Section 2

From data to topological representation

There are many instances of real data that can be described using topology. The basic example are:

1. 2, 3, higher dimensional images, (pictures, static or dynamic ct scans...)

2. Point clouds – e.g. charateristics of patients with Covid 19.

3. Time series

4. ...

We need to represent all those data in the way that allows to compute their topology

Data $\rightarrow$ Combinatorial representation $\rightarrow$ Chain complex $\rightarrow$ (persistent) homology

Let us quickly discuss what this is for various type of data.

Images naturally correspond to so called **cubical complexes**.

What is a cubical complex? Let us start by asking a more fundamental question: What is **a cube**?

"Before the time began, there was the Cube. We do not know where it comes from, only that it holds the power... "

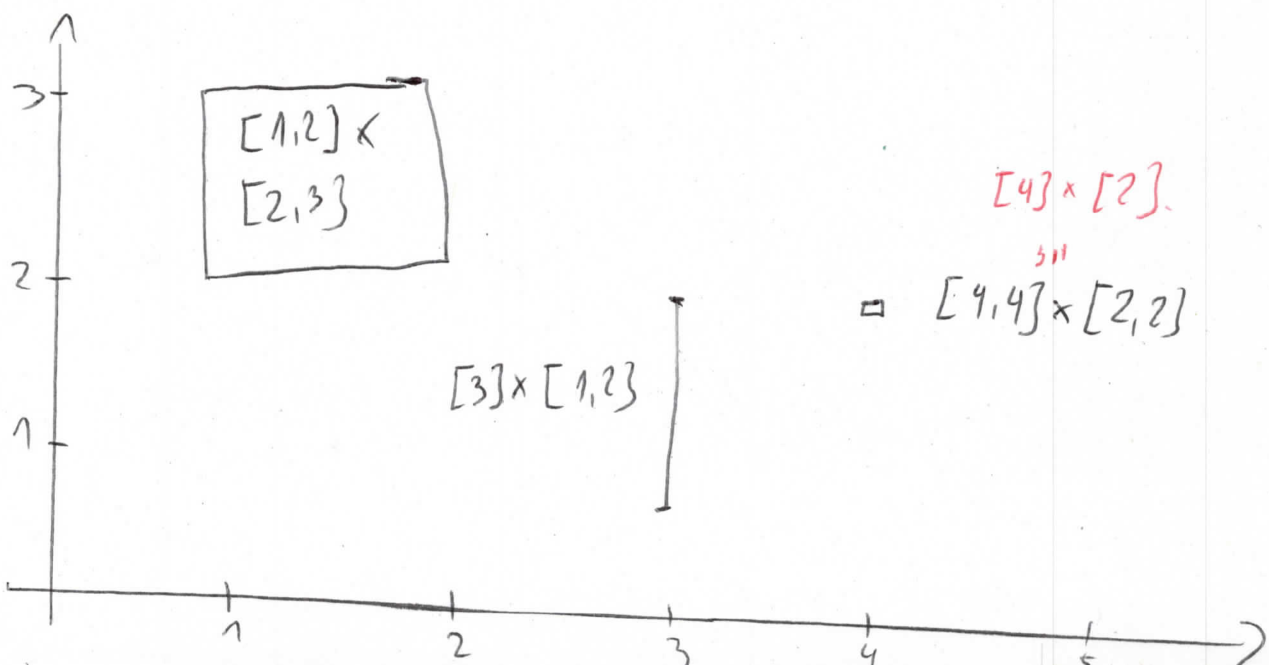
Optimus Prime - Transformers

More seriously,

Elementary interval is an interval of a form:

$$\left.\begin{array}{l}[n, n+1] \\ [n, n]\end{array}\right\} \text{ for } n \in \mathbb{N}.$$

A cube is a Cartesian product of elementary intervals



Cubes, especially those that decompose a rectangle in the ambient space, can be very effectively stored in a computer in a form of bitmap.

That include:

(1) For black & white (tresholded) images:

B & W image

0	1	0
1	0	1
0	1	0

black = 1
white = 0

Bitmap

$$[\underbrace{0,1,0,1,0,1,0,1,0}_{\text{vector of 9 values in lexicographical order}}],\ \underset{\text{number of rows}}{3},\ \overset{\text{number of columns}}{3}$$

One bit (not byte) per cube

(2) For greyscale or RGB images similarly, but we need to store more...

8	239	11
251	242	228
9	236	3

grayscale values, integers between
0 & 255

↓

[9, 236, 3, 251, 242, 228, 8, 239, 11] + sizes of picture

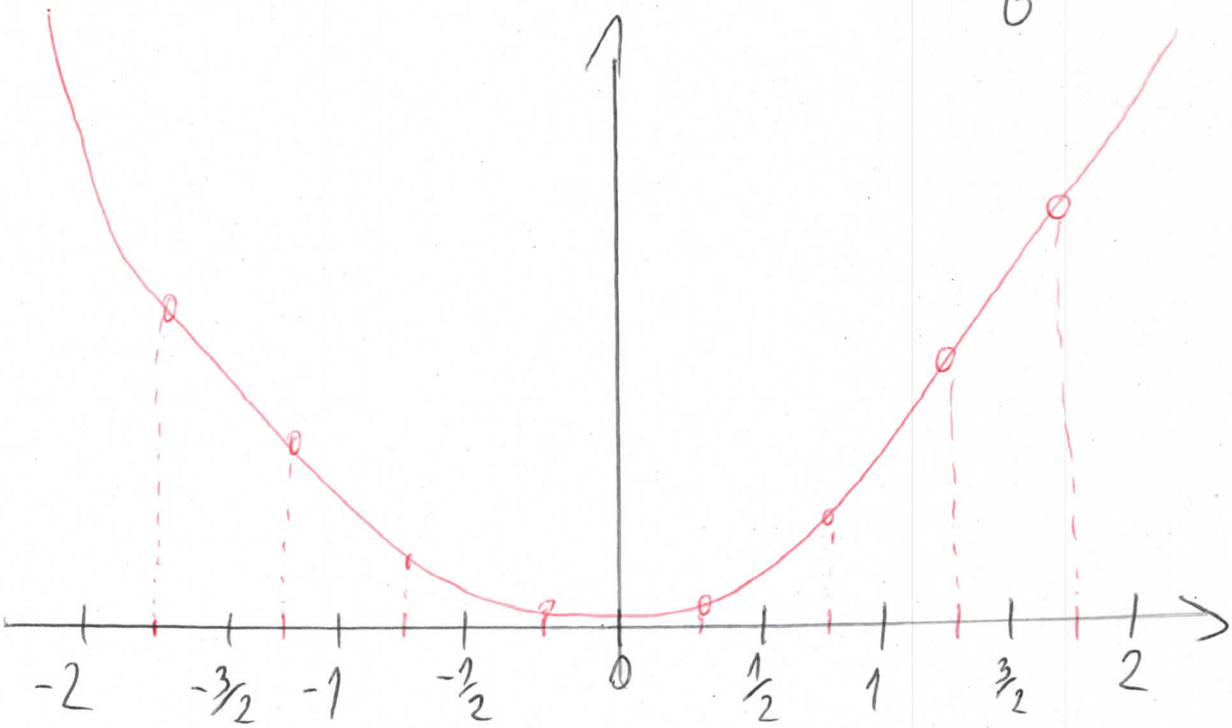
One integer per cube
For RGB we will store three integers per cube...

Cubical complexes were introduced & promoted (in oppose to simplicial complexes) by Jean-Paul Serre in the middle of 20'th century. Later they were made popular with so called interval arythmetic

But, we will not get too deep to this topic in this tutorial. Let me just say that the cubical complexes allow to approximate functions on rectangular domains.

Let us consider a simple example:

$y = x^2$, domain $= [-2, 2]$, number of cubes $= 8$:



$\left(-\frac{7}{4}\right)^2$	$\left(-\frac{5}{4}\right)^2$	$\left(-\frac{3}{4}\right)^2$	$\left(-\frac{1}{4}\right)^2$	$\left(\frac{1}{4}\right)^2$	$\left(\frac{3}{4}\right)^2$	$\left(\frac{5}{4}\right)^2$	$\left(\frac{7}{4}\right)^2$

Let us now play with two dimensional images & their persistence.

Let us do it on the google colab file i will share with you now

< Let us play with python >

Having the 2d example of a cubical complex we can play & observe the concept of stability of persistence.

Formally, in this contex:

Let $\Omega \subseteq \mathbb{R}^n$ and $f, g : \Omega \to \mathbb{R}$.

Let P = persistent homology of Ω with a filtration f

Q = persistent homology of Ω with filtration g.

Then $d_B(P, Q) \leq \|f - g\|_\infty$

Bottleneck distance between diagrams P & Q

L^∞ distance between functions of Ω, i.e.

$\max_{x \in \Omega} |f(x) - g(x)|$.

We will explain soon what the Bottleneck distance is, but first let us pertub the bitmap we get & see if the stability theorem can be observed in practice.

Cubical complexes can also be used to approximate values of function.

In our 3rd numerical experiment let us see how one can infere topology of a point cloud (in low dimensional space) by using Kernel Density Estimator. For that let us:

1. Sample a collection of points from our favorite set (in my case, I will be boring & will sample from a unit circle).
2. Compute a Kernel Density Estimator (KDE) of these points (for more information, consult statistic courses - in general this technique allows ~~for~~ to obtain a possible density function (possible distribution) from which the input collection of points is sampled.
3. Approximate KDE on a fixed size grid.

Comment 1

This approximation is not rigorous i.e. we do not know how far away (in, say L^p or L^∞ norm) the function is from its partially constant approximation.

But.. we can make it rigorous using rigorous arithmetic, for instance interval arithmetic.

[Dlotko, Wanner, Rigorous cubical approximation and persistent homology of continuous functions]

Comment 2

The number of cubes in a cubical complex meshing a square $\underbrace{n \times n \times \dots \times n}_{k}$ grows exponentially with k.

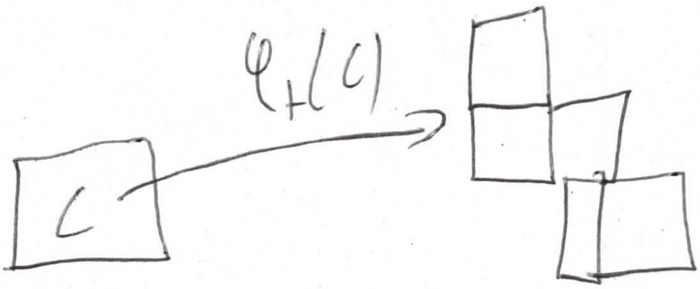
It is therefore no possible to use those techniques in higher dimensions.

Bonus Section

Topology meet Dynamics

Using rigorous arithmetic, like interval arithmetic, it is possible to approximate dynamics.
To be precise, we know where the point, after time T, may be mapped:



More precisely, given a space X and a function

$$\Phi : \mathbb{R} \times X \to X$$

continuous dynamical system

We require that:

(1) $\Phi(0, x) = x$ ← we do not move in zero time

(2) $\Phi(t_2, \Phi(t_1, x)) = \Phi(t_1 + t_2, x)$ ie

Using a computational technique called rigorous arithmetic, having interval arithmetic as the most popular, we can estimate (over estimate) where an image of a cube is. In some cases however that allows to detect invariant sets of dynamics.

A set $M \subseteq X$ is invariant with respect to Φ if $\Phi(t, X) \subseteq X$ for every t

(we can speak about forward/backward invariant set, but will skip it here).

Examples of invariant sets: fixed points, periodic orbits, ...

Suppose we find a set I such that for every $x \in \partial I$ the trajectory at x either enters I, or exit I. Such a set is called insulating neighborhood.

Let $E = \{ x \in \partial I \mid \text{trajectory at } x \text{ leaves } I \}$

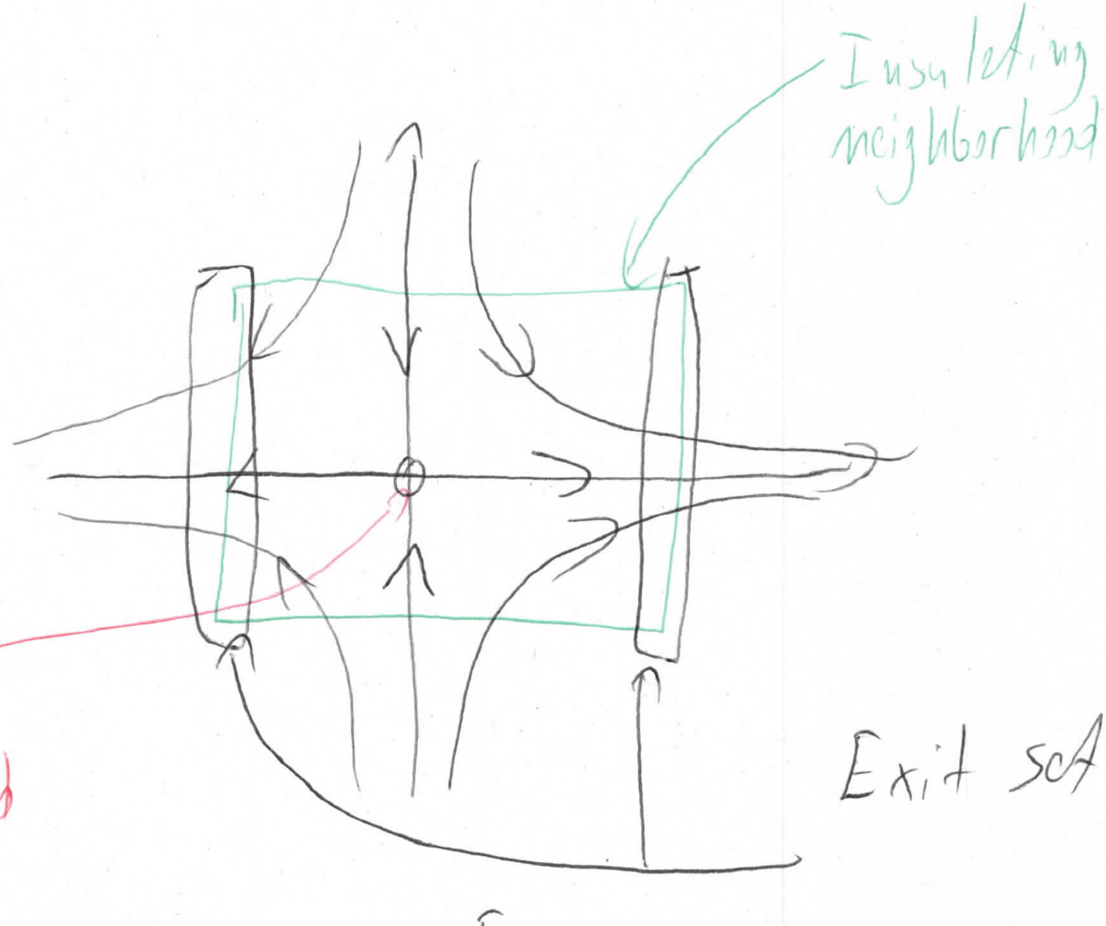
$\uparrow$ this is an exit set.

Then the celebrated result of homotopical Conley Index states that if

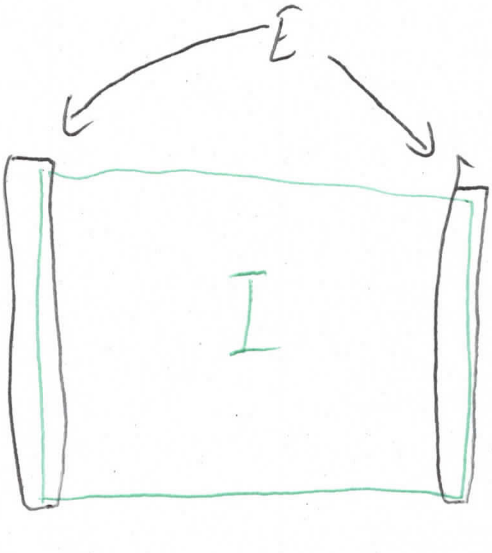
$H_*(I, E)$ is nontrivial, then $\mathrm{Inv}(I) \neq \emptyset$

meaning that there is something invariant in I. It can be fixed point, periodic orbit, ...

E_3:



Therefore



$H(I, E) \cong H(\bullet)$

$\Downarrow$ Conley

$\mathrm{Inv}(I) \neq \emptyset$

It is a very nice example how topology and dynamics go pair-to-pair (or, should I write hand-to-hand?)

Topology help dynamics to find invariant sets. Soon we will see that dynamics help topology to find, for instance and among others, Betti numbers.