Information Geometry and AdS/CFT

University of Cape Town - Department of Mathematics and Applied Mathematics,
Quantum Gravity and Strings Laboratory

2rd July 2015, IHEP Beijing - Tslil Clingman, Emanuel Malek, Jeff Murugan, JS

Dr Jonathan Shock | The Laboratory for Quantum Gravity & Strings | Department of Mathematics & Applied Mathematics | University of Cape Town | South Africa |
jon.shock@gmail.com | tel: +27-21-650-3644 |fax: +27-21-650-2334

Outline

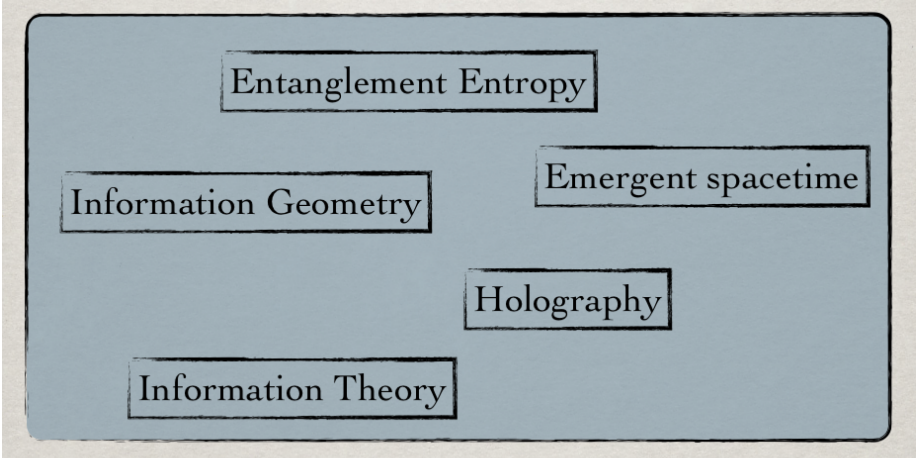
Some references

- Methods of information geometry - Shun'ichi Amari and Hiroshi Nagaoka
- Information geometry for neural networks - Daniel Wagenaar
- Information geometry blog entries - John Baez
- Statistical Inference, Occam's Razor, and Statistical Mechanics on the Space of Probability Distributions - Vijay Balasubramanian
- Relative Entropy and Proximity of Quantum Field Theories - Balasubramanian et al.
- Statistical Inference and String Theory - Jonathan Heckman
- Multi-Instanton Calculus and the AdS/CFT Correspondence in N=4 Superconformal Field Theory - Dorey et al.

Motivation

- AdS/CFT: We would like to be able to derive the dual gravitational description of a given strongly coupled field theory.
- There are hints that information theory might point us in the right direction
- There are also hints that instantons know something about the dual geometry to a given QFT.
- Can we combine these to make a generative AdS/CFT algorithm?

Information theory and holography



Caveats

- The following is very much speculative and we have some way to go, but together with colleagues in Cape Town, I believe that this is a direction worth investigating. At the very least it is the confluence of a number of very interesting subjects!

A digression

- To build up to information geometry let's remind ourselves a little about entropic ideas.
- The first few slides will be familiar to anyone at my talk last year, but it quickly diverges...

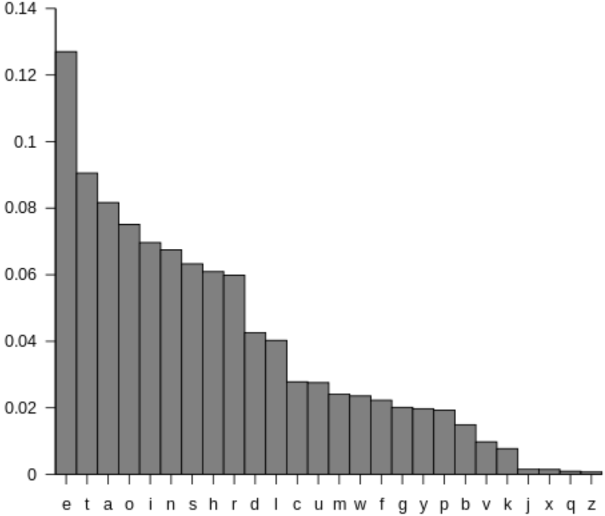
Shannon Entropy

- The information we learn from an event:
  - $I_i = -\log P_i$
- We learn more from an unlikely event than from a likely one.
- The entropy in a system is the probabilistically weighted sum of all possible information:

$$S_{Shannon} = -\sum_i P_i \log P_i \tag{1}$$

eg. calculate the entropy inherent in a language in terms of the average amount of information you'll get when looking at random letters in a book.

The entropy of English

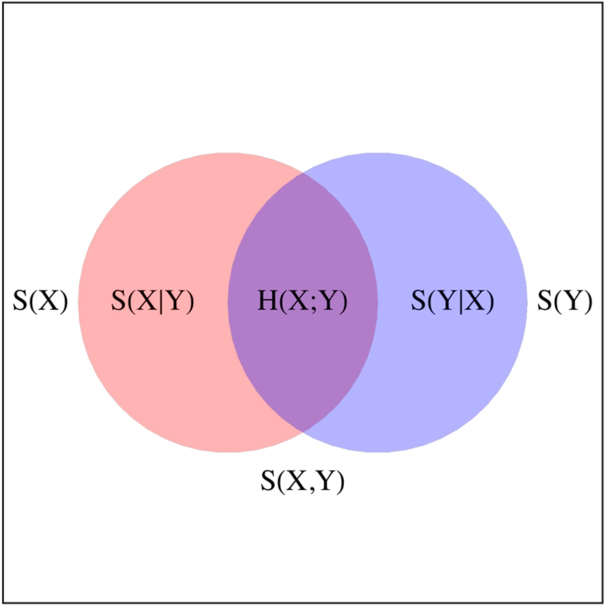


$S_{English} \sim 4.1 bits/symbol$

The entropy of a thermodynamic system

- Given partial knowledge (ie. the macroscopic thermodynamic variables), we can say how much information we would get from learning about the precise microstate and can measure the thermodynamic entropy positing that the microstates are ergodic (a reasonable assumption under the guise of statistical mechanics).
- Entropy is related to the energy in a system which cannot be used to do work (This can be taken in both a thermodynamic sense and a purely information theoretic sense).

Mutual Information



Mutual Information

- When two systems are correlated, how much can we learn from one system by making a measurement in the other?
- Joint Entropy:

$$S(X,Y) = -\sum_{i,j} P(x_i,y_j)\log P(x_i,y_j) \tag{2}$$

- Conditional Entropy:

$$S(Y|X) = -\sum_{i,j} P(x_i,y_j)\log P(y_j|x_i) \tag{3}$$

- Mutual Information:

$$H(X;Y) = S(Y) - S(Y|X) \tag{4}$$

This is roughly how much information is shared

A measure of distance in statistical ensembles

- How about the difference between the joint entropy and the mutual information.

$$\begin{aligned} D(X,Y) &= S(X,Y) - H(X;Y) = S(X|Y) + S(Y|X) \\ &= -\sum_{x\in X, y\in Y} P(x,y)\log(P(x|y)P(y|x)) \end{aligned} \tag{5}$$

  Defines a distance between the two sources (satisfies all metric requirements).
  Distance between two distributions over the same source:
- The Kullback-Leibler pseudo distance (or relative entropy)

$$D[p(x)||q(x)] = \left\langle \log\frac{p(x)}{q(x)} \right\rangle_p \tag{6}$$

- The information lost when q is used to approximate p.

The Fisher Metric

- How about measuring an infinitesimal distance using the KL distance?

$$D_{KL}(P(x;\zeta)||P(x;\zeta+\Delta\zeta) \sim \frac{1}{2}\sum_i\sum_j g_{ij}(\zeta)\Delta\zeta^i\Delta\zeta^j \tag{7}$$

where:

$$g_{ij} = \sum_x P(x;\zeta)\frac{\partial}{\partial\zeta^i}\log P(x,\zeta)\frac{\partial}{\partial\zeta^j}\log P(x,\zeta) \tag{8}$$

or the continuous version:

$$g_{ij} = \int dx P(x;\zeta)\frac{\partial}{\partial\zeta^i}\log P(x,\zeta)\frac{\partial}{\partial\zeta^j}\log P(x,\zeta) \tag{9}$$

This of course can be written as the second derivative of the expectation value $\langle\partial_i \log P \partial_j \log P\rangle$.

- This is the Fisher Metric.

A simple example

How about looking at the Fisher metric for the Gaussian distribution? Can we guess what it might look like from the beginning?

$$P(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{10}$$

The Fisher Metric tells us about how difficult it is to tell apart two distributions. Clearly the absolute value of μ won't make a difference, but the absolute value of σ will. For a given μ two distributions which have large σ will be harder to tell apart than two which have small σ.

A simple example

Let's look at one entry in the metric explicitly:

$$
\begin{aligned}
g_{\mu\mu} &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\partial_\mu \left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right)^2 \\
&= \left\langle \frac{(x-\mu)^2}{\sigma^2} \right\rangle = \frac{1}{\sigma^2}
\end{aligned}
\tag{11}
$$

Cross-terms vanish, and $g_{\sigma\sigma}$ is very similar:

$$
ds^2_{Gaussian} = \frac{d\sigma^2/2 + d\mu^2}{\sigma^2} \tag{12}
$$

The space is hyperbolic: Gaussians in some way are related to Euclidean AdS...not such a big deal...

Reversing the procedure

- Can we start with a metric and figure out what the corresponding probability distribution is?
- It turns out that most of the information about the distribution is lost when we calculate the expectation value.
- There is an infinite degeneracy when going in the other direction.
- In fact one can write a multi-dimensional PDF:

$$P(X;\theta) = f(x^j - h^j) \tag{13}$$

and then:

$$g_{ab} = \partial_a h^j \partial_b h^k \delta_{jk} \tag{14}$$

- The fisher metric can be written as a pullback

The PDF of a spherical metric

More accurately A PDF of a spherical metric:

$$P(x, y, z; \theta, \phi) = \frac{1}{(2\pi)^{-\frac{3}{2}}} e^{-\frac{1}{2}\left((x-\cos\theta\sin\phi)^2+(y-\sin\theta\sin\phi)^2+(z-\cos\phi)^2\right)} \quad (15)$$

- Take a quantum harmonic oscillator
- $$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x) e^{-x^2/2}$$

  where $H_n(x)$ are the Hermite polynomials.
- The general sum of the first n eigenfunctions gives the metric on $\mathbb{S}^n$.

Constructing Probability Density Functions for Arbitrary Metric Tensors via the Fisher Information Metric - T Clingman, J Murugan, JS - arXiv:1504.03184.

So this is information geometry

- Defining a natural geometry on the spaces of probability distributions.
- So?
- This allows for better machine learning algorithms which can take geodesics in the paths of parameter space as defined on the information manifold.
- Information and geometry are somehow intimately linked in the AdS/CFT correspondence - could this be a clue?

What has this to do with AdS/CFT

- The Geometry and Topology of moduli spaces - Hitchin
- Studied instantons of $SU(2)$ YM on R^4.
- Self-dual, finite energy, euclidean solutions, corresponding in the Lorentzian picture to tunnelling solutions between vacua.
- Look at single instanton solutions (labeled by a topological charge which corresponds to the winding of the global part of the gauge symmetry as $r \to \infty$)

$$F_A = \frac{\lambda^2 dx \wedge d\bar{x}}{(\lambda^2 + |x-a|^2)^2} \tag{16}$$

  where x is a quaternion valued variable: $x = x_0 + ix_1 + jx_2 + kx_3$.
- This is a 5-parameter solution.
- classically this theory is conformal (of course that is broken QMically).

The moduli space of SU(2)

- The moduli space is given by:

$$\mathcal{M} = \left\{ (\lambda^2, a) \in \mathcal{R}^5 | \lambda^2 > 0 \right\} \tag{17}$$

- generally one defines an L^2 metric on this space.
- The metric is conformally flat - not conformally invariant.
- The L^2 metric is not complete - there are certain singularities in the metric related to concentrated instantons

An alternative metric?

- Can we define a metric which has the same symmetries as the underlying gauge theory and is everywhere smooth?
- We can define a family of probability distributions:

$$\rho(X,\theta) \sim F_A^2(X,\theta) \sim \mathcal{L} \tag{18}$$

  With appropriate normalisation. This is essentially the instanton density. $(\theta = \{\lambda^2, a\})$
- Now we can define a metric:

$$g_{AB} \sim \int \sqrt{g}\, d^4x F_A^2(X,\theta)\partial_A \log F_A^2(X,\theta)\partial_B \log F_A^2(X,\theta) \tag{19}$$

so what is it for SU(2)?

- It turns out that for $SU(2)$ instantons on R^4, the metric defined from the information theoretic point of view is Euclidean AdS_5. This can be shown very simply from symmetry arguments, or from a relatively simple explicit calculation.
- Symmetry is powerful and so perhaps this isn't greatly surprising.
- Is there anything more subtle going on? $hep - th/0108122$ - Blau, Narain, Thompson...

Some more powerful statements

- Blau et al:
- Perturb the space-time on which the instantons are defined:

$$\delta_h tr(F_A \wedge F_A) \tag{20}$$

  where $\delta_h g_{\mu\nu} = h_{\mu\nu}$ is the variation of the boundary metric. This variation gives rise to a variation of the bulk metric.
- That leads precisely to the equation for the AdS_5 boundary to bulk graviton propagator.
- This means that hidden in the instanton moduli space in R^4 for $SU(2)$ instantons is dynamical gravity in AdS_5.
- In essence, one can show that the variation of the boundary metric and the instantons thereon leads to the Einstein Equations for the bulk:
- $R_{AB} = -4G_{AB}$.

Some more powerful statements

- Moreover, for the perturbed metric, the perturbed instanton density is the boundary to bulk massless scalar propagator:

$$\Box F^2 = 0 \tag{21}$$

And at finite temperature?

- We can look at caloron solutions (instantons at finite temperature): http://arxiv.org/pdf/hep-th/0507082.pdf Soo-Jong Rey and Yasuaki Hikida.
- Only numerical solutions exist for the Fisher metric.
- There is a horizon, but there are some rather strange numerical artefacts which are still not cleared up.

So is this AdS/CFT?

- Not yet, by any means. This was for SU(2) Yang Mills and it shows that somehow Einstein gravity on AdS_5 is a natural measure on the moduli space of instantons. There also seems to be a correlation function $\leftrightarrow$ bulk propagator correspondence going on.
- This is all at the linearised level - going beyond linear is tricky.
- It's not clear how the higher form fields would come about
- How about the compact part of the space?
- What happens with supersymmetry
- Are there any hints of something special happening at large N?

Well, maybe SU(2) is relevant...

The large N saddle point approximation of $SU(N)$ Yang Mills has a moduli space which is that of $SU(2)$.
Integrating out the gauge degrees of freedom leaves us with $SU(2)$.

Instantons in holography

- $hep - th/9901128$ - Multi-Instanton Calculus and the AdS/CFT Correspondence in N=4 Superconformal Field Theory - Dorey et al
- Instanton contributions to fermionic correlators in large N SYM can be mapped to $D(-1)$ instantons in $AdS_5 \times S^5$. The instantons see the space.
- How about multi-instantons? The space collapses to a single copy of $AdS_5 \times S^5$.
- This is found by going to next order in N and seeing that the degeneracy in $(AdS_5 \times S^5)^k$ is lifted to a single copy of the space.
- Each multi-instanton contribution sees the same space

Let's start somewhere simpler

- Can we understand where the internal space comes from?
- In fact Dorey et al already see the compact space but we want to show why the geometry is dynamical.
- Start with a simpler model: $\mathcal{C}P^N$.

$$S = \int d^2x D_\mu \phi^\dagger \cdot D_\mu \phi \tag{22}$$

where

$$D_\mu \phi = \partial_\mu \phi - A_\mu \phi \,, \tag{23}$$

is the $U(1)$ gauge covariant derivative and the (auxiliary) gauge field is given by

$$A_\mu = \phi^\dagger \cdot \partial_\mu \phi \,. \tag{24}$$

instanton solutions

Introducing complex coordinates

$$x_\pm = x_1 \pm i x_2 \tag{25}$$

(anti-)instanton solutions satisfy

$$D_\mp \phi = 0\,. \tag{26}$$

Thus we can express any instanton solutions in terms of a rational holomorphic function $f(x_+)$

$$\phi(x) = \frac{f(x_+)}{|f(x_+)|} \tag{27}$$

Anti-instantons are obtained by using rational anti-holomorphic functions. The instanton charge is given by the degree of the zero of $f(x_+)$.

The most general $k = 1$ instanton is thus given by

$$\phi = e^{i\theta} \frac{(x_+ - a)u + \lambda v}{\sqrt{|x_+ - a|^2 + \lambda^2}}, \tag{28}$$

where a labels the position of the instanton, λ its size and u, v satisfy

$$u^\dagger \cdot u = v^\dagger \cdot v = 1, \qquad u^\dagger \cdot v = 0, \tag{29}$$

and give the orientations of the instanton in the $(N+1)$-dimensional complex vector space. This instanton solution has $2N + 1$ moduli corresponding to the two positions a, the size λ, and the $2N - 2$ possible orientations of v relative to u. We will now set $\theta = 0$ and $u = (1, 0, \ldots, 0)$ without loss of generality.

the moduli space

- If we use the Lagrangian density, as the probability density, then we cannot get any information about the internal global symmetries. The space will simply be AdS_3. Can we define a vector valued probability density?
- Conjecture a current.
- We use the coset formulation of the CP^N model: J^μ takes values in

$$\frac{U(N+1)}{U(N) \times U(1)} \tag{30}$$

The technicalities

- Write out the generators of $U(N+1)$.
- Find those elements which form representations of $U(N)$ and $U(1)$ and define the orthogonal generators to this space.
- Exponentiate to find the group elements g.
- Define the current $J_\mu = g^\dagger dg|_m$ where the projection is onto the space orthogonal to $U(N) \times U(1)$.
- Equate an appropriate group element with the original instanton solution and match parameters.
- write the coset current evaluated on the instanton solution
- Calculate:

$$\int dx_1 \, dx_2 \partial_{(i} J^{\mu\dagger}_\alpha \partial_{b)} J^\alpha_\mu \tag{31}$$

The result

- The spatial moduli are the same as when we used the lagrangian, so we still get AdS_3. However, we now have moduli related to the global internal directions.
- We can thus show that for CP^N the information metric is given by:

$$AdS_3 \times S^{2N-1} \tag{32}$$

and SUSY

- Using a super current as the probability density we can get a compact space. In fact from $\mathcal{N} = 4$ $SU(2)$, we should be able to get $AdS_5 \times S^5$.
- The big question is whether the perturbation of the boundary leads to Einstein gravity in all ten dimensions. That is still not known.

Why do we care?

- The hope is that if we can find the instantons of non-scale free theories we may be able to derive their dual holographic geometries using the Fisher metric.
- Clearly we must also figure out how to get the other fields in an information theoretic manner...
- And see if we must always use a coset model formulation to derive the metric.

Conclusions

- Information geometry links information theory and geometric structures
- Instanton contributions to correlation functions seem to map out the geometry felt by $D(-1)$ instantons in a specific geometry
- The Fisher Metric seems to be the natural metric on the moduli space to encode the holographic geometry
- The geometry is, at least to leading order Einstein gravity.
- How can we see the F_5?
- What can we understand about deformations?
- What constraints can we put on the correspondence using information geometry?

Thank You!