Universes as Bigdata:

or, Machine-Learning Mathematical Structures

YANG-HUI HE

Dept of Mathematics, City, **University of London**

Merton College, **University of Oxford**

School of Physics, **NanKai University**

Cambridge Accelerate Programme:

Winter School on ML & Beyond, Jan, 2021

Unification: Scientific Reductionism

- Isaac Newton: Universal Gravitation and motion

  Principia Mathematica Philosophiæ Naturalis (1687)

- James Clerk Maxwell: Electricity + Magnetism

  Treatise on Electricity and Magnetism (1873)

- Albert Einstein: Gravity and Motion + Structure of 4-dimensional space-time
  - *Zur Elektrodynamik bewegter Körper* (1905)
  - *Die Formale Grundlage der Allgemeinen Relativitätstheorie* (1914)

- Planck, Bohr, Heisenberg, CN Yang, Feynmann, ... (1901 -)

  All elementary particles $\rightsquigarrow$ Quantum Field Theory, only 3 forces:

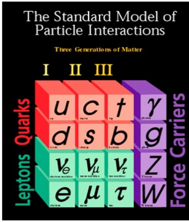
  Strong, weak nuclear, eletro-magnetic

- 400 years of mathematical physics: Gotten down to TWO! ...

Two Pillars of Modern Physics

1. MACROSCOPIC (General Relativity) GR
    - motions of galaxies and stars
    - Gravitational Waves: 2018, new era for cosmology $p \sim 99.99994\%$

2. MICROSCOPIC (Quantum Field Theory) QFT
    - elementary particles $= SU(3)_S \times [SU(2) \times U(1)]_{EW}$ gauge theory
    - experimentally verified to 19 digits!
    - Higgs Boson: 2014 at LHC, CERN, last piece to SM!



- Biggest Question in Fundamental Science: 1 + 2 does not work

  Unrenormalizability of gravity; Einstein's & Hawking's dream of

  Theory of Everything ToE

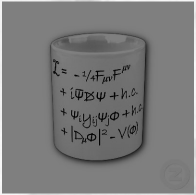
The Geometrization Programme

- Algebraic/differential geometry/topology : the right language for physics (realized by C20th)
  - Space-Time = GR = Differential Geometry (Riemann-Einstein)
  - Interactions = particles = QFT = Algebraic Geometry + Group Theory (Weyl, Weil, CN Yang-SS Chern, Wigner, Atiyah, 't Hooft ...)

    see CN.Yang-ML.Ge-YHH, *Topology & Physics*, WS 2019
  - **String theory: brain-child of gauge-gravity geometrization tradition** (Witten, Maldacena, ...)

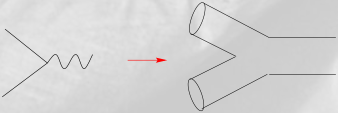
    since 1970s; manages to give a ToE. ...with a caveat
- THIS TALK: C21st A new exciting era: for synergy with (pure & computational) geometry, group theory, combinatorics, number theory, algorithms, data science.

Paradigm Shift: Points $\rightsquigarrow$ Strings

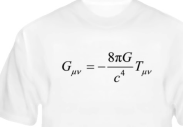


Standard Model

- Universe gives a hint: Planck Length
  $L_{Planck} = \sqrt{\frac{\hbar G_{Newton}}{c^3}} \sim 10^{-35} m$
- Fund constituents 1-dim, NOT points (0-dim)? $\rightsquigarrow$ strings, size L_{Planck}
- PREMISE: All particles, all of space-time, all of reality are vibration modes of an OPEN or CLOSED string



- WORKS only in 10 dimensions



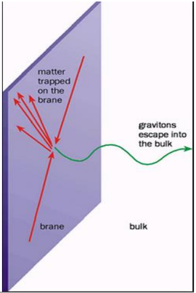
General Relativity

Superstring Theory 9+1 d

Unified theory of quantum gravity

I. 6 Large Dim

AdS/CFT
Brane World

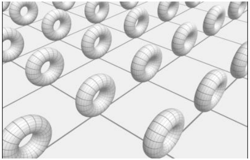


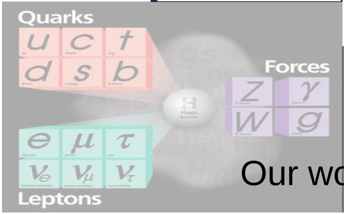
1. Reduce Dim: 10 = 6+4

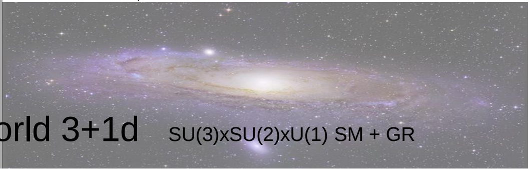
2. Break SUSY

II. 6 small dim

Compactification



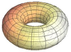




Our world 3+1d   SU(3)xSU(2)xU(1) SM + GR

$10 = 4 + 3 \times 2$

- String Phenomenology [Candelas-Horowitz-Strominger-Witten]: 1985
  - Standard Solution, e.g., $\mathbb{R}^{3,1} \times X$, X is not just 6-manifold, but a Complex 3-fold, Ricci-flat (vacuum Einstein), Kähler (SUSY, $g_{\mu\bar{\nu}} = \partial_\mu \bar{\partial}_{\bar{\nu}} K$)
- *mathematicians independently thinking of same problem*:
  - Euler, Gauss, Riemann Σ: $\dim_{\mathbb{R}} = 2, i.e., \dim_{\mathbb{C}} = 1$ (in fact Kähler)
  - Trichtomy classification of (compact orientable) surfaces [Riemann surfaces/complex algebraic curves] Σ Euler number $\chi(\Sigma)$, genus $g(\Sigma)$

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

Calabi-Yau

- Want to generalize:

$\chi(\Sigma) = 2 - 2g(\Sigma) =$	$= [c_1(\Sigma)] \cdot [\Sigma] =$	$= \frac{1}{2\pi} \int_\Sigma R =$	$= \sum\limits_{i=0}^{2} (-1)^i h^i(\Sigma)$
Topology	Algebraic Geom.	Differential Geom.	Index/(co-)cohomology
Invariants	Characteristic class	Curvature	Betti Numbers

- HIGH DIMENSION: HARD, Luckily, for our class of **Kähler** manifolds:
  - CONJECTURE [Eugenio Calabi , 1954, 1957]: M compact Kähler (g, ω), $([R] = [c_1(M)])_{H^{1,1}(M)} \Rightarrow \exists!(\tilde{g}, \tilde{\omega})$ s.t. $([\omega] = [\tilde{\omega}])_{H^2(M;\mathbb{R})}$, $Ricci(\tilde{\omega}) = R$.

Rmk: $c_1(M) = 0 \Leftrightarrow$ Ricci-flat (rmk: Ricci-flat familiar in GR long before strings)

  - THEOREM [S-T Yau , 1977-8; Fields 1982] Existence Proof
  - Calabi-Yau: Kähler and Ricci-flat (Strominger & Yau were neighbours at IAS)

The Inevitability of Algebraic Geometry

- Construct CY3?: CY Examples

  realize as zero sets of polynomials, **Algebraic Geometry** e.g.,

  $\{(p,q)|p^2+q^2-1=0\} \subset \mathbb{R}^2$ is a circle (1-real dimension)

- Complexify and Projectivize (Projective algebraic variety)
  - Cubic equation in $\mathbb{CP}^2$: e.g. CY1 $= T^2$ $\{(x,y,z)|x^3+y^3+z^3=0\} \subset \mathbb{CP}^2$ (elliptic curve); $\dim_{\mathbb{C}} = 2-1 = 1$
  - TMH: Homogeneous Eq in $\mathbb{CP}^n$, degree $= n+1$ is Calabi-Yau of $\dim_{\mathbb{C}} = n-1$

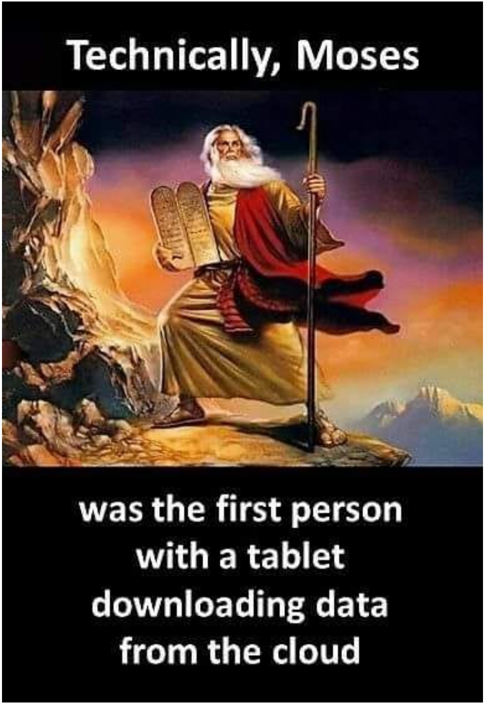
- **An Early Physical Challenge to Algebraic Geometry**
  - Particle content in [CHSW]

$$\left.\begin{array}{ll} \text{Generation} & h^1(X,TX) = h^{2,1}_{\bar\partial}(X) \\ \text{Anti-Generation} & h^1(X,TX^*) = h^{1,1}_{\bar\partial}(X) \end{array}\right\} \begin{array}{l} \text{Net-gen: } \chi = 2(h^{1,1}-h^{2,1}) \\ = \text{Euler Number } (X) \end{array}$$

- 1986 Question: Are there Calabi-Yau threefolds with $\chi = \pm 6$?

The First Data-sets in Mathematical Physics/Geometry

- [Candelas-A. He-Hübsch-Lutken-Schimmrigk-Berglund] (1986-1990)
  - CICYs (complete intersection CYs) multi-deg polys in products of $\mathbb{CP}^{n_i}$ CICYs
  - Problem: *classify all configuration matrices*; employed the best computers at the time (**CERN supercomputer**); q.v. magnetic tape and dot-matrix printout in Philip's office
  - 7890 matrices, 266 Hodge pairs $(h^{1,1}, h^{2,1})$, 70 Euler $\chi \in [-200, 0]$
- [Candelas-Lynker-Schimmrigk, 1990]
  - Hypersurfaces in Weighted P4
  - 7555 inequivalent 5-vectors w_i, 2780 Hodge pairs, $\chi \in [-960, 960]$
- [Kreuzer-Skarke, mid-1990s - 2000] Reflexive Polytopes
  - Batyrev-Borisov: Hypersurfaces in (Reflexive, Gorenstein Fano) Toric 4-folds
  - 6-month running time on dual Pentium SGI machine
  - at least 473,800,776, with 30,108 distinct Hodge pairs, $\chi \in [-960, 960]$



The age of data science in mathematical physics/string theory not as recent as you might think

of course, experimental physics had been decades ahead in data-science/machine-learning

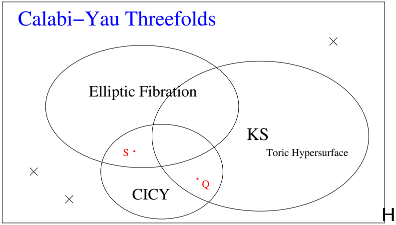
After 40 years of research by mathematicians and physicists

......

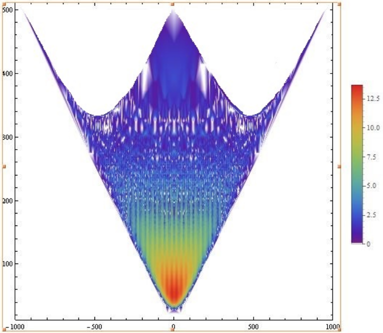
The Compact CY3 Landscape

cf. YHH, *The Calabi-Yau Landscape: from Geometry, to Physics, to Machine-Learning,* 1812.02893, [Springer, to appear]

Vienna (KS, Knapp,...), Penn (Ovrut, Cvetic, Donagi, Pantev ...), Oxford/London (Candelas, Constantin, Lukas, Mishra, YHH, ...), MIT (Taylor, Johnson, Wang, ...), Northeastern/Wits (Halverson, Long, Nelson, Jejjala, YHH), Virginia Tech (Anderson, Gray, SJ Lee, ...), Utrecht (Grimm ...), CERN (Weigand, ...), Cornell (MacAllister, Stillman), Munich (Luest, Vaudravange), Uppsala (Larfors, Seong) ...

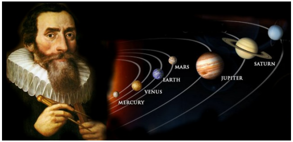


Georgia O'Keefe on Kreuzer-Skarke



Horizontal $\chi = 2(h^{1,1} - h^{2,1})$ vs. Vertical $h^{1,1} + h^{2,1}$

The Geometric Origin of our Universe

- Each CY3 (+ bundles, discrete symmetries) X gives a 4-D universe
  - The geometry (algebraic geometry, topology, differential geometry etc.) of X determines the physical properties of the 4-D world
  - particles and interactions $\sim$ cohomology theory; masses $\sim$ metric; Yukawa $\sim$ Triple intersections/integral of forms over X
- 

  Ubi materia, ibi geometria

  – Johannes Kepler (1571-1630)

- Our Universe: (1) probabilistic/anthropic?
  (2) Sui generis/selection rule? Triadophilia
  (3) one of multi-verse ?

cf. *Exo-planet/Habitable Zone search*

The Landscape Explosion & Vacuum Degeneracy Problem

meanwhile . . . LANDSCAPE grew rapidly with

- D-branes Polchinski 1995
- M-Theory/G_2 Witten, 1995
- F-Theory/4-folds Katz-Morrison-Vafa, 1996
- AdS/CFT Maldacena 1998
- Flux-comp. Kachru-Kallosh-Linde-Trivedi, 2003, Denef-Douglas 2005-6: $10^{\gg 500}$ possibilities . . .

- String theory trades one hard-problem [quantization of gravity] by another [looking for the right compactification] (in many ways a richer and more interesting problem, especially for the string/maths community)
- Is String Theory ToE? (don't know), Is String Theory THE ONE branch of fundamental science constantly generating new dialogues between physics, maths, CS, data-science etc. ? (YES!)

Where we stand ...

The Good: Last 10-15 years: large collaborations of physicists, computational mathematicians (cf. SageMATH, GAP, Bertini, MAGMA, Macaulay2, Singular) have bitten the bullet computed many geometrical/physical quantities and compiled them into various databases **Landscape Data** ($10^{9\sim10}$ entries typically)

The Bad: Generic computation HARD: dual cone algorithm (exponential), triangulation (exponential), Gröbner basis (double-exponential) ...e.g., how to construct stable bundles over the $\gg 473$ million KS CY3? Sifting through for SM computationally impossible ...

The ???: **Borrow new techniques from "Big Data" revolution**

A Wild Question

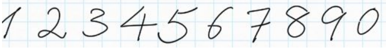
- Typical Problem in String Theory/Algebraic Geometry:

$$\begin{array}{ccc} INPUT & & OUTPUT \\ \boxed{\text{integer tensor}} & \longrightarrow & \boxed{\text{integer}} \end{array}$$

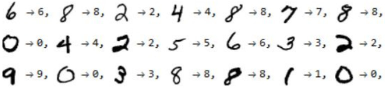
- Q: Can (classes of problems in computational) Algebraic Geometry be "learned" by AI ? , i.e., can we
  - machine-learn the string landscape?
  - ML algebraic geometry?
- [YHH 1706.02714] Deep-Learning the Landscape, *PLB 774, 2017* (*Science*, Aug, vol 365 issue 6452): Experimentally, it seems so for many situations in geometry and beyond.

A Prototypical Question

- Hand-writing Recognition, e.g., my 0 to 9 is different from yours:



- How to set up a bijection that takes these to $\{1, 2, \ldots, 9, 0\}$? Find a clever Morse function? Compute persistent homology? Find topological invariants? ALL are inefficient and too sensitive to variation.
- What does your iPhone/tablet do? What does Google do? Machine-Learn
  - Take large sample, take a few hundred thousand (e.g. NIST database)



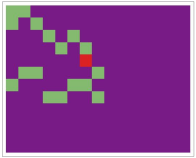
...

$28 \times 28 \times (RGB)$

NN Doesn't Care/Know about Alg. Geo. (YHH 1706.02714)

- Hodge Number of a Complete Intersection CY is the association rule, e.g.

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}, \qquad h^{1,1}(X) = 8 \quad \rightsquigarrow$$



$$\longrightarrow 8$$

CICY is 12×15 integer matrix with entries $\in [0, 5]$ is simply represented as a 12×15 pixel image of 6 colours Proper Way ; ML in matter of seconds/minutes

- Cross-Validation:
  - Take samples of $X \to h^{1,1}$
  - train a NN, or SVM
  - Validation on *unseen* $X \to h^{1,1}$

Deep-Learning String Geometry: Early Experiments

2017 String Landscape YHH (1706.02714); Calabi-Yau Volume Seong-Krefl (1706.03346); Line-Bundles Ruehle (1706.07024); Vacuum Selection Carifio-Halverson-Krioukov-Nelson (1707.00655)

Hodge Numbers YHH (1706.02714) Bull-YHH-Jejjala-Mishra (1806.03121, 1903.03113), Krippendorf-Syvaeri [2003.13679] Erbin-Finotello (2007.13379); YHH-Lukas [2009.02544] (~ 0.99)

Distinguishing Elliptic Fibrations YHH-SJ Lee (1904.08530) (~ 0.99)

ML now a standard programme in string community since 2017

Progress in String Theory

q.v., Review YHH 2011.14442 "Universes as Big Data" (IJMPA)

Success Stories in String/Algebraic Geometry 2017-

- q.v., Bundle Cohomology Ruehle, Brodie-Constantin-Lukas, Larfors-Schneider, Otsuka-Takemoto, Klaewer-Schlechter
- q.v., Kreuzer-Skarke Dataset Halverson, Long, Nelson; McCallister-Stillman
- Calabi-Yau metric: improves Donaldson alg. for numerical CY metric by 10-100 times Ashmore–YHH–Ovrut '19, q.v. Anderson, Gray, Krippendorf, Raghuram, Ruehle; Douglas–Lakshminarasimhan–Qi, '20
- q.v., MSSM from orbifold models Parr-Vaudrevange-Wimmer; q.v. Particle Masses Gal-Jejjala-Pena-Mishra . . .
- TDA on cosmology and phrase transitions Cole-Shiu
- q.v. Knot invariants: Jejjala-Kar-Parrikar, Craven-Jejjala-Kar Gukov-Halverson-Ruehle-Sułkowski, using NLP
- q.v. DEEP CONNECTIONS NN = everything

Machine Learning Mathematical Structures

Why stop at string/geometry?

q.v. Review Paper: YHH 2101.06317

How does one *DO* mathematics, I ?

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

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

- Automated Theorem Proving (ATP) The practicing mathematician hardly ever worries about Gödel
  - Newell-Simon-Shaw [1956] Logical Theory Machine:

    proved subset of *Principia* theorems
  - Type Theory [1970s] Martin-Löf, Coquand, . . . Coq interactive proving system: 4-color (2005); Feit-Thompson Thm (2012); Lean (2013)
  - Univalent Foundation / Homotopy Type Theory [2006-] Voevodsky

We can call this Bottom-up Mathematics

How does one *DO* mathematics, I ?

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

  Gödel [1931] Incompleteness ; Church-Turing [1930s] Undecidability
- Automated Theorem Proving (ATP) The practicing mathematician hardly ever worries about Gödel
  - Newell-Simon-Shaw [1956] Logical Theory Machine: proved subset of *Principia* theorems
  - Type Theory [1970s] Martin-Löf, Coquand, ... Coq interactive proving system: 4-color (2005); Feit-Thompson Thm (2012); Lean (2013)
  - Univalent Foundation / Homotopy Type Theory [2006-] Voevodsky

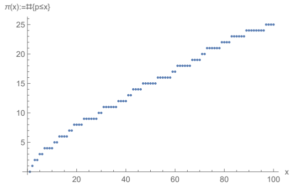
We can call this Bottom-up Mathematics

How does one do mathematics, II ?

- Late C20th - increasing rôle of computers: 4-color [Appel-Haken-Koch 1976]; Classif. Finite Simple Groups [Galois 1832 - Gorenstein et al. 2008] ...
- Buzzard: "Future of Maths" 2019: already plenty of proofs unchecked (incorrect?) in the literature, MUST use computers for proof-checking; XenaProject, Lean establish database of mathematical statements
- Davenport: ICM 2018 "Computer Assisted Proofs".
- Hale & Buzzard: Foresee within 10 years AI will help prove "early PhD" level lemmas, all of undergrad-level maths formalized;
- Szegedy: more extreme view, computers $>$ humans @ chess (1990s); @ Go (2018); @ Proving theorems (2030)

How does one *DO* mathematics, III ?

- Historically, Maths perhaps more Top-Down: practice before foundation
  - Countless examples: calculus before analysis; algebraic geometry before Bourbaki, permutation groups / Galois theory before abstract algebra . . .
  - A lot of mathematics starts with intuition, experience, and experimentation
- The best neural network of C18-19th? brain of Gauß ; e.g., age 16

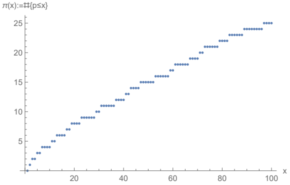


(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

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

How does one *DO* mathematics, III ?

- Historically, Maths perhaps more Top-Down: practice before foundation
  - Countless examples: calculus before analysis; algebraic geometry before Bourbaki, permutation groups / Galois theory before abstract algebra ...
  - A lot of mathematics starts with intuition, experience, and experimentation
- The best neural network of C18-19th? brain of Gauß ; e.g., age 16

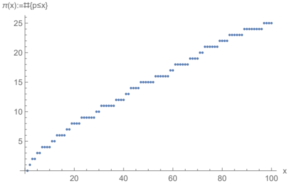


(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

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

How does one *DO* mathematics, III ?

- Historically, Maths perhaps more Top-Down: practice before foundation
  - Countless examples: calculus before analysis; algebraic geometry before Bourbaki, permutation groups / Galois theory before abstract algebra ...
  - A lot of mathematics starts with intuition, experience, and experimentation
- 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

Question

- To extend the analogy: AlphaGo is top-down (need to see human games); even AlphaZero is not bottom-up (need to generate samples of games)
- In tandem with the bottom-up approach of Coq, Lean, Xena ... how to put in a little intuition and human results? If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Perfect for (unsupervised & supervised) machine-learning; focus on labeled case because it encodes WHAT is interesting to calculate (if not how).

Question

- To extend the analogy: AlphaGo is top-down (need to see human games); even AlphaZero is not bottom-up (need to generate samples of games)
- In tandem with the bottom-up approach of Coq, Lean, Xena ... how to put in a little intuition and human results? If I gave you 100,000 cases of

$$\text{e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g. } \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Perfect for (unsupervised & supervised) machine-learning; focus on labeled case because it encodes WHAT is interesting to calculate (if not how).

Question

- To extend the analogy: AlphaGo is top-down (need to see human games); even AlphaZero is not bottom-up (need to generate samples of games)
- In tandem with the bottom-up approach of Coq, Lean, Xena ... how to put in a little intuition and human results? If I gave you 100,000 cases of

$$\text{e.g.} \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g.} \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Perfect for (unsupervised & supervised) machine-learning; focus on labeled case because it encodes WHAT is interesting to calculate (if not how).

Question

- To extend the analogy: AlphaGo is top-down (need to see human games); even AlphaZero is not bottom-up (need to generate samples of games)
- In tandem with the bottom-up approach of Coq, Lean, Xena ... how to put in a little intuition and human results? If I gave you 100,000 cases of

$$\text{e.g.} \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix}, \quad \text{or, labeled data e.g.} \begin{pmatrix} 5&3&4&3&5&1&4&4&1&2\\ 5&0&4&5&2&4&4&2&2&4\\ 1&1&2&2&0&4&1&4&5&0\\ 5&0&1&1&0&2&0&5&0&1\\ 2&5&0&1&1&3&2&3&0&3\\ 3&2&2&3&0&0&2&2&1&0\\ 2&2&5&1&4&4&0&0&1&2\\ 5&0&0&0&4&5&0&4&1&1\\ 4&3&4&3&3&1&0&0&2&5\\ 2&0&5&0&3&0&4&4&1&5 \end{pmatrix} \longrightarrow 3$$

- Q: Is there a pattern? Can one conjecture & then prove a formula?
- Q: What branch of mathematics does it come from?
- Perfect for (unsupervised & supervised) machine-learning; focus on labeled case because it encodes WHAT is interesting to calculate (if not how).

Mathematical Data: perfect for mining

- Mathematical Data is more structured than "real world" data, much less susceptible to noise; Outliers even more interesting, e.g. Sporadics, Exceptionals, . . .
- Last 10-20 years: large collaborations of computational mathematicians, physicists, CS (cf. SageMATH, GAP, Bertini, MAGMA, Macaulay2, Singular, Pari, Wolfram, . . .) computed and compiled vast data
  - links
  - Generic computation *HARD*
  - mining provides some level of "intuition" & is based on "experience"

Methodology

Bag of Tricks Hilbert's Programme of *Finitary Methods*, Landau's *theoretical minimum*, Migdal's *Mathmagics* . . .

IMO Grand Challenge (2020-) Good set of concrete problems to try on AI

Standard Supervised ML Methods Regressor & Classifiers

- NN: MLPs; CNNs; RNNs, . . . (gentle tuning of architecture and hyper-parameters)
- SVM, Bayes, Decision Trees, PCA, Clustering, . . .
- ML: emergence of complexity via connectivity $\rightsquigarrow$ Intution (?)

This Talk: Status Report of Experiments in the last couple of years

- all standard methods $\simeq$ same performance
- $\sim$ 20-80 split; training on 20 (precision, Matthews' ϕ or R^2)

Representation/Group Theory

- ML Algebraic Structures (GAP DB) [YHH-MH. Kim 1905.02263, YHH-Jejjala-Mishra-Sharnoff, to appear]
  - When is a Latin Square (Sudoku) the Cayley (multiplication) table of a finite group? Bypass quadrangle thm (0.95, 0.9)
  - Can one look at the Cayley table and recognize a finite simple group?
    - bypass Sylow and Noether Thm; (0.97, 0.95) rmk: can do it via character-table T, but getting T not trivial
    - **SVM:** space of finite-groups (point-cloud of Cayley tables) seems to exist a hypersurface separating simple/non-simple
- ML Lie Structure Chen-YHH-Lal-Majumder [2011.00871] Weight vector $\rightarrow$ length of irrep decomp / tensor product: (0.97, 0.93); (train on small dim, predict high dim: (0.9, 0.8))
- [Chen-YHH-Lal-Zas 2006.16114]: even/odd/reflection sym (>0.99); distinguishing CFT 3pt functions (>0.99); Fourier coefficients / conformal block presence (>0.97) . . .

Combinatorics, Graph/Quivers

- [YHH-ST. Yau 2006.16619] Wolfram Finite simple graphs DB
  - ML standard graph properties:

    ?acyclic (0.95, 0.96); ?planar (0.8, 0.6); ?genus $>, =, < 0$ (0.8, 0.7); ?∃ Hamilton cycles (0.8, 0.6); ?∃ Euler cycles (0.8, 0.6)

    (Rmk: NB. Only "solving" the likes of traveling salesman stochastically)
  - spectral bounds ($R^2 \sim 0.9$) ...
  - Recognition of Ricci-Flatness (0.9, 0.9) (todo: find new Ricci-flat graphs);
- [Bao-Franco-YHH-Hirst-Musiker-Xiao 2006.10783]: categorizing different quiver mutation (Seiberg-dual) classes (0.9 - 1.0, 0.9)

Number Theory: A Classical Reprobate?

Arithmetic (prime numbers are Difficult!)

- [YHH 1706.02714, 1812.02893:]
  - Predicting primes $2 \to 3,\ 2,3 \to 5,\ 2,3,5 \to 7$; no way
  - fixed window of $(\text{yes/no})_{1,2,\ldots,k}$ to $(\text{yes/no})_{k+i}$ for some i; ML PRIMES problem (0.7, 0.8) NOT random! (prehaps related to AKS algorithm [2002], PRIMES is in P)
  - Sarnak's challenge: same window $\to$ Liouville Lambda (0.5, 0.001) Truly random (no simple algorithm for Lambda)
- [Alessandretti-Baronchelli-YHH 1911.02008] (*New Scientist*, feature, Dec 2019) ML/TDA@Birch-Swinnerton-Dyer Ш and Ω ok with regression & decision trees: RMS < 0.1; Weierstrass $\to$ rank: random

Number Theory: A Modern Hope?

Arithmetic Geometry (Surprisingly Good)

- [Hirst-YHH-Peterken 2004.05218]: adjacency+permutation triple of dessin d'enfants (Grothendieck's Esquisse for $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$) ; predicting transcendental degree (0.92, 0.9)
- YHH-KH Lee-Oliver arithmetic of curves
  - 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]

Clearly useful for maths and physics

- looking for new conjectures e.g.,
  - '19 YHH-Kim: separating hyperplane - simple/non-simple groups; open
  - '19 Brodie-Constantin-Lukas: exact formulae for cohomo surf.; proved.
  - '20 YHH-Lee-Oliver: L-coefs and integer pt./torsion on ell; proved.
  - '20 Craven-Jejjala-Par: Jones poly best-fit function; open
  - ...
- speed up computations and accuracies e.g.,
  - computing/estimating (top.inv., charges, etc) MUCH FASTER
  - '19 Ashmore-YHH-Ovrut: speed up Donaldson alg@CY metric 10-100
  - '20 Douglas et al., Anderson et al. accuracy improvement on Donaldson 10-100 times
  - ...

The other Extreme (?) View-Point

On the other hand, what is analyticity?

- n-th pime $= \left\lfloor \frac{n! \bmod (n+1)}{n} \right\rfloor (n-1) + 2$ (not efficient)
- bundle-cohomology:

Bott for Projective space:

$$h^q(\mathbb{P}^n, (\wedge^p T\mathbb{P}^n) \otimes \mathcal{O}(k)) = \begin{cases} \binom{k+n+p+1}{p}\binom{k+n}{n-p} & q=0 \quad k > -p-1, \\ 1 & q = n-p \quad k = -n-1, \\ \binom{-k-p-1}{-k-n-1}\binom{-k-n-2}{p} & q=n \quad k < -n-p-1, \\ 0 & \text{otherwise} \end{cases}$$

e.g. $(2,4)$-CY3 hypersurface:

$$h^q(X, \mathcal{O}_X(-k,m)) = \begin{cases} (k+1)\binom{m}{3} - (k-1)\binom{m+3}{3} & q=0 \quad k < \frac{(1+2m)(6+m+m^2)}{3(2+3m(1-m))} \\ (k-1)\binom{m+3}{3} - (k+1)\binom{m}{3} & q=1 \quad k > \frac{(1+2m)(6+m+m^2)}{3(2+3m(1-m))} \\ 0 & \text{otherwise} \end{cases}$$

- ...
- better suited for a computer programme any way

An Inherent Hierarchy?

- In decreasing precision/increasing difficulty:

$$\begin{array}{lc} & \text{numerical} \\ \text{string theory} \rightarrow & \text{algebraic geometry over } \mathbb{C} \sim \text{arithmetic geometry} \\ & \text{algebra} \\ \text{string theory} \rightarrow & \text{combinatorics} \\ & \text{analytic number theory} \end{array}$$

- **Categorical Theory**
  - suggested by & in prog. w/ B. Zilber, Merton Prof. of Logic, Ox
  - major part of Model Theory: Morley-Shelah Categoricity Thm
  - Hart-Hrushovski-Laskowski Thm: 13 classes (levels) of iso-classes $I(T, k)$ of a theory T in first order logic over some cardinality k.

Semantics vs Syntax

Alpha Go	$\to$	Alpha Zero
ML	$\to$	Voevodsky's Dream; Automated Thm Pf

  - Renner et al., PRL/Nature News, 2019: ML (*SciNet*, *autoencoder*)
  - Lample-Charton, 2019: ML Symolic manipulations in mathematics
  - Tegmark et al., 2019 AI Feynman, symb regressor
  - Raayoni et al. 2020 Ramanujan-Machine

Sophia (Hanson Robotics, HK)

1st non-human citizen (2017, Saudi)

1st non-human with UN title (2017)

1st String Data Conference (2017)

THANK YOU

- Paolo Di Vecchia: *String theory is a piece of 21st century physics that happened to fall into the 20th century ...*
- Edward Witten: *piece of 21st century mathematics that happened to begin in the 20th century ...*
- Alexander Kaspryzyk; Kevin Buzzard: *Letting AI/ML do mathematics could well be the standard for the 22nd century ...*

THANK YOU

- Paolo Di Vecchia: *String theory is a piece of 21st century physics that happened to fall into the 20th century . . .*
- Edward Witten: *piece of 21st century mathematics that happened to begin in the 20th century . . .*
- Alexander Kaspryzyk; Kevin Buzzard: *Letting AI/ML do mathematics could well be the standard for the 22nd century . . .*

THANK YOU

- Paolo Di Vecchia: *String theory is a piece of 21st century physics that happened to fall into the 20th century ...*
- Edward Witten: *piece of 21st century mathematics that happened to begin in the 20th century ...*
- Alexander Kaspryzyk; Kevin Buzzard: *Letting AI/ML do mathematics could well be the standard for the 22nd century ...*

Digressions

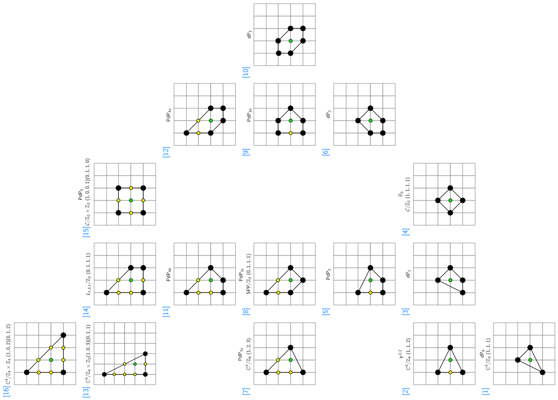
$\chi(\Sigma) = 2$	$\chi(\Sigma) = 0$	$\chi(\Sigma) < 0$
Spherical	Ricci-Flat	Hyperbolic
$+$ curvature	0 curvature	$-$ curvature
Fano	Calabi-Yau	General Type
- Euler, Gauss, Riemann, Bourbaki, Atiyah-Singer ... $\rightsquigarrow$ generalize

$$\chi(\Sigma) = 2 - 2g(\Sigma) = [c_1(\Sigma)] \cdot [\Sigma] = \frac{1}{2\pi} \int_\Sigma R = \sum_{i=0}^{2} (-1)^i h^i(\Sigma)$$

- CONJECTURE [E. Calabi, 1954, 1957] / Thm [ST. Yau, 1977-8] 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 $Ricci(\tilde{\omega}) = R$.
- Strominger & Yau were neighbours at IAS in 1985: CHSW named Ricci-Flat Kähler as Calabi-Yau Back

16 Reflexive Polygons

Back to Reflexives



classify convex lattice polytopes with single interior point and all faces are distance 1 therefrom (up to $SL(n;\mathbb{Z})$)

Kreuzer-Skarke: 4319 reflexive polyhedra, 473,800,776 reflexive 4-polytopes,

Skarke: next number is at least 185,269,499,015.

Heterotic Comp: Recent Development

- E_6 GUTs a toy, $SU(5)$ and $SO(10)$ GUTs and SM: general embedding
  - Instead of TX, use (poly-)stable holomorphic vector bundle V
  - Gauge group$(V) = G = SU(n)$, $n = 3, 4, 5$, gives $H =$ Commutant(G, E_8):

$E_8 \to G \times H$	Breaking Pattern
$SU(3) \times E_6$	$248 \to (1,78) \oplus (3,27) \oplus (\overline{3},\overline{27}) \oplus (8,1)$
$SU(4) \times SO(10)$	$248 \to (1,45) \oplus (4,16) \oplus (\overline{4},\overline{16}) \oplus (6,10) \oplus (15,1)$
$SU(5) \times SU(5)$	$248 \to (1,24) \oplus (5,\overline{10}) \oplus (\overline{5},10) \oplus (10,5) \oplus (\overline{10},\overline{5}) \oplus (24,1)$

- MSSM: $H \xrightarrow{\text{Wilson Line}} SU(3) \times SU(2) \times U(1)$
- Issues in low-energy physics $\sim$ Precise questions in Alg Geo of (X, V)
  - Particle Content $\sim$ (tensor powers) V Bundle Cohomology on X
  - LE SUSY $\sim$ Hermitian Yang-Mills connection $\sim$ Bundle Stability
  - Yukawa $\sim$ Trilinear (Yoneda) composition
  - Doublet-Triplet splitting $\sim$ representation of fundamental group of X

Back to Landscape

Various Databases

- Kreuzer-Skarke: http://hep.itp.tuwien.ac.at/~kreuzer/CY/
  - new PALP: Braun-Walliser: ArXiv 1106.4529
  - Triang: Altmann-YHH-Jejjala-Nelson: http://www.rossealtman.com/
- CICYs: resurrected Anderson-Gray-YHH-Lukas, http://www-thphys.physics.ox.ac.uk/projects/CalabiYau/cicylist/index.html
- q.v. other databases of interesting to the math/physics community:

  Graded Rings/Varieties: Brown, Kasprzyk, et al. http://www.grdb.co.uk/

  Finite Groups/Rings: GAP https://www.gap-system.org/

  Modular Forms: Sutherland, Cremona et al. https://www.lmfdb.org/

  Knots & Invariants: KnotAtlas http://katlas.org/ Return

  ...

Progress in String Theory

Back to ML/Maths

Major International Annual Conference Series

1986- First "Strings" Conference

2002- First "StringPheno" Conference

2006 - 2010 String Vacuum Project (NSF)

2011- First "String-Math" Conference

2014- First String/Theoretical Physics Session in SIAM Conference

2017- First "String-Data" Conference

A Single Neuron: The Perceptron

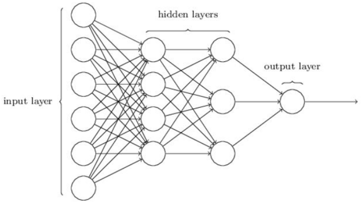
- began in 1957 (!!) in early AI experiments (using CdS photo-cells)
- DEF: Imitates a neuron: activates upon certain inputs, so define
  - Activation Function $f(z_i)$ for input tensor z_i for some multi-index i;
  - consider: $f(w_i z_i + b)$ with w_i weights and b bias/off-set;
  - typically, $f(z)$ is sigmoid, Tanh, etc.
- Given training data: $D = \{(x_i^{(j)}, d^{(j)}\}$ with input x_i and **known output** $d^{(j)}$, minimize

$$SD = \sum_j \left(f(\sum_i w_i x_i^{(j)} + b) - d^{(j)} \right)^2$$

  to find optimal w_i and $b \rightsquigarrow$ "learning", then check against Validation Data
- Essentially (non-linear) regression

The Neural Network: network of neurons $\rightsquigarrow$ the "brain"

- DEF: a connected graph, each node is a perceptron (*Implemented on Mathematica* $\geq$ *11.1 / TensorFlow-Keras on Python*)
  1. adjustable weights/bias;
  2. distinguished nodes: 1 set for input and 1 for output;
  3. iterated training rounds.



Simple case: forward directed only,

called multilayer perceptron

Many Layers : DEEP Learning

$\boxed{\text{Connectivity} \rightsquigarrow \text{Emergence of Complexity}}$

- Essentially how brain learns complex tasks; **apply to our Landscape Data**

Back to Landscape

CICYs

$$M = \left[\begin{array}{c|cccc} n_1 & q_1^1 & q_1^2 & \dots & q_1^K \\ n_2 & q_2^1 & q_2^2 & \dots & q_2^K \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n_m & q_m^1 & q_m^2 & \dots & q_m^K \end{array} \right]_{m \times K}$$

– Complete Intersection Calabi-Yau (CICY) 3-folds

– K eqns of multi-degree $q_j^i \in \mathbb{Z}_{\geq 0}$

embedded in $\mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_m}$

– $c_1(X) = 0 \rightsquigarrow \sum_{j=1}^{K} q_r^j = n_r + 1$

– M^T also CICY

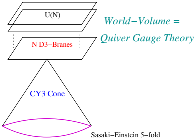
- The Quintic $Q = [4|5]^{1,101}_{-200}$ (or simply $[5]$);
- CICYs Central to string pheno in the 1st decade [Distler, Greene, Ross, et al.]

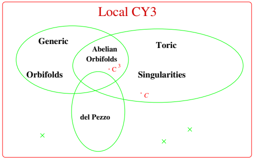
  E_6 GUTS unfavoured; Many exotics: e.g. 6 entire anti-generations

Back to CICYs

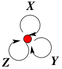
AdS/CFT as a Quiver Rep/Moduli Variety Corr.

a 20-year prog. joint with **A. Hanany**, S. Franco, B. Feng, et al.





D-Brane Gauge Theory
(SCFT encoded as quiver)
$\longleftrightarrow$
Vacuum Space as affine Variety

- ($\mathcal{N} = 4$ SYM) ( , $W = \text{Tr}([x, y], z)$) $\longleftrightarrow \mathbb{C}^3 = \text{Cone}(S^5)$ [Maldacena]

- THM [(P) Feng, Franco, Hanany, YHH, Kennaway, Martelli, Mekareeya, Seong, Sparks, Vafa, Vegh, Yamazaki, Zaffaroni . . . (M) R. Böckland, N. Broomhead, A. Craw, A. King, G. Musiker, K. Ueda . . .] (coherent component of) representation variety of a quiver is toric CY3 iff quiver + superpotential graph dual to a bipartite graph on T^2 Back to Landscape

combinatorial data/lattice polytopes $\longleftrightarrow$ gauge thy data as quivers/graphs

Computing Hodge Numbers $\mathcal{O}(e^{e^d})$

- 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:

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

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

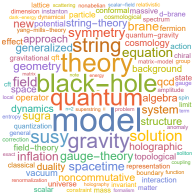
Back to ML

Deep Connections

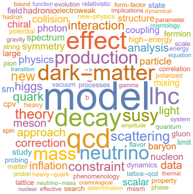
- K. Hashimoto 2019-: AdS/CFT = Boltzmann Machine;
- Halverson-Maiti-Stoner 2020: QFT = NN;
- de Mello-Koch 2020: NN = RG;
- Vanchurin 2008: Universe = NN.

Back to ML Maths

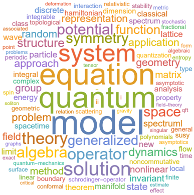
ArXiv Word-Clouds



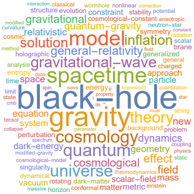
hep-th



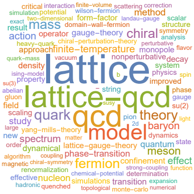
hep-ph



math-ph



hep-lat



hep-lat

[Back to Word2Vec]

Classifying Titles

Compare, + non-physics sections, non-science (Times), pseudo-science (viXra)

Actual \ Word2Vec + SVM	1	2	3	4	5
1	40.2	6.5	8.7	24.0	20.6
2	7.8	65.8	12.9	9.1	4.4
3	7.5	11.3	72.4	1.5	7.4
4	12.4	4.4	1.0	72.1	10.2
5	10.9	2.2	4.0	7.8	75.1

1 : hep-th
2 : hep-ph
3 : hep-lat
4 : gr-qc
5 : math-ph

Actual \ NN	1	2	3	4	5	6	7	8	9	10
viXra-hep	11.5	47.4	6.8	13.	11.	4.5	0.2	0.3	2.2	3.1
viXra-qgst	13.3	14.5	1.5	54.	8.4	1.8	0.1	1.1	2.8	3.

6: cond-mat, 7: q-fin, 8: stat, 9: q-bio, 10: Times of India Back to Main

Explicit Examples of Calabi-Yau Spaces

An interesting sequence: 1,2, ??? ...

$\dim_{\mathbb{C}} = 1$

Torus $T^2 = S^1 \times S^1$

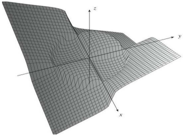
QFT in $10 - 2 = 8d$

$\dim_{\mathbb{C}} = 2$

(1) 4-Torus $T^4 = S^1 \times S^1 \times S^1 \times S^1$

(2) K3 surface

QFT in $10 - 4 = 6d$

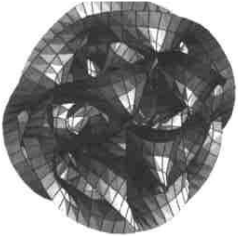


$\dim_{\mathbb{C}} = 3$

Unclassified ???

(Yau's Conjecture: Finite Number)

Desired QFT in $10 - 6 = 4d$



CalabiYau the Game

游戏介绍 （from https://www.9k9k.com/shouyou/klbq/）
《卡拉比丘》是一款宏大世界观的动作游戏。采用5V5的战斗模式，玩家需要选择自己的阵营，操控英雄探索地图，与队友密切协作，战胜敌人玩家即可获得比赛的胜利，更有上百位美少女英雄等待你的召唤！

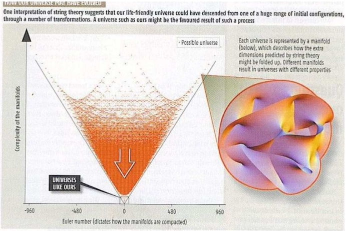
Back to CY

Triadophilia

Exact (MS)SM Particle Content from String Compactification

- [Braun-YHH-Ovrut-Pantev, Bouchard-Cvetic-Donagi 2005] first exact MSSM
- [Anderson-Gray-YHH-Lukas, 2007-] use alg./comp. algebraic geo & sift
- Anderson-Gray-Lukas-Ovrut-Palti ~ 200 in 10^{10} MSSM Stable Sum of Line Bundles over CICYs (Oxford-Penn-Virginia 2012-)

  Constantin-YHH-Lukas '19: 10^{23} exact MSSMs (by extrapolation on above set)?



A Special Corner [New Scientist, 5/1/2008 feature]

Candelas-de la Ossa-YHH-Szendroi

"Triadophilia: A Special Corner of the Landscape" ATMP, 2008

Back to String Landscape