Title: Lessons for quantum gravity from quantum information theory

Speakers: Daniel Harlow

Collection: Quantum Gravity 2020

Date: July 13, 2020 - 1:15 PM

URL: http://pirsa.org/20070003

Abstract: Gravity is unique among the other forces in that within general relativity we are able to do calculations which, when properly interpreted, give us information about non-perturbative quantum gravity. A classic example is Bekenstein and Hawking's calculation of the entropy of a black hole, and a more recent example is the calculation of the ``Page curve" for certain evaporating black holes. A common feature of both of these calculations is that they compute entropies without using von Neumann's formula S=-Tr(\rho \log \rho). In this strange situation where we are able to compute entropies without understanding the details of the states for which they are the entropy, quantum information theory is a powerful tool that lets us extract information about those states. In this talk I'll review aspects of these developments, emphasizing in particular the role of quantum extremal surfaces and quantum error correction.

Lessons for quantum gravity from quantum information

Daniel Harlow

MIT

July 13, 2020

Why is quantum gravity hard?

There are various reasons which are sometimes given:

- **Non-renormalizability?** No: e.g. the pion Lagrangian is also non-renormalizable but has long since been UV-completed into QCD.
- **Diffeomorphism invariance?** No: e.g. the particle worldline theory and Jackiw-Teitelboim gravity are both diffeomorphism-invariant, but can be quantized in the usual way without difficulty.
- **Black holes?** Yes! The real difficulties seems to arise only in theories of quantum gravity which are rich enough to have black holes with non-trivial microstates. One way to think about this is that black holes create an obstruction to formulating quantum gravity in terms of local fields: they prevent the usual approach to testing short-distance physics by scattering, and postulating a local field description seems to lead to "too many" black hole degrees of freedom.

Our understanding of all the other forces is based on local fields, so if we give them up then we need to create some new kind of mathematical structure. How might we find it?

- **Experiment:** This would be the best, but so far it seems difficult. I have some hopes for cosmology, for example one thing we do know is that whatever theory of quantum gravity we invent it needs to be able to accomodate something like inflation and something like $\Lambda > 0$ today.
- **Toy Models:** Consider simpler theories with unrealistic properties such as $d \neq 4$, $\Lambda < 0$, unbroken SUSY, but which still capture some aspects (e.g. the existence of black holes) of real quantum gravity. Examples: string theory, AdS/CFT.
- **Semiclassical techniques:** Use perturbative quantum gravity to guess at general principles which might suggest the way forward. Great if you are smart, but what if you get the wrong principles?

The remarkable progress of the last decade has come primarily from using tools from quantum information theory to relate the second two strategies.

A new formula for entropy

More concretely, the recent progress has in large part come from the development of a new formula for von Neumann entropy in gravitational systems: Ryu/Takayanagi, Hubeny/Rangamani/Takayanagi, Faulkner/Lewkowycz/Maldacena, Dong, Engelhardt/Wall

$$S(R) = \min_{X_R} \left[\operatorname*{ext}_{X_R} \left(\frac{\mathrm{Area}(X_R) + \ldots}{4G} + S_{bulk}(r(X_R)) \right) \right].$$

We'll call this the *quantum extremal surface formula*. Among its many interesting consequences are the following:

- **Entanglement wedge reconstruction:** explains how in AdS/CFT the dual CFT can (at least sometimes) see into the interiors of black holes.
- **No global symmetries:** entanglement wedge reconstruction can be used to exclude bulk global symmetries in AdS/CFT.
- **Page curve:** the QES formula can account for the unitarity of black hole evaporation.

In the rest of this talk, I'll review the QES formula and a bit about these applications.

Euclidean gravity magic

In quantum gravity we are very lucky to have access to an oracle: the Euclidean gravity path integral.
In ordinary quantum systems the Euclidean path integral is fairly boring, in the sense that we obtain it directly from the operator formalism via formulas like

$$Z(\beta) \equiv \mathrm{Tr}(e^{-\beta H}) = \int \mathcal{D}\phi|_{\mathrm{periodic}} e^{-S_E(\phi)}.$$

In quantum gravity however we do not have a non-perturbative picture of the Hilbert space in terms of the metric and matter fields, so the equality here is not clear.
The amazing thing however is that if we *assume* that there is some Hilbert space for which this equality holds, then by evaluating the path integral we are able to learn a remarkable amount of things about that Hilbert space!

The canonical example of this is the Euclidean Schwarzschild geometry

$$ds^2 = \frac{r - r_s}{r} d\tau^2 + \frac{r}{r - r_s} dr^2 + r^2 d\Omega^2,$$

whose Euclidean action leads immediately to the Bekenstein-Hawking formula

Gibbons/Hawking, Hawking/Page

$$S = \frac{A}{4G}.$$

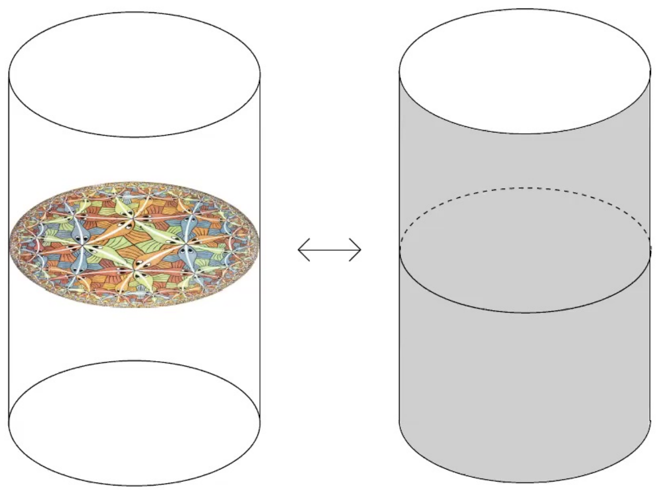
Somehow the Euclidean path integral knows the number of microstates of a black hole, even though those microstates cannot be described in terms of the semiclassical Hilbert space of general relativity!

The reason that this is not an obvious contradiction is that in this geometry the τ circle contracts to zero size at $r = r_s$, so it wouldn't be a valid contribution to the partition function if we tried to interpret it as a trace in terms of the metric variables.

Nonetheless, in every situation where we can actually test this formula (e.g. SUSY black holes in string/M theory, AdS/CFT, SYK model), it works!

AdS/CFT

The Bekenstein-Hawking entropy is typically thought of as coarse-grained entropy, which is an inherently vague notion, but at least in AdS/CFT we can also think of it (at sufficiently high temperature) as the fine-grained entropy of the thermal state $\rho = \frac{1}{Z(\beta)} e^{-\beta H}$.



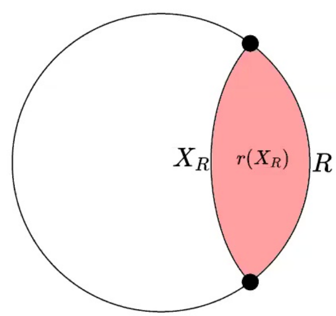
What AdS/CFT buys us is a non-perturbatively precise picture of the Hilbert space and dynamics of quantum gravity in asymptotically-AdS space: $|\psi\rangle_{bulk} \leftrightarrow |\psi\rangle_{CFT}$.

The Ryu-Takayanagi formula

Fourteen years ago, Ryu and Takayanagi made a remarkable conjecture: that for any CFT spatial subregion R and any "reasonable" state ρ on the CFT Hilbert space with a moment of time-reflection symmetry, the fine-grained von Neumann entropy of the reduced state ρ_R can be computed on the bulk side of the correspondence via

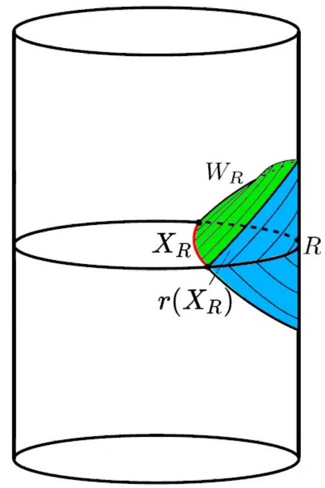
$$S(\rho_R) \equiv -\mathrm{Tr}(\rho_R \log \rho_R) = \min_{X_R} \frac{\mathrm{Area}(X_R)}{4G} + O(G^0),$$

where the minimum over X_R is taken over codimension-two surfaces which are anchored to R in the sense that $\partial X_R = \partial R$, homologous to R in the sense that there is a spatial surface $r(X_R)$ such that $\partial r(X_R) = X_R \cup R$, and X_R lies within the hypersurface which is invariant under the time-reflection symmetry.



The quantum extremal surface formula

The RT formula was soon generalized to remove the assumption of time-reflection symmetry by Hubeny, Rangamani, and Takayanagi, and then Faulkner/Lewkowycz/Maldacena and Engelhardt/Wall proposed a further generalization to extend it to higher orders in G:



$$S(\rho_R) = \min_{X_R}\left[\operatorname{ext}_{X_R}\left(\frac{\mathrm{Area}(X_R)+\dots}{4G} + S_{bulk}(r(X_R))\right)\right].$$

The bulk domain of dependence of $r(X_R)$, W_R, will be important in what follows and is called the **entanglement wedge of** R.

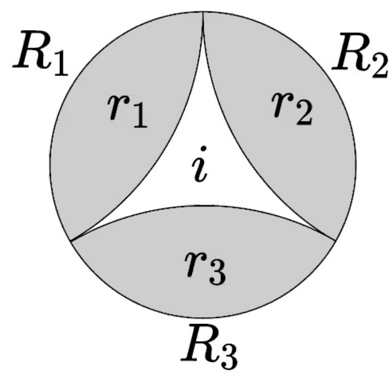
The QES formula has been checked in a number of situations, but it still may seem surprising that such a remarkable formula is true. In fact there are two independent ways of understanding it:

- **Euclidean path integral:** It turns out that the Gibbons-Hawking calculation of black hole entropy can be generalized to justify the QES formula. Lewkowycz/Maldacena, Faulkner/Lewkowycz/Maldacena, Dong/Lewkowycz/Rangamani, Dong/Lewkowycz Another example of Euclidean gravity magic!
- **Quantum error correction:** The embedding of bulk effective field theory into AdS/CFT can be thought of a quantum error-correcting code Almheiri/Dong/Harlow. These are protocols which were invented to protect quantum computers against decoherence. It can then be shown that *any* quantum error correcting code must obey a version of the QES formula! Pastawski/Yoshida/Harlow/Preskill, Hayden/Nezami/Qi/Thomas/Walter/Yang, Harlow

I'll spend the rest of the talk reviewing a few of the things we can learn from the QES formula.

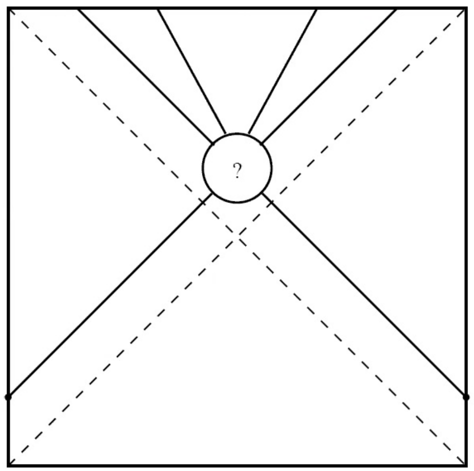
Subregion duality and quantum error correction

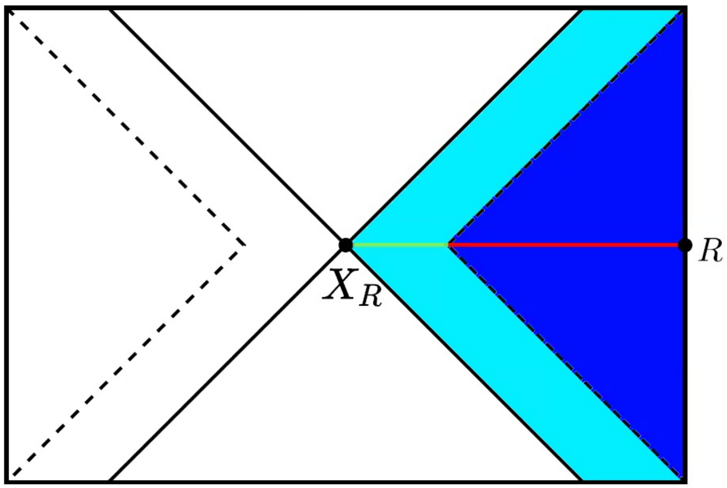
The QES formula has important consequences for **subregion duality**, which was a conjecture that the CFT observables in a boundary spatial subregion R should have complete information about some bulk subregion W_R. At first it was not clear which subregion W_R should be, but using the QES formula we were able to show that all observables within the entanglement wedge W_R are accessible to a CFT observer who can only measure operators in R. Dong/Harlow/Wall



This statement is called **entanglement wedge reconstruction**, and it implies a remarkable redundancy in how the bulk is encoded into the boundary: it has the features of a **quantum error correcting code**. Almheiri/Dong/Harlow The detailed mathematical connection between the QES formula and subregion duality proceeds using the mathematical techniques developed for studying such codes.

Perhaps the most interesting consequence of entanglement wedge reconstruction is that it implies that, at least in some cases, the boundary CFT degrees of freedom know what is going on behind black hole horizons:





We'll return to this soon.

No global symmetries in quantum gravity

It is an old piece of lore that in quantum gravity there can be no global symmetries. For continuous global symmetries there is a heuristic argument for this based on the Bekenstein-Hawking entropy formula, and all known compactifications of string theory have neither discrete nor continuous global symmetries. Recently Hirosi Ooguri and I were able to use entanglement wedge reconstruction to argue that indeed there are no bulk global symmetries in AdS/CFT. Harlow/Ooguri

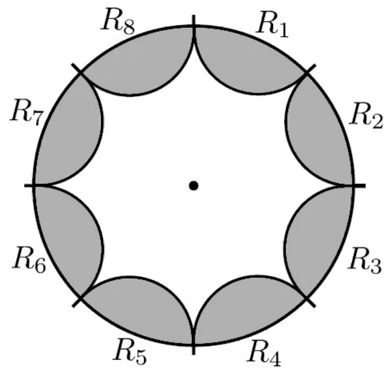
Indeed recall that in quantum field theory a global symmetry is one which acts faithfully on the set of local operators at each point x:

$$U^\dagger(g) O_i(x) U(g) = \sum_j D_{ij}(g) O_j(x).$$

The essential feature of global symmetry is that if you want to know how much charge there is in a region of space, you have to go visit each point and see:

$$Q[R] = \int_R J^0.$$

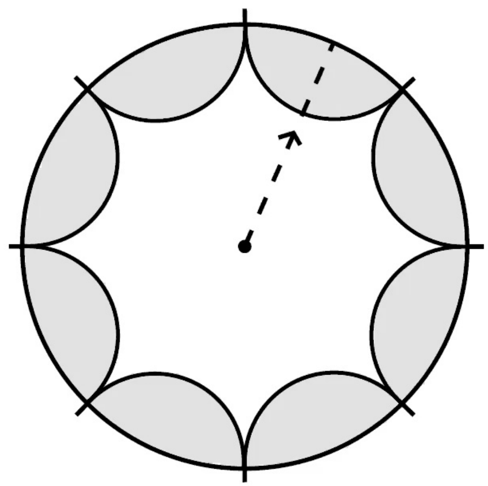
The idea of our proof is that since a global symmetry in the bulk implies a global symmetry in the boundary, we can break up the total boundary charge into a sum of charges for disjoint regions $R_1, R_2, \ldots$:



$$Q = Q[R_1] + Q[R_2] + \ldots.$$

By assumption there should be a bulk object which is charged under our putative bulk global symmetry, so let's put it in the center of the space.
Each of the $Q[R_i]$ can only detect charge in the entanglement wedge of R_i, but none of these contain the operator! Therefore it couldn't have been charged in the first place; a contradiction.

This contradiction is avoided for gauge symmetries, since there are no charged local operators and a charged line always reaches into one of the R_i:

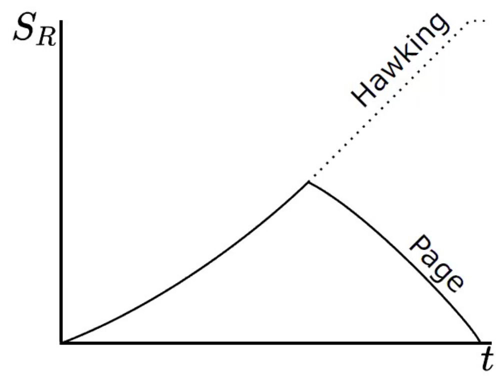


Indeed all known examples of AdS/CFT have gauge symmetries in the bulk, so it is good that we didn't rule them out!

The black hole information problem

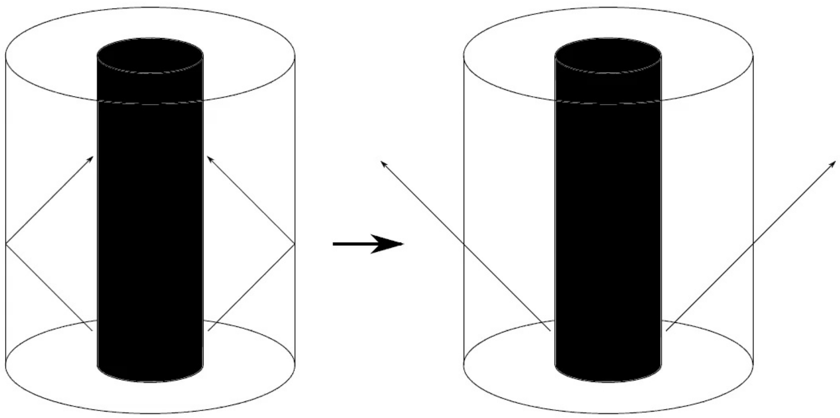
- The black hole information problem is one of the great contradictions of theoretical physics: like the infrared catastrophe of classical thermodynamics or the inconsistency of Maxwell's equations without the displacement current, it tells us that our current understanding of the world is incomplete.
- Since Hawking's seminal work in 1975, we have gone through long stretches of seemingly little progress towards resolving it, interrupted by occasional periods where major developments take place.
- I believe we are currently seeing one of those exciting periods, which was initiated last year with two remarkable papers Penington, Almheiri/Engelhardt/Marolf/Maxfield, and which has continued throughout the year in a sequence of interesting results from various groups.

The basic idea of the black hole information problem is that Hawking's calculation of the rate of particle production by black holes also tells us the quantum state of those particles: they are (basically) in a mixed thermal state. As time proceeds there are more and more of them, so the von Neumann entropy of the radiation cloud grows with time. Eventually the radiation is all that is left, so a pure state has evolved to a mixed state and quantum mechanics has been violated.



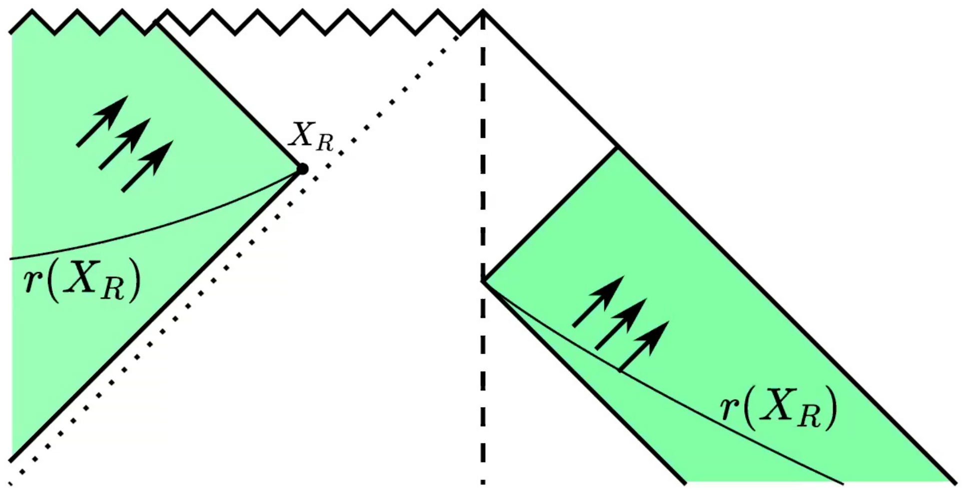
This can be represented graphically using the "Page curve", which shows the von Neumann entropy of the radiation as a function of time.
Reliably showing that the Page curve bends downwards is one of the key requirements for solving the information problem.

Since AdS/CFT is a non-perturbative theory of quantum gravity, at some level it is clear that we should be able to compute the Page curve of an evaporating black hole with in AdS space.
Unfortunately the black holes which we understand the best in AdS/CFT are those whose size is at least as big as the AdS curvature radius, and these black holes do not evaporate since their Hawking radiation is reflected off of the boundary and re-absorbed.



The first key idea of Penington, Almheiri/Engelhardt/Marolf/Maxfield was to extend the system by allowing bulk matter fields other than the metric to propagate "out" of the AdS boundary in a computable way, allowing even big black holes to evaporate.

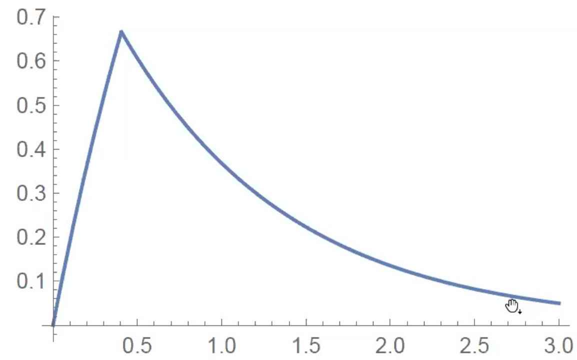
The second key idea was to *not* compute the entropy of the radiation directly, but to instead compute it using the quantum extremal surface formula. Moreover they realized that, despite working perturbatively in G, it is possible to have a quantum extremal surface which is not near any classical extremal surface! This can happen because the short distance singularity in S_{bulk} can be balanced against the $1/G$ in the area term. The new quantum extremal surface for an evaporating black hole looks like this:



$$S(\rho_R) = \min_{X_R}\left[\underset{X_R}{\text{ext}}\left(\frac{\text{Area}(X_R) + \dots}{4G} + S_{bulk}(r(X_R))\right)\right],$$

This surface gives an entropy which is approximately just the horizon area, so it decreases as the black hole evaporates.

In a few simple models this calculation can be done in complete detail, leading to explicit formulas for the Page curve! Here is an example:



The increasing part of the curve arises from a situation where there is no “island” contribution, while the decreasing part arises once the island configuration should in the previous figure becomes dominant.

Conclusion

We've learned a lot already, but there are many things still to consider:

- What is the experience of an infalling observer in a typical black hole microstate?
- What level of "approximate" global symmetry is allowed in quantum gravity?
- How can we move beyond AdS/CFT to more realistic cosmologies and the big bang?

Thanks!