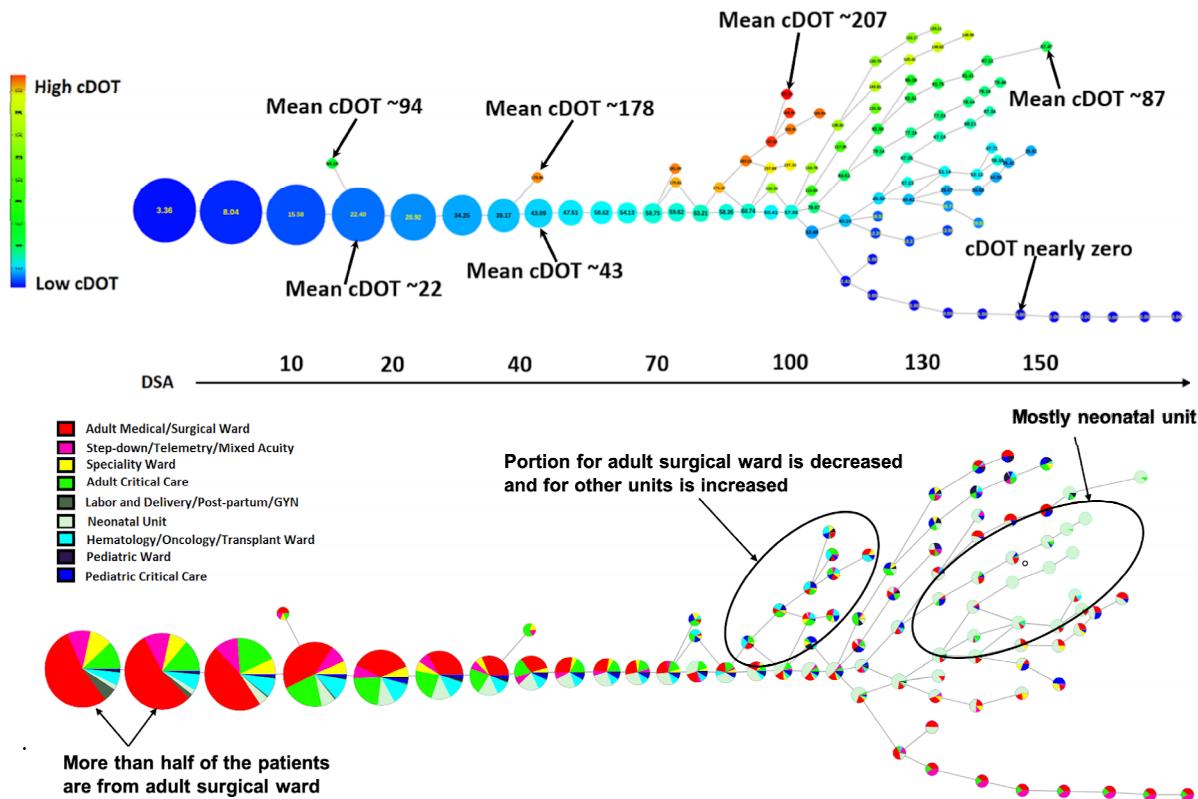
If a patient gets one dose of an agent (antibiotic) that is counted as 1 Day On therapy (DOT).

cDOT : cumulative Days On Therapy  
 DSA : Days Since Admission.



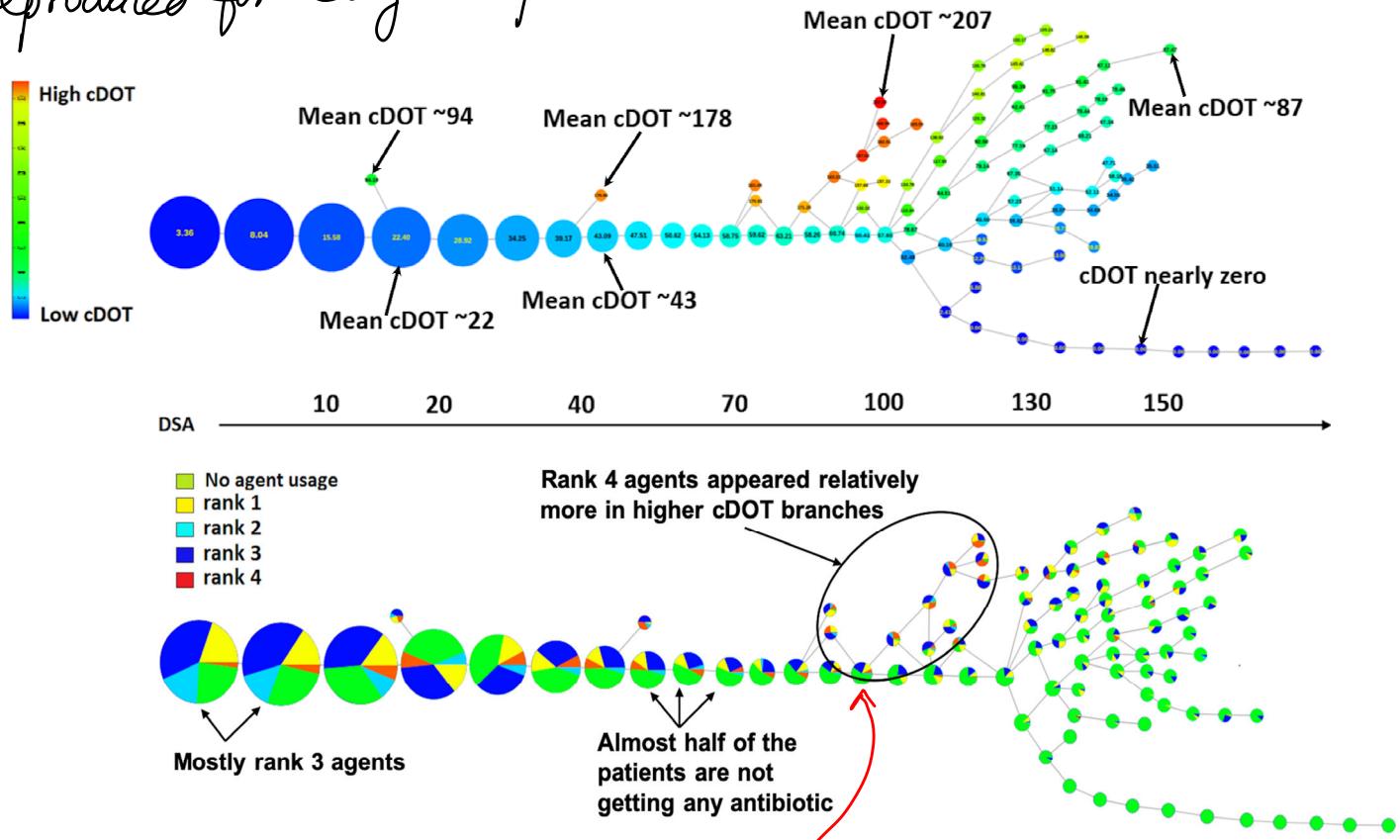
This plot by itself is not very informative or insightful (even if we were to use color...)

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.



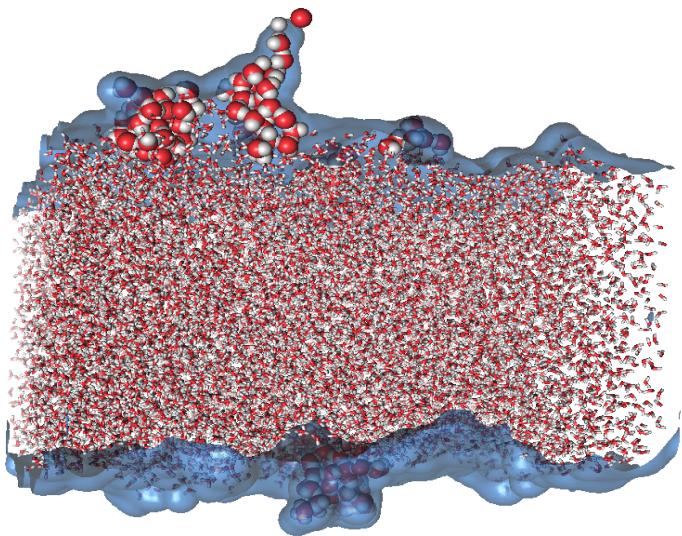
The high cDOT + high rank sub-branches had more patients in the other (higher risk) wards. Similarly, the much higher DSA group (120+) on the right end with relatively smaller cDOT values turned out to be patients in neonatal ICUs.

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

## 2. Interface features in Chemistry

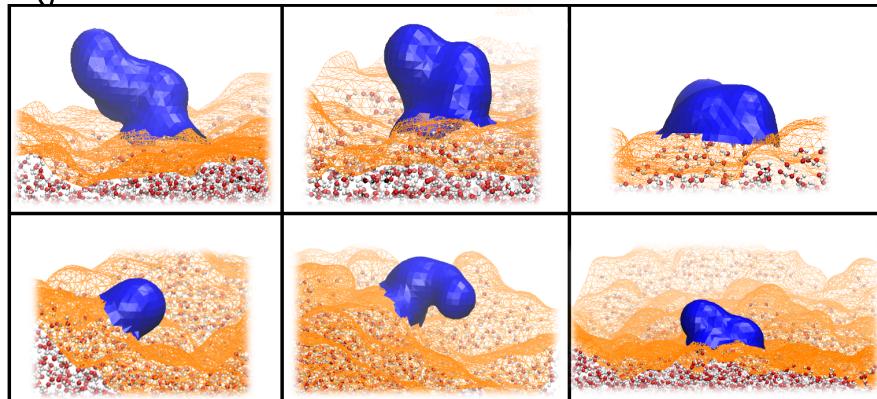
<https://doi.org/10.1021/acs.jctc.0c00260> ( <https://doi.org/10.26434/chemrxiv.11988048.v1>)

→ 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!

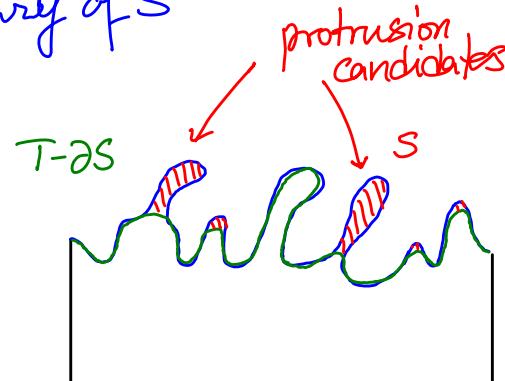
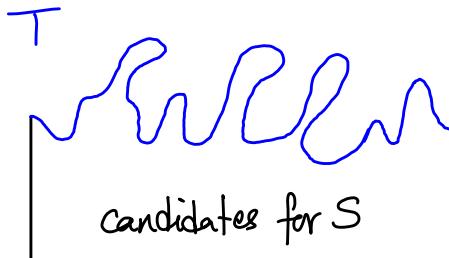
We use the notion of multiscale flat norm of surface  $T$ :

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

scale parameter

$\xrightarrow{\text{3D volume}}$   $\xrightarrow{\text{boundary of } S}$

Illustration in 2D:



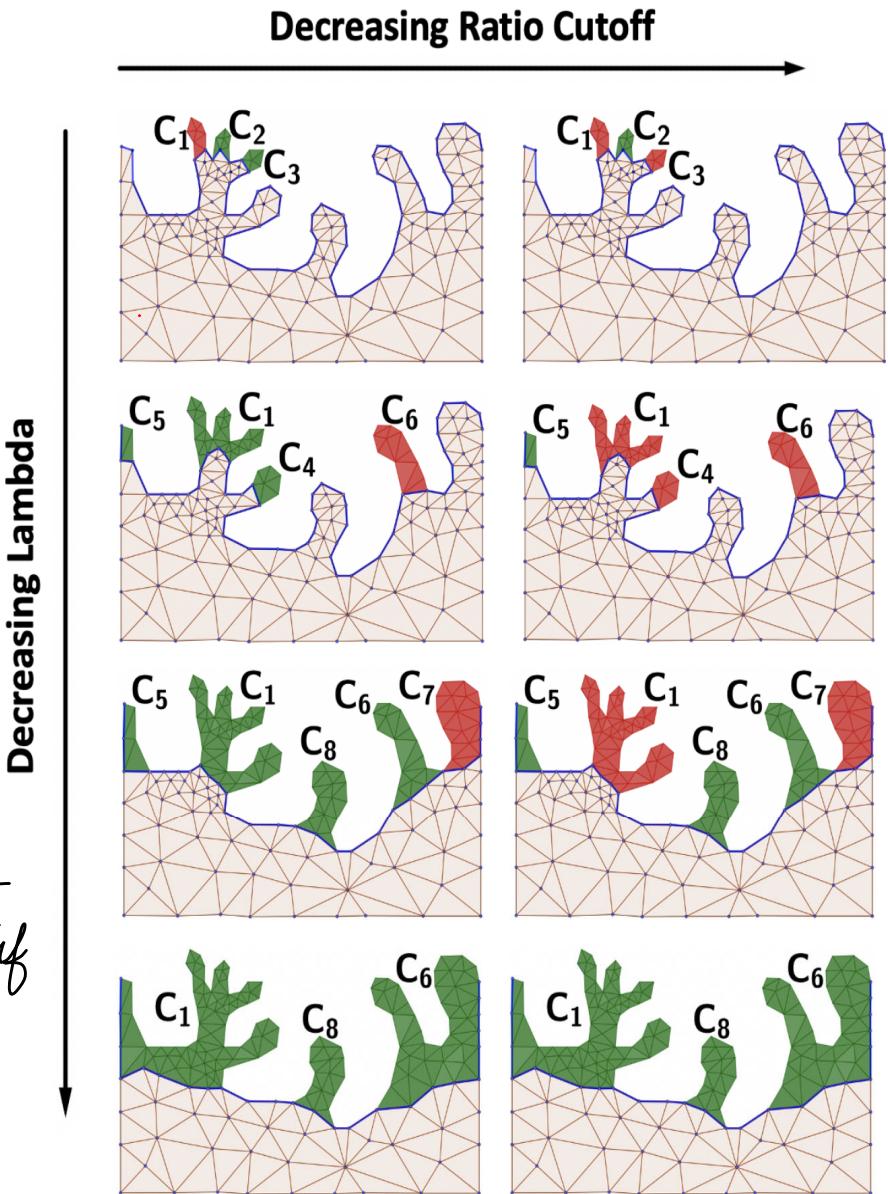
We keep track of connected components in  $S$  as  $\lambda \downarrow$ . We relabel them and also keep track of merging behavior.

We also track the ratio  $\frac{\text{vol}(C)}{\text{vol}(B(\lambda))}$  for each

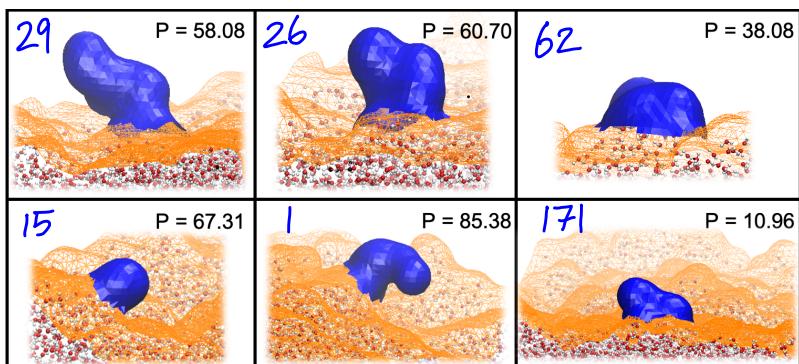
component  $C$ , where  $B(\lambda)$  is the ball with radius  $\lambda$  and  $\text{Vol}(C)$  is the volume of component  $C$ .

A component  $C$  is "alive" at ratio cutoff  $r$  and scale  $\lambda$  if

$$\frac{\text{Vol}(C)}{\text{Vol}(B(\lambda))} > r.$$



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



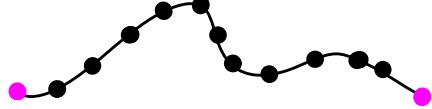
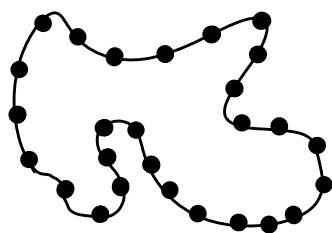
It turned out that all six of these features were protrusions! The probabilities (as %'s) along with their ranks among 195 candidate features (lower rank  $\Rightarrow$  more likely to be a protrusion) are shown here.

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 (01/15/2026)

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)  
↑ index set  
infinite or finite
3.  $\emptyset, X \in T$  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  $X \setminus A$ ) 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  $\text{int } A$  of  $A \subseteq X$ :  $\text{int } A = \bigcup^{\text{union}} (\text{open sets contained in } A)$

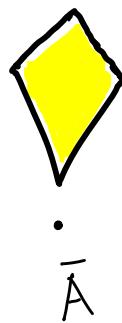
closure  $\bar{A}$  of  $A \subseteq X$ :  $\bar{A} = \bigcap^{\text{intersection}} (\text{closed sets containing } A)$ .  
 minimal closed set that contains  $A$ .

boundary  $\partial A$  of  $A \subseteq X$ :  $\partial A = \bar{A} - \text{int } A$ .

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

### Examples

1.

 $A \subseteq X$  $\text{int } A$  $\bar{A}$  $\partial A$ 

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

$$A = \{a, b\}, \text{int } A = \bigcup \{\emptyset, \{b\}, \{a, b\}\} = \{a, b\} = A.$$

We can specify other topologies on  $X$ , e.g.,  $T_3 = 2^X$ , where each set in  $T_3$  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  $x \in \overset{\circ}{A}$ .

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  $f: A \rightarrow B$  is a rule that assigns exactly one  $b \in B$  for every  $a \in A$ .

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

$\text{im } f$ : image of  $f = \{b \in B \mid f(a) = b \text{ for some } a \in A\} = \{f(a) \mid a \in A\}$ .  
 $\text{im } f$  is also called the range of  $f$ . Note that  $\text{im } f \subseteq \text{cod } f$ .

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

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

A continuous function is also called a map.

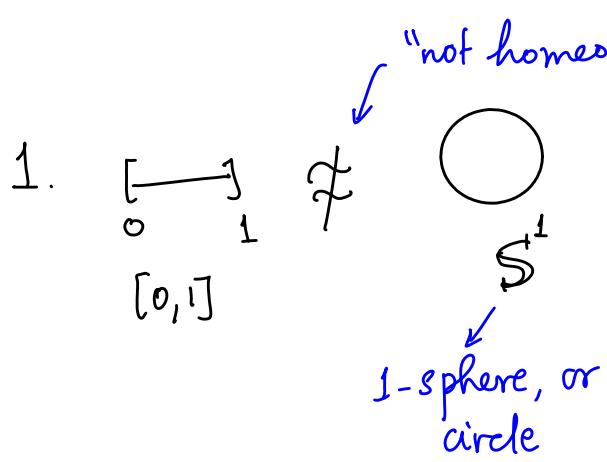
**Def** A homeomorphism  $f: \mathbb{X} \rightarrow \mathbb{Y}$  is a bijective function such that both  $f$  and  $f^{-1}$  are continuous.

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

We also say that  $\mathbb{X}$  and  $\mathbb{Y}$  have the same topological type.

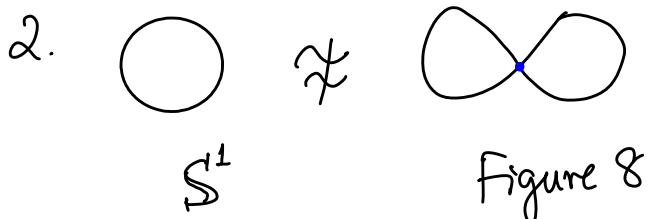
Examples

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.



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

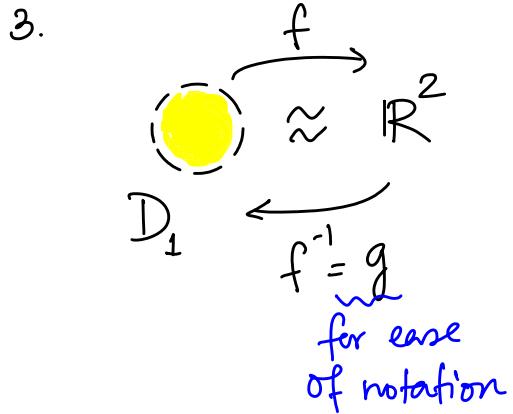
But the inverse of any map that takes both end points of  $[0, 1]$  to one point in  $S^1$  is not bijective.



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

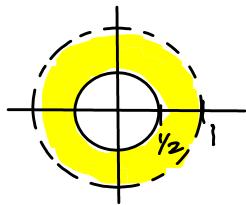
Also, we could map  $S^1$  to one of the two circles in figure-8, but not both.

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



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

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



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

Euclidean norm

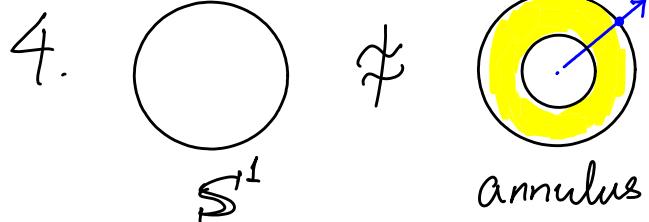
$g$  maps all of  $D_1$  (in  $\mathbb{R}^2$ ) to fit within  $D_{\frac{r_2}{2}} = \{\bar{x} \in \mathbb{R}^2 \mid \|\bar{x}\| < \frac{1}{2}\}$ , and then fits all of  $\mathbb{R}^2$  outside  $D_1$  within the half open annulus with radii  $\frac{r_2}{2}$  and 1.

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

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

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  $D$ . We stretch the open disc so as to fill the entire plane, and hence it is called an infinite stretch.

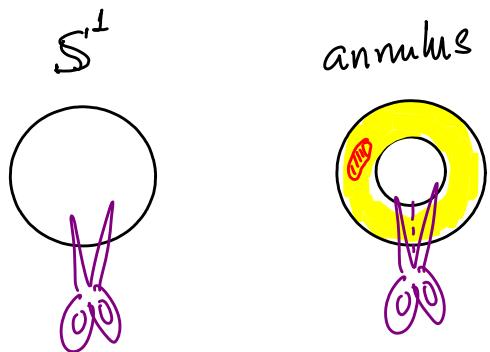
Usually, we try to define the continuous maps  $f$  and  $f^{-1}$  to show that two spaces  $X$  and  $Y$  are homeomorphic. At the same time, the intuition (geometric when possible) is also important to grasp. On the other hand, to show that  $X$   $\not\cong$   $Y$ , it is often sufficient to identify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



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 ( $\mathbb{R}'$ ). 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.



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 (red) 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!