What is Persistent Homology?

What This Presentation Will Do

- I will give an intuitive explanation of persistent homology, the main technique of topological data analysis
    - Topological data analysis studies the shape (qualitative features) of datasets with tools from topology
- My project itself will be coding an algorithm to compute persistent homology, but this presentation will largely explain what persistent homology is
- This corresponds to what I’ve done in the project so far - next steps are to understand and program the algorithm
- I’ll start by provide some background about topology - the presentation aims to capture key concepts intuitively rather than working through the material technically

Topology

Some classic continuous deformations (homeomorphisms)

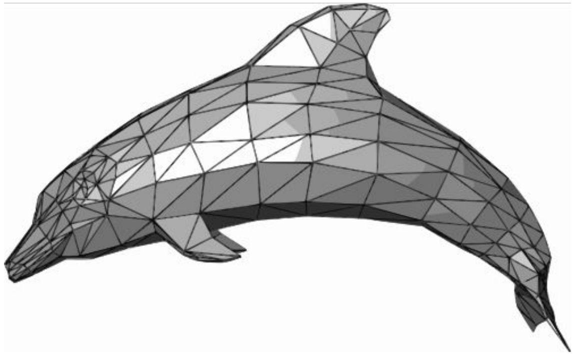
Gif source: https://en.wikipedia.org/wiki/Topology

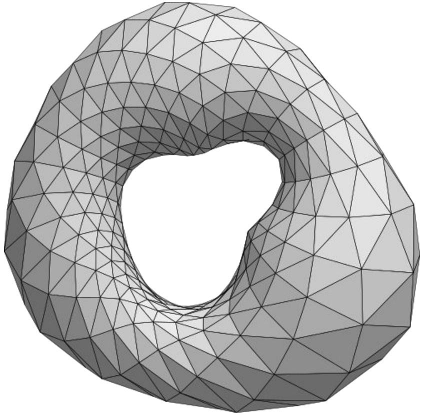
Topology

- Topology studies qualitative features of spaces
- Typically these are topological properties, which are properties that are invariant under homeomorphism
- A common type of example is loops/holes
  - A circle is a 1-dimensional hole, a sphere a 2-dimensional hole
- We can detect these with something called homology, which misses a lot of features that homeomorphism can detect

Triangulation

- Homology is defined via a *triangulation*, or a homeomorphism from the topological space of interest to something called a *simplicial complex*





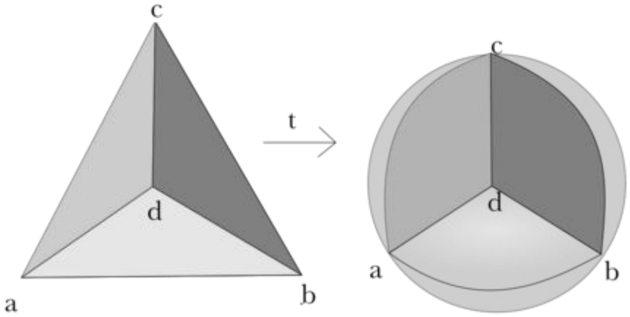


Image source: https://en.wikipedia.org/wiki/Triangulation_(topology)

Triangulation: Simplices

- Simplicial complexes are made up of simplices (or simplexes)
- We can (a little loosely) define simplices inductively: a 0-simplex is a point, and a k+1 simplex is taken by adding a vertex to a k-simplex and connecting it to all the points of the k-simplex

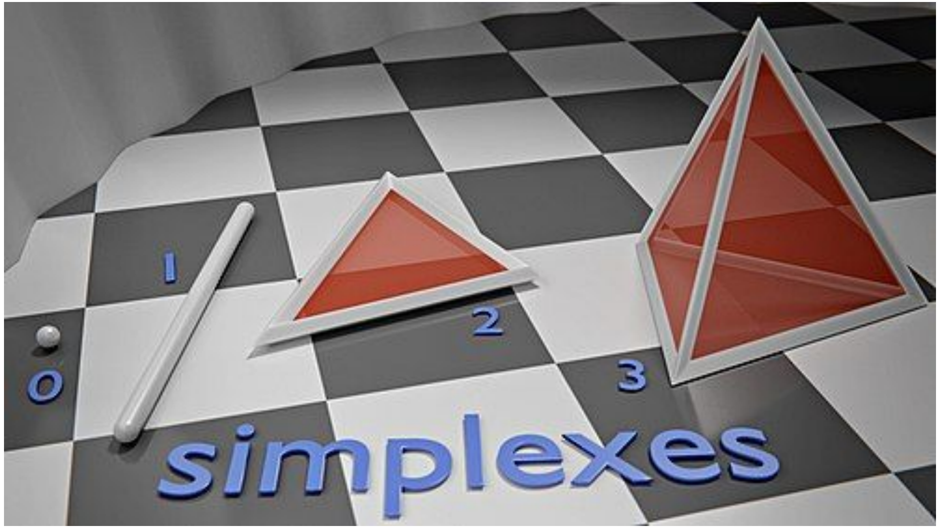


Image source: https://en.wikipedia.org/wiki/Simplex

Triangulation: Simplicial Complex

- Simplicial complexes are made by attaching a bunch of simplices at their faces (subsimplices)

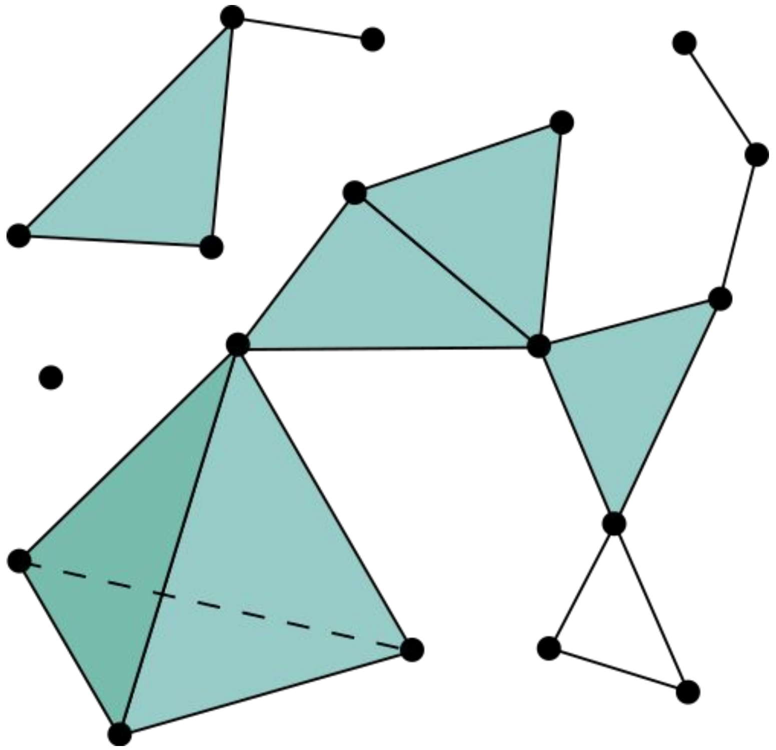


Image source: https://en.wikipedia.org/wiki/Simplicial_complex

Cycles and Boundaries

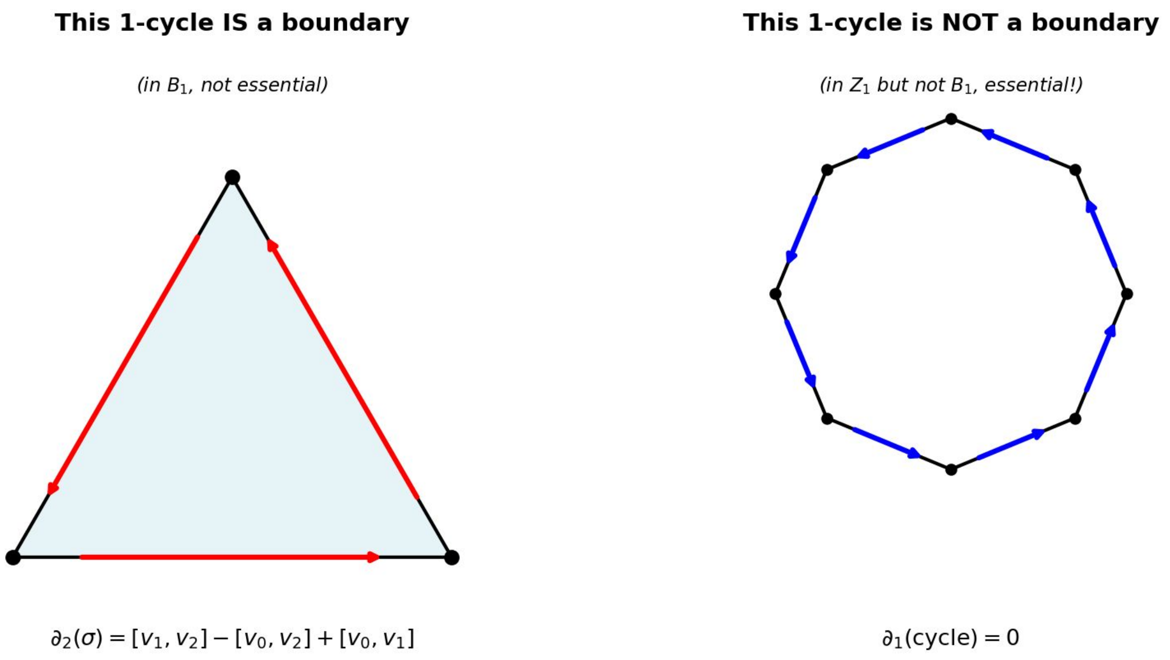


Image source: Claude

Cycles and Boundaries

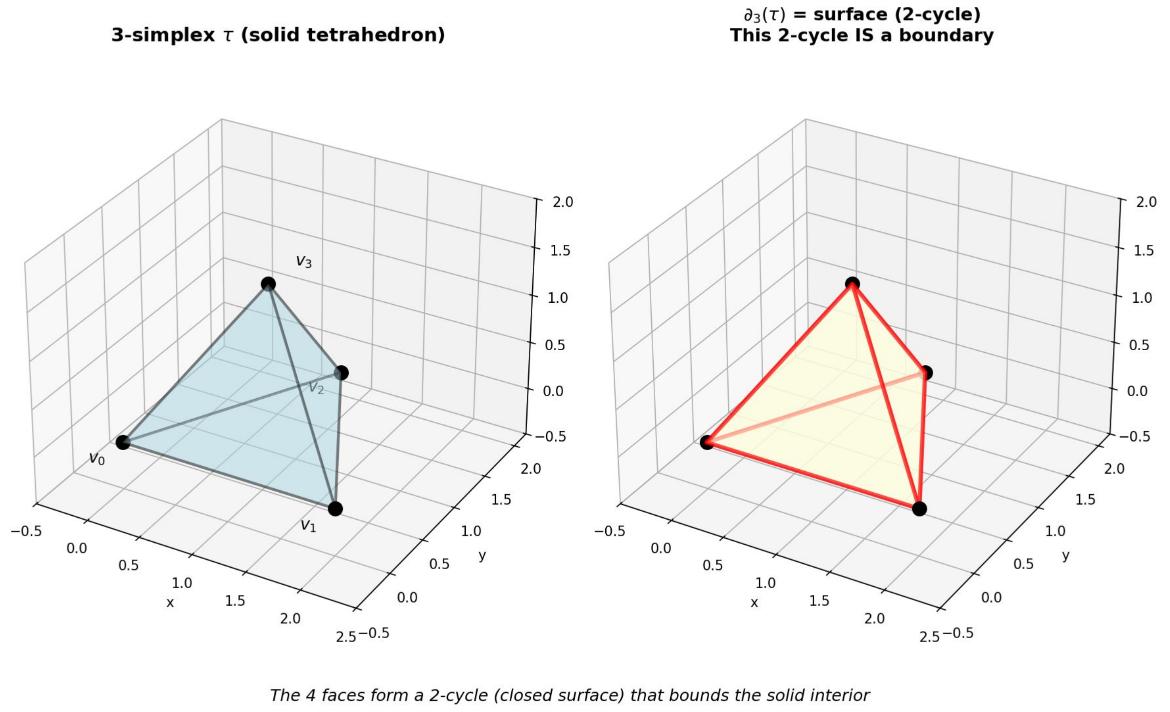


Image source: Claude

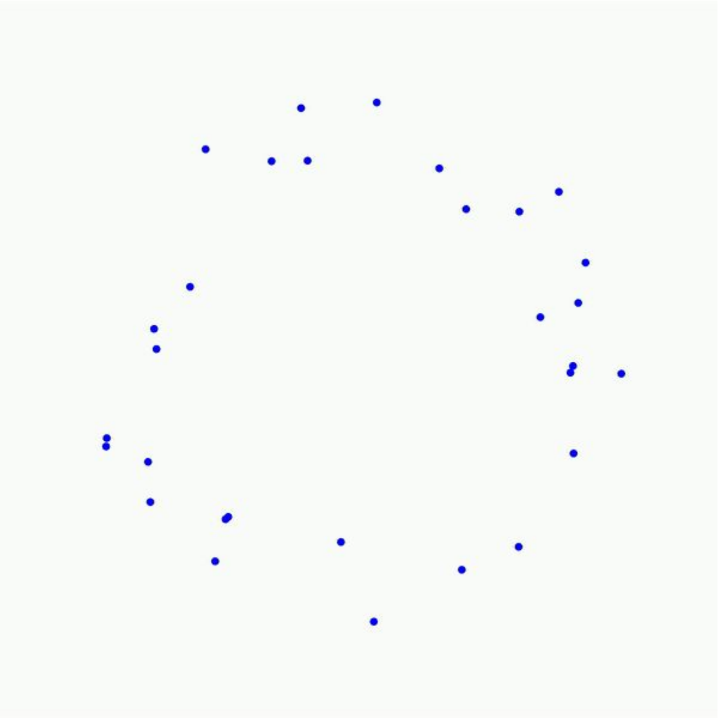
Homology

- Key insight: loops/holes in a space correspond to cycles that are **not boundaries** in that space's triangulation
- Something called a *homology group* keeps track of the structure of cycles that are not boundaries
- You will have different homology groups for different dimensions, corresponding to the differing dimensions of loops

Recap

- This project is in topological data analysis, which analyzes data by computing qualitative features of the data with topological tools
- We are specifically looking for loops/holes in data, which we can find by triangulating our space, then finding homology groups by finding the boundaries and cycles of the corresponding complex
- Two hanging questions: how do we triangulate a data set, and what is *persistent* homology? The answer to the first will prompt the answer to the second

Vietoris-Rips Complex



- Given a distance d, we will include a simplex of a set of points iff all points in the set have distance at most d
- We can see that increasing d will increase the connectivity of our complex
- Low d -> seeing loops that are just noise
- High d -> not seeing any loops
- How we can detect underlying shape in light of this?

Gif source: https://lewtun.github.io/hepml/lesson06_persistent-homology/

Persistent Homology

- Persistent homology tracks the “persistence” of holes in the data across varying values of d
- This produces a “barcode” diagram that allows us to visually inspect the structure of the data

Persistence Barcodes

- H1 is the group of 1-d loops (like a circle), H2 the group of 2-d loops (like a sphere), H0 unfortunately disanalogously the set of connected components

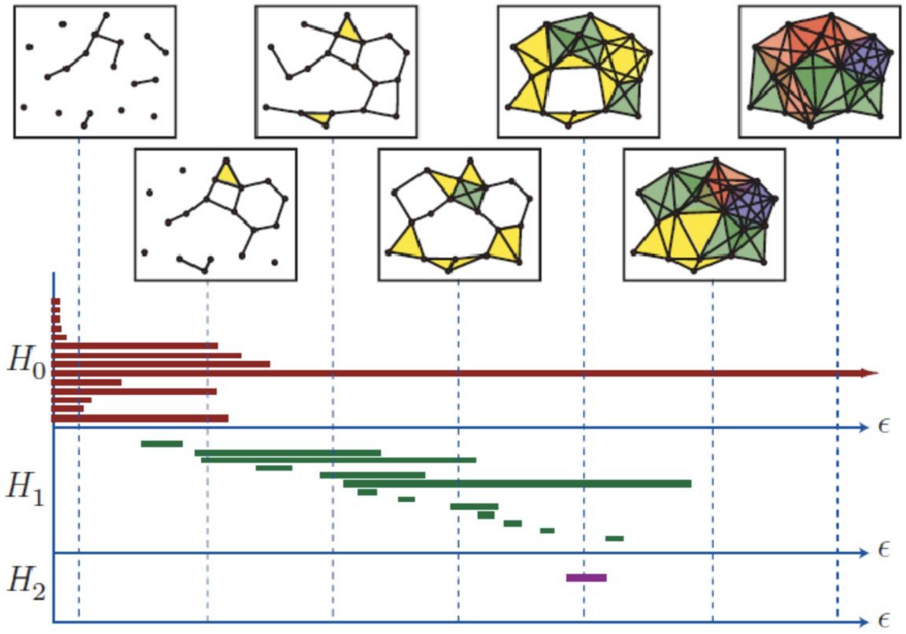


Image source: “Persistent Homology for Random Fields and Complexes”, https://math.uchicago.edu/~shmuel/PH-RF.pdf

Questions?