The AI Mathematician

YANG-HUI HE

何楊輝

London Institute of Mathematical Sciences, Royal Institution

Merton College, University of Oxford

ICMS 2024, Durham

Classical Algebraic Geometry & Modern Computer Algebra: Innovative Software Design and its Applications

How does one DO mathematics?

Bottom-Up as a formal logical system

Top-Down as a creative/intuitive art

Meta-Mathematics as a language

Review, YHH: A Triumvirate of AI Driven Theoretical Discovery, 2405.19973 to appear *Nature Rev. Phys*

Review, YHH: Machine-Learning Mathematical Structures, 2101.06317 *IJMSDS '21*

C20th: Computers in Mathematics

- speed-up in computations & modelling: goes without saying
- crucial to increasing number of important theorems
  - 4-color [Appel-Haken-Koch 1976]
  - Kepler Conjecture [Hales 1998, formal check + acceptance 2017]
  - Classification of Finite Simple Groups [Galois 1832 - Gorenstein et al. 2008]
  - ...

Bottom-Up

Russell-Whitehead *Principia Mathematica* [1910s] programme (since at least Frege, even Leibniz) to axiomatize mathematics,

but . . .

Bottom-Up

Russell-Whitehead *Principia Mathematica* [1910s] programme (since at least Frege, even Leibniz) to axiomatize mathematics,

but . . . Gödel [1931] Incompleteness ; Church-Turing [1930s] Undecidability

but . . .

Bottom-Up

Russell-Whitehead *Principia Mathematica* [1910s] programme (since at least Frege, even Leibniz) to axiomatize mathematics,

but ... Gödel [1931] Incompleteness ; Church-Turing [1930s] Undecidability

but ... "The practicing mathematician hardly ever worries about Gödel" - MH Kim

Bottom-Up

Russell-Whitehead *Principia Mathematica* [1910s] programme (since at least Frege, even Leibniz) to axiomatize mathematics,

but . . . Gödel [1931] Incompleteness ; Church-Turing [1930s] Undecidability

but . . . "The practicing mathematician hardly ever worries about Gödel" - MH Kim

Automated Theorem Proving (ATP) a long tradition

- Newell-Simon-Shaw [1956] Logical Theory Machine $\rightsquigarrow$ **proved some *Principia***
- H. Wang [1961] Proving thm by pattern recognition
- Type Theory [1970s] Martin-Löf, Coquand
- Univalent Foundation / Homotopy Type Theory [2006-] Voevodsky

Bottom-Up (won't say much)

- Coq interactive proving system: 4-color (2005); Feit-Thompson Thm (2012);
- Lean (2013-) all of undergraduate maths
- Davenport: ICM 2018 "Computer Assisted Proofs"
- Buzzard: ICM 2022: XenaProject (Lean)
- over-optimistic view Szegedy (DeepMind): computers > humans @ chess (1990s); @ Go (2018); @ Proving theorems (2030)

Meta-Mathematics (won't say much)

2018 [YHH-Jejjala-Nelson] 1807.00735: $\sim 10^6$ titles of hep-th, hep-ph, gr-qc, math-ph, hep-lat from ArXiv 1989-2017 $\Rightarrow$ Word2Vec LLM

- Subfields on ArXiv has own linguistic particulars
- Science (ArXiv) / Pseudo-science (viXra) syntactically distinguishable

2019 Tshitoyan et al., **Nature** July : 3.3. million materials-science abstracts; uncovers structure of periodic table, predicts discoveries of new thermoelectric materials years in advance, and suggests as-yet unknown materials

2022 ChatGPT has passed the Turing Test

2023-24 LLM for Maths, DeepMind's FunSearch OpenAI's QStar, to appear; Meta-AI's LLama, Deepmind's AlphaGeo (53%)

Meta-Mathematics (won't say much)

2018 [YHH-Jejjala-Nelson] 1807.00735: $\sim 10^6$ titles of hep-th, hep-ph, gr-qc, math-ph, hep-lat from ArXiv 1989-2017 $\Rightarrow$ Word2Vec LLM

- Subfields on ArXiv has own linguistic particulars
- Science (ArXiv) / Pseudo-science (viXra) syntactically distinguishable

2019 Tshitoyan et al., **Nature** July : 3.3. million materials-science abstracts; uncovers structure of periodic table, predicts discoveries of new thermoelectric materials years in advance, and suggests as-yet unknown materials

2022 ChatGPT has passed the Turing Test

2023-24 LLM for Maths, DeepMind's FunSearch OpenAI's QStar, to appear; Meta-AI's LLama, Deepmind's AlphaGeo (53%) AlphaGeo (84%); AlphaProof (LLM+Lean)

Top-Down Mathematics: intuition, experience, experimentation

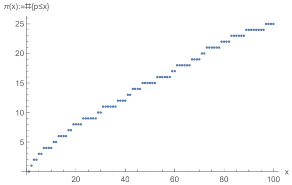
- In practice, Maths is Top-Down: practice before ($<$) foundation Countless eg:

  calculus $<$ analysis; alg geometry $<$ Bourbaki, permutations / Galois theory $<$ abstract algebra ...

- The best neural network of C18-19th?

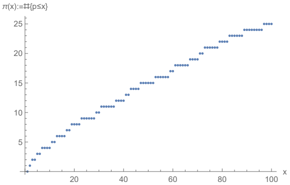
Top-Down Mathematics: intuition, experience, experimentation

- In practice, Maths is Top-Down: practice before ($<$) foundation Countless eg: calculus $<$ analysis; alg geometry $<$ Bourbaki, permutations / Galois theory $<$ abstract algebra ...
- The best neural network of C18-19th? brain of Gauß ; e.g., age 16



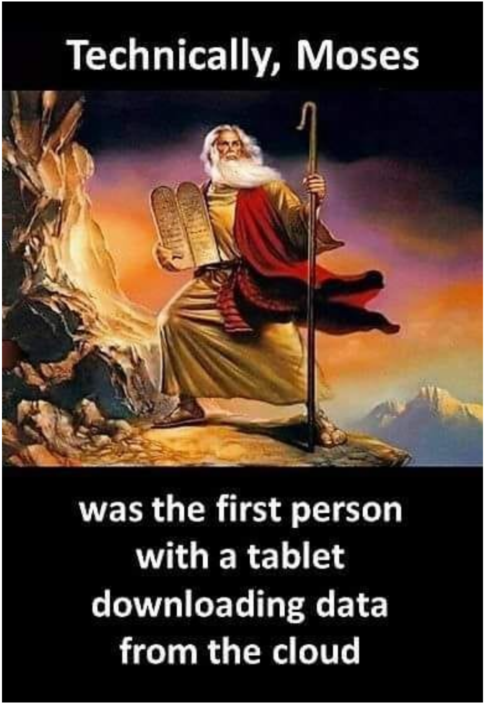
(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

- In practice, Maths is Top-Down: practice before (<) foundation Countless eg: calculus < analysis; alg geometry < Bourbaki, permutations / Galois theory < abstract algebra ...
- The best neural network of C18-19th? brain of Gauß ; e.g., age 16



(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

- BSD computer experiment of Birch & Swinnerton-Dyer [1960's] on plots of rank r & N_p on elliptic curves



The age of data science in mathematics/theoretical physics not as recent as you might think

Pattern Recognition: Human Eye

- $[0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, \ldots]$

Pattern Recognition: Human Eye

- $[0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, \ldots]$

  multiple of 3 or not.

- **[1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0 . . .]**

Pattern Recognition: Human Eye

- $[0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, \ldots]$

  multiple of 3 or not.

- **[1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0 ...]**

  Prime or Not for odd integers.

- **[1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, ...]**

Pattern Recognition: Human Eye

- $[0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, \ldots]$

  multiple of 3 or not.

- **[1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0 . . .]**

  Prime or Not for odd integers.

- **[1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, . . .]**

  Even/Odd of number of prime factors (Liouville Lambda)

What is Mathematics

A mathematician, like a painter or a poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas... *G. Hardy, A Mathematician's Apology*

What is Mathematics

A mathematician, like a painter or a poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas... *G. Hardy, A Mathematician's Apology*

One (only?) sure thing that AI can do better than humans is pattern detection.

Pattern Recognition: Machine-Learning

- Binary Classification of a Binary Vector (sliding window of, say, length 100); supervised learning: predict next one, e.g., Prime/Not becomes:

$$
\begin{array}{ccc}
\{0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,\ldots,0\} & \longrightarrow & 1 \\
\{1,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,\ldots,1\} & \longrightarrow & 0 \\
\{0,0,0,1,0,0,0,0,0,1,0,0,1,0,0,0,0,1,1,0,0,\ldots,0\} & \longrightarrow & 1 \\
\ldots & \ldots & \ldots
\end{array}
$$

Pattern Recognition: Machine-Learning

- Binary Classification of a Binary Vector (sliding window of, say, length 100); supervised learning: predict next one, e.g., Prime/Not becomes:

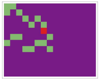
$$\begin{array}{ccc}
\{0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,\ldots,0\} & \longrightarrow & 1 \\
\{1,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,\ldots,1\} & \longrightarrow & 0 \\
\{0,0,0,1,0,0,0,0,0,1,0,0,1,0,0,0,0,1,1,0,0,\ldots,0\} & \longrightarrow & 1 \\
\ldots & \ldots & \ldots
\end{array}$$

- pass to standard classifiers: SVW, Bayes, Nearest Neighbour; NN of the form $\mathbb{R}^{100} \xrightarrow{\text{linear}} \mathbb{R}^{20} \xrightarrow{\tanh} \mathbb{R}^{20} \xrightarrow{\textbf{Round}\sum} \mathbb{Z}$, your kitchen sink, ...
- take 50,000 samples, 20-80 cross-validation, record (precision, MCC)
- similar performance for most: Mod3: (1.0, 1.0); PrimeQ, after balancing: (0.8, 0.6); Liouville Λ: (0.5, 0.001)

- A typical calculation:

$$X = \begin{pmatrix} 1&1&0&0&0&0&0&0\\ 1&0&1&0&0&0&0&0\\ 0&0&0&1&0&1&0&0\\ 0&0&0&0&1&0&1&0\\ 0&0&0&0&0&0&2&0\\ 0&1&1&0&0&0&0&1\\ 1&0&0&0&0&1&1&0\\ 0&0&0&1&1&0&0&1 \end{pmatrix} \longrightarrow \boxed{\text{What Bourbaki teaches us}} \longrightarrow h^{2,1}(X) = 22$$

- [YHH 1706.02714] Deep-Learning the Landscape, *PLB 774, 2017*; (cf. Feature in *Science* , Aug, vol 365 issue 6452, 2019): think of it as an image processing problem



$$\longrightarrow \boxed{\text{What Machine-Learning teaches us}} \longrightarrow 22$$

Thank you! Since 2017-

my fantastic students

Jiakang Bao, Elli Heyes, Ed Hirst

Tejas Acharya, Daatta Aggrawal, Malik Amir, Kieran Bull, Lucille Calmon, Siqi Chen, Suvajit Majumder, Maks Manko, Toby Peterken, Juan Pérez-Ipiña, Max Sharnoff, Yan Xiao

my wonderful collaborators

Physics: **Guillermo Arias-Tamargo, David Berman, Heng-Yu Chen, Andrei Constantin, Sebastián Franco, Vishnu Jejjala, Seung-Joo Lee, Andre Lukas, Shailesh Lal, Brent Nelson, Diego Rodriguez-Gomez, Zaid Zaz**

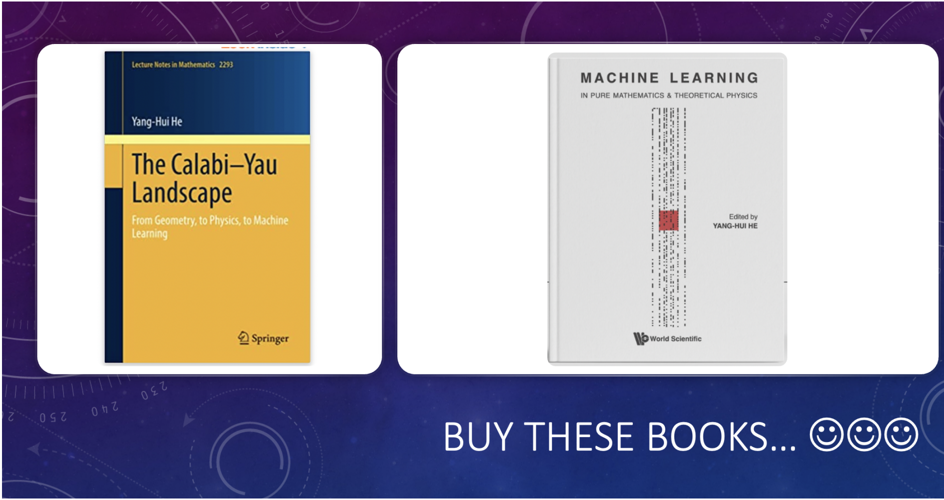
Algebraic Geometry: **Anthony Ashmore, Challenger Mishra, Rehan Deen, Burt Ovrut**

Number Theory: **Laura Alessandretti, Andrea Baronchelli, Kyu-Hwan Lee, Tom Oliver, Alexey Pozdnyakov, Drew Sutherland, Eldar Sultanow**

Representation Theory: **Mandy Cheung, Pierre Dechant, Minhyong Kim, Jianrong Li, Gregg Musiker**

Combinatorics: **Johannes Hofscheier, Alexander Kasprzyk, Shiing-Tung Yau**

Please buy

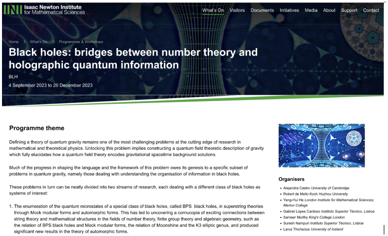


Lecture Notes in Mathematics 2293
Yang-Hui He
The Calabi–Yau Landscape
From Geometry, to Physics, to Machine Learning
Springer
MACHINE LEARNING
IN PURE MATHEMATICS & THEORETICAL PHYSICS
Edited by
YANG-HUI HE
World Scientific
BUY THESE BOOKS... ☺☺☺

The Birch Test

With Buzzard, Klemm, Nampuri, et al, inspired by a talk by Birch, we (half-jokingly) formulated the *Birch Test* (cf. chatGPT passed Turing test in 2022)

YHH, M. Burtsev, *Nature*, Jan 2024.


Isaac Newton Institute for Mathematical Sciences

What's On · Visitors · Documents · Initiatives · Media · About · Support · Contact

Home > What's On > Programmes & Workshops

Black holes: bridges between number theory and holographic quantum information

BLH

4 September 2023 to 20 December 2023

Programme theme

Defining a theory of quantum gravity remains one of the most challenging problems at the cutting edge of research in mathematical and theoretical physics. Unlocking this problem implies constructing a quantum field theoretic description of gravity which fully elucidates how a quantum field theory encodes gravitational spacetime background solutions.

Much of the progress in shaping the language and the framework of this problem owes its genesis to a specific subset of problems in quantum gravity, namely those dealing with understanding the organisation of information in black holes.

These problems in turn can be neatly divided into two streams of research, each dealing with a different class of black holes as systems of interest:

1. The enumeration of the quantum microstates of a special class of black holes, called BPS black holes, in superstring theories through Mock modular forms and automorphic forms. This has led to uncovering a cornucopia of exciting connections between string theory and mathematical structures in the fields of number theory, finite group theory and algebraic geometry, such as the relation of BPS black holes and Mock modular forms, the relation of Moonshine and the K3 elliptic genus, and produced significant new results in the theory of automorphic forms.

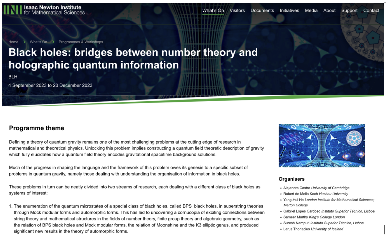
Organisers

- Alejandra Castro *University of Cambridge*
- Robert de Mello Koch *Huzhou University*
- Yang-Hui He *London Institute for Mathematical Sciences; Merton College*
- Gabriel Lopes Cardoso *Instituto Superior Técnico, Lisboa*
- Sameer Murthy *King's College London*
- Suresh Nampuri *Instituto Superior Técnico, Lisboa*
- Larus Thorlacius *University of Iceland*

The Birch Test

With Buzzard, Klemm, Nampuri, et al, inspired by a talk by Birch, we (half-jokingly) formulated the *Birch Test* (cf. chatGPT passed Turing test in 2022)

YHH, M. Burtsev, *Nature*, Jan 2024.


Isaac Newton Institute for Mathematical Sciences
What's On | Visitors | Documents | Initiatives | Media | About | Support | Contact
Home > What's On > Programmes & Workshops

Black holes: bridges between number theory and holographic quantum information

BLH

4 September 2023 to 20 December 2023

Programme theme

Defining a theory of quantum gravity remains one of the most challenging problems at the cutting edge of research in mathematical and theoretical physics. Unlocking this problem implies constructing a quantum field theoretic description of gravity which fully elucidates how a quantum field theory encodes gravitational spacetime background solutions.

Much of the progress in shaping the language and the framework of this problem owes its genesis to a specific subset of problems in quantum gravity, namely those dealing with understanding the organisation of information in black holes.

These problems in turn can be neatly divided into two streams of research, each dealing with a different class of black holes as systems of interest:

1. The enumeration of the quantum microstates of a special class of black holes, called BPS black holes, in superstring theories through Mock modular forms and automorphic forms. This has led to uncovering a cornucopia of exciting connections between string theory and mathematical structures in the fields of number theory, finite group theory and algebraic geometry, such as the relation of BPS black holes and Mock modular forms, the relation of Moonshine and the K3 elliptic genus, and produced significant new results in the theory of automorphic forms.

Organisers

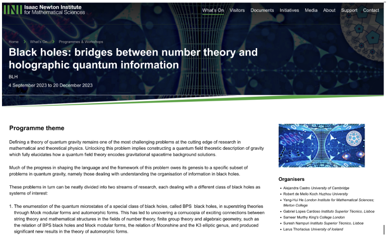
- Alejandra Castro *University of Cambridge*
- Robert de Mello Koch *Huzhou University*
- Yang-Hui He *London Institute for Mathematical Sciences; Merton College*
- Gabriel Lopes Cardoso *Instituto Superior Técnico, Lisboa*
- Sameer Murthy *King's College London*
- Suresh Nampuri *Instituto Superior Técnico, Lisboa*
- Larus Thorlacius *University of Iceland*

- (Automaticity) be generated by AI

The Birch Test

With Buzzard, Klemm, Nampuri, et al, inspired by a talk by Birch, we (half-jokingly) formulated the *Birch Test* (cf. chatGPT passed Turing test in 2022)

YHH, M. Burtsev, *Nature*, Jan 2024.


Isaac Newton Institute for Mathematical Sciences

What's On Visitors Documents Initiatives Media About Support Contact

Home > What's On > Programmes & Workshops

Black holes: bridges between number theory and holographic quantum information

BLH

4 September 2023 to 20 December 2023

Programme theme

Defining a theory of quantum gravity remains one of the most challenging problems at the cutting edge of research in mathematical and theoretical physics. Unlocking this problem implies constructing a quantum field theoretic description of gravity which fully elucidates how a quantum field theory encodes gravitational spacetime background solutions.

Much of the progress in shaping the language and the framework of this problem owes its genesis to a specific subset of problems in quantum gravity, namely those dealing with understanding the organisation of information in black holes.

These problems in turn can be neatly divided into two streams of research, each dealing with a different class of black holes as systems of interest:

1. The enumeration of the quantum microstates of a special class of black holes, called BPS black holes, in superstring theories through Mock modular forms and automorphic forms. This has led to uncovering a cornucopia of exciting connections between string theory and mathematical structures in the fields of number theory, finite group theory and algebraic geometry, such as the relation of BPS black holes and Mock modular forms, the relation of Moonshine and the K3 elliptic genus, and produced significant new results in the theory of automorphic forms.

Organisers

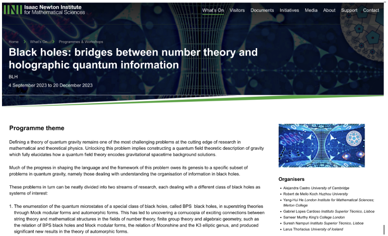
- Alejandra Castro *University of Cambridge*
- Robert de Mello Koch *Huzhou University*
- Yang-Hui He *London Institute for Mathematical Sciences; Merton College*
- Gabriel Lopes Cardoso *Instituto Superior Técnico, Lisboa*
- Sameer Murthy *King's College London*
- Suresh Nampuri *Instituto Superior Técnico, Lisboa*
- Larus Thorlacius *University of Iceland*

- (**A**utomaticity) be generated by AI
- (**I**nterpretability) concrete enough to be a conjecture

The Birch Test

With Buzzard, Klemm, Nampuri, et al, inspired by a talk by Birch, we (half-jokingly) formulated the *Birch Test* (cf. chatGPT passed Turing test in 2022)

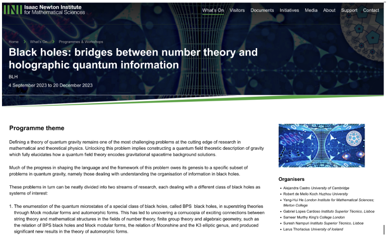
YHH, M. Burtsev, *Nature*, Jan 2024.



- (**A**utomaticity) be generated by AI
- (**I**nterpretability) concrete enough to be a conjecture
- (**N**on-Triviality) for the community to work on it

The Birch Test

With Buzzard, Klemm, Nampuri, et al, inspired by a talk by Birch, we (half-jokingly) formulated the *Birch Test* (cf. chatGPT passed Turing test in 2022) YHH, M. Burtsev, *Nature*, Jan 2024.



Isaac Newton Institute for Mathematical Sciences

What's On | Visitors | Documents | Initiatives | Media | About | Support | Contact

Home > What's On > Programmes & Workshops

Black holes: bridges between number theory and holographic quantum information

BLH

4 September 2023 to 20 December 2023

Programme theme

Defining a theory of quantum gravity remains one of the most challenging problems at the cutting edge of research in mathematical and theoretical physics. Unlocking this problem implies constructing a quantum field theoretic description of gravity which fully elucidates how a quantum field theory encodes gravitational spacetime background solutions.

Much of the progress in shaping the language and the framework of this problem owes its genesis to a specific subset of problems in quantum gravity, namely those dealing with understanding the organisation of information in black holes.

These problems in turn can be neatly divided into two streams of research, each dealing with a different class of black holes as systems of interest:

1. The enumeration of the quantum microstates of a special class of black holes, called BPS black holes, in superstring theories through Mock modular forms and automorphic forms. This has led to uncovering a cornucopia of exciting connections between string theory and mathematical structures in the fields of number theory, finite group theory and algebraic geometry, such as the relation of BPS black holes and Mock modular forms, the relation of Moonshine and the K3 elliptic genus, and produced significant new results in the theory of automorphic forms.

Organisers

- Alejandra Castro *University of Cambridge*
- Robert de Mello Koch *Huzhou University*
- Yang-Hui He *London Institute for Mathematical Sciences; Merton College*
- Gabriel Lopes Cardoso *Instituto Superior Técnico, Lisboa*
- Sameer Murthy *King's College London*
- Suresh Nampuri *Instituto Superior Técnico, Lisboa*
- Larus Thorlacius *University of Iceland*

- (Automaticity) be generated by AI
- (Interpretability) concrete enough to be a conjecture
- (Non-Triviality) for the community to work on it
- make Birch happy

Candidates for Birch Test, 2017 -

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

2019 ML Graphs Euler/Hamilton same accuracy [YHH-Yau] (not interpretable)

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

2019 ML Graphs Euler/Hamilton same accuracy [YHH-Yau] (not interpretable)

2022 ML Knots new knot invariants relations from saliency [Davies et al] (not non-trivial enough)

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

2019 ML Graphs Euler/Hamilton same accuracy [YHH-Yau] (not interpretable)

2022 ML Knots new knot invariants relations from saliency [Davies et al] (not non-trivial enough)

2023 ML Fanos Fano dimension from quantum period [Coates-Kaspryzyk-Veneziale] (not interpretable)

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

2019 ML Graphs Euler/Hamilton same accuracy [YHH-Yau] (not interpretable)

2022 ML Knots new knot invariants relations from saliency [Davies et al] (not non-trivial enough)

2023 ML Fanos Fano dimension from quantum period [Coates-Kaspryzyk-Veneziale] (not interpretable)

2024 GA CY4 GA finds new CY4 [Berglund-YHH-Heyes-Hirst-Jejjala-Lukas] (not non-trivial enough)

. . .

Candidates for Birch Test, 2017 -

2017-8 ML Topological Invariants Calabi-Yau Hodge numbers using NN [YHH] (not interpretable)

2018 ML Simple Groups SVM Separation of simple/non-simple group [YHH-Kim] (not precise enough)

2019 ML Graphs Euler/Hamilton same accuracy [YHH-Yau] (not interpretable)

2022 ML Knots new knot invariants relations from saliency [Davies et al] (not non-trivial enough)

2023 ML Fanos Fano dimension from quantum period [Coates-Kaspryzyk-Veneziale] (not interpretable)

2024 GA CY4 GA finds new CY4 [Berglund-YHH-Heyes-Hirst-Jejjala-Lukas] (not non-trivial enough)

. . .

2022 Murmuration Phenomenon A new pattern in the primes, relation to BSD and a bias in L-coefficients of elliptic curves [YHH-Lee-Oliver-Podznyakov, 2022, YHH-Lee-Oliver-Podznyakov-Sutherland, 2024] made Buzzard/Birch almost happy (still completely since human intervention was needed)

ML Experiments in Number Theory

- [YHH 1706.02714, 1812.02893:]
  - Predicting primes $2 \to 3,\ 2,3 \to 5,\ 2,3,5 \to 7$; no way
  - PrimeQ: (0.7, 0.8); Sarnak's Challenge of Liouville Lambda (0.5, 0.001)
- [Alessandretti-Baronchelli-YHH 1911.02008] ML/TDA@BSD, naive attempt
- Arithmetic Geometry: A Modern Hope? YHH-KH Lee-Oliver
  - 2010.01213: Complex Multiplication, Sato-Tate $(0.99 \sim 1.0, 0.99 \sim 1.0)$
  - 2011.08958: Number Fields: rank and Galois group (0.97, 0.9)
  - 2012.04084: BSD from Euler coeffs, integer points, torsion (0.99, 0.9);

    Tate-Shafarevich $Ш$ (0.6, 0.8) [Hardest quantity of BSD]

AI-Driven Mathematical Discovery: Murmuration

YHH, Kyu-Hwan Lee, Tom Oliver, Alexey Pozdnyakov (2204.10140), 2022-

Quanta Feature 2024:

- E an elliptic curve, local zeta-function & L-function:

$$Z(E/\mathbb{F}_p;T) = \exp\left(\sum_{k=1}^{\infty} \frac{\#E(\mathbb{F}_{p^k})T^k}{k}\right) = \frac{L_p(E,T)}{(1-T)(1-pT)} \ ;$$
$$L_p(E,T) = 1 - a_p T + pT^2; \quad a_p = p + 1 - \#E\left(\mathbb{F}_p\right).$$

Fix N and define vector $v_L(E) = (a_{p_1}, \ldots, a_{p_N}) \in \mathbb{Z}^N$;

$\sim 10^5$ balanced data from www.lmfdb.org; 50-50 cross validation.

- Labeled data: $v_L(E) \longrightarrow$ rank, torsion, ... ([Birch-Swinnerton-Dyer:])

$$L(E,s) := \prod_p L^{-1}(E, T := p^{-s}); \qquad \frac{L^{(r)}(E,1)}{r!} \stackrel{???}{=} \frac{|Ш|\Omega \mathrm{Reg} \prod_p c_p}{(\#E(\mathbb{Q})_{\mathrm{tors}})^2},$$

r=rank; Ш=Shafarevich group; Reg=regulator; c_p=Tamagawa; tors=Torsion

Important Lesson: HYBRID human-AI math

Importance of Representation

(Alessandretti-Baronchelli-YHH 1911.02008, *New Scientist* feature 2019 used Weierstrass coefficients of elliptic curves: useless in predicting any of the BSD quantities

needed insights from Oliver+Lee to use a_p coefficients

Importance of Human Interpretation

Murmurations of elliptic curves: YHH, Lee, Oliver, Pozdnyakov, 2204.10140

A new mathematical phenomenon

The Murmuration Phenomenon

Q: YHH, Lee, Oliver, Pozdnyakov on HLOP results from 2020 - 22: WHY is ML so good at telling ranks apart by looking at a_p coefficients?? e.g., PCA:

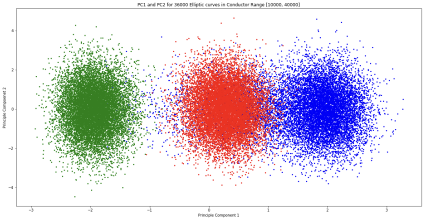


Figure 2: A plot of PC1 (x-axis) against PC2 (y-axis) for elliptic curves in the balanced dataset of $36,000$ randomly chosen elliptic curves with rank $r_E \in \{0,1,2\}$ and conductor $N_E \in [10000, 40000]$. The blue (resp. red, green) points are the images of the vectors $v_L(E)$ corresponding to the elliptic curves in our dataset with rank 0 (resp. 1, 2) under a map $\mathbb{R}^{1000} \to \mathbb{R}^2$ constructed using PCA.

Murmuration function

construct a vertical average (rank r, conductor range $[N_1, N_2]$, n-th prime p_n)

$$f_r(n) := \frac{1}{\#\mathcal{E}_r[N_1, N_2]} \sum_{E \in \mathcal{E}_r[N_1,N_2]} a_{p_n}(E)$$

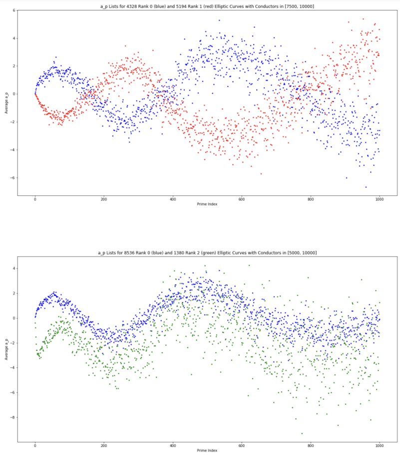


Figure 1: (Top) Plots of the functions $f_0(n)$ (blue) and $f_1(n)$ (red) for $1 \leq n \leq 1000$ and $[N_1, N_2] = [7500, 10000]$. (Bottom) Plots of the functions $f_0(n)$ (blue) and $f_2(n)$ (green) for $1 \leq n \leq 1000$ and $[N_1, N_2] = [5000, 10000]$. Further details are given in Example 1.

Murmurations: An interesting Phenomenon

- To appear [HLOP + Sutherland]
  - Does not work if ordered by height (Weierstrass coef)
  - Take dyadic conductor range: $[N^x, N^{x+1}]$: scale invariant (indep of x)
  - Taking more data $(10^{7\sim 8})$ at high N: converges to oscillatory curve
- A General Phenomenon that reflects biases in distribution of primes
  - L-function for Dirichlet characters (Lee-Oliver-Podznyakov 2023)
  - Zubrilina, Cowan: for weight 2 modular forms (2023)
  - conference at ICERM in July

    https://icerm.brown.edu/events/htw-23-ma/

Mathematical Conjectures: Three Centuries

Conjecture Formulation

C19th Gauss's eyes on $\pi(x) \sim \int_2^x \frac{dx}{\ln(x)}$

C20th Birch + Swinnerton-Dyer on the EDSAC-2 computer@Cambridge

C21st AI guided human intuition:

Knots $\rightsquigarrow$ New Expressions for Invariants (DeepMind)

LMFdD $\rightsquigarrow$ Murmuration Conjectures (YHH-Lee-Oliver-Poznyakov)

New Matrix Multiplication (DeepMind)

Mathematical Conjectures: Three Centuries

Conjecture Formulation

C19th Gauss's eyes on $\pi(x) \sim \int_2^x \frac{dx}{\ln(x)}$

C20th Birch + Swinnerton-Dyer on the EDSAC-2 computer@Cambridge

C21st AI guided human intuition:

Knots $\rightsquigarrow$ New Expressions for Invariants (DeepMind)

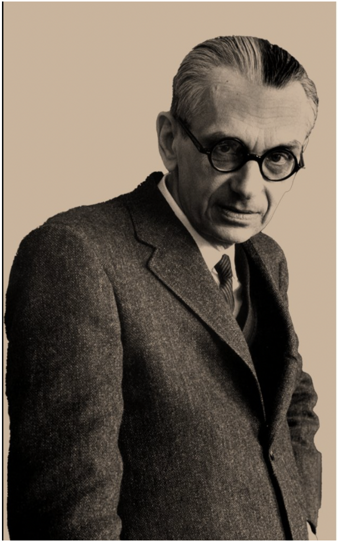
LMFdD $\rightsquigarrow$ Murmuration Conjectures (YHH-Lee-Oliver-Poznyakov)

New Matrix Multiplication (DeepMind)

The future of mathematics is a combination of

- Bottom-up ATP using AI
- Top-Down machine-guided human intuition using AI
- Mathematics as LLM using AI

THANK YOU



Either **MATHEMATICS** is too big for the human mind, or the human mind is more than a machine.

- KURT GÖDEL
(1906-1978)

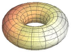
Initial Motivation: A Classic Problem (since 1736)

- Trichtomy classification of (connected compact orientable) surfaces Σ

Euler: topological classification of $\dim_{\mathbb{R}} = 2$ Euler number $\chi(\Sigma)$, genus $g(\Sigma)$

Gauss: relates topology to metric geometry

Riemann: complexify $\rightsquigarrow$ Riemann surfaces or complex curves: $\dim_{\mathbb{C}} = 1$

		 ...
$g(\Sigma) = 0$	$g(\Sigma) = 1$	$g(\Sigma) > 1$
$\chi(\Sigma) = 2$	$\chi(\Sigma) = 0$	$\chi(\Sigma) < 0$
Spherical	Ricci-Flat	Hyperbolic
$+$ curvature	0 curvature	$-$ curvature

Classical Results for Riemann Surface Σ

$\chi(\Sigma) = 2 - 2g(\Sigma) =$	$= [c_1(\Sigma)] \cdot [\Sigma] =$	$= \frac{1}{2\pi} \int_\Sigma R =$	$= \sum_{i=0}^{2} (-1)^i h^i(\Sigma)$
Topology	Algebraic Geometry	Differential Geometry	Index Theorem (co-)Homology
Invariants	Characteristic classes	Curvature	Betti Numbers

Going up in Complex Dimension

- $\dim_{\mathbb{R}} > 2$ manifolds extremely complicated
- Luckily, for a special class of complex manifolds called **Kähler**

$$g_{\mu\bar{\nu}} = \partial_\mu \partial_{\bar{\nu}} K(z, \bar{z})$$

all Σ in $\dim_{\mathbb{C}} = 1$ automatically Kähler

- CONJECTURE [E. Calabi, 1954, 1957]: M compact Kähler manifold (g, ω) and $([R] = [c_1(M)])_{H^{1,1}(M)}$.
Then $\exists!(\tilde{g}, \tilde{\omega})$ such that $([\omega] = [\tilde{\omega}])_{H^2(M;\mathbb{R})}$ and $Ricci(\tilde{\omega}) = R$.

Rmk: $c_1(M) = 0 \Leftrightarrow$ Ricci-flat (rmk: Ricci-flat familiar to physicists through GR)

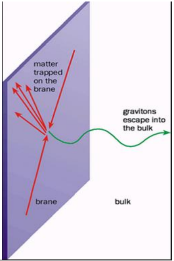
- THEOREM [S-T Yau, 1977-8; Fields 1982] Existence Proof

Superstring Theory 9+1 d

Unified theory of quantum gravity

I. 6 Large Dim

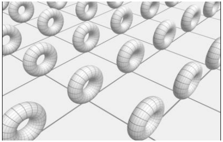
AdS/CFT
Brane World



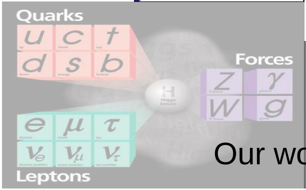
1. *Reduce Dim: 10 = 6+4*
2. *Break SUSY*

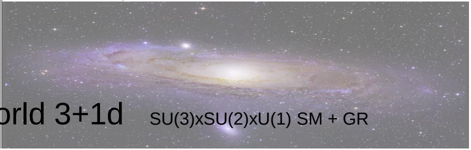
II. 6 small dim

Compactification



String





Our world 3+1d   SU(3)xSU(2)xU(1) SM + GR

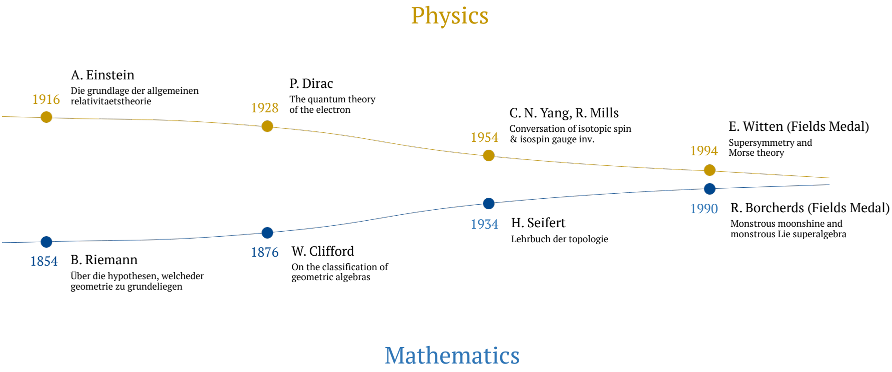
Phenomenology [Candelas-Horowitz-Strominger-Witten]: 1985

simplest solution of 6 extra dimensions: Ricci-Flat, Kähler $\dim_{\mathbb{C}} = 3$

When Physics meets Maths

- Strominger was next door to Yau in 1986 at the IAS, physicists called Ricci-Flat, Kähler manifolds, CHSW called these Calabi-Yau manifolds
- GEOMETRIZATION PROGRAMME: Historically, the right language of physics is increasingly geometrical:
  - Gravity/Space-time $\rightsquigarrow$ GR $\rightsquigarrow$ Differential geometry;
  - Particle physics/Standard Model $\rightsquigarrow$ Gauge Theory/Yang-Mills $\rightsquigarrow$ Algebraic geometry (bundles/connections) + group theory (Lie and Finite groups);
  - Condensed matter physics of topological insulators $\rightsquigarrow$ algebraic topology; ...
  - String theory is a brain-child of this tradition
- TAKE-HOME MESSAGE: Whenever physics and maths converge and generate new ideas, the right things are happening

The Confluence of Maths and Physics



The Confluence of Maths and Physics

1959

2010

The Unreasonable Effectiveness of Mathematics in the Natural Sciences

Richard Courant Lecture in Mathematical Sciences delivered at New York University, May 11, 1959

EUGENE P. WIGNER
Princeton University

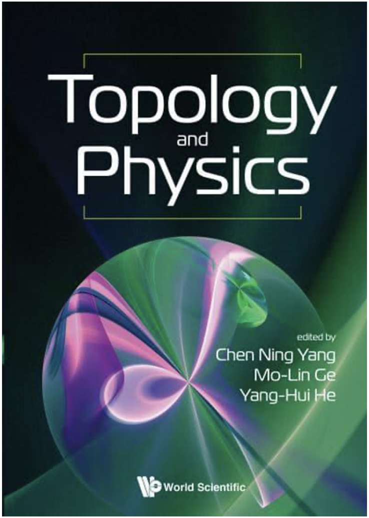
Phil. Trans. R. Soc. A (2010) **368**, 913–926

Geometry and physics

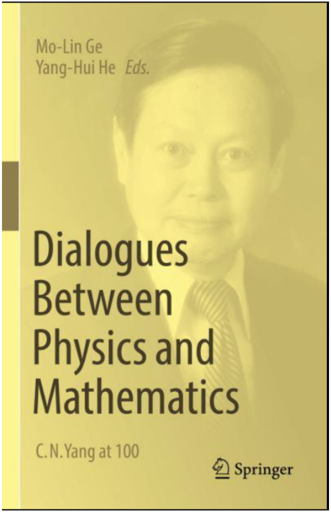
BY MICHAEL ATIYAH[1], ROBBERT DIJKGRAAF[2,*] AND NIGEL HITCHIN[3]

[1]*School of Mathematics, University of Edinburgh, Edinburgh EH9 3JZ, UK*
[2]*Institute for Theoretical Physics, University of Amsterdam, Valckenierstraat 65, 1018 Amsterdam, The Netherlands*
[3]*Mathematical Institute, University of Oxford, 24–29 St Giles, Oxford OX1 3LB, UK*

"One may be tempted to invert Wigner's comment and marvel at 'the unreasonable effectiveness of physics in mathematics.'"



CN Yang, ML Ge & YH He, ed, World Scientific, 2019 contributions: Atiyah, Dijkgraaf, Kim, Penrose, Witten, et al.



ML Ge & YH He, ed, Springer-Nature, 2022 contributions: Drinfeld, Leggett, Manin, Penrose, Polyakov, Wilczek, Witten, et al.

Progress in String Theory: Start Dates of Annual Series

1986- "Strings" Conference

2002- "StringPheno" Conference

2006 - 2010 String Vacuum Project (NSF)

2008 - ISGT Integrability in String/Gauge

2011- "String-Math" Conference (2020 - , M-theory & Maths Workshop)

2012- "Amplitudes"

2014- String/Theoretical Physics Session in SIAM Conference

2017- "String-Data" Conference

Back to ML for Maths

Computing Hodge Numbers: Sketch

- Recall Hodge decomposition $H^{p,q}(X) \simeq H^q(X, \wedge^p T^\star X) \rightsquigarrow$

$$H^{1,1}(X) = H^1(X, T_X^\star), \qquad H^{2,1}(X) \simeq H^{1,2} = H^2(X, T_X^\star) \simeq H^1(X, T_X)$$

- Euler Sequence for subvariety $X \subset A$ is short exact:

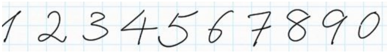
$$0 \to T_X \to T_M|_X \to N_X \to 0$$

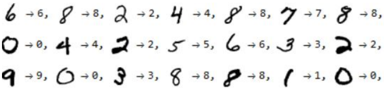
- Induces long exact sequence in cohomology:

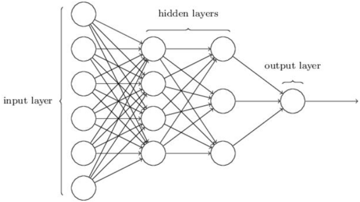
$$\begin{array}{ccccccccc} 0 & \to & \cancel{H^0(X, T_X)} \overset{0}{\nearrow} & \to & H^0(X, T_A|_X) & \to & H^0(X, N_X) & \to \\ & \to & \boxed{H^1(X, T_X)} & \overset{d}{\to} & H^1(X, T_A|_X) & \to & H^1(X, N_X) & \to \\ & \to & H^2(X, T_X) & \to & \ldots & & & \end{array}$$

- Need to compute Rk(d), cohomology and $H^i(X, T_A|_X)$ (Cf. Hübsch)

Back to AG

- Bijection from  to $\{1, 2, \ldots, 9, 0\}$?
- Take large sample, take a few hundred thousand (e.g. NIST database)





- Data = Training Data $\sqcup$ Validation Data

  Test trained NN on validations data to see accuracy performance

Universal Approximation Theorems

Large Depth Thm: (Cybenko-Hornik) For every continuous function $f : \mathbb{R}^d \to \mathbb{R}^D$, every compact subset $K \subset \mathbb{R}^d$, and every $\epsilon > 0$, there exists a continuous function $f_\epsilon : \mathbb{R}^d \to \mathbb{R}^D$ such that $f_\epsilon = W_2(\sigma(W_1))$, where σ is a fixed continuous function, $W_{1,2}$ affine transformations and composition appropriately defined, so that $\sup_{x \in K} |f(x) - f_\epsilon(x)| < \epsilon$.

Large Width Thm: (Kidger-Lyons) Consider a feed-forward NN with n input neurons, m output neuron and an arbitrary number of hidden layers each with $n + m + 2$ neurons, such that every hidden neuron has activation function φ and every output neuron has activation function the identity. Then, given any vector-valued function f from a compact subset $K \subset \mathbb{R}^m$, and any $\epsilon > 0$, one can find an F, a NN of the above type, so that $|F(x) - f(x)| < \epsilon$ for all $x \in K$.

ReLU Thm: (Hanin) For any Lebesgue-integral function $f : \mathbb{R}^n \to \mathbb{R}$ and any $\epsilon > 0$, there exists a fully connected ReLU NN F with width of all layers less than $n + 4$ such that $\int_{\mathbb{R}^n} |f(x) - F(x)| dx < \epsilon$.

Back to NN@Alg Geo