## "NoStat" technologia naukowa dla statystyków

dr Sebastian Zając

# Instytut Statystyki i Demografii Zakład Analizy Historii Zdarzeń i Analiz Wielowymiarowych

22.10.2019





https://sebastianzajac.pl

### Education

### na początku było ... "Ekono co ? "

- 2002-2005 Licencjat Modelowanie szeregów czasowych za pomocą procesów ARMA i ARIMA.
- 2005-2007 Mgr Topologiczne i geometryczne metody w klasycznej i kwantowej teorii pola.

### "NoStat"

chaos  $\rightarrow$  statystyka  $\rightarrow$  chaos deterministyczny chaos liczbowy  $\rightarrow$  teoria mnogości  $\rightarrow$  teoria kategorii chaos  $\rightarrow$  statystyka  $\rightarrow$  mechanika kwantowa  $\rightarrow$  QFT

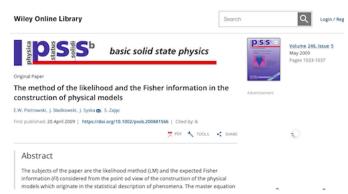
$$P_{\psi}(\phi) = |\langle \phi | \psi \rangle|^2.$$



## ciekawa literatura QM i Statystyka

- A. Einstein, B. Podolsky, N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" Physical Review 47 (1935) 777-780.
- G. Birkhoff, J von Neumann, "The Logic of Quantum Mechanics", J. Ann. of Math., 37, 823 (1936).
- J. Bell "On the Einstein Podolsky Rosen Paradox" Physics 1.3 (1964)
- Aspect, Alain (1976). "Proposed experiment to test the nonseparability of quantum mechanics". Physical Review D. 14 (8): 1944–1951.
- P. Billingsley "Probability and Measure" NY: John Wiley & Sons, Inc. 1979
- Benjamin Feintzeig Hidden Variables and Commutativity in Quantum Mechanics (2013)





"The method of the likelihood and the Fisher information in the construction of physical models."

E. W. Piotrowski, J. Sładkowski, J.Syska, S.Z. Physica Status Solidi B vol. 246 no. 5 (5.2009).

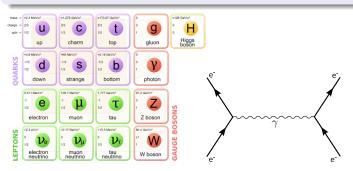


## PHD - Fizyka

### Rozprawa doktorska

"Oscylacje akceleratorowych neutrin z uwzględnieniem ich niestandardowych oddziaływań" Prof. dr hab. Marek Zrałek.

"We can't solve problems by using the same kind of thinking we used when we created them." (Albert Einstein)

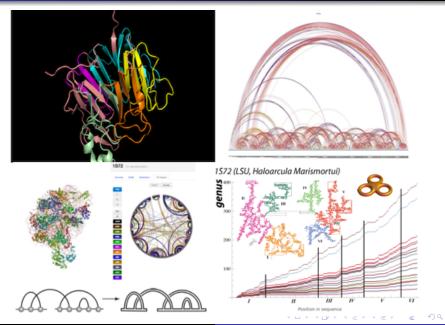


### Fizyka do dziś...

- Jak na podstawie modeli niestatystycznych wybrać, który model realizuje dane i które rozszerzenie jest istotne?
- 7 publikacji dotyczących fizyki neutrin oddziaływania i zastosowanie dyskretnych grup symetrii do mas i mieszania leptonów.
- udział w 2 grantach, prezentacje na 5 międzynarodowych konferencjach neutrinowych.



# Poza Fizyką .... ale z fizykami na UW



## Gdzie znajdziesz diagramy chordowe?

### **Physics**



"Physics is like sex: sure, it may give some practical results, but that's not why we do it."

Richard Feynman



"Category Theory is like sex: it may give some practical results, but that's not why we do it. "  $\,$  - Sebastian Zając

### Server danych

http://genus.fuw.pl



nature > scientific reports > articles > article



#### **SCIENTIFIC REPORTS**

Article | Open Access | Published: 03 December 2018

### Genus trace reveals the topological complexity and domain structure of biomolecules

Sebastian Zając, Cody Geary, Ebbe Sloth Andersen, Pawel Dabrowski-Tumanski, Joanna I. Sulkowska & Piotr Sulkowski 🏴

Scientific Reports 8, Article number: 17537 (2018) | Download Citation & 1014 Accesses | 7 Altmetric | Metrics 30

#### Abstract

The structure of bonds in biomolecules, such as base pairs in RNA chains or native interactions in proteins, can be presented in the form of a chord diagram. A given biomolecule is then characterized by the genus of an auxiliary two-dimensional surface associated to such a diagram. In this work we introduce the notion of the genus trace, which describes dependence of genus on the choice of a subchain of a given backbone chain. We find that the genus trace encodes interesting physical and biological

Download PDF		
Sections	Figures	References
Abstract		
Introduction		
Genus and Bio	omolecules	
Genus Trace fo	or Biomolecules	

Discussion

Materials and Methods

Acknowledgements

Author information

Ethics declarations

Additional information

Electronic supplementary material Rights and permissions

About this article

Comments



### Nucleic Acids Research

Section browse ▼ Advance articles Submit ▼ Purchase About ▼ Issues All Nucleic Acids Resea **Article Contents** Genus for biomolecules 3 Paweł Rubach, Sebastian Zajac, Borys Jastrzebski, Joanna I Sulkowska, Abstract Piotr Sułkowski 🐯 INTRODUCTION Nucleic Acids Research, gkz845, https://doi.org/10.1093/nar/gkz845 DISCUSSION Published: 04 October 2019 Article history v FUNDING J PDF ■ Split View 66 Cite Permissions <! Share v REFERENCES Comments (0) Abstract The 'Genus for biomolecules' database (http://genus.fuw.edu.pl) collects information about topological structure and complexity of proteins and RNA chains, which is captured by the genus of a given chain and its subchains. For each biomolecule, this information is shown in the form of a genus trace plot, as well as a genus matrix diagram. We assemble such information for all and RNA structures deposited in the Protein Data Bank (PDB). This database

presents also various statistics and extensive information about the biological

## Poza Fizyką .... ale z matematykami z UKSW

## The category of chord diagrams

We fix three positive integers k,  $\ell$ , and m and a multifunction

$$\varphi \colon \ell \times \ell \to m$$

The objects of  $\mathfrak{C}_{k,arphi}$  are structures of the form

$$\mathbb{S} = \langle S, <^S, \{B_i^S\}_{i < k}, \{N_i^S\}_{i < \ell}, \{E_i\}_{i < m} \rangle,$$

#### where:

- (D1)  $\langle S, <^S \rangle$  is a finite linearly ordered set.
- (D2)  $\{B_i^S\}_{i < k}$  and  $\{N_i^S\}_{i < \ell}$  are partitions of S.
- (D3)  $B_{i_0} < B_{i_1}$  whenever  $i_0 < i_1 < k$ .
- (D4)  $\langle S, E_i \rangle$  is a graph for every i < m.
- (D5)  $E_{i_0} \cap E_{i_1} = \emptyset$  whenever  $i_0 \neq i_1$ .
- (D6) If  $x \in N_{i_0}$ ,  $y \in N_{i_1}$  and  $\langle x, y \rangle \in E_j$ , then  $j \in \varphi(i_0, i_1)$ .

The sets  $B_i$  are called *backbones*, while the sets  $N_i$  are *types of nodes* and  $E_i$  are *types of edges*.

## Poza Fizyką .... ale z matematykami z UKSW

A  $\mathfrak{C}_{k,\varphi}$ -morphism from  $\mathbb{S}$  to  $\mathbb{T}=\langle T,<^T,\{B_i^T\}_{i< k},\{N_i^T\}_{i< \ell},\{E_i\}_{i< m}\rangle$  is a mapping  $f\colon S\to T$  that preserves the linear orderings (that is,  $x<^Sy\implies f(x)<^Tf(y)$ ) and satisfies for every  $x,y\in S$ :

(M1) 
$$f(x) \in B_i^T \iff x \in B_i^S$$

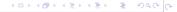
(M2) 
$$f(x) \in N_i^T \iff x \in N_i^S$$

(M3) 
$$\langle f(x), f(y) \rangle \in E_i^T \iff \langle x, y \rangle \in E_i^S$$
.

Informally, a  $\mathfrak{C}_{k,\varphi}$ -arrow is a mapping that preserves the structure of  $\mathbb{S}$ , "adding" new vertices and new edges of various types.

In the language of model theory,  $\mathfrak{C}_{k,arphi}$ -arrows are called *embeddings*.

It is clear that  $\mathfrak{C}_{k,\varphi}$  forms a category.



### Na SGH i w biznesie

- Nowy przedmiot dla studentów SGH Analizy danych w czasie rzeczywistym.
- Twórca portalu do analiz plików JPK
- Współtwórca darmowej i płatnej wersji biblioteki w pythonie do automatycznego generowania modeli scoringowych -Advanced Scorecard Builder.
- Programowanie i przetwarzanie danych oraz statystyka w SAS.
- Badania statutowe Metody doboru zmiennych z wykorzystaniem narzędziami machine i deep learning.

## SGH wydarzenia

- 28.XI.2019 Konferencja Analityka dla Biznesu
- 4.XII.2019 roku w Sali 2b w budynku C , odbędzie się promocja książki **Modelowanie dla biznesu**.

https://businessintelligence.pl/pl/wizualizacja-wynikow-modelowania-z-qlik-sense/



## Poza Światem











### Wyjaśnienie

"NoSTAT" czytamy i tłumaczymy jako **Not Only Stat**!

Dziękuję za uwagę!