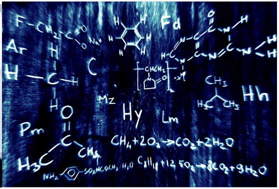
The shape of data in biology

Heather Harrington

Mathematical Institute
University of Oxford

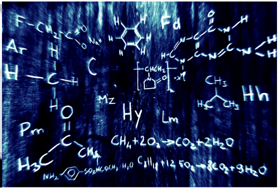
21 September 2016

What do gossip, guitar hero and chemical reactions have in common?



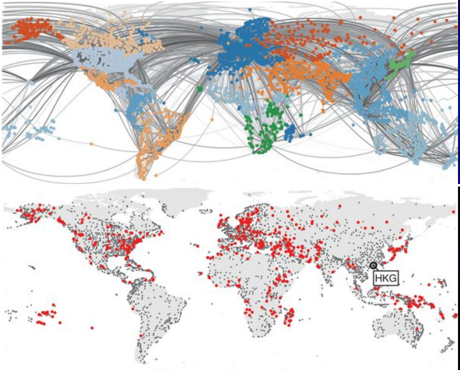
What do gossip, guitar hero and chemical reactions have in common?

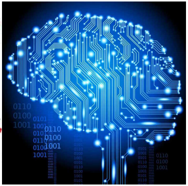




Can mathematics help us find out?

1. How do processes (eg gossip, epidemics) spread?
2. How do we learn (to play guitar hero)?
3. How do cells make decisions (via chemical reactions)?



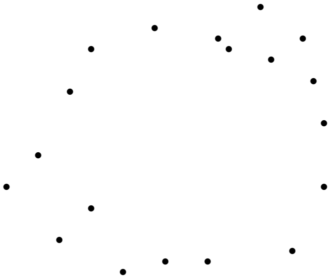
Brockman and Helbing (2013) Science

To study these problems requires models and data.

Data

Models

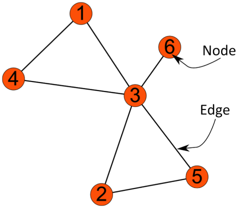
Data



Models

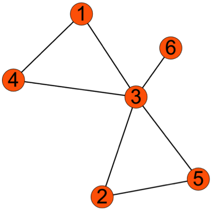
Data

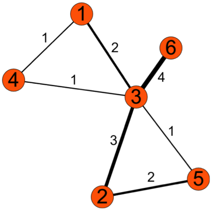
Models



Data

Models





Data

Models

Data

Models

$$y = m \cdot x + b$$

Data

Models

$$y = m \cdot x + b$$

Jon is going punting. The cost to initially hire a punt is £10. The hourly charge is £5/hour. If Jon hired a punt for 4 hours, how much did he pay?

Data

Models

$$y = m \cdot x + b$$

Jon is going punting. The cost to initially hire a punt is £10. The hourly charge is £5/hour. If Jon hired a punt for 4 hours, how much did he pay?

$$y = 5 \cdot x + 10$$

Data

Models

$$y = m \cdot x + b$$

Jon is going punting. The cost to initially hire a punt is £10. The hourly charge is £5/hour. If Jon hired a punt for 4 hours, how much did he pay?

$$y = 5 \cdot x + 10$$

$$y = 5 \cdot 4 + 10$$

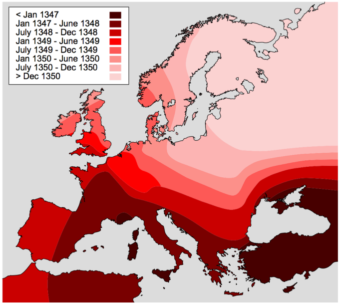
Spreading processes

Social contagion

- Information diffusion (innovations, memes, marketing)
- Belief and opinion (voting, political views, civil unrest)
- Behavior and health

Epidemic contagion

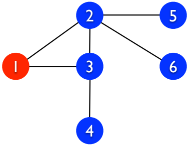
- Epidemiology for networks (social networks, technology)
- Preventing epidemics (immunization, malware)



Black death. Marvel et al (2014) arxiv 1310.2636

Complex contagion

Adoption of a contagion requires multiple contacts with the contagion

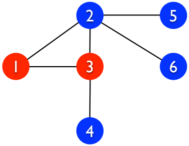


Timestep 0

Node	1	2	3	4	5	6
Activation time t	0					

Complex contagion

Adoption of a contagion requires multiple contacts with the contagion

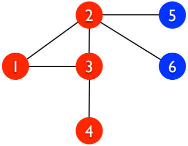


Timestep 1

Node	1	2	3	4	5	6
Activation time t	0		1			

Complex contagion

Adoption of a contagion requires multiple contacts with the contagion

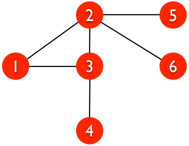


Timestep 2

Node	1	2	3	4	5	6
Activation time t	0	2	1	2		

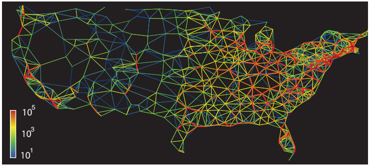
Complex contagion

Adoption of a contagion requires multiple contacts with the contagion

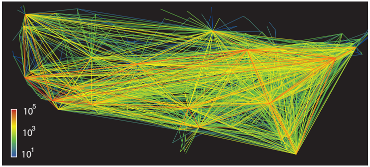


Timestep 3

Node	1	2	3	4	5	6
Activation time t	0	2	1	2	3	3

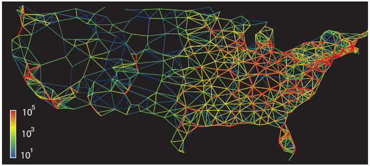


$+$

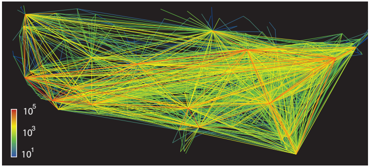


$=$

Balcan et al (2009) PNAS



+



=

Balcan et al (2009) PNAS

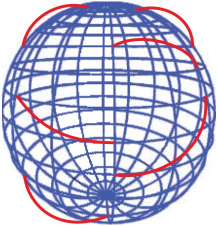
We consider two types of connections:

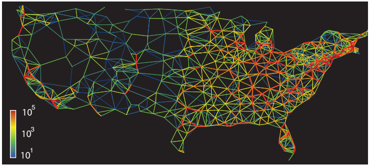
Geometric connections

+

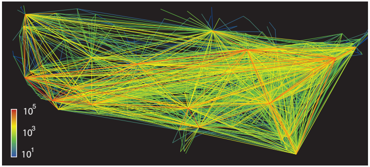
Non-geometric connections

=





$+$



$=$

Balcan et al (2009) PNAS

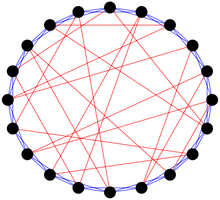
We consider two types of connections:

Geometric connections

$+$

Non-geometric connections

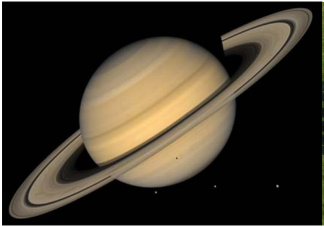
$=$



Topology is concerned with the global properties of space

Topology is concerned with the global properties of space






H is for homology

One can stretch or bend a shape, but not tear or glue it.
Homology allows one to distinguish shapes (even stretching or bending).

Home » About Us » Life in Oxford Mathematics » Oxford Mathematics Alphabet

H is for Homology

The Authors

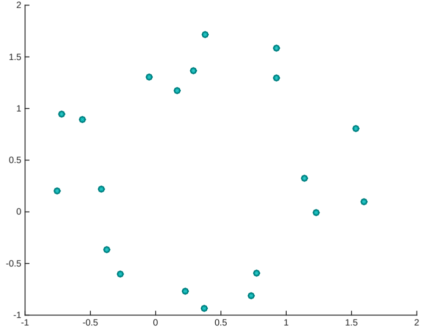
Bernadette Stolz and Barbara Mahler are DPhil students in the

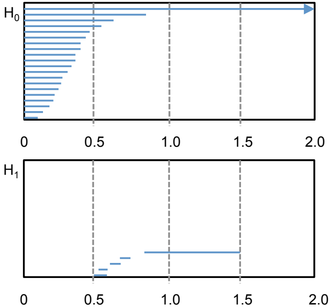
H is for Homology

A life belt, a coffee cup, a jumping ball, a beach ball - what do these objects have in common? What sets them apart? It is questions like these that are considered in the mathematical field called *topology*. A method to study these questions is given by the theory of homology.

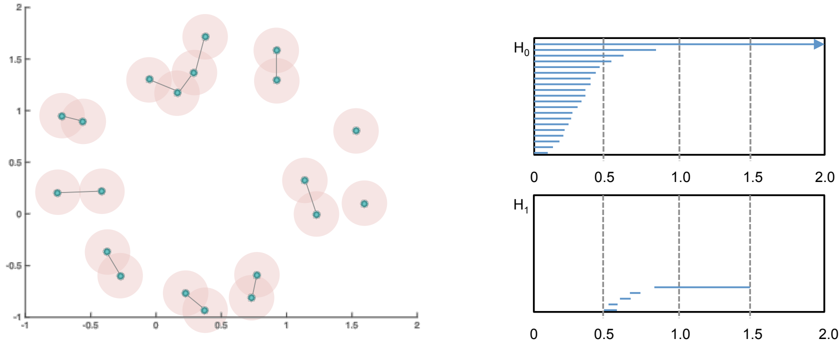
`https://www.maths.ox.ac.uk/about-us/life-oxford-mathematics/oxford-mathematics-alphabet/h-homology`

How to compute topology from a sample of the space?

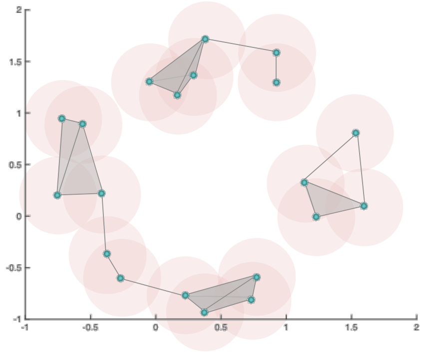


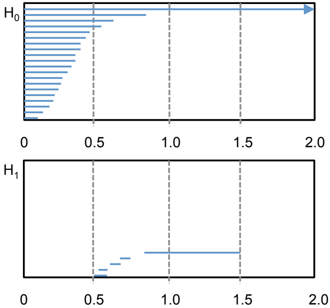


How to compute topology from a sample of the space?

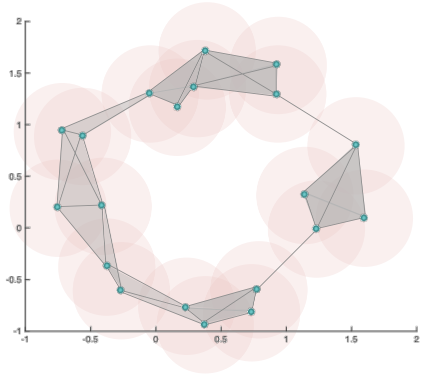


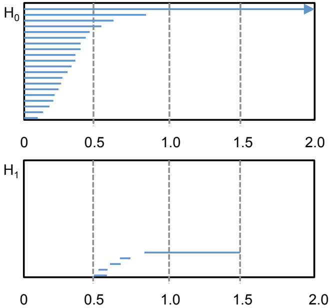
How to compute topology from a sample of the space?



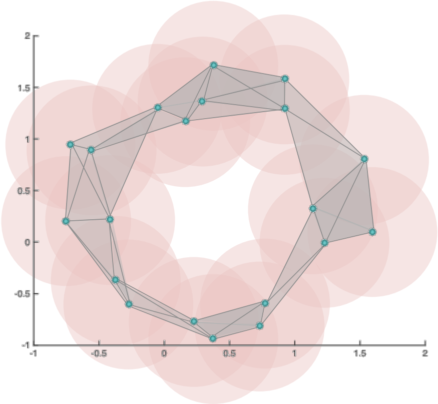


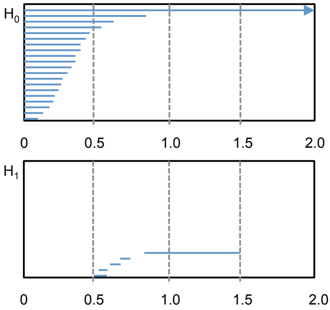
How to compute topology from a sample of the space?



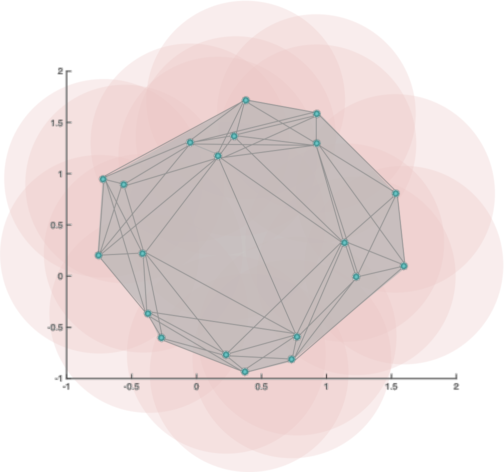


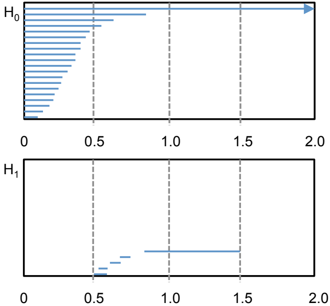
How to compute topology from a sample of the space?

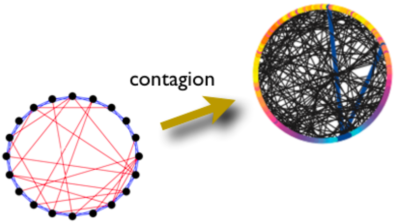


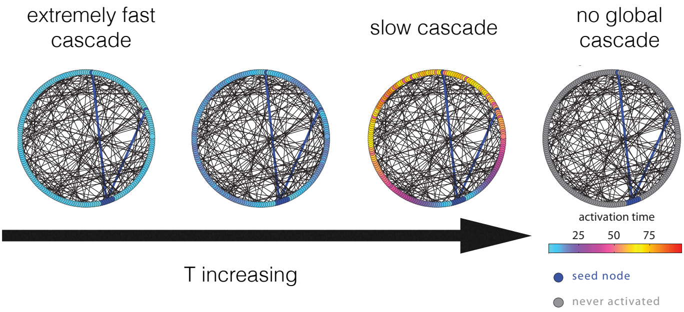


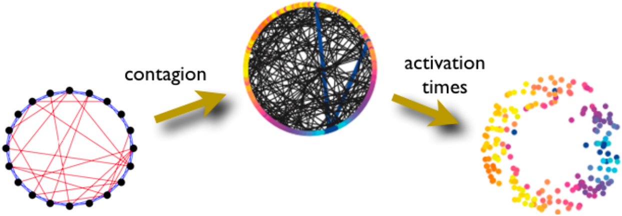
How to compute topology from a sample of the space?

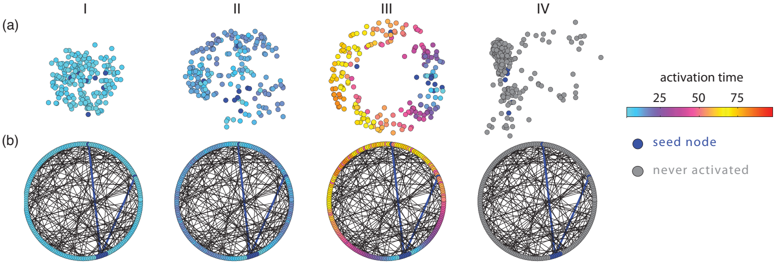




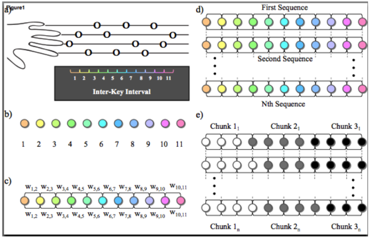

contagion


extremely fast cascade
slow cascade
no global cascade
activation time
25
50
75
seed node
never activated
T increasing

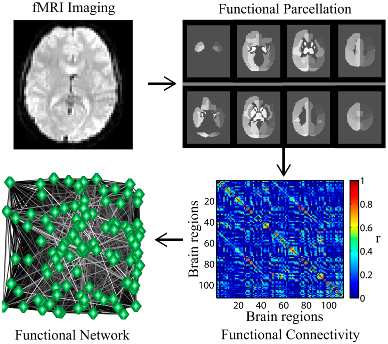

contagion
activation
times


I
II
III
IV
(a)
(b)
activation time
25
50
75
seed node
never activated

How do we learn?



- 100 billion neurons in the brain
- Create a functional network

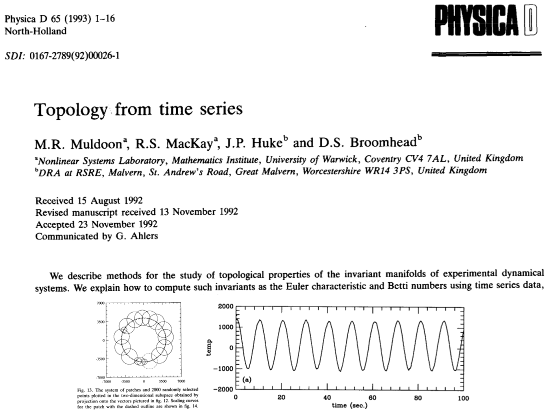


Bassett et al. (2011) *PNAS*.

British mathematician: Sir Christopher Zeeman

Two-page paper available on LMS website.
In search bar: "topological theory of the brain"

British mathematicians: Muldoon, MacKay, Huke, Broomhead



Physica D 65 (1993) 1–16
North-Holland

SDI: 0167-2789(92)00026-1

PHYSICA D

Topology from time series

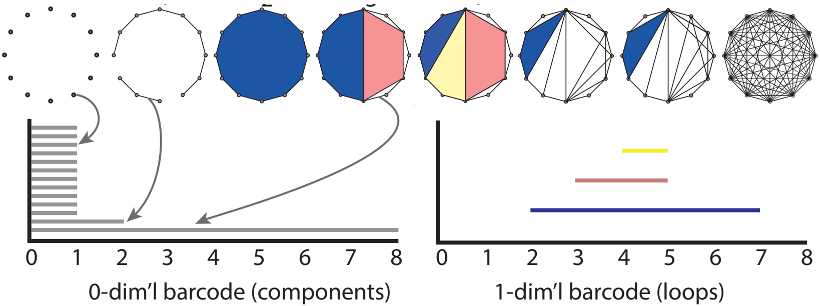
M.R. Muldoon[a], R.S. MacKay[a], J.P. Huke[b] and D.S. Broomhead[b]

[a]Nonlinear Systems Laboratory, Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
[b]DRA at RSRE, Malvern, St. Andrew's Road, Great Malvern, Worcestershire WR14 3PS, United Kingdom

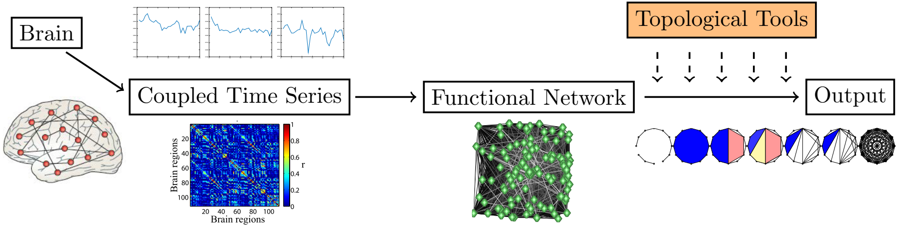
Received 15 August 1992
Revised manuscript received 13 November 1992
Accepted 23 November 1992
Communicated by G. Ahlers

We describe methods for the study of topological properties of the invariant manifolds of experimental dynamical systems. We explain how to compute such invariants as the Euler characteristic and Betti numbers using time series data,

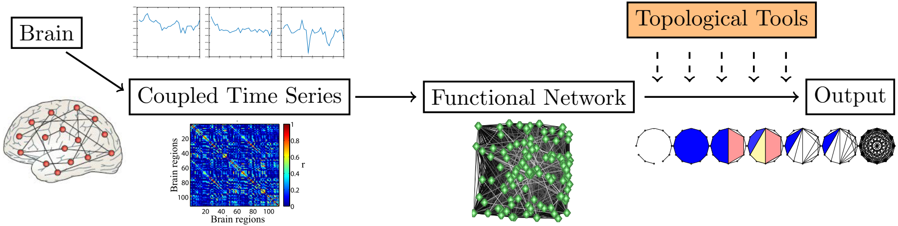
Fig. 13. The system of patches and 2000 randomly selected points plotted in the two-dimensional subspace obtained by projection onto the vectors pictured in fig. 12. Scaling curves for the patch with the dashed outline are shown in fig. 14.

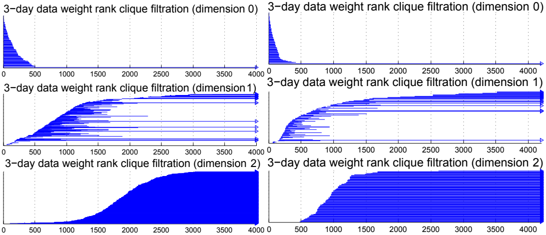

0
1
2
3
4
5
6
7
8
0-dim'l barcode (components)
0
1
2
3
4
5
6
7
8
1-dim'l barcode (loops)

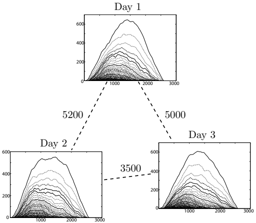
Topology for neuronal networks



Topology for neuronal networks

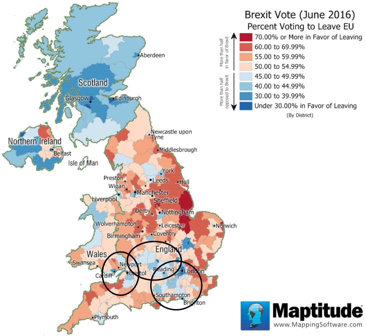


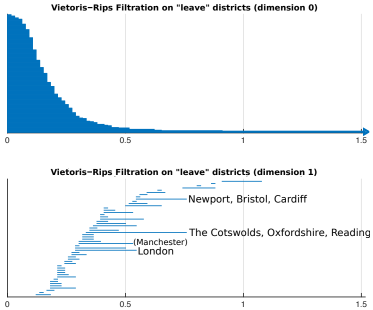




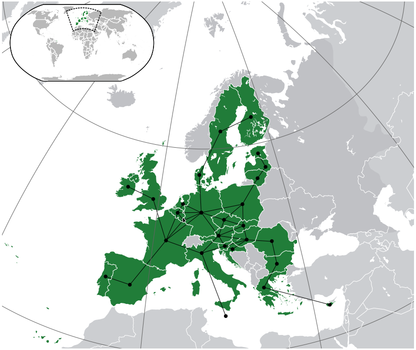
Bassett et al. (2011) *PNAS.*   B Stolz, HAH, MA Porter. arXiv:1605.00562

POINT CLOUD Topological data analysis of Brexit (UK)

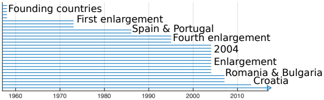




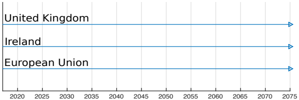
NETWORK Topological data analysis of Brexit (EU)



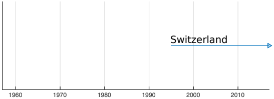
EU "pre-Brexit" weight rank clique filtration (dimension 0)



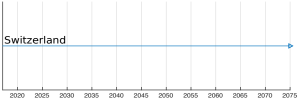
EU "post-Brexit" weight rank clique filtration (dimension 0)



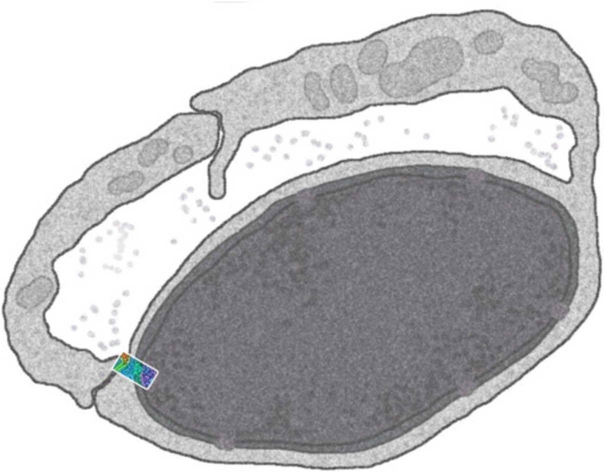
EU "pre-Brexit" weight rank clique filtration (dimension 1)



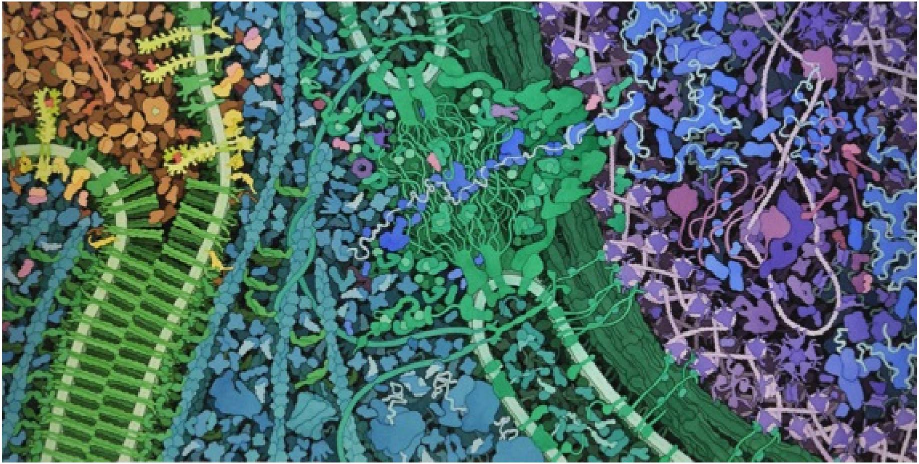
EU "post-Brexit" weight rank clique filtration (dimension 1)



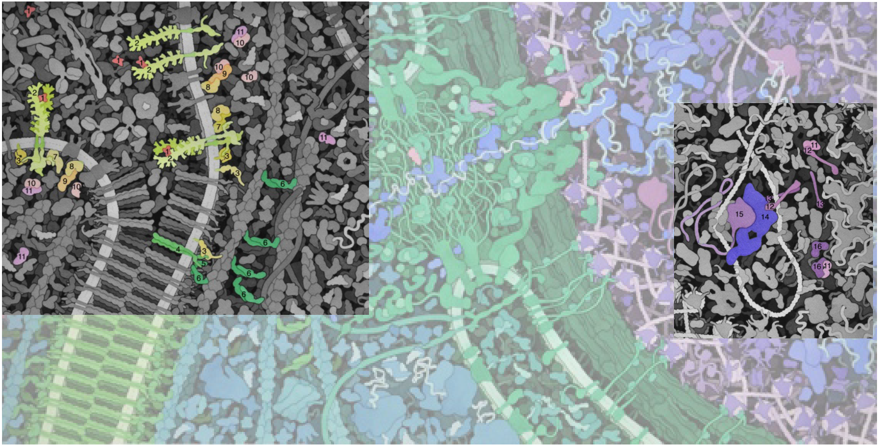
How do cells make decisions?



How do cells make decisions?

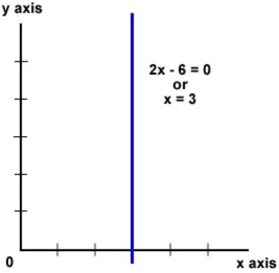


David Goodsell

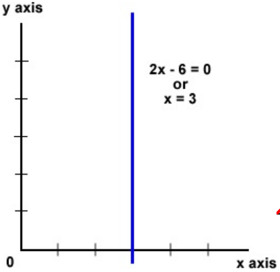


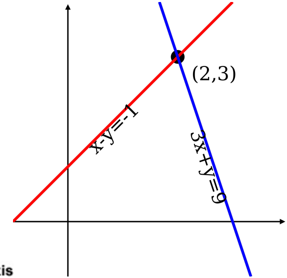
David Goodsell

Linear equations

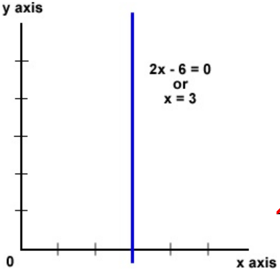


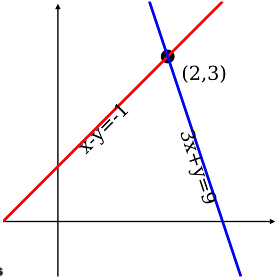
Linear equations

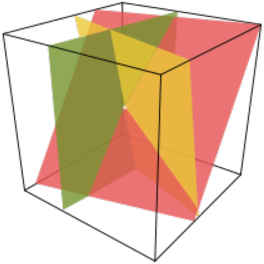




Linear equations





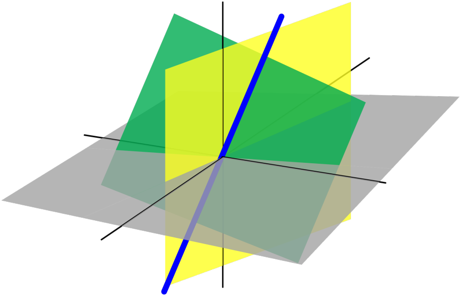


Linear equations

Suppose we are given 2 equations:
$x + y - z = 0$ and $2x + 3y + 2z = 0$.

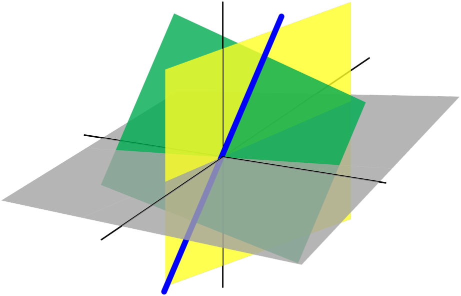
Linear equations

Suppose we are given 2 equations:
$x + y - z = 0$ and $2x + 3y + 2z = 0$.



Linear equations

Suppose we are given 2 equations:
$x + y - z = 0$ and $2x + 3y + 2z = 0$.



$$
\begin{array}{r}
2 \\
x + y - z \overline{)\, 2x + 3y + 2z} \\
\underline{2x + 2y - 2z} \\
y + 4z
\end{array}
$$

Now we can write

$$y + 4z = 0 \implies y = -4z$$

and substitute,

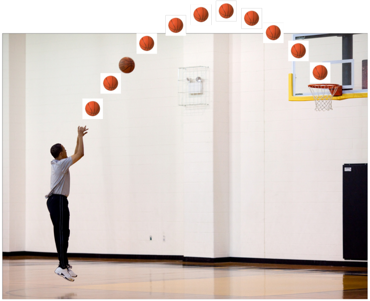
$$
\begin{aligned}
x + y - z &= 0 \\
x - 4z - z &= 0 \\
\implies x &= 5z
\end{aligned}
$$

Nonlinear equations

$$ax^2 + bx + c = 0$$

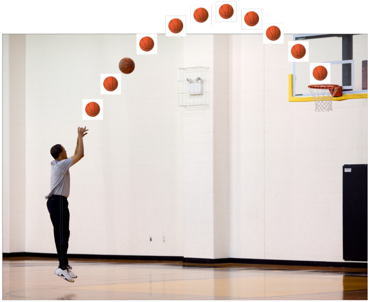
Nonlinear equations

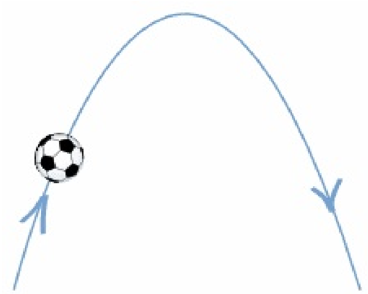
$$ax^2 + bx + c = 0$$



Nonlinear equations

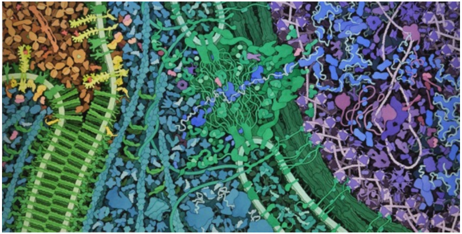
$$ax^2 + bx + c = 0$$



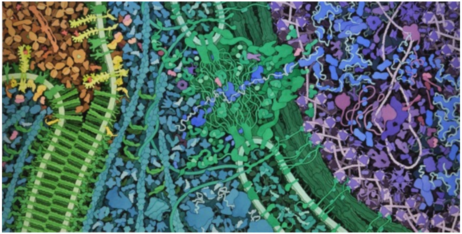


$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

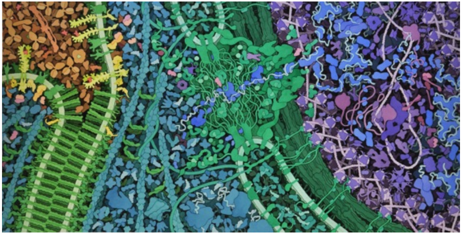
Nonlinear equations



Nonlinear equations



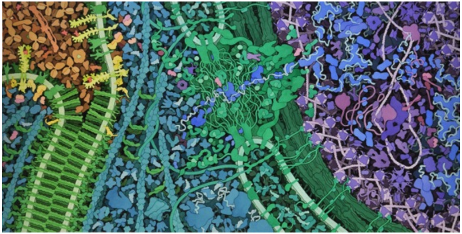
Nonlinear equations



Chemical reaction system

$$x + y \xrightarrow{a} 2x$$
$$3x \xrightarrow{b} y + 2z$$
$$z \xrightarrow{c} x \ , \ z \xrightarrow{d} y$$

Nonlinear equations



Chemical reaction system

$$x + y \xrightarrow{a} 2x$$

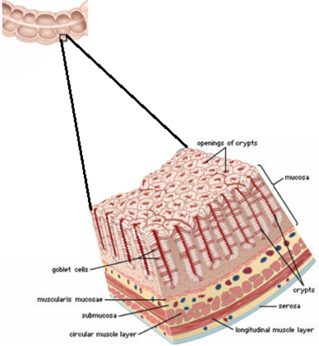
$$3x \xrightarrow{b} y + 2z$$

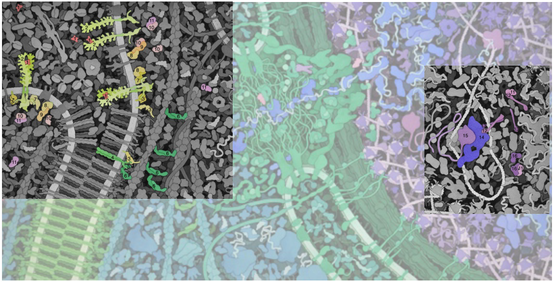
$$z \xrightarrow{c} x \ , \ z \xrightarrow{d} y$$

$$axy - 3bx^3 + cz = 0, \quad -axy + bx^3 + dz = 0, \quad 2bx^3 - cz - dz = 0$$

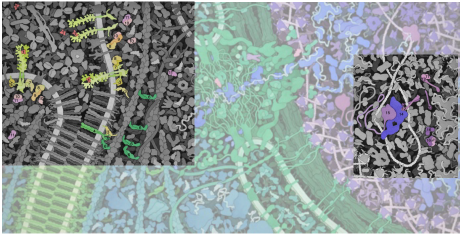
Mathematical models

	LINEAR	NON-LINEAR
One species (x)	EX1: $x - 4 = 0$	EX3: $ax^2 + bx + c = 0$
More species (x,y,z)	EX2: $x + y - z = 0$ $2x + 3y + 2z = 0$	EX4: $axy - 3bx^3 + cz = 0$ $-axy + bx^3 + dz = 0$ $2bx^3 - cz - dz = 0$





David Goodsell



$$x_1 \underset{k_2}{\overset{k_1}{\rightleftarrows}} x_2$$

$$x_2 + x_4 \underset{k_4}{\overset{k_3}{\rightleftarrows}} x_{14} \xrightarrow{k_5} x_2 + x_5$$

$$x_5 + x_8 \underset{k_7}{\overset{k_6}{\rightleftarrows}} x_{16} \xrightarrow{k_8} x_4 + x_8$$

$$x_4 + x_{10} \underset{k_{10}}{\overset{k_9}{\rightleftarrows}} x_{18} \xrightarrow{k_{11}} x_4 + \emptyset$$

$$\emptyset \xrightarrow{k_{12}} x_{10}$$

$$x_{10} \xrightarrow{k_{13}} \emptyset$$

$$x_3 + x_6 \underset{k_{15}}{\overset{k_{14}}{\rightleftarrows}} x_{15} \xrightarrow{k_{16}} x_3 + x_7$$

$$x_7 + x_9 \underset{k_{18}}{\overset{k_{17}}{\rightleftarrows}} x_{17} \xrightarrow{k_{19}} x_6 + x_9$$

$$x_6 + x_{11} \underset{k_{21}}{\overset{k_{20}}{\rightleftarrows}} x_{19} \xrightarrow{k_{22}} x_6 + \emptyset$$

$$x_{11} \xrightarrow{k_{23}} \emptyset$$

$$x_{11} + x_{12} \underset{k_{25}}{\overset{k_{24}}{\rightleftarrows}} x_{13}$$

$$x_2 \underset{k_{27}}{\overset{k_{26}}{\rightleftarrows}} x_3$$

$$x_5 \underset{k_{29}}{\overset{k_{28}}{\rightleftarrows}} x_7$$

$$x_{10} \underset{k_{31}}{\overset{k_{30}}{\rightleftarrows}} x_{11}$$

Models

$$x_1 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} x_2$$

$$x_2 + x_4 \underset{k_4}{\overset{k_3}{\rightleftharpoons}} x_{14} \xrightarrow{k_5} x_2 + x_5$$

$$x_5 + x_8 \underset{k_7}{\overset{k_6}{\rightleftharpoons}} x_{16} \xrightarrow{k_8} x_4 + x_8$$

$$x_4 + x_{10} \underset{k_{10}}{\overset{k_9}{\rightleftharpoons}} x_{18} \xrightarrow{k_{11}} x_4 + \emptyset$$

$$\emptyset \xrightarrow{k_{12}} x_{10}$$

$$x_{10} \xrightarrow{k_{13}} \emptyset$$

$$x_3 + x_6 \underset{k_{15}}{\overset{k_{14}}{\rightleftharpoons}} x_{15} \xrightarrow{k_{16}} x_3 + x_7$$

$$x_7 + x_9 \underset{k_{18}}{\overset{k_{17}}{\rightleftharpoons}} x_{17} \xrightarrow{k_{19}} x_6 + x_9$$

$$x_6 + x_{11} \underset{k_{21}}{\overset{k_{20}}{\rightleftharpoons}} x_{19} \xrightarrow{k_{22}} x_6 + \emptyset$$

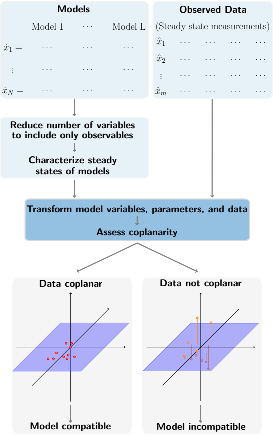
$$x_{11} \xrightarrow{k_{23}} \emptyset$$

$$x_{11} + x_{12} \underset{k_{25}}{\overset{k_{24}}{\rightleftharpoons}} x_{13}$$

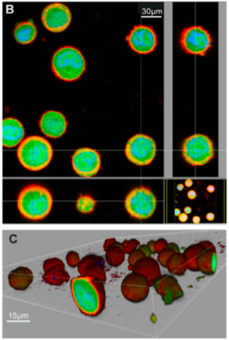
$$x_2 \underset{k_{27}}{\overset{k_{26}}{\rightleftharpoons}} x_3$$

$$x_5 \underset{k_{29}}{\overset{k_{28}}{\rightleftharpoons}} x_7$$

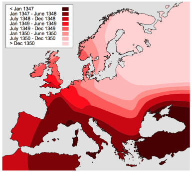
$$x_{10} \underset{k_{31}}{\overset{k_{30}}{\rightleftharpoons}} x_{11}$$

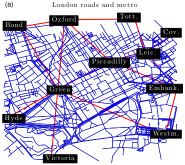


Data



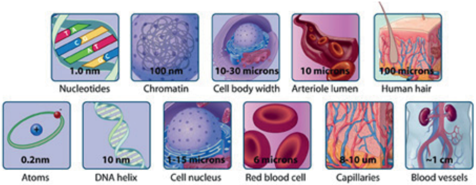
Computational algebra and topology is useful for biology!

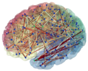


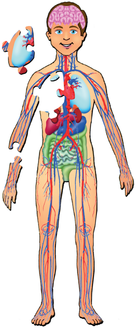


Contagions/Epidemic spread

Population







Molecules $\longleftarrow$ Organelles $\longleftarrow$ Cells $\longleftarrow$ Tissues $\longleftarrow$ Organs $\longleftarrow$ Organism