Deep Learning Topology

Vishnu Jejjala
University of the Witwatersrand

LOGML Summer School
10 July 2024

Topology

- Topology is about shapes
- Topology is also about functions between shapes
- Two shapes are topologically equivalent if they can be continuously deformed into one another, *e.g.*, by bending, twisting, stretching, crumpling, but not by tearing, cutting, gluing

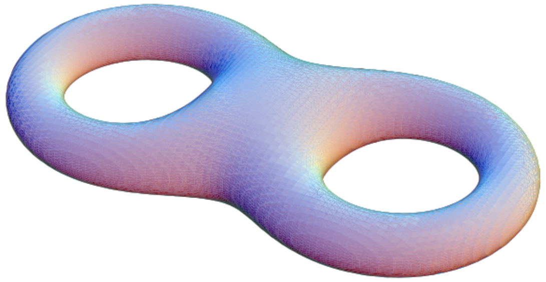
Formally: There exists a **homeomorphism** $f: X \to Y$

— f is one-to-one and onto

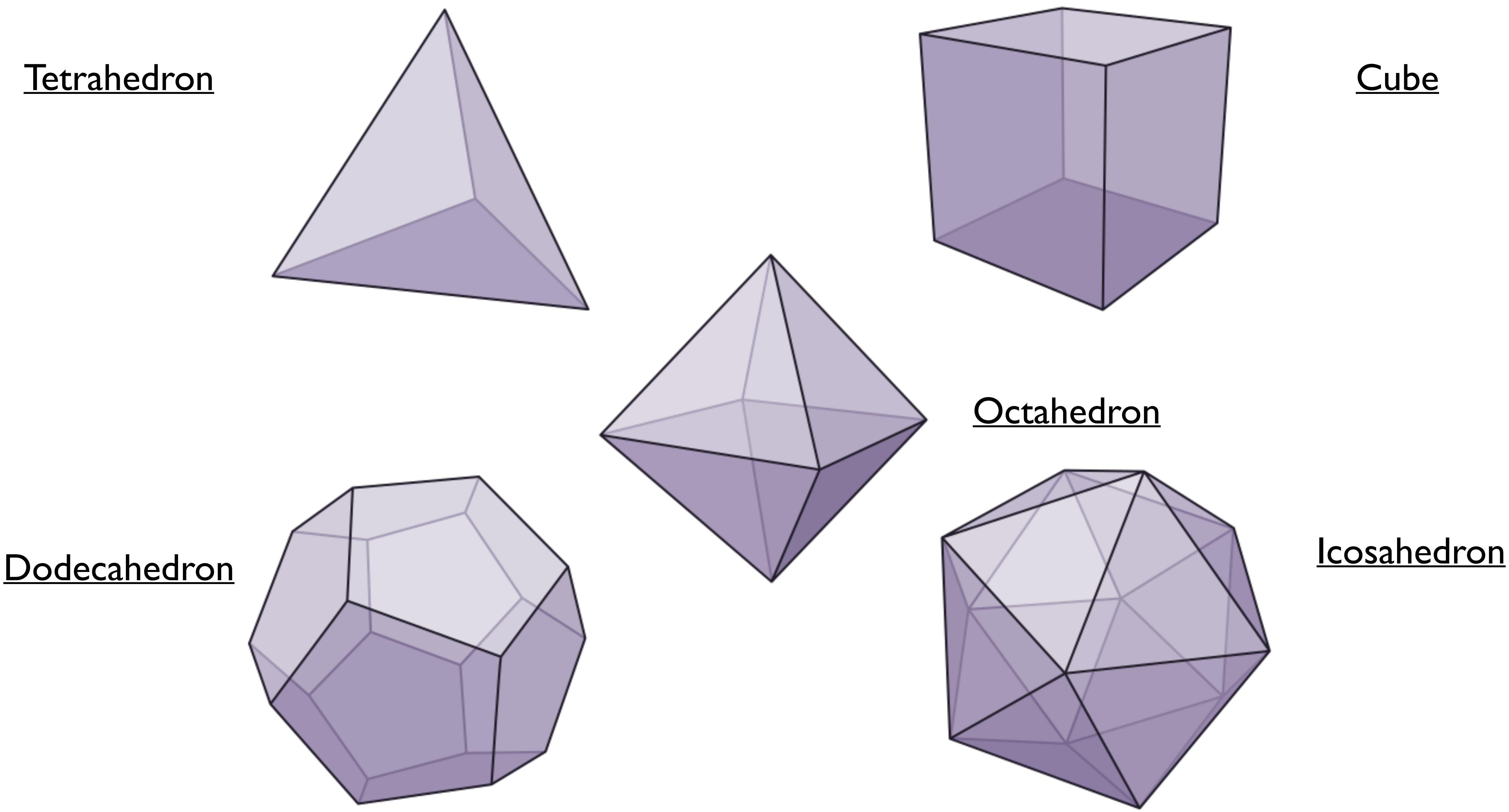
— f and f^{-1} are continuous

- You have an intuitive understanding of topology already





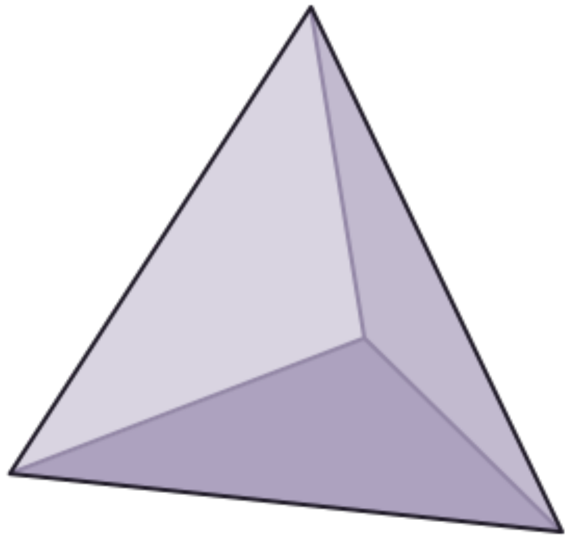

Euler Characteristic



- **Platonic solids:** convex, regular polyhedra in 3d

Theaetetus (c. 369 BCE)

Euler Characteristic

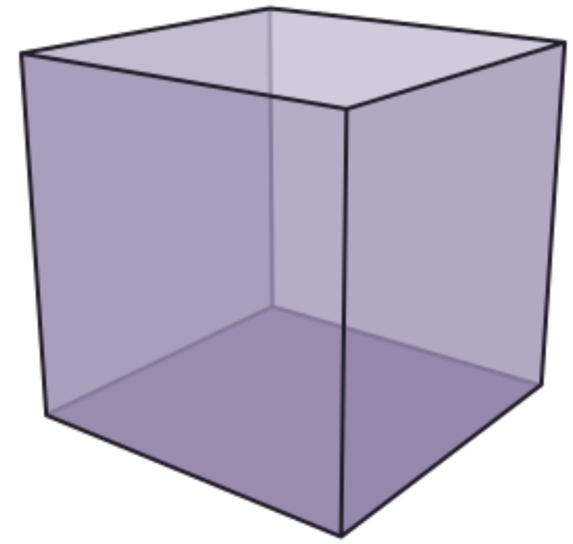


Tetrahedron

$V = 4$

$E = 6$

$F = 4$

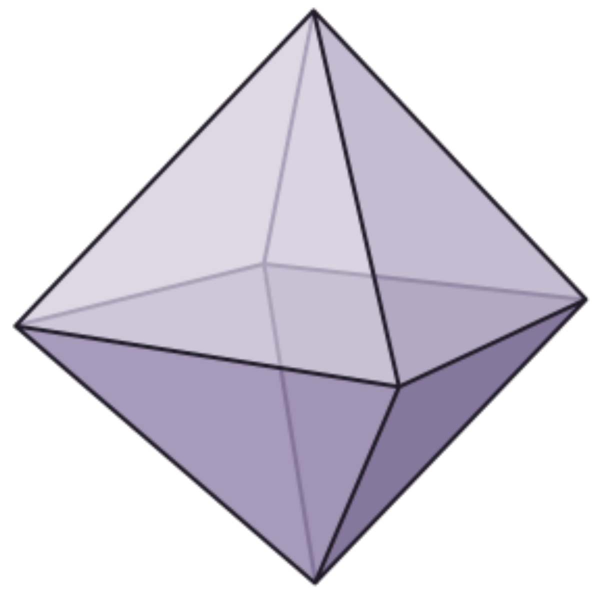


Cube

$V = 8$

$E = 12$

$F = 6$

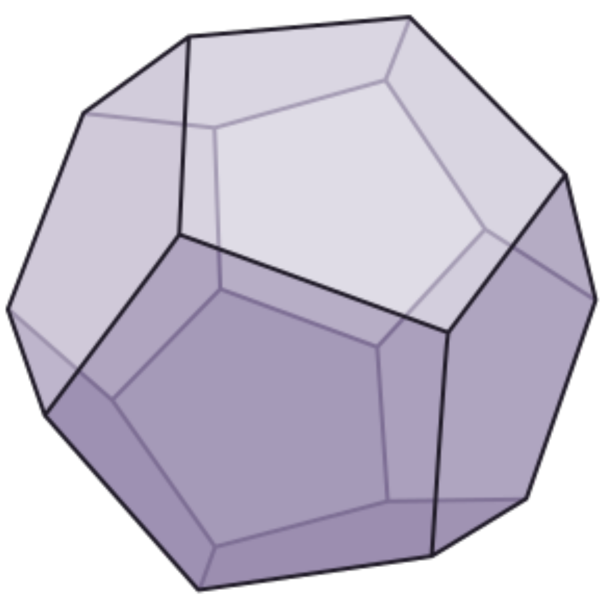


Octahedron

$V = 6$

$E = 12$

$F = 8$

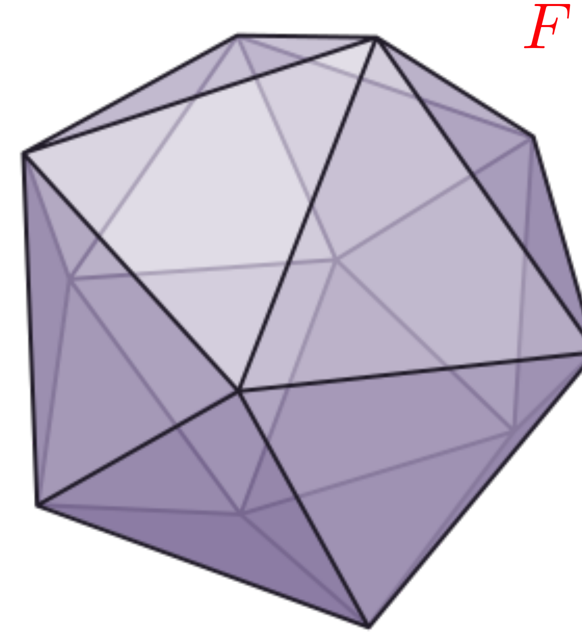


Dodecahedron

$V = 20$

$E = 30$

$F = 12$



Icosahedron

$V = 12$

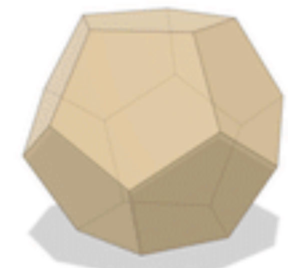
$E = 30$

$F = 20$

$$\boxed{\chi = V - E + F = 2}$$

Maurolico (1537, unpublished)
Euler (1752)

Euler Characteristic

	Tetrahedron	Cube	Octahedron	Dodecahedron	Icosahedron
Platonic Solid					
Spherical Polyhedron					

$$\chi = 2 - 2g$$

$$g = \text{genus}$$

- The Euler characteristic and genus are topological invariants

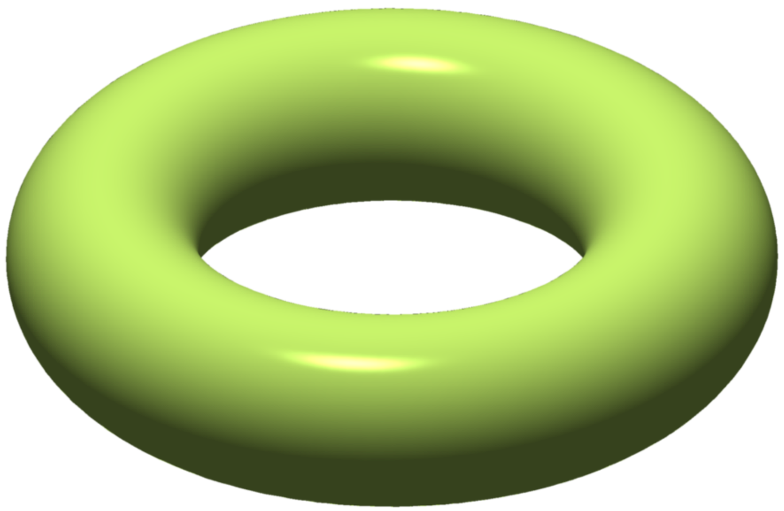
Euler Characteristic

genus 0

genus 1

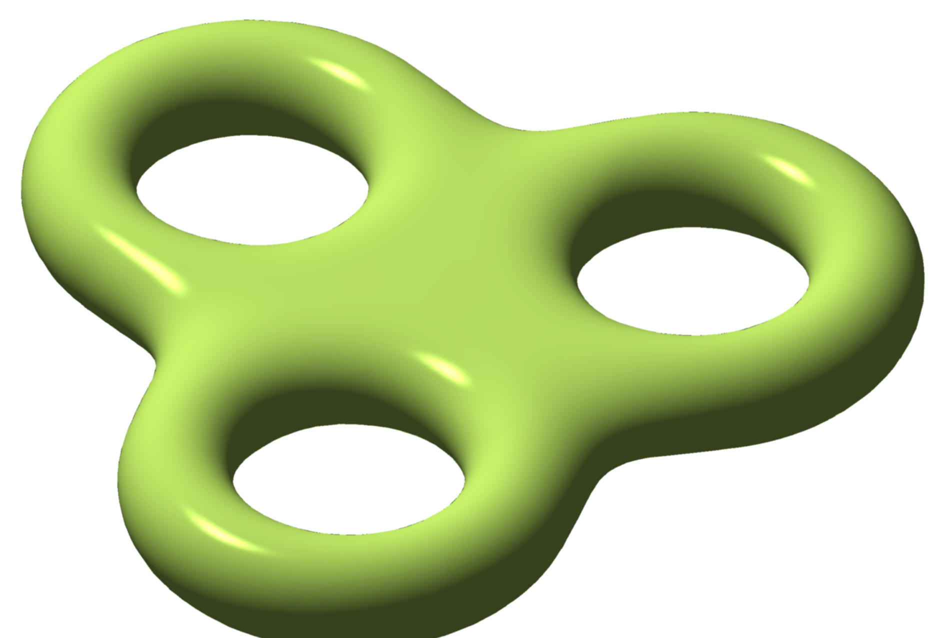
topological invariants

A Topologist's Morning



coffee

trousers

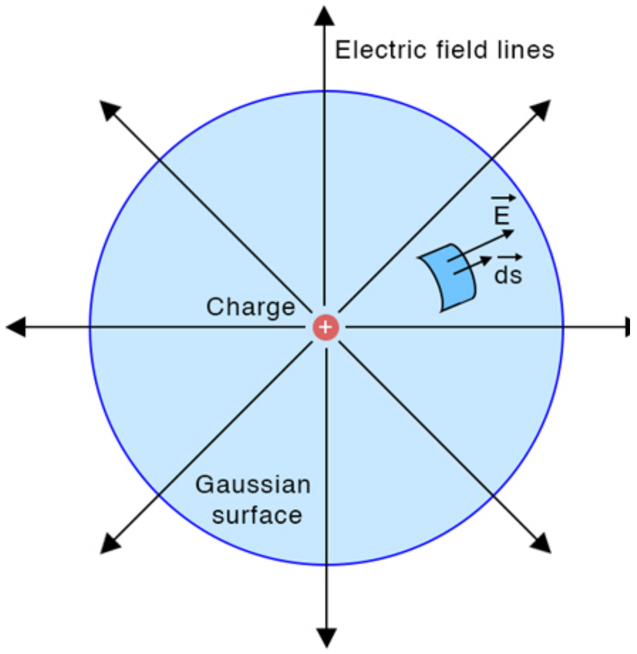


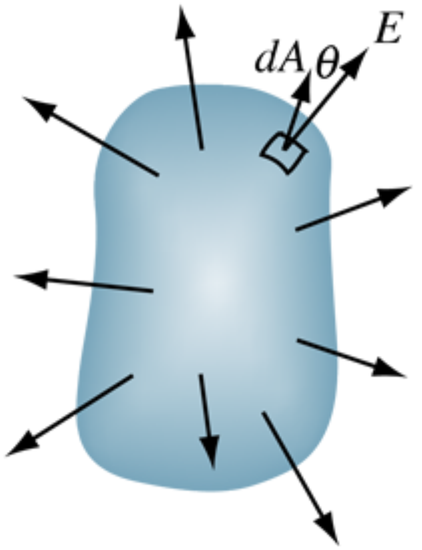
shirt

Topology in Physics

- Gauss's law

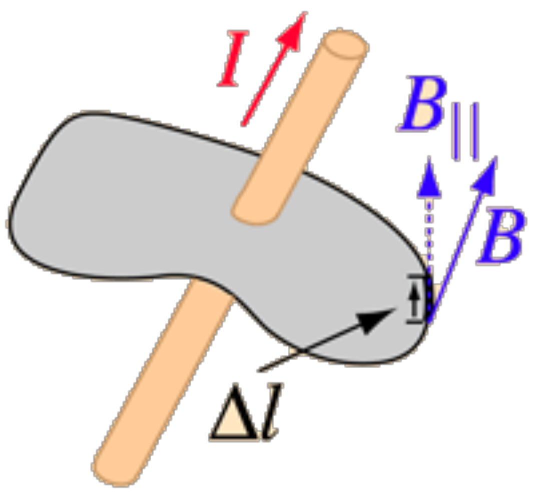
$$\Phi_E = \int_\Sigma \vec{\mathbf{E}} \cdot \mathrm{d}\vec{\mathbf{A}}$$





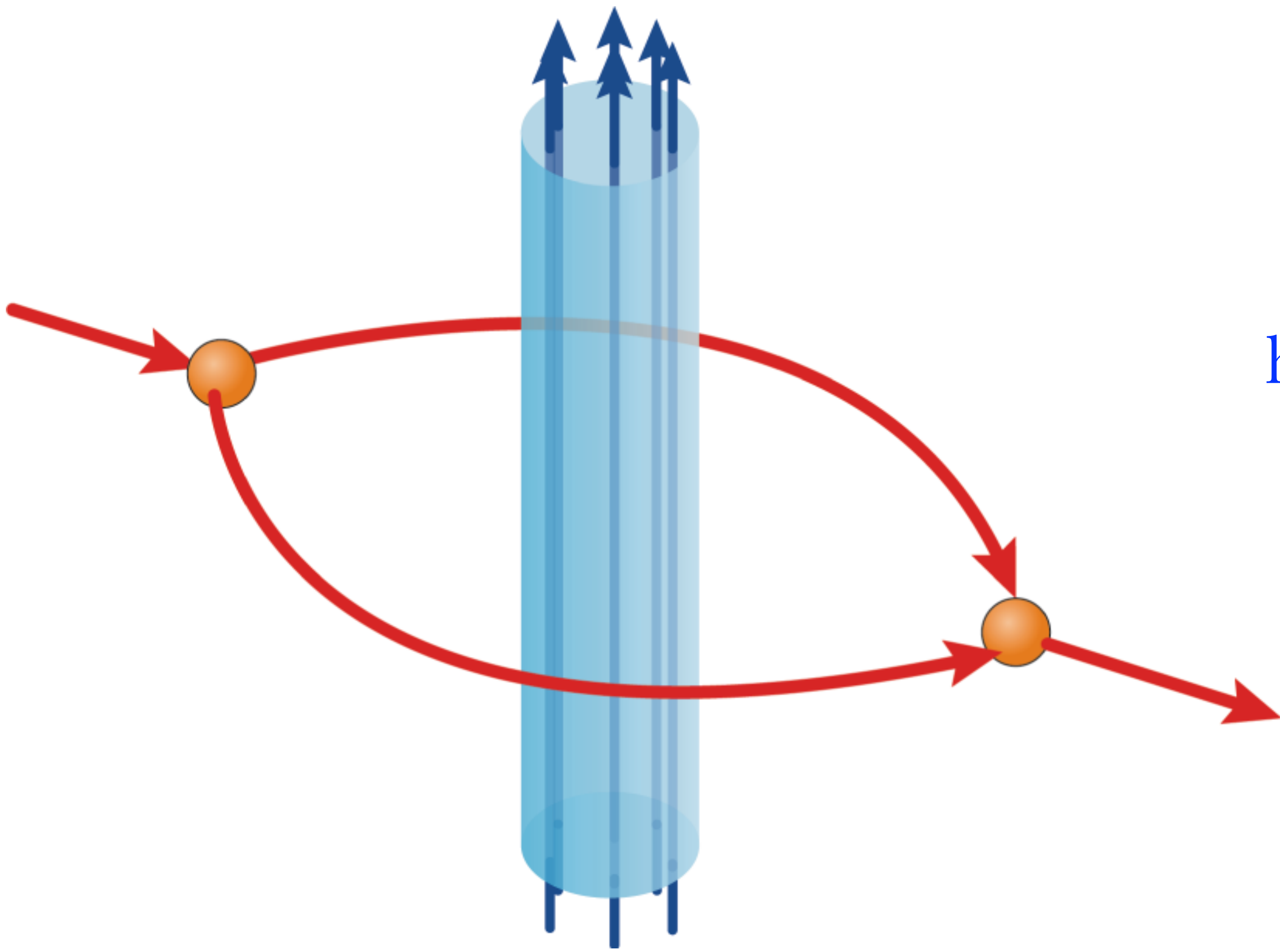
- Ampère's law

$$\oint_C \vec{\mathbf{B}} \cdot \mathrm{d}\vec{\ell} = \int\int (\vec{\mathbf{J}} + \partial_t \vec{\mathbf{E}}) \cdot \mathrm{d}\vec{\mathbf{S}}$$



Topology in Physics

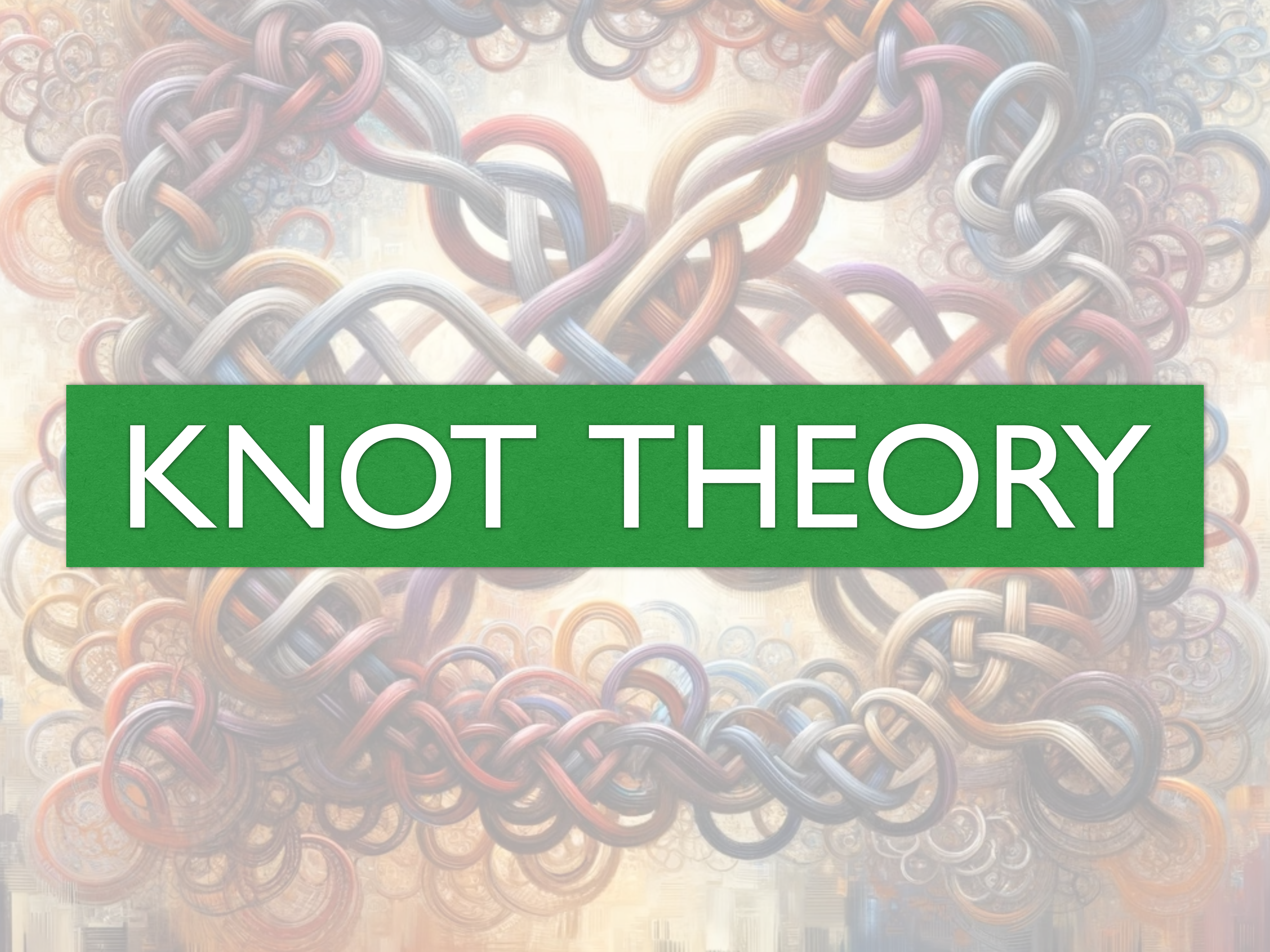
- Aharonov–Bohm effect



holonomy of gauge field
around closed loop

$$\Delta\varphi = \frac{e}{\hbar}\int_{\gamma} \vec{\mathbf{A}} \cdot \mathrm{d}\vec{\mathbf{x}}$$
$$= \frac{e\Phi_B}{\hbar}$$
$$= 2\pi n\ , \quad n \in \mathbb{Z}$$

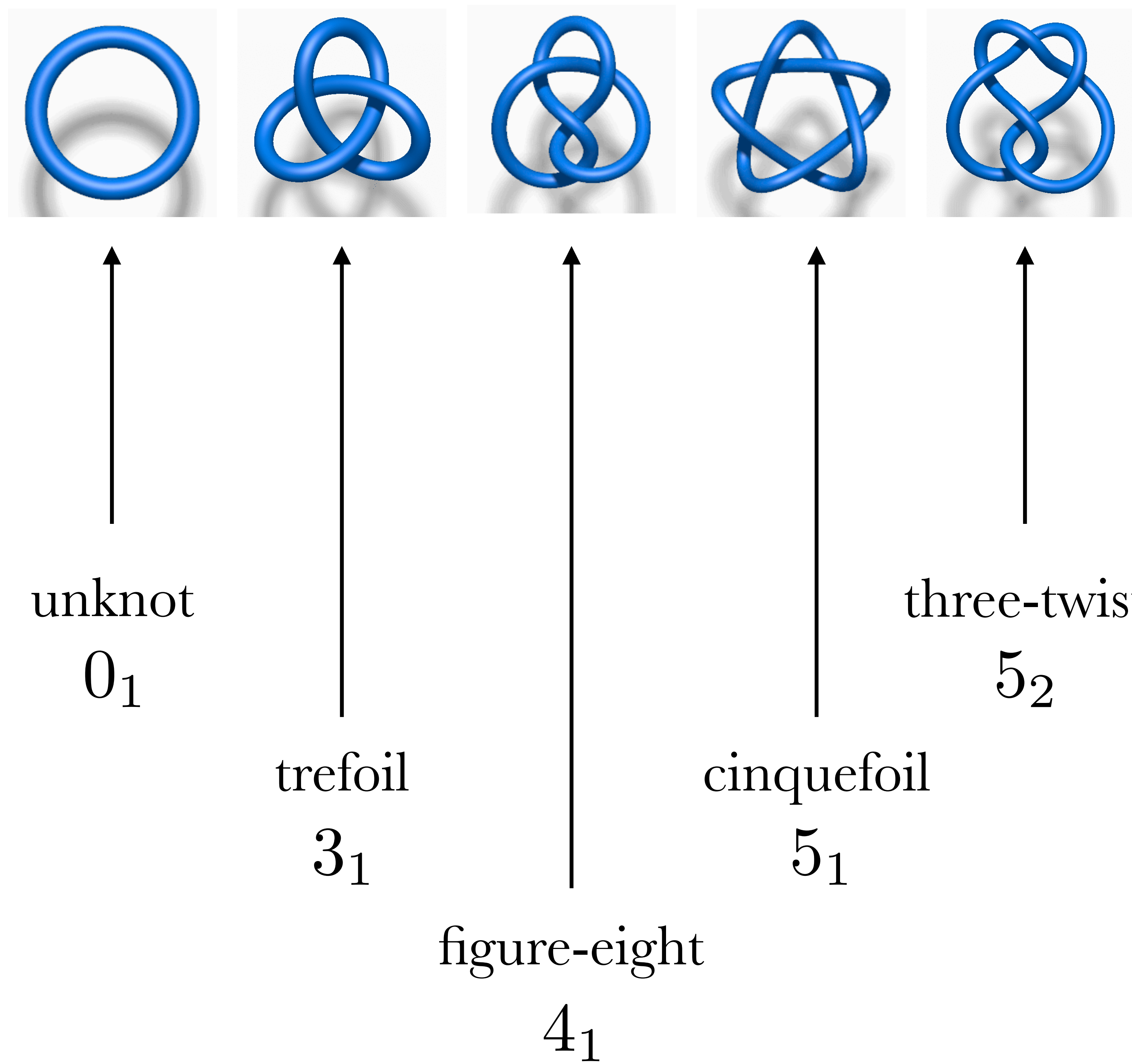
- Condensed matter systems, *e.g.*, quantum Hall effect



KNOT THEORY

Dramatis Personae

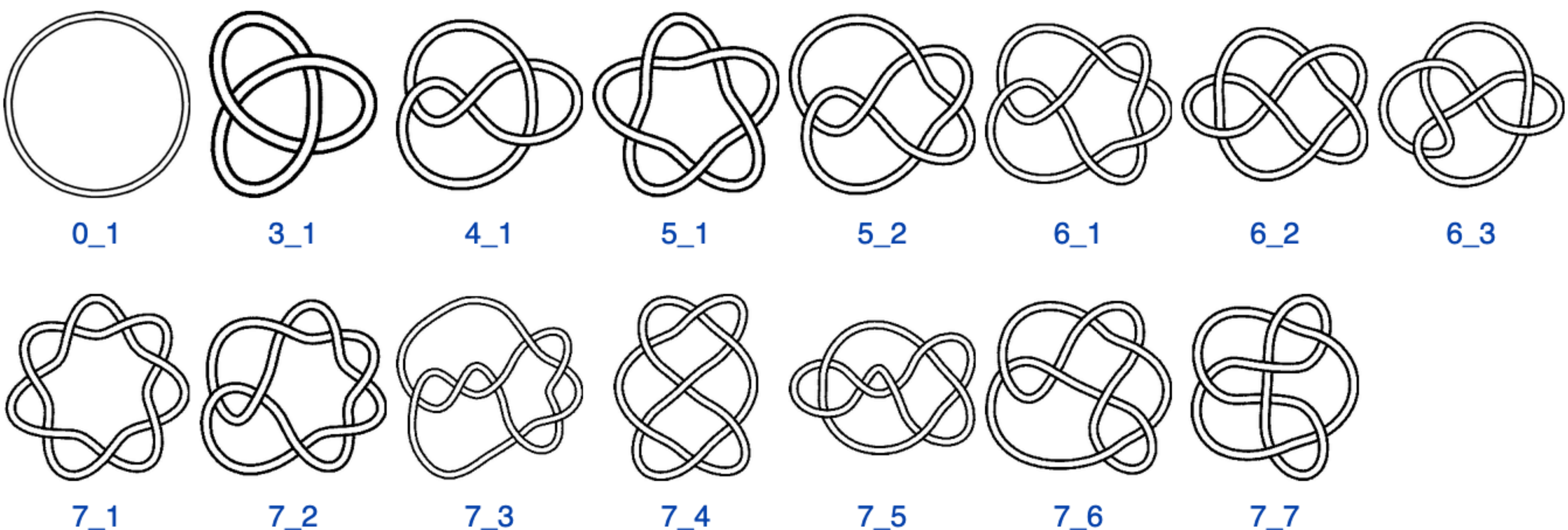
Knot: $S^1 \subset S^3$; *e.g.*,



Dramatis Personae

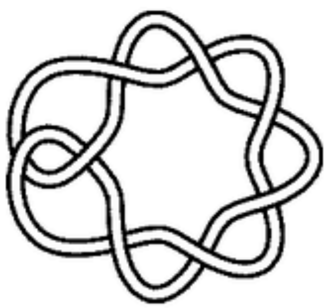
Knot: $S^1 \subset S^3$

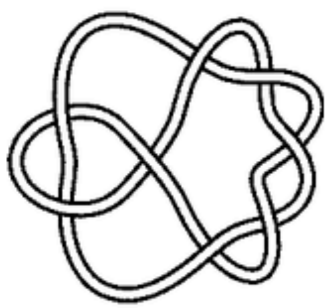
- Organized by crossing number, a proxy for complexity

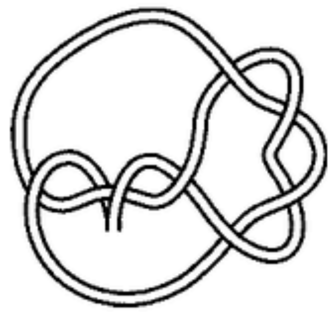


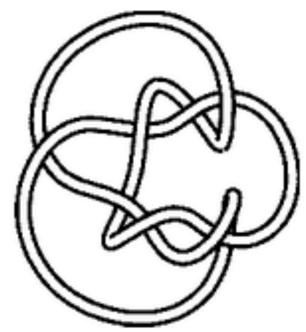
Dramatis Personae

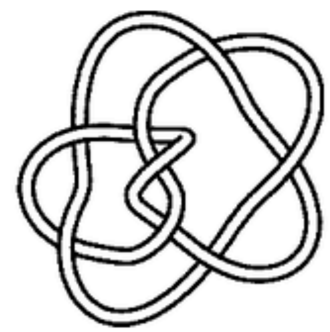
Knot: $S^1 \subset S^3$

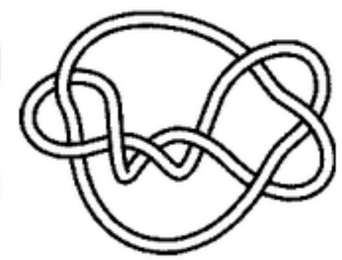

8_1


8_2

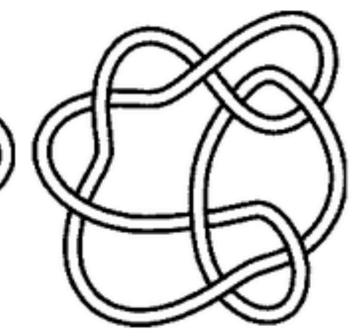

8_3


8_4


8_5


8_6

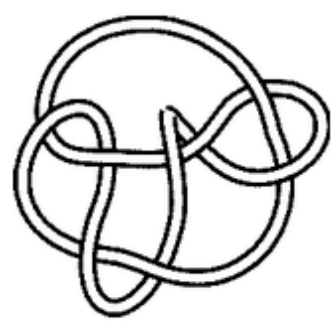

8_7

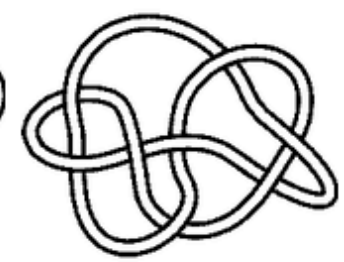

8_8

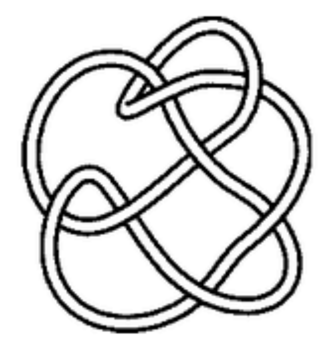

8_9


8_10


8_11

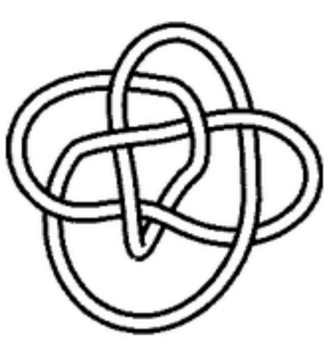

8_12

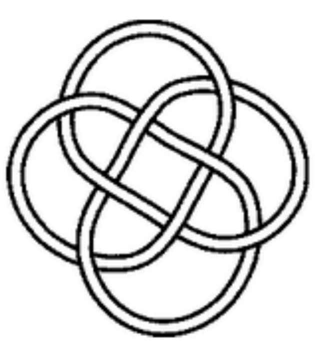

8_13

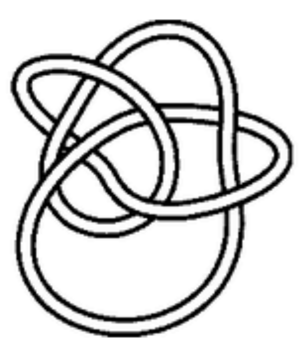

8_14

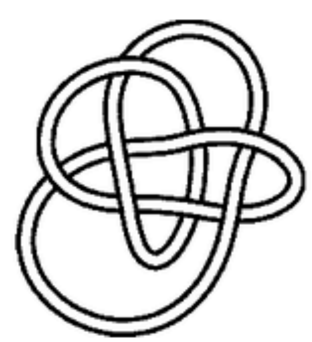

8_15

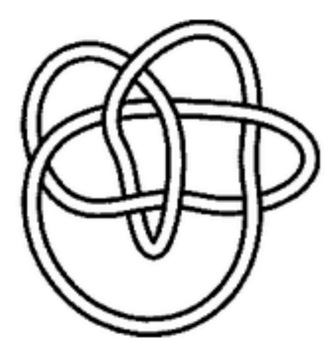

8_16


8_17


8_18

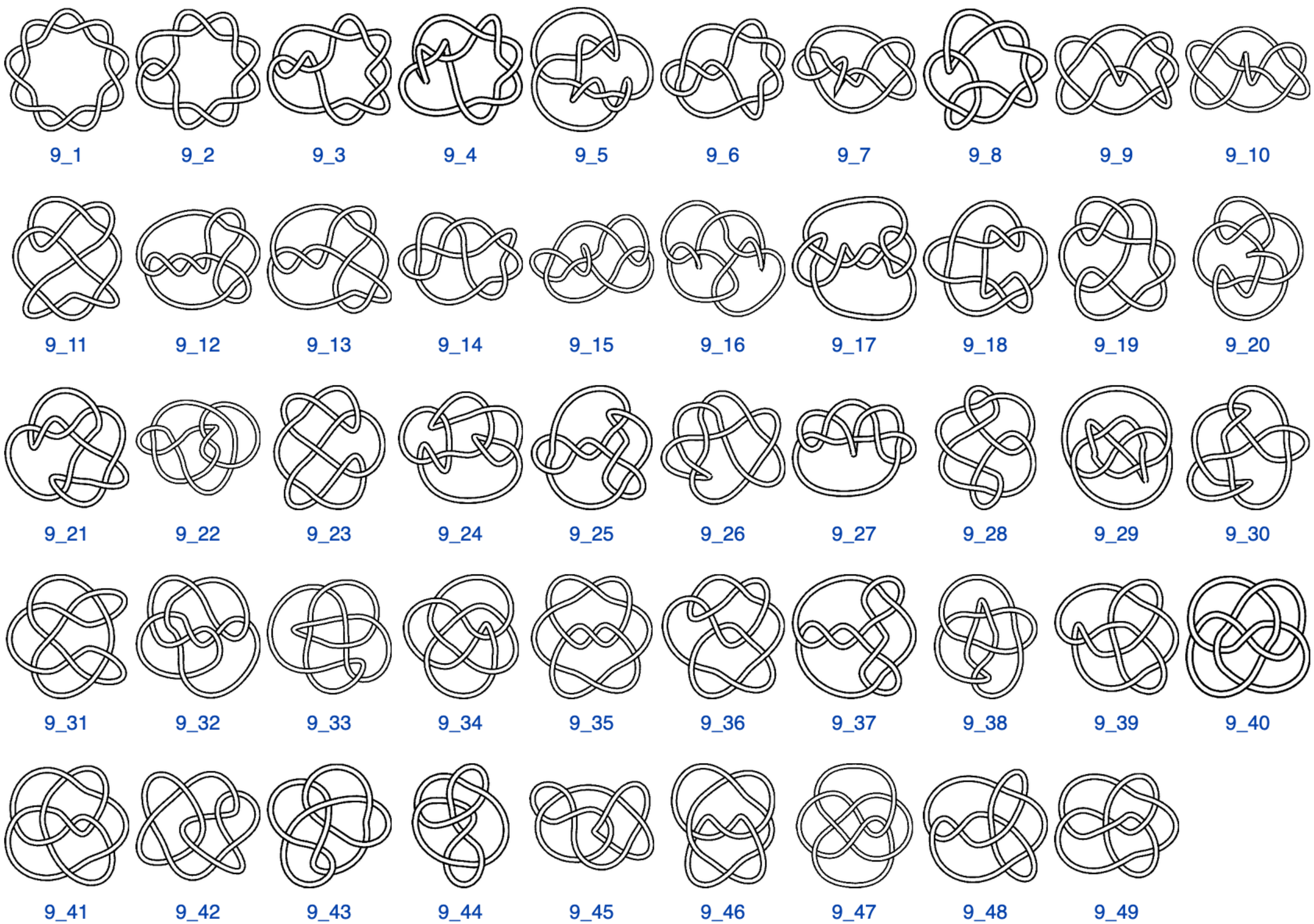

8_19


8_20


8_21

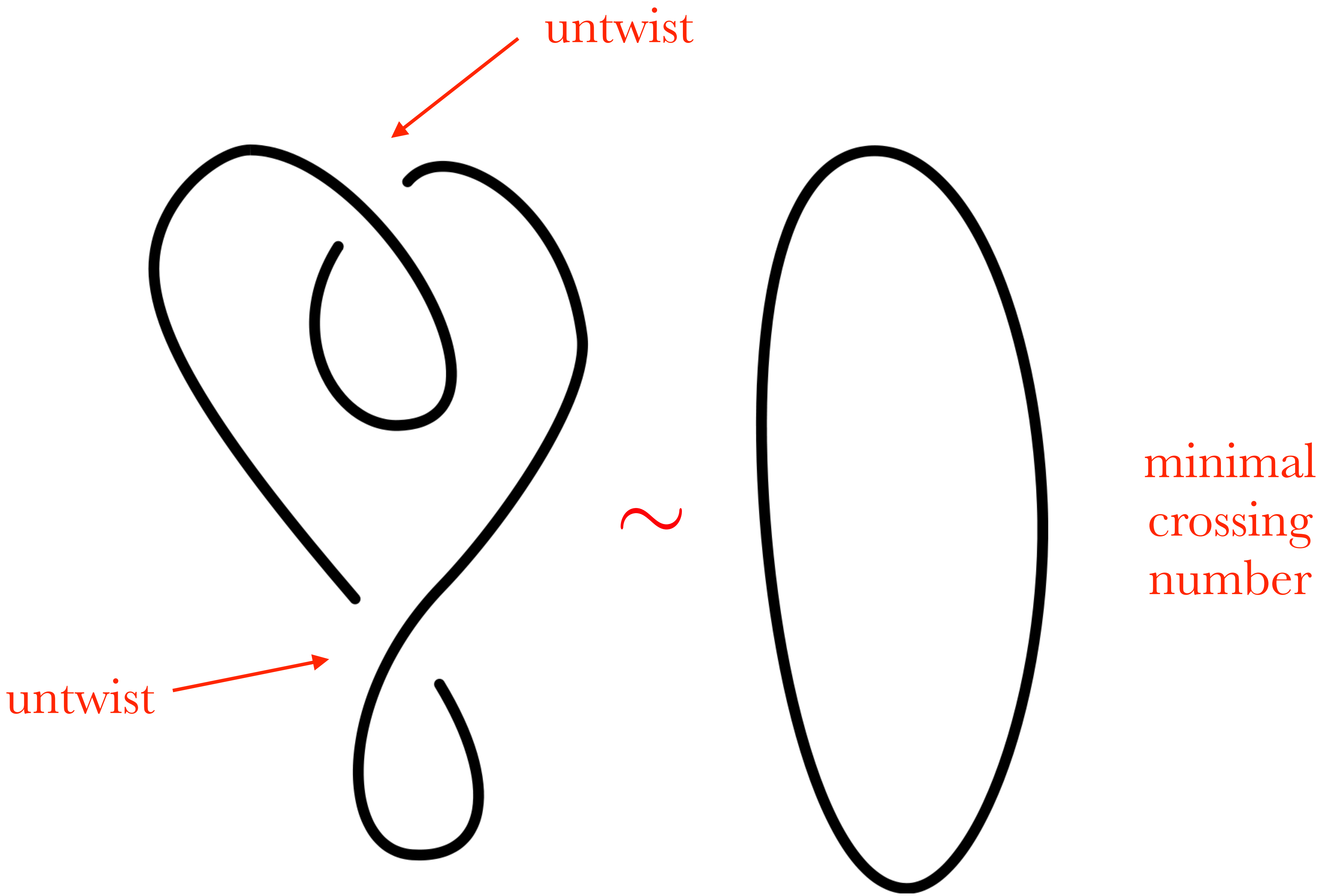
Dramatis Personae

Knot: $S^1 \subset S^3$



Dramatis Personae

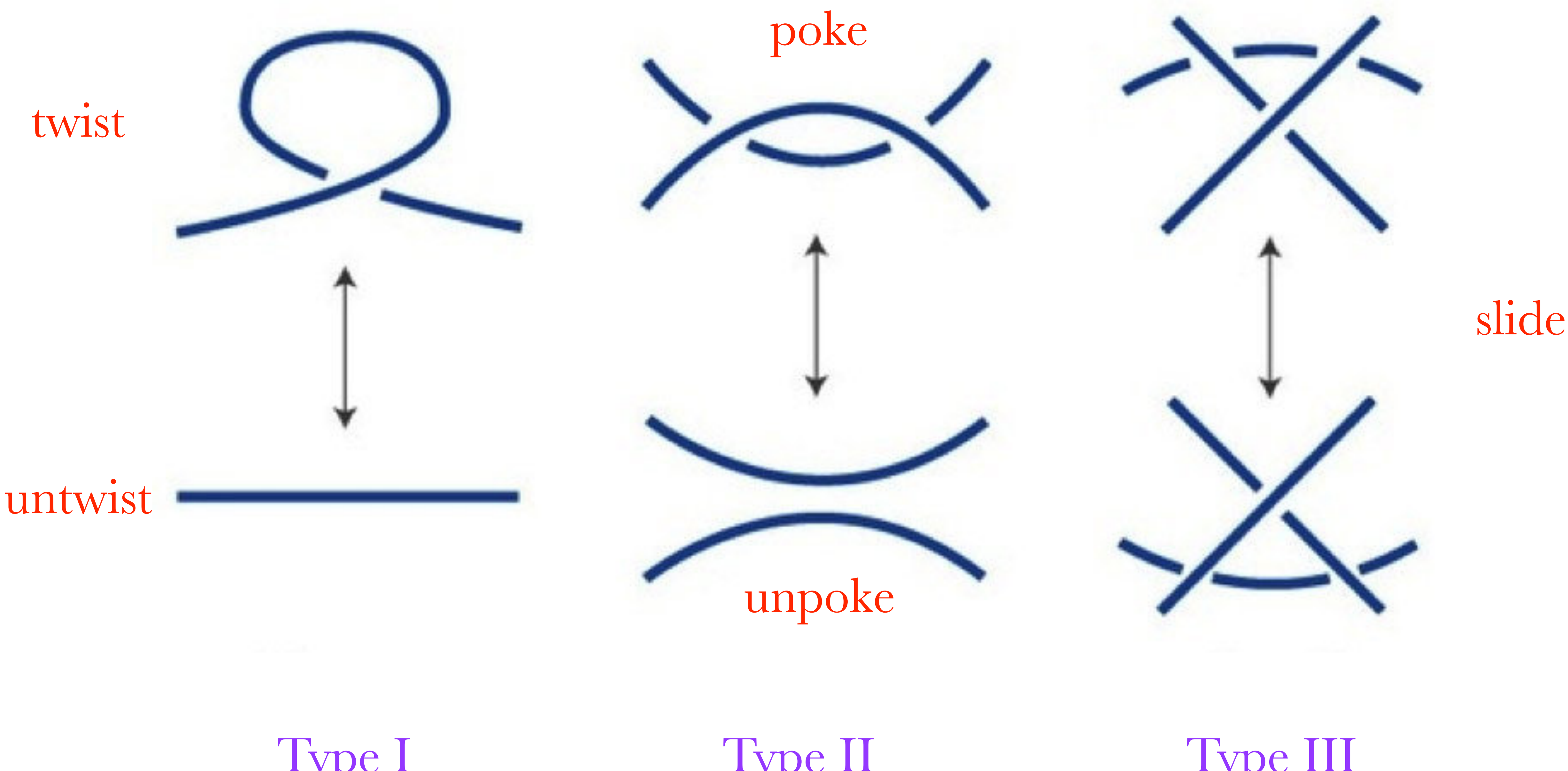
Knot: $S^1 \subset S^3$



Dramatis Personae

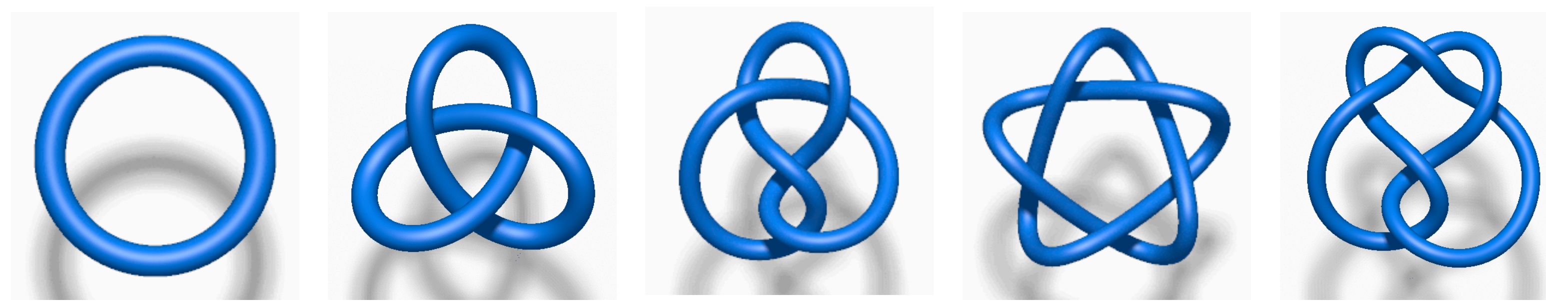
Knot: $S^1 \subset S^3$

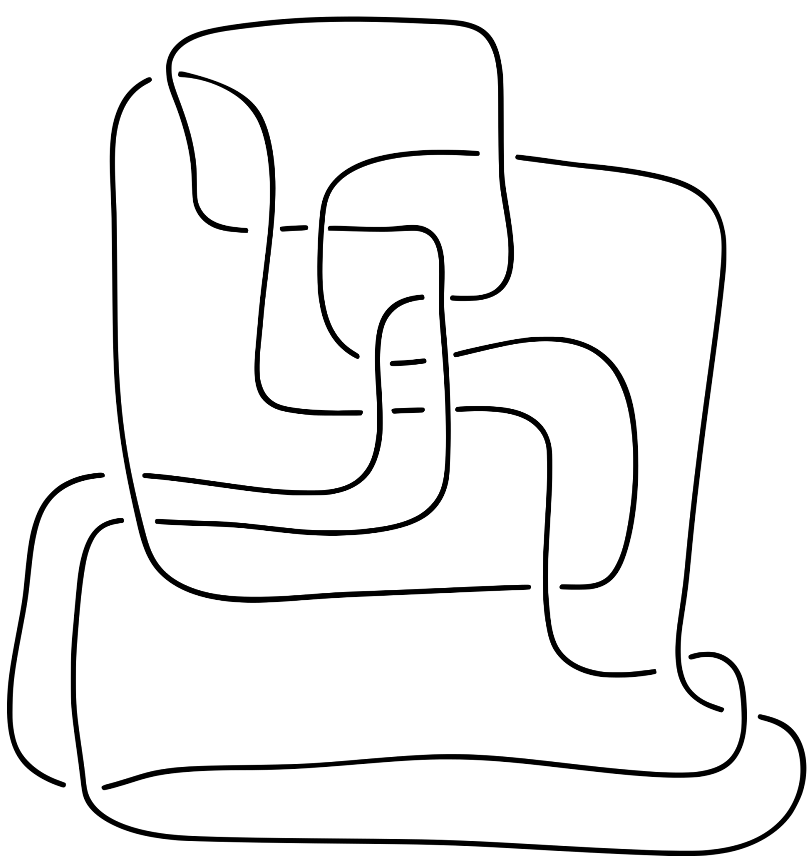
Reidemeister moves



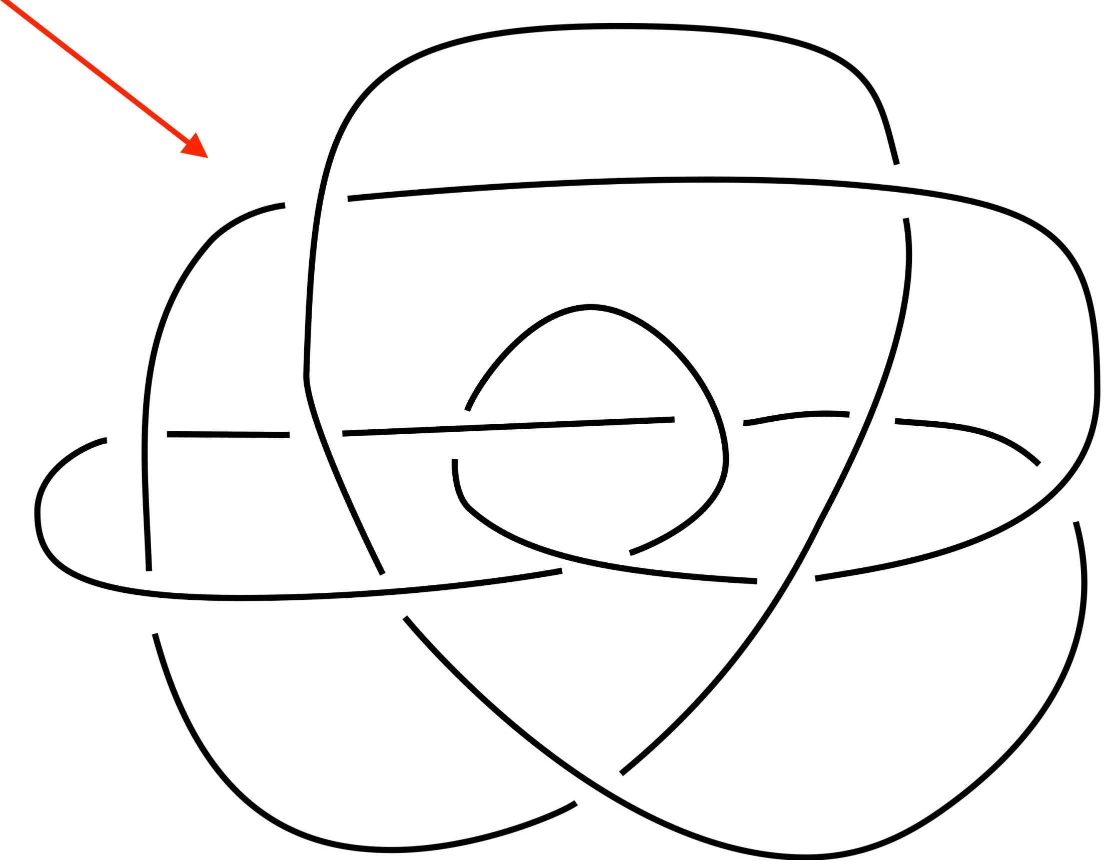
Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,





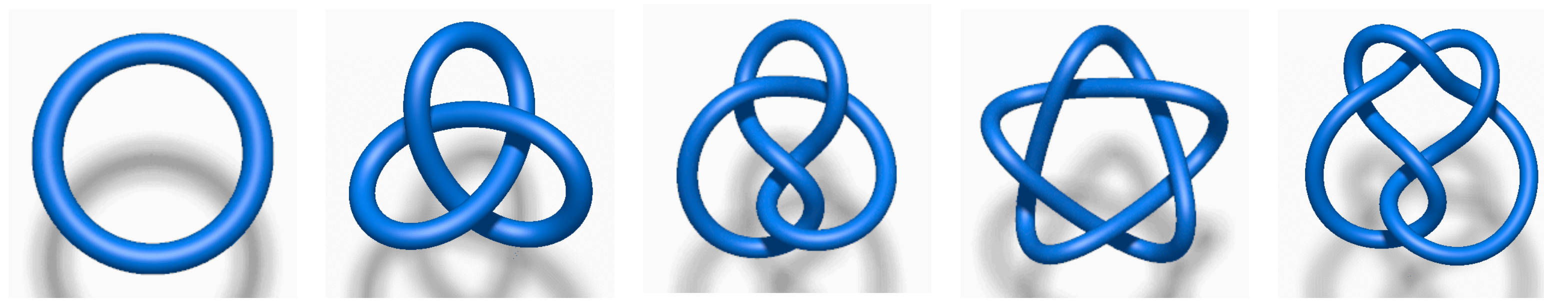
Thistlethwaite unknot

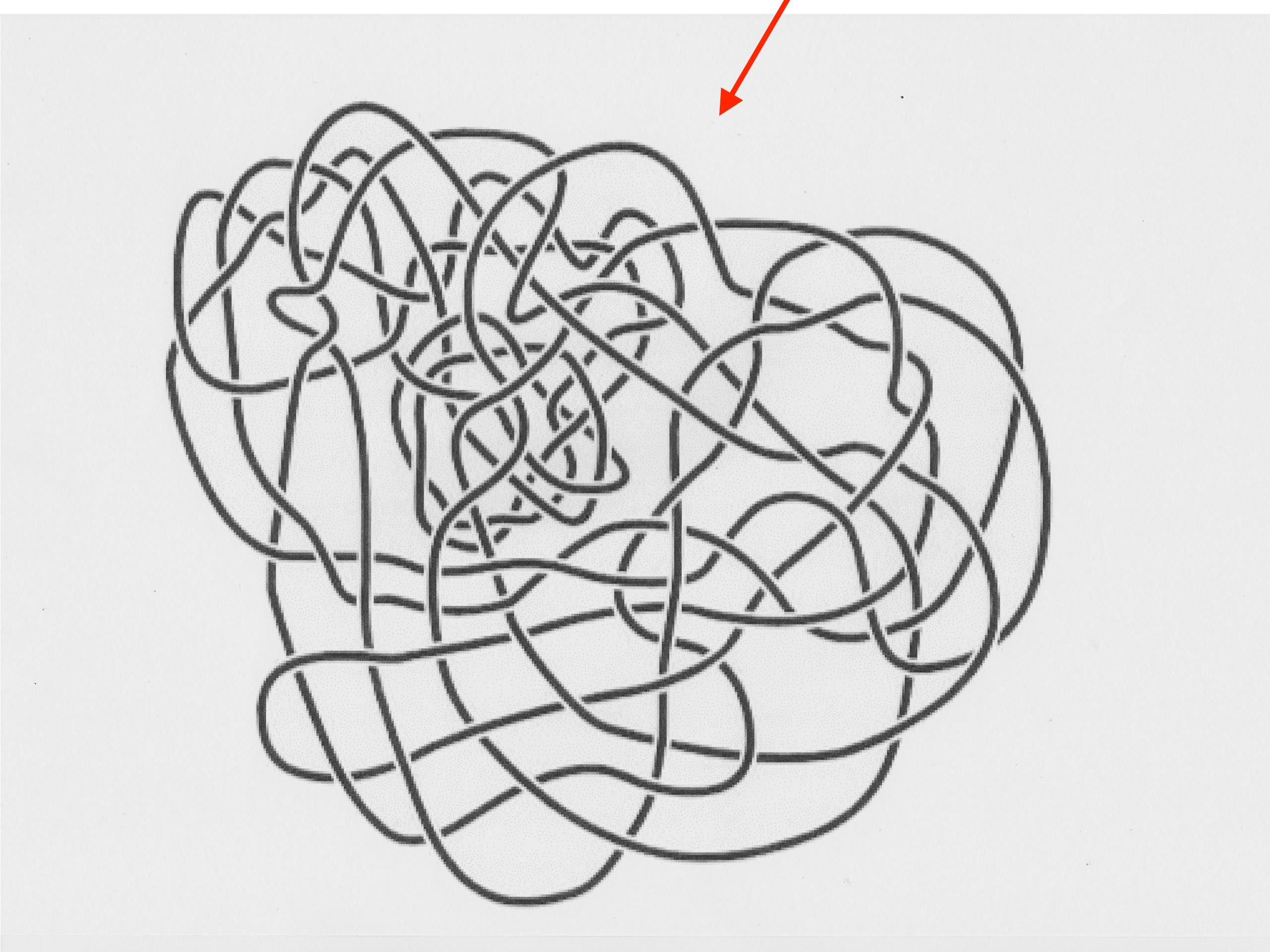


Ochiai unknot

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,





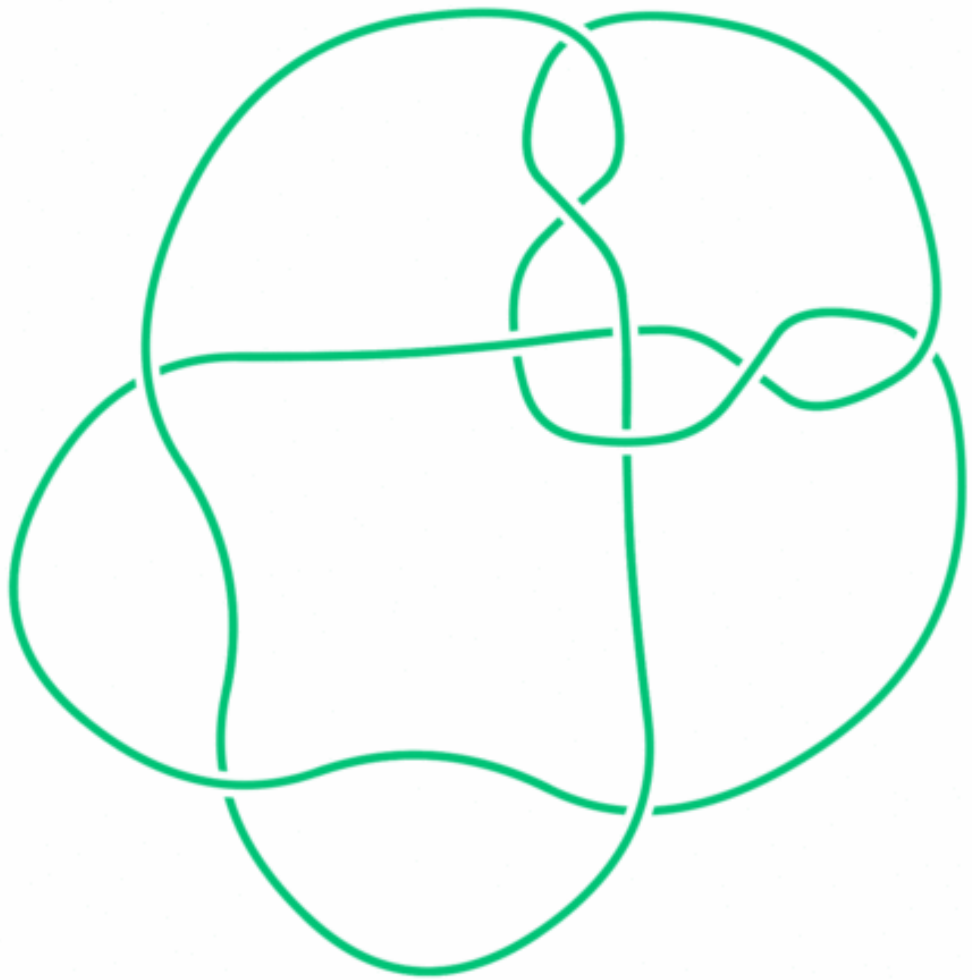
Burton (2015)

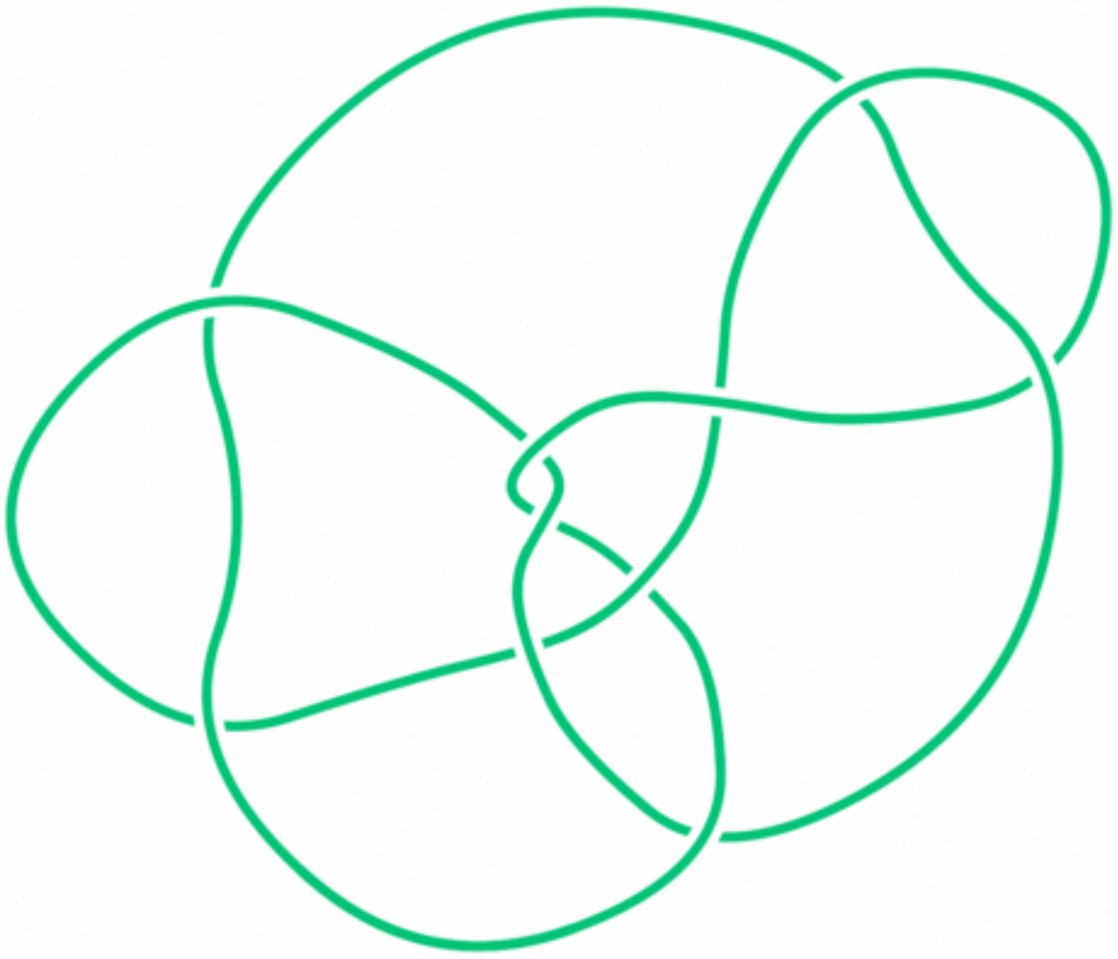
Haken or Gordian unknot

Dramatis Personae

Knot: $S^1 \subset S^3$

Rolfsen table

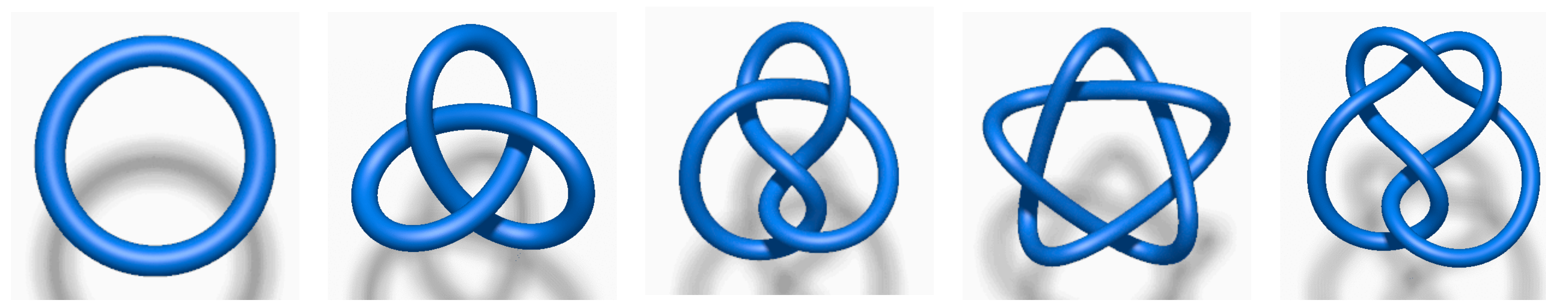




10_{161} ⟵ Perko pair (1973) ⟶ 10_{162}

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,



Jones polynomial: $J(K;q) = (-q^{\frac{3}{4}})^{w(K)} \dfrac{\langle K \rangle}{\langle \bigcirc \rangle}$

$$\langle \times \rangle = q^{\frac{1}{4}} \langle \asymp \rangle + \frac{1}{q^{\frac{1}{4}}} \langle)(\rangle$$

$w(K) =$ overhand $-$ underhand

$J(\bigcirc;q) = 1$

Jones (1985)

topological invariant: independent of how knot is drawn

Question: how to calculate these?

Answer: quantum field theory!

Topological Invariants

- On a manifold $\mathcal{M}$ with metric $g_{\mu\nu}$, a topological invariant enjoys:

$$\frac{\delta}{\delta g_{\mu\nu}} \langle \mathcal{O}_1 \ldots \mathcal{O}_n \rangle = 0$$

- In Chern–Simons theory, the operators are Wilson loops

$$U_R(\gamma) = \mathrm{tr}_R \, \mathcal{P} \exp \left(i \oint_\gamma A \right)$$

- The colored Jones polynomial is a knot invariant in 3d:

$$J_n(K; q = e^{2\pi i/(k+2)}) = \frac{\int_{\mathcal{U}} [DA] \; U_n(K) e^{iS_{\mathrm{CS}}(A)}}{\int_{\mathcal{U}} [DA] \; U_n(0_1) e^{iS_{\mathrm{CS}}(A)}} = \frac{\langle U_n(K) \rangle}{\langle U_n(0_1) \rangle}$$

$$S_{\mathrm{CS}}(A) = \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right) , \quad Z(\mathcal{M}) = \int_{\mathcal{U}} [DA] \; e^{iS_{\mathrm{CS}}(A)}$$

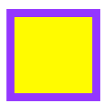
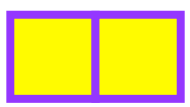
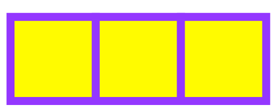
$$k \in \mathbb{Z}$$

simplest non-trivial quantum field theory

Witten (1989)

Group Theory

- Irreducible representations of $SU(2)$: $R(g_1 g_2) = R(g_1)R(g_2)$

	defining (fundamental)	2×2	matrices
	adjoint	3×3	matrices
		4×4	matrices
$\vdots$			

- Algebra: $[S_i, S_j] = i\hbar\epsilon_{ijk}S_k$

$$S_i = \frac{\hbar}{2}\sigma_i\ , \qquad \sigma_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

spin-$\frac{1}{2}$

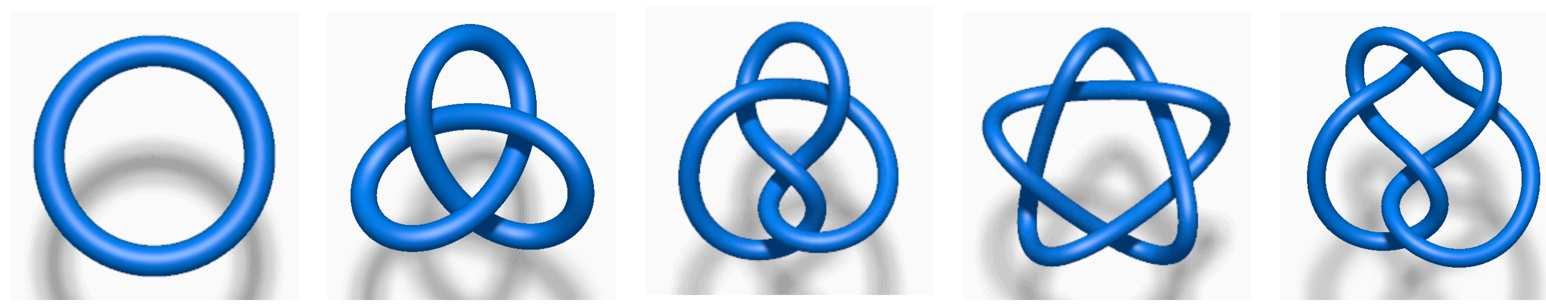
$$S_1 = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} & 1 & \\ 1 & & 1 \\ & 1 & \end{pmatrix}, \quad S_2 = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} & -i & \\ i & & -i \\ & i & \end{pmatrix}, \quad S_3 = \hbar\begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix}$$

spin-1

- Colored Jones polynomials correspond to traces in different representations

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,



Jones polynomial: $J(K;q) = (-q^{\frac{3}{4}})^{w(K)} \dfrac{\langle K \rangle}{\langle \bigcirc \rangle}$

$$\langle \times \rangle = q^{\frac{1}{4}} \langle \asymp \rangle + \frac{1}{q^{\frac{1}{4}}} \langle \,)(\, \rangle$$

$w(K) = \text{overhand} - \text{underhand}$

vev of Wilson loop operator along K in
$\square$ for $SU(2)$ Chern–Simons on S^3

Jones (1985)
Witten (1989)

$$J_2(4_1;q) = q^{-2} - q^{-1} + 1 - q + q^2 \ , \quad q = e^{\frac{2\pi i}{k+2}}$$

Open question: Is unknot unique knot for which $J_2(q) = 1$?

Open question: Unknot recognition problem — is this NP?

Dehn (1910)
Turing (1954)
Haken (1961)
Lackenby (2021)

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,

Jones polynomial: $J(K;q) = (-q^{\frac{3}{4}})^{w(K)} \dfrac{\langle K \rangle}{\langle \bigcirc \rangle}$

$$\langle \times \rangle = q^{\frac{1}{4}} \langle \asymp \rangle + \frac{1}{q^{\frac{1}{4}}} \langle)(\rangle$$

$w(K) = \text{overhand} - \text{underhand}$

vev of Wilson loop operator along K in

$\square$ for $SU(2)$ Chern–Simons on S^3

Jones (1985)
Witten (1989)

$$J_2(4_1;q) = q^{-2} - q^{-1} + 1 - q + q^2 \ , \quad q = e^{\frac{2\pi i}{k+2}}$$

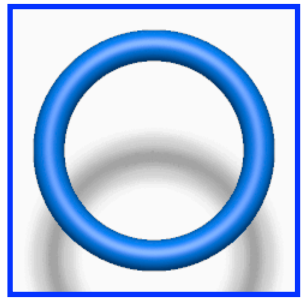
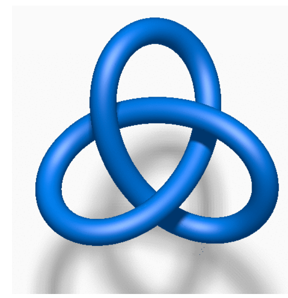
Hyperbolic volume: volume of $S^3 \backslash K$ is another knot invariant in 3d

computed from tetrahedral decomposition of knot complement

Thurston (1978)
Mostow (1968)

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,

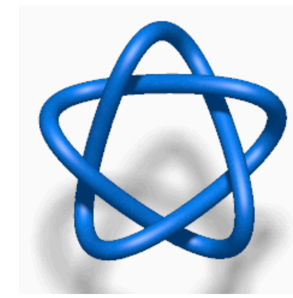
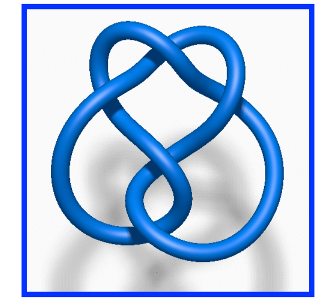



Table A1

# Crossings	# Knots	# Torus	# Sat.
0	1		
3a	1	1	
4a	1		
5a	2	1	
6a	3		
7a	7	1	
8a	18		
8n	3	1	
9a	41	1	
9n	8		
10a	123		
10n	42	1	
11a	367	1	
11n	185		
12a	1,288		
12n	888		
13a	4,878	1	
13n	5,110		2
14a	19,536		
14n	27,436	1	2
15a	85,263	1	
15n	168,030	1	6
16a	379,799		
16n	1,008,906	1	10

1701936 knots up to 16 crossings

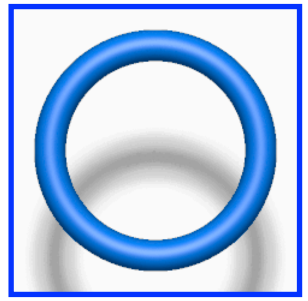
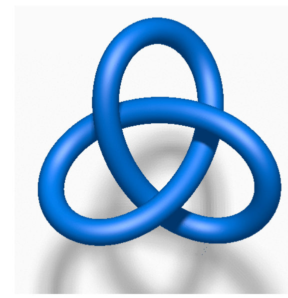
All but 32 are hyperbolic

Hoste, Thistlewaite, Weeks (1998)

Conjecture [Adams]: proportion of hyperbolic knots approaches 1 as crossing number goes to ∞

Dramatis Personae

Knot: $S^1 \subset S^3$; *e.g.*,

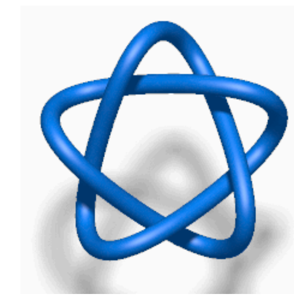
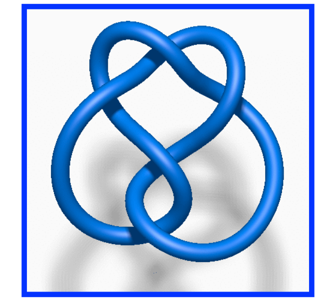



Table A1

# Crossings	# Knots	# Torus	# Sat.
0	1		
3a	1	1	
4a	1		
5a	2	1	
6a	3		
7a	7	1	
8a	18		
8n	3	1	
9a	41	1	
9n	8		
10a	123		
10n	42	1	
11a	367	1	
11n	185		
12a	1,288		
12n	888		
13a	4,878	1	
13n	5,110		2
14a	19,536		
14n	27,436	1	2
15a	85,263	1	
15n	168,030	1	6
16a	379,799		
16n	1,008,906	1	10

1701936 knots up to 16 crossings

All but 32 are hyperbolic

Hoste, Thistlewaite, Weeks (1998)

Conjecture [Adams]: proportion of hyperbolic knots approaches 1 as crossing number goes to ∞

Contradicts another **conjecture**: crossing number of composite knot not less than that of each of its factors

Malyutin (2016)

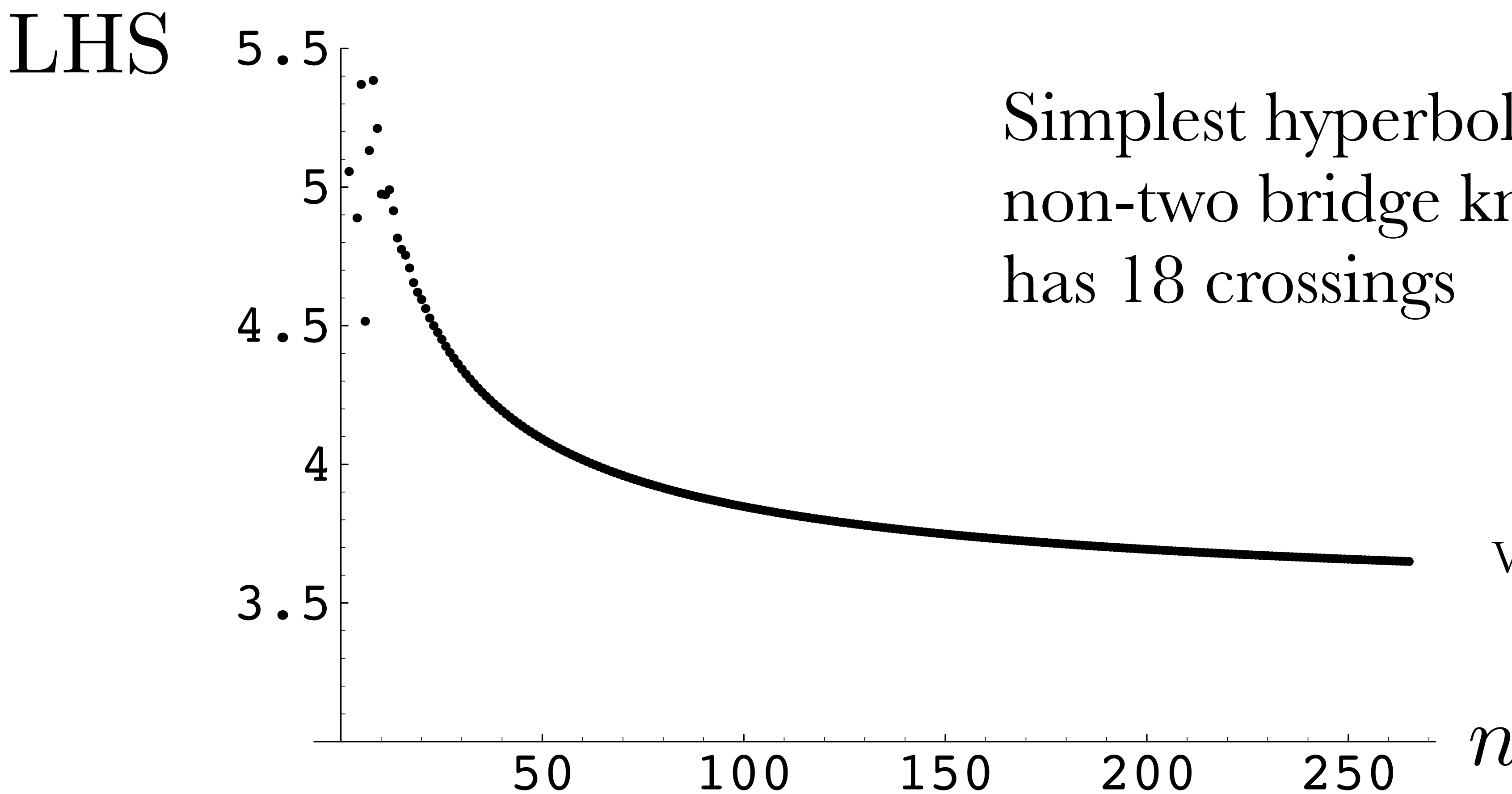
Dramatis Personae

Volume conjecture: $\lim_{n\to\infty} \frac{2\pi \log|J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$

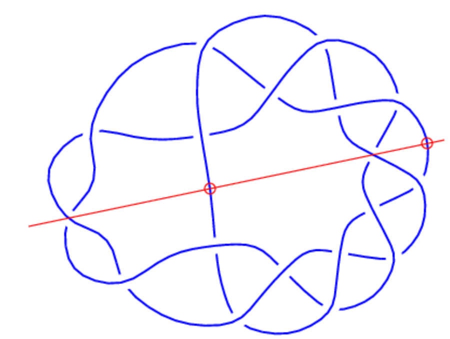
Kashaev (1997)
Murakami x 2 (2001)
Gukov (2005)

$\omega_n = e^{\frac{2\pi i}{n}}$

In fact, we take $n, k \to \infty$



Simplest hyperbolic non-two bridge knot, has 18 crossings



$\sigma_1^2(\sigma_1\sigma_2)^8$

Behavior is not monotonic!

Garoufalidis, Lan (2004)

Dramatis Personae

Volume conjecture: $$\lim_{n\to\infty} \frac{2\pi \log |J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$$

$$\omega_n = e^{\frac{2\pi i}{n}}$$

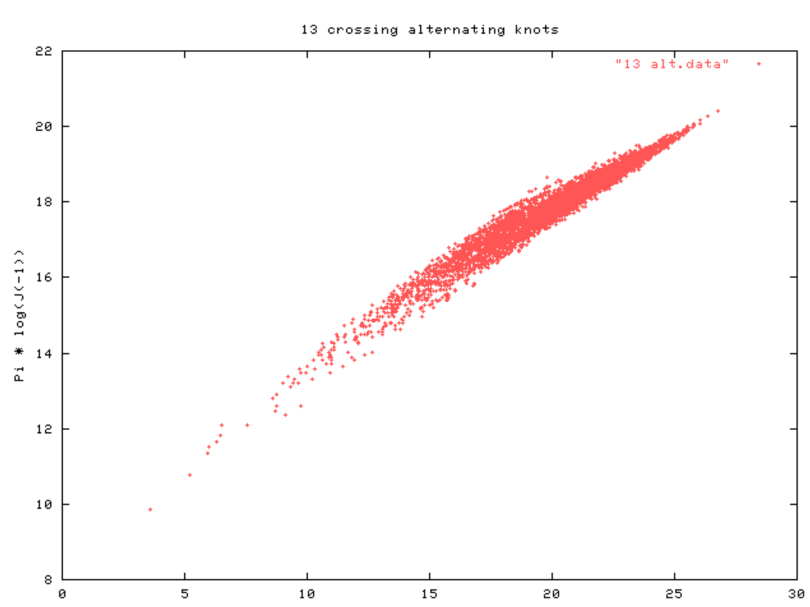
Kashaev (1997)
Murakami x 2 (2001)
Gukov (2005)

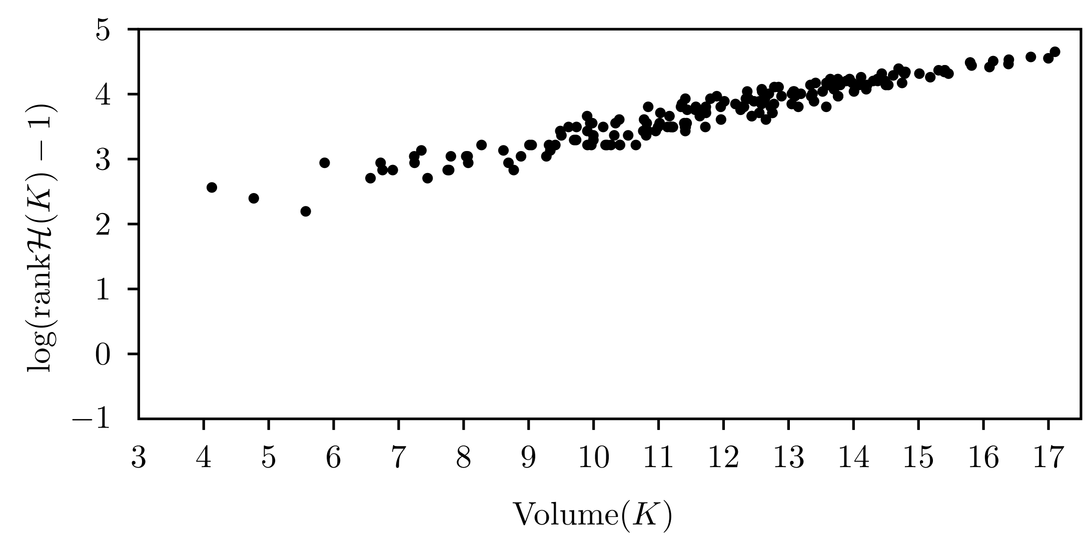
Khovanov homology: a homology theory $\mathcal{H}_K$ whose graded Euler characteristic is $J_2(K;q)$; explains why coefficients are integers

Khovanov (2000)
Bar-Natan (2002)

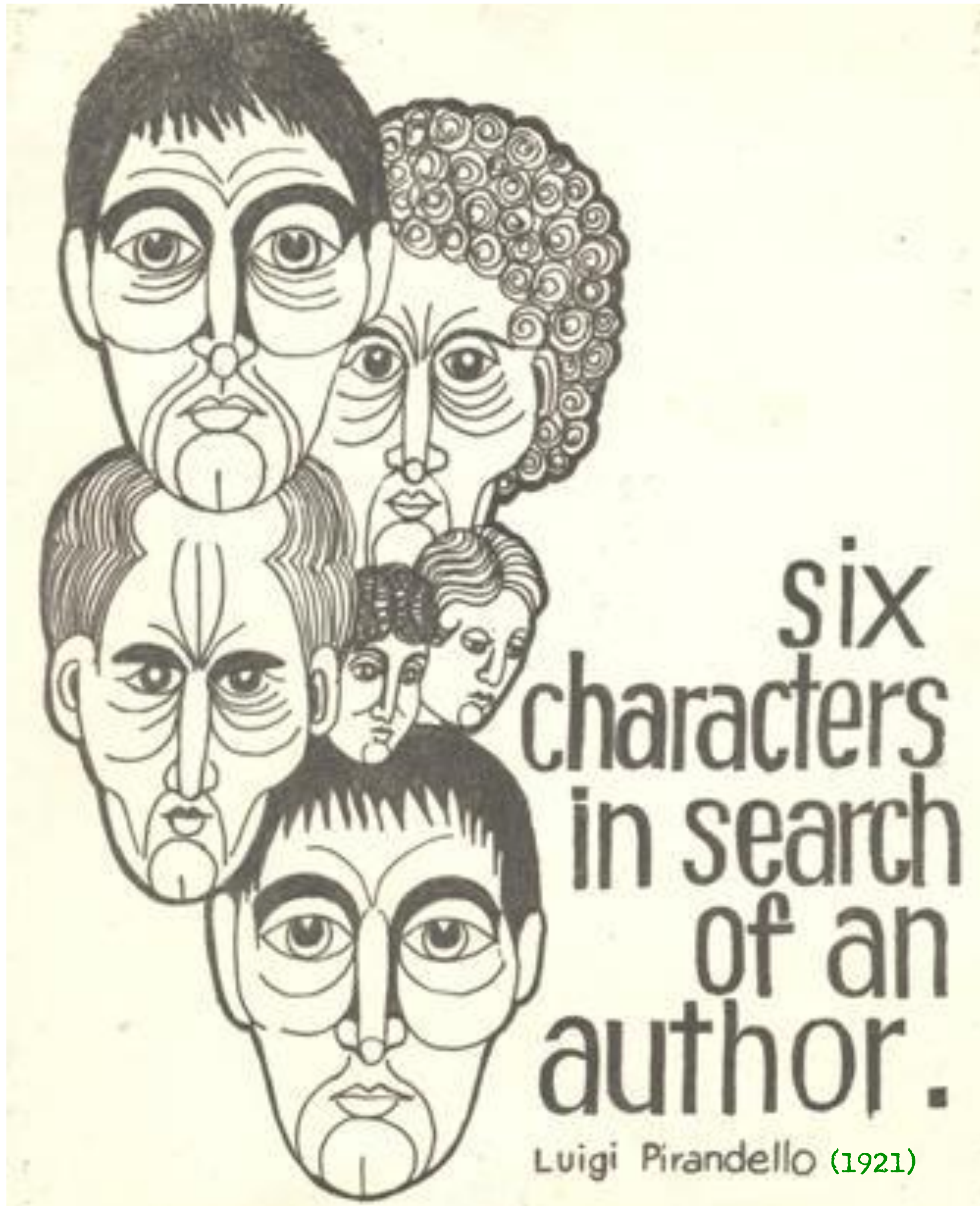
$$\log |J_2(K;-1)| \ , \ \ \log(\mathrm{rank}(\mathcal{H}_K) - 1) \propto \mathrm{Vol}(S^3 \setminus K)$$

Dunfield (2000)
Khovanov (2002)

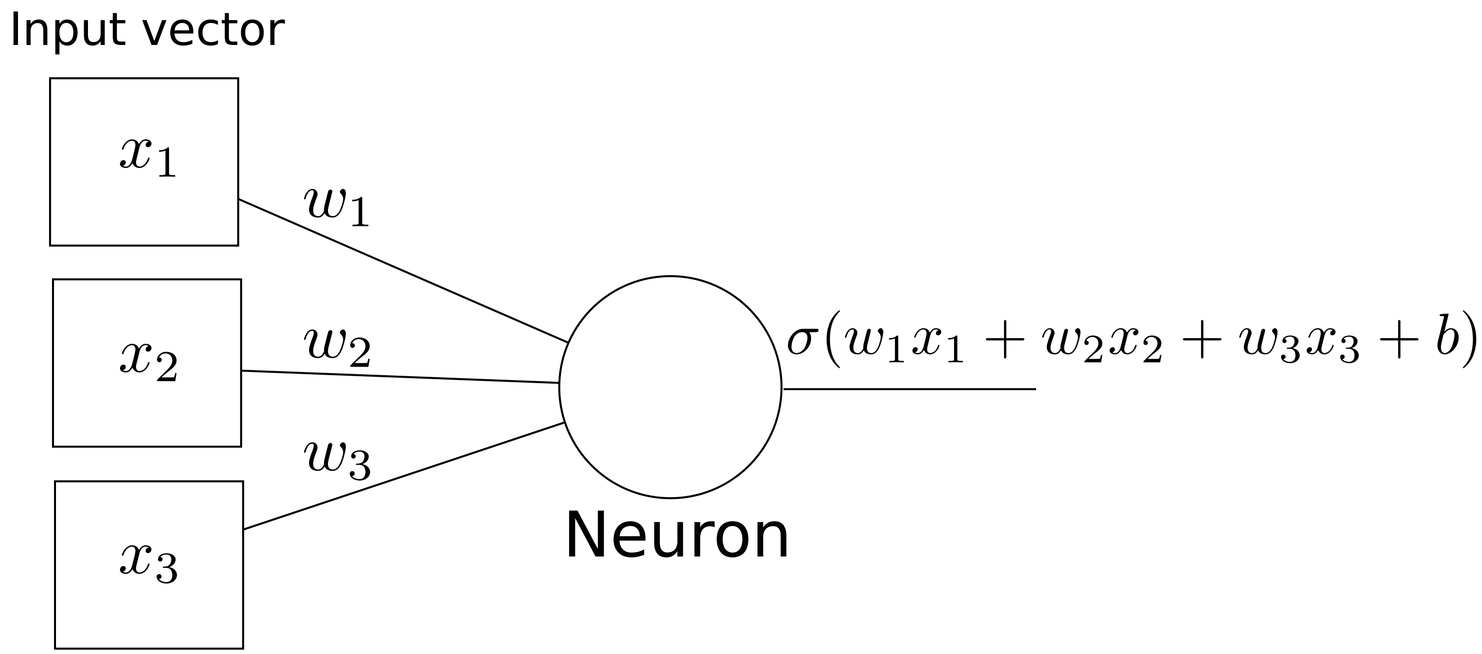




Dramatis Personae



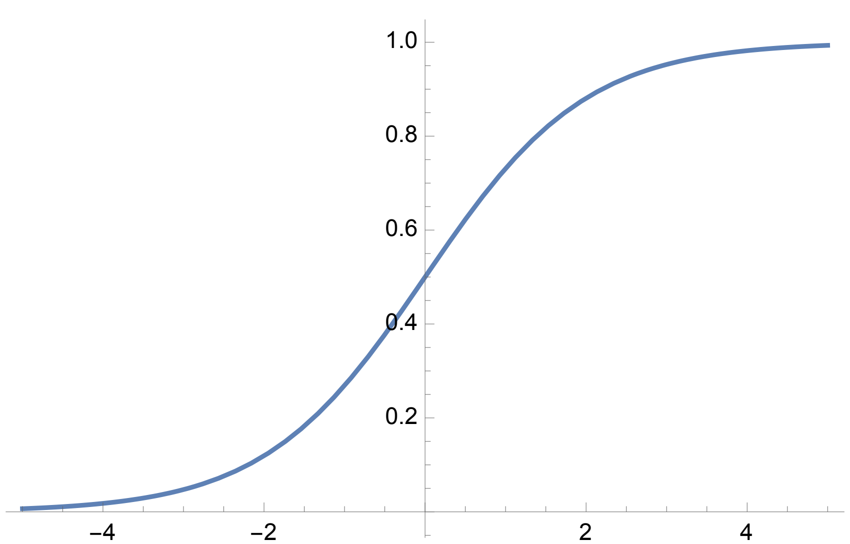
Feedforward Neural Networks



Rosenblatt (1957)

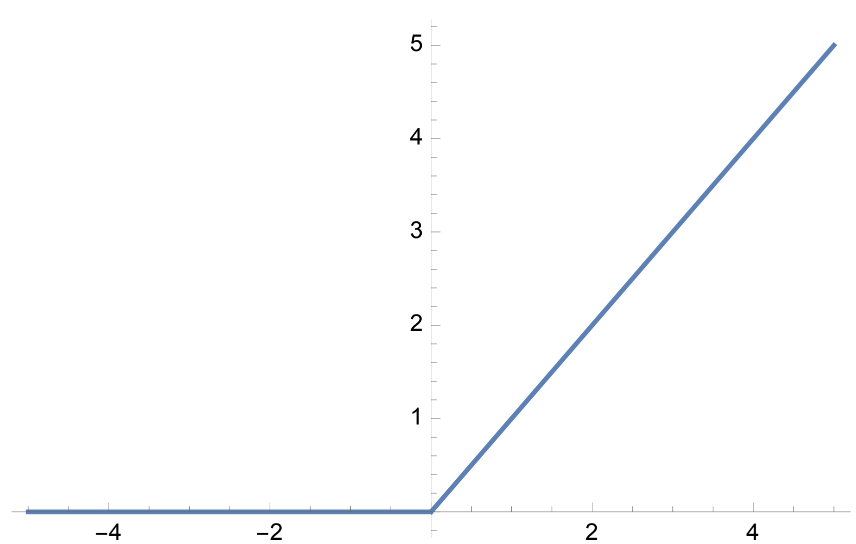
non-linearity

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



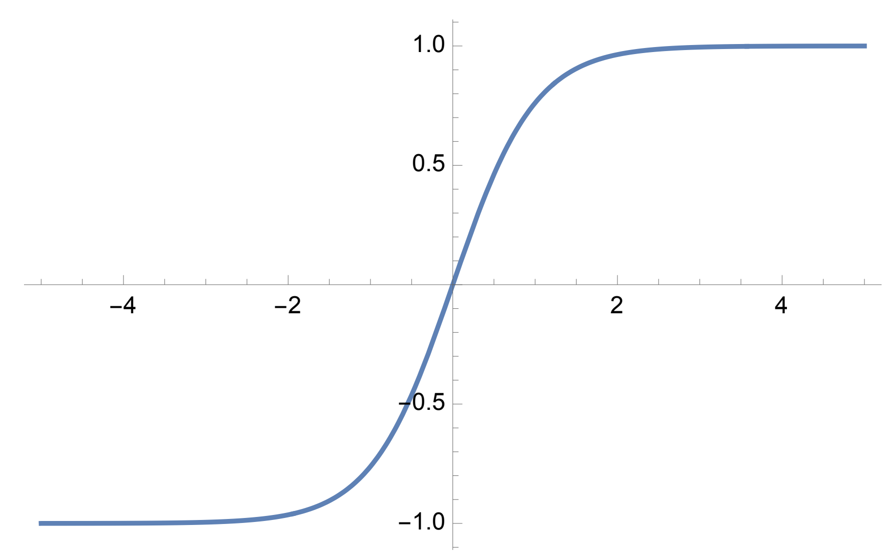
logistic sigmoid

$$\sigma(x) = \max(0, x)$$



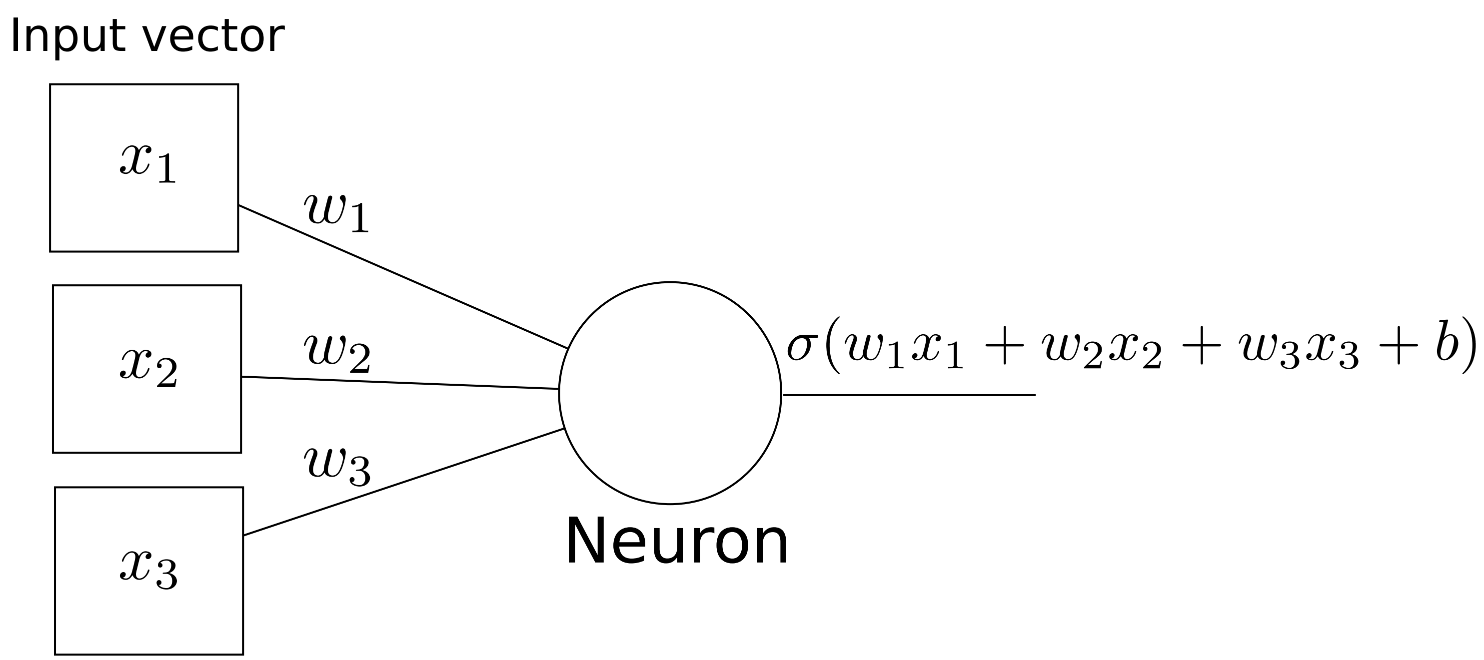
ReLU

$$\sigma(x) = \tanh x$$

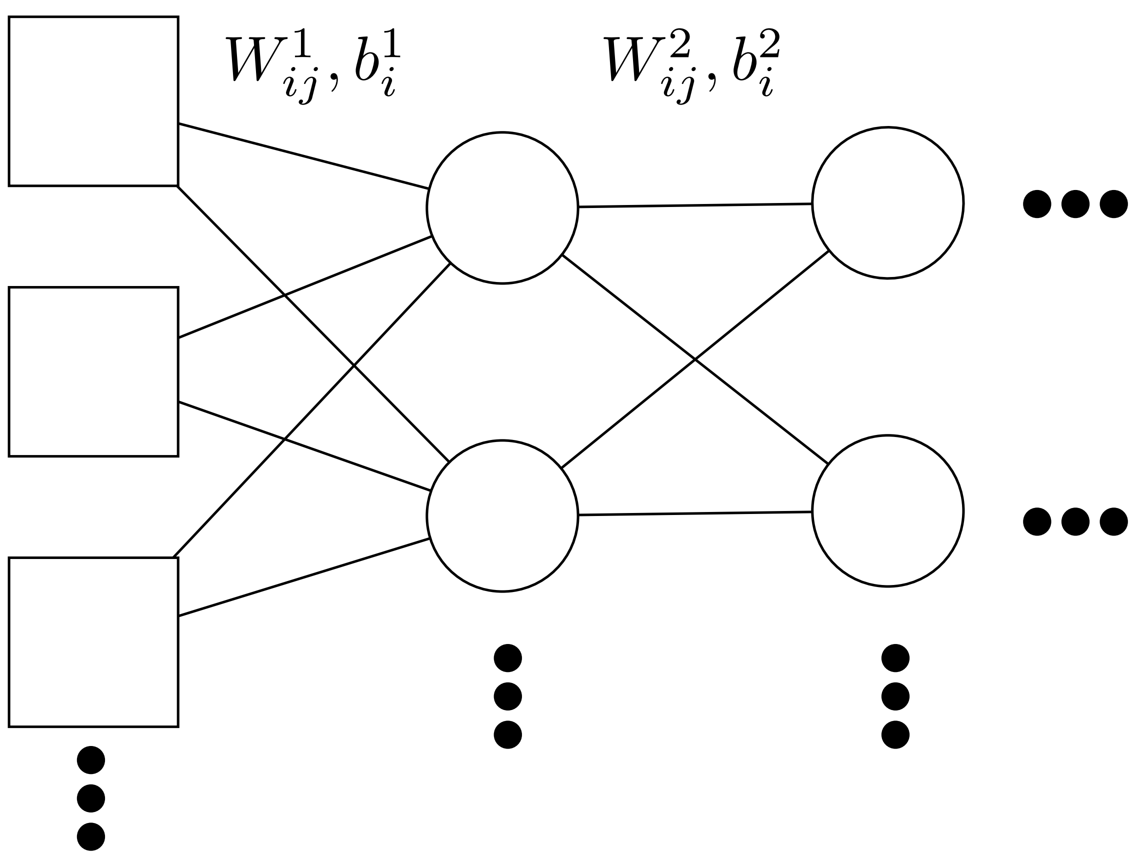


tanh

Feedforward Neural Networks



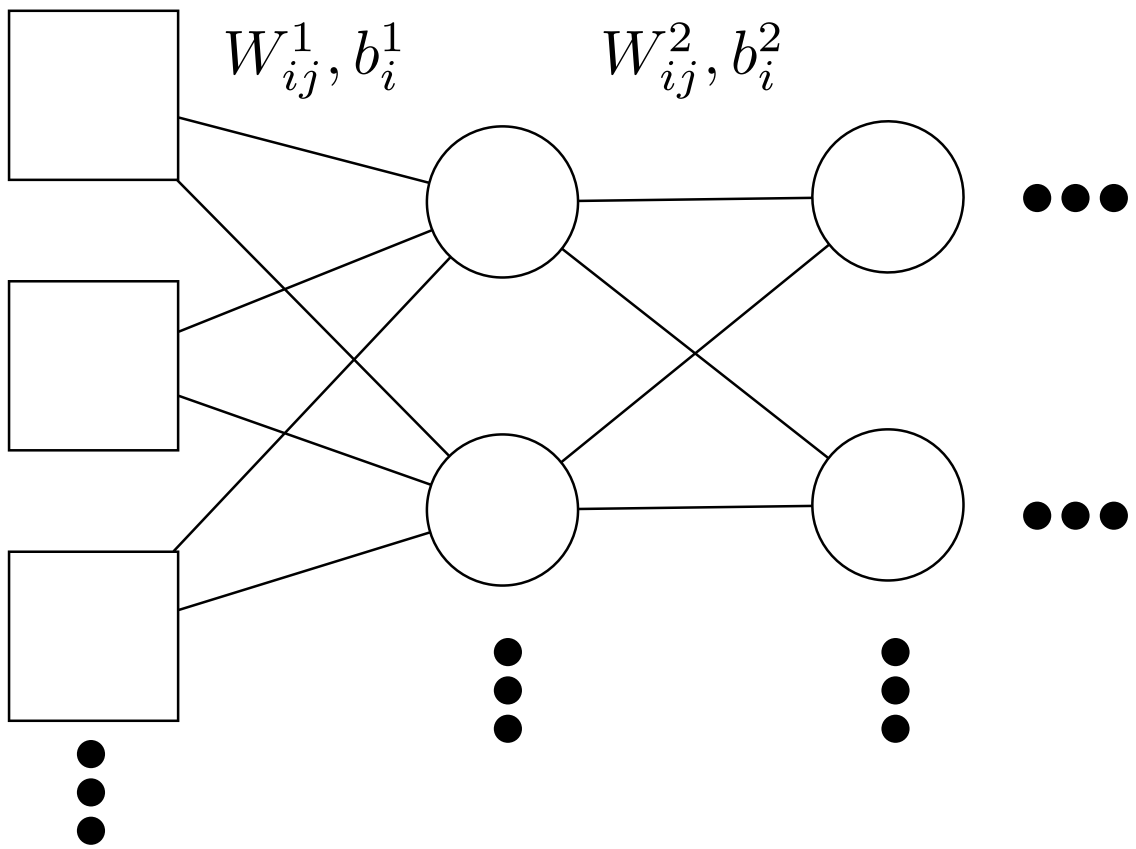
Rosenblatt (1957)



Feedforward Neural Networks

$$f_i : \quad \mathbb{R}^{n_{i-1}} \to \mathbb{R}^{n_i}$$
$$\mathbf{v}_{i-1} \mapsto \mathbf{v}_i = \sigma(\mathbf{v}'_i) \qquad \mathbf{v}'_i = W^i \cdot \mathbf{v}_{i-1} + \mathbf{b}_i$$

$$\mathbf{v}_\ell = f_\ell(f_{\ell-1}(\cdots(f_1(\mathbf{v}_0)\cdots))) \; , \qquad a_{\text{out}} := \sum_{m=1}^{n_\ell} \mathbf{v}_\ell^m$$



Feedforward Neural Networks

$$\begin{array}{llll} f_i : & \mathbb{R}^{n_{i-1}} & \to & \mathbb{R}^{n_i} \\ & \mathbf{v}_{i-1} & \mapsto & \mathbf{v}_i = \sigma(\mathbf{v}'_i) \end{array}$$

number of neurons in i-th layer

$$\mathbf{v}'_i = W^i \cdot \mathbf{v}_{i-1} + \mathbf{b}_i$$

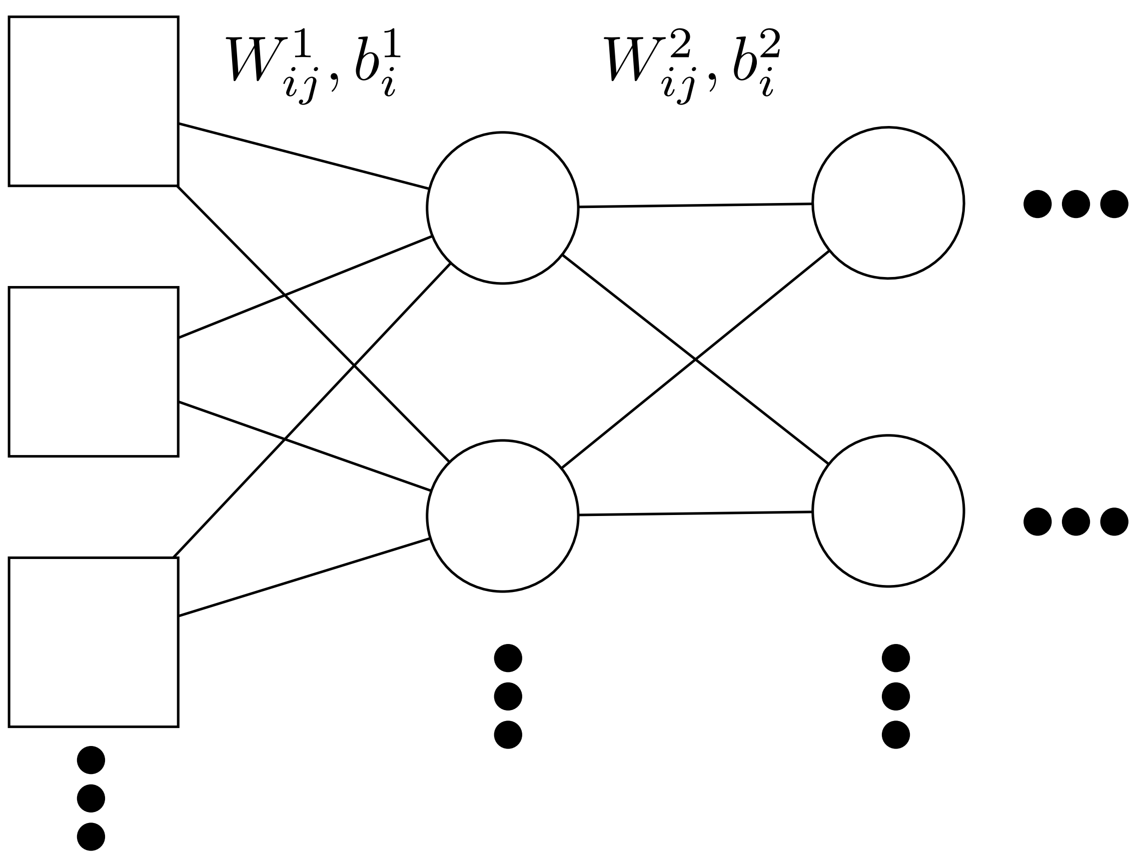
weight matrix

bias vector

$$\mathbf{v}_\ell = f_\ell(f_{\ell-1}(\cdots(f_1(\mathbf{v}_0)\cdots))) \;, \qquad a_{\text{out}} := \sum_{m=1}^{n_\ell} \mathbf{v}_\ell^m$$

number of layers

compare prediction
to ground truth



Feedforward Neural Networks

$$\begin{aligned} f_i : \quad \mathbb{R}^{n_{i-1}} &\to \mathbb{R}^{n_i} \\ \mathbf{v}_{i-1} &\mapsto \mathbf{v}_i = \sigma(\mathbf{v}'_i) \end{aligned}$$

number of neurons in i-th layer

$$\mathbf{v}'_i = W^i \cdot \mathbf{v}_{i-1} + \mathbf{b}_i$$

weight matrix

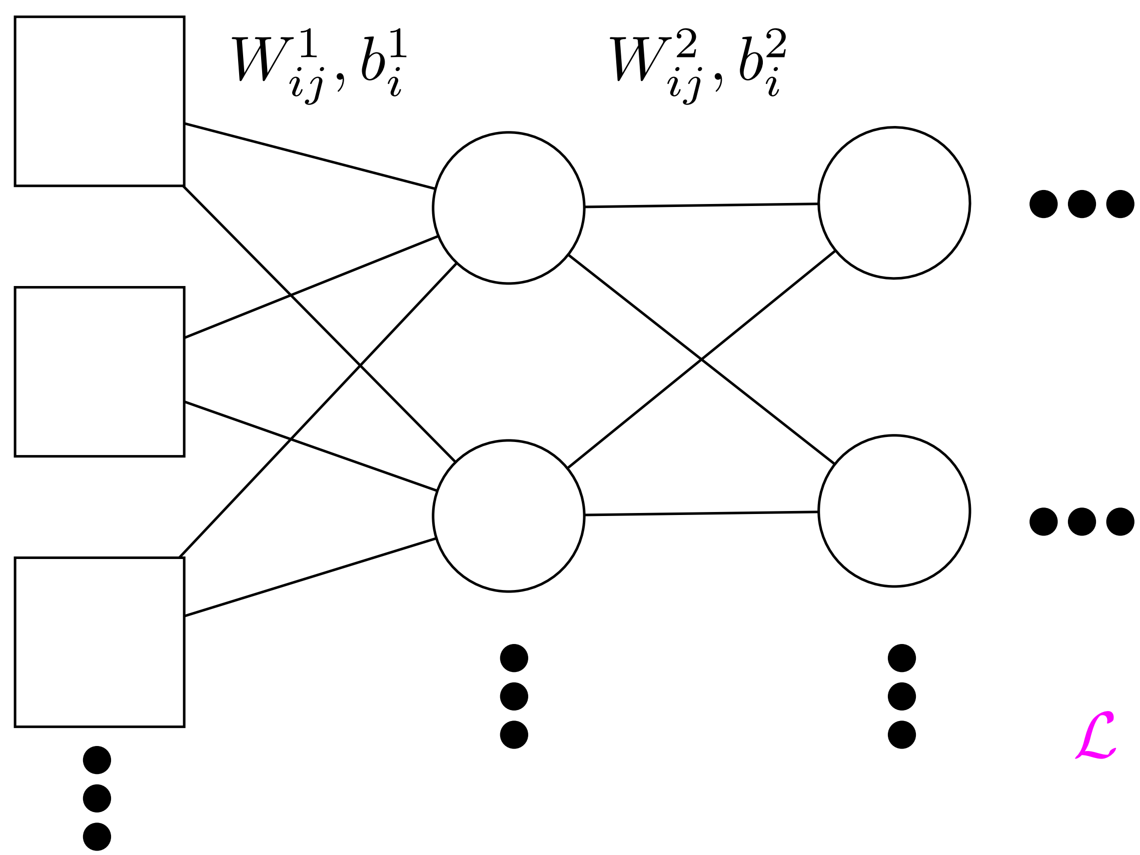
bias vector

$$\mathbf{v}_\ell = f_\ell(f_{\ell-1}(\cdots(f_1(\mathbf{v}_0)\cdots))) \ ,$$

number of layers

$$a_{\text{out}} := \sum_{m=1}^{n_\ell} \mathbf{v}_\ell^m$$

compare prediction
to ground truth

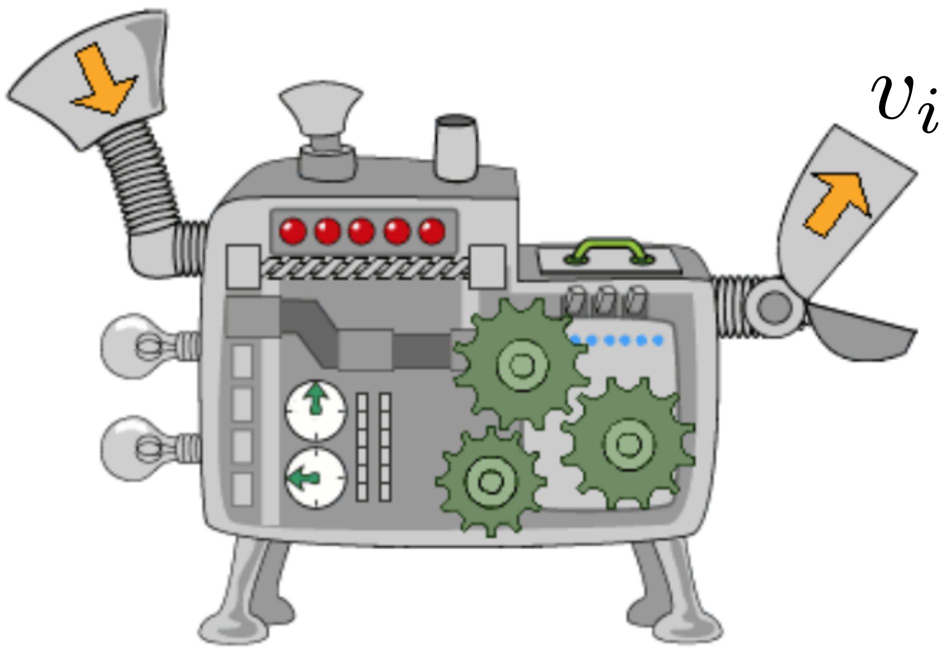


tune weights and biases
to minimize **loss function**
during **training**

$$\mathcal{L} = \sum_{\text{batch}} |\text{predicted} - \text{actual}|^2$$

Neural Network

J_i



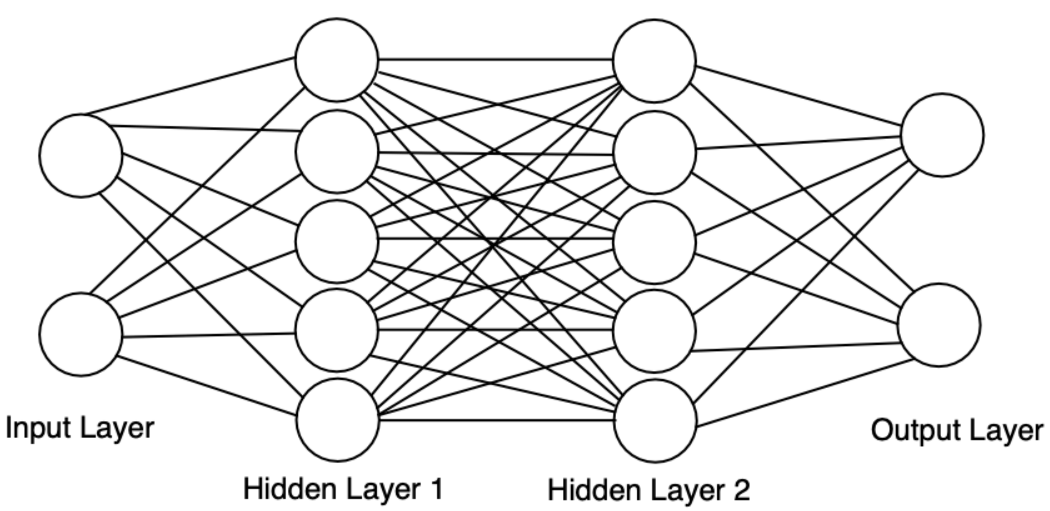
v_i

$$\{J_1, \ldots, J_n\} \longrightarrow \{v_1, \ldots, v_n\}$$

$$J_i \in T$$

$$\{J'_1, \ldots, J'_m\} \longrightarrow ???$$

$$J'_i \in T^c$$



$\vec{J}$ $\quad 100 \times 18 \quad 100 \times 100 \quad \sum_{a=1}^{100}$

12000 hyperparameters

Jones polynomials are represented as 18-vectors

$$\vec{J}_K = (\text{min}, \text{max}, \text{coeffs}, 0, \ldots, 0)$$

Two layer neural network in Mathematica

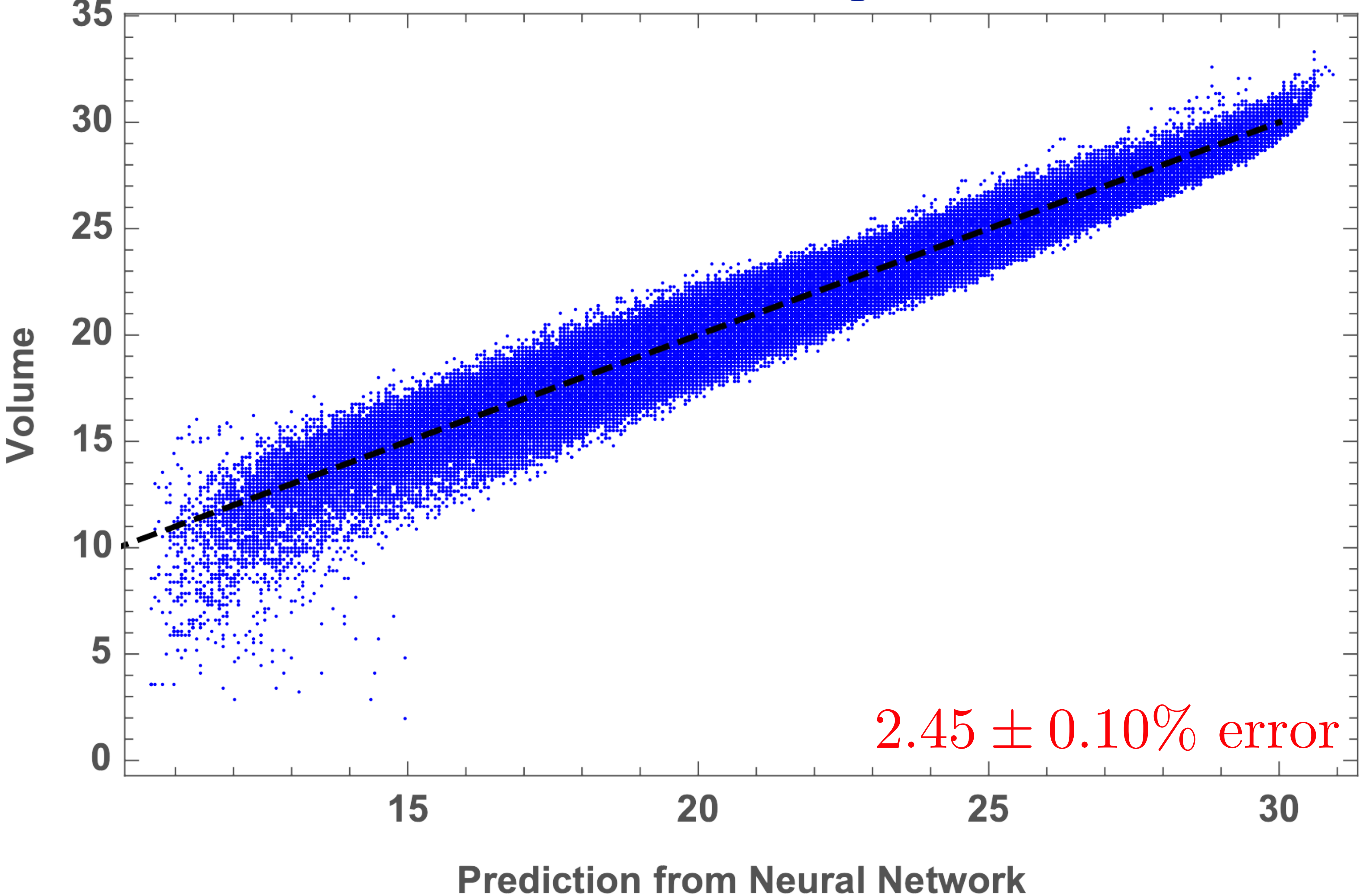
$$f_\theta(\vec{J}_K) = \sum_a \sigma\left(W^2_\theta \cdot \sigma(W^1_\theta \cdot \vec{J}_K + \vec{b}^1_\theta) + \vec{b}^2_\theta\right)^a$$

Logistic sigmoids for the hidden layers

Mean squared loss function

VJ, Kar, Parrikar (2019)

Volume from Jones



trained on 10% of the $313,209$ knots up to 15 crossings

VJ, Kar, Parrikar (2019)

Mathematica Code

knot run.nb

```
In[•]:= f[frac_] :=
  Module[{y, list, listcomp, joneslist, vollist, joneslistcomp, vollistcomp,
    test, aa, KnotNet0, TrainedKnotNet0},
   y = frac;
   list = Sort[RandomSample[Range[1, len], IntegerPart[len y]]];
   listcomp = Complement[Table[i, {i, 1, len}], list];
   joneslist = Table[joneses[[list[[i]]]], {i, 1, Length[list]}];
   vollist = Table[volumes[[list[[i]]]], {i, 1, Length[list]}];
   joneslistcomp = Table[joneses[[listcomp[[i]]]], {i, 1, Length[listcomp]}];
   vollistcomp = Table[volumes[[listcomp[[i]]]], {i, 1, Length[listcomp]}];
   KnotNet0 = NetChain[{LinearLayer[100], ElementwiseLayer[LogisticSigmoid],
      LinearLayer[100], ElementwiseLayer[LogisticSigmoid], SummationLayer[]},
     "Input" → {maxlen}];
   TrainedKnotNet0 = NetTrain[KnotNet0, joneslist → vollist] ;
   test = TrainedKnotNet0[joneses];
   Show[ListPlot[Table[{test[[i]], volumes[[i]]}, {i, 1, len}],
     AxesLabel → {"Predicted Volume", "Volume"}],
    Plot[x, {x, 0, 40}, PlotRange → Full, PlotStyle → {Black, Dashed}]]
  ]
f[.1]
```

Result

$$v_i = f(J_i) + \text{small corrections}$$

- J_i does not uniquely identify a knot

  e.g., 4_1 and K11n19 have same Jones polynomial, different volumes

- $174,619$ unique Jones polynomials

  2.83% average spread in volumes for a Jones polynomial

  intrinsic mitigation against overfitting

- Same applies to 1,701,903 hyperbolic knots up to 16 crossings

  841,139 unique polynomials

  (database compiled from **Knot Atlas** and `SnapPy`)

VJ, Kar, Parrikar (2019)

Result

$$v_i = f(J_i) + \text{small corrections}$$

- **Universal Approximation Theorem:** feedforward neural network, sigmoid activation function, single hidden layer with finite number of neurons can approximate continuous functions on compact subsets of $\mathbb{R}^n$

Cybenko (1989)
Hornik (1991)
cf. Kidger, Lyona (2019)

- Surprise here is simplicity of architecture that does the job

- We want a **not** machine learning knot result

Entr'acte

$$v_i = f(J_i) + \text{small corrections}$$

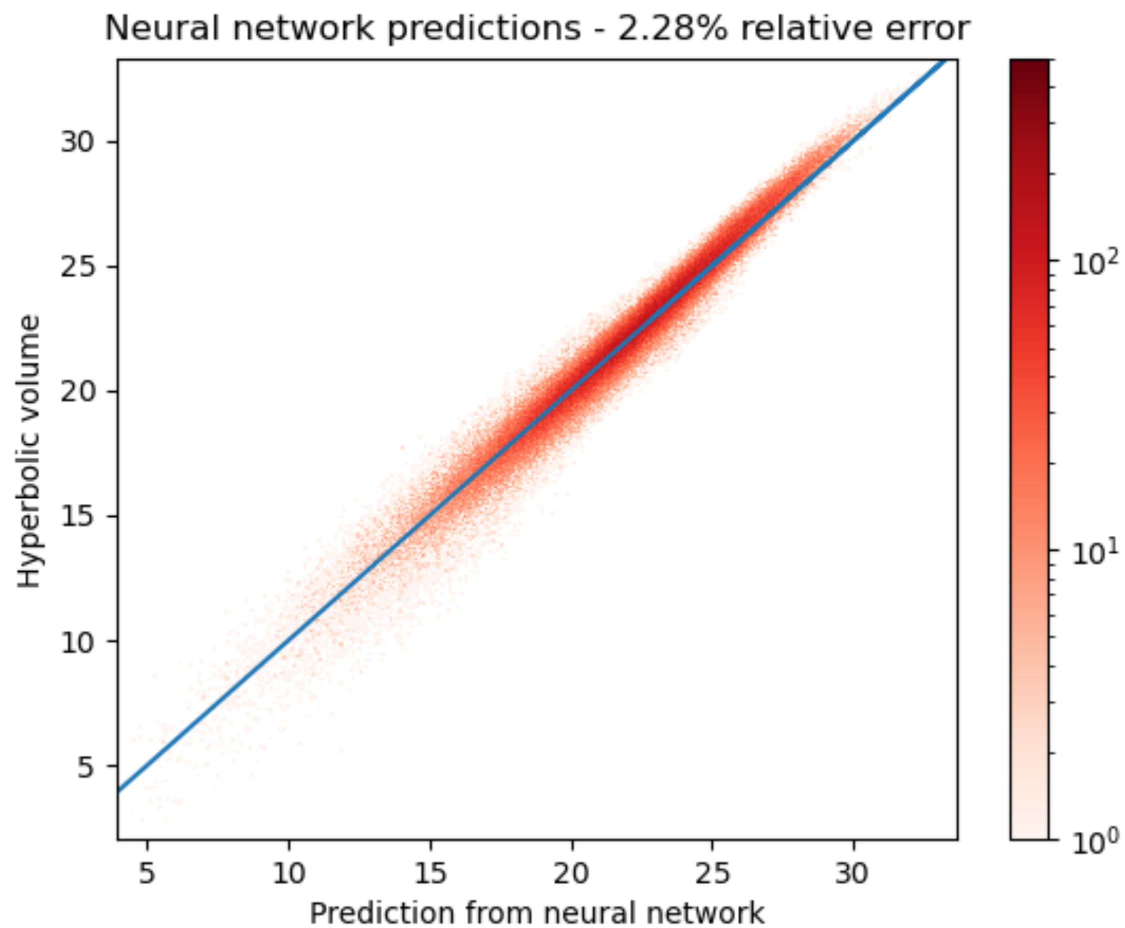
We seek to reverse engineer the neural network to obtain an analytic expression for the volume as a function of the Jones polynomial

To interpret the formula, we use machinery of analytically continued Chern–Simons theory

No Degrees Needed

- Suppose we drop the degrees and provide only the coefficients; Jones polynomial is no longer recoverable from the input vector

- Results are unchanged!

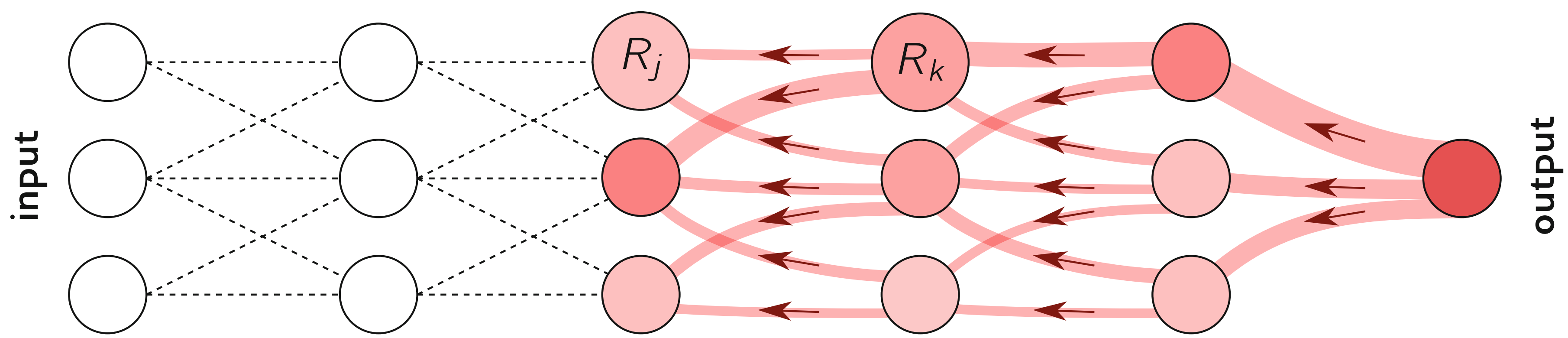


__N.B.__: we have switched to `Python 3` using `GPU-Tensorflow` with `Keras` wrapper
two hidden layers, 100 neurons/layer, ReLu activation, mean squared loss, `Adam` optimizer

Craven, VJ, Kar (2020)

Layer-wise Relevance Propagation

- To determine which inputs carry the most weight, propagate backward starting from output layer employing a conservation property



Montavon et al. (2019)

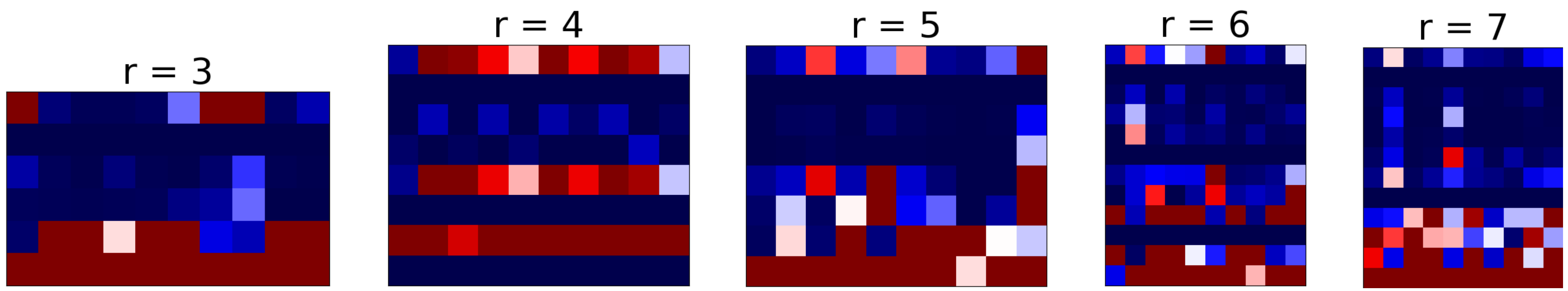
- Compute relevance score for a neuron using activations, weights, and biases

$$R_j^{m-1} = \sum_k \frac{a_j^{m-1} W_{jk}^m + N_{m-1}^{-1} b_k^m}{\sum_l a_l^{m-1} W_{lk}^m + b_k^m} R_k^m \quad , \quad \sum_k R_k^m = 1$$

j^{th} neuron in layer $m-1$

Craven, VJ, Kar (2020)

Layer-wise Relevance Propagation



- Each column is a single input corresponding to evaluations of the Jones polynomial at phases $e^{\frac{2\pi i p}{r+2}}$, $0 \leq 2p \leq r+2$, $p \in \mathbb{Z}$

- Ten different knots

- We show the relevances (red is most relevant) and notice that the same input features light up

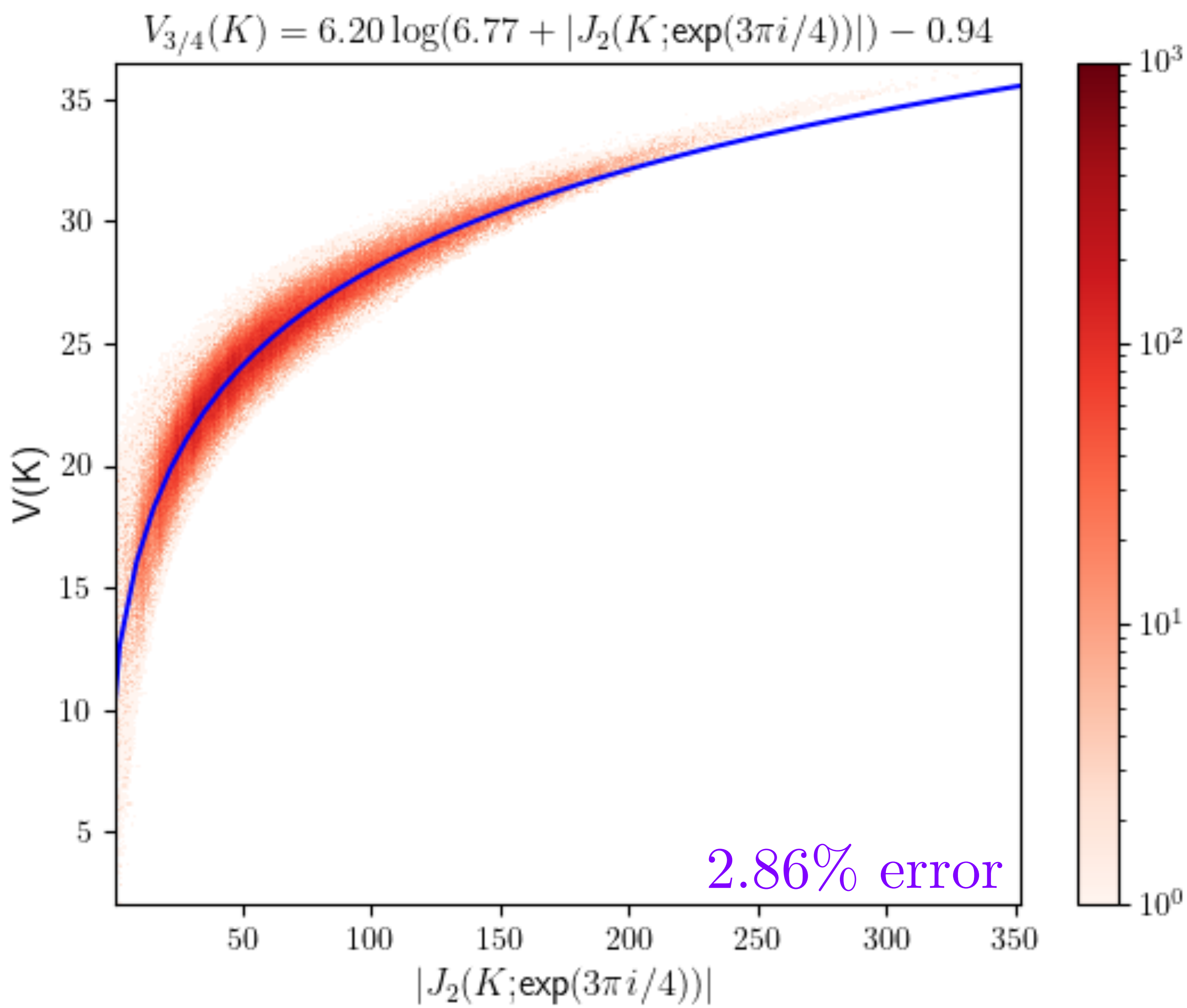
Craven, VJ, Kar (2020)

Phenomenological Function

$$\omega = e^{\frac{3\pi i}{4}}$$

$$\Updownarrow$$

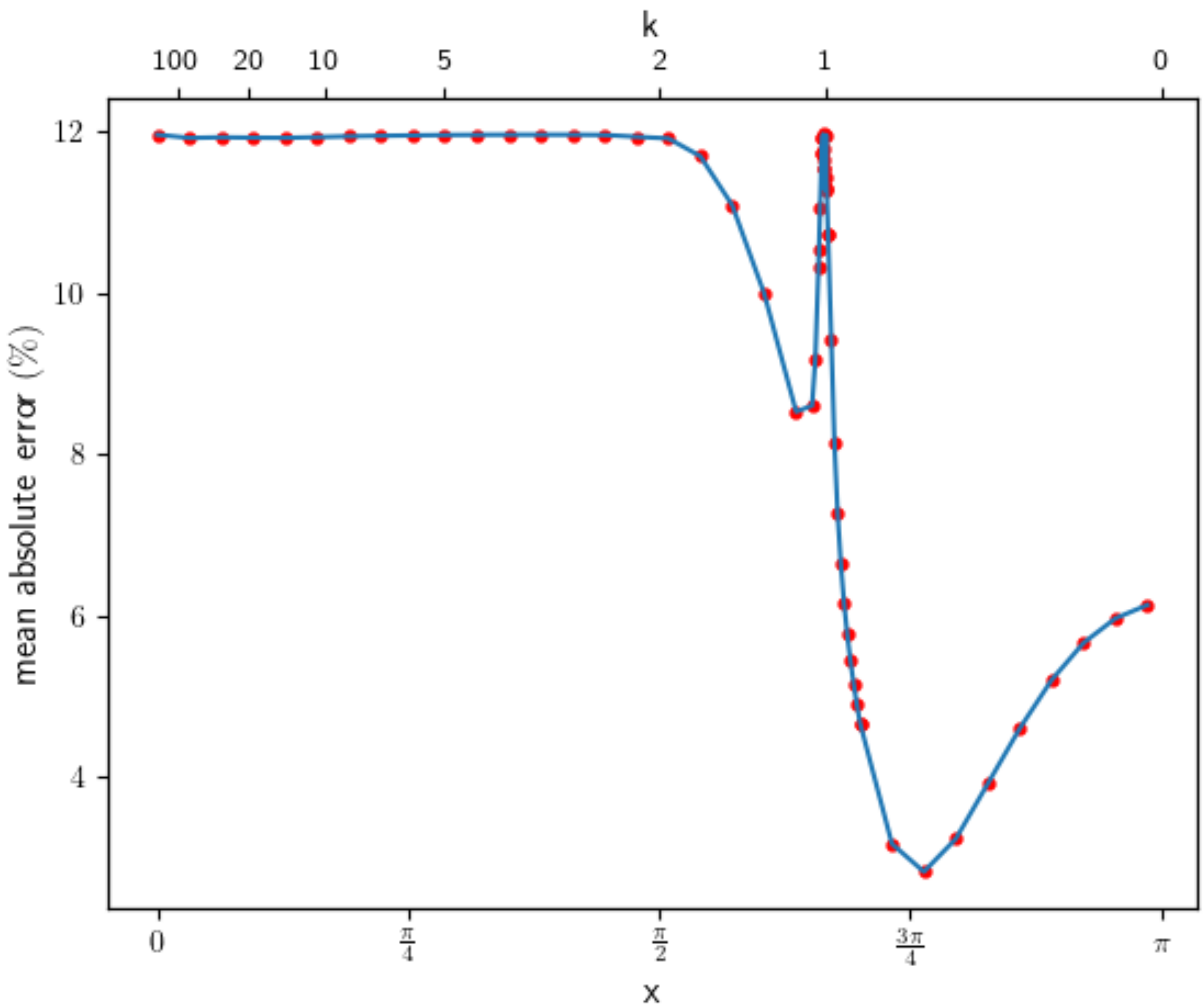
$$k = \frac{2}{3}$$



- 1,701,903 knots up to 16 crossings
- Behavior explainable in Witten's analytic continuation of Chern–Simons theory

Craven, VJ, Kar (2020)

The Shape of Things



Craven, VJ, Kar (2020)

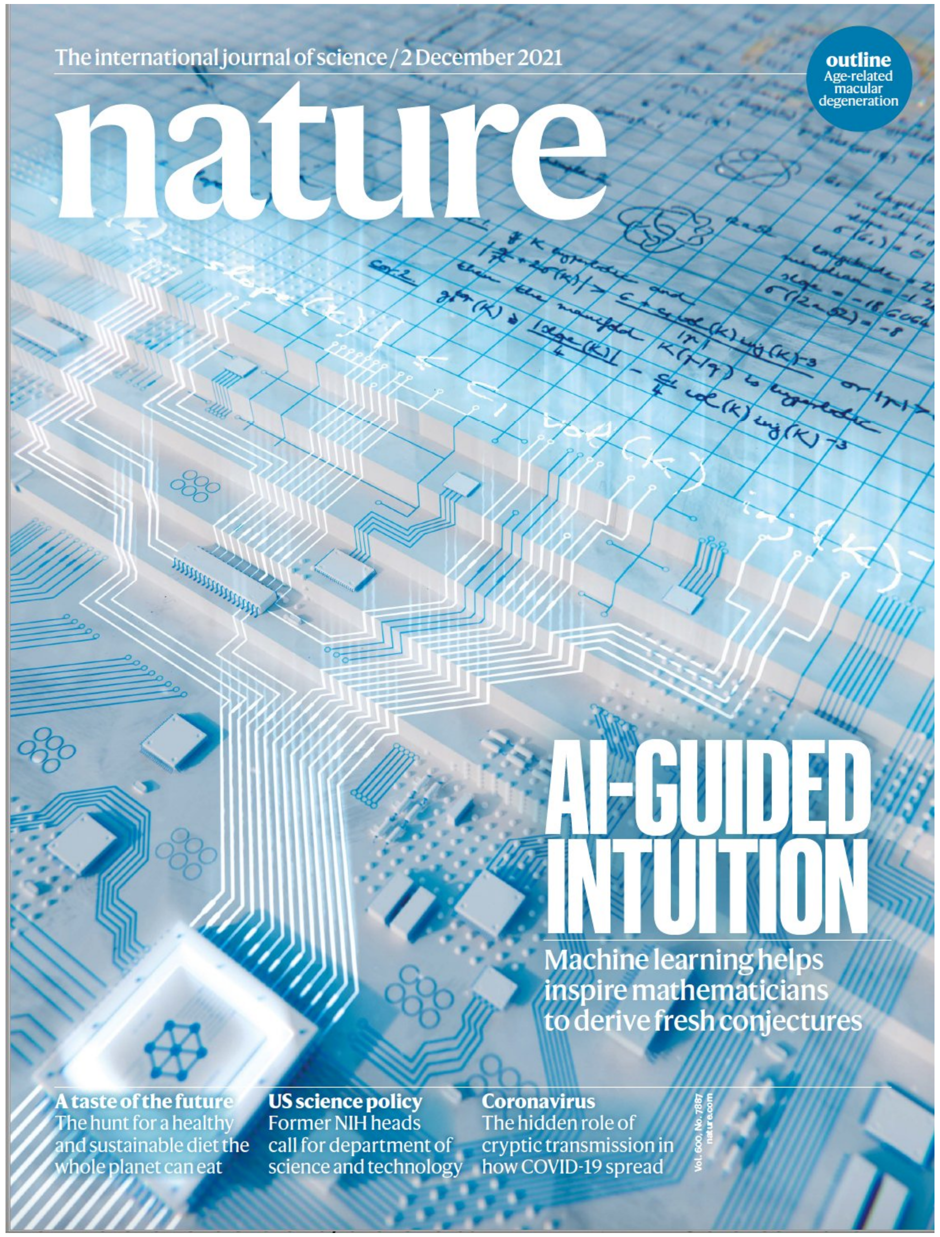
Explanation

The approximation formula works well for levels k for which $\mathcal{A}_+$ makes a contribution to the Chern–Simons path integral, and its accuracy increases with fraction of knots in dataset that receive such a contribution

$$Z \sim e^{iS(\mathcal{A}_+)}\left(1 - e^{2\pi i k}\right)$$

Craven, VJ, Kar (2020)

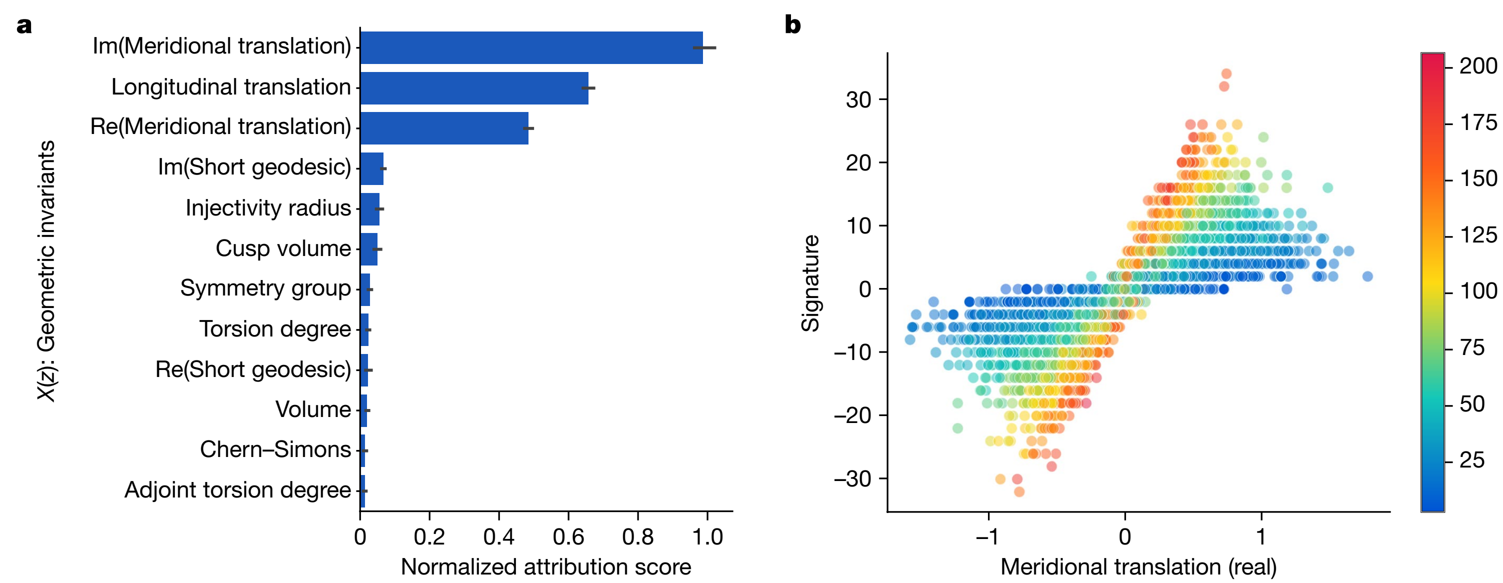
Other Experiments



- Obtain a theorem in knot theory from saliency analysis of ML experiments

Theorem: There exists a constant c such that for every hyperbolic knot,

$$|2\sigma(K) - \mathrm{slope}(K)| \leq c\,\mathrm{Vol}(K)\,\mathrm{inj}(K)^{-3}$$



Davies, et al. (2021)

- See also: Hughes (2016)

  Levit, Hajij, Sazdanovic (2019)

  Gukov, Halverson, Ruehle, Sulkowski (2020)

  Craven, Hughes, VJ, Kar (2021, 2022)

  Gukov, Halverson, Manolescu, Ruehle (2023)

Poincaré Conjecture

- A homotopy *n*-sphere has the same homotopy and homology groups as S^n

- **Conjecture:** Any homotopy *n*-sphere is homeomorphic, PL-isomorphic, diffeomorphic to S^n

  — homeomorphic is always true

  — piecewise linear isomorphic is true for $n \neq 4$

  — $n = 1$ is trivial

  — $n = 2$ proved by Poincaré (1907) and Koebe (1907)

  — $n \geq 5$ wins Fields Medal for Smale (1966); *cf.* Newman (1966)

  — $n = 4$ wins Fields Medal for Freedman (1986)

  — $n = 3$ wins Fields Medal for Perelman (2006, declined)

SPC4

- **Conjecture:** Any manifold homotopy equivalent to S^n is also diffeomorphic to S^n

isomorphism of smooth manifolds, *i.e.*, f, f^{-1} continuously differentiable

- False generally, but true if $n = 1, 2, 3, 5, 6, 12, 56, 61$

- $n = 7$: 28 different smooth structures on the sphere — Milnor (1956)

Remark: These are interpreted as gravitational instantons — Witten (1985)

dimension	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
# structures	1	1	1	?	1	1	28	2	8	6	992	1	3	2	16256	2	16	16

Observation: spikes at $n = 3 \bmod 4$

Kervaire, Milnor (1963)

SPC4

- **Conjecture:** Any manifold homotopy equivalent to S^n is also diffeomorphic to S^n

  isomorphism of smooth manifolds, *i.e.*, f, f^{-1} continuously differentiable

- False generally, but true if $n = 1, 2, 3, 5, 6, 12, 56, 61$

- $n = 7$: 28 different smooth structures on the sphere Milnor (1956)

  Remark: These are interpreted as gravitational instantons Witten (1985)

- $n = 4$: answer unknown; equivalent to PL-isomorphic statement
  [also unknown for $n = 126$]

- If certain knots have certain topological invariants, this supplies counterexamples to SPC4

$$S^3 = \partial B_4 \ , \quad K = \partial \Sigma$$

$$K \subset S^3 \ , \quad \Sigma \subset B_4$$

SPC4: Any smooth manifold B with $\partial B = S^3$ homotopy equivalent to B_4 is also diffeomorphic to B_4

CALABI-YAU MANIFOLDS

The Forces of Nature

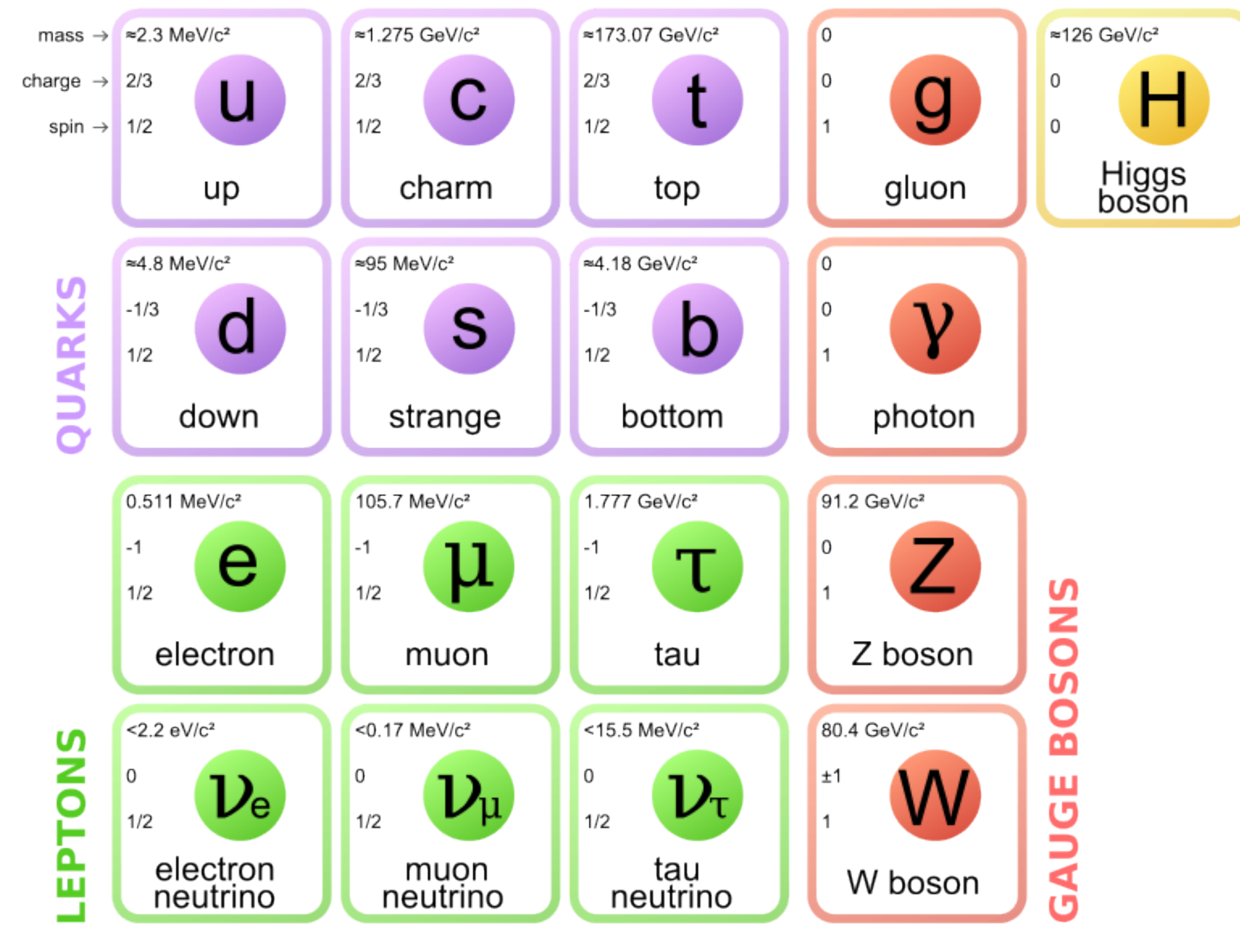
- General relativity (gravity) is geometry of spacetime in presence of energy

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G_N}{c^4} T_{\mu\nu}$$

- Standard Model of particle physics

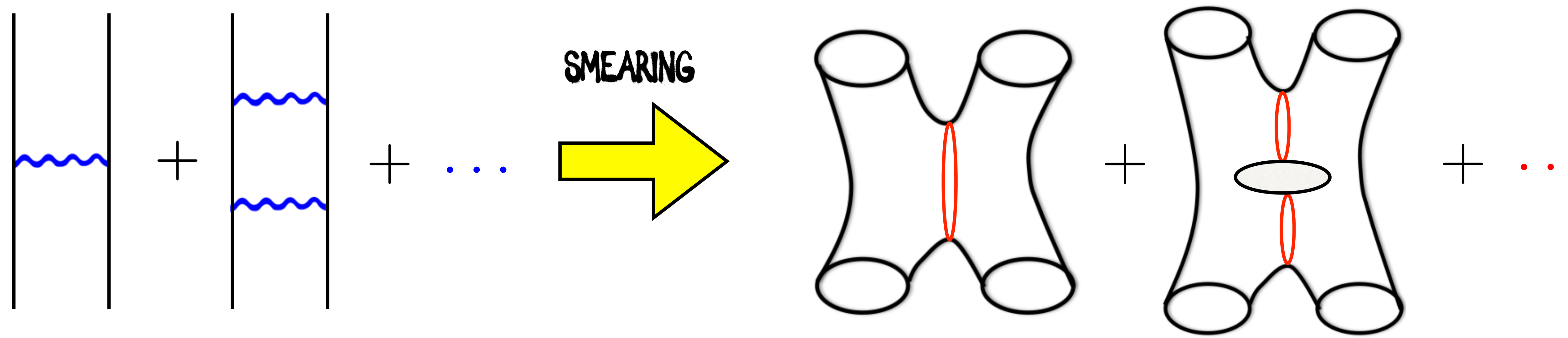
electromagnetism
weak force
strong force
Higgs effect

$$\alpha_{\rm exp}^{-1} = 137.035999139(31)$$
$$\alpha_{\rm th}^{-1} = 137.035999173(35)$$



String Theory

Gravity as a QFT | **Gravity from String Theory**



$d = 3 + 1: \ \infty$ | $d = 9 + 1: \ \text{finite}$

- Map from string worldsheet to target space (spacetime)
  $X^\mu: \ \Sigma \to \mathcal{M}$

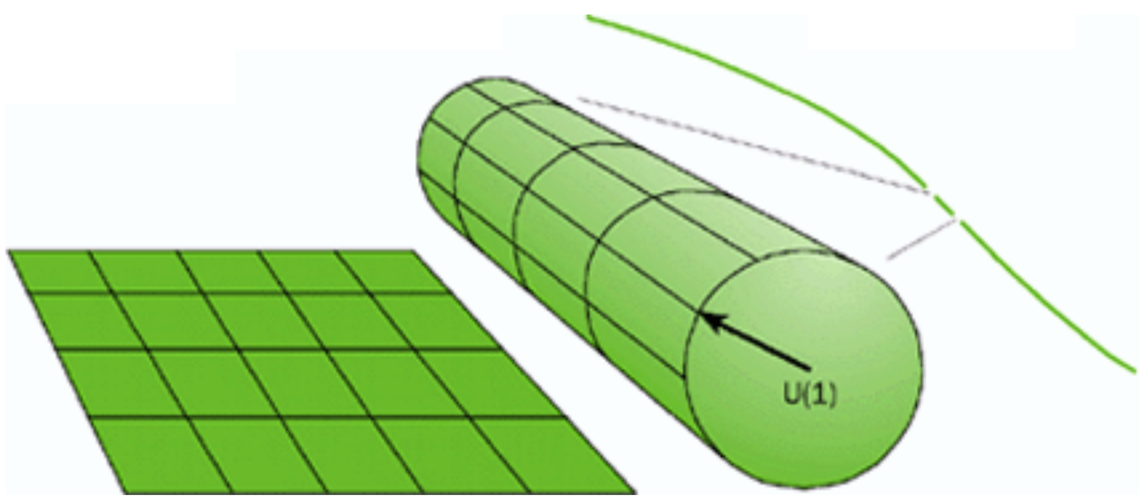
- X^μ are fields on string worldsheet; consistency of this QFT gives Einstein equations in ten dimensional spacetime

Compactification

- Non-gravitational interactions are not encoded as geometry

  Theorem [Coleman–Mandula]**:** symmetry group in 4 dimensions is Poincaré x internal

- **Clever loophole:** internal symmetries may arise from higher dimensional geometry



$\mathbb{R}^{1,3} \times S^1$

Kaluza–Klein: 5d Einstein equations give 4d Einstein + Maxwell equations

$S^1 \longleftrightarrow U(1)$

Geometric Engineering

12 dimensions F-theory

T^2

10 dimensions Heterotic $E_8 \times E_8$ Type IIB

$\vdots$

CY_3 CY_4

4 dimensions 4d $\mathcal{N} = 1$ effective field theory

e.g., minimal supersymmetric Standard Model

The Unreal World

- The objective is to obtain the real world from a string compactification

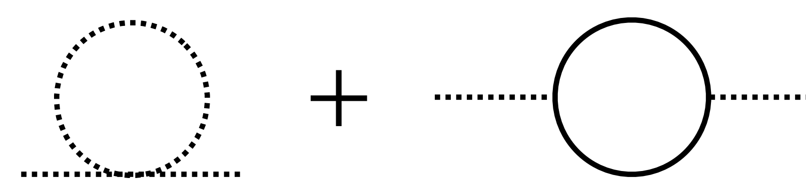
- We would happily settle for a modestly unreal world

$$\boxed{\mathcal{N} = 1 \text{ supersymmetry in 4 dimensions}}$$

No experimental evidence so far!

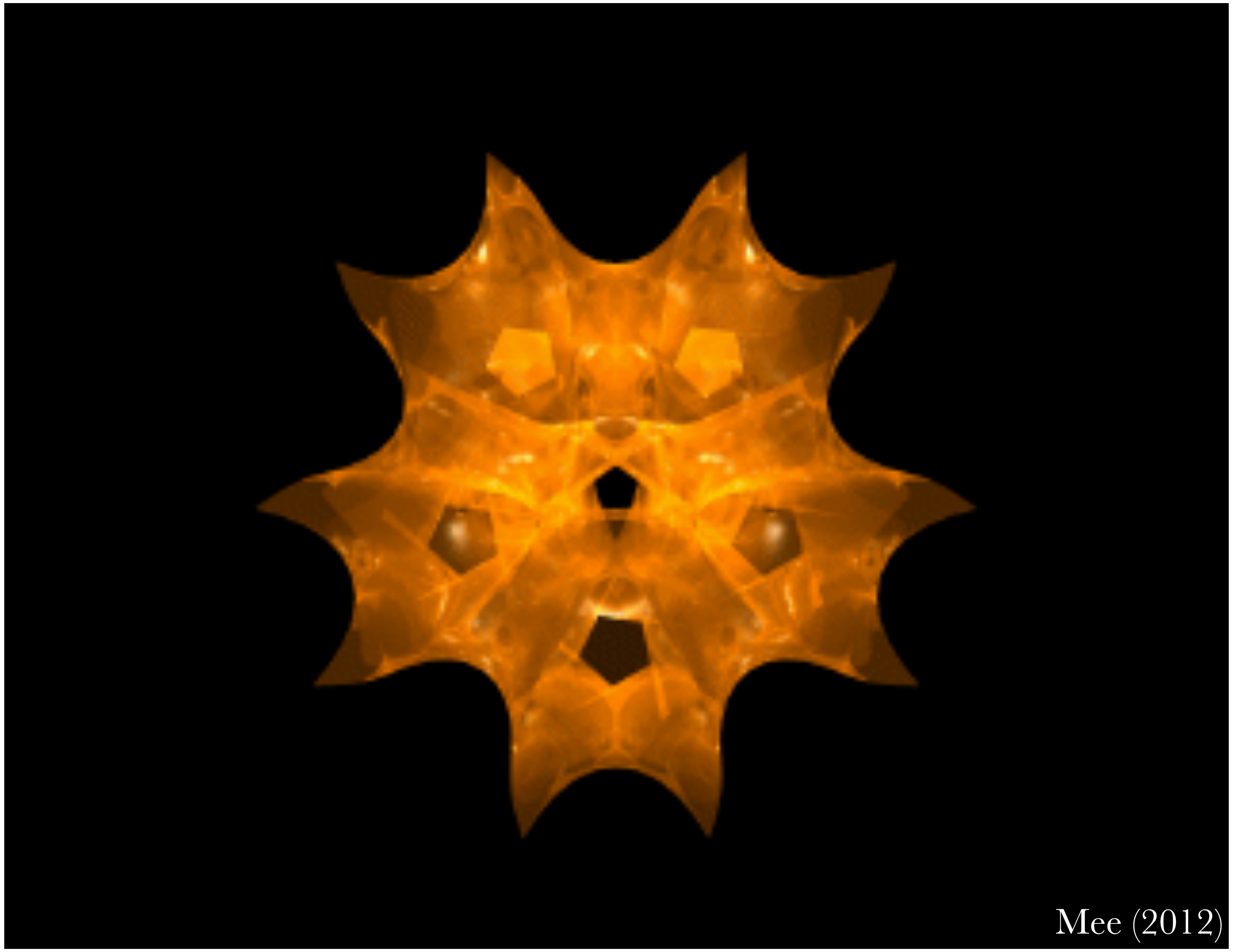
$$Q|\lambda\rangle \sim |\lambda \pm \frac{1}{2}\rangle$$

$$|\text{boson}\rangle \longleftrightarrow |\text{fermion}\rangle$$



$$m_H \ll m_P$$

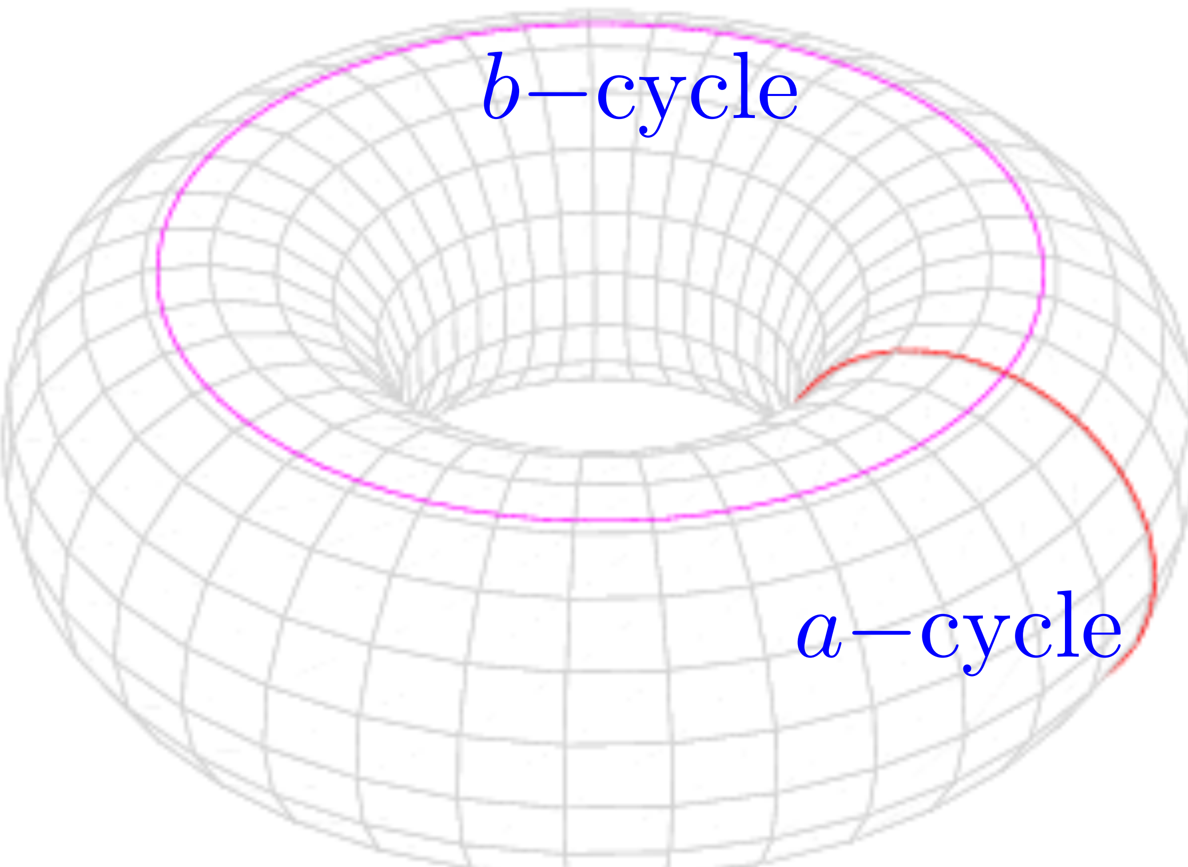
Calabi–Yau



$u^5 + v^5 + x^5 + y^5 + z^5 = 0 \subset \mathbb{P}^4$

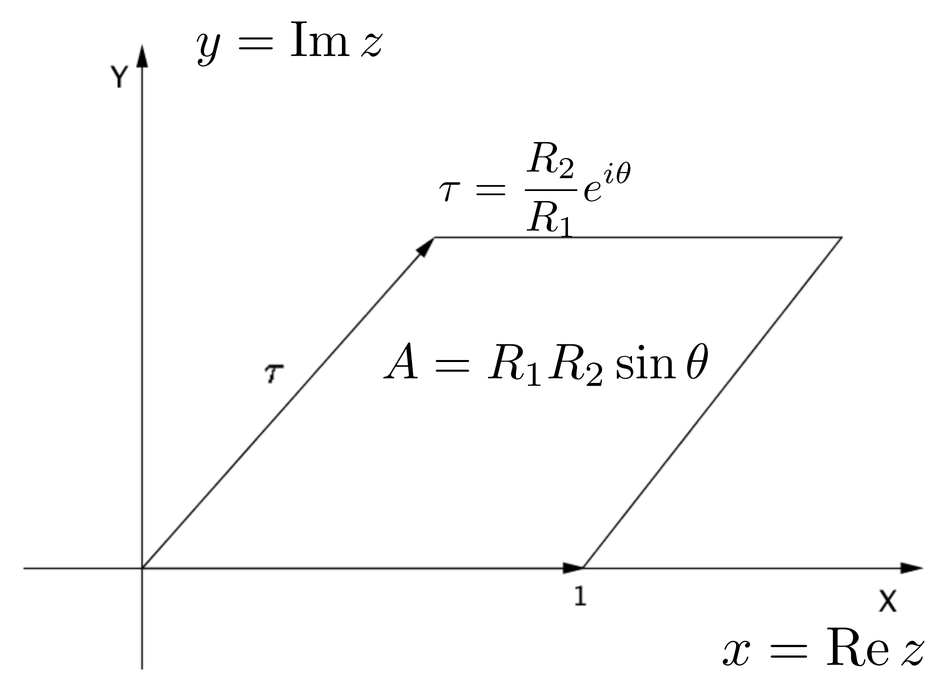
- Complex, Kähler manifold that admits Ricci-flat metric
- Numerical approximations of metric using neural networks

Torus



Flat, but has non-trivial homotopy

There are non-contractible cycles



Kähler parameter: area A **size**

complex structure parameter: τ **shape**

1-fold: $z_0^3 + z_1^3 + z_2^3 = 0 \subset \mathbb{P}^2$

$$\begin{array}{ccc} & 1 & \\ 1 & & 1 \\ & 1 & \end{array}$$

$\chi = 2 - 2g = 0$

Calabi–Yau n-folds

2-folds (K3):

$$
\begin{array}{ccccc}
 & & 1 & & \\
 & 0 & & 0 & \\
1 & & 20 & & 1 \\
 & 0 & & 0 & \\
 & & 1 & &
\end{array}
\qquad \chi = 24
$$

3-folds:

$$
\begin{array}{ccccccc}
 & & & 1 & & & \\
 & & 0 & & 0 & & \\
 & 0 & & h^{1,1} & & 0 & \\
1 & & h^{2,1} & & h^{2,1} & & 1 \\
 & 0 & & h^{1,1} & & 0 & \\
 & & 0 & & 0 & & \\
 & & & 1 & & &
\end{array}
\qquad \chi = 2(h^{1,1} - h^{2,1})
$$

2-cycles, 4-cycles

3-cycles

- Learn topological invariants of 7890 Calabi–Yau threefolds realized as complete intersections of homogeneous polynomials in projective space

Bull, He, VJ, Mishra (2018, 2019)

Toric CYs

- Starting from a reflexive polytope, one can realize a Calabi–Yau geometry as a smooth hypersurface in a toric variety

Batyrev (1993)
Batyrev, Borisov (1994)

- 4319 3d polytopes give K3; 473,800,776 4d polytopes yield toric Calabi–Yau threefolds with 30,108 unique pairs of Hodge numbers; more than 185,269,499,015 5d maximal polytopes yield toric Calabi–Yau fourfolds

Kreuzer, Skarke (1997, 1998, 2003)
Schöller, Skarke (2019)

- New formula for Euler characteristic in 4d (*i.e.*, for threefolds)

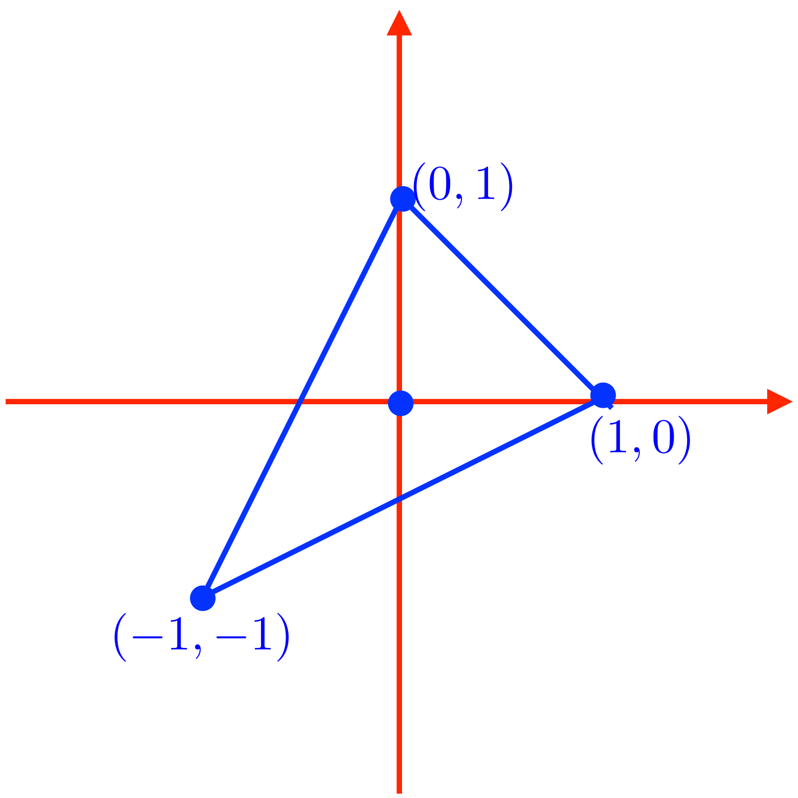
$$\chi(\mathcal{M}) = 2\big(\ell(2\Delta) - \ell(2\Delta^\circ)\big) + 18\big(\ell(\Delta^\circ) - \ell(\Delta)\big)$$

Berglund, Campbell, VJ (2021)

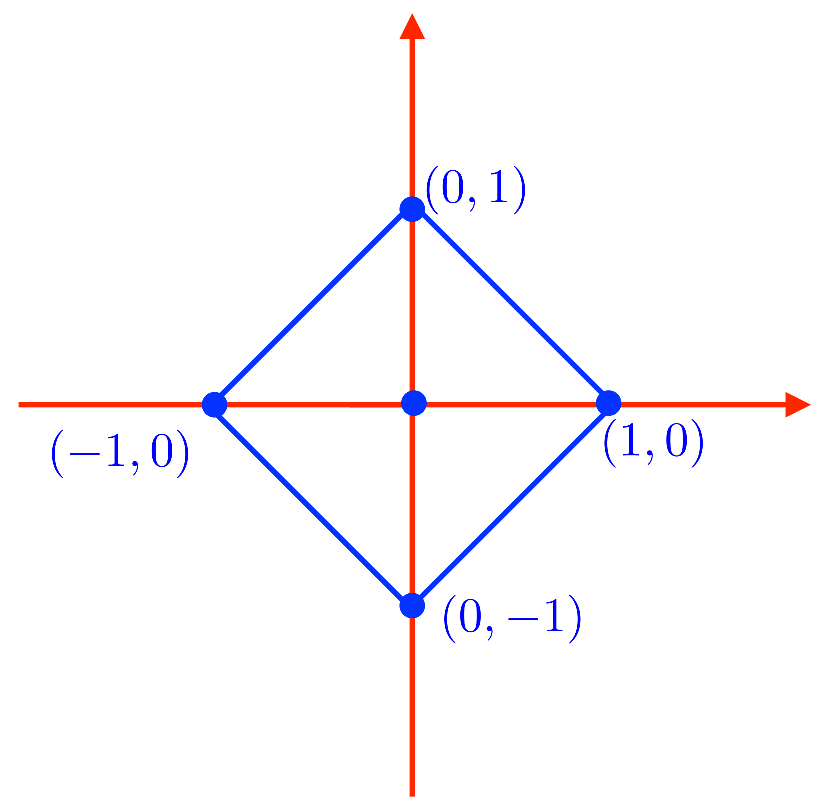
- New reflexive polytopes in 5d with new Hodge numbers, designer properties

Berglund, He, Heyes, Hirst, VJ, Lukas (2023)

2d Reflexive Polytopes



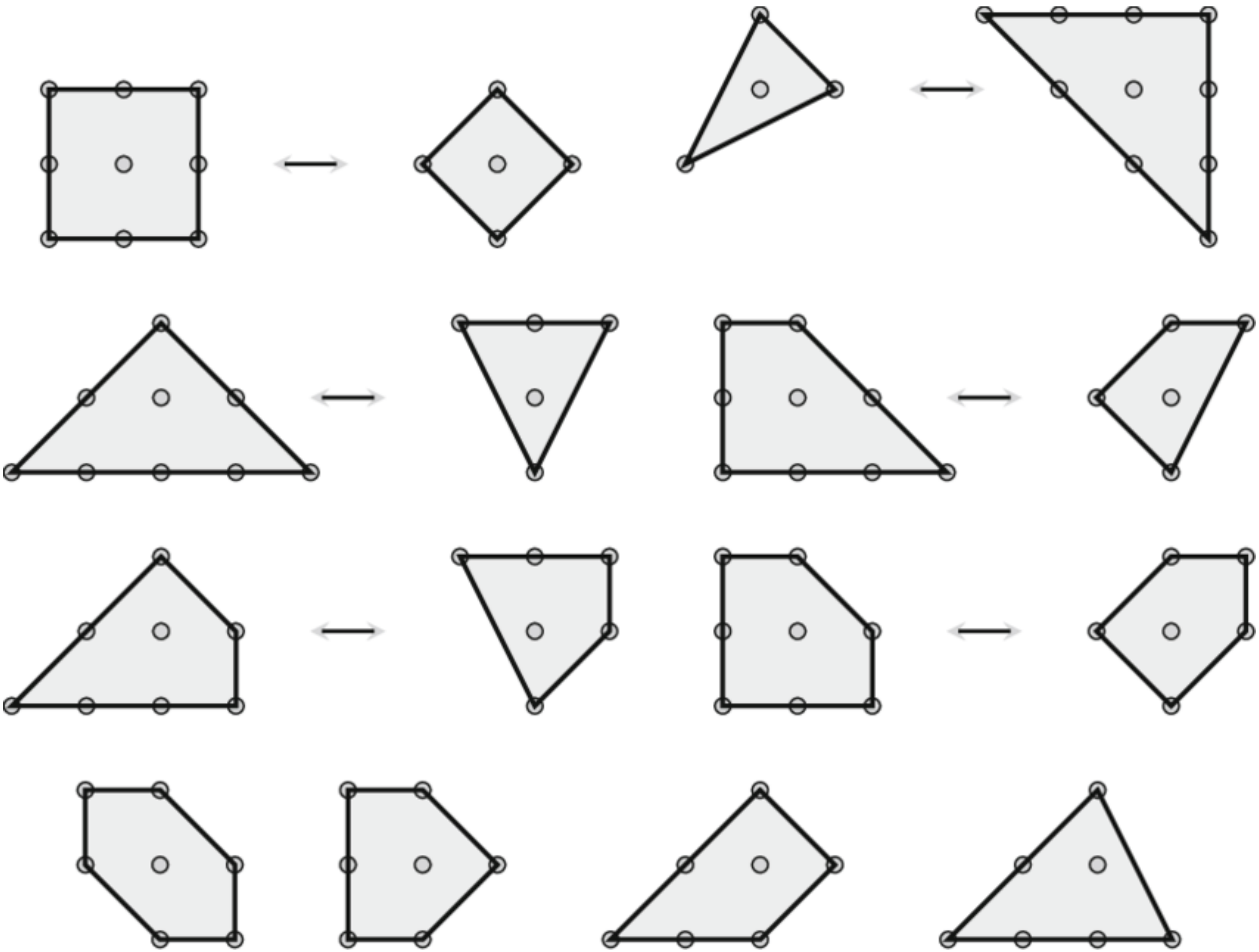
$$x+y=1$$
$$x-2y=1$$
$$-2x+y=1$$



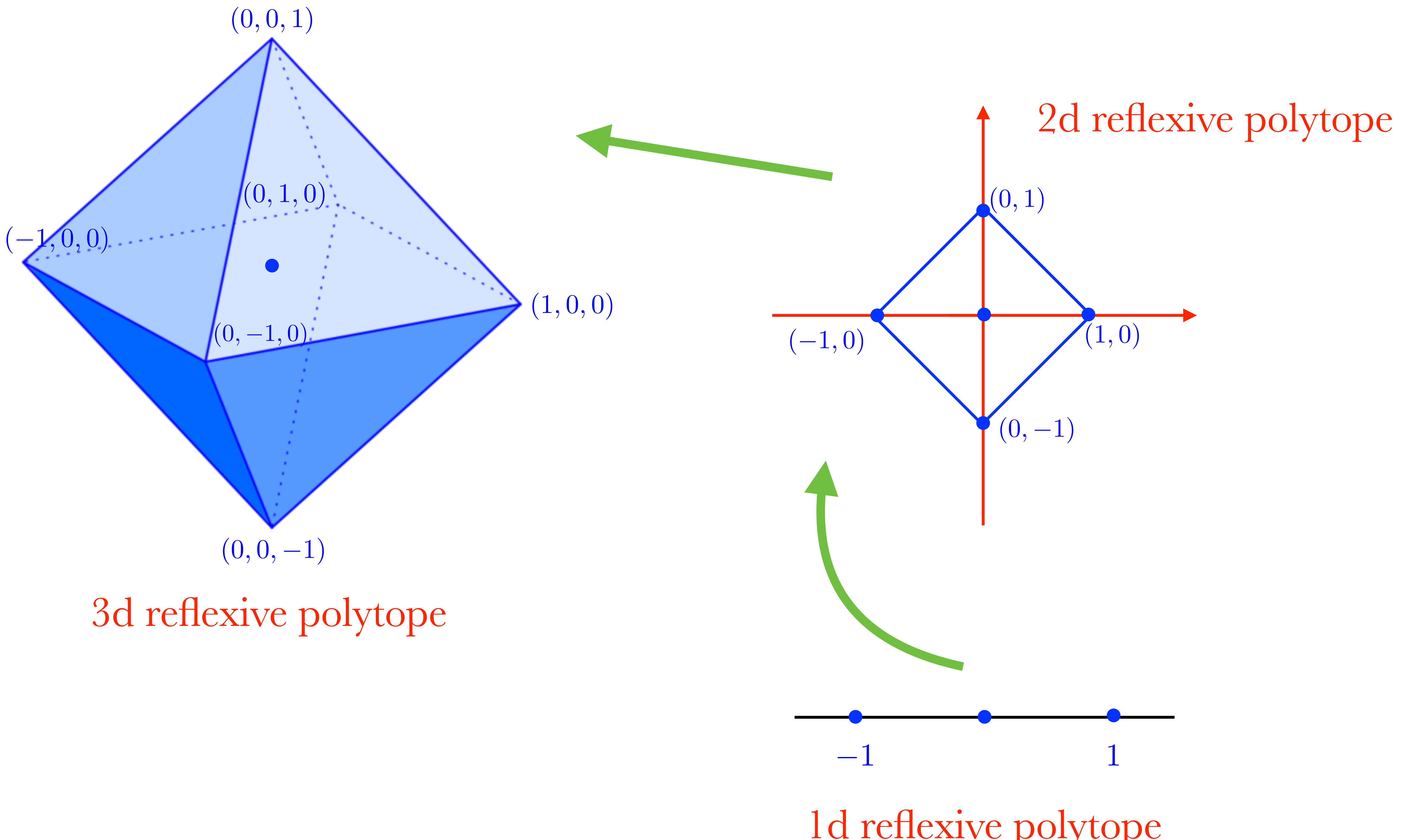
$$x+y=1$$
$$x-y=1$$
$$-x+y=1$$
$$-x-y=1$$

$$d_i(\Delta)=1$$

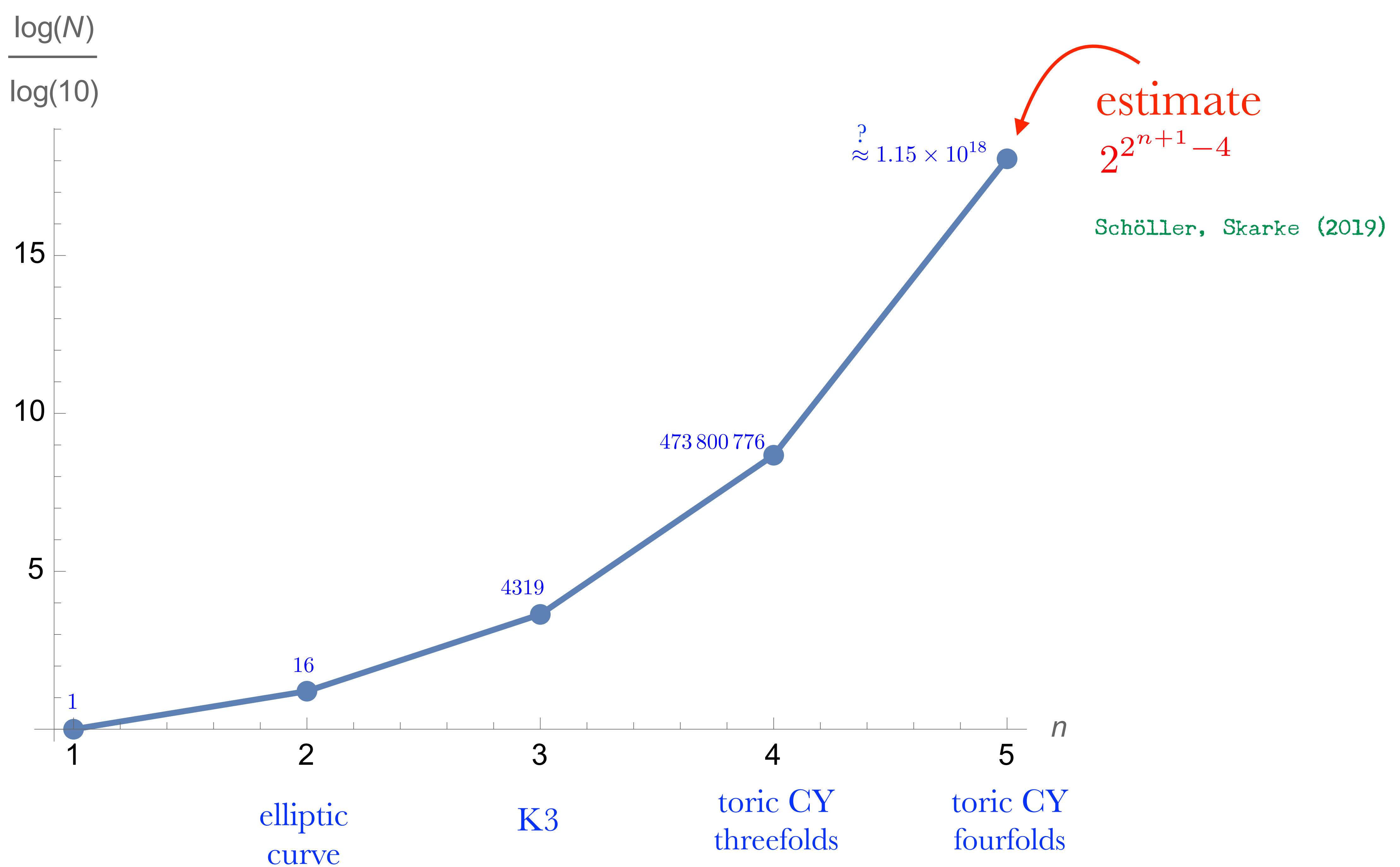
2d Reflexive Polytopes



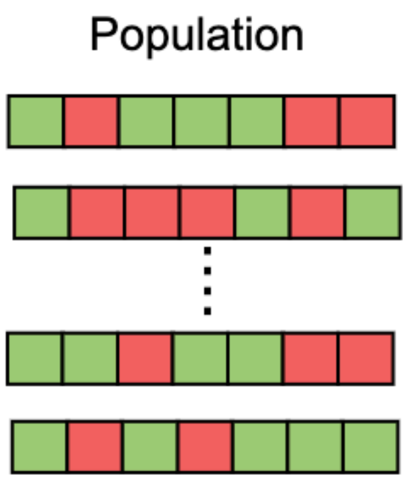
3d Reflexive Polytopes



n-d Reflexive Polytopes

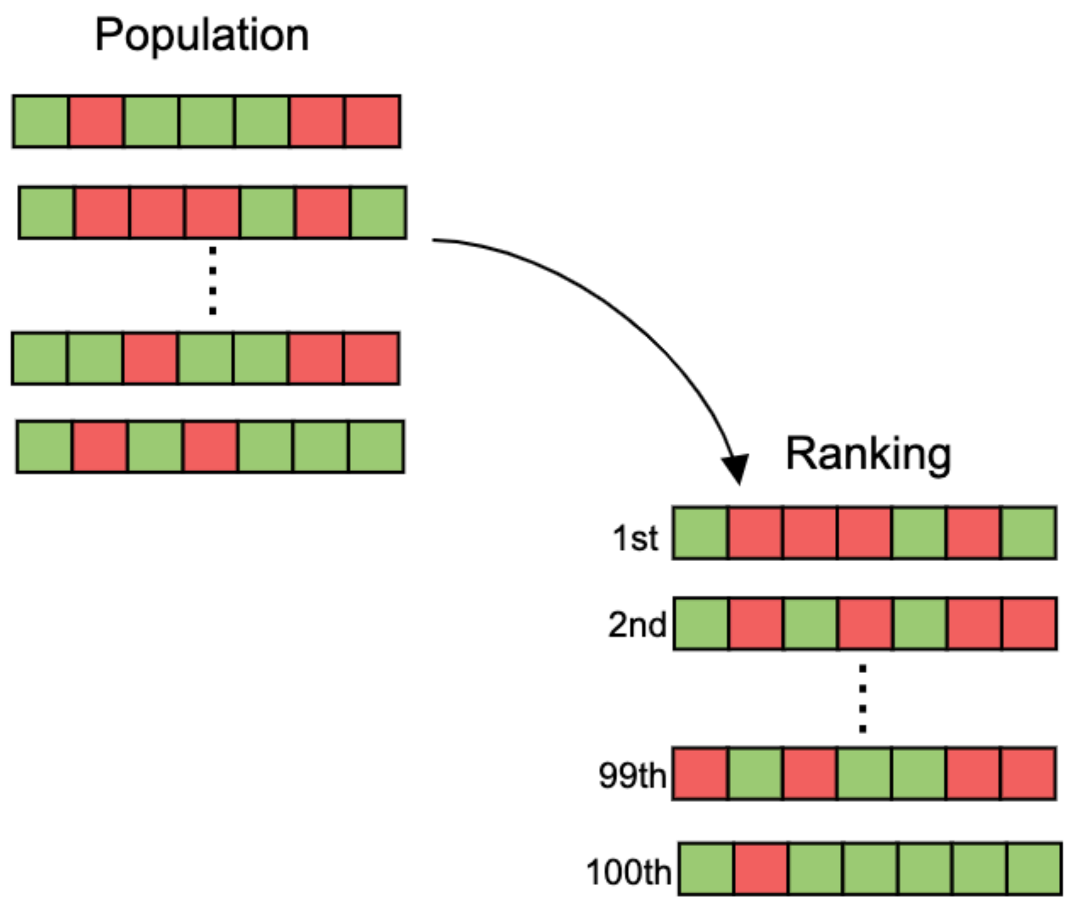


Genetic Algorithm



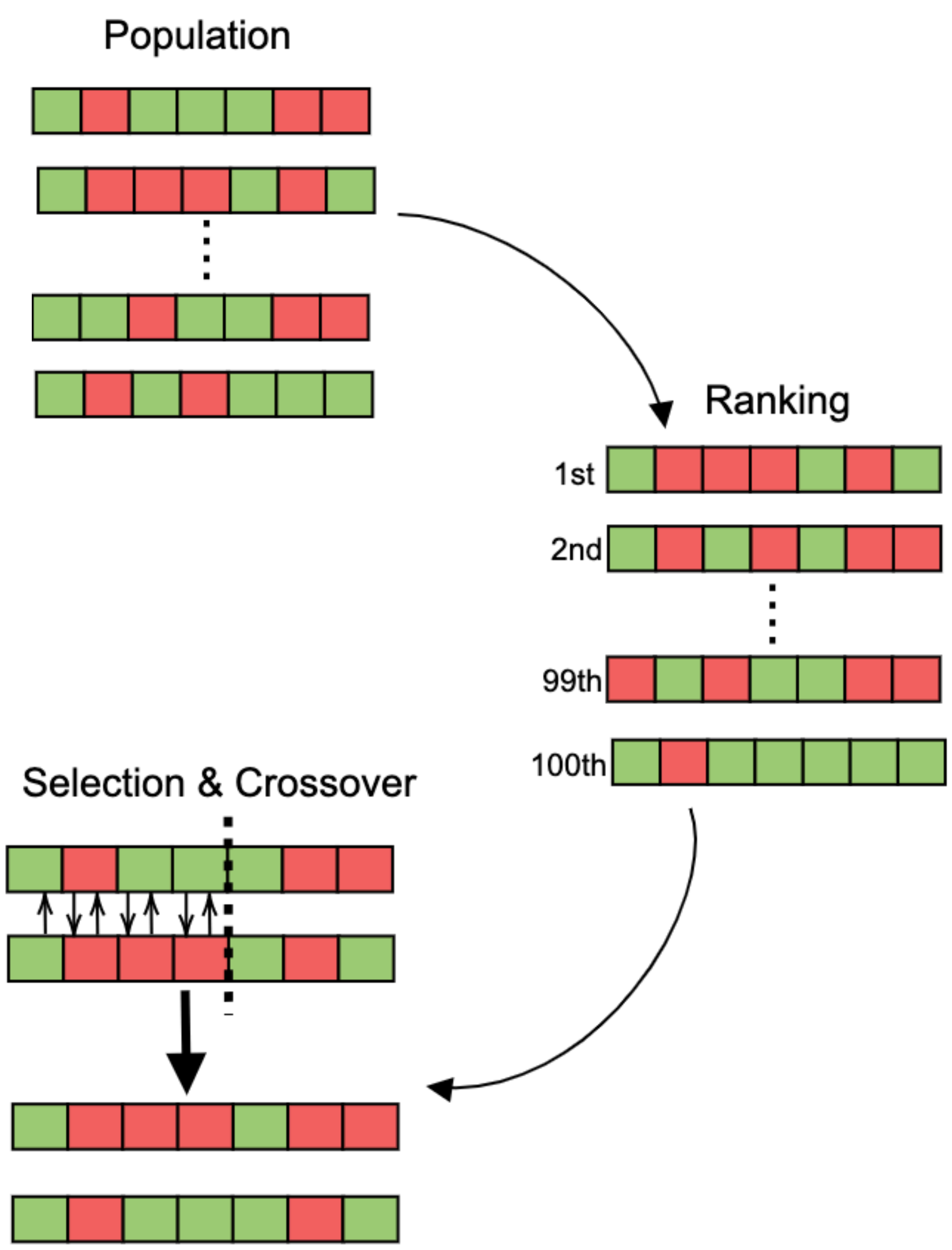
Darwin (1859)
Turing (1950)
Barricelli (1954)
Holland (1992)

Genetic Algorithm



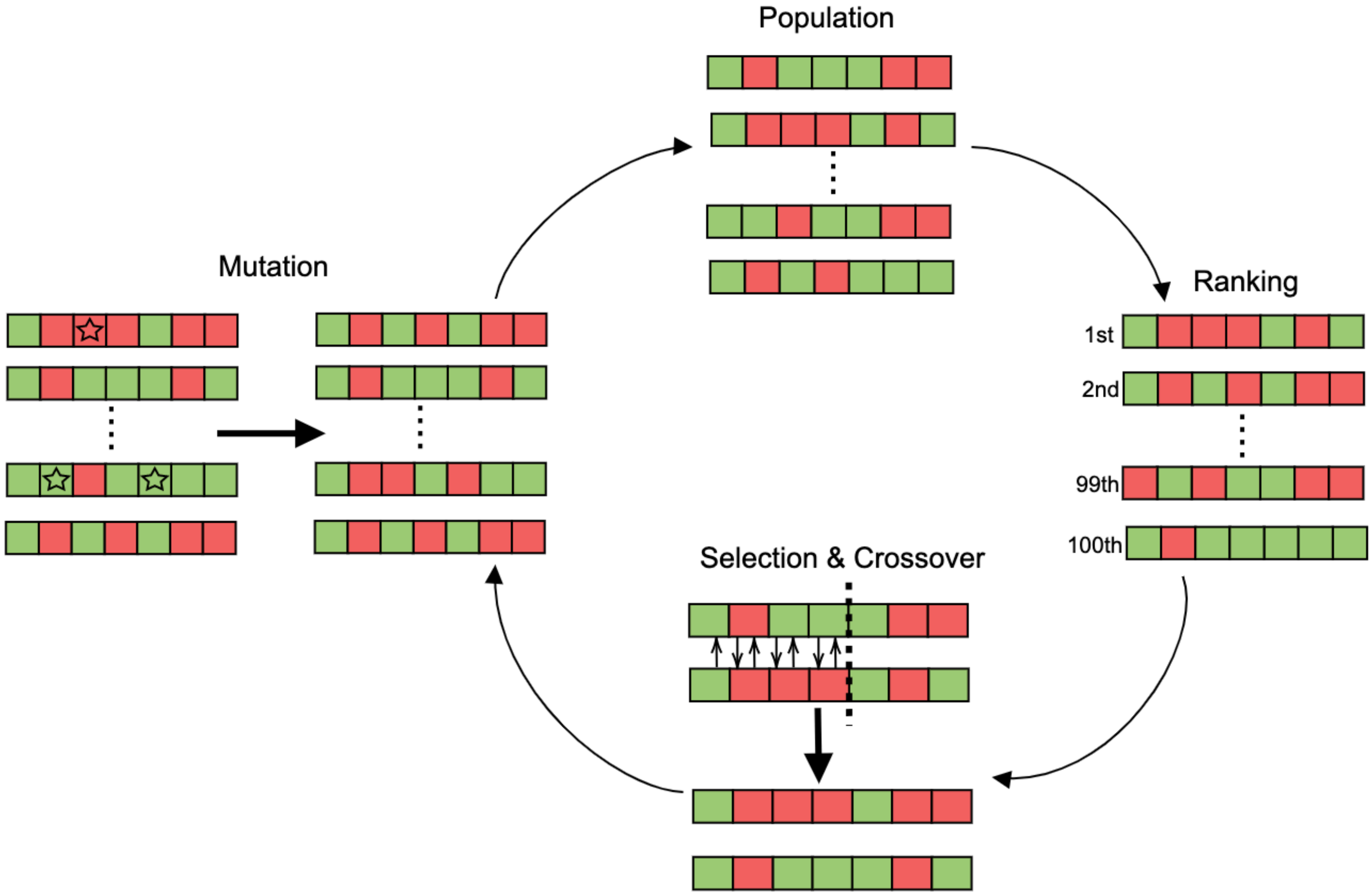
Darwin (1859)
Turing (1950)
Barricelli (1954)
Holland (1992)

Genetic Algorithm



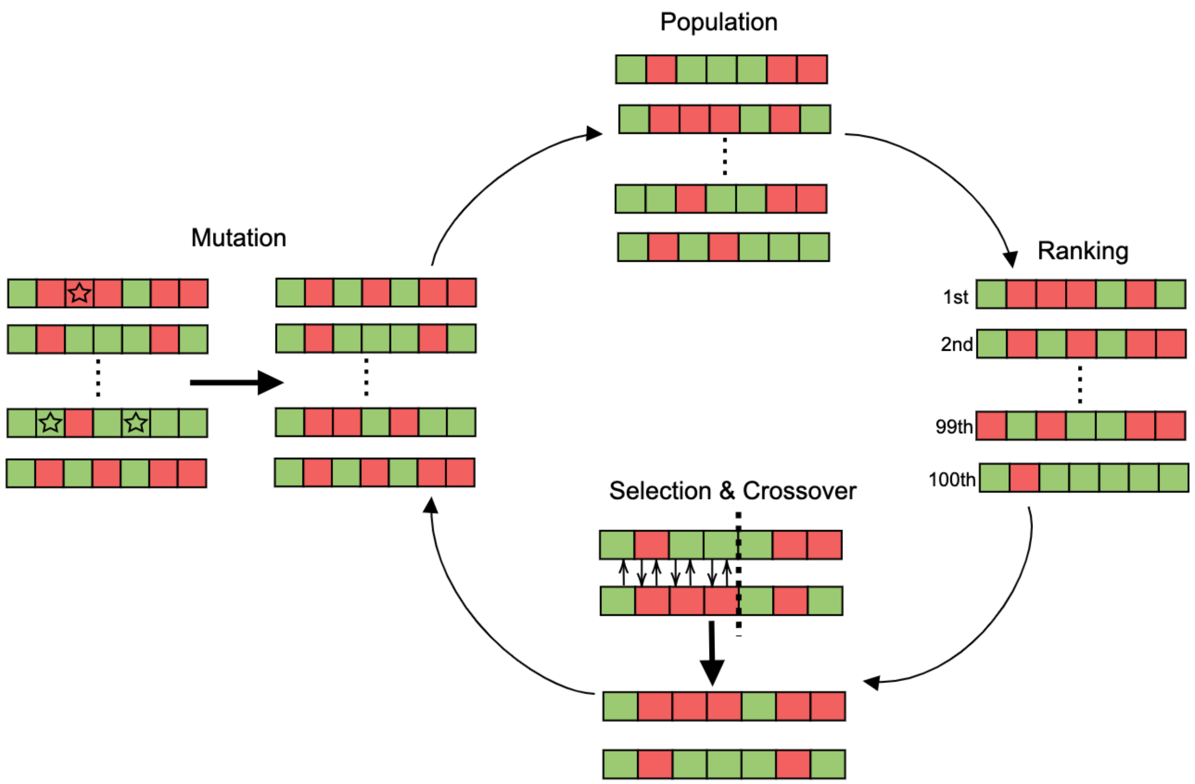
Darwin (1859)
Turing (1950)
Barricelli (1954)
Holland (1992)

Genetic Algorithm



Darwin (1859)
Turing (1950)
Barricelli (1954)
Holland (1992)

Genetic Algorithm



1. Create a random population

2. Evolve the population

3. Extract all generated reflexive polytopes ← these optimize cost function

4. Remove redundancy in this list

5. Repeat until all reflexive polytopes with desired properties realized

Berglund, He, Heyes, Hirst, VJ, Lukas (2023)

Fitness Function

- How to rank polytopes?

- Let $\Delta = \{m \in M_{\mathbb{R}} \mid \langle u_i, m \rangle \leq a_i \ , \ i = 1, \ldots, k\}$, where $u_i \in N_{\mathbb{R}} \ , \ a_i \in \mathbb{R}$

- Reflexive iff IP and $a_i = 1 \ \forall \ i$

- Fitness function: $f(\Delta) = w_1(\mathrm{IP}(\Delta) - 1) - \frac{w_2}{k} \sum_{i=1}^{k} |d_i(\Delta) - 1|$

  $w_1 \ , \ w_2 \in \mathbb{R}_{\geq 0}$ are weights

  $\mathrm{IP}(\Delta) = 1$ if Δ enjoys IP property and 0 otherwise

  fitter polytopes have larger $f(\Delta)$

Berglund, He, Heyes, Hirst, VJ, Lukas (2023)

5d Reflexive Polytopes

- $h^{1,1} = 1:$ $$V = \begin{pmatrix} 0 & 4 & -1 & 2 & -1 & -2 \\ 3 & -1 & -4 & 9 & 0 & -3 \\ 2 & 2 & -2 & 2 & -2 & 2 \\ 1 & -3 & 0 & 3 & 0 & -1 \\ 3 & -5 & 1 & -1 & 1 & -1 \end{pmatrix}$$ $h^{2,1} = 0$, $h^{3,1} = 111$, $h^{2,2} = 492$, $\chi = 720$

- **Conjecture:** There are 15 reflexive polytopes with $h^{1,1} = 1$

- $h^{1,1} = 2:$ $$V = \begin{pmatrix} 0 & -2 & -4 & -1 & 10 & 8 & -2 \\ 4 & 0 & 2 & -1 & -6 & -4 & 0 \\ -1 & 1 & -3 & 0 & 1 & 3 & 1 \\ 2 & -2 & 0 & 0 & -2 & 0 & 1 \\ 2 & 0 & -2 & 1 & -6 & -2 & 3 \end{pmatrix}$$ $h^{2,1} = 0$, $h^{3,1} = 111$, $h^{2,2} = 496$, $\chi = 726$

- *N.B.*: Calabi–Yau fourfolds inherit Hodge numbers from reflexive polytope. These are new geometries!

Berglund, He, Heyes, Hirst, VJ, Lukas (2023)

5d Reflexive Polytopes

- 11d supergravity compactified on a fourfold to break $\mathcal{N}=2$ to $\mathcal{N}=1$ in 3d needs χ divisible by 24, 224, 504

Berg, Haack, Samtleben (2003)

- $$\chi = 2016 : \quad V = \begin{pmatrix} -1 & -1 & 0 & -1 & 0 & 1 & -1 & 0 \\ 1 & 1 & 2 & 0 & -1 & 2 & 0 & 0 \\ -1 & 3 & 2 & 3 & -1 & -1 & 2 & 0 \\ 0 & 4 & 3 & 4 & -2 & 0 & 3 & 1 \\ -3 & 4 & 0 & 4 & 0 & 0 & 2 & -2 \end{pmatrix} \qquad \begin{aligned} h^{1,1} &= 331 \\ h^{2,1} &= 9 \\ h^{3,1} &= 6 \\ h^{2,2} &= 1374 \end{aligned}$$

- We can do targeted searches for desired phenomenology

Berglund, He, Heyes, Hirst, VJ, Lukas (2023)

- Use reinforcement learning to construct Calabi–Yau manifolds explicitly

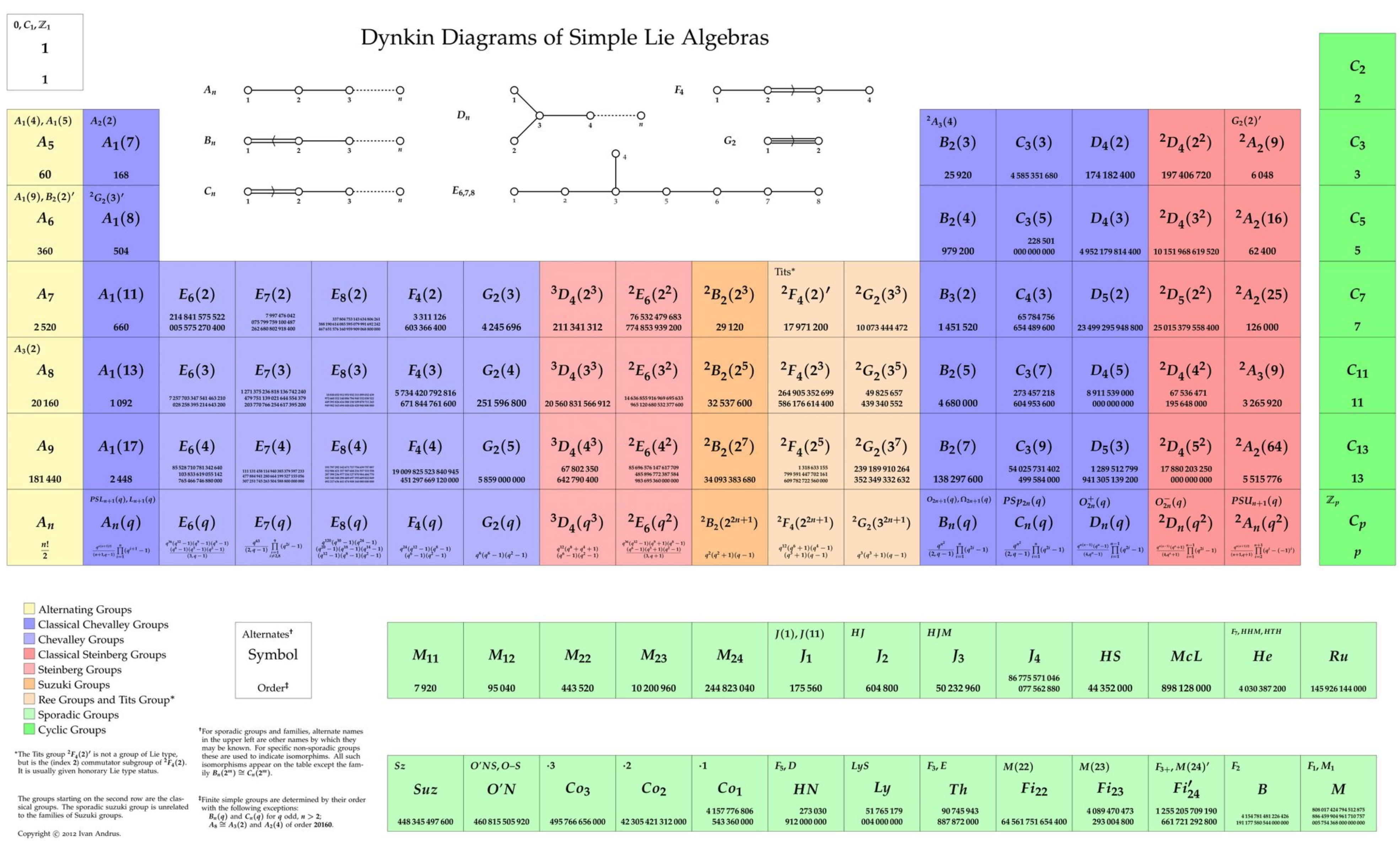
Berglund, Butbaia, He, Heyes, Hirst, VJ (2024)

PROSPECTUS

Challenge

- Utility of computers for mathematics

The Periodic Table Of Finite Simple Groups

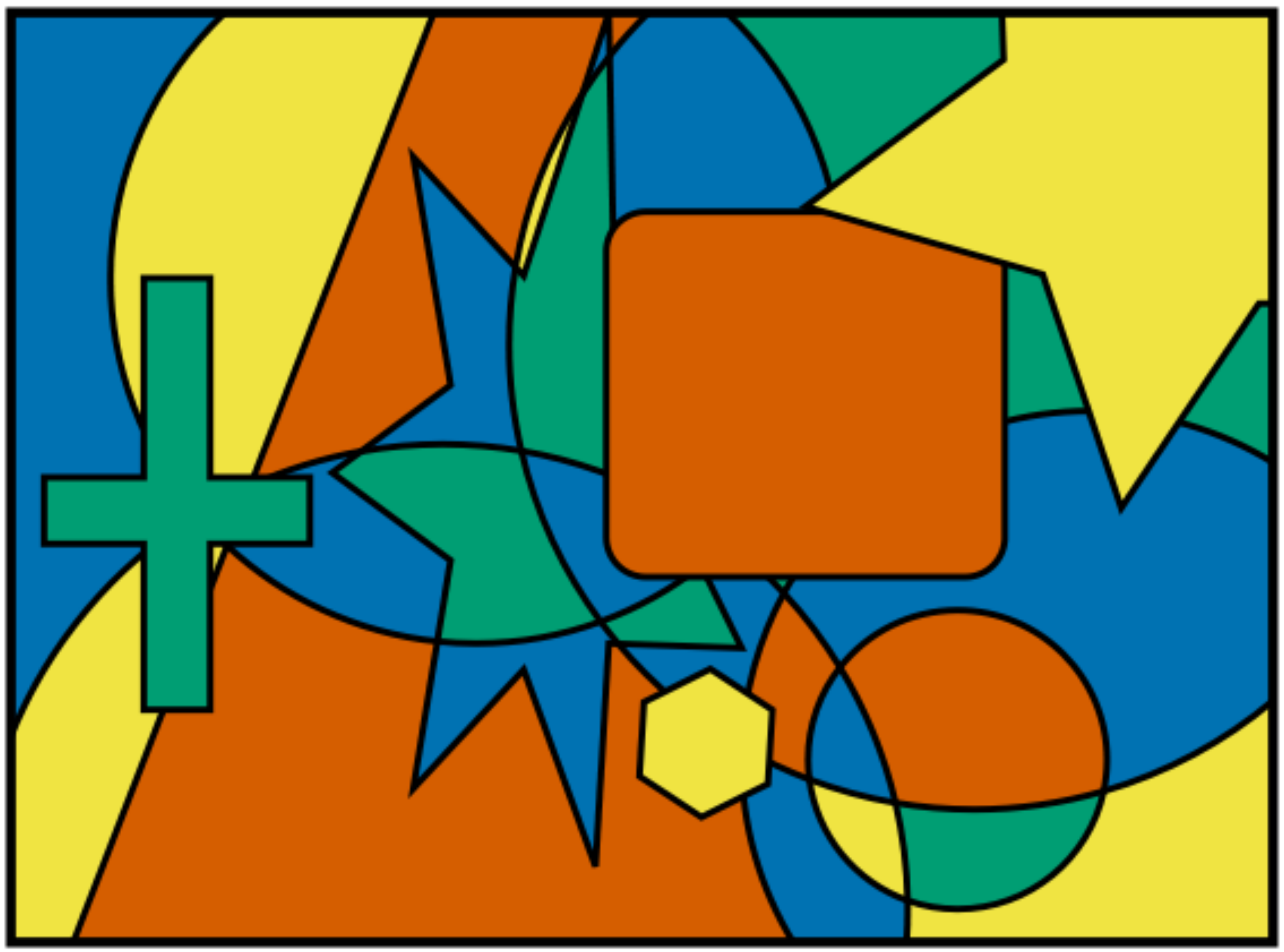


Gorenstein (1982, 1983)
Ashbacher, Lyons, Smith (2011)

Challenge

- Utility of computers for mathematics

Four color theorem



Appel, Haken (1976)

Challenge

Applications [edit]

Theorems proved with the help of computer programs [edit]

Inclusion in this list does not imply that a formal computer-checked proof exists, but rather, that a computer program has been involved in some way. See the main articles for details.

- Four color theorem, 1976
- Mitchell Feigenbaum's universality conjecture in non-linear dynamics. Proven by O. E. Lanford using rigorous computer arithmetic, 1982
- Connect Four, 1988 – a solved game
- Non-existence of a finite projective plane of order 10, 1989
- Double bubble conjecture, 1995[2]
- Robbins conjecture, 1996
- Kepler conjecture, 1998 – the problem of optimal sphere packing in a box
- Lorenz attractor, 2002 – 14th of Smale's problems proved by Warwick Tucker using interval arithmetic
- 17-point case of the Happy Ending problem, 2006
- NP-hardness of minimum-weight triangulation, 2008
- Optimal solutions for Rubik's Cube can be obtained in at most 20 face moves, 2010
- Minimum number of clues for a solvable Sudoku puzzle is 17, 2012
- In 2014 a special case of the Erdős discrepancy problem was solved using a SAT-solver. The full conjecture was later solved by Terence Tao without computer assistance.[3]
- Boolean Pythagorean triples problem solved using 200 terabytes of data in May 2016.[4]
- Applications to the Kolmogorov-Arnold-Moser theory[5][6]
- Kazhdan's property (T) for the automorphism group of a free group of rank at least five
- Schur number five, the proof that S(5) = 161 was announced in 2017 by Marijn Heule and took up 2 petabytes of space[7][8]
- Keller's conjecture in dimension 7 the only remaining case in 2020 with a 200 gigabyte proof[9][10][11]

Theorems for sale [edit]

In 2010, academics at The University of Edinburgh offered people the chance to "buy their own theorem" created through a computer-assisted proof. This new theorem would be named after the purchaser.[12][13] This service now appears to no longer be available.

Challenge

- How does a black box learn semantics without knowing syntax?

  — Knowing that there are approximate functions can we find analytic expressions by opening the black box?

- Computers are creative — *cf.*, new fuseki, jōseki in go `AlphaZero (2017)`

- Can artificial intelligence do interesting research?

  — Proofs in real analysis `Ganesalingam, Gowers (2013)`

  — Proof assistants `Voevodsky (2014)`

  — Lean, Xena Project, Liquid Tensor Experiment `Buzzard (2019)` `Clausen, Scholze (2020)` `Commelin (2021)` `Tao (2023)`

- Logical reasoning + natural language — how do we collaborate effectively?

  — does silicon replace carbon? when?

THANK YOU!

Department:
Science and Innovation
REPUBLIC OF SOUTH AFRICA

NRF
National
Research
Foundation