Black Holes, Holography, and Quantum Information

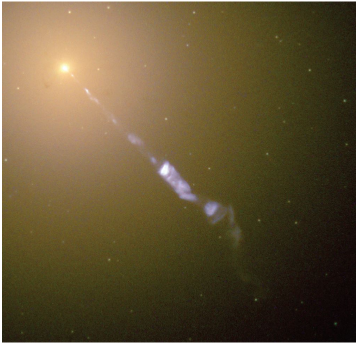
Daniel Harlow

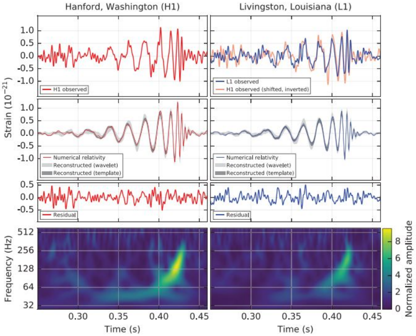
Massachusetts Institute of Technology

August 31, 2017

Black Holes

Black holes are the most extreme objects we see in nature!





Classically we understand them quite well, but quantum mechanically they are a major obstacle to formulating a satisfactory theory of gravity.

This tension is often expressed by way of the *black hole information paradox*.

- Quantum field theory in curve spacetime predicts that black holes formed from collapse radiate, in a manner which is uncorrelated with how they were formed. They must therefore evaporate, which seems to lead to a violation of quantum unitarity. Hawking
- Five years ago this paradox was rephrased in a striking new way: the *firewall paradox* of Almheiri, Polchinski, Marolf, and Sully. AMPS

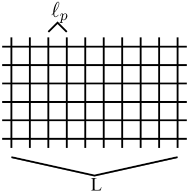
The AMPS paradox is quite subtle, with many implicit assumptions, and unfortunately I cannot yet resolve it for you here.

Thinking about it however has led to many exciting developments in the last few years: entanglement, the emergence of spacetime, quantum error correction, chaos, etc.

I can best illustrate some of this by formulating the tension between quantum mechanics and gravity in a simpler way, which has its own related paradoxes. These I *will* be able to resolve in a way that is hopefully illuminating!

The basic problem black holes introduce for quantum gravity is that they prevent us from defining the localized observables that are essential to quantum field theory.

Say we want to use a network of rods to define the locations of points at distances of order ℓ_p:



Not a black hole:

$$L \gg r_S \sim GM = \ell_p \left(\frac{L}{\ell_p}\right)^3 \frac{m_{rod}}{m_p} \implies m_{rod} \ll m_p$$

Rods localized:

$$\ell_p \gg \Delta x > \frac{1}{m_{rod}\Delta v} \gg \frac{1}{m_{rod}} \implies m_{rod} \gg m_p.$$

Black Hole Thermodynamics

But if gravity is not going to be a local theory, what is it going to be? Fortunately for us, general relativity and quantum field theory gives us some guidance on what might need to change:

- **Area Theorem:** The dynamics of GR imply that horizon area always increases, reminiscent of second law of thermodynamics.
- **Hawking Radiation:** Quantum field theory in curved spacetime tell us that black holes radiate at a temperature consistent with the idea that
$$S = \frac{A}{4G}.$$

This suggests that in quantum gravity we should think of the number of degrees of freedom in a spatial region as being subextensive!

Holography

Inspired by the Bekenstein-Hawking formula, 't Hooft and Susskind proposed the *holographic principle*:

- A theory of quantum gravity in d spatial dimensions should really be formulated as a local theory in a lower number of dimensions.

This may seem crazy, but in fact we now have two explicit examples of this:

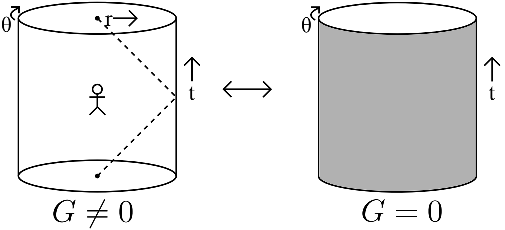
- The BFSS model uses matrix quantum mechanics to describe 11 dimensional flat space. Banks/Fischler/Shenker/Susskind
- The AdS/CFT correspondence uses conformal field theory in d dimensions to describe gravity in (at least) $d+1$ dimensions. Maldacena

How can this be true!?

In the past few years we have understood much better where this extra dimension comes from in AdS/CFT, I'll give a sketch of this in the rest of the talk. I'll return to black holes at the end.

AdS/CFT

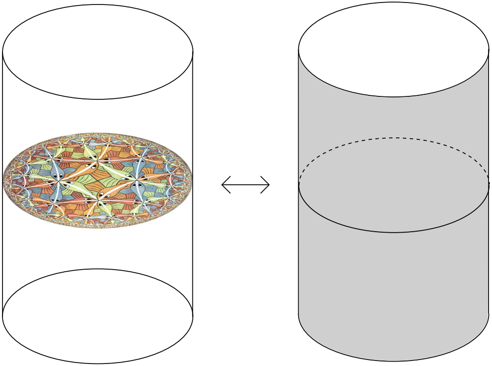
AdS/CFT says that quantum gravity in asymptotically AdS space is equivalent to conformal field theory on the boundary:



$$ds^2 = -\left(1+\frac{r^2}{L^2}\right)dt^2 + \frac{dr^2}{1+\frac{r^2}{L^2}} + r^2 d\Omega^2$$

$$\to r^2\left(-\frac{dt^2}{L^2} + d\Omega^2\right)$$

This correspondence is a quantum correspondence:



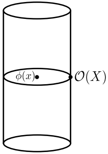
- $|\psi_{bulk}\rangle \longleftrightarrow |\psi_{boundary}\rangle$
- $H, J, \dots \longleftrightarrow H, J, \dots$
- $\lim_{r\to\infty} r^{\Delta}\phi(r,t,\Omega) \longleftrightarrow \mathcal{O}(t,\Omega)$.
- Vacuum perturbations $\longleftrightarrow$ low-energy states
- Big black holes $\longleftrightarrow$ high-energy states

First Puzzle

In quantum field theory, causality is enforced by locality:

$$[\mathcal{O}(X), \mathcal{O}(Y)] = 0 \qquad (X - Y)^2 > 0.$$

We can consider this in the bulk as well:

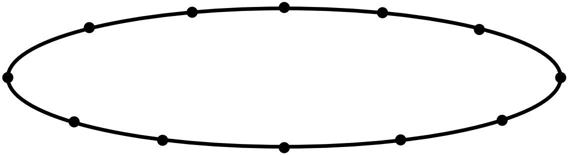


x and X are spacelike-separated in the bulk, so we expect that

$$[\phi(x), \mathcal{O}(X)] = 0.$$

But actually this is impossible in quantum field theory!

The problem is that in a quantum field theory, any operator that commutes with all local operators at a fixed time must be trivial. For example consider a chain of Pauli spins:



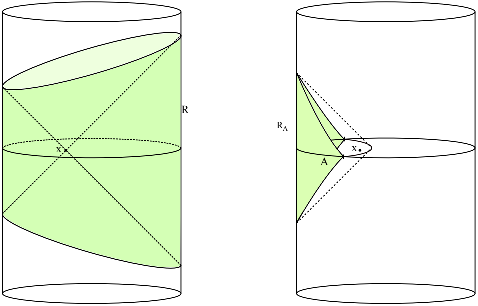
The set of products of Pauli operators, eg

$$Z_1 X_4 Y_7 \dots,$$

gives a basis for all operators, so an operator which commutes with all of the individual Pauli operators must be proportional to the identity. This is a basic expression of the local structure of the Hilbert space in a QFT. But then how can we get an extra dimension to emerge?

Subregion Duality

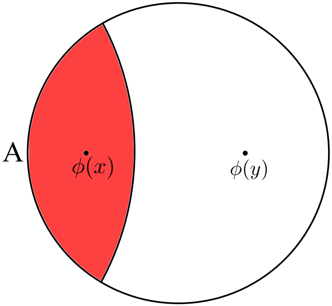
To describe the second puzzle, we need to develop the dictionary between the two sides a bit more:



Bulk operators that are not near the boundary can also be represented as operators in the CFT, via formulas like

$$\phi(x) = \int_R dX\ K(x;X)\mathcal{O}(X) + \ldots .$$

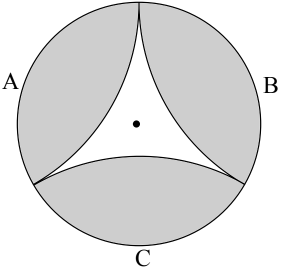
This is perhaps more intuitive if we look from above at a single time-slice:



The operator $\phi(x)$ can be represented using only the "Pauli operators" in A, but the operator $\phi(y)$ cannot.

Second Puzzle

Now say we split the boundary into three regions:



The operator in the center has no representation on A, B, or C, but it does have a representation either on AB, AC, or BC!
Where is the information?

Quantum Error Correction

- It was understood in the last two years that these puzzles, and the correspondence more generally, can be understood by re-interpreting AdS/CFT as a *Quantum Error Correcting Code*. Almheiri/Dong/Harlow, Harlow/Pastawski/Preskill/Yoshida (HaPPY!), Hayden/Nezami/Qi/Thomas/Walter/Yang
- Quantum error correcting codes were invented to protect quantum computers from decoherence, the basic idea is to protect a quantum message by encoding it nonlocally in the entanglement between many degrees of freedom. Shor, Gottesman
- There is a beautiful general theory of quantum error correction we could study, but we will instead focus on a simple example to illustrate how it works.

Three qutrits

Say that I want to send you a "single qutrit" state:

$$|\psi\rangle = \sum_{i=0}^{2} C_i |i\rangle.$$

If I just send it to you, it might get lost or corrupted. So the idea is to instead send you three qutrits in the state

$$|\widetilde{\psi}\rangle = \sum_{i=0}^{2} C_i |\widetilde{i}\rangle,$$

where $|\widetilde{i}\rangle$ is a basis for a special subspace of the full 27-dimensional Hilbert space, which is called the *code subspace*.

Explicitly, we take Cleve/Gottesman/Lo

$$
\begin{aligned}
|\widetilde{0}\rangle &= \frac{1}{\sqrt{3}}\left(|000\rangle + |111\rangle + |222\rangle\right) \\
|\widetilde{1}\rangle &= \frac{1}{\sqrt{3}}\left(|012\rangle + |120\rangle + |201\rangle\right) \\
|\widetilde{2}\rangle &= \frac{1}{\sqrt{3}}\left(|021\rangle + |102\rangle + |210\rangle\right).
\end{aligned}
$$

- Note that this subspace is symmetric between the three qutrits, and each state is highly entangled.
- This entanglement leads to the interesting property that in any state in the subspace, the density matrix on any one of the qutrits is maximally mixed, ie is given by $\frac{1}{3}\left(|0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|\right)$.
- In other words, any single qutrit has *no information* about the encoded state $|\widetilde{\psi}\rangle$.
- This leads to the remarkable fact that we can *completely recover* the quantum state from any two of the qutrits!

To see this explicitly, we can define a two-qutrit unitary operation U_{12} that acts as

$$\begin{array}{lll} |00\rangle \to |00\rangle & |11\rangle \to |01\rangle & |22\rangle \to |02\rangle \\ |01\rangle \to |12\rangle & |12\rangle \to |10\rangle & |20\rangle \to |11\rangle \\ |02\rangle \to |21\rangle & |10\rangle \to |22\rangle & |21\rangle \to |20\rangle \end{array} .$$

- It is easy to see then that we have

$$U_{12}|\widetilde{i}\rangle = |i\rangle_1|\chi\rangle_{23},$$

with $|\chi\rangle \equiv \frac{1}{\sqrt{3}}(|00\rangle + |11\rangle + |22\rangle)$.

- This then gives us

$$U_{12}|\widetilde{\psi}\rangle = |\psi\rangle_1 \otimes |\chi\rangle_{23},$$

so we can recover the state!

- By symmetry there must also exist U_{13} and U_{23}.

This is reminiscent of our "ABC" example of the operator in the center, but there we talked about operators instead of states. We can easily remedy this.
Say we have a single-qutrit operator O

$$O|i\rangle = \sum_j (O)_{ji}|j\rangle.$$

We can always find a three-qutrit operator $\widetilde{O}$ that implements this operator on the code subspace:

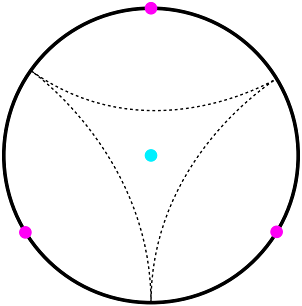
$$\widetilde{O}|\widetilde{i}\rangle = \sum_j (O)_{ji}|\widetilde{j}\rangle.$$

Generically this operator will have nontrival support on all three qutrits, but using our U_{12} we can define

$$O_{12} \equiv U_{12}^{\dagger} O_1 U_{12},$$

which acts nontrivially only on the first two but still implements O on the code subspace.

Thus by using the entanglement in the code subspace, we can reproduce the puzzling behavior of subregion duality that we saw earlier:



- Three “physical” qutrits are local CFT degrees of freedom on the boundary
- One “logical” qutrit is a field in the center of the bulk
- The bulk point is in the subregion wedge of any two of the boundary points, so its information is well-protected

We can also make contact with the commutator puzzle:

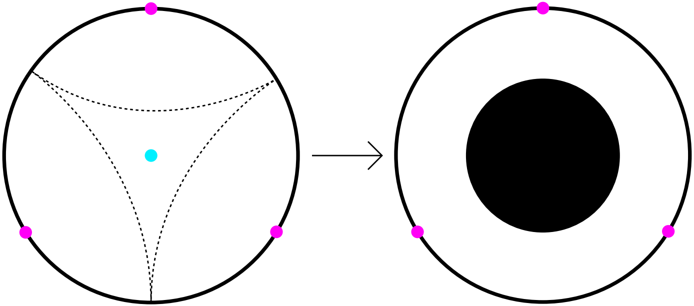
- Consider

$$\langle\widetilde{\psi}|[\widetilde{O}, X_3]|\widetilde{\phi}\rangle,$$

  where X_3 is some operator on the third qutrit and $|\widetilde{\phi}\rangle$, $|\widetilde{\psi}\rangle$ are arbitrary states in the code subspace.
- Since $\widetilde{O}$ always acts either to the left on a state in the code subspace, we can replace it by O_{12}. But then the commutator is zero! This would have worked for X_1 or X_2 as well, so we see that on the code subspace $\widetilde{O}$ commutes with all "local" operators.
- It is because we are working only in the code subspace that we are able to circumvent the algebraic puzzle we discussed before.

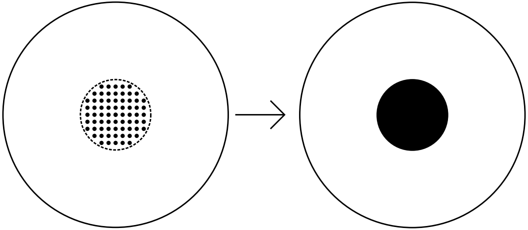
But what about the rest of the states? There is a a 24-dimensional subspace orthogonal to the code subspace, what about bulk locality in those states?

This is where gravity comes to the rescue: these states are the microstates of a black hole that has swallowed our bulk point!

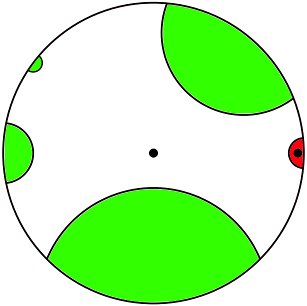


The point is far enough behind the horizon that we no longer need to account for it.

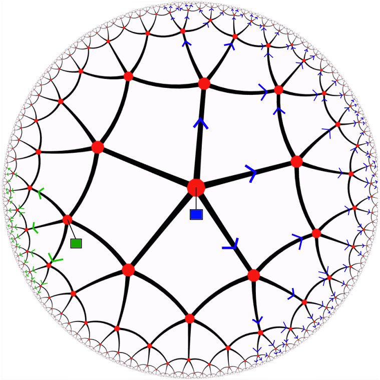
- We thus come back to our initial discussion: any attempt to probe the locality of spacetime too precisely *does* break down due to black hole formation! The full theory of quantum gravity, given by the CFT in this case, simply has no notion of locality beyond this point.
- This may seem a bit arbitrary, since what is a black hole in this model anyway? But actually this conclusion can be generalized to the full AdS/CFT correspondence, where we can see from a general theorem in quantum error correction that the breakdown of the error correcting properties of the code parametrically happens in the same place we expect black hole formation in the bulk.

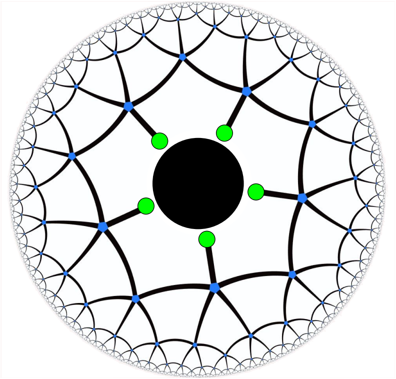


- One can pursue this much further, using the general theory of error correcting codes and the physics of the bulk to learn more about what kind of error correcting code AdS/CFT realizes. But I don't have time to discuss that today, so let me just mention two things.
- In general, we can understand the emergence of the radial direction as a measure of how well information is protected:



One can construct generalizations of the three-qutrit code that have a volume's worth of bulk degrees of freedom, and illustrate these lessons in an exactly-soluble setting using "tensor networks": HaPPY





We are learning more all the time, and I am excited to see what else will come!

Thanks!