Machine-Learning the Landscape

YANG-HUI HE

London Institute of Mathematical Sciences, Royal Institution of Great Britain

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

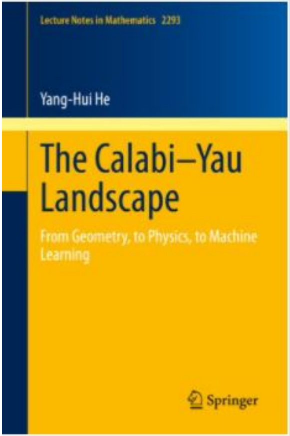
Merton College, **University of Oxford**

School of Physics, **NanKai University**

Kavli Asian Winter School on Strings, Particles and Cosmology

Jan 2022

Further Reading



YHH, *The CY Landscape: from Geometry to Physics, to ML, 2021, Springer LNM 2293* (the 224pp version)

YHH, "Universes as Big Data" 2011.14442 (the 20pp version)

PLAN

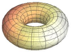
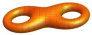
Part I Introducing CY manifolds as a microcosm of the string/mathematics landscape

Part II Machine-Learning 101

Part III Having fun

A Classic Problem in Mathematics

- 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

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

Euler number $\chi(\Sigma)$, genus $g(\Sigma)$

Classical Results for Riemann Surface Σ

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

- First Chern Class $c_1(\Sigma)$
- Rank of (co-)homology group (Betti Number) $h^i(\Sigma)$
- Complexifies (Künneth) $h^i = \sum\limits_{j+k=i} h^{j,k}$, Hodge Number $h^{j,k}$

Calabi-Yau

- $\dim_{\mathbb{C}} > 1$ extremely complicated (high-dim geometry hard: cf. Poincaré Conjecture/Perelman Thm/Thurston-Hamilton Prog)
- Luckily, for our class of **Kähler** complex manifolds: Recall Defs
- 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 in GR long before strings)

- THEOREM [S-T Yau, 1977-8; Fields 1982] Existence Proof
- Calabi-Yau: Kähler and Ricci-flat

Calabi-Yau

- $\dim_{\mathbb{C}} > 1$ extremely complicated (high-dim geometry hard: cf. Poincaré Conjecture/Perelman Thm/Thurston-Hamilton Prog)
- Luckily, for our class of **Kähler** complex manifolds: Recall Defs
- 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 in GR long before strings)

- THEOREM [S-T Yau, 1977-8; Fields 1982] Existence Proof
- Calabi-Yau: Kähler and Ricci-flat

Calabi-Yau

- $\dim_{\mathbb{C}} > 1$ extremely complicated (high-dim geometry hard: cf. Poincaré Conjecture/Perelman Thm/Thurston-Hamilton Prog)
- Luckily, for our class of **Kähler** complex manifolds: Recall Defs
- 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 in GR long before strings)

- THEOREM [S-T Yau, 1977-8; Fields 1982] Existence Proof
- Calabi-Yau: Kähler and Ricci-flat

String Phenomenology

- Superstring: unifies QM + GR in 10 dimensions: X^{10}
- We live in some M^4 (assume maximally symmetric)

$$R_{\mu\nu\rho\lambda} = \frac{R}{12}(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\lambda}g_{\nu\rho}), \qquad R \begin{cases} = 0 & \text{Minkowski} \\ > 0 & \text{de Sitter (dS)} \\ < 0 & \text{anti-de Sitter (AdS)} \end{cases}$$

- $10 = 4 + 6$: **two scenarios**
  1. SMALL: compactification $X^{10} \simeq M^4 \times X^6$
  2. LARGE: brane-world trapped on a 3-brane in 10-D
- supersymmetry at intermediate scale (between string and EW)
- want: classical vacuum of string theory on X^{10} preserves $\mathcal{N} = 1$ SUSY in M^4

Heterotic Compactification

[Candelas-Horowitz-Strominger-Witten] (1986): $\delta_{SUSY} S_{Het} = 0$

- $S \sim \int d^{10}x \sqrt{g} e^{-2\Phi} \left[R + 4\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2}|H_3'|^2) - \frac{1}{g_s^2} \mathrm{Tr}\, |F_2|^2 \right]$+SUSY)

gravitino	$\delta_\epsilon \Psi_{M=1,\ldots,10} = \nabla_M \epsilon - \frac{1}{4} H_M^{(3)} \epsilon$
dilatino	$\delta_\epsilon \lambda = -\frac{1}{2}\Gamma^M \partial_M \Phi\ \epsilon + \frac{1}{4} H_M^{(3)} \epsilon$
adjoint YM	$\delta_\epsilon \chi = -\frac{1}{2} F^{(2)} \epsilon$
Bianchi	$dH^{(3)} = \frac{\alpha'}{4}[\mathrm{Tr}(R \wedge R) - \mathrm{Tr}(F \wedge F)]$

- Assume $H^{(3)} = 0$ (can generalise) $\rightsquigarrow$ Killing spinor equation:

  $\delta_\epsilon \Psi_{M=1,\ldots,10} = \nabla_M \epsilon = 0 = \nabla_M \xi(x^{\mu=1,\ldots,4}) \eta(y^{m=1,\ldots,6})$

- External 4D Space: $[\nabla_\mu, \nabla_\nu]\xi(x) = \frac{1}{4} R_{\mu\nu\rho\sigma} \Gamma^{\rho\sigma} \xi(x) = 0 \rightsquigarrow R = 0 \Rightarrow M$ is Minkowski (of course, should be looking for dS, but to 1st order)
- Internal 6D Space: $R_{mn} = 0$ (but not necessarily max symmetric)

Mille Viæ ducunt homines Romam ...

- X^6 as a spin 6-manifold: holonomy group is $SO(6) \simeq SU(4)$
  - want covariant constant spinor: largest possible is $SU(4) \to SU(3)$ with $4 \to 3 \oplus 1 \Rightarrow X^6$ has $SU(3)$ holonomy
  - Thus $\epsilon(x^{1,\ldots,4}, y^{1,\ldots,6}) = \xi_+ \otimes \eta_+(y) + \xi_- \otimes \eta_-(y)$ with $\eta_+^* = \eta_-$ and ξ constant
- Define $J_m^n = i\eta_+^\dagger \gamma_m^n \eta_+ = -i\eta_-^\dagger \gamma_m^n \eta_-$, check: $J_m^n J_n^p = -\delta_m^n$
- Can show $\boxed{X^6 \text{ is a Kähler manifold of } \dim_{\mathbb{C}} = 3, \text{ with } SU(3) \text{ holonomy}}$
- Three other SUSY variation equations (recall $H^{(3)} = 0$ by choice)
  - choose constant dilation $\Phi \rightsquigarrow \delta_\epsilon = 0$
  - choose $R = F$ (spin connection for gauge field): Bianchi satisfied
  - Also $R = 0$ so $\delta_\epsilon \chi = 0$

Special Holonomy

- For a Riemannian, spin manifold M of real dimension d, holonomy is $Spin(d)$ as double cover of $SO(d)$ *generically*, but could have *special holonomy*

Holonomy $\mathcal{H} \subset$	Manifold Type (IFF)
$U(d/2)$	Kähler
$SU(d/2)$	Calabi-Yau
$Sp(d/4)$	Hyper-Kähler
$Sp(d/4) \times Sp(1)$	Quaternionic-Kähler

- X^6 **is Calabi-Yau**
- no-where vanishing holomorphic 3-form: $\Omega^{(3,0)} = \frac{1}{3!}\Omega_{mnp} dz^m \wedge dz^n \wedge dz^p$ with $\Omega_{mnp} := \eta_-^T \, \gamma^{[m}\gamma^n\gamma^{p]} \, \eta_-$ check: $d\Omega = 0$ but not exact; $\Omega \wedge \bar{\Omega} \sim$ Volume form

Summary

Some equivalent Definitions for X^6 Calabi-Yau Threefold

- Kähler, $c_1(TX) = 0$
- Kähler, vanishing Ricci curvature
- Kähler, holonomy $\subset SU(n)$
- Kähler, nowhere vanishing global holomorphic 3-form (volume)
- Covariant constant spinor
- Canonical bundle (sheaf) $K_X := \bigwedge^n T_X^* \simeq \mathcal{O}_X$
- low-energy SUSY in 4D from string compactification

Some Topological Properties I

- **Hodge Numbers** $h^{p,q}(X) = \dim H^{p,q}_{\bar\partial}(X)$
  - Hodge decomposition and Betti Numbers: $b_k = \sum\limits_{p+q=k} h^{p,q}(X)$
  - Complex conjugation $\rightsquigarrow h^{p,q} = h^{q,p}$
  - Hodge star (Poincaré) $\rightsquigarrow h^{p,q} = h^{n-p,n-q}$
  - Hodge Diamond:

$$\begin{array}{ccccccc}
 & & & h^{0,0} & & & \\
 & & h^{0,1} & & h^{0,1} & & \\
 & h^{0,2} & & h^{1,1} & & h^{0,2} & \\
h^{0,3} & & h^{2,1} & & h^{2,1} & & h^{0,3} \\
 & h^{0,2} & & h^{1,1} & & h^{0,2} & \\
 & & h^{0,1} & & h^{0,1} & & \\
 & & & h^{0,0} & & &
\end{array}$$

  - Compact, connected, Kähler: $h^{0,0} = 1$ (constant functions)
  - If simply-connected:

$$\pi_1(X) = 0 \rightsquigarrow H_1(X) = \pi_1(X)/[\ ,\ ] = 0 \rightsquigarrow h^{1,0} = h^{0,1} = 0$$

Some Topological Properties II

- Finally, CY3 has $h^{3,0} = h^{0,3} = 1$ [unique holomorphic 3-form], also $h^{p,0} = h^{3-p,0}$ by contracting $(p,0)$-form with $\bar{\Omega}$ to give $(p,3)$-form, then use Poincaré duality to give $(3-p,0)$-form
- 2-topological numbers for a (connected, simply connected) CY3:

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{1,1} & & 0 & \\ 1 & & h^{2,1} & & h^{2,1} & & 1 \\ & 0 & & h^{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array}$$

(Kähler, Complex-Structure) : $(h^{1,1}, h^{2,1})$

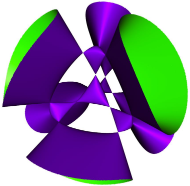
$\chi(X) = 2(h^{1,1} - h^{2,1})$

- Moduli Space of CY3 locally: $\mathcal{M} \simeq \mathcal{M}^{2,1} \times \mathcal{M}^{1,1}$

Explicit Examples of Calabi-Yau Manifolds

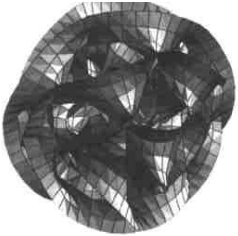
- $d = 1$ Torus $T^2 = S^1 \times S^1$

- $d = 2$ K3



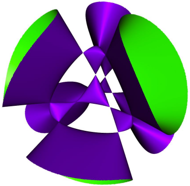
; 4-torus: $T^4 = (S^1)^4$

- $d = 3$ CY3: Unclassified, billions known

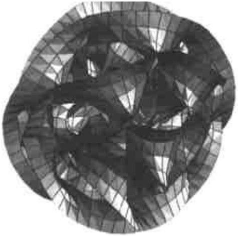


Explicit Examples of Calabi-Yau Manifolds

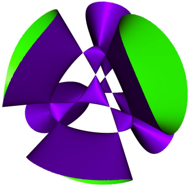
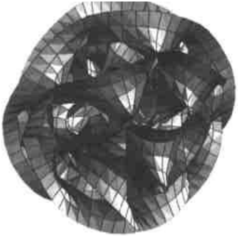
- $d = 1$ Torus $T^2 = S^1 \times S^1$ 

- $d = 2$ K3  ; 4-torus: $T^4 = \left(S^1\right)^4$

- $d = 3$ CY3: Unclassified, billions known



Explicit Examples of Calabi-Yau Manifolds

- $d = 1$ Torus $T^2 = S^1 \times S^1$ 
- $d = 2$ K3  ; 4-torus: $T^4 = \left(S^1\right)^4$
- $d = 3$ CY3: Unclassified, billions known 

As Projective Varieties

- Embed X into $\mathbb{P}^n$ as **complete intersection** of K polynomials

$$n = K + 3$$

- Canonical bundle $\mathcal{K}_X \simeq \wedge^{\dim(X)} T^*_X$; algebraic condition for Calabi-Yau: $K_X \simeq \mathcal{O}_X$ (indeed $c_1(TX) = 0$)
- **Adjunction formula** for subvariety $X \subset A$: $\mathcal{K}_X = (K_A \otimes N^*)|_X$
- Recall $K_{A=\mathbb{P}^n} \simeq \mathcal{O}_{\mathbb{P}^n}(-n-1)$ and $K_X \simeq \mathcal{O}_X$, thus:

$$\text{degree}(X) = n + 1$$

- Find only 5 solutions. These all have $h^{1,1}(X) = 1$, inherited from the 1 Kähler class of $\mathbb{P}^n$; called cyclic Calabi-Yau threefolds

Cyclic Manifolds

Intersection	$\mathcal{A}$	Configuration	$\chi(X)$	$h^{1,1}(X)$	$h^{2,1}(X)$	$d(X)$	$\tilde{c}_2(TX)$
Quintic	$\mathbb{P}^4$	$[4\|5]$	-200	1	101	5	10
Quadric and quartic	$\mathbb{P}^5$	$[5\|2\ 4]$	-176	1	89	8	7
Two cubics	$\mathbb{P}^5$	$[5\|3\ 3]$	-144	1	73	9	6
Cubic and 2 quadrics	$\mathbb{P}^6$	$[6\|3\ 2\ 2]$	-144	1	73	12	5
Four quadrics	$\mathbb{P}^7$	$[7\|2\ 2\ 2\ 2]$	-128	1	65	16	4

- Euler numbers quite large, $d(X)$ is volume normalisation
- used standard **matrix configuration notation**
- most famous example: Quintic 3-fold $[4|5]$

$$\{\sum_{i=0}^{4} x_i^5 = 0\} \subset \mathbb{P}^4_{[x_0:\ldots x_4]}$$

written as Fermat quintic, also has $h^{2,1}(X) = 101$ deformation parameters

Part I

Strings and the Compact Calabi-Yau Landscape

Triadophilia: A 40-year search

- A 2-decade Problem: [Candelas-Horowitz-Strominger-Witten] (1986)
  - $E_8 \supset SU(3) \times SU(2) \times U(1)$ Natural Gauge Unification
  - Mathematically succinct
  - Witten: "still the best hope for the real world"
- CY3 X, tangent bundle $SU(3) \Rightarrow E_6$ GUT: commutant $E_8 \to SU(3) \times E_6$ (generalize later)
  - Particle Spectrum:

Generation	$n_{27} = h^1(X, TX) = h^{2,1}_{\bar\partial}(X)$
Anti-Generation	$n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar\partial}(X)$

- Net-generation: $\chi = 2(h^{1,1} - h^{2,1})$
- Question: Are there Calabi-Yau threefolds with Euler character ± 6?
- Strominger was visiting Yau at the IAS in 1986-7

Triadophilia: A 40-year search

- A 2-decade Problem: [Candelas-Horowitz-Strominger-Witten] (1986)
  - $E_8 \supset SU(3) \times SU(2) \times U(1)$ Natural Gauge Unification
  - Mathematically succinct
  - Witten: "still the best hope for the real world"
- CY3 X, tangent bundle $SU(3) \Rightarrow E_6$ GUT: commutant $E_8 \to SU(3) \times E_6$ (generalize later)
  - Particle Spectrum: Generation $n_{27} = h^1(X, TX) = h^{2,1}_{\bar{\partial}}(X)$; Anti-Generation $n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar{\partial}}(X)$
- Net-generation: $\chi = 2(h^{1,1} - h^{2,1})$
- Question: Are there Calabi-Yau threefolds with Euler character ± 6?
- Strominger was visiting Yau at the IAS in 1986-7

Triadophilia: A 40-year search

- A 2-decade Problem: [Candelas-Horowitz-Strominger-Witten] (1986)
  - $E_8 \supset SU(3) \times SU(2) \times U(1)$ Natural Gauge Unification
  - Mathematically succinct
  - Witten: "still the best hope for the real world"
- CY3 X, tangent bundle $SU(3) \Rightarrow E_6$ GUT: commutant $E_8 \to SU(3) \times E_6$ (generalize later)
  - Particle Spectrum: Generation $n_{27} = h^1(X, TX) = h^{2,1}_{\bar{\partial}}(X)$
    Anti-Generation $n_{\overline{27}} = h^1(X, TX^*) = h^{1,1}_{\bar{\partial}}(X)$
- Net-generation: $\chi = 2(h^{1,1} - h^{2,1})$
- Question: Are there Calabi-Yau threefolds with Euler character ± 6?
- Strominger was visiting Yau at the IAS in 1986-7

Complete Intersection Calabi-Yau (CICY) 3-folds

- immediately: Quintic Q in $\mathbb{P}^4$ is CY3, recall: $Q_{\chi}^{h^{1,1},h^{2,1}} = Q_{-200}^{1,101}$ so too may generations (even with quotient $-200 \notin 3\mathbb{Z}$)
- [Candelas-A. He-Hübsch-Lutken-Schimmrigk-Berglund] (1986-1990)
  - dim(Ambient space) - #(defining Eq.) = 3 (complete intersection)

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

– K eqns of multi-degree $q_j^i \in \mathbb{Z}_{\geq 0}$ embedded in $\mathbb{P}^{n_1} \times \ldots \times \mathbb{P}^{n_m}$

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

– M^T also CICY

- Famous Examples

The First Data-sets in Mathematical Physics/Geometry I

- 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 from 1×1 to max(row) $= 12$, max(col) $= 15$; with $q_j^i \in [0,5]$
  - 266 distinct Hodge pairs $(h^{1,1}, h^{2,1}) = (1,65), \ldots, (19,19)$
  - 70 distinct Euler $\chi \in [-200, 0]$ (all negative)
  - [V. Braun, 1003.3235] : 195 have freely-acting symmetries (quotients), 37 different finite groups (from $\mathbb{Z}_2$ to $\mathbb{Z}_8 \rtimes H_8$)

- Rmk: Integration pulls back to ambient product of projective space A

$$\int_X \cdot = \int_A \mu \wedge \cdot \ , \qquad \mu := \bigwedge_{j=1}^{K} \left(\sum_{r=1}^{m} q_r^j J_r \right) \ .$$

Topological Quantities

- Chern classes of CICY

$$
\begin{array}{rcl}
c_1^r(T_X) & = & 0 \\
c_2^{rs}(T_X) & = & \frac{1}{2}\left[-\delta^{rs}(n_r+1)+\sum_{j=1}^{K} q_j^r q_j^s\right] \\
c_3^{rst}(T_X) & = & \frac{1}{3}\left[\delta^{rst}(n_r+1)-\sum_{j=1}^{K} q_j^r q_j^s q_j^t\right]
\end{array}
$$

- Triple intersection numbers: $d_{rst} = \int_X \cdot = \int_A J_r \wedge J_s \wedge J_t$
- Euler number: $\chi(X) = \text{Coefficient}(c_3^{rst} J_r J_s J_t \cdot \mu, \quad \prod_{r=1}^{m} J_r^{n_r})$
- As always, computing individual terms $(h^{1,1}, h^{2,1})$ hard even though $h^{1,1} - h^{2,1} = \frac{1}{2}\chi$ (index theorem)

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:

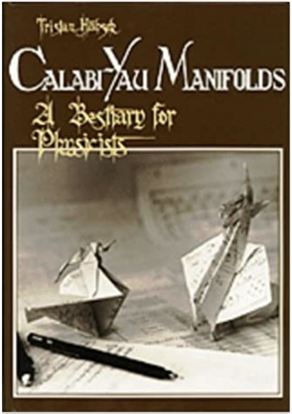
$$0 \to T_X \to T_M|_X \to N_X \to 0$$

- Induces long exact sequence in cohomology:

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

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

A Classic

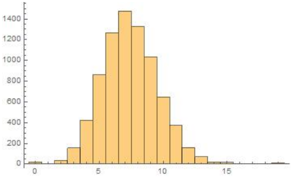


T. Hübsch, *CY Manifolds: a bestiary for physicists, 1992, WS*

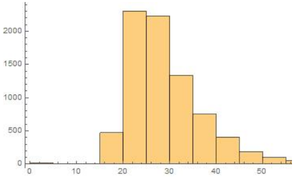
first book to introduce Algebraic Geometry to physicists

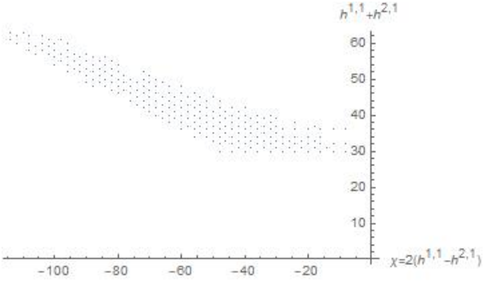
Distribution

$h^{1,1}$



$h^{2,1}$

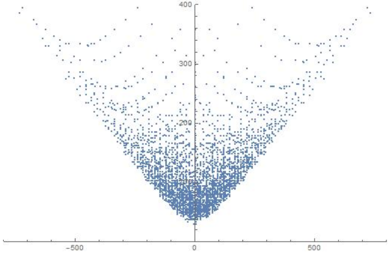




The First Data-sets in Mathematical Physics/Geometry II

[Candelas-Lynker-Schimmrigk, 1990] Hypersurfaces in Weighted $\mathbb{P}^4$

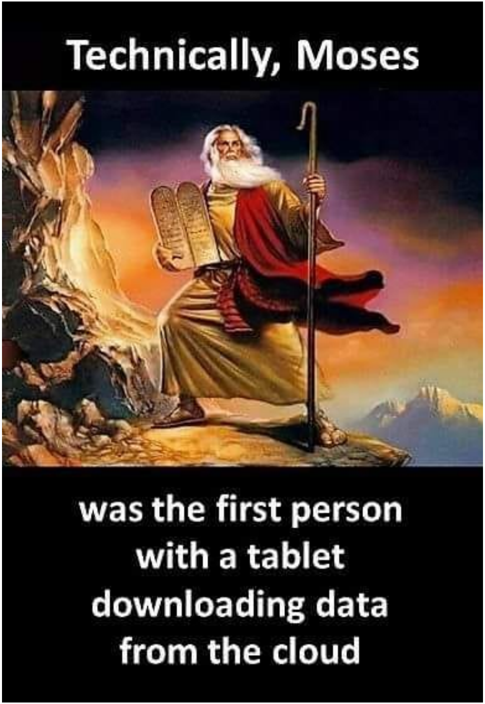
- generic homog deg $= \sum_{i=0}^{4} w_i$ polynomial in $W\mathbb{P}^4_{[w_0:w_1:w_2:w_3:w_4]} \simeq (\mathbb{C}^5 - \{0\})/(x_0, x_1, x_2, x_3, x_4) \sim (\lambda^{w_0} x_0, \lambda^{w_1} x_1, \lambda^{w_2} x_2, \lambda^{w_3} x_3, \lambda^{w_4} x_4)$
- specified by a single integer 5-vector: w_i
- Rmk: ambient WP4 is singular (need to resolve)



7555 inequivalent 5-vectors w_i

2780 Hodge pairs

$\chi \in [-960, 960]$



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

Elliptically Fibered CY3: [Gross, Morrison-Vafa, 1994]

- X elliptically fibered over some base B: as Weierstraß model in $\mathbb{P}^2_{[x:y:z]}$-bundle over B (g_2, g_3 complex structure coeff)

$$zy^2 = 4x^3 - g_2 x z^2 - g_3 z^3$$

  x, y, z, g_2, g_3 must be sections of powers of some line bundle $\mathcal{L}$ over B

- Specifically (x, y, z, g_2, g_3) are global sections of $(\mathcal{L}^{\oplus 2}, \mathcal{L}^{\oplus 3}, \mathcal{O}_B, \mathcal{L}^{\oplus 4}, \mathcal{L}^{\oplus 6})$
- $c_1(TX) = 0 \Rightarrow \mathcal{L} \simeq K_B^{-1} \Rightarrow B$ highly constrained :
  1. del Pezzo surface $d\mathbb{P}_{r=1,\ldots,9}$: $\mathbb{P}^2$ blown up at r points
  2. Hirzebruch surface $\mathbb{F}_{r=0,\ldots 12}$: $\mathbb{P}^1$-bundle over $\mathbb{P}^1$
  3. Enriques surface $\mathbb{E}$: involution of K3
  4. Blowups of $\mathbb{F}_r$

Elliptically Fibered CYn

- Belief (Conjecture?): VAST majority of CYn are elliptic fibrations
- Kollar Conjecture: A CY n-fold $\mathcal{M}$ is elliptic iff there exists a $(1,1)$-class $D \in H^2(\mathcal{M}, \mathbb{Q})$ s.t. for every algebraic curve C
  - $D \cdot C \geq 0; \quad D^{n-1} \neq 0; \quad D^n = 0$
  - Oguiso, Wilson: True for $n = 3$ if D is effective or $D \cdot c_2(\mathcal{M}) \neq 0$
- Anderson-Gao-Gray-Lee-Lukas: 99.33% (all but 53) of the $7,868$ CICY3; 99.95% (all but 462) of $905,684$ CICY4
- Huang-Taylor: KS-dataset (see shortly)
- Quintic is not, Schön is

Tour de Force: The Kreuzer-Skarke Dataset

- Generalize WP4, take Toric Variety $A(\Delta_n)$ and consider hypersurface therein
- $A(\Delta_n)$ is special: it is constructed from a reflexive polytope Lattice Polytopes
- THM [Batyrev-Borisov, '90s] anti-canonical divisor in $X(\Delta_n)$ gives a smooth Calabi-Yau $(n-1)$-fold as hypersurface:

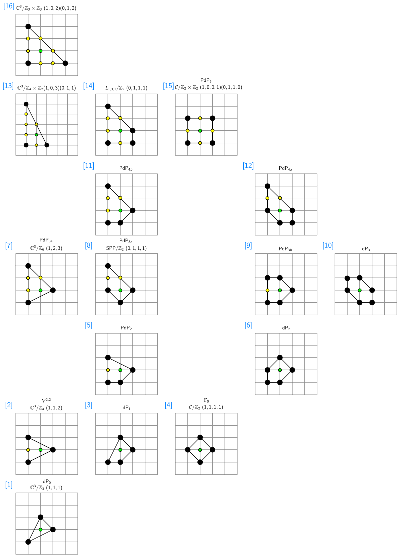
$$0 = \sum_{\mathbf{m}\in\Delta} C_{\mathbf{m}} \prod_{\rho=1}^{k} x_\rho^{\langle \mathbf{m}, \mathbf{v}_\rho\rangle + 1} , \qquad \Delta^\circ = \{\mathbf{v} \in \mathbb{R}^4 \mid \langle \mathbf{m}, \mathbf{v}\rangle \geq -1 \ \forall \mathbf{m} \in \Delta\}$$

$\mathbf{v}_\rho$ vertices of Δ.

- Simplest case: $A = \mathbb{P}^4$ and we have quintic $[4|5]$ again.

$$\Delta : \begin{array}{ccl} \mathbf{m}_1 & = & (-1,-1,-1,-1), \\ \mathbf{m}_2 & = & (\ 4,-1,-1,-1), \\ \mathbf{m}_3 & = & (-1,\ 4,-1,-1), \\ \mathbf{m}_4 & = & (-1,-1,\ 4,-1), \\ \mathbf{m}_5 & = & (-1,-1,-1,\ 4)\ , \end{array} \qquad \Delta^\circ : \begin{array}{ccl} \mathbf{v}_1 & = & (1,0,0,0), \\ \mathbf{v}_2 & = & (0,1,0,0), \\ \mathbf{v}_3 & = & (0,0,1,0), \\ \mathbf{v}_4 & = & (0,0,0,1), \\ \mathbf{v}_5 & = & (-1,-1,-1,-1)\ . \end{array}$$

Reflexive Polygons: 16 special elliptic curves



- THM (classical): All Δ_2 are $GL(2;\mathbb{Z})$ equivalent to one of the 16
- $\rightarrow$ #vertices: 3, . . . , 6
- $\uparrow$ #lattice points: 4, . . . , 10
- 4 self-dual
- 5 smooth $X(\Delta_2)$ = toric del Pezzo surfaces: $dP_{0,1,2,3}$, $\mathbb{P}^1 \times \mathbb{P}^1$ (smooth toric Fano surfaces)

Known Classification Results

- $GL(n;\mathbb{Z})$-equivalence classes of reflexive Δ_n finite for each n
- Kreuzer[†]-Skarke (Using PALP) [1990s]: a fascinating sequence

dimension	1	2	3	4	...
# Reflexive Polytopes	1	16	4319	473,800,776	...
# Regular	1	5	18	124	...

- $n \geq 5$ still not classified; generating function also not known
- Smooth ones known for a few more dimensions (Kreuzer-Nill, Øbro, Paffenholz): $\{1,\ 5,\ 18,\ 124,\ 866,\ 7622,\ 72256,\ 749892,\ 8229721\ldots\}$
- $n = 2, 3$ built into SAGE

Kreuzer-Skarke

- Kreuzer†-Skarke 1997-2002: 473,800,776 Δ_4
  - AT LEAST this many CY3 hypersurfaces in $A(\Delta_4)$: CY3 depends on triangulation (resolution) of Δ, but Hodge numbers only depend on Δ_4 (Batyrev-Borisov):

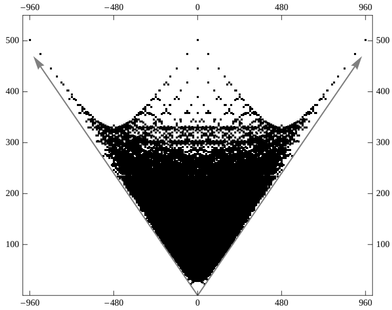
$$\begin{aligned} h^{1,1}(X) &= \ell(\Delta^\circ) - \sum_{\mathrm{codim}\theta^\circ=1} \ell^\circ(\theta^\circ) + \sum_{\mathrm{codim}\theta^\circ=2} \ell^\circ(\theta^\circ)\ell^\circ(\theta) - 5; \\ h^{1,2}(X) &= \ell(\Delta) - \sum_{\mathrm{codim}\theta=1} \ell^\circ(\theta) + \sum_{\mathrm{codim}\theta=2} \ell^\circ(\theta)\ell^\circ(\theta^\circ) - 5\ . \end{aligned}$$

  - Dual polytope $\Delta \leftrightarrow \Delta^\circ$ = mirror symmetry
- Vienna group (KS, Knapp,...), Oxford group (Candelas, Lukas, YHH, ...), MIT group (Taylor,Johnson, Wang, ...), Northeastern/Wits Collab (Nelson, Jejjala, YHH), Virginia Tech (Anderson, Gray, Lee, ...) Tsinghua/London/Oxford Collab (Yau, Seong, YHH)

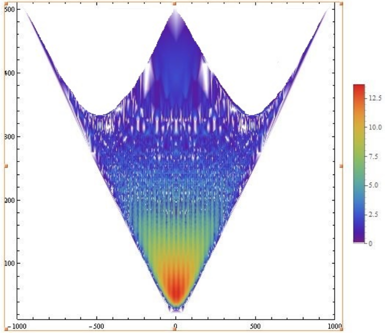
Georgia O'Keefe

30,108 distinct Hodge pairs, $\chi \in [-960, 960]$;

$(h^{1,1}, h^{2,1}) = (27, 27)$ dominates: 910113 instances



In Philip's Office



YHH (1308.0186)

Refined Structure in KS Data

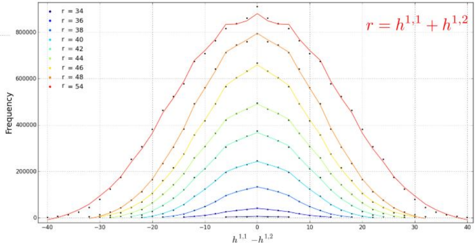
- DATABASES:

  `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`

  `http://www.rossealtman.com/`

- Altman-Gray-YHH-Jejjala-Nelson 2014-17 triangulate Δ_4 (orders more than 1/2-billion): up to $h^{1,1} = 7$
- Candelas-Constantin-Davies-Mishra 2011-17 special small Hodge numbers
- Taylor, Johnson, Wang et al. 2012-17 elliptic fibrations
- YHH-Jejjala-Pontiggia 2016 distribution of Hodge, χ, Pseudo-Voigt

KS stats



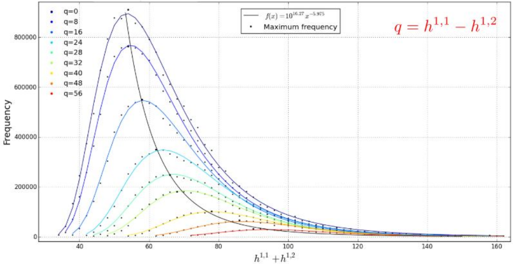
Pseudo-Voigt distribution

sum of Gaussian and Cauchy

$$(1-\alpha)\frac{A}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} + \alpha\frac{A}{\pi}\left[\frac{\sigma^2}{(x-\mu)^2+\sigma^2}\right]$$

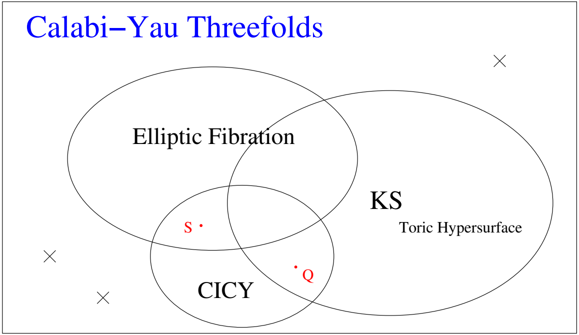
Planck distribution

$$\frac{A}{x^n}\frac{1}{e^{b/(x-c)}-1}$$



He, VJ, Pontiggia (2015)

The Compact CY3 Landscape

- 

  40 years of research by mathematicians and physicists; 10^{10} data-points (and growing)

- **OPEN CONJECTURES:**

  Yau: Topological type of CY in any dim is FINITE

  Reid's Fantasy: All CY3 are connected by conifold-like transitions

CY3 Compactification: Recent Development

- E_6 GUTs less favourable, $SU(5)$ and $SO(10)$ GUTs: general embedding
  - Instead of TX, use (poly-)stable holomorphic vector bundle V
  - LE particles $\sim$ massless modes of V-twisted Dirac Operator: $\not{\nabla}_{X,V}\Psi = 0$
  - massless modes of $\not{\nabla}_{X,V} \xleftrightarrow{1:1}$ V-valued cohomology groups
- 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	$\to$	$(1,78) \oplus (3,27) \oplus (\bar{3},\overline{27}) \oplus (8,1)$
$SU(4) \times SO(10)$	248	$\to$	$(1,45) \oplus (4,16) \oplus (\bar{4},\overline{16}) \oplus (6,10) \oplus (15,1)$
$SU(5) \times SU(5)$	248	$\to$	$(1,24) \oplus (5,\overline{10}) \oplus (\bar{5},10) \oplus (10,5) \oplus (\overline{10},\bar{5}) \oplus (24,1)$

- Particle content

Decomposition	**Cohomologies**
$\mathrm{SU}(3) \times \mathrm{E}_6$	$n_{27} = h^1(V)$, $n_{\overline{27}} = h^1(V^*) = h^2(V)$, $n_1 = h^1(V \otimes V^*)$
$\mathrm{SU}(4) \times \mathrm{SO}(10)$	$n_{16} = h^1(V)$, $n_{\overline{16}} = h^2(V)$, $n_{10} = h^1(\wedge^2 V)$, $n_1 = h^1(V \otimes V^*)$
$\mathrm{SU}(5) \times \mathrm{SU}(5)$	$n_{10} = h^1(V^*)$, $n_{\overline{10}} = h^1(V)$, $n_5 = h^1(\wedge^2 V)$, $n_{\bar{5}} = h^1(\wedge^2 V^*)$, $n_1 = h^1(V \otimes V^*)$

- Further to SM: $H \xrightarrow{\text{Wilson Line}} SU(3) \times SU(2) \times U(1)$

Ubi Materia, Ibi Geometria

- Issues in low-energy physics $\sim$ Precise questions in Alg Geo of (X, V)
  - Particle Content $\sim$ (tensor powers) V Equivariant 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
- e.g., for $\pi_1(X) = \mathbb{Z}_3 \times \mathbb{Z}_3$ WL:

Cohomology	Representation	Multiplicity	Name
$[\alpha_1^2 \alpha_2 \otimes H^1(X,V)]^{inv}$	$(\mathbf{3},\mathbf{2})_{1,1}$	3	left-handed quark
$[\alpha_1^2 \otimes H^1(X,V)]^{inv}$	$(\mathbf{1},\mathbf{1})_{6,3}$	3	left-handed anti-lepton
$[\alpha_1^2 \alpha_2^2 \otimes H^1(X,V)]^{inv}$	$(\overline{\mathbf{3}},\mathbf{1})_{-4,-1}$	3	left-handed anti-up
$[\alpha_2^2 \otimes H^1(X,V)]^{inv}$	$(\overline{\mathbf{3}},\mathbf{1})_{2,-1}$	3	left-handed an ti-down
$[H^1(X,V)]^{inv}$	$(\mathbf{1},\mathbf{2})_{-3,-3}$	3	left-handed lepton
$[\alpha_1 \otimes H^1(X,V)]^{inv}$	$(\mathbf{1},\mathbf{1})_{0,3}$	3	left-handed anti-neutrino
$[\alpha_1 \otimes H^1(X,\wedge^2 V)]^{inv}$	$(\mathbf{1},\mathbf{2})_{3,0}$	1	up Higgs
$[\alpha_1^2 \otimes H^1(X,\wedge^2 V)]^{inv}$	$(\mathbf{1},\mathbf{2})_{-3,0}$	1	down Higgs

A Heterotic Standard Model

- [Braun-YHH-Ovrut-Pantev] (hep-th/0512177, 0601204)

$$
\begin{tikzcd}
 & X \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\
B_1 \arrow[dr, "\beta_1"'] & & B_2 \arrow[dl, "\beta_2"] \\
 & \mathbb{P}^1 &
\end{tikzcd}
$$

- $X_0^{19,19}$ double-fibration over dP_9 $\quad \pi_1(X) = \mathbb{Z}_3 \times \mathbb{Z}_3$
- V stable $SU(4)$ bundle: Generalised Serre Constrct
- Couple to $\mathbb{Z}_3 \times \mathbb{Z}_3$ Wilson Line
- Matter $= \mathbb{Z}_3 \times \mathbb{Z}_3$-Equivariant cohomology on $X_0^{3,3}$

- Exact $SU(3) \times SU(2) \times U(1) \times U(1)_{B-L}$ spectrum:

  $\boxed{\text{No exotics; no anti-generation; 1 pair of Higgs; RH Neutrino}}$

- $SU(5) \to SU(3) \times SU(2) \times U(1)$ version [Bouchard-Cvetic-Donagi]

  same manifold

- $X_0^{19,19}$ is a CICY! Obvervatio Curiosa

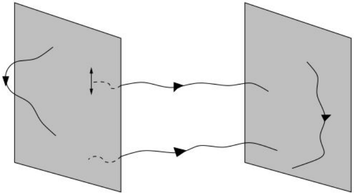
Algorithmic Compactification

- Searching the MSSM, *Sui Generis?*
  - $\sim 10^7$ Spectral Cover bundles [Donagi, Friedman-Morgan-Witten, 1996-8] over elliptically fibered CY3 (2005-9), [Donagi-YHH-Ovrut-Pantev-Reinbacher, Gabella-YHH-Lukas,. . .]
  - $\sim 10^5$ (Monad) Bundles over all CICYs [Anderson-Gray-YHH-Lukas, 2007-9]
  - Monad Bundles over KS YHH-Kreuzer-Lee-Lukas 2010-11: ~ 200 in 10^5 3-gens
  - culminating in .. Stable Sum of Line Bundles over CICYs (Oxford-Penn-Virginia 2012-) Anderson-Gray-Lukas-Ovrut-Palti ~ 200 in 10^{10} MSSM
- meanwhile . . . LANDSCAPE grew with D-branes Polchinski 1995, M-Theory/G_2 Witten, 1995, F-Theory/4-folds Katz-Morrison-Vafa, 1996, AdS/CFT Maldacena 1998, Flux-compactification Kachru-Kallosh-Linde-Trivedi, 2003, . . .

Digression

D-branes, Type II & Non-Compact CY

- D-branes Dirichlet Boundary conditions for open strings;
- D-brane world-volumes: Dp has $p+1$-D w.v.



$D1, D3, \ldots, D9$ of dimensions $1+1, \ldots, 9+1$;

DYNAMICAL: Carry charges $(2, 4, \ldots, 10$ forms) $\int_{Dp} Q^{(p+1)}$

- i.e., Open strings carry charges (Chan-Paton factors) $\Rightarrow$

  D-branes = Supports of Sheafs (strictly: D-brane = object in $D^b(Coh)$)
- important property: GAUGE ENHANCEMENT
  - i.e., *world-volume sees a $U(1)$-bundle*
  - Bringing together (stack) n parallel D-branes $U(1)^n \to U(n)$

Another $10 = 4 + 6$

- $\boxed{\text{SUMMARY}}$ Type IIB: 10D, Closed Strings, Open Strings/Dp-Branes, p odd
- $\mathbb{R}^{1,9} \simeq \mathbb{R}^{1,3}$(world-volume of D3) $\times X^6$(transverse non-compact CY3)
- SIMPLEST CASE: transverse CY3 $= \mathbb{C}^3$
  - Original Maldacena's AdS/CFT (1997):

    $\mathcal{N} = 4$ U(n) SYM on 4D world-volume of n D3s
  - R-symmetry $SU(4) \simeq SO(6)$ of S^5 in $AdS_5 \times S^5$
  - Gauge Fields A^μ: $\mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^n)$
  - Matter Fields $\mathcal{R} = \mathbf{4}, \mathbf{6}$: Adjoint (Weyl) fermions $\Psi^{\mathbf{4}}_{IJ}$: $\mathbf{4} \otimes \mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^n)$

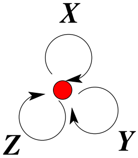
    Bosons $\Phi^{\mathbf{6}}_{IJ}$: $\mathbf{6} \otimes \mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^n)$

A Geometer's AdS/CFT

- $\boxed{\text{Rep. Variety(Quiver)} \sim \text{VMS(SUSY QFT)} \sim \text{affine/singular variety}}$

  e.g $\mathcal{N}=1$ Quiver variety = vacuum of F- & D-flatness = non-compact CY3

- $\mathcal{N}=4$ $U(N)$ Yang-Mills
  - 3 adjoint fields X, Y, Z with superpotential $W = \text{Tr}(XYZ - XZY)$



  - N D3-branes (w.v. is $\mathcal{N}=4$ in $\mathbb{R}^{3,1}$) $\perp \mathbb{R}^6$

    $\simeq \mathbb{C}^3$ = Vacuum Moduli Space
  - VMS $\simeq$ affine non-compact CY3 by construction
  - QUIVER = Finite graph (label = rk(gauge factor)) + relations from W
    - Matter Content: Nodes + arrows
    - Relations (F-Terms): $D_i W = 0 \rightsquigarrow [X,Y] = [Y,Z] = [X,Z] = 0$
  - Here $\mathbb{C}^3$ is real cone over S^5 (simplest Sasaki-Einstein 5-manifold), others?

Orbifolds (V-manifolds)

- Orbifolds: next best thing to $\mathbb{C}^3$ (Satake 60's);
- Transverse CY3 $\simeq \mathbb{C}^3/\{\Gamma \subset SU(k)\}$ that admit crepant resolution, i.e., resolve to Calabi-Yau; Γ discrete finite subgroup of holonomy $SU(k)$; $k = 2, 3$
- Γ-Projection: $\gamma A^\mu \gamma^{-1} = A^\mu$ and $\Psi_{IJ} = R(\gamma)\gamma\Psi_{IJ}\gamma^{-1}$; i.e.,
  - Gauge Group $U(n) \Rightarrow \prod_i U(N_i)$
  - Matter fields decompose as

$$\begin{aligned}(\mathcal{R} \otimes \hom(\mathbb{C}^n, \mathbb{C}^n))^\Gamma &= \bigoplus_{i,j} \mathcal{R} \otimes \left(\mathbb{C}^{N_i} \otimes \mathbb{C}^{N_j *} \otimes \mathbf{r_i} \otimes \mathbf{r_j^*}\right)^\Gamma \\ &= \bigoplus_{i,j} a_{ij}^{\mathcal{R}} \left(\mathbb{C}^{N_i} \otimes \mathbb{C}^{N_j *}\right),\end{aligned}$$

where $\boxed{\mathcal{R} \otimes \mathbf{r}_i = \bigoplus_j a_{ij}^{\mathcal{R}} \mathbf{r}_j}$

  - $a_{ij}^{\mathbf{4}}$ bi-fundamental fermions: $(N_i, \bar{N}_j)$ of $SU(N_i) \times SU(N_j)$
  - $a_{ij}^{\mathbf{6}}$ bi-fundamental bosons: $(N_i, \bar{N}_j)$ of $SU(N_i) \times SU(N_j)$

Quivers

-

	Parent	$\xrightarrow{\Gamma}$	Orbifold Theory
SUSY	$\mathcal{N}=4$	$\rightsquigarrow$	$\mathcal{N}=2,\ \text{for } \Gamma \subset SU(2)$ $\mathcal{N}=1,\ \text{for } \Gamma \subset SU(3)$ $\mathcal{N}=0,\ \text{for } \Gamma \subset \{SU(4) \simeq SO(6)\}$
Gauge Group	$U(n)$	$\rightsquigarrow$	$\prod\limits_i U(N_i), \qquad \sum\limits_i N_i \dim \mathbf{r}_i = n$
Fermion Boson	$\Psi^{\mathbf{4}}_{IJ}$ $\Phi^{\mathbf{6}}_{IJ}$	$\rightsquigarrow$ $\rightsquigarrow$	$\Psi^{ij}_{f_{ij}}$ $\Phi^{ij}_{f_{ij}} \qquad \mathcal{R} \otimes \mathbf{r}_i = \bigoplus\limits_j a^{\mathcal{R}}_{ij} \mathbf{r}_j$

$I, J = 1, \ldots, n;\ f_{ij} = 1, \ldots, a^{\mathcal{R}=\mathbf{4},\mathbf{6}}_{ij}$

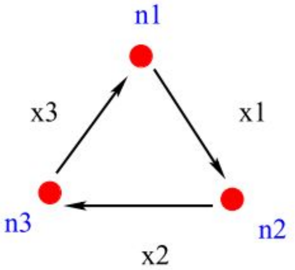
- In physics: Douglas & Moore (9603167), $\mathbb{C}^2/\mathbb{Z}_n$; Johnson & Meyers (9610140) Formalised in Lawrence, Nekrasov & Vafa, (9803015);

Quivers: Finite Graphs with Representation

- A Graphical way to represent this data
  - Node $i \sim$ gauge factor $U(N_i)$
  - Arrow $i \to j \sim$ bi-fundamental $(N_i, \bar{N}_j)$
- e.g.

Adjacency Matrix

$$A_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$



- Gabriel: 1970s: $x_1 \in \mathrm{Hom}(\mathbb{C}^{n_1}, \mathbb{C}^{n_2})$, etc.

McKay Correspondence

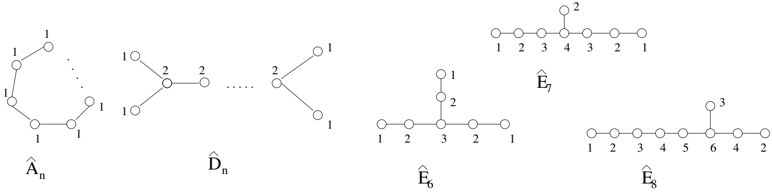
- Take the $\mathbb{C}^2/(\Gamma \subset SU(2)) \times \mathbb{C}$ case: Discrete Finite Subgroups of $SU(2)$
- F. Klein (1884) (double covers of those of $SO(3)$, i.e., symmetry groups of the Platonic solids)

Group	Name	Order
$A_n \simeq \mathbb{Z}_{n+1}$	Cyclic	$n+1$
D_n	Binary Dihedral	$2n$
E_6	Binary Tetrahedral	24
E_7	Binary Octahedral (Cube)	48
E_8	Binary Icosahedral (Dodecadedron)	120

- McKay (1980) Take the Clebsch-Gordan decomposition for $\mathcal{R}$ = fundamental **2** representation of $SU(2)$

ADE-ology

- $\mathbf{2} \otimes \mathbf{r}_i = \bigoplus_j a_{ij}^{\mathbf{2}} \mathbf{r}_j$ and treat $a_{ij}^{\mathbf{2}}$ as adjacency matrix
- McKay Quivers (rmk: Cartan matrix symmetric $\rightsquigarrow$ graph unoriented)
- QUIVERS = DYNKIN DIAG. OF CORRESPONDING AFFINE LIE ALGEBRA!!



Geometrical McKay

- Geometrically: González-Springberg & Verdier (1981)

  Crepant Resolution $K3 \to \mathbb{C}^2/\Gamma$

$$\begin{aligned} A_n : &\quad xy + z^n = 0 \\ D_n : &\quad x^2 + y^2 z + z^{n-1} = 0 \\ E_6 : &\quad x^2 + y^3 + z^4 = 0 \\ E_7 : &\quad x^2 + y^3 + yz^3 = 0 \\ E_8 : &\quad x^2 + y^3 + z^5 = 0 \end{aligned}$$

- Intersection matrix of -2 exceptional curves in the blowup $\rightsquigarrow$ Quiver
- Bridgeland-King-Reid (1999) Use Fourier-Mukai: McKay as an auto-equivalence in $\mathcal{D}^b(\mathrm{coh}(\widetilde{X/G})) = \mathcal{D}^b(\mathrm{coh}^G(X))$

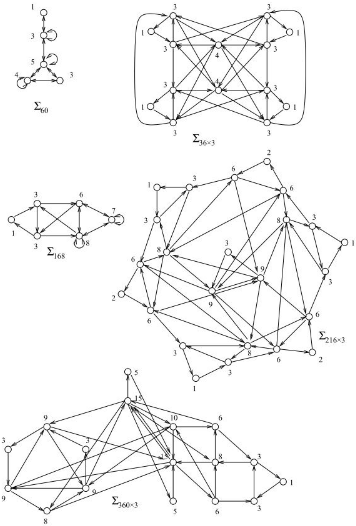
CY3 case: $\mathbb{C}^3/(\Gamma \subset SU(3))$

- *McKay Quiver* $\Rightarrow \mathcal{N}=2$ *SUSY gauge theory on 4D world-volume*
- $\mathcal{N}=1$ SUSY: Need discrete finite groups $\Gamma \subset SU(3)$
- Classification: Blichfeldt (1917)

Infinite Series	$\Delta(3n^2), \Delta(6n^2)$
Exceptionals	$\Sigma_{36\times 3}, \Sigma_{60\times 3}, \Sigma_{168\times 3}, \Sigma_{216\times 3}, \Sigma_{360\times 3}$

- Gives chiral $\mathcal{N}=1$ gauge theories in 4D wv of D3-probe
- most phenomenologically interesting
- Hanany & YHH hep-th/9811183
- Rmk: Crepant Resolutions to CY3 and Generalised McKay (Reid, Ito et al.) not as well established

SU(3) quivers and $\mathcal{N} = 1$ gauge theories



$\Gamma \subset SU(3)$	Gauge Group
$\widehat{A_n} \cong \mathbb{Z}_{n+1}$	(1^{n+1})
$\mathbb{Z}_k \times \mathbb{Z}_{k'}$	$(1^{kk'})*$
$\widehat{D_n}$	$(1^4, 2^{n-3})$
$\widehat{E_6} \cong \mathcal{T}$	$(1^3, 2^3, 3)$
$\widehat{E_7} \cong \mathcal{O}$	$(1^2, 2^2, 3^2, 4)$
$\widehat{E_8} \cong \mathcal{I}$	$(1, 2^2, 3^2, 4^2, 5, 6)$
$E_6 \cong T$	$(1^3, 3)$
$E_7 \cong O$	$(1^2, 2, 3^2)$
$E_8 \cong I$	$(1, 3^2, 4, 5)$
$\Delta_{3n^2} (n = 0 \textbf{ mod } 3)$	$(1^9, 3^{\frac{n^2}{3}-1})*$
$\Delta_{3n^2} (n \neq 0 \textbf{ mod } 3)$	$(1^3, 3^{\frac{n^2-1}{3}})*$
$\Delta_{6n^2} (n \neq 0 \textbf{ mod } 3)$	$(1^2, 2, 3^{2(n-1)}, 6^{\frac{n^2-3n+2}{6}})*$
Σ_{168}	$(1, 3^2, 6, 7, 8)*$
Σ_{216}	$(1^3, 2^3, 3, 8^3)$
$\Sigma_{36\times3}$	$(1^4, 3^8, 4^2)*$
$\Sigma_{216\times3}$	$(1^3, 2^3, 3^7, 6^6, 8^3, 9^2)*$
$\Sigma_{360\times3}$	$(1, 3^4, 5^2, 6^2, 8^2, 9^3, 10, 15^2)*$

DICTIONARY: Quivers & Gauge Theory

$$S = \int d^4x \, [\int d^2\theta d^2\bar{\theta} \; \Phi_i^\dagger e^V \Phi_i + \left(\frac{1}{4g^2} \int d^2\theta \; \mathrm{Tr}\, \mathcal{W}_\alpha \mathcal{W}^\alpha + \int d^2\theta \; W(\Phi) + \text{c.c.} \right)]$$

W = superpotential $\qquad V(\phi_i, \bar{\phi_i}) = \sum_i \left| \frac{\partial W}{\partial \phi_i} \right|^2 + \frac{g^2}{4} (\sum_i q_i |\phi_i|^2)^2$

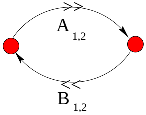
- Encode into QUIVER (rep of finite labelled graph with relations):

k nodes, dim vec $(N_1, \ldots, N_k)$	$\prod_{j=1}^k U(N_j)$ gauge group
Arrow $i \to j$	bi-fund X_{ij} field $(\square, \overline{\square})$ of $U(N_i) \times U(N_j)$
Loop $i \to i$	adjoint ϕ_i field of $U(N_i)$
Cycles	Gauge Invariant Operator
2-cycles	Mass-terms
$W = \sum c_i$ cycles$_i$	Superpotenital
Relations	Jacobian of $W(\phi_i, X_{ij})$

- VACUUM $\sim V(\phi_i, \bar{\phi_i}) = 0 \Rightarrow \begin{cases} \frac{\partial W}{\partial \phi_i, X_i} = 0 & \text{F-TERMS} \\ \sum\limits_i q_i |\phi_i|^2 + q_k |X_k| = 0 & \text{D-TERMS} \end{cases}$

Another Famous Example: Conifold

- $SU(N) \times SU(N)$ gauge theory with 4 bi-fundamental fields



QUIVER

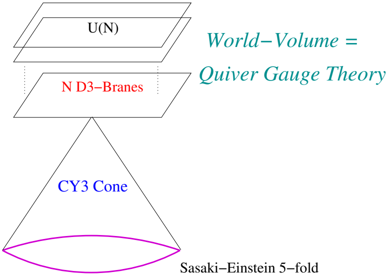
	$SU(N)$	$SU(N)$
$A_{i=1,2}$	$\square$	$\overline{\square}$
$B_{j=1,2}$	$\overline{\square}$	$\square$

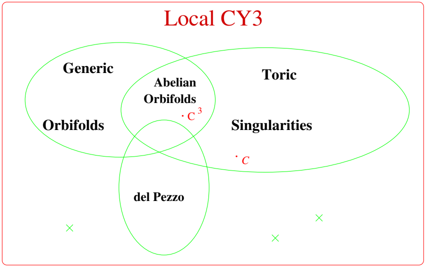
$W = \mathrm{Tr}(\epsilon_{il}\epsilon_{jk} A_i B_j A_l B_k)$

- D3-branes transverse to the conifold singularity = $(\{uv = wz\} \subset \mathbb{C}^4)$ = VMS (Klebanov-Witten 1999] $\mathcal{N} = 1$ "conifold" Theory)
- # gauge factors = $N_g = 2$; # fields = $N_f = 4$; # terms in W = $N_w = 2$
- Observatio Curiosa: $N_g - N_f + N_w = 0$, as with $\mathbb{C}^3$, true for almost all known cases in AdS_5/CFT_4

The Landscape of Affine (Singular) CY3

- 2 decade programme of the School of A. Hanany:



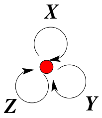


- Orbifolds: $\mathbb{C}^3/(\Gamma \subset SU(3))$ Generalized McKay Correspondence (Hanany-YHH, 98); Fano (del Pezzo): $dP_{0,\ldots,8}$ (w/ Hanany,Feng, Franco, et al. 98 - 00); LARGEST FAMILY by far Toric: e.g., conifold, $Y^{p,q}$, $L^{p,q}$ ...

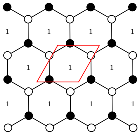
$\mathcal{M}$ Toric CY3 $\longleftrightarrow$ Bipartite Graph on T^2

Feng, Franco, Hanany, YHH, Kennaway, Martelli, Mekareeya, Seong, Sparks, Vafa, Vegh, Yamazaki, Zaffaroni ...

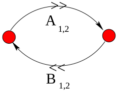
- $N_g - N_f + N_w = 0$ is Euler relation for a tiling of torus
- $\text{Jac}(W)$ = binomial ideal (toric): bipartite Notation for Toric Cones



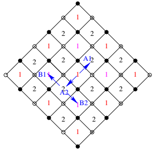
$\rightarrow$



$$W = \text{Tr}(XYZ - XZY)$$



$\rightarrow$



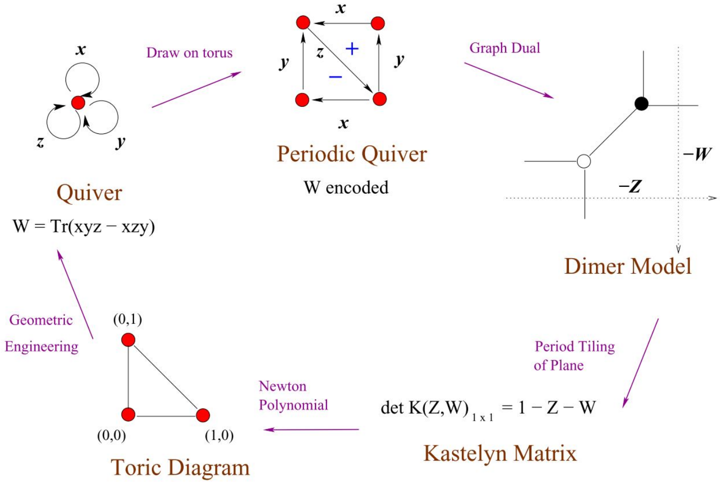
$$W = \text{Tr}(\epsilon_{il}\epsilon_{jk} A_i B_j A_l B_k)$$

Toric CY3, Mirror Symmetry & Bipartite Tilings

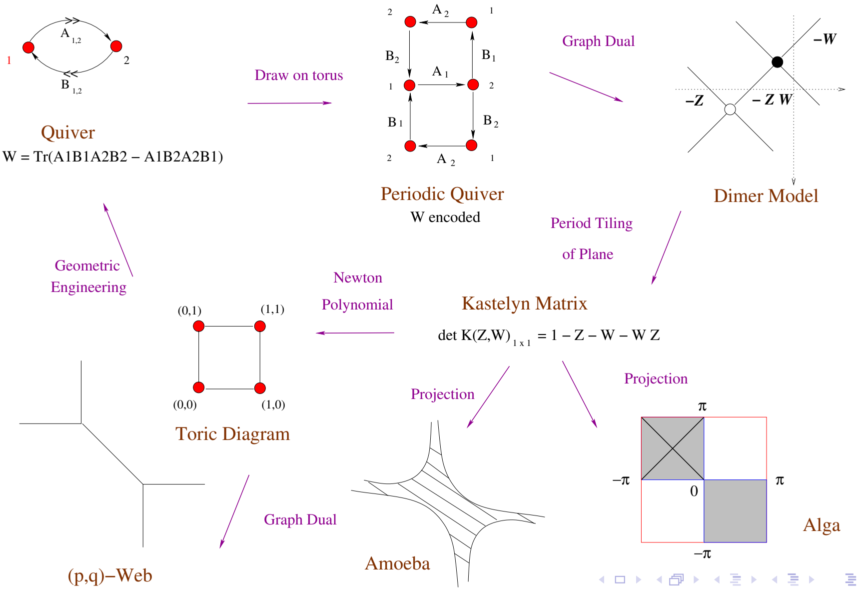
- Mirror Symmetry [Strominger-Yau-Zaslow; Hori-Vafa]
  D3-brane on CY3 $\rightsquigarrow$ D6-branes wrapping 3-cycles in mirror CY3
- [Feng-Kennaway-YHH-Vafa] torus T^2 lives in T^3 of mirror symmetry;
  Tropical Geometry
- THEOREM: [R. Böckland, N. Broomhead, A. Craw, A. King, K. Ueda . . .]
  The (coherent component of) VMS as representation variety of a quiver is an affine (non-compact, possibly singular) toric Calabi-Yau variety of complex dimension 3 $\Leftrightarrow$ the quiver + superpotential is graph dual to a bipartite graph drawn on T^2
- Rmk: Each $\Rightarrow$ SCFT in 3+1-d

SUMMARY: $\mathbb{C}^3$, Hexagonal Tilings, SYM

$\mathcal{N}=1$ SYM = D3-branes transverse to $\mathbb{C}^3 = \mathcal{C}(S^5)$ = hexagonal bipartite tiling



SUMMARY: Conifold and Square Tilings



The String Landscape

Vacuum Degeneracy

Perhaps the biggest theoretical challenge to string theory:

selection criterion??? metric on the landscape???

- Douglas (2003): Statistics of String vacua
- Kachru-Kallosh-Linde-Trivedi (2003): type II/CY estimates of 10^{500}
- Taylor-YN Wang (2015-7): F-theory estimates 10^{3000} to 10^{10^5}
- Basically: Combinatorial geometry usually tends exponentially

  e.g., Kreuzer-Skarke (2000s): Reflexive polytopes up to $SL(n;\mathbb{Z})$:

  1, 16, 4319, 473800776, ???

  Altman-Carifio-Halverson-Nelson (2018): estimated 10^{10^4} triangulations

  Altman-Gray-YHH-Jejjala-Nelson (2014): brute-force: $\sim 10^6$ up to $h^{1,1} = 6$

Searching the Standard Model

SM places some constraints but still not enough:

- Braun-YHH-Ovrut; Bouchard-Cvetic-Donagi (2005): exact MSSM particles
- Gmeiner-Blumenhagen-Honecker-Lüst-Weigand (2005):1 in 10^9 in D-brane MSSM modles
- Candelas-de la Ossa-YHH-Szendroi (2007): Triadophilia $\Rightarrow$ "des res"?
- Anderson-Gray-Lukas-Palti (2012-3): Het line bundle MSSM: 200 in 10^{10}

Recent estimates

- Constatin-YHH-Lukas; Deen-YHH-SJ Lee-Lukas (2018-9) MSSM from heterotic line bundles: 10^{23} from CICYs; 10^{723} from KS
- Cvetic-Halverson-Lin-Liu-Tian (2019): 10^{15} F-theory MSSMs

WWJD

What Would JPython/AI Do?

YHH, 1706.02714, PLB 774, 2017

(Feature article, M. Hutchinson, *Science*, Vol 365, July, 2019)

SUMMARY: Algorithms and Datasets in String Theory

- Growing databases and computational algorithms motivated by string theory
- Archetypical Problems
  - Classify configurations (typically integer matrices: polyotope, adjacency, ...)
  - Compute geometrical quantity algorithmically
    - toric $\rightsquigarrow$ combinatorics;
    - quotient singularities $\rightsquigarrow$ rep. finite groups;
    - generically $\rightsquigarrow$ ideals in polynomial rings;
    - Numerical geometry (homotopy continuation);
    - Cohomolgy (spectral sequences, Adjunction, Euler sequences)
- Typical Problem in String Theory/Algebraic Geometry:

$$\underset{\boxed{\text{integer tensor}}}{\overset{INPUT}{}} \longrightarrow \underset{\boxed{\text{integer}}}{\overset{OUTPUT}{}}$$

Where we stand . . .

The Good — Last 10-15 years: several international groups have bitten the bullet

Oxford, London, Vienna, Blacksburg, Boston, Johannesburg, Munich, . . .

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 MSSM not possible . . .

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

A Wild Question

- Typical Problem in String Theory/Algebraic Geometry:

$$\overset{INPUT}{\boxed{\text{integer tensor}}} \longrightarrow \overset{OUTPUT}{\boxed{\text{integer}}}$$

- Q: Can (classes of problems in computational) Algebraic Geometry be "learned" by AI ? , i.e., can we "machine-learn the landscape?"
- 1706.02714 Deep-Learning the Landscape, PLB 774, 2017:

  Experimentally, it seems to be the case for many situations

2017: String Theory enters the Machine-Learning Era

YHH (1706.02714);

Krefl-Seong (1706.03346);

Ruehle (1706.07024)

Carifio-Halverson-Krioukov-Nelson (1707.00655)

Sophia: Hanson Robotics, HongKong

- Beginning of String_Data Progress in String Theory
- How can ML and modern data-science help with the vacuum degeneracy problem??
- Meanwhile ... Sophia becomes a "human" citizen (in Saudi Arabia)

2017: String Theory enters the Machine-Learning Era

YHH (1706.02714);

Krefl-Seong (1706.03346);

Ruehle (1706.07024)

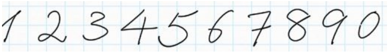
Carifio-Halverson-Krioukov-Nelson (1707.00655)

Sophia: Hanson Robotics, HongKong

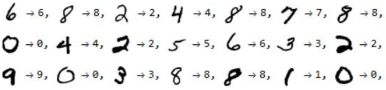
- Beginning of String_Data Progress in String Theory
- How can ML and modern data-science help with the vacuum degeneracy problem??
- Meanwhile . . . Sophia becomes a “human” citizen (in Saudi Arabia)

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? <u>ALL are inefficient and too sensitive to variation.</u>
- What does your iPhone/tablet do? What does Google do? Machine-Learn
  - Take large sample, take a few hundred thousand (e.g. NIST database)



$\ldots$

$28 \times 28 \times (RGB)$

Part II: Machine Learning

A Brief Introduction to the Novice

Back in Kindergarten...

- Given a set of data-points (now called point cloud $\vec{x}_i \in \mathbb{R}^n$), we were taught to do 2 types of things
  1. Plot them and see if there are any patterns (if n is small), known distribution? components?
  2. Consider $n-1$ as independent variables and 1 as dependent, find best-fit function $x_n = f(x_{i=1,\ldots,n-1})$ by regression (typically linear) （線性）回歸 .
- Now, we have more sophisticated generalizations/names:
  1. Unsupervised machine-learning 非監督機器學習
  2. Supervised machine-learning 監督性機器學習

A Long History (contrary to what you might think)

(cf. Goodfellow, Bengio, Courville, "Deep-Learning", 2006, MIT Press [GBC])

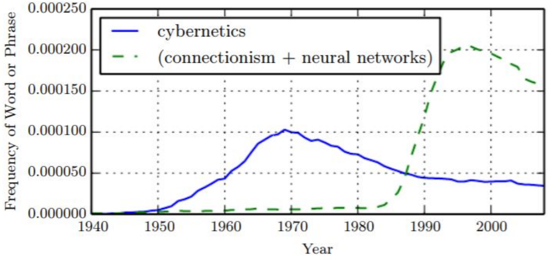
- 1940 - 60: Cybernetics 控制論

  The Perceptron 感知器 1957 (!!) in early AI (using CdS photo-cells)

- 1980 - 90: Connectionism 聯合主義

  (Artificial) Neural Networks (NN) （人工）神經網絡

- 2006: Deep Learning 深度（機器）學習



(Fig 1.7 of [GBC]; from all texts on *Googlebooks*)

A Key Mathematical Tool

- Gradient/Steepest Descent 梯度下降優化 : Find the (local) minimum $\vec{x}^*$ of a function $f(\vec{x})$ [Cauchy, 1847]

$$\vec{x}_{n+1} := \vec{x}_n - \epsilon \nabla f(\vec{x}_n) \ , \qquad \text{iterate } n = 1, 2, 3, \ldots$$

  $f(\vec{x}_{n+1}) \leq \vec{x}_n$; learning rate ϵ and initial value $\vec{x}_0$ are hyper-parametres

- Stochstatic Gradient Descent 隨機梯度下降 : Typically f is a cost function
  - of form $f = \sum\limits_{\mathcal{D}} f_i$ summed over the data $\mathcal{D}$ where $|\mathcal{D}|$ is huge
  - Take random samples $\mathcal{D}' \subset \mathcal{D}$ and sum over $\mathcal{D}'$: mini- batch size $|\mathcal{D}|$

Basic Types

- Discrete — Classifier
  Continuous — Regressor

- Unsupervised
  - Clustering (e.g., nearest neiboughrs, k-Means, . . .)
  - Autoencoders
  - GAN (Generative Adversarial Networks)
  - PCA (Principal Component Analysis) PCA . . .

  Supervised (labeled data)
  - Perceptron
  - SVM Support Vector Machine
  - Neural Network
  - Bayesian Classifiers, Decision Trees, . . .

A Single Neuron: The Perceptron 神經元：感知器

- 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;
- Given Training data: $\mathcal{D} = \{(x_i^{(j)}, d^{(j)}\}$ with input x_i and **known output** $d^{(j)}$, minimize some cost/loss function to find optimal w_i and $b \leadsto$ "learning", then check against Validation Data
- Just (non-linear) regression
- supervision because of association (teaching) $x_i \to d$

Common Activation Functions 激活函數

- Logistic Sigmoid: $(1+e^{-x})^{-1}$
- Hyperbolic tangent: $\tanh(x) = \frac{e^x+e^{-x}}{e^x-e^{-x}}$
- Softplus: $\log(1+e^x)$, a "softened" version of ReLu (Rectified Linear Unit): $\max(0,x)$
- Softmax: $x_i \to \frac{e^{x_i}}{\sum_i e^{x_i}}$
- Parametric ReLu: $R(x) = \begin{cases} x\ , & x \geq 0 \\ \alpha x\ , & x < 0 \end{cases}$
- Maxout: $x_i \to \max_i x_i$
- Linear/Identity: $x_i \to x_i$

  (rmk: the weights and bias will make it $x_i \to w_{ij}x_j + b_i$)

Common Cost/Loss Functions 代價函數

(supervised) dataset $\mathcal{D} = \{x_i^{(j)} \longrightarrow d^{(j)}\}_{j=1,2,\ldots,N}$

Training set and validation set: $\mathcal{D} = \mathcal{T} \sqcup \mathcal{V}$, $|\mathcal{T}| = n, \ |\mathcal{V}| = N - n$

Best-fit function/predictor $f(x)$ trained on $\mathcal{T}$:

- When output is continuous (best-fit function), typically use SEL (squared-error-loss)

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

- When output is discrete (categorical classification problem), typically use XC (cross-entropy)

$$XC := -\frac{1}{n} \sum_j \left[d^{(j)} \log f(x^{(j)}) + (1 - d^{(j)}) \log(1 - f(x^{(j)})) \right]$$

Measures for Goodness of Fit/Performance: Continuous

On validation dataset $\mathcal{V} = \{x_i^{(j)} \longrightarrow d^{(j)}\}_{j=1,2,\ldots,m=N-n}$;

Predicted values: $\{x_i^{(j)} \longrightarrow \hat{d}^{(j)}\}_j$

- Need to compare $\hat{d}$ and d pairwise and have a measure of how good the predictor is 決定係數

Coefficient of Determination $\quad R^2 := 1 - \dfrac{SS_{\text{res}}}{SS_{\text{tot}}}$

$$\text{Data Variance} = SS_{\text{tot}} := \sum_j (d^{(j)} - \overline{d^{(j)}})^2, \quad \overline{d^{(j)}} := \text{mean}$$

$$\text{Residual sum of squares} = SS_{\text{res}} := \sum_j (d^{(j)} - \hat{d}^{(j)})^2,$$

- bad fit $= 0 \leq R^2 \leq 1 =$ perfect fit
- Also do a scatter-plot of $(d^{(j)}, \ \hat{d}^{(j)})$, needs to be close to $y = x$ line

Measures for Performance: Discrete

- Categorial Classification (K classes)
- When $K = 2$, called binary classification, denote $d_i \in \{0, 1\}$
  - confusion matrix: $C :=$

		Actual	
		True (1)	False (0)
Predicted Classification	True (1)	True Positive (tp)	False Positive (fp)
	False (0)	False Negative (fn)	True Negative (tn)

  - True/False positive rate TPR/FPR:

$$\text{TPR} := \frac{tp}{tp+fn} \ , \qquad \text{FPR} := \frac{fp}{fp+tn} \ ,$$
$$\text{Accuracy} \ \frac{tp+tn}{tp+tn+fp+fn} \ , \qquad \text{Precision} := \frac{tp}{tp+fp} \ .$$

- want accuracy (% agreement) and precision to be close to 1 but these are not good enough in discounting fp and fn.

Further Performance Measure: Discrete

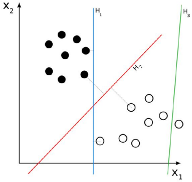
- F1-Score $F := \frac{2}{\frac{1}{TPR} + \frac{1}{Precision}} \in [0,1]$
  - Harmonic mean between true positives and precision
  - closer to 1 the better the prediction
- Matthews' Correlation Coefficient
  $\phi := \sqrt{\frac{\chi^2}{m}} = \frac{tp \cdot tn - fp \cdot fn}{\sqrt{(tp+fp)(tp+fn)(tn+fp)(tn+fn)}} \in [-1,1]$
  - -1 anti-correlation; 0 random; 1 perfect correlation
  - generalize to K-category classification (for $K \times K$ confusion matrix)

$$\phi := \frac{\sum\limits_k \sum\limits_l \sum\limits_m C_{kk}C_{lm} - C_{kl}C_{mk}}{\sqrt{\sum\limits_k (\sum\limits_l C_{kl})(\sum\limits_{k'|k' \neq k} \sum\limits_{l'} C_{k'l'})} \sqrt{\sum\limits_k (\sum\limits_l C_{lk})(\sum\limits_{k'|k' \neq k} \sum\limits_{l'} C_{l'k'})}}$$

- rmk: everything so far, perceptron included, is just old-fashioned regression

Support Vector Machines 向量支持器

- a classic example of supervised learning: find hyperplanes which separate labeled categories
- e.g., binary classification: given $(\vec{x}_i \to y_i)_{i=1,\dots,N}$ with $\vec{x}_i \in \mathbb{R}^n$, $y_i = \pm 1$



- find 2 hyperplanes so that

  $\vec{x}_i \cdot \vec{w} + b \geq 1$ if $y_i = 1$;

  $\vec{x}_i \cdot \vec{w} + b \leq -1$ if $y_i = -1$

- distance between 2 hyperplanes is $2/\|\vec{w}\|$, which we need to maximize

- i.e., have optimization problem: (combining the 2 hyperplanes)

$$\min_{\vec{w},b} \frac{1}{2} \|\vec{w}\|^2 , \quad \text{constraint: } y_i(\vec{x}_i \cdot \vec{w} + b) \geq 1$$

Support Vector Machines

- Solution: $\vec{w} = \sum_i \alpha_i y_i \vec{x}_i$ for some $\alpha_i \in \mathbb{R}$ such that

  $\alpha_i \neq 0$ only for $\vec{x}_i$ on the margins of hyperplane

  which are the support vectors
- Generalizations
  - In case not separable, add slack:

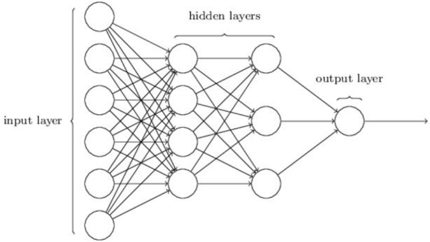
    $\min_{\vec{w},b} \frac{1}{2} \|\vec{w}\|^2 + c \sum_i \xi_i,$ constraint: $y_i(\vec{x}_i \cdot \vec{w} + b) + \xi_i \geq 1, \xi_i \geq 0;$
  - In case not linear/hyperplane, add kernel: SVM hyperplane replaced by

$$\sum_i \alpha_i y_i K(\vec{x}, \vec{x}_i) + b = 0 \ ,$$

common kernel, Gaussian $K(s,t) = \exp(-\gamma \|s-t\|^2)$

Multi-Layer Perceptron (MLP) 多層感知器

- MAGIC: put many neurons together and let connectionism do the magic



Simplest case:

forward directed only,

called multilayer perceptron

or feedforward Neural Network

前饋神經網絡

- Typical layers (depth = # layers (hence the name deep learning)):
  - (fully-connected) linear layer from $m \to n$ nodes: $m \times n$ matrix of linear fnc
  - node-wise activation function (from the list before)
  - summation layer
- Width: $\sim$ # neurons per layer

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 $\int_{\mathbb{R}^n} |f(x) - F(x)| dx < \epsilon$.

Some Technicalities

(*Implemented on Mathematica 11.1 + / TensorFlow-Keras on Python*)

- Regularization 正規化
  - if it performs well on $\mathcal{T}$ but not so well on $\mathcal{V}$, possible overfitting
  - L1 or L2 Reg: add $\lambda \|w\|^{i=1,2}$ to cost function to ensure weight doesn't become too large
  - Dropout: randomly delete neurons
  - Data Enhancement: add equivalent representations of the training data (e.g., Cayley table of finite group, add any row/column permutation)
  - early stoppping: if validation error gets increasingly worse, stop training
- In computing gradient descent for layer i, backward propagation 反向傳播: reduces computation for ∇f_i, i.e., *chain rule* $\nabla(f_i(g_{i-1})) = \nabla f_i(\nabla g_{i-1})$

Beyond MLP: Two important NN Types

- CNN (convulutional NN) 卷積神經網絡
  - perfect for image processing:

    Convolution Layer $\to$ Non-linear Layer $\to$ Pooling Layer
  - Convolution: $(L \star K)_{i,j} = \sum\limits_{m,n} L_{i+m,j+n} K_{m,n}$
  - Pooling: compare neighbours, e.g., $\max_{i,j=m,n\pm 1} L_{i,j}$
- RNN (recurrence NN) 循環神經網絡
  - perfect for series prediction
  - essentially MLP + arrows going backwards so that outputs of one layer can be fed back $\rightsquigarrow$ memory
- general NN a mixture of MLP, CNN, NN, and indeed any direct graph of neurons.

Deep-Learning the String Landscape

Warmup: Hypersurfaces in $W\mathbb{P}^4$

Simplest Data Structure $[w_1 : w_2 : w_3 : w_4 : w_5] \longrightarrow h^{1,1}$

[YHH 1706.02714] Oftentimes, questions in pheno are **qualitative**, e.g.,

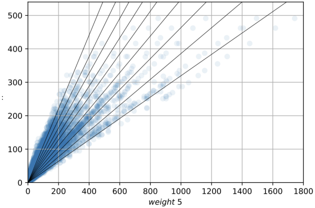
- large # complex structure how many have, say, $h^{2,1} > 50$?

  [Candelas-Lynker-Schimmrigk] Landau-Ginzburg methods: many hours; using Euler sequence/Adjunction Distributions

- Standard method: take partial training and validation data, s.t., $D = T \sqcup V$
  - train NN with random 2000/7555 inputs ($\sim 1/4$ only)
  - use the trained NN to predict value for the remaining UNSEEN 7555 - 2000
  - Get $\sim 91.8\%$ precision, **Cosine Distance** $d_C = 0.91$, **Matthew Coefficient** $\phi = 0.84$ in less than 20 sec on regular laptop! Learning Curve Training Curve

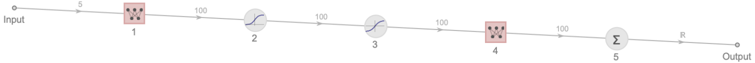
Detailed Study: Berman-YHH-Hirst 2112.06350

- clustering shows that the most significant dependence is on w_5



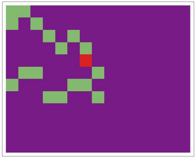
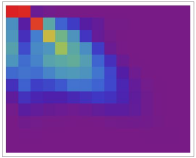
- 5-fold cross-validation on predicting $h^{1,1}$ from w_i gets $R^2 > 0.95$

- Simple architecture of NN: e.g., 5-layer MLP



CICYs: a Colourful Example

- An image = a matrix (pixels) with entries denoting shade/colour; NN really good at images (e.g. hand-writing) [RMK: not using a convolutional NN here]
- CICY is a (padded) 12×15 matrix with 6 colours $\rightsquigarrow$ CICY is an image

(a)  (b) 

(a) typical CICY;

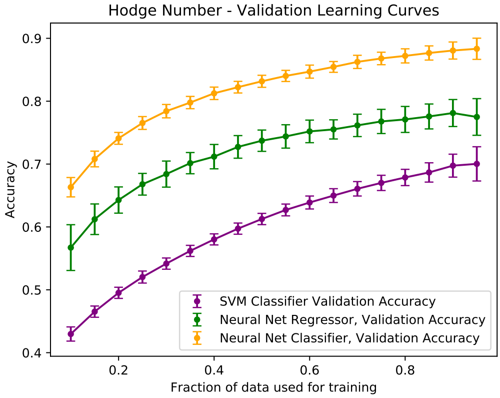
(b) average CICY

- Initial binary classifier e.g. in learning large number of Kahler parametres $h^{1,1} > 5$: learns 4000 samples ($< 50\%$) in ~ 5 min; validate against 7890-4000: 97% accuracy, $d_C = 0.98$, $\phi = 0.87$.

CICYs: Detailed Analysis

Bull-YHH-Jejjala-Mishra (1806.03121, 1903.03113)

- **TensorFlow** Python's implementation of NNs and DL
- Compare NNs with Decision Trees, Support Vector Machines, etc

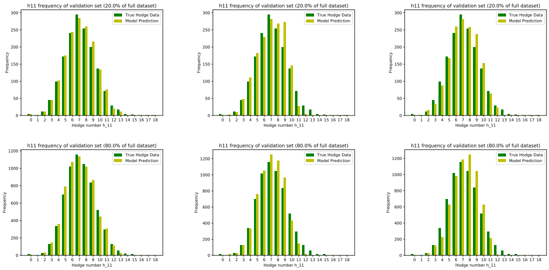


Can one learn the FULL information on Hodge numbers?

$h^{1,1} \in [0, 19]$ so can set up 20-channel NN classifer, regressor, as well as SVM

CICYs: Comparative Studies

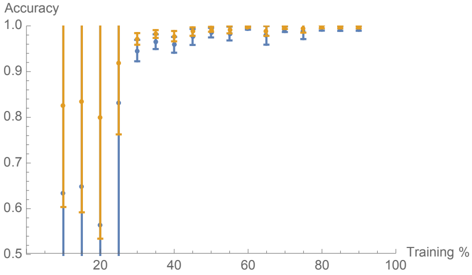
$h^{1,1}$ for NN, Regressor, SVM at 20 and 80% training



Massive improvement: Krippendorf-Syvaeri [2003.13679] Erbin-Finotello (2007.13379; 2007.15706 Google Inception NN) YHH-Lukas [2009.02544] Larfors-Lukas-Ruehle-Schneider (2111.01436); Erbin-Finotello-Schneider-Tamaazousti (2108.02221) $> 99.96\%$ precision using more sophiscated NN (e.g., Google Inception CNN)

Distinguishing Elliptic Fibrations

- [YHH-SJ. Lee 1904.08530]: test in CICY which are elliptically fibred (bypass Oguiso-Kollar-Wilson Theorem/Conjecture)
- Explicit computation by finding divisor D by Anderson et al. very expensive; AI achieves in seconds:



- A control test: let a random set have property "1" and complementary set, "0", get 50% precision and $\phi \sim 0$ (complete guessing)

Take-Home Lesson

- GENERAL: ANY algebraic variety can be represented as a tensor and hence pixelated image
- much of computational algebraic geometry = no different than an image-recognition problem
- all of (computational) algebraic geometry = finding (co-)kernels of integer matrices: thus is perfectly adapt for ML

String/Algebraic Geometry: 2018-

- q.v., Bundle Cohomology (Ruehle, Brodie-Constantin-Lukas, Larfors-Schneider, Otsuka-Takemoto, Klaewer-Schlechter)
- q.v., Kreuzer-Skarke Dataset (Halverson, Long, Nelson; McCallister-Stillman, Berglund-Campbell-Jejjala)
- q.v., Calabi-Yau volumes in AdS/CFT (Krefl-Seong)
- q.v., MSSM from orbifold models (Parr-Vaudrevange-Wimmer)
- q.v. Particle Masses Gal-Jejjala-Pena-Mishra ...
- q.v. Knot invariants: Jejjala-Kar-Parrikar, Craven-Jejjala-Kar Gukov-Halverson-Ruehle-Sułkowski, using NLP

String/Algebraic Geometry: 2018-

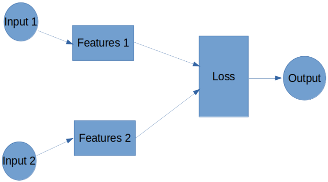
- q.v. Ashmore-YHH-Ovrut+Calmon, Douglas-Lakshminarasimhan-Qi, Jejjala-Pena-Mishra, Anderson-Gerdes-Gray-Krippendorf-Raghuram Numerical CY Metrics
- Otsuka-Takemoto; Deen-YHH-Lee-Lukas Distinguishing Heterotic SMs
- q.v. DEEP CONNECTIONS
  - K. Hashimoto: AdS/CFT = Boltzmann Machine;
  - Halverson-Maiti-Stoner: QFT = NN;
  - de Mello-Koch: NN = RG;
  - Vanchurin 2008: Universe = NN.
- What about the vacuum degeneracy problem?

One-Shot Learning

Fei-Fei Li et al. (2002 -)

- Estimated that by 6, a child has learnt all $10 \sim 30 \times 10^3$ object categories

  NOT done by sampling % of cases in each category
- could not have supervise learnt everything in the standard way!
- Knowledge Transfer: having seen lots of horses and a single bird, would recognize a chicken is closer to a bird than to a horse
- a SINGLE representative in a category suffices, or at most a handful $\rightsquigarrow$ Few-Shot Learning



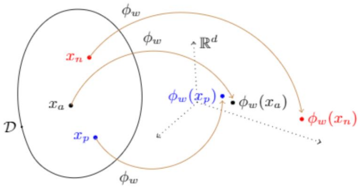
Loss $= \mathcal{L}(w) :=$

$$\max\{d_w(x_a, x_p) - d_w(x_a, x_n) + 1, 0\}$$

$$d_w(x_1, x_2) := (\phi_w(x_1) - \phi_w(x_2))^2$$

ϕ representation by features network (FN)

FN: represents the data by mapping to $\mathbb{R}^3$, say: $\phi : \mathcal{D} \to \mathbb{R}^3$:



a anchor point for the class;

p close-by; n far-apart

FN some appropriately chosen NN

SNN returns a similarity score $\in [0, \infty)$

where 0 means identical

CICYs as Representative Landscape

CICY3
- classified by Candelas, Dale, Green, Hubsch, Lutken (1988-9)
- 7890 configurations, $h^{1,1} \in [1, 19]$; $h^{2,1} \in [15, 101]$

  (m, K) ranges from $(1, 1)$ to $(12, 15)$

CICY4
- classified by Gray, Haupt, Lukas (2013-4)
- 905684 configurations, $h^{1,1} \in [1, 24]$; $h^{2,1} \in [1, 33]$;

  $h^{3,1} \in [20, 426]; h^{2,2} \in [204, 1752]$

  (m, K) ranges from $(1, 1)$ to $(16, 20)$

Can we One-Shot learn the String Landscape?

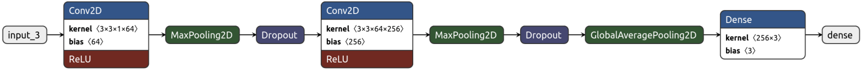
q.v. 2111.04761, YHH, Shailesh Lal, M. Zaid Zas

Methodology

Labelled Data of the form $(q^i_j) \longrightarrow h^{1,1}$ where similarity is

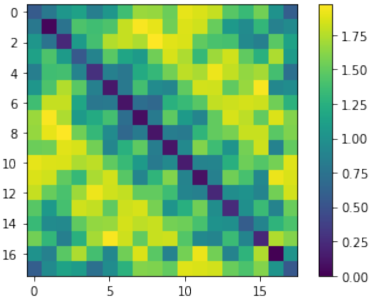
$$q^{(A)} \sim q^{(B)} \quad \text{iff} \quad h^{(A)} = h^{(B)}$$

- Represent each CICY as pixelated image (after normalization), and use CNN as FN (tried other architectures like Inception and MLP):

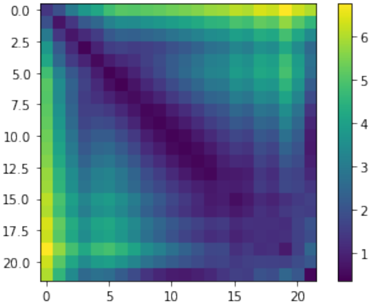


- trained on 3% of CICY3 and 0.6% of CICY4 (mostly just few per class of $h^{1,1}$): Few-Shot ML hundreds to extrapolate to hundreds of thousands
- Standard ADAM optimizer @ learning-rate of 0.01

Mean Similarity Scores on Pairs

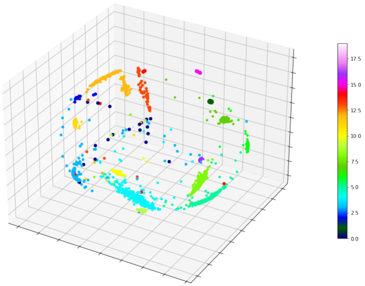


$CICY3$

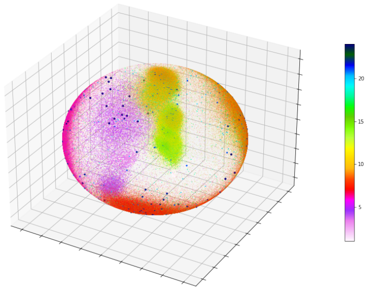


$CICY4$

Clustering of CICY by $h^{1,1}$? ...



$CICY3$



$CICY4$

- Two-birds with one stone
  1. Few-shot ML of the landscape
  2. The similarity score gives a distance measure on the landscape
- This reduction + distance: a step toward a vacuum selection principle given the complexity of the landscape

Machine Learning Mathematics

Why stop at string/geometry?

q.v. Review YHH 2101.06317

A Zealot's Perspective

Q: We have seen that algebraic geometry (over $\mathbb{C}$) is a tensor manipulation / image recognition problem,

how much of mathematics is not?

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

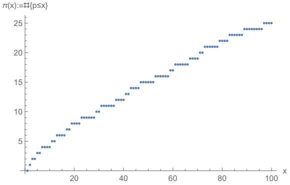
We can call this Bottom-up Mathematics

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

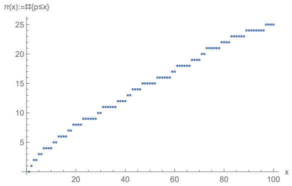


(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

- 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

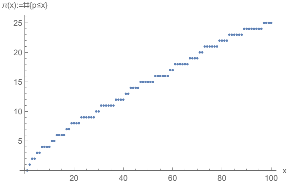


(w/o computer and before complex analysis [50 years before Hadamard-de la Vallée-Poussin's proof]): PNT $\pi(x) \sim x/\log(x)$

- 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
  - 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 ML Regressor & Classifiers (w/ NO KNOWLEDGE of the maths)

- 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 (?)

will give Status Report of Experiments in the last couple of years

- focus on supervised ML ("knows where to get to")
- 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,]
  - 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 $\to$ length of irrep decomp / tensor product: (0.97, 0.93); (train on small dim, predict high dim: (0.9, 0.8))

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 (or $x/\log(x)$-scaled) window of $(\text{yes/no})_{1,2,\ldots,k}$ to $(\text{yes/no})_{k+i}$ for some i (in binary); 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]

  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]

An Inherent Hierarchy?

- In decreasing precision/increasing difficulty:

numerical

string theory $\rightarrow$ algebraic geometry over $\mathbb{C}$ $\sim$ arithmetic geometry

algebra

string theory $\rightarrow$ combinatorics

analytic number theory

- **Categorical Theory**
  - suggested by & in prog. w/ B. Zilber, Merton Prof. of Logic, Oxford
  - 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.

Meta-mathematics/physics?

[YHH-Jejjala-Nelson] "hep-th" 1807.00735

- Word2Vec: [Mikolov et al., '13] NN which maps words in sentences to a vector space **by context** (much better than word-frequency, quickly adopted by Google); maximize (partition function) over all words with sliding window ($W_{1,2}$ weights of 2 layers, C_α window size, D # windows)

$$Z(W_1, W_2) := \frac{1}{|D|} \sum_{\alpha=1}^{|D|} \log \prod_{c=1}^{C_\alpha} \frac{\exp([\vec{x}_c]^T \cdot W_1 \cdot W_2)}{\sum\limits_{j=1}^{V} \exp([\vec{x}_c]^T \cdot W_1 \cdot W_2)}$$

- We downloaded all $\sim 10^6$ titles of hep-th, hep-ph, gr-qc, math-ph, hep-lat from ArXiv since the beginning (1989) till end of 2017 Word Cloud
  (rmk: Ginzparg has been doing a version of linguistic ML on ArXiv)
  (rmk: abs and full texts in future)

Subfields on ArXiv has own linguistic particulars

- Linear Syntactical Identities

  bosonic + string-theory = open-string

  holography + quantum + string + ads = extremal-black-hole

  string-theory + calabi-yau = m-theory + g2

  space + black-hole = geometry + gravity . . .

- binary classification (Word2Vec + SVM) of formal (hep-th, math-ph, gr-qc) vs phenomenological (hep-ph, hep-lat) : 87.1% accuracy (5-fold classification 65.1% accuracy). ArXiv classifications

- Cf. **Tshitoyan et al.**, "Unsupervised word embeddings capture latent knowledge from materials science literature", **Nature** July, 2019: 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

Please submit

- Special Collection in AACA, Birkhäuser, *Dechant, YHH, Kaspryzyk, Lukas, ed*:
  `https://www.springer.com/journal/6/updates/18581430`
- Special Volume in JSC, Springer, *Hauenstein, YHH, Kotsireas, Mehta, Tang, ed.*
  `https://www.journals.elsevier.com/journal-of-symbolic-computation/call-for-papers/algebraic-geometry-and-machine-learning`
- ML in theoretical physics & pure maths, Book, WS, *YHH, ed.*
- Int. J. Data Science in the Mathematical Sciences, WS, *YHH et al., ed.*

THANK YOU

Go and try your favourite problem

Some Rudiments & Nomenclature

A sequence of specializations:

- M Riemannian: positive-definite symmetric metric
- M Complex Riemannian: have (p,q)-forms with p-holomorphic and q-antiholomorphic indices: $d = \partial + \bar{\partial}$ (with $\partial^2 = \bar{\partial}^2 = \{\partial, \bar{\partial}\} = 0$)
- M Hermitian: complex Riemannian and can tranform $g_{mn} = g_{\bar{m}\bar{n}} = 0$
- M Kähler: Hermitian with Kähler form $\omega := i g_{m\bar{n}} dz^m \wedge dz^{\bar{n}}$ such that $d\omega = 0$ ($\Rightarrow \partial_m g_{n\bar{p}} = \partial_n g_{m\bar{p}}$; $g_{m\bar{n}} = \partial\bar{\partial} K(z,\bar{z})$ for some scalar K)

Cohomology:

- On Riemannian M: can define Laplacian on p-forms (Hodge star $\star(dx^{\mu_1} \wedge \ldots \wedge dx^{\mu_p}) := \frac{\epsilon^{\mu_1 \cdots \mu_n}}{(n-p)!\sqrt{|g|}} g_{\mu_{p+1}\nu_{p+1}} \cdots g_{\mu_n \nu_n} dx^{\nu_{p+1}} \wedge \ldots \wedge dx^{\nu_n}$)

$$\Delta_p = dd^\dagger + d^\dagger d = (d + d^\dagger)^2, \qquad d^\dagger := (-1)^{np+n+1} \star d \star$$

  Harmonic p-Form $\Delta_p A^p = 0 \overset{1:1}{\longleftrightarrow} H^p_{deRham}(X)$

- On Hermitian M: Dolbeault Cohomology $H^{p,q}_{\bar\partial}(X)$: cohomology on $\bar\partial$ (similarly ∂) and $\Delta_\partial := \partial\partial^\dagger + \partial^\dagger\partial$ and similarly $\Delta_{\bar\partial}$

- On Kähler M: $\Delta = 2\Delta_\partial = 2\Delta_{\bar\partial}$, Hodge decomposition:

$$H^i(M) \simeq \bigoplus_{p+q=i} H^{p,q}(M)$$

Back to Calabi-Yau

Covariant Constant Spinor

- Define $J_m^n = i\eta_+^\dagger \gamma_m^n \eta_+ = -i\eta_-^\dagger \gamma_m^n \eta_-$, check: $J_m^n J_n^p = -\delta_m^n$
- (X^6, J) is thus almost-complex
- But η covariant constant $\rightsquigarrow \nabla_m J_n^p = 0 \rightsquigarrow \nabla N_{mn}^p = 0$

  Nijenhuis tensor $N_{mn}^p := J_m^q \partial_{[q} J_{n]}^p - (m \leftrightarrow n)$
- (X^6, J) is thus complex ($J_m^n = i\delta_m^n,\ J_{\bar{m}}^{\bar{n}} = i\delta_{\bar{m}}^{\bar{n}},\ J_{\bar{m}}^n = J_m^{\bar{n}} = 0$ for some local coordinates $(z, \bar{z})$; transition functions holomorphic)
- Define $J = \frac{1}{2} J_{mn} dx^m \wedge dx^n$ ($J_{mn} := J_m^k g_{kn}$) check:

  $dJ = (\partial + \bar{\partial})J = 0$
- (X^6, J) is thus Kähler
- summary $\boxed{X^6 \text{ is a Kähler manifold of } \dim_{\mathbb{C}} = 3, \text{ with } SU(3) \text{ holonomy}}$

Back to Het

Famous CICYs

- The Quintic $Q = [4|5]^{1,101}_{-200}$ (or simply $[5]$);
- Yau-Tian Manifold: $TY = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}^{14,23}_{-18}$
  - no CICY has $\chi = \pm 6$
  - TY has freely-acting $\mathbb{Z}_3 \rightsquigarrow (TY/\mathbb{Z}_3)^{6,9}_{-6}$;
  - central to early string pheno [Distler, Greene, Ross, et al.]
- Schön Manifold: $S = \begin{pmatrix} 1 & 1 \\ 3 & 0 \\ 0 & 3 \end{pmatrix}^{19,19}_{0}$ has $\mathbb{Z}_3 \times \mathbb{Z}_3$ freely acting symmetry
  - explored more recently;
  - The quotient is $M^0_{3,3}$.

Back to CICYs

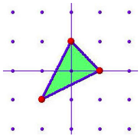
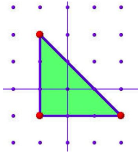
Reflexive Polytopes: Rudiments

- Convex Lattice Polytope Δ (use Δ_n to emphasize dim n)
  - DEF1 (Vertex Rep): Convex hull of set S of k lattice points $p_i \in \mathbb{Z}^n \subset \mathbb{R}^n$

$$\mathrm{Conv}(S) = \left\{ \sum_{i=1}^{k} \alpha_i p_i | \alpha_i \geq 0, \ \sum_{i=1}^{k} \alpha_i = 1 \right\}$$

  - DEF2 (Half-Plane Rep): intersection of integer inequalities $A \cdot \underline{x} \geq \underline{b}$
  - {extremal pts = vertices, edges, 2-faces, 3-faces, ..., (n-1)-faces = facets, Δ}
  - $n = 2$ polygons, $n = 3$ polyhedra, ...
- Polar Dual: $\Delta^\circ = \{\underline{v} \in \mathbb{R}^n \mid \underline{m} \cdot \underline{v} \geq -1 \ \ \forall \underline{m} \in \Delta\}$
- Reflexive Δ: if Δ° is also convex lattice polytope
  - in general, vertices of Δ° are rational, not integer
  - duality: $(\Delta^\circ)^\circ = \Delta$
  - if further $\Delta = \Delta^\circ$, self-dual/self-reflexive

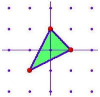
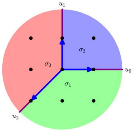
Reflexive Polytope: example

Δ_2		Vertices : $(1,0),\ (0,1),\ (-1,-1)$ Facets : $\begin{cases} -x-y \geq -1 \\ 2x-y \geq -1 \\ -x+2y \geq -1 \end{cases}$
Δ_2°		Vertices : $(-1,2),\ (-1,-1),\ (2,-1)$ Facets : $\begin{cases} -x-y \geq -1 \\ x \geq -1 \\ y \geq -1 \end{cases}$

THM: Reflexive $\Leftrightarrow$ single interior lattice point

(set to origin; all facets = hyperplanes of distance 1 away)

Toric Variety from Δ_n

- Face Fan $\Sigma(\Delta) \equiv \{\sigma = \text{pos}(F) \big| F \in \text{Faces}(\Delta)\}$ with
  $\text{pos}(F) \equiv \left\{\sum_i \lambda_i \underline{v}_i \big| \underline{v}_i \in F \ , \ \lambda_i \geq 0\right\}$
- e.g. $\Delta_2 =$  $\Rightarrow \Sigma(\Delta_2) =$ 

- $\Sigma(\Delta_n)$ then defines a compact Toric variety $X(\Delta_n)$ of $\dim_{\mathbb{C}} = n$
- $X(\Delta)$ called Gorenstein Fano, i.e., $-K_X$ is Cartier and ample, i.e., $\mathcal{O}(-K_X)$ is line bundle and X is positive curvature
- THM: $X(\Delta)$ smooth $\Leftrightarrow$ generators of every cone σ is part of $\mathbb{Z}$-basis, i.e., $\det(\text{gens}(\sigma)) = \pm 1$ Back to KS CY3

Observatio Curiosa

- Penn group *purely abstract*, but $X_0^{19,19} = \begin{pmatrix} 1 & 1 \\ 3 & 0 \\ 0 & 3 \end{pmatrix}$, Tian-Yau: $\begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$
- TRANSPOSES!!
- Why should the best manifold from 80's be so-simply related to the best manifold from completely different data-set and construction 20 years later ??
- Two manifolds are conifold transitions and vector bundles thereon transgress to one another ([Candelas-de la Ossa-YHH-Szendroi, 2008])
- **Connectedness of the Heterotic Landscape**
    - All CICY's are related by conifold transitions
    - Reid Conjecture: All CY3 are connected
    - Proposal: All (stable) vector bundles on all CY3 transgress

Back to Compactifications

A Computational Approach

- Northeastern/Witts/Notre Dame/Cornell Collaboration: Programme to study the computational algebraic geometry of $\mathcal{M}$: joint with M. Stillman, D. Grayson, H. Schenck (Macaulay 2), J. Hauenstein (Bertini), B. Nelson, V. Jejjala

  1. n-fields: start with polynomial ring $\mathbb{C}[\phi_1, \ldots, \phi_n]$
  2. D = set of k GIO's: a ring map $\mathbb{C}[\phi_1, \ldots, \phi_n] \xrightarrow{D} \mathbb{C}[D_1, \ldots, D_k]$
  3. Now incorporate superpotential: F-flatness

     $\langle f_{i=1,\ldots,n} = \frac{\partial W(\phi_i)}{\partial \phi_i} = 0 \rangle \simeq$ ideal of $\mathbb{C}[\phi_1, \ldots, \phi_k]$
  4. Moduli space = image of the ring map

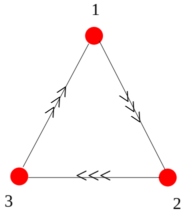
     $$\boxed{\frac{\mathbb{C}[\phi_1,\ldots,\phi_n]}{\{F=\langle f_1,\ldots,f_n\rangle\}} \xrightarrow{D=GIO} \mathbb{C}[D_1, \ldots, D_k], \quad \mathcal{M} \simeq \mathrm{Im}(D)}$$

- Image is an ideal of $\mathbb{C}[D_1, \ldots, D_k]$, i.e.,

  $\mathcal{M}$ explicitly realised as an affine variety in $\mathbb{C}^k$

Abelian Quotient: $\mathcal{M} = \mathbb{C}^3/\Gamma$

- All abelian orbifolds are toric.
- Archetypal example: $\mathbb{C}^3/\mathbb{Z}_3$ with action $(1,1,1) \rightsquigarrow U(1)^3$ quiver theory



$$W = \epsilon_{\alpha\beta\gamma} X_{12}^{(\alpha)} X_{23}^{(\beta)} X_{31}^{(\gamma)}, \qquad X_{12}^{(\alpha)}, X_{23}^{(\beta)}, X_{31}^{(\gamma)}, \alpha, \beta, \gamma = 1, 2, 3$$

Adjacency Matrix: $A = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & 3 \\ 3 & 0 & 0 \end{pmatrix}$

Incidence Matrix: $d = \begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{pmatrix}$

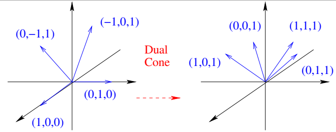
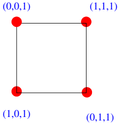
- loops: $3^3 = 27$ GIOs; arrows: 3×3 fields
- Moduli space: 27 quadrics in $\mathbb{C}^{10}$, explicit equations for

  $\mathbb{C}^3/\mathbb{Z}_3 \leftarrow Tot(\mathcal{O}_{\mathbb{P}^2}(-3))$

Back to Toric Quivers

Notation for Affine Toric Variety

Back to Toric Quivers

Def		Example (Conifold)
Comb.:	Convex Cone $\sigma \in \mathbb{Z}^d \rightsquigarrow$ Dual Cone $\sigma^\vee \rightsquigarrow X =$ Spec$_{Max}\mathbb{C}[S_\sigma = x_i^{\mathrm{gen}(\sigma^\vee)\cap\mathbb{Z}^d}]$ Toric Diagram $= S_\sigma$	 $S_\sigma = \langle a = z, c = yz, b = xyz, d = xz \rangle$ $ab = cd$ in $\mathbb{C}^4[a,b,c,d]$
Symp:	Generalise $\mathbb{P}^n$: a $(\mathbb{C}^*)^{q-d}$ action on $\mathbb{C}^q_{[x_i]}$ $x_i \mapsto \lambda_a^{Q_{i=1\ldots q}^{a=1\ldots q-d}} x_i$ with Relations: $\sum\limits_{i=1}^{d} Q_i^a v_i = 0$ Toric Diagram $= v_i$	 $Q = [-1,-1,1,1]$ $\mathbb{C}^*$ on $\mathbb{C}^4 \rightsquigarrow$ ker$Q = G_t =$ $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$
Comp:	Binomial Ideal $\langle \prod p_i = \prod q_j \rangle$	$ab = cd$ in $\mathbb{C}^4$

Progress in String Theory

Back to ML Landscape

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

Principle Component Analysis 主成分分析

- INPUT: $\vec{x}^{(i)} \in \mathbb{R}^n$, n large, $i = 1, \ldots, m$;
- OUTPUT: $\vec{c}^{(i)} = f(\vec{x}) \in \mathbb{R}^\ell$, and $g : \vec{x} \simeq g(f(\vec{x}))$
  - $\ell \ll n$ to help with the curse of dimension
  - try linear encoding: $g(\vec{c}) = D_{n\times\ell} \cdot \vec{c}$, with $D^T D = \mathbb{I}$ thus $\vec{c} = D^T \vec{x}$
  - Cost Function: $||\vec{x} - g(\vec{c})||$; Need to find $D_{n\times\ell}$ s.t., minimize

$$\sqrt{\sum_i \left|\vec{x}^{(i)} - DD^T \vec{x}^{(i)}\right|^2} \ , \qquad \text{s.t. } D^T D = \mathbb{I}_\ell \ .$$

- ℓ gives the ℓ-th Principal Component

Back to ML

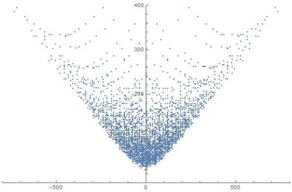
Some Jargon

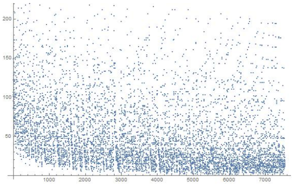
- Epoch: (training round) 訓練輪 1 complete cycle where the NN has seen $\mathcal{T}$
- Batch: 批量

  $\mathcal{T}$ (since $|\mathcal{T}|$ is often too large) is divided into batches (mini-batches) to be passed through the NN
- iterations: need to iterate in order to pass all through all of $\mathcal{T}$
- Hence $|\mathcal{T}| =$ Batch size $\times \#$ Iterations
- Often need to sample $\mathcal{T}$ from $\mathcal{D}$ and pass through multiple epochs

Back to NN

Hodge Plots for WP4

(a) 

(b) 

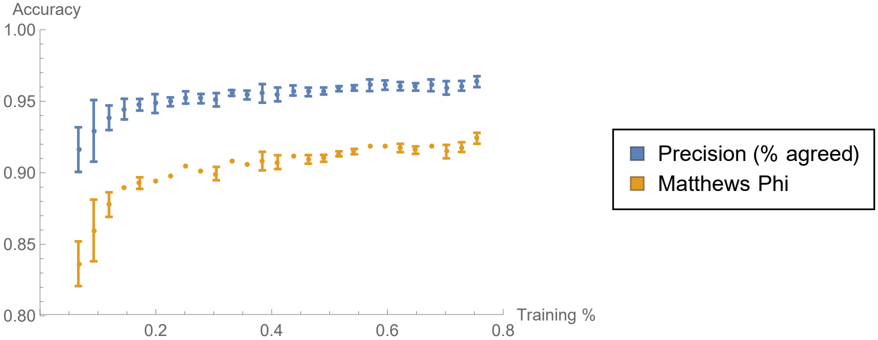
(a)Mirror plot of

$(\chi, h^{1,1} + h^{2,1})$

(b)Distribution of$h^{2,1}$

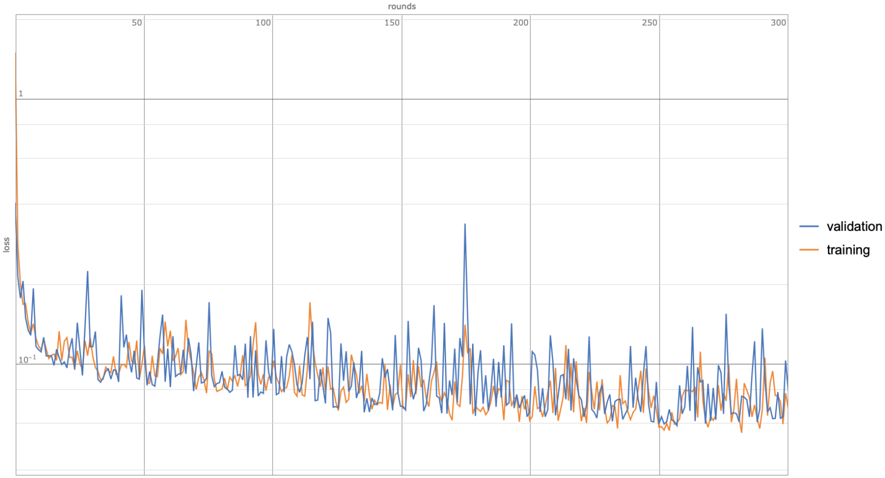
Return

Learning Curve: Deciding Large $h^{2,1}$ WP4



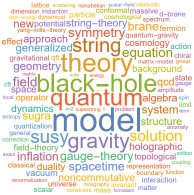
Return

Training Curve: Deciding Large $h^{2,1}$ WP4

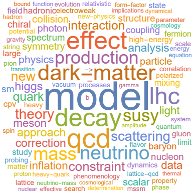


Return

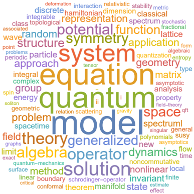
ArXiv Word-Clouds



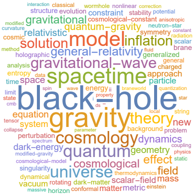
hep-th



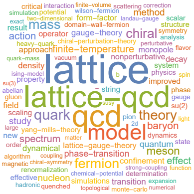
hep-ph



math-ph



gr-qc



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