Math-for-industry
Education & Research Hub

ISSN2188-1200

International Project Research-Workshop (I)

WORKSHOP on Mathematics for Industry

Basis of Mathematics in nanomedicine structures and life sensing

Editors: Osamu Saeki, Wojciech Domitrz, Stanisław Janeczko, Marcin Zubilewicz, Michał Zwierzyński

九州大学マス·フォア·インダストリ研究所

MI Lecture Note Vol.95 : Kyushu University

International Project Research-Workshop (I)

WORKSHOP on Mathematics for Industry

Basis of Mathematics in nanomedicine structures and life sensing

Editors Osamu Saeki, Wojciech Domitrz, Stanisław Janeczko,
Marcin Zubilewicz, Michał Zwierzyński

About MI Lecture Note Series

The Math-for-Industry (MI) Lecture Note Series is the successor to the COE Lecture Notes, which were published for the 21st COE Program “Development of Dynamic Mathematics with High Functionality,” sponsored by Japan’s Ministry of Education, Culture, Sports, Science and Technology (MEXT) from 2003 to 2007. The MI Lecture Note Series has published the notes of lectures organized under the following two programs: “Training Program for Ph.D. and New Master’s Degree in Mathematics as Required by Industry,” adopted as a Support Program for Improving Graduate School Education by MEXT from 2007 to 2009; and “Education-and-Research Hub for Mathematics-for-Industry,” adopted as a Global COE Program by MEXT from 2008 to 2012.

In accordance with the establishment of the Institute of Mathematics for Industry (IMI) in April 2011 and the authorization of IMI’s Joint Research Center for Advanced and Fundamental Mathematics-for-Industry as a MEXT Joint Usage / Research Center in April 2013, hereafter the MI Lecture Notes Series will publish lecture notes and proceedings by worldwide researchers of MI to contribute to the development of MI.

October 2022
Kenji Kajiwara
Director, Institute of Mathematics for Industry

WORKSHOP on Mathematics for Industry
Basis of Mathematics in nanomedicine structures and life sensing

MI Lecture Note Vol.95, Institute of Mathematics for Industry, Kyushu University
ISSN 2188-1200
Date of issue: March 14, 2024
Editors: Osamu Saeki, Wojciech Domitrz, Stanisław Janeczko, Marcin Zubilewicz,
Michał Zwierzyński
Publisher:
Institute of Mathematics for Industry, Kyushu University
Graduate School of Mathematics, Kyushu University
Motooka 744, Nishi-ku, Fukuoka, 819-0395, JAPAN
Tel +81-(0)92-802-4402, Fax +81-(0)92-802-4405
URL https://www.imi.kyushu-u.ac.jp/

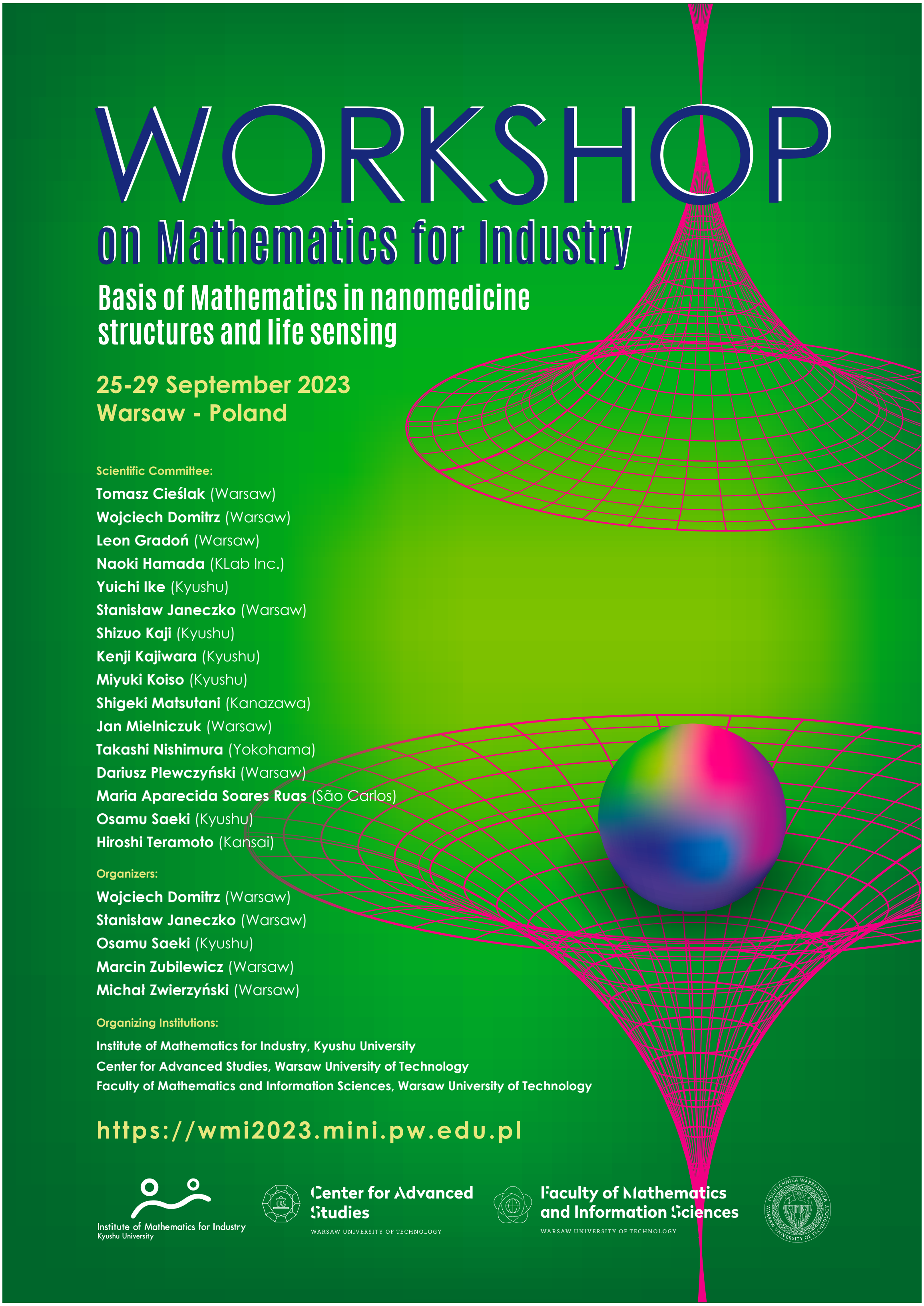
Preface

The "WORKSHOP on Mathematics for Industry 2023 – Basis of Mathematics in nanomedicine structures and life sensing" convened during September 25–29, 2023, at Warsaw University of Technology, Poland, under the joint auspices of the Faculty of Mathematics and Information Science, Warsaw University of Technology; Center for Advanced Studies, Warsaw University of Technology; and Institute of Mathematics for Industry, Kyushu University, with the support of the Excellence Initiative: Research University Programme at the Warsaw University of Technology. With the participation of approximately 70 attendees, including researchers, students, and PhD candidates, the workshop served as a nexus for interdisciplinary dialogue and collaboration between the realms of mathematics and applied sciences.

The workshop program encompassed 25 individual talks and 5 mini-courses, each comprising 3 lectures, spanning a spectrum of topics such as topological data analysis, medical imaging methods, human genome models, big data, machine learning, cryptography, information geometry, convex optimization, physical models of elastic/plastic bodies and fluids and material engineering. Delivered by experts from Polish and Japanese institutions, the presentations illuminated the symbiotic relationship between abstract mathematical constructs and real-world engineering challenges, thereby fostering innovation and knowledge exchange. The accompanying booklet contains comprehensive materials from the workshop prepared by the speakers, including detailed summaries, presentation slides and references, providing a valuable resource for continued study of the concepts presented during the event, with hope that it will not only facilitate the exploration of novel research directions, but also catalyze the establishment of international collaborations between academic environments in Poland and Japan with the goal of leveraging mathematical methodologies to address pressing industrial concerns and societal needs.

This work was supported by Institute of Mathematics for Industry, Joint Usage/Research Center in Kyushu University (FY2023 Workshop(I) "WORKSHOP on Mathematics for Industry 2023 – Basis of Mathematics in nanomedicine structures and life sensing" (2023b004)).

February 2024


WORKSHOP
on Mathematics for Industry
Basis of Mathematics in nanomedicine structures and life sensing
25-29 September 2023
Warsaw - Poland
Scientific Committee:
Tomasz Cieślak (Warsaw)
Wojciech Domitrz (Warsaw)
Leon Gradoń (Warsaw)
Naoki Hamada (KLab Inc.)
Yuichi Ike (Kyushu)
Stanisław Janeczko (Warsaw)
Shizuo Kaji (Kyushu)
Kenji Kajiwara (Kyushu)
Miyuki Koiso (Kyushu)
Shigeki Matsutani (Kanazawa)
Jan Mielniczuk (Warsaw)
Takashi Nishimura (Yokohama)
Dariusz Plewczyński (Warsaw)
Maria Aparecida Soares Ruas (São Carlos)
Osamu Saeki (Kyushu)
Hiroshi Teramoto (Kansai)
Organizers:
Wojciech Domitrz (Warsaw)
Stanisław Janeczko (Warsaw)
Osamu Saeki (Kyushu)
Marcin Zubilewicz (Warsaw)
Michał Zwierzyński (Warsaw)
Organizing Institutions:
Institute of Mathematics for Industry, Kyushu University
Center for Advanced Studies, Warsaw University of Technology
Faculty of Mathematics and Information Sciences, Warsaw University of Technology
https://wmi2023.mini.pw.edu.pl
Institute of Mathematics for Industry
Kyushu University
Center for Advanced Studies
WARSAW UNIVERSITY OF TECHNOLOGY
Faculty of Mathematics and Information Sciences
WARSAW UNIVERSITY OF TECHNOLOGY


Centrum Studiów Zaawansowanych
Politechniki Warszawskiej
Centrum Studiów Zaawansowanych
Politechniki Warszawskiej

WORKSHOP on Mathematics for Industry 2023

Programme

Mini courses (green) Registration (red)

	Monday (25.09)	Tuesday (26.09)	Wednesday (27.09)	Thursday (28.09)	Friday (29.09)
8:00 – 9:00	**Registration (up to 11:00)**	**8:30 – 9:00** *Lecture: Paweł Józiak*	**8:15 – 9:00** *Mini course: Shunsuke Ichiki 2*	**8:30 – 9:00** *Lecture: Piotr Borowik*	*Lecture: Kenji Kajiwara*
9:00 – 9:10	**Opening of the Workshop**				
9:15 – 10:00	*Mini course: Jan Mielniczuk 1*	*Mini course: Jan Mielniczuk 2*	*Mini course: Dariusz Plewczyński 2*	*Mini course: Jan Mielniczuk 3*	*Mini course: Shunsuke Ichiki 3*
10:15 – 11:00	*Mini course: Shunsuke Ichiki 1*	*Mini course: Dariusz Plewczyński 1*	*Mini course: Paweł Dłotko 3*	*Mini course: Arimura Hidetaka 1*	*Mini course: Dariusz Plewczyński 3*
11:00 – 11:30	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break
11:30 – 12:15	*Mini course: Paweł Dłotko 1*	*Mini course: Paweł Dłotko 2*	*Lecture: Yuichi Ike*	*Mini course: Arimura Hidetaka 2*	*Mini course: Arimura Hidetaka 3*
12:30 – 13:00	*Lecture: Naoki Hamada*	*Lecture: Tomasz Cieślak*	*Lecture: Przemysław Grzegorzewski*	*Lecture: Shigeki Matsutani*	*Lecture: Hiroshi Teramoto*
13:00 – 15:00	Lunch	Lunch	Lunch	Lunch	Lunch
15:00 – 15:30	*Lecture: Przemysław Biecek*	*Lecture: Leon Gradoń*		*Lecture: Zbigniew Peradzyński*	*Lecture: Osamu Saeki*
15:45 – 16:15	*Lecture: Mariusz Niewęgłowski*	*Lecture: Karol Ćwieka*		*Lecture: Konrad Kisiel*	*Lecture: Naomichi Nakajima*
16:15 – 16:45	Coffee break	Coffee break		Coffee break	Coffee break
16:45 – 17:15	*Lecture: Lucía Ivonne Hernández Martínez*	*Lecture: Toshizumi Fukui*		*Lecture: Shizuo Kaji*	*Lecture: Bartosz Kołodziejek*
17:30 – 18:00	*Lecture: Stanisław Janeczko*	*Lecture: Miyuki Koiso*		*Lecture: Takashi Nishimura*	*Lecture: Konstanty Junosza-Szaniawski*
18:00 – 22:00	**Dinner @ MaIS Faculty**			**Dinner @ MaIS Faculty**	

Contents

Introduction to Topological Data Analysis

Paweł Dłotko
Dioscuri Centre in Topological Data Analysis, IMPAN, Poland

In this mini-course we will explore both theoretical and practical foundations of Topological Data Analysis (TDA) — a field with a number of applications in physical, natural and social sciences in the intersection between algebraic topology, computational geometry and computational methods. We will cover the basic tools of TDA including discretization of spaces (in the form of various point cloud-based simplicial, cubical and general CW-complexes), algorithms to compute homology and persistent homology and applications of those. We will also explore TDA tools of visualization, like mapper and ball mapper algorithms. Moreover we will present new tools of Euler Characteristic curves and profiles and show how they can be applied to standard statistics. All the concepts will be illustrated with real examples. You will also be required to perform computations on a number of toy and real-world datasets.

References

[1] Edelsbrunner, Harrer (2011), *Computational Topology: An Introduction*
[2] P. Dłotko, *Computational and applied topology, tutorial*, `https://arxiv.org/abs/1807.08607`

Introduction to Topological Data Analysis

Paweł Dłotko, Dioscuri Centre in TDA, IMPAN,

WORKSHOP on Mathematics for Industry 2023

Politechnika Warszawska, MINI, 25-27 September 2023

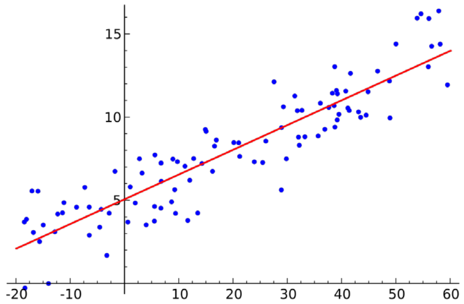
Topological Data Analysis

- Persistent homology,
- Conventional mapper,
- Ball mapper,
- Discrete Morse theory (if time permits),
- TopoTests (alternative option),
- On a very intuitive level,
- with a number of practical examples.

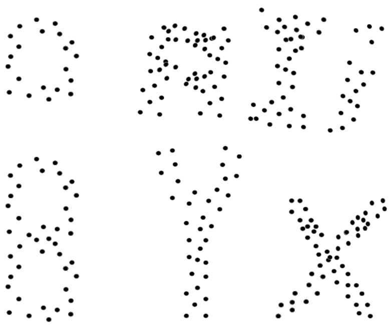
The credo

Data have shape,
shape has meaning,
meaning brings value.

We all know this story



Trap of models

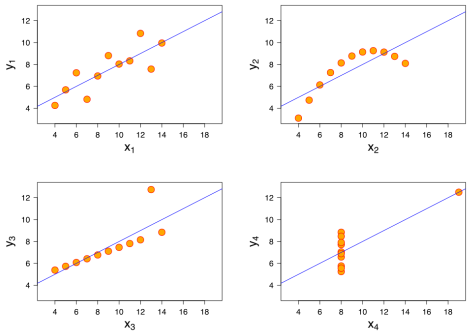


It is not possible to adjust an algebraic model to any possible shape of the data – over-fitting.

Topology and statistics, together

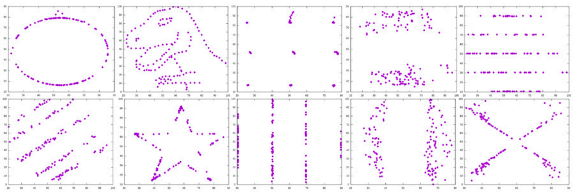
- Statistics provide a vast collection of tools to summarize properties of point clouds.
- However, there are numerous examples (line Anscombe's quartet and Datasaurus dataset presented below) of point clouds with the same descriptive statistics, but very different shape.
- This is why, in statistics, we should always *visualize* the considered dataset.
- It is however not possible to visualize high dimensional data.
- That is where the tools from topology came into rescue – topological tools we discuss in this tutorial allow us to estimate if two datasets have similar shape.

Anscombe's Quartet



Anscombe's Quartet; Same statistics, different shapes
Anscombe, "Graphs in Statistical Analysis", American Statistician, 1973.

Datasaurus Dataset

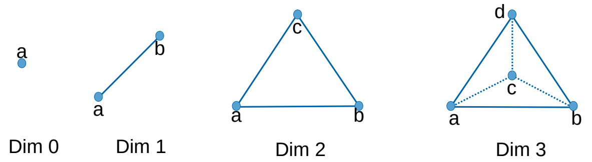


Datazaurus Dozen, Alberto Cairo,
`http://www.thefunctionalart.com/2016/08/download-datasaurus-never-trust-summary.html`

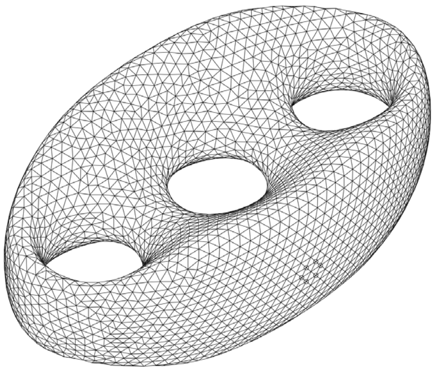
TDA pipeline

Simplicial complexes

- $\mathcal{K}$ is an abstract simplicial complex iff for every $A \in \mathcal{K}$ and $B \subset A$, $B \in \mathcal{K}$.
- Each abstract simplicial complex has its geometrical realization built from simplices.
- In this case, simplices consist of points in a general position.



Sample simplicial complexes



Source: Wikipedia, typical use FEM-like methods.

Let the data tell you the story

Topological data analysis:

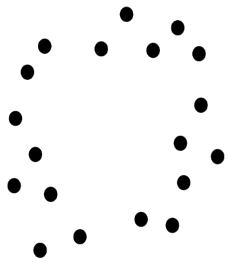
- Persistent homology – point-cloud based homology.
- Accurate network models to examine landscapes of data,

! Stable.

!! No black boxes.

!!! We do not enforce *any* models of data.

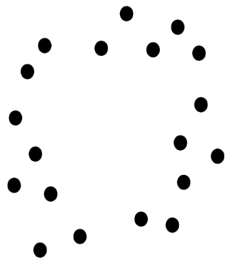
What do you see?



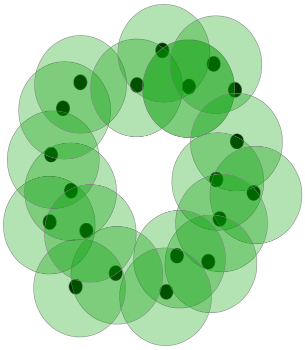
What do you see?

- We may say that we see a circle,
- But we really see is 19 points...
- ...that may be sampled from a probability distribution supported at a circle.
- Persistent homology is a tool to make this observation precise.
- To do so, we need to construct a *filtered complex* of the point cloud.
- The filtered complex is a nested sequence of subcomplexes - a way of building a model by adding a sequence of simplices in a number of steps.

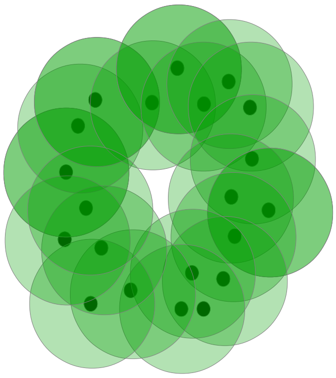
What do you see?



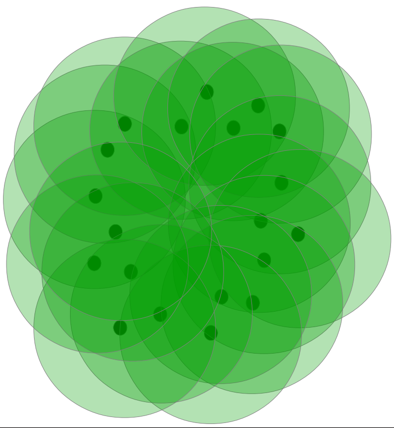
What do you see?



What do you see?



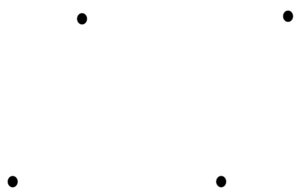
What do you see?



Simplicial complexes built from point clouds

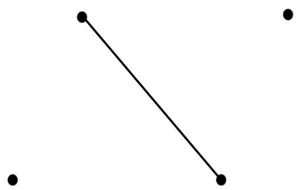
- $P = \{p_1, \dots, p_n\}$, a finite point cloud with a metric d.
- We need a finite, combinatorial representation of the union of balls.
- Rips complex at level ϵ consists of simplices supported in $p_0, \dots, p_n$ if $B(p_i, \frac{\epsilon}{2}) \cap B(p_j, \frac{\epsilon}{2}) \neq \emptyset$ for every $i, j \in \{0, \dots, n\}$.
- Čech complex at level ϵ consists of simplices supported in $p_0, \dots, p_n$ iff $\bigcap_{i=0}^{n} B(p_i, \frac{\epsilon}{2}) \neq \emptyset$.

Filtration of Rips complex



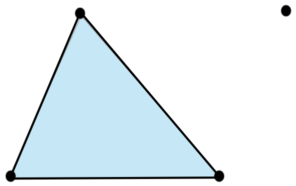
4 vertices

Filtration of Rips complex



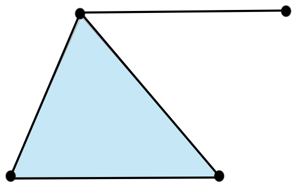
4 vertices, 1 edge

Filtration of Rips complex



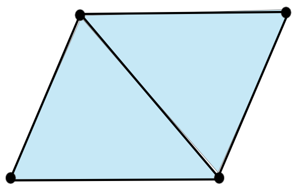
4 vertices, 3 edges, 1 triangle

Filtration of Rips complex



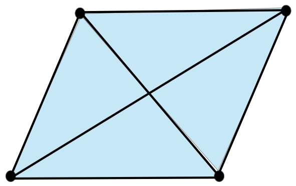
4 vertices, 4 edges, 1 triangle

Filtration of Rips complex



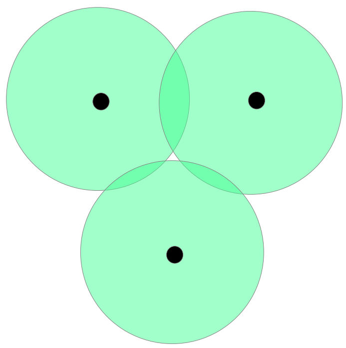
4 vertices, 5 edges, 2 triangles

Filtration of Rips complex

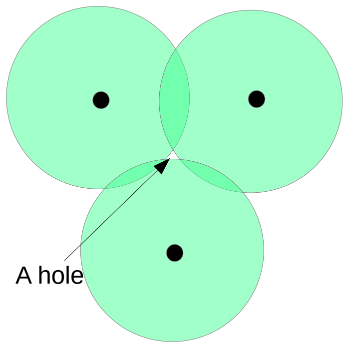


4 vertices, 6 edges, 4 triangles, 1 tetrahedra

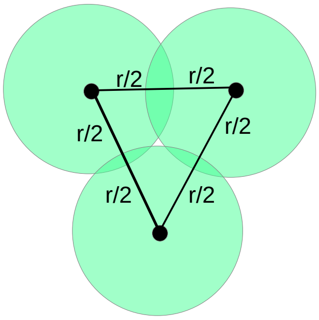
Rips vs Čech



Rips vs Čech



Rips vs Čech

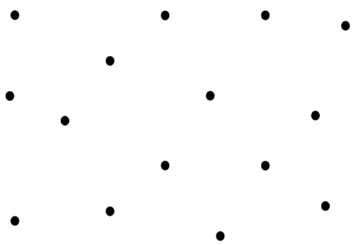


In this case Rips complex is a triangle with a boundary, the Čech complex is the boundary of a triangle

Čech complex is topologically accurate

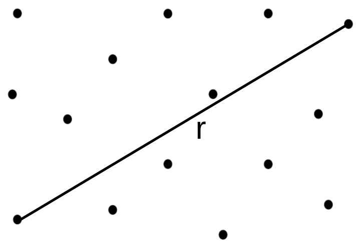
- $\bigcup_{p \in P} B(p, \frac{\epsilon}{2})$ is topologically equivalent to the Čech complex based on those balls.
- Meaning, there exist a continuous deformation from one into another.
- No tearing, no gluing.

Rips and Čech complexes can grow large



If all points get connected by edges in the complex, we witness so-called *combinatorial explosion*. You will encounter it when using Rips complexes.

Rips and Čech complexes can grow large



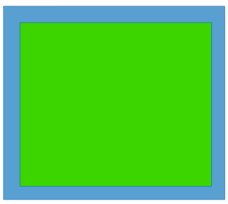
For N points, $\binom{N}{1}$ vertices, $\binom{N}{2}$ edges, $\binom{N}{3}$ triangles, ...
This is why we always limit the level (ϵ) and the maximal dimension of simplices in the complex.

Alpha complexes

Intersecting $B(x, r)$, for $x \in X$ with Voronoi cells of X allows to build much smaller complexes that preserve homotopy type of $\bigcup_{x \in X} B(x, r)$.

Be careful with distances (in high dimensions)

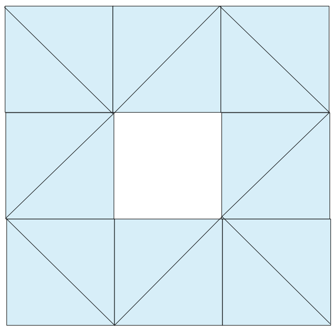
1. Concentration of measure $(1 - 2\epsilon)^n$,
2. Points in dimension d close to be of the same distance $\frac{d}{3}$ from each other in l^1 distance,
3. Manifold hypothesis.



From complexes to parameter dependent homology

Simplicial complex → Chains, Cycles, Boundaries → Homology groups

Homology

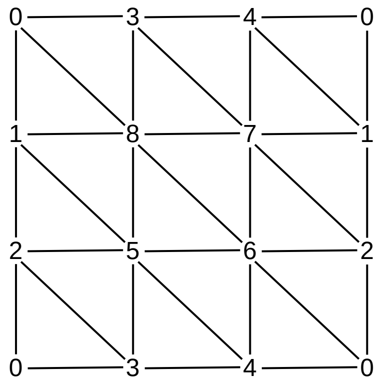


One connected component, one hole in dimension 1.

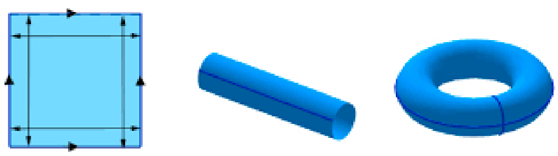
Practical exercise 1

- Please go to `https://github.com/dioscuri-tda/tutorials`,
- Open PH_intro_to_homology and play with triangulation of a torus.
- What are the homology groups of this triangulation?

Triangulation of a torus



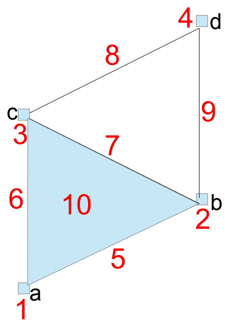
Triangulation of a torus



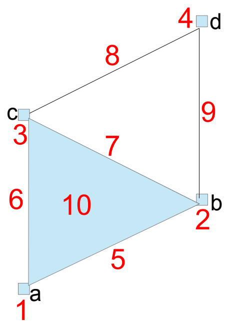
Persistent homology, under the hood

- Let us order simplices according to the minimal ϵ for which they appear (filtration).
- Algorithm to compute (persistent) homology is a version of Gaussian elimination.
- If we run it for a prefix of filtration, we will get homology of the complex composed by simplices in that prefix (a subcomplex of the final complex).
- Analyzing the structure of zero and non-zero columns in the reduced matrix allows us to find generators that are created and which become trivial as we move along the filtration.

Persistence matrix algorithm

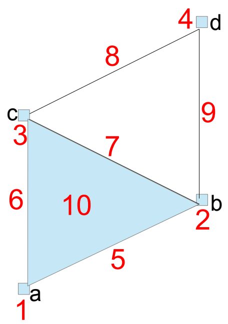


Persistence matrix algorithm



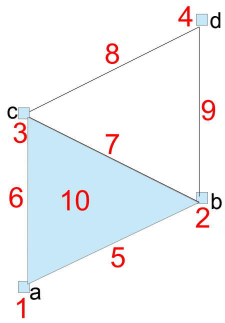
	ab	ac	bc	cd	bd	abc
a	1	1				
b	1		1		1	
c		1	1	1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



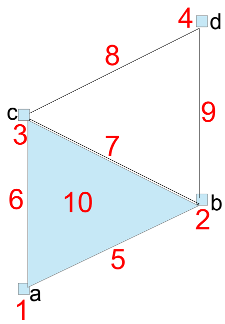
	ab	ac	bc	cd	bd	abc
a	1	1				
b	1		1		1	
c		1	1	1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



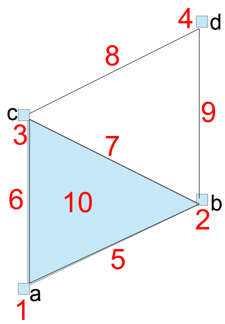
	ab	ac	bc	cd	bd	abc
a	1	1				
b	1		1		1	
c		1	1	1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



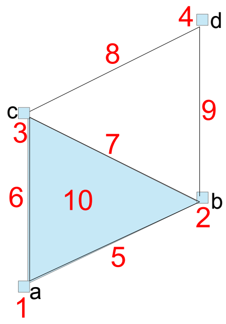
	ab	ac	bc	cd	bd	abc
a	1	1				
b	1		1		1	
c		1	1	1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



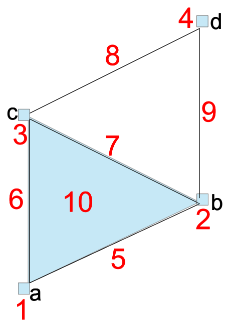
	ab	ac	bc+ ac	cd	bd	abc
a	1	1	1			
b	1		1		1	
c		1		1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



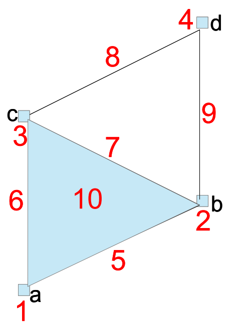
	ab	ac	bc+ ac+ ab	cd	bd	abc
a	1	1				
b	1				1	
c		1		1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



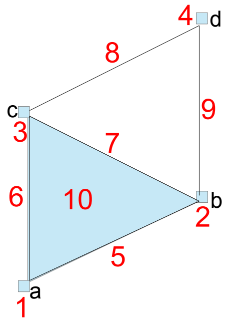
	ab	ac	bc+ ac+ ab	cd	bd	abc
a	1	1				
b	1				1	
c		1		1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



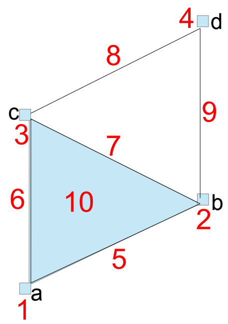
	ab	ac	bc+ ac+ ab	cd	bd	abc
a	1	1				
b	1				1	
c		1		1		
d				1	1	
ab						1
ac						1
bc						1

Persistence matrix algorithm



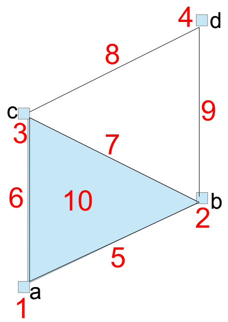
	ab	ac	bc+ac+ab	cd	bd+cd	abc
a	1	1				
b	1				1	
c		1		1	1	
d				1		
ab						1
ac						1
bc						1

Persistence matrix algorithm



	ab	ac	bc+ac+ab	cd	bd+cd+ac	abc
a	1	1			1	
b	1				1	
c		1		1		
d				1		
ab						1
ac						1
bc						1

Persistence matrix algorithm

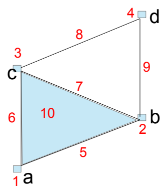


	ab	ac	bc+ac+ab	cd	bd+cd+ac+ab	abc
a	1	1				
b	1					
c		1		1		
d				1		
ab						1
ac						1
bc						1

Interpretation of reduced matrix

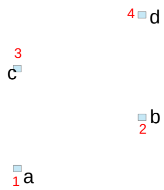
1. The reduced matrix gives the persistence intervals.
2. If the column is zero, then it creates a new homology class.
3. If the column is nonzero, then it kills a homology class.

Persistence matrix algorithm



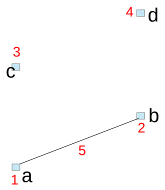
	a	b	c	d	ab	ac	bc+ ac+ ab	cd	bd+ cd+ ac+ ab	abc
a					1	1				
b					1					
c						1		1		
d								1		
ab										1
ac										1
bc										1

Persistence matrix algorithm



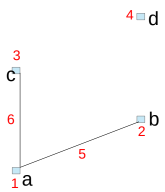
	a	b	c	d						
a										
b										
c										
d										
ab										
ac										
bc										

Persistence matrix algorithm



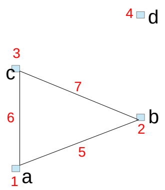
	a	b	c	d	ab					
a					1					
b					1					
c										
d										
ab										
ac										
bc										

Persistence matrix algorithm



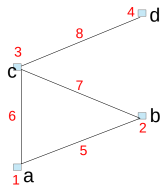
	a	b	c	d	ab	ac				
a					1	1				
b					1					
c						1				
d										
ab										
ac										
bc										

Persistence matrix algorithm



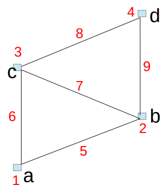
	a	b	c	d	ab	ac	bc+ ac+ ab			
a					1	1				
b					1					
c						1				
d										
ab										
ac										
bc										

Persistence matrix algorithm



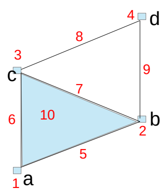
	a	b	c	d	ab	ac	bc+ ac+ ab	cd		
a					1	1				
b					1					
c						1		1		
d								1		
ab										
ac										
bc										

Persistence matrix algorithm



	a	b	c	d	ab	ac	bc+ ac+ ab	cd	bd+ cd+ ac+ ab	
a					1	1				
b					1					
c						1		1		
d								1		
ab										
ac										
bc										

Persistence matrix algorithm

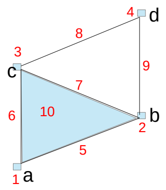


	a	b	c	d	ab	ac	bc+ ac+ ab	cd	bd+ cd+ ac+ ab	abc
a					1	1				
b					1					
c						1		1		
d								1		
ab										1
ac										1
bc										1

Dim 0: [2,5], [3,6], [4,8]

Dim 1: [7.10]

Persistence matrix algorithm



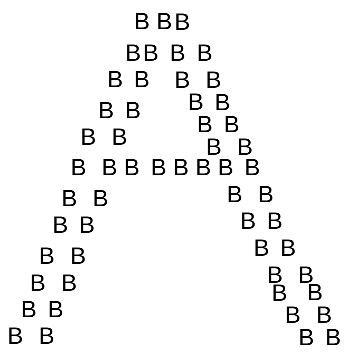
	a	b	c	d	ab	ac	bc+ ac+ ab	cd	bd+ cd+ ac+ ab	abc
a					1	1				
b					1					
c						1		1		
d								1		
ab										1
ac										1
bc										1

Dim 0: [2,5], [3,6], [4,8], [1,inf]   Dim 1: [7.10], [9,inf]

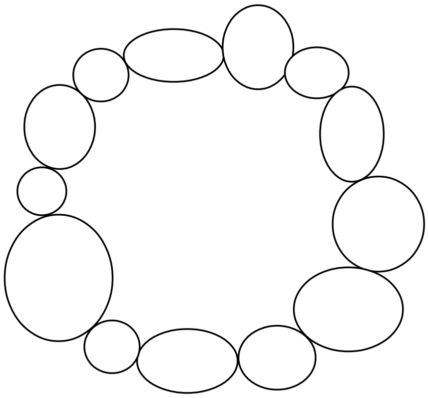
Invariance

- Persistent homology is a rigorous way of quantifying closed *shapes*,
- ... like connected components, cycles, voids and more.
- No matter if they are embedded in two or a million dimensional space,
- No matter if they are rotated, stretched or transformed in any other way.
-
-

Lots of **B**, or a single **A**?



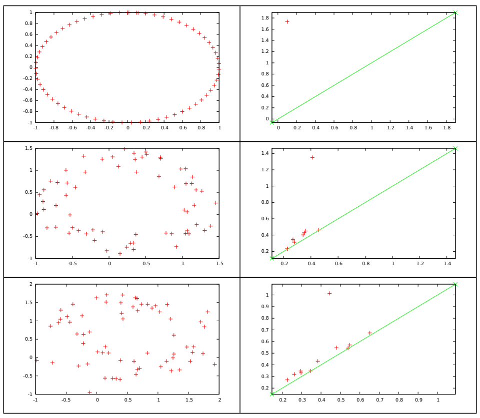
Lots of small circles, or a large one?



Multiscale

- Persistent homology is a rigorous way of quantifying closed *shapes*,
- ... like connected components, cycles, voids and more.
- No matter if they are embedded in two or a million dimensional space,
- No matter if they are rotated, stretched or transformed in any other way.
- Multi-scale,
-

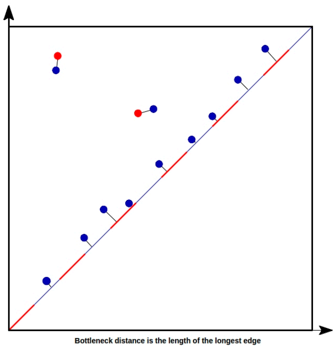
Robustness



Robustness

- Persistent homology is a rigorous way of quantifying closed *shapes*,
- ... like connected components, cycles, voids and more.
- No matter if they are embedded in two or a million dimensional space,
- No matter if they are rotated, stretched or transformed in any other way.
- Multi-scale,
- Robust.

Distances between diagrams

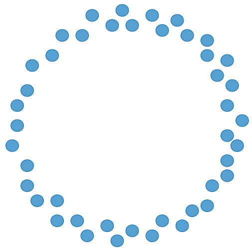


Optimal matchings between points of two persistence diagrams allow us to define standard distances between them – bottleneck (length of the longest edge in the matching) and p-Wasserstein (sum of lengths of matching lines to the power q) to the power $\frac{1}{q}$.

Practical exercise 2

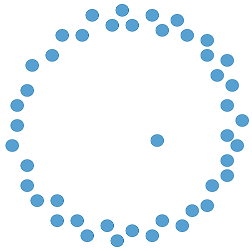
- Let us go back to our jupyter-notebooks exercises.
- Open PH_persistence_simple_point_cloud,
- Compute persistent homology of a point cloud sampled from a circle (without and with a considerable amount of noise).

Warning, outliers!



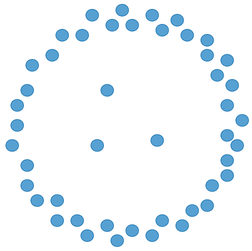
Outlayers can be a problem, filtration weighted by a distance to measure estimators

Warning, outliers!



outliers can be a problem, filtration weighted by a distance to measure estimators

Warning, outlayers!

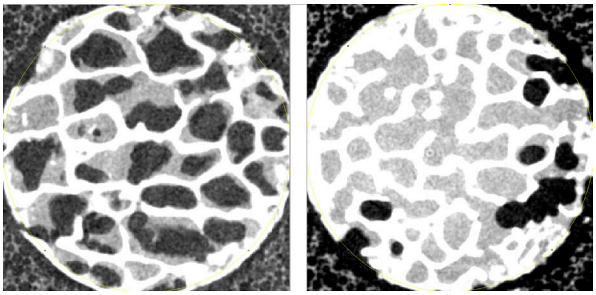


outliers can be a problem, filtration weighted by a distance to measure estimators

Not only point clouds....

- If you work with:
  - Pixel / voxel / cubical data,
  - Time series,
  - Correlation and similarity measures,
  - ...
- you may still use similar ideas and track connected components and holes emerging and disappearing.

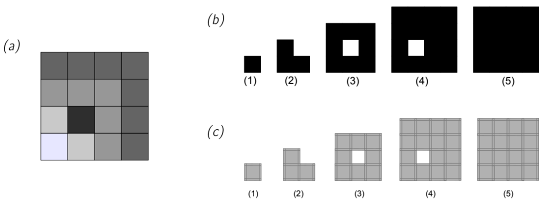
Apply to digital images



Left – osteoporotic, right – normal bone (vertebrae).
Not only density, but mostly structure is responsible for osteoporotic fractures.

What is a cubical persistence?

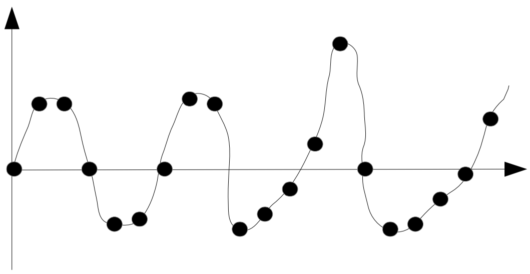
- Sub-level sets of a function.
- Cubes enter from lower to highest function/filtration value.
- We track changes in homology of sub-level sets.



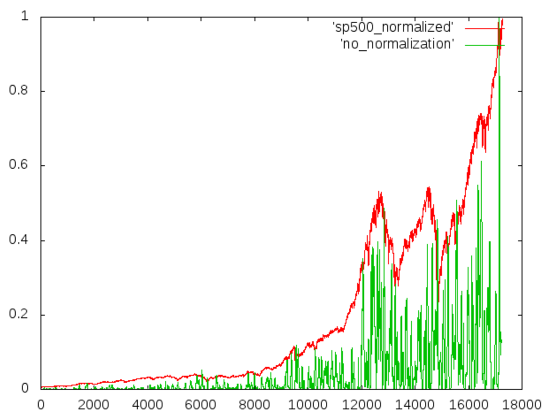
Practical exercise 3

- Digital images are partially-constant discretization of functions.
- Let us go back to our exercises.
- Open PH_distance_from_circle,
- In this exercise we will construct a cubical approximation of a function $f : [-2,2]^2 \to \mathbb{R}$. $f(x,y)$ is a distance from (x,y) to a unit circle $x^2 + y^2 = 1$.
- Let us visualize it as an image, and let us compute persistent homology of the corresponding cubical complex.

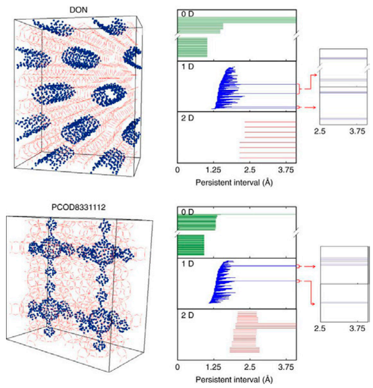
Persistence for time series analysis



S&P-500 and crashes



Persistence–based descriptors of nanoporous materials



Lee, Bathel, Dłotko, Mossavit, Smit, Hess, Quantifying similarity of pore-geometry in nanoporous materials, Nature Communications, 15396

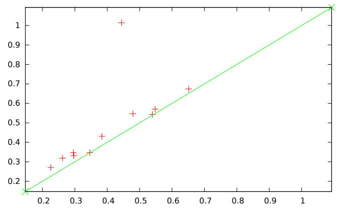
And more...

- We do not have time to cover all this ground.
- But, there are numerous resources for further work:
  - https://arxiv.org/abs/1807.08607
  - https://www.maths.ed.ac.uk/~v1ranick/papers/edelcomp.pdf
  - https://gudhi.inria.fr/tutorials/
  - and many more...

Persistent homology

- We have robust,
- multi scale,
- coordinate–free,
- compressed,
- tool to detect connected components, cycles, voids and their generalizations.
- It can be interfaced in various ways with standard stat. and ML tools.

Persistent homology, the output

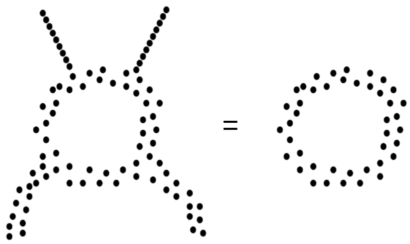


- Muti set of points in $\mathbb{R}^2$.
- Variable size, not ideal representation to interface with ML/AI and statistics $\rightarrow$ persistence representations, embeddings, ...
- We need to embed persistence diagrams into a Hilbert space (vectorize them).
- That makes topological/statistical inference - hypothesis testing, confidence intervals,... possible.

Homology and persistent homology, biased collection of resources

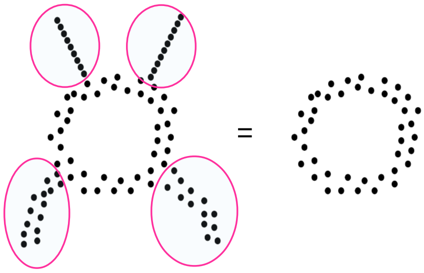
- Edelsbrunner and Harer, Computational Topology, An Introduction, AMS.
- Kaczynski, Mischaikow, Mrozek, Computational Topology, Springer 2003.
- Dłotko, Applied and Computational Topology, Tutorial
- Multiple youtube videos.

Persistence is nice, but, what about flares?



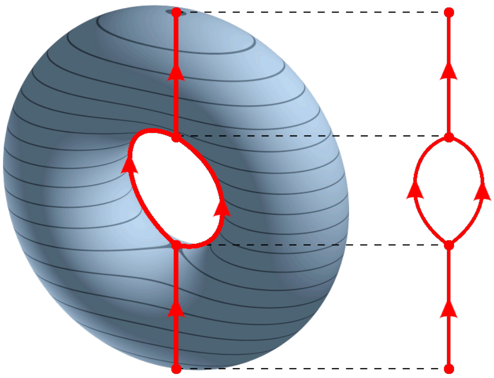
Persistence homology of those two point clouds will be very similar, as they both have one connected component and one hole.

But, what about flares?



But, oftentimes the information in the *flares* may be important (it may for instance carry information about anomalies).

Reeb graph

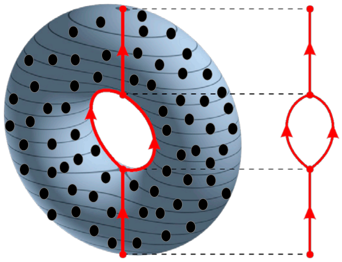


source: Wikipedia

Reeb graph, formally

- Input: M, $f : M \to \mathbb{R}$.
- We define an equivalence relation $x \mathrel{R} y$ iff:
  - $f(x) = f(y)$,
  - x and y belong to the same connected component of $f^{-1}(x)$.
- $M/_{R}$.

Conventional Mapper algorithm

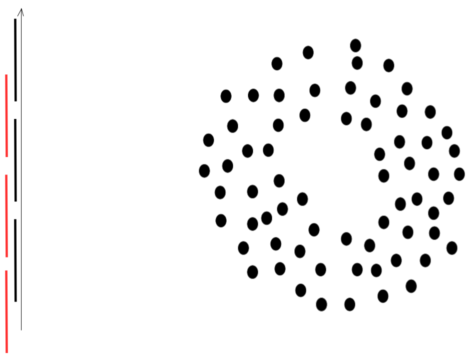


Conventional mapper graph is an attempt to define Reeb graph for discrete point cloud instead of a manifold.

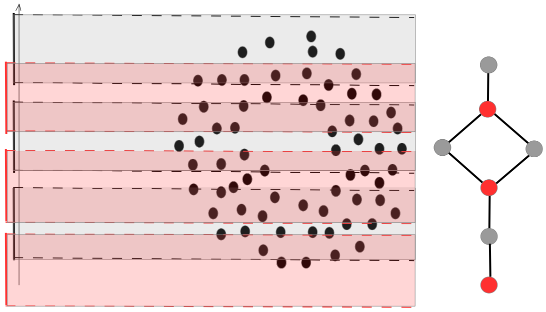
Mapper algorithm, idea

- Input: finite collection of points sampled from M, $f : M \to \mathbb{R}$.
- We define a relation x R y iff:
  - $f(x)$ is close to $f(y)$,
  - x and y belong to the same cluster ...

Conventional Mapper algorithm



Conventional Mapper algorithm



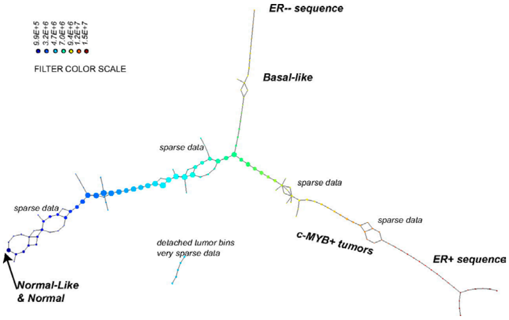
Mapper algorithm, formally

- Input: finite collection of points sampled from M, $f : M \to \mathbb{R}$.
- Cover of the range of f with overlapping boxes.
- Fix a clustering algorithm
- We define a relation x R y iff:
  - $f(x)$ and $f(y)$ belong to the same element I of a cover of the range of f,
  - x and y belong to the same cluster in $f^{-1}(I)$.
- Vertices of Mapper graph corresponds to the clusters,
- An edge is placed between two vertices if the corresponding clusters have nonempty intersection.

Mapper algorithm, coloring

- Vertices of the Mapper graph may be colored by an average value of an objective function on points covered by clusters.
- Fix a point cloud X and an objective function $f : X \to \mathbb{R}$.
- Each vertex of the Mapper graph correspond a subset (cluster) of points from X.
- Typically the value of the vertex will be an average value of f on the corresponding cluster.

Mapper is the most well known tool of TDA



Nicolau, Levine, Carlsson, *Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival*, PNAS 2011.

Practical exercise 1

- Let us play with Mapper algorithm!
- Go to `https://github.com/dioscuri-tda/tutorials`
- Let us start from something simple – open Mapper_concentric_circles
- In this exercise we will generate two concentric circles in a plane.
- We will use projection to the y coordinate as a lens function,
- And a DBSCAN with certain parameters as a clustering algorithm.
- What is the Mapper graph we obtain?

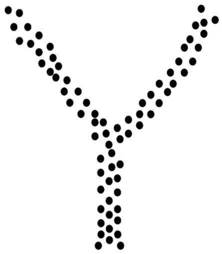
Practical exercise 2

- Let us play with something more advanced, let us consider standard Boston property dataset.
- Please open Mapper_boston_dataset
- It contains 13 variables, we want to understand its relation to prices of properties in Boston area (in '1970).
- Here we will use t-distributed stochastic neighbor embedding as a filtering function.
- We will be able to experiment with numerous clustering methods as well.
- Obtained mapper graphs will be colored by the average price of a property in a given cluster.
- This is not the last time we see Boston Property Dataset!

Ball Mapper algorithm

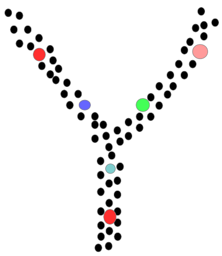
- As the last part of our schedule, we will play with Ball Mapper algorithm.
- As you might have noticed, it is not always trivial to choose the *lens function* as well as *clustering algorithm* in standard Mapper construction.
- The idea of Ball Mapper is intuitively explained in the following slides.

Ball Mapper algorithm



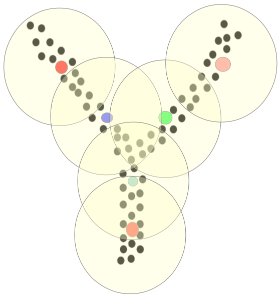
Take a point cloud X

Ball Mapper algorithm



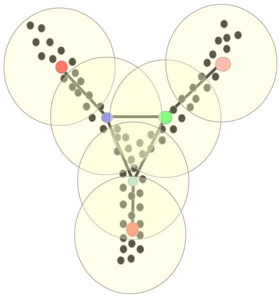
Given $\epsilon > 0$, select subset of points $N \subset X$ such that for every $x \in X$ there exists $n \in N$ such that $d(x, n) \leq \epsilon$ (we call N an ϵ-net)

Ball Mapper algorithm



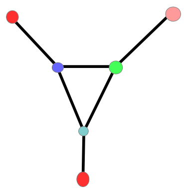
Consequently $X \subset \bigcup_{n \in N} B(n, \epsilon)$, i.e. $\{B(n, \epsilon), n \in N\}$ cover X.

Ball Mapper algorithm



Take one dimensional nerve of that cover (an abstract graph whose vertices correspond to $B(n, \epsilon)$, and edges to nonempty intersections of balls)

Ball Mapper algorithm

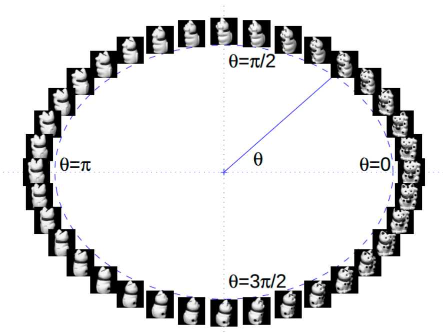


This way we obtain a Ball Mapper graph of X with radius ϵ. Vertices of the graph can be colored analogously to those of standard Mapper graph.

Network based landscapes of data

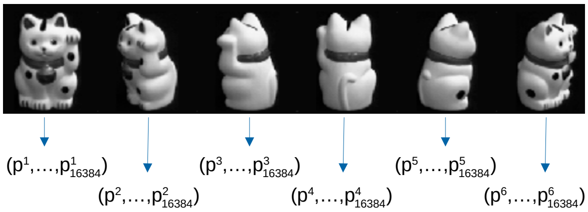
Meet the Lucky Cat

Network based landscapes of data



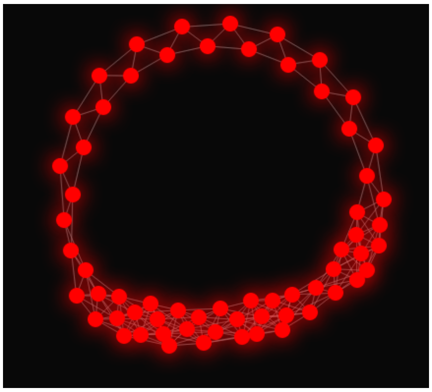
$128 \times 128 = 16384$ dimensional space

From a gray scale image to a point



Gray scale images converted to vectors in high dimensional space

Network based landscapes of data



$128 \times 128 = 16384$ dimensional space

Practical exercise 1

- Please open BM_basic_circle.
- In this proof-of-concept example we will generate a collection of points sampled from a unit circle $x^2 + y^2 = 1$.
- And built a Ball Mapper graph based on it.
- Do we see what we expected to see?

Practical exercise 2

- In our second example we will re-visit already known Boston Property Dataset.
- Please open BM_Boston_property
- This time we will use Ball Mapper to examine the structure of the 13 dimensional point cloud, and the distribution of the explanatory variable (price of properties) on the top of it.
- We will use tools from the Ball Mapper implementations to recognize which coordinates makes most statistical differences between the regions of the graph.

Basic stats

Topology and hypotehesis testing

Basic stats

- **One-sample problem**: We are given a data sample $X = \{x_1, x_2, \ldots, x_n\}, x_i \in R^d$ and cumulative distribution function $F : R^d \to [0, 1]$. Does the data X follow the distribution F: $X \sim F$?

$$H_0 : X \sim F \quad \text{vs.} \quad H_1 : X \nsim F$$

- **Two-sample problem**: We are given two samples $X_1 \sim F_1$ and $X_2 \sim F_2$ and want to test hypothesis that X_1 and X_2 were drawn from the same (unknown) distribution

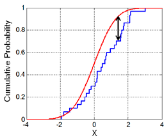
$$H_0 : F_1 = F_2 \quad \text{vs.} \quad H_1 : F_1 \neq F_2$$

Testing, for one-sample problem

Available methods depends on the data dimension

- 1-D: plenty of available tests: e.g. Kolmogorov-Smirnov, Cramer-von Mises, Andersonâ€“Darling, Chi-squared, Shapiro-Wilks
- 2-D: theoretical results for Kolmogorov-Smirnov and Cramer-von Mises, some implementations available in `python` and `R`
- d-D: Kolmogorov-Smirnov should work but no implementation available, critical values of test statistics unknown, impractical in higher dimensions

Kolmogorov-Smirnov test



Here, K-S will be used as benchmark

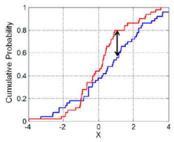
- one-sample: $D_n = \sup_x |F_n(x) - F(x)|$
- two-sample: $D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|$

Testing, for one-sample problem

Available methods depends on the data dimension

- 1-D: plenty of available tests: e.g. Kolmogorov-Smirnov, Cramer-von Mises, Andersonâ€“Darling, Chi-squared, Shapiro-Wilks
- 2-D: theoretical results for Kolmogorov-Smirnov and Cramer-von Mises, some implementations available in `python` and `R`
- d-D: Kolmogorov-Smirnov should work but no implementation available, critical values of test statistics unknown, impractical in higher dimensions

Kolmogorov-Smirnov test



Here, K-S will be used as benchmark

- one-sample: $D_n = \sup_x |F_n(x) - F(x)|$
- two-sample: $D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|$

TopoTests, one-sample problem, input

We are given a data sample $X = \{x_1, x_2, \ldots, x_n\}, x_i \in R^d$ and cumulative distribution function $F : R^d \to [0, 1]$.

One sample TopoTests

Input: sample $X = \{x_1, x_2, \ldots, x_n\}, x_i \in R^d$ and CDF $F : R^d \to [0, 1]$.

Step 1: $E_F(\chi(n, r))$, the Blueprint of F

- draw n-element samples $X'_1, X'_2, \ldots, X'_M$ from F
- for each sample X'_i compute its ECC $\chi(C_r(X'_i))$
- $$\frac{1}{M}\sum_{i=1}^{M} \chi(C_r(X'_i)) \xrightarrow[M\to\infty]{a.s.} E_F(\chi(n, r))$$

One sample TopoTests

Input: sample $X = \{x_1, x_2, \dots, x_n\}, x_i \in R^d$ and CDF $F : R^d \to [0, 1]$.

Step 2: variation form $E_F(\chi(n, r))$

- draw a new set of m-element samples $Y_1', Y_2', \dots, Y_m'$ from F
- Calculate sup distance between $\chi(C_r(Y_i')), i = 1, \dots, m$ and average ECC
- determine the threshold value t_α as a $(1-\alpha)$'th quantile of $\{d_i\}_{i=1}^m$, where α is required level of statistical significance

TopoTests

Input: sample $X = \{x_1, x_2, \dots, x_n\}, x_i \in R^d$ and CDF $F : R^d \to [0, 1]$.

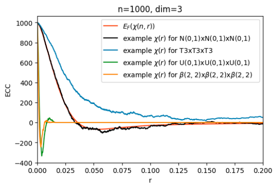
Step 3: Actual testing

- compute the ECC for sample data X: $\chi(C_r(X))$
- compute the l_∞ between $\chi(C_r(X))$ and $E_F(\chi(n, r))$

$$D = \sup_{r \in \mathbb{R}} |\chi(C_r(X)) - E_F(\chi(n, r))|$$

- reject H_0 if $D > t_\alpha$
- it is possible to get p-value as well

For the two-sample problem the procedure is slightly different but the idea remains.



TopoTests – properties

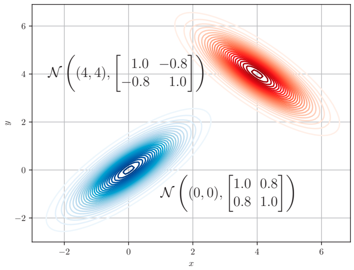
Design and goals

- general method: works regardless of the data dimension and form of probability distribution function we are testing against
- computationally feasible in higher dimensions
- theoretical results derived (no ML-like approach)
- in fact it is framework not one particular test
- outperforms baseline methods i.e. Kolmogorov-Smirnov test

TopoTests – Translational & rotational invariance

The test is not sensitive to:

- change of location parameter
- rotation
- reflection
- components reordering (c.f. $\mathcal{N} \times \Gamma$ vs. $\Gamma \times \mathcal{N}$)
- This can be tested using standard moments (after topotest is done)



Theoretical guarantees

Type II erorr (false negative, fail to reject H_0 when it is false

For fixed significance level α, probability of type II error goes to 0 exponentially with number of points sampled

$$P(\text{type II error}) \leq\sim e^{-n^2} \to 0$$

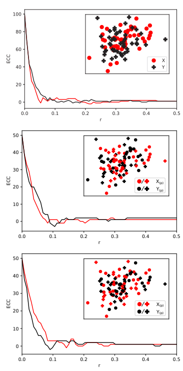
(Technical details swapped under the rug)

Two-sample problem

- $X \sim F, Y \sim G, |X| = n, |Y| = m,$

$$H_0 : F = G \quad \text{vs.} \quad H_1 : F \neq G$$

- compute distance D between ECC curves on X and Y
- data samples are pooled $Z = X \cup Y$
- split Z randomly into $X_{(p)}$ and $Y_{(p)}$ of same sizes
- compute distance $d_{(p)}$ between ECC build on $X_{(p)}$ and $Y_{(p)}$
- p-value is obtained as $p = \sum_p I\left(d_{(p)} > D\right)/N$

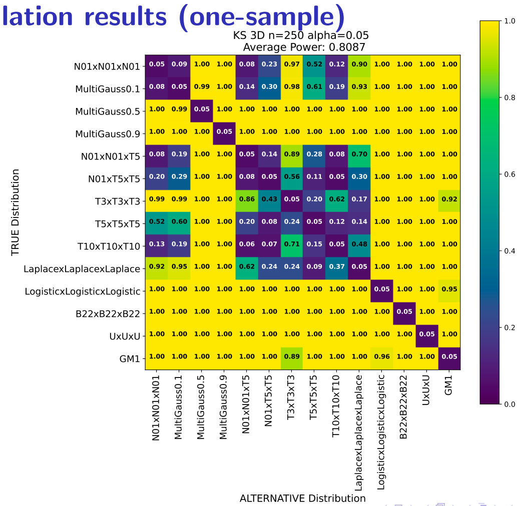


Simulation results (one-sample)

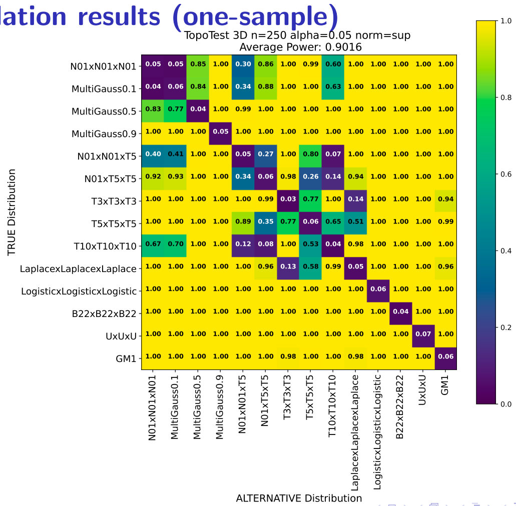
Test Power: probability that H_0 is correctly rejected when H_1 is true

- samples sizes 100–5000 data points
- test power estimated using 1000 MC replications
- power compared with KS ($d \leq 3$)
- α on diagonal is expected
- distributions easy to confuse with normal:
  - t-Student with $\nu = \{3, 5, 10\}$ DoF
  - MVN non-diagonal Σ matrix
  - Cartesian products with $\mathcal{N}(0,1)$ marginals
- TopoTests yielded higher power than KS in most of the cases
- Heavy MC simulations powered by Google.

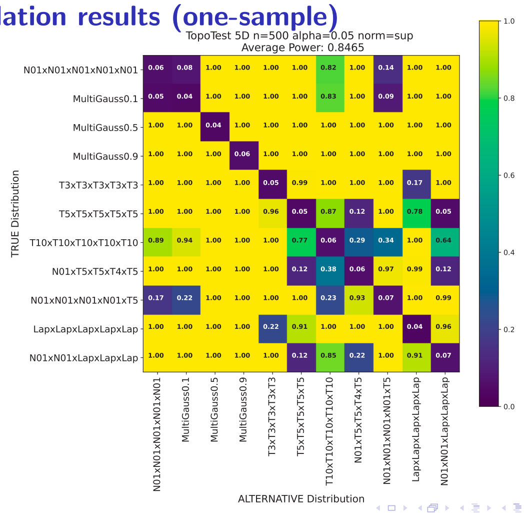
Simulation results (one-sample)



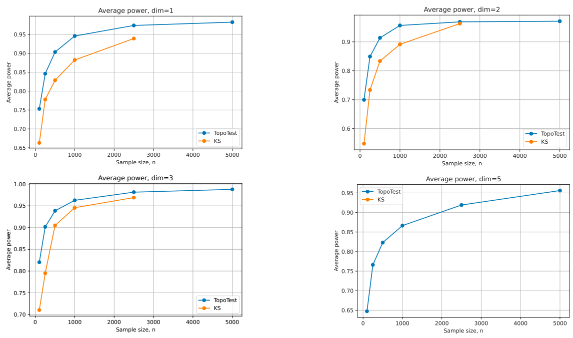
Simulation results (one-sample)



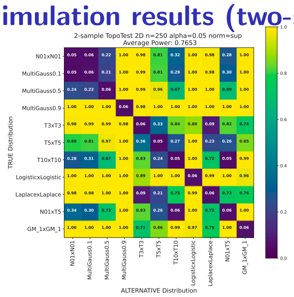
Simulation results (one-sample)

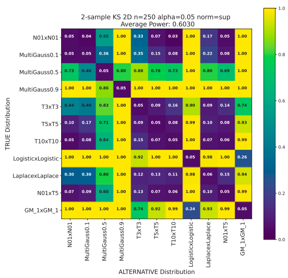


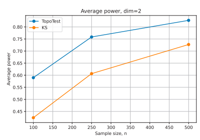
Simulation results (one-sample)



Simulation results (two-sample)







- TopoTests still outperforms the KS ($d = 2, n = 250$, 0.765 vs. 0.603)
- very expensive method

TopoTests, take home message

- There are multiple papers where topological techniques are used to show differences in distributions
- Usually they work
- We shown an important case, where it works, is comparable or better than state of the art in low dimension and have no competitions in high dimensions
- Not only that, we have theoretical guarantee for that
- Those guarantees does not depend on the fact that we started from point clouds
- We hope that this meta–observation will open up new opportunities in applied topology

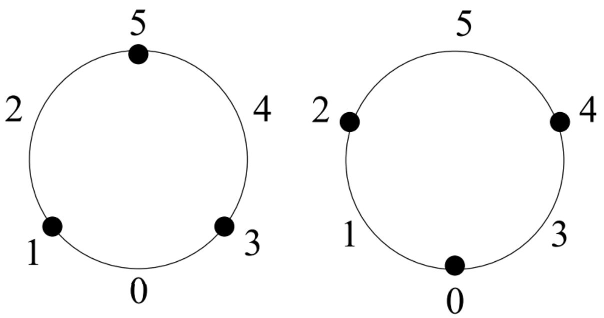
Every mathematician has a secret weapon. Mine is Morse theory.

Raoul Bott

Discrete Morse Theory

1. Let us now have a look at a Discrete Morse Theory.
2. $\mathcal{K}$ - finite regular CW complex.
3. $f : \mathcal{K} \to \mathbb{R}$, constant on every cell, is a discrete Morse function if for every $\alpha^p \in \mathcal{K}$:
   3.1 $\#\{\beta^{p+1} > \alpha^p | f(\beta^{p+1}) \leq f(\alpha^p)\} \leq 1$
   3.2 $\#\{\gamma^{p-1} < \alpha^p | f(\gamma^{p-1}) \geq f(\alpha^p)\} \leq 1$
4. Simplex is critical if both (1) $= 0$ and (2) $= 0$.
5. For any simplex conditions (1) and (2) cannot be both $= 1$ ($\implies$ define discrete gradient).

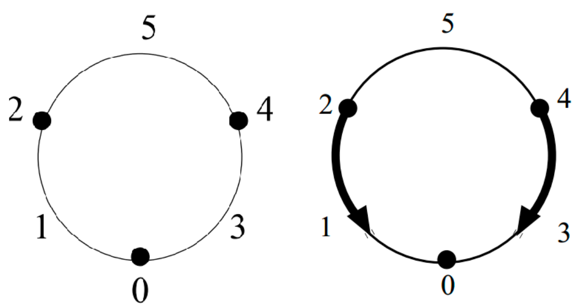
Which of them is discrete Morse function?



Why it is called discrete Morse Theory

1. Suppose $\mathcal{K}$ is a cell complex with a discrete Morse function. Then $\mathcal{K}$ is homotopy equivalent to a CW complex with exactly one cell of dimension p for each critical simplex of dimension p (we will construct this complex soon).
2. If there are no critical simplices a with $f(a) \in (a, b]$, then $\mathcal{K}(b)$ is homotopy equivalent to $\mathcal{K}(a)$. (In fact, $\mathcal{K}(b)$ collapses to $\mathcal{K}(a)$).
3. If there is a single critical simplex a with $f(a) \in (a, b]$ then $\mathcal{K}(b)$ is homotopy equivalent to $\mathcal{K}(a)$ with a handle of dimension $dim(a)$ glued.
4. Morse inequities hold.
5. Gradient of a function is more convenient to use then a function itself.

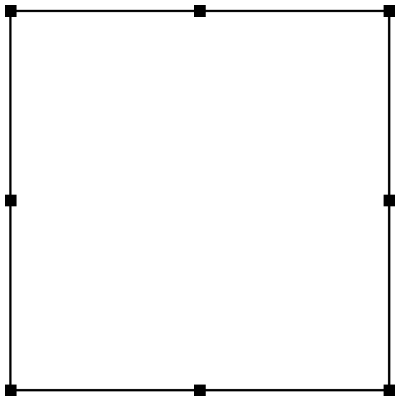
Discrete Gradient



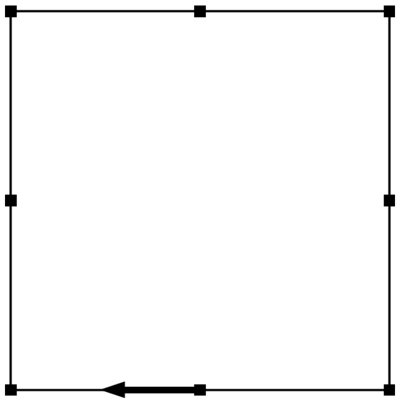
How discrete Morse functions are usually constructed?

1. We almost never assign the values. Gradient is sufficient.
2. It will be represented by arrows.
3. Every simplex can be either tail of head of exactly one arrow.
4. The vector field is curl-free (i.e. there are no loops).
5. Critical cells of Morse functions = cells which are unpaired.

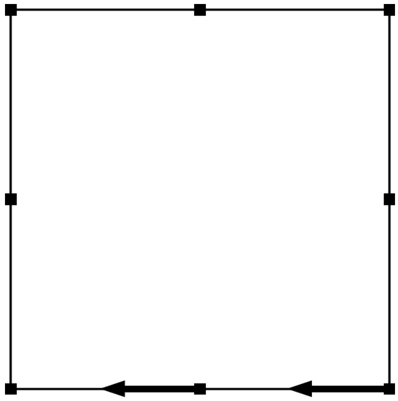
Illustration



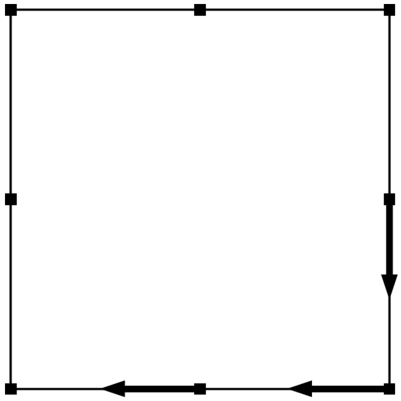
Illustration



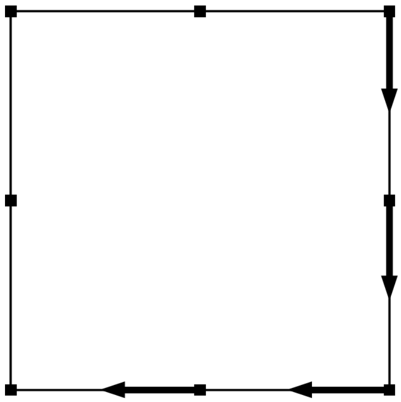
Illustration



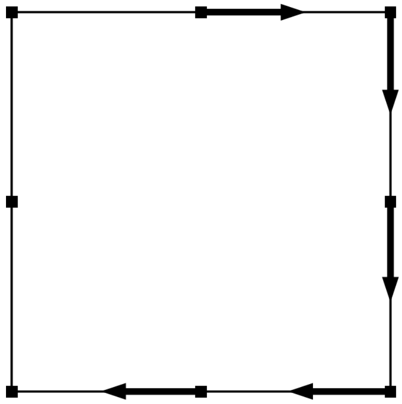
Illustration



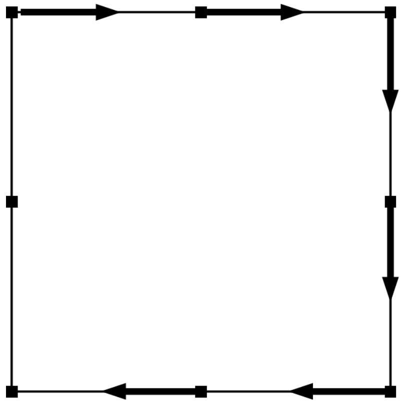
Illustration



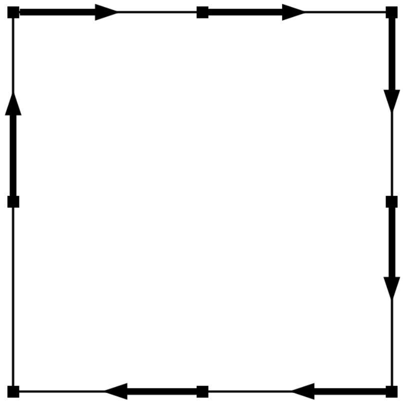
Illustration



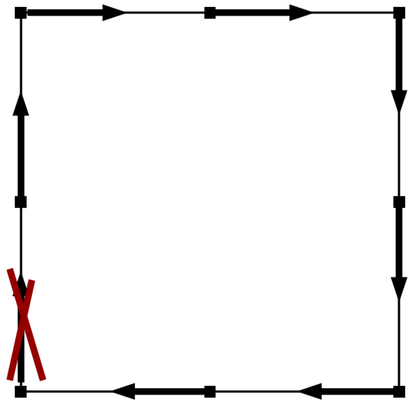
Illustration



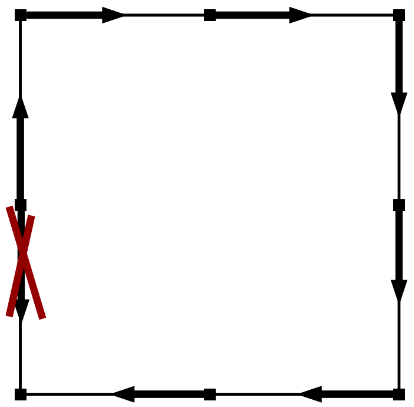
Illustration



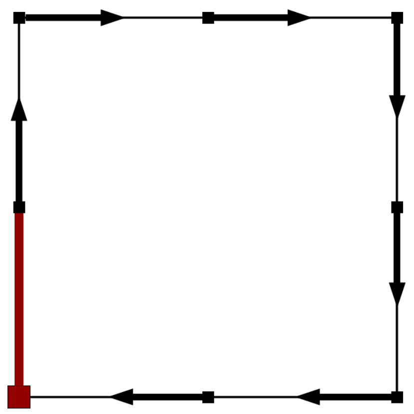
Illustration



Illustration

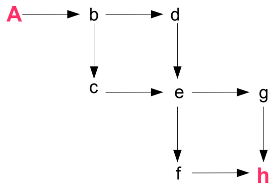


Illustration



The Morse complex

- Cells of Morse complex = critical cells of discrete vector field.
- Boundary relation computed by using gradient paths.
- Over $\mathbb{Z}_2$ – $\kappa(A, h)$ =number of gradient paths from A to h mod 2.
- Morse complex (over integers) and the initial complex are homotopically equivalent.
- Homology of a complex and its Morse complex - isomorphic.
- $\kappa(A, h) = 0$.



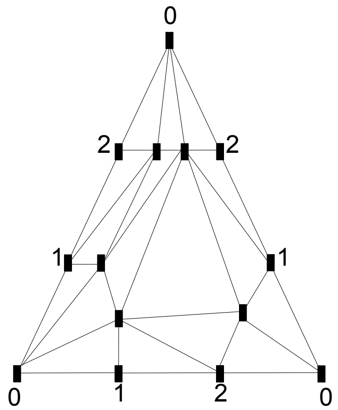
Iterated Morse Complex

- By iterating construction of a Morse complex we can obtain both (field) homology and persistence.
- Let us concentrate first on standard homology.
- Homology over a field $\implies$ pairing between A, B can be made iff $\kappa(A, B) \neq 0$ (Dmitry Kozlov).
- Algorithm to construct Morse complex – a functor $\mathbb{M} : \mathbb{C} \to \mathbb{C}$.
- $\mathbb{C}$ category of chain complexes.
- **Assumption:** if there are some Morse pairings in C, at least one of them is made in $\mathbb{M}(C)$ (vitality).
- E.g. $\mathbb{M}$ procedure search for a single possible pairing and do it.

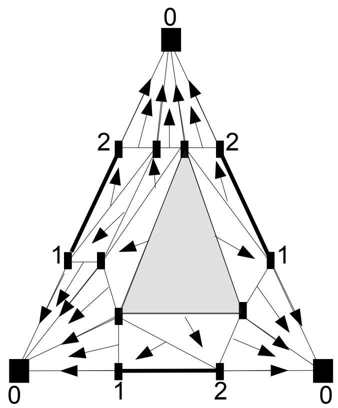
Iterated Morse Complex and homology

- Let us apply $\mathbb{M}$ iteratively.
- Homology is preserved, homotopy type is not.
- $\exists_{n \in \mathbb{N}} \mathbb{M}^n(C) = \mathbb{M}^{n+1}(C) = \ldots =: \mathbb{M}^\infty(C)$ – *Iterated Morse complex*.
- $\beta_i(C) = \#\{$ cells in $\mathbb{M}^\infty(C)$ of dimension $i\}$.
- Generators can be obtained from this procedure.

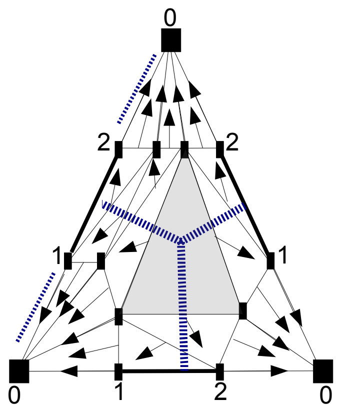
The Dounce hat.



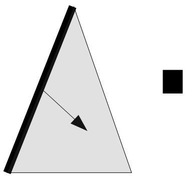
First iteration pairings



Boundary on the first iteration



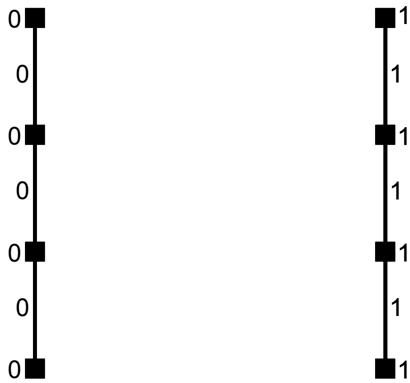
$\mathbb{M}^1(\mathcal{K})$ with Morse pairings on it



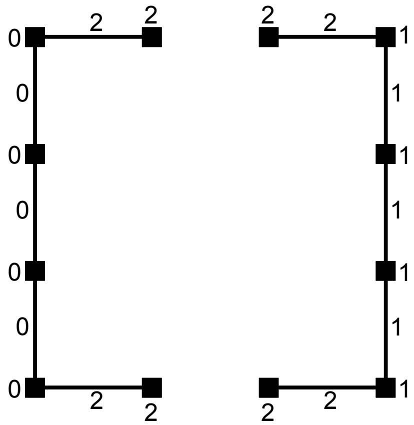
Morse complex for persistence

- C – chain complex with filtration $g : C \to \mathbb{Z}$.
- s.t. $a, B \in C$, $a < B \implies g(a) \leq g(B)$.
- Morse pairing $v : C \nrightarrow C$ is *compatible with filtration* if $g(a) = g(v(a))$ for every paired a.
- **Assumption:** $\mathbb{M}$ constructs only a vector fields compatible with filtration.
- Persistence of C and $\mathbb{M}(C)$ are the same.

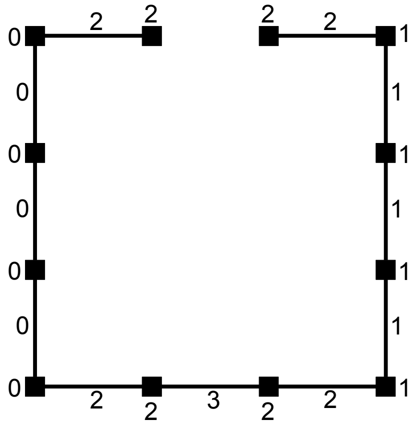
Filtered complex



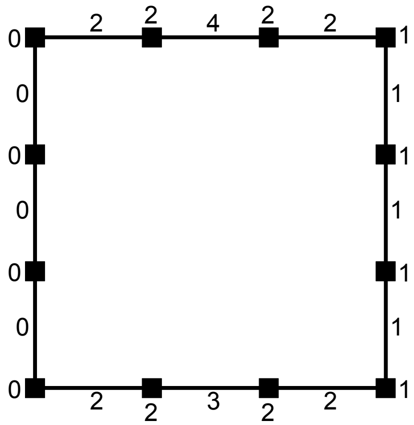
Filtered complex



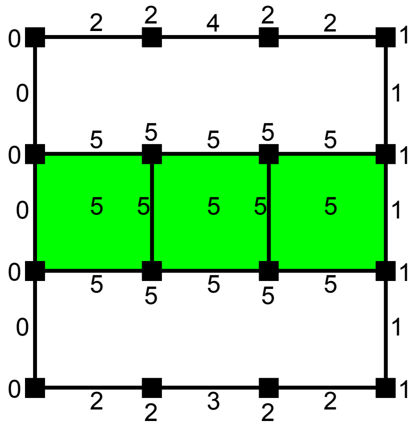
Filtered complex



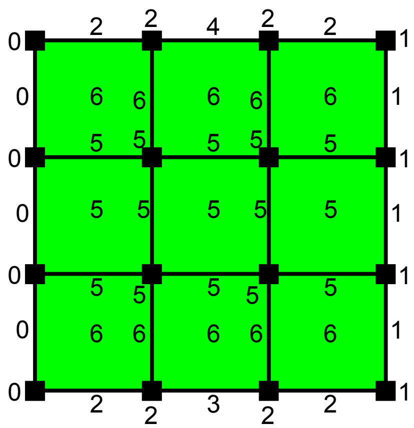
Filtered complex



Filtered complex



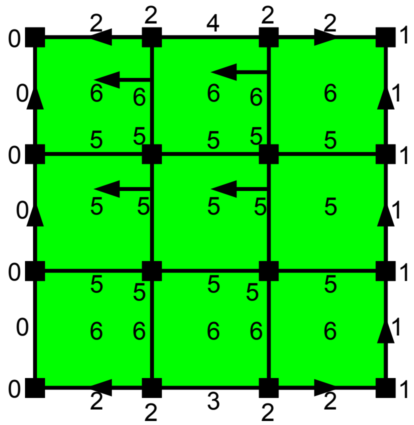
Filtered complex



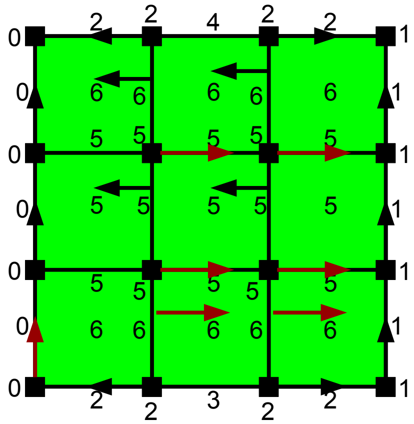
Iterated Morse Complex for persistence

- Dimension 0 – $[0, \infty)$, $[1, 3]$.
- Dimension 1 – $[4, 6]$, $[5, 6]$.

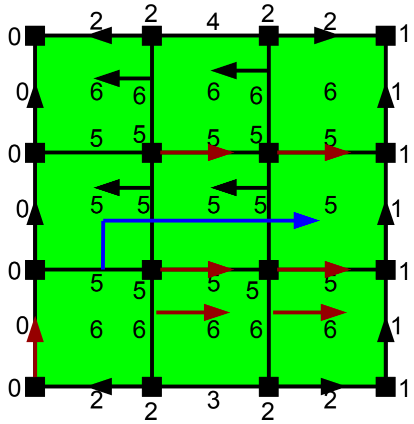
First iteration



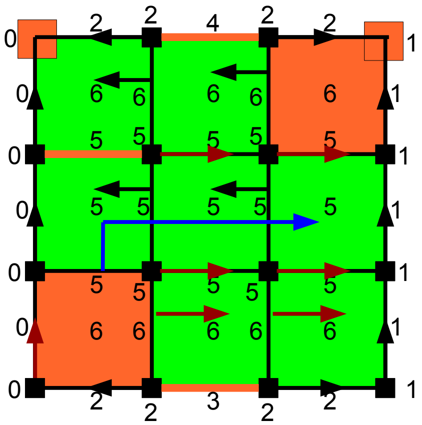
Second iteration



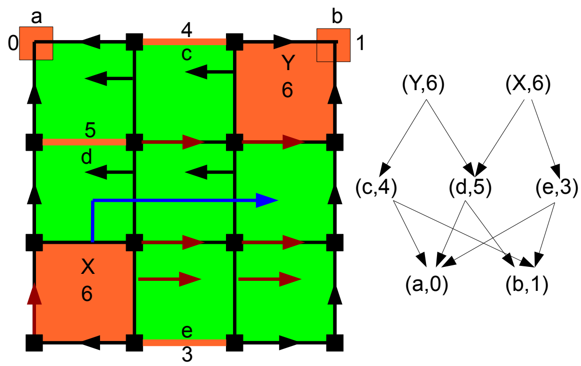
Final iteration



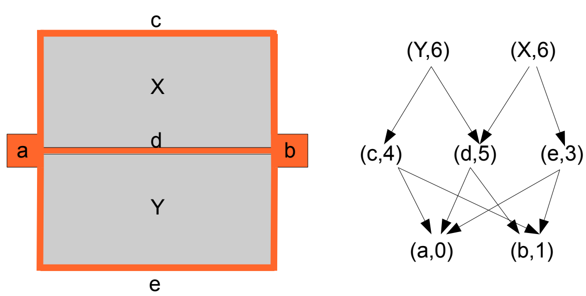
Critical cells



Critical cells



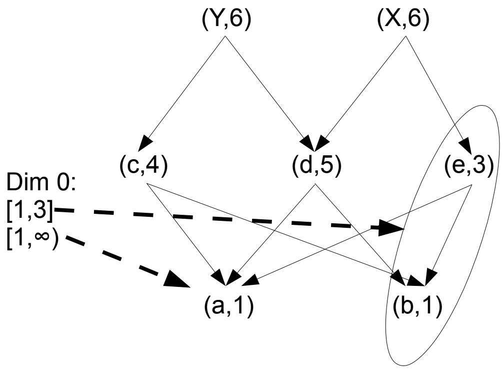
Critical cells



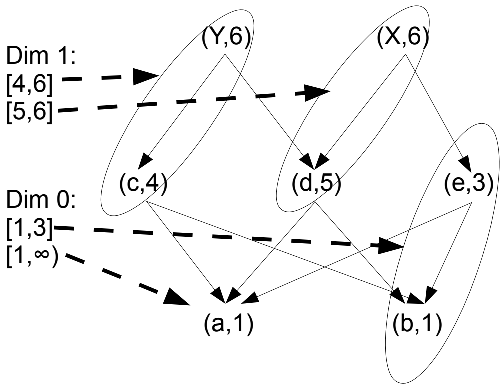
Observations

- $A \in M^{\infty}(C)$, and $B_1, \ldots, B_n$ be in boundary of A in $M^{\infty}(C)$.
- $g(A) > g(B_1), \ldots, g(B_n)$.
- $M^{\infty}(C)$ is the minimal cell complex (w.r.t number of cells) with the same persistence as C.

Observations



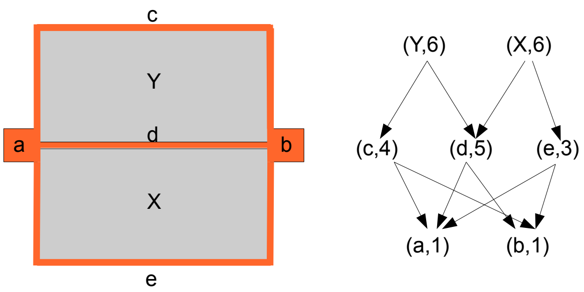
Observations



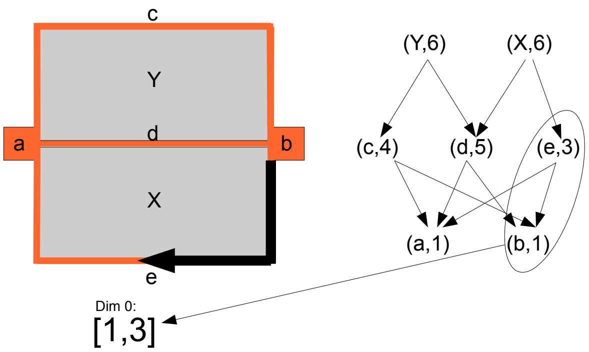
Persistent Homology via DMT

- Based on Morse theory one can obtain persistent intervals.
- No need to change representation for one suitable for matrix algorithm.
- Unlike the simplification phase, cells of *different filtration value* are paired and nonzero persistent intervals are reported.
- Pairings between cells of different filtration value – allowed (to some extent).

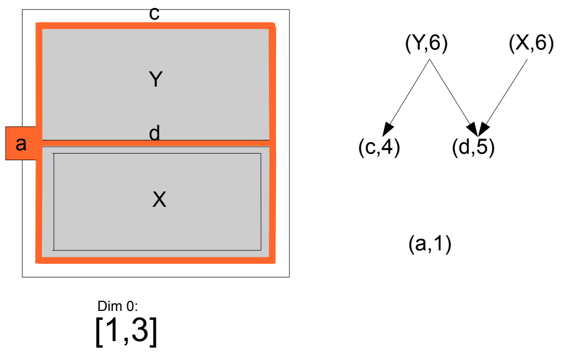
Level 0



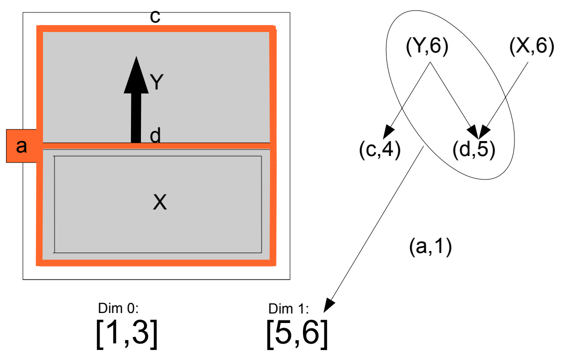
Level 3



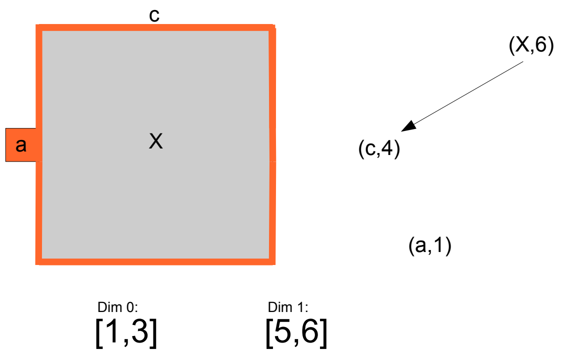
Level 3



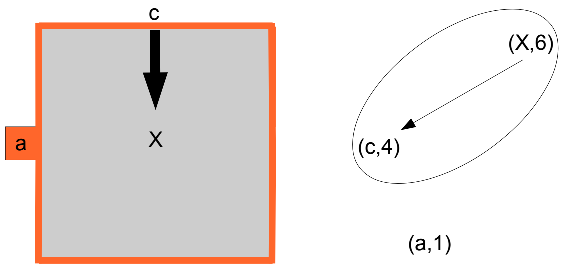
Level 6



Level 6



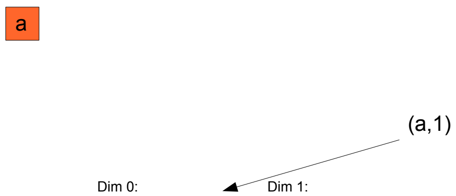
Level 6



Dim 0:
[1,3]

Dim 1:
[5,6], [4,6]

Level ∞



Dim 0:
[1,3], [1,∞)

Dim 1:
[5,6], [4,6]

The story begins here?

1. We have barely scratch the surface,
2. there are many more invariants,
3. and more applications.
4. Hence, I would like to invite you to Topological Data Analysis!

Collaborators

This research has been conducted over several years with the invaluable contributions of numerous great collaborators:

1. Phd students: Ahmad Farhad, Davide Gurnari, Niklas Hellmer, Jakub Malinowski, Jan Senge,
2. Postdocs: John Harvey, Tak-Shing Chan (Swansea), Michal Lipinski, Bartosz Naskrecki, Justyna Signerska-Rynkowska, Anastasios Stefanou, Rafal Topolnicki (Dioscuri),
3. Collaborators: Senja Barthel, Hubert Wagner, Simon Rudkin, Radmila Sazdanovic, Berend Smit, Alex Smith, Ruben Specogna, Lukasz Stettner and others

Support

I gratefully acknowledge support by Dioscuri program initiated by the Max Planck Society, jointly managed with the National Science Centre (Poland), and mutually funded by the Polish Ministry of Science and Higher Education and the German Federal Ministry of Education and Research.

Thank you for your time!

Dioscuri Centre in Topological Data Analysis
@Facebook

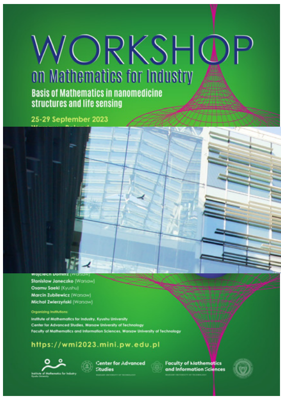
DIOSCURI
CENTRE IN TOPOLOGICAL DATA ANALYSIS

Paweł Dłotko
pdlotko @ impan.pl
pdlotko @ gmail
pawel_dlotko @ skype

Medical imaging signatures with topology for cancer

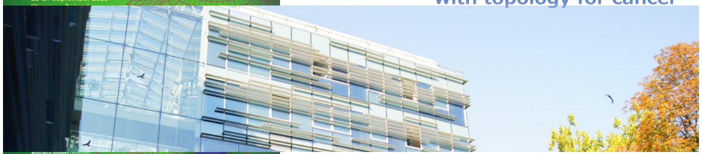
Hidetaka Arimura
Faculty of Medical Sciences, Kyushu University, Japan

What the author is interested in is the connection between medicine and mathematics. A human body is equivalent to a tube or donut (without considering holes of nose and eyes). The central hole is a digestive system. The body is covered by surface tissue (epithelial cells). The epithelial cells exposed to the outside world might have gene mutations, thereby resulting in cancer cells. On the other side, the heterogeneity of pixel values in medical images (computed tomography, magnetic resonance imaging, positron emission tomography, etc) would reflect biological tumor heterogeneity, which could be related to the degree of malignancy and patients' prognoses. We have attempted to develop novel medical imaging signatures, which are defined as sets of features calculated based on mathematical models from medical images, for prediction of the degree of malignancy and patients' prognoses. As results, the author's group has shown several data that the topological imaging signatures could be superior to conventional ones in terms of the prediction. The topological image features are derived from Betti number maps (b_0, b_1, and b_2) within cancer regions of medical images. The assumption that the author has thought through (not twisting things around) is that the b_0, b_1, b_2 features may characterize high tumor cell density areas, scattered dead cell areas (necrotic tissues), cancer blood vessels (angiogenesis), respectively. The author will present the basics of topological image features and the applications to lung cancer and head and neck cancer.



Workshop on mathematics for industry
25-29 September 2023 (Warsaw)
Basis of mathematics in nanomedicine structures and life sensing

Medical imaging signatures with topology for cancer

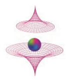


1. Background and radiomics

Hidetaka Arimura, PhD
Professor (medical Physics)
Division of Medical Quantum Science
Department of Health Sciences, Faculty of Medical Sciences
Kyushu University

I am a medical physicist, not a mathematician

Overall outline



✓1st: Background and radiomics

2nd: Medical background for topological radiomics

3rd: Applications of topological radiomics

Outline

- ✓ Nature of medical physics
- ✓ Association in shape between human body and topology
- ✓ What we are doing now (radiomics)
- ✓ Mathematical models beyond conventional radiomics
- ✓ Summary (1)

4

Nature of medical physics

5

What is mathematics?

- ✓ Mathematics is the art of giving the same name to different things (Henri Poincar´e)
- ✓ Mathematics is the structure of abstract reasoning (Richard Philips Feynman)

Abstract science ?

6

What is physics?

- ✓ Basic science that understands and describes concrete natural phenomena by using mathematics that can explain them
- ✓ Basically, the natural phenomena could be theoretically predicted in the macroscopic world, but probabilistically predicted in the microscopic world (quantum mechanics).

Concrete science ?

7

What is *medicine*?

- ✓ Science of uncertainty and an art of probability [William Osler (1849-1919) , Principle and Practice of Medicine]
- ✓ Inherent uncertainty in health care [The Lancet 2010; 375: 1666]

Uncertain science ?

8

What is *medical* physics (my field)?

Applied science that could describe natural phenomena related to human bodies with uncertainties (due to thermal motion or dynamic metabolic activity?) using mathematics that can be used for diagnosis and therapy

Concrete, but
Uncertain science?
with abstract spice ? ?

9

What we can get in cancer properties

Big pictures on human body and diseases, because you can only predict softly something with uncertainties

フッターを追加 10

"What is essential is visible to the heart. It is invisible to the eye." (modified from an original version)

(By a fox in "The Little Prince")

11

"The truth is invisible to the eye. However, mathematics can reveal and express its appearance."

(By a mathematician in a novel of "The Doctor Loves Equations" written by Youko Ogawa)

12

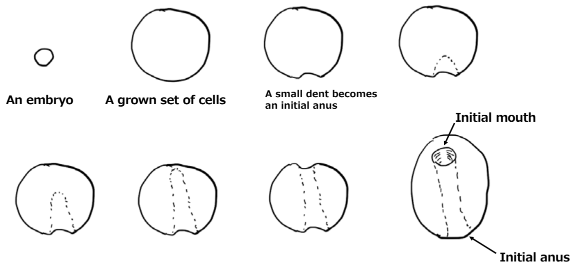
"The book of the universe is written in mathematics"

(Galileo Galilei)

Association in shape between human body and topology

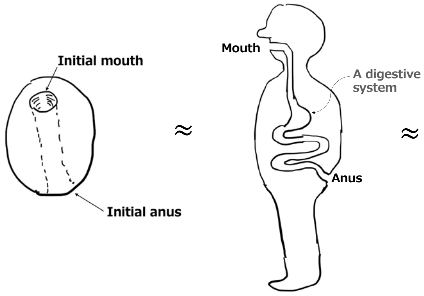
17

What is a human body?

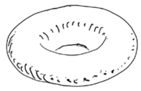


18

What is a human body basically?



Our bodies are covered by surface tissue (epithelial cells).



Epithelial cells of the digestive system exposed to the outside world

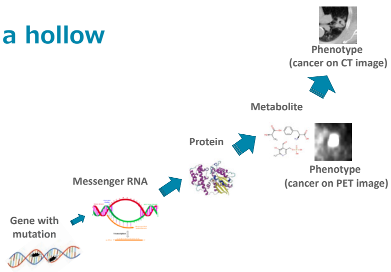
They might be cancerous due to gene mutations such as colon cancer.

20

Respiration system: a hollow

Epithelial cells of lung, which may be exposed to the smoke or chemical things.

They might be lung cancer.

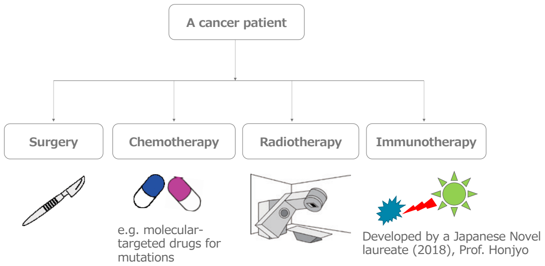


Modified from http://omics.jp/about-omics

フッターを追加 21

What we are doing now (radiomics)

How are cancer patients treated?

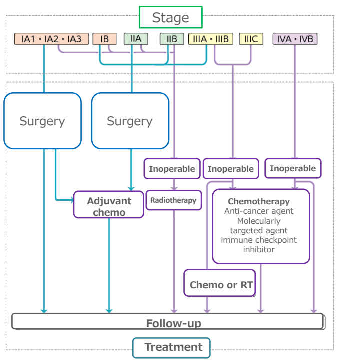


26

How to decide each treatment approach : NSCLC

If a patient is operable with a stage IA tumor, this patient will receive surgery, but if not, this patient could receive radiotherapy.

National Cancer Research Center, Cancer Information Service
https://ganjoho.jp/public/cancer/lung/treatment.html

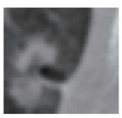


27

However, we have an issue: patients with similar profiles treated by radiotherapy may have different survival times

Homogeneous

Inhomogeneous



Gender	Male	Female
Histology	SCC	SCC
Age	61	68
Overall stage	Ⅰ (T1N0M0)	Ⅰ (T1N0M0)
Survival time (day)	1357	237

Issue: "One-size-fits-all" treatment procedures could be unsuitable for some patients.

28

Outcomes of surgery and radiotherapy are almost the same, but they are not perfect

Non–small-cell lung cancer (NSCLC) Stage I

Onishi, et al. Int J Radiat Oncol Biol Phys. 2011

Table 3. Comparison of 5-y overall survival rate between surgical series and SBRT

Clinical stage	United States (1)	Japanese National Cancer Center (2)	Japanese National Survey (3)	SBRT
IA	61	71	77	76
IB	40	44	60	64

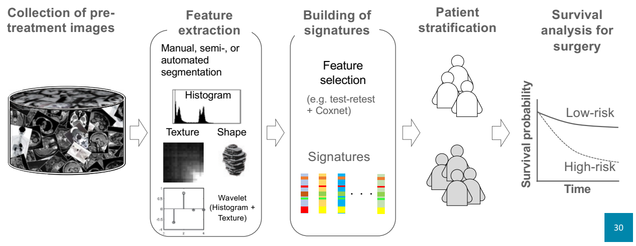
Abbreviation: SBRT = stereotactic body radiotherapy. Values are percentages.

A treatment method has been effective for some patients, but the method is not always effective for all patients.

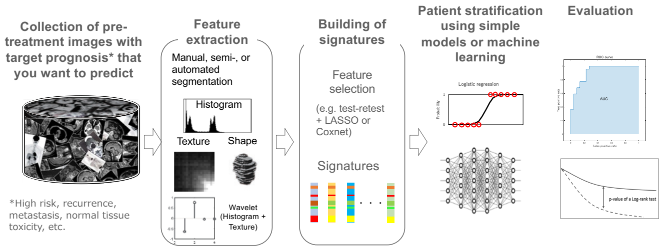
My definition of radiomics

Radiomics: one of omics research fields where a set of medical image features related to patients' prognoses are considered "radiome" like genome, transcriptome, proteome, metabolome, to improve diagnosis, treatment, and prognosis

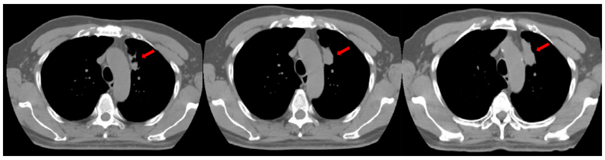
Application: Selection of patients for a more appropriate treatment strategy by prediction of of patients' prognoses obtained from pre-treatment medical images (except molecularly targeted drugs for tumors with gene mutations)



Workflow of radiomics with AI



Prediction of cancer prognosis* (expected outcome) after treatment from medical images



Prediction: progression (left to right) and partial response (right to left)

*Prognosis: expected outcome or outlook related to a medical condition or treatment

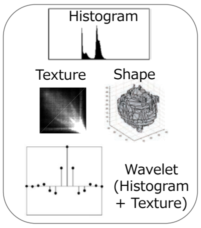
33

Mathematical models beyond conventional radiomics

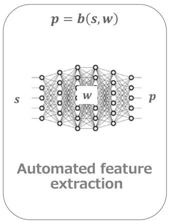
35

There are disadvantages in conventional radiomics and deep learning

Radiomics features, y (hand crafted)



Automatically learned features using deep learning (DL)



Disadvantage of conventional radiomis features:

- ✓ Interscanner variability
- ✓ Vulnerability for various imaging protocols

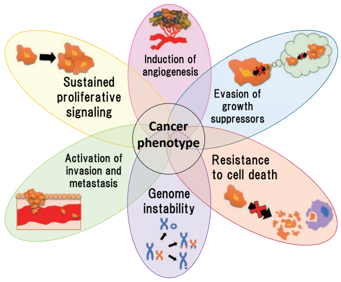
Disadvantage of DL:

- ✓ Necessity of huge number of data

36

What are requirements of better features in radiomics from medical point of view?

- ✓ Image features should reflect one or some of cancer phenotypes
- ✓ Image features should be mathematically invariant



PLoS ONE 2016;11(7): e0157416. doi:10.1371/journal.pone.0157416

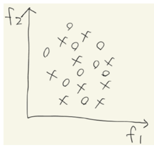
39

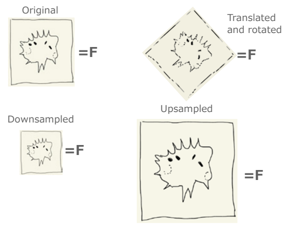
What are requirements of better features in radiomics from AI (pattern recognition) point of view?

- ✓ Features should be similar to those for objects in a same category, but they should be different from those for objects in different categories [Duda 2000].
- ✓ Features should be invariant to irrelevant transformations of the inputs such as translation, rotation, and scale.

Good features

Bad features





Duda, R.O.; Hart, P.E.; Stork, D.G. Pattern classification, 2nd ed.; Wiley-Interscience: New York, NY, USA, 2000.

40

Topology is one of theories that we want

Topologists can eat a coffee cup with trying to drink coffee from donut!

They can not differentiate a coffee cup and donut.

41

Hypothesis: Classification of cancer patterns into several categories depending on Betti numbers

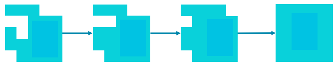
- ❑ Mathematical classification of objects by simplifying connectivity

[G. Rote, in: Eff. Comput. Geom. Curves Surfaces, 2006: 277–312.]

- ❑ Betti number: Invariant value
  - b0: Number of connected components
  - b1: Number of holes

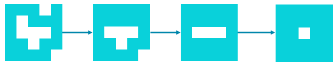
Betti numbers after continuous deformation

Cancer pattern C



- ❑ b0 = 1
- ❑ b1 = 0

Cancer pattern A



- ❑ b0 = 1
- ❑ b1 = 1

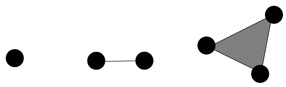
Cancer patterns with prognostic information could be classified into several categories based on Betti numbers with intrinsic geometrical patterns

42

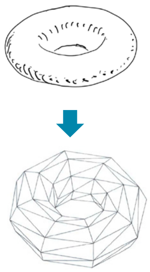
Computational topology

Decomposition of a shape: a set of simplices (many triangles)

= Vertices (Points) + Edges (Lines) + Faces

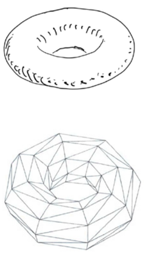


Vertex (point) Side (line) Face



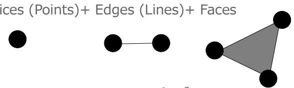
フッターを追加 43

Computational topology



Decomposition of a shape: a set of simplices (many triangles)

= Vertices (Points)+ Edges (Lines)+ Faces



Homology group: $H_n(X) = \frac{ker\partial_n}{im\partial_{n+1}}$

$b_n = rank\,(H_n(X))$

$= rank\,(ker\partial_n) - rank\,(im\partial_{n+1})$

$rank\,(ker\partial_n)$: n-dimensional cycles

$rank\,(im\partial_{n+1})$: n-dimensional boundaries

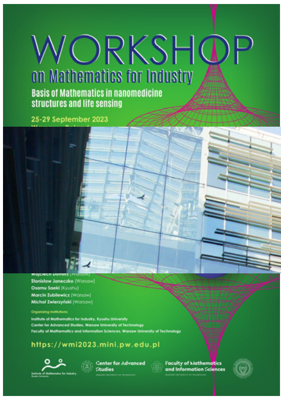
フッターを追加 44

Summary (1)

- ✓ We try to find significant image features to characterize intrinsic cancer geometry
- ✓ We assume that the significant image features may be mathematically invariant.
- ✓ We believe that the image features could be calculated from topology

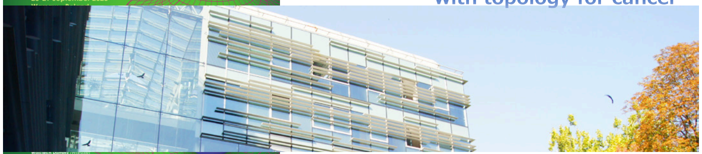
Any question or comments ?

46



Workshop on mathematics for industry
25-29 September 2023 (Warsaw)
Basis of mathematics in nanomedicine structures and life sensing

Medical imaging signatures with topology for cancer



2. Medical background for topological radiomics

Hidetaka Arimura, PhD
Professor (medical Physics)
Division of Medical Quantum Science
Department of Health Sciences, Faculty of Medical Sciences
Kyushu University

Overall outline

1st: Background and radiomics

✓ 2nd: Medical background for topological radiomics

3rd: Applications of topological radiomics

2

Outline

✓Are there relationships among cancer properties, cancer geometry and prognosis?

✓Hole analysis (topological radiomics) as an explainable AI

✓Summary (2)

3

Hole analysis (topological radiomics) as an explainable AI

16

Association between biological components of cancer and topology

Medical targets

- ✓ Human body
- ✓ Cancer

Biological components in histological images

- ✓ Cells
- ✓ Necrosis
- ✓ Gland duct
- ✓ Vessel

Topological components in CT images

- ✓ Connected components
- ✓ Holes (tubes)
- ✓ Cavities

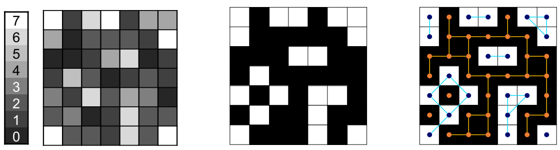
17

Definition of connected components and background within a kernel

8 neighborhood

4 neighborhood

- Center pixel for a connected component in 8 neighborhood
- Neighboring pixel for a connected component in 8 neighborhood
- Center pixel for background in 4 neighborhood
- Neighboring pixel for background in 4 neighborhood



An original image

Binary image obtained by a threshold value of 3

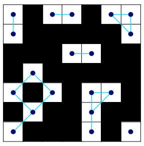
Definition of connected components (white) and background (black)

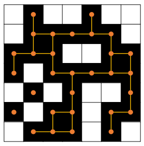
18

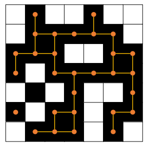
Definitions of connected components and holes within a kernel

Number of connected components (white) with a pixel of 1, b0 = 7

Number of holes =
(Number of all connected background)
– (Number of connected background contacting with edges)







Number of all connected background (black): 3

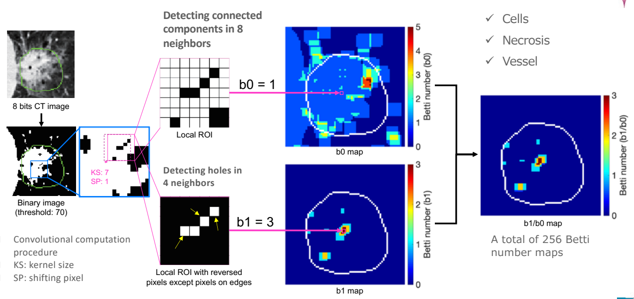
Number of connected background contacting with edges: 2

b0 = 7

b1 = 3 – 2 = 1

19

2D Betti maps



- ✓ Cells
- ✓ Necrosis
- ✓ Vessel

A total of 256 Betti number maps

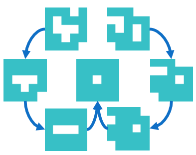
- ❑ Convolutional computation procedure
- ❑ KS: kernel size
- ❑ SP: shifting pixel

Ninomiya K, Arimura H, et al., Phys Med. 2020; 69:90-100

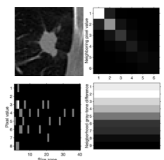
20

Purpose

To develop a novel image features based on topologically invariant Betti numbers for prognosis prediction of non-small cell lung cancer patients



×



Homology

Radiomics

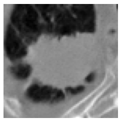
21

Patients' information

Number of patients	205
Age (yrs)	43-91 (median: 69)
Follow-up	135
Censored	70
Gender	
Male	147
Female	58
Stage	
I	38 (18.5 %)
II	22 (10.7 %)
IIIA	52 (25.4 %)
IIIB	92 (44.9 %)
Unavailable	1 (0.5 %)

- ❑ All the patients received radiotherapy or chemo-radiotherapy
- ❑ Free-breathing PET-CT

CT image

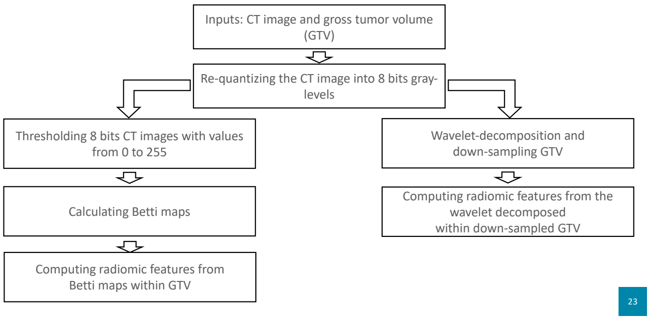


Gross tumor volume (GTV)

[Aerts et al, Data from NSCLC-radiomics. The Cancer Imaging Archive (TCIA). https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI]

22

Overall scheme: computation of image features

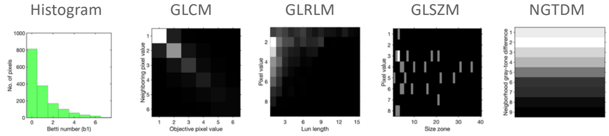


23

Betti number (BN)-based features calculation

- ❑ 13,824 BN-based features were obtained by applying 54 calculation methods based on histogram and texture features to 256 Betti maps
- ❑ 14 histogram-, 9 gray level co-occurrence matrix (GLCM)-, 13 gray level run-length matrix (GLRLM)-, 13 gray level size-zone matrix (GLSZM)-, and 4 neighborhood gray-tone difference matrix (NGTDM)-based features

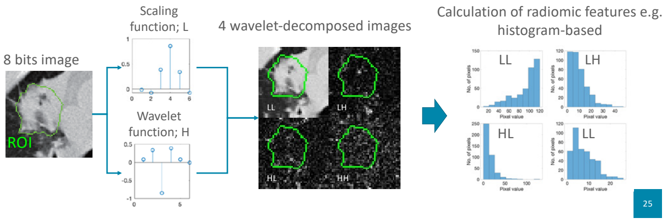
Histogram GLCM GLRLM GLSZM NGTDM



24

Wavelet decomposition and feature calculation

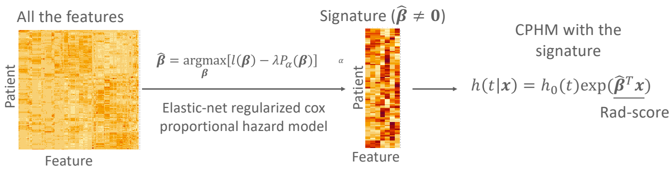
- A total of 216 features were calculated from 4 wavelet-decomposed images
- Coiflet1 mother wavelet was applied for the wavelet decomposition



Construction of radiomic signatures and Cox proportional hazard model (CPHM)

Radiomic signature was constructed from features up to 13

Radiomic scores (rad-scores) were produced using CPHM



$\boldsymbol{\beta}$:Parameter vector
$\widehat{\boldsymbol{\beta}}$:Optimum parameter vector
$\boldsymbol{\beta}^{(0)}$:Zero vector

$l(\cdot)$:log partial likelihood
$h(\cdot)$:CPHM
$h_0(\cdot)$:Baseline hazard (Kaplan-Meier curve)

$P_\alpha(\cdot)$:Elastic-net regularization term
t:Survival time
$\boldsymbol{x}$:Feature vector

Simon et al, Journal of Statistical Software, Vol. 39(5) 1-13, 2011; Sun et al, Statistica Sinica. 24 1433-1459, 2014

Optimization of parameters in Cox proportional hazard model

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}}[kl(\boldsymbol{\beta}) - \lambda P_\alpha(\boldsymbol{\beta})]$$

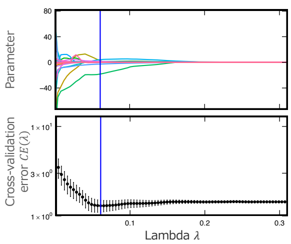
$$P_\alpha(\boldsymbol{\beta}) = \alpha\|\boldsymbol{\beta}\|_1 + \frac{1-\alpha}{2}\|\boldsymbol{\beta}\|_2^2$$

Cross-validation error $CE(\lambda)$ of s-fold cross-validation

$$CE(\lambda) = -\frac{1}{n}\sum_{i=1}^{s}\{l(\widehat{\boldsymbol{\beta}}_\lambda^i) - l^i(\widehat{\boldsymbol{\beta}}_\lambda^i)\}$$

k:scaling factor
$l(\cdot)$:log partial likelihood
$\boldsymbol{\beta}$:Parameter vector
λ:Lagrange multiplier
$P_\alpha(\cdot)$:Penalty term
n: Number of patients in a training dataset

$\|\cdot\|_1$:L1 norm
$\|\cdot\|_2$:L2 norm
$l^i(\cdot)$:log partial likelihood calculated from a dataset without the ith patient
$\widehat{\boldsymbol{\beta}}_\lambda^i$:Parameter vector optimized from a dataset without the ith patient



- 5-fold cross validation test based on random numbers
- 100 times repetition
- Significant features selected by more than 90 times

Friedman J, et al. J Stat Softw. 2010;33: 1–22

Equation of partial likelihood

in failure/censoring time. We wish to find β which maximizes

$$L(\beta) = \prod_{i=1}^{m} \frac{e^{x_{j(i)}^{\top}\beta}}{\sum_{j \in R_i} e^{x_j^{\top}\beta}}$$

Simon et al, Journal of Statistical Software, Vol. 39(5) 1-13, 2011

subject to our constraint: $\alpha \sum |\beta_i| + (1-\alpha) \sum \beta_i^2 \le c$. Maximizing the partial likelihood is equivalent to maximizing a scaled log partial likelihood,

$$\frac{2}{n}\ell(\beta) = \frac{2}{n}\left[\sum_{i=1}^{m} x_{j(i)}^{\top}\beta - \log\left(\sum_{j \in R_i} e^{x_j^{\top}\beta}\right)\right]$$

We scale by a factor of $2/n$ for convenience. Hence, if we consider the Lagrangian formulation, our problem becomes

$$\hat{\beta} = \operatorname{argmax}_{\beta}\left[\frac{2}{n}\left(\sum_{i=1}^{m} x_{j(i)}^{\top}\beta - \log\left(\sum_{j \in R_i} e^{x_j^{\top}\beta}\right)\right) - \lambda P_{\alpha}(\beta)\right] \qquad (1)$$

where,

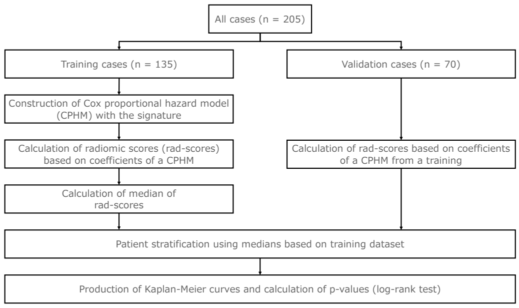
$$\lambda P_{\alpha}(\beta) = \lambda\left(\alpha \sum_{i=1}^{p} |\beta_i| + \frac{1}{2}(1-\alpha)\sum_{i=1}^{p} \beta_i^2\right)$$

is known as the elastic net penalty. It is a mixture of the ℓ_1 (lasso) and ℓ_2 (ridge regression) penalties. The lasso penalty (Tibshirani 1996) tends to choose only a few nonzero coefficients. While often desirable, this can cause problems. If two predictors are very correlated, the lasso will pick one and entirely ignore the other.

フッターを追加

28

Evaluation method



29

Results: constituent features in the signatures

Betti number (BN)	Wavelet-decomposition (WD)
GLSZM_SZLGE_62	GLRLM_LRE_HL
GLSZM_SZE_28	Hist_Mean_HL
GLSZM_SZE_108	GLRLM_LRHGE_LL
GLRLM_SRHGE_4	GLSZM_ZP_LL
GLSZM_SZHGE_95	GLSZM_LZE_LL
GLRLM_SRHGE_94	GLSZM_LZLGE_LL
GLRLM_SRHGE_111	
GLSZM_SZHGE_100	
GLSZM_SZHGE_102	
GLSZM_LZHGE_95	

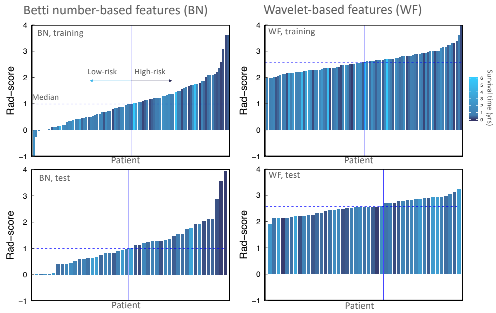
GLSZM: Gray level size zone matrix
GLRLM: Gray level run length matrix
SZLGE: Small zone low gray level emphasis
SZE: Small zone emphasis
SRHGE: Short run high gray level emphasis
SZHGE: Small zone high gray level emphasis
LZHGE: Large zone high gray level emphasis

Hist: Histogram
LRE: Long run emphasis
LRHGE: Long run high gray level emphasis
ZP: Zone percentage
LZE: Large zone emphasis
LZLGE: Large zone low gray level emphasis

30

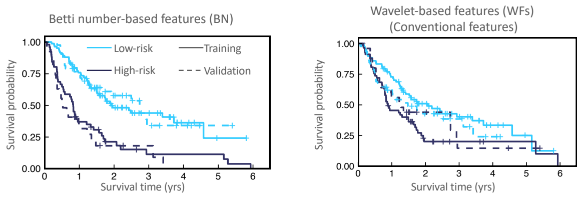
Results: rad-scores and survival time

Rad-score $=\hat{\beta}^T x$



31

Results: p-values of Kaplan-Meier curves



	p-value (hazard ratio, 95% confidence interval)	
	Training (n=135)	Validation (n=70)
BN	6.7×10^{-6} (0.41, 0.26–0.65)	3.4×10^{-5} (0.32, 0.16–0.62)
WF	5.9×10^{-3} (0.57, 0.37–0.88)	6.7×10^{-1} (0.88, 0.48–1.60)

32

Relapse predictability of topological signature on pretreatment planning CT images of stage I non-small cell lung cancer patients before treatment with stereotactic ablative radiotherapy

(Kodama T, Arimura H, et al. *Thorac Cancer*. 2022)

Takumi Kodama[1], Hidetaka Arimura[2], Yuko Shirakawa[3]
Kenta Ninomiya[4], Tadamasa Yoshitake[5], Yoshiyuki Shioyama[6]

[1]Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University
[2]Division of Medical Quantum Science, Department of Health Sciences, Faculty of Medical Sciences, Kyushu University
[3]National Hospital Organization Kyushu Cancer Center
[4]Sanford Burnham Prebys Medical Discovery Institute
[5]Department of Clinical Radiology, Graduate School of Medical Sciences, Kyushu University
[6]Ion Beam Therapy Center, SAGA HIMAT Foundation

KYUSHU UNIVERSITY

33

Background

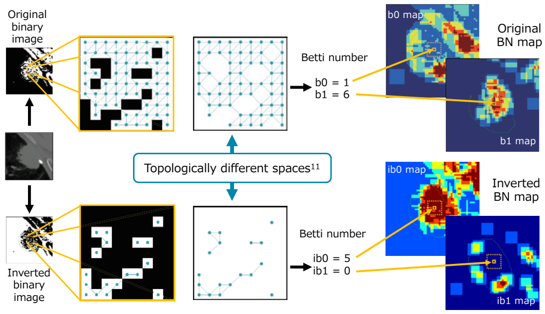
- Surgery is a first treatment option for patients with stage I non-small cell lung cancer (NSCLC), and stereotactic ablative radiotherapy (SABR) is recommended for inoperable patients[1]
- Outcomes of surgery and SABR for stage I NSCLC patients were comparable[2,3]
- Probabilities of locoregional recurrence (LRR) after surgery and SABR were also comparable[2,3]

Prediction of the cancer relapse before treatment is important to select a more appropriate therapy

1. Lackner RP. *JNCCN-Journal Natl Compr Cancer Netw.* 2017; 15(4): 504-535.
2. Ackerson BG. *Lung Cancer.* 2018; 125: 185-191.
3. Onishi H. *Int J Radiat Oncol Biol Phys.* 2011; 81(5): 1352-8.

34

Inverted Betti number maps



11. Hatcher A. Tsinghua University Press; 2005.

35

Purpose

This study aimed to explore the predictability of topological signatures linked to the locoregional relapse (LRR) and distant metastasis (DM) on pretreatment planning computed tomography images of stage I non-small cell lung cancer (NSCLC) patients before treatment with stereotactic ablative radiotherapy (SABR).

36

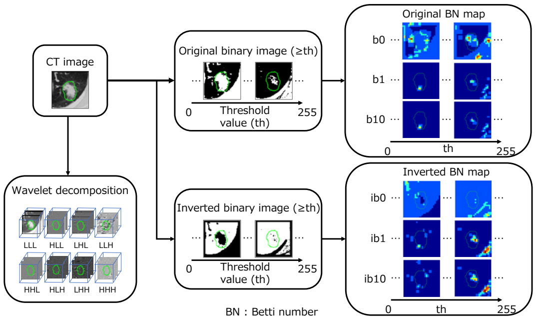
Clinical information

Planning computed tomography images of 125 stage I NSCLC patients treated with SABR at Kyushu University Hospital were employed

	Training (n = 65)	Test (n = 60)	p-value
Age (year; min - max (median))	60 - 91 (78)	60 - 89 (79)	0.586 (U test)
Gender			
Male	44	46	0.359 (chi-square)
Female	21	14	
Stage			
IA	43	34	0.365 (chi-square)
IB	22	26	
Histology			
Adenocarcinoma	41	36	0.348 (chi-square)
Squamous cell carcinoma	21	23	
Large cell carcinoma	3	1	
Component			
Solid	52	44	0.335 (chi-square)
Part solid	7	12	
GGO	6	4	
Tumor volume (mm^3; min - max (median))	1484 - 57374 (8841)	1015 - 45444 (9201)	0.698 (U test)
Prognosis			
LRR	18	16	1.00 (chi-square)
LRR free	47	44	
TTLRR (month; min - max (median))	3.6 - 162.2 (32.2)	5.3 - 158.2 (29.8)	0.886 (U test)

37

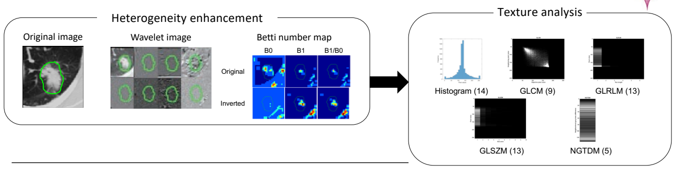
Betti maps for multiple thresholding pixel values



BN : Betti number

38

Imaging feature calculation

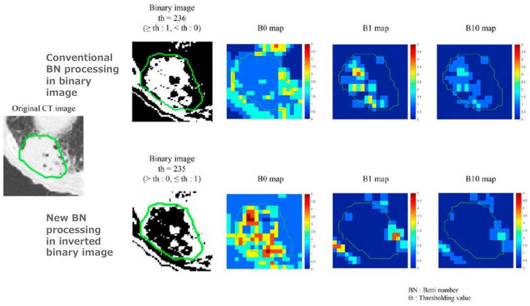


Imaging feature	Heterogeneity enhancement	No. of features
WFs	Wavelet decomposition	486
BFs	Original BN mapping	41526
iBFs	Original & inverted BN mapping	82998
BWFs	Combination (WFs + BFs)	41958
iBWFs	Combination (WFs + iBFs)	83430

GLCM : Gray-level co-occurrence matrix
GLRLM : Gray-level run-length matrix
GLSZM : Gray-level size-zone matrix
NGTDM : Neighbor gray-tone difference matrix

39

Betti number (BN) and inverted BN



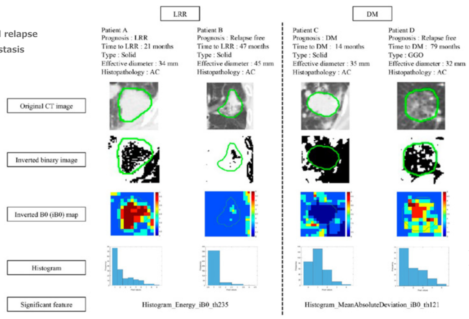
- ✓ Cells
- ✓ Necrosis
- ✓ Vessel

Kodama T, Arimura H, et al. Thorac Cancer. 2022 Aug;13(15):2117-2126

40

Patients with relapse and relapse free linked to significant features of inverted BN map feature (iBF) for LRR and DM

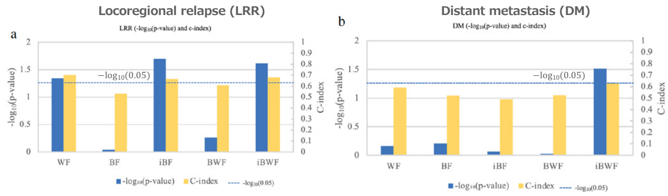
LRR: Locoregional relapse
DM: Distant metastasis



Kodama T, Arimura H, et al. Thorac Cancer. 2022 Aug;13(15):2117-2126

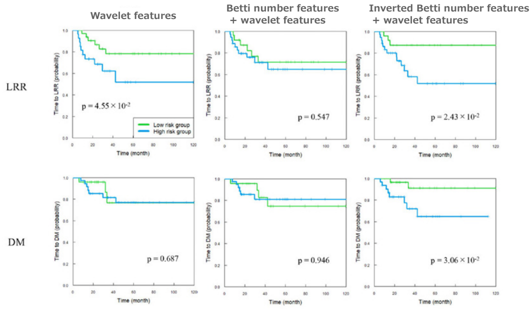
41

Predictability of 5 signatures for locoregional relapse (LRR) and distant metastasis (DM)



42

Prognostic power of inverted BN: Kaplan–Meier curves of time to locoregional relapse (LRR) and distant metastasis (DM)



Kodama T, Arimura H, et al. Thorac Cancer. 2022 Aug;13(15):2117-2126

43

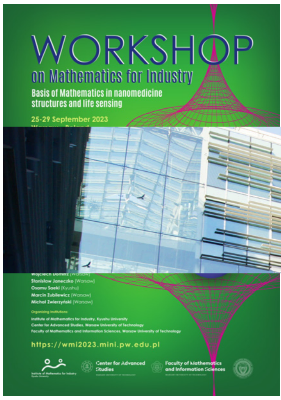
Summary (2)

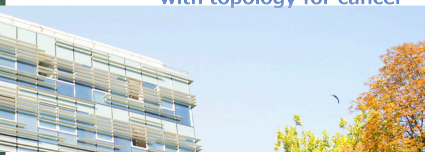
✓Heterogeneous intensities within cancer could come from cancer properties (cell density, necrosis, angiogenesis, gland ducts, etc)

✓Hole analysis (topological radiomics) can be an explainable AI

Any question or comments ?

44





Workshop on mathematics for industry
25-29 September 2023 (Warsaw)
Basis of mathematics in nanomedicine structures and life sensing

Medical imaging signatures with topology for cancer

3. Applications of topological radiomics

Hidetaka Arimura, PhD
Professor (medical Physics)
Division of Medical Quantum Science
Department of Health Sciences, Faculty of Medical Sciences
Kyushu University

Overall outline

1st: Background and radiomics

2nd: Medical background for topological radiomics

✓ 3rd: Applications of topological radiomics

2

Outline

- ✓ Robust radiogenomics approach to identification of EGFR mutations among patients with NSCLC from three different countries using topologically invariant Betti numbers (Ninomiya K, Arimura H, PLOS ONE 2021)
- ✓ Three-dimensional topological radiogenomics of epidermal growth factor receptor Del19 and L858R mutation subtypes on computed tomography images of lung cancer patients (Ninomiya K, Arimura H, Comput Methods Programs Biomed. 2023)
- ✓ Topology-based radiomic features for prediction of parotid gland cancer malignancy grade in magnetic resonance images (Ikushima K, Arimura H, et al. Magnetic Resonance Materials in Physics, Biology and Medicine 2023)
- ✓ Can Persistent Homology Features Capture More Intrinsic Information about Tumors from 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Images of Head and Neck Cancer Patients? (Le QC, Arimura H, et al. Metabolites 2022)
- ✓ Summary (3)

3

Robust radiogenomics approach to identification of EGFR mutations among patients with NSCLC from three different countries using topologically invariant Betti numbers

(Ninomiya K, Arimura H, PLOS ONE 2021)

Kenta Ninomiya[1], Hidetaka Arimura[2*], Wai Yee Chan[3*]
Kentaro Tanaka[4], Shinichi Mizuno[2], Nadia Fareeda Muhammad Gowdh[3]
Nur Adura Yaakup[3], Chong-Kin Liam[5], Chee-Shee Chai[6]
Kwan Hoong Ng[3]

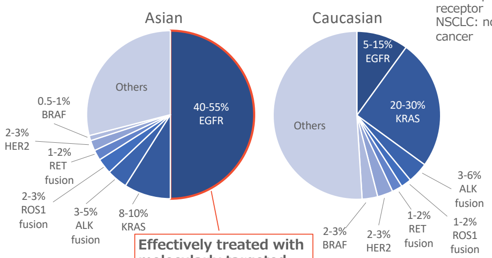
1: Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University
2: Department of Health Sciences, Faculty of Medical Sciences, Kyushu University
3: Department of Biomedical Imaging, Faculty of Medicine, University of Malaya
4: Respiratory Medicine, Kyushu University Hospital
5: Department of Medicine, Faculty of Medicine, University of Malaya
6: Department of Medicine, Faculty of Medicine and Health Science, University Malaysia Sarawak

九州大学 KYUSHU UNIVERSITY
UNIVERSITY OF MALAYA
UNIVERSITI MALAYSIA SARAWAK

Ninomiya K, Arimura H, PLOS ONE 2021

High prevalence rate of EGFR mutations in NSCLC among Asian

EGFR: epidermal growth factor receptor
NSCLC: non-small cell lung cancer



Effectively treated with molecularly targeted drugs; tyrosine kinase inhibitors

Diagnosis for detection of EGFR is performed by "wet" biopsy, i.e., needle biopsy (tissue test) or liquid biopsy (blood test)

[1] K. Li, Onco. Rep.. 37 (2017) 1347–1358.

5

Precision medicine

- ✓ Precision medicine: a form of medicine that uses information about a person's own genes or proteins to prevent, diagnose, or treat diseases
- ✓ Patients with tumors that share a same gene mutation can receive a same drug that targets that mutation, no matter the type of cancer

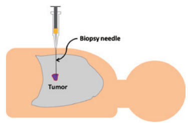
(https://www.cancer.gov/publications/dictionaries/cancer-terms/def/precision-medicine)

6

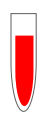
Issues on "wet" biopsy for cancer treatment

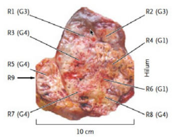
- Some patients may refuse the invasive needle biopsy due to concerns about pneumothorax [refusing rate: around 30-40% (Fukui T, Thoracic Cancer 2019;10:501–7)]
- Liquid biopsy (circulating tumour DNA: ctDNA) has not shown high sensitivities [sensitivity: around 50% (Uchida J, Clin Chem 2015;61(9):1191–1196)]
- A single biopsy of heterogeneous tumors could lead to under- or over-estimation of omics information [Gerlinger M, et al. N Engl J Med 2012;366:883-92]
- Normal tissue should not be sampled

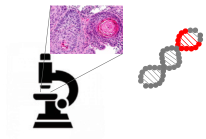
Invasive needle biopsy



Liquid biopsy



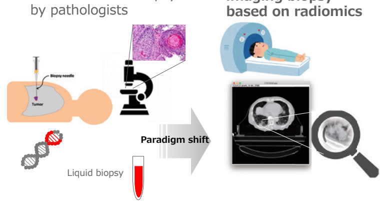




7

Imaging biopsy or "dry" biopsy

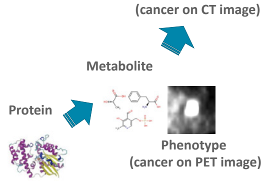
- Non-invasive virtual biopsy that extracts what are equivalent to information obtained from conventional invasive "wet" biopsies, from medical images for cancer treatment
- Computational processes to characterize tumors as well as normal tissues by extracting intrinsic information from medical images

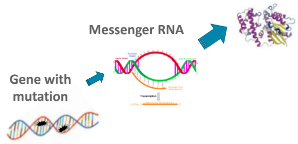


8

What do I assume in radiomics?

- ✓ Since the medical images could be terminals of cascades from our genes, the images may include genetic and protein information.
- ✓ Genetic and protein information are considered as including prognostic information, and thus medical images may also include them.





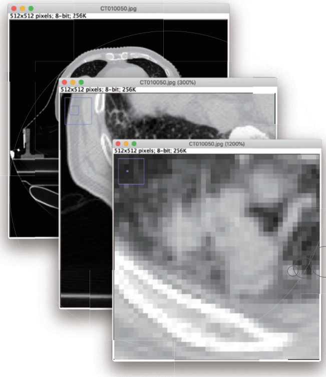
Front Immunol. 2016 Feb 29;7:44. doi: 10.3389/fimmu.2016.00044.

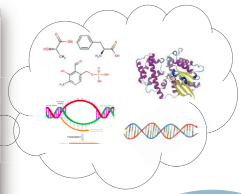
Modified from http://omics.jp/about-omics

http://organicavirtuale.altervista.org/www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/nucacids.htm

9

Nano- or micro-order information might be obtained from medical images …





10

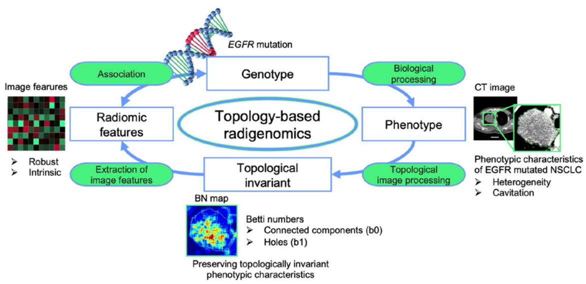
Characteristics of EGFR mutated NSCLC

- Promoted cellular proliferation, differentiation, and migration of non-small cell lung cancer (NSCLC)
- Intra-tumor heterogeneity on contrast enhanced (CE) CT images[1]
- Presence of ground-glass opacity (GGO) causing low intensity cavitation[1-3]

[1] Y. Liu, Radiology. 280 (2016) 271–280. [2] H. Zhang, Int. J. Clin. Oncol. 24 (2019) 649–659. [3] J. Dai, Lung Cancer. 98 (2016) 22–28.

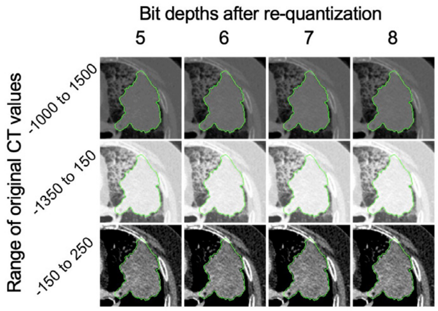
11

Our assumption



15

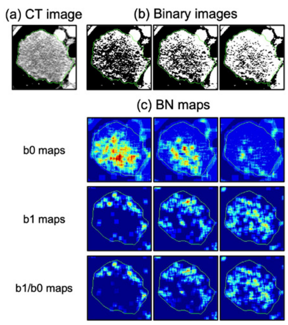
q-bit images generated from CT images



q-bit images generated from CT images using four ranges of Hounsfield units (HU) of CT images (-1000 to 1500 HU, -1350 to 150 HU [lung range window], and -150 to 250 HU [mediastinal range window]) with four bit-depths after re-quantization (5, 6, 7, and 8 bits).

16

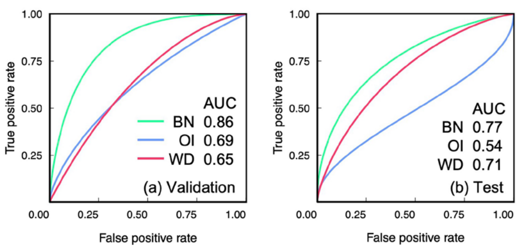
Representative images of CT images, binary images, and Betti number maps



Representative images of
(a) computed tomography (CT) images,
(b) binary images, and
(c) Betti number (BN) maps

17

Receiver operating characteristic (ROC) curves for identification of epidermal growth factor receptor mutants



Receiver operating characteristic (ROC) curves for identification of epidermal growth factor receptor mutants using Betti number (BN)-, original image (OI)- and wavelet decomposition (WD)-based models with area under the ROC curves (AUC) in (a) the validation and (b) the test procedures

18

Three-dimensional topological radiogenomics of epidermal growth factor receptor Del19 and L858R mutation subtypes on computed tomography images of lung cancer patients

(Ninomiya K, Arimura H, Comput Methods Programs Biomed. 2023)

Kenta Ninomiya[1], Hidetaka Arimura[2*], Wai Yee Chan[3*]
Kentaro Tanaka[4], Shinichi Mizuno[2], Nadia Fareeda Muhammad Gowdh[3]
Nur Adura Yaakup[3], Chong-Kin Liam[5], Chee-Shee Chai[6]
Kwan Hoong Ng[3]

1: Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University
2: Department of Health Sciences, Faculty of Medical Sciences, Kyushu University
3: Department of Biomedical Imaging, Faculty of Medicine, University of Malaya
4: Respiratory Medicine, Kyushu University Hospital
5: Department of Medicine, Faculty of Medicine, University of Malaya
6: Department of Medicine, Faculty of Medicine and Health Science, University Malaysia Sarawak

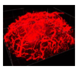
九州大学 KYUSHU UNIVERSITY
UNIVERSITY OF MALAYA
UNIVERSITI MALAYSIA SARAWAK

Relationship between medical evidences and topology

- ✓ Outputs of the AI should be explainable
- ✓ Inputs to AI should be based on medical evidences

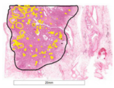
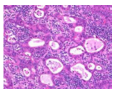
Medical targets

- ✓ Human body
- ✓ Cancer

Biological components in histological images

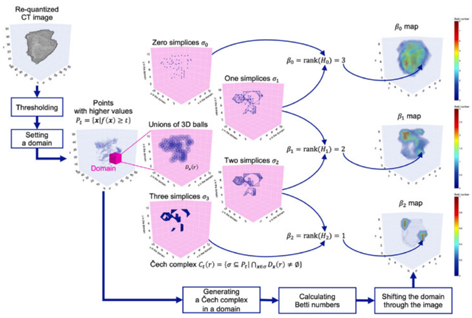
- ✓ Cells
- ✓ Necrosis
- ✓ Gland duct
- ✓ Vessel

Topological components in CT images

- ✓ Islands 
- ✓ Holes (tubes) 
- ✓ Cavities 

3D topological analysis



Topology-based radiomic features for prediction of parotid gland cancer malignancy grade in magnetic resonance images

(Ikushima K, Arimura H, et al. Magnetic Resonance Materials in Physics, Biology and Medicine 2023)

Kojiro Ikushima[1,2], Hidetaka Arimura[3], Ryuji Yasumatsu[4], Hidemi Kamezawa[5], Kenta Ninomiya[6]

1 Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
2 Department of Radiological Technology, Yamaguchi University Hospital, 1-1-1 Minami-kogushi, Ube, Yamaguchi 755-8505, Japan
3 Division of Quantum Radiation Science, Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
4 Department of Otorhinolaryngology-Head and Neck Surgery, Faculty of Medicine, Kindai University, 377-2, Onohigashi, Sayama, Osaka 589-0014, Japan
5 Department of Radiological Technology, Faculty of Fukuoka Medical Technology, Teikyo University, 6-22 Misaki-machi, Omuta, Fukuoka 836-8505, Japan
6 Sanford Burnham Prebys Medical Discovery Institute, 10901, North Torrey Pines Road, La Jolla, CA 92037, USA

KYUSHU UNIVERSITY

26

Parotid gland salivary

- ✓ The three major salivary glands are the parotid gland, submandibular gland, and sublingual glands.
- ✓ Salivary glands produce saliva, and the **gland ducts (tunnels)** function as conduits for delivering the saliva to the oral cavity.[1]
- ✓ Parotid gland cancer (PGC) is a rare form of cancer, accounting for approximately 5% of all head and neck cancers.[2]

1. Amano O, et al. Acta Histochem Cytochem 2012
2. Chang JW, et al. Otolaryngol Head Neck Surg 2015
3. Ho K, et al. Head & Neck Oncology 2011

27

Parotid gland cancer (PGC)

- ✓ The overall survival was significantly worse in patients with high grade cancer than in patients with low to intermediate grade cancer[1].
- ✓ The treatment approaches for parotid cancer depend on the malignancy grade of PGC[2].
- ✓ The malignancy grade is determined by using an invasive fine-needle aspiration cytology (FNAC).
- ✓ Therefore, quantitative and noninvasive approaches such as radiomics are preferable for assessing PGC malignancy.

[1]Honda K et al. *Am J Otolaryngol*, 2018, 39(1), 65-70
[2]Nishikado A et al. *Int J Clin Oncol*, 23(4), 615-624, 2018.

Overall scheme

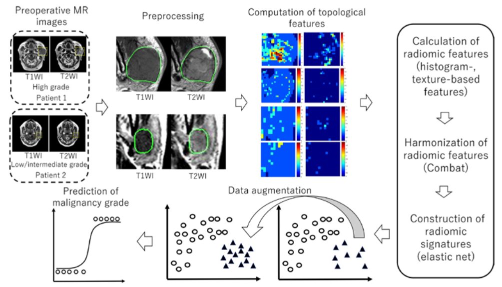
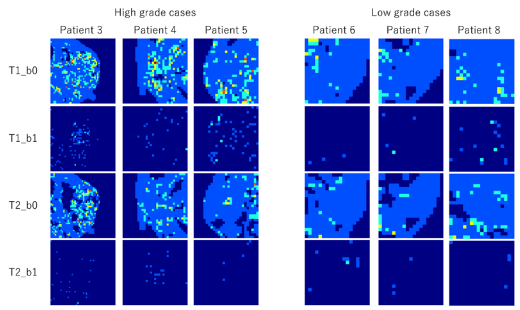


Fig. 1 An overall scheme of the proposed approach. The ROI for a PGC region are shown by green lines. *MR* magnetic resonance, *T1WI* T1-weighted image; T2WI, T2-weighted image

Betti maps from T1 and T2 weighted MR images for 6 patients with high- and low/intermediate-grade PGC



Scatter plots of the most feasible features obtained from topological and conventional features

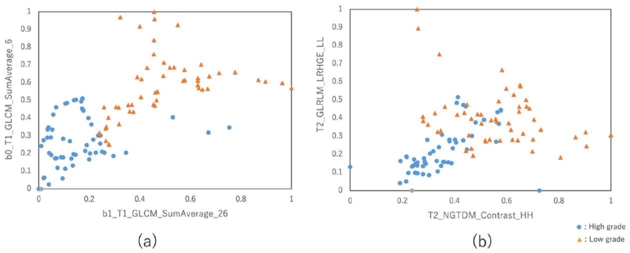


Fig. 7 Scatter plots of radiomic features for 112 training cases obtained from conventional and proposed approaches: **a** topological features and **b** conventional features

Can Persistent Homology Features Capture More Intrinsic Information about Tumors from 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Images of Head and Neck Cancer Patients?

(Le QC, Arimura H, et al. Metabolites 2022)

Quoc Cuong LE [1], Hidetaka ARIMURA [2], Takumi KODAMA [3], Yutaro KABATA [4]

[1] Ho Chi Minh City Oncology Hospital, Ho Chi Minh City, Vietnam
[2] Faculty of Medical Sciences, Kyushu University, Japan
[3] Graduate school of Medical Sciences, Kyushu University, Japan
[4] School of Information and Data Sciences, Nagasaki University, Japan

BỆNH VIỆN UNG BƯỚU THÀNH PHỐ HỒ CHÍ MINH

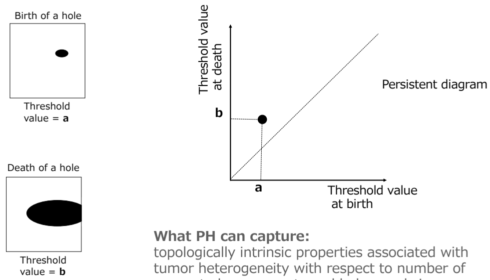
KYUSHU UNIVERSITY

Purpose

This study investigated the feasibility of using PH features for prognostic prediction of patients with HN cancer by using PET/CT images.

This is the first study to examine the potential of PH features on PET/CT images for prognostic prediction of patients with HN cancer.

Persistent homology with considering size



What PH can capture:
topologically intrinsic properties associated with tumor heterogeneity with respect to number of connected components and holes, and size

PH images from PH diagrams

Let $B(b_k, d_k)$ be a PH diagram in the birth–death coordinates $(b_k, d_k), k\epsilon\{1,2,\ldots,N\}$ where N denotes the number of pairs. $B(b_k, d_k)$ was mapped into $B(b_k, p_k = d_k - b_k)$ in the birth–persistence coordinates. Next, $B(b_k, p_k)$ was transformed into PH images, $\rho_B(x, y)$.

$$\rho_B(x,y) = \sum_{k=1}^{N} w_P(p_k)\, G_{B(b_k,p_k)}(x,y)$$

The Gaussian distribution can be expressed as:

$$G_{B(b_k,p_k)}(x,y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x-b_k)^2+(y-p_k)^2}{2\sigma^2}\right]\times B(b_k,p_k)$$

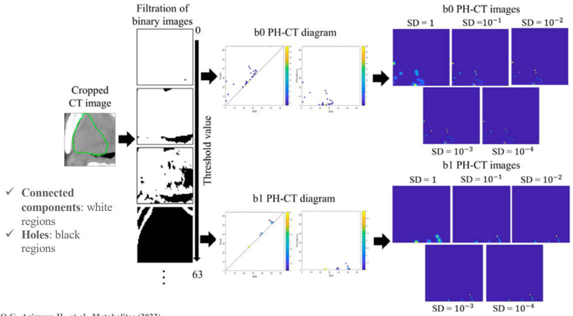
where σ denotes the SD, and x and y are the row and column of a pixel on the PH image, respectively.

A linear weighting function that can adjust the importance of pairs in different regions can be expressed as:

$$w_P(p_k) = \begin{cases} 0, & p_k \le 0 \\ \frac{p_k}{P}, & 0 < p_k < P \\ 1, & p_k \ge P \end{cases}$$

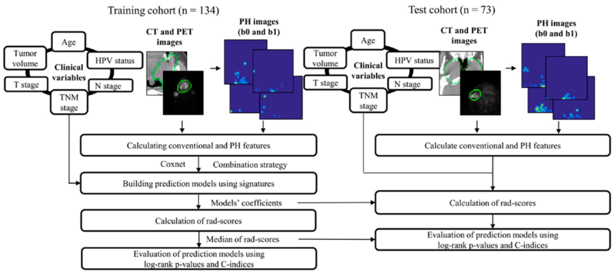
where P is the depth of the quantized CT or PET images (e.g., P = 255 if the images were 8-bit depth).

Persistent homology images



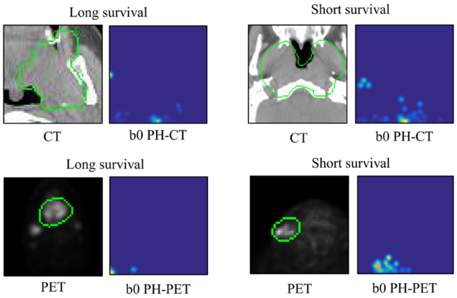
Le Q.C., Arimura H., et al., Metabolites (2022)

Overall workflow



PH: persistence homology, HPV: human papilloma virus; CT: computed tomography; PET: positron emission tomography; rad-score: radiomic score; the green circle represents gross tumor volume.

b0 PH-CT and PH-PET images for long- and short-survival patients

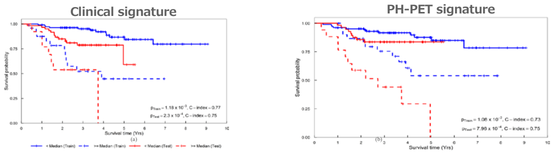


39

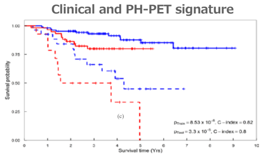
Kaplan–Meier curves obtained from three types of signatures

Clinical signature

PH-PET signature



Clinical and PH-PET signature



PH-PET: persistent homology-positron emission tomography

41

Investigation of repeatability of persistent homology features for patients with lung cancer based on computed tomography images

(Le QC, Arimura H, et al. Metabolites 2022)

Quoc Cuong LE [1], Hidetaka ARIMURA [2], Takumi KODAMA [3], Yutaro KABATA [4]

[1] Ho Chi Minh City Oncology Hospital, Ho Chi Minh City, Vietnam
[2] Faculty of Medical Sciences, Kyushu University, Japan
[3] Graduate school of Medical Sciences, Kyushu University, Japan
[4] School of Information and Data Sciences, Nagasaki University, Japan

BỆNH VIỆN UNG BƯỚU THÀNH PHỐ HỒ CHÍ MINH

KYUSHU UNIVERSITY

Materials and methods

Materials

RIDER dataset*
(Lung cancer)

Test scans ⟷ Re-test scans

- Same patients
- Same imaging protocols
- 15 minutes apart

Methods

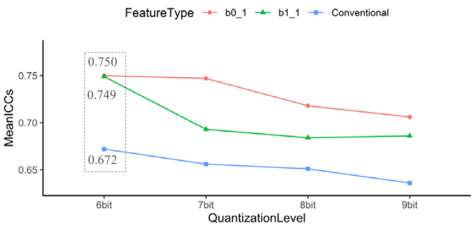
Intraclass correlation coefficients** (ICCs) of image features between two scans:
- A measure of repeatability
- Reflection of both degree of correlation and agreement between measurements

ICCs	< 0.5	0.5 – 0.75	0.75 – 0.9	> 0.9
	Low	Moderate	Good	Excellent

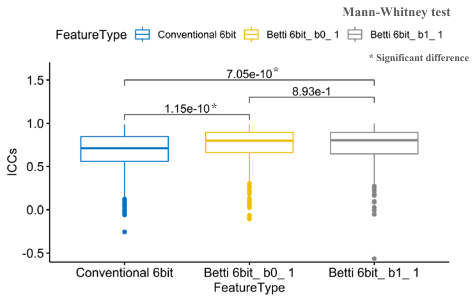
* Zhao B, James L.P, Moskowitz, et al . *Radiology* 2009
** Koo T.K., Li M.Y. *J. Chiropr. Med. 2016*

43

Mean ICCs of image features between two scans for different quantization levels



Comparison of ICCs among three signatures

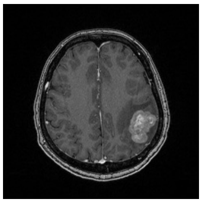


Evidences of usefulness of topology to characterize cancer and adverse events

1. Egashira M, Arimura H, Kobayashi K, Moriyama K, Kodama T, Tokuda T, Ninomiya K, Okamoto H, Igaki H. Magnetic Resonance-Based Imaging Biopsy with Signatures Including Topological Betti Number Features for Prediction of Primary Brain Metastatic Sites. Phys Eng Sci Med. 2023 (accepted on July 21, 2023)
2. Ikushima K, Arimura H, Yasumatsu R, Kamezawa H, Ninomiya K. Topology-based radiomic features for prediction of parotid gland cancer malignancy grade in magnetic resonance images. MAGMA. 2023 Apr 20. doi: 10.1007/s10334-023-01084-0. Epub ahead of print. PMID: 37079154.
3. Ninomiya K, Arimura H, Tanaka K, Chan WY, Kabata Y, Mizuno S, Gowdh NFM, Yaakup NA, Liam CK, Chai CS, Ng KH. Three-dimensional topological radiogenomics of epidermal growth factor receptor Del19 and L858R mutation subtypes on computed tomography images of lung cancer patients. Comput Methods Programs Biomed. 2023 Jun;236:107544. doi: 10.1016/j.cmpb.2023.107544. Epub 2023 Apr 13. PMID: 37148668.
4. Le QC, Arimura H, Ninomiya K, Kodama T, Moriyama T. Can Persistent Homology Features Capture More Intrinsic Information about Tumors from 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Images of Head and Neck Cancer Patients? Metabolites. 2022 Oct 14;12(10):972. doi: 10.3390/metabo12100972. PMID: 36295874; PMCID: PMC9610853.
5. Ninomiya K, Arimura H, Chan WY, et al. Robust radiogenomics approach to the identification of EGFR mutations among patients with NSCLC from three different countries using topologically invariant Betti numbers. PLoS One. 2021;16(1):e0244354. doi: 10.1371/journal.pone.0244354.
6. Ninomiya K, Arimura H, Yoshitake T, et al. Synergistic combination of a topologically invariant imaging signature and a biomarker for the accurate prediction of symptomatic radiation pneumonitis before stereotactic ablative radiotherapy for lung cancer: A retrospective analysis. PLoS One. 2022 Jan 31;17(1):e0263292. doi: 10.1371/journal.pone.0263292.
7. Kodama T, Arimura H, Shirakawa Y, Ninomiya K, Yoshitake T, Shioyama Y. Relapse predictability of topological signature on pretreatment planning CT images of stage I non-small cell lung cancer patients before treatment with stereotactic ablative radiotherapy. Thorac Cancer. 2022 Aug;13(15):2117-2126.
8. Ninomiya K, Arimura H. Homological radiomics analysis for prognostic prediction in lung cancer patients. Phys Med. 2020 Jan;69:90-100. doi: 10.1016/j.ejmp.2019.11.026. Epub 2019 Dec 16. PMID: 31855844.

フッターを追加 46

Feasibility for Prediction of Primary Cancer Sites of Brain Metastases Based on Hessian Index Images (Moriyama K, Arimura H, Medical Imaging and Information Sciences 2022, in Japanese)



Physical examinations, such as invasive needle biopsy, biochemical examinations, and medical imaging (e.g., MRA, CT) to identify primary sites

47

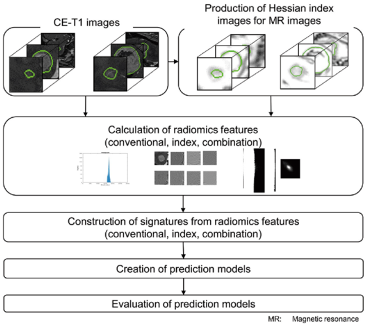
Issues of current invasive examinations for brain metastases

- ✓ Invasive biopsy may not always identify primary sites owing to different cells in brain metastases from those in primary sites, and there could be a risk of tract recurrence after stereotactic needle biopsy of brain metastases
- ✓ Approximately 15% of brain metastases remain unidentified
- ✓ Invasive biopsy can impose a large burden on patients; in particular, it is unsuitable for patients in poor condition

48

Necessary of noninvasive technologies that aid the current invasive examinations or alternative approaches to overcome the issues

Prediction model for primary cancer sites



49

What is the Hessian index to characterize tumor heterogeneity? [Le Q, Arimura H, Sci Rep 2020]

The Hessian matrix of each voxel of a 3D image $I(x,y,z)$ can be given by

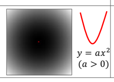
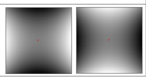
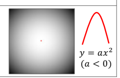
$$\boldsymbol{H} = \begin{pmatrix} I_{xx}(x,y,z,\sigma) & I_{xy}(x,y,z,\sigma) & I_{xz}(x,y,z,\sigma) \\ I_{yx}(x,y,z,\sigma) & I_{yy}(x,y,z,\sigma) & I_{yz}(x,y,z,\sigma) \\ I_{zx}(x,y,z,\sigma) & I_{zy}(x,y,z,\sigma) & I_{zz}(x,y,z,\sigma) \end{pmatrix},$$

where

$$I_{x^i y^j z^k}(x,y,z,\sigma) = \frac{\partial^2}{\partial x^i \partial y^j \partial z^k}[G_\sigma(x,y,z) * I(x,y,z)] = \left[\frac{\partial^2}{\partial x^i \partial y^j \partial z^k} G_\sigma(x,y,z)\right] * I(x,y,z),$$

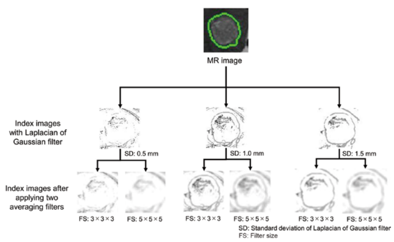
$i, j, k \in N$; $i + j + k = 2$, $G_\sigma(x,y,z)$ is the Gaussian filter with a SD σ, and $*$ denotes convolution

λ_1, λ_2, and λ_3 ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) : first, second, and third eigenvalues of $\boldsymbol{H}$, each of which represents heterogeneity of a tumor
i_H ($0 \leq i_H \leq 3$) : Hessian index (number of negative eigenvalues of $\boldsymbol{H}$)

Hessian index, i_H	0 (Concave)	1 and 2 (Saddle)	3 (Convex)
Approximation of each point within a tumor using quadratic functions	 $y = ax^2$ $(a > 0)$		 $y = ax^2$ $(a < 0)$

50

Results and conclusion



- ✓309 patients (610 T1-weighted contrast-enhanced magnetic resonance images) who have BM were chosen for calculating image features and constructing a light gradient boosting machine model for identification of primary cancer sites.
- ✓The proposed model achieved higher AUCs of 0.77 and 0.66 in a training and test, respectively
- ✓The proposed approach could have a potential for identifying primary cancer sites, but it has already been improved.

51

Summary (3)

- ✓ We believe that the ability of topology for characterizing cancer geometry was found.
- ✓ The ability is extended to prediction of patients' progranoses

Any question or comments ?

52

Singularity theory and its applications to strongly convex multiobjective optimization problems

Shunsuke Ichiki

Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, Japan

A multiobjective optimization problem is a problem to optimize multiple objectives, such as cost, quality, safety and environmental impact in the industrial world. In this mini-course, I would like to introduce theoretical applications of "singularity theory of differentiable mappings", which is a branch of geometry, to strongly convex multiobjective optimization problems.

For this purpose, we first introduce some of basic notions of singularity theory. We also discuss a result called a "parametric transversality theorem", which is an important and fundamental tool in singularity theory for investigating generic mappings. Then, as an application, we give a transversality theorem on linear perturbations. Next, we explain some basic notions of multiobjective optimization and introduce a property of the Pareto set (i.e. the set of optimal solutions) of a strongly convex multiobjective optimization problem from the viewpoint of topology. Finally, based on them, we introduce theoretical applications of singularity theory to strongly convex multiobjective optimization problems.

Singularity theory and its applications to strongly convex multiobjective optimization problems

Shunsuke Ichiki

Tokyo Institute of Technology

WORKSHOP on Mathematics for Industry
Warsaw 25-29 September, 2023

Multiobjective optimization

- X : a set
- $f = (f_1, \ldots, f_\ell) : X \to \mathbb{R}^\ell$: a mapping
- $L = \{1, \ldots, \ell\}$
- $x \in X$: a Pareto solution of f
  $\overset{\text{def}}{\Leftrightarrow}$ there does not exist another point $y \in X$ such that $f_i(y) \leq f_i(x)$ for all $i \in M$ and $f_j(y) < f_j(x)$ for at least one index $j \in M$.
  $\Leftrightarrow$ for any $x' \in X$, either (a) or (b) holds.
  (a) $\forall i \in L$, $f_i(x) = f_i(x')$.
  (b) $\exists i \in L$ s. t. $f_i(x) < f_i(x')$.
- $X^*(f) = \{\, x \in X \mid x$: a Pareto solution of $f \,\}$: the Pareto set of f
- The set $f(X^*(f))$ is called the Pareto front of f.

Multiobjective optimization

- $f = (f_1, \ldots, f_\ell) : X \to \mathbb{R}^\ell$, $L = \{\, 1, \ldots, \ell \,\}$
  The problem of determining $X^*(f)$ is called the problem of minimizing f.
  For a non-empty subset $I = \{\, i_1, \ldots, i_k \,\}$ of L $(i_1 < \cdots < i_k)$, set
  $$f_I = (f_{i_1}, \ldots, f_{i_k}).$$
  The problem of determining $X^*(f_I)$ is called a subproblem of the problem of minimizing f.
- $$\Delta^{\ell-1} = \left\{ (w_1, \ldots, w_\ell) \in \mathbb{R}^\ell \;\middle|\; \sum_{i=1}^{\ell} w_i = 1,\ w_i \geq 0 \right\}.$$
- We also denote a face of $\Delta^{\ell-1}$ for a non-empty subset I of L by
  $$\Delta_I = \{\, (w_1, \ldots, w_\ell) \in \Delta^{\ell-1} \mid w_i = 0 \ (i \notin I) \,\}.$$

Multiobjective optimization

Definition 1 (Simplicial problems, Weakly simplicial problems)

$f : X \to \mathbb{R}^\ell$ $(X \subset \mathbb{R}^m)$, $L = \{1, \ldots, \ell\}$, $r \in \mathbb{Z}_{\geq 0}$ or $r = \infty$

- The problem of minimizing f is C^r simplicial
  $\overset{\text{def}}{\Leftrightarrow}$ $\exists \Phi : \Delta^{\ell-1} \to X^*(f) \subset \mathbb{R}^m$: a C^r mapping
  s. t. $\forall I \subset L$ $(I \neq \emptyset)$, both
  $\Phi|_{\Delta_I} : \Delta_I \to X^*(f_I)$ and
  $f|_{X^*(f_I)} : X^*(f_I) \to f(X^*(f_I))$
  are C^r diffeomorphisms.
- The problem of minimizing f is C^r wealkly simplicial
  $\overset{\text{def}}{\Leftrightarrow}$ $\exists \phi : \Delta^{\ell-1} \to X^*(f) \subset \mathbb{R}^m$: a C^r mapping
  s. t. $\forall I \subset L$ $(I \neq \emptyset)$, $\phi(\Delta_I) = X^*(f_I)$.

Definition 2

X : a convex set of $\mathbb{R}^m$

1. $f : X \to \mathbb{R}$: a strongly convex function
   $\overset{\text{def}}{\Leftrightarrow}$ $\exists \alpha > 0$ s. t. $\forall x, y \in X$, $\forall t \in [0, 1]$,
   $$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x - y\|^2,$$
   where $\|z\|$ is the Euclidean norm of $z \in \mathbb{R}^m$.
   (α : a convexity parameter of f.)
2. $f = (f_1, \ldots, f_\ell) : X \to \mathbb{R}^\ell$: a strongly convex mapping
   $\overset{\text{def}}{\Leftrightarrow}$ f_i is strongly convex for any $i \in L$.

strongly convex $\Longrightarrow$ strictly convex $\Longrightarrow$ convex

Proposition 3

f is strongly convex with a convexity parameter $\alpha > 0$
$\Longleftrightarrow$ *$\exists g : X \to \mathbb{R}$: convex s.t. $f(x) = g(x) + \frac{\alpha}{2}\|x\|^2$*

i.e. (a strongly convex function with $\alpha > 0$) = (a convex function) $+\frac{\alpha}{2}\|x\|^2$

Some results on strongly convex multiobjective optimization problems

Theorem 4

- *$f = (f_1, \ldots, f_\ell) : \mathbb{R}^m \to \mathbb{R}^\ell$: a strongly convex C^r mapping, where $1 \leq r \leq \infty$*

Then, we have the following:

1. *The problem of minimizing f is C^{r-1} weakly simplicial.*
2. *Moreover, if $\operatorname{rank} df_x = \ell - 1$ $(\forall x \in X^*(f))$, then this problem is C^{r-1} simplicial.*

- The case $2 \leq r \leq \infty$: N. Hamada, K. Hayano, S. Ichiki, Y. Kabata and H. Teramoto, *Topology of Pareto sets of strongly convex problems*, SIAM Journal on Optimization, **30** (2020), no. 3, 2659–2686.
- The case $r = 1$: N. Hamada, S. Ichiki, *Simpliciality of strongly convex problems*, Journal of the Mathematical Society of Japan, **73** (2021), no. 3, 965–982.

Some results on strongly convex multiobjective optimization problems

Without differentiability, on C^0 weak simpliciality, the following result is obtained.

Theorem 5

- *$f = (f_1, \ldots, f_\ell) : \mathbb{R}^m \to \mathbb{R}^\ell$: a strongly convex mapping*

Then, the problem of minimizing f is C^0 weakly simplicial.

- Y. Mizota, N. Hamada and S. Ichiki, *All unconstrained strongly convex problems are weakly simplicial*, available from arXiv:2106.12704.

Remark 1

The result has recently been applied to engineering, and the application was introduced by the following talk:
Naoki Hamada, Brief Introduction to Topology for Multi-objective Optimization

Dropping the rank assumption of the theorem

Remark 2

For the proof, we use the mapping $x^* : \Delta^{\ell-1} \to X^*(f)$ defined by

$$x^*(w) = \arg\min_{x \in \mathbb{R}^m} \left(\sum_{i=1}^{\ell} w_i f_i(x) \right).$$

Remark 3

- $f = (f_1, f_2) : \mathbb{R} \to \mathbb{R}^2$ $(f_1(x) = f_2(x) = x^2)$
  Then, we have the following.
  - f : a strongly convex mapping of class C^∞
  - Since $\operatorname{rank} df_0 = 0$ and $X^*(f) = \{0\}$, f does not satisfy the assumption of the theorem.

  Since $X^*(f) = \{0\}$, f is not C^0 simplicial.

A theoretical application of Singularity Theory to multiobjective optimization

Proposition 6 (I)

- *$f : \mathbb{R}^m \to \mathbb{R}^\ell$ $(m \geq \ell)$: a C^2 mapping*
- *$\Sigma = \{\pi \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell) \mid \exists x \in \mathbb{R}^m$ s. t. $\operatorname{rank} d(f+\pi)_x \leq \ell - 2\}$*

If $m - 2\ell + 4 > 0$, then Σ has Lebesgue measure zero.

Lemma 7

- *$f : \mathbb{R}^m \to \mathbb{R}^\ell$: a strongly convex mapping*

Then, $\forall \pi \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell)$, $f + \pi : \mathbb{R}^m \to \mathbb{R}^\ell$ is also strongly convex.

Theorem 8 (Hamada, Hayano, Kabata, Teramoto, I)

- *$f : \mathbb{R}^m \to \mathbb{R}^\ell$ $(m \geq \ell)$: a strongly convex C^r mapping, where $2 \leq r \leq \infty$*
- *Let Σ be the set defined by*

$$\{\pi \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell) \mid \text{The problem of minimizing } f+\pi \text{ is not } C^{r-1} \text{ simplicial}\}.$$

If $m - 2\ell + 4 > 0$, then Σ has Lebesgue measure zero.

Proposition 9 (I)

- $f : \mathbb{R}^m \to \mathbb{R}^\ell$ $(m \geq \ell)$: *a* C^2 *mapping.*
- $\Sigma = \{\, \pi \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell) \mid \exists x \in \mathbb{R}^m \text{ s. t. } \operatorname{rank} d(f + \pi)_x \leq \ell - 2 \,\}$

If $m - 2\ell + 4 > 0$, *then for any non-negative real number* s *satisfying*

$$s > m\ell - (m - 2\ell + 4),$$

it follows that $\mathcal{H}^s(\Sigma) = 0$, *and thus,*

$$\begin{cases} \Sigma = \varnothing & \text{if } \ell = 1, \\ \dim_H \Sigma \leq m\ell - (m - 2\ell + 4) & \text{if } \ell \geq 2. \end{cases}$$

- S. Ichiki, *A refined version of parametric transversality theorems*, Journal of Geometric Analysis, **32** (2022), no. 9, Paper No. 234, 14 pp.

Theorem 10 (I)

- $f : \mathbb{R}^m \to \mathbb{R}^\ell$ $(m \geq \ell)$: *a strongly convex* C^r *mapping* $(2 \leq r \leq \infty)$
- *Let* Σ *be the set defined by*

  $\{\, \pi \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^\ell) \mid$ *The problem of minimizing* $f + \pi$ *is not* C^{r-1} *simplicial* $\}$

If $m - 2\ell + 4 > 0$, *then*

$$\begin{cases} \Sigma = \varnothing & \text{if } \ell = 1, \\ \dim_H \Sigma \leq m\ell - (m - 2\ell + 4) & \text{if } \ell \geq 2. \end{cases}$$

Example 11

- $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x_1, x_2) = (x_1^2 + x_2^2, x_1^2 + x_2^2) \leftarrow C^\infty$ strongly convex
- Let Σ be the set defined by
  $\{\, \pi \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2) \mid$ The problem of minimizing $f + \pi$ is not C^∞ simplicial $\}$

By the above theorem, $\dim_H \Sigma \leq 2$.
By a direct calculation, $\Sigma = \{\, \pi = (\pi_1, \pi_2) \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2) \mid \pi_1 = \pi_2 \,\}$.
Since $\dim_H \Sigma = 2$, we cannot improve the evaluation "≤ 2".

- S. Ichiki, *A refined version of parametric transversality theorems*, Journal of Geometric Analysis, **32** (2022), no. 9, Paper No. 234, 14 pp.

Explanatory Model Analysis

Przemysław Biecek
MI^2 Data Lab, Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland

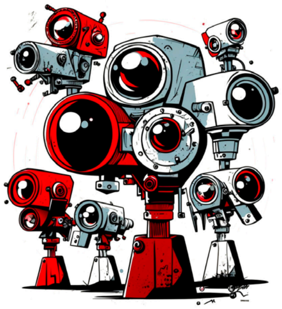
Shapley values currently stand as the most widely employed technique for conducting Explanatory Model Analysis (EMA) and achieving Explainable Artificial Intelligence (XAI). Ongoing efforts are focused on crafting modifications and extensions to adapt this method to address the diverse challenges posed by a wide array of applications. In this presentation, I will illustrate instances where Shapley values, and by extension, techniques utilized in explainable artificial intelligence, prove effective in distinguishing models exhibiting distinct behaviors, even if their performance appears identical at first glance. Subsequently, I will present a proposal for an iterative model analysis process utilizing Shapley values. Drawing inspiration from Rashomon perspectives, this approach, termed Shapley Lenses, provides a more nuanced perspective on predictive models. The insights derived from predictive models can then be leveraged to construct subsequent iterations of models with enhanced interpretability.

MI REDTEAM

Explanatory Model Analysis

aka Explainable AI

Przemysław Biecek
/ˈpʂɛ.mɛk/



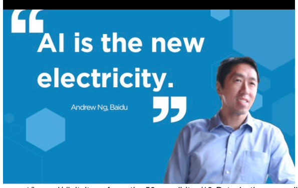
WORKSHOP on Mathematics for Industry 2023 Przemyslaw Biecek September 25, 2023

MI REDTEAM

Why?

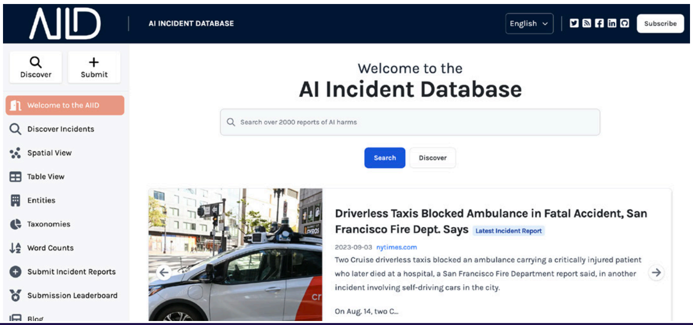
MI REDTEAM

Hype for AI is growing



https://www.slideshare.net/jaypod/digitaltransformation50soundbites/19-Data_is_the_new_oilClive
https://www.newworldai.com/forget-the-hype-what-every-business-leader-needs-to-know-about-artificial-intelligence-now/

but AI is broken



https://incidentdatabase.ai/

but AI is broken

Image	Question	Answer
	Does this person value marriage?	No
	Does this person like algebra?	No
	Does this person enjoy sculpture?	Yes

Table 2: Example of stereotypical question-image pairs.

https://www.slideshare.net/jaypod/digitaltransformation50soundbites/19-Data_is_the_new_oilClive
https://www.newworldai.com/forget-the-hype-what-every-business-leader-needs-to-know-about-artificial-intelligence-now/
https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

but AI is broken

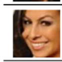
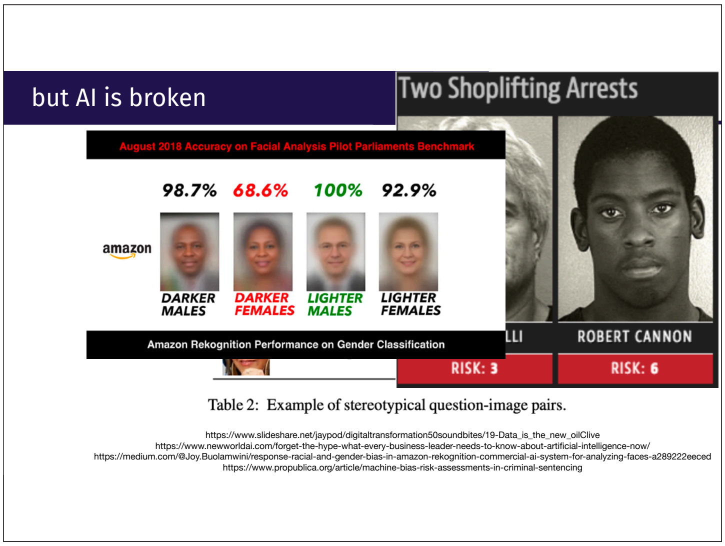
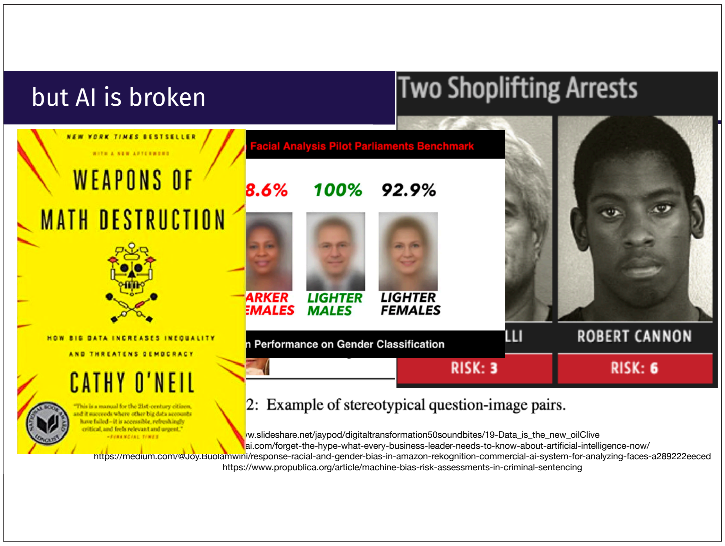
Image	Question
	Does this person v
	Does this person l
	Does this person e

Table 2: Example of stereotypical question-image pairs.

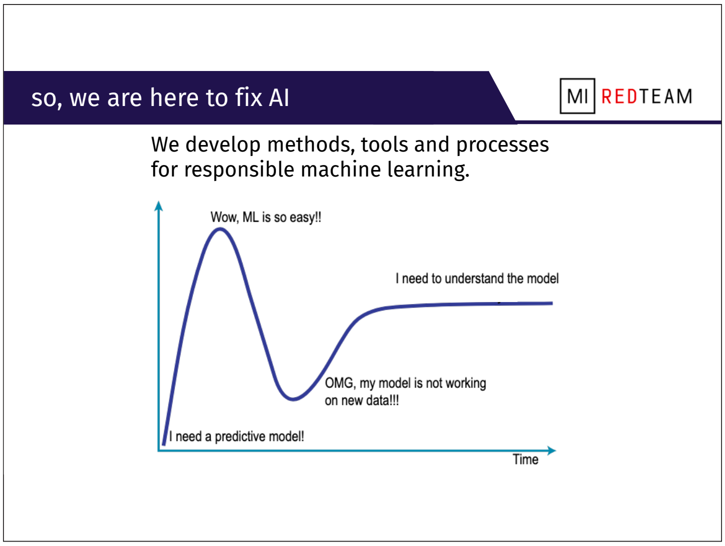
https://www.slideshare.net/jaypod/digitaltransformation50soundbites/19-Data_is_the_new_oilClive
https://www.newworldai.com/forget-the-hype-what-every-business-leader-needs-to-know-about-artificial-intelligence-now/
https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing



but AI is broken
Two Shoplifting Arrests
August 2018 Accuracy on Facial Analysis Pilot Parliaments Benchmark
98.7% 68.6% 100% 92.9%
amazon
DARKER MALES
DARKER FEMALES
LIGHTER MALES
LIGHTER FEMALES
Amazon Rekognition Performance on Gender Classification
ROBERT CANNON
RISK: 3
RISK: 6
Table 2: Example of stereotypical question-image pairs.
https://www.slideshare.net/jaypod/digitaltransformation50soundbites/19-Data_is_the_new_oilClive
https://www.newworldai.com/forget-the-hype-what-every-business-leader-needs-to-know-about-artificial-intelligence-now/
https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing



but AI is broken
Two Shoplifting Arrests
NEW YORK TIMES BESTSELLER
WEAPONS OF MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY
CATHY O'NEIL
100% 92.9%
LIGHTER MALES
LIGHTER FEMALES
ROBERT CANNON
RISK: 3
RISK: 6
https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing



so, we are here to fix AI
MI REDTEAM
We develop methods, tools and processes
for responsible machine learning.
Wow, ML is so easy!!
I need to understand the model
OMG, my model is not working
on new data!!!
I need a predictive model!
Time

How?

Shapley values

Problem A:

Set S with n players cooperating in a game can earn a reward v(S).

How to divide this reward fairly among the players?

Shapley values

A VALUE FOR n-PERSON GAMES

L. S. Shapley

18 March 1952

THEOREM. A unique value function ϕ exists satisfying Axioms 1 - 3, for games with finite carriers; it is given by the formula

$$\phi_i[v] = \sum_{S \subseteq N} \gamma_n(s)\,[v(S) - v(S-(i))] \quad (\text{all } i \in U), \tag{13}$$

where N is any finite carrier of v.

Proof. (13) follows from (11), (12), and Lemma 1. We note that (13), like (10), does not depend on the particular finite carrier N ; the ϕ of the theorem is therefore well defined. By its derivation it is clearly the only value function which could satisfy the axioms. That it does in fact satisfy the axioms is easily

Shapley values

$$\phi_j = \sum_{S \subseteq P/\{j\}} \frac{|S|!(|P|-|S|-1)!}{|P|!}(v(S \cup \{j\}) - v(S))$$

Shapley, Lloyd S. A Value for n-Person Games. Princeton University Press. 1952

Shapley values

MI REDTEAM

- **Efficiency**: all contributions sum up to the final reward

$$\sum_j \phi_j = v(P)$$

- **Symmetry**: if players i and j contributed in the same way to each coalition then they get the same reward

$$\forall_S v(S \cup \{i\}) = v(S \cup \{j\}) \Rightarrow \phi_i = \phi_j$$

- **Dummy**: if player i does not contribute then its reward is 0

$$\forall_S v(S \cup \{i\}) = v(S) \Rightarrow \phi_i = 0$$

- **Additivity**: reward in sum of games v_1 and v_2 is sum of rewards

$$\forall_S v(S) = v_1(S) + v_2(S) \Rightarrow \phi_i = \phi_{1,i} + \phi_{2,i}$$

Shapley values

$$\phi_j = \sum_{S \subseteq P/\{j\}} \frac{|S|!(|P| - |S| - 1)!}{|P|!}(v(S \cup \{j\}) - v(S))$$

Shapley, Lloyd S. A Value for n-Person Games. Princeton University Press. 1952

Shapley values for ML models

MI REDTEAM

Problem B:

In machine learning, we train a function f(x) : $R^p \rightarrow R$ that calculates predictions based on p variables.

How to quantify the effect of each variable on the final prediction?

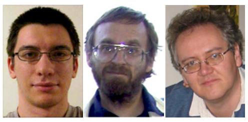
Shapley values for ML models

MI REDTEAM

Explaining instance classifications with interactions of subsets of feature values

E. Štrumbelj *, I. Kononenko, M. Robnik Šikonja

University of Ljubljana, Faculty of Computer and Information Science, Tržaška cesta 25, 1000 Ljubljana, Slovenia



Theorem 1 *For the game $\langle N, v\rangle$ there exists a unique solution ϕ, which satisfies axioms 1 to 4 and it is the Shapley value:*

$$Sh_i(v) = \sum_{S \subseteq N \setminus \{i\}, s=|S|} \frac{(n-s-1)!s!}{n!}(v(S \cup \{i\}) - v(S)), \quad i = 1, \ldots, n.$$

Proof For a detailed proof of this theorem refer to Shapley's paper (1953). ∎

Erik Strumbelj, Igor Kononenko, Marko Robnik-Sikonja. Explaining instance classifications with interactions of subsets of feature values. Data & Knowledge Engineering, 2009

Shapley values for ML models

MI REDTEAM

Example:

We have a complex predictive model f(x) that predicts the probability of surviving the Titanic disaster.

How do we calculate contributions of individual variables to the prediction for a single passenger.

Here: an eight-year-old boy travelling 1st class.

Explanatory Model Analysis. Przemyslaw Biecek, Tomasz Burzykowski. 2021. CRC. https://ema.drwhy.ai

Shapley values for ML models

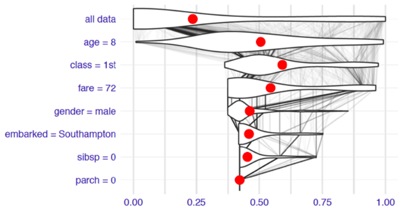
MI REDTEAM

Example:

We have a complex predictive model f(x) that predicts the probability of surviving the Titanic disaster.

How do we calculate contributions of individual variables to the prediction for a single passenger.

Here: an eight-year-old boy travelling 1st class.



$$f(\underline{X})|X^1 = x_*^1, \ldots, X^j = x_*^j$$

Explanatory Model Analysis. Przemyslaw Biecek, Tomasz Burzykowski. 2021. CRC. https://ema.drwhy.ai

Shapley values for ML models

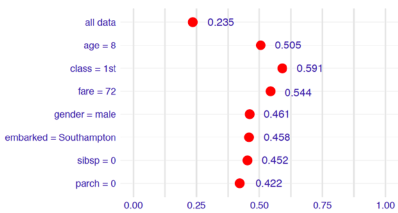
MI REDTEAM

Example:

We have a complex predictive model f(x) that predicts the probability of surviving the Titanic disaster.

How do we calculate contributions of individual variables to the prediction for a single passenger.

Here: an eight-year-old boy travelling 1st class.



$$v(j, \underline{x}_*) = E_{\underline{X}}\{f(\underline{X})|X^1 = x_*^1, \ldots, X^j = x_*^j\}$$

Explanatory Model Analysis. Przemyslaw Biecek, Tomasz Burzykowski. 2021. CRC. https://ema.drwhy.ai

Shapley values for ML models

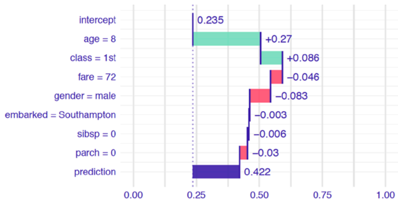
MI REDTEAM

Example:

We have a complex predictive model f(x) that predicts the probability of surviving the Titanic disaster.

How do we calculate contributions of individual variables to the prediction for a single passenger.

Here: an eight-year-old boy travelling 1st class.



$$f(\underline{x}_*) = v_0 + \sum_{j=1}^{p} v(j, \underline{x}_*)$$

Explanatory Model Analysis. Przemyslaw Biecek, Tomasz Burzykowski. 2021. CRC. https://ema.drwhy.ai

SHAP as an unification of LIME, DeepLIFT, Layer-Wise Relevance Propagation

MI REDTEAM

A Unified Approach to Interpreting Model Predictions

Scott M. Lundberg
Paul G. Allen School of Computer Science
University of Washington
Seattle, WA 98105
slund1@cs.washington.edu

Su-In Lee
Paul G. Allen School of Computer Science
Department of Genome Sciences
University of Washington
Seattle, WA 98105
suinlee@cs.washington.edu

Definition 1 Additive feature attribution methods *have an explanation model that is a linear function of binary variables:*

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i, \tag{1}$$

where $z' \in \{0,1\}^M$, *M is the number of simplified input features, and* $\phi_i \in \mathbb{R}$.

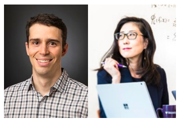
Lundberg, Scott M, Su-In Lee. A Unified Approach to Interpreting Model Predictions. NeurIPS. 2017

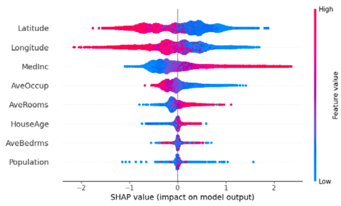
TreeSHAP as quick SHAP for Tree Ensembles

MI REDTEAM

Consistent Individualized Feature Attribution for Tree Ensembles

Scott M. Lundberg, Gabriel G. Erion, and Su-In Lee
University of Washington
{slund1,erion,suinlee}@uw.edu





Lundberg, Scott M,Gabriel G. Erion, Su-In Lee. Consistent Individualized Feature Attribution for Tree Ensembles. NeurIPS. 2019

SurvSHAP(t) for survival models

MI REDTEAM

Problem C:

In survival modelling, the model output is
a survival function S(t).

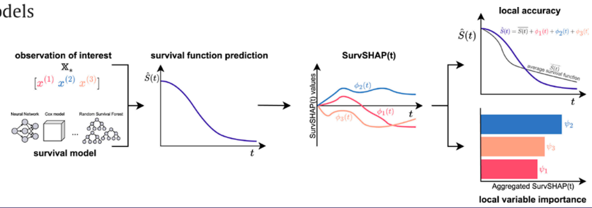
How to quantify the effect of each variable on the final prediction (which is a function)?

SurvSHAP(t) for survival models

MI REDTEAM

Knowledge-Based Systems
Volume 262, 28 February 2023, 110234

SurvSHAP(t): Time-dependent explanations of machine learning survival models



Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

How to calculate SurvSHAP(t) values

MI REDTEAM

Contribution of variable ***d*** in time point ***t*** for the patient ***x***:

$$\phi_t(\mathbf{x}_*, d) = \frac{1}{|\Pi|}\sum_{\pi\in\Pi} e_{t,\mathbf{x}_*}^{\text{before}(\pi,d)\cup\{d\}} - e_{t,\mathbf{x}_*}^{\text{before}(\pi,d)}$$

$$e_{t,\mathbf{x}_*}^{D} = \mathbb{E}[\hat{S}(t,\mathbf{x})|\mathbf{x}^D = \mathbf{x}_*^D]$$

Local variable importance of variable ***d*** for the patient ***x***:

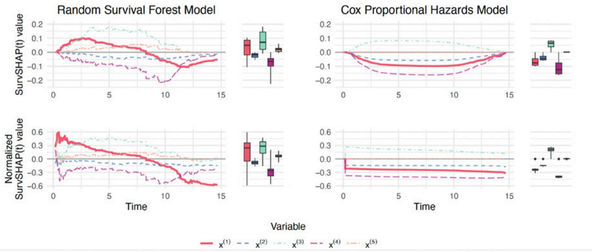
$$\psi(\mathbf{x}_*, d) = \int_0^{t_{max}} |\phi_t(\mathbf{x}_*, d)|\,\mathrm{d}w(t)$$

Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

SurvSHAP(t) can detect time-dependent variable effects

MI REDTEAM

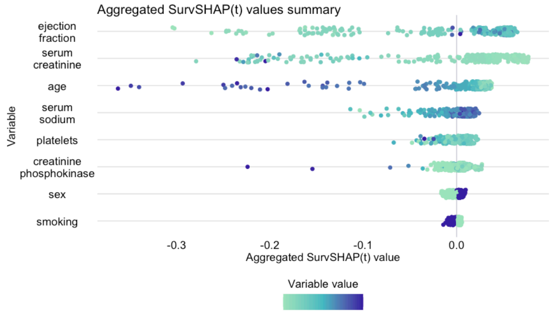
$$h(t) = h_0(t) \cdot \exp\{\boxed{[-0.9 + 0.1t + 0.9\log(t)]x^{(1)}} \\ +0.5x^{(2)} - 0.2x^{(3)} + 0.1x^{(4)} + 10^{-6}x^{(5)}\}$$



Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

SurvSHAP(t) show global variable importance

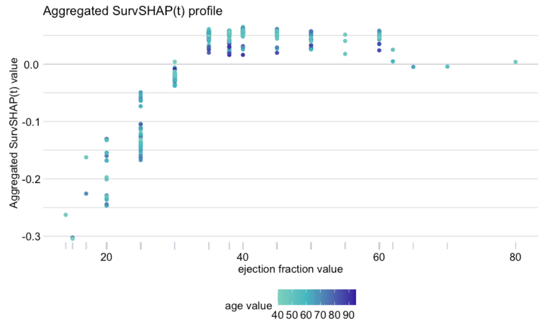
MI REDTEAM



Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

SurvSHAP(t) shows the dependence of variable attributions on its values

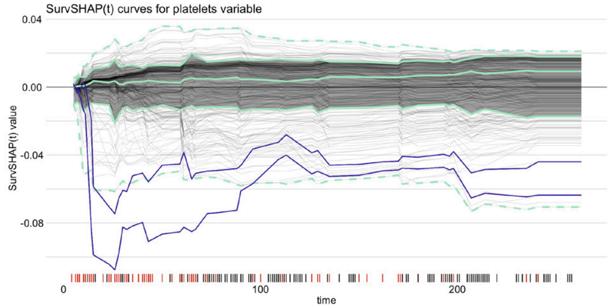
MI REDTEAM



Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

SurvSHAP(t) can be analyzed using functional data analysis technique

MI REDTEAM



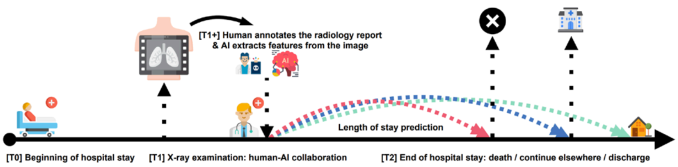
Mateusz Krzyziński, Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek.
SurvSHAP(t): Time-dependent explanations of machine learning survival models. Knowledge-Based Systems. 2023

Application of SurvSHAP(t)

MI REDTEAM

Hospital Length of Stay Prediction Based on Multi-modal Data towards Trustworthy Human-AI Collaboration in Radiomics

Hubert Baniecki[1,2][0000−0001−6661−5364], Bartlomiej Sobieski[2], Przemysław Bombiński[2,3], Patryk Szatkowski[2,3], and Przemysław Biecek[1,2][0000−0001−8423−1823]

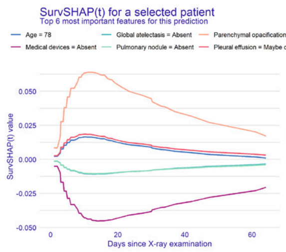


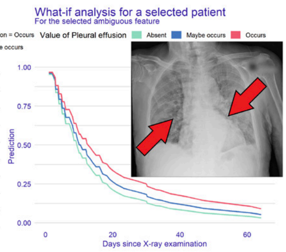
Hubert Baniecki, Bartlomiej Sobieski, Przemyslaw Bombinski, Patryk Szatkowski, Przemyslaw Biecek. Hospital Length of Stay Prediction Based on Multi-modal Data towards Trustworthy Human-AI Collaboration in Radiomics. AIME. 2023

Application of SurvSHAP(t)

MI REDTEAM

Hospital Length of Stay Prediction Based on Multi-modal Data towards Trustworthy Human-AI Collaboration in Radiomics





Hubert Baniecki, Bartlomiej Sobieski, Przemyslaw Bombinski, Patryk Szatkowski, Przemyslaw Biecek. Hospital Length of Stay Prediction Based on Multi-modal Data towards Trustworthy Human-AI Collaboration in Radiomics. AIME. 2023

survex: open software for XAI+surv

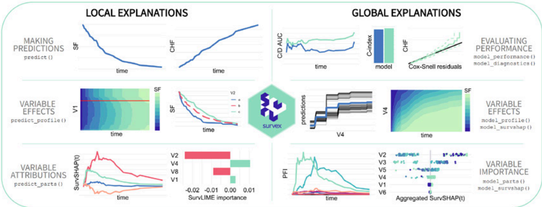
MI REDTEAM

survex: an R package for explaining machine learning survival models

Mikołaj Spytek,[1] Mateusz Krzyziński,[1] Sophie Hanna Langbein,[2,3] Hubert Baniecki,[1,4] Marvin N. Wright[2,3,5] and Przemysław Biecek[1,4*]

[1]MI2.AI, Warsaw University of Technology, Poland, [2]Leibniz Institute for Prevention Research and Epidemiology – BIPS, Germany, [3]Faculty of Mathematics and Computer Science, University of Bremen, Germany, [4]MI2.AI, University of Warsaw, Poland and [5]Section of Biostatistics, Department of Public Health, University of Copenhagen, Denmark

*Corresponding author. przemyslaw.biecek@pw.edu.pl



When?

MI REDTEAM

Rashomon set

MI REDTEAM

Statistical Science
2001, Vol. 16, No. 3, 199–231

Statistical Modeling: The Two Cultures

Leo Breiman

8. RASHOMON AND THE MULTIPLICITY OF GOOD MODELS

Rashomon is a wonderful Japanese movie in which four people, from different vantage points, witness an incident in which one person dies and another is supposedly raped. When they come to testify in court, they all report the same facts, but their stories of what happened are very different.

What I call the Rashomon Effect is that there is often a multitude of different descriptions [equations $f(\mathbf{x})$] in a class of functions giving about the same minimum error rate. The most easily understood example is subset selection in linear regression. Suppose there are 30 variables and we want to find the best five variable linear regressions. There are about 140,000 five-variable subsets in competition. Usually we pick the one with the lowest residual sum-of-squares (RSS), or, if there is a test set,

Picture 1

$$y = 2.1 + 3.8x_3 - 0.6x_8 + 83.2x_{12} - 2.1x_{17} + 3.2x_{27},$$

Picture 2

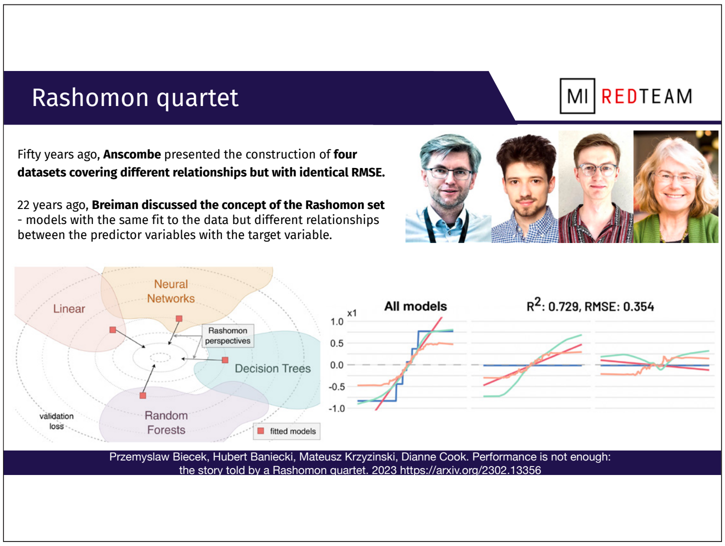
$$y = -8.9 + 4.6x_5 + 0.01x_6 + 12.0x_{15} + 17.5x_{21} + 0.2x_{22},$$

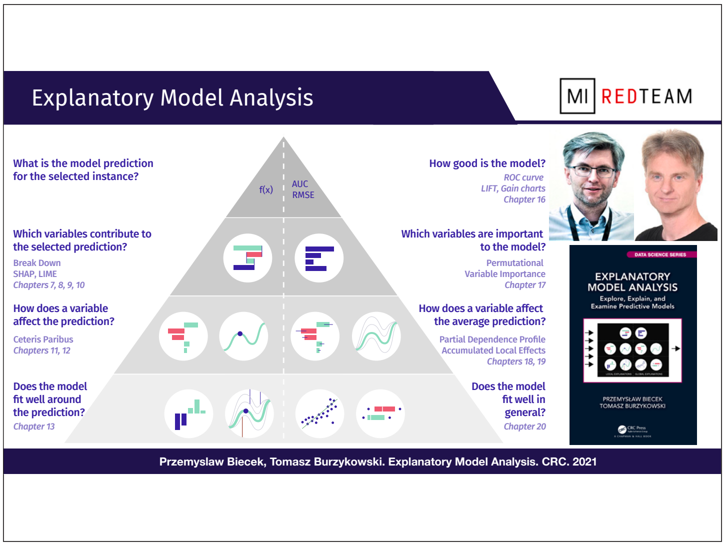
Picture 3

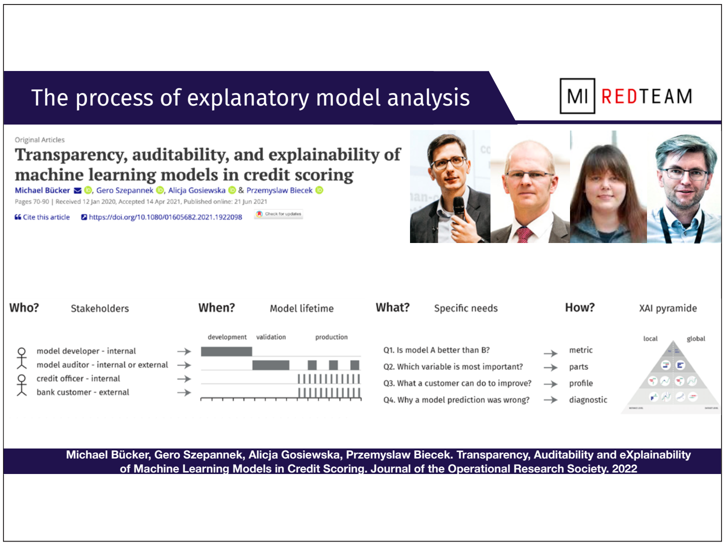
$$y = -76.7 + 9.3x_2 + 22.0x_7 - 13.2x_8 + 3.4x_{11} + 7.2x_{28}.$$

Which one is better? The problem is that each one tells a different story about which variables are important.

Leo Breiman. Statistical Modeling: The Two Cultures. Statistical Science. 2001

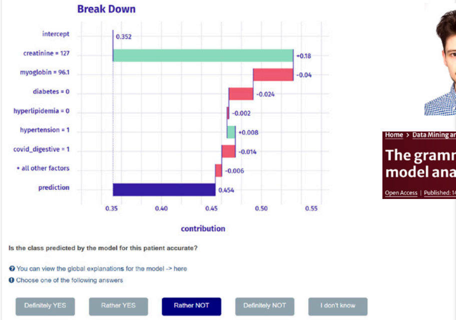

Rashomon quartet
MI REDTEAM
Fifty years ago, Anscombe presented the construction of four datasets covering different relationships but with identical RMSE.
22 years ago, Breiman discussed the concept of the Rashomon set - models with the same fit to the data but different relationships between the predictor variables with the target variable.
Linear
Neural Networks
Rashomon perspectives
Decision Trees
Random Forests
validation loss
fitted models
All models
R^2: 0.729, RMSE: 0.354
x1
1.0
0.5
0.0
-0.5
-1.0
Przemyslaw Biecek, Hubert Baniecki, Mateusz Krzyzinski, Dianne Cook. Performance is not enough: the story told by a Rashomon quartet. 2023 https://arxiv.org/2302.13356


Explanatory Model Analysis
MI REDTEAM
What is the model prediction for the selected instance?
f(x)
AUC
RMSE
How good is the model?
ROC curve
LIFT, Gain charts
Chapter 16
Which variables contribute to the selected prediction?
Break Down
SHAP, LIME
Chapters 7, 8, 9, 10
Which variables are important to the model?
Permutational Variable Importance
Chapter 17
How does a variable affect the prediction?
Ceteris Paribus
Chapters 11, 12
How does a variable affect the average prediction?
Partial Dependence Profile
Accumulated Local Effects
Chapters 18, 19
Does the model fit well around the prediction?
Chapter 13
Does the model fit well in general?
Chapter 20
EXPLANATORY MODEL ANALYSIS
Explore, Explain, and Examine Predictive Models
PRZEMYSLAW BIECEK
TOMASZ BURZYKOWSKI
Przemyslaw Biecek, Tomasz Burzykowski. Explanatory Model Analysis. CRC. 2021


The process of explanatory model analysis
MI REDTEAM
Original Articles
Transparency, auditability, and explainability of machine learning models in credit scoring
Michael Bücker, Gero Szepannek, Alicja Gosiewska & Przemyslaw Biecek
Pages 70-90 | Received 12 Jan 2020, Accepted 14 Apr 2021, Published online: 21 Jun 2021
Cite this article
https://doi.org/10.1080/01605682.2021.1922098
Who? Stakeholders
When? Model lifetime
What? Specific needs
How? XAI pyramide
model developer - internal
model auditor - internal or external
credit officer - internal
bank customer - external
development
validation
production
Q1. Is model A better than B?
Q2. Which variable is most important?
Q3. What a customer can do to improve?
Q4. Why a model prediction was wrong?
metric
parts
profile
diagnostic
local
global
Michael Bücker, Gero Szepannek, Alicja Gosiewska, Przemyslaw Biecek. Transparency, Auditability and eXplainability of Machine Learning Models in Credit Scoring. Journal of the Operational Research Society. 2022

EMA process validated with user-studies

MI REDTEAM



Hubert Baniecki, Dariusz Parzych, Przemyslaw Biecek. The grammar of interactive explanatory model analysis. Data Mining and Knowledge Discovery. 2023

EMA process validated with user-studies

MI REDTEAM

Q_4: Which of the following aspects had the greatest impact on your decision making in the presented patient case?

Answer	Frequency
Break-down explanation (1st screen)	16.7%
Ceteris Paribus "What-if?" explanation (2nd screen)	27.5%
Shapley Values explanation or/and an additional Ceteris Paribus "What-if?" explanation (3rd screen)	35.3%
Comparison of the local explanations with the global explanations	19.2%
My answer was random, I ran out of information to make a decision	0.5%
Other (three descriptive answers in total: a Permutational Importance explanation, both Ceteris Paribus explanations, a high residual value)	0.8%

Table 5 Frequency of answers for Q_4 averaged across 12 cases times 30 participants.

Hubert Baniecki, Dariusz Parzych, Przemyslaw Biecek. The grammar of interactive explanatory model analysis. Data Mining and Knowledge Discovery. 2023

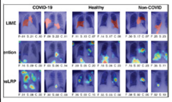
MI REDTEAM

Who?

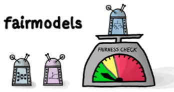
MI2.AI is here to fix AI

MI REDTEAM

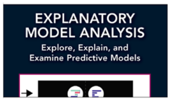
Papers



Software



Books



Teaching

Linear instability of Prandtl spirals

Tomasz Cieślak
Institute of Mathematics, Polish Academy of Sciences, Poland

We review a recent result with P.Kokocki and W.Ożański stating that the union of three or more uniformly distributed Prandtl spirals is linearly unstable as a solution to the Birkhoff-Rott equation. First, a linearization of the Birkhoff-Rott equation around the Prandtl spirals is found. Next, a perturbation leading to the instability is shown. Notice that, unlike for the flat sheet, the unstable modes grow only algebraically in time. In our talk we partially answer the question of Helmholtz from his famous 1868 paper on discontinuous flows.

① NO. DATE

TOMASZ CIEŚLAK

Kościelisko, 13.09.2023

LINEAR INSTABILITY OF PRANDTL SPIRALS.

The talk concerns vortex sheets. Similarly to other discontinuous flows introduced in 1868 classical paper of Hermann von Helmholtz

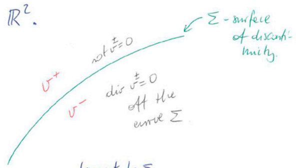
Other objects introduced in the paper were wakes or cavities. With important followup of Kirchhoff '1869 & Levi-Civita '1907.

KOKUYO

---

②

VORTEX SHEET IS A FLOW IN $\mathbb{R}^2$.



$$\left(v_+ - v_-\right)\big|_{\Sigma} \cdot \vec{\tau} = \gamma$$ [density of a measure supported on Σ]

tangent to Σ

$$\left(v_+ - v_-\right)\big|_{\Sigma} \cdot \vec{n} = 0 \qquad \operatorname{div} v^\pm = 0$$
$$\operatorname{rot} v^\pm = 0$$

Off the curve $v^\pm$ satisfies 2D Euler:

$$\begin{cases} v_t + (v \cdot \nabla) v + \nabla p = 0 \\ \operatorname{div} v = 0 \end{cases}$$

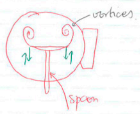
And $(p_+ - p_-)\big|_{\Sigma} = 0$

---

③ NO. DATE

~~Helmholtz ideas.~~

Kaffelöffel Experiment.

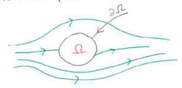


- Generation of vorticity from initially potential flow. Against Helmholtz theorem. Due to nonsmooth flows.
- Important point in the discussion of d'Alembert's paradox
- Helmholtz suggests instability of vortex sheets. Including spirals.
- Kelvin arrives at the Kelvin-Helmholtz instability 1894

KOKUYO

Kutta-Żukowski criterion. (4)

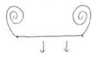
Lift around the body immersed in a potential flow.



Emphasizes the role of circulation:

$$\oint_{\partial\Omega} v\cdot\tau = \int_{\Omega} \omega \qquad [\omega = \operatorname{rot} v]$$

Prandtl introduces the model of spirals generating the circulation/vorticity.



$$Z_m(t,\theta) = t^{\mu} e^{a(\theta-\theta_m)} e^{i\theta}$$

$$\Gamma_m(t,\theta) = g_m t^{2\mu-1} e^{2a(\theta-\theta_m)}$$

$0=\theta_0<\theta_1<\dots<\theta_m<2\pi$.
$a>0$, $g_m\in\mathbb{R}\setminus\{0\}$,
$\mu\in\mathbb{R}$.

Birkhoff-Rott equation. (5)

NO.
DATE

Young Birkhoff (son of the ergodic thm. guy).

complex conjugate

$$\partial_t Z_m(t,\theta)^{*} = \frac{1}{2\pi i}\int \sum_{k=0}^{M-1} \frac{d\Gamma_k}{Z_m(t,\Gamma_m)-Z_k(t,\Gamma_k)}.$$

In self-similar variables related Prandtl spirals B-R takes the form:

$$\mu + \frac{(1-2\mu)(a+i)}{2a} = \left(\frac{1}{2\pi i}\int_{\mathbb{R}} \sum_{k=0}^{M-1} \frac{2ag_k e^{2a\sigma}\, d\sigma}{1-e^{(a+i)\sigma+i(\theta_k-\theta_m)}}\right)^{*}$$

The lack of integrability of the kernel when $\sigma\to+\infty$.

Remarkable observation by Elling/Gnann (2019):

$$\frac{e^{2a\sigma}}{1-e^{(a+i)\sigma+i\Delta}} = -e^{(a-i)\sigma-i\Delta} - e^{-2i\sigma-2i\Delta} + \frac{e^{-2i\sigma-2i\Delta}}{1-e^{(a+i)\sigma+i\Delta}};$$

$$\sum_{k=0}^{M-1} g_k e^{-i\theta_k} = \sum_{k=0}^{M-1} g_k e^{-2i\theta_k} = 0.$$

KOKUYO

$M\geq 3$ uniformly distributed Prandtl spirals solve B-R equation. (6)

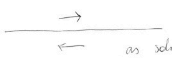
[! $M\geq 1$ uniformly distributed Prandtl spirals solve 2D Euler in a weak sense; Cieślak-Kokocki-Ożański 2021].

BACK TO THE QUESTION OF HELMHOLTZ!

The union of $M\geq 3$ uniformly distributed Prandtl spirals is linearly unstable As solutions to B-R !!! [Kelvin-Helmholtz instability]. The growth of instabilities is algebraic in time!

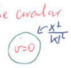
Comment. CKO '23.

Majda/Bertozzi show that the flat sheet



is linearly unstable as solution to B-R (with exponential in time growth of instabilities).

Murray/Wilcox 2022 show that the circular sheet flow is linearly stable !!!



7

HOW DO WE COMPUTE IT?

THE IDEA OF KELVIN-HELMHOLTZ INSTABILITIES. TAKE $Z_m(t,\Gamma_m)+\varepsilon\zeta_m(t,\Gamma_m)$, EXPAND IN ε. NEGLECT HIGHER ORDER TERMS, SHOW INSTABILITY OF THE LINEARIZATION.

FIRST STEP:

$$\mathrm{p.v.}\int_{\mathbb{R}}\sum_{k=0}^{M-1}\frac{d\Gamma_k}{Z_m(t,\Gamma_m)-Z_k(t,\Gamma_k)+\varepsilon\left(\zeta_m(t,\Gamma_m)-\zeta_k(t,\Gamma_k)\right)}=$$

$$=-\varepsilon\,\mathrm{f.p.}\int_{\mathbb{R}}\sum_{k=0}^{M-1}\frac{\zeta_m(t,\Gamma_m)-\zeta_k(t,\Gamma_k)}{[Z_m(t,\Gamma_m)-Z_k(t,\Gamma_k)]^2}\,d\Gamma_k+O(\varepsilon^2).$$

f.p. is the Hadamard finite part

THE LINEARIZED B-R:

$$t\,\partial_t\zeta_m(t,\theta)^*-\frac{2\mu-1}{2a}\,\partial_\theta\zeta_m(t,\theta)^*=$$

$$=-\mathcal{H}_m\zeta\text{, where}$$

8

$$\mathcal{H}_m\zeta:=\frac{ag}{\pi i}\,\mathrm{f.p.}\int_{\mathbb{R}}\sum_{k=0}^{M-1}\frac{\left(\zeta_m(\theta)-\zeta_k(\sigma+\theta+\theta_k-\theta_m)\right)e^{2a\sigma-2i\theta}}{\left[1-e^{(a+i)\sigma}e^{i(\theta_k-\theta_m)}\right]^2}\,d\sigma.$$

ANOTHER ALGEBRAIC IDENTITY:

Cieślak-Kokocki-Ożański (2023)

$$\frac{e^{2a\sigma}}{\left(1-e^{(a+i)\sigma}e^{i\Delta}\right)^2}=e^{-2i\sigma-2i\Delta}+e^{-a\sigma-3i\sigma-3i\Delta}$$

$$-\frac{e^{-a\sigma-3i\sigma-3i\Delta}}{1-e^{(a+i)\sigma+i\Delta}}-\frac{e^{-2i\sigma-2i\Delta}}{1-e^{(a+i)\sigma+i\Delta}}$$

$$+\frac{e^{-2i\sigma-2i\Delta}}{\left(1-e^{(a+i)\sigma+i\Delta}\right)^2}\text{, hence we obtain}$$

integrability of $\mathcal{H}_m$ for $\sigma\to+\infty$! ($M\geqslant 3$).

9

THE EXACT FORM OF UNSTABLE MODES:

$$\zeta_m(t,\theta):=X_m(t)\zeta_m^+(t,\theta)e^{i\theta}$$

$$+Y_m(t)\zeta_m^-(t,\theta)e^{i\theta}\text{, where}$$

$$\zeta_m^\pm(t,\theta):=e^{\pm i\alpha\ln\frac{\Gamma_m(t,\theta)}{g_m}}=e^{\pm i\alpha(2a(\theta-\theta_m)+(2\mu-1)\ln t)}$$

Then, taking $X_i(t)=X_j(t)$, $Y_i(t)=Y_j(t)$,

one obtains: $\partial_t\left(t^{i\frac{1-2\mu}{2a}}X(t)\right)=-t^{-1+i\frac{1-2\mu}{2a}}\cdot$

$$\cdot\left(c_0^*-c^-(\alpha)^*\right)Y^*(t);$$

$$\partial_t\left(t^{i\frac{1-2\mu}{2a}}Y(t)\right)=-t^{-1+i\frac{1-2\mu}{2a}}\left(c_0^*-c^+(\alpha)^*\right)^*X^*(t).$$

For some choices of c_0 & $c^\pm(\alpha)$

$|X(t)|,|Y(t)|\geqslant Ct^\delta$.

10

$$C_0 := \frac{ga(a-i)}{(a+i)^2} \coth(\pi A/M), \qquad A := -\frac{2ai}{a+i},$$

$$C^+(\alpha) := \frac{ga^2(2i\alpha+1)}{(a+i)^2} \coth(\pi B_+(\alpha)/M),$$

$$C^-(\alpha) := \frac{ga^2(-2i\alpha+1)}{(a+i)^2} \coth(\pi B_-(\alpha)/M),$$

$$B_\pm(\alpha) := -a(\pm 2i\alpha+1)\frac{1+ai}{1+a^2}.$$

THANK YOU!

DZIĘKUJĘ!

GRACIAS!

Pseudospheres from singularity theory view-point with a classification of 2-soliton surfaces

Toshizumi Fukui

Department of Mathematics, Saitama University, Japan

(joint work with Yutaro Kabata)

We discuss pseudospheres in the Euclidean 3-space with taking care about their singularity types and Backlünd transformations. We investigate a classification of 2-soliton surfaces by noting how the ridge lines appear.

Pseudospheres from singularity theory view point with a classification of 2-soliton surfaces (j/w with Yutaro Kabata)

Toshi Fukui (Saitama University)

16:45–17:15, 26 September, 2023

Workshop for Mathmatics for Industry
25–29 September, 2023
Warsaw University of Technology

Surfaces in $\mathbb{R}^3$

$\varphi : \mathbb{R}^2 \longrightarrow M = \varphi(\mathbb{R}^2) \subset \mathbb{R}^3$, C^∞

$$E = \langle \varphi_u, \varphi_u \rangle, \ \ F = \langle \varphi_u, \varphi_v \rangle, \ \ G = \langle \varphi_v, \varphi_v \rangle$$
$$L = \langle \varphi_{uu}, \boldsymbol{\nu} \rangle, \ \ M = \langle \varphi_{uv}, \boldsymbol{\nu} \rangle, \ \ N = \langle \varphi_{vv}, \boldsymbol{\nu} \rangle$$

where ν is a unit normal.
The first fundamental form

$$I = E\, du^2 + 2F\, du\, dv + G\, dv^2$$

The second fundamental form

$$II = L\, du^2 + 2M\, du\, dv + N\, dv^2$$

Chebyshev' net

A pseudosphere is a surface with constant negative Gauss curvatures. We can assume that they have Gauss curvature -1 up to similarity transformations.
For a surface with $K = -1$, we can take the asymptotic coordinate (u, v) with the following fundamental forms:

$$I = du^2 + 2\cos\phi\, du\, dv + dv^2$$
$$II = 2\sin\phi\, du\, dv$$

where ϕ is the asymptotic angle.
Gauss-Coddazi equation becomes sine-Gordon equation:

$$\phi_{uv} = \sin\phi$$

Curvature coordinate

The curvature coordinate is given by

$$x = \frac{u+v}{2}, \quad y = \frac{u-v}{2}.$$

The fundamental forms are

$$I = \cos^2 \tfrac{\phi}{2}\, dx^2 + \sin^2 \frac{\phi}{2}\, dy^2$$
$$II = \frac{1}{2} \sin\phi\, (dx^2 - dy^2)$$

The principal curvatures are

$$\tan\frac{\phi}{2}, \text{ and } -\cot\frac{\phi}{2}$$

4 / 16

Ridge and flecnodal

Let v_i denote a principal vector of a surface and let κ_i denote the corresponding principal curvature of a surface.

A point P on a surface is v_i-ridge if $v_i\kappa_i(P) = 0$.

A point P on a surface is flecnodal if there is a line with at least 4 point contact with the surface at P.

5 / 16

1. The level sets of ϕ, κ_1 and κ_2 containing P are equal.
2. The differentials of the principal curvatures are given as follows:

$$\partial_x\kappa_1 = \frac{\phi_x}{1+\cos\phi}, \qquad \partial_y\kappa_1 = \frac{\phi_y}{1+\cos\phi},$$
$$\partial_x\kappa_2 = \frac{\phi_x}{-1+\cos\phi}, \qquad \partial_y\kappa_2 = \frac{\phi_y}{-1+\cos\phi}.$$

So ∂_x-ridge (resp. ∂_y-ridge) is given by $\phi_x = 0$ (resp. $\phi_y = 0$). (A level of ϕ has a horizontal (or vertical) tangent.)
Flecnodal point on pseudosphere is given by $\phi_u\phi_v = 0$. (i.e., A level of ϕ has a diagonal (or anti-diagonal) tangent.)

6 / 16

Backlünd transformation

We say $\tilde{\phi}$ is Backlünd transformation of ϕ if

$$\left(\frac{\tilde{\phi}+\phi}{2}\right)_u = \lambda \sin\frac{\tilde{\phi}-\phi}{2}, \quad \left(\frac{\tilde{\phi}-\phi}{2}\right)_v = \lambda^{-1}\sin\frac{\tilde{\phi}+\phi}{2}. \quad (1)$$

where $\lambda = \tan\theta/2$. θ is in the next sheet.
If ϕ is a solution of sine-Gordon equation, so is $\tilde{\phi}$.

$$\{\text{sol. of sine-Gordon}\} \xrightarrow{BT} \{\text{sol. of sine-Gordon}\}$$

Geometric BT

We say

$$M \longrightarrow \widetilde{M}, \quad p \longmapsto \tilde{p},$$

is Gemetric BT, if

- The line $\overline{p\tilde{p}}$ is in T_pM and also in $T_{\tilde{p}}\widetilde{M}$.
- $d(p,\tilde{p})$ is constant $(= r)$.
- the unit normals $\boldsymbol{\nu}_p$ and $\tilde{\boldsymbol{\nu}}_{\tilde{p}}$ has a constant angle θ, that is $\langle \boldsymbol{\nu}_p, \tilde{\boldsymbol{\nu}}_{\tilde{p}}\rangle = \cos\theta$.

Geometric BT between $K=-1$ surfaces is given by

$$\tilde{\varphi} = \varphi + r\left(\frac{\cos\tilde{\phi}/2}{\cos\phi/2}\varphi_x + \frac{\sin\tilde{\phi}/2}{\sin\phi/2}\varphi_y\right), \quad r = \sin\theta$$

and it preserves Chebyshev's nets.

Bianchi's permutability

If ϕ_i $(i=1,2)$ satisfies

$$\left(\frac{\phi_i+\phi}{2}\right)_u = \lambda_i \sin\frac{\phi_i-\phi}{2}, \quad \left(\frac{\phi_i-\phi}{2}\right)_v = \lambda_i^{-1}\sin\frac{\phi_i+\phi}{2},$$

and $\tilde{\phi}$ satisfies

$$(\lambda_2-\lambda_1)\tan\frac{\tilde{\phi}-\phi}{4} = (\lambda_2+\lambda_1)\tan\frac{\phi_2-\phi_1}{4},$$

then

$$\left(\frac{\tilde{\phi}+\phi_1}{2}\right)_u = \lambda_2 \sin\frac{\tilde{\phi}-\phi_1}{2}, \quad \left(\frac{\tilde{\phi}-\phi_1}{2}\right)_v = \lambda_2^{-1}\sin\frac{\tilde{\phi}+\phi_1}{2}.$$

$$\left(\frac{\tilde{\phi}+\phi_2}{2}\right)_u = \lambda_1 \sin\frac{\tilde{\phi}-\phi_2}{2}, \quad \left(\frac{\tilde{\phi}-\phi_2}{2}\right)_v = \lambda_1^{-1}\sin\frac{\tilde{\phi}+\phi_2}{2}.$$

Soliton

0-soliton $\xrightarrow{BT}$ 1-soliton $\xrightarrow{BT}$ 2-soliton

$\phi = 0$

$\phi_\lambda = 4\tan^{-1}(\lambda u + \lambda^{-1} v)$

$\xi_i = \lambda_i u + \lambda_i^{-1} v$

$\phi_{\lambda_1,\lambda_2} = 4\tan^{-1}\left(\frac{\lambda_1+\lambda_2}{\lambda_2-\lambda_1} \cdot \frac{\sinh\frac{\xi_1-\xi_2}{2}}{\cosh\frac{\xi_1+\xi_2}{2}}\right)$

line

Beltrami's pseudosphere

Dini's pseudosphere

2-soliton surfaces

10 / 16

Singular locus of φ

Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ be a Chevyshev net for a pseudosphere with $K = -1$. Let ϕ denote the asymptotic angle. Then

$$\begin{aligned} I &= du^2 + 2\cos\phi\, du\, dv + dv^2 \\ II &= 2\sin\phi\, du\, dv \end{aligned}$$

Remark that the singular locus of φ is defined by

$$\Sigma : \sin\phi = 0, \quad \text{i.e.,} \quad \phi = k\pi, \quad k \in \mathbb{Z}.$$

For 2-soliton surface, we have $k = 0, \pm 1$.

11 / 16

Criteria of singularities

Let C denote the curvature line through P whose principal direction is null direction at P.

1. Assume that ϕ is nonsingular at P, i.e., the singular locus of φ is nonsingular at P.
   1.1 φ is cuspidal edge at P if and only if Σ and C intersect transversely at P.
   1.2 φ is swallowtail at P if and only if Σ has 2-point contact with C at P.
2. Assume that ϕ has a Morse singularity at P.
   2.1 φ is cuspidal beaks at P if and only if the Hessian of ϕ is positive.
   2.2 φ is cuspidal lips at P if and only if the Hessian of ϕ is negative.

12 / 16

Flecnodal and ridge on a 2-soliton surface

On pseudospheres, we have
∂_u-flecnodal line ($\phi_u = 0$),
∂_v-flecnodal line ($\phi_v = 0$),
∂_x-ridge line ($\phi_x = 0$),
∂_y-ridge line ($\phi_y = 0$)
and, on 2-soliton surfaces, they are

$$\frac{\cosh \xi_2}{\cosh \xi_1} = \frac{\lambda_2}{\lambda_1},\ \frac{\lambda_1}{\lambda_2},\ \frac{\lambda_2 + \lambda_2^{-1}}{\lambda_1 + \lambda_1^{-1}},\ \frac{\lambda_2 - \lambda_2^{-1}}{\lambda_1 - \lambda_1^{-1}},\ \text{respectively.}$$

Here $\xi_i = \lambda_i u + v/\lambda_i$, $i = 1, 2$.

When $\lambda_2 \to \lambda_1 = \lambda$,

$$\phi_{\lambda,\lambda} = \lim_{\lambda' \to \lambda} \phi_{\lambda,\lambda'} = 4 \tan^{-1} \frac{-\eta}{\cosh \xi},$$

where $\xi = \lambda u + \lambda^{-1} v + c$ and $\eta = \lambda u - \lambda^{-1} v$.
The ∂_u-flecnodal, ∂_v-flecnodal, ∂_x-ridge and ∂_y-ridge are defined by

$$\eta \tanh \xi = 1,\ -1,\ \frac{\lambda - \lambda^{-1}}{\lambda + \lambda^{-1}},\ \frac{\lambda + \lambda^{-1}}{\lambda - \lambda^{-1}},\ \text{respectively.}$$

Classification of 2-soliton

The result in this section should compare the classification of 2-soliton surfaces (Popov). They show four types for generic 2-soliton surfaces. The correspondence between their classification and our results is summarized as follows:

Type	$\lambda_1\lambda_2$	μ	flecnodal	∂_x-ridge	∂_y-ridge
1	$+$	$+$	exist	exist	exist
2	$+$	$-$	exist	exist	not exist
3	$-$	$+$	not exists	not exist	not exist
4	$-$	$-$	not exists	not exist	exist

$\mu = (\lambda_1^2 - 1)(\lambda_2^2 - 1)$.

Breather surfaces

For $\lambda \in \mathbb{C}$, with $\mathrm{Re}\,\lambda \neq 0$, $\mathrm{Im}\,\lambda \neq 0$, we have

$$\phi_{\lambda,\bar{\lambda}} = -4\tan^{-1}\left(\cot\arg\lambda \cdot \frac{\sin \mathrm{Im}\,\xi}{\cosh \mathrm{Re}\,\xi}\right)$$

where $\xi = \lambda u + v/\lambda$.
The ∂_u-flecnodal, ∂_v-flecnodal, ∂_x-ridge and ∂_y-ridge are defined by

$$\frac{(\tanh \mathrm{Re}\,\xi)(\tan \mathrm{Im}\,\xi)}{\tan\arg\lambda} = 1, \quad -1, \quad \frac{|\lambda| - |\lambda|^{-1}}{|\lambda| + |\lambda|^{-1}}, \quad \frac{|\lambda| + |\lambda|^{-1}}{|\lambda| - |\lambda|^{-1}},$$

respectively.

16 / 16

Formation of nanostructured functional particles with the spray-drying method

Leon Gradoń

Faculty of Chemical and Process Engineering,
Warsaw University of Technology, Poland

The structure of matter, on both an atomic and macroscopic scale, is a result of the interplay between the requirements of the physical forces operating between the individual parts and the mathematical requirements of space-filling. Nanoparticles with well-defined chemical composition can act as a building block for the construction of functional structures, such as highly ordered aggregates, as well as porous and hollow aggregates. A spray drying technique is used for the production of crystal-like structures with nanoparticle building blocks. When spray-drying uniform spherical particles tightly packed aggregates with either simple or broken symmetry were formed using polystyrene particles with varied zeta potential as templates, it is also possible to form highly ordered porous and hollow aggregates from inorganic colloidal particles potentially useful for controlled drug delivery and catalysis. The process by which organized mesoporous silica particles are formed by the spray-drying method was examined using elementary laws of topology.

WARSAW UNIVERSITY OF TECHNOLOGY
Faculty of Chemical and Process Engineering

Formation of nanostructured functional particles with the spray-drying method

Leon Gradon
Faculty of Chemical and Process Engineering,
Warsaw University of Technology Warsaw, Poland

Stanisław Janeczko
Faculty of Mathematics and Information Sciences
Warsaw University of Technology Warsaw, Poland

Ratna Balgis, Takahashi Ogi, Kikuo Okuyama
Department of Chemical Engineering, Hiroshima University,
Higashi-Hiroshima, Japan

Contents

1. Introduction
2. Principles of self-assembly
3. Shapes of the structures
4. Examples of nanostructures applications
5. Principle of spray-drying process
6. Examples of produced templates
7. Topographical structures for challenging aspects of nanocatalysis
8. Conclusions

"There is plenty of room at the bottom"

Richard P. Feynman

(there is a room for great development even in the microscopic world)

The structure of matter, on both an atomic and macroscopic scale, is a result of the interplay between the requirements of the physical forces operating between the individual parts and the mathematical requirements of space filling.

Self-assembly

Reasons for interest in self-assembly:

1) Humans are attracted by the appearance of order from disorder.
2) Living cells self-assemble $\Rightarrow$ stimulation for the design of non-living systems.
3) Self-assembly is one of the few practical strategies for making ensembles of nanostructures.

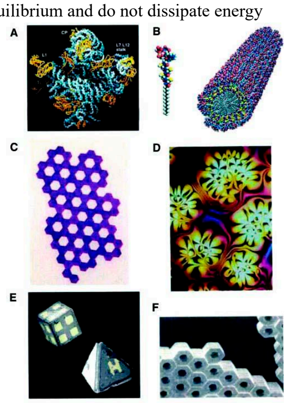
$\boxed{\text{It will therefore be an essential part of nanotechnology.}}$

4) Manufacturing and robotics will benefit from applications of self-assembly.
5) Self-assembly is common to many dynamic and multicomponent systems:
   - smart materials
   - self-healing structures
   - netted sensors
   - computer networks

Types of self-assembly

<u>**Static self-assembly (S)**</u>

S – involves systems that are at global or local equilibrium and do not dissipate energy



Examples of static self-assembly

(A) Crystal structure of a ribosome

(B) Self-assembled peptideamphiphile nanofibers.

(C) An array of millimetersized polymeric plates assembled at a water/perfluorodecalin interface by capillary interactions.

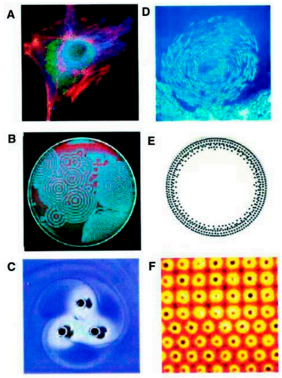
(D) Thin film of a nematic liquid crystal on an isotropic substrate.

(E) Micrometersized metallic polyhedra folded from planar substrates.

(F) A three-dimensional aggregate of micrometer plates assembled by capillary forces.

Dynamic self-assembly (D)

D – interaction responsible for the formation of structures or pattern between components only occur if the system is dissipating energy



Examples of dynamic self-assembly

(A) An optical micrograph of a cell with fluorescently labeled cytoskeleton and nucleus; microtubules (~24 nm in diameter) are colored red.
(B) Reaction-diffusion waves in a Belousov-Zabatinski reaction in a 3.5-inch Petri dish.
(C) A simple aggregate of three millimeter-sized, rotating, magnetized disks interacting with one another via vortex-vortex interactions.
(D) A school of fish.
(E) Concentric rings formed by charged metallic beads 1 mm in diameter rolling in circular paths on a dielectric support.
(F) Convection cells formed above a micropatterned metallic support.
The distance between the centers of the cells is ~2 mm.

Self-assembly reflects information coded as: shape, surface properties, charge, polaribility, etc.

The design of components that organize themselves into desired patterns and function is the key to application of self-assembly.

The components must be able to move with respect to one another. Their steady-state positions balance attraction and repulsion.

Self-assembly requires that the components are mobile. It takes place in fluid phases or on the smooth surfaces.

Molecular self-assembly involves: non-covalent or weak covalent interactions, i.e. van der Waals, electrostatic, hydrophobic, hydrogen and coordinative bonds.

Self-assembly of meso- or macroscopic objects: interactions are selected and tailored include gravity, external electromagnetic fields, capillary, entropic interactions.

Using shape for self-assembly

Major milestones towards the goal of self-assembly:

1) Making the building blocks.

2) Understanding and controlling the interactions.

3) Predicting the consequence of many components interacting in a prescribed environment.

4) Identify components and interactions that will organize to form a desired product (reverse self-assembly).

5) Knowing how to use self-assembly.

What is “shape”?

- The idea of the shape is used for the purpose of understanding its effect on self-assembly.
- It defines the shape of an object as the ensemble of the geometries of all interactions elicited by that object.
- By this definition an object could have multiple shapes, depending on the particular interaction of interest.
- Challenge of self-assembly is thus to understand how these different shapes of the same objects contribute to its assembly.

Templates

A brute force approach to create nearly arbitral shapes uses templates.

Template is a sacrificial mold in which material is grown or deposited, e.g. micelles, membrane, colloid crystals, zeolites, and block copolymers.

Instabilities

This approach aims to create a highly symmetric yet metastable structure (spherical colloid coated with a metal).

Under the stimulus the structure “relaxes” toward one of its ground states by breaking its own symmetry, e.g. stimulus heat, shell devotes leading to the formation of a lower symmetry, stimulus-stretch metastable conformation fold into functional shape (proteins).

The reciprocity in material technology

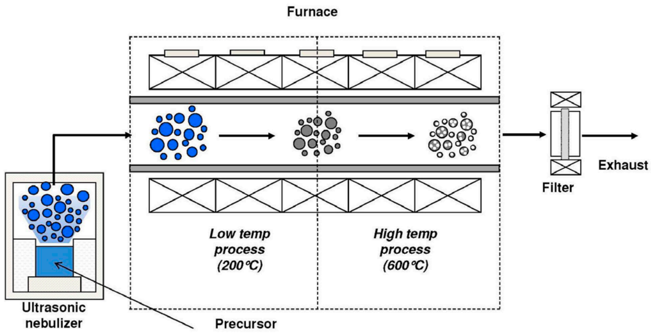
means ← goal

processing → structure → properties → performance

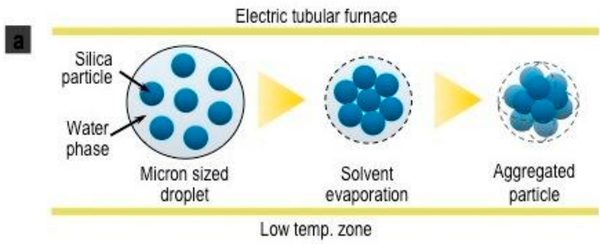
→

cause and effect

Principle of spray-drying process



Principles of spheres arrangements in the spray-drying process



Close-packing of spheres in Euclidean space:

$$\{S\} = \{(S_1, p_1 ... (S_N, p_N)\}$$

Two spheres (S_i, p_i) and (S_j, p_j) of radius r are in contact, i.e.:

$$\text{dist}\,(p_i, p_j) = 2\text{r}$$

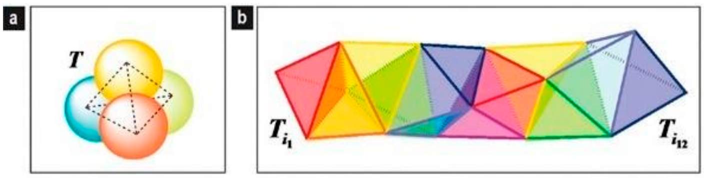
Cluster of spheres is weakly tetrahedral, T, if for each sphere (S_{i1}, p_{i1}) there exist three spheres (S_{i2}, p_{i2}), (S_{i3}, p_{i3}) and (S_{i4}, p_{i4}).

Such the distance $\quad \text{dist}\,(S_{ik}, p_{il}) = 2r \quad \text{if } 1 \le k, l \le 4$

Tetrahedral nano-cluster
(cluster which consists of tetrahedrals)

For every two tetrahedra T_{i_l}, T_{i_k} there exist an ordered chain: $\{T_{i_l}\}_{l=1...k}$

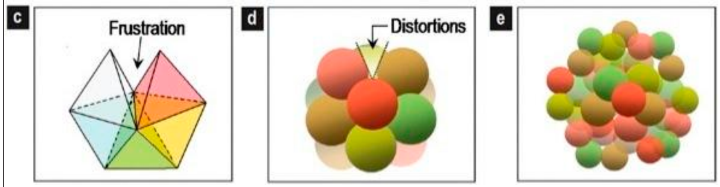
That $T_{i_n}, T_{i_{n+1}}$ have common face, $n = 1 \ldots k - 1$

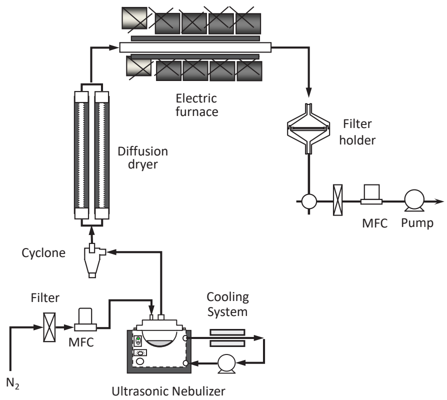


- Tetrahedron is a basic unit of the tight packing by equal spheres.
- Distortion associated with tetrahedral packing.
- 13 spheres icosahedron have small distortion.
- 12 spheres arranged symmetrically around one sphere are not packed in perfectly way.
- Distance a between spheres: $a > 2r$
- Elementary property of icosahedron gives a relation:

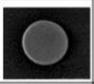
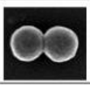
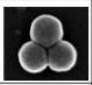
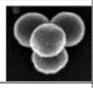
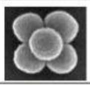
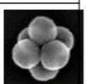
$$a = 8r / \left(10 + 2\sqrt{5}\right)^{1/2}$$

Compact packing of spheres

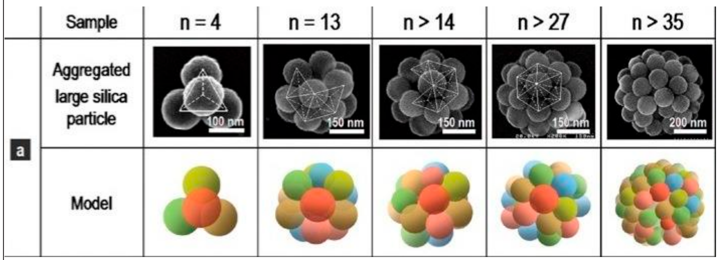


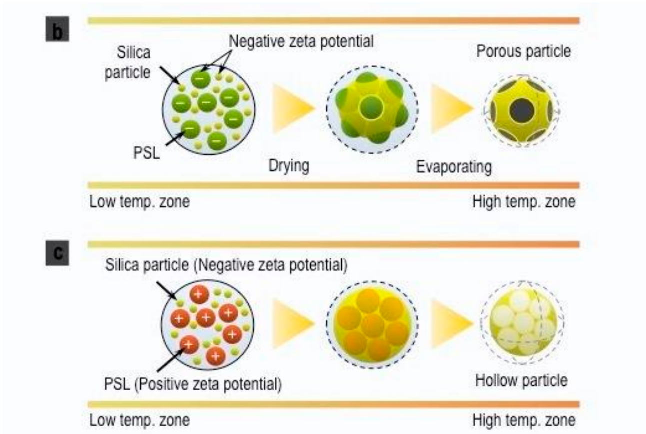


Organization of spheres in the droplet

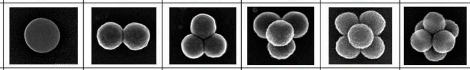
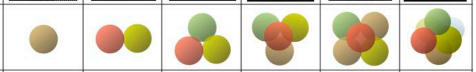
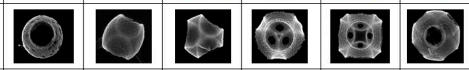
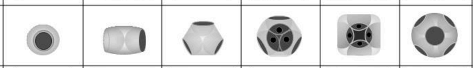
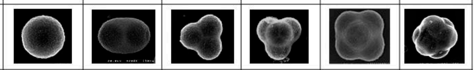
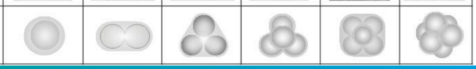
Sample	a	b	c	d	e	f
n	one	two	three	four	five	six
Silica particle						
Model						

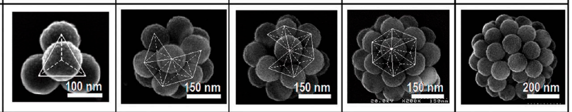
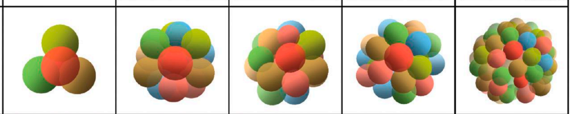
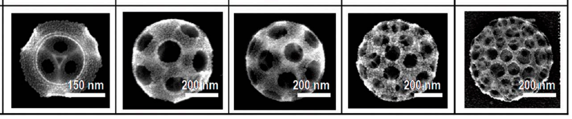
Organization of spheres in the droplet





Examples of produced templates

Sample	a	b	c	d	e	f
n	one	two	three	four	five	six
Silica particle						
Model						
Porous particle						
Model						
Hollow particle						
Model						

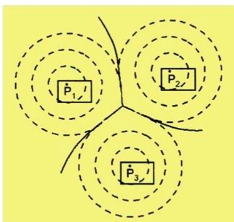
	Sample	n = 4	n = 13	n > 14	n > 27	n > 35
a	Aggregated large silica particle					
	Model					
b	Porous particle					

Cell structure on the plane

$P = \{P_1, \ldots P_N\}$ – cell centers on R^2

$d_k(x) = d(x, P_k)$ – Euclidean distance function from P_k

$f_k : R^2 \rightarrow R$ – cell structure function



Global competition squared distance function:

$$\hat{d}(x) = \min\left\{f_1(x)d_1^{\,2}(x), \ldots f_N(x)d_N^{\,2}(x)\right\}$$

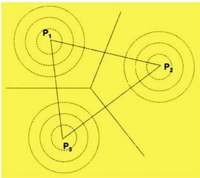
Conflict sets arrangements and the intercells patterns

System of competition organizing centers

$$P = \{P_1, P_2 \ldots . P_n\} = \{(P_1, f_1), (P_2, f_2) \ldots . (P_n, f_n)\}$$

Cells corresponding to the system

$$CP_i = \{x \in R^2 : \forall_{k \in (1 \ldots n)}\ f_i(x)\, d_i^2(x) \le f_k(x)\, d_k^2(x)\}$$

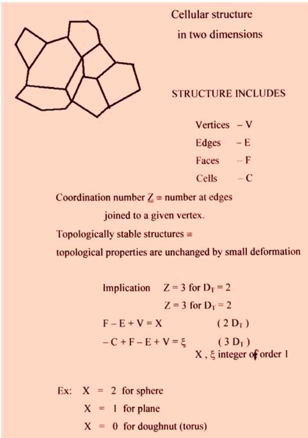


Interface curve (generalized Voroni diagram)

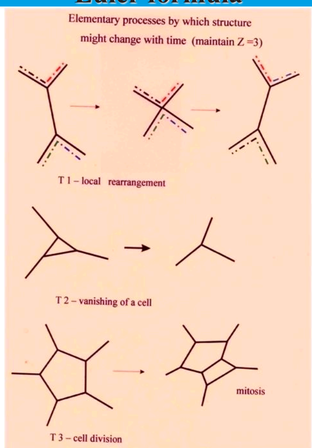
$$VP = \{ x \in R^2 : \exists_{ifj}\ S \text{ that } \forall_k f_i d^2(x, p_i) = f_j\, d^2(x, p_j) = f_j\, d^2(x, p_j) \le f_k\, d^2(x, p_k)\}$$

assumption: $f_i \cong 1$, $i = 1, \ldots N$

Euler formula



Euler formula



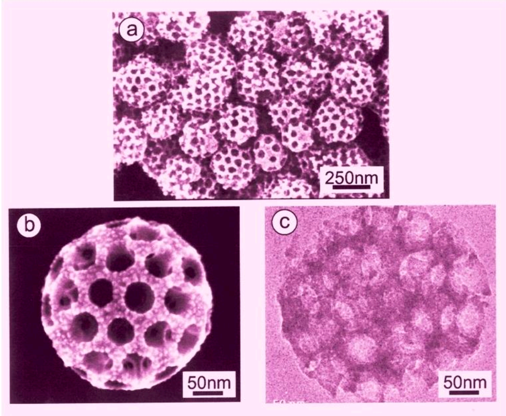
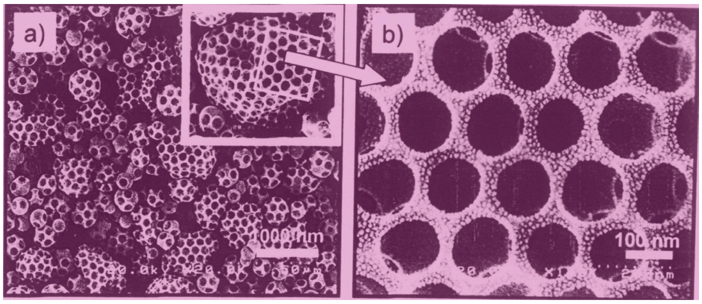
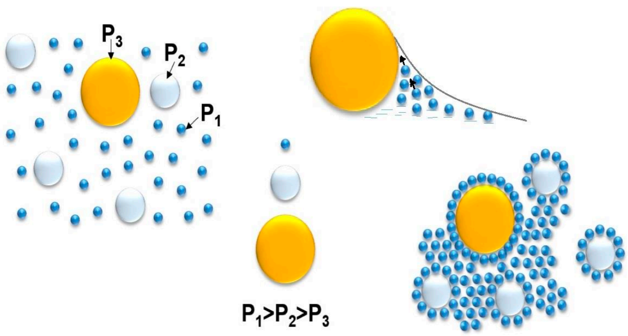
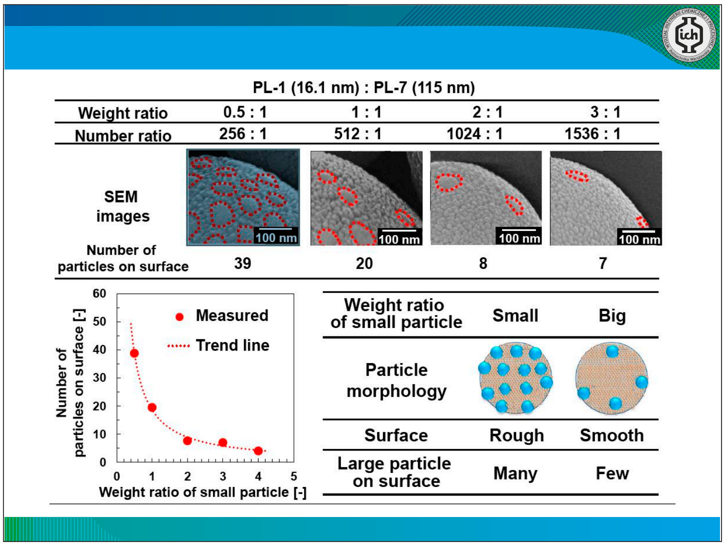
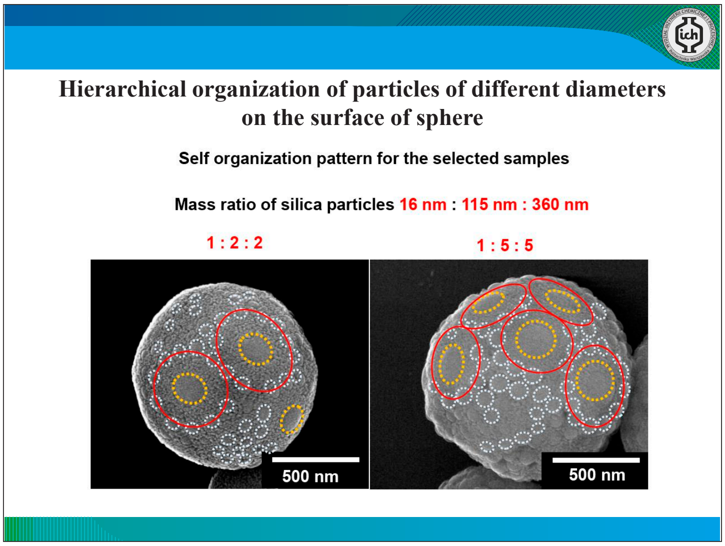


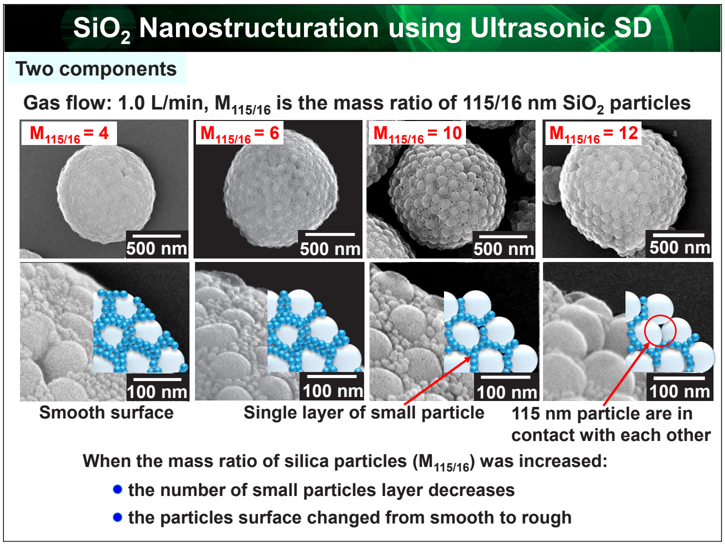
Fig. 8






PL-1 (16.1 nm) : PL-7 (115 nm)
Weight ratio 0.5 : 1 1 : 1 2 : 1 3 : 1
Number ratio 256 : 1 512 : 1 1024 : 1 1536 : 1
SEM images
100 nm
Number of particles on surface 39 20 8 7
Measured
Trend line
Number of particles on surface [-]
Weight ratio of small particle [-]
Weight ratio of small particle Small Big
Particle morphology
Surface Rough Smooth
Large particle on surface Many Few

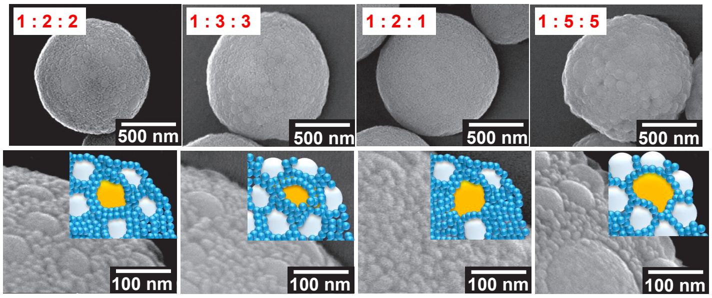

Hierarchical organization of particles of different diameters on the surface of sphere
Self organization pattern for the selected samples
Mass ratio of silica particles 16 nm : 115 nm : 360 nm
1 : 2 : 2
1 : 5 : 5
500 nm
500 nm


SiO_2 Nanostructuration using Ultrasonic SD
Two components
Gas flow: 1.0 L/min, $M_{115/16}$ is the mass ratio of 115/16 nm SiO_2 particles
$M_{115/16} = 4$
$M_{115/16} = 6$
$M_{115/16} = 10$
$M_{115/16} = 12$
500 nm
100 nm
Smooth surface
Single layer of small particle
115 nm particle are in contact with each other
When the mass ratio of silica particles ($M_{115/16}$) was increased:
the number of small particles layer decreases
the particles surface changed from smooth to rough

SiO_2 Nanostructuration using Ultrasonic SD

Three components

Mass ratio of SiO_2 particles 16 nm : 115 nm : 360 nm



- **360 nm particles were surrounded by 115 nm particles**
- **When the mass ratio of large particles was increased, surface morphology changed from smooth to concave-convex**

Conclusions

- The aerosol assisted spray-drying process is an useful method for production of developed and desired space-forms made of nanoparticles.
- Mesoporous nanostructured particles were produced using PSL particles as a template material for organizing nanoparticles around them.
- The composition of the cells on the surface of sphere is described using the concept of conflict set arrangement.
- Stationary state of cell configuration on the sphere has equal infinitesimal of cell boundaries in real vortex and they are equal $2\pi/3$.
- The signs of zeta potential of the template particles and colloid particles used in the spray drying process define the structure of the final product, which could be either hollow or porous.

Thank you for your attention
☺

On comparing distributions with imprecise data

Przemysław Grzegorzewski

Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland

One of the most fundamental problems in mathematical statistics is the comparison of two or more distributions that characterize the underlying populations. Classical tests applied there are constructed with pretty specific assumptions concerning the distributions, like normality, exponentiality, etc. However, in reality, these assumptions are often not met. The problem becomes much more difficult when the output of an experiment consists of data that are imprecise, or vague. There we need a model that allows us to grasp both aspects of uncertainty that appear in such data: randomness, associated with the data generation mechanism, and fuzziness, connected with data imprecision. To cope with this problem Puri and Ralescu (1986) introduced a fuzzy random variable.

On the other hand, in analyzing fuzzy data from the statistical perspective we immediately come upon some key obstacles, like the nonlinearity associated with the fuzzy number arithmetic, the lack of a universally accepted total ranking, the lack of suitable probability distribution models, or no limit theorems for random mechanisms producing fuzzy data which could be directly applied in statistical inference. Therefore, statistical tests with imprecise data usually cannot be generalized straightforwardly from their classical prototypes.

We show that some of the aforementioned difficulties in test construction can be overcome by using permutation-based nonparametric procedures. Combining these with a distance-based approach or a dominance credibility index gives us some interesting goodness-of-fit and location tests, respectively.

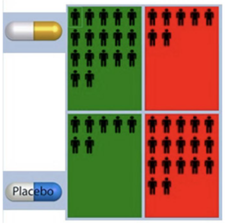
On comparing distributions with imprecise data

Przemysław Grzegorzewski

Faculty of Mathematics and Information Science,
Warsaw University of Technology

Workshop on Mathematics for Industry
Warsaw, September 27, 2023

Motivations



$$\begin{cases} H_0: \text{no treatment effect} \\ H_1: \text{new treatment effect} \end{cases}$$

$X_1, \ldots, X_n$ i.i.d. $\mathrm{N}(\mu_1, \sigma_1)$ and $Y_1, \ldots, Y_m$ i.i.d. $\mathrm{N}(\mu_2, \sigma_2)$

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases} \quad \text{or} \quad \begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \ > \mu_2 \end{cases}$$

Here we can use the well-known parametric tests.

$X_1, \ldots, X_n$ i.i.d. $F = ?$ and $Y_1, \ldots, Y_m$ i.i.d. $G = ?$

$$\begin{cases} H_0: F = G \\ H_1: F \neq G \end{cases} \quad \text{or} \quad \begin{cases} H_0: F = G, \\ H_1: X \overset{st}{>} Y. \end{cases}$$

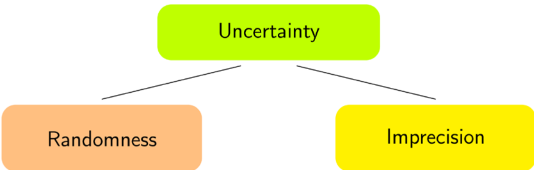
Here we can use some nonparametric tests.

Example

The Gamonedo cheese is a kind of a blue cheese produced in Asturias. In quality control experts (tasters) express their perceptions about

- visual parameters (shape, rind, appearance),
- texture parameters (hardness and crumbliness),
- olfactory-gustatory parameters (smell intensity, smell quality, flavour intensity, flavour quality and aftertaste),
- an overall impression of the cheese.

(Ramos-Guajardo A.B., et al., 2019)



Outline:

- How do we model imprecise data?
  - fuzzy numbers
  - fuzzy random variables
  - pro and cons
- How do we compare distributions with imprecise data?
  - distance-based goodness-of-fit permutation tests
  - tests based on the credibility degree of dominance

Fuzzy numbers

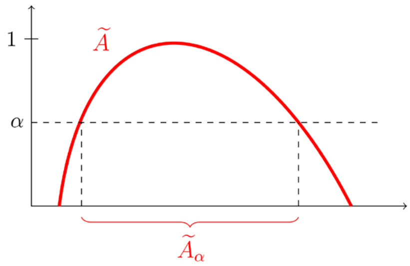
A **fuzzy number** is identified by a mapping $\widetilde{A} : \mathbb{R} \to [0,1]$, called a membership function, such that its α-cuts

$$\widetilde{A}_\alpha = \begin{cases} \{x \in \mathbb{R} : \widetilde{A}(x) \geqslant \alpha\} & \text{if } \alpha \in (0,1], \\ cl\{x \in \mathbb{R} : \widetilde{A}(x) > 0\} & \text{if } \alpha = 0, \end{cases}$$

are nonempty compact intervals for each $\alpha \in [0,1]$, where cl denotes the closure operator.

A fuzzy number is completely characterized by its membership function $\widetilde{A}(x)$ or by a family of its α-cuts $\{\widetilde{A}_\alpha\}_{\alpha\in[0,1]}$.

Let $\mathbb{F}(\mathbb{R})$ denote the family of all fuzzy numbers.

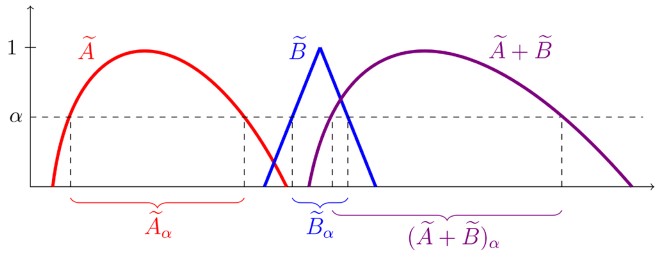


Each α-cut of a fuzzy number is a closed interval $\widetilde{A}_\alpha$.

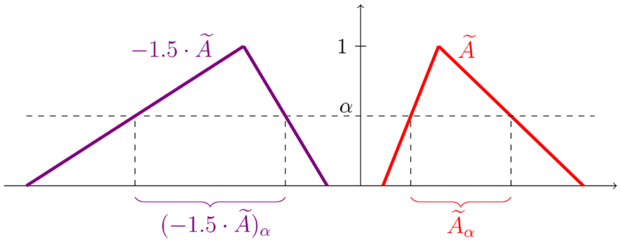
$\widetilde{A}_0 = \operatorname{supp}(\widetilde{A})$ is called the **support** and $\widetilde{A}_1 = \operatorname{core}(\widetilde{A})$ is known as the **core** of fuzzy number $\widetilde{A}$, respectively.

Arithmetic in $\mathbb{F}(\mathbb{R})$

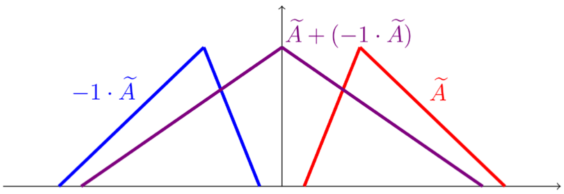
$$\left(\widetilde{A} + \widetilde{B}\right)_\alpha = \left[\inf \widetilde{A}_\alpha + \inf \widetilde{B}_\alpha, \sup \widetilde{A}_\alpha + \sup \widetilde{B}_\alpha\right], \quad \forall \alpha \in [0,1]$$



$$(\theta \cdot \widetilde{A})_\alpha = \begin{cases} [\theta \inf \widetilde{A}_\alpha,\, \theta \sup \widetilde{A}_\alpha] & \text{if } \theta > 0 \\ [\theta \sup \widetilde{A}_\alpha,\, \theta \inf \widetilde{A}_\alpha] & \text{if } \theta < 0 \end{cases}, \qquad \forall \alpha \in [0,1]$$



Note $(\mathbb{F}(\mathbb{R}), +, \cdot)$ has not linear but semilinear structure since $\widetilde{A} + (-1 \cdot \widetilde{A}) \neq \mathbb{1}_{\{0\}}$.



Moreover, the Minkowski difference does not satisfy, in general, the addition/subtraction property that $(\widetilde{A} + (-1)\widetilde{B}) + \widetilde{B} = \widetilde{A}$.

Let λ denote a normalized measure associated with a continuous distribution with support in $[0,1]$ and let $\gamma > 0$.

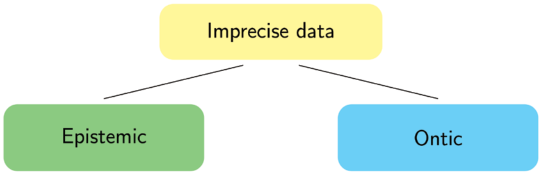
Then for any $\widetilde{A}, \widetilde{B} \in \mathbb{F}(\mathbb{R})$ we define a metric D_γ^λ as follows

$$D_\gamma^\lambda(\widetilde{A}, \widetilde{B}) = \sqrt{\int_0^1 \left[(\operatorname{mid} \widetilde{A}_\alpha - \operatorname{mid} \widetilde{B}_\alpha)^2 + \gamma(\operatorname{spr} \widetilde{A}_\alpha - \operatorname{spr} \widetilde{B}_\alpha)^2\right] d\lambda(\alpha)},$$

where $\operatorname{mid} \widetilde{A}_\alpha = \frac{1}{2}(\inf \widetilde{A}_\alpha + \sup \widetilde{A}_\alpha)$, $\operatorname{spr} \widetilde{A}_\alpha = \frac{1}{2}(\sup \widetilde{A}_\alpha - \inf \widetilde{A}_\alpha)$.

(Gil et al., 2002; Trutschnig et al., 2009)

Whatever (λ, γ) is chosen D_γ^λ is invariant to translations and rotations. Moreover, $(\mathbb{F}(\mathbb{R}), D_\gamma^\lambda)$ is a separable metric space and for each fixed λ all metrics D_γ^λ are topologically equivalent.



Fuzzy random variables

Fuzzy random variables (random fuzzy numbers) integrate **randomness** (associated with data generation) and **fuzziness** (associated with data nature).

Definition (Puri M.L., Ralescu D., 1986)

Let $(\Omega, \mathcal{A}, P)$ be a probability space. A mapping $\widetilde{X} : \Omega \to \mathbb{F}(\mathbb{R})$ is a fuzzy random variable (**random fuzzy number**) if for all $\alpha \in [0, 1]$ the α-level function is a compact random interval.

In other words, $\widetilde{X}$ is a fuzzy random variable if and only if $\widetilde{X}$ is a Borel measurable function w.r.t. the Borel σ-field generated by the topology induced by D_γ^λ.

The **Aumann-type mean** of a fuzzy random variable $\widetilde{X}$ is the fuzzy number $E(\widetilde{X}) \in \mathbb{F}(\mathbb{R})$ such that for each $\alpha \in [0, 1]$ the α-cut $\big(E(\widetilde{X})\big)_\alpha$ is equal to the Aumann integral of $\widetilde{X}_\alpha$, i.e.

$$\big(E(\widetilde{X})\big)_\alpha = \big[\mathbb{E}(\mathrm{mid}\,\widetilde{X}_\alpha) - \mathbb{E}(\mathrm{spr}\,\widetilde{X}_\alpha), \mathbb{E}(\mathrm{mid}\,\widetilde{X}_\alpha) + \mathbb{E}(\mathrm{spr}\,\widetilde{X}_\alpha)\big].$$

The D_γ^λ-**Fréchet-type variance** $V(\widetilde{X})$ is a non-negative real number such that

$$\begin{aligned} V(\widetilde{X}) &= \mathbb{E}\Big(\big[D_\gamma^\lambda(\widetilde{X}, E(\widetilde{X}))\big]^2\Big) \\ &= \int_0^1 \mathrm{Var}(\mathrm{mid}\,\widetilde{X}_\alpha) d\lambda(\alpha) + \gamma \int_0^1 \mathrm{Var}(\mathrm{spr}\,\widetilde{X}_\alpha) d\lambda(\alpha). \end{aligned}$$

Given a fuzzy sample $\widetilde{\mathbb{X}} = (\widetilde{X}_1, \ldots, \widetilde{X}_n)$ we can determine its various characteristics, like the average $\overline{\widetilde{X}} \in \mathbb{F}(\mathbb{R})$ defined by its α-cuts

$$\overline{\widetilde{X}}_\alpha = \Big[\frac{1}{n}\sum_{i=1}^{n} \operatorname{mid}(\widetilde{X}_i)_\alpha - \frac{1}{n}\sum_{i=1}^{n} \operatorname{spr}(\widetilde{X}_i)_\alpha,$$
$$\frac{1}{n}\sum_{i=1}^{n} \operatorname{mid}(\widetilde{X}_i)_\alpha + \frac{1}{n}\sum_{i=1}^{n} \operatorname{spr}(\widetilde{X}_i)_\alpha\Big],$$

or the sample variance $S^2 \in \mathbb{R}$ given by

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} D_\gamma^\lambda(\widetilde{X}_i, \overline{\widetilde{X}})^2.$$

Note

In contrast to the statistical analysis of numerical data one should be aware of the following problems typical for fuzzy data:

- problems with subtraction and division of fuzzy numbers;
- the lack of universally accepted total ranking between fuzzy numbers;
- there are not yet realistic suitable models for the distribution of random fuzzy numbers;
- there are not yet Central Limit Theorems for random fuzzy numbers that can be directly applied for making inference.

Conclusion

No straightforward generalizations of the classical statistical tests for fuzzy data exist.

Permutation ANOVA for r.f.n.

Suppose, we observe independently $p \geqslant 2$ **fuzzy random samples** drawn from populations with unknown distributions, i.e.

$$\widetilde{\mathbb{X}}_1 = (\widetilde{X}_{11}, \ldots, \widetilde{X}_{1n_1})$$
$$\vdots$$
$$\widetilde{\mathbb{X}}_p = (\widetilde{X}_{p1}, \ldots, \widetilde{X}_{1n_p}).$$

We want to verify the null hypothesis that all p samples come from the same distribution, i.e.

$$H_0 : \widetilde{\mathbb{X}}_1 \stackrel{d}{=} \ldots \stackrel{d}{=} \widetilde{\mathbb{X}}_p$$

against the alternative hypothesis $H_1 : \neg H_0$ that at least two population distributions differ.

Let $\widetilde{\mathbb{V}} = \widetilde{\mathbb{X}}_1 \uplus \ldots \uplus \widetilde{\mathbb{X}}_p$, where $\uplus$ stands for vector concatenation, so that the p samples are pooled into one, i.e. $\widetilde{V}_i = \widetilde{X}_{1i}$ if $1 \leqslant i \leqslant n_1$, $\widetilde{V}_i = \widetilde{X}_{2i}$ if $n_1 + 1 \leqslant i \leqslant n_1 + n_2$ and so on until $\widetilde{V}_i = \widetilde{X}_{p,i-(n_1+\ldots n_{p-1})}$ if $n_1 + \ldots n_{p-1} + 1 \leqslant i \leqslant N$.

Now, let $\widetilde{\mathbb{V}}^*$ denote a permutation of the initial dataset $\widetilde{\mathbb{V}}$.
Then

$$\begin{aligned}
\widetilde{\mathbb{X}}_1^* &= (\widetilde{X}_{11}^*, \ldots, \widetilde{X}_{1n_1}^*) \longleftarrow (\widetilde{V}_1^*, \ldots, \widetilde{V}_{n_1}^*) \\
\widetilde{\mathbb{X}}_2^* &= (\widetilde{X}_{21}^*, \ldots, \widetilde{X}_{2n_2}^*) \longleftarrow (\widetilde{V}_{n_1+1}^*, \ldots, \widetilde{V}_{n_1+n_2}^*) \\
&\vdots \\
\widetilde{\mathbb{X}}_p^* &= (\widetilde{X}_{p1}^*, \ldots, \widetilde{X}_{pn_p}^*) \longleftarrow (\widetilde{V}_{N-n_p+1}^*, \ldots, \widetilde{V}_N^*).
\end{aligned}$$

If H_0 holds we expect that all p sample means would not differ to much from the overall sample mean.

Thus to decide whether the distance between the observed sample means is large enough to conclude as significant we consider the following test statistic

$$T(\widetilde{\mathbb{V}}^*) = \sum_{i=1}^{p} n_i \cdot D_\theta^\lambda(\overline{\widetilde{X}_i^*}, \overline{\overline{\widetilde{X}}})^2,$$

where

$$\overline{\widetilde{X}_i^*} = \frac{1}{n_i} \sum_{j=n_1+\ldots+n_{i-1}+1}^{n_1+\ldots+n_i} \widetilde{V}_j^*.$$

Obviously, $\overline{\overline{\widetilde{X}^*}} = \frac{1}{p}\sum_{i=1}^{p} \overline{\widetilde{X}_i^*} = \frac{1}{N}\sum_{i=1}^{N} \widetilde{V}_i^* = \overline{\overline{\widetilde{X}}}$ for any $\widetilde{\mathbb{V}}^*$.

(Grzegorzewski P., 2020)

For a given realization of a fuzzy sample $\widetilde{\mathbb{v}} = \widetilde{\mathbb{x}}_1 \uplus \ldots \uplus \widetilde{\mathbb{x}}_p$ we compute the observed test statistic

$$t_0 = T(\widetilde{\mathbb{v}}) = \sum_{i=1}^{k} n_i \cdot D_\theta^\lambda\big(\overline{\widetilde{x}_i^*}, \overline{\overline{\widetilde{x}}}\big)^2.$$

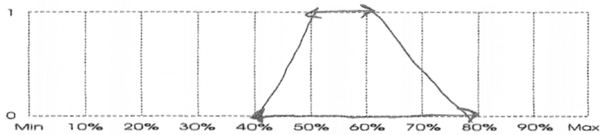
The p-value of our test is defined as the proportion of cases when the test statistic values are greater or equal to the observed experimental value $t_0 = T(\widetilde{\mathbb{v}})$.

We repeat the whole procedure, i.e. we draw a permutation and compute a value of the test statistic $T(\widetilde{\mathbb{v}}^*)$ B times (usually about 1000). Then the approximate p-value of our test is given by

$$\text{p-value} \simeq \frac{1}{B} \sum_{B=1}^{B} \mathbb{1}(T(\widetilde{\mathbb{v}}_b^*) \geqslant t_0).$$

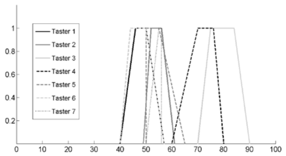
Example (cont.)

So far, the experts provide an ordinal number ranging from 1 to 7 to describe their perceptions about different cheese characteristics. Recently, the tasters were proposed to express their subjective perceptions about the quality of the Gamonedo cheese by using fuzzy numbers.

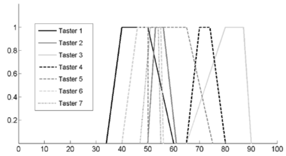


Opinion of a taster expressed by means of a trapezoidal fuzzy set

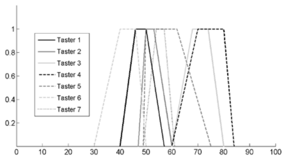
(Ramos-Guajardo A.B., et al., 2019)



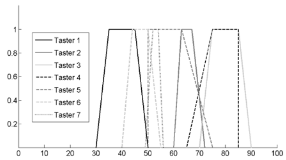
Appearance



Smell quality



Flavour quality



Overall impression

We consider some data given in Ramos-Guajardo A.B. et al.(2019) to compare the opinions of the three experts about the overall impression of the Gamonedo cheese. We have three independent fuzzy samples of sizes $n_1 = 40$, $n_2 = 38$ and $n_3 = 42$, coming from the unknown distributions.

Opinion	Expert 1	Expert 2	Expert 3
1	$(65, 75, 85, 85)$	$(50, 50, 63, 75)$	$(60, 63, 67, 72)$
2	$(35, 37, 44, 50)$	$(39, 47, 52, 60)$	$(53, 58, 63, 68)$
3	$(66, 70, 75, 80)$	$(60, 70, 85, 90)$	$(43, 47, 54, 58)$
4	$(70, 74, 80, 84)$	$(50, 56, 64, 74)$	$(70, 76, 83, 86)$
5	$(65, 70, 75, 80)$	$(39, 45, 53, 57)$	$(54, 60, 65, 70)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Our problem is to check whether there is a general agreement between these experts.

To reach the goal we verify the following null hypothesis

$$H_0 : \widetilde{\mathbb{X}}_1 \stackrel{d}{=} \widetilde{\mathbb{X}}_2 \stackrel{d}{=} \widetilde{\mathbb{X}}_3,$$

stating there is no significant difference between experts' opinions, against $H_1 : \neg H_0$ that their opinions on the cheese quality differ.

Substituting data into formula for T we obtain $t_0 = 2259.436$.

Then, after generating $M = 10\,000$ random permutations we have obtained the p-value of 0.0011. Hence, we may conclude that there is no general agreement between experts' opinion on the overal impression of the Gamonedo cheese.

Other tests based on distances

- Energy distance test (Grzegorzewski P., Gadomska O., 2022)

$$T_{en}(\widetilde{\mathbb{X}}, \widetilde{\mathbb{Y}}) = = \frac{nm}{n+m}\left[\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} D_\gamma^\lambda(\widetilde{X}_i, \widetilde{Y}_j) - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} D_\gamma^\lambda(\widetilde{X}_i, \widetilde{X}_j) - \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} D_\gamma^\lambda(\widetilde{Y}_i, \widetilde{Y}_j)\right].$$

- Nearest neighbor test (Grzegorzewski P., Gadomska O., 2022)

$$T_{knn}(\widetilde{\mathbb{X}}, \widetilde{\mathbb{Y}}) = \frac{1}{kN}\sum_{i=1}^{N}\sum_{j=1}^{k} I_j(\widetilde{V}_i),$$

where $\widetilde{V} = \widetilde{X} \uplus \widetilde{Y}$ and

$$I_k(\widetilde{V}_i) = \begin{cases} 1, & \text{if } \widetilde{V}_i \text{ and } \mathrm{NN}_k(\widetilde{V}_i) \text{ belong to the same sample,} \\ 0, & \text{if } \widetilde{V}_i \text{ and } \mathrm{NN}_k(\widetilde{V}_i) \text{ belong to different samples,} \end{cases}$$

The generalized Mann-Whitney test for fuzzy data

Let $\mathbb{X} = (X_1, \ldots, X_n)$ and $\mathbb{Y} = (Y_1, \ldots, Y_m)$ denote independent samples from two populations F and G, respectively.

We consider the following testing problem

$$\begin{cases} H_0 : F = G, \\ H_1 : X \stackrel{st}{>} Y. \end{cases}$$

The Mann-Whitney test statistic is given by

$$U(\mathbb{X}, \mathbb{Y}) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathbb{1}(X_i > Y_j).$$

Our goal: to generalize the Mann-Whitney test for fuzzy data.

Consider the possibility and necessity measures (Dubous & Prade, 1983) for ranking fuzzy numbers $\widetilde{A}$ and $\widetilde{B}$:

$$\begin{aligned}\mathrm{Pos}(\widetilde{A} \succ \widetilde{B}) &= \sup_{x>y} \min\{\widetilde{A}(x), \widetilde{B}(y)\},\\ \mathrm{Nes}(\widetilde{A} \succ \widetilde{B}) &= 1 - \mathrm{Pos}(\widetilde{A} \preceq \widetilde{B})\\ &= 1 - \sup_{x\leqslant y} \min\{\widetilde{A}(x), \widetilde{B}(y)\}.\end{aligned}$$

Obviously, $\mathrm{Nes}(\widetilde{A} \succ \widetilde{B}) > 0$ implies that $\mathrm{Pos}(\widetilde{A} \succ \widetilde{B}) = 1$.

Following Liu (2004) we aggregate both measures by the following index

$$\mathrm{Cr}(\widetilde{A} \succ \widetilde{B}) = \frac{\mathrm{Pos}(\widetilde{A} \succ \widetilde{B}) + \mathrm{Nes}(\widetilde{A} \succ \widetilde{B})}{2},$$

to obtain the credibility degree that $\widetilde{A}$ is larger than $\widetilde{B}$.

Lemma 1

For any trapezoidal fuzzy numbers $\widetilde{A} = \mathit{Tra}(a_1, a_2, a_3, a_4)$ and $\widetilde{B} = \mathit{Tra}(b_1, b_2, b_3, b_4)$ the credibility degree that $\widetilde{A}$ is larger than $\widetilde{B}$ is given by the following formula

$$\mathit{Cr}(\widetilde{A} \succ \widetilde{B}) = \begin{cases} 0, & a_4 \leqslant b_1 \text{ and } a_3 < b_2,\\ \frac{a_4 - h(a_4,b_1)}{2(a_4-a_3)}, & a_4 > b_1 \text{ and } a_3 < b_2,\\ \frac{1}{2}, & a_3 \geqslant b_2, a_4 \geqslant b_1 \text{ or } a_2 \leqslant b_3, a_1 \leqslant b_4,\\ 1 - \frac{h(a_1,b_4)-a_1}{2(a_2-a_1)}, & a_1 < b_4 \text{ and } a_2 > b_3,\\ 1, & b_4 \leqslant a_1 \text{ and } a_2 > b_3,\end{cases}$$

where

$$\begin{aligned} h(a_4, b_1) &= \frac{a_4 b_2 - b_1 a_3}{b_2 - b_1 + a_4 - a_3},\\ h(a_1, b_4) &= \frac{b_4 a_2 - a_1 b_3}{b_4 - b_3 + a_2 - a_1}.\end{aligned}$$

Lemma 2

For any triangular fuzzy numbers $\widetilde{A} = (l_A, c_A, r_A)$ and $\widetilde{B} = (l_B, c_B, r_B)$ the credibility degree that $\widetilde{A}$ is larger than $\widetilde{B}$ is given by the following formula

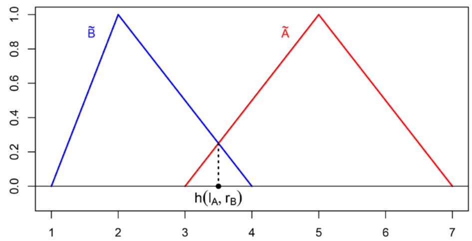
$$\mathit{Cr}(\widetilde{A} \succ \widetilde{B}) = \begin{cases} 0, & r_A \leqslant l_B \text{ and } c_A \neq c_B,\\ \frac{h(r_A,l_B)-r_A}{2(c_A-r_A)}, & c_A < c_B \text{ and } r_A > l_B,\\ \frac{1}{2}, & c_A = c_B,\\ 1 - \frac{h(l_A,r_B)-l_A}{2(c_A-l_A)}, & c_A > c_B \text{ and } l_A < r_B,\\ 1, & r_B \leqslant l_A \text{ and } c_A \neq c_B,\end{cases}$$

where

$$\begin{aligned} h(r_A, l_B) &= \frac{r_A c_B - l_B c_A}{c_B - l_B - (c_A - r_A)},\\ h(l_A, r_B) &= \frac{l_A c_B - r_B c_A}{c_B - r_B - (c_A - l_A)}.\end{aligned}$$

Example

Consider triangular fuzzy numbers $\widetilde{A} = (3, 5, 7)$ and $\widetilde{B} = (1, 2, 4)$.



$$\mathrm{Cr}(\widetilde{A} \succ \widetilde{B}) = 1 - \frac{h(l_A, r_B) - l_A}{2(c_A - l_A)} = 1 - \frac{3.5 - 3}{2 \cdot 2} = \frac{7}{8}.$$

Let $\widetilde{\mathbb{X}} = (\widetilde{X}_1, \ldots, \widetilde{X}_n)$ and $\widetilde{\mathbb{Y}} = (\widetilde{Y}_1, \ldots, \widetilde{Y}_m)$ denote independent samples, each consisting of i.i.d. random fuzzy numbers.

We want to verify

$$\begin{cases} H_0 : \widetilde{X} \overset{d}{=} \widetilde{Y}, \\ H_1 : \widetilde{X} \succ \widetilde{Y}. \end{cases}$$

Using the credibility index for each pair of observations from both samples we obtain the following test statistic

$$U_{CR}(\widetilde{\mathbb{X}}, \widetilde{\mathbb{Y}}) = \sum_{i=1}^{n} \sum_{j=1}^{m} Cr(\widetilde{X}_i \succ \widetilde{Y}_j).$$

To decide whether to reject or not the null hypothesis H_0 we design a permutation test.

(Grzegorzewski P. and Zacharczuk M., 2023)

Algorithm 1: The generalized Mann-Whitney test for fuzzy data

Data: Fuzzy samples $\widetilde{\mathbb{x}} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ and $\widetilde{\mathbb{y}} = (\widetilde{y}_1, \ldots, \widetilde{y}_m)$

begin

$u_0 \longleftarrow \sum_{i=1}^{n} \sum_{j=1}^{m} Cr(\widetilde{x}_i \succ \widetilde{y}_j)$;

Pool the data: $\widetilde{\mathbb{w}} = \widetilde{\mathbb{x}} \uplus \widetilde{\mathbb{y}}$;

for $b = 1$ to B **do**

Take a permutation $\widetilde{\mathbb{w}}^* = (\widetilde{w}_1^*, \ldots, \widetilde{w}_{n+m}^*)$ of $\widetilde{\mathbb{w}}$;

$\widetilde{\mathbb{x}}^* = (\widetilde{x}_1^*, \ldots, \widetilde{x}_n^*) \longleftarrow (\widetilde{w}_1^*, \ldots, \widetilde{w}_n^*)$;

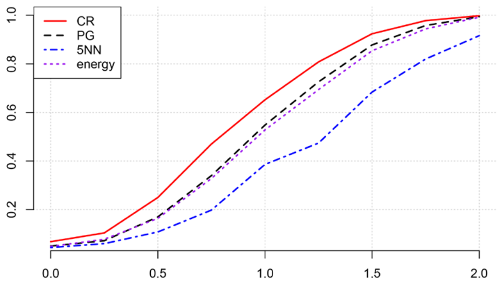
$\widetilde{\mathbb{y}}^* = (\widetilde{y}_1,^* \ldots, \widetilde{y}_m^*) \longleftarrow (\widetilde{w}_{n+1}^*, \ldots, \widetilde{w}_{n+m}^*)$;

$U_{CR} \longleftarrow \sum_{i=1}^{n} \sum_{j=1}^{m} Cr(\widetilde{x}_i^* \succ \widetilde{y}_j^*)$;

end

p-value $\longleftarrow \frac{1}{B} \sum_{b=1}^{B} \mathbb{1}\big(U_{CR}(\widetilde{\mathbb{x}}_b^*, \widetilde{\mathbb{y}}_b^*) \geqslant u_0\big)$.

end



Power comparison for the increasing difference in location.

The p-sample ($p \geqslant 2$) location problem

More generally, we observe $p \geqslant 2$ independent samples

$$\begin{aligned} \mathbb{X}_1 = (X_{11}, \ldots, X_{1n_1}) &\quad \sim \quad F_1 \\ &\vdots \\ \mathbb{X}_p = (X_{p1}, \ldots, X_{pn_p}) &\quad \sim \quad F_p. \end{aligned}$$

We want to verify the hypotheses

$$\begin{cases} H_0 : F_1 = \ldots = F_p \\ H_1 : F \leqslant F_2 \leqslant \ldots \leqslant F_p, \end{cases}$$

where at least one inequality is strict.

The generalized Jonkheere-Terpstra test for fuzzy data

More generally, we observe $p \geqslant 2$ independent fuzzy samples:
$\widetilde{\mathbb{X}}_1 = (\widetilde{X}_{11}, \ldots, \widetilde{X}_{1n_1}), \ldots, \widetilde{\mathbb{X}}_p = (\widetilde{X}_{p1}, \ldots, \widetilde{X}_{pn_p})$.
We want to verify

$$\begin{cases} H_0 : \widetilde{X}_1 \stackrel{d}{=} \widetilde{X}_2 \stackrel{d}{=} \ldots \stackrel{d}{=} \widetilde{X}_p, \\ H_1 : \widetilde{X}_1 \succ \widetilde{X}_2 \succ \ldots \succ \widetilde{X}_p. \end{cases}$$

The generalized Jonkheere-Terpstra test statistic:

$$\begin{aligned} J_{CR} &= \sum\sum_{1 \le i < j \le p} U_{CR}(\widetilde{\mathbb{X}}_i, \widetilde{\mathbb{X}}_j) \\ &= \sum\sum_{1 \le i < j \le p} \sum_{r=1}^{n_i} \sum_{s=1}^{n_j} Cr(\widetilde{X}_{ir} \succ \widetilde{X}_{js}). \end{aligned}$$

(Grzegorzewski P. and Zacharczuk M., 2023)

Conclusions and further research

- Due to certain difficulties with fuzzy modeling statistical tests with imprecise data usually cannot be generalized straightforwardly from their classical prototypes.
- Some of those difficulties in test constructions might be solved by applying nonparametric tests based of permutations.
- Permutation tests require extremely limited assumptions, i.e. *exchangeability* (we can exchange the labels of the observations under H_0 without affecting the results).
- The credibility index might appear useful for some test constructions, especially for situations connected with the dominance relation.

and this is the end

Thank you for your attention :)

Brief Introduction to Topology for Multi-objective Optimization

Naoki Hamada
Machine Learning Group, KLab Inc.

A broad range of scientific and engineering tasks, including data analysis, product design, modeling, planning, and management, can be formulated in multi-objective optimization problems. Recent developments in convex analysis and data science using topology have brought a new paradigm for solving and analyzing multi-objective optimization problems. In this talk, several applications of topology to multi-objective optimization will be presented. We will show how the topology of convex analysis can be applied to a sparse modeling task, generalizing the regularization path of the elastic net and efficiently tuning its two hyper-parameters simultaneously. To extend this idea beyond the convexity assumption, we introduce a statistical test using persistent homology and the Poincaré conjecture whether the hyper-parameter tuning method works.

WMI 2023

Brief Introduction to Topology for Multiobjective Optimization

Naoki Hamada (KLab Inc.)

From the viewpoint of an industrial researcher...

I will talk about

- **Relation between mobile game industry and mathematics**
  - How cutting-edge techs are applied to game products
  - Why mobile games are matter for mathematics
- **Topology, multiobjective optimizaiton and sparse modeling**
  - A new math theory developed in collaboration with academia
  - An application to mobile games

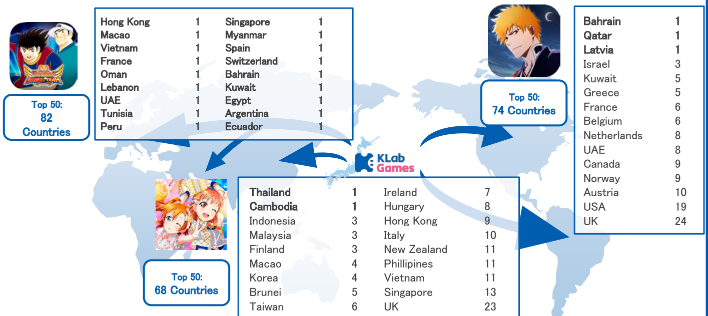
Mobile Game Industry and Mathematics

Company Overview

KLab

Company Name	KLab Inc.
Founded	August 1, 2000
Capital	5,457.9 million yen (February 2023)
Stock Exchange	Tokyo Stock Exchange Prime Market (3656)
CEO	Hidekatsu Morita (President and CEO) Yosuke Igarashi (Vice Chairman)
Offices in Japan	Headquarters (Tokyo, Roppongi Hills Mori Tower) Osaka Office, Fukuoka Office, Sendai Office
Affiliated Companies	Global Gear Inc. BLOCKSMITH&Co.
International Office	KLab China Inc.
KLab Group Employees	541 (Full-time employees as of December 2022)

App Store Top Sales Rankings



Top 50: 82 Countries

Hong Kong	1	Singapore	1
Macao	1	Myanmar	1
Vietnam	1	Spain	1
France	1	Switzerland	1
Oman	1	Bahrain	1
Lebanon	1	Kuwait	1
UAE	1	Egypt	1
Tunisia	1	Argentina	1
Peru	1	Ecuador	1

Top 50: 74 Countries

Bahrain	1
Qatar	1
Latvia	1
Israel	3
Kuwait	5
Greece	5
France	6
Belgium	6
Netherlands	8
UAE	8
Canada	9
Norway	9
Austria	10
USA	19
UK	24

Top 50: 68 Countries

Thailand	1	Ireland	7
Cambodia	1	Hungary	8
Indonesia	3	Hong Kong	9
Malaysia	3	Italy	10
Finland	3	New Zealand	11
Macao	4	Phillipines	11
Korea	4	Vietnam	11
Brunei	5	Singapore	13
Taiwan	6	UK	23

*Data compiled by KLab using research from App Annie statistics available from each game's release date to September, 2021.

Mobile games are a good test bed for mathematical science

Ideal conditions for science

- **Big data**
  - Millions of active users ➡ TB-scale data / day
  - Sensors on phones: touch, mic, camera, GPS, gyro, …
- **Open problems**
  - Applications: Game AI, Game Dev/Ops/Ads, …
  - Challenges: Real-time ML on edge devices, rapid MLOps, …
- **Rapid research cycle**
  - Res team deploys every week and gets immediate feedback
  - Not mission critical: we can do trials-and-errors in products

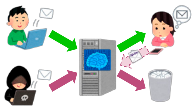
Mission of Machine Learning Group in KLab

Research cutting-edge ML technologies to create unprecedented gaming experiences

Math ⟵ Data Science — HPC ⟶ Game

Math	Data Science	HPC	Game
Singularity Theory Optimization	Machine Learning Data Analysis	Supercomputer Cloud	Game Dev/Ops Game AI
with IMI, Kyushu U.	with RIKEN AIP	with RIIT, Kyushu U.	in KLab

Applications of ML in KLab



Spam filter for in-game chat
(Google Cloud Day: Digital '22)

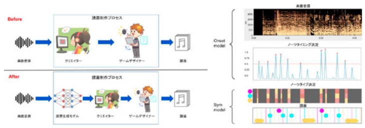
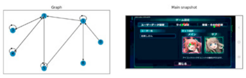


Chart generation for rhythm games
(CEDEC2021, SIGGRAPH Asia 2021, AAAI-23)



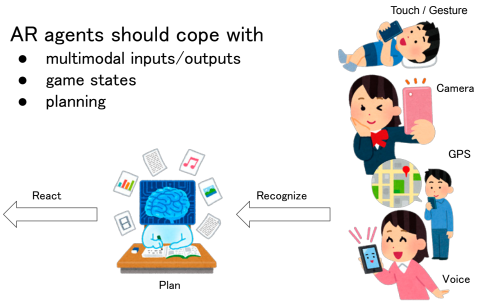
Autoplay for UI testing
(JSAI2022 Award)

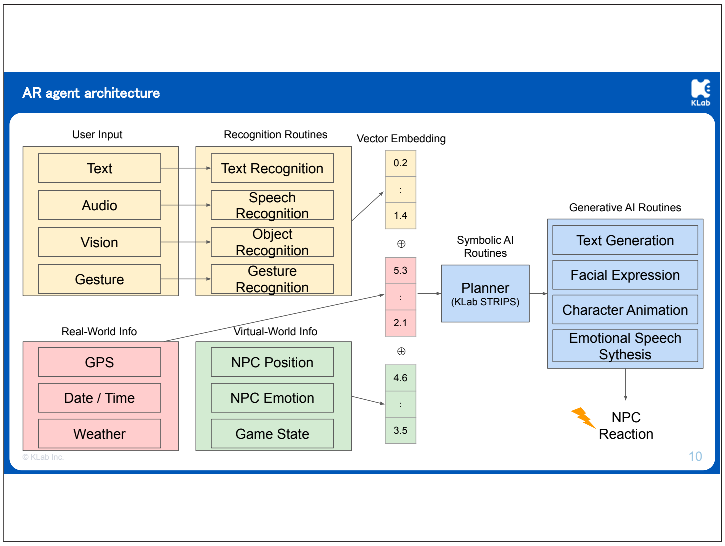
3D motion retrieval
(IBIS2022)

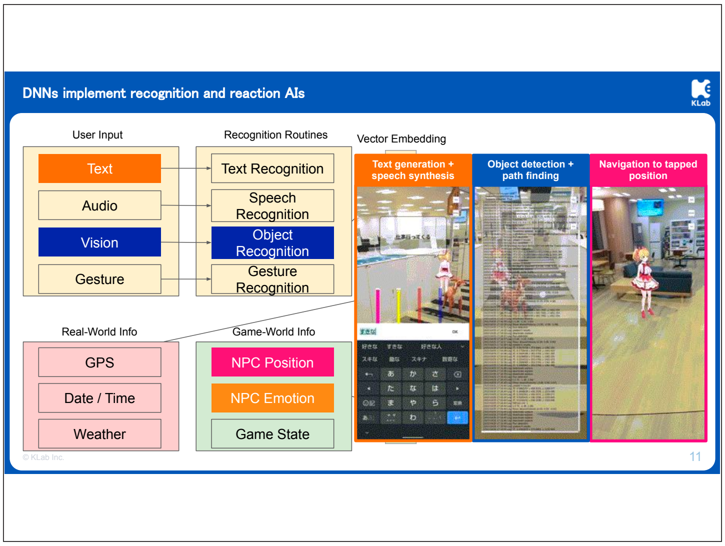
and more

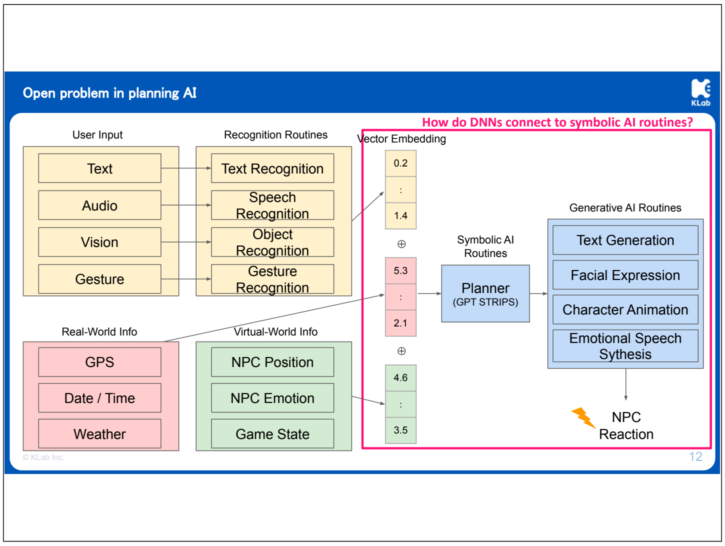
Artificial General Intelligence (AGI) for AR games

AR agents should cope with
- multimodal inputs/outputs
- game states
- planning

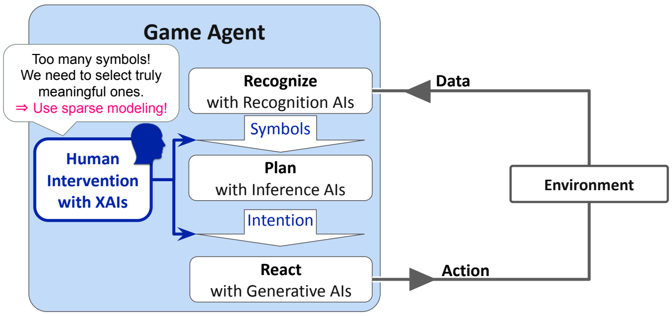



AR agent architecture
User Input
Text
Audio
Vision
Gesture
Recognition Routines
Text Recognition
Speech Recognition
Object Recognition
Gesture Recognition
Vector Embedding
0.2
1.4
⊕
5.3
2.1
⊕
4.6
3.5
Symbolic AI Routines
Planner
(KLab STRIPS)
Generative AI Routines
Text Generation
Facial Expression
Character Animation
Emotional Speech Sythesis
NPC Reaction
Real-World Info
GPS
Date / Time
Weather
Virtual-World Info
NPC Position
NPC Emotion
Game State
10


DNNs implement recognition and reaction AIs
User Input
Text
Audio
Vision
Gesture
Recognition Routines
Text Recognition
Speech Recognition
Object Recognition
Gesture Recognition
Vector Embedding
Text generation + speech synthesis
Object detection + path finding
Navigation to tapped position
Real-World Info
GPS
Date / Time
Weather
Game-World Info
NPC Position
NPC Emotion
Game State
11


Open problem in planning AI
How do DNNs connect to symbolic AI routines?
User Input
Text
Audio
Vision
Gesture
Recognition Routines
Text Recognition
Speech Recognition
Object Recognition
Gesture Recognition
Vector Embedding
0.2
1.4
⊕
5.3
2.1
⊕
4.6
3.5
Symbolic AI Routines
Planner
(GPT STRIPS)
Generative AI Routines
Text Generation
Facial Expression
Character Animation
Emotional Speech Sythesis
NPC Reaction
Real-World Info
GPS
Date / Time
Weather
Virtual-World Info
NPC Position
NPC Emotion
Game State
12

Our approach: Human-In-The-Neuro



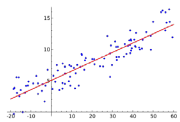
© KLab Inc. 13

Topology, Multiobjective Optimization and Sparse Modeling

© KLab Inc. 14

Elastic net and regularization path

Linear regression: $y = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n + \zeta$



Elastic net:

$$\text{minimize } E(\theta) = \frac{1}{2M}\underline{\|X\theta - y\|^2} + \lambda\left(\underline{\alpha|\theta|} + \frac{1-\alpha}{2}\underline{\|\theta\|^2}\right)$$

OLS error (goodness-of-fit)

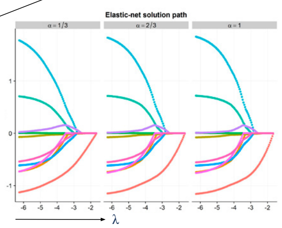
L1-regularization (sparseness)

L2-regularization (robustness)

- Optimize with different λ's and fixed α

  → Regularization path for α

- To tune two hyper-params (λ, α), we need to compute many regularization paths for different α's



© KLab Inc. 15

From elastic net problem to weakly simplicial problem

Our strategy

1. Show all multi–objective strongly convex problems are weakly simplicial
2. Reformulate the elastic net problem to a multi–objective strongly convex problem
3. Extend the regularization path on that problem and approximate it by a Bezier simplex

Multi–objective optimization problems

Optimize multiple functions:

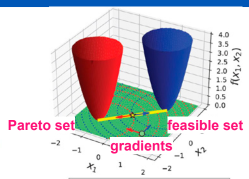
$$\text{minimize} \begin{array}{l} f_1(x_1, x_2) = (x_1+1)^2 + (x_2+1)^2 \\ f_2(x_1, x_2) = (x_1-1)^2 + (x_2-1)^2 \end{array}$$

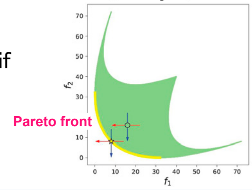
$$\text{subject to } -2 \le x_i \le 2 \quad (i = 1,2)$$

Goal: find Pareto set (rather than a point)

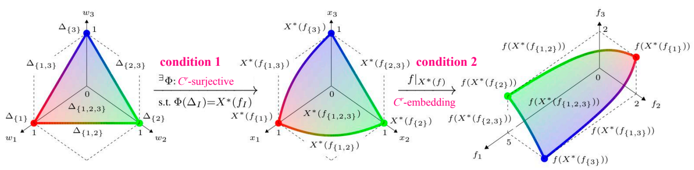
Definition: A point x is a Pareto solution if there is no y such that

- $f_i(y) \le f_i(x)$ for all i,
- $f_j(y) < f_j(x)$ for some j.





Weakly Simplicial Problems [Kobayashi+ 2019]



(a) Simplex Δ^2. (b) Pareto set $X^*(f)$. (c) Pareto front $f(X^*(f))$.

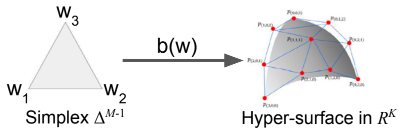
Figure 1: A weakly simplicial $f = (f_1, f_2, f_3) : \mathbb{R}^3 \to \mathbb{R}^3$. An M-objective problem f is C^r-*weakly simplicial* if two conditions are satisfied: (i) there exists a C^r-surjective map $\Phi : \Delta^{M-1} \to X^*(f)$ such that $\Phi(\Delta_I) = X^*(f_I)$ for all $I \subseteq \{1, \dots, M\}$; (ii) the restriction $f|_{X^*(f)} : X^*(f) \to \mathbb{R}^M$ is a C^r-embedding (and thus so is $f \circ \Phi : \Delta^{M-1} \to \mathbb{R}^M$).

where an (M - 1)-simplex is $\Delta^{M-1} = \left\{ (t_1, \dots, t_M) \in \mathbb{R}^M \,\middle|\, \sum_{m=1}^{M} t_m = 1,\ t_m \ge 0 \right\}$

an I-subsimplex is $\Delta_I^{M-1} = \{ (t_1, \dots, t_M) \in \Delta^{M-1} \mid t_m = 0 \ (m \notin I) \}$ for each $I \subseteq \{1, \dots, M\}$

A Bezier simplex approximates the Pareto set of any weakly simplicial problem

A Bezier simplex is a polynomial map of degree *D*:



$$b(w) = \sum_{d \in \mathbb{N}_D^M} \binom{D}{d} w^d p_d \qquad \mathbb{N}_D^M = \left\{ (d_1, \dots, d_M) \in \mathbb{N}^M \,\middle|\, \sum_{m=1}^{M} d_m = D \right\}$$

Theorem 1 *Let $\phi : \Delta^{M-1} \to \mathbb{R}^K$ be a continuous map. There exists an infinite sequence of Bézier simplices $\boldsymbol{b}^{(i)} : \Delta^{M-1} \to \mathbb{R}^K$ such that*

$$\lim_{i \to \infty} \sup_{\boldsymbol{t} \in \Delta^{M-1}} |\boldsymbol{\phi}(\boldsymbol{t}) - \boldsymbol{b}^{(i)}(\boldsymbol{t})| = 0.$$

Bezier simplex fitting [Tanaka+ 2020]

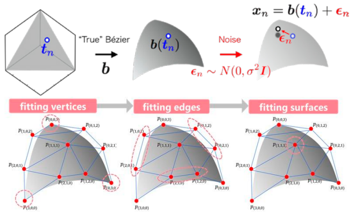


Table 2: Comparison of asymptotic risks of the all-at-once R_N^{AAO} vs. the inductive skeleton with the optimal subsample ratio R_N^{ISK} (M: the dimension of the Bézier simplex, D: the degree of the Bézier simplex, N: the sample size). The winner is shown in bold.

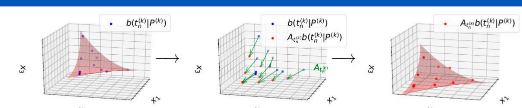
	$D=2$		$D=3$	
	R_N^{AAO}	R_N^{ISK}	R_N^{AAO}	R_N^{ISK}
$M=2$	$3.0/N$	$\mathbf{2.91}/N$	$4.0/N$	$\mathbf{3.73}/N$
$M=3$	$6.0/N$	$\mathbf{5.50}/N$	$\mathbf{10.0}/N$	$10.650/N$
$M=4$	$10.0/N$	$\mathbf{8.67}/N$	$\mathbf{20.0}/N$	$23.88/N$
$M=5$	$15.0/N$	$\mathbf{11.99}/N$	$\mathbf{35.0}/N$	$44.76/N$
$M=6$	$21.0/N$	$\mathbf{15.17}/N$	$\mathbf{56.0}/N$	$73.14/N$
$M=7$	$28.0/N$	$\mathbf{18.07}/N$	$\mathbf{84.0}/N$	$107.57/N$
$M=8$	$36.0/N$	$\mathbf{20.68}/N$	$\mathbf{120.0}/N$	$146.21/N$

$$R_N^{\mathrm{AAO}} \xrightarrow{p} \frac{\sigma^2 L}{N} \sum_{A,B} \Sigma_{AB} \Sigma_{AB}^{-1} = \frac{\sigma^2 L}{N} \binom{D+M-1}{D} \quad (N \to \infty)$$

$$R_N^{\mathrm{ISK}} \xrightarrow{p} \sum_{\boldsymbol{d}_A, \boldsymbol{d}_B \in \mathbb{N}_D^M} \Sigma_{AB} \left(\bigoplus_{i,j=0}^{M-1} \boldsymbol{X}^{(i)(j)} \right)_{AB} \qquad \boldsymbol{X}^{(i)(j)} := \sigma^2 L \sum_{\substack{m \le i \\ m \le j}} \sum_{\substack{m \le k_1 < \cdots < k_\heartsuit < i \\ m \le l_1 < \cdots < l_\spadesuit < j}} \frac{(-1)^{\heartsuit + \spadesuit}}{N^{(m)}} \boldsymbol{\Lambda}_\heartsuit \boldsymbol{\Lambda}_{(m)} \boldsymbol{\Lambda}_\spadesuit$$

[Tanaka+ 2020] Asymptotic Risk of Bézier Fitting

Bezier Flow: A Surface-wise Gradient Descent Method [Sannai+ 2022]

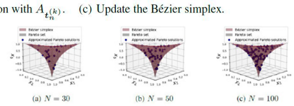


(a) Generate solutions on the Bézier simplex. (b) Update each solution with $A_{t_n^{(k)}}$. (c) Update the Bézier simplex.

Algorithm 2 Surface-wise Gradient Descent Method

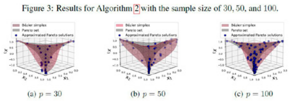
Set $k \leftarrow 1$ and the initial control point $\boldsymbol{P}^{(k)}$.
while $k \le K$ **do**
  Draw $\{\boldsymbol{t}_n^{(k)}\}_{n=1}^N$ for which each $\boldsymbol{t}_n^{(k)}$ is drawn i.i.d. from the uniform distribution on Δ^{M-1}.
  Obtain $\{\boldsymbol{b}(\boldsymbol{t}_n^{(k)}|\boldsymbol{P}^{(k)})\}_{n=1}^N$ by (1).
  Update $\{\boldsymbol{b}(\boldsymbol{t}_n^{(k)}|\boldsymbol{P}^{(k)})\}_{n=1}^N$ by (10).
  Update control points by (11).
  $k \leftarrow k+1$.
end while
return $\boldsymbol{P}^{(K+1)}$.

Theorem 6.5. *Assume that $\alpha^{(k)} \in (0,1]$ for all $k \in [K]$. Then, Algorithm 2 is PAC uniformly stable.*



(a) $N=30$ (b) $N=50$ (c) $N=100$

Figure 3: Results for Algorithm 2 with the sample size of 30, 50, and 100.



(a) $p=30$ (b) $p=50$ (c) $p=100$

Figure 4: Results for NSGA-II and the all-at-once with the population size of 30, 50, and 100.

[Sannai+ 2022] Bézier Flow: a Surface-wise Gradient Descent Method for Multi-objective Optimization

Strongly convex mappings

$f : \mathbb{R}^n \to \mathbb{R}$ is strongly convex $\overset{\text{def}}{\iff} \exists \alpha > 0, \forall x, y \in \mathbb{R}^n$, $\forall t \in [0, 1]$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x - y\|^2$$

where $\|z\|$ is the Euclidean norm of $z \in \mathbb{R}^n$. The constant α is called a convexity parameter of the function f.
A mapping $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ is strongly convex if all component functions are strongly convex.

Previous work

Previous work shows that if f is C^r strongly convex where $1 \leq r \leq \infty$, then the problem of minimizing f is C^{r-1} weakly simplicial.

references	strongly convex	weakly simplicial
[Hamada+ 2020]	C^∞	C^∞
[Hamada+ 2020]	C^r	C^{r-1} $(r \geqq 2)$
[Hamada+ 2021]	C^1	C^0

[Hamada+ 2020] N. Hamada, K. Hayano, S. Ichiki, Y. Kabata, and H. Teramoto, Topology of Pareto sets of strongly convex problems, SIAM J. Optim. 30, no. 3, 2659–2686.
[Hamada+ 2021] N. Hamada and S. Ichiki, Simpliciality of strongly convex problems, J. Math. Soc. Japan. 73, no. 3, 965–982.

If f is C^0 strongly convex, what happens?

Outline of proof

We can define a mapping $x^* : \Delta^{m-1} \to X^*(f)$ for any strongly convex mapping $f : \mathbb{R}^m \to \mathbb{R}^n$ as follows:

$$x^*(w) = \arg\min_{x \in \mathbb{R}^n} \left(\sum_{i=1}^m w_i f_i(x) \right),$$

where $\arg\min_{x \in \mathbb{R}^n} \left(\sum_{i=1}^m w_i f_i(x)\right)$ is the unique minimizer of $\sum_{i=1}^m w_i f_i$.

Theorem 2

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a strongly convex mapping. Then, the mapping $x^* : \Delta^{m-1} \to X^*(f)$ is surjective and continuous.

Thus, the problem of minimizing f is weakly simplicial.

From elastic net problem to weakly simplicial problem

Our strategy

1. Show all multi–objective strongly convex problems are weakly simplicial
2. Reformulate the elastic net problem to a multi–objective strongly convex problem
3. Extend the regularization path on that problem and approximate it by a Bezier simplex

© KLab Inc.

25

Multi–objective elastic net

Single–objective strongly convex problem

$$\text{minimize } E(\theta) = \underbrace{\frac{1}{2M}\|X\theta - y\|^2}_{f_1(\theta)\text{: convex}} + \lambda\left(\alpha\underbrace{|\theta|}_{f_2(\theta)\text{: convex}} + \frac{1-\alpha}{2}\underbrace{\|\theta\|^2}_{f_3(\theta)\text{: strongly convex}}\right)$$

⬇

Multi–objective strongly convex problem

$$\underset{\theta\in\mathbb{R}^n}{\text{minimize}}\ \tilde{f}(\theta) := (\tilde{f}_1(\theta), \tilde{f}_2(\theta), \tilde{f}_3(\theta))$$

$$\text{where } \underbrace{\tilde{f}_i(\theta)}_{\text{strongly convex}} = \underbrace{f_i(\theta)}_{\text{convex}} + \underbrace{\varepsilon f_3(\theta)}_{\text{strongly convex}} \quad (i = 1, 2, 3)$$

© KLab Inc.

26

From elastic net problem to weakly simplicial problem

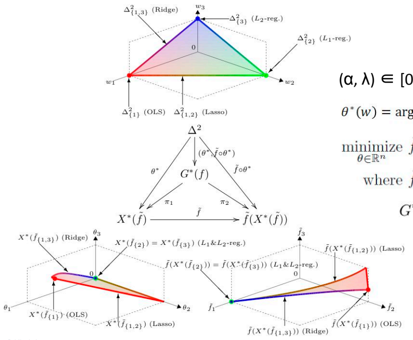
Our strategy

1. Show all multi–objective strongly convex problems are weakly simplicial
2. Reformulate the elastic net problem to a multi–objective strongly convex problem
3. Extend the regularization path on that problem and approximate it by a Bezier simplex

© KLab Inc.

27

Regularization Map: 2–Parameter Extension of Regularization Path



$(\alpha, \lambda) \in [0, \infty)^2 \to (w1, w2, w3) \in \Delta^2$

$\theta^*(w) = \operatorname{argmin}_\theta\ w_1 \tilde f_1(\theta) + w_2 \tilde f_2(\theta) + w_3 \tilde f_3(\theta)$

$$\underset{\theta \in \mathbb{R}^n}{\text{minimize}}\ \tilde f(\theta) := (\tilde f_1(\theta), \tilde f_2(\theta), \tilde f_3(\theta))$$

$$\text{where } \tilde f_i(\theta) = f_i(\theta) + \varepsilon f_3(\theta) \quad (i = 1, 2, 3)$$

$$G^*(\tilde f) = \{ (\theta, \tilde f(\theta)) \in \mathbb{R}^{n+3} \mid \theta \in X^*(\tilde f) \}$$

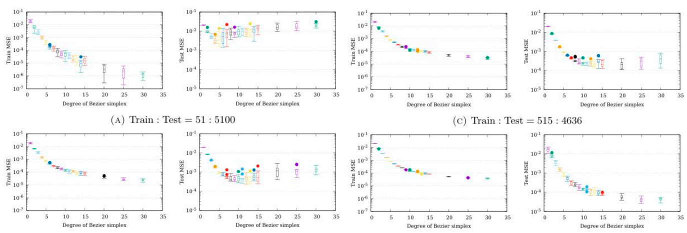
Experiments

1. For each dataset, train 5,151 elastic nets with different weights on Δ^2
2. Split a set of trained elastic nets into training/test sets for Bezier simplex fitting
3. Train a Bezier simplex with training set and evaluate its error with test set

	Attributes		
Dataset	Predictors	Responses	Instances
Blog Feedback[a] [2]	280	1	60,021
Fertility[b] [7]	9	1	100
Forest Fires[c] [4]	12	1	517
QSAR Fish Toxicity[d] [3]	6	1	908
Residential Building[e] [17]	103	2	372
Slice Localization[f] [8]	385	2	53,500
Wine[g] [1]	11	1	178
Yacht Hydrodynamics[h] [16]	6	1	308

[a] https://archive.ics.uci.edu/ml/datasets/BlogFeedback
[b] https://archive.ics.uci.edu/ml/datasets/Fertility
[c] https://archive.ics.uci.edu/ml/datasets/Forest+Fires
[d] https://archive.ics.uci.edu/ml/datasets/QSAR+fish+toxicity
[e] https://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set
[f] https://archive.ics.uci.edu/ml/datasets/Relative+location+of+CT+slices+on+axial+axis
[g] https://archive.ics.uci.edu/ml/datasets/wine
[h] https://archive.ics.uci.edu/ml/datasets/Yacht+Hydrodynamics

A Bezier Simplex Converges on the Regularization Map as n → ∞



(A) Train : Test = 51 : 5100
(B) Train : Test = 257 : 4894
(C) Train : Test = 515 : 4636
(D) Train : Test = 5100 : 51

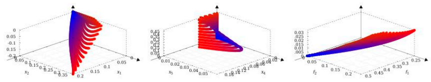
FIGURE 2. Degree versus MSE on QSAR Fish Toxicity.

The Optimal Degree of a Bezier Simplex is Not Sensitive to Datasets

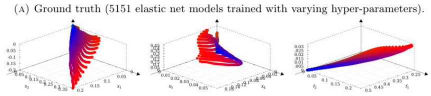
TABLE 2. Optimal degree d^* and its approximation error (average $\pm$ standard deviation over 10 trials).

	Large sample		Small sample	
Dataset	d^*	Test MSE	d^*	Test MSE
Blog Feedback	30	5.21E-04 ± 4.28E-04	1	5.62E-03 ± 1.26E-04
Fertility	30	4.71E-05 ± 1.34E-05	3	7.56E-03 ± 1.82E-03
Forest Fires	30	5.52E-05 ± 3.08E-05	3	7.17E-03 ± 1.11E-03
QSAR Fish Toxicity	25	4.16E-05 ± 1.09E-05	4	3.66E-03 ± 1.41E-03
Residential Building	25	3.55E-04 ± 2.55E-04	3	6.94E-03 ± 7.20E-04
Slice Localization	30	5.95E-04 ± 4.38E-04	3	8.83E-03 ± 1.60E-03
Wine	30	6.71E-05 ± 1.42E-05	3	7.00E-03 ± 5.63E-04
Yacht Hydrodynamics	30	6.75E-05 ± 4.32E-05	3	3.51E-03 ± 3.62E-04

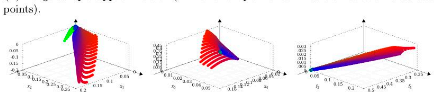
Regularization Map: Ground Truth vs. Approximation



(A) Ground truth (5151 elastic net models trained with varying hyper-parameters).



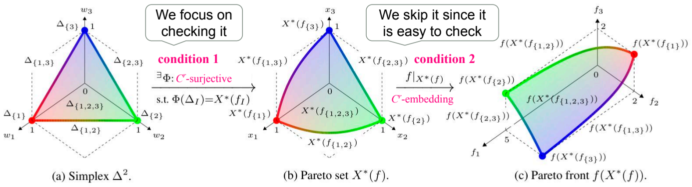
(B) Large sample approximation (A Bézier simplex of $d = 25$ trained with 5100 data points).



(C) Small sample approximation (A Bézier simplex of $d = 4$ trained with 51 data points)

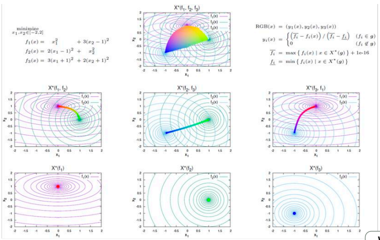
Generalization and Poincaré Conjecture

Generalization to non–strongly convex cases: Statistical test of simplicity



- When the simplicity of the problem is unknown, we need to estimate it from data
- We statistically test Condition 1 from a sample of Pareto set and front

Make the condition 1 easier: condition 1 ⇔ conditions 1a & 1b



We focus on checking it

We skip it since it is easy to check

- Cond 1a : $X^*(f_I) \simeq \Delta^{|I|-1}\ (\forall I)$
- Cond 1b : $\operatorname{int} X^*(f_I) \cap \operatorname{int} X^*(f_J) = \emptyset\ (\forall I, J \text{ s.t. } I \neq J)$

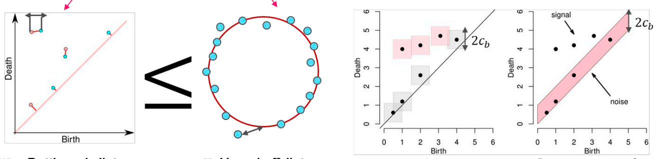
[Hamada 2017] Simple problems: The simplicial gluing structure of Pareto sets and Pareto fronts. In GECCO 2017, pp. 315-316.

Homology Estimation Using Confidence Set of Persistent Diagram

Stability theorem of persistence diagrams:

$$\mathbb{P}(W_\infty(\widehat{\mathcal{P}}, \mathcal{P}) > c_b) \le \mathbb{P}\left(\mathbf{H}(\mathcal{S}_n, \mathbb{M}) > c_b\right) \le \underline{\alpha} + O\left(\frac{b}{n}\right)^{1/4}$$ for $n \to \infty$

α-confidence interval of noise



W_∞: Bottleneck distance

H: Hausdorff distance

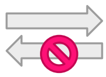
$\mathbb{M}$ は未知なので、S_n のブートストラップにより c_b を推定する

[Fasy+ 2014] Confidence sets for persistence diagrams. Ann. Statist. Vol. 42, No. 6, pp. 2301-2339.

There is a missing link...

In general,

Topological spaces
X and Y
are homeomorphic



Homologies
$H_*(X)$ and $H_*(Y)$
are isomorphic

We will seek a condition
where the converse is true.

In a certain assumption, homology implies homeomorphism

Theorem 1. *Let M be an n-dimensional compact C^∞-manifold with boundary such that M and ∂M are both simply connected. If M is homologous to an n-dimensional simplex Δ^n, i.e.,*

$$H_q(M) \cong H_q(\Delta^n) \cong \begin{cases} \mathbb{Z} & (q = 0), \\ 0 & (q \neq 0), \end{cases}$$

then M is homeomorphic to Δ^n.

To show the theorem, we uses the following lemmas.

Lemma 1 (Corollary of Hurewicz's theorem)**.** *Let M, N be simply connected CW-complexes and $f : M \to N$ be a continuous mapping. If the homomorphism between homology $f_* : H_*(M) \to H_*(N)$ introduced by f is isomorphism, then f is homology equivalence.*

⇔ n-dimensional Poincaré conjecture

Lemma 2 (h-cobordism theorem)**.** *Let M, N be a simply connected n-dimensional closed C^∞-manifold. Let W be a simply connected $(n+1)$-dimensional compact C^∞-manifold such that $\partial W = M \cup N$ (disjoint union). If the inclusion mappings*

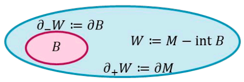
$$M \hookrightarrow W, \quad N \hookrightarrow W$$

are both homotopy equivalence, then W is homeomorphic to $M \times [0, 1]$ (see remark 1). This homeomorphism can be the identity map on $M \times \{0\}$. W is called the h-cobordism between M and N.

Proof

Proof of Theorem 1. Let M be an n-dimensional compact C^∞-manifold. Let B be a C^∞-submanifold of M that is homeomorphic to Δ^n. Let W be $M - \mathrm{int}B$ and $\partial_+ W := \partial M$, $\partial_- W := \partial B$.

From the assumption of the homology of M, we have



$$H_q(M, \partial M) \cong \begin{cases} \mathbb{Z} & (q = n), \\ 0 & (q \neq n). \end{cases}$$

It implies $H_q(W, \partial_+ W) = 0$ for all q. Substituting this to the long exact sequence

$$\cdots \to H_{q+1}(W, \partial_+ W) \to H_q(\partial_+ W) \xrightarrow{i_*} H_q(W) \to H_q(W, \partial_+ W) \to \cdots$$

we see that $i_* : H_q(\partial_+ W) \to H_q(W)$ is an isomorphism. Since W is compact by its construction, lemma 1 shows that $i : \partial_+ W \hookrightarrow W$ is a homotopy equivalence. The same argument repeats for $\partial_- W$. By lemma 2, W is an h-cobordism between $\partial_+ W$ and $\partial_- W$.

Remember we have assumed $\partial_+ W$ and $\partial_- W$ are simply connected. Thus, lemma 2 ensures that there exists a homeomorphism $\phi : W \to \partial_- W \times [0, 1]$ such that $\phi|_{\partial_- W} = \mathrm{id} \times \{0\}$. Since $(\phi \cup \mathrm{id}) : M \to (W \cup B)$ is homeomorphism, and $(W \cup B)$ is homeomorphic to Δ^n by its construction, M is homeomorphic to Δ^n. □

Takeaways

Mobile games bridge between math and industry

Math ← Data Science — HPC → Game

Math	Data Science	HPC	Game
Singularity Theory Optimization	Machine Learning Data Analysis	Supercomputer Cloud	Game Dev/Ops Game AI
with IMI, Kyushu U.	with RIKEN AIP	with RIIT, Kyushu U.	in KLab

We always welcome collaboration with you!

40

KLab

41

Persistent Homology and Machine Learning

Yuichi Ike
Institute of Mathematics for Industry, Kyushu University, Japan

Persistent homology is a central tool in topological data analysis. It encodes the topological features of given data into persistence diagrams, which are multisets in the two-dimensional space. In connection with machine learning, persistence diagrams have been used as an input of machine learning algorithms as feature vectors and are effectively applied in material science and medical science. Recently, many techniques have been developed to incorporate persistence diagrams into loss functions for controlling the topology of parameters. In this talk, I will start with the basics of persistent homology and some applications. Then I would like to discuss several recent developments in optimizing TDA-based loss functions and their applications in dimensionality reduction or visualization.

Persistent Homology and Machine Learning

WORKSHOP on Mathematics for Industry

2023-09-27

Yuichi Ike
Institute of Mathematics for Industry, Kyushu University

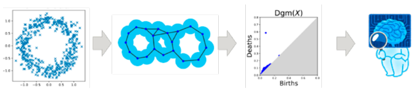
Joint work with
Mathieu Carrière, Frédéric Chazal, Marc Glisse,
Hariprasad Kannan, Théo Lacombe, Martin Royer, Yuhei Umeda
(Collaboration with Inria and Fujitsu)

1/26

Outline

1. Persistent Homology and Applications

- "Shape" of data and persistence diagrams
- Typical applications of persistent homology



2. PH-based Loss Functions

- Differentiability of persistence diagrams
- Applications of PH-based loss functions

2/26

Persistent Homology and Applications

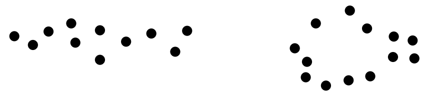
- Extracting the shape of data
- Persistent homology and persistence diagrams (PDs)
- Some applications

3/26

Idea of persistent homology (PH)

Topological Data Analysis (TDA)

■Method to extract topological features of data



Without hole One hole

Q. How to extract the "topology" of a discrete point cloud?

4/26

Idea of persistent homology (PH)

Topological Data Analysis (TDA)

■Method to extract topological features of data

Without hole One hole

Q. How to extract the "topology" of a discrete point cloud?

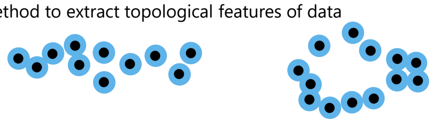
■Idea1: Consider **the union of balls centered at data points**

5/26

Idea of persistent homology (PH)

Topological Data Analysis (TDA)

■Method to extract topological features of data



Without hole One hole

Q. How to extract the "topology" of a discrete point cloud?

■Idea1: Consider **the union of balls centered at data points**

6/26

Idea of persistent homology (PH)

Topological Data Analysis (TDA)

■Method to extract topological features of data

Without hole **One hole**

Q. How to extract the "topology" of a discrete point cloud?

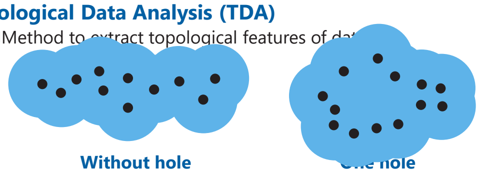
■Idea1: Consider **the union of balls centered at data points**
-> How to adjust the radius value of balls?

7/26

Idea of persistent homology (PH)

Topological Data Analysis (TDA)

■Method to extract topological features of data



Without hole

Q. How to extract the "topology" of a discrete point cloud?

■Idea1: Consider **the union of balls centered at data points**
-> How to adjust the radius value of balls?

■Idea2: Consider all radii and **track the evolution**: **persistent homology**
-> Can **distinguish noise and essential topological features**

8/26

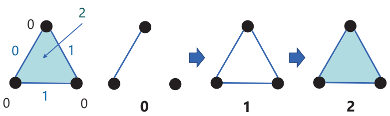
Filtrations on simplicial complexes

Filtration: increasing family of subcomplexes

■**Simplicial complex**: collection $K \subset 2^V$ s.t. $\sigma \in K, \tau \subset \sigma \Rightarrow \tau \in K$

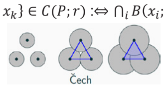
■$\mathcal{K} = (K_r)_r, K_r \subset K$ filtration of K: $\Leftrightarrow K_r \subset K_s$ $(r \leq s)$ and $\cup_r K_r = K$

$\leftrightarrow$ Function $f: K \to \mathbb{R}$ s.t. $\sigma \subset \tau \Rightarrow f(\sigma) \leq f(\tau)$, $K_r = \{\sigma \in K \mid f(\sigma) \leq r\}$



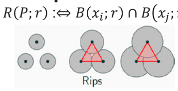
Čech filtration

$\{x_0, \dots x_k\} \in C(P; r) :\Leftrightarrow \cap_i B(x_i; r) \neq \emptyset$



Rips filtration

$\{x_0, \dots x_k\} \in R(P; r) :\Leftrightarrow B(x_i; r) \cap B(x_j; r) \neq \emptyset$ $(\forall i, j)$



9/26

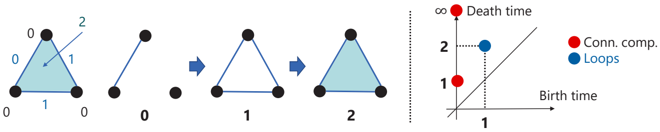
PH and persistence diagrams (PDs)

Filtration: increasing family of subcomplexes

■**Simplicial complex**: collection $K \subset 2^V$ s.t. $\sigma \in K, \tau \subset \sigma \Rightarrow \tau \in K$

■$\mathcal{K} = (K_r)_r, K_r \subset K$ filtration of K: $\Leftrightarrow$ $K_r \subset K_s$ $(r \leq s)$ and $\cup_r K_r = K$

$\leftrightarrow$ Function $f: K \to \mathbb{R}$ s.t. $\sigma \subset \tau \Rightarrow f(\sigma) \leq f(\tau)$, $K_r = \{\sigma \in K \mid f(\sigma) \leq r\}$



Persistent homology of $\mathcal{K} = (K_r)_r$ is the family

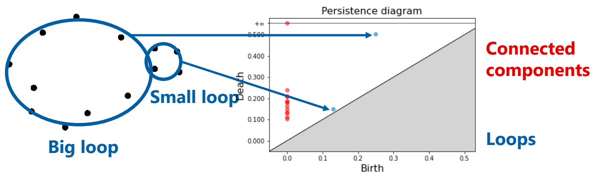
$$\cdots \to H_n(K_r) \to H_n(K_s) \to H_n(K_t) \to \cdots \quad (r \leq s \leq t)$$

⇝**Persistence diagram (PD)**: encodes the birth and death time of each homology class

How to use PDs

■Points far from the diagonal express essential shapes while those near the diagonal are regarded as noises

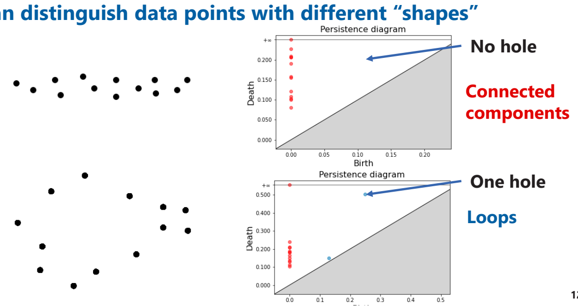
■We can analyze which type of shape is represented by a point in PD



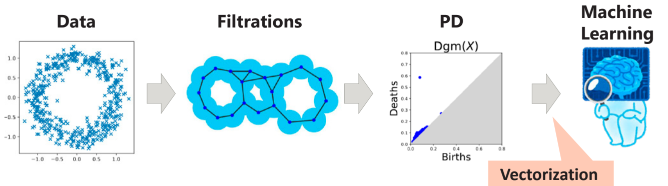
How to use PDs

More directly …

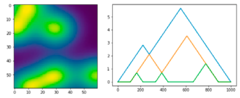
We can distinguish data points with different "shapes"



Analysis with persistent homology



Example
- Point clouds generated by simulation
- Graphs
- Image data



e.g., persistence image, persistence landscape

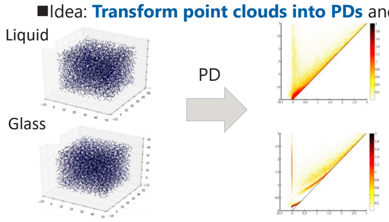
Analysis of silica glass

Kusano, Fukumizu, and Hiraoka: Persistence weighted Gaussian kernel for topological data analysis, ICML2016

■cf. Nakamura et al.: Description of medium-range order in amorphous structures by persistent homology.

■Estimate the temperature that SiO_2 changes from liquid to glass state

■Idea: **Transform point clouds into PDs** and analyze them



\+ Vectorization of PDs (They used kernel method)

Figures from K. Fukumizu, Persistence Weighted Gaussian Kernel for Topological Data Analysis

Application to graph classification

PH extracts **some global structure of graphs**, which can be used for classification

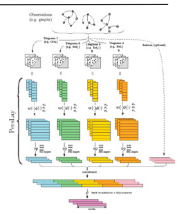
■Need to find **suitable filtrations on graphs**
e.g., degree function, Heat Kernel Signature

PersLay, M. Carrière, F. Chazal, I., T. Lacombe, M. Royer, Y. Umeda, AISTATS 2020

■Proposed a new architecture for graph classification

■PersLay (NN vectorization) + one-layer NN

More and more studies to **combine PH and Machine Learning**



Dataset	SV[1]	RetGK* [2]	FGSD [3]	GCNN [4]	GIN [5]	PersLay Mean	PersLay Max
REDDIT5K	—	56.1	47.8	52.9	57.0	55.6	56.5
REDDIT12K	—	48.7	—	46.6	—	47.7	49.1
COLLAB	—	81.0	80.0	79.6	80.1	76.4	78.0
IMDB-B	72.9	71.9	73.6	73.1	74.3	71.2	72.6
IMDB-M	50.3	47.7	52.4	50.3	52.1	48.8	52.2
COX2*	78.4	80.1	—	—	—	80.9	81.6
DHFR*	78.4	81.5	—	—	—	80.3	80.9
MUTAG*	88.3	90.3	92.1	86.7	89.0	89.8	91.5
PROTEINS*	72.6	75.8	73.4	76.3	75.9	74.8	75.9
NCI1*	71.6	84.5	79.8	78.4	82.7	73.5	74.0
NCI109*	70.5	—	78.8	—	—	69.5	70.1

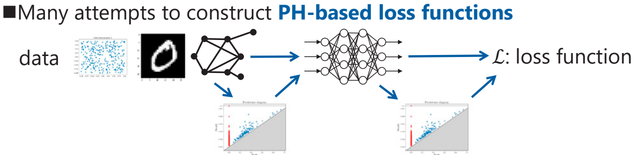
PH-based Loss Functions

- Applications of PH-based loss functions
- Differentiability and convergence of PH-based functions
- Applications

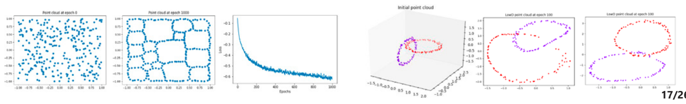
16/26

PH-based loss functions

■Many attempts to construct **PH-based loss functions**



■**Brüel-Gabrielsson et al., A Topology Layer for Machine Learning, AISTATS2020**: deformation of point clouds, topological generative models
■**Moor et al., Topological Autoencoders, ICML2020**: Topology-preserving AE



17/26

Parametrized filtrations and PDs

Recall: A filtration of a simplicial complex K
$\leftrightarrow$ a vector $f \in \mathbb{R}^K$ s.t. $\sigma \subset \tau \Rightarrow f_\sigma \le f_\tau$

$$\mathrm{Filt}_K := \{f \in \mathbb{R}^K \mid \sigma \subset \tau \Rightarrow f_\sigma \le f_\tau\}$$

A **parametrized filtration**: a function $F: A \to \mathrm{Filt}_K$, where $A \subset \mathbb{R}^d$

■**Rips filtration** $F: (\mathbb{R}^d)^N \to \mathbb{R}^{|\Delta_N|}$, $F_\sigma(x) := {}^1\!/_2 \max_{i,j\in\sigma} \|x_i - x_j\|$
■**Parameters in ML** $f_\theta: \mathbb{R}^d \to \mathbb{R}^D$ $(\theta \in \Theta)$, $P \subset \mathbb{R}^d$: finite subset
$F: \Theta \to \mathbb{R}^{|\Delta_N|}$, $F_\sigma(\theta) := {}^1\!/_2 \max_{i,j\in\sigma} \|f_\theta(x_i) - f_\theta(x_j)\|$

The PD is a vector in $\mathbb{R}^{|K|}$: $(p_1, \dots, p_m, e_1, \dots, e_n)$, $p_i \in \mathbb{R}^2, e_i \in \mathbb{R}$
$\Rightarrow$ The assignment is viewed as $\mathbf{Pers}\colon \mathbf{Filt}_K \to \mathbb{R}^{|K|}$ **persistence map**
$\mathbf{Pers} \circ \boldsymbol{F}\colon \boldsymbol{A} \to \mathbb{R}^{|K|}$ **parametrized PD**

18/26

Optimization of PH-based functions

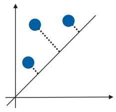
■Q. How can we optimize PH-based loss functions?

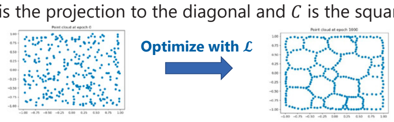
■A. Usually just apply **gradient descent**

■For a **differentiable function** $\mathcal{L}: A \to \mathbb{R}, A \subset \mathbb{R}^d$, update the parameter by
$$x_{k+1} = x_k - \alpha_k \nabla \mathcal{L}(x_k),$$
where α_k is the learning rate at step k

■Toy example: optimize a point set to **maximize "# of loops"**

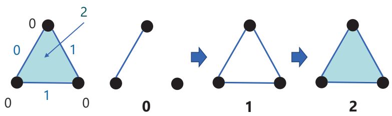
■For the 1st PD $D_1(P)$, consider
$$\mathcal{L}(P) = -\sum_{p \in D_1(P)} \|p - \pi_\Delta(p)\|_\infty^2 + d(P, C),$$
where π_Δ is the projection to the diagonal and C is the square

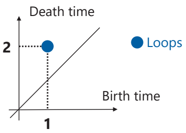




Differentiability of persistence map

■How to compute PDs from filtrations?





1. Find **the pairs of birth and death simplices** $\{(\sigma_{b_i}, \sigma_{d_i})\}_i$ (combinatorial)

   e.g.  : birth : death

2. **Associate the filtration value** to each pair $\{(F(\sigma_{b_i}), F(\sigma_{d_i}))\}_i$

   e.g. (1,2)

■F is smoothly parametrized => can consider $x \mapsto \{(\nabla_x F(\sigma_{b_i}), \nabla_x F(\sigma_{d_i}))\}_i$
in the area where the order of simplices does not change

PH-based functions and subdifferential

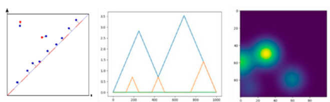
Function of PDs: a **permutation invariant function** E: $\mathbb{R}^{|K|} \to \mathbb{R}$
$$E(p_{\alpha(1)}, \dots, p_{\alpha(m)}, e_{\beta(1)}, \dots, e_{\beta(n)}) = E(p_1, \dots, p_m, e_1, \dots, e_n)$$

e.g.

■Distance between PDs,

■Persistence landscape,

■Persistence image, ...



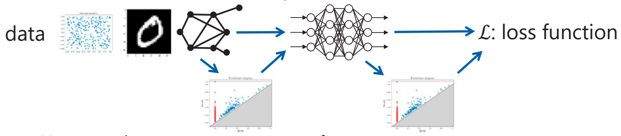
If F and E are in a good class, then so is $\mathcal{L} = E \circ \mathrm{Pers} \circ F$: $A \to \mathbb{R}$ and it is **differentiable a.e.**

=> can define the subdifferential
$$\partial \mathcal{L}(z) = \mathrm{Conv}\left\{\lim_{z_i \to z} \nabla \mathcal{L}(z_i) : \mathcal{L} \text{ is differentiable at } z_i\right\}$$

Figures from GUDHI Library webpage

Convergence of PH-based functions

We can apply stochastic (sub)gradient descent to optimize **PH-based loss functions** using automatic differentiation



- However, there was no guarantee of convergence

Carrière, Chazal, Glisse, I., Kannan, and Umeda, Optimizing persistent homology based functions, ICML2021

- Proved the **almost surely convergence of stochastic subgradient descent** for **a wide class of PH-based loss functions**
- The class includes almost all the PH-based functions in the literature

22/26

Convergence result

- **Theorem** K simplicial compex, $F\colon A \to \mathbb{R}^{|K|}$ parametrized family of filtration, $E\colon \mathbb{R}^{|K|} \to \mathbb{R}$ function of PDs, $\mathcal{L} = E \circ \mathrm{Pers} \circ F\colon A \to \mathbb{R}$
  Assume that **F and E are in a good class** (definable) and **$\mathcal{L}$ is locally Lipschitz**.

  Consider the sequence obtained by
  $$x_{k+1} = x_k - \alpha_k(y_k + \xi_k), \qquad y_k \in \partial\mathcal{L}(x_k),$$
  where α_k: learning rate and ξ_k: random variable s.t.
  1. $\alpha_k \geq 0,\ \sum_k \alpha_k = \infty, \sum_k \alpha_k^2 < \infty$;
  2. $\sup_k \|x_k\| < \infty$ almost surely;
  3. For $\mathcal{F}_k = \sigma(x_j, y_j, \xi_j, j < k)$, there exists a function $p\colon \mathbb{R}^d \to \mathbb{R}$ that is bounded on any bounded set s.t. for any k almost surely
  $$\mathbb{E}[\,\xi_k | \mathcal{F}_k] = 0, \qquad \mathbb{E}[|\xi_k|^2 | \mathcal{F}_k] < p(x_k).$$

  Then **$(x_k)_k$ converges to a critical point of $\mathcal{L}$ almost surely**.

cf. Davis et al., Stochastic subgradient method converges on tame functions, 2020

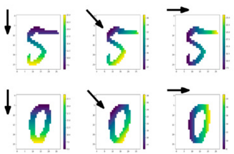
23/26

Applications: Filtration learning

Learn filtrations to give PDs for a classification task

- Consider the toy task to classify the MNIST images with 0th PDs + RF
- For a linear function f to some direction, consider
$$\mathcal{L}(f) = \sum_l \frac{\sum_{y_i = y_j = l} d\left(D_0(I_i, f), D_0(I_j, f)\right)}{\sum_{y_i = l} d\left(D_0(I_i; f), D_0(I_j; f)\right)}.$$

Optimize $\mathcal{L}(f)$ to find the best direction.



Dataset	Baseline	Before	After	Difference	Dataset	Baseline	Before	After	Difference
vs01	100.0	61.3	99.0	**+37.6**	vs26	99.7	98.8	98.2	-0.6
vs02	99.4	98.8	97.2	-1.6	vs28	99.1	96.8	96.8	0.0
vs06	99.4	87.3	98.2	**+10.9**	vs29	99.1	91.6	98.6	**+7.0**
vs09	99.4	86.8	98.3	**+11.5**	vs34	99.8	99.4	99.1	-0.3
vs16	99.7	89.0	97.3	**+8.3**	vs36	99.7	99.3	99.3	-0.1
vs19	99.6	84.8	98.0	**+13.2**	vs37	98.9	94.9	97.5	**+2.6**
vs24	99.4	98.7	98.7	0.0	vs57	99.7	90.5	97.2	**+6.7**
vs25	99.4	80.6	97.2	**+16.6**	vs79	99.1	85.3	96.9	**+11.5**

Classification accuracy before and after filter optimization

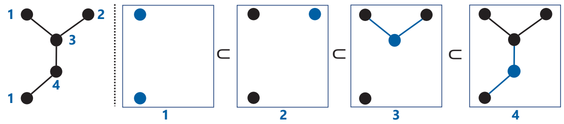
24/26

Applications: Graph Filtration Learning

Hofer et al., Graph Filtration Learning, ICML2020

■**Learn a filter function of graphs end-to-end**

■Recall that **a function on vertices** gives a filtration of a graph



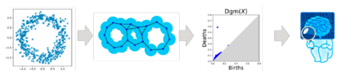
■Vectorization of the resulting PDs is used for classification

■Parametrized vertex filter function **can be implemented by GNN** and **learned thanks to the differentiability and the convergence result**

Summary

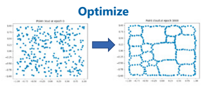
1. Persistent Homology and Applications

- ●Extract "topology" of data as persistence diagrams (PDs)
- ●We can use PD as input of machine learning (ML)
- ●Applications: material science, graph classification, ...



2. PH-based Loss Functions

- ●Many attempts to **combine PH and ML**
- ●Proved the **convergence of SSGD** for a PH-based loss functions
- ●Developing thanks to differentiability of PDs
  - ● Filtration Leaning, Topologically Regularized Embeddings, ...



Thank you for your attention!

Exotic shapes of nano-spherical structures - new DNA coding

Stanisław Janeczko
Center for Advanced Studies,
Warsaw University of Technology, Poland
(joint work with Hassan Babiker)

The simplest naturally ordered tetrahedral packing consists of an ordered sequence of regular tetrahedra glued together face to face as with the linear packing of a tetrahedral helix.Such tetrahedral structures are called *tetrahedral chains*.

Any tetrahedral chain consists of the three types of simplest configurations of four consecutive tetrahedra called *tetrahedral units*. Two of these types are left and right tetrahedral short spirals, U, D, and the third type, F, is a flat configuration of four tetrahedra. The structure of a tetrahedral chain in D, F, U elementary units is written as a word like $UUDFUD....$

The three strands of the left or right oriented tetrahedral helix form a spiral with irrational slope. This is the reason for the effective density of tetrahedral chains and nonexistence of closed tetrahedral chains in Euclidean space.

Let us assume that the gluing process of tetrahedra is ordered along a chain and each step of this process is realized by reflection in a particular face of adjacent tetrahedron. To each tetrahedron we assign four reflections $R_i, i = 1, \ldots, 4$, in the configurational three dimensional space V. Reflections R_i in V are represented by four corresponding reflect-morphisms $\bar{R}_i, i = 1, \ldots, 4$, acting in the space of regular tetrahedra $\mathcal{T}$ through a reflectional transformation of their vertices. In $V, \dim V = n$, any tetrahedral chain of length $n + 1$ is uniquely represented by an initial tetrahedron T and an ordered sequence of n reflect-morphisms

$$\bar{R}_{i_1}, \ldots, \bar{R}_{i_n}, \quad i_k \neq i_{k+1}, k = 1, \ldots, n-1.$$

The fact that a tetrahedral chain is so rigid in 3-space and regular tetrahedra can not tile the space gives rise to several questions. The main question we consider is the recognition of combinatorial and algebraic structures of tetrahedral chains. We want to investigate their geometric properties and determine what kind of information is contained in the chain invariants of orthogonal transformations and re-numberings. We use the parametrization of the chains by sequences of ordered reflections in barycentric coordinates and find their combinatorial structure. Periodicity along a chain is based on the structure of sequences of admissible triplets of integers and their cycling properties. The corresponding numerical invariants and an indexing role of a binary tetrahedral group defines the complete coding properties in dimension three.

Exotic shapes of nano-spherical structures - new DNA coding

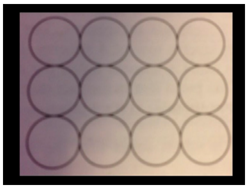
Stanislaw Janeczko

Faculty of Mathematics and Information Sciences, Warsaw University of Technology

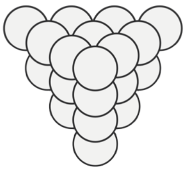
Workshop on Mathematics for Industry
Kyushu - Warsaw 25-29 September 2023

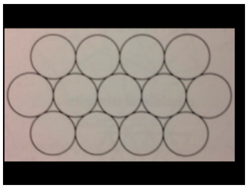
Sphere packings

- Square packing, face-centered cubic packing



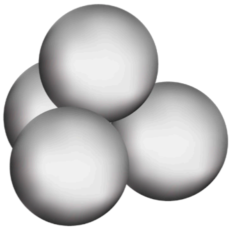
- Barrow boy's packing, cell is a rhombic dodecahedron



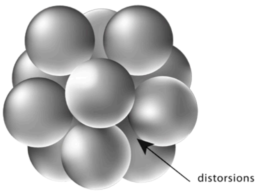


- Hexagonal packing, the third layer sits exactly above the first layer.

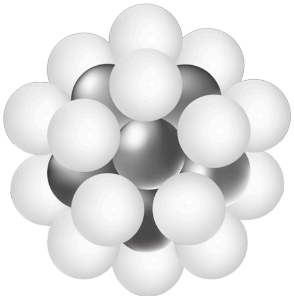
Sphere packing



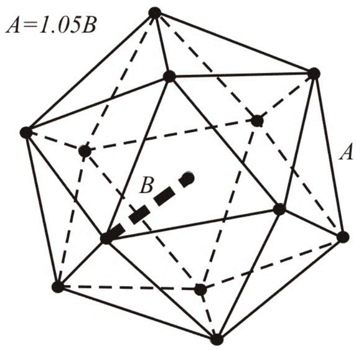
Sphere packing



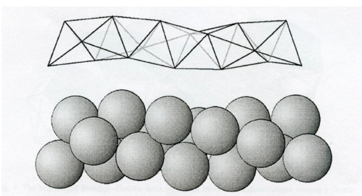
Sphere packing



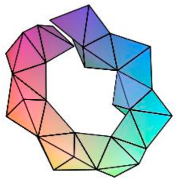
Icosahedron



Tetrahedral chains



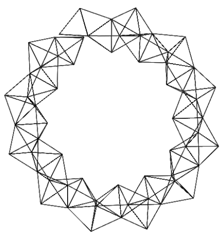
Tetrahedral chains



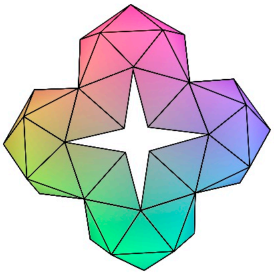
H. Steinhaus, 1957; J.H. Mason, 1972

Tetrahedral chains

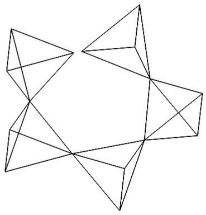
Tetrahedral chains



Almost closed tetrahedral chains



Dual tetrahedral chains



Tetrahedra in barycentric coordinates

$$T \equiv \{p_1, p_2, p_3, p_4\}, \{(S_1, p_1), \ldots, (S_4, p_4)\}$$

$\mathcal{T}-$regular tetrahedra, $\| p_i - p_j \| = \| p_k - p_l \|, i \neq j, k \neq l$

$$\mathcal{T} \subset V \otimes U^*, \quad U \equiv \mathbb{R}^4$$

V - configurational affine space, $dim V = 3$

U - barycentric coordinates $(\alpha_1, \ldots, \alpha_4) \in U$

$H = \{\Sigma_{i=1}^4 \alpha_i = 1\}$ - canonical affine hyperplane

$T \in \mathcal{T}, T = \Sigma_{i=1}^{4} p_i \otimes e_i^*$

Barycentric coordinate map $\mathbb{T} : H \to V$:

$\mathbb{T}(\alpha) = \Sigma_{i=1}^{4} p_i \otimes e_i^*(\alpha) = \Sigma_{i=1}^{4} \alpha_i p_i$,

$\alpha = \Sigma_{i=1}^{4} \alpha_i e_i \in H$, and geometrically

$T = \mathbb{T}(H \cap \{\alpha_i \geq 0\})$

$F : V \to V$ affine mapping.

F lifts to a linear mapping

$$M : (U, H) \to (U, H)$$

preserving the hyperplane H

M is defined uniquely by the commuting diagram

$$\mathbb{T}(M(\bullet)) = F(\mathbb{T}(\bullet))$$

$F(p_i) = \Sigma_{j=1}^{4} \alpha_{ji} p_j$ in barycentric coordinates α_{ji}.

Then

$$\sum_{i=1}^{4} \sum_{j=1}^{4} \alpha_{ji} p_j \otimes e_i^* = \sum_{j=1}^{4} p_j \otimes (\sum_{i=1}^{4} \alpha_{ji} e_i^*) = \sum_{j=1}^{4} p_j \otimes M^*(e_j^*).$$

Generation of tetrahedral chain

s_i center of S_i, $s_i = \frac{1}{3}(\Sigma_{j=1}^{4} p_j - p_i)$

Four orthogonal reflections by S_i

$$R_i(p) = p - 2\frac{(p - s_i | s_i - p_i)}{(s_i - p_i | s_i - p_i)}(s_i - p_i)$$

$$R_i(p_j) = p_j + 2\delta_{ij}(\frac{1}{3}\sum_{k \neq i} p_k - p_j), \quad j = 1, \ldots, 4$$

$\{T^{(i)}\}_{i=0}^{n}$ tetrahedral chain

$$
\begin{aligned}
T^{(0)} &= T, \\
T^{(1)}_{i_1} &= \bar{R}_{i_1} T, \\
T^{(2)}_{i_1 i_2} &= \bar{R}_{i_2} \bar{R}_{i_1} T, \quad i_1 \neq i_2, \\
& \ldots \quad \ldots \quad \ldots \\
T^{(n)}_{i_1 i_2 \ldots i_n} &= \bar{R}_{i_n} \ldots \bar{R}_{i_2} \bar{R}_{i_1} T, \quad i_{k+1} \neq i_k, k = 1, \ldots, n-1.
\end{aligned}
$$

$\bar{R}_i : \mathcal{T} \to \mathcal{T}$ twist morphisms, defined by R_i.

Representation in barycentric coordinates

$\bar{R}_i : \mathcal{T} \to \mathcal{T}, \quad \bar{R}_i(v \otimes u^*) = v \otimes M_i^* u^*$

$$
M_1 = \begin{pmatrix} -1 & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^T, M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & -1 & \frac{2}{3} & \frac{2}{3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^T,
$$

$$
M_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & -1 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}^T, M_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & -1 \end{pmatrix}^T.
$$

$\bar{R}_i$ is represented by transpose of M_i

EXAMPLE

$$
\bar{R}_1\left(\sum_{i=1}^{4} p_i \otimes e_i^*\right) = \sum_{i=1}^{4} p_i^{(1)_1} \otimes e_i^*,
$$

where

$$
\begin{pmatrix} p_1^{(1)_1} \\ p_2^{(1)_1} \\ p_3^{(1)_1} \\ p_4^{(1)_1} \end{pmatrix} = \begin{pmatrix} -1 & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix},
$$

$$
T^{(n)}_{i_1 \ldots i_n} = \bar{R}_{i_n} \ldots \bar{R}_{i_1} T.
$$

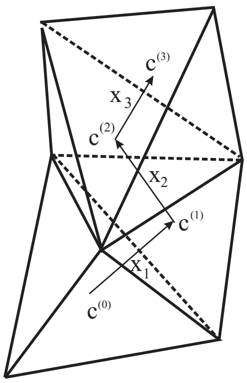
Coding in triplets of consecutive steps

$$\begin{aligned} T_k^{(r+1)} &= \bar{R}_k T^{(r)} \\ T_{kj}^{(r+2)} &= \bar{R}_j \bar{R}_k T^{(r)} \\ T_{kji}^{(r+3)} &= \bar{R}_i \bar{R}_j \bar{R}_k T^{(r)}. \end{aligned}$$

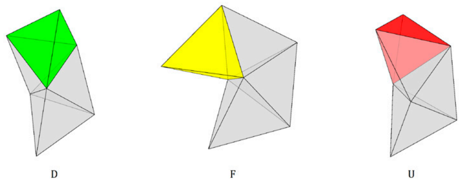
$$U, D, F: \quad T_{kji}^{(r+3)} = \bar{R}_i \bar{R}_j \bar{R}_k T^{(r)}$$

$$\begin{aligned} F:&\quad T^{(r+3)}; \quad det(x_{r+1}, x_{r+2}, x_{r+3}) = 0 \\ U:&\quad T^{(r+3)}; \quad det(x_{r+1}, x_{r+2}, x_{r+3}) > 0 \\ D:&\quad T^{(r+3)}; \quad det(x_{r+1}, x_{r+2}, x_{r+3}) < 0 \end{aligned}$$

Shape orientation



Basic units



Tetrahedral chains: $DDUF \ldots UDFFD$.

Combinatorial codes for U, D, F

Admissible triplets parametrizing U D F:

$$(k,i,j), 1 \leq i,j,k \leq 4, \quad k \neq j \neq i$$

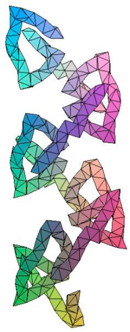
EXAMPLE

$UUDFD$

$$(3,4,2) \to (4,2,1) \to (2,1,4) \to (1,4,1) \to (4,1,3).$$

$$T^{(7)}_{3421413} = R_3 R_1 R_4 R_1 R_2 R_4 R_3 T$$

UUUUUUUUUUFFUUUUUUUUUUFUFUUUUUUUUUUFFUUUUUUUUUUFF



UUUUUUUUUUFFUUUUUUUUUUFUFUUUUUUUUUUFFUUUUUUUUUUFF

Classification of admissible triplets

u	d	f
$det(x_1,x_2,x_3) = 32\sqrt{3}/243$	$det(x_1,x_2,x_3) = -32\sqrt{3}/243$	$det(x_1,x_2,x_3) = 0$
(k,j,i)	(k,j,i)	(k,j,i)
$(3,2,1)$	$(4,2,1)$	$(1,2,1)$
$(4,3,1)$	$(2,3,1)$	$(1,3,1)$
$(2,4,1)$	$(3,4,1)$	$(1,4,1)$
$(4,1,2)$	$(3,1,2)$	$(2,1,2)$
$(1,3,2)$	$(4,3,2)$	$(2,3,2)$
$(3,4,2)$	$(1,4,2)$	$(2,4,2)$
$(2,1,3)$	$(4,1,3)$	$(3,1,3)$
$(4,2,3)$	$(1,2,3)$	$(3,2,3)$
$(1,4,3)$	$(2,4,3)$	$(3,4,3)$
$(3,1,4)$	$(2,1,4)$	$(4,1,4)$
$(1,2,4)$	$(3,2,4)$	$(4,2,4)$
$(2,3,4)$	$(1,3,4)$	$(4,3,4)$

U-chains period
$(3,2,1) \to (2,1,4) \to (1,4,3) \to (4,3,2)$
$(4,3,1) \to (3,1,2) \to (1,2,4) \to (2,4,3)$
$(2,4,1) \to (4,1,3) \to (1,3,2) \to (3,2,4)$
$(3,4,2) \to (4,2,1) \to (2,1,3) \to (1,3,4)$
$(4,1,2) \to (1,2,3) \to (2,3,4) \to (3,4,1)$
$(4,2,3) \to (2,3,1) \to (3,1,4) \to (1,4,2)$
$(1,4,3) \to (4,3,2) \to (3,2,1) \to (2,1,4)$
$(1,2,4) \to (2,4,3) \to (4,3,1) \to (3,1,2)$
$(1,3,2) \to (3,2,4) \to (2,4,1) \to (4,1,3)$
$(2,1,3) \to (1,3,4) \to (3,4,2) \to (4,2,1)$
$(2,3,4) \to (3,4,1) \to (4,1,2) \to (1,2,3)$
$(3,1,4) \to (1,4,2) \to (4,2,3) \to (2,3,1)$

D-chains period
$(2,1,4) \to (1,4,3) \to (4,3,2) \to (3,2,1)$
$(3,1,2) \to (1,2,4) \to (2,4,3) \to (4,3,1)$
$(4,1,3) \to (1,3,2) \to (3,2,4) \to (2,4,1)$
$(4,2,1) \to (2,1,3) \to (1,3,4) \to (3,4,2)$
$(1,2,3) \to (2,3,4) \to (3,4,1) \to (4,1,2)$
$(2,3,1) \to (3,1,4) \to (1,4,2) \to (4,2,3)$
$(4,3,2) \to (3,2,1) \to (2,1,4) \to (1,4,3)$
$(2,4,3) \to (4,3,1) \to (3,1,2) \to (1,2,4)$
$(3,2,4) \to (2,4,1) \to (4,1,3) \to (1,3,2)$
$(1,3,4) \to (3,4,2) \to (4,2,1) \to (2,1,3)$
$(3,4,1) \to (4,1,2) \to (1,2,3) \to (2,3,4)$
$(1,4,2) \to (4,2,3) \to (2,3,1) \to (3,1,4)$

Combinatorial structure

$$\mathbb{I} = \{(\alpha, \beta) \in \Delta \times \Delta : \alpha \neq \beta\}$$

$$\Delta = \{1, 2, 3, 4\}$$

Uniquely defined mappings

$$L_u, L_d, L_f : \mathbb{I} \to \Delta, \quad \#\mathbb{I} = 12$$

and bijections

$$\mathcal{L}_u, \mathcal{L}_d, \mathcal{L}_f : \mathbb{I} \to \mathbb{I},$$

$\mathcal{L}_*(i_1, i_2) = (i_2, L_*(i_1, i_2))$, $* = u, d, f$.

$\mathcal{L}$ - sequence for tetrahedral chain

Example

$$DUUFD \quad \longrightarrow \quad \mathcal{L}_d\mathcal{L}_f\mathcal{L}_u\mathcal{L}_d\mathcal{L}_d$$

Any periodic tetrahedral chain is characterized by cycling composition of a numerical representation of its period

Compositions of $\mathcal{L}_*-$sequences form the indexing space for tetrahedral chains

The indexing space is a binary tetrahedral subgroup of S_{12}

generated by three elements $\mathcal{L}_u$, $\mathcal{L}_d$, $\mathcal{L}_f$ with the relations

$$\mathcal{L}_u^3 = id, \quad \mathcal{L}_d^3 = id, \quad \mathcal{L}_f^2 = id, \quad (\mathcal{L}_u\mathcal{L}_d)^2 = id.$$

Geometric characteristics

-proper tetrahedral chains

n	3	4	5	6	7	8	9	10	11	12	13
A_n	1	3	9	26	76	218	628	1802	5146	14670	41734

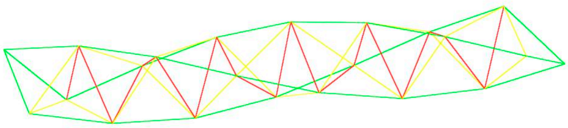
-branching order $0 \leq b \leq 3$

-vertex order $P(p)$, $\Sigma_{p \in V_{C_n}} P(p) = 4n$

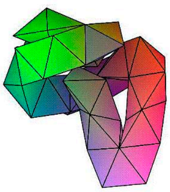
-clustering function

$$Cl(C_n) = \sum_{p \in V_{C_n}} max(0, P(p) - 4)$$

Tetrahelix

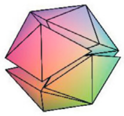
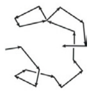


Zero branching order



Proper chains sharing one common vertex

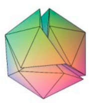
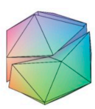
$b\backslash n$	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	0	1	2	6	9	19	38	49	69	79	71	34	6
2	0	0	1	4	6	10	24	46	78	113	137	153	132	85	36	6	0
3	2	4	6	9	16	27	38	48	55	56	50	35	22	12	2	0	0
total	2	4	7	13	22	38	64	100	142	188	225	237	223	176	109	40	6

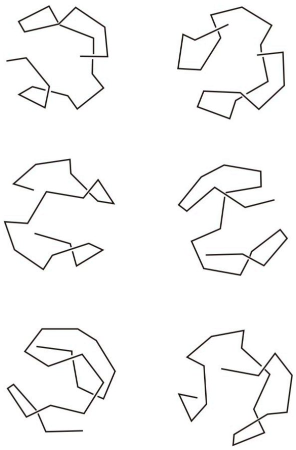



Ico-clusters

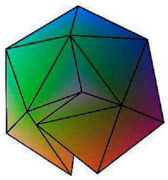
$$FFUFFDUDUDFFUFFDU, FFUFFDUDUDUDFFU$$
$$UFFDFFUDUDUFFDFFU, UFFDUDFFUFFDUDF$$
$$UDFFUFFDUDFFUFFDU, UDFFUFFDUDUDFFU$$

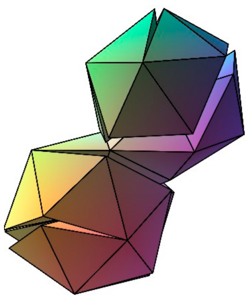


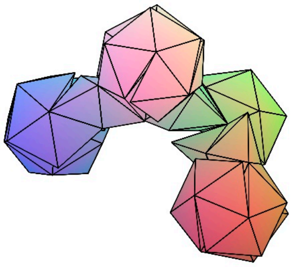



Smallest unit $b = 1$

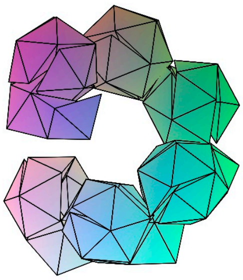




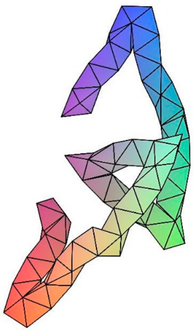
Clustering folding



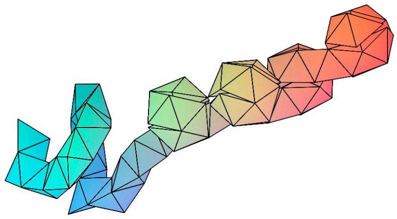
Clustering folding



Big periodic

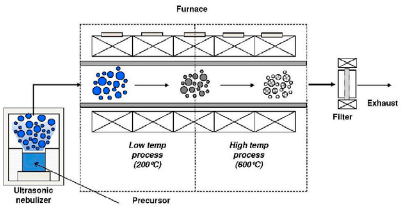


Mixed clustering folding

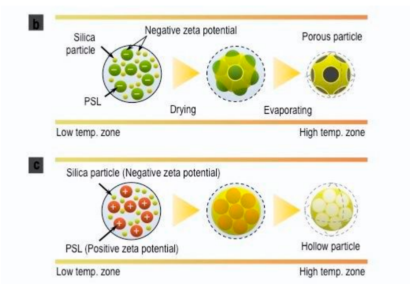


H. Babiker, S. Janeczko, **Combinatorial representation of tetrahedral chains,** *Communications in Information and Sciences,* Vol. 15, No. 3, (2015), 331-359

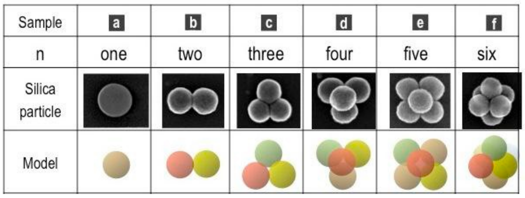
Nano-blood particles



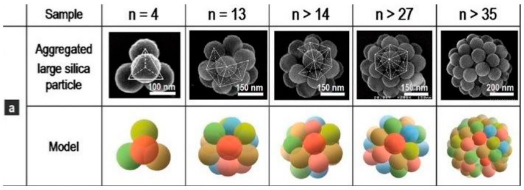
Spray technology



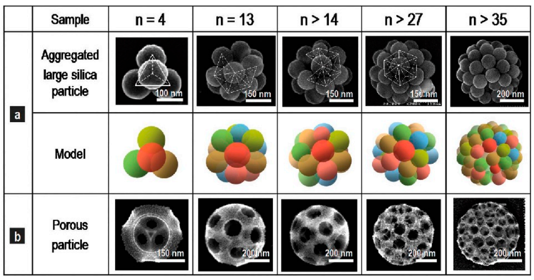
Silica particles



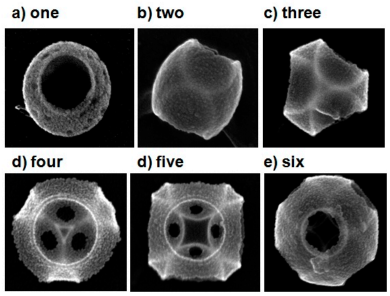
Large silica particles



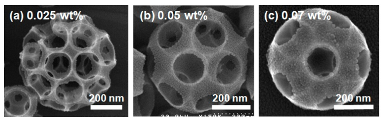
Porous particles



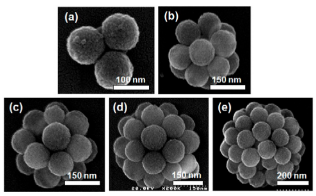
Stable porous particles



Porous particles



Hollow particles



S.Y. Lee, L. Gradon, S. Janeczko, F. Iskandar, K. Okuyama, **Formation of Highly Ordered Nanostructures by drying Micrometer Colloidal Droplets,** *ACS Nano Journal,* Vol. 4, No. 8, (2010), 4717-4724

L. Gradon, S. Janeczko, M. Abdullah, F. Iskandar, K. Okuyama, **Self-Organization Kinetics of Mesoporous Nanostructured Particles,** *AIChE Journal* Vol. 50, No. 10, (2004), 2583-2593.

Geometrical coding of DNA sequences

I

{ FFUFFDUDUDUDFFUFF,{ A,C} } ACACTCTACTACTATACAC
{ FFUFFDUDUDUDFFUFF,{ A,T} } ATATGTGATGATGAGATAT
{ FFUFFDUDUDUDFFUFF,{ A,G} } AGAGCGCAGCAGCACAGAG
{ FFUFFDUDUDUDFFUFF,{ C,A} } CACAGAGCAGCAGCGCACA
{ FFUFFDUDUDUDFFUFF,{ C,T} } CTCTATACTACTACACTCT
{ FFUFFDUDUDUDFFUFF,{ C,G} } CGCGTGTCGTCGTCTCGCG
{ FFUFFDUDUDUDFFUFF,{ T,A} } TATACACTACTACTCTATA
{ FFUFFDUDUDUDFFUFF,{ T,C} } TCTCGCGTCGTCGTGTCTC
{ FFUFFDUDUDUDFFUFF,{ T,G} } TGTGAGATGATGATATGTG
{ FFUFFDUDUDUDFFUFF,{ G,A} } GAGATATGATGATGTGAGA
{ FFUFFDUDUDUDFFUFF,{ G,C} } GCGCACAGCAGCAGAGCGC
{ FFUFFDUDUDUDFFUFF,{ G,T} } GTGTCTCGTCGTCGCGTGT

II

{ FFUFFDUDUDFFUFFDU,{ A,C} } ACACTCTACTACACTCTAC
{ FFUFFDUDUDFFUFFDU,{ A,T} } ATATGTGATGATATGTGAT
{ FFUFFDUDUDFFUFFDU,{ A,G} } AGAGCGCAGCAGAGCGCAG
{ FFUFFDUDUDFFUFFDU,{ C,A} } CACAGAGCAGCACAGAGCA
{ FFUFFDUDUDFFUFFDU,{ C,T} } CTCTATACTACTCTATACT
{ FFUFFDUDUDFFUFFDU,{ C,G} } CGCGTGTCGTCGCGTGTCG
{ FFUFFDUDUDFFUFFDU,{ T,A} } TATACACTACTATACACTA
{ FFUFFDUDUDFFUFFDU,{ T,C} } TCTCGCGTCGTCTCGCGTC
{ FFUFFDUDUDFFUFFDU,{ T,G} } TGTGAGATGATGTGAGATG
{ FFUFFDUDUDFFUFFDU,{ G,A} } GAGATATGATGAGATATGA
{ FFUFFDUDUDFFUFFDU,{ G,C} } GCGCACAGCAGCGCACAGC
{ FFUFFDUDUDFFUFFDU,{ G,T} } GTGTCTCGTCGTGTCTCGT

III

{ UDFFUFFDUDUDFFUFF,{ A,C} } ACTATACACTACTATACAC
{ UDFFUFFDUDUDFFUFF,{ A,T} } ATGAGATATGATGAGATAT
{ UDFFUFFDUDUDFFUFF,{ A,G} } AGCACAGAGCAGCACAGAG
{ UDFFUFFDUDUDFFUFF,{ C,A} } CAGCGCACAGCAGCGCACA
{ UDFFUFFDUDUDFFUFF,{ C,T} } CTACACTCTACTACACTCT
{ UDFFUFFDUDUDFFUFF,{ C,G} } CGTCTCGCGTCGTCTCGCG
{ UDFFUFFDUDUDFFUFF,{ T,A} } TACTCTATACTACTCTATA
{ UDFFUFFDUDUDFFUFF,{ T,C} } TCGTGTCTCGTCGTGTCTC
{ UDFFUFFDUDUDFFUFF,{ T,G} } TGATATGTGATGATATGTG
{ UDFFUFFDUDUDFFUFF,{ G,A} } GATGTGAGATGATGTGAGA
{ UDFFUFFDUDUDFFUFF,{ G,C} } GCAGAGCGCAGCAGAGCGC
{ UDFFUFFDUDUDFFUFF,{ G,T} } GTCGCGTGTCGTCGCGTGT

1

Geometrical coding of DNA sequences

IV

{ UDFFUFFDUDFFUFFDU,{ A,C} } ACTATACACTACACTCTAC
{ UDFFUFFDUDFFUFFDU,{ A,T} } ATGAGATATGATATGTGAT
{ UDFFUFFDUDFFUFFDU,{ A,G} } AGCACAGAGCAGAGCGCAG
{ UDFFUFFDUDFFUFFDU,{ C,A} } CAGCGCACAGCACAGAGCA
{ UDFFUFFDUDFFUFFDU,{ C,T} } CTACACTCTACTCTATACT
{ UDFFUFFDUDFFUFFDU,{ C,G} } CGTCTCGCGTCGCGTGTCG
{ UDFFUFFDUDFFUFFDU,{ T,A} } TACTCTATACTATACACTA
{ UDFFUFFDUDFFUFFDU,{ T,C} } TCGTGTCTCGTCTCGCGTC
{ UDFFUFFDUDFFUFFDU,{ T,G} } TGATATGTGATGTGAGATG
{ UDFFUFFDUDFFUFFDU,{ G,A} } GATGTGAGATGAGATATGA
{ UDFFUFFDUDFFUFFDU,{ G,C} } GCAGAGCGCAGCGCACAGC
{ UDFFUFFDUDFFUFFDU,{ G,T} } GTCGCGTGTCGTGTCTCGT

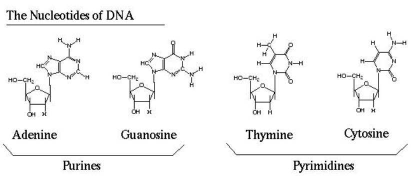
V

{ UFFDUDFFUFFDUDFFU,{ A,C} } ACTCTACTCTATACTATAC
{ UFFDUDFFUFFDUDFFU,{ A,T} } ATGTGATGTGAGATGAGAT
{ UFFDUDFFUFFDUDFFU,{ A,G} } AGCGCAGCGCACAGCACAG
{ UFFDUDFFUFFDUDFFU,{ C,A} } CAGAGCAGAGCGCAGCGCA
{ UFFDUDFFUFFDUDFFU,{ C,T} } CTATACTATACACTACACT
{ UFFDUDFFUFFDUDFFU,{ C,G} } CGTGTCGTGTCTCGTCTCG
{ UFFDUDFFUFFDUDFFU,{ T,A} } TACACTACACTCTACTCTA
{ UFFDUDFFUFFDUDFFU,{ T,C} } TCGCGTCGCGTGTCGTGTC
{ UFFDUDFFUFFDUDFFU,{ T,G} } TGAGATGAGATATGATATG
{ UFFDUDFFUFFDUDFFU,{ G,A} } GATATGATATGTGATGTGA
{ UFFDUDFFUFFDUDFFU,{ G,C} } GCACAGCACAGAGCAGAGC
{ UFFDUDFFUFFDUDFFU,{ G,T} } GTCTCGTCTCGCGTCGCGT

VI

{ UFFDFFUDUDUFFDFFU,{ A,C} } ACTCTATACTACTCTATAC
{ UFFDFFUDUDUFFDFFU,{ A,T} } ATGTGAGATGATGTGAGAT
{ UFFDFFUDUDUFFDFFU,{ A,G} } AGCGCACAGCAGCGCACAG
{ UFFDFFUDUDUFFDFFU,{ C,A} } CAGAGCGCAGCAGAGCGCA
{ UFFDFFUDUDUFFDFFU,{ C,T} } CTATACACTACTATACACT
{ UFFDFFUDUDUFFDFFU,{ C,G} } CGTGTCTCGTCGTGTCTCG
{ UFFDFFUDUDUFFDFFU,{ T,A} } TACACTCTACTACACTCTA
{ UFFDFFUDUDUFFDFFU,{ T,C} } TCGCGTGTCGTCGCGTGTC
{ UFFDFFUDUDUFFDFFU,{ T,G} } TGAGATATGATGAGATATG
{ UFFDFFUDUDUFFDFFU,{ G,A} } GATATGTGATGATATGTGA
{ UFFDFFUDUDUFFDFFU,{ G,C} } GCACAGAGCAGCACAGAGC
{ UFFDFFUDUDUFFDFFU,{ G,T} } GTCTCGCGTCGTCTCGCGT

1



A = Adenine, T = Thymine, C = Cytosine, G = Guanine

How to measure data diversity and why it is important?

Paweł Józiak

Faculty of Mathematics and Computer Science,
Warsaw University of Technology, Poland

In Machine Learning we often hear about patterns that algorithms overfit to. To prevent it, a *high quality data*, a bunch of data that is *curated* needs to be prepared. I will discuss what tools are available, other than manual labor, in order to tell whether the dataset is *diverse*, and how we used the knowledge gained through it in order to prepare a highly diverse (and thus highly challenging) *Document Understanding Dataset and Evaluation (DUDE)* in the domain of DocumentAI, a field at the boundary of Natural Language Processing and Computer Vision. Joint work with Jordy Van Landeghem, Rubén Tito, Łukasz Borchmann, Michał Pietruszka, Rafał Powalski, Dawid Jurkiewicz, Mickaël Coustaty, Bertrand Ackaert, Ernest Valveny, Matthew Blaschko, Sien Moens, Tomasz Stanisławek.

References

[1] Jordy Van Landeghem, Rubén Tito, Łukasz Borchmann, Michał Pietruszka, Rafał Powalski, Dawid Jurkiewicz, Mickaël Coustaty, Bertrand Ackaert, Ernest Valveny, Matthew Blaschko, Sien Moens, Tomasz Stanisławek. *Document Understanding Dataset and Evaluation (DUDE).* Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 19528-19540

How to measure data diversity and why it is important?

Paweł Józiak

Faculty of Mathematics and Information Science
WARSAW UNIVERSITY OF TECHNOLOGY

&

snowflake

26 IX 2023
Warsaw University of Technology
Mathematics in Industry

Plan of the talk

1. Data from ML practitioner's perspective
2. Current approaches
3. Our approach
4. Conclusions

Plan of the talk

1. Data from ML practitioner's perspective
2. Current approaches
3. Our approach
4. Conclusions

Data in Computer Vision

Images have natural representation as quaternionic matrices (CMYK) of *width* × *height* size.

- Pros: naturality, robust representation.
- Cons: uneven & possibly hight dimensionality.

Scale images to standarized sizes, apply RoI-pooling.

- uniform dimensionality,
- dimensionality reduction, but makes method prone to small perturbations.

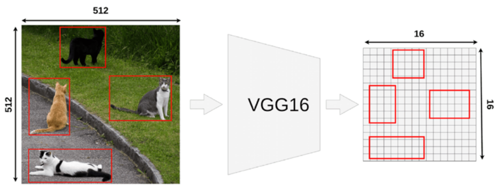
Data in Computer Vision

Images have natural representation as quaternionic matrices (CMYK) of *width* × *height* size.

- Pros: naturality, robust representation.
- Cons: uneven & possibly hight dimensionality.

Scale images to standarized sizes, apply RoI-pooling.

- uniform dimensionality,
- dimensionality reduction, but makes method prone to small perturbations.



Data in Natural Language Processing

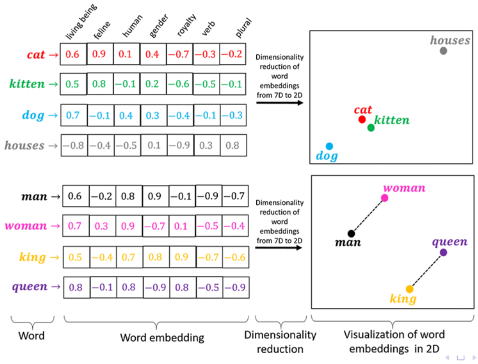
One-dimensional chain of tokens (words).

- **T**erm **f**requency **I**nverse **d**ocument **f**requency.
- Contextual embeddings (Word2Vec, Glove)

Data in Natural Language Processing

One-dimensional chain of tokens (words).

- **T**erm **f**requency **I**nverse **d**ocument **f**requency.
- Contextual embeddings (Word2Vec, Glove)

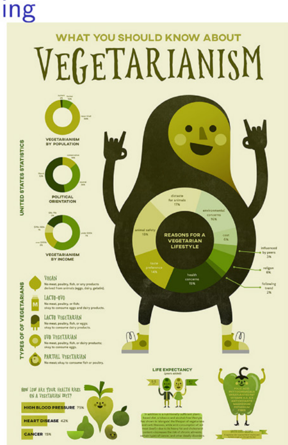


Data in Document Understanding

Data in Document Understanding

Data in Document Understanding

Data in Document Understanding



Plan of the talk

1. Data from ML practitioner's perspective
2. Current approaches
3. Our approach
4. Conclusions

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \ldots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \ldots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \ldots, n_d) = P(N)\frac{N!}{n_1!n_2!\ldots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\ldots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \ldots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \ldots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \ldots, n_d) = P(N)\frac{N!}{n_1!n_2!\ldots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\ldots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 8 / 18

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \ldots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \ldots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \ldots, n_d) = P(N)\frac{N!}{n_1!n_2!\ldots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\ldots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 8 / 18

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \dots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \dots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \dots, n_d) = P(N)\frac{N!}{n_1!n_2!\dots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\dots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \dots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \dots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \dots, n_d) = P(N)\frac{N!}{n_1!n_2!\dots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\dots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 8 / 18

Simpson diversity measure

- Let $\pi\colon 2^{[d]} \to [0,1]$ be a probability distribution on $[d] = \{1, 2, \dots, d\}$.
- Denote $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$ the Fisher concetration.
- Let $x_1, \dots, x_N$ is a sample drawn from π and let $n_i = |\{j \in [N] \mid x_j = i\}|$.
- $\hat{\lambda} = \dfrac{\sum_{i=1}^{d} n_i(n_i - 1)}{N(N-1)}$ is an unbiased estimator of λ.
- If now N is random, the above still holds under factorization assumption

$$P(n_1, n_2, \dots, n_d) = P(N)\frac{N!}{n_1!n_2!\dots n_d!}\pi(1)^{n_1}\pi(2)^{n_2}\dots\pi(d)^{n_d}$$

- $Var(\hat{\lambda}) = \dfrac{4}{N}(\sum_{i=1}^{d} \pi(i)^3 - \lambda^2) + O(N^{-2})$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 8 / 18

PAMI places 1

- Consider $(X, \Sigma, \mathbb{P})$ and an embedding function $f\colon X \to \mathbb{R}^d$; by abuse of notation: $\mathbb{P} = f_* \circ \mathbb{P}$.
- The sizes $n_j = |\{i : x_i = j\}|$ no longer makes sense $\to$ we can still ask if x and x' look similar (wrt to some similarity function d).
- E.g. we can ask if x and x' look similar, or at least: whether (x, x') look more similar than (y, y').

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity function $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

PAMI places 1

- Consider $(X, \Sigma, \mathbb{P})$ and an embedding function $f\colon X \to \mathbb{R}^d$; by abuse of notation: $\mathbb{P} = f_* \circ \mathbb{P}$.
- The sizes $n_j = |\{i : x_i = j\}|$ no longer makes sense $\to$ we can still ask if x and x' look similar (wrt to some similarity function d).
- E.g. we can ask if x and x' look similar, or at least: whether (x, x') look more similar than (y, y').

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity function $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

PAMI places 1

- Consider $(X, \Sigma, \mathbb{P})$ and an embedding function $f\colon X \to \mathbb{R}^d$; by abuse of notation: $\mathbb{P} = f_* \circ \mathbb{P}$.
- The sizes $n_j = |\{i : x_i = j\}|$ no longer makes sense $\to$ we can still ask if x and x' look similar (wrt to some similarity function d).
- E.g. we can ask if x and x' look similar, or at least: whether (x, x') look more similar than (y, y').

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity function $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

PAMI places 1

- Consider $(X, \Sigma, \mathbb{P})$ and an embedding function $f\colon X \to \mathbb{R}^d$; by abuse of notation: $\mathbb{P} = f_* \circ \mathbb{P}$.
- The sizes $n_j = |\{i : x_i = j\}|$ no longer makes sense $\to$ we can still ask if x and x' look similar (wrt to some similarity function d).
- E.g. we can ask if x and x' look similar, or at least: whether (x, x') look more similar than (y, y').

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity function $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

PAMI places 1

- Consider $(X, \Sigma, \mathbb{P})$ and an embedding function $f\colon X \to \mathbb{R}^d$; by abuse of notation: $\mathbb{P} = f_* \circ \mathbb{P}$.
- The sizes $n_j = |\{i : x_i = j\}|$ no longer makes sense $\to$ we can still ask if x and x' look similar (wrt to some similarity function d).
- E.g. we can ask if x and x' look similar, or at least: whether (x, x') look more similar than (y, y').

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity function $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

PAMI places 2

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x, x') < d(y, y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0, \infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Fisher concentration

For $\pi\colon 2^{[d]} \to [0, 1]$ probability distribution define $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$

Let $X = Y = [n]$ and $p_Y = \delta_1$. Then

$$D_X(Y) = 1 - P(d(y, y') < d(x, x'))$$
$$= 1 - P(x \neq x') = P(x = x') = \sum_{x=1}^{n} p_X(x)^2 = \lambda_X.$$

PAMI places 2

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x,x') < d(y,y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0,\infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Fisher concentration

For $\pi\colon 2^{[d]} \to [0,1]$ probability distribution define $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$

Let $X = Y = [n]$ and $p_Y = \delta_1$. Then

$$D_X(Y) = 1 - P(d(y,y') < d(x,x'))$$

$$= 1 - P(x \neq x') = P(x = x') = \sum_{x=1}^{n} p_X(x)^2 = \lambda_X.$$

PAMI places 2

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x,x') < d(y,y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0,\infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Fisher concentration

For $\pi\colon 2^{[d]} \to [0,1]$ probability distribution define $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$

Let $X = Y = [n]$ and $p_Y = \delta_1$. Then

$$D_X(Y) = 1 - P(d(y,y') < d(x,x'))$$

$$= 1 - P(x \neq x') = P(x = x') = \sum_{x=1}^{n} p_X(x)^2 = \lambda_X.$$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 10 / 18

PAMI places 2

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x,x') < d(y,y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0,\infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Fisher concentration

For $\pi\colon 2^{[d]} \to [0,1]$ probability distribution define $\lambda = \sum_{i=1}^{d} \pi(i)^2 \in [\frac{1}{d}, 1]$

Let $X = Y = [n]$ and $p_Y = \delta_1$. Then

$$D_X(Y) = 1 - P(d(y,y') < d(x,x'))$$

$$= 1 - P(x \neq x') = P(x = x') = \sum_{x=1}^{n} p_X(x)^2 = \lambda_X.$$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 10 / 18

PAMI places 3

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x,x') < d(y,y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0,\infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Relative diversity

Diversity of set X_1 relative to a family of sets $X_2, \ldots, X_n$:

$$\mathrm{Div}_{X_2,\ldots,X_n}(X_1) = 1 - \mathbb{P}(d(x_1,x_1') < \min_{2 \le i \le n} d(x_i,x_i'))$$

for a similarity fn $d\colon (\bigcup_{i=1}^n X_i)^2 \to (0,\infty)$, with $x_i, x_i' \in X_i$.

PAMI places 3

Relative diversity

Diversity of set X relative to set Y:

$$\mathrm{Div}_Y(X) = 1 - \mathbb{P}(d(x,x') < d(y,y'))$$

for a similarity fn $d\colon (X \cup Y)^2 \to (0,\infty)$, with $x, x' \in X$ and $y, y' \in Y$.

Relative diversity

Diversity of set X_1 relative to a family of sets $X_2, \ldots, X_n$:

$$\mathrm{Div}_{X_2,\ldots,X_n}(X_1) = 1 - \mathbb{P}(d(x_1,x_1') < \min_{2 \le i \le n} d(x_i,x_i'))$$

for a similarity fn $d\colon (\bigcup_{i=1}^n X_i)^2 \to (0,\infty)$, with $x_i, x_i' \in X_i$.

Plan of the talk

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\le i\le n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \cdots \sum_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1')) \cdots \mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \prod_{j=2}^{n} \sum_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\le i\le n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \cdots \sum_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1')) \cdots \mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \prod_{j=2}^{n} \sum_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\le i\le n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \cdots \sum_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1')) \cdots \mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum_{x_1,x_1'\in X_1} \prod_{j=2}^{n} \sum_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2} \cdots \binom{|X_n|}{2}}$$

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\leq i\leq n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\cdots\sum\limits_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1'))\dots\mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\prod\limits_{j=2}^{n}\sum\limits_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\leq i\leq n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\cdots\sum\limits_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1'))\dots\mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\prod\limits_{j=2}^{n}\sum\limits_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 13 / 18

Embeddings

- Use a learnable representation $f : X \to \mathbb{R}^d$.
- calculate the cosine similarity: $d(x, x') = 1 - \arccos(f(x), f(x'))$.
- Use uniform distribution on X, calculation easier than with push-forward to $f(X)$ as a measure on (discrete subset of) $\mathbb{R}^d$.
- Calculate

$$Div_{X_2,\dots,X_n}(X_1) = 1 - p(d(x_1, x_1') < \min_{2\leq i\leq n} d(x_i, x_i'))$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\cdots\sum\limits_{x_n,x_n'\in X_n} \mathbb{1}_{(0,d(x_2,x_2'))}(d(x_1,x_1'))\dots\mathbb{1}_{(0,d(x_n,x_n'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

$$= 1 - \frac{\sum\limits_{x_1,x_1'\in X_1}\prod\limits_{j=2}^{n}\sum\limits_{x_j,x_j'\in X_j} \mathbb{1}_{(0,d(x_j,x_j'))}(d(x_1,x_1'))}{\binom{|X_1|}{2}\dots\binom{|X_n|}{2}}$$

Paweł Józiak (MiNI PW, Snowflake) On data diversity measures. Warsaw, 26 IX 2023 13 / 18

Results

Natural candidates for these representations

- Visual: ResNet, VGG etc Neural Networks
- Textual: TfIdf, word2vec etc statistical vectorization techniques

	ResNet	TfIdf
DUDE	0.82	0.95
DocVQA	0.76	0.93
VisualMRC	0.83	0.99
InfographicsVQA	0.86	0.94
TAT-DQA	0.73	0.15

Plan of the talk

The dataset is challenging.

Model type	ANSL test score	ANLS diagnostic score
Big Bird	26.27	30.67
BERT-large	25.48	32.18
Longformer	27.14	33.45
T5-base	19.65-41.8	25.62-44.95
ChatGPT	-	35.07-41.89
GPT3	-	43.95-47.04
T5-2D-base	37.1-42.1	40.5-45.73
T5-2D-large	46.06	48.14
HiVT5	23.06	22.33
LayoutLMv3	20.31	25.27
Human	-	74.76

Dataset	Human score	best models
DocVQA	98.11	87.05-90.16;
TAT-DQA	84.1	70.3-76.8;
InfographicVQA	97.18	52.58-61.2
VisualMRC	-	56-57.2

The dataset is challenging.

Model type	ANSL test score	ANLS diagnostic score
Big Bird	26.27	30.67
BERT-large	25.48	32.18
Longformer	27.14	33.45
T5-base	19.65-41.8	25.62-44.95
ChatGPT	-	35.07-41.89
GPT3	-	43.95-47.04
T5-2D-base	37.1-42.1	40.5-45.73
T5-2D-large	46.06	48.14
HiVT5	23.06	22.33
LayoutLMv3	20.31	25.27
Human	-	74.76

Dataset	Human score	best models
DocVQA	98.11	87.05-90.16;
TAT-DQA	84.1	70.3-76.8;
InfographicVQA	97.18	52.58-61.2
VisualMRC	-	56-57.2

References

- E. H. Simpson, *Measurement of Diversity*, Nature **163**, `https://doi.org/10.1038%2F163688a0`.
- B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, A. Torralba, *Places: A 10 Million Image Database for Scene Recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence **40**, `https://doi.org/10.1109/tpami.2017.2723009`.
- J. Van Landeghem, R. Tito, Ł. Borchmann, M. Pietruszka, P. Józiak, R. Powalski, D. Jurkiewicz, M. Coustaty, B. Ackaert, E. Valveny, M. Blaschko, S. Moens, T. Stanisławek, *Document Understanding Dataset and Evaluation (DUDE)*, to appear in: International Conference on Computer Vision 2023, `https://arxiv.org/abs/2305.08455`.

Thank you

Cryptographic protocol verification - results of EPW project

Konstanty Junosza-Szaniawski

Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland

Cryptographic protocols are fundamental to cybersecurity, necessitating assurance that these protocols are devoid of flaws. Among the various tools available for the verification of cryptographic protocols, ProVerif stands out. ProVerif models protocols using Horn formulas and verifies the security properties through the satisfiability of corresponding logical formulas. However, the complexity of modeling protocols and their properties in ProVerif is time-consuming and requires a high level of knowledge. To address this, we have developed a translator from the AnB language, which describes protocols from a global perspective, to ProVerif syntax. This translator simplifies the modeling process, enabling easy verification of key security properties with ProVerif, such as secrecy, forward secrecy, weak secrecy, indistinguishability, authentication, non-replay authentication, and key compromise impersonation. Our translator is a principal outcome of the project "Experimental Platform for Automatic Validation of Crypto Algorithms and Verification of Crypto Protocols" (EPW), funded by The National Centre for Research and Development under the grant CYBERSECIDENT/456962/III/NCBR/2020.

epw

Cryptographic protocol verification results of EPW project

Elżbieta Andrukiewicz, Daniel Waszkiewicz
National Institute of Telecommunications, Poland
Tomasz Brengos, Anna Cichocka, Konstanty Junosza-Szaniawski, Adam Komorowski, Agata Pilitowska, Hubert Grochowski,
Warsaw University of Technology, Poland

Supported by The National Centre for Research and Development
CYBERSECIDENT/456962/III/NCBR/2020

Wydział Matematyki i Nauk Informacyjnych · Narodowe Centrum Badań i Rozwoju · Instytut Łączności · NASK

R+D Project EPW – brief presentation

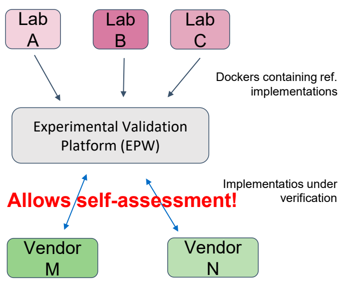
epw

Experimental platform for automatic validation of cryptoalgorithms and verification of cryptoprotocols (acronym: EPW)

- Project financed under the national program sponsored by National Centre of Research and Development „Cybersecurity and e-Identity" (CyberSecIdent)
- The Consortium consists of 3 R&D Polish entities:
  - National Institute of Telecommunications – State Research Institute (Consortium Leader)
  - NASK – State Research Institute
  - Warsaw University of Technology, Mathematics and Information Science Faculty
- timeframe: July 2020 – December 2023

EPW original concept

epw



Formal verification of cryptoalgorithms

Equivalence checking (1)

Equivalence checking' (EC) for two implementations means that two implementations return the same output for every input

- **First strategy** is to perform tests for a random set of inputs and check if the outputs are the same
- Main drawback - low assurance of the result.

Equivalence checking (2)

Equivalence checking' (EC) for two implementations means that two implementations return the same output for every input

- **Second strategy** is to use a formal approach.
- In general, computational complexity of formal equivalence checking is the *co-NP* class.
- For two implementations the CNF (**Conjunctive Normal Form) formulas** $F_1(X) = Y_1$ oraz $F_2(X) = Y_2$ shall be created.
- Next, $Y_1 \oplus Y_2 \neq 0$ constraints shall be added.
- If SAT-solver returns *UNSAT* it means two implementations are equivalent.

Equivalence checking

epw

Example:

Task: Encode

$x_1+x_2+x_3+x_4 = 0$ (where + denote addition modulu 2)

as CNR formula.

Solution:

$[(\sim x_1 \wedge \sim x_2 \wedge \sim x_3) \Rightarrow \sim x_4] \wedge [x_1 \wedge \sim x_2 \wedge \sim x_3) \Rightarrow x_4] \wedge \ldots$

$(x_1 \vee x_2 \vee x_3 \vee \sim x_4) \wedge (\sim x_1 \vee x_2 \vee x_3 \vee x_4) \wedge \ldots$

8 clauses with 4 literals

Equivalence checking

epw

Example2:

Task: Encode

$$x_1+x_2+x_3+x_4 = 0$$
$$x_3+x_4+x_5 = 0$$

as CNR formula.

Solution: 8 clauses with 4 literals and 4 clauses with 3 literals

Equivalence checking

epw

Example2:

Task: Encode

$$x_1+x_2+x_3+x_4 = 0$$
$$x_3+x_4+x_5 = 0$$

as CNR formula.

Solution: 8 clauses with 4 literals and 4 clauses with 3 literals

Adding second equality to the first one we obtain

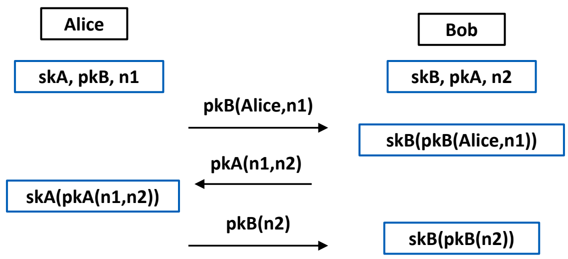
$$x_1+x_2+ \quad x_5 = 0$$
$$x_3+x_4+x_5 = 0$$

What gives us 8 clauses with 3 literals

Formal verification of cryptoprotocols

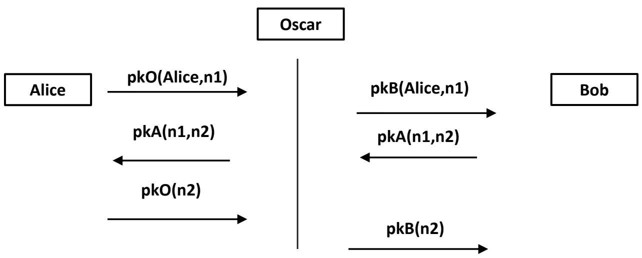
Example of crypto protocol

Needham-Schroeder protocol

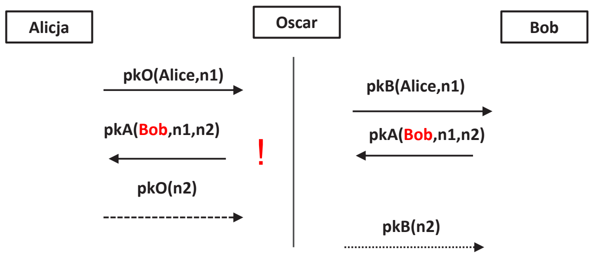


Example of crypto protocol

Needham-Schroeder protocol flaw –
man-in-the middle attack



Modification: Needham-Schroeder-Lowe



ProVerif how does it work

Models the protocol and desired properties as Horn formula:

Let att(x) denote formula saying that an information x is known to the attacker:

- att(x) for every public information x (including information sent on public channel)
- att(f) ^ att(x) => att(f(x)) for any function f and any argument of f.
  e.g. att(pkB)^att(Alice)^att(n1) => att(pkB(Alice,n1))
- If by protocol agent receives message x and responds with resp(x) then:
  e.g. att(x)=> att(resp(x))

If the above formula ^ ~ att(s) is UNSAT then s is a secret to the attacker.

```
(* ProVerif File - Auto Generated - EPW Project *)
free c: channel.
const A: bitstring.
const B: bitstring.
free k1: bitstring [private].
free k2: bitstring [private].
free n1: bitstring [private].

fun enc(bitstring, bitstring): bitstring.
reduc forall messageEnc: bitstring, keyEnc : bitstring;
  dec(enc(messageEnc,keyEnc),keyEnc) = messageEnc.
fun pk(bitstring): bitstring.
fun aenc(bitstring, bitstring): bitstring.
reduc forall messageAenc: bitstring, keyAenc : bitstring;
  adec(aenc(messageAenc,pk(keyAenc)), keyAenc) = messageAenc.
fun exp(bitstring, bitstring): bitstring.
fun concat(bitstring, bitstring): bitstring.
reduc forall x1: bitstring, y1: bitstring;  fst(concat(x1, y1)) = x1.
reduc forall x1: bitstring, y1: bitstring;  snd(concat(x1, y1)) = y1.
fun getSymmetricKey(bitstring): bitstring.
fun getPublicKey(bitstring): bitstring.
event Send(bitstring, bitstring, bitstring).
event Recv(bitstring, bitstring, bitstring).

query secret n1.
let processA(A: bitstring, B: bitstring, n1: bitstring, k1: bitstring) =
 phase 1;
 in(c, var1 : bitstring);
 phase 2;
 out(c, enc(n1, dec(var1, k1)));
 0.
let processB(A: bitstring, B: bitstring, k1: bitstring, k2: bitstring) =
 phase 1;
 out(c, enc(k2, k1));
 phase 2;
 in(c, var2 : bitstring);
 0.
process processA(A, B, n1, k1) | processB(A, B, k1, k2)
 | out(c,k1)
NORMAL  SimpleExample.pv
```

Applied Pi-Calculus in ProVerif

Sophisticated tools

Difficult to understand for non-specialists

User-friendly automated formal verification platform

epw

- Proverif, Tamarin et. al. are complicated tools, necessary extensive code written by expert
- Ongoing work on ISO/IEC 29128 (multipart) will require application of formal methods in crypto protocol verification

Our aim: to simplify the process of protocol writing with formal proof of correctness

Alice and Bob language

epw

```
Types:
    Agent A, B;
    SymmetricKey k1, k2;
  Number n1;

Knowledge:
    A: A, B, n1, k1;
    B: A, B, k1, k2;

Actions:
    B -> A : {| k2 |} k1
    A -> B : {| n1 |} k2

Goals:
    forwardSecret n1 with compromised k1

NORMAL    SimpleExample.anb
```

Easy to read and write

Building a formal bridge

epw

Alice and Bob Language

Applied Pi-calculus In ProVerif

Practical goals

Translator from AnB language with cryptographic primitives to Proverif's Pi-calculus

Use Case: EAP-AKA

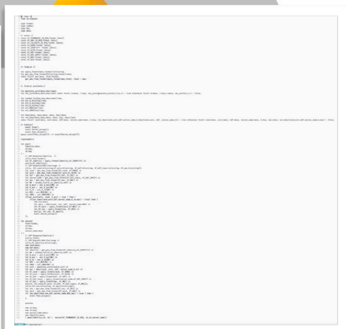
```
Types:
    Agent Peer, Server;
    PublicKey keyPEER, keySERVER;
    Number NAI, RAND, AUTN, KDF, KDFINPUT, MAC, MAC2, RES;
    Const Success;

Knowledge:
    Server: Server, Peer, NAI, RAND, AUTN, KDF, KDFINPUT, MAC, keySERVER;
    Peer: Server, Peer, NAI, MAC2, keyPEER, RES;

Actions:
    Peer -> Server : pk(keyPEER)
    Server -> Peer : pk(keySERVER)
    Peer -> Server : { NAI } pk(keySERVER)
    Server -> Peer : { RAND, AUTN, KDF, KDFINPUT, MAC } pk(keyPEER)
    Peer -> Server : { RES, MAC2 }  pk(keySERVER)
    Server -> Peer : Success

Goals:
    NAI secret of Server, Peer
```

NORMAL EAP.anb



Number of lines of code comparison

Protocol\Language	Alice and Bob	Proverif
Simple Example	15	40
Needham-Schoeder	18	54
Otway-Rees	21	64

Practical goals

epw

Translator from AnB language with cryptographic primitives to Proverif's Pi-calculus

Formally verified translator

epw

F. Montesi "Choreographic Programming"
F. Montesi "Introduction to Choreographies"

$$C ::= \mathsf{p} \to \mathsf{q}; C \mid \mathbf{0}$$

Figure 1: Simple choreographies, syntax.

$$\frac{}{\mathsf{p} \to \mathsf{q}; C \to C}\ \textsf{Com}$$

Figure 2: Simple choreographies, semantics.

Properties verified

epw

Formally verified* AnB to ProVerif translator with automatic checks:

- Secrecy
- Forward secrecy
- Weak secrecy
- Indistinguishability
- Authentication
- No-replay Authentication
- Key Compromise Impersonation

Automatic formal checks handled by dockerized Proverif, available on our Platform

*Work in progress

Challenges*:

- Formally verifiable translations to other protocol languages
- Extended list of security properties

* currently under consideration

Thank you

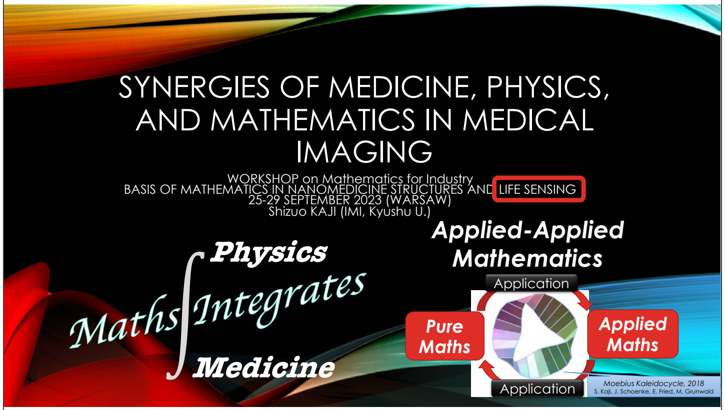
Synergies of medicine, physics, and mathematics in medical imaging

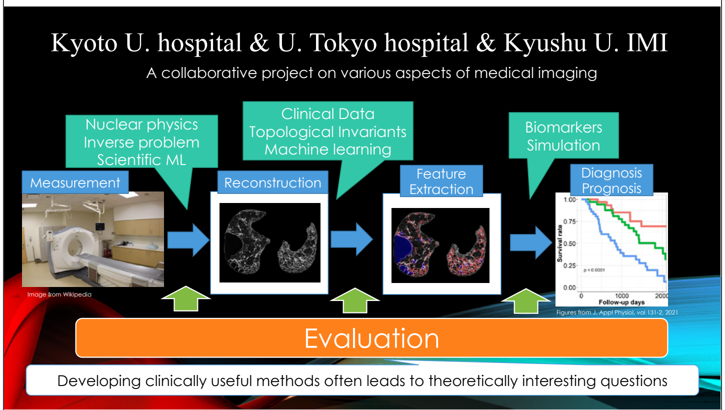
Shizuo Kaji
Institute of Mathematics for Industry, Kyushu University, Japan

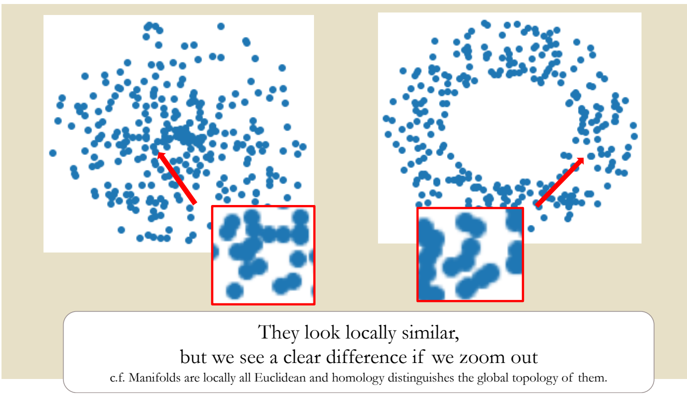
Medical imaging provides detailed visual representations of internal structures and functions of the human body and plays a pivotal role in diagnosing, monitoring, and treating various medical conditions. Mathematical disciplines intersect with medical imaging in multifaceted ways, encompassing:

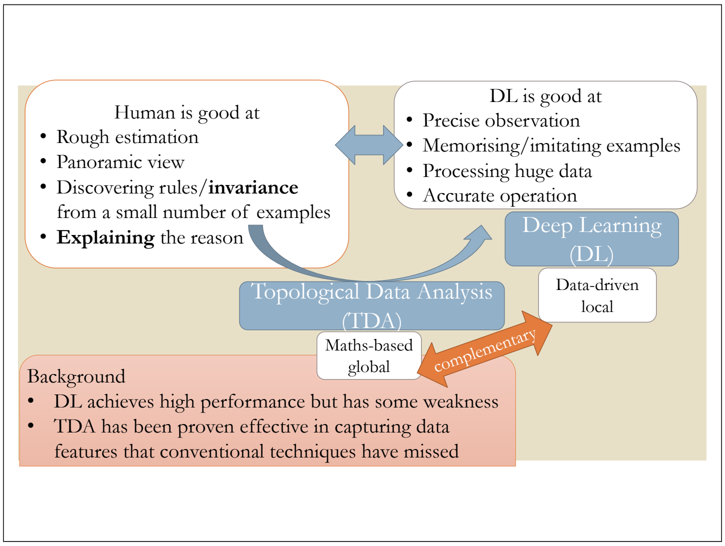
- **Image reconstruction** involves the transformation of raw measurements across diverse modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound into coherent, human-interpretable images.
- **Image enhancement and information Extraction** aim at refining image quality while extracting vital information embedded within.
- **Quantitative analysis** unveils deeper insight into the heterogeneity and progression of diseases in an objective and reproducible manner.

We will present some of our collaborative endeavours, bridging the expertise of medical doctors, medical physicists, and the realm of mathematics. Our work showcases applications of machine learning and topology that fortify and enrich the field of medical imaging.


SYNERGIES OF MEDICINE, PHYSICS, AND MATHEMATICS IN MEDICAL IMAGING
WORKSHOP on Mathematics for Industry
BASIS OF MATHEMATICS IN NANOMEDICINE STRUCTURES AND LIFE SENSING
25-29 SEPTEMBER 2023 (WARSAW)
Shizuo KAJI (IMI, Kyushu U.)
Physics
Maths
Integrates
Medicine
Applied-Applied Mathematics
Application
Pure Maths
Applied Maths
Application
Moebius Kaleidocycle, 2018


Kyoto U. hospital & U. Tokyo hospital & Kyushu U. IMI
A collaborative project on various aspects of medical imaging
Nuclear physics
Inverse problem
Scientific ML
Clinical Data
Topological Invariants
Machine learning
Biomarkers
Simulation
Measurement
Reconstruction
Feature Extraction
Diagnosis
Prognosis
Evaluation
Developing clinically useful methods often leads to theoretically interesting questions


They look locally similar,
but we see a clear difference if we zoom out
c.f. Manifolds are locally all Euclidean and homology distinguishes the global topology of them.


Human is good at
Rough estimation
Panoramic view
Discovering rules/invariance from a small number of examples
Explaining the reason
DL is good at
Precise observation
Memorising/imitating examples
Processing huge data
Accurate operation
Deep Learning (DL)
Data-driven local
Topological Data Analysis (TDA)
Maths-based global
complementary
Background
DL achieves high performance but has some weakness
TDA has been proven effective in capturing data features that conventional techniques have missed

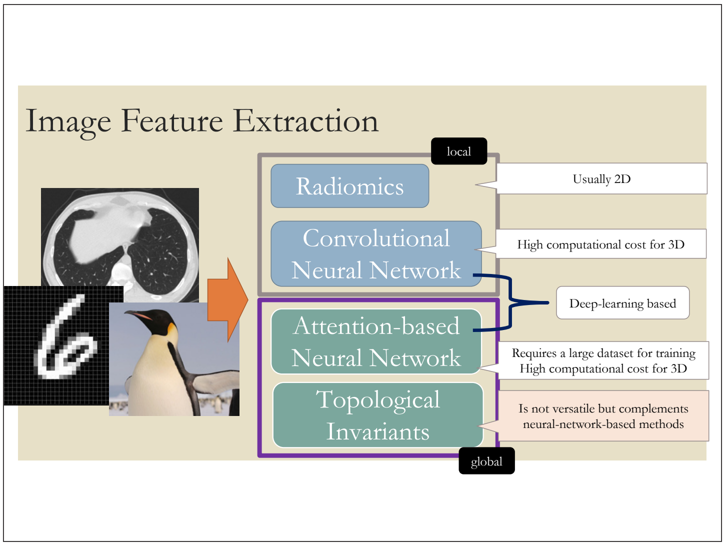
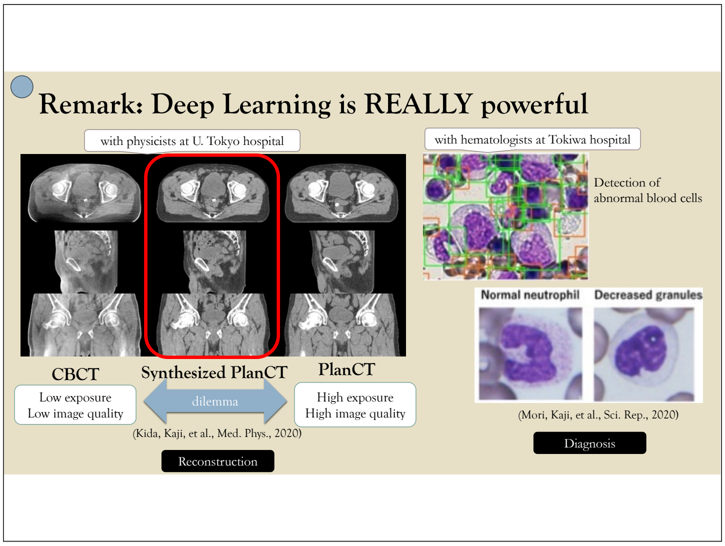


Image Feature Extraction
local
Radiomics
Usually 2D
Convolutional Neural Network
High computational cost for 3D
Deep-learning based
Attention-based Neural Network
Requires a large dataset for training
High computational cost for 3D
Topological Invariants
Is not versatile but complements neural-network-based methods
global

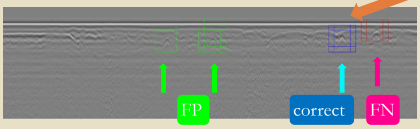

Remark: Deep Learning is REALLY powerful
with physicists at U. Tokyo hospital
with hematologists at Tokiwa hospital
Detection of abnormal blood cells
Normal neutrophil
Decreased granules
CBCT
Synthesized PlanCT
PlanCT
Low exposure
Low image quality
dilemma
High exposure
High image quality
(Kida, Kaji, et al., Med. Phys., 2020)
(Mori, Kaji, et al., Sci. Rep., 2020)
Reconstruction
Diagnosis

Remark: **DL and TDA are NOT competitors but collaborators**

Example: Sinkhole detection in ground penetrating radar image
(with S. Choi, T. Kim, GK Engineering)

Difficulty

1. No big data (data acquisition is costly)
2. False Negative (overlook) is critical



By combining CNN and TDA, we achieved comparable performance with human experts with only **40 labelled volumes**

Why it worked well: Topology helped to distinguish sinkholes from pipes, which were reflected in H_1

Topological Features of Image

Pipeline

2D/3D image

Filtered complex

Sequence of topological objects

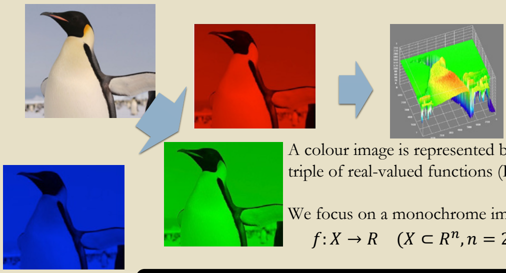
Topological invariants

Features

Results+Interpretation

Domain knowledge + Machine learning

Image = Function on a metric space



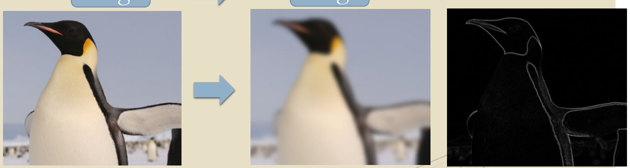
A colour image is represented by a triple of real-valued functions (R,G,B)

We focus on a monochrome image
$f: X \to R \quad (X \subset R^n, n = 2,3)$

An image processing/analysis method is an operation on the space of functions

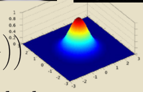
Image processing = Operation on functions

image → image



Example: Convolution

Example of a kernel:
$h(x,y) = c\exp\left(-\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)\right)$



$$h * f(x,y) = \int\int h(x-s, y-t) f(s,t)\, ds\, dt$$

Topological Image Analysis
Function → Space → "Numbers"

topological space X
function $f: X \to R$

Each threshold value gives rise to the **sub-level set** {x | f(x)<a}

Function ➔ Space ➔ "Numbers"

For each threshold a, we have a space
X(f,a) := {x | f(x)<a}

We can compute topological invariants of X(f,a) to obtain image features

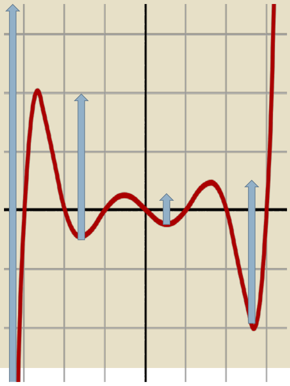
Q: How to choose a threshold?

A: We do not choose. Use them all!

Persistent homology (PH)

- Extension of homology defined for functions over topological spaces
- For each topological feature(cycle), the threshold values with which it was born and destroyed are recorded

Remark:
We can also view PH as a **"continuous relaxation" of homology**.
Homology is a discrete quantity that is sometimes problematic.
(e.g., homology can change abruptly with small variation in the input)



Persistent homology (formal definition)

Increasing sequence of spaces $\emptyset \subset X_{t_1} \subset X_{t_2} \subset \cdots \subset X_{t_m} = X$

Apply the homology functor
(with coefficients in F_2)

PH is by definition the sequence of F_2-vector spaces (for each dimension d)

$$H_d(\emptyset) \to \cdots \to H_d(X_{t_{i-1}}) \to H_d(X_{t_i}) \to \cdots \to H_d(X_{t_{j-1}}) \to H_d(X_{t_j}) \to \cdots \to H_d(X)$$

The sequence decomposes into the direct sum of "intervals" having the form

$$0 \longrightarrow \cdots \longrightarrow 0 \longrightarrow \mathbf{F}_2 \xrightarrow{\mathrm{Id}} \cdots \xrightarrow{\mathrm{Id}} \mathbf{F}_2 \longrightarrow 0 \longrightarrow \cdots \longrightarrow 0$$

which correspond to cycles
(= generators = topological features)

represented by $(t_i, t_j) \in \mathbb{R}^2$

2D Example

0	0	0	0	0
0	0	1	2	0
1	1	1	2	0
0	1	0	0	0

$0 < t \leq 1$

0	0	0	0	0
0	0	1	2	0
1	1	1	2	0
0	1	0	0	0

$1 < t \leq 2$

0	0	0	0	0
0	0	1	2	0
1	1	1	2	0
0	1	0	0	0

$2 < t$

$PH_0 = \{(0,1], (0,\infty)\}$ (islands)
$PH_1 = \{(1,2]\}$ (holes)

A cycle of the form (a,b] is represented by a point $(a, b) \in R^2$

Software for Persistent Homology computation for image and volumetric data

- **Cubical Ripser (**K-Sudo-Ahara, 2021)
  - Open-source (MIT license), Available at my github repository https://github.com/shizuo-kaji/CubicalRipser_3dim/
  - Capable of computing persistent homology of time-series, image, volumetric data
  - One of the fastest program for computing persistent homology of cubical complexes
  - The only program which can handle two major constructions of cubical complexes
  - Python binding that works nicely with Numpy (including DICOM converters)
  - Tutorial (run on Google Colab): google “shizuo TDA tutorial”

2D or 3D image → Sublevel sets by sweeping thresholds → Persistent homology

PH as a feature

Input: Function (over a topological space)

1D

2D, 3D

Output：Persistence Diagram (finite points in R^2)

Red: PH_0 cycles
Blue: PH_1 cycles

Applications to CT analysis

We cannot include the details here
since some materials are not publishable online.
Contact Kaji for a full copy of the slides.

Summary

- Topology (persistent homology) provides a way to extract image/volume features that are not easy to obtain by conventional method
- Global and invariant features encoded by persistent homology (PH) complement those (mainly local) features obtained by deep learning (DL) and can be used in conjunction to boost performance
- PH-based image analysis has some advantages:
  - robust and easily **transferable** (⇔ DL needs re-training)
  - **interpretable** (⇔ DL is often a blackbox)
  - 3D (⇔ many conventional analyses are 2D slice-based)

Plasticity – Modeling and mathematical analysis

Konrad Kisiel
Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland
(joint work with Krzysztof Chełmiński)

Systems of equations describing an inelastic response of metals, with the fundamental assumption of small deformations, consist of linear partial differential equations coupled with nonlinear differential inclusions (or ordinary differential equations) for the vector of internal variables. The partial differential equations result from general mechanical laws. The differential inclusions are experimental, and depend on the kind of considered materials. One of the main assumptions needed in known existence theories is so-called safe-load condition. This kind of assumption is an indirect assumption on regularity of data. Our main goal is to present a method to obtaining existence of solutions, where the safe-load condition can be replaced by an assumption abouth the size of the set of addmissible stresses.

Plasticity – Modeling and mathematical Analysis

Krzysztof Chełmiński, Konrad Kisiel
Faculty of Mathematics and Information Science
Warsaw University of Technology

Workshop on Mathematics for Industry

Warszawa, 25-29 September, 2023

1

Table of contents

2

1. Theory of inelastic deformations - short introduction

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with a smooth boundary $\partial\Omega$.

Balance of momentum

$$\rho u_{tt}(x,t) \;=\; \mathrm{div}_x\, T(x,t) + F(x,t)\,, \qquad \rho u_{tt} \sim 0 \;\; \text{(quasistatic case)}$$

$(u,T) : \Omega \times (0,T_e) \to \mathbb{R}^3 \times \mathcal{S}^3$ – (the displacement vector, the stress tensor)
$F : \Omega \times (0,T_e) \to \mathbb{R}^3$ – the given external force, $\rho > 0$ – the mass density

Elastic constitutive relation

$$T(x,t) = \mathcal{D}\Big(\varepsilon(x,t) - \varepsilon^p(x,t)\Big)$$

$\varepsilon = \frac{1}{2}(\nabla u + \nabla^T u)$ – the linearized strain tensor
$\varepsilon^p : \Omega \times (0,T_e) \to \mathcal{S}^3$ – the plastic strain tensor, $\mathcal{D} : \mathcal{S}^3 \to \mathcal{S}^3$ – the elasticity tensor (symmetric, > 0)

Inelastic constitutive relation

$$\varepsilon^p_t(x,t) \in f\Big(\varepsilon(x,t), \varepsilon^p(x,t)\Big)$$

$f : D(f) \subset \mathcal{S}^3 \times \mathcal{S}^3 \to \mathcal{P}(\mathcal{S}^3)$ – a given constitutive multifunction

3

Models of premonotone type

Prof. Dr. Dr. h.c. Hans-Dieter Alber in the monograph *Materials with memory* LNM 1998 has defined a very large class of models: models of premonotone type.

Definition 1
A model is called of premonotone type if the inelastic constitutive relation is in the form

$$\varepsilon_t^p \in g\Big(-\rho\nabla_{\varepsilon^p}\psi(\varepsilon,\varepsilon^p)\Big)$$

where $\psi(\varepsilon,\varepsilon^p)=\frac{1}{2}\mathcal{D}(\varepsilon-\varepsilon^p)\cdot(\varepsilon-\varepsilon^p)$ is the free energy function and $g:D(g)\subset\mathcal{S}^3\to\mathcal{P}(\mathcal{S}^3)$ is a given inelastic multifunction satisfying:

$$\forall z\in D(g)\qquad g(z)\cdot z\geq 0\qquad (*)$$

If we additionally assume that $g(0)\ni 0\quad(*)\quad\Leftrightarrow$ monotonicity at the point 0. All models used in practice are of premonotone type.

Models of monotone type $\Leftrightarrow$ g is additionally monotone

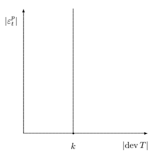
4

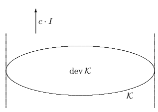
2. Elasto - perfect plasticity (the Prandtl–Reuss model)

$$\varepsilon_t^p(x,t)\in\partial I_{\mathcal{K}}\big(T(x,t)\big)\,,\quad \mathcal{K}=\operatorname{dev}\mathcal{K}\times\{c\cdot\mathbb{I}:c\in\mathbb{R}\}$$

where $\operatorname{dev}T=T-1/3\,(\operatorname{tr}T)\cdot\mathbb{I}$. Moreover, $\operatorname{dev}\mathcal{K}$ is convex with $0\in\operatorname{int}(\mathcal{K})$.

Hencky flow rule $\operatorname{dev}\mathcal{K}=B(0,k)\quad\Leftrightarrow\quad\forall\,S\in\mathcal{K}\quad|\operatorname{dev}S|\leq k\,.$





$$S\in\partial I_{\mathcal{K}}(T)\Leftrightarrow(S,T-\tau)\geq 0\quad\forall\,\tau\in\mathcal{K}$$
$$\partial I_{\mathcal{K}}(T)\ \text{ is monotone and }\ 0\in\partial I_{\mathcal{K}}(0)$$

5

3. Safe-load condition

Definition 2 (quasistatic case)

The given data F,g_N satisfy the safe-load condition if there exists g_D^* such that the unique solution (u^*,T^*) of the linear system

$$\begin{aligned}\operatorname{div}_x T^*(x,t)&=-F(x,t)\\ T^*(x,t)&=\mathcal{D}\varepsilon(u^*(x,t))\\ u^*(x)_{|\Gamma_D}&=g_D^*(x,t)\,,\quad T^*(x)\cdot n(x)_{|\Gamma_N}=g_N(x,t)\,.\end{aligned}$$

have the regularity:
$u^*\in\mathbb{W}^{1,\infty}(\mathbb{H}^1)$, $T^*\in\mathbb{W}^{1,\infty}(\mathbb{L}^2)$ and there exists $\delta>0$ such that

$$\{T^*+\sigma:|\sigma|\leq\delta\}\subset D(g)$$

and there exist uniformly bounded in $\mathbb{L}^\infty(\mathbb{L}^2)$ selections of the sets $g(T^*+\sigma)$.

For the Prandtl-Reuss model with the Hencky flow rule this condition is equivalent to: there exists $\delta>0$ such that $|\operatorname{dev}T^*|\leq k-\delta\,.$

6

Theorem 1
If the given data satisfy the safe-load condition then the sequences $\{\varepsilon_t^{p,k}\}, \{\varepsilon_t^k\}$ from a "good enough" approximation are bounded in the space $\mathbb{L}^\infty(\mathbb{L}^1)$.

Remark 1
Without any additional geometrical conditions for g the strains are weakly relatively compact in the space $\mathbb{L}^\infty(\mathcal{M})$ where $\mathcal{M}$ is the space containing bounded measures.

Remark 2
C. Johnson in 1976 was the first mathematician, which has formulated the safe-load condition for the Prandtl-Reuss model. The condition of Johnson is a little bit weaker as presented in this lecture.

The Johnson safe-load condition for the Prandtl-Reuss model

There exists a stress field S^* such that

$-\operatorname{div} S^* = F\,,\ S^* \cdot n = g_N$ and $\exists\, \delta > 0 \quad S^* + B(0,\delta) \subset \mathcal{K} \Leftrightarrow \quad |\operatorname{dev} S^*| \le k - \delta\,.$

7

4. Energy estimates without safe-load condition

Let us consider for simplicity the quasistatic Prandtl-Reuss model .

$$\begin{aligned} -\operatorname{div}_x T &= F, \\ T &= \mathcal{D}\left(\varepsilon - \varepsilon^p\right), \\ \varepsilon_t^p &\in \partial I_{\mathcal{K}}\left(T\right), \end{aligned}$$

Our approach is to modify only the inelastic constitutive equation and consider the following problem

$$\begin{aligned} -\operatorname{div}_x T^\lambda &= F, \\ T^\lambda &= \mathcal{D}\left(\varepsilon^\lambda - \varepsilon^{p,\lambda}\right), \\ \varepsilon_t^{p,\lambda} &= \mathcal{M}^\lambda\left(T^\lambda\right), \end{aligned}$$

where $\mathcal{M}^\lambda : \mathcal{S}^3 \to \mathcal{S}^3$ denotes the Yosida approximation of the maximal-monotone operator $\partial I_{\mathcal{K}}$.

8

Let us recall the definition of the space $LD(\Omega)$.

Definition 3

$$LD\left(\Omega\right) = \left\{u \in \mathbb{L}^1\left(\Omega;\mathbb{R}^3\right) : \varepsilon\left(u\right) \in \mathbb{L}^1\left(\Omega;\mathcal{S}^3\right)\right\}$$

$LD\left(\Omega\right)$ is the Banach space equipped with the standard norm

$$\|u\|_{LD(\Omega)} = \|u\|_{\mathbb{L}^1(\Omega)} + \|\varepsilon\left(u\right)\|_{\mathbb{L}^1(\Omega)}\,.$$

Theorem 2
Assume that $\Omega \subset \mathbb{R}^3$ is open, bounded and $\partial\Omega \in C^1$. Then, there exists a bounded linear operator

$$\gamma\colon LD\left(\Omega\right) \to \mathbb{L}^1\left(\partial\Omega;\mathbb{R}^3\right),$$

such that $\gamma\left(u\right) = u_{|\partial\Omega}$ for every $\varphi \in LD\left(\Omega\right) \cap C^0\left(\overline{\Omega}\right)$. Hence

$$\exists\, C_{LD} > 0 \quad \forall\, u \in LD\left(\Omega\right) \qquad \|\gamma\left(u\right)\|_{\mathbb{L}^1(\partial\Omega)} \leqslant C_{LD}\,\|u\|_{LD(\Omega)}.$$

Moreover, the following embedding theorem holds,

$$\exists\, C_{ELD} > 0 \quad \forall\, u \in LD\left(\Omega\right) \qquad \|u\|_{\mathbb{L}^{3/2}(\Omega)} \leqslant C_{ELD}\,\|u\|_{LD(\Omega)}.$$

9

We observed that in order to obtain proper energy estimates it is enough to assume the admissibility of the Neumann boundary data and the external force, which means

Definition 4 (Admissibility of forces)
We say that in the dynamical case the Neumann boundary data g_N is admissible if

$$C_{LD} \|g_N\|_{\mathbb{L}^\infty(0,T_e;\mathbb{L}^\infty(\Gamma_N))} < C^*,$$

where C_{LD} is a positive constant from the trace theorem in the space $LD(\Omega)$. The constant C^* depends on the maximal monotone inelastic multifunction only (for the Prandtl-Reuss model with the Hencky flow rule C^* is equal to the yield constant k.)

We say that in the quasi-static case the Neumann boundary data g_N and the external force F are admissible if

$$C_{ELD}\|F\|_{\mathbb{L}^\infty(0,T_e;\mathbb{L}^3(\Omega))} + C_{LD} \|g_N\|_{\mathbb{L}^\infty(0,T_e;\mathbb{L}^\infty(\Gamma_N))} < C^*,$$

where the constant C_{ELD} is from the embedding theorem for the space $LD(\Omega)$ and the constant C^* is the same as in the dynamical case.

10

Theorem 3 Assume that the data are regular enough and boundary data g_N is admissible or in the quasistatic case g_N and F are admissible. Then there exists a positive constant C, independent of λ, such that in the dynamical case

$$\mathcal{E}\left(u_t^\lambda, \varepsilon^\lambda, \varepsilon^{p,\lambda}\right) \ , \int_0^t \int_\Omega \varepsilon_t^{p,\lambda} \cdot T^\lambda, \ \mathcal{E}\left(u_{tt}^\lambda, \varepsilon_t^\lambda, \varepsilon_t^{p,\lambda}\right) \ , \left\|\varepsilon_t^{p,\lambda}\right\|_{\mathbb{L}^\infty(\mathbb{L}^1)} \le C.$$

where $2\mathcal{E}\left(u_t^\lambda, \varepsilon^\lambda, \varepsilon^{p,\lambda}\right) = \int_\Omega (\rho |u_t^\lambda|^2 + \mathcal{D}(\varepsilon^\lambda - \varepsilon^{p,\lambda}) \cdot (\varepsilon^\lambda - \varepsilon^{p,\lambda}))\, dx$
and in the quasistatic case

$$\mathcal{E}\left(\varepsilon^\lambda, \varepsilon^{p,\lambda}\right) \ , \int_0^t \int_\Omega \varepsilon_t^{p,\lambda} \cdot T^\lambda, \ \mathcal{E}\left(\varepsilon_t^\lambda, \varepsilon_t^{p,\lambda}\right) \ , \left\|\varepsilon_t^{p,\lambda}\right\|_{\mathbb{L}^\infty(\mathbb{L}^1)} \le C.$$

where $2\mathcal{E}\left(\varepsilon^\lambda, \varepsilon^{p,\lambda}\right) = \int_\Omega \mathcal{D}(\varepsilon^\lambda - \varepsilon^{p,\lambda}) \cdot (\varepsilon^\lambda - \varepsilon^{p,\lambda})\, dx$

10

Literature used in the lecture

- H.-D. Alber – *Materials with memory* – Lecture Notes in Math. v. 1682, Springer, Berlin Heidelberg New York, 1998
- C.Johnson – *Existence theorems for plasticity problems* – J. M. P. Appl. 1976
- R. Temam – *A generalized Norton-Hoff model and the Prandtl-Reuss law of plasticity* – ARMA, 1986
- K. Ch. – *Coercive approximation of viscoplasticity and plasticity* – Asymptotic Anal., 2001
- K. K., – *Dynamical problems in the theory of inelastic deformations* – PhD thesis, 2018
- K. K., K. Ch. – *On strong solutions of viscoplasticity without safe-load conditions* – J. Diff. Equ., 2020
- K. K, K. Ch. – *Prandtl-Reuss dynamical elasto-perfect plasticity without safe-load conditions* – Nonlinear Analysis TMA, 2020
- K. K, K. Ch. – *Quasistatic viscoplasticity without safe load conditions* – J. Diff. Equ., 2021

11

Developable surfaces with curved folds and applications

Miyuki Koiso
Institute of Mathematics for Industry, Kyushu University, Japan

A developable surface is a surface which is isometric to a planar region, that is, there exists a continuous bijective mapping from the surface to a planar region which preserves the length of every curve. If the considered surface is smooth, then it is developable if and only if its Gaussian curvature vanishes everywhere. Moreover, in this case, the surface can be continuously and isometrically deformed until the planar region. In this talk, we discuss developable surfaces with curved folds, which are naturally appear as origami works and have many applications in manufacturing objects. We discuss intrinsic and extrinsic singular points (such as vertices and points in edges), curvatures at each singular point, and the existence and nonexistence of continuous isometric deformations from such a surface to a planar region. We also discuss applications and discretization of these objects.

WORKSHOP on Mathematics for Industry 2023

Developable surfaces with curved folds and applications*

Miyuki Koiso (Kyushu University, Japan)

Collaborated with:
J. Mitani (information science), T. Homma (architecture),
Y. Yokosuka (architecture), T. Kitahata (physics),
M. Yasumoto (discrete geometry), Y. Jikumaru (geometry)

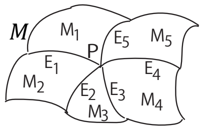
September 26, 2023, Warsaw University of Technology, Poland

*This work is supported by JST CREST Grant Number JPMJCR1911 and JSPS KAKENHI Grant Number JP20H01801.

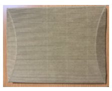
1

Plan of the talk

We consider oriented piecewise smooth (PW smooth) surfaces $M = \sum_i M_i$ in $\mathbb{E}^3$. Here M is a 2-dimensional manifold, each M_i is a smooth surface with boundary, and locally the number of M_i is finite.



- Developable surfaces
- A specific class of PW smooth developable surfaces called "pillow boxes", and a variational problem for them.



A double flat rectangle

Bend, and fold along curves

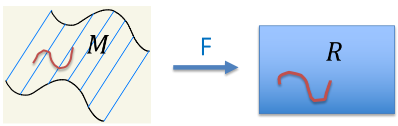
A pillow box

- Continuous isometric deformations from pillow boxes to planar regions
- Application

2

Developable surfaces

Def. 1. A PW-smooth surface M is said to be developable if it is isometric to a planar region R (that is, there exists a continuous bijective mapping F from M onto R that preserves the length of each curve).

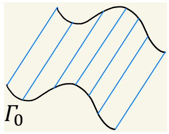


Remark 1. It is well-known that a smooth surface M is developable if and only if the Gaussian curvature $K(p)$ of M vanishes at any point $p \in M$.

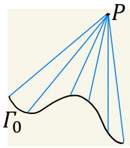
3

Smooth developable surfaces
(smooth surfaces with 0-Gaussian curvature)

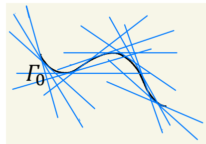
Fact 1. Smooth developable surfaces in $\mathbb{E}^3$ are the following:
(1) cylinders, (2) cones, (3) tangent developable surfaces.



cylinder



cone



tangent developable surface

Since developable surfaces can be constructed by bending a flat sheet, they are important in manufacturing objects from sheet metal, cardboard, and plywood (consists of three or more layers of veneer).

4

A variational problem for developable surfaces
"Find the optimal pillow box ! "

What is a pillow box?

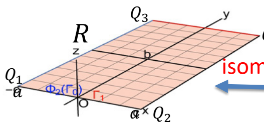
A double flat rectangle
(topologically, 2-spere S^2)
made of paper

Fold along curves

Pillow box

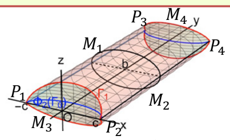
Def. 2 (Pillow box). A pillow box is a compact PW-smooth surface without boundary with genus 0 that consists of four parts of (generalized) cylinders and that is isometric to a double rectangle.

Q: For a given double rectangle, find the pillow box with the maximal volume. A: We will give a (rigorous) answer.



A double rectangle

isometric

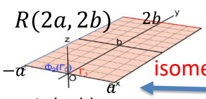


Pillow box

5

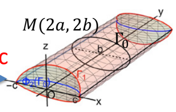
Existence and uniqueness of the optimal pillow box

Theorem 1 (K): For any given double rectangle $R(2a, 2b)$ with side lengths $2a, 2b$ (see the picture below) there exists a unique pillow box $M(2a, 2b)$ (which we call the optimal pillow box) that encloses the largest volume. It has an explicit representation using elliptic integrals. It consists of four (generalized) cylinders (of C^∞class) of which the base curves (the top and the bottom half of Γ_0 and two blue curves in the picture below right) are congruent and they are elastic curves.



A double rectangle

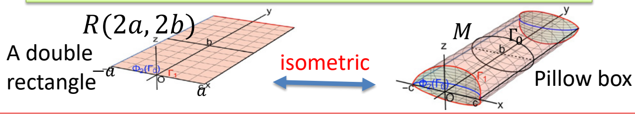
isometric



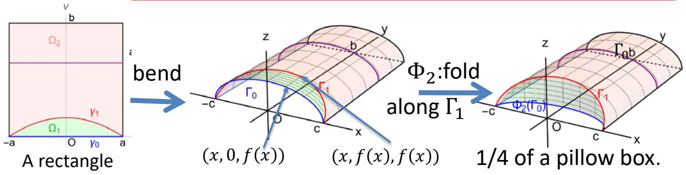
Pillow box

Remark 2.
(1) $\lim_{b\to\infty} M(2a, 2b) =$ a right circular cylinder with radius $2a/\pi$.
(2) $\lim_{a\to\infty} M(2a, 2b) =$ two parallel rectangles with width b and infinite length.

Outline of the proof of Theorem 1 (I) --- Step 1---



Step1. We observe that, for any pillow box M, the base curves of the four cylinders of which M consists are all congruent. We denote one of them by $\Gamma_0 : z = f(x)\ (-c \leq x \leq c)$ (see the picture below), and it is sufficient to study only 1/4 of the pillow box.



Therefore, our problem becomes a problem for plane curves!

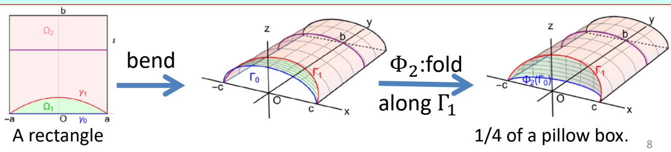
7

Outline of the proof of Theorem 1 (2) --- Step 2---

Step2. We consider the following variational problem for plane curves. For a given surface area, we maximize the enclosed volume of the pillow box given by a plane curve $\Gamma_0 : z = f(x)$. Using the method of Lagrange multiplier, we derive the Euler-Lagrange equation for $\Gamma_0 : z = f(x)$ which gives a critical point of the functional "Area $+\mu\cdot$Volume".
The result is the following ODE:

$$(1 + (f')^2)^{-\frac{3}{2}} f'' = \left(\frac{2\mu}{b}\right) f - \mu. \qquad \cdots (1)$$

This equation means that the curvature κ of Γ_0 is a linear function of the height, which implies that Γ_0 is an elastic curve.



8

Outline of the proof of Theorem 1 (3) --- Step 3, 4, 5---

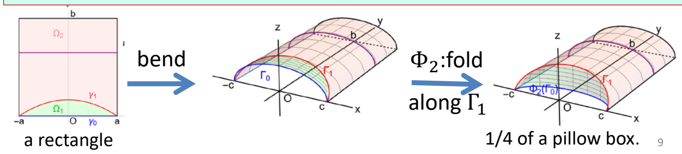
Step3. We derive the boundary condition for our ODE:

$$(1 + (f')^2)^{-\frac{3}{2}} f'' = \left(\frac{2\mu}{b}\right) f - \mu. \qquad \cdots (1)$$

in order that the solution gives a (local) maximum of volume. The result is: the curve $\Gamma_0 : z = f(x)$ must be orthogonal to the xy plane.

Step4. We solve our ODE (1) for the curve $\Gamma_0 : z = f(x)$ under the boundary condition that the curve Γ_0 is orthogonal to the xy plane.

Step5 (final step). We prove the existence of the (global) maximum of the volume of pillow boxes for any given double rectangle $R(2a, 2b)$.



9

Representation of the optimal pillow box (I) --- base curves---

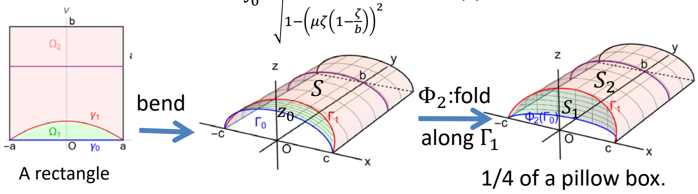
The base curve $\Gamma_0 : z = f(x)$ of the optimal pillow box is represented as follows.

$$\begin{cases} x = -I_\mu(z) + c, & 0 \le z \le z_0, \quad (0 \le x \le c) \\ x = I_\mu(z) - c, & 0 \le z \le z_0, \quad (-c \le x \le 0) \end{cases} \quad \cdots (2)$$

where, $I_\mu(z) := \int_0^z \frac{-\mu\zeta\left(1-\frac{\zeta}{b}\right)}{\sqrt{1-\left(\mu\zeta\left(1-\frac{\zeta}{b}\right)\right)^2}} d\zeta > 0, (0 < z < b)$, $z_0 := \frac{b}{2}\left(1 - \sqrt{1 - \frac{4}{b|\mu|}}\right)$,

$c := I_\mu(z_0)$. $\mu\ (< 0)$ is the curvature of Γ_0 at the end points that is determined by the following.

$$a = \int_0^{z_0} \frac{d\zeta}{\sqrt{1-\left(\mu\zeta\left(1-\frac{\zeta}{b}\right)\right)^2}} \quad \cdots (3)$$



Representation of the optimal pillow box (II) --- surface and volume ---

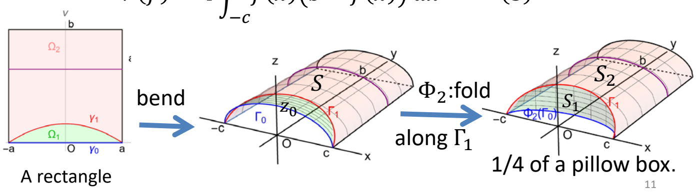
Let $\Gamma_0 : z = f(x)$ be the base curve of the optimal pillow box given in the previous slide.

The parts S_1, S_2 of the ¼ of the optimal pillow box are represented as

$$\begin{cases} S_1 = \{(x, f(x), z); -c \le x \le c, 0 \le z \le f(x)\} \\ S_2 = \{(x, y, f(x)); -c \le x \le c, f(x) \le y \le b\} \end{cases} \quad \cdots (4)$$

Hence, the volume $V(f)$ of the optimal pillow box is

$$V(f) = 4\int_{-c}^{c} f(x)\big(b - f(x)\big)\, dx \quad \cdots (5)$$

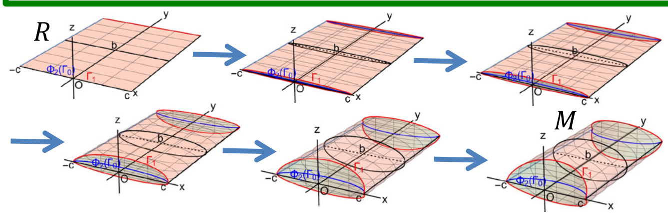


11

Continuous isometric (i.e. not expanding,not contracting) deformation from a planar double rectangle to a pillow box

For application, it is important to obtain the explicit representation from a planar region to a developable surface.

We can deform the initial double rectangle R to any given pillow box M which is isometric to R continuously and isometrically. However, the crease pattern (the red curves in the pictures below) is changed, which is not good for application.



An isometric deformation from R to M

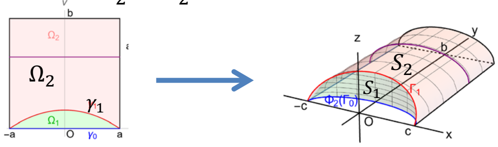
12

Isometric deformation from the single rectangle to 1/2 of the pillow box without changing crease pattern! (I)

The crease Γ_1 of a pillow box is represented as $(\eta(s), \zeta(s), \zeta(s))$, $(0 \leq s \leq L)$, where s is arc-length parameter of Γ_1. Set

$$\varphi_t(s) = \int_0^s \sqrt{1-(1+t^2)(\zeta'(s))^2}\, ds - c,$$
$$C_t(s) = (\varphi_t(s), \zeta(s), t\zeta(s)), \quad 0 \leq t \leq 1,$$
$$q_t(s,\tau) = C_t(s) + \tau \cdot (0,1,0), \quad 0 \leq \tau \leq b - \zeta(s).$$

Then, $C_0=\gamma_1$, $C_1=\Gamma_1$, and q_t gives an isometric deformation from Ω_2 to S_2.



Isometric deformation from the single rectangle to 1/2 of the pillow box without changing crease pattern! (II)

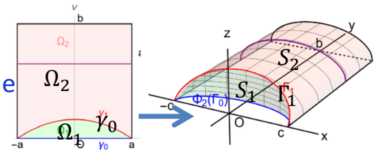
Next, set $p_t(s,\tau) = C_t(s) - \tau\beta_t(s)$,

$$0 \leq t \leq 1, \quad 0 \leq s \leq L, \quad -\zeta(s) \leq \tau \leq 0,$$

Where $C_t(s) = (\varphi_t(s), \zeta(s), t\zeta(s))$, $\beta_t(s) = \left(0, \frac{t^2-1}{t^2+1}, \frac{-2t}{t^2+1}\right)$.

Then, p_t gives an isometric deformation from Ω_1 to S_1.

p_t with q_t (in the previous page) gives an isometric deformation from a rectangle to 1/4 of the pillow box.



By extending the above deformation using the reflection with respect to the plane $\{y = b\}$, we obtain an isometric deformation from a single rectangle to 1/2 of the pillow box.

Future works

- For application, it is important to discuss "good" discretization of surfaces with curved folds.
- Discuss continuous isometric deformations from a general developable surface with curved folds to planar regions.

Summary

- We gave the definition of developable surfaces.
- We gave the existence, uniqueness, and representation formula of the optimal pillow box.
- We gave a continuous isometric deformation (concretely) from a planar region to a pillow box.
- We mentioned an application to architecture and discretization in the talk in the workshop. Because this work is in progress, its details are not included in this article.

16

Learning Permutation Symmetry of a Gaussian Vector

Bartosz Kołodziejek

Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland

The study of hidden structures in data presents challenges in modern statistics and machine learning. We introduce a Bayesian model selection approach, which allows to identify permutation subgroup symmetries in Gaussian vectors. In other words, given a finite iid sample of a p-dimensional Gaussian vector $Z = (Z_1, \ldots, Z_p)^\top$, we are looking for a permutation subgroup Γ acting on $\{1, \ldots, p\}$ such that

$$(Z_i)_{i=1}^p \text{ and } (Z_{\sigma(i)})_{i=1}^p \text{ have the same distributions}$$

for any $\sigma \in \Gamma$. We also find the maximum likelihood estimate of the covariance matrix in a Gaussian model obeying such symmetry restrictions. The talk is based on [1] and [2].

References

[1] Graczyk, P., Ishi, H., Kołodziejek, B. and Massam, H. (2022) *Model selection in the space of Gaussian models invariant by symmetry*. Ann. Statist. 50, no. 3, pp. 1747-1774.

[2] Graczyk, P., Ishi, H. and Kołodziejek, B. (2022) Graphical *Gaussian models associated to a homogeneous graph with permutation symmetries*, Physical Sciences Forum, 5(1), 20, pp. 1-9.

Learning Permutation Symmetry of a Gaussian Vector

Bartosz Kołodziejek

Warsaw University of Technology

Workshop on Mathematics for Industry
Warsaw 2023

29.09.2023

Talk is based on

Graczyk, Ishi, K., Massam
Model selection in the space of Gaussian models invariant by symmetry.
Annals of Statistics (2022)
and

Graczyk, Ishi, K.
Graphical Gaussian models associated to a homogeneous graph with permutation symmetries.
Proceedings of MaxEnt2022 (Physical Sciences Forum (2022))

This is an ongoing project.

R package: **gips**: Gaussian Model Invariant by Permutation Symmetry
`https://cran.r-project.org/package=gips`

Chojecki, Morgen, K.
Learning permutation symmetries with gips in R
arXiv:2307.00790

Program

1. Colored graphical Gaussian models.
2. Bayesian model selection when the graph is known.
3. Sketch of the main argument.
4. Main theoretical results and main message.
5. Some simulations.

Gaussian graphical models

- Assume that $Z = (Z_1, \dots, Z_p)^\top$ follows a Gaussian centered distribution with covariance matrix Σ.
- Let $K = \Sigma^{-1}$ be its precision/concentration matrix.
- Crucial fact: one has for $i \neq j$

$$K_{ij} = 0 \quad \iff \quad Z_i \text{ and } Z_j \text{ given } (Z_k)_{k \neq i,j}.$$

- We can define a undirected graph $G = (V, E)$ with $V = \{1, \dots, p\}$ and

$$\{i, j\} \in E \quad \iff \quad K_{ij} \neq 0.$$

- Graph G encodes the conditional independence structure of Z.
- Model selection problem: based on a iid sample $Z^{(1)}, \dots, Z^{(n)}$ find graph G - frequentist (e.g. GLASSO) and Bayesian methods.
- Knowledge about graph G significantly improves the usual estimator of Σ and gives a nice interpretation.
- This is not only a representation of a problem: many algorithms from graph theory are important in this setting.

Colored graphical models

- Colored graphical model is a special type of a graphical model.
- Apart from the conditional independence structure, **symmetry restrictions** are imposed on the concentration or partial correlation matrices.
- These symmetries can represented by a colored graph.
- Three types of such models (RCON, RCOR, **RCOP**) were introduced by Höjsgaard and Lauritzen (JRSSB, 2008) to describe situations where some entries of concentration or partial correlation matrices are approximately equal.
- Motivation: Imposing symmetry reduces the number of parameters to estimate. This is especially useful when parsimony is needed, i.e. $p \gg n$.

Colored graphical Gaussian models

- Let $G = (V, E)$ be a undirected graph with $V = \{1, \dots, p\}$.

$$\mathcal{P}_G = \{K \in \mathrm{Sym}^+(p; \mathbb{R}) \colon K_{ij} = 0 \quad \text{iff} \quad i \nsim j\},$$

Statistical model is $\{\mathrm{N}_p(0, K^{-1}) \colon K \in \mathcal{P}_G\}$.

Colored graphical Gaussian models

- Let $G = (V, E)$ be a undirected graph with $V = \{1, \dots, p\}$.

$$\mathcal{P}_G = \{K \in \mathrm{Sym}^+(p; \mathbb{R}) \colon K_{ij} = 0 \quad \text{iff} \quad i \nsim j\},$$

Statistical model is $\{\mathrm{N}_p(0, K^{-1}) \colon K \in \mathcal{P}_G\}$.

- For a permutation subgroup Γ on V, we define the space of concentration matrices **invariant under** Γ:

$$\mathrm{RCOP}_G(\Gamma) = \{K \in \mathcal{P}_G \colon K_{ij} = K_{\sigma(i)\sigma(j)} \text{ for all } \sigma \in \Gamma\}.$$

- Clearly, one requires that zeros are preserved, i.e.

$$i \sim j \qquad \iff \qquad \sigma(i) \sim \sigma(j) \quad \text{for all } \sigma \in \Gamma,$$

which implies that $\Gamma \subset \mathrm{Aut}(G)$.

- **Nice algebraic structure of RCOP and nice interpretation:**

$$Z \sim \mathrm{N}_p(0, \Sigma)$$

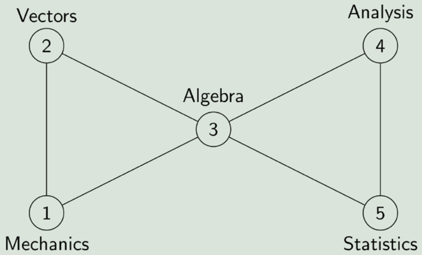
$$\Sigma^{-1} \in \mathrm{RCOP}_G(\Gamma) \qquad \iff \qquad Z \overset{d}{=} (Z_{\sigma(i)})_i \quad \text{for all } \sigma \in \Gamma.$$

- **Interpretation**: If the distribution of Z is invariant under the subgroup Γ and $(i, j) \in \Gamma$, then Z_i and Z_j play a symmetrical role.
- E.g.: some genes may have similar functions or groups of genes may have similar interactions or regulatory mechanisms.
- **Central problem = model selection**: given a centered Gaussian iid sample $Z^{(1)}, \dots, Z^{(n)}$ find a subgroup Γ under which the distribution of Z is most likely (in a Bayesian setting) invariant.

- There are no other model selection procedures which search among RCOP models.
- There is a number of articles dealing with model selection within RCON models:
  - Gehrmann. *Symmetry*, 2011.
  - Gao and Massam. *J. Comput. Graph. Statist.*, 2015.
  - Massam, Li and Gao. *Biometrika*, 2018.
  - Li, Gao and Massam. *J. Stat. Comput. Simul.*, 2020.
  - Li, Sun, Wang and Gao. *Stat. Anal. Data Min.*, 2021.
  - Ranciati, Roverato and Luati. *J. R. Stat. Soc. Ser. C.* , 2021.

  Typically, both the graph G and the coloring $\mathcal{C}$ are assumed unknown.
  Methods: both Bayesian and penalized likelihood methods.

Example (Graphical model)



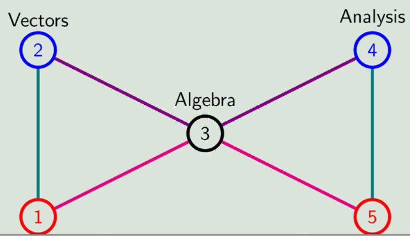
If $K \in \mathcal{P}_G = \{K \in \mathrm{Sym}^+(p;\mathbb{R}) \colon K_{ij} = 0 \quad \text{iff} \quad i \nsim j\}$, then

$$K = \begin{pmatrix} x_{11} & x_{21} & x_{31} & 0 & 0 \\ x_{21} & x_{22} & x_{32} & 0 & 0 \\ x_{31} & x_{32} & x_{33} & x_{43} & x_{53} \\ 0 & 0 & x_{43} & x_{44} & x_{54} \\ 0 & 0 & x_{53} & x_{54} & x_{55} \end{pmatrix}.$$

Example (Colored graphical model)

There are 10 subgroups of $\mathrm{Aut}(G)$, they correspond to 7 different colorings.
Let $\Gamma = \langle (1 \quad 5)\,(2 \quad 4) \rangle = \{\mathrm{id}, (1 \quad 5)\,(2 \quad 4)\}$.
If $K \in \mathrm{RCOP}_G(\Gamma) = \{K \in \mathcal{P}_G \colon K_{ij} = K_{\sigma(i)\sigma(j)}$ for all $\sigma \in \Gamma\}$, then

$$K = \begin{pmatrix} a & d & e & 0 & 0 \\ d & b & f & 0 & 0 \\ e & f & c & f & e \\ 0 & 0 & f & b & d \\ 0 & 0 & e & d & a \end{pmatrix},$$



Bayesian model

- Fixed graph G is chordal, i.e. each cycle in G has a chord (there are no induced cycles of length $\geqslant 4$).
- We assume that $K = \Sigma^{-1}$ and the subgroup Γ are random.
- $Z_1, \ldots, Z_n$ given $\{K, \Gamma\}$ are i.i.d. $\mathrm{N}_p(0, K^{-1})$.
- Γ is uniform on **(a subfamily of)** subgroups of $\mathrm{Aut}(G)$.
- $K|\Gamma = \gamma$ is the Diaconis-Ylvisaker conjugate prior on $\mathrm{RCOP}_G(\gamma)$:

$$f_{K|\Gamma=\gamma}(k) = \frac{1}{I_G^{\gamma}(\delta, D)} \mathrm{Det}\,(k)^{(\delta-2)/2}\, e^{-\frac{1}{2}\mathrm{Tr}[D\cdot k]} \mathbf{1}_{\mathrm{RCOP}_G(\gamma)}(k).$$

- By standard argument, we have the posterior distribution:

$$\boxed{\mathbb{P}(\Gamma = \gamma | Z_1, \ldots, Z_n) \propto \frac{I_G^{\gamma}(\delta + n, D + \sum_{i=1}^n Z_i \cdot Z_i^\top)}{I_G^{\gamma}(\delta, D)}.}$$

- Bayesian paradigm: choose the model with the highest posterior probability:

$$\hat{\Gamma} = \arg\max_{\Gamma} \frac{I_G^{\Gamma}(\delta + n, D + \sum_{i=1}^n Z_i \cdot Z_i^{\top})}{I_G^{\Gamma}(\delta, D)}.$$

- Caution: as we will see, the state space is very big for large p.
- When p is large, we have to resort to MCMC methods.
  We can define a irreducible Markov chain on **(a subfamily of)** subgroups of $\mathrm{Aut}(G)$.
- We have to compute Gamma-like integrals over the colored cones:

$$I_G^{\Gamma}(\delta, D) = \int_{\mathrm{RCOP}_G(\Gamma)} \mathrm{Det}\,(k)^{(\delta-2)/2}\, e^{-\frac{1}{2}\mathrm{Tr}[D\cdot k]} dk.$$

Example

Let $G = K_3$ be the full graph on $V = \{1, 2, 3\}$ and let $\Gamma = \langle(13)\rangle$. We have

$$\mathrm{RCOP}_{K_3}(\Gamma) = \left\{ \begin{pmatrix} \alpha & \beta & \Delta \\ \beta & \gamma & \beta \\ \Delta & \beta & \alpha \end{pmatrix} : 2\beta^2\Delta + \alpha^2\gamma > 2\alpha\beta^2 + \Delta^2\gamma,\ \alpha\gamma > \beta^2 \right\}$$

and therefore

$$I_{K_3,\Gamma}(\delta, D) = \iiiint_{2\beta^2\Delta+\alpha^2\gamma>2\alpha\beta^2+\Delta^2\gamma,\ \alpha\gamma>\beta^2}$$

$$\mathrm{Det}^{(\delta-2)/2} \begin{pmatrix} \alpha & \beta & \Delta \\ \beta & \gamma & \beta \\ \Delta & \beta & \alpha \end{pmatrix} e^{-\mathrm{Tr}\left[D\cdot \begin{pmatrix} \alpha & \beta & \Delta \\ \beta & \gamma & \beta \\ \Delta & \beta & \alpha \end{pmatrix} \right]}$$

$$d\alpha\, d\gamma\, d\beta d\,\Delta.$$

Such integrals were known only if $\Gamma = \{\mathrm{id}\}$.

Sketch of the main argument

- Let $R(\sigma)$ be a permutation matrix of $\sigma \in \mathfrak{S}_p$.
- $R\colon \Gamma \mapsto \mathrm{GL}(p; \mathbb{R})$ satisfies

$$R(\sigma \circ \sigma') = R(\sigma) \cdot R(\sigma'), \qquad \sigma, \sigma' \in \mathfrak{S}_p.$$

- In other words, R is a representation of group Γ.
- Observe that for any $\sigma \in \mathfrak{S}_p$,

$$R(\sigma) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

- The space $W_0 = \mathbb{R}(1, \ldots, 1)^{\top}$ is a Γ invariant subspace for any subgroup Γ, that is, $\forall \sigma \in \Gamma$,

$$\forall w \in W_0 \qquad R(\sigma) w \in W_0.$$

- Similarly for $W_0^{\perp} = \{x \in \mathbb{R}^p\colon \sum_{i=1}^p x_i = 0\}$.

Sketch of the main argument

- Let orthogonal matrix U_Γ be constructed from a basis of W_0 (one column) and any basis of $W_0^\perp$. Then,

$$U_\Gamma^\top R(\sigma) U_\Gamma = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$$

- Define

$$\begin{aligned}\mathcal{Z}_\Gamma &= \{\, x \in \mathrm{Sym}(p;\mathbb{R})\,;\, x_{ij} = x_{\sigma(i),\sigma(j)} \text{ for all } \sigma \in \Gamma \,\} \\ &= \{\, x \in \mathrm{Sym}(p;\mathbb{R})\,;\, R(\sigma)\cdot x = x \cdot R(\sigma) \text{ for all } \sigma \in \Gamma \,\}.\end{aligned}$$

  and recall that $\mathrm{RCOP}_\Gamma(G) = \mathcal{P}_G \cap \mathcal{Z}_\Gamma$.
- Then $U_\Gamma^\top \mathcal{Z}_\Gamma U_\Gamma$ coincides with

$$\{\, y \in \mathrm{Sym}(p;\mathbb{R})\,;\, [U_\Gamma^\top R(\sigma) U_\Gamma]\cdot y = y \cdot [U_\Gamma^\top R(\sigma) U_\Gamma] \,\}.$$

- Block decomposition of $U_\Gamma^\top R(\sigma) U_\Gamma$ implies block decomposition of $y \in U_\Gamma^\top \mathcal{Z}_\Gamma U_\Gamma$.

Sketch of the main argument

- In general, there exist many proper Γ-invariant subspaces of $W_0^\perp$. Finding them is a classical problem and is not easy.
- Formally, $\mathcal{Z}_\Gamma$ is the set of symmetric intertwining operators of the representation $(R, \mathbb{R}^p)$.
- This implies the existence of a orthogonal matrix U_Γ such that $U_\Gamma^\top \mathcal{Z}_\Gamma U_\Gamma$ coincides with

$$\left\{ \begin{pmatrix} M_{\mathbb{K}_1}(x_1)^{\oplus k_1/d_1} & & \\ & \ddots & \\ & & M_{\mathbb{K}_L}(x_L)^{\oplus k_L/d_L} \end{pmatrix} : \begin{array}{r} x_i \in \mathrm{Herm}(r_i;\mathbb{K}_i) \\ i = 1,2,\ldots,L \end{array} \right\},$$

  where consecutive blocks correspond to irreducible components in a decomposition of $(R, \mathbb{R}^p)$.
- Each block corresponds to a **uncolored model**.

Theoretical results and the main message

- We have explicit formulas for normalizing constants I_G^Γ when G is a decomposable graph.
- These formulas depend on so-called **structure constants**.
  In principle, we know how to find these constants: "just" find irreducible representations over reals of Γ, which is a classical problem.
- However, this is generally computationally impossible for big p.
- We therefore identify a **good subfamily** of subgroups for which we can find these structure constants efficiently with p-polynomial complexity.
- We restrict our search space to models corresponding to that **good subfamily**.

Good subfamily = Cyclic subgroups

- Cyclic subgroups = groups generated by one permutation.
- A distribution is invariant under Γ **if and only if** it is invariant under any cyclic subgroup of Γ.
- Easy to interpret and seem rich enough.
- When G is sparse, then $\mathrm{Aut}(G)$ is small and contains mostly cyclic subgroups.
- We can use a permutation random walk to travel through cyclic subgroups: $\sigma_n = \sigma_{n-1} \circ \tau_n$, where $(\tau_n)_n$ are i.i.d. transpositions.

p	#subgroups of $\mathfrak{S}_p$	$\#\mathrm{RCOP}_{K_p}(\Gamma)$	#cyclic groups
1	1	1	1
2	2	2	2
3	6	5	5
4	30	22	17
5	156	93	67
6	1 455	739	362
7	11 300	4 508	2 039
$\vdots$			
18	$7 \cdot 10^{18}$	?	$7 \cdot 10^{14}$

The MLE of Σ in $\mathrm{RCOP}_{K_p}(\Gamma)$

Assume that G is complete:

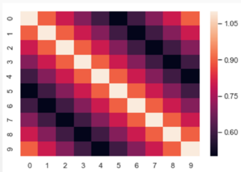
- We have "if and only if" condition on n for the MLE to exist.
- E.g.: if $V = \{1, \ldots, p\}$ and $\Gamma = \langle (1, 2, \ldots, p) \rangle$, then the MLE exists already for $n = 1$!
- If the graph is complete, then the MLE (if exists) under $\mathrm{RCOP}_G(\Gamma)$ model is given by

$$\hat{\Sigma}_{\mathrm{RCOP}(\Gamma)} = \pi_\Gamma \left(\hat{\Sigma} \right),$$

  where
  - $\hat{\Sigma}$ is the usual empirical covariance matrix,
  - π_Γ is the projection onto the colored matrix space: it averages entries corresponding to the same color.
- This results in improved estimation properties.

Simulations

- We choose $p = 10$ and $n = 20$,
- Let Σ_0 be a symmetric circulant matrix of the form



- We sample $Z_1, \ldots, Z_n$ from $\mathrm{N}_p(0, \Sigma_0)$.
- The distribution of Z is invariant under $\Gamma_0 = \langle (1, 2, \ldots, p) \rangle$.
- **We construct a Markov chain on permutations** and use it to travel through cyclic subgroups.
- The usual hyperparameters are $\delta = 3$, $D = I_p$.
- We iterate Metropolis-Hastings algorithm 100 000 times.
- We do this 100 times to assess variability of the procedure.

Tabela: Cyclic subgroups which were chosen by M-H algorithm

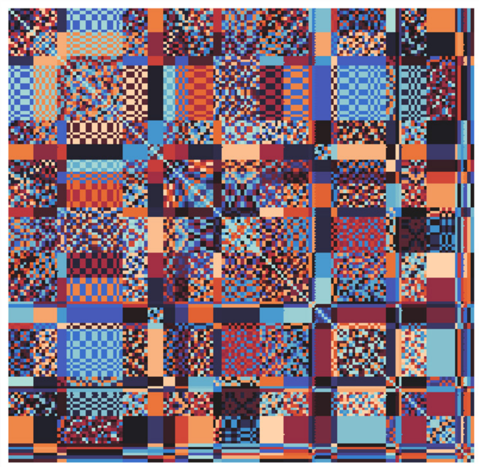
generator of a cyclic group	#most visited	ARI
(1, 2, 3, 4, 5, 6, 7, 8, 9,10)	25	1.00
(1, 3, 5, 7, 9)(2, 4, 6, 8,10)	13	0.60
(1, 2, 4, 3, 5, 6, 7, 9, 8, 10)	3	0.43
(1, 2, 4, 3, 5, 6, 7, 8, 9, 10)	2	0.46
(1, 3, 2, 4, 5, 6, 8, 7, 9, 10)	2	0.43
(1, 3, 5, 9, 2, 6, 8, 10, 4, 7)	2	0.43
(1, 4, 3, 5, 2, 6, 9, 8, 10, 7)	2	0.35
(1, 4, 5, 7, 8)(2, 3, 6, 9, 10)	2	0.24
(1, 8, 10, 9)(2, 7)(3, 5, 4, 6)	2	0.19
(1, 2, 10, 3)(4, 9)(5, 8, 6, 7)	2	0.19

- ARI = adjusted Rand index is a similarity measure comparing given coloring with the true one. ARI$\in [-1, 1]$
- For $n = p = 10$, the results were only slightly worse.

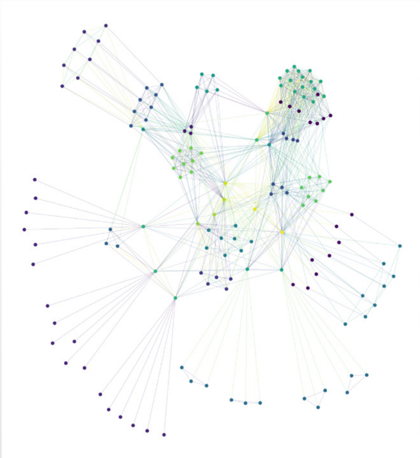
Real data example: $p = 150$

- Breast cancer data set: $p = 150$ genes and $n = 58$ samples.
- Cardinality of the search space is about 10^{250}.
- We iterate Metropolis-Hastings algorithm 150 000 times.
- The cyclic subgroup $\hat{\Gamma}$ with highest estimated posterior probability (7.1%) is of order 720.
- We have $\dim \mathrm{RCOP}_{K_p}(\hat{\Gamma}) = 844$ vs 11325 parameters of unrestricted model.
- The MLE for Σ exists for this model.

The color pattern of the space of $p \times p$ matrices from $\mathrm{RCOP}_{K_p}(\hat{\Gamma})$.



Thresholding the partial correlation



Thank you for your attention

Graczyk, Ishi, K., Massam
Model selection in the space of Gaussian models invariant by symmetry.
Annals of Statistics (2022)

Chojecki, Morgen, K.
Learning permutation symmetries with gips in R
arXiv:2307.00790

Supercoiled structure of DNA and hyperelliptic functions

Shigeki Matsutani
Institute of Science and Engineering, Kanazawa University, Japan

The geometry of DNA has a helical structure as well as a more global supercoiled structure. The geometry of this supercoiled structure is dominated by weak elastic forces, but its geometry has not yet been mathematically described. Geometric models that minimize its elastic energy, known as elasticae (elastic curves), cannot describe the shape of DNA, even if three-dimensional effects are considered. Since 1997, the speaker has been working to mathematically represent this shape by considering finite temperature effects [1]. It is known from elementary considerations that the shape of elastic curves under a finite temperature can be described by the hyperelliptic solution of the modified KdV equation, which is a nonlinear integrable equation, in the two-dimensional plane, and of the nonlinear Schrodinger equation in the three-dimensional space. However, Abelian function theory, including hyperelliptic function theory, had not reached the level where hyperelliptic function solutions could be specifically described and concretely treated at all as of 1997. Therefore, the speaker, together with late Emma Previato since 2003, has restructured the Abelian function theory to the level of elliptic function theory, and has also developed related theories [2]. With Previato, he obtained certain shapes in 2022, albeit incomplete [3]. Although incomplete means that it does not fully satisfy the reality condition, we were able to produce mathematically shapes that have some features of the supercoiled structure of DNA, albeit tentatively. This talk will describe the results obtained in 2022 and the process that led to them.

The speaker has been studied novel devies and materials mathematically in research and development of devices and materials for 27 years in Canon Inc. The usefulness of mathematics, including the theory of singularity, in modern society will be briefly discussed.

References

[1] S. Matsutani, *Statistical mechanics of elastica on a plane: origin of the MKdV hierarchy* , J. Phys. A: Math. & Gen., **31** (1998) 2705-2725.

[2] S. Matsutani, E. Previato, *The Weierstrass sigma function in higher genus and applications to integrable equations*, (in preparation).

[3] S. Matsutani, E. Previato, *An algebro-geometric model for the shape of supercoiled DNA* Physica D **430** (2022) 133073

Supercoiled structure of DNA and hyperelliptic functions

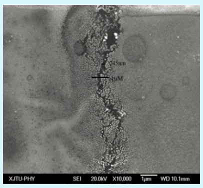
WORKSHOP on Mathematics for Industry 2023
September 28, (Thursday) 2023

Shigeki Matsutani
Kanazawa University

Menu

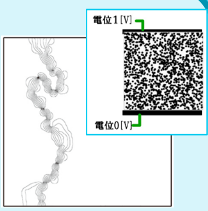
We live in an age in which we can create a new reality by translating the real world into mathematical language and investigating it.

Real World



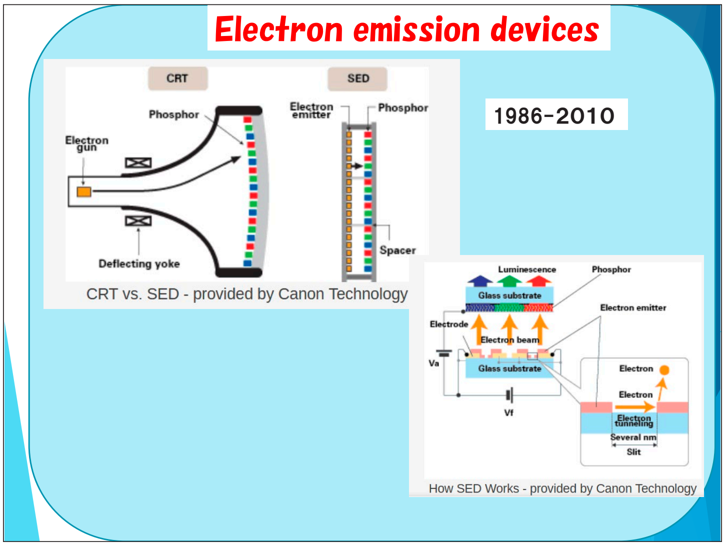
Point Process
Quasi-conf
Conformal

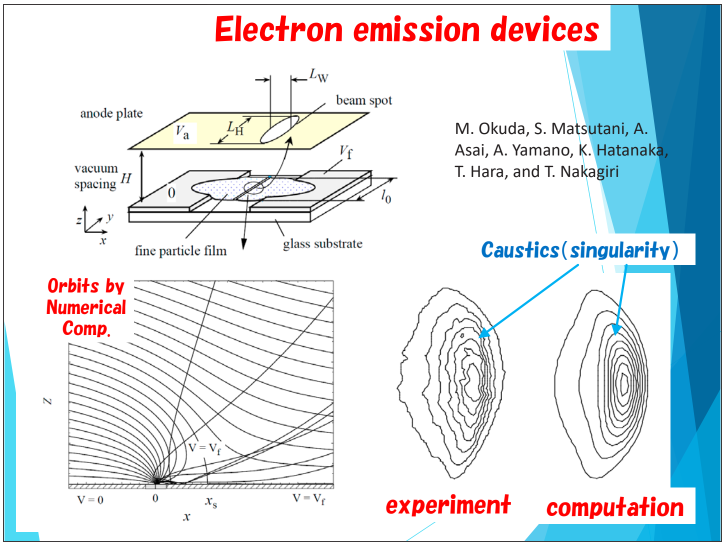
Math World

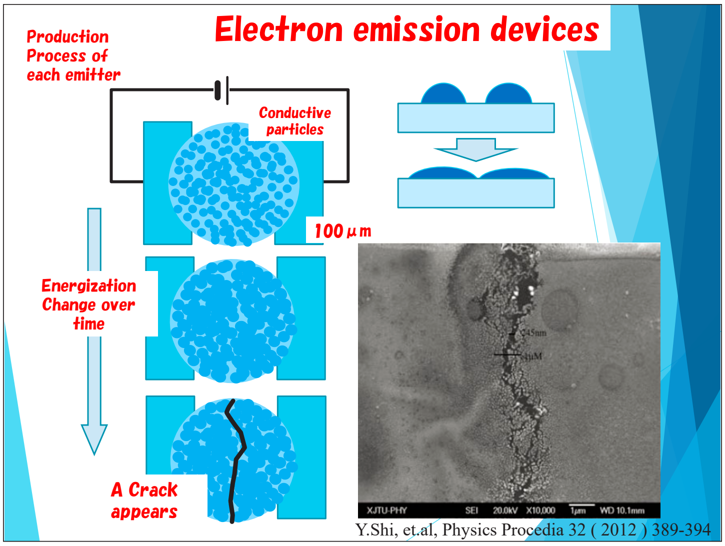


How to do this?

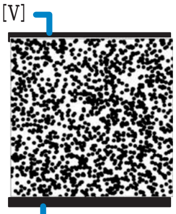
Y.Shi, et.al, Physics Procedia 32 (2012) 389-394


Electron emission devices
CRT
SED
Phosphor
Electron gun
Deflecting yoke
Electron emitter
Phosphor
Spacer
CRT vs. SED - provided by Canon Technology
1986-2010
Luminescence
Phosphor
Glass substrate
Electron emitter
Electrode
Electron beam
Va
Glass substrate
Vf
Electron
Electron
Electron tunneling
Several nm
Slit
How SED Works - provided by Canon Technology


Electron emission devices
anode plate
beam spot
vacuum spacing
fine particle film
glass substrate
M. Okuda, S. Matsutani, A. Asai, A. Yamano, K. Hatanaka, T. Hara, and T. Nakagiri
Caustics(singularity)
Orbits by Numerical Comp.
V = 0
V = Vf
experiment
computation


Electron emission devices
Production Process of each emitter
Conductive particles
100μm
Energization Change over time
A Crack appears
Y.Shi, et.al, Physics Procedia 32 (2012) 389-394

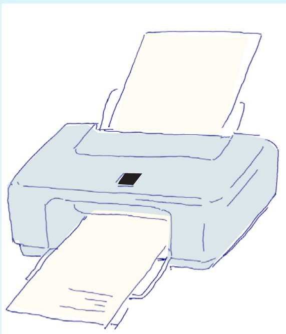
Electron emission devices



1. Scatter the conductive particles as a Monte-Carlo computation.
2. Solve the Dirichlet-Neumann problem,
   $$\mathrm{div}(\gamma \cdot \mathrm{grad}\ u) = 0$$
   numerically.

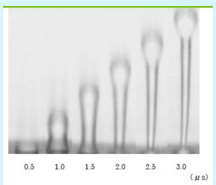
●conductive particle
□high resistance matter

Computational Fluid Dynamics



We live in an age in which we can create a new reality by translating the real world into mathematical language and investigating it.

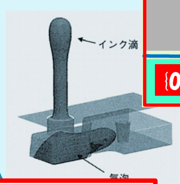
Real World

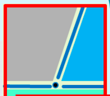


浅井朗　ながれ 24(2005)

Stratification
Inf. Lie. Alg.
Phase field

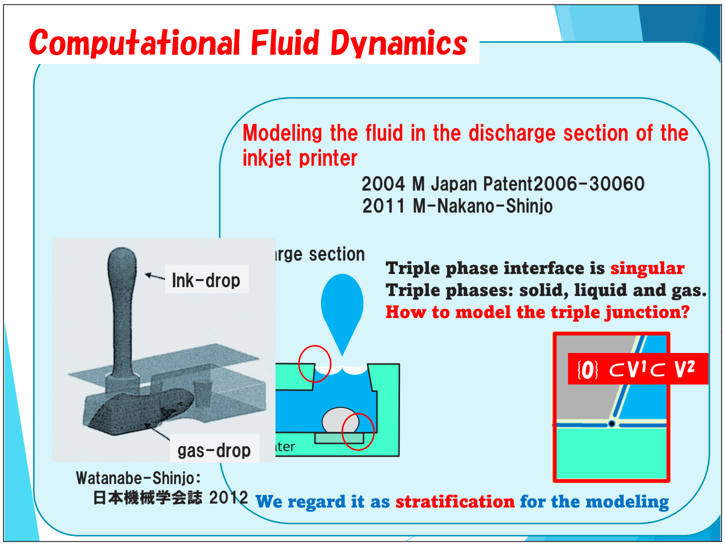
Math World

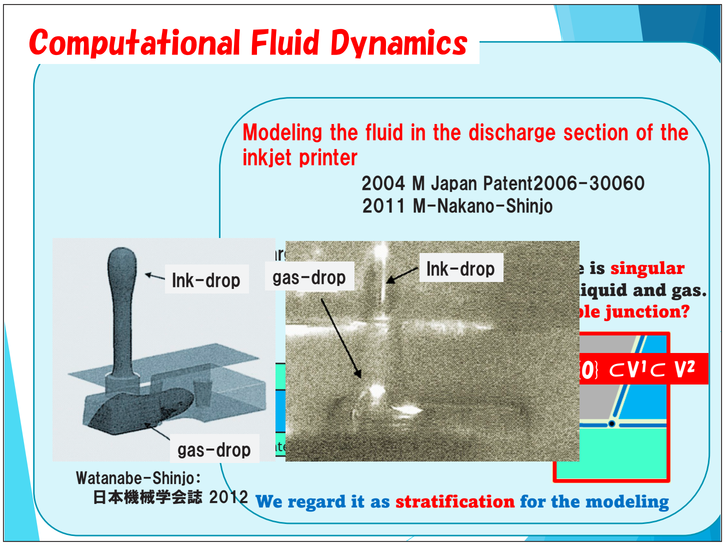


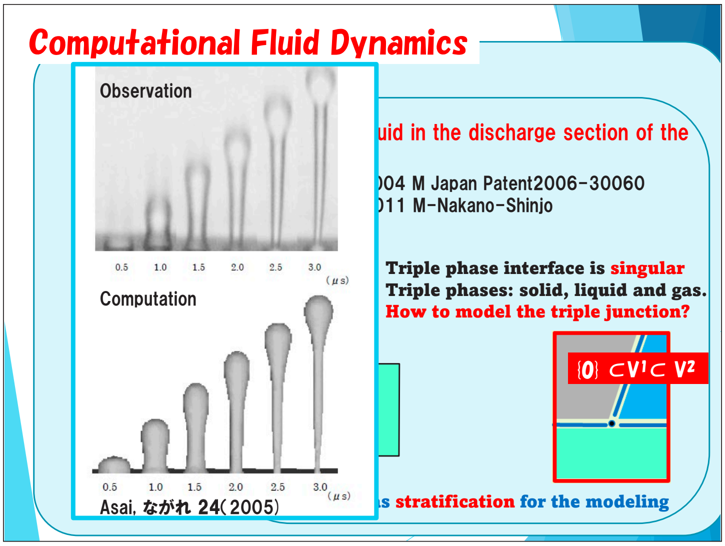


$\{0\} \subset V^1 \subset V^2$

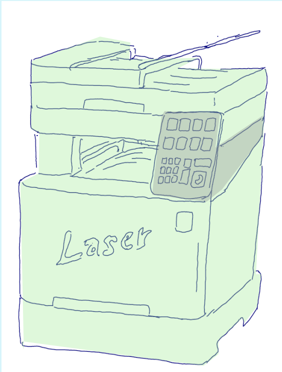
How to do this?


Computational Fluid Dynamics
Modeling the fluid in the discharge section of the inkjet printer
2004 M Japan Patent2006-30060
2011 M-Nakano-Shinjo
Ink-drop
gas-drop
Triple phase interface is singular
Triple phases: solid, liquid and gas.
How to model the triple junction?
$\{0\} \subset V^1 \subset V^2$
Watanabe-Shinjo:
日本機械学会誌 2012
We regard it as stratification for the modeling


Computational Fluid Dynamics
Modeling the fluid in the discharge section of the inkjet printer
2004 M Japan Patent2006-30060
2011 M-Nakano-Shinjo
Ink-drop
gas-drop
gas-drop
Ink-drop
Watanabe-Shinjo:
日本機械学会誌 2012
We regard it as stratification for the modeling

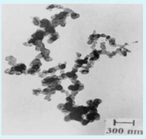

Computational Fluid Dynamics
Observation
0.5 1.0 1.5 2.0 2.5 3.0 (μs)
Computation
0.5 1.0 1.5 2.0 2.5 3.0 (μs)
Asai, ながれ 24(2005)
Triple phase interface is singular
Triple phases: solid, liquid and gas.
How to model the triple junction?
$\{0\} \subset V^1 \subset V^2$

Nano-matrials



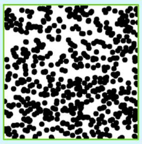
We live in an age in which we can create a new reality by translating the real world into mathematical language and investigating it.

Real World



Point Process
Quasi-conf
Conformal

Math World



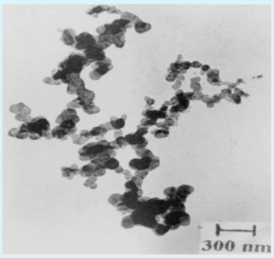
How to do this?

Preparation and Some Properties of a Nanocomposite of Polyacrylonitrile with Acetylene Black Arjun Maity and Mukul Biswas
Polymer Journal, Vol. 36 (2004) No. 10 pp.812-816

Nano-matrials

Material design of key components of LBPs
Mathematical modeling using percolation theory.
2015M-Shimosako

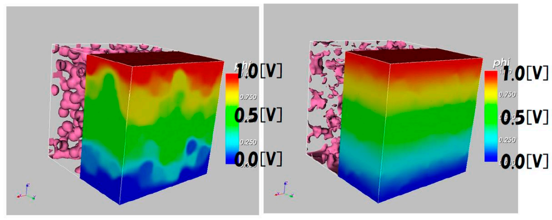
How to control the high resistivity by mixing conductive carbon nanoparticles with a high-resistivity polymer matrix to preserve its property for a long time.



Preparation and Some Properties of a Nanocomposite of Polyacrylonitrile with Acetylene Black Arjun Maity and Mukul Biswas
Polymer Journal, Vol. 36 (2004) No. 10 pp.812-816

Nano-matrials

Material design of key components of LBPs
Mathematical modeling using percolation theory.
2015M-Shimosako



3D potential distribution
Computational simulation

3D potential distribution
Computational simulation

We live in an age in which we can create a new reality by translating the real world into mathematical language and investigating it.

Real World

Point Process
Quasi-conf

Math World

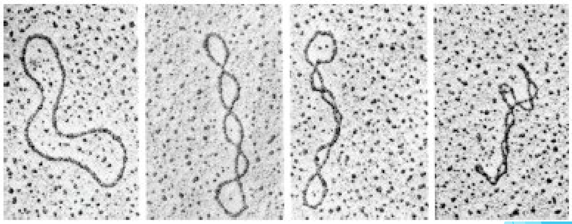
Perspectives from industry,
1. Cutting-edge technology requires cutting-edge mathematics.
2. Some of the mathematics for technology is so profound.

How to do this?

Menu

1. Self Introduction
2. Continuation of Self-Introduction
3. **Supercoiled structure of DNA**
4. Elastic curves
5. Statistical Mechanics of Elastic Curves
6. Excited states of elastic curves and the MKdV equation
7. MKdV hyperelliptic curve solution for genus 2
   7.1 Review of the case of genus 1
   7.2 Review of the case of genus 2
8. Future task

Supercoil in DNA

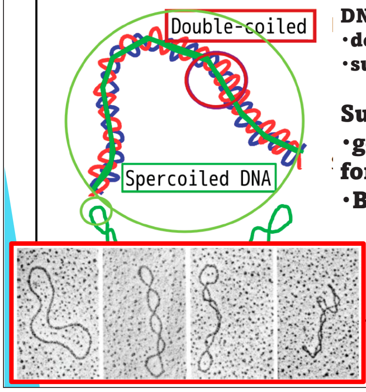


Electron microscope image of DNA

http://www.udel.edu/chem/bahnson/chem645/websites/Sapra/Supercoiling.html

Supercoil in DNA

Elastica の統計力学



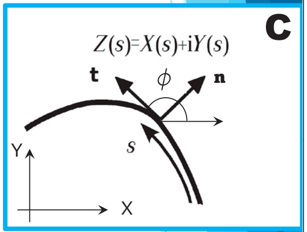
DNA forms
- **double coil structure and**
- **supercoil structure.**

Supercoil structure is
- **governed by weak elastic forces.**
- **But not an elastic curve**

Menu

What is an elastica (elastic curve)?

Elastica is a curve on a (complex) plane determined by elastic forces.



What is an elastica (elastic curve)?

$Z : N \to \mathbb{C}$:analytic immersion $(|\partial_s Z| = 1)$.

$N = S^1$ or $= [0, 1]$

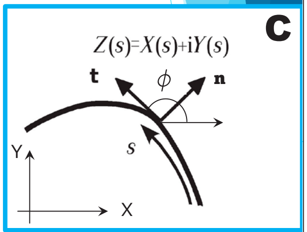
s : arclength

$Z(s) = X(s) + iY(s)$,

$t = \partial_s Z = e^{i\phi}$,

$(\phi \in C^\omega(N, R))$

$= \cos\phi + i \sin\phi$



$k := \partial_s \phi$: curvature; $k = 1/$[curvature radius].

What is an elastica (elastic curve)?

Curvature & Frenet-Serret relation

$t := \partial_s Z$:tangential vector, $n := i\,\partial_s Z$:normal vector

$\partial_s t = k\,n$, $\quad \partial_s n = -k\,t$, $\quad (\partial_s^2 Z = ik\partial_s Z)$

$k := \partial_s \phi$: curvature; $k = 1/$[curvature radius].

What is an elastica (elastic curve)?

Curvature & Frenet-Serret relation

t :=$\partial_s Z$:tangential vector, n:=i $\partial_s Z$:normal vector

∂_s t = k n, ∂_s n = −k t, ($\partial_s^2 Z$ = ik$\partial_s Z$)

$k := \partial_s \phi$: curvature; k = 1/[curvature radius].

Elastica Problem (Jacob/Daniel Bernoulli-Euler (1691-1744))

Determine the shape of the elastic curve that exists on the plane mathematically!

⇔ Find the shape that minimizes the energy $\boxed{\mathcal{E}[Z] = \frac{1}{2}\int_N k(s)^2\, ds}$ **under the iso-arc length.**

Infinitesimal isometric deformation

Infinitesimal variation: $Z_\varepsilon(s_\varepsilon) = Z(s) + \mathrm{i}\varepsilon(s)\partial_s Z$

$$\partial_s Z_\varepsilon = (1 - \varepsilon k(s) + \mathrm{i}\partial_s\varepsilon)\partial_s Z$$

$$ds_\varepsilon^2 = d\overline{Z_\varepsilon} dZ_\varepsilon = (1 - 2\varepsilon k + O(\varepsilon^2))ds^2$$

$-\mathrm{i}\partial_{s_\varepsilon} \log \partial_{s_\varepsilon} Z_\varepsilon$:

$$k_\varepsilon = k + (k^2 + \partial_s^2)\varepsilon + O(\varepsilon^2)$$

$$k_\varepsilon^2 ds_\varepsilon = (k^2 + (k^3 + 2k\partial_s^2)\varepsilon + O(\varepsilon^2))ds$$

Infinitesimal isometric deformation

$$\frac{\delta(2\mathcal{E}_\varepsilon - a\int_N ds_\varepsilon)}{\delta\varepsilon(s)} = k^3 + 2\partial_s^2 k + 2ak = 0$$

Static modified KdV equation

$$ak + \frac{1}{2}k^3 + \partial_s^2 k = 0$$

Elastica Problem is to find the shape of the curve whose curvature obeys the static modified KdV equation

It is a prototype of the nonlinear integrable system.

Infinitesimal isometric deformation

$$\frac{\delta(2\mathcal{E}_\varepsilon - a\int_N ds_\varepsilon)}{\delta\varepsilon(s)} = k^3 + 2\partial_s^2 k + 2ak = 0$$

$$ak + \frac{1}{2}k^3 + \partial_s^2 k = 0$$

$$(\partial_s k)^2 + \frac{1}{4}k^4 + ak^2 + b = 0$$

Infinitesimal isometric deformation

$$(\partial_s k)^2 + \frac{1}{4}k^4 + ak^2 + b = 0$$

$$x(s) := \frac{\mathrm{i}}{4a}\partial_s k + \frac{1}{8}k^2 + \frac{1}{12}a$$

$$y(s) := \partial_s x$$

$$y^2 = (x - e_1)(x - e_2)(x - e_3)$$

Elliptic curve

Elliptic curve

$$y^2 = (x - e_1)(x - e_2)(x - e_3)$$

$$e_1 = -\frac{1}{6}a \qquad a^2 - b = 16$$

$$e_2 = \frac{1}{12}a + \frac{1}{4}\sqrt{b}$$

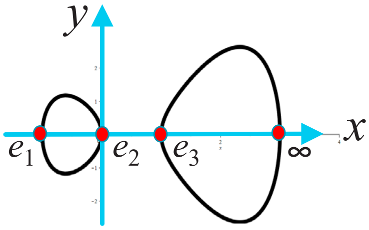
$$e_3 = \frac{1}{12}a - \frac{1}{4}\sqrt{b}$$

$$a = 2(e_2 + e_3 - 2e_1)$$

$$b = -(e_2 - e_3)^2$$

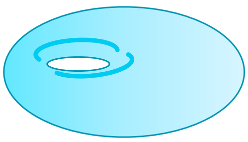
Elliptic curve

$$y^2 = (x - e_1)(x - e_2)(x - e_3)$$



x,y are regarded as complex numbers

$$x \to x + x'\sqrt{-1}, \quad y \to y + y'\sqrt{-1}$$



Meromorphic function over elliptic curve
= Elliptic function

Elastica is expressed by the elliptic functions!

$$\partial_s Z(s) = \sqrt{-1}(\wp(s + u_0) - e_1)$$

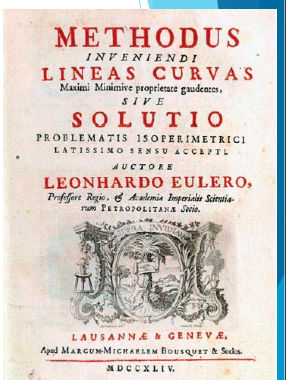
$$Z(s) = \sqrt{-1}(-\zeta(s + u_0) - e_1 s) + Z_0$$

Elastica (Elastic curve):

Euler's solutions

$$s = \int^X \frac{a^2 dX}{\sqrt{a^4 - (\alpha + \beta X + \gamma X^2)^2}}$$

$$Y = \int^X \frac{(\alpha + \beta X + \gamma X^2) dX}{\sqrt{a^4 - (\alpha + \beta X + \gamma X^2)^2}}$$

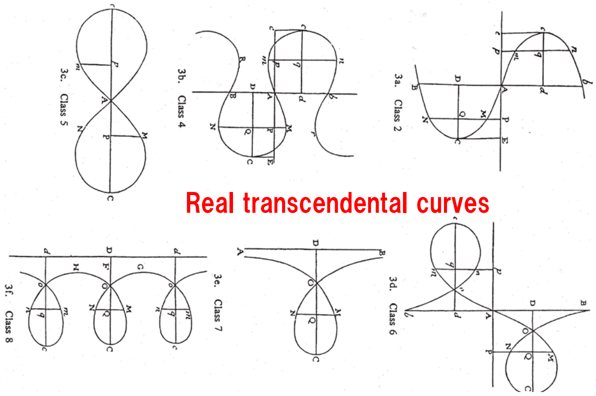


METHODUS
INVENIENDI
LINEAS CURVAS
Maximi Minimive proprietate gaudentes,
SIVE
SOLUTIO
PROBLEMATIS ISOPERIMETRICI
LATISSIMO SENSU ACCEPTI
AUCTORE
LEONHARDO EULERO,
Professore Regio, & Academiae Imperialis Scientiarum PETROPOLITANAE Socio.
LAUSANNAE & GENEVAE,
Apud MARCUM-MICHAELEM BOUSQUET & Socios.
MDCCXLIV.

It is a solution of MKdV equation from a modern point of view.

$$ak + \frac{1}{2}k^3 + \partial_s^2 k = 0$$

Elastica (Elastic curve):



Euler's sketch by numerical comp.(1744)

Menu

Statistical mechanics of elastica

$$\mathcal{Z}[\beta] = \int_{\mathbb{M}} DZ \exp(-\beta \mathcal{E}[Z])$$

$$\mathcal{M}_{S^1} := \{Z : S^1 \hookrightarrow \mathbb{C} \mid Z \in C^{\omega}(S^1, \mathbb{C}), |dZ/ds| = 1\},$$

$\mathrm{pr}_1 : \mathcal{M}_{S^1} \to \mathbb{M} := \mathcal{M}_{S^1}/\sim,\quad \sim$: eulidean move

The geometric structure of the parameter space (moduli) of a shape is unknown:

Find orbits with iso-energy.

E.Previato2015 SM 1997

Menu

1. Self Introduction
2. Continuation of Self-Introduction
3. Supercoiled structure of DNA
4. Elastic curves
5. Statistical Mechanics of Elastic Curves
6. **Excited states of elastic curves and the MKdV equation**
7. MKdV hyperelliptic curve solution for genus 2
   - 7.1 Review of the case of genus 1
   - 7.2 Review of the case of genus 2
8. Future task

Formulation of iso-energetic geometry

$$\mathcal{Z}[\beta] = \int_0^\infty \mathrm{Vol}(\mathcal{M}_E)\, \mathrm{e}^{-\beta E}\, dE$$

$$\mathcal{M}_E := \left\{ Z : S^1 \hookrightarrow \mathbb{C} \;\middle|\; \begin{array}{l} \textbf{Analytic, isometric} \\ E = \mathcal{E}[Z] \end{array} \right\}$$

$$\mathcal{E}[Z] = \frac{1}{2} \oint k(s)^2\, ds$$

Find orbits with the iso-energy.

E.Previato2015 SM 1997

MKdV equation

$$\partial_t k + \frac{3}{2} k^2 \partial_s k + \partial_s^3 k = 0$$

Solutions of the MKdV equation preserve

$\mathcal{E}[Z] = \frac{1}{2} \oint k(s)^2\, ds$ **for the time-development**

$$\frac{\partial \mathcal{E}}{\partial t} = \int k \partial_t k\, ds = -\int \partial_s \left(\frac{3}{8} k^4 - \frac{1}{2} (\partial_s k)^2 \right) ds$$

The time t is not physical time but a parameter in the moduli of the immersion of the curve

MKdV equation

$$\partial_t k + \frac{3}{2}k^2\partial_s k + \partial_s^3 k = 0$$

MKdV equation contains the static MKdV equation of elastica as t=s.
$\Leftrightarrow$ **It is a natural generalization of elastica**

$$ak + \frac{1}{2}k^3 + \partial_s^2 k = 0$$

Statistical mechanics of elastica

$$\mathcal{Z}[\beta] = \int_{\mathbb{M}} DZ \exp(-\beta\mathcal{E}[Z])$$

$\mathcal{M}_{S^1} := \{Z : S^1 \hookrightarrow \mathbb{C} \mid Z \in \mathcal{C}^\omega(S^1, \mathbb{C}), |dZ/ds| = 1\},$

$\mathrm{pr}_1 : \mathcal{M}_{S^1} \to \mathbb{M} := \mathcal{M}_{S^1}/\sim, \quad \sim$: eulidean move

Find orbits with iso-energy

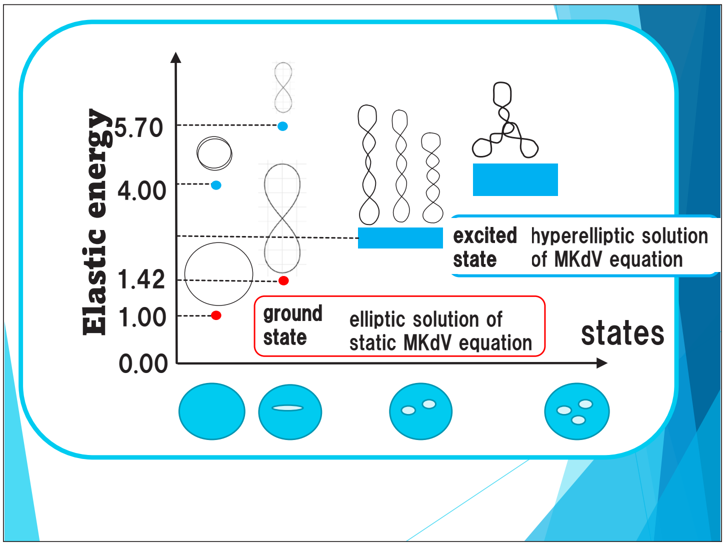
$\Leftrightarrow$ **Find higher-order solutions of the MKdV equation!**

$\partial_t k + \frac{3}{2}k^2\partial_s k + \partial_s^3 k = 0$

MKdV equation

$$\partial_t k + \frac{3}{2}k^2\partial_s k + \partial_s^3 k = 0$$

1. The MKdV equation has hyperelliptic function solutions.
2. Due to the higher genus of hyperelliptic curves (compact Riemann surfaces), the solutions are expected to express more complicated (elastic) curves.



Statistical mechanics of elastica

Assign the appropriate topology in the parameter space of the geometry (moduli),formulate the above integral in terms of the measures determined from the Boltzmann weights of the Euler-Bernoulli energy functional, and perform the integration.

1. **Construct hyperelliptic solutions to the MKdV equation of higher genus.**
2. **Extract "real" part of hyperelliptic Jacobi variety as the moduli of "real" hyperelliptic curves over C.**

M 1997, M-Onishi 2001, M-Previato 2015

Statistical mechanics of elastica

Assign the appropriate topology in the parameter space of the geometry (moduli),formulate the above integral in term Boltzm functio

1. **To Construct solutions to the MKdV equation**
2.

If hyperelliptic function theory had the same level of sophistication and concreteness as Weierstrass' elliptic function theory, this problem would be solved! but it is not at that level.

Reconstruct the theory of hyperelliptic (Abelian) functions to have the same level as the theory of elliptic functions.

M

The main theme is a reconstruction of Abelian function theory with E.Previato

Menu

1. Self Introduction
2. Continuation of Self-Introduction
3. Supercoiled structure of DNA
4. Elastic curves
5. Statistical Mechanics of Elastic Curves
6. Excited states of elastic curves and the MKdV equation
7. **MKdV hyperelliptic curve solution for genus 2**
   - **7.1 Review of the case of genus 1**
   - **7.2 Review of the case of genus 2**
8. Future task

MKdV solutions of genus two

- **Review the genus one case**
- **Step to genus two**

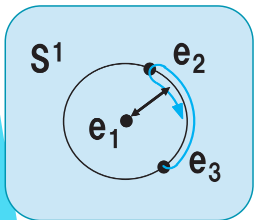
S.M., and Emma Previato,
An algebro-geometric model for the shape of supercoiled DNA
Physica D, 2022

Review the genus one case:

$$y^2 = (x-e_1)(x-e_2)(x-e_3)$$

$$ds = \Re du \qquad u = \int_\infty^{(x,y)} \frac{dx}{2y}$$

$$\partial_s Z = (x-e_1) = \mathrm{e}^{\mathrm{i}\phi} \qquad |\partial_s Z| = 1$$

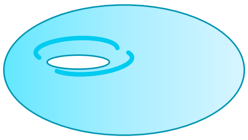


They satisfy the SMKdV eq.

$$\boxed{\partial_s \phi + \frac{1}{8}(\partial_s \phi)^3 + \frac{1}{4}\partial_s^3 \phi = 0}$$

$$\boxed{ak + \frac{1}{2}k^3 + \partial_s^2 k = 0.}$$

Review the genus one case:



Meromorphic function over elliptic curve
=Elliptic function

Elastica is expressed by the elliptic functions!

$$\partial_s Z(s) = \sqrt{-1}(\wp(s+u_0) - e_1)$$

$$Z(s) = \sqrt{-1}(-\zeta(s+u_0) - e_1 s) + Z_0$$

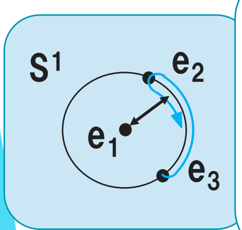
$$\zeta(u) = \frac{d}{du}\log\sigma(u), \qquad \wp(u) = -\frac{d^2}{du^2}\log\sigma(u)$$

Review the genus one case:

$$y^2 = (x-e_1)(x-e_2)(x-e_3)$$

$$ds = \Re du \qquad u = \int_\infty^{(x,y)} \frac{dx}{2y}$$

$$\partial_s Z = (x-e_1) = \mathrm{e}^{\mathrm{i}\phi} \qquad |\partial_s Z| = 1$$



$$\phi = 2\varphi \qquad e_{ab} := e_a - e_b$$

$$k := \frac{2\mathrm{i}\sqrt[4]{e_{ba}e_{ca}}}{\sqrt{e_{ba}} - \sqrt{e_{ca}}}$$

$$du = \frac{2k\,d\varphi}{\sqrt{1-k^2\sin^2\varphi}}$$

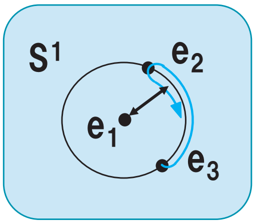
$$a=1,\ b=2,\ c=3$$

Review the genus one case:

$$y^2 = (x-e_1)(x-e_2)(x-e_3)$$

$$ds = \Re du \qquad u = \int_\infty^{(x,y)} \frac{dx}{2y}$$

$$\partial_s Z = (x-e_1) = \mathrm{e}^{\mathrm{i}\phi} \qquad |\partial_s Z| = 1$$



Num.Comp.

$$x := \mathrm{e}^{\mathrm{i}\phi} + e_1$$

$$\delta s := \frac{(\partial x/\partial\phi)\delta\phi}{2y}$$

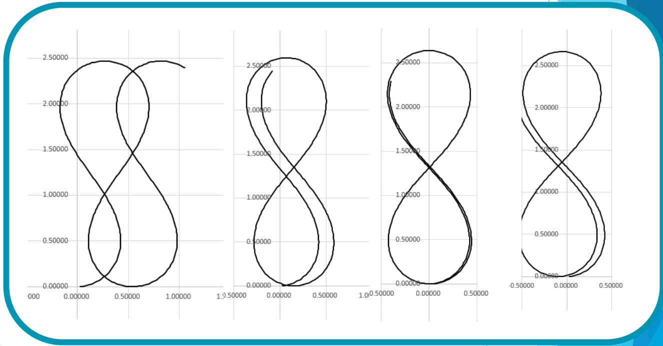
$$s := s + \delta s$$

$$\phi := \phi + \delta\phi$$

$$Z := Z + (x - e_1)\delta s$$

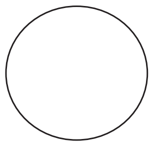
Euler method

a=1.2 a=1.25

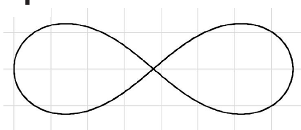


Closed elasticae

g=0



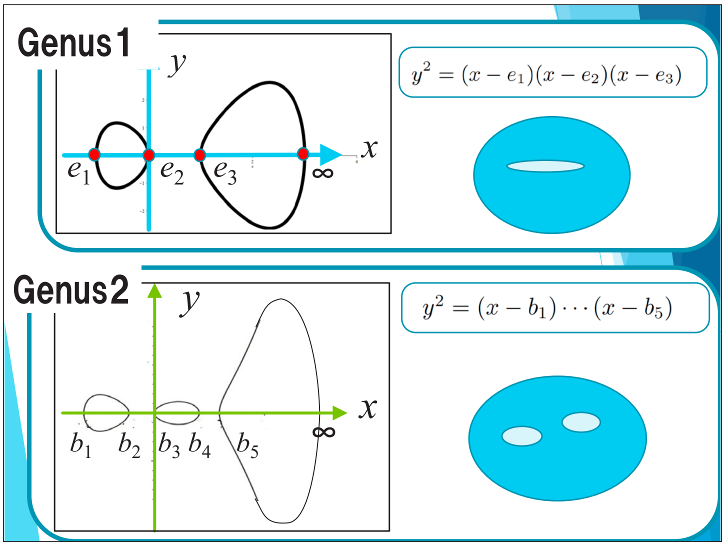
g=1

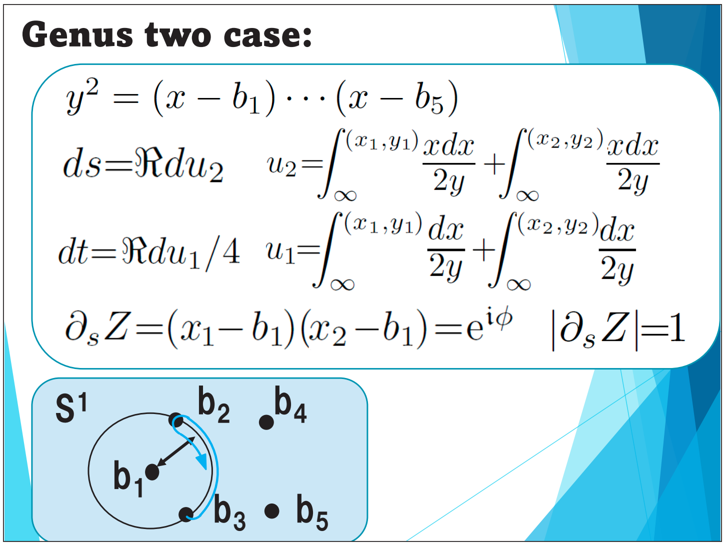


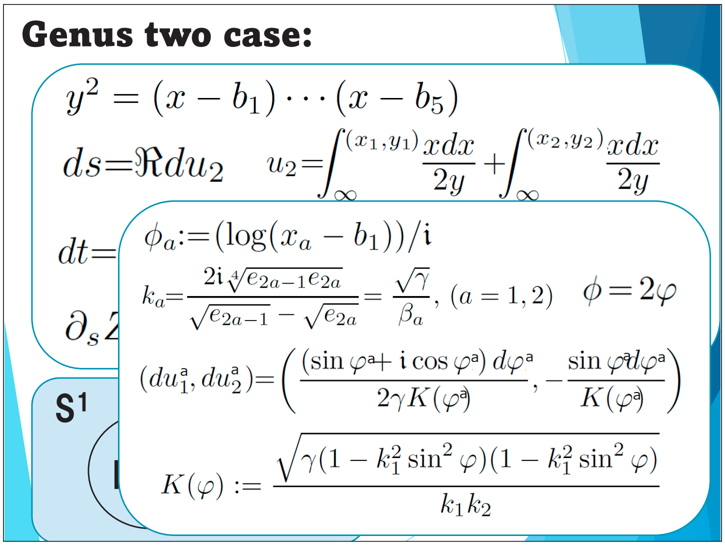
MKdV solutions of genus two

- Review the genus one case
- **Step to genus two**

S.M., and Emma Previato,
An algebro-geometric model for the shape of supercoiled DNA
Physica D, 2022


Genus 1
$y^2=(x-e_1)(x-e_2)(x-e_3)$
y
x
e_1 e_2 e_3 ∞
Genus 2
$y^2=(x-b_1)\cdots(x-b_5)$
y
x
b_1 b_2 b_3 b_4 b_5 ∞


Genus two case:
$y^2=(x-b_1)\cdots(x-b_5)$
$ds=\Re du_2 \quad u_2=\int_\infty^{(x_1,y_1)}\frac{xdx}{2y}+\int_\infty^{(x_2,y_2)}\frac{xdx}{2y}$
$dt=\Re du_1/4 \quad u_1=\int_\infty^{(x_1,y_1)}\frac{dx}{2y}+\int_\infty^{(x_2,y_2)}\frac{dx}{2y}$
$\partial_s Z=(x_1-b_1)(x_2-b_1)=e^{i\phi} \quad |\partial_s Z|=1$
S^1
b_1 b_2 b_3 b_4 b_5


Genus two case:
$y^2=(x-b_1)\cdots(x-b_5)$
$ds=\Re du_2 \quad u_2=\int_\infty^{(x_1,y_1)}\frac{xdx}{2y}+\int_\infty^{(x_2,y_2)}\frac{xdx}{2y}$
$\phi_a:=(\log(x_a-b_1))/i$
$k_a=\frac{2i\sqrt[4]{e_{2a-1}e_{2a}}}{\sqrt{e_{2a-1}}-\sqrt{e_{2a}}}=\frac{\sqrt{\gamma}}{\beta_a},\ (a=1,2) \quad \phi=2\varphi$
$(du_1^a, du_2^a)=\left(\frac{(\sin\varphi^a+i\cos\varphi^a)\,d\varphi^a}{2\gamma K(\varphi^a)}, -\frac{\sin\varphi^a d\varphi^a}{K(\varphi^a)}\right)$
$K(\varphi):=\frac{\sqrt{\gamma(1-k_1^2\sin^2\varphi)(1-k_1^2\sin^2\varphi)}}{k_1k_2}$
S^1

Genus two case:

$$y^2 = (x-b_1)\cdots(x-b_5)$$

$$ds = \Re du_2 \quad u_2 = \int_\infty^{(x_1,y_1)} \frac{x dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{x dx}{2y}$$

$$dt = \Re du_1/4 \quad u_1 = \int_\infty^{(x_1,y_1)} \frac{dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{dx}{2y}$$

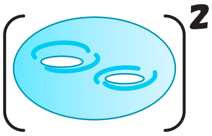
$$\partial_s Z = (x_1 - b_1)(x_2 - b_1) = e^{i\phi} \quad |\partial_s Z| = 1$$

MKdV equation/C

$$(4\partial_{u_1} - a\partial_{u_2})\phi + \frac{1}{2}(\partial_{u_2}\phi)^3 + \partial_{u_2}^3\phi = 0$$

$$a := \textstyle\sum_{i=2}^5 b_i - 2b_1$$

Genus two case:



Meromorphic function over hyperelliptic curve
=Hyperelliptic function

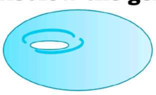
Shape of elastica is determined by hyperelliptic functions

$$\partial_{u_2} Z = b_1^2 - \wp_{22} b_1 + \wp_{21},$$

$$Z = b_1^2 u_2 + \zeta_2 b_1 - \zeta_1 + Z_0$$

$$\zeta_i(u) := \frac{\partial}{\partial u_i} \log\sigma(u), \quad \wp_{ij}(u) := -\frac{\partial^2}{\partial u_i \partial u_j} \log\sigma(u)$$

Review the genus one case:



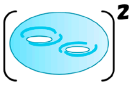
Meromorphic function over elliptic curve
=Elliptic function

Elastica is expressed by the elliptic functions!

$$\partial_s Z(s) = \sqrt{-1}(\wp(s+u_0) - e_1)$$

$$Z(s) = \sqrt{-1}(-\zeta(s+u_0)$$

$$\zeta(u) = \frac{d}{du}\log\sigma(u), \qquad \wp(u) =$$



Meromorphic function over hyperelliptic curve
=Hyperelliptic function

Shape of elastica is determined by hyperelliptic functions

$$\partial_{u_2} Z = b_1^2 - \wp_{22} b_1 + \wp_{21},$$

$$Z = b_1^2 u_2 + \zeta_2 b_1 - \zeta_1 + Z_0$$

$$\zeta_i(u) := \frac{\partial}{\partial u_i} \log\sigma(u), \quad \wp_{ij}(u) := -\frac{\partial^2}{\partial u_i \partial u_j} \log\sigma(u)$$

It is a direct generalisation.

Genus two case:

MKdV equation/C

$$(4\partial_{u_1} - a\partial_{u_2})\phi + \frac{1}{2}(\partial_{u_2}\phi)^3 + \partial_{u_2}^3\phi = 0$$

$$a := \sum_{i=2}^{5} b_i - 2b_1$$

$$ds = \Re du_2 \qquad \phi_{\mathrm{r}} := \Re\phi$$

$$dt = \Re du_1/4 \qquad \phi_{\mathrm{i}} := \Im\phi$$

MKdV equation with gauge field/R

$$(\partial_t - A\partial_s)\phi_{\mathrm{r}} + \frac{1}{2}(\partial_s\phi_{\mathrm{r}})^3 + \partial_s^3\phi_{\mathrm{r}} = 0$$

$$A := a - 3(\partial_s\phi_{\mathrm{i}})^2/2$$

Genus two case:

MKdV equation with gauge field/R

$$(\partial_t - A\partial_s)\phi_{\mathrm{r}} + \frac{1}{2}(\partial_s\phi_{\mathrm{r}})^3 + \partial_s^3\phi_{\mathrm{r}} = 0$$

$$A := a - 3(\partial_s\phi_{\mathrm{i}})^2/2$$

$$\Im du_1 = \Im du_2 = 0$$

$$d\phi_{\mathrm{i}} = 0 \quad d\partial_s\phi_{\mathrm{i}} = 0$$

MKdV equation/R

Genus two case:

MKdV equation with gauge field/R

$$(\partial_t - A\partial_s)\phi_{\mathrm{r}} + \frac{1}{2}(\partial_s\phi_{\mathrm{r}})^3 + \partial_s^3\phi_{\mathrm{r}} = 0$$

$$A := a - 3(\partial_s\phi_{\mathrm{i}})^2/2$$

$$\Im du_1 = \Im du_2 = 0$$

$$d\phi_{\mathrm{i}} = 0 \quad d\partial_s\phi_{\mathrm{i}} = 0$$

MKdV equation/R

4=Dim. of parameter space of hyperelliptc curve g=2 $\longleftrightarrow$ 4 conditions $\longleftrightarrow$ Dim of sol. of MKdV (t&s) =2

Genus two case:

MKdV equation with gauge field/R

$$(\partial_t - A\partial_s)\phi_r + \frac{1}{2}(\partial_s\phi_r)^3 + \partial_s^3\phi_r = 0$$

$$A := a - 3(\partial_s\phi_i)^2/2$$

$$\Im du_1 = \Im du_2 = 0$$

$$d\phi_i = 0 \quad d\partial_s\phi_i = 0$$

MKdV equation/R

4=Dim. of parameter space of hyperelliptc curve g=2

ill-posed

Dim of sol. of MKdV (t&s) =2

Genus two case:

MKdV equation/R

$$\Im du_1 = \Im du_2 = 0$$

$$d\phi_i = 0$$

4=Dim. of parameter space of hyperelliptc curve g=2

4 conditions

ill-posed

Dim of sol. of MKdV (t&s) =2

Need to adjust the parameters (b_i) of the curve itself to find the situation where the conditions are degenerate
⇒ Extremely difficult
(only possible with more than genus 3?)

First, loosen the conditions, Investigate the properties of g = 2

Genus two case:

MKdV equation/R

$$\Im du_1 = \Im du_2 = 0$$

$$d\phi_i = 0$$

We conclude that in this stage, the hyperelliptic curves X with genus two cannot exhibit the generalized elastica well because we cannot extract the real parts in both X and its Jacobi variety J_X over $\mathbb{C}$.
(S.M. ,E. Previato, Physica D, 2022)

(only possible with more than genus 3?)

First, loosen the conditions, Investigate the properties of g = 2

Genus two case:

MKdV equation/C

3 conditions

$$\phi_a := (\log(x_a - b_1))/\mathrm{i}$$

$$\Im(d\phi_a) = 0 \qquad a=1.2$$

$$d\partial_s \phi_\mathrm{i} = 0$$

$$d\partial_s\phi_\mathrm{i} = \frac{\partial(\partial_s\phi_\mathrm{i})}{\partial\phi_{1\mathrm{r}}} d\phi_{1\mathrm{r}} + \frac{\partial(\partial_s\phi_\mathrm{i})}{\partial\phi_{2\mathrm{r}}} d\phi_{2\mathrm{r}} = 0$$

$$d\phi_{1\mathrm{r}} = \frac{\partial(\partial_s\phi_\mathrm{i})}{\partial\phi_{2\mathrm{r}}} d\eta, \qquad d\phi_{2\mathrm{r}} = -\frac{\partial(\partial_s\phi_\mathrm{i})}{\partial\phi_{1\mathrm{r}}} d\eta$$

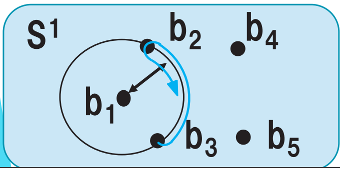
Genus two case:

$$y^2 = (x-b_1)\cdots(x-b_5)$$

$$ds = \Re du_2 \qquad u_2 = \int_\infty^{(x_1,y_1)} \frac{x dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{x dx}{2y}$$

$$dt = \Re du_1/4 \qquad u_1 = \int_\infty^{(x_1,y_1)} \frac{dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{dx}{2y}$$

$$\partial_s Z = (x_1 - b_1)(x_2 - b_1) = \mathrm{e}^{\mathrm{i}\phi} \qquad |\partial_s Z| = 1$$



Genus two case:

$$y^2 = (x-b_1)\cdots(x-b_5)$$

$$ds = \Re du_2 \qquad u_2 = \int_\infty^{(x_1,y_1)} \frac{x dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{x dx}{2y}$$

$dt=$

$\partial_s Z$

S¹

$$\phi_a := (\log(x_a - b_1))/\mathrm{i}$$

$$k_a = \frac{2\mathrm{i}\sqrt[4]{e_{2a-1}e_{2a}}}{\sqrt{e_{2a-1}} - \sqrt{e_{2a}}} = \frac{\sqrt{\gamma}}{\beta_a}, \ (a=1,2) \qquad \phi = 2\varphi$$

$$(du_1^a, du_2^a) = \left(\frac{(\sin\varphi^a + \mathrm{i}\cos\varphi^a)\, d\varphi^a}{2\gamma K(\varphi^a)}, -\frac{\sin\varphi^a d\varphi^a}{K(\varphi^a)} \right)$$

$$K(\varphi) := \frac{\sqrt{\gamma(1-k_1^2\sin^2\varphi)(1-k_1^2\sin^2\varphi)}}{k_1 k_2}$$

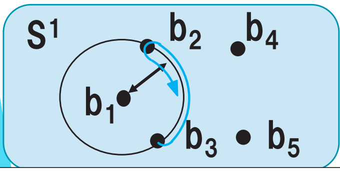
Genus two case:

$$y^2 = (x-b_1)\cdots(x-b_5)$$

$$ds = \Re du_2 \qquad u_2 = \int_\infty^{(x_1,y_1)} \frac{x dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{x dx}{2y}$$

$$dt = \Re du_1/4 \qquad u_1 = \int_\infty^{(x_1,y_1)} \frac{dx}{2y} + \int_\infty^{(x_2,y_2)} \frac{dx}{2y}$$

$$\partial_s Z = (x_1 - b_1)(x_2 - b_1) = \mathrm{e}^{\mathrm{i}\phi} \qquad |\partial_s Z| = 1$$



Genus two case:

$$y^2 = (x-b_1)\cdots(x-b_5)$$

$ds = \Re d$

$dt = \Re du$

$\partial_s Z = ($

Num.Comp.

$$x_a := \mathrm{e}^{\mathrm{i}\phi_{ar}} + b_1, \quad a = 1, 2$$

$$\delta\phi_{ar} := \varepsilon_{ab} \frac{\partial(\partial_s \phi_i)}{\partial \phi_{br}} \delta\eta$$

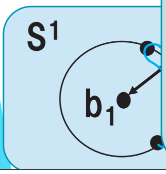
$$\delta s := \sum_a \frac{x(\partial x_a/\partial \phi_a)\delta\phi_{ar}}{2y_a}$$

$$s := s + \delta s$$

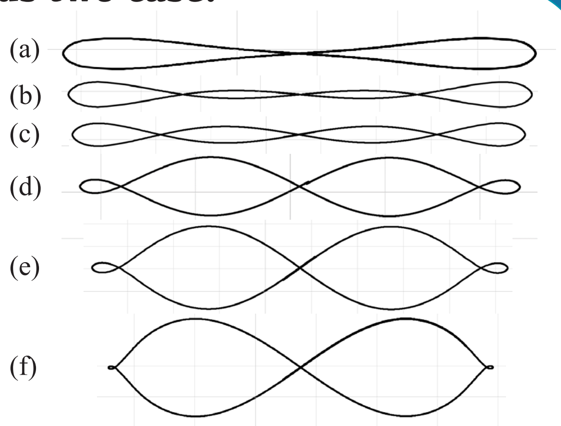
$$\phi_{ar} := \phi_{ar} + \delta\phi_{ar}$$

$$Z := Z + (x_1 - b_1)(x_2 - b_1)\delta s$$

Euler method



Genus two case:

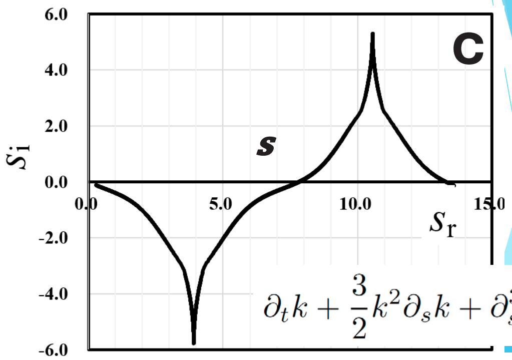


They are solutions of MKdV equation over C but not over R

M. Previato Physica D, 2022

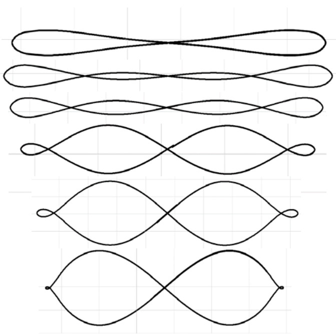
Genus two case:

The orbit s on the complex plane



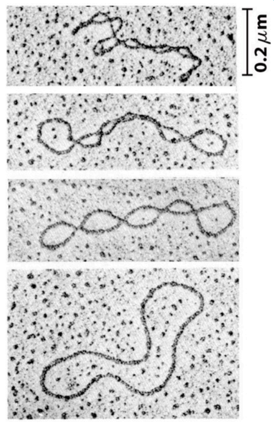
$$\partial_t k + \frac{3}{2} k^2 \partial_s k + \partial_s^3 k = 0$$

M. Previato
Physica D, 2022



$g = 2$

S.Matsutani E.Previato, 2021



Electron Micrographs of DNA

Fundamentals of Biochemistry: Life at the Molecular Level, 3rd Ed. D. Voet, J. Voet, C. Pratt

Future tasks:

1. **to investigate the cases of g > 3.**
2. **to evaluate the moduli space of generalized elastica analytically and numerically**
3. **to extend them to a generalized elasitca in R^3 by finding the hyperelliptic solutions of non-linear Schrodinger equation in our novel approach with Emma Previato.**

Thank you!

Information geometry of positive measures

Naomichi Nakajima
Faculty of Science and Engineering, Waseda University, Japan

Information geometry brings a united geometric insight into various aspects of statistical science, machine learning and so on by regarding the parameter space of a statistical model as a Riemannian manifold equipped with the Fisher-Rao metric. The dually flat structure on a Riemannian manifold introduced by Amari-Nagaoka takes a central role in information geometry. It is known that the space of probability distributions on a finite set naturally has the dually flat structure. For this space, Amari has characterized the dually flat structure from the viewpoint of statistics through defining the space of positive measures simply by removing the normalization condition. On the other hand, we have developed the counterpart for the space of transition probabilities of a given Markov chain, which may provide a new geometric insight into Markov chains. In this presentation, I would like to talk about Amari's theory and our theory for Markov chains.

Information geometry of positive measures

Naomichi Nakajima
Waseda University, Japan

Sep. 29th, 2023
WORKSHOP on Mathematics for Industry
@Warsaw University of Technology

supp.: JSPS KAKENHI Grant Number 22KJ0052, 22KK0034

1 / 23

Summary of my talk

- **Information geometry** brings a united geometric insight on various fields such as statistics, machine learning, optimization theory and so on. In information geometry, a statistical model is regarded as a Riemannian manifold endowed with *the Fisher-Rao metric* and *two kinds of affine connections* satisfying a certain duality, called **a statistical manifold**.
- **A dually flat manifold** is a statistical mfd with flat connections, that takes a central role in information geometry, introduced by Amari-Nagaoka.
- Regarding dually flat structures, there is a well established theory of **positive measures** on a finite set S due to Amari. It investigates dually flat structures of the space $\mathcal{P}(S)$ of probability distributions on S in terms of some "asymmetric distance function" on $\mathcal{P}(S)$, called a divergence.

2 / 23

Summary of my talk

- On the other hand, information geometry of Markov chains has been studied by Nagaoka and others using the dually flat structure of the space of transition probabilities.
- In comparison with information geometry of $\mathcal{P}(S)$, roughly speaking, the studies above are on information geometry of the space of conditional probabilities.

▶ Main topic of my talk.
We will investigate the counterpart for a Markov chain of Amari's theory of positive measures. This study does not only investigate information geometry of the specific model, a Markov chain, but also suggests a new direction of statistics of conditional probabilities.

3 / 23

Contents

- Backgrounds
  - Statistical manifolds, dually flat manifolds and divergences
  - Amari's theory of positive measures on a finite set [Amari]
  - Information geometry of transition probabilities of a given Markov chain [Nagaoka]
- Our theory for transition probabilities [N]

[N] *The space of positive transition measures of a Markov model*, in preparation.

[Amari] S. Amari, *α-divergence is unique, belonging to both f-divergence and Bregman divergence classes*, IEEE Trans. Inform. Theory **55** (2009), 4925–4931.

[Nagaoka] H. Nagaoka, *The exponential family of Markov chains and its information geometry*, Proceedings of The 28th Symposium on Information Theory and Its Applications (SITA2005) (2005).

Statistical manifolds, dually flat manifolds and divergences

Let (M, h) be a pseudo-Riem. mfd and ∇ a torsion-free affine connection of TM.

- The triplet (M, h, ∇) is a **statistical mfd** if the **cubic tensor** $C := \nabla h$ is totally symmetric. Then C is called the Amari-Chentsov tensor [3, 4].
- **The dual connection** ∇^* of ∇ w.r.t. h is defined by

$$Xh(Y,Z) = h(\nabla_X Y, Z) + h(Y, \nabla^*_X Z) \quad (X, Y, Z \in \mathfrak{X}(M))$$

Also, an "asymmetric distance" $\rho : M \times M \to \mathbb{R}$ induces (h, ∇, ∇^*) on M as follows. For vector fields $X_1, \cdots, X_k, Y_1, \cdots, Y_l$ on M, define the function

$$\rho[X_1 \cdots X_k | Y_1 \cdots Y_l] : M \to \mathbb{R},$$
$$\rho[X_1 \cdots X_k | Y_1 \cdots Y_l](r) = (X_1)_p \cdots (X_k)_p (Y_1)_q \cdots (Y_l)_q (\rho(p,q))|_{p=q=r}.$$

Statistical manifolds, dually flat manifolds and divergences

We call ρ a contrast function if it satisfies

$$\begin{cases} \text{(i)} & \rho[-|-](r) = \rho(r,r) = 0, \\ \text{(ii)} & \rho[X|-](r) = \rho[-|X](r) = 0, \\ \text{(iii)} & -\rho[X|Y] : \text{pseudo-Riemannian metric on } M. \end{cases}$$

We call ρ a weak contrast function if it satisfies only (i) and (ii) ([N.-Ohmoto2021]).

$$\begin{cases} h(X,Y) := -\rho[X|Y], \\ h(\nabla_X Y, Z) = \rho[XY|Z], \quad h(\nabla^*_X Y, Z) = \rho[Z|XY]. \end{cases}$$

$\rho \xrightarrow{\text{[Eguchi1992]}} (M, h, \nabla, \nabla^*)$

$\rho \in$ Contrast Functions — multi : 1 — Statistical Manifolds $\ni (M, h, \nabla, \nabla^*)$

$\exists \rho \xleftarrow[\text{restores } (h,\nabla,\nabla^*)]{\text{[Matumoto1993]}} (M, h, \nabla, \nabla^*)$

Statistical manifolds, dually flat manifolds and divergences

For a statistical mfd (M, h, ∇), ∇ is flat $\iff$ its dual connection ∇^* is flat.

Definition (Amari-Nagaoka [3, 4])

A statistical mfd (M, h, ∇, ∇^*) is a **dually flat mfd** if ∇ is flat. Then we also call (h, ∇, ∇^*) the dually flat structure of M.

We write $\theta = (\theta_1, \cdots, \theta_n)$ for ∇-affine coords. Put $\partial_i := \frac{\partial}{\partial \theta_i}$. Then there exists a potential function $f(\theta)$ on θ s.t.

1. the metric h is locally given by the Hessian matrix of $f(\theta)$: $h(\partial_i, \partial_j) = \partial_i \partial_j f$,
2. the gradient map $\eta = (\eta_1, \cdots, \eta_n)$ $(\eta_i := \frac{\partial f}{\partial \theta_i})$ gives ∇^*-affine coordinates, called the dual coordinates of θ,

Another definition (Hessian structure [Shima]):
Given a $(M, h, (\nabla, \theta), f(\theta))$ with $h = \partial_i \partial_j f$
$\rightsquigarrow$ define the dual flat connection and the dual coord $(\nabla^*, \eta = (\eta_i))$ by $\eta_i := \frac{\partial f}{\partial \theta_i}$

Statistical manifolds, dually flat manifolds and divergences

A dually flat mfd (M, h, ∇, ∇^*) has the canonical contrast function $\mathcal{D} : M \times M \to \mathbb{R}$, called **the Bregman divergence**:

$$\mathcal{D}(p, q) = f(\theta(p)) - f(\theta(q)) + \frac{\partial f}{\partial \theta}(\theta(q))^T(\theta(q) - \theta(p)) \ \ (p, q \in M),$$

where $f(\theta)$ is a potential function of M. (strictly speaking, $\mathcal{D}$ is defined on an open neighborhood of the diagonal set of M)

Remark:

- The definition of $\mathcal{D}$ is independent of the choice of $(\theta, \ f(\theta))$.
- $\mathcal{D}$ restores the dually flat structure (h, ∇, ∇^*), i.e.,

$$\begin{cases} h(X, Y) = \mathcal{D}[X|Y], \\ h(\nabla_X Y, Z) = -\mathcal{D}[XY|Z], \ \ h(\nabla^*_X Y, Z) = -\mathcal{D}[Z|XY]. \end{cases}$$

Example: the space of discrete distributions

- $S = \{0, 1, \cdots, n\}$: a finite set
- $\mathcal{P}(S) = \{(p_0, p_1, \cdots, p_n) \in \mathbb{R}^{n+1} \mid p_i > 0$ and $\sum_{i=0}^n p_i = 1\}$

We call $\mathcal{P}(S)$ the space of discrete distributions on S. Take a system of coordinates $(p_1, \cdots, p_n)$ $(p_0 = 1 - p_1 - \cdots p_n)$. We regard it as flat coordinates $(\nabla, \eta = (\eta_i)_{i=1}^n)$:

$$\eta_i := p_i \ (\textbf{the expectation parameters}).$$

Then

$$\varphi(\eta) = \textstyle\sum_{i=1}^n p_i \log p_i$$

is a convex function, known as the negative entropy in statistics.

Hence the metric h is defined by

$$h(\tfrac{\partial}{\partial \eta_i}, \tfrac{\partial}{\partial \eta_j}) = \tfrac{\partial^2 \varphi}{\partial \eta_i \partial \eta_j}.$$

Therefore, $(\mathcal{P}(S), h, (\nabla, \eta), \varphi(\eta))$ is a dually flat mfd (Hessian mfd).

Example: the space of discrete distributions

Importantly, the Bregman divergence $\mathcal{D} : \mathcal{P}(S) \times \mathcal{P}(S) \to \mathbb{R}$ induced by φ is **the KL-divergence** on $\mathcal{P}(S)$, i.e.,

$$\mathcal{D}(p, q) = \sum_{i=0}^{n} p_i \log \frac{p_i}{q_i} =: KL[p, q],$$

where $p = (p_0, \cdots, p_n), \ \ q = (q_0, \cdots, q_n) \in \mathcal{P}(S)$.

- We consider the following problem: are there any other contrast functions to derive a dually flat structure of $\mathcal{P}(S)$?
- Of course, for example, we consider a quadratic function as a potential function, and then it derives another dually flat structure of $\mathcal{P}(S)$.
- We are interested in the dually flat structure with "statistical invariance", which is a certain condition required from statistics.

The space of positive measures on a finite set

- Amari has introduced the space $\bar{\mathcal{P}}(S)$ of positive measures on S as an extended space of $\mathcal{P}(S)$ and investigated the problem above by finding the Bregman and **F-divergence** on $\bar{\mathcal{P}}(S)$ suitably.
- An F-divergence $\mathcal{D}_F$ on $\bar{\mathcal{P}}(S)$ is a contrast function, and it is known that the statistical manifold structure induced by $\mathcal{D}_F$ of $\bar{\mathcal{P}}(S)$ satisfies statistical invariance.
- In [Amari], Amari has shown that the KL-divergence $\mathcal{D}_{KL}$ on $\bar{\mathcal{P}}(S)$ is the only contrast function such that
  - it is both a Bregman divergence and an F-divergence,
  - it and its restriction to $\mathcal{P}(S)$ induce the dually flat structures of $\bar{\mathcal{P}}(S)$ and $\mathcal{P}(S)$, respectively.

The space of positive measures on a finite set and F-divergences

- $S = \{0, 1, \cdots, n\}$: a finite set
- $\bar{\mathcal{P}}(S) = \{(p_0, p_1, \cdots, p_n) \in \mathbb{R}^{n+1} \mid p_i > 0\} \supset \mathcal{P}(S) = \{p_i > 0 \text{ and } \sum_{i=0}^{n} p_i = 1\}$

We call $\bar{\mathcal{P}}(S)$ the space of positive measures on S.
Given a strictly convex function $F : (0, \infty) \to \mathbb{R}$ with

$$F(1) = F'(1) = 0 \text{ and } F''(1) = 1,$$

called a standard convex function [Amari], the function $\mathcal{D}_F : \bar{\mathcal{P}}(S) \times \bar{\mathcal{P}}(S) \to \mathbb{R}$ defined by

$$\mathcal{D}_F(p, q) = \sum_{i=0}^{n} p_i F\left(\frac{q_i}{p_i}\right)$$

is called the F-divergence on $\bar{\mathcal{P}}(S)$, where $p = (p_0, \cdots, p_n), \ \ q = (q_0, \cdots, q_n)$.

The space of positive measures on a finite set and F-divergences

In the case where $F(t) = -\log t + (t-1)$, the F-divergence $\mathcal{D}_F$ is the KL-divergence on $\bar{\mathcal{P}}(S)$:

$$\mathcal{D}_F(p,q) = \sum_{i=0}^{n} p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=0}^{n} q_i - \sum_{i=0}^{n} p_i.$$

- In fact, $\bar{\mathcal{P}}(S)$ has the dually flat structure; its flat coordinates are $\eta = (p_0, \cdots, p_n)$ and the potential function $\varphi(\eta)$ is given by

$$\varphi(\eta) = \sum_{i=0}^{n} p_i \log p_i.$$

- For $p, q \in \mathcal{P}(S)$, it holds that $\sum_{i=0}^{n} p_i = \sum_{i=0}^{n} q_i = 1$, which yields

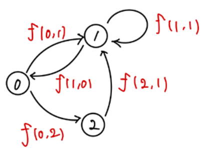
$$\mathcal{D}_F(p,q) = \sum_{i=0}^{n} p_i \log\left(\frac{p_i}{q_i}\right) = KL[p,q].$$

13 / 23

Information geometry of the space of transition probabilities

- Setting:
  - $\mathcal{X} = \{0, 1, \cdots, d\}$: a finite set
  - $\mathcal{E} \subset \mathcal{X} \times \mathcal{X}$: a subset
    $\rightsquigarrow$ We regard $(\mathcal{X}, \mathcal{E})$ as a direct graph.
  - $\mathcal{F}^+ = \{f : \mathcal{E} \to \mathbb{R} \mid f(x,y) > 0 \text{ for any } (x,y) \in \mathcal{E}\}$
  - $\mathcal{W} = \{w \in \mathcal{F}^+ \mid \sum_{y:(x,y)\in\mathcal{E}} w(x,y) = 1 \text{ for any } x \in \mathcal{E}\} \subset \mathcal{F}^+$

We call $w \in \mathcal{W}$ a transition probability on $\mathcal{E}$ (the word “transition probability” comes from Markov chains).



14 / 23

Information geometry of the space of transition probabilities

We assume that $\mathcal{E}$ is strongly connected, that is, for any $x, y \in \mathcal{X}$ there exist $(x_1, x_2), (x_2, x_3), \cdots, (x_{N-1}, x_N) \in \mathcal{E}$ such that $x_1 = x, x_N = y$ $(N \geq 2)$.
By this assumption, for every $f \in \mathcal{F}^+$ we can apply the Perron-Frobenius theorem to

$$A(f) = [a_{ij}(f)]_{0\leq i,j\leq d}, \quad a_{ij}(f) = \begin{cases} f(i,j) & (i,j) \in \mathcal{E} \\ 0 & (i,j) \notin \mathcal{E} \end{cases}$$

Then we get a unique real value $r(f) > 0$ and vector $\mu_f = (\mu_f(0), \cdots, \mu_f(d))^T$ satisfying

- $r(f)$ is the Perron-Frobenius root, which is an eigenvalue of $A(f)$,
- μ_f is a left eigenvector associated with $r(f)$ such that $\mu_f(i) > 0$ for any i, and $\sum_{i=0}^{d} \mu_f(i) = 1$.

We call the vector μ_f the stationary distribution for f.

15 / 23

Information geometry of the space of transition probabilities

We consider the following two spaces:

$$\overline{M} = \{\boldsymbol{\eta} = (\eta_{xy})_{(x,y)\in\mathcal{E}} \in \mathbb{R}^{|\mathcal{E}|} \mid \eta_{xy} > 0\},$$
$$M = \{\boldsymbol{\eta} = (\eta_{xy})_{(x,y)\in\mathcal{E}} \in \overline{M} \mid \textstyle\sum_{(x,y)\in\mathcal{E}} \eta_{xy} = 1 \text{ and } \sum_{y:(x,y)\in\mathcal{E}} \eta_{xy} = \sum_{y:(y,x)\in\mathcal{E}} \eta_{yx} \text{ for any } x \in \mathcal{X}\}.$$

In [Nagaoka], it is shown that $\mathcal{W}$ is a dually flat manifold, and its expectation parameter space is M.

Theorem ([Nagaoka])

1. The mapping $T : \mathcal{W} \to M,\ w \mapsto (\mu_w(x)w(x,y))_{(x,y)\in\mathcal{E}}$ is a diffeomorphism.
2. There exists a convex function $\varphi : M \to \mathbb{R}$; the Bregman divergence $\mathcal{D} : \mathcal{W} \times \mathcal{W} \to \mathbb{R}$ induced by φ is
$$\mathcal{D}(w_1, w_2) = \sum_{(x,y)\in\mathcal{E}} \mu_{w_1}(x) w_1(x,y) \log \frac{w_1(x,y)}{w_2(x,y)}.$$

16 / 23

Positive transition measures on $\mathcal{W}$ (our work)

- Aim: We construct the counterpart of Amari's picture in $(\mathcal{P}(S), \bar{\mathcal{P}}(S))$ for $\mathcal{W}$.
- Main results:
  - We extend $\mathcal{W}$ to the bigger space $\mathcal{F}^+$.
  - We define an F-divergence on $\mathcal{F}^+$ and a diffeomorphism $\bar{T}$ between $\mathcal{F}^+$ and $\overline{M}$.
  - We give a divergence that is both a Bregman divergence and an F-divergence.
  - Actually, the potential function $\bar{\varphi}$ has a 1-dimensional kernel of its Hessian matrix at every point of $\overline{M}$, thus we take a hyperplane section $\tilde{M}$ in $\overline{M}$ so that a genuine dually flat structure is defined on it. That induces a hypersurface $\tilde{\mathcal{W}}$ in $\mathcal{F}^+$.

$$\begin{tikzcd}
\mathcal{F}^+ \arrow[r, "\sim", "\bar{T}"'] & \overline{M} \arrow[dr, "\bar{\varphi}"] & \\
\cup & \cup & \\
\tilde{\mathcal{W}} \arrow[r, "\sim", "\bar{T}|_{\tilde{\mathcal{W}}}"'] \arrow[d] & \tilde{M} \arrow[r, "\bar{\varphi}|_{\tilde{M}}"'] \arrow[d] & \mathbb{R} \arrow[d, equal] \\
\mathcal{W} \arrow[r, "\sim", "T"'] & M \arrow[r, "\varphi"'] & \mathbb{R}
\end{tikzcd}$$

17 / 23

Positive transition measures on $\mathcal{W}$ (our work)

Definition ([N])

Let $F : (0,\infty) \to \mathbb{R}$ be a strictly convex function with $F(1) = F'(1) = 0$ and $F''(1)$.
We define the F-divergence on $\mathcal{F}^+$ as $\mathcal{D}_F : \mathcal{F}^+ \times \mathcal{F}^+ \to \mathbb{R}$,
$$\mathcal{D}_F(f,g) = \textstyle\sum_{(x,y)\in\mathcal{E}} \mu_f(x) f(x,y) F\left(\frac{g(x,y)}{r(g)} \Big/ \frac{f(x,y)}{r(f)}\right).$$

Proposition ([N])

The F-divergence $\mathcal{D}_F$ has the following properties:

1. $\mathcal{D}_F(f,g) \geq 0$.
2. $\mathcal{D}_F(f,g) = 0$ if and only if $g = af$ for some $a > 0$.
3. $\mathcal{D}_F$ is a weak contrast function on $\mathcal{F}^+$. Let h_F denote the symmetric $(0,2)$-tensor on $\mathcal{F}^+$ induced by $\mathcal{D}_F$.
4. The null space of h_F at $f \in \mathcal{F}^+$ is the tangent space of the halfline $\{af \mid a > 0\} \subset \mathcal{F}^+$.

18 / 23

Positive transition measures on $\mathcal{W}$ (our work)

We set

$$\bar{T} : \mathcal{F}^+ \to \overline{M}, \quad f \mapsto (\mu_f(x) f(x,y))_{(x,y)\in\mathcal{E}}.$$

We also set for $\boldsymbol{\eta} = (\eta_{xy})_{(x,y)\in\mathcal{E}} \in \overline{M}$

$$r(\boldsymbol{\eta}) := \sum_{(x,y)\in\mathcal{E}} \eta_{xy}.$$

Lemma ([N])

$\bar{T}$ has the following properties:

1. $\bar{T}$ is a diffeomorphism, and $\bar{T}|_{\mathcal{W}} = T : \mathcal{W} \xrightarrow{\sim} M$.
2. $\bar{T}(af) = a\bar{T}(f)$ for $f \in \mathcal{F}^+$ and $a > 0$.
3. $r(f) = r(\boldsymbol{\eta})$ with $\bar{T}(f) = \boldsymbol{\eta}$.

Positive transition measures on $\mathcal{W}$ (our work)

Theorem ([N])

Let $F(t) = -\log t + (t-1)$. Then the F-divergence is the Bregman divergence given by the following potential function on $\overline{M}$:

$$\bar{\varphi}(\boldsymbol{\eta}) = \sum_{(x,y)\in\mathcal{E}} \eta_{xy} \log \eta_{xy} - \sum_{x\in\mathcal{X}} \eta_x \log \eta^x. \qquad (1)$$

For $w_1, w_2 \in \mathcal{W}$ we see

$$\begin{aligned} \mathcal{D}_F(w_1,w_2) &= \sum_{(x,y)\in\mathcal{E}} \mu_{w_1}(x) w_1(x,y) F\left(\frac{w_2(x,y)}{w_1(x,y)}\right) \\ &= \sum_{(x,y)\in\mathcal{E}} \mu_{w_1}(x) w_1(x,y) \log \frac{w_1(x,y)}{w_2(x,y)} + \sum_{(x,y)\in\mathcal{E}} \mu_{w_1}(x)(w_2(x,y) - w_1(x,y)) \\ &= \sum_{(x,y)\in\mathcal{E}} \mu_{w_1}(x) w_1(x,y) \log \frac{w_1(x,y)}{w_2(x,y)} : \text{ the divergence by Nagaoka} \end{aligned}$$

Positive transition measures on $\mathcal{W}$ (our work)

We see that the Hessian matrix of $\bar{\varphi}$ at every point $\boldsymbol{\eta} \in \overline{M}$ has the 1-dimensional kernel spanned by the numerical vector $\boldsymbol{\eta} \in \mathbb{R}^{|\mathcal{E}|} \cong T_{\boldsymbol{\eta}}\overline{M}$.

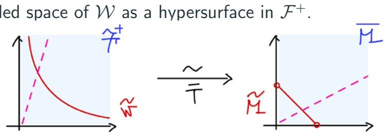
Therefore, by imposing only the normalization condition $\sum_{(x,y)\in\mathcal{E}} \eta_{xy} = 1$ on $\overline{M}$, we have the hyperplane section $\tilde{M}$ in $\overline{M}$ so that $\bar{\varphi}$ is strictly convex on it:

$$\tilde{M} := \{\boldsymbol{\eta} = (\eta_{xy}) \in \overline{M} \mid r(\boldsymbol{\eta}) = 1\}.$$

Using the relation $r(f) = r(\boldsymbol{\eta})$ with $\bar{T}(f) = \boldsymbol{\eta}$, we get the genuine dually flat manifold

$$\tilde{\mathcal{W}} = \{f \in \mathcal{F}^+ \mid r(f) = 1\},$$

which is an extended space of $\mathcal{W}$ as a hypersurface in $\mathcal{F}^+$.



Positive transition measures on $\mathcal{W}$ (our work)

Theorem ([N])

The hypersurface $\tilde{\mathcal{W}}$ has the dually flat structure induced by the potential function $\tilde{\varphi} := \bar{\varphi}|_{\tilde{M}}$ on $\tilde{M}$; the restriction of this dually flat structure to $\mathcal{W}$ restores the dually flat structure of [Nagaoka]. We call $\tilde{\mathcal{W}}$ the space of **positive transition measures**. Moreover F-divergences on $\tilde{\mathcal{W}}$ are written as

$$\mathcal{D}_F(f,g) = \sum_{(x,y)\in\mathcal{E}} \mu_f(x) f(x,y) F\left(\frac{g(x,y)}{f(x,y)}\right) \quad (f,g \in \tilde{\mathcal{W}}).$$

$$\begin{tikzcd}
\mathcal{F}^+ \arrow[r, "\sim", "\bar{T}"'] & \overline{M} \arrow[dr, "\bar{\varphi}"] & \\
\tilde{\mathcal{W}} \arrow[u, phantom, "\cup"] \arrow[r, "\sim", "\bar{T}|_{\tilde{\mathcal{W}}}"'] \arrow[d] & \tilde{M} \arrow[u, phantom, "\cup"] \arrow[r, "\bar{\varphi}|_{\tilde{M}}"'] \arrow[d] & \mathbb{R} \arrow[d, equal] \\
\mathcal{W} \arrow[r, "\sim", "T"'] & M \arrow[r, "\varphi"'] & \mathbb{R}
\end{tikzcd}$$

22 / 23

Summary and future plans

- We have defined the class of F-divergences on $\mathcal{F}^+$ and given a divergence which is both a Bregman divergence and an F-divergence. Moreover, we have given a dually flat manifold $\tilde{\mathcal{W}}$ which is an extension of $\mathcal{W}$ by analyzing the kernels of the potential function $\bar{\varphi}$ on $\overline{M}$.
- In order to completely establish the counterpart of Amari's theory for the pair $(\mathcal{W}, \tilde{\mathcal{W}})$, we need some discussions from the view point of statistics.
- In the first place, the "statistical invariance" for conditional probabilities must be discussed.
- Then, F-divergences should be characterized by the statistical invariance above.
- Besides, a divergence on $\tilde{\mathcal{W}}$ which is both a Bregman divergence and an F-divergence may be uniquely determined under certain conditions.

23 / 23

Thank you for your attention!

23 / 23

Reference i

[1] N. Nakajima and T. Ohmoto, *The dually flat structure for singular models*, Info. Geom., **4** (2021), 31–64.

[2] S. Amari, *Information geometry of the EM and em algorithms for neural networks*, Neural Networks, **8** (1995), 1379–1408.

[3] S. Amari, *Information Geometry and Its Application*, Applied Math. Sci., 194, Springer(2016)

[4] S. Amari and H. Nagaoka, *Methods of Information Geometry*, A.M.S., Oxford Univ. Press (2000).

[5] S. Eguchi, *Geometry of minimum contrast*, Hiroshima Math. J., **22** (1992), 631–647.

Reference ii

[6] T. Matumoto, *Any statistical manifold has a contrast function – On the C^3-functions taking the minimum at the diagonal of the product manifold*, Hiroshima Math. J., **23** (1993), 327–332.

[7] H. Nagaoka, *The exponential family of Markov chains and its information geometry*, Proceedings of The 28th Symposium on Information Theory and Its Applications (2005).

[8] H. Shima, *The geometry of Hessian Structures*, World Scientific (2007).

[9] J. Takeuchi, *Fisher information determinant and stochastic complexity for Markov models*, 2009 IEEE International Symposium on Information Theory (2009), 1894–1898.

Multivariate Hawkes processes with graphs

Mariusz Niewęgłowski
Faculty of Mathematics and Information Science,
Warsaw University of Technology, Poland

A very interesting and important class of stochastic processes was introduced by Alan Hawkes in [1]. These processes, called now Hawkes processes, are meant to model self-exciting and mutually-exciting random phenomena that evolve in time. The self-exciting phenomena are modeled as univariate Hawkes processes, and the mutually-exciting phenomena are modeled as multivariate Hawkes processes. The Hawkes processes have been applied to modeling in meany areas of science, including: insurance, finance, seismology and neurology. In this talk we provide some results on markovianity of the Generalized Multivariate Hawkes Processes (GMHP) introduced in our earlier papers. GMHP are multivariate marked point processes that add an important feature to the family of the (classical) multivariate Hawkes processes: they allow for explicit modelling of simultaneous occurrence of excitation events coming from different sources, i.e. caused by different coordinates of the multivariate process. We propose that this structure of mutual excitations is specified in terms of the excitation graph. We provide results which show that under some conditions on its kernels the intensities of GMHP's are functions of a Markov processes. Moreover we show that it is possible to compute their Laplace transform by means of system of ODE's. The talk is based on [4].

References

[1] A.G. Hawkes, "Spectra of Some Self-Exciting and Mutually Exciting Point Processes", *Biometrika* 58(1):83–90, 1971.

[2] T.R. Bielecki, J. Jakubowski, M. Niewęgłowski, "Construction and Simulation of Generalized Multivariate Hawkes Processes", *Methodology and Computing in Applied Probability* (2022) 24:2865-2896

[3] T.R. Bielecki, J. Jakubowski, M. Niewęgłowski, "Multivariate Hawkes processes with simultaneous occurrence of excitation events coming from different sources", *Stochastic Models* (2022) *online*

[4] T.R. Bielecki, J. Jakubowski, M. Niewęgłowski, "Markovianization of Multivariate Hawkes processes", *preprint*

Multivariate Hawkes processes

Mariusz Niewęgłowski

Faculty of Mathematics and Information Science
Warsaw University of Technology
based on joint work
Tomasz R. Bielecki, Illinois Institute of Technology,
Jacek Jakubowski, Warsaw University

WORKSHOP on Mathematics for Industry
Warsaw, Poland 2023

Hawkes Processes

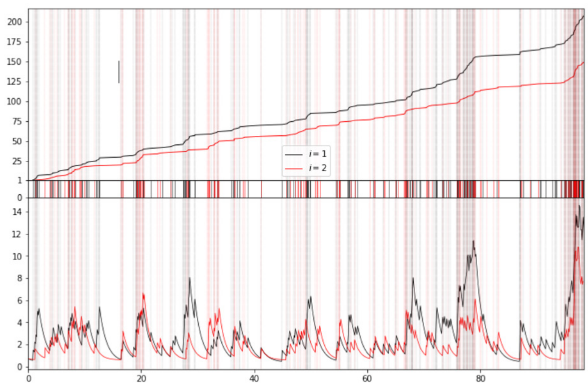
- Hawkes processes (self-exciting point process, 1971 Allan Hawkes) one dimensional point process (counting) N, defined by intensity

$$\lambda(t) = \lim_{\Delta t \downarrow 0} \frac{\mathbb{P}(N_{t+\Delta t} - N_t = 1 | \mathcal{F}^N_{t-})}{\Delta t} = \eta(t) + \int_{(0,t)} w(t-s) dN_s$$
$$= \eta(t) + \sum_{n: T_n < t} w(t - T_n),$$

where η non-negative function–background intensity, w non-negative function–impact function.
- Multivariate Hawkes Process: (mutually-exciting point processes), $N = (N^1, \dots, N^d)$ where N^i, $i = 1, \dots, d$, is a point process with the intensity given by

$$\lambda_i(t) = \eta_i(t) + \sum_{j=1}^{d} \int_{(0,t)} w_{i,j}(t-s) dN^j(s), \quad t \geq 0,$$

Trajectory of 2-dimensional Hawkes process



Applications

Y. Ogata
Space-time point-process models for earthquake occurrences.
Ann Inst Stat Math 50(2):379–402.(1998)

A. Bonnet, Ch. Dion-Blanc, F. Gindraud, S. Lemler,
Neuronal network inference and membrane potential model using multivariate Hawkes processes.
Journal of Neuroscience Methods, Volume 372, (2022).

Y. Bessy-Roland, A. Boumezoued, & C. Hillairet,
Multivariate Hawkes process for cyber insurance.
Annals of Actuarial Science, 15(1), 14-39, (2021).

M. Morariu-Patrichi & M.S. Pakkanen
State-dependent Hawkes processes and their application to limit order book modelling.
Quantitative Finance, 22:3, 563-583, (2022)

Introduction

- Goal: Provide framework for tractable specification of multivariate Hawkes processes (with common event times).
- What makes model tractable ?
  - Statistical methods.
  - Explicit formula for some distribution-related quantities.
  - Numerical methods for computations of such quantities.
  - Markov property.
- N-univariate Hawkes process is not a Markov process !
- Markovianization Problem: Find a Markov process X, function g such that $\lambda(t) = g(t, X_t)$ and (X, N) is a Markov process.
- Let $\eta = const$, $w(t) = ae^{-bt}$,

$$X(t) := \int_0^t ae^{-b(t-s)} dN_s$$

then (X, N) is a Markov process.

MPP-Marked Point Process

- Let us consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.
- Marked Point process N

$$N = (T_n, X_n)_{n \in \mathbb{Z}},$$

where $(T_n)_{n \in \mathbb{Z}}$ satisfies

$$T_n \leq T_{n+1}, \quad |T_n| < \infty \Rightarrow T_n < T_{n+1}$$

and (X_n) sequence of random variables, called marks, with values in $(E^\partial, \mathcal{E}^\partial)$ (∂ - point external to E)

$$X_n = \partial \Leftrightarrow |T_n| = \infty, \quad X_n \in E \Leftrightarrow |T_n| < \infty$$

- The explosion time of N, say T_∞, is defined as

$$T_\infty := \lim_{n \to \infty} T_n.$$

Random measures and MPP

- We associate with the N an integer valued random measure on $(\mathbb{R} \times E, \mathcal{B}(\mathbb{R}) \otimes \mathcal{E})$:

$$N(dt, dx) := \sum_{n \in \mathbb{Z}} \delta_{(T_n, X_n)}(dt, dx) \mathbb{1}_{\{|T_n| < \infty\}}$$

- Filtration $\mathbb{F}^N = (\mathcal{F}_t^N, \ t \geq 0)$ generated by N (completed)

$$\mathcal{F}_t^N = \sigma(N((s, r] \times A) \, : \, 0 \leq s < r \leq t, \ A \in \mathcal{E}), \quad t \geq 0.$$

Multivariate Mark space

- Let $(E_i, \mathcal{E}_i)$, $i = 1, 2, \ldots, d$, be some non-empty Borel measurable spaces. We extend $(E_i, \mathcal{E}_i)$

$$E_i^\Delta := E_i \cup \Delta, \qquad \mathcal{E}_i^\Delta = \sigma(\mathcal{E}_i, \{\Delta\}),$$

where Δ is a dummy mark.

- Then, we define a multivariate mark space, say E^Δ by

$$E^\Delta := E_1^\Delta \times E_2^\Delta \times \ldots \times E_d^\Delta \setminus (\Delta, \Delta, \ldots, \Delta).$$

σ-field $\mathcal{E}$ on E^Δ,

$$\mathcal{E}^\Delta := \left\{ A \cap E^\Delta : A \in \otimes_{i=1}^d \mathcal{E}_i^\Delta \right\}.$$

Motivation for $d = 2$

$N^1 =$

t_m^1	x_m^1
00:25	12.34
00:45	10.45
01:30	15.54
02:25	11.64
03:11	10.82
03:59	9.91
04:21	7.64
05:05	10.99
06:15	12.99
09:05	11.21

, $N^2 =$

t_m^2	x_m^2
01:54	3.49
03:11	5.78
03:45	4.31
03:59	3.95
04:35	7.91
06:15	9.99
09:05	8.74

, $N =$

t_n	x_n^1	x_n^2
00:25	12.34	Δ
00:45	10.45	Δ
01:30	15.54	Δ
01:54	Δ	3.49
02:25	11.64	Δ
03:11	10.82	5.78
03:45	Δ	4.31
03:59	9.91	3.95
04:21	7.64	Δ
04:35	Δ	7.91
05:05	10.99	Δ
06:15	12.99	9.99
09:05	11.21	8.74

Multivariate Marked Hawkes process

Definition

Let N^0 be random measure on $(\mathbb{R}_- \times E^\Delta, \mathcal{B}(\mathbb{R}_-) \otimes \mathcal{E}^\Delta)$, $\mathbb{G}$ a given filtration and a pair of kernels η, f satisfying

1. η is a finite kernel from $(\Omega \times [0,\infty), \mathcal{P}^{\mathbb{G}})$ to $(E^\Delta, \mathcal{E}^\Delta)$
2. f is a finite kernel from $(\Omega \times \mathbb{R}_+ \times \mathbb{R} \times E^\Delta, \mathcal{P}^{\mathbb{G}} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{E}^\Delta)$ to $(E^\Delta, \mathcal{E}^\Delta)$ and satisfies $f(t,s,x,A) = 0$ for $s > t$.

We call MPP N with multivariate mark space E^Δ a generalized $\mathbb{G}$-doubly stochastic multivariate Hawkes process (GDSMHP) directed by (η, f) with initial condition N^0 if $N = N^0$ on $(\mathbb{R}_- \times E^\Delta, \mathcal{B}(\mathbb{R}_-) \otimes \mathcal{E}^\Delta)$ and $\mathbb{G} \vee \mathbb{F}^N$-compensator of N on $(\mathbb{R}_+ \times E^\Delta, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}^\Delta)$, is of the form

$$\nu(\omega, dt, dy) = \mathbb{1}_{]\!]0,T_\infty[\![}\kappa(\omega, t, dy)dt,$$

where

$$\kappa(t, dy) = \eta(t, dy) + \int_{(-\infty,t)\times E^\Delta} f(t,s,x,dy)N(ds,dx).$$

Auxiliary notation

- By $2^{[d]}$ we denote all non-empty subsets of $[d] := \{1,\dots,d\}$.
- For generic $\mathcal{I} \in 2^{[d]}$ we let $\mathcal{I}^c := [d] \setminus \mathcal{I}$ and we set

$$E^{\mathcal{I}} = \bigtimes_{i=1}^{d} A_i, \quad where \quad A_i = \begin{cases} E_i & \text{if } i \in \mathcal{I}, \\ \{\Delta\} & \text{otherwise,} \end{cases}$$

- Let $(i_1, \dots, i_{d_{\mathcal{I}}})$ be the ordered sequence of elements of $\mathcal{I}$ we denote

$$\begin{aligned} E_{\mathcal{I}} &= \bigtimes_{j=1}^{d_{\mathcal{I}}} E_{i_j}, \\ x_{\mathcal{I}} &= (x_{i_1}, x_{i_2}, \dots, x_{i_{d_{\mathcal{I}}}}) \in E_{\mathcal{I}}, \\ \mathrm{d}x_{\mathcal{I}} &= \mathrm{d}x_{i_1}\, \mathrm{d}x_{i_2} \dots \mathrm{d}x_{i_{d_{\mathcal{I}}}}, \\ \delta_{\Delta^{\mathcal{I}^c}}(\mathrm{d}y_{\mathcal{I}^c}) &= \otimes_{i\in\mathcal{I}^c}\delta_\Delta(\mathrm{d}y_i) \end{aligned}$$

- $E^{\mathcal{I}} \subset E^\Delta$, $E^\Delta = \bigcup_{\mathcal{I}\in 2^{[d]}} E^{\mathcal{I}}$.

$\mathcal{I}$-idiosyncratic coordinate

Definition

For a random measure $N(\mathrm{d}u, \mathrm{d}x)$ on $(\mathbb{R} \times E^\Delta)$ and a set $\mathcal{I} \in 2^{[d]}$ we define a random measure $N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}s, \mathrm{d}x_{\mathcal{I}})$ on $(\mathbb{R} \times E_{\mathcal{I}})$ by setting

$$N^{\mathrm{id}}_{\mathcal{I}}((s,t] \times A) = N((s,t] \times \Gamma^{\mathcal{I}}(A)), \quad A \in E_{\mathcal{I}}$$

where $\Gamma^{\mathcal{I}} : E_{\mathcal{I}} \to E^{\mathcal{I}}$ is a lifting mapping defined by

$$[\Gamma^{\mathcal{I}}(x_{\mathcal{I}})]_i = \begin{cases} x_i & \text{if } i \in \mathcal{I}, \\ \Delta & \text{otherwise,} \end{cases} \qquad i \in [d].$$

We call $N^{\mathrm{id}}_{\mathcal{I}}$ - the $\mathcal{I}$-idiosyncratic coordinate process.

N can be represented in the form

$$N((s,t] \times A) = \sum_{\mathcal{J}\in 2^{[d]}} N((s,t] \times (A \cap E^{\mathcal{J}})) = \sum_{\mathcal{J}\in 2^{[d]}} N^{\mathrm{id}}_{\mathcal{J}}((s,t] \times (\Gamma^{\mathcal{J}})^{-1}(A \cap E^{\mathcal{J}}))$$

Illustration for $d = 2$

$N =$

t_n	x_n^1	x_n^2
00:25	12.34	Δ
00:45	10.45	Δ
01:30	15.54	Δ
01:54	Δ	3.49
02:25	11.64	Δ
03:11	10.82	5.78
03:45	Δ	4.31
03:59	9.91	3.95
04:21	7.64	Δ
04:35	Δ	7.91
05:05	10.99	Δ
06:15	12.99	9.99
09:05	11.21	8.74

$N^{\mathrm{id}}_{\{1\}} =$

t_m^1	x_m^1
00:25	12.34
00:45	10.45
01:30	15.54
02:25	11.64
04:21	7.64
05:05	10.99

$N^{\mathrm{id}}_{\{2\}} =$

t_m^2	x_m^2
01:54	3.49
03:45	4.31
04:35	7.91

$N^{\mathrm{id}}_{\{1,2\}} =$

$t_m^{1,2}$	x_m^1	x_m^2
03:11	10.82	5.78
03:59	9.91	3.95
06:15	12.99	9.99
09:05	11.21	8.74

Lemma

1. Every kernel η from a measurable space $(\Omega \times \mathbb{R}_+, \mathcal{A})$ to $(E^\Delta, \mathcal{E}^\Delta)$ can be uniquely written as

$$\eta(t, \mathrm{d}y) = \sum_{\mathcal{J} \in 2^{[d]}} \eta_{\mathcal{J}}(t, \mathrm{d}y_{\mathcal{J}}) \otimes \delta_{\Delta^{\mathcal{J}^c}}(\mathrm{d}y_{\mathcal{J}^c})$$

where $\eta_{\mathcal{J}}$ are kernels from $(\Omega \times \mathbb{R}_+, \mathcal{A})$ to $(E_{\mathcal{J}}, \mathcal{E}_{\mathcal{J}})$ such that

$$\eta_{\mathcal{J}}(t, A_{\mathcal{J}}) = \eta(t, \Gamma^{\mathcal{J}}(A_{\mathcal{J}})) \quad \text{for } A_{\mathcal{J}} \in \mathcal{E}_{\mathcal{J}}.$$

2. Every kernel f from $(\Omega \times \mathbb{R}_+ \times \mathbb{R} \times E^\Delta, \mathcal{A} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{E}^\Delta)$ to $(E^\Delta, \mathcal{E}^\Delta)$ can be uniquely written as

$$f(t, s, x, \mathrm{d}y) = \sum_{\mathcal{I}, \mathcal{J} \in 2^{[d]}} f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, \mathrm{d}y_{\mathcal{J}}) \otimes \delta_{\Delta^{\mathcal{J}^c}}(\mathrm{d}y_{\mathcal{J}^c}) \mathbb{1}_{E^{\mathcal{I}}}(x),$$

where $f_{\mathcal{I},\mathcal{J}}$ are kernels from $(\Omega \times \mathbb{R}_+ \times \mathbb{R} \times E_{\mathcal{I}}, \mathcal{A} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{E}_{\mathcal{I}})$ to $(E_{\mathcal{J}}, \mathcal{E}_{\mathcal{J}})$ such that

$$f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, A_{\mathcal{J}}) = f(t, s, \Gamma^{\mathcal{I}}(x_{\mathcal{I}}), \Gamma^{\mathcal{J}}(A_{\mathcal{J}})) \quad \text{for } A_{\mathcal{J}} \in \mathcal{E}_{\mathcal{J}}.$$

Introducing graph(-ic)

- Suppose that directing kernels (η, f) are defined by means of a given $\mathbb{M} \subset \mathbb{V} \subset 2^{[d]}$, $\mathbb{A} \subset \mathbb{V} \times \mathbb{V}$ and families of non-zero kernels $\{\eta_{\mathcal{J}} : \mathcal{J} \in \mathbb{M}\}$, $\{f_{\mathcal{I},\mathcal{J}} : (\mathcal{I}, \mathcal{J}) \in \mathbb{A}\}$ by following formula

$$\eta(t, dy) = \sum_{\mathcal{J} \in \mathbb{M}} \eta_{\mathcal{J}}(t, \mathrm{d}y_{\mathcal{J}}) \otimes \delta_{\Delta^{\mathcal{J}^c}}(\mathrm{d}y_{\mathcal{J}^c}),$$

$$f(t, s, x, dy) = \sum_{(\mathcal{I}, \mathcal{J}) \in \mathbb{A}} f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, \mathrm{d}y_{\mathcal{J}}) \otimes \delta_{\Delta^{\mathcal{J}^c}}(\mathrm{d}y_{\mathcal{J}^c}) \mathbb{1}_{E^{\mathcal{I}}}(x).$$

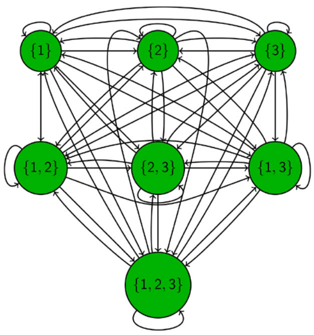
- We call $G = ((\mathbb{V}, \mathbb{A}), \mathbb{M})$ an excitations graphic.
- We call $\mathbb{M}$ set of exogenous sources of excitations.
- For a given $\mathcal{J} \in \mathbb{V}$ we define the *parents of $\mathcal{J}$* in G

$$\mathrm{Pa}_G(\mathcal{J}) = \{\mathcal{I} \in \mathbb{V} : (\mathcal{I}, \mathcal{J}) \in \mathbb{A}\},$$

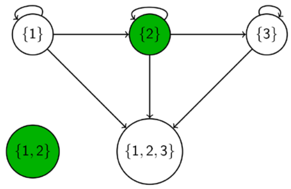
for a given $\mathcal{I} \in \mathbb{V}$ we define the set of *ancestors of $\mathcal{I}$* in G

$$\mathrm{An}_G(\mathcal{I}) = \{\mathcal{J} \in \mathbb{V} : (\mathcal{I}, \mathcal{J}) \in \mathbb{A}\}.$$

Graphical description



$d = 3, |\mathbb{A}| = 49, |\mathbb{M}| = 8$



$d = 3, |\mathbb{A}| = 8, |\mathbb{M}| = 2$

Proposition

The Hawkes intensity kernel of $\mathbb{G}$-DSGMHP N with initial condition N^0 directed by such (η, f) is of the form

$$\kappa(t, dy) = \sum_{\mathcal{J} \in \mathbb{M}} \eta_{\mathcal{J}}(t, dy_{\mathcal{J}}) \otimes \delta_{\Delta_{\mathcal{J}^c}}(dy_{\mathcal{J}^c}) + \sum_{(\mathcal{I},\mathcal{J}) \in \mathbb{A}} \int_{(-\infty,t) \times E_{\mathcal{I}}} f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, dy_{\mathcal{J}}) \otimes \delta_{\Delta_{\mathcal{J}^c}}(dy_{\mathcal{J}^c}) N^{id}_{\mathcal{I}}(ds, dx_{\mathcal{I}}).$$

The $\mathbb{G} \vee \mathbb{F}^N$-intensity kernel of the random measure $N^{id}_{\mathcal{K}}$ is given by

$$\kappa^{id}_{\mathcal{K}}(t, dy_{\mathcal{K}}) = \mathbb{1}_{\mathbb{M}}(\mathcal{K}) \eta_{\mathcal{K}}(t, dy_{\mathcal{K}}) + \sum_{\mathcal{I} \in \mathrm{Pa}_G(\mathcal{K})} \int_{(-\infty,t) \times E_{\mathcal{I}}} f_{\mathcal{I},\mathcal{K}}(t, s, x_{\mathcal{I}}, dy_{\mathcal{K}}) N^{id}_{\mathcal{I}}(ds, dx_{\mathcal{I}}).$$

Structural Assumption (1)

Respective components $\eta_{\mathcal{J}}$ and $f_{\mathcal{I},\mathcal{J}}$ satisfy

1. For every $\mathcal{J} \in \mathbb{M}$ the kernel $\eta_{\mathcal{J}}(t, dy_{\mathcal{J}})$ takes form

$$\eta_{\mathcal{J}}(t, dy_{\mathcal{J}}) = \widetilde{\eta}_{\mathcal{J}}(t) Q_{\mathcal{J}}(dy_{\mathcal{J}}),$$

where $(\widetilde{\eta}_{\mathcal{J}}(t))$ is a $\mathbb{G}$-predictable stochastic process, $Q_{\mathcal{J}}$ is a probability measure on $(E_{\mathcal{J}}, \mathcal{E}_{\mathcal{J}})$

2. For every $(\mathcal{I}, \mathcal{J}) \in \mathbb{A}$ the kernel $f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, dy_{\mathcal{J}})$ takes form

$$f_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}, dy_{\mathcal{J}}) = \widetilde{f}_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}) R_{\mathcal{I},\mathcal{J}}(dy_{\mathcal{J}})$$

where $(\widetilde{f}_{\mathcal{I},\mathcal{J}}(t, s, x_{\mathcal{I}}))$ is a $\mathbb{G}$-predictable mapping, $R_{\mathcal{I},\mathcal{J}}$ is a probability measure on $(E_{\mathcal{J}}, \mathcal{E}_{\mathcal{J}})$.

Proposition

Assume that structural assumption holds. Then

1. *The Hawkes kernel of N has the form*

$$\kappa(t,\mathrm{d}y) = \sum_{\mathcal{I}\in\mathbb{M}} \widetilde{\eta}_{\mathcal{I}}(t)Q_{\mathcal{I}}(\mathrm{d}y_{\mathcal{I}})\otimes\delta_{\Delta^{\mathcal{I}^c}}(\mathrm{d}y_{\mathcal{I}^c}) + \sum_{(\mathcal{I},\mathcal{J})\in\mathbb{A}} \lambda_{\mathcal{I},\mathcal{J}}(t)R_{\mathcal{I},\mathcal{J}}(\mathrm{d}y_{\mathcal{J}})\otimes\delta_{\Delta^{\mathcal{J}^c}}(\mathrm{d}y_{\mathcal{J}^c}),$$

where

$$\lambda_{\mathcal{I},\mathcal{J}}(t) = \int_{(-\infty,t)\times E_{\mathcal{I}}} \widetilde{f}_{\mathcal{I},\mathcal{J}}(t,s,x_{\mathcal{I}})N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}s,\mathrm{d}x_{\mathcal{I}}).$$

2. *Fix $\mathcal{K}\in 2^{[d]}$. The $\mathbb{F}$-compensator of the random measure $N^{\mathrm{id}}_{\mathcal{K}}$, say $\kappa^{\mathrm{id}}_{\mathcal{K}}(t,\mathrm{d}y_{\mathcal{K}})\,\mathrm{d}t$, is given by*

$$\kappa^{\mathrm{id}}_{\mathcal{K}}(t,\mathrm{d}y_{\mathcal{K}})dt = \mathbb{1}_{\mathbb{M}}(\mathcal{K})\widetilde{\eta}_{\mathcal{K}}(t)Q_{\mathcal{K}}(\mathrm{d}y_{\mathcal{K}})\,\mathrm{d}t + \sum_{\mathcal{I}\in\mathrm{Pa}_G(\mathcal{K})} \lambda_{\mathcal{I},\mathcal{K}}(t)R_{\mathcal{I},\mathcal{K}}(\mathrm{d}y_{\mathcal{K}})\,\mathrm{d}t.$$

Proposition

In particular, intensity process of $N^{\mathrm{id}}_{\mathcal{K}}$ is given by

$$\Lambda^{\mathrm{id}}_{\mathcal{K}}(t) := \kappa^{\mathrm{id}}_{\mathcal{K}}(t,E_{\mathcal{K}}) = \widetilde{\eta}_{\mathcal{K}}(t)+\lambda^{\mathrm{id}}_{\mathcal{K}}(t), \quad \text{where} \quad \lambda^{\mathrm{id}}_{\mathcal{K}}(t) := \sum_{\mathcal{I}\in\mathrm{Pa}_G(\mathcal{K})} \lambda_{\mathcal{I},\mathcal{K}}(t).$$

Definition

We say that a Markov process (X,Y) (possibly time inhomogeneous) with a state space $(S,\mathcal{S}) = (S_1\times S_2, \mathcal{S}_1\otimes\mathcal{S}_2)$ is a *markovianization* of $\mathbb{G}$-doubly stochastic Hawkes process N directed by (η,f) if

$$\widetilde{\eta}_{\mathcal{I}}(t) = \widehat{\eta}_{\mathcal{I}}(t,Y(t-)), \quad \lambda_{\mathcal{I},\mathcal{K}}(t) = \widehat{\lambda}_{\mathcal{I},\mathcal{K}}(t,X(t-)),$$

for some measurable functions $\{\widehat{\eta}_{\mathcal{I}}:\mathbb{R}_+\times S_2\to\mathbb{R}_+ : \mathcal{I}\in\mathbb{M}\}$ and $\left\{\widehat{\lambda}_{\mathcal{I},\mathcal{K}}:\mathbb{R}_+\times S_1\to\mathbb{R}_+ : (\mathcal{I},\mathcal{K})\in\mathbb{A}\right\}$. We call Y the *exogenous factor* process if it is $\mathbb{G}$-adapted and X *endogenous factor* process if it is $\mathbb{F}^N$-adapted.

Structural Assumption (2)

- For every $\mathcal{I}\in\mathbb{M}$

$$\widetilde{\eta}_{\mathcal{I}}(t) = \mu_{\mathcal{I}}(t) + \beta_{\mathcal{I}}(t)\int_{(0,t)\times\mathbb{R}} \phi_{\mathcal{I}}(t-s,s,x)M_{\mathcal{I}}(\mathrm{d}s,\mathrm{d}x),$$

  where $\beta_{\mathcal{I}}$, $\mu_{\mathcal{I}}$ are non-negative deterministic functions on $\mathbb{R}_+$, whereas $\phi^{\mathcal{I}}:\mathbb{R}_+\times\mathbb{R}\to\mathbb{R}_+$, and $(M_{\mathcal{I}})_{\mathcal{I}}$ are independent Poisson r.m. such that the $\mathbb{F}$-compensator of $M_{\mathcal{I}}$ is $P_{\mathcal{I}}(\mathrm{d}x)\theta_{\mathcal{I}}dt$ for $\theta_{\mathcal{I}}\geq 0$, $P_{\mathcal{I}}$-probability measure.

- For every $(\mathcal{I},\mathcal{J})\in\mathbb{A}$

$$f_{\mathcal{I},\mathcal{J}}(t,s,x_{\mathcal{I}}) = \alpha_{\mathcal{I},\mathcal{J}}(t)\psi_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}})$$

  where $\alpha_{\mathcal{I},\mathcal{J}}$, is a non-negative deterministic function on $\mathbb{R}_+$, whereas $\psi_{\mathcal{I},\mathcal{J}}:\mathbb{R}_+\times\mathbb{R}\times E_{\mathcal{I}}\to\mathbb{R}_+$.

- The above assumption implies that the $\widetilde{\eta}_{\mathcal{J}}(t)$ and $\lambda_{\mathcal{I},\mathcal{J}}$ can be written as

$$\widetilde{\eta}_{\mathcal{J}}(t) = \mu_{\mathcal{J}}(t) + \beta_{\mathcal{J}}(t) Y_{\mathcal{J}}(t-), \quad \lambda_{\mathcal{I},\mathcal{J}}(t) = \alpha_{\mathcal{I},\mathcal{J}}(t) X_{\mathcal{I},\mathcal{J}}(t-), \quad t \geq 0$$

where

$$Y_{\mathcal{J}}(t) := \int_{(0,t]\times\mathbb{R}} \phi_{\mathcal{I}}(t-s,s,x) M_{\mathcal{I}}(\mathrm{d}s, \mathrm{d}x)$$
$$X_{\mathcal{I},\mathcal{J}}(t) := \int_{(-\infty,t]\times E_{\mathcal{I}}} \psi_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}}) N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}s, \mathrm{d}x_{\mathcal{I}}),$$

- First step: Provide conditions for Markovian dynamics of these processes
- Note that the intensity kernel of $N^{\mathrm{id}}_{\mathcal{I}}$ is given by

$$\kappa^{\mathrm{id}}_{\mathcal{I}}(t, \mathrm{d}y_{\mathcal{I}}) = \mathbb{1}_{\mathbb{M}}(\mathcal{I}) \widetilde{\eta}_{\mathcal{I}}(t) Q_{\mathcal{I}}(\mathrm{d}y_{\mathcal{I}}) + \sum_{\mathcal{K}\in \mathrm{Pa}_G(\mathcal{I})} \lambda_{\mathcal{K},\mathcal{I}}(t) R_{\mathcal{K},\mathcal{I}}(\mathrm{d}y_{\mathcal{I}}).$$

Exponential case generalized

Theorem

Suppose that $\psi_{\mathcal{I},\mathcal{J}}$ satisfies linear ODE (in first variable)

$$\psi^{(n)}_{\mathcal{I},\mathcal{J}}(t,s,z_{\mathcal{I}}) = g^{-1}_{\mathcal{I},\mathcal{J}} + g^{0}_{\mathcal{I},\mathcal{J}} \psi^{(0)}_{\mathcal{I},\mathcal{J}}(t,s,z_{\mathcal{I}}) + g^{1}_{\mathcal{I},\mathcal{J}} \psi^{(1)}_{\mathcal{I},\mathcal{J}}(t,s,z_{\mathcal{I}}) + \ldots + g^{(n-1)}_{\mathcal{I},\mathcal{J}} \psi^{(n-1)}_{\mathcal{I},\mathcal{J}}(t,s,z_{\mathcal{I}})$$

with initial conditions

$$\psi^{(0)}_{\mathcal{I},\mathcal{J}}(0,s,z_{\mathcal{I}}) = \overline{\psi}^{0}_{\mathcal{I},\mathcal{J}}(s,z_{\mathcal{I}}), \; \ldots \;, \; \psi^{(n-1)}_{\mathcal{I},\mathcal{J}}(0,s,z_{\mathcal{I}}) = \overline{\psi}^{n-1}_{\mathcal{I},\mathcal{J}}(s,z_{\mathcal{I}}),$$

where $\psi^{(i)}$ denotes the derivative of i-th order in first variable. Let $\overline{X}_{\mathcal{I},\mathcal{J}}$ be a $\mathbb{R}_{n+1}$ valued process given by

$$\overline{X}_{\mathcal{I},\mathcal{J}}(t) = \int_{(-\infty,t]\times E_{\mathcal{I}}} \overline{\psi}_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}}) N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}s, \mathrm{d}x_{\mathcal{I}}),$$

where

$$\overline{\psi}_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}}) = \left[1, \psi^{(0)}_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}}), \ldots, \psi^{(n-1)}_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}})\right]'.$$

Theorem (cont'd)

Then $\overline{X}_{\mathcal{I},\mathcal{J}} = (\overline{X}^{k}_{\mathcal{I},\mathcal{J}})^{n+1}_{k=1}$ solves SDE on $\mathbb{R}_+$

$$\mathrm{d}\overline{X}_{\mathcal{I},\mathcal{J}}(t) = G_{\mathcal{I},\mathcal{J}} \overline{X}_{\mathcal{I},\mathcal{J}}(t)\,\mathrm{d}t + \overline{\psi}_{\mathcal{I},\mathcal{J}}(0,t,z_{\mathcal{I}}) N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}t, \mathrm{d}z_{\mathcal{I}}),$$
$$\overline{X}_{\mathcal{I},\mathcal{J}}(0) = \int_{(-\infty,0]\times E_{\mathcal{I}}} \overline{\psi}_{\mathcal{I},\mathcal{J}}(-s,s,x_{\mathcal{I}}) N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}s, \mathrm{d}x_{\mathcal{I}}),$$

where $G_{\mathcal{I},\mathcal{J}} \in \mathbb{R}_{n+1,n+1}$ and $\overline{\psi}_{\mathcal{I},\mathcal{J}}(t,z_{\mathcal{I}}) \in \mathbb{R}_{n+1}$ are given by

$$G_{\mathcal{I},\mathcal{J}} = \begin{bmatrix} 0 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 1 & 0 & & 0 & 0 \\ 0 & 0 & 0 & 1 & & 0 & 0 \\ & & & & & & \\ 0 & 0 & 0 & 0 & & 1 & 0 \\ 0 & 0 & 0 & 0 & & 0 & 1 \\ g^{-1}_{\mathcal{I},\mathcal{J}} & g^{0}_{\mathcal{I},\mathcal{J}} & g^{1}_{\mathcal{I},\mathcal{J}} & g^{2}_{\mathcal{I},\mathcal{J}} & & g^{n-2}_{\mathcal{I},\mathcal{J}} & g^{n-1}_{\mathcal{I},\mathcal{J}} \end{bmatrix}.$$

Moreover

$$\lambda_{\mathcal{I},\mathcal{J}}(t) := \alpha_{\mathcal{I},\mathcal{J}}(t) \overline{X}^{2}_{\mathcal{I},\mathcal{J}}(t-).$$

Lemma

Suppose that for every $\mathcal{I} \in \mathbb{M}$ $\phi_{\mathcal{I}}$ satisfies linear ODE (in first variable)

$$\phi_{\mathcal{I}}^{(m)}(t,s,x) = h_{\mathcal{I}}^{-1} + h_{\mathcal{I}}^{0}\phi_{\mathcal{I}}^{(0)}(t,s,x) + h_{\mathcal{I}}^{1}\phi_{\mathcal{I}}^{(1)}(t,s,x) + \ldots + h_{\mathcal{I}}^{(m-1)}\phi_{\mathcal{I}}^{(m-1)}(t,s,x)$$

with initial conditions

$$\phi_{\mathcal{I}}^{(0)}(0,s,x) = \overline{\phi}_{\mathcal{I}}^{0}(s,x), \ \ldots \ , \ \phi_{\mathcal{I}}^{(m-1)}(0,s,x) = \overline{\phi}_{\mathcal{I}}^{m-1}(s,x),$$

and let

$$\overline{Y}_{\mathcal{I}}(t) = \int_{(0,t]\times\mathbb{R}} \overline{\phi}_{\mathcal{I}}(t-s,s,x)M_{\mathcal{I}}(\mathrm{d}s,\mathrm{d}x),$$

where

$$\overline{\phi}_{\mathcal{I},\mathcal{J}}(t-s,s,x_{\mathcal{I}}) = \left[1, \phi_{\mathcal{I},\mathcal{J}}^{(0)}(t-s,s,x_{\mathcal{I}}), \ldots, \phi_{\mathcal{I},\mathcal{J}}^{(m-1)}(t-s,s,x_{\mathcal{I}})\right]'.$$

Then $\overline{Y}_{\mathcal{I}}$ is a Markov process which solves SDE

$$\mathrm{d}\overline{Y}_{\mathcal{I}}(t) = H_{\mathcal{I}}\overline{Y}_{\mathcal{I}}(t)\,\mathrm{d}t + \int_{\mathbb{R}} \overline{\phi}_{\mathcal{I}}(0,t,x)M_{\mathcal{I}}(\mathrm{d}t,\mathrm{d}x), \quad \overline{Y}_{\mathcal{I}}(0) = 0_{m+1}.$$

Moreover

$$\widetilde{\eta}_{\mathcal{I}}(t) = \mu_{\mathcal{I}}(t) + \beta_{\mathcal{I}}(t)\overline{Y}_{\mathcal{I}}^{2}(t)$$

Vectorizations of $(\overline{X}_{\mathcal{I},\mathcal{J}})_{(\mathcal{I},\mathcal{J})\in\mathbb{A}}$

- We first let σ to be a bijection
$$\sigma : \mathbb{A} \to [|\mathbb{A}|] = \{1, \ldots, |\mathbb{A}|\}.$$
- Then, for $(\mathcal{I},\mathcal{J}) \in \mathbb{A}$ w define vector $\mathrm{c}_{\mathcal{I},\mathcal{J}} \in \mathbb{R}_{|\mathbb{A}|}$ by formula
$$(\mathrm{c}_{\mathcal{I},\mathcal{J}})_i = \begin{cases} 1 & \text{if } i = \sigma(\mathcal{I},\mathcal{J}), \\ 0 & \text{otherwise.} \end{cases}$$
- and the stacked vector
$$X = \sum_{(\mathcal{I},\mathcal{J})\in\mathbb{A}} \mathrm{c}_{\mathcal{I},\mathcal{J}} \otimes \overline{X}_{\mathcal{I},\mathcal{J}}.$$
where $\otimes$ denotes Kornecker product of vectors.
- We have
$$\overline{X}_{\mathcal{I},\mathcal{J}}^{k} = X^{\mathrm{i}(\mathcal{I},\mathcal{J},k)}, \ \text{where} \quad \mathrm{i}(\mathcal{I},\mathcal{J},k) := (\sigma(\mathcal{I},\mathcal{J}) - 1)(n+1) + k.$$

Vectorization of $(Y_{\mathcal{I}})_{\mathcal{I}\in\mathbb{M}}$

- We let τ be a bijection

$$\tau : \mathbb{M} \to [|\mathbb{M}|] = \{1, \dots, |\mathbb{M}|\},$$

- for $\mathcal{I} \in \mathbb{M}$ let $c_{\mathcal{I}}$ be a vector $c_{\mathcal{I}} \in \mathbb{R}_{|\mathbb{M}|}$ defined by formula

$$(c_{\mathcal{I}})_i = \begin{cases} 1 & \text{if } i = \tau(\mathcal{I}), \\ 0 & \text{otherwise.} \end{cases}$$

- Now the stacked vector $\overline{Y}$ is defined by

$$Y = \sum_{\mathcal{I}\in\mathbb{M}} c_{\mathcal{I}} \otimes \overline{Y}_{\mathcal{I}}$$

- $$\overline{Y}^k_{\mathcal{I}} = Y^{\mathrm{j}(\mathcal{I},k)}, \quad \text{where} \quad \mathrm{j}(\mathcal{I},k) := (\tau(\mathcal{I}) - 1)(m+1) + k,$$

- Then (X, Y) solves system of SDE

$$\mathrm{d}X(t) = GX(t)\,\mathrm{d}t + \sum_{\mathcal{I}\in\mathrm{Pa}_G} \int_{E_{\mathcal{I}}} \psi_{\mathcal{I}}(t, z_{\mathcal{I}}) N^{\mathrm{id}}_{\mathcal{I}}(\mathrm{d}t, \mathrm{d}z_{\mathcal{I}}),$$
$$\mathrm{d}Y(t) = HY(t)\,\mathrm{d}t + \sum_{\mathcal{I}\in\mathbb{M}} \int_{\mathbb{R}} c_{\mathcal{I}} \otimes \overline{\phi}_{\mathcal{I}}(t,x) M_{\mathcal{I}}(\mathrm{d}t, \mathrm{d}x)$$

- where $\mathrm{Pa}_G = \{\mathcal{I} \in \mathbb{V} : \mathrm{An}(\mathcal{I}) \neq \emptyset\}$ and

$$\psi_{\mathcal{I}}(t, z_{\mathcal{I}}) = \sum_{\mathcal{J}\in\mathrm{An}_G(\mathcal{I})} c_{\mathcal{I},\mathcal{J}} \otimes \overline{\psi}_{\mathcal{I},\mathcal{J}}(0, t, z_{\mathcal{I}})$$

- and

$$G := \sum_{(\mathcal{I},\mathcal{J})\in\mathbb{A}} c_{\mathcal{I},\mathcal{J}} \otimes c'_{\mathcal{I},\mathcal{J}} \otimes G_{\mathcal{I},\mathcal{J}}$$
$$H := \sum_{\mathcal{I}\in\mathbb{M}} c_{\mathcal{I}} \otimes c'_{\mathcal{I}} \otimes H_{\mathcal{I}}.$$

Theorem

Then, the process (X, Y) is a markovianization of a $\mathbb{G}$-doubly stochastic Hawkes process N directed by (η, f) i.e. it holds that

$$\widetilde{\eta}_{\mathcal{I}}(t) = \mu_{\mathcal{I}}(t) + \beta_{\mathcal{I}}(t) Y^{\mathrm{i}(\mathcal{I},2)}(t-), \qquad t \geq 0.$$
$$\lambda_{\mathcal{I},\mathcal{J}}(t) = \alpha_{\mathcal{I},\mathcal{J}}(t) X^{\mathrm{i}(\mathcal{I},\mathcal{J},2)}(t-)$$

The generator of (X, Y) is given by

$$\begin{aligned}
&\mathcal{A}v(t,x,y) \\
&= \frac{\partial v}{\partial t} + \sum_{j=1}^{|\mathbb{A}|(n+1)} \Big(\sum_{k=1}^{|\mathbb{A}|(n+1)} G^{j,k} x^k \Big) \frac{\partial v}{\partial x^j} + \sum_{i=1}^{|\mathbb{M}|(m+1)} \Big(\sum_{j=1}^{|\mathbb{M}|(m+1)} H^{i,j} y^j \Big) \frac{\partial v}{\partial y^i} \\
&+ \sum_{\mathcal{I}\in\mathbb{M}} (\mu_{\mathcal{I}}(t) + \beta_{\mathcal{I}}(t) y^{\mathrm{i}(\mathcal{I},2)}) \int_{E_{\mathcal{I}}} \Big(v(t, x + \psi_{\mathcal{I}}(t, z_{\mathcal{I}}), y) - v(t,x,y) \Big) Q_{\mathcal{I}}(\mathrm{d}z_{\mathcal{I}}) \\
&+ \sum_{(\mathcal{K},\mathcal{I})\in\mathbb{A}} x^{\mathrm{i}(\mathcal{K},\mathcal{I},2)} \alpha_{\mathcal{K},\mathcal{I}}(t) \int_{E_{\mathcal{I}}} \Big(v(t, x + \psi_{\mathcal{I}}(t, z_{\mathcal{I}}), y) - v(t,x,y) \Big) R_{\mathcal{K},\mathcal{I}}(\mathrm{d}z_{\mathcal{I}}) \\
&+ \sum_{\mathcal{I}\in\mathbb{M}} \theta_{\mathcal{I}} \int_{\mathbb{R}} \Big(v(t,x,y + c_{\mathcal{I}} \otimes \overline{\phi}_{\mathcal{I}}(t,z)) - v(t,x,y) \Big) P_{\mathcal{I}}(\mathrm{d}z)
\end{aligned}$$

Extending $\overline{X}$

- Let us consider

$$n(t) = n(0) + \sum_{\mathcal{I}} \int_0^t \int_{E_{\mathcal{I}}} \mathrm{e}_{\mathcal{I}} N_{\mathrm{id}}^{\mathcal{I}}(\mathrm{d}t, \mathrm{d}z_{\mathcal{I}})$$

$$L(t) = L(0) + \sum_{\mathcal{I}} \int_0^t \int_{E_{\mathcal{I}}} \xi^{\mathcal{I}}(z_{\mathcal{I}}) N_{\mathrm{id}}^{\mathcal{I}}(\mathrm{d}t, \mathrm{d}z_{\mathcal{I}}),$$

  where $\mathrm{e}_{\mathcal{I}} \in \mathbb{R}_d$ are vectors defined by

$$(\mathrm{e}_{\mathcal{I}})_i = \begin{cases} 1 & \text{if } i \in \mathcal{I}, \\ 0 & \text{otherwise.} \end{cases}$$

  and where $\xi^{\mathcal{I}} : E_{\mathcal{I}} \to \mathbb{R}_d$.
- The process N^i is the counting process of i-th coordinate.
- (X, Y, N, L) is also a Markov process under generalized exponential assumption.

Theorem

Joint Laplace transform of $X(T), Y(T), n(T) - n(t), L(T) - L(t)$ is given by

$$\mathbb{E}(e^{-(u,X(T))-(v,Y(T))-(w,n(T)-n(t))-(z,L(T)-L(t))}|\mathcal{F}_t)$$
$$= e^{A(t,T)-(B(t,T),X(t))-(C(t,T),Y(t))}$$
$$u \in \mathbb{R}_{|\mathbb{A}|(n+1)}, v \in \mathbb{R}_{|\mathbb{M}|(m+1)}, w \in \mathbb{R}_d, z \in \mathbb{R}_d.$$

where A, B, C solve following system of ODE's:

$$\partial_t C(t,T) = -H'C(t,T) + r(t, B(t,T), w, z), \qquad C(T,T) = v,$$
$$\partial_t B(t,T) = -G'B(t,T) + q(t, B(t,T), w, z), \qquad B(T,T) = u,$$
$$\partial_t A(t,T) = -\sum_{\mathcal{I} \in \mathbb{M}} \Big\{ \theta_{\mathcal{I}}(L_{P_{\mathcal{I}}}(\widehat{K}_{\mathcal{I}} C(t,T)) - 1)$$
$$+ \mu_{\mathcal{I}}(t)\big[e^{-\sum_{i\in\mathcal{I}} w_i} L_{Q_{\mathcal{I}}}(t, K_{\mathcal{I}} B(t,T)) - 1\big] \Big\}, \qquad A(T,T) = 0.$$

with

$$q(t,x,w,z) = \sum_{(\mathcal{K},\mathcal{I}) \in \mathbb{A}} \mathrm{c}_{\mathcal{K},\mathcal{I}} \otimes \mathrm{e}_{2,n+1} \cdot \alpha_{\mathcal{K},\mathcal{I}}(t)\big(e^{-\sum_{i\in\mathcal{I}} w_i} L_{R_{\mathcal{K},\mathcal{I}}}(t, K_{\mathcal{I}} x, z) - 1\big)$$

$$r(t,x,w,z) = \sum_{\mathcal{I} \in \mathbb{M}} \mathrm{c}_{\mathcal{I}} \otimes \mathrm{e}_{2,m+1} \cdot \beta_{\mathcal{I}}(t)\big(e^{-\sum_{i\in\mathcal{I}} w_i} L_{Q_{\mathcal{I}}}(t, K_{\mathcal{I}} x, z) - 1\big),$$

$$L_{Q_{\mathcal{I}}}(t,v,z) = \int_{E_{\mathcal{I}}} e^{-\left(v, \sum_{\mathcal{J}\in \mathrm{An}(\mathcal{I})} \mathrm{a}_{\mathcal{I},\mathcal{J}} \otimes \overline{\psi}_{\mathcal{I},\mathcal{J}}(t, z_{\mathcal{I}})\right) - (z, \xi_{\mathcal{I}}(z_{\mathcal{I}}))} Q_{\mathcal{I}}(\mathrm{d}z_{\mathcal{I}}), \quad v \in \mathbb{R}$$

$$L_{R_{\mathcal{K},\mathcal{I}}}(t,v,z) = \int_{E_{\mathcal{I}}} e^{-\left(v, \sum_{\mathcal{J}\in \mathrm{An}(\mathcal{I})} \mathrm{a}_{\mathcal{I},\mathcal{J}} \otimes \overline{\psi}_{\mathcal{I},\mathcal{J}}(t, z_{\mathcal{I}})\right) - (z, \xi_{\mathcal{I}}(z_{\mathcal{I}}))} R_{\mathcal{K},\mathcal{I}}(\mathrm{d}z_{\mathcal{I}}),$$

$$L_{P_{\mathcal{I}}}(t,u) = \int_{\mathbb{R}} e^{-\left(u, \overline{\phi}_{\mathcal{I}}(t,z)\right)} P_{\mathcal{I}}(\mathrm{d}z), \quad u \in \mathbb{R}_n.$$

Bibliography

Y. Bessy-Roland, A. Boumezoued, and C. Hillairet.
Multivariate hawkes process for cyber insurance.
Annals of Actuarial Science, 15(1):14–39, 2021.

T. R. Bielecki, J. Jakubowski, and M. Nieweglowski.
Construction and Simulation of Generalized Multivariate Hawkes Processes
Methodology and Computing in Applied Probability, 2022.

A. Boumezoued.
Population viewpoint on Hawkes processes.
Advances in Applied Probability, 48(2), 463–480, (2016).

P. Brémaud and L. Massoulié.
Stability of nonlinear Hawkes processes.
Ann. Probab., 24(3):1563–1588, 1996.

Bibliography

L. Cui, A. Hawkes, H. Yi
An elementary derivation of moments of Hawkes processes.
Advances in Applied Probability, 52(1), 102-–137, (2020).

T.J. Liniger.
Multivariate hawkes processes.
PhD thesis, ETH Zurich, 2009.

Thank you for your attention !!!

On envelopes created by circle families in the plane

Takashi Nishimura
Faculty of Environment and Information Sciences,
Yokohama National University, Japan
(joint work with Yongqiao Wang)

Envelopes of planar curve families have fascinated many pioneers since the dawn of differential analysis. In most typical cases, straight line families have been studied. However, even for envelopes created by straight line falimies, to our surprize, there were several unsolved problems until very recently. In my talk at WAAS, recently discovered answers to these problems were explained.

On the other hand, circle families in the plane are non-negligible families because the envelopes of them have already had important applications to Industry. In this talk, firstly, as one of important applications of envelopes of circle families to Industry, the so-called "Mohr failure envelope" is introduced. After that, a general theory for envelopes of circle families shall be explained.

On envelopes created by circle families in the plane (a joint work with Yongqiao Wang)

Takashi Nishimura
(Yokohama National University)

Reference

[WN] Yongqiao Wang and T.N., *Envelopes created by circle families in the plane*, preprint.
(available at https://arxiv.org/abs/2301.04478)

1

§1. Soil Mechanics

Circle families in the plane are non-negligible families because the envelopes of them have already had important applications. As one of application of circle family in the plane, Let me first explain the so-called Mohr failure envelope in the field "Soil Mechanics".

2

In analysis of the stability of soil masses, the shear strength τ_f of a soil at a point on a particular plane is expressed as a linear function of the effective normal stress σ_f at failure:

$$\tau_f = \sigma_f \tan \varphi + c,$$

where φ and c are the *angle of shearing resistance* and *cohesion intercept* respectively. A method using *Mohr circles* to obtain the shear strength parameters φ and c can be found (for instance) in "R.F. Craig, *Craig's soil mechanics, Seventh edition*, Taylor and Francis Group Press, New York, 2004. ISBN: 9780415332941". A brief description of this method is given as follows.

3

The stress state of a soil can be represented by a Mohr circle which is defined by the effective principal stresses σ_1 and σ_2. The center and the radii of the Mohr circle are $(\frac{\sigma_1+\sigma_2}{2}, 0)$ and $\frac{\sigma_1-\sigma_2}{2}$, respectively. By experiments, one can obtain some values of effective principal stresses σ_1 and σ_2 at failure. The Mohr circles in terms of effective principal stress are drawn in Figure 1.

4

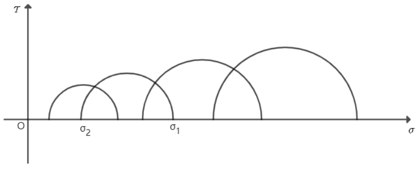


Figure 1

5

The envelope created by Mohr circles is called the *Mohr failure envelope* which may be a slightly curved curve. Then the shear strength parameters φ and c can be obtained by approximating the curved envelope to a straight line, namely the slope of the straight line equals $\tan\varphi$ and the intercept of straight line on the vertical axis is c (see Figure 2).

6

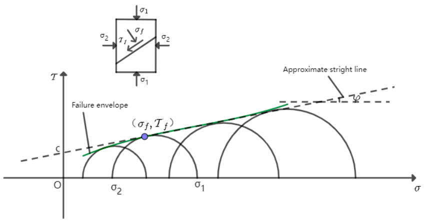


Figure 2

7

The so-called "liquefaction phenomenon" is one of contemporary important problems especially in the country where people can not avoid large-scale earthquakes. Therefore, Mohr failure envelope is a significant notion for industry.

In order to understand the mechanism of "liquefaction phenomenon" well and in order to find an effective measure against real liquefaction phenomena, it seems important to construct a general theory of the envelopes created by circle families.

8

§2. Envelopes of circle families

For a point P of $\mathbb{R}^2$ and a positive number λ, the circle $C_{(P,\lambda)}$ centered at P with radius λ is naturally defined as follows, where the dot in the center stands for the standard scalar product.

$$C_{(P,\lambda)} = \left\{(X,Y) \in \mathbb{R}^2 \,\middle|\, ((X,Y)-P)\cdot((X,Y)-P) = \lambda^2\right\}.$$

For a curve $\gamma : I \to \mathbb{R}^2$ and a positive function $\lambda : I \to \mathbb{R}_+$, the circle family $\mathcal{C}_{(\gamma,\lambda)}$ is naturally defined as follows. Here, $\mathbb{R}_+$ stands for the set consisting of positive real numbers.

$$\mathcal{C}_{(\gamma,\lambda)} = \left\{C_{(\gamma(t),\lambda(t))}\right\}_{t\in I}.$$

9

It is reasonable to assume that at each point $\gamma(t)$ the normal vector to the curve γ is well-defined. Thus, we easily reach the following definition.

Definition 1 A curve $\gamma : I \to \mathbb{R}^2$ is called a *frontal* if there exists a mapping $\nu : I \to S^1$ such that the following identity holds for each $t \in I$, where S^1 is the unit circle in $\mathbb{R}^2$.

$$\frac{d\gamma}{dt}(t)\cdot\nu(t) = 0.$$

For a frontal γ, the mapping $\nu : I \to S^1$ given above is called the *Gauss mapping* of γ.

Hereafter, the curve $\gamma : I \to \mathbb{R}^2$ for a circle family $\mathcal{C}_{(\gamma,\lambda)}$ is assumed to be a frontal.

10

In this talk, the following is adopted as the definition of an envelope created by a circle family.

Definition 2 Let $\mathcal{C}_{(\gamma,\lambda)}$ be a circle family. A mapping $f : I \to \mathbb{R}^2$ is called an *envelope* created by $\mathcal{C}_{(\gamma,\lambda)}$ if the following two hold for any $t \in I$.

(1) $\frac{df}{dt}(t)\cdot(f(t)-\gamma(t)) = 0$.

(2) $f(t) \in C_{(\gamma(t),\lambda(t))}$.

11

Problem 1 *Let $\gamma : I \to \mathbb{R}^2$ be a frontal with Gauss mapping $\nu : I \to S^1$ and let $\lambda : I \to \mathbb{R}_+$ be a positve function.*

(1) Find a necessary and sufficient condition for the circle family $\mathcal{C}_{(\gamma,\lambda)}$ to create an envelope in terms of γ, ν and λ.

(2) Suppose that the circle family $\mathcal{C}_{(\gamma,\lambda)}$ creates an envelope. Then, find a parametrization of the envelope created by $\mathcal{C}_{(\gamma,\lambda)}$ in terms of γ, ν and λ.

12

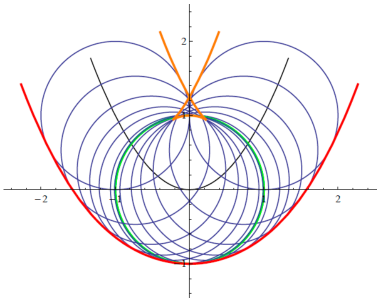
Example 1 *Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be the mapping defined by $\gamma(t) = \left(t^3, t^6\right)$. Set $\nu(t) = \frac{1}{\sqrt{4t^6+1}}\left(-2t^3, 1\right)$. It is clear that the mapping γ is a frontal with Gauss mapping $\nu : \mathbb{R} \to S^1$. Let $\lambda : \mathbb{R} \to \mathbb{R}_+$ be the constant function defined by $\lambda(t) = 1$. Then, it seems that the circle family $\mathcal{C}_{(\gamma,\lambda)}$ creates envelopes. Thus, we can expect that the created envelopes can be obtained by the well-known method.*

13

Set $F(x,y,t) = \left(x - t^3\right)^2 + \left(y - t^6\right)^2 - 1$. Then, we have the following.

$$
\begin{aligned}
& \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \exists t \text{ s.t. } F(x,y,t) = \frac{\partial F}{\partial t}(x,y,t) = 0\right\} \\
= & \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \exists t \text{ s.t. } \left(x - t^3\right)^2 + \left(y - t^6\right)^2 - 1 = -6t^2\left(x - t^3\right) - 12t^5\left(y - t^6\right) = 0\right\} \\
= & \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \exists t \text{ s.t. } \left(x - t^3\right)^2 + \left(y - t^6\right)^2 - 1 = t^2\left(\left(x - t^3\right) + 2t^3\left(y - t^6\right)\right) = 0\right\} \\
= & \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 = 1\right\} \\
& \qquad \bigcup \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \left(x - t^3\right)^2 + \left(y - t^6\right)^2 - 1 = 0,\ x = t^3 - 2t^3\left(y - t^6\right)\right\} \\
= & \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 = 1\right\} \\
& \qquad \bigcup \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \left(-2t^3\left(y - t^6\right)\right)^2 + \left(y - t^6\right)^2 = 1,\ x = t^3\left(1 - 2y + 2t^6\right)\right\} \\
= & \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2 + y^2 = 1\right\} \\
& \qquad \bigcup \left\{\left(t^3 \mp \frac{2t^3}{\sqrt{4t^6+1}}, t^6 \pm \frac{1}{\sqrt{4t^6+1}}\right) \in \mathbb{R}^2 \;\middle|\; t \in \mathbb{R}\right\}.
\end{aligned}
$$

14



15

In order to solve Problem 1, we prepare several terminologies which can be derived from a frontal $\gamma : I \to \mathbb{R}^2$ with Gauss mapping $\nu : I \to S^1$ and a positive function $\lambda : I \to \mathbb{R}_+$. For a frontal $\gamma : I \to \mathbb{R}^2$ with Gauss mapping $\nu : I \to S^1$, following "T. Fukunaga and M. Takahashi, *Existence and uniqueness for Legendre curves*, Journal of Geometry, **104** (2013), 297–307", set

$$\mu(t) = J(\nu(t)),$$

where J is the anti-clockwise rotation by $\pi/2$. Then we have a moving frame $\{\mu(t), \nu(t)\}_{t\in I}$ along the frontal γ . Set

$$\ell(t) = \frac{d\nu}{dt}(t) \cdot \mu(t), \quad \beta(t) = \frac{d\gamma}{dt}(t) \cdot \mu(t).$$

16

The following definition is the key of this talk.

Definition 3 ([WN], KEY DEFINITION) Let $\gamma : I \to \mathbb{R}^2$, $\lambda : I \to \mathbb{R}_+$ be a frontal with Gauss mapping $\nu : I \to S^1$ and a positive function respectively. Then, the circle family $\mathcal{C}_{(\gamma,\lambda)}$ is said to be *creative* if there exists $\tilde{\nu} : I \to S^1$ such that the following identity holds for any $t \in I$.

$$\frac{d\lambda}{dt}(t) = -\beta(t)\,(\tilde{\nu}(t) \cdot \mu(t))\,.$$

17

By definition, any family of concentric circles with smoothly expanding radii is not creative, and it is clear that such the circle family does not create an envelope.

Theorem 1 ([WN]) *Let $\gamma : I \to \mathbb{R}^2$ be a frontal with Gauss mapping $\nu : I \to S^1$ and let $\lambda : I \to \mathbb{R}_+$ be a positive function. Then, the following hold.*

(1) The circle family $\mathcal{C}_{(\gamma,\lambda)}$ creates an envelope if and only if $\mathcal{C}_{(\gamma,\lambda)}$ is creative.

(2) Suppose that the circle family $\mathcal{C}_{(\gamma,\lambda)}$ creates an envelope $f : I \to \mathbb{R}^2$. Then, the created envelope f is represented as follows.

$$f(t) = \gamma(t) + \lambda(t)\tilde{\nu}(t).$$

18

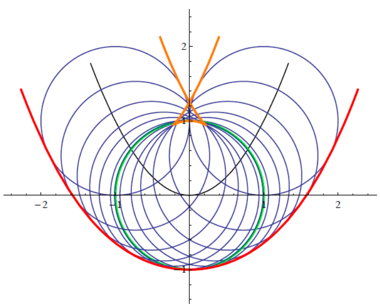
Example 2 We examine Example 1 by applying Theorem 1. In Example 1, $\gamma : \mathbb{R} \to \mathbb{R}^2$ is given by $\gamma(t) = \left(t^3, t^6\right)$. Thus, we can say that $\nu : \mathbb{R} \to S^1$ and $\mu : \mathbb{R} \to S^1$ are given by $\nu(t) = \frac{1}{\sqrt{4t^6+1}}\left(-2t^3, 1\right)$ and $\mu(t) = \frac{1}{\sqrt{4t^6+1}}\left(-1, -2t^3\right)$ respectively. Moreover, the radius function $\lambda : \mathbb{R} \to \mathbb{R}$ is the constant function defined by $\lambda(t) = 1$. Thus,

$$\frac{d\lambda}{dt}(t) = 0.$$

By calculation, we have

$$\beta(t) = \frac{d\gamma}{dt}(t) \cdot \mu(t) = \frac{-3t^2(1 + 4t^6)}{\sqrt{4t^6 + 1}}.$$

19



20

Therefore, the unit vector $\tilde{\nu}(t) \in S^1$ satisfying

$$\frac{d\lambda}{dt}(t) = -\beta(t)\,(\tilde{\nu}(t) \cdot \mu(t))$$

exists and it must have the form

$$\tilde{\nu}(t) = \pm\nu(t) = \frac{\pm 1}{\sqrt{4t^6+1}}\left(-2t^3, 1\right).$$

Hence, by the assertion (1) of Theorem 1, the circle family $\mathcal{C}_{(\gamma,\lambda)}$ creates an envelope $f : \mathbb{R} \to \mathbb{R}^2$.

21

By the assertion (2) of Theorem 1, f is parametrized as follows.

$$\begin{aligned} f(t) &= \gamma(t) + \lambda(t)\tilde{\nu}(t) \\ &= \left(t^3, t^6\right) \pm \frac{1}{\sqrt{4t^6+1}}\left(-2t^3, 1\right) \\ &= \left(t^3 \mp \frac{2t^3}{\sqrt{4t^6+1}}, t^6 \pm \frac{1}{\sqrt{4t^6+1}}\right). \end{aligned}$$

22

Thank you for your listening!

23

Calcium waves sustained by calcium influx through mechanically activated channels in the cell membrane

Zbigniew Peradzyński

Military Technological University and Institute of Fundamental Technological Research PAS, Poland

(joint work with Bodan Kaźmierczak and Sławomir Białecki)

The work is devoted to the mathematical modeling of fast calcium waves propagating in some cells. According to the suggestion of biologists, this type of waves exists due to the complicated mechanisms of the influx of calcium from the extracellular space through mechanically operated calcium channels placed in the cell membrane. A change in the concentration of calcium in the cell causes the reorganization of the network composed of actin-myosin filaments. Under the influence of local forces exerted by these fibers, ion channels in the cell membrane are opened. At the same time, excess calcium is pumped out of the cell by several types of pumps located in the cell membrane. All this together leads to the possibility of wave propagation in the form of homoclinic pulses of calcium concentration. We start from the construction of the model in 3-D. Then we derive 1-D nonlocal approximation, which as it turns out, can be still approximated by a FitzHugh Nagumo type of system. The theoretical model will also be supported by numerical calculations.

Calcium waves sustained by calcium influx through mechanically activated channels in the cell membrane

Zbigniew Peradzyński * ,Bogdan Kaźmierczak, Sławomir Białecki**,**
* Warsaw Military University of Technology (earlier in Faculty of Mathematics, Informatics and Mechanics, University of Warsaw),
**Institute of Fundamental Technological Research.
Workshop on Mathematics for Industry 2023, Warsaw

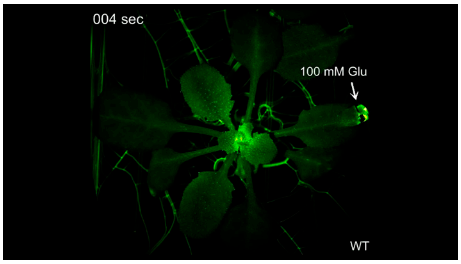
Provocative question:

Can plants be aware of the danger?

Please see the video:
https://www.youtube.com/watch?app=desktop&&v=7-3yFcZSyvo

„"Supplying glutamate directly to the tip of one leaf creates a strong wave of calcium across the entire plant, visualized by fluorescent light. This video is part of research by UW–Madison botany professor Simon Gilroy that shows how waves of calcium crisscrossing a plant help it respond to attacks by preparing for future threats.
The work was published in Science in September of 2018".

It turns out that plants or their parts can communicate with each other (e.g by sending sigals calcium waves), preparing thus for unpleasant consequences



By waves we mean travelling waves, special solutions: $\boldsymbol{u} = \boldsymbol{U}(x - ct)$ to Reaction-Diff. equations (c-const)

- Waves are usually associated with the wave equation or with hyperbolic systems. However hyperbolic equations are almost nonexisting in biology. One predominantly encounters parabolic equations or semilinear parabolic systems – Reaction-Diffsion Systems.
- The travelling waves in R-D eqs are appearing as an interplay between the diffusion and nonlinearity.

Single reaction –diffusion equation

$$\frac{\partial}{\partial t} u = D\Delta u + F(u)$$

If u(t,x) – density of individuals, F(u) = ru(1-u/K), then one can speak of a simple model in population dynamics. The diffusive term reflects the fact that individuals are moving erratically. The reaction term F(u) is responsible for the birth and death processes.

Here travelling wave solutions are heteroclinic fronts. As F is monostable, because u=0 is unstable equilibrim, there are solutions for an arbitrary speed $\geq c_0$.

Bistable case; the wave speed is uniquely determined!

F(u) has two stable: u_1, u_3 and one unstable (u_2) equilibrium.

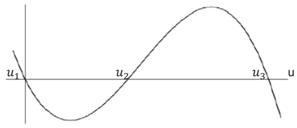


Fig. An example of a bistable source function:

$$F(u) = -A(u - u_1)(u - u_2)(u - u_3)$$

An example of a travelling front

The following bistable reaction diffusion equation with a cubic (**bistable)** source term

$$\frac{\partial}{\partial t}\boldsymbol{u} = \boldsymbol{D}\frac{\partial^2}{\partial x^2}\,\boldsymbol{u} \;-\boldsymbol{A}\,\boldsymbol{u}(\boldsymbol{u}-\boldsymbol{a})(\boldsymbol{u}-\mathbf{1})$$

has (D=1, A=1) following travelling front solutions

$$u = \frac{1}{1+\exp\left(\frac{\pm x - vt}{\sqrt{2}}\right)}$$

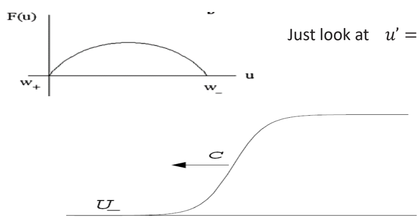
where $v = \sqrt{2}\,(\frac{1}{2} - a)$ defines the propagation speed.

Monostable reaction term – waves can propagate with an arbitrary speed grater then some v_0. The minimal speed makes physical sense)

The case of $F(u)$= ru(1-u/K) is a good example of a monostable reaction term. It has two zeros:

Unstable state u=0 and stable state u=K

Just look at $u' = ru(1 - K/u)$



Theory based on single reaction diffusion equation predicts travelling waves in the form of heteroclinic fronts, joining two stable (in the bistable case) equilibria of the source term, whereas observed experimentally calcium waves are of homoclinic type. Thus, such simplified theory describes properly only the front part of the wave. To obtain the shape of a homoclinic, the additional equation for "recovery variable" is usually added.

In the proposed here theory for CICI waves this additional equation appears in a natural way.

Ecology, Population dynamics

Reaction Diffusion System (interacting species)

$$\frac{\partial}{\partial t}u_1 = D_1\frac{\partial^2}{\partial x^2}u_1 + ru_1(1 - A_1 \cdot u)$$
$$\frac{\partial}{\partial t}u_2 = D_2\frac{\partial^2}{\partial x^2}u_2 + ru_2(1 - A_2 \cdot u)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\frac{\partial}{\partial t}u_n = D_n\frac{\partial^2}{\partial x^2}u_n + ru_n(1 - A_n \cdot u)$$

The matrix $A = (A_1, A_2, \ldots, A_m)$ describes the interactions between the species. If the entries are positive we have the case of species competing for food.

MONOTONE SYSTEMS

Definition. The system

$$\frac{\partial}{\partial t}\boldsymbol{u_i} = \boldsymbol{D_i}\frac{\partial^2}{\partial x^2}\,\boldsymbol{u_i} + \boldsymbol{F_i}(\boldsymbol{u_1}, \ldots, \boldsymbol{u_n}),\ \ \boldsymbol{i} = \mathbf{1}, \ldots, \boldsymbol{n}$$

is called monotone if $\frac{\partial F_i}{\partial u_j} > 0$ for $\boldsymbol{i} \neq \boldsymbol{j},\ \ \boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \ldots, \boldsymbol{n}$

is satisfied for all u. Such systems arise in numerous application in chemical kinetics and populations dynamics.

The maximum principle appears to be valid for monotone systems. Its applicability allows us to formulate the results on wave existence, stability and velocity similar to those for the scalar equation.

Comments on a multistable case

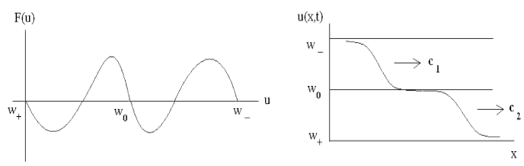


Fig. 2. Bistable nonlinearity with a stable intermediate zero w_0 (left). System of two waves, the speed c_2 of the lower wave is greater than the speed c_1 of the upper one (right).

Calcium waves were discovered in 1977 on medaka fish egg.

John C. Gilkey, Lionel f. Jaffe, Ellis B. Ridgway, and George T. Reynolds *„A FREE CALCIUM WAVE TRAVERSES THE ACTIVATING EGG OF THE MEDAKA, ORYZIAS LA TIPES”,,* Journ. Cell Biology" Vol. 76, 1978

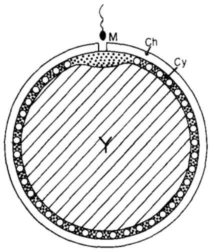
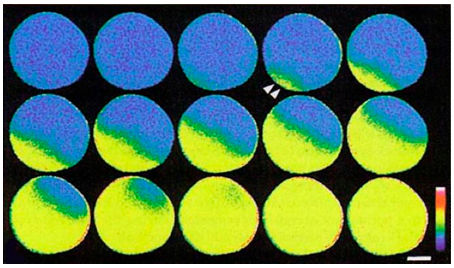
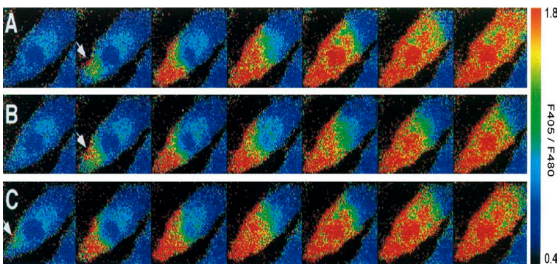


FIGURE 1 Diagram of unfertilized medaka egg (1.2-mm diameter). A sperm will cross the chorion (*Ch*) via the micropyle (*M*), enter the cytoplasm (*Cy*) and initiate a wave of cortical vesicle secretion. Vesicles are indicated by small circles. The bulk of the egg is occupied by a membrane-bounded yolk compartment (*Y*). The cytoplasmic thickness (0.03 mm) is exaggerated, and oil droplets are omitted for clarity.

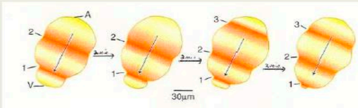
- Signals can be transmitted by various means – calcium concentration waves among the others. After the fertilization of an egg the wave sprading on its surface is generated, which changes the status of an egg. The second sperm can not enter the egg.



The calcium wave through moving amoebae. Speed 15 μm/s. (L. Jaffe)



Deformation accompanying calcium waves on the surface of fertilized egg



From L. Jaffe

Calcium waves (first seen on the fertilizing medaka egg) turned out to be quite common. They can propagate both in individual cells and in tissues. The range of their speed: 1nm/s – 30 cm/s (nearly a billion fold) falling into four speed –based groups (after L. Jaffe)

In our lecture we will be interestet in CICI fast waves (see diagram below).

CICI WAVES

The mechanism of propagation of **CICR** waves is based on autocatalytic release of calcium from the internal stores (e.g. endoplasmic reticulum) located in the cells.

CICI waves. According to L. Jaffe this cannot explain the speed of the second group of „fast waves" . Their speed can be by two orders higher. Such waves are also observed in cells not having internal stores of calcium.

Thus: Stretch-activated ion channels in the membrane are responsible for the calcium delivery from the extracellular space.

CELL is extremally complex! (Nobel Prize 2013).

The cell membrane is equipped with

a) ion channels (MECHANICALLY, chemically or electrically controlled) through which ions are admitted into the cell interior.

b) There are pumps in the membrane - at least two types:

- ATP type - efficient at low Ca^{++} concentrations
- sodium-calcium exchangers; very efficient at high Ca^{++} concentrations.

Thanks to them, balance in the cell can be restored.

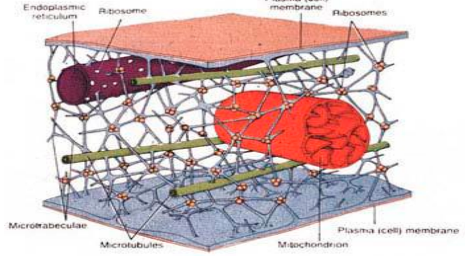
Mechanically operated ion channels (stretch activated) are opened when the membrane is stretched.

Inside the cel we have

1. Cytoplasm
2. Actin filaments
3. Internal stores of calcium (endoplasic reticulum)
4. Other important ingredients as: ion channels and ion pumps located in the cell membrane.

- As the Ca concentration increases, the filaments are increasingly connected by myosin bridges and the filament network contracts.
- The filaments also serve as routes along which various materials in bags (vesicles) are transported by appropriate motors. (F=2.7 pN). See for example : https://learn.genetics.utah.edu/content/cells/vesicles/

Model of a cell



(from the lecture by Kizylvova)

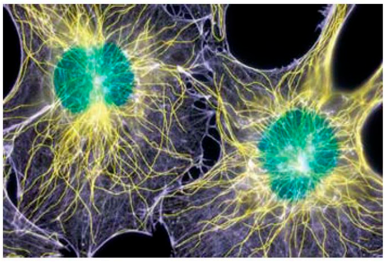
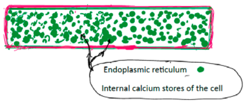


Figure 1: *Actin filaments (dark and thin) and microtubules (bright and thick)* –

Coming back to Ca waves

There are already well known and well researched CICR waves i.e. „Calcium Induced Calcium Released" waves (L. Jaffe) . The simplest theoretical description is based on single reaction diffusion equation with a bistable source term. For a small excess of calcium above the equilibrium concentration, calcium is absorbed into internal stores. After exceeding a certain threshold value (the second zero of the source function) calcium is released from the internal stores of the cell in an autocathalitic reaction, untill its concentration reaches the next equilibrium value (the third zero of source function).



Lionel Jaffe Hypothesis

According to L. Jaffe, the CICR mechanism cannot be responsible for high speed of CICI waves (see diagram).

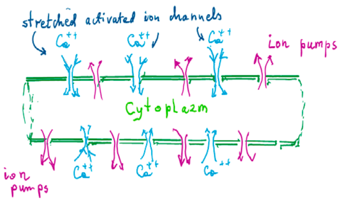
It is known that:

Stretching the membrane activates the ion channels and calcium can enter the cell from the extracellular space.

Hypothesis: when the calcium concentration grows the actin-myosin network is reorganized – the filament network contracts. Consequently, filaments are pulling the membrane. Mechanically stimulated channels are opened and calcium enters the cell. This mechanism (calcium induced calcium influx) supports the wave propagation.

Hypothetical CICI Waves – the subject of our modelling

- **Accorfing to L. Jaffe in this case calcium from the extracellular space enters the cell through mechanicaly activated ion channels located in the cell membrane. In the extracellular space Ca^{++} concentration is by two orders higher than in the cell internal stores. The channels are opened when the membrane is stretched.**



Calcium pumps

Calcium pumps are ion transporters found in the cell membrane. They are responsible for active transport of calcium out of the cell, keeping the intracellular calcium concentration 10 000 times lower than the extracellular. The plasma membrane Ca^{2+} ATPase and sodium-calcium calcium exchanger are the main regulators of intracellular Ca^{2+} concentration. The first type is efficient at low Ca concentration, whereas the second type is extremely efficient at higher concentrations.

They also seem to play the crucial role in supporting the CICI Waves!

Assumptions.

1. The contraction of the actomyosin network results in appearing of so called “traction forces”. However, the effect of contraction following the increase of calcium concentration appears with some delay –relaxation time is needed to form the myosin bridges

2. The calcium can enter from the intercellular space through the mechanically stimulated ion channels located in the cell membrane

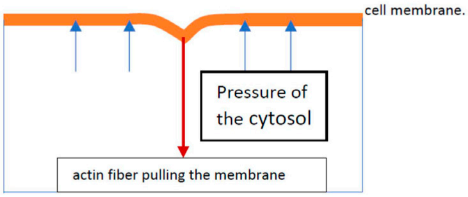
3. The mechanical stimulation of the membrane is caused by the actomyosin network - cortex. The fibers of the cortex as well as the rest of actomyosin network in the cell are subject to the contraction whenever the calcium concentration in the cell cytoplasm increases.

As the calcium concentration increases, the myosin filaments become more and more connected through the increasing number of myosin Bridges. This leads to the contraction of the filament network.

This contraction influences the shape of the cell. If we imagine the ideal cell of a cylindrical shape, then the cell radius will be reduced. Therefore, at first glance, we should not expect any stretching of the cell membrane.

This is however macroscopic view. Microscopically the membrane will be very unsmooth. Funnel-shaped depressions will appear under the influence of pulling forces, in places where the filaments are anchored. So in spite of this that the average radiuce gets smaller we will have the membrane stretching as its shape becames more complex.

When the wave passes, the cell radius shrinks. So how can we have stretching ? locally we expect the following picture



Suppose, the ion channels are openned whenever the membrane is streched. Then permanent stretch :

High calcium concentration over a long period of time would lead to the cell death. Therefore, a permanent state of stretch should not result in a continuous influx of calcium.

Experiment: oscillatory stretching leads to Ca^{++} influx proportional to the amplitude and oscillations frequency.

This suggests that the calcium influx should rather be related to the speed of membrane stretching !

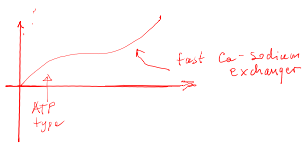
H.1. Therefore, if **n** is an internal unit vector normal to the cell membrane and **F** is the force acting on unit membrane area, then the calcium influx (flux per unit area) is proportional to the **positive part of the time derivative of the force** acting on the unit surface.

$$Ca^{++}\,influx \;\sim\; \left[\frac{\partial}{\partial t}(\boldsymbol{n}\cdot\boldsymbol{F})\right]_{+}$$

Positive part, because only stretching counts. One can show that otherwise the Ca concentration may become negative !

Taking into account the pumps p(c)

$p(c)$:



$$total\; Ca^{++}\,influx \;\sim\; \left[\frac{\partial}{\partial t}(\boldsymbol{n}\cdot\boldsymbol{F})\right]_{+} - p(u)\text{ -}$$

This is the boundary condition for the Ca diffusion equation.

Now we arrived at the **MECHANICAL PROBLEM**:

Determine the forces acting on the membrane ; i.e. forces resulting from the actin filaments attached to it.

In principle two approaches seem to be possible:

a) **Calculate the distribution of forces on each filament of the contracting network due to the appearance of myosin bridges. In particular those anchored in the membrane. Then find the shape of deformed membrane.**

This seems hopelessly difficult !

Continuum mechanical approach ?

b) In mathematical biology (Murray, *Mathematical Biology*) , the cell is often treated as an elastic (or viscoelastic) body, and the forces associated with the contraction (traction forces) are expressed by the traction tensor. This description is very similar to termo-elasticity. Ca^{++}concentration plays the role of the temperature (in fact $-T$).

Applying this idea, we arrive at a system of three equations.

The system consists of

1. The equation of motion of the viscoelastic body, i.e cytoplasm with the filament network. The equation of motion (linear approximation) for the displacement vector $u(t,x)$ must be equipped with proper boundary conditions. Under the influence of traction forces the membrane is deflected. So basically, we should know the elasticity of the membrane. However, to estimate the forces acting on the membrane, one can assume that the membrane is stiff and not deformed. In such a case we have simple b-dry condition: $u(R) = 0$

Let us remind that if the initial position of a material point is x and it position changes to $\tilde{x}$ then $u(x) = \tilde{x} - x$.

2. Relaxation equation for the traction tensor $\widehat{T}$ with a given equilibrium form $\widehat{T}^*(c)$. We have $\widehat{T}(t,x) = \widehat{T}^*(c(t,x))$ for very slow changes of the concentration $c(t,x)$.

3. The diffusion equation for calcium concentration $c(t,x)$ and nonlinear boundary condition expressing the influx of calcium (by ion channels and ion pumps) caused by positive part of time derivative of traction forces acting on the membrane.

In fact, the diffusion of calcium in the cell is quite a complicated process because of the buffers - proteins that can attach and release calcium ions. This can be described by a system of equations for the diffusion reaction. If we use one equation as here, D should be treated as the effective diffusion coefficient.

Treating (idealized) cell as an Infinite cylinder we could try to solve:

(1) $\boldsymbol{\rho \frac{\partial^2}{\partial t^2} u - \nu_2 \Delta \dot{u} + (\nu_1 + \nu_2)\nabla \mathrm{div}\dot{u}} = \mu \, \Delta \, \boldsymbol{u} + (\boldsymbol{\mu} + \boldsymbol{\lambda})\boldsymbol{\nabla div u} + \boldsymbol{div}\, \widehat{\boldsymbol{T}}(\boldsymbol{c})$

with b-dry condition: $\boldsymbol{u}(t,R) = 0$

(2) $\frac{\partial}{\partial t}\widehat{\boldsymbol{T}} = \beta[\widehat{\boldsymbol{T}}^*(\boldsymbol{c}) - \widehat{\boldsymbol{T}}]$, where $\widehat{\boldsymbol{T}}^*(\boldsymbol{c})$ - known (e.g. linear)

(3) $\frac{\partial}{\partial t} c = D\Delta c$ inside the cell

$D\frac{\partial}{\partial r}c(t,R,z) = Q\left[\frac{d}{dt}\sigma_{rr}(t,R,z)\right]^+ - p(u)$ on the b-dry

suplied by initial conditions for u, T, c .

Comment. The first equation, the equation of motion can be simplified by omitting the dynamical term $\boldsymbol{\rho \frac{\partial^2}{\partial t^2} u}$ and possibly the viscouse terms $\boldsymbol{\nu_2 \Delta \dot{u} + (\nu_1 + \nu_2)\nabla \mathrm{div}\dot{u}}$.

Then one obtains an eliptic system for the displacement $\boldsymbol{u}(t,x)$.

In principle it is possible to solve the above system numerically. For reasons discussed below, we decided on a slightly roundabout but simpler route.

In presented here equations we assumed the medium to be isotropic. However, the anisotropy, can be important as it can greatly influence the speed of waves. Indeed, the network structure - the way the filaments are connected, affects the transfer of forces acting on the membrane through the interconnected fibers.

Depending on the way the filament network is interconnected, calcium channels may be opened in places more or less distant from the front of the wave of increased calcium concentration. Thus, we should solve systems with different degree of anisotropy.

To avoid all these complications, we chose a slightly different modeling route.

Intermediate way, Here $\widehat{\boldsymbol{T}} = \tau \mathrm{I}$

Instead, we chose the intermediate solution. By solving the equations of mechanical equilibrium,

$$\boldsymbol{\mu}\, \Delta\, \boldsymbol{u} + (\boldsymbol{\mu} + \boldsymbol{\lambda}) \boldsymbol{\nabla div u} + \boldsymbol{div}\, \widehat{\boldsymbol{T}}(\boldsymbol{c}) = \mathbf{0}$$

assuming that the solution is independent of the axial variable, and for isotropic traction tensor $\widehat{\boldsymbol{T}} = \tau \mathbf{I}$ we can estimate the forces acting on the membrane as

$$\sigma_{rr}(t, R) = \frac{1}{\pi R^2} \int_0^R \tau(t, r) 2\pi r dr$$

Since the Ca influx is proportional to time derivative of σ_{rr}

$$\frac{\partial}{\partial t} \sigma_{rr}(t, R) = \frac{1}{\pi R^2} \int_0^R \frac{\partial}{\partial t} \tau(t, r) 2\pi r dr$$

we have $\frac{\partial}{\partial t} \tau = \beta\, [\tau^*(c) - \tau]$, so

$$\frac{\partial}{\partial t} \sigma_{rr}(t, R) = \frac{\beta}{\pi R^2} \int_0^R [\tau^*(c) - \tau]\, 2\pi r dr$$

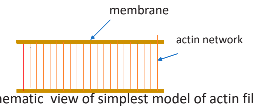
Smearing the force (interconnected filaments)

The previous step do not include transmition of force from one point to another by interconnected filaments. To take this into account we introduce a kind of smearing out of forces acting on the membrane through an averaging integral operator (convolution wit K_σ)

$$D\frac{\partial}{\partial r}c(t,R.z) = A\left\{K_\sigma * \left[\frac{2}{R^2}\int_0^R \left(\tau^*(c(t,r,z)) - \tau(t,r,z)\right) r dr\right]^+ - p(c)\right\}$$

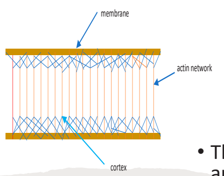
where in numerical simmulations we took $K_\sigma = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{z^2}{2\sigma^2})$.

This non-local mechanism embodies the idea of L. Jaffe



Schematic view of simplest model of actin fibers network in 2D. When Ca concentration increases the fibers contract pulling the membrane. This arrangment of fibers corresponds to completely anisotropic case (no myosin bridges between filaments). The force is not transfered between filaments – local mechanism.

$$K \sim \delta(x)$$



Nonlocal mechanism of propagation

- This mechanizm is nonlocal. The filaments are interconnected by myosin bridges. Their number grows with Ca concentration.
- The force that appears in one place is transferred by the tangled fibers to other neighboring ones. Thus the channels are openned ahead of the propagating wave of Ca^{++} concentration. This accelerates the wave propagation.

Numerical computations

All numerical computations were done for the diffusion coefficient **D**=1.
The source term:

$[\boldsymbol{K}(\boldsymbol{0,25u+0.1u^2-\tau})]_{+}-\boldsymbol{p(u)}$ where

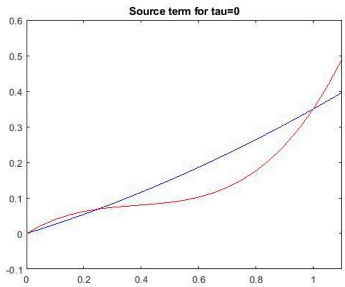
$\boldsymbol{p(u)=u(u^2-1.15u+0.5)}$

For K=id and $\tau\equiv 0$ the source term takes form

$$u(u-0{,}25)(u-1)$$

Eq. $\frac{\partial}{\partial t}u=\frac{\partial^2}{\partial x^2}u-u(u-0{,}25)(u-1)$ has heteroclinic solutions (travelling fronts) of the form

Source term for $\tau=0$
For $\tau=0$ we must have bistable case!



3D NUMERICAL SIMMULATIONS I

Assuming cylindical symmetry we solved numerically the system :

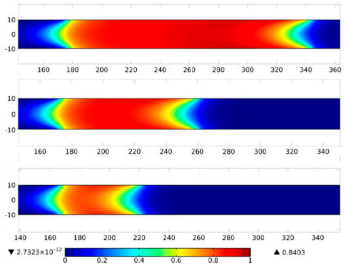
$$\frac{\partial}{\partial t}\boldsymbol{c}=\boldsymbol{D}\Delta\boldsymbol{c}\quad \text{in } \Omega\,,$$

$$\boldsymbol{D}\,\boldsymbol{n}\cdot\nabla\mathbf{c}=\boldsymbol{A}\{\left[\boldsymbol{K}_{\sigma}\frac{\partial}{\partial t}\boldsymbol{\tau}\right]_{+}-\boldsymbol{p}(\boldsymbol{c})\}\qquad \text{on } \partial\Omega,$$

$$\frac{\partial}{\partial t}\boldsymbol{\tau}=\boldsymbol{\beta}[\boldsymbol{\tau}^{*}(\mathbf{c})-\boldsymbol{\tau}]\qquad \text{in } \Omega$$

Numerically determined travelling homoclinic waves (moving to the right)

Ca^{++} concentration (for different σ)



Numerical computations

All numerical computations were done for the diffusion coefficient **D**=1.
The source term:

$[\boldsymbol{K}(\mathbf{0},\mathbf{25}\boldsymbol{u}+\mathbf{0}.\mathbf{1}\boldsymbol{u}^{\mathbf{2}}-\boldsymbol{\tau})]_{+}-\boldsymbol{p}(\boldsymbol{u})$ where

$\boldsymbol{p}(\boldsymbol{u})=\boldsymbol{u}(\boldsymbol{u}^{\mathbf{2}}-\mathbf{1}.\mathbf{15}\boldsymbol{u}+\mathbf{0}.\mathbf{5})$

For K=id and $\tau\equiv 0$ the source term takes form

$$u(u-0{,}25)(u-1)$$

Eq. $\frac{\partial}{\partial t}u=\frac{\partial^2}{\partial x^2}u-u(u-0{,}25)(u-1)$ has heteroclinic solutions (travelling fronts) of the form

ONE DIMENSIONAL APPROXIMATION

Averaging our diffusion equation with respect to r :
and similarly, the equation for the traction, we arrive at the one dimensional problem

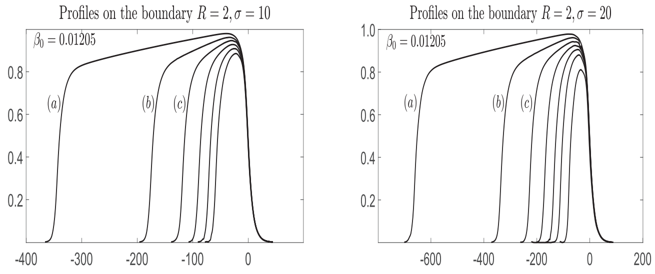
$$\frac{\partial}{\partial t}\boldsymbol{u}=\boldsymbol{D}\frac{\partial^2}{\partial x^2}\boldsymbol{u}+\frac{2A}{R}\boldsymbol{\beta}\mathbf{K}_2*[\boldsymbol{\tau}^*(\mathbf{u})-\boldsymbol{\tau}]-\boldsymbol{p}(\boldsymbol{u})$$

$$\frac{\partial}{\partial \mathrm{t}}\boldsymbol{\tau}=\boldsymbol{\beta}[\boldsymbol{\tau}^*(\mathbf{u})-\boldsymbol{\tau}]$$

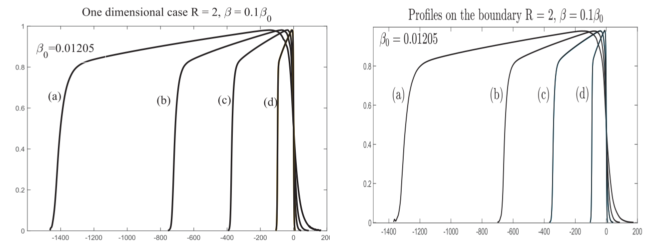
where

$$u(t,x)=\frac{1}{\pi R^2}\int_0^R 2\pi r\,c(t,x,r)dr$$

Waves profiles at r=R, (R=2) for different β: (a) β=0,1$\boldsymbol{\beta_0}$, (b) β=0,2$\boldsymbol{\beta_0}$, (c) β=0,3$\boldsymbol{\beta_0}$ etc. where the reference $\boldsymbol{\beta}$ is $\boldsymbol{\beta_0 = 0,01205}$. On the left for σ=10. On the right for σ=20.



On the left: wave profiles and wave velocities in 1-D simulations for A=1, and (a) σ=40, (b) σ=20, (c) σ=10, (d) σ=0 and for β=0.001205 (=0.1 β_0). On the right: 3D simulations for A=1, R=2, and β=0,001205 and the same values of σ. Propagation velocities with respect to the heteroclinic case ($v_0 = \sqrt{2}/4$) are: (a)13.7,(b) 6.96, (c) 3.63 , (d) 0.978



Fitzhugh –Nagumo type of approximation

The influence of the variance σ of K_σ on the wave velocity.

Expanding :

$\mathbf{K_2} * \boldsymbol{\tau}^*(\mathbf{u})$ we arrive to easier, local system of equations

$$\frac{\partial}{\partial t}\boldsymbol{w} = \frac{\partial^2}{\partial z^2}\left(\boldsymbol{Dw} + \frac{\boldsymbol{A}}{\boldsymbol{R}}\boldsymbol{\sigma}^2\,\boldsymbol{\tau}^*(\boldsymbol{w})\right) + \frac{\boldsymbol{2A}}{\boldsymbol{R}}\{[\boldsymbol{\tau}^*(\boldsymbol{w}) - \boldsymbol{\tau}]^+ - \boldsymbol{p}(\boldsymbol{w})\}$$

$$\frac{\partial}{\partial t}\boldsymbol{\tau} = \boldsymbol{\beta}[\boldsymbol{\tau}^*(\boldsymbol{w}) - \boldsymbol{\tau}]$$

with larger diffusivity. The wave velocity is $\sim\sqrt{diffusivity}$

F-N model is simple and gives good wave speed.

This F-N model we studied (with J. Napiokowska) for a particular shape of the source term step like $\boldsymbol{\tau}^*(\boldsymbol{w})$ and linear $\boldsymbol{p}(\boldsymbol{w})$.

- In this case the existence of homoclinic waves is proven for some range of $\beta < \beta_0$,
- For $\beta > \beta_0$ there are no homoclinic waves.
- There are two solutions for given $\beta < \beta_0$. Narrow one unstable and wider which is stable.

Conclusions

1. It seems that the idea of F. Jaffe works

a) Wave velocity grows as σ . σ – range of mechanical interactions due to actin-myosin fiber network.

b) The concentration of Ca in extracellular space is 100 times bigger than in endoplasmic reticulum, so flux through ion channel can be quite high. Again, wave velocity grows as $\sqrt{Source}$

2. 1-D approximation seems to work quite well ! It well reproduces the 3-D simulations.

REFERENCES

[1] L.F. Jaffe, "Stretch-activated calcium channels relay fast calcium waves propagated by calcium-induced calcium inux", Biol. Cell 99, 175{184 (2007).

[2] J.C. Gilkey, L.F. Jaffe, E.B. Ridgway, G.T. Reynolds GT, "A free calcium wave traverses the activating egg of the Medaka, Oryzias latipes", J. Cell Biol. 76, 448-466 (1978).

[3]Z. Peradzynski, B. Kaźmierczak, S. Białecki, Modelling fast CICI Waves – to be published.

[4] Murray JD. Mathematical Biology, Springer, Berlin, 2nd edition, 1993.

[5] W.S. Nishitani, Molecular mechanisms involved in cell response to mechanical forces, phd dissertation, University of Illinois at Urbana-Champaign, 2011.

[6] K. Okeyo, et al., "Actomyosin contractility spatiotemporally regulates actin network dynamics in migrating cells", J. Biomech. 42(15), 2540-8 (2009).

[7] Z. Peradzynski, "Diffusion of calcium in biological tissues and accompanying mechanochemical effects", Arch. Mech. 62, 423-440 (2010).

[8] B. Kaźmierczak, Z. Peradzyński, Calcium waves with fast buffers and mechanical effects, J. Math. Biol., DOI 10.1007/s00285-009-0323-2, 2010.

Thank you for your attention

and

the organizers for the invitation.

Generalization of Reeb spaces and application to data visualization

Osamu Saeki
Institute of Mathematics for Industry, Kyushu University, Japan

In many cases, data sets can be considered to be discrete samples of differentiable maps between manifolds. For a differentiable multivariate function into $\mathbb{R}^p$ with $p \geq 2$, its Reeb space is the space of connected components of its fibers. This is a generalization of the notion of Reeb graphs for univariate functions in the case of $p = 1$. It has been known that Reeb spaces are often very useful for visualizing the given multivariate function. In this talk, we generalize the Reeb space in such a way that it captures more of the topological features of the fibers, not only their connected components. This theoretical part essentially relies on the global singularity theory of differentiable maps between manifolds developed mainly by the author. Such techniques have been used for efficiently visualize large scale data. If time permits, we will also discuss an application to multi-objective optimization problems.

Generalization of Reeb Spaces and Application to Data Visualization

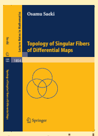
Osamu Saeki
(Institute of Mathematics for Industry,
Kyushu Univ., Japan)

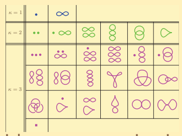
September 29, 2023
WORKSHOP on
Mathematics for Industry
Warsaw University of Technology

Who am I?

Got PhD in **Mathematics** in 1992.
"On 4-manifolds homotopy equivalent to the 2-sphere"

Main interest: **Singularity Theory**, 3- and 4-Dimensional **Topology**
I proposed the Theory of **Singular Fibers** of Differentiable Maps.



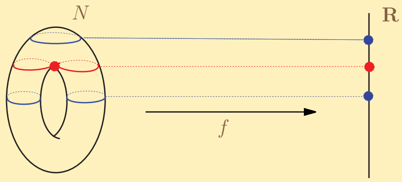


My recent interests include collaboration with computer scientists on enhancing **visualization of multi-variate data** from the viewpoint of **topology** or **singularity theory**.

§1. Fiber

Setting

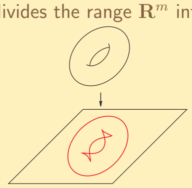
N^n : C^∞ **manifold** (e.g. bdd domain in $\mathbf{R}^n$), $f : N^n \to \mathbf{R}^m$ C^∞ **map**
We can write $f = (f_1, f_2, \ldots, f_m)$ **multi-variate data**
We assume f is **generic** (C^∞ stable, C^0 stable, finite codimension, etc.)
We are interested in the topology of **fibers** $f^{-1}(y)$, $y \in \mathbf{R}^m$.
Generically, $\dim f^{-1}(y) = n - m$. We usually assume $n \geq m$.



We can grasp global feature of data by chasing fibers (or level sets).
We have **singular fibers** (or **critical level sets**) where **topological transitions of fibers** occur.

Singular points and Jacobi set

$f : N^n \to \mathbf{R}^m$ $(n \geq m)$ C^∞ map
For $p \in N^n$, consider the **differential** $df_p : T_pN^n \to T_{f(p)}\mathbf{R}^m$.
The set of singular points $J(f) = \{p \in N^n \mid \mathrm{rank}\ df_p < m\}$ is called the **Jacobi set** of f. Generically, $\dim J(f) = m - 1$.
Jacobi set image $f(J(f))$ divides the range $\mathbf{R}^m$ into some regions.



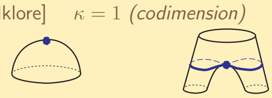
Topology of fibers changes along the Jacobi set image.
Singular fiber is a fiber $f^{-1}(y)$ with $y \in f(J(f))$.
It is important to know topological changes of fibers near a singular fiber.

Classification of singular fibers

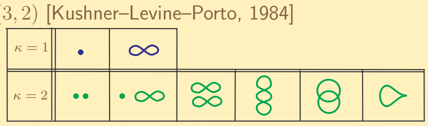
For certain dimensions, we can **classify singular fibers** of generic maps.

Example 1.1 *Classification results for (n, m) with $n - m = 1$.*
For simplicity, we assume the domain N^n is **orientable**.
We will ignore regular fiber components.

1. $(n, m) = (2, 1)$ [Folklore] $\kappa = 1$ *(codimension)*

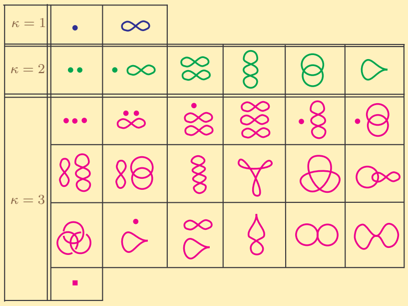


2. $(n, m) = (3, 2)$ [Kushner–Levine–Porto, 1984]



Singular fibers for $(n, m) = (4, 3)$

3. $(n, m) = (4, 3)$ [S.]

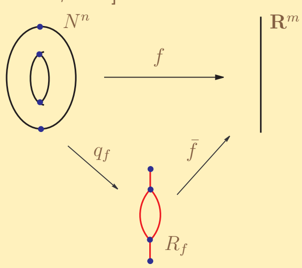


4. $(n, m) = (5, 4)$ [Yamamoto–S.]

§2. Reeb Space

Reeb space

For a C^∞ map $f : N^n \to \mathbf{R}^m$, $n > m$, the space R_f obtained by contracting each connected component of a fiber to a point is called the **Reeb space** of f [Edelsbrunner–Harer–Patel, 2008].



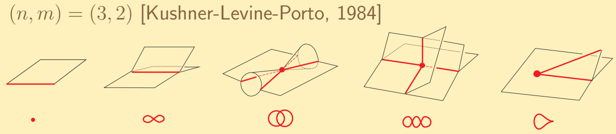
When $m = 1$, it is also called the **Reeb graph**.

Local structures of Reeb spaces

Classification of fibers
⇒ Characterization of local structures of Reeb spaces

Example 2.1 *1.* $(n, m) = (2, 1)$

2. $(n, m) = (3, 2)$ [Kushner-Levine-Porto, 1984]



3. $(n, m) = (4, 3)$ [Hiratuka, 2001]

Applications

An example of an application of Reeb graphs:
[Takahashi–Takeshima–Fujishiro, 2004] **Topological Volume Skeletonization and its Application to Transfer Function Design**
See Fig. 4 of [STSWKCDY].

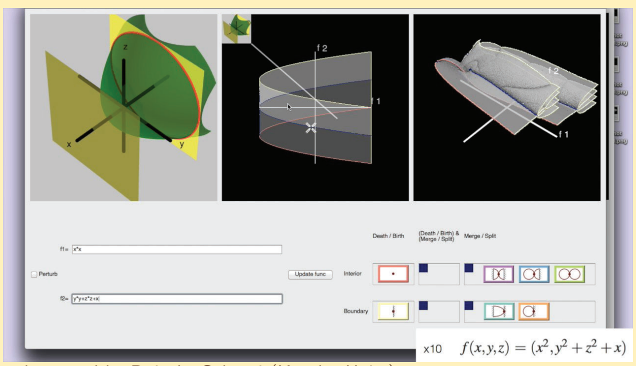
Explicit example: Atom collision
See Fig. 5 of [STSWKCDY].

Application of singular fibers: Hurricane Isabel data
See Fig. 15 of [STSWKCDY].

Reference:

[STSWKCDY] O. Saeki, S. Takahashi, D. Sakurai, Hsiang-Yun Wu, K. Kikuchi, H. Carr, D. Duke, and T. Yamamoto, Visualizing multivariate data using singularity theory, The Impact of Applications on Mathematics, Proceedings of Forum "Math-for-Industry" 2013, pp.51–65, Springer, 2014.

User interface

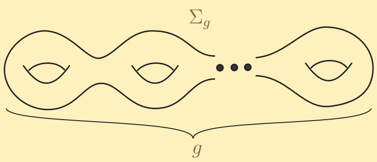


Implemented by Daisuke Sakurai (Kyushu Univ.)

§3. Case with $n-m=2$ and Beyond

Regular fibers for $n-m=2$

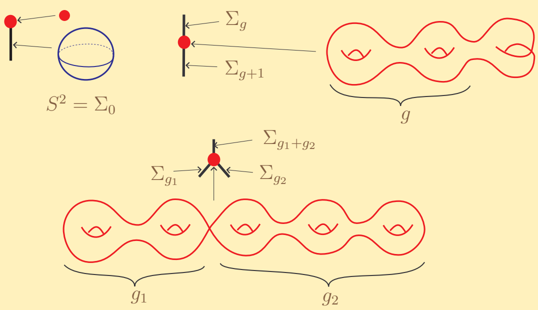
How about the case $n-m=2$?
We assume the domain N^n is orientable, compact, and w/o boundary.
Regular fibers are **closed orientable surfaces**.



Σ_g : closed orientable surface of genus g

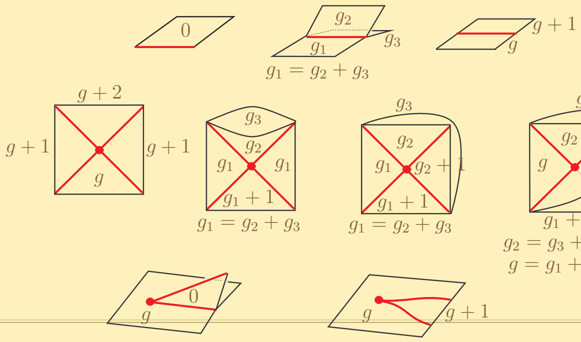
Reeb graph for $(n,m)=(3,1)$

Case with $(n,m)=(3,1)$.



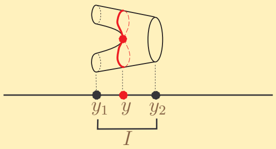
Reeb space for $(n, m) = (4, 2)$

Each regular "stratum" has its own **label** (genus of the corresponding regular fiber component). [Furuya, 1986]



0-th Homology

Reeb space describes the **connected components of fibers** and their "adjacencies".

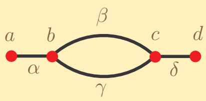


$$\begin{tikzcd}
H_0(f^{-1}(y_1)) \arrow[r] \arrow[d, "\cong"'] & H_0(f^{-1}(I)) \arrow[d, "\cong"'] & H_0(f^{-1}(y_2)) \arrow[l] \arrow[d, "\cong"'] \\
\mathbf{Z} \oplus \mathbf{Z} & H_0(f^{-1}(y)) \arrow[d, "\cong"'] & \mathbf{Z} \\
 & \mathbf{Z} &
\end{tikzcd}$$

Reeb diagram

This idea can be extended to homology groups of any dimension (or homotopy groups, if you want).

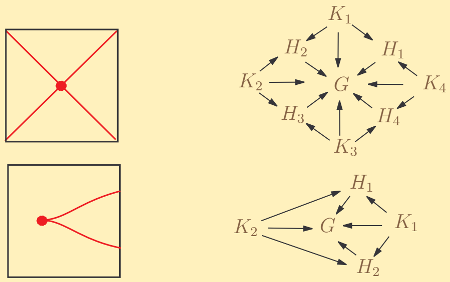
$$H_i(f^{-1}(y_1)) \to H_i(f^{-1}(y)) \leftarrow H_i(f^{-1}(y_2))$$



$$\begin{tikzcd}
 & & & H_i(\beta) \arrow[dl] \arrow[dr] & & & \\
H_i(a) & H_i(\alpha) \arrow[l] \arrow[r] & H_i(b) & & H_i(c) & H_i(\delta) \arrow[l] \arrow[r] & H_i(d) \\
 & & & H_i(\gamma) \arrow[ul] \arrow[ur] & & &
\end{tikzcd}$$

$\Longrightarrow$ Notion of **Reeb diagram**.

Categorical formulation

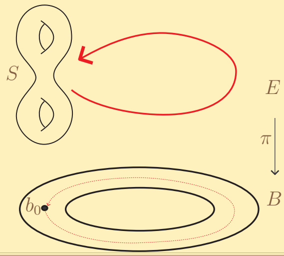


In a certain **categorical formulation** of a Reeb space, this can be considered to be a **functor**.

Remark 3.1 This makes sense if each stratum is contractible.

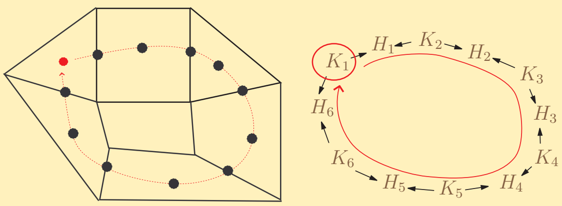
Monodromy

If a stratum is not contractible, we need to consider **monodromy**.
Suppose $\pi : E \to B$ is a C^∞ fiber bundle with fiber S.
$\mathrm{MCG}(S) = \pi_0(\mathrm{Diff}_+(S))$ **mapping class group**
Associated to π is the **monodromy** $\pi_1(B, b_0) \to \mathrm{MCG}(S)$.
This measures the "twist" of the fibers along a loop in the base B.



Monodromy in Reeb diagram

Given a generic map $f : N^n \to \mathbf{R}^m$, we can subdivide $\mathbf{R}^m$ (or the Reeb space R_f) so that each stratum is contractible.
In this case, the **monodromy is hidden** in the Reeb diagram.



$$\pi_1(B, b_0) \to \mathrm{MCG}(S) \to \mathrm{Aut}(H_*(S))$$

Problem 3.2 *Formulate all these, including monodromy! Category theory? How to compute Reeb diagram and/or monodromy?*

Possible application

Let us consider a **multi-ojective optimization problem**.
Such a problem can be formulated in terms of a C^∞ muti-function $f = (f_1, f_2, \ldots, f_m) : N^n \to \mathbf{R}^m$.

For example, given a **bench-mark problem** of multi-optimization, we can evaluate its complexity or certain characteristics in terms of its **Reeb space**, or more generally, its **Reeb diagram**.

Ending

Thank you for your attention!

Acknowledgements to collaborators:
Takahiro Yamamoto (Tokyo Gakugei Univ.),
Daisuke Sakurai (Kyushu Univ.),
Shigeo Takahashi (Univ. of Aizu),
Hamish Carr & David Duke (Univ. of Leeds),
Hsiang-Yun Wu (St. Pölten Univ. of Applied Sciences),
Keisuke Kikuchi

MI レクチャーノートシリーズ刊行にあたり

本レクチャーノートシリーズは、文部科学省 21 世紀 COE プログラム「機能数理学の構築と展開」（H15-19 年度）において作成した COE Lecture Notes の続刊であり、文部科学省大学院教育改革支援プログラム「産業界が求める数学博士と新修士養成」（H19-21 年度）および、同グローバル COE プログラム「マス・フォア・インダストリ教育研究拠点」（H20-24 年度）において行われた講義の講義録として出版されてきた。平成 23 年 4 月のマス・フォア・インダストリ研究所（IMI）設立と平成 25 年 4 月の IMI の文部科学省共同利用・共同研究拠点として「産業数学の先進的･基礎的共同研究拠点」の認定を受け、今後、レクチャーノートは、マス・フォア・インダストリに関わる国内外の研究者による講義の講義録、会議録等として出版し、マス・フォア・インダストリの本格的な展開に資するものとする。

2022 年 10 月
マス・フォア・インダストリ研究所
所長　梶原 健司

International Project Research-Workshop (I)

WORKSHOP on Mathematics for Industry
Basis of Mathematics in nanomedicine structures and life sensing

発　行　2024年 3 月 14 日
編　集　Osamu Saeki, Wojciech Domitrz, Stanisław Janeczko, Marcin Zubilewicz, Michał Zwierzyński
発　行　九州大学マス・フォア・インダストリ研究所
九州大学大学院数理学府
〒819-0395 福岡市西区元岡744
九州大学数理・IMI 事務室
TEL 092-802-4402　FAX 092-802-4405
URL https://www.imi.kyushu-u.ac.jp/

印　刷　城島印刷株式会社
〒810-0012 福岡市中央区白金 2 丁目 9 番 6 号
TEL 092-531-7102　FAX 092-524-4411

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note	Mitsuhiro T. NAKAO Kazuhiro YOKOYAMA	Computer Assisted Proofs - Numeric and Symbolic Approaches - 199pages	August 22, 2006
COE Lecture Note	M.J.Shai HARAN	Arithmetical Investigations - Representation theory, Orthogonal polynomials and Quantum interpolations- 174pages	August 22, 2006
COE Lecture Note Vol.3	Michal BENES Masato KIMURA Tatsuyuki NAKAKI	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2005 155pages	October 13, 2006
COE Lecture Note Vol.4	宮田　健治	辺要素有限要素法による磁界解析 - 機能数理学特別講義 21pages	May 15, 2007
COE Lecture Note Vol.5	Francois APERY	Univariate Elimination Subresultants - Bezout formula, Laurent series and vanishing conditions - 89pages	September 25, 2007
COE Lecture Note Vol.6	Michal BENES Masato KIMURA Tatsuyuki NAKAKI	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2006 209pages	October 12, 2007
COE Lecture Note Vol.7	若山　正人 中尾　充宏	九州大学産業技術数理研究センター キックオフミーティング 138pages	October 15, 2007
COE Lecture Note Vol.8	Alberto PARMEGGIANI	Introduction to the Spectral Theory of Non-Commutative Harmonic Oscillators 233pages	January 31, 2008
COE Lecture Note Vol.9	Michael I.TRIBELSKY	Introduction to Mathematical modeling 23pages	February 15, 2008
COE Lecture Note Vol.10	Jacques FARAUT	Infinite Dimensional Spherical Analysis 74pages	March 14, 2008
COE Lecture Note Vol.11	Gerrit van DIJK	Gelfand Pairs And Beyond 60pages	August 25, 2008
COE Lecture Note Vol.12	Faculty of Mathematics, Kyushu University	Consortium “MATH for INDUSTRY” First Forum 87pages	September 16, 2008
COE Lecture Note Vol.13	九州大学大学院 数理学研究院	プロシーディング「損保数理に現れる確率モデル」 — 日新火災・九州大学 共同研究2008年11月 研究会 — 82pages	February 6, 2009

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.14	Michal Beneš, Tohru Tsujikawa Shigetoshi Yazaki	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2008 77pages	February 12, 2009
COE Lecture Note Vol.15	Faculty of Mathematics, Kyushu University	International Workshop on Verified Computations and Related Topics 129pages	February 23, 2009
COE Lecture Note Vol.16	Alexander Samokhin	Volume Integral Equation Method in Problems of Mathematical Physics 50pages	February 24, 2009
COE Lecture Note Vol.17	矢嶋　徹 及川　正行 梶原　健司 辻　英一 福本　康秀	非線形波動の数理と物理　66pages	February 27, 2009
COE Lecture Note Vol.18	Tim Hoffmann	Discrete Differential Geometry of Curves and Surfaces 75pages	April 21, 2009
COE Lecture Note Vol.19	Ichiro Suzuki	The Pattern Formation Problem for Autonomous Mobile Robots —Special Lecture in Functional Mathematics— 23pages	April 30, 2009
COE Lecture Note Vol.20	Yasuhide Fukumoto Yasunori Maekawa	Math-for-Industry Tutorial: Spectral theories of non-Hermitian operators and their application 184pages	June 19, 2009
COE Lecture Note Vol.21	Faculty of Mathematics, Kyushu University	Forum "Math-for-Industry" Casimir Force, Casimir Operators and the Riemann Hypothesis 95pages	November 9, 2009
COE Lecture Note Vol.22	Masakazu Suzuki Hoon Hong Hirokazu Anai Chee Yap Yousuke Sato Hiroshi Yoshida	The Joint Conference of ASCM 2009 and MACIS 2009: Asian Symposium on Computer Mathematics Mathematical Aspects of Computer and Information Sciences 436pages	December 14, 2009
COE Lecture Note Vol.23	荒川　恒男 金子　昌信	多重ゼータ値入門　111pages	February 15, 2010
COE Lecture Note Vol.24	Fulton B.Gonzalez	Notes on Integral Geometry and Harmonic Analysis 125pages	March 12, 2010
COE Lecture Note Vol.25	Wayne Rossman	Discrete Constant Mean Curvature Surfaces via Conserved Quantities 130pages	May 31, 2010
COE Lecture Note Vol.26	Mihai Ciucu	Perfect Matchings and Applications 66pages	July 2, 2010

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.27	九州大学大学院 数理学研究院	Forum “Math-for-Industry” and Study Group Workshop Information security, visualization, and inverse problems, on the basis of optimization techniques　100pages	October 21, 2010
COE Lecture Note Vol.28	ANDREAS LANGER	MODULAR FORMS, ELLIPTIC AND MODULAR CURVES LECTURES AT KYUSHU UNIVERSITY 2010　62pages	November 26, 2010
COE Lecture Note Vol.29	木田　雅成 原田　昌晃 横山　俊一	Magma で広がる数学の世界　157pages	December 27, 2010
COE Lecture Note Vol.30	原　　　隆 松井　　卓 廣島　文生	Mathematical Quantum Field Theory and Renormalization Theory 201pages	January 31, 2011
COE Lecture Note Vol.31	若山　正人 福本　康秀 高木　　剛 山本　昌宏	Study Group Workshop 2010 Lecture & Report　128pages	February 8, 2011
COE Lecture Note Vol.32	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2011 “TSUNAMI-Mathematical Modelling” Using Mathematics for Natural Disaster Prediction, Recovery and Provision for the Future　90pages	September 30, 2011
COE Lecture Note Vol.33	若山　正人 福本　康秀 高木　　剛 山本　昌宏	Study Group Workshop 2011 Lecture & Report　140pages	October 27, 2011
COE Lecture Note Vol.34	Adrian Muntean Vladimír Chalupecký	Homogenization Method and Multiscale Modeling　72pages	October 28, 2011
COE Lecture Note Vol.35	横山　俊一 夫　　紀恵 林　　卓也	計算機代数システムの進展　210pages	November 30, 2011
COE Lecture Note Vol.36	Michal Beneš Masato Kimura Shigetoshi Yazaki	Proceedings of Czech-Japanese Seminar in Applied Mathematics 2010 107pages	January 27, 2012
COE Lecture Note Vol.37	若山　正人 高木　　剛 Kirill Morozov 平岡　裕章 木村　正人 白井　朋之 西井　龍映 栄　伸一郎 穴井　宏和 福本　康秀	平成23年度 数学・数理科学と諸科学・産業との連携研究ワークショップ　拡がっていく数学　～期待される “見えない力” ～ 154pages	February 20, 2012

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.38	Fumio Hiroshima Itaru Sasaki Herbert Spohn Akito Suzuki	Enhanced Binding in Quantum Field Theory　204pages	March 12, 2012
COE Lecture Note Vol.39	Institute of Mathematics for Industry, Kyushu University	Multiscale Mathematics; Hierarchy of collective phenomena and interrelations between hierarchical structures　180pages	March 13, 2012
COE Lecture Note Vol.40	井ノ口順一 太田　泰広 筧　　三郎 梶原　健司 松浦　　望	離散可積分系・離散微分幾何チュートリアル2012　152pages	March 15, 2012
COE Lecture Note Vol.41	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2012 “Information Recovery and Discovery”　91pages	October 22, 2012
COE Lecture Note Vol.42	佐伯　　修 若山　正人 山本　昌宏	Study Group Workshop 2012 Abstract, Lecture & Report　178pages	November 19, 2012
COE Lecture Note Vol.43	Institute of Mathematics for Industry, Kyushu University	Combinatorics and Numerical Analysis Joint Workshop　103pages	December 27, 2012
COE Lecture Note Vol.44	萩原　　学	モダン符号理論からポストモダン符号理論への展望　107pages	January 30, 2013
COE Lecture Note Vol.45	金山　　寛	Joint Research Workshop of Institute of Mathematics for Industry (IMI), Kyushu University “Propagation of Ultra-large-scale Computation by the Domain-decomposition-method for Industrial Problems (PUCDIP 2012)” 121pages	February 19, 2013
COE Lecture Note Vol.46	西井　龍映 栄　伸一郎 岡田　勘三 落合　啓之 小磯　深幸 斎藤　新悟 白井　朋之	科学・技術の研究課題への数学アプローチ —数学モデリングの基礎と展開—　325pages	February 28, 2013
COE Lecture Note Vol.47	SOO TECK LEE	BRANCHING RULES AND BRANCHING ALGEBRAS FOR THE COMPLEX CLASSICAL GROUPS　40pages	March 8, 2013
COE Lecture Note Vol.48	溝口　佳寛 脇　　隼人 平坂　　貢 谷口　哲至 島袋　　修	博多ワークショップ「組み合わせとその応用」　124pages	March 28, 2013

シリーズ既刊

Issue	Author／Editor	Title	Published
COE Lecture Note Vol.49	照井　　章 小原　功任 濱田　龍義 横山　俊一 穴井　宏和 横田　博史	マス・フォア・インダストリ研究所　共同利用研究集会 II 数式処理研究と産学連携の新たな発展　137pages	August 9, 2013
MI Lecture Note Vol.50	Ken Anjyo Hiroyuki Ochiai Yoshinori Dobashi Yoshihiro Mizoguchi Shizuo Kaji	Symposium MEIS2013: Mathematical Progress in Expressive Image Synthesis　154pages	October 21, 2013
MI Lecture Note Vol.51	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2013 “The Impact of Applications on Mathematics”　97pages	October 30, 2013
MI Lecture Note Vol.52	佐伯　　修 岡田　勘三 髙木　　剛 若山　正人 山本　昌宏	Study Group Workshop 2013 Abstract, Lecture ＆ Report 142pages	November 15, 2013
MI Lecture Note Vol.53	四方　義啓 櫻井　幸一 安田　貴徳 Xavier Dahan	平成25年度　九州大学マス・フォア・インダストリ研究所 共同利用研究集会　安全・安心社会基盤構築のための代数構造 ～サイバー社会の信頼性確保のための数理学～　158pages	December 26, 2013
MI Lecture Note Vol.54	Takashi Takiguchi Hiroshi Fujiwara	Inverse problems for practice, the present and the future　93pages	January 30, 2014
MI Lecture Note Vol.55	栄　伸一郎 溝口　佳寛 脇　　隼人 渋田　敬史	Study Group Workshop 2013 数学協働プログラム Lecture & Report 98pages	February 10, 2014
MI Lecture Note Vol.56	Yoshihiro Mizoguchi Hayato Waki Takafumi Shibuta Tetsuji Taniguchi Osamu Shimabukuro Makoto Tagami Hirotake Kurihara Shuya Chiba	Hakata Workshop 2014 ~ Discrete Mathematics and its Applications ~　141pages	March 28, 2014
MI Lecture Note Vol.57	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2014: “Applications + Practical Conceptualization + Mathematics = fruitful Innovation”　93pages	October 23, 2014
MI Lecture Note Vol.58	安生健一 落合啓之	Symposium MEIS2014: Mathematical Progress in Expressive Image Synthesis　135pages	November 12, 2014

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.59	西井　龍映 岡田　勘三 梶原　健司 髙木　　剛 若山　正人 脇　　隼人 山本　昌宏	Study Group Workshop 2014 数学協働プログラム Abstract, Lecture & Report　196pages	November 14, 2014
MI Lecture Note Vol.60	西浦　　博	平成26年度九州大学 IMI 共同利用研究・研究集会（I） 感染症数理モデルの実用化と産業及び政策での活用のための新たな展開　120pages	November 28, 2014
MI Lecture Note Vol.61	溝口　佳寛 Jacques Garrigue 萩原　　学 Reynald Affeldt	研究集会 高信頼な理論と実装のための定理証明および定理証明器 Theorem proving and provers for reliable theory and implementations (TPP2014)　138pages	February 26, 2015
MI Lecture Note Vol.62	白井　朋之	Workshop on “β-transformation and related topics”　59pages	March 10, 2015
MI Lecture Note Vol.63	白井　朋之	Workshop on “Probabilistic models with determinantal structure” 107pages	August 20, 2015
MI Lecture Note Vol.64	落合　啓之 土橋　宜典	Symposium MEIS2015: Mathematical Progress in Expressive Image Synthesis　124pages	September 18, 2015
MI Lecture Note Vol.65	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2015 “The Role and Importance of Mathematics in Innovation”　74pages	October 23, 2015
MI Lecture Note Vol.66	岡田　勘三 藤澤　克己 白井　朋之 若山　正人 脇　　隼人 Philip Broadbridge 山本　昌宏	Study Group Workshop 2015 Abstract, Lecture & Report 156pages	November 5, 2015
MI Lecture Note Vol.67	Institute of Mathematics for Industry, Kyushu University	IMI-La Trobe Joint Conference “Mathematics for Materials Science and Processing” 66pages	February 5, 2016
MI Lecture Note Vol.68	古庄　英和 小谷　久寿 新甫　洋史	結び目と Grothendieck-Teichmüller 群 116pages	February 22, 2016
MI Lecture Note Vol.69	土橋　宜典 鍛治　静雄	Symposium MEIS2016: Mathematical Progress in Expressive Image Synthesis　82pages	October 24, 2016
MI Lecture Note Vol.70	Institute of Mathematics for Industry, Kyushu University	Forum “Math-for-Industry” 2016 “Agriculture as a metaphor for creativity in all human endeavors” 98pages	November 2, 2016
MI Lecture Note Vol.71	小磯　深幸 二宮　嘉行 山本　昌宏	Study Group Workshop 2016 Abstract, Lecture & Report　143pages	November 21, 2016

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.72	新井　朝雄 小嶋　　泉 廣島　文生	Mathematical quantum field theory and related topics　133pages	January 27, 2017
MI Lecture Note Vol.73	穴田　啓晃 Kirill Morozov 須賀　祐治 奥村　伸也 櫻井　幸一	Secret Sharing for Dependability, Usability and Security of Network Storage and Its Mathematical Modeling　211pages	March 15, 2017
MI Lecture Note Vol.74	QUISPEL, G. Reinout W. BADER, Philipp MCLAREN, David I. TAGAMI, Daisuke	IMI-La Trobe Joint Conference Geometric Numerical Integration and its Applications　71pages	March 31, 2017
MI Lecture Note Vol.75	手塚　　集 田上　大助 山本　昌宏	Study Group Workshop 2017 Abstract, Lecture & Report　118pages	October 20, 2017
MI Lecture Note Vol.76	宇田川誠一	Tzitzéica 方程式の有限間隙解に付随した極小曲面の構成理論 ―Tzitzéica 方程式の楕円関数解を出発点として―　68pages	August 4, 2017
MI Lecture Note Vol.77	松谷　茂樹 佐伯　　修 中川　淳一 田上　大助 上坂　正晃 Pierluigi Cesana 濵田　裕康	平成29年度　九州大学マス・フォア・インダストリ研究所 共同利用研究集会（I） 結晶の界面，転位，構造の数理　148pages	December 20, 2017
MI Lecture Note Vol.78	瀧澤　重志 小林　和博 佐藤憲一郎 斎藤　　努 清水　正明 間瀬　正啓 藤澤　克樹 神山　直之	平成29年度　九州大学マス・フォア・インダストリ研究所 プロジェクト研究　研究集会（I） 防災・避難計画の数理モデルの高度化と社会実装へ向けて 136pages	February 26, 2018
MI Lecture Note Vol.79	神山　直之 畔上　秀幸	平成29年度　AIMaP チュートリアル 最適化理論の基礎と応用　96pages	February 28, 2018
MI Lecture Note Vol.80	Kirill Morozov Hiroaki Anada Yuji Suga	IMI Workshop of the Joint Research Projects Cryptographic Technologies for Securing Network Storage and Their Mathematical Modeling　116pages	March 30, 2018
MI Lecture Note Vol.81	Tsuyoshi Takagi Masato Wakayama Keisuke Tanaka Noboru Kunihiro Kazufumi Kimoto Yasuhiko Ikematsu	IMI Workshop of the Joint Research Projects International Symposium on Mathematics, Quantum Theory, and Cryptography　246pages	September 25, 2019
MI Lecture Note Vol.82	池森　俊文	令和２年度　AIMaP チュートリアル 新型コロナウイルス感染症にかかわる諸問題の数理 145pages	March 22, 2021

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.83	早川健太郎 軸丸　芳揮 横須賀洋平 可香谷　隆 林　　和希 堺　　雄亮	シェル理論・膜理論への微分幾何学からのアプローチと その建築曲面設計への応用　49pages	July 28, 2021
MI Lecture Note Vol.84	Taketoshi Kawabe Yoshihiro Mizoguchi Junichi Kako Masakazu Mukai Yuji Yasui	SICE-JSAE-AIMaP Tutorial Advanced Automotive Control and Mathematics　110pages	December 27, 2021
MI Lecture Note Vol.85	Hiroaki Anada Yasuhiko Ikematsu Koji Nuida Satsuya Ohata Yuntao Wang	IMI Workshop of the Joint Usage Research Projects Exploring Mathematical and Practical Principles of Secure Computation and Secret Sharing　114pages	February 9, 2022
MI Lecture Note Vol.86	濱田　直希 穴井　宏和 梅田　裕平 千葉　一永 佐藤　寛之 能島　裕介 加葉田雄太朗 一木　俊助 早野　健太 佐伯　　修	2020年度採択分　九州大学マス・フォア・インダストリ研究所 共同利用研究集会 進化計算の数理　135pages	February 22, 2022
MI Lecture Note Vol.87	Osamu Saeki, Ho Tu Bao, Shizuo Kaji, Kenji Kajiwara, Nguyen Ha Nam, Ta Hai Tung, Melanie Roberts, Masato Wakayama, Le Minh Ha, Philip Broadbridge	Proceedings of Forum "Math-for-Industry" 2021 -Mathematics for Digital Economy-　122pages	March 28, 2022
MI Lecture Note Vol.88	Daniel PACKWOOD Pierluigi CESANA, Shigenori FUJIKAWA, Yasuhide FUKUMOTO, Petros SOFRONIS, Alex STAYKOV	Perspectives on Artificial Intelligence and Machine Learning in Materials Science, February 4-6, 2022　74pages	November 8, 2022

シリーズ既刊

Issue	Author／Editor	Title	Published
MI Lecture Note Vol.89	松谷　茂樹 落合　啓之 井上　和俊 小磯　深幸 佐伯　　修 白井　朋之 垂水　竜一 内藤　久資 中川　淳一 濵田　裕康 松江　　要 加葉田雄太朗	2022年度採択分 九州大学マス・フォア・インダストリ研究所 共同利用研究集会 材料科学における幾何と代数 III　356pages	December 7, 2022
MI Lecture Note Vol.90	中山　尚子 谷川　拓司 品野　勇治 近藤　正章 石原　　亨 鍛治　静雄 藤澤　克樹	2022年度採択分 九州大学マス・フォア・インダストリ研究所 共同利用研究集会 データ格付けサービス実現のための数理基盤の構築　58pages	December 12, 2022
MI Lecture Note Vol.91	Katsuki Fujisawa Shizuo Kaji Toru Ishihara Masaaki Kondo Yuji Shinano Takuji Tanigawa Naoko Nakayama	IMI Workshop of the Joint Usage Research Projects Construction of Mathematical Basis for Realizing Data Rating Service 610pages	December 27, 2022
MI Lecture Note Vol.92	丹田　　聡 三宮　　俊 廣島　文生	2022年度採択分　九州大学マス・フォア・インダストリ研究所 共同利用研究集会 時間・量子測定・準古典近似の理論と実験 ～古典論と量子論の境界～　150pages	Janualy 6, 2023
MI Lecture Note Vol.93	Philip Broadbridge Luke Bennetts Melanie Roberts Kenji Kajiwara	Proceedings of Forum “Math-for-Industry” 2022 -Mathematics of Public Health and Sustainability-　170pages	June 19, 2023
MI Lecture Note Vol.94	國廣　　昇 池松　泰彦 伊豆　哲也 穴田　啓晃 縫田　光司	2023年度採択分　九州大学マス・フォア・インダストリ研究所 共同利用研究集会 現代暗号に対する安全性解析・攻撃の数理　260pages	Janualy 11, 2024

Institute of Mathematics for Industry
Kyushu University

九州大学マス・フォア・インダストリ研究所
九州大学大学院 数理学府
〒819-0395 福岡市西区元岡744　TEL 092-802-4402　FAX 092-802-4405
URL http://www.imi.kyushu-u.ac.jp/