Learning CY metrics

FABIAN RUEHLE

Physics meets ML

April 21st, 2021

Based on:

[Anderson, Gray, Gerdes, Krippendorf, Raghuram, FR: 2012.04656]

[Larfors, Lukas, FR: 1805.08499]

[Ashmore, FR: 2103.07472]

Outline

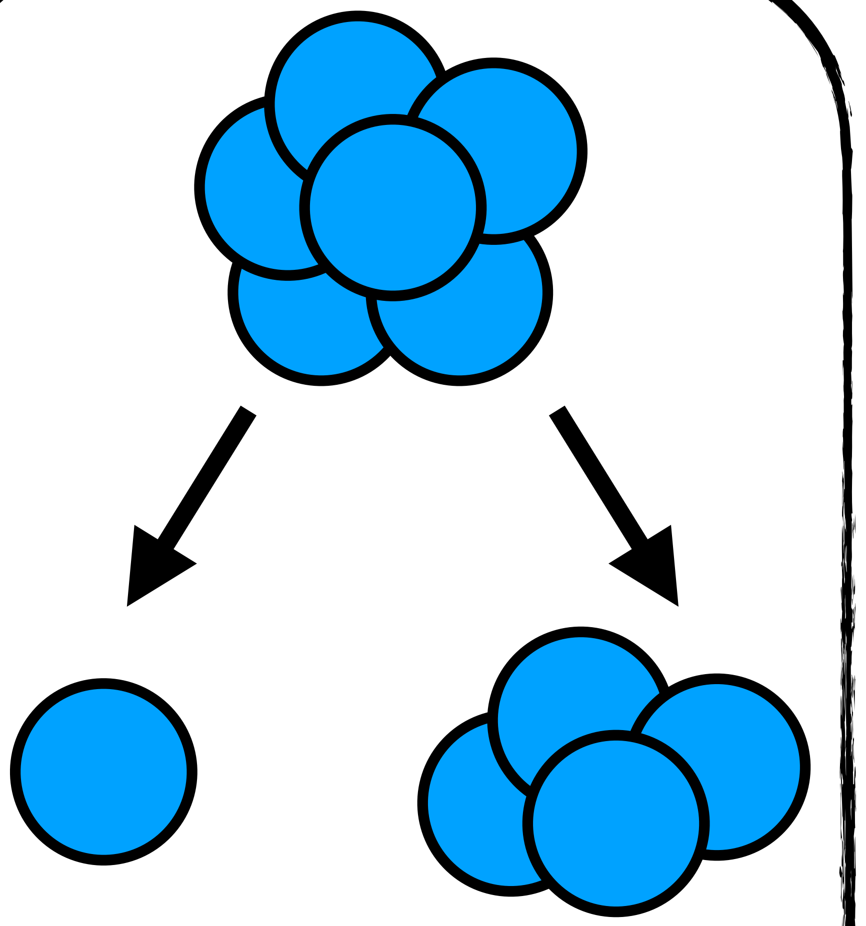
- Motivation for String Theory
- Metrics
- Calabi-Yau Manifolds
- ML for CY Metrics

Fundamental interactions

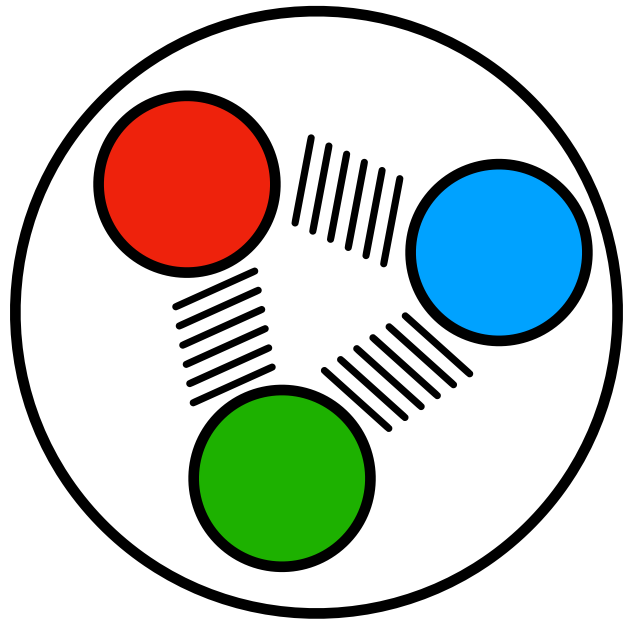
Many observations in our Universe can be explained with just four fundamental forces

Gravity

EM



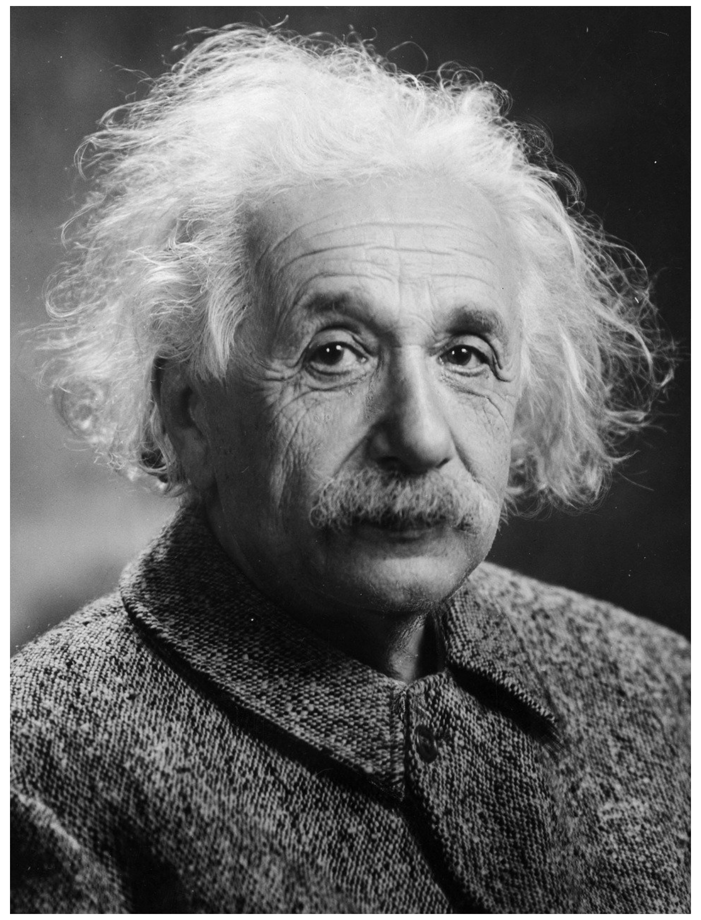
Weak



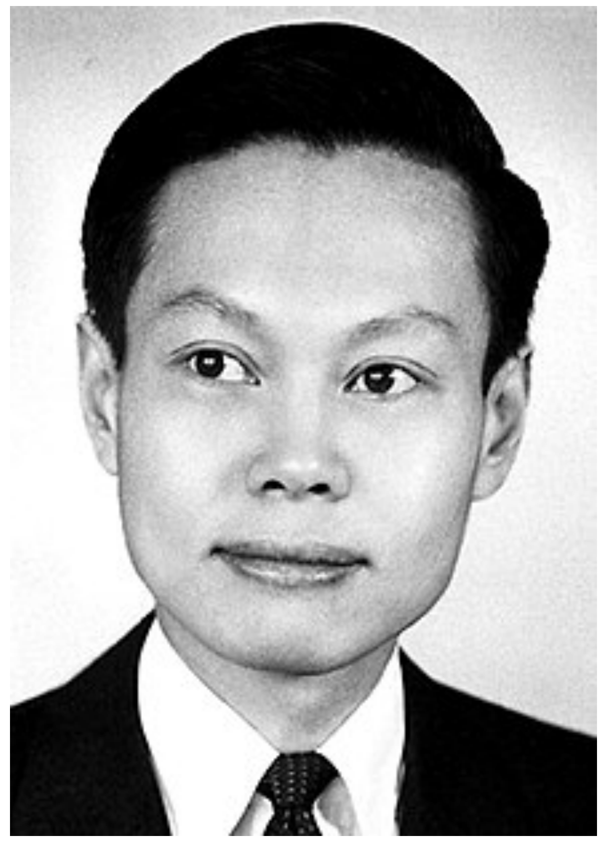
Strong

Fundamental interactions

Two theories to describe these four forces.



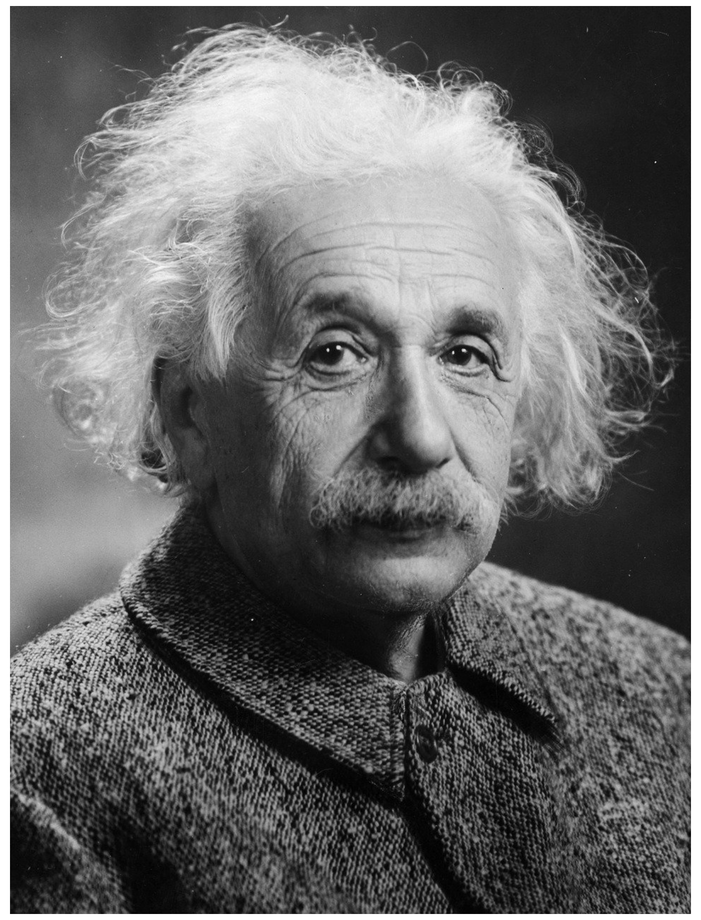
General Relativity



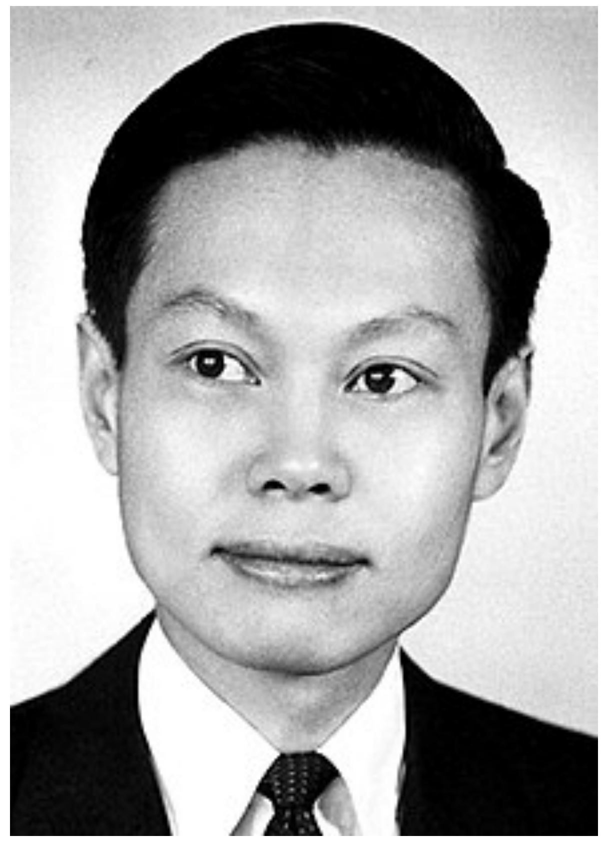
Quantum Field Theory (Yang-Mills Theory)

[Image Source: wikipedia]

Fundamental interactions



+



=

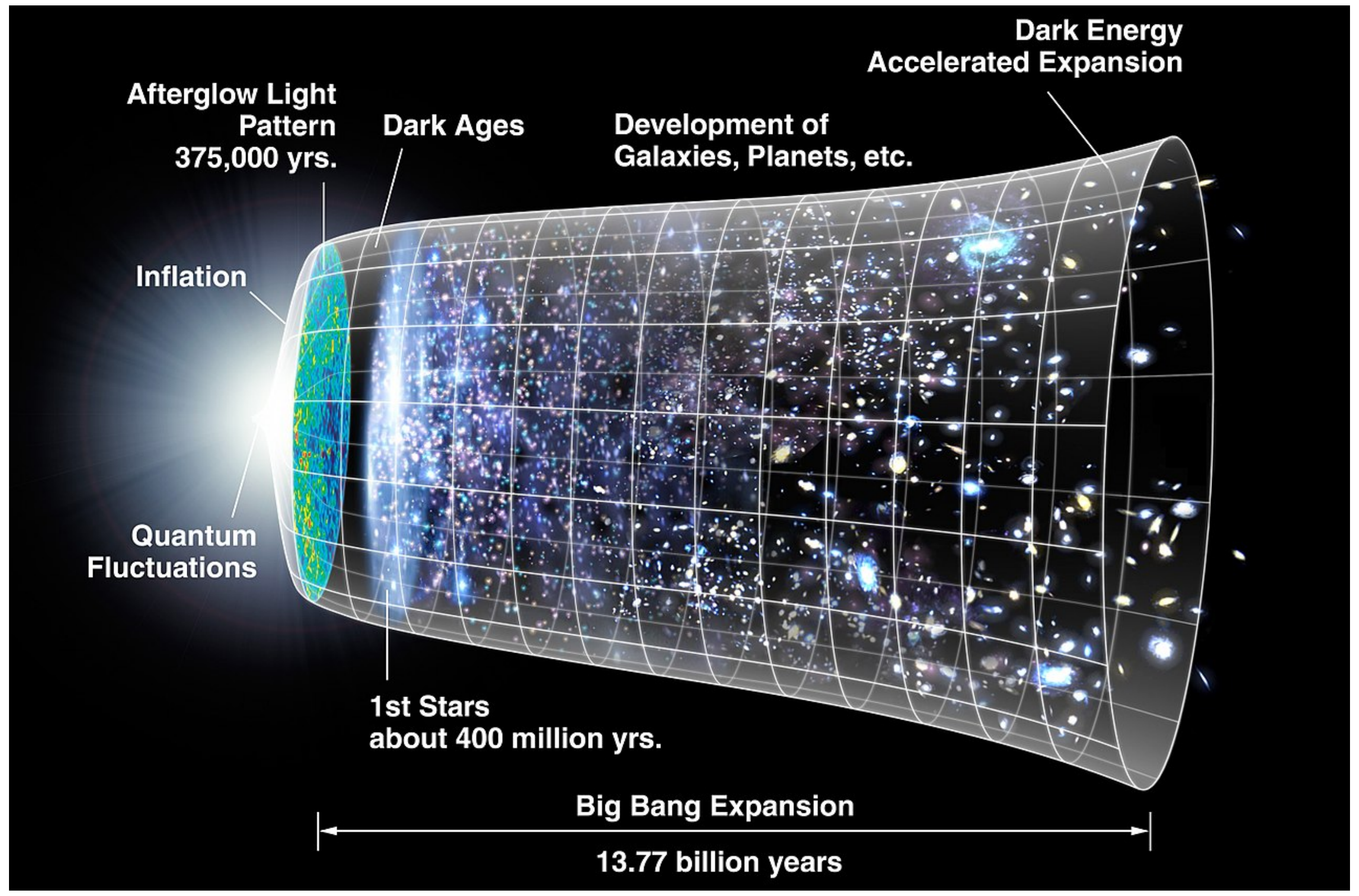
∞

[Image Source: wikipedia]

"… so don't combine them"

Combining the descriptions

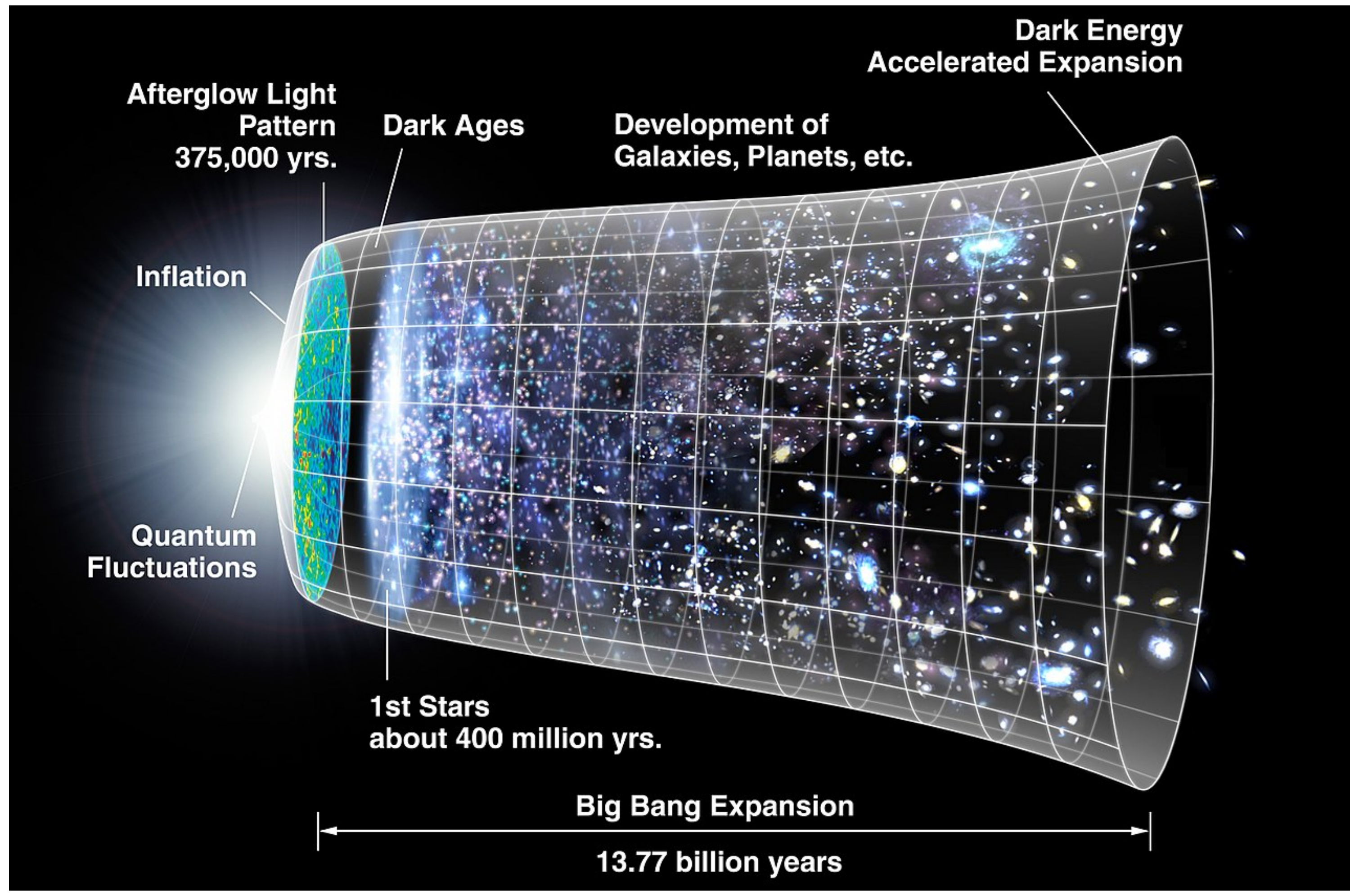
…however, we need a unified description to study physics at high energies



What happened at the Big Bang?
What is Dark Energy?

Combining the descriptions

…however, we need a unified description to study physics at high energies

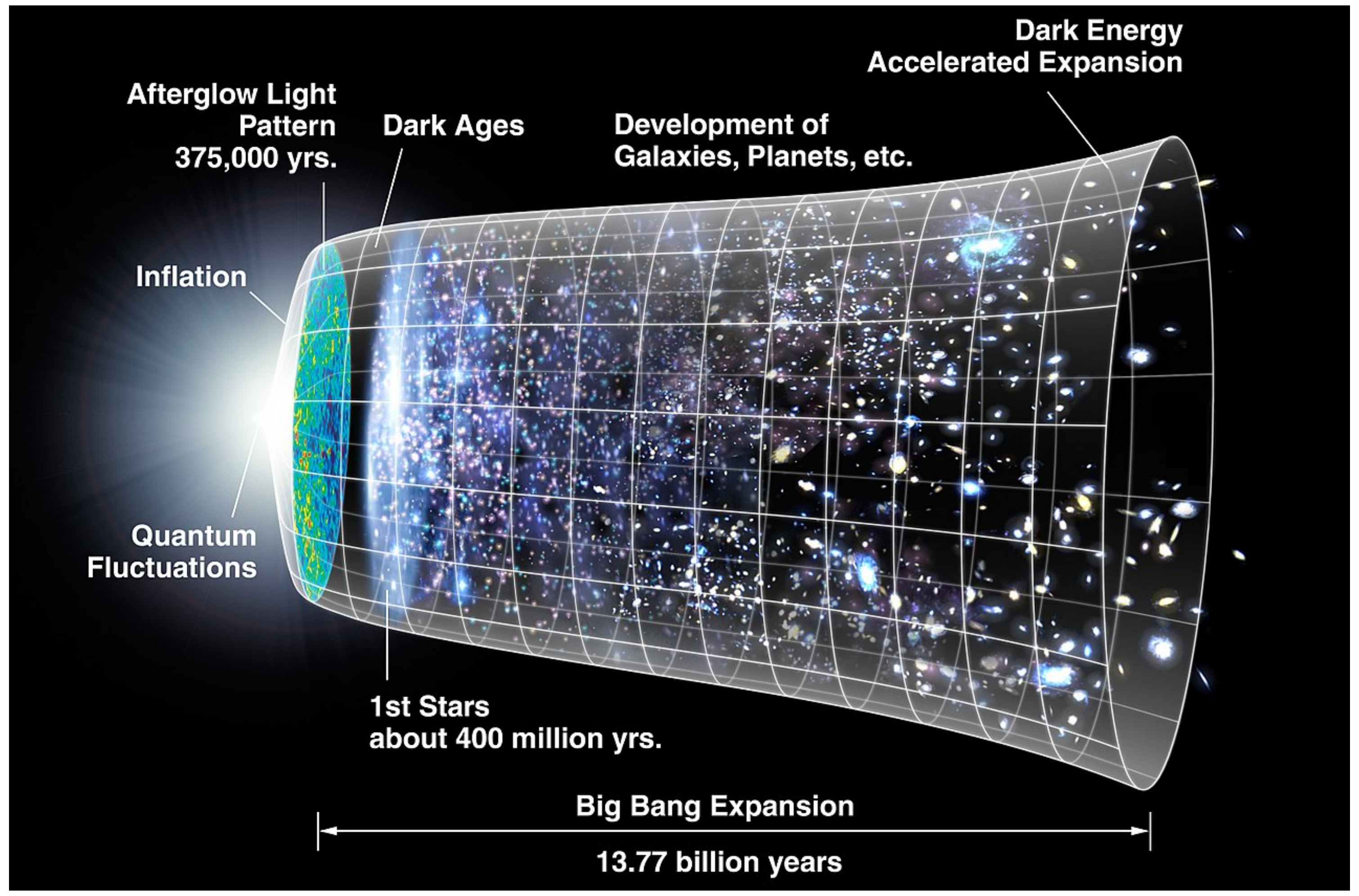


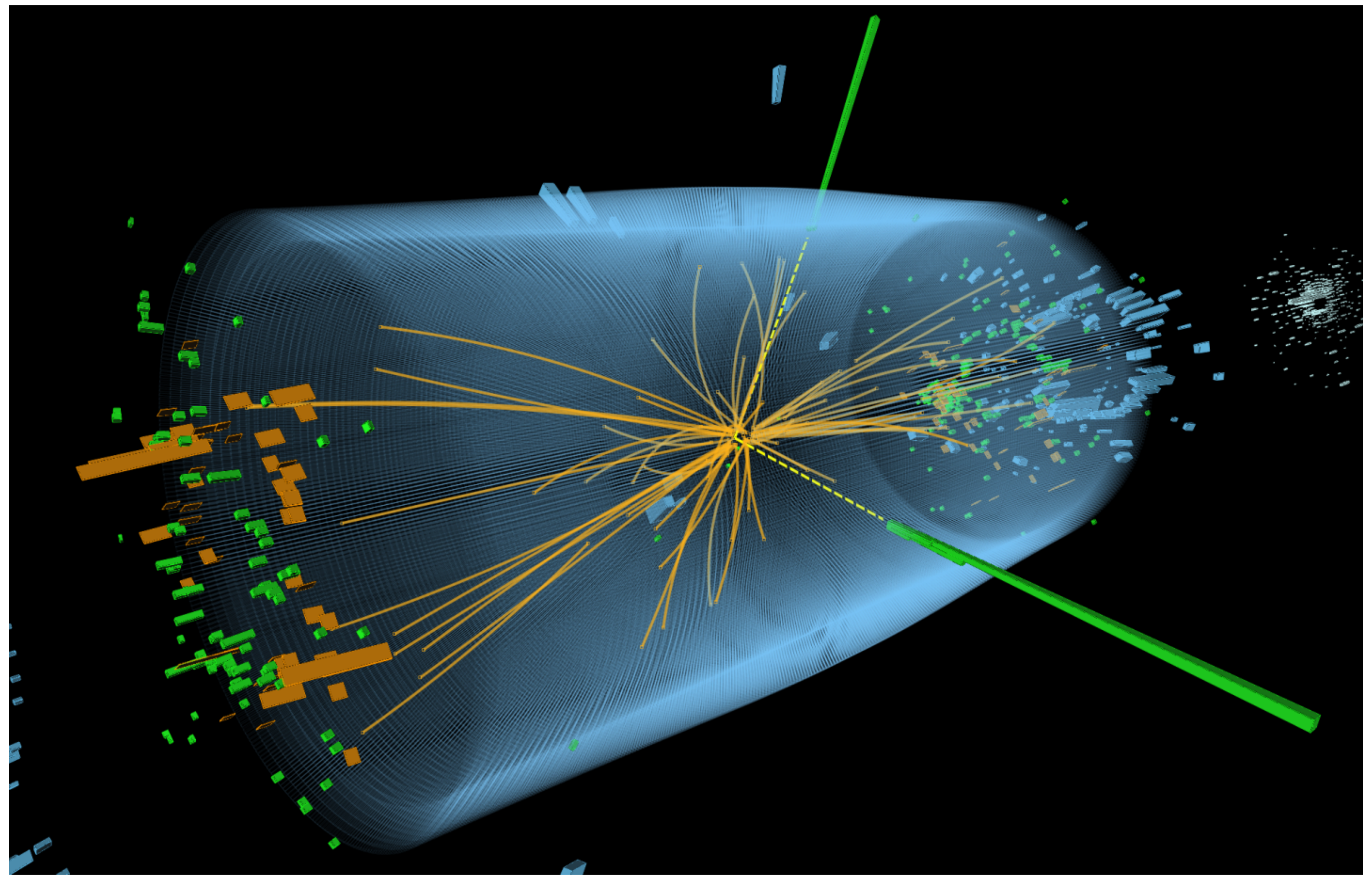
What happened at the Big Bang?
What is Dark Energy?

What happens at the event horizon of a black hole?

Combining the descriptions

…however, we need a unified description to study physics at high energies

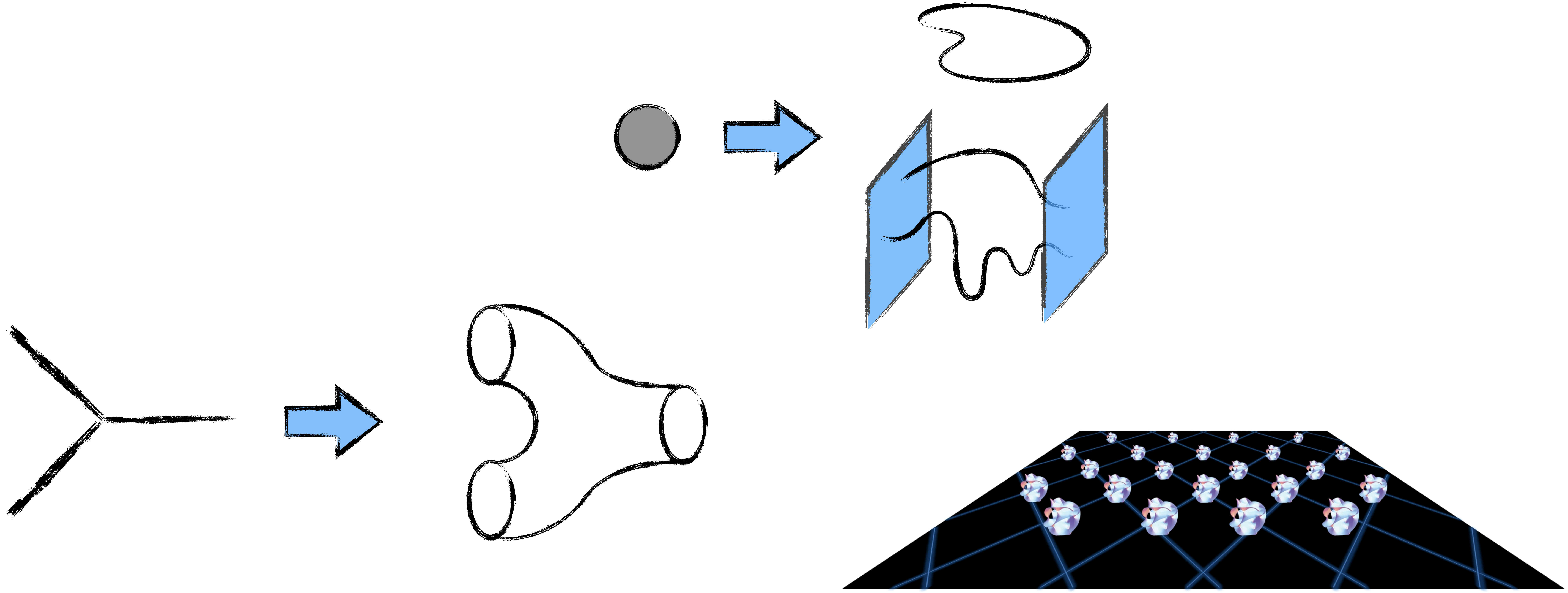




What happened at the Big Bang? What is Dark Energy?

What happens at the event horizon of a black hole?

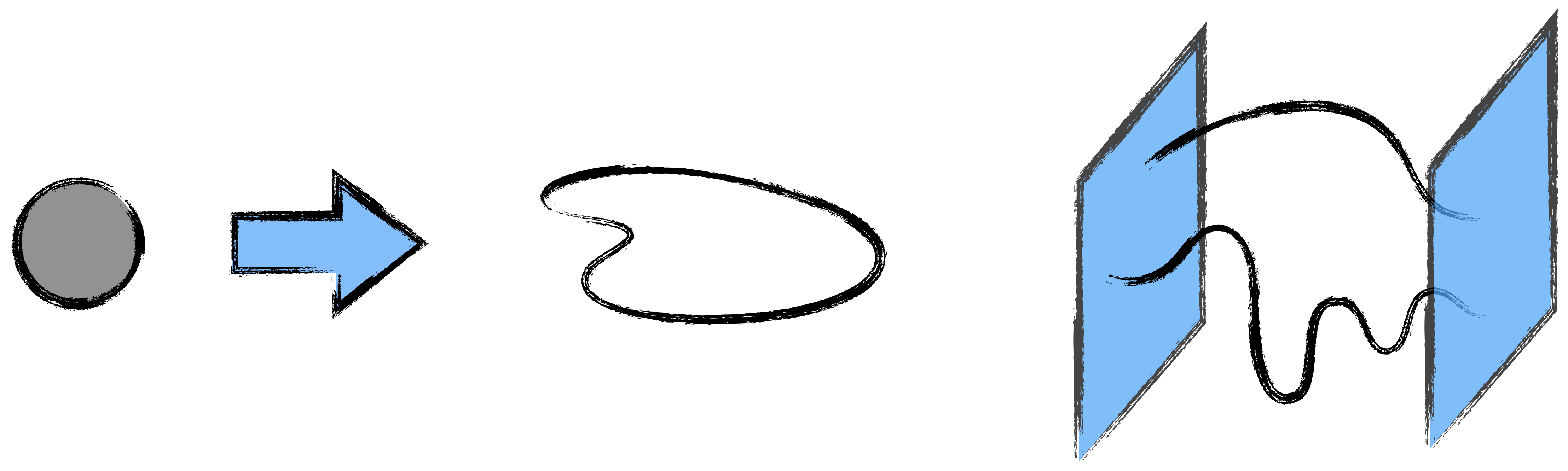
Do all particles and forces come from a Grand Unified Theory?



Introduction to String Theory

String Theory

- One promising candidate for a unified description of General Relativity and Quantum Field Theory: **String Theory**
- Basic assumption: Fundamental constituents of the particles that mediate the four forces and of all matter are not-point-like, but one-dimensional, extended strings



String Theory - Compactifications

This has far-reaching consequences!

- Requires ten space-time dimensions

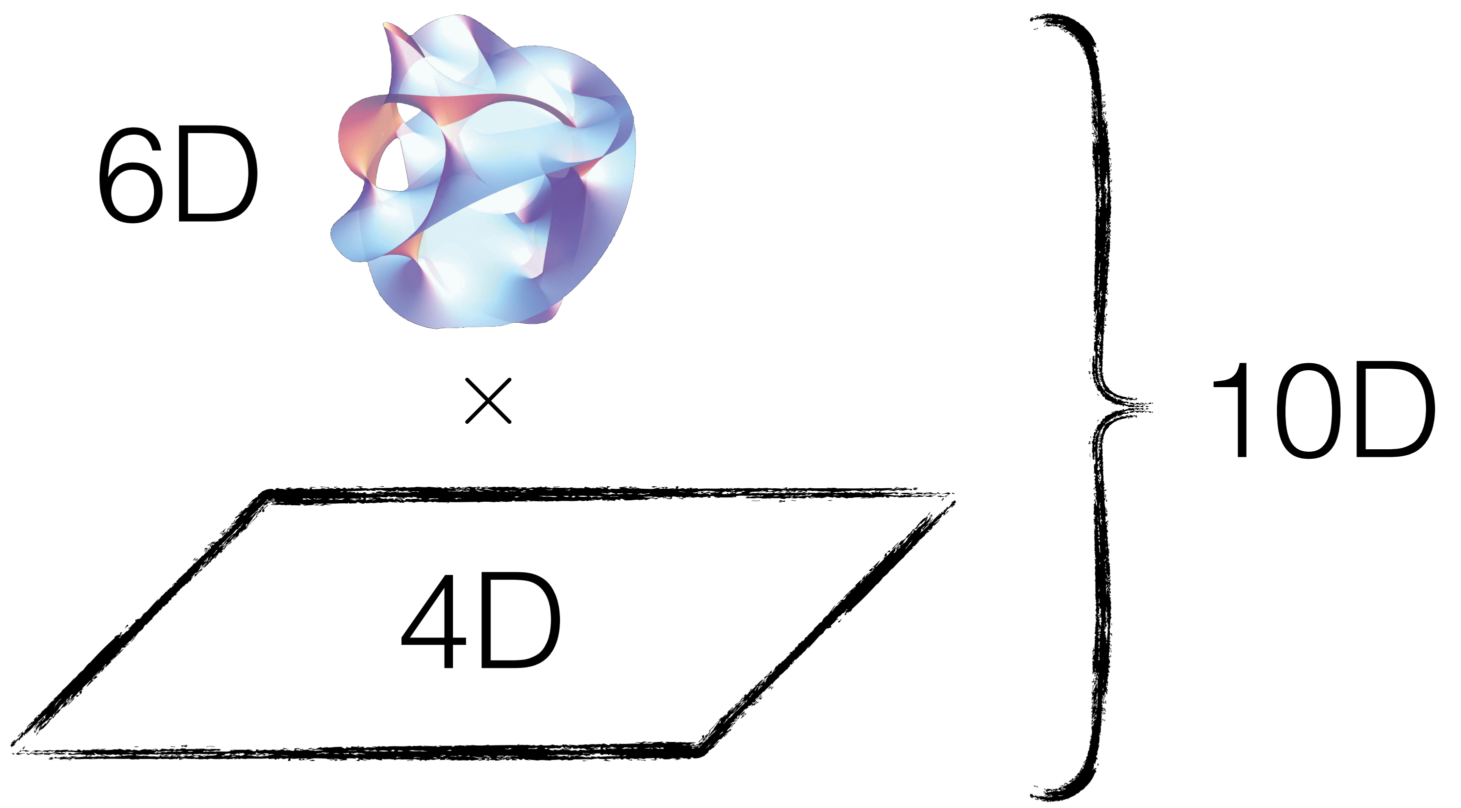
"… last time I checked the world was 3D"

String Theory - Compactifications

This has far-reaching consequences!

- Requires ten space-time dimensions

- We only observe 3, so 6 have to be small to evade detection $\Rightarrow$ compactifications



- The equations of motion put stringent constraints on the compactification geometry. The compact 6D space has to be a Calabi-Yau manifold

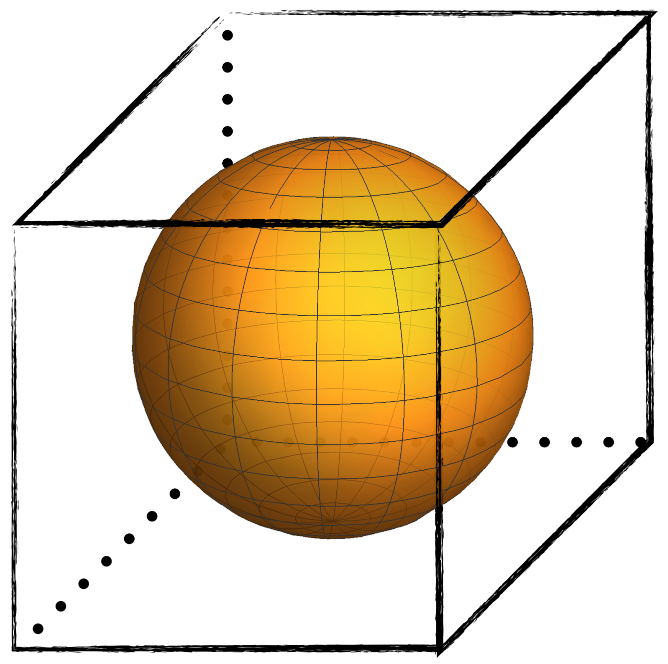
“…ok, but how do we describe the extra dimensions?”

Description of compact extra dimensions

- It is often easier to describe spaces inside larger ambient spaces:

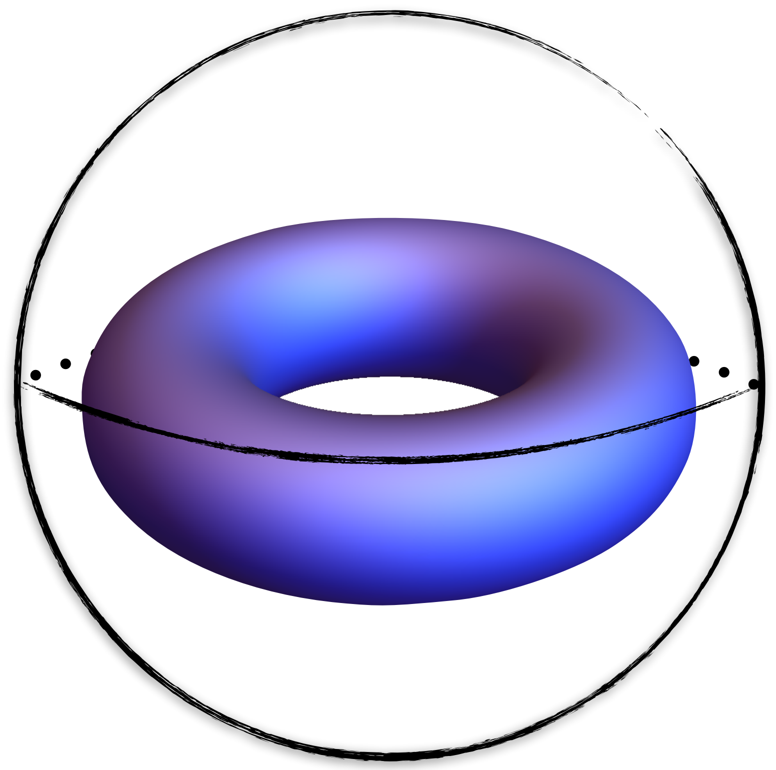
Sphere S^2 inside $\mathbb{R}^3$: $x^2 + y^2 + z^2 = R^2$

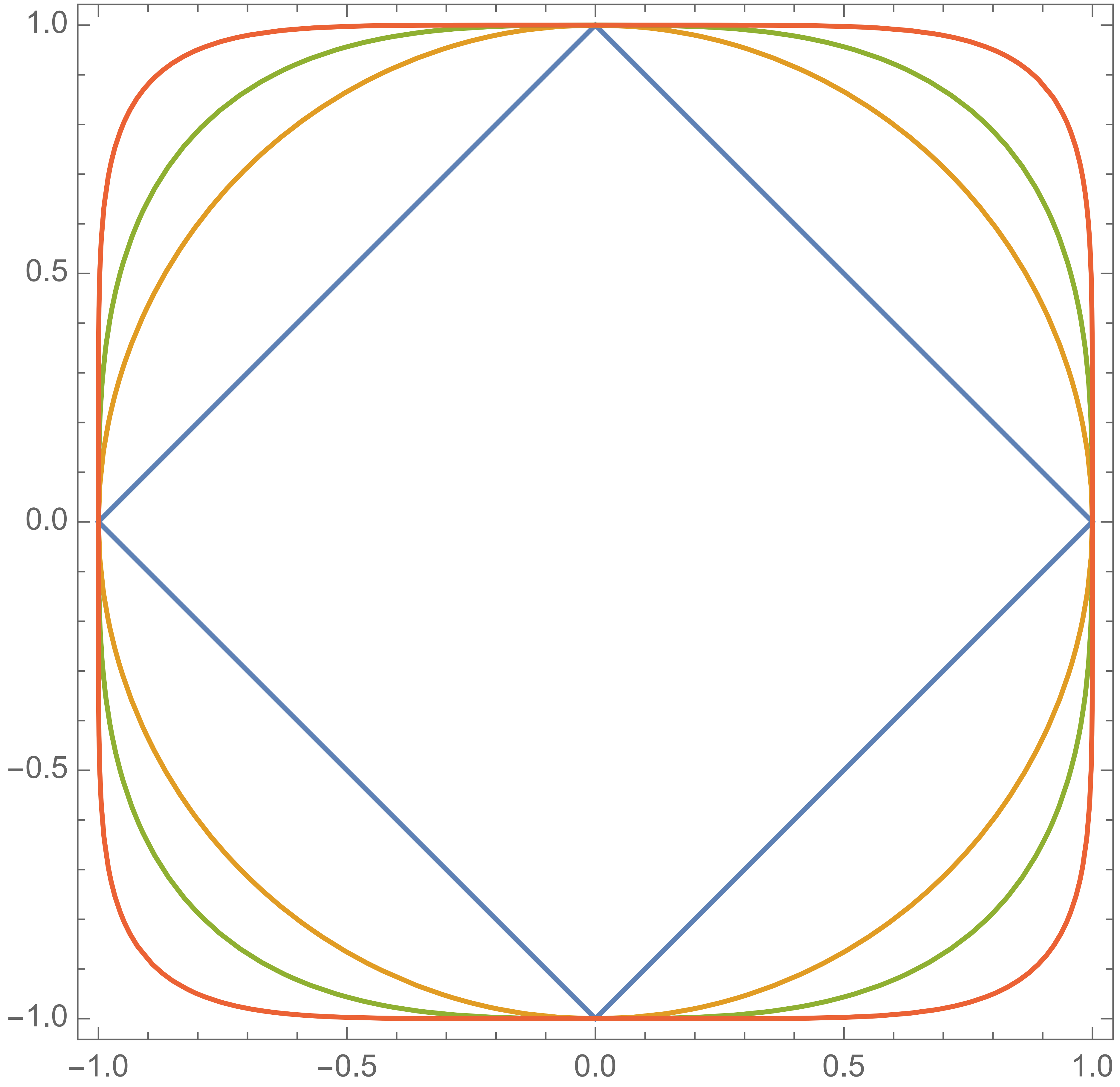
Radius



Torus T^2 inside $\mathbb{P}^2$: $x^3 + y^3 + z^3 = \psi\, xyz$

Shape



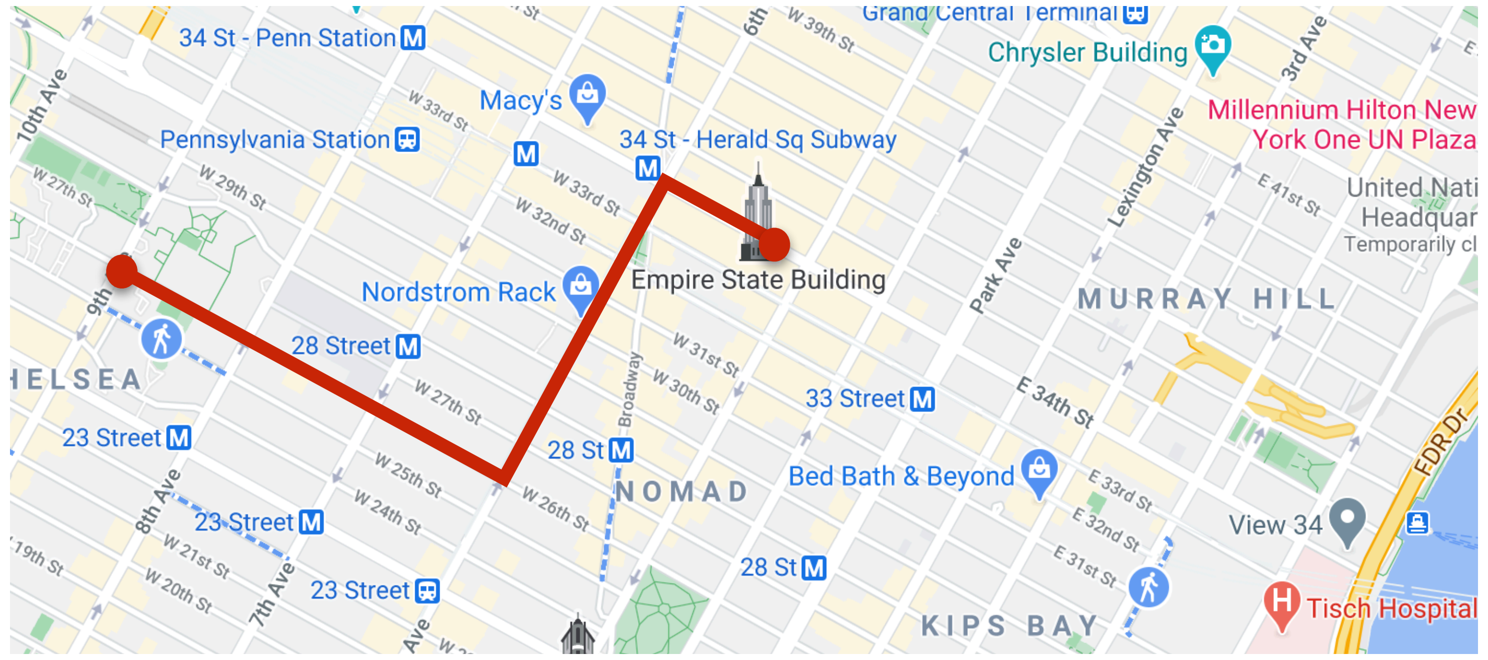


Metrics

Metrics

- Metrics measure distances, but the choice is not unique

[Source: wikipedia]

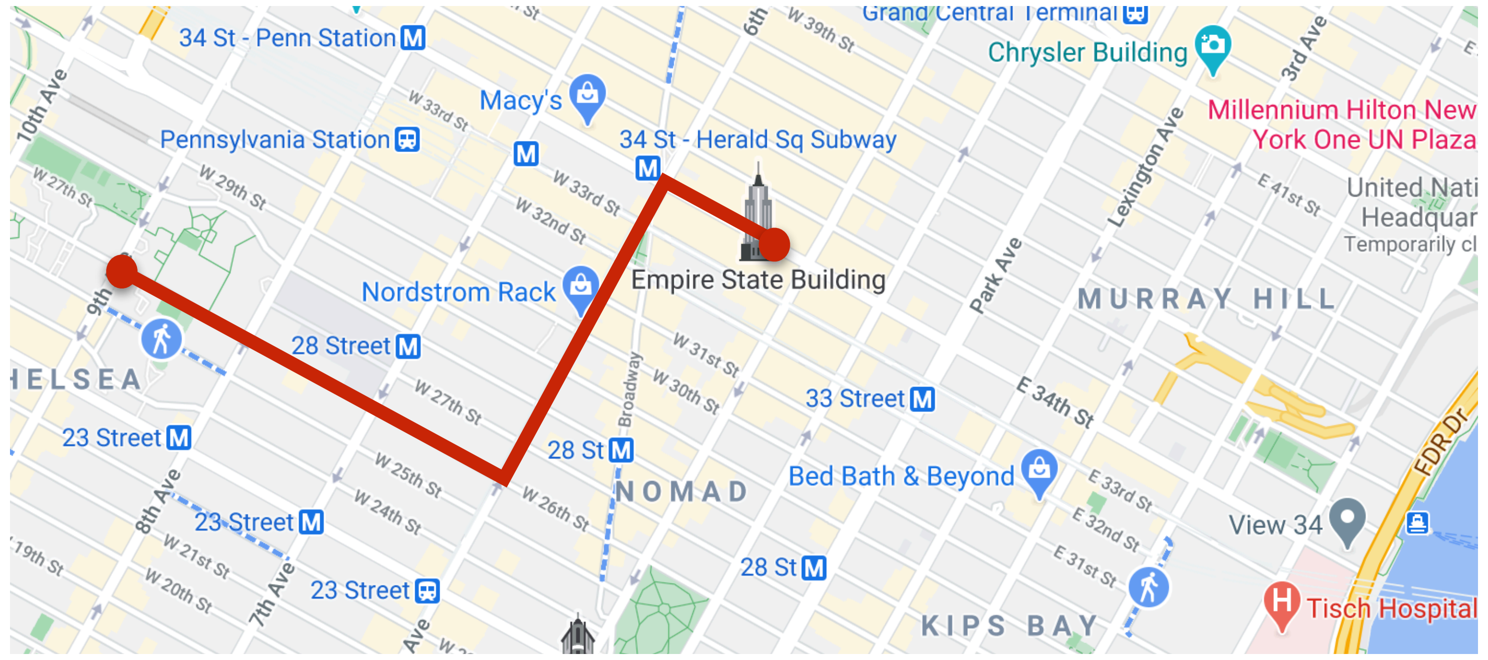


[Source: google maps]

Metrics

- Metrics measure distances, but the choice is not unique

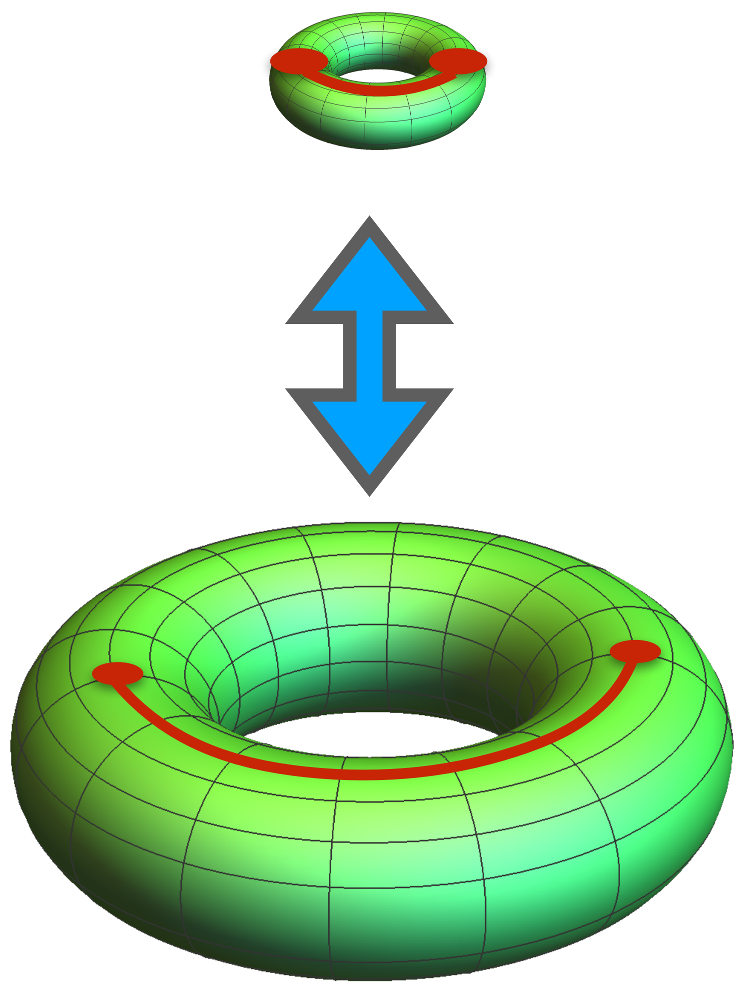
[Source: wikipedia]

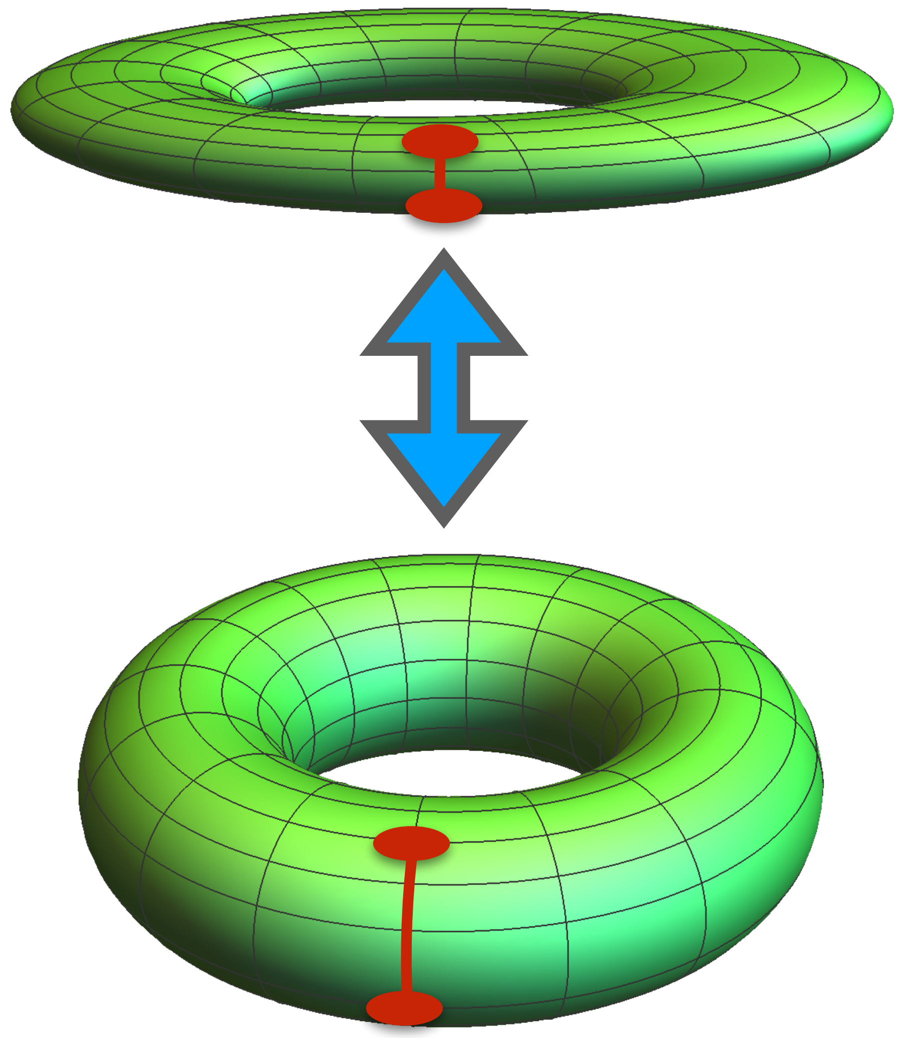


[Source: google maps]

- If space is curved, metric depends on the point you are at. It also depends on volume/shape

[Photo Credit: Jimmy Chin]

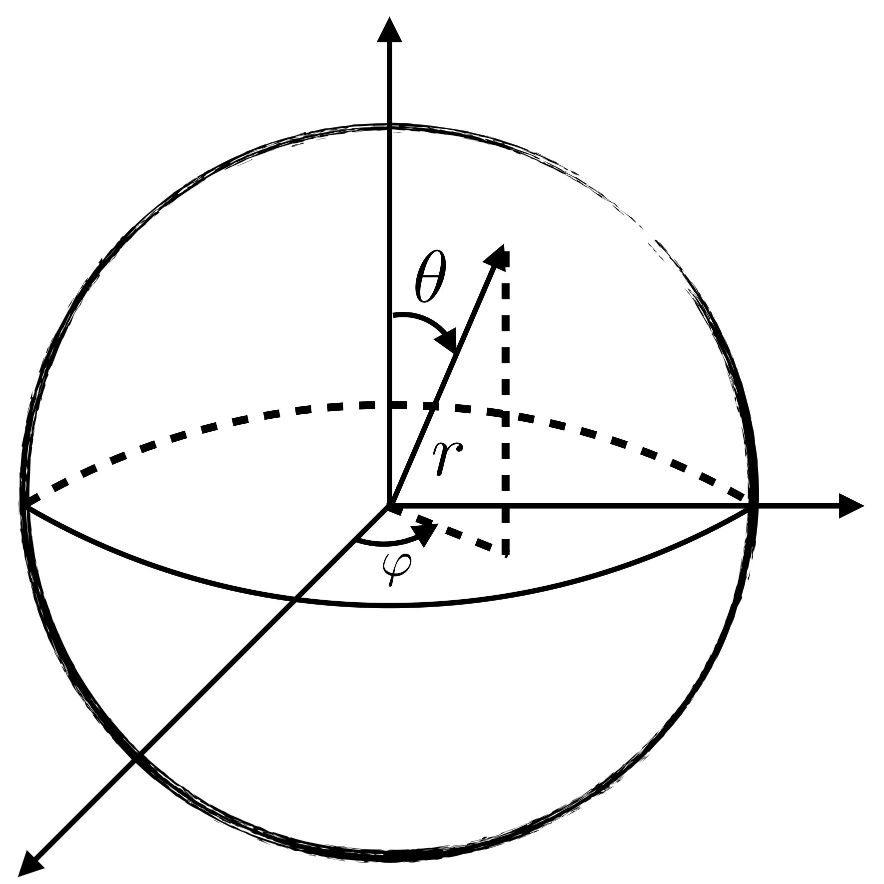




Metrics and integration

- Consider e.g. spherical coordinates

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2 \quad \Rightarrow \quad g_{ab} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}$$



- From this we can compute the volume of the 3D ball B^3

$$\int_{B^3} dV \quad = \quad \int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\theta dr\, d\theta\, d\varphi \quad , \qquad dV = \sqrt{\det g}\, dr\, d\theta\, d\varphi$$

Think of a metric g as a function

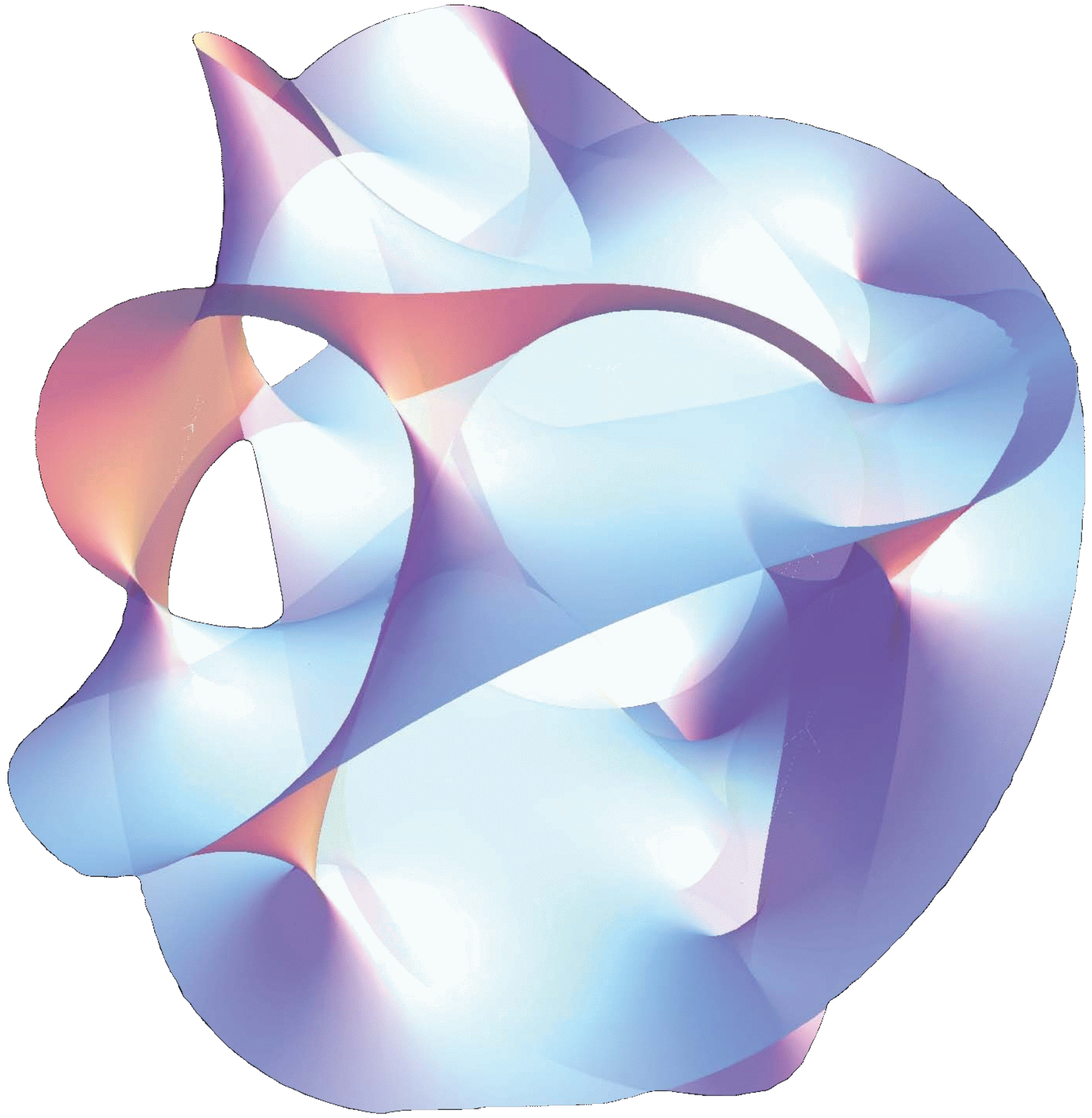
$$g: \quad \text{position} \times \text{volume} \times \text{shape} \rightarrow \mathbb{R}^{d \times d}$$

and optimize a NN to represent this function

Calabi-Yau Metrics

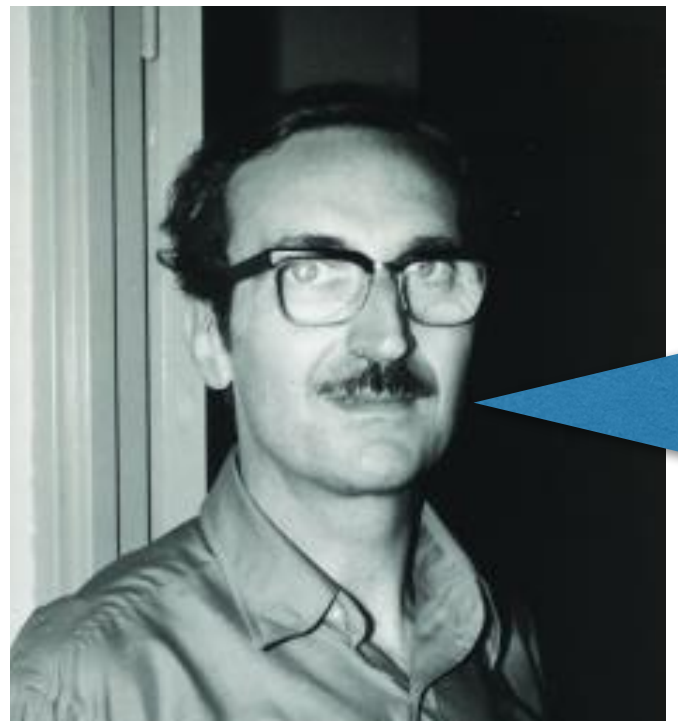
- Calabi-Yau spaces are spaces on which a metric exists that is "flat enough", i.e. their Ricci tensor vanishes

$$
\begin{aligned}
R_{ij} = & -\frac{1}{2}\sum_{a,b=1}^{n}\left(\frac{\partial^2 g_{ij}}{\partial x^a \partial x^b} + \frac{\partial^2 g_{ab}}{\partial x^i \partial x^j} - \frac{\partial^2 g_{ib}}{\partial x^j \partial x^a} - \frac{\partial^2 g_{jb}}{\partial x^i \partial x^a}\right) g^{ab} \\
& + \frac{1}{2}\sum_{a,b,c,d=1}^{n}\left(\frac{1}{2}\frac{\partial g_{ac}}{\partial x^i}\frac{\partial g_{bd}}{\partial x^j} + \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jd}}{\partial x^b} - \frac{\partial g_{ic}}{\partial x^a}\frac{\partial g_{jb}}{\partial x^d}\right) g^{ab} g^{cd} \\
& - \frac{1}{4}\sum_{a,b,c,d=1}^{n}\left(\frac{\partial g_{jc}}{\partial x^i} + \frac{\partial g_{ic}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^c}\right)\left(2\frac{\partial g_{bd}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^d}\right) g^{ab} g^{cd} \\
= 0 &
\end{aligned}
$$



Introduction to Calabi-Yau manifolds

CY metrics - Existence and Uniqueness

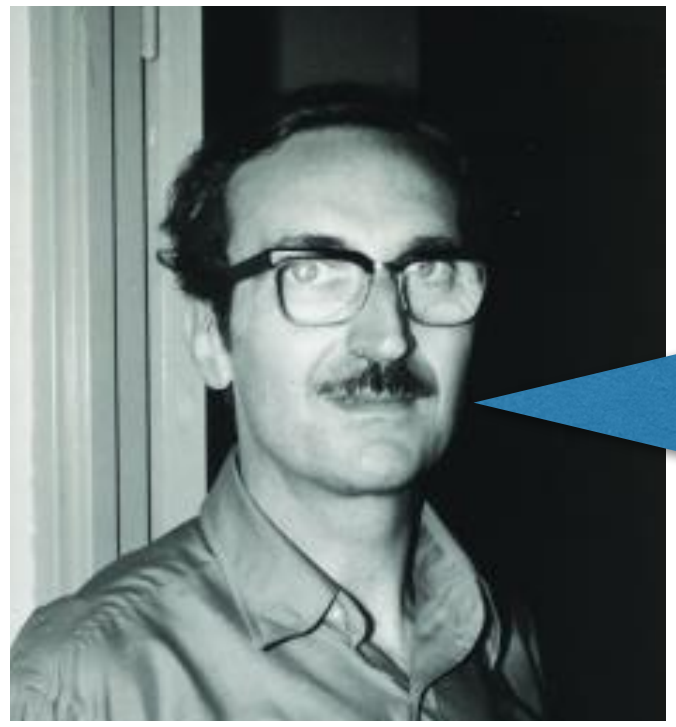


If some topological conditions are satisfied, I conjecture a Ricci-flat metric exists. I cannot prove that it exists, but I know it will be unique.

[Calabi `54]

[Image Source: wikipedia]

CY metrics - Existence and Uniqueness



[Calabi `54]

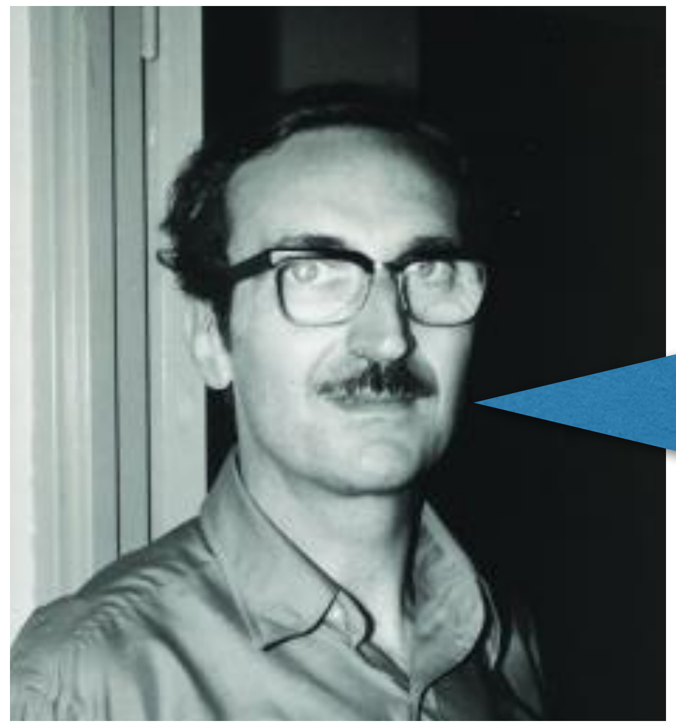
If some topological conditions are satisfied, I conjecture a Ricci-flat metric exists. I cannot prove that it exists, but I know it will be unique.

I can prove that such a metric exists, but I can neither tell you what it looks like nor how to construct it.

[Yau `77]

[Image Source: wikipedia]

CY metrics - Existence and Uniqueness



[Calabi `54]

If some topological conditions are satisfied, I conjecture a Ricci-flat metric exists. I cannot prove that it exists, but I know it will be unique.

I can prove that such a metric exists, but I can neither tell you what it looks like nor how to construct it.

[Yau `77]

[Fields `82]

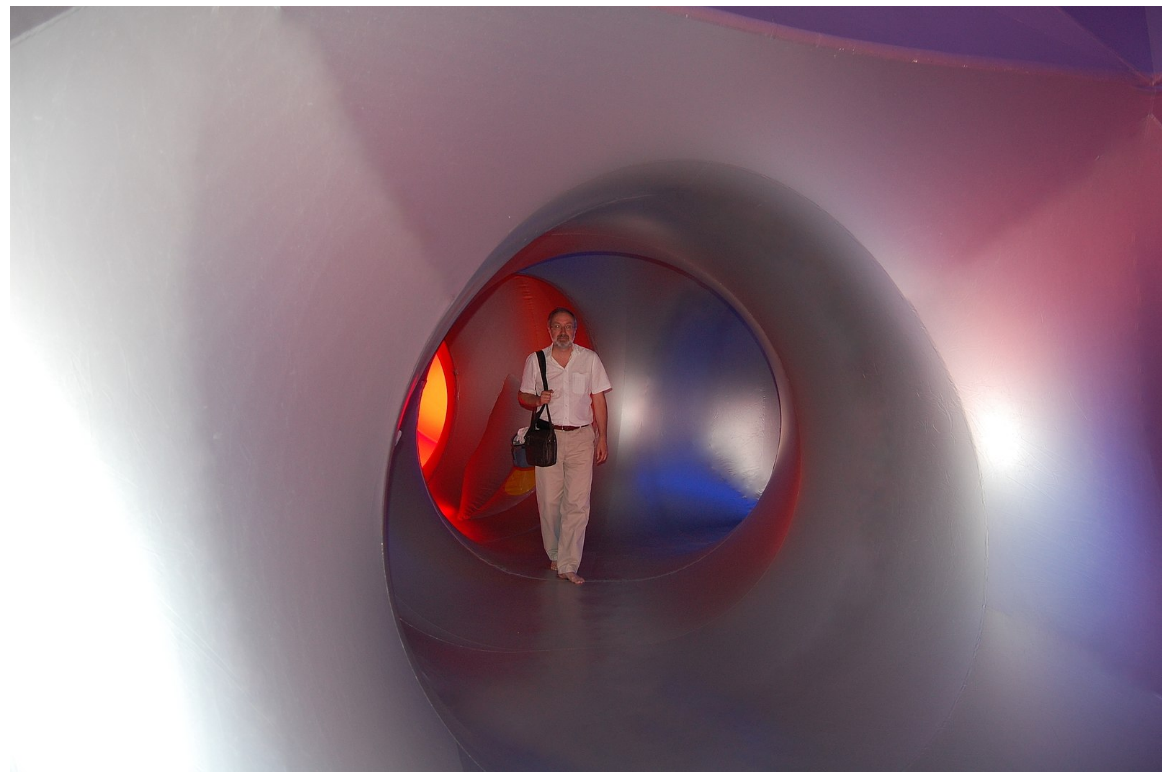
Great, take my medal!

[Image Source: wikipedia]

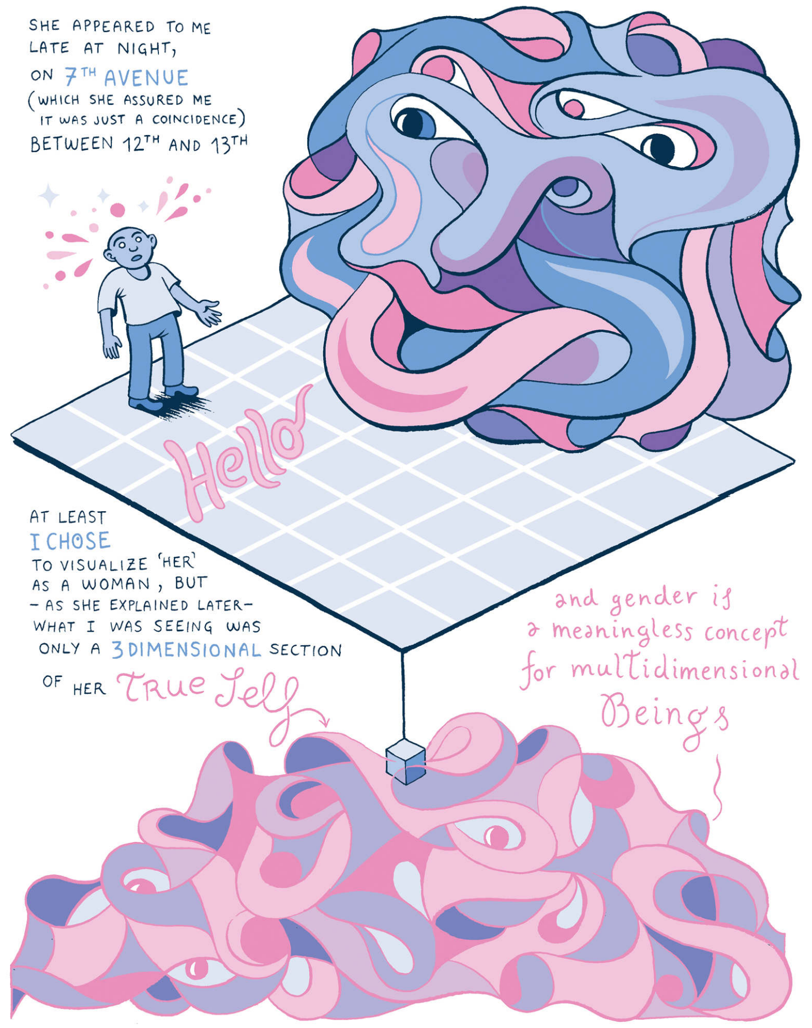
… even now, 40 years later, we still don’t know the geometry of the space-time we live in

Calabi-Yaus in pop culture

[Source: Half-Life 2]



[Source: wikipedia]

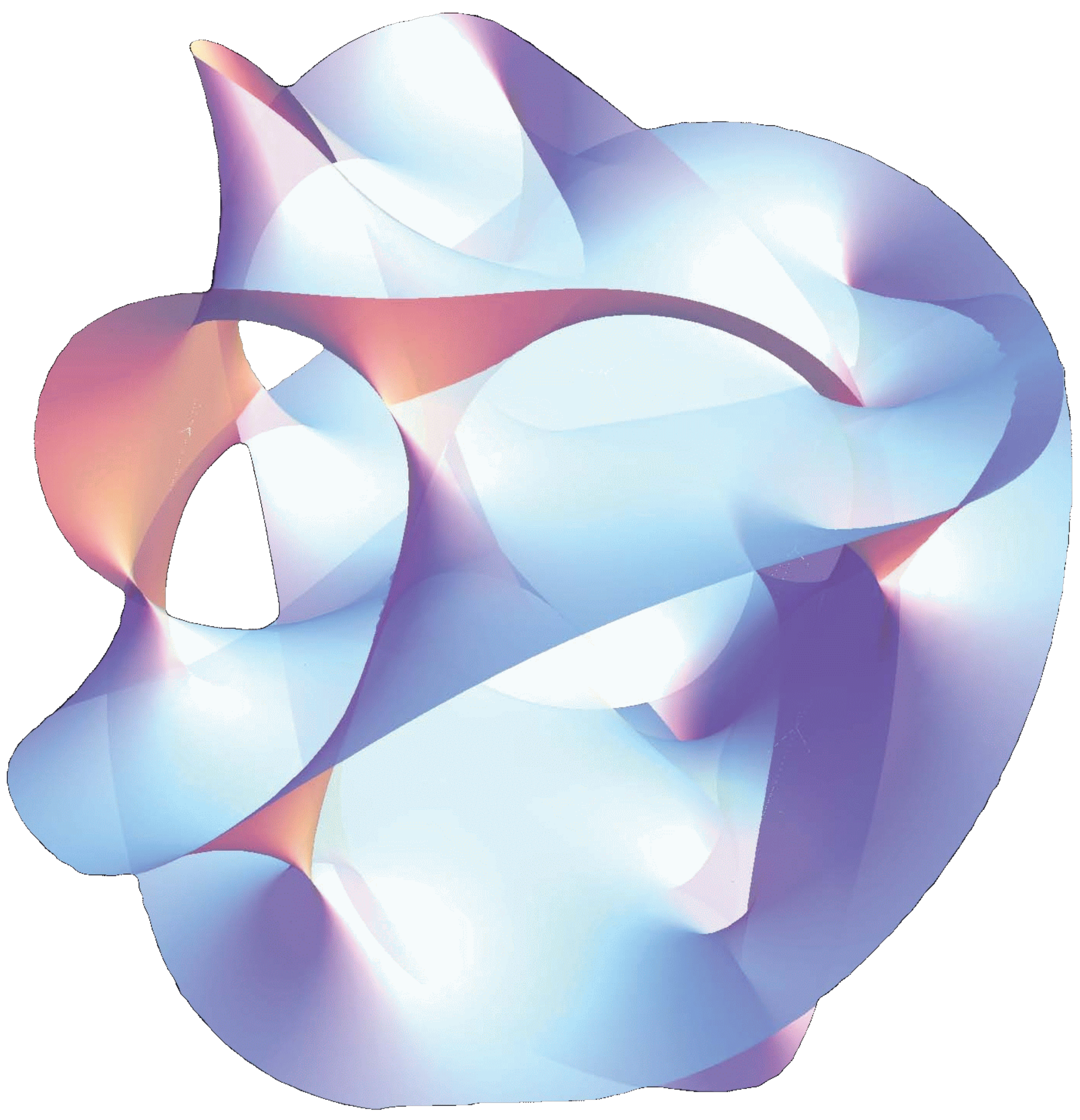


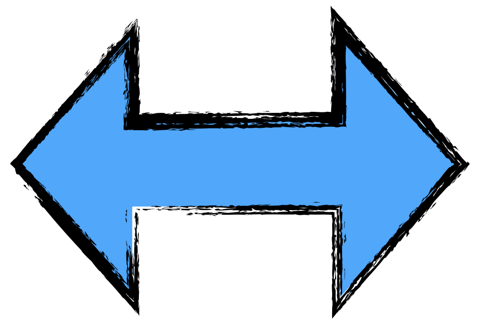
[Source: matteofarinella.com]

[Source: soundcloud.com]

[Source: tfwiki.net]

Calabi-Yau manifolds - Properties

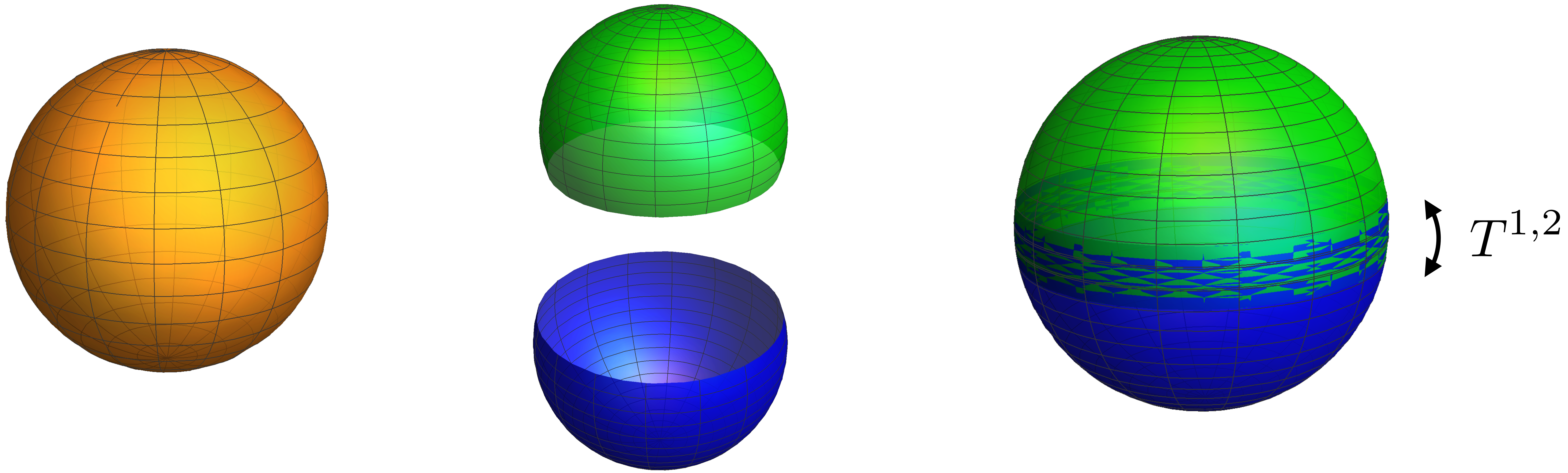




- Complex
- Kähler
- Ricci-flat

CY Property 1 - Complex

- In general manifolds cannot be covered by a single patch
- On each patch, one can choose a local description, coordinate system, etc. But one must make sure that the descriptions can be matched on the overlap and everything can be patched to a complex manifold globally (e.g. choice $i=\sqrt{-1}$ vs $i=-\sqrt{-1}$, …)



CY Property 2 - Kähler

- The space must be Kähler
- This means that the metric can be written in terms of derivatives of a real, scalar function called the **Kähler potential** K

$$g_{a\bar{b}} = \frac{\partial}{\partial z^a}\frac{\partial}{\partial \bar{z}^{\bar{b}}} K \quad , \qquad J = \frac{i}{2}\sum_{a<b} g_{a\bar{b}}\, \varepsilon^{a\bar{b}}\, dz^a d\bar{z}^{\bar{b}} \quad , \quad z = x + iy\,, \quad \bar{z} = x - iy$$

Kähler potential

Kähler form

- In general, integrating the metric to find the Kähler potential is hard. So one can either start with a Kähler potential and derive the metric, or one has to solve the differential equation $\frac{\partial J}{\partial z^a} = \frac{\partial J}{\partial \bar{z}^{\bar{b}}} = 0$.

CY Property 3 - Ricci-flat

- The CY metric needs to have Ricci tensor 0
- The Ricci tensor involves two derivatives of the metric, which itself involves 2 derivatives of K
- This fourth-order partial differential equation is extremely hard to solve
- We can improve on this. On a CY, one can write down

$$J = \frac{i}{2}\sum_{a<b} g_{a\bar{b}}\, \varepsilon^{a\bar{b}}\, dz^a d\bar{z}^{\bar{b}} \qquad \Rightarrow \qquad J^3 = -\frac{i}{8}\sqrt{\det g}\, dz_1\, d\bar{z}_1\, dz_2\, d\bar{z}_2\, dz_3\, d\bar{z}_3$$

$$\Omega = \left(\frac{\partial p}{\partial z_4}\right)^{-1} dz_1\, dz_2\, dz_3 \quad \Rightarrow \quad |\Omega|^2 = \left|\frac{\partial p}{\partial z_4}\right|^{-2} dz_1\, dz_2\, dz_3\, d\bar{z}_1\, d\bar{z}_2\, d\bar{z}_3$$

- Since the volume form is unique (up to a constant): $J^3 = \kappa|\Omega|^2$
- This leads to a second-order PDE of Monge-Ampere type

CY manifolds - simplest example

Let us illustrate the construction of CY manifolds. We start in 2D, i.e. with a torus.

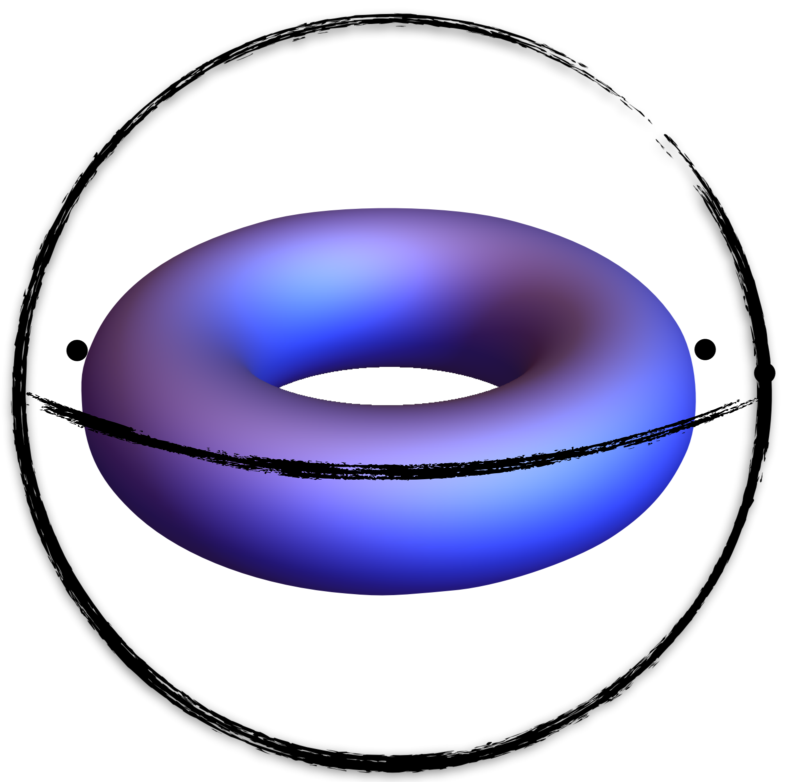
- First, we start with a complex Kähler ambient space $\mathbb{P}^2$

$$\mathbb{P}^2 = \{(z_0, z_1, z_2) \in \mathbb{C}^3 \mid (z_0, z_1, z_2) \sim (\lambda z_0, \lambda z_1, \lambda z_2)\,, \quad \lambda \in \mathbb{C}\}$$

- This ambient space is clearly complex. It is also Kähler, and a canonical metric is the Fubini-Study metric derived from the Kähler potential

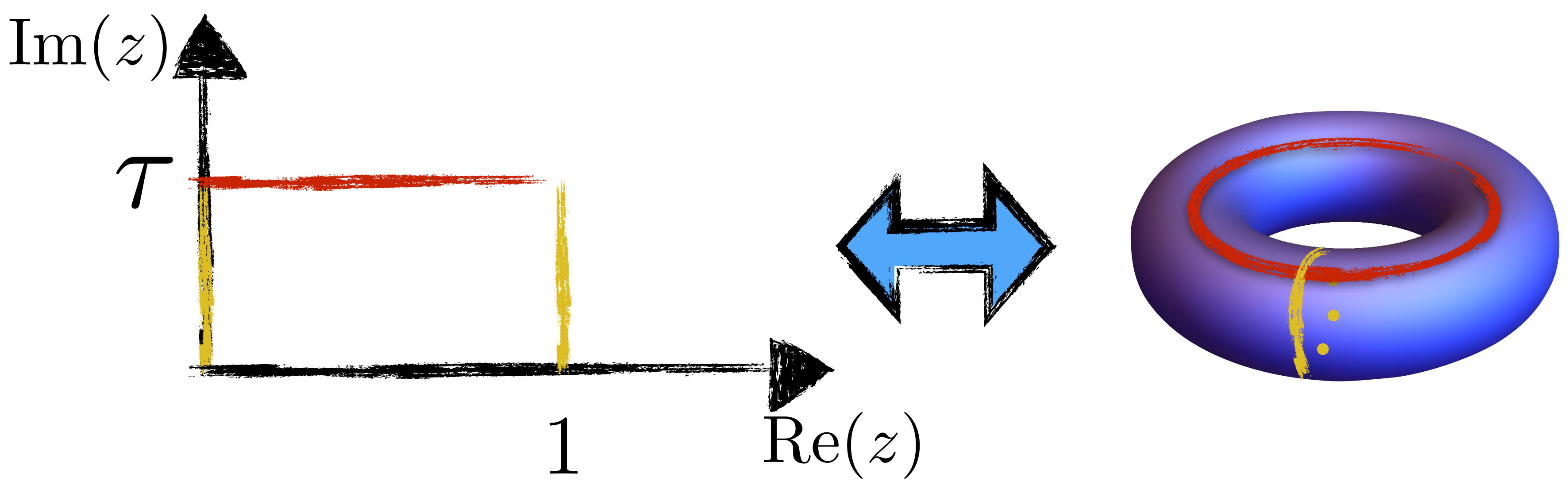
$$K = \log(|z_0|^2 + |z_1|^2 + |z_2|^2)$$

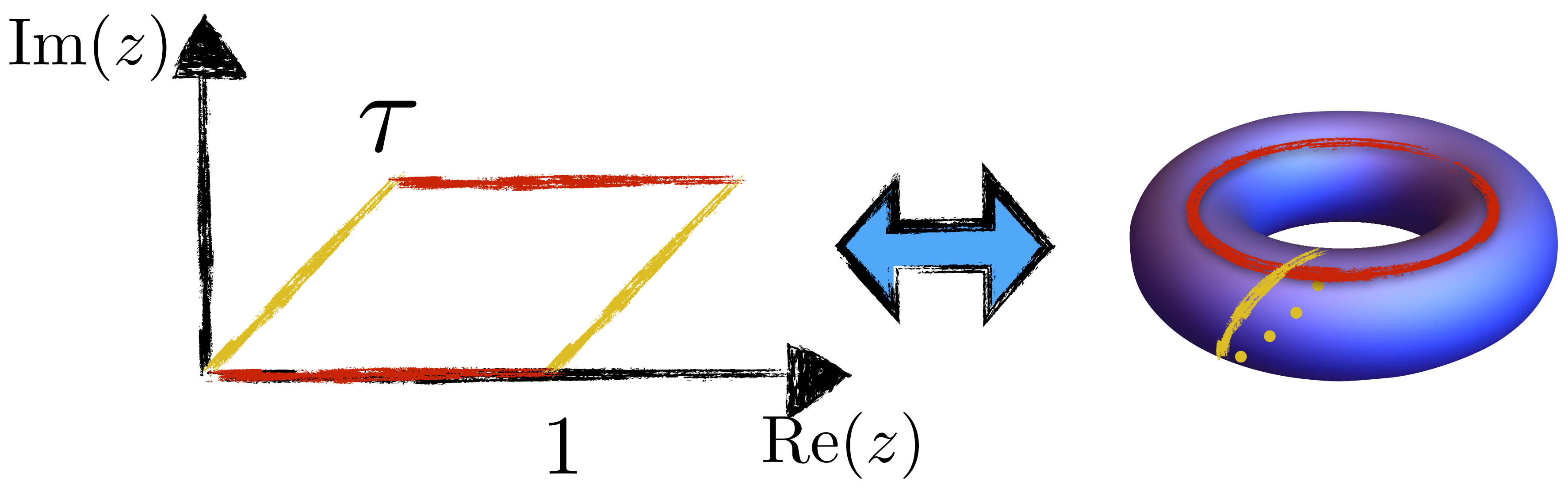
- Now we describe the torus as a complex 1D surface inside $\mathbb{P}^2$



CY manifolds - simplest example

- $\mathrm{tr}(R) = 0$ iff the equation is homogeneous of degree 3 in z_0, z_1, z_2
- Take e.g. $p = z_0^3 + z_1^3 + z_2^3 + \psi z_0 z_1 z_2 \overset{!}{=} 0$
- The torus is a 1D CY manifold. But inheriting the ambient FS metric does not lead to a Ricci-flat metric on the torus.
- However, the torus can also be described as $\{z \in \mathbb{C} \mid z \sim z + 1 \sim z + \tau\}$



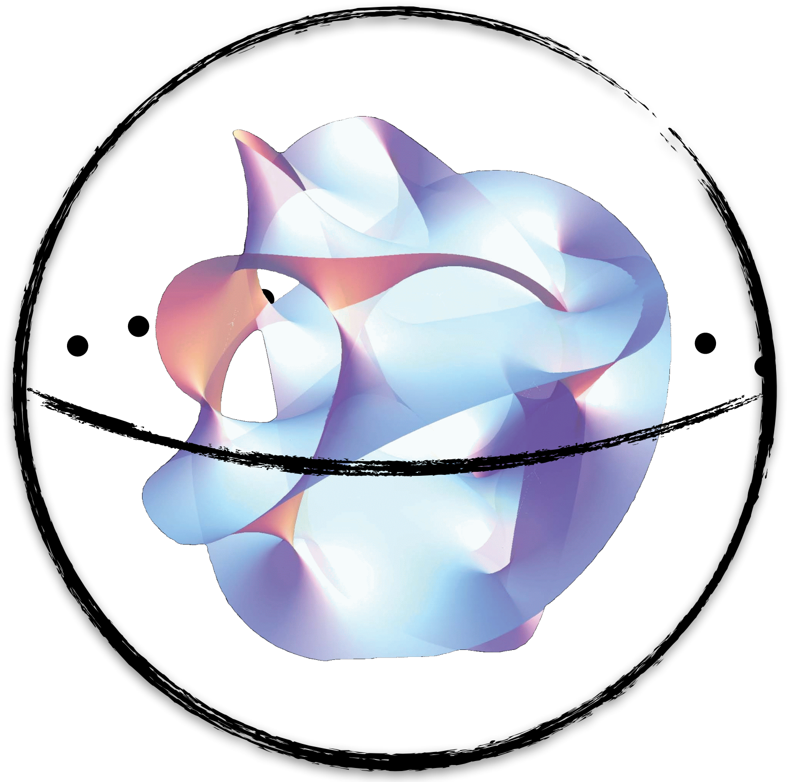


- The parameter τ describes the tilting (shape) of the torus and is related to the parameter ψ. These parameters are called complex structure parameters and are a priori not fixed.

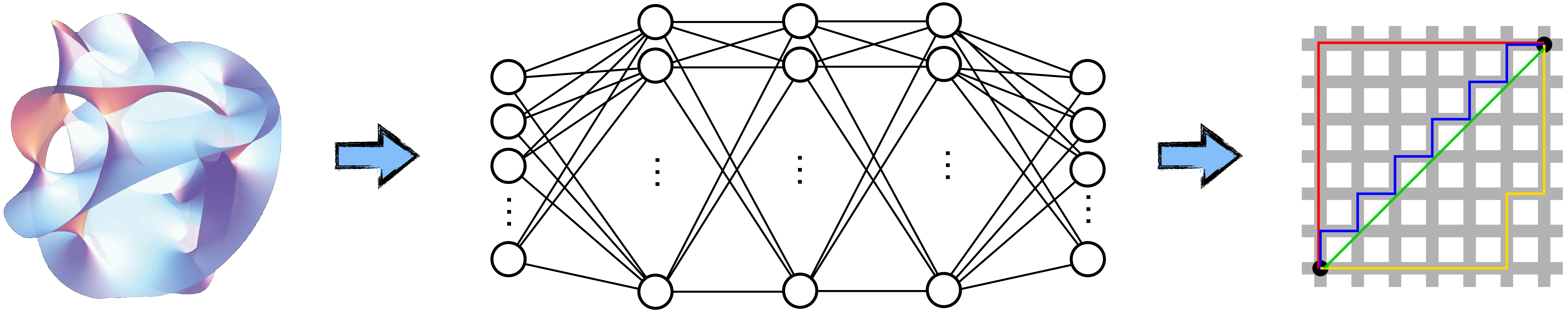
CY manifolds - simplest 6D example

- The same construction carries over to 3-folds. Start with a quintic equation in $\mathbb{P}^4$:

$$p = z_0^5 + z_1^5 + z_2^5 + z_3^5 + z_4^5 + \psi z_0 z_1 z_2 z_3 z_4 \stackrel{!}{=} 0$$



- Again, this manifold is complex, Kähler and has $\mathrm{tr}(R) = 0$, so a Ricci-flat metric does exist by Yau's theorem
- The parameter ψ encodes the complex structure of the CY manifold
- While we do not know the Ricci-flat metric, we can readily construct Ω



ML for Calabi-Yau metrics

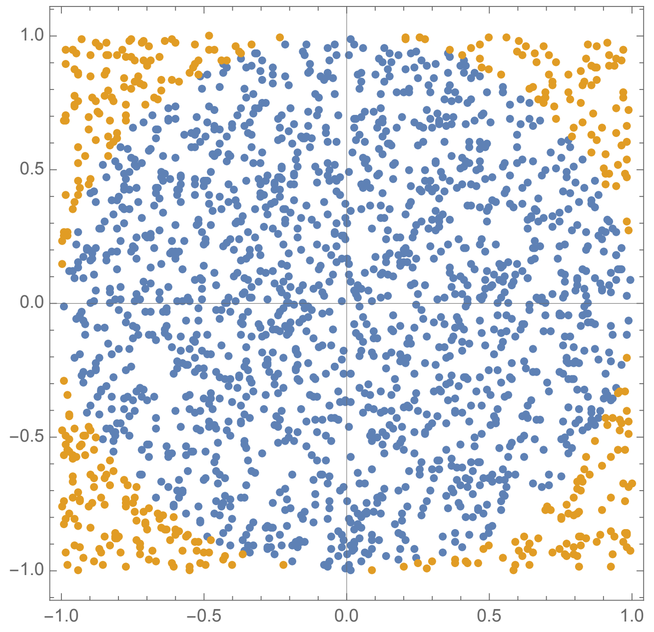
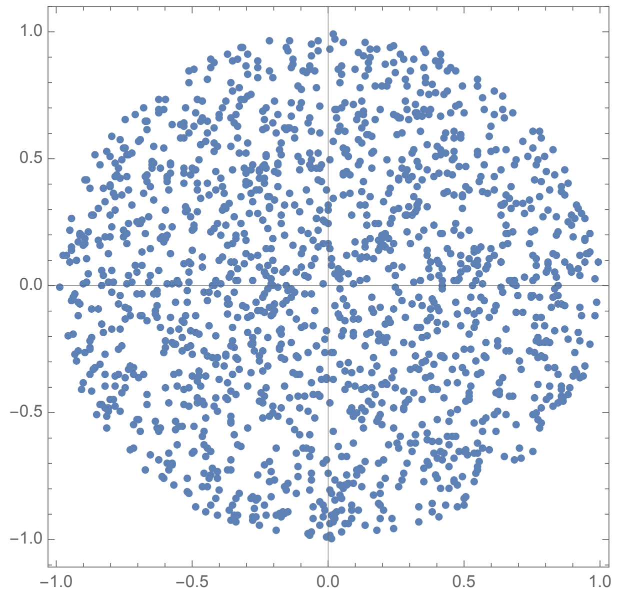
Steps

1. Find points on CY
2. Optimize NN
   - Overlap
   - Kahler
   - Ricci-flat
3. Repeat for different shapes and volumes
4. CY metric

1. Sampling points

Random sampling depends on the metric. Take a disk $x = r\cos\varphi\,,\;\; y = r\sin\varphi$

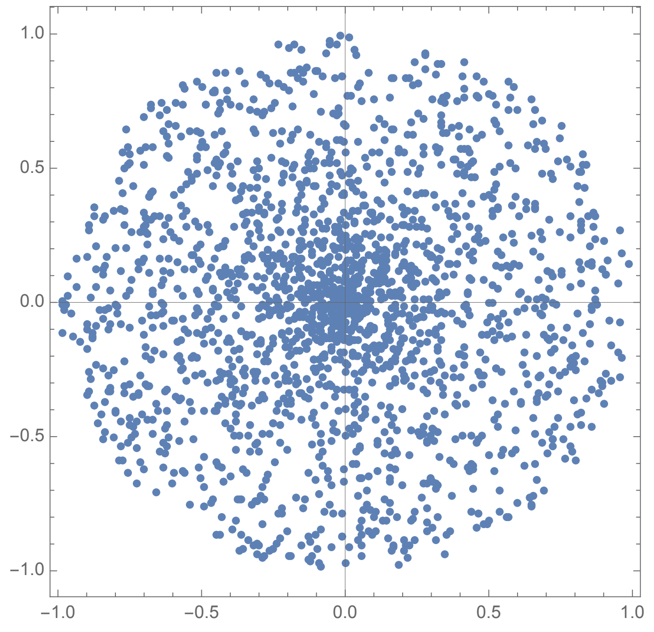
Rejection sampling

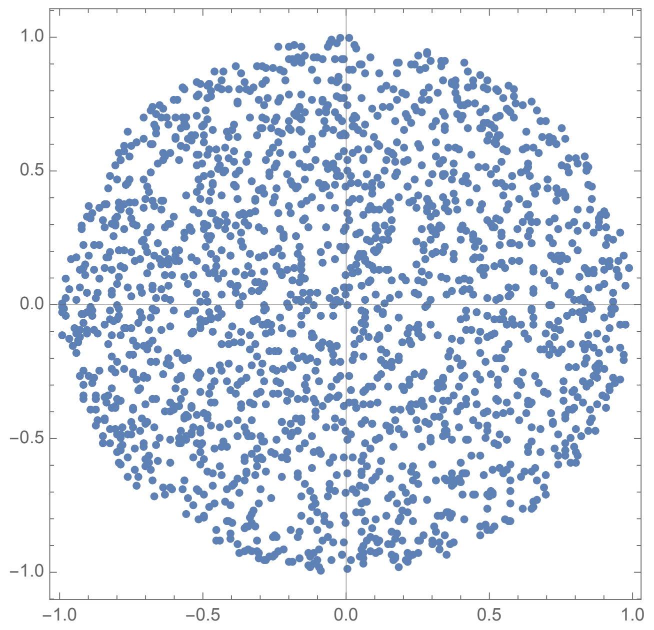
$x \in [-1,1]\,,\; y \in [-1,1]$

Reject samples that are not on the manifold

Random sampling



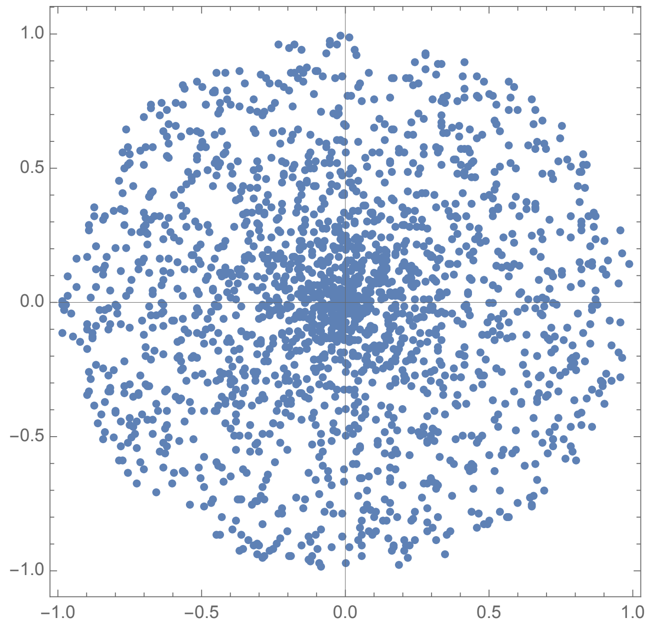
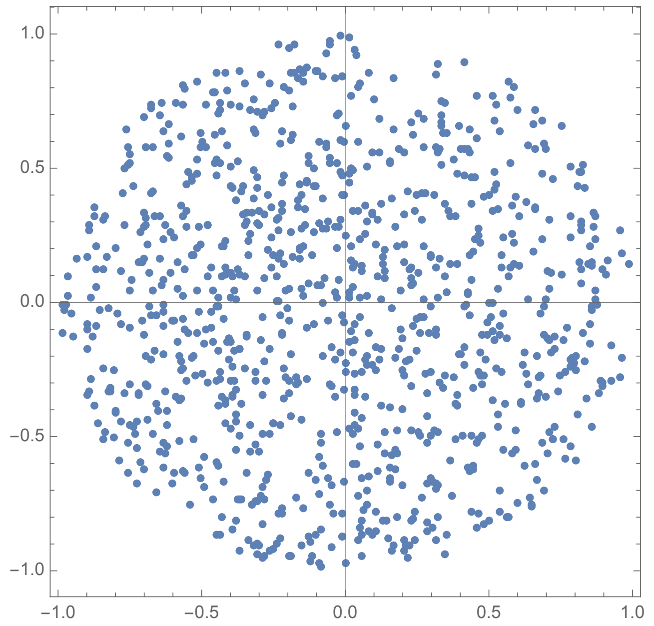
$r \in [0,1]\,,\; \varphi \in [0,2\pi]$



$r^2 \in [0,1]\,,\; \varphi \in [0,2\pi]$

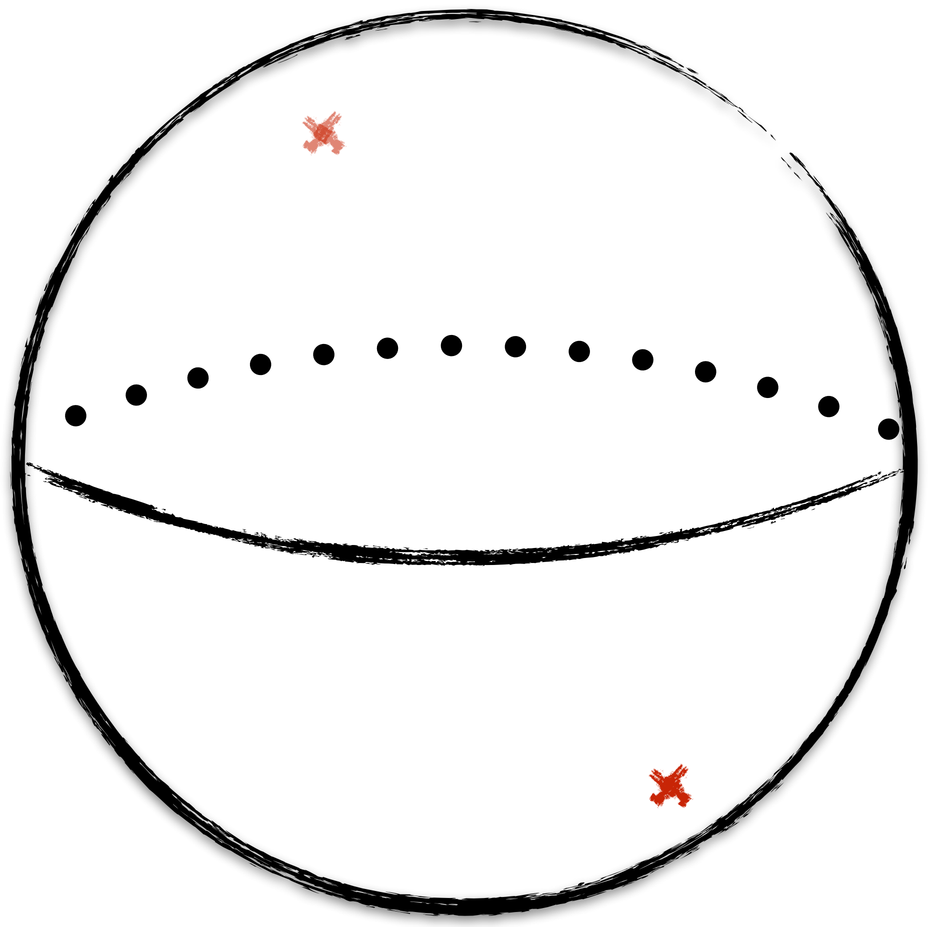
Sample w.r.t. correct measure

Resampling

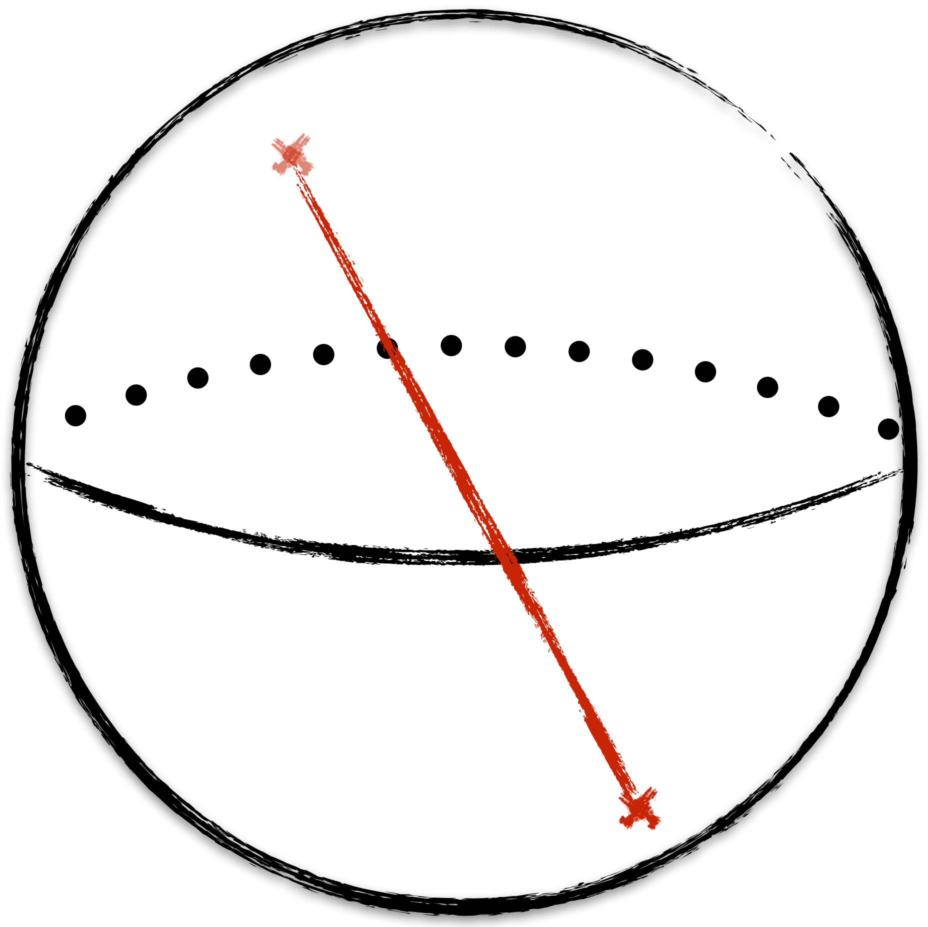



- Find points w.r.t. wrong measure
- Assign weight to each point
- Draw points with prior given by these weights
- Alternatively, use all points and weight the loss by the weights

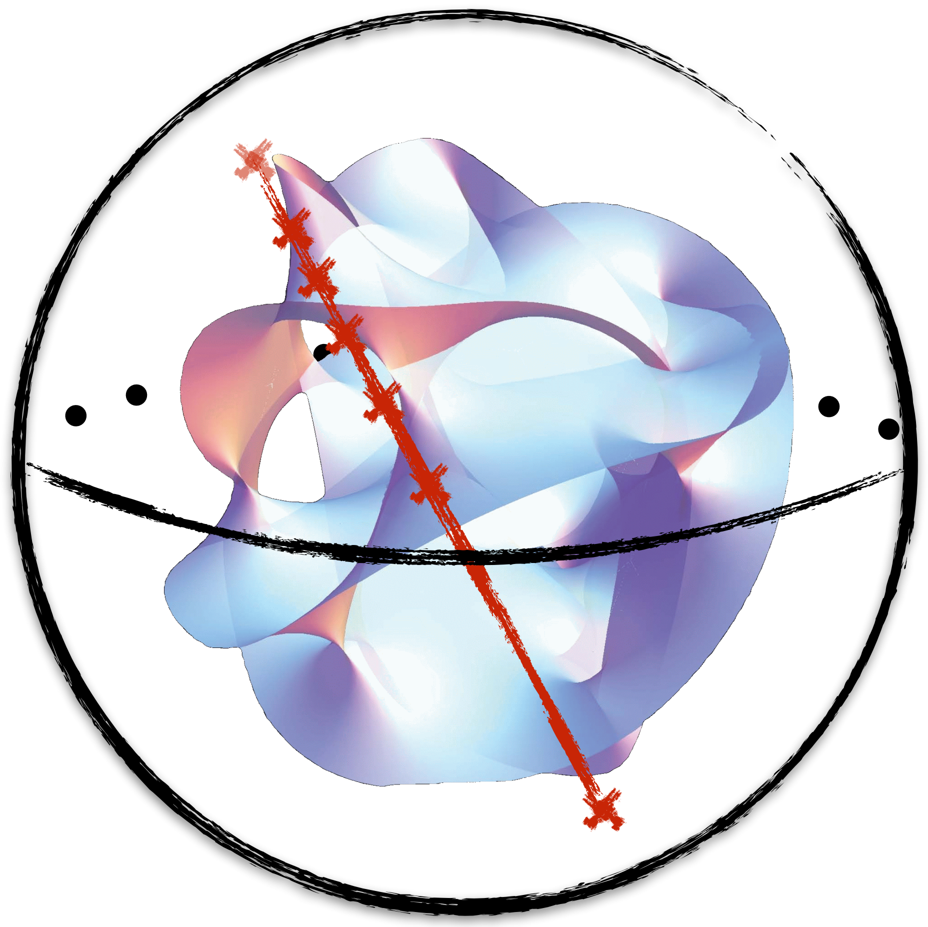
1. Find points on CY



Sample points from a sphere inside $\mathbb{P}^4$



Find the line through these ambient points



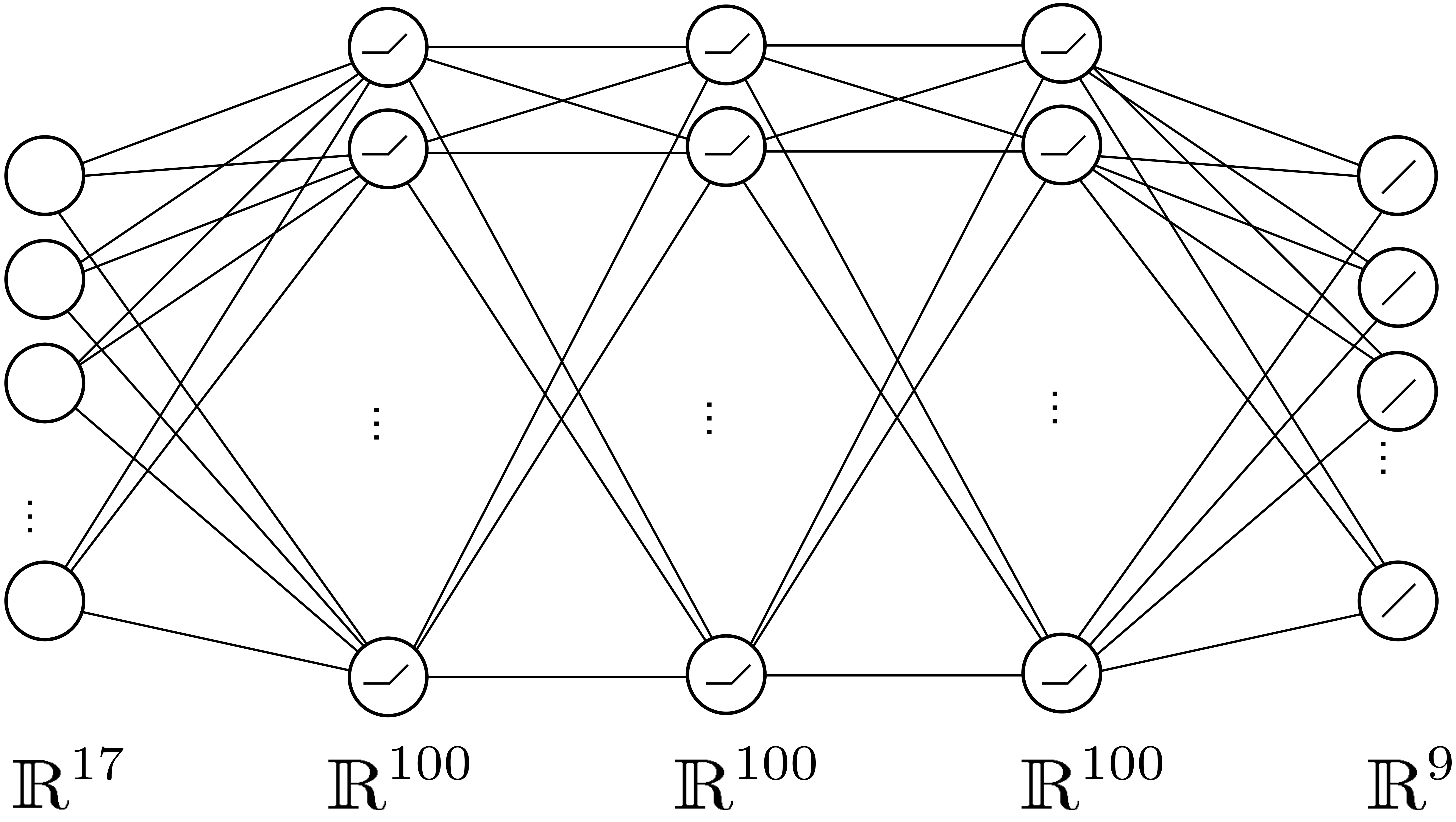
Intersect the CY eqn with the line ⇒ 5 solns.

Then, the weights for a "uniform" sampling on the CY are given by $w = \dfrac{|\Omega|^2}{J_{\mathrm{FS}}^3}$

2. Setup the NN

- Input of the NN:
  - The point on the CY (5 complex coordinates $\Rightarrow$ 10 input nodes)
  - The patch in which to evaluate the NN (a 5D vector in $[0,1]^{\times 5}$ with all 0 and a 1 at position i to indicate the patch $z_i \neq 0$)
  - The point in CS space at which to compute the metric (2 real coordinates $\Rightarrow$ 2 input nodes)
- Output of the NN: The independent component of the hermitian metric (3 real dof's on the diagonal, 3 complex off-diagonal dof's $\Rightarrow$ 9 output nodes)

2. Setup the NN



Leaky ReLU activation function

Identity activation function

2. Setup the NN

1. Take the pullback of the FS metric and learn the CY metric as a correction:

$$g_{\mathrm{CY}} = g_{\mathrm{FS}}(\mathbb{1} + g_{\mathrm{NN}})$$

2. Tune the NN parameters such that the metric satisfies at all points:

(i) Overlap: $g^{(i)}_{\mathrm{CY}} = T_{i,j}\ g^{(j)}_{\mathrm{CY}}\ T^{\dagger}_{i,j} \quad \Rightarrow \quad L_1 = \sum_{\mathrm{points}} \sum_{i,j} \sum_{a,b} \left[(g^{(i)}_{\mathrm{CY}} - T_{i,j}\ g^{(j)}_{\mathrm{CY}}\ T^{\dagger}_{i,j})_{ab} \right]^2$

(ii) Kähler: $\partial_a J = \overline{\partial}_{\bar{b}} J = 0 \ \Rightarrow L_2 = \sum_{\mathrm{points}} \sum_{a} \sum_{i,\bar{j}} \left[\partial_a J_{i\bar{j}} - \partial_i J_{a\bar{j}} \right]^2 + \sum_{\mathrm{points}} \sum_{\bar{b}} \sum_{i,\bar{j}} \left[\partial_{\bar{b}} J_{i\bar{j}} - \partial_{\bar{j}} J_{a\bar{b}} \right]^2$

(iii) Monge-Ampere equation: $J^3 = \kappa |\Omega|^2 \quad \Rightarrow \quad L_3 = \sum_{\mathrm{points}} \sigma^2 \quad , \ \sigma := \left| 1 - \frac{J^3}{|\Omega|^2} \right|$

3. The overall loss is $L_{\mathrm{tot}} = \lambda_1 L_1 + \lambda_2 L_2 + \lambda_3 L_3$
   We perform a box search over the hyperparameters λ_i and choose $\lambda_3 = 10,\ \lambda_1 = \lambda_2 = 1$

2. Setup the NN

- The result does not depend crucially on the NN architecture/activation function (ReLU, Tanh, GELU)
- We train the NN with 50’000 points per ψ for $\mathrm{Re}(\psi), \mathrm{Im}(\psi) \in [-100, 100]$.
    - We use a batch size of 900 and ADAM optimizer, train for 20 epochs
    - Adding more points does improve the accuracy
- We use an evaluation set of 5’000 points
- We also tried training longer and including measures to prevent overfitting (batch normalization, dropout layers, early stopping), but this did not improve the result
- We implemented this in PyTorch and Tensorflow
- We need derivatives of the NN w.r.t. input as well as w.r.t. parameters
    - No direct backward hooks in PyTorch 1.6
    - Bug in TF 2.4 with 2 gradient tapes and pfor

2. Setup the NN

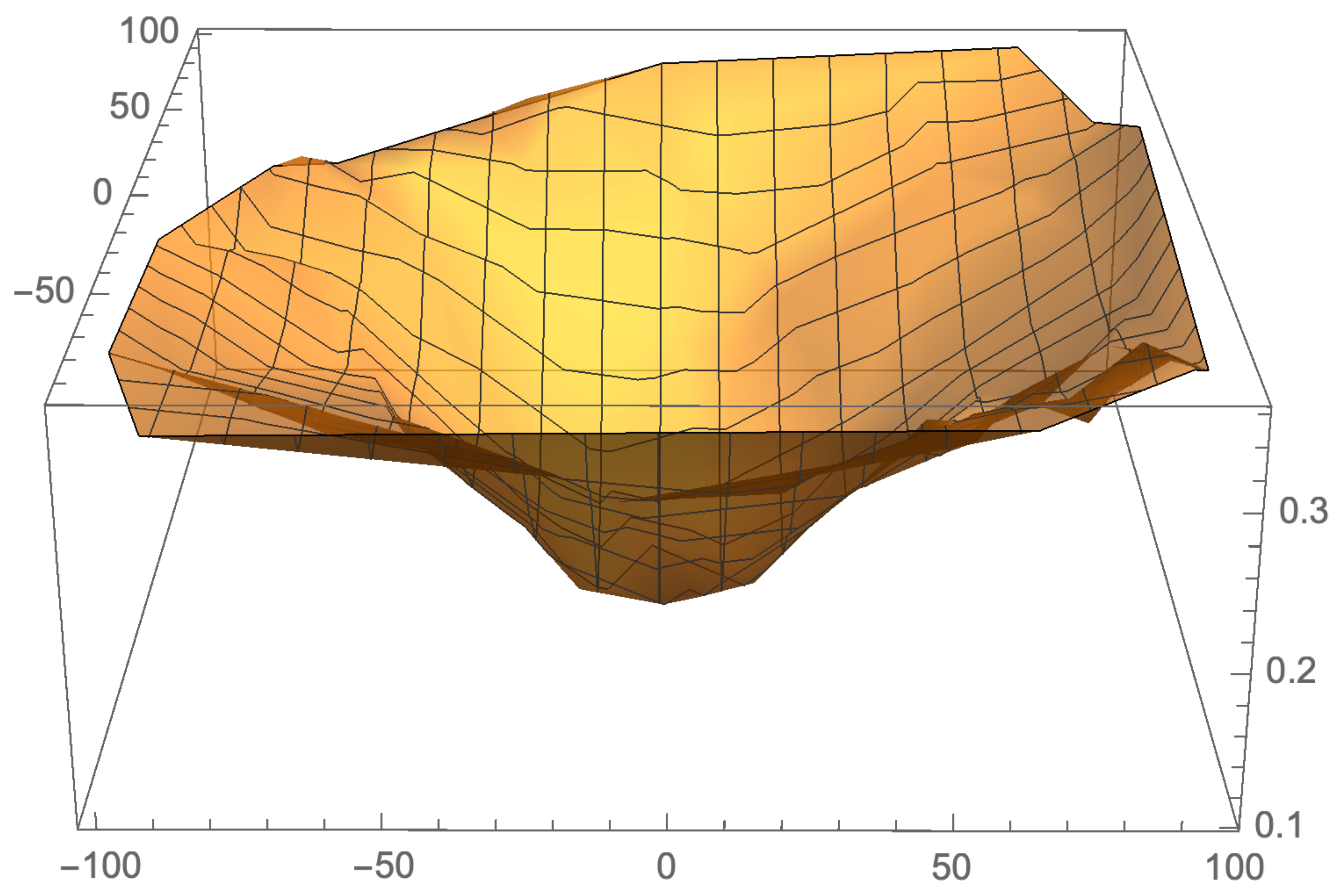
- We need derivatives of the NN w.r.t. input as well as w.r.t. parameters
  - No direct backward hooks in PyTorch 1.6

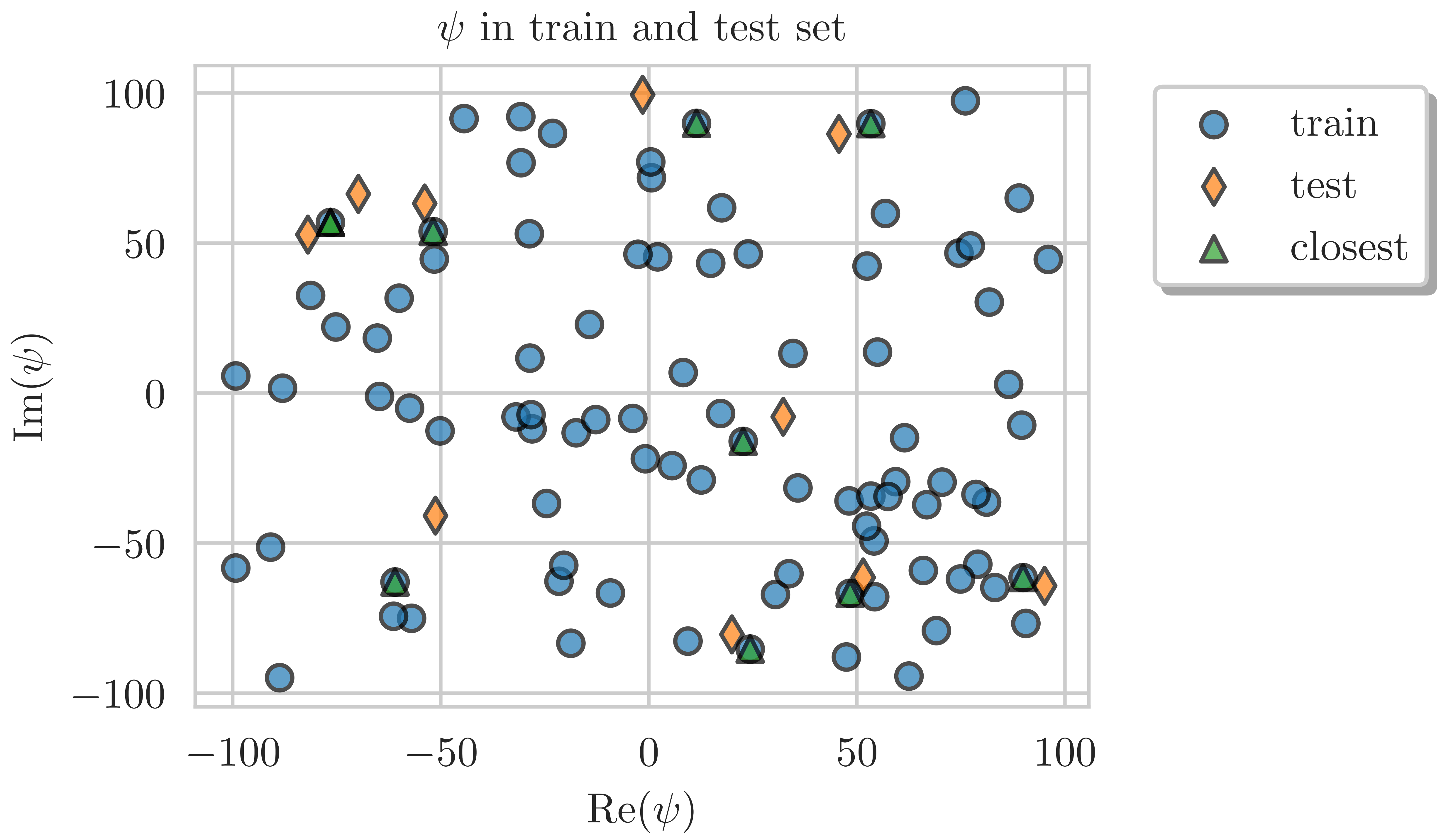
```
def get_jacobian_reim(self, y, x):
    jacobian = []

    for i in range(y.shape[-1]):
        v = torch.zeros_like(y)
        v[:, i] = 1.
        dy_i_dx = torch.autograd.grad(y, x, grad_outputs=v, create_graph=True, retain_graph=True)[0]
        dy_i_dx = dy_i_dx[:, :2 * self.dim_a]
        jacobian.append(dy_i_dx)

    jacobian = torch.stack(jacobian, dim=2).requires_grad_()
    return self.jacs_to_re_im(jacobian)
```

  - Bug in TF 2.4 with 2 gradient tapes and pfor

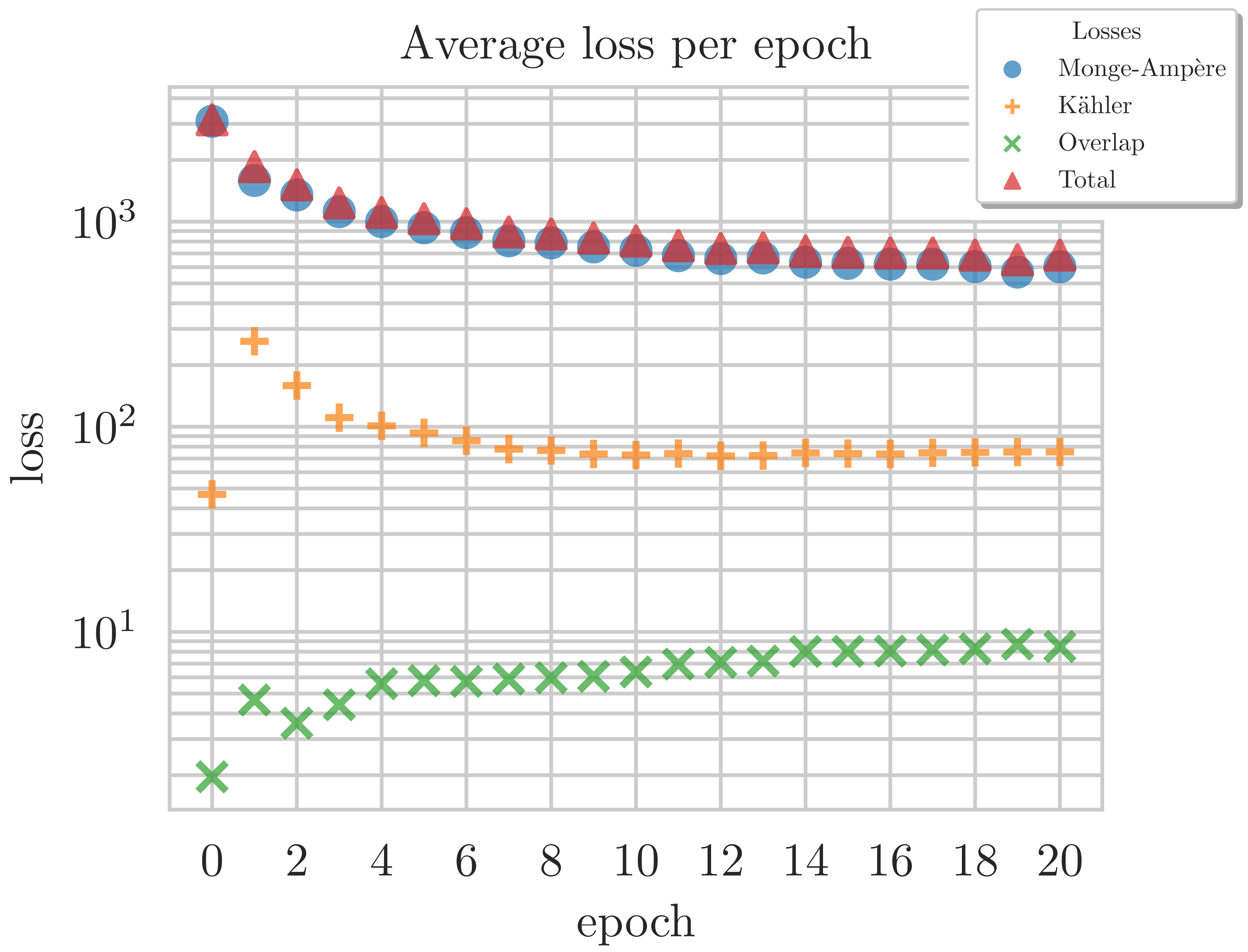
```
def get_jacobian_reim(self, y, x):
with tf.GradientTape(persistent=True, watch_accessed_variables=True) as t1:
    with tf.GradientTape(persistent=True, watch_accessed_variables=True) as t2:
        t2.watch(x)
        y_pred = fsmodel.pullbacks(x)[:,:,3] * y
    batch_jac = t2.batch_jacobian(y_pred, x)
```

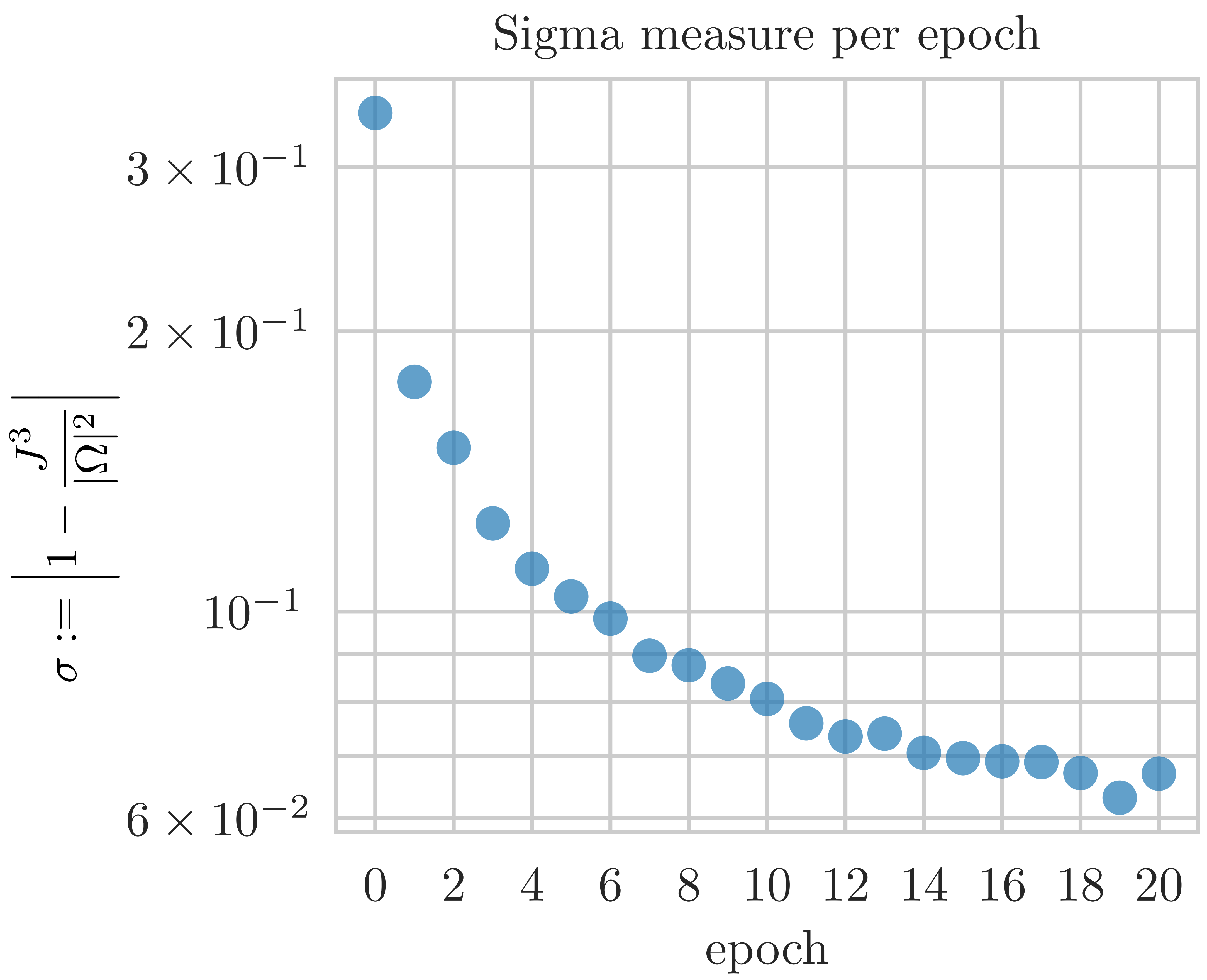
3. Learn the metric





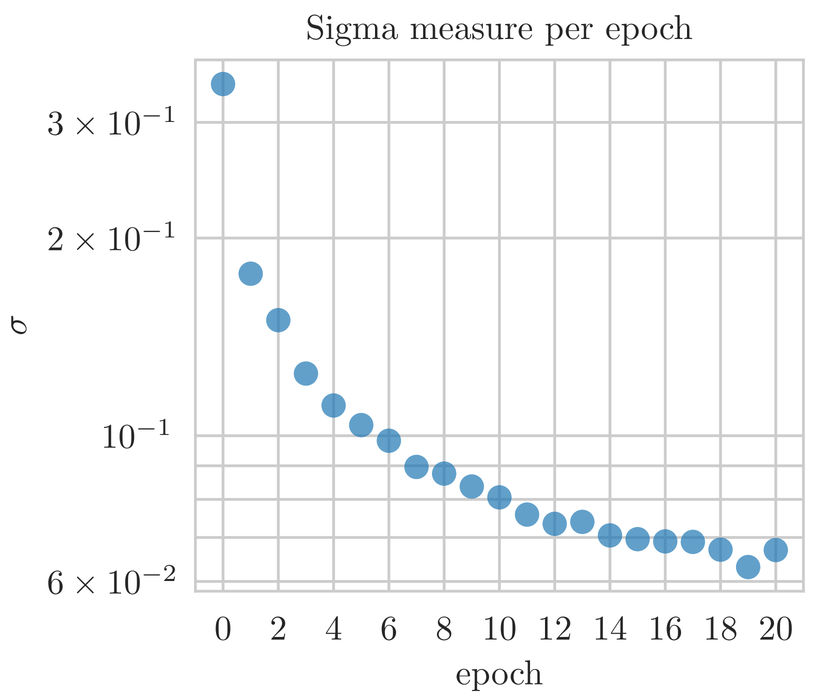
3. Learn the metric



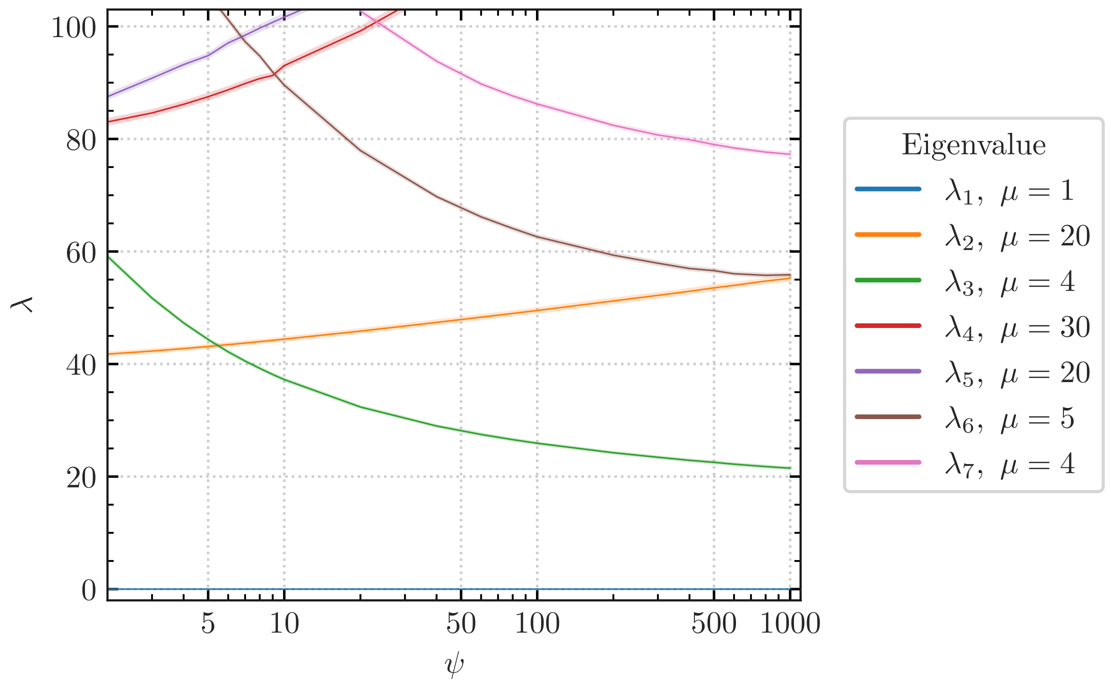


Conclusions

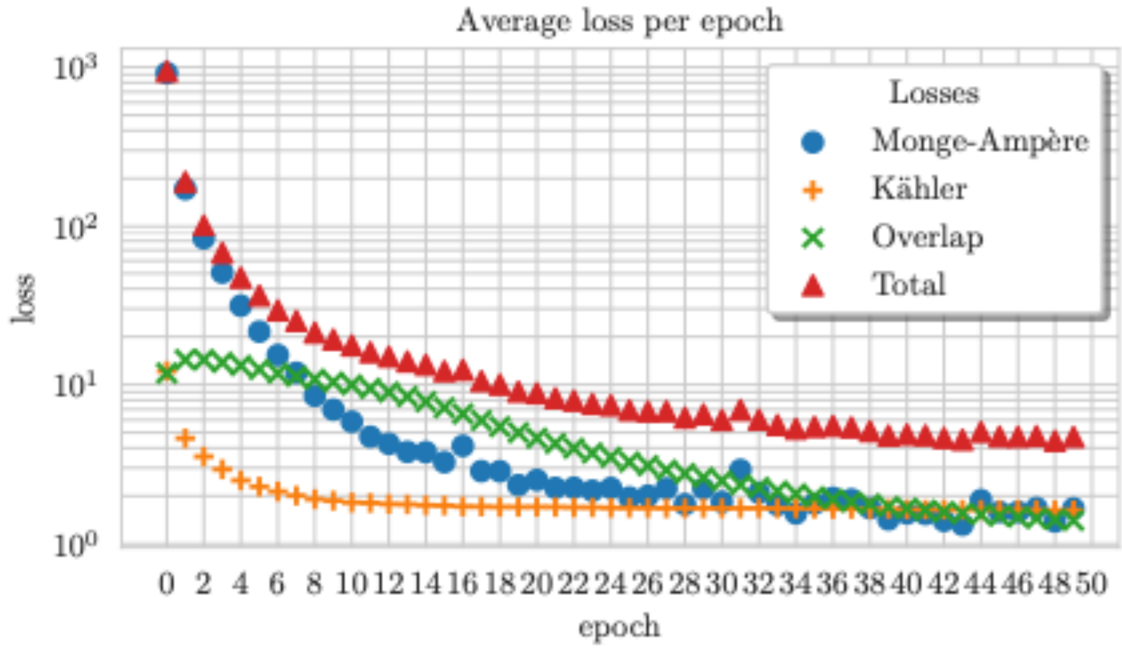
- NNs can learn the moduli- (and position-) dependent CY metric by optimizing the underlying differential equations



- This has many applications in Physics (massive string modes, swampland conjectures, Yukawa couplings) and Mathematics (SYZ conjecture)

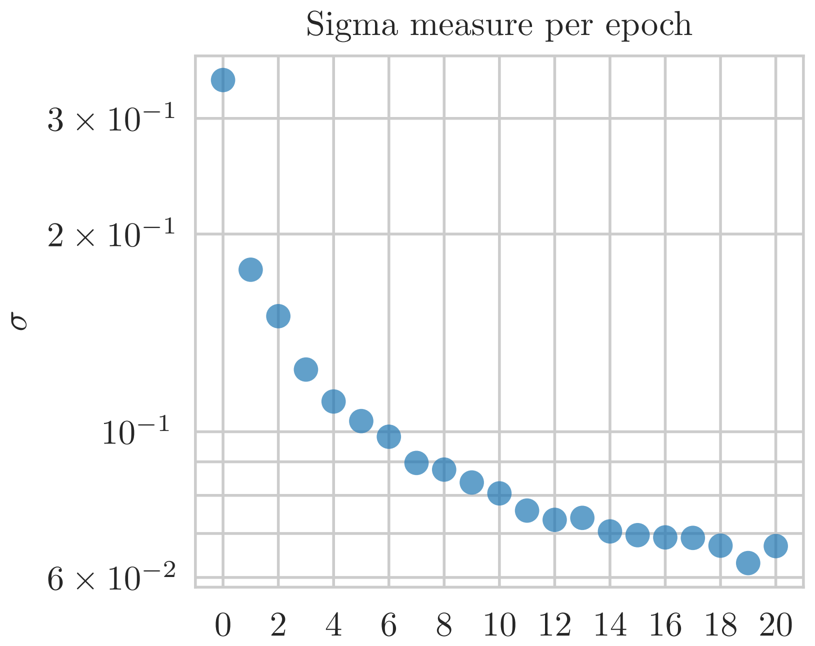


- The methods generalize to SU(n)/G_2 structure



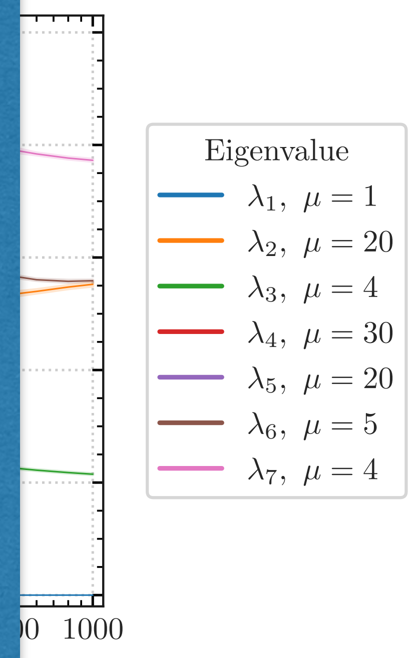
Conclusions

- NNs can learn the moduli- (and position-) dependent CY metric by optimizing the underlying differential equations



- This h
string
coupl

Thank you for your attention!



- The methods generalize to SU(n)/G_2 structure