Black Hole Information Paradox

Delon Shen

Notes for Suvrat Raju's Black Hole Information Paradox course at ICTP(well really online) during Spring 2021. Course website can be found here which contains notes, assignments, and links to lecture videos. This course closely follow a review posted here. If you have any comments let me know at hi@delonshen.com.

Lecture 1: Introduction and Two-Point QFT Correlators

January 13,2021

The main organization of this course

(a) Hawking's Original Paradox → Thermalization and exponentially small corrections.

(b) Paradoxes about interior of evaporating Black Holes → holography of information, islands and page curve.

(c) Paradoxes about large Black Holes in AdS/CFT→ Mirror operators, state-dependence, and firewalls/furrballs

Lets start by talking about **Hawking Radiation**, it's the effect that underlies the information paradox. Take a black hole in asymptotically flat space. This black hole radiates with a temperature $\propto$ surface gravity. We should also recall that hawking radiation relies on short distance QFT physics and on global late-time properties of the black hole geometry. The interesting thing is that the derivation for Hawking's radiation also implies the existence of the entangled modes across horizons. So what are the common derivations of hawking radiation(TODO (a) is in appendix of review paper and (b) might be in wald)?

(a) Hawking's original derivation

(b) Rindler ↔ Minkowski Bogolivlov transformation

In this course we'll consider a different derivation from both of these

Lets take a second to step back from black hole and look at Quantum Fields near a null surface. We'll apply what we learn here to black holes later. What we want to show is that across any null surface in a smooth state (TODO smooth state who?) we can isolate a "local" QFT (which we'll define in a bit) with universal entanglement. This is useful because we'll find that in a black hole spacetime local degrees of freedom near the horizon gives global modes in blackhole geometry.

First lets define what we mean by a smooth metric around some point. Consider a point in some $D = d + 1$ space and let this point be the origin. We have U, V, two null coordinates, and $d - 1$ transverse coordinates. A metric is smooth around some point if around some point we can locally choose some coordinates so the metric takes the following form. (think light cone variant Kruskal coordinates in arbitrary dimensions?)

$$ds^2 = -dUdV + \delta_{\alpha\beta} dy^\alpha dy^\beta + \dots$$

Where $dUdV$ are two null coordinates and α, β is over $d - 1$ indices and where the $\dots$ terms vanish near origin.

$$figure$$

We also want to make an additional demand. Consider a scalar field ϕ and points near $U = 0$. If we're still thinking in terms of Kruskal coordinates this means we're thinking of things close to eachother on each side of the horizon? In the limit where x_1 approaches x_2 for any nonsingular state the two point correlation function (Wightman function?) becomes.

$$\langle \phi(x_1)\phi(x_2) \rangle = \frac{N}{|x_1 - x_2|^{d-1}} + \dots$$

We also impose the following scales

(a) $|x_1 - x_2| \ll \ell_{\text{curvature}}$

(b) $|x_1 - x_2| \ll \frac{1}{m}$

(c) $|x_1 - x_2| \gg \ell_{\text{Pl}}$ or any UV scale where EFT breaks down.

These length scales give us the normalization if we consider a free field (e.g. $\mathcal{L} = 1/2(\partial_\mu \phi)^2$).

$$N = \frac{\Gamma(d-1)}{2^d \pi^{d/2}\Gamma(d/2)} \Rightarrow \langle \phi(x_1)\phi(x_2)\rangle = \frac{\Gamma(d-1)}{2^d \pi^{d/2}\Gamma(d/2)} \frac{1}{|x_1 - x_2|^{d-1}} + \dots$$

Because of the length scales we assume we can say that the structure of the two point function is universal (TODO what in the world.) Before we continue lets look at a few things that will be useful

$$|x_1 - x_2|^2 = -\delta U \delta V + \delta_{\alpha\beta}\delta y^\alpha \delta y^\beta \qquad \delta O = O_1 - O_2$$

Also if we grind through some calculations we'll find that

$$\langle \partial_{U_1}\phi(x_1)\partial_{U_2}\phi(x_2)\rangle = -\frac{d^2-1}{4}\frac{N(\delta V)^2}{|x_1 - x_2|^{d+3}} + \dots$$

Taking $\delta V \to 0$, e.g. we take δV to be the smallest separation, then we find that

$$\lim_{\delta V \to 0} \frac{(\delta V)^2}{(-\delta U \delta V + \delta y^\alpha \delta y^\beta \delta_{\alpha\beta})^{(d+3)/2}} \neq 0$$

It's not zero since it does receive a contribution when $y^\alpha = 0$. To see this we can do an integral over all the transverse separations.

$$\int \frac{(\delta V)^2}{(\delta U \delta V + \delta y^\alpha \delta y^\beta \delta_{\alpha\beta})^{(d+3)/2}} d^{d-1}\delta y^\alpha$$

We'll also take the $|x_1 - x_2|$ is positive. Now with the substitution $\delta\tilde{y}^\alpha = \delta y^\alpha / \sqrt{-\delta U \delta V}$ we get

$$\frac{1}{(\delta V)^2}\int \frac{d\delta\tilde{y}^\alpha}{[1 + \delta\tilde{y}^\alpha \delta\tilde{y}^\alpha]^{(d-3)/2}}$$

What happens in the end is that all factors of δV cancel. In the notes Suvrat says that for $\delta y^\alpha \neq 0$ the integral vanishes. I think this might be due to symmetry, there's a ring of δy^α with appropriate sign that cancels out when we do the integral. But since $\delta y^\alpha = 0$ doesn't have this cancellation it remains finite. In the end we get

$$\boxed{\lim_{\delta V \to 0} \langle \partial_{U_1}\phi(x_1)\partial_{U_2}\phi(x_2)\rangle = -\frac{1}{4\pi}\frac{\delta^{d-1}(\delta y^\alpha)}{(U_1 - U_2 - i\epsilon)^2}}$$

Something to note: the reason we did an integral was to pick up the coefficient of the delta function. How do we sniff out the presence of a delta function. We use the property $\int \delta(x)dx = 1$ with the fact that the integral vanishes for $x \neq 0$. Thus if the integral we considered above gave a finite answer then we know the two point correlation function of the derivatives of the states would be proportional to a delta function.

The next step which will happen in the next lecture would to define modes as approximately

$$\int \partial_\mu \phi(-U)^{i\omega}$$

What is this doing? It's picking up the right moving modes with constant V. (TODO huh?)

Two-point function normalization

First we'll consider a free field governed by the lagrangian in $D = 3+1$ flat space time.

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{m^2}{2}\phi^2$$

Lecture 2: Entangled Modes accross Null Surfaces

January 14,2021

Last time we mentioned that we can extract right moving modes accross the null surface $U = 0$ with an integral (this is an approximate expression, we'll get a more precise integral in a bit)

$$\int \partial_U \phi (-U)^{i\omega} dU \qquad \int \partial_U \phi U^{i\omega} dU$$

Note that we can't integrate over a large region in U because this would violate our limit of x_1 approaching x_2 which we derived the form of the two point function for in the last lecture. So then we know that we should integrate over a small region of U instead(with respect to the length scales defined in the last lecture).

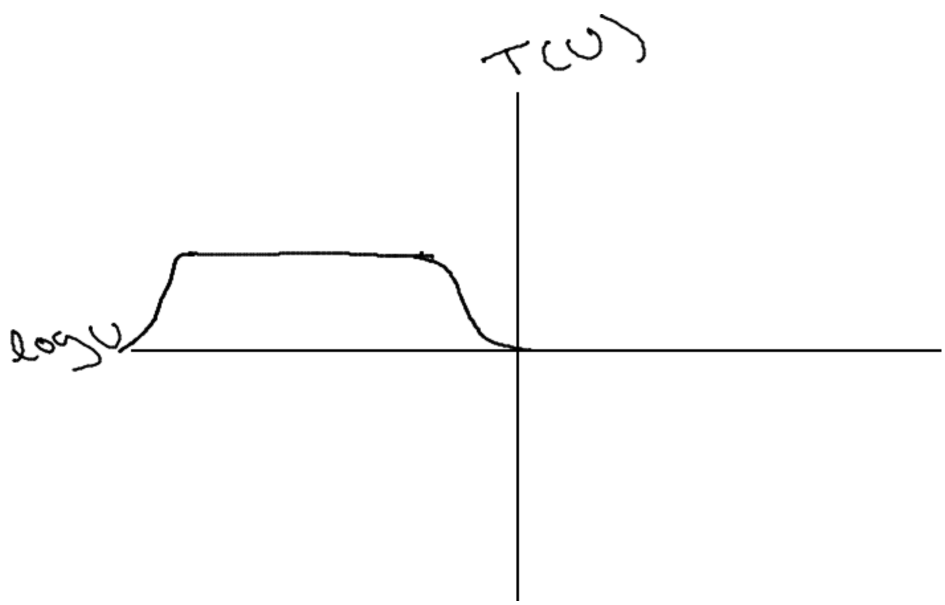
We'll start by introducing a smearing function (TODO what?) $T(U)$ with the following properties

(a) $T(U)$ dies off smoothly near $U \to 0$

(b) Support in interval $[U_l, U_r]$ where $\ell_{\mathrm{UV}} \ll U_r, U_l \ll \ell_{\mathrm{curv}}$ and $\frac{U_r}{U_\ell} \gg 1$ and $U_l, U_r > 0$

(c) We normalize the semaring function $\int T(U)^2 dU/U = 2\pi$. Note that $dU/U = d\log U$.

(d) $T(U)$ is flat for a large range of $\log U$.



The next thing we need to do is define what's happening in the transervse direction by integrating over a volume Vol in the transverse direction which is smaller than a cube of the curvature scale. From this we can write a more precise expression for the mode.

$$a_{\omega_0} = \int (\partial_U \phi(U, V = 0, y^\alpha))(-U)^{-i\omega_0} T(-U) dU \frac{d^{d-1}y^\alpha}{\sqrt{\pi\omega_0 \mathrm{Vol}}}$$

We can similarly define a similar integral for the other side of the null surface. Note we let $V = -\epsilon$ to ensure that we're considering points that are spacelike separated.

$$\tilde{a}_{\omega_0} = \int (\partial_U \phi(U, V = -\epsilon, y^\alpha))(U)^{i\omega_0} T(U) dU \frac{d^{d-1}y^\alpha}{\sqrt{\pi\omega_0 \mathrm{Vol}}}$$

The $T(U)$ insures we only integrate over a small region in U and we have some normalization factors put in that aren't motivated from what I can tell.

Let's now compute the two point function of a and $\tilde{a}$ and we will find that this will only depend on the short distance field correlator that we found last lecture. First spelling things out

$$\langle a\tilde{a}\rangle = \frac{1}{\pi \mathrm{Vol}\omega_0} \int dU_1 dU_2 \langle \partial_{U_1}\phi(U_1, V = 0, y_1)\partial_{U_2}\phi(U_2, V = -\epsilon, y_2)\rangle \times$$
$$\times (-U_1)^{-i\omega_0} U_2^{i\omega_0} T(-U_1)T(U_2) d^{d-1}y_1 d^{d-1}y_2$$

Remember that the correlator gives a delta function

$$= -\frac{1}{4\pi^2\omega_0} \int \frac{1}{(U_1 - U_2)^2} \left(\frac{U_2}{-U_1}\right)^{i\omega_0} T(-U_1)T(U_2) dU_1 dU_2$$

To do this integral we need the identity

$$\frac{1}{U_1 - U_2}^2 = \frac{1}{(-U_1)U_2} \int_{-\infty}^{\infty} \frac{\omega e^{-\pi\omega}}{1 - e^{-2\pi\omega}} (U_2/(-U_1))^{-i\omega} d\omega$$

When $U_1 < 0$ and $U_2 > 0$. Assume that $|U_1| > |U_2|$. This lets do a contour integral where there are poles at $\omega = in$ where $n \in \mathbb{N}$. From here we can sum the residuals. What are the residuals at the poles? Well mathematica can tell us and so can Suvrat.

$$\frac{1}{(U_1 - U_2)^2} = \frac{1}{|U_1|U_2} \sum_{n=1} -n(-1)^n (U_2/|U_1|)^n$$

We can also plug in the identity (before computing the residuals) into $\langle a\tilde{a}\rangle$ to get

$$\langle a\tilde{a}\rangle = \frac{1}{4\pi^2\omega_0} \int \frac{dU_1}{U_1}\frac{dU_2}{U_2} (U_2/(-U_1))^{-i(\omega-\omega_0)} \frac{\omega e^{-\pi\omega}}{1 - e^{-2\pi\omega}} T(-U_1)T(U_2) d\omega$$
$$= \int T(-U_1)(1/(-U_1))^{i(\omega_0-\omega)} dU_1/U_1 \times \int T(U_2) U_2^{i(\omega_0-\omega)} dU_2/U_2 \times \int \frac{\omega e^{-\pi\omega}}{1 - e^{-2\pi\omega}} d\omega$$

Now note that we can rewrite this in terms of $\log(U)$ since $dU/U = d\log U$.

$$= \int T(-U_1) e^{-i(\log[-U_1])(\omega_0-\omega)} d\log U_1 \times \int T(U_2) e^{i(\log U_2)(\omega_0-\omega)} d\log U_2 \times \int \frac{\omega e^{-\pi\omega}}{1 - e^{-2\pi\omega}} d\omega$$

The first two integrals are fourier transforms of T. Namely $S(\gamma) = \frac{1}{2\pi}\int_0^\infty T(U) U^{-i\gamma} dU/U$

$$= \frac{1}{\omega_0} \int \frac{\omega e^{-\pi\omega}}{1 - e^{-2\pi\omega}} |S(\omega - \omega_0)|^2 d\omega$$

Now note that since we said T is very flat for a large range of U then we know that the fourier trasnform of T has to be very big at $T = 0$ and thus the fourier transform of T becomes basically a delta function. This gives us finally

$$\boxed{\langle a\tilde{a}\rangle = \frac{e^{-\pi\omega_0}}{1-e^{-2\pi\omega_0}} + \dots} \tag{1}$$

There are similar calculations we can do to find

$$\langle aa^\dagger\rangle = \frac{1}{1-e^{-2\pi\omega_0}} \quad \langle\tilde{a}\tilde{a}^\dagger\rangle = \frac{1}{1-e^{-2\pi\omega_0}} \quad [a,\tilde{a}] = 0 \quad [a,a^\dagger] = [\tilde{a},\tilde{a}^\dagger] = 1 \quad \langle a^\dagger\tilde{a}\rangle = \langle a^\dagger a^\dagger\rangle = 0 \tag{2}$$

Something to note is that in some quantum field theories $\langle i_\omega \tilde{i}_{\omega'}\rangle = e^{-\pi\omega}1 - e^{-2\pi\omega}\delta(\omega-\omega')$. However in this case that is not true. Here we have (TODO how?) $\langle\tilde{a}_{\omega_0}c_{\omega_0'}\rangle \approx 0$ where $\omega_0 \neq \omega_0'$.

In the special case where we have a spacetime with spherical symmetry

$$ds^2 = -dUdV + r_0^2 d\Omega_{d-1}^2 + \dots$$

In this we can derive a analgous form of the modes where

$$a = \frac{r_0^{d-1}}{\sqrt{\pi\omega_0}}\int \partial_U\phi(U,V=0,\Omega)(-U)^{-i\omega_0}T(-U)dUY_l^*(\Omega)d\Omega \qquad \tilde{a} = \dots$$

Where Y_l are our spherical harmonic functions. These satisfy all the same correlator and commutation relation as we found before.

Now moving to a different topic. We've been writing $\langle\dots\rangle$ for correlators but we need to describe what state we're calculating these correlators for. Say we are in a state $|\psi\rangle$. What we want to show is that $\tilde{a}|\psi\rangle \propto a^\dagger|\psi\rangle$. Thinking about this geometrically this means that $\tilde{a}$ and $a^\dagger$ are parallel to eachother. To prove this consider the decomposition of $\tilde{a}|\psi\rangle$ (TODO: is that actually a complete set?)

$$\tilde{a}|\psi\rangle = c_1 a|\psi\rangle + c_2 a^\dagger|\psi\rangle + |\chi\rangle$$

Where $|\chi\rangle$ is orthogonal to $a|\psi\rangle$ and $a^\dagger|\psi\rangle$. From here we can use the correlators we have in (2) to get $c_1 = 0$. Similarly we can use (1) to get

$$\frac{e^{-\pi\omega_0}}{1-e^{-2\pi\omega_0}} = \langle\psi|a\tilde{a}|\psi\rangle = c_2\langle\psi|aa^\dagger|\psi\rangle + \langle\psi|a|\chi\rangle = \frac{c_2}{1-e^{-2\pi\omega_0}} \Rightarrow c_2 = e^{-\pi\omega_0}$$

Finally we can find $|\chi\rangle$ through the following. From (2) we have $[\tilde{a},\tilde{a}^\dagger] = 1$. This means

$$\langle\tilde{a}\tilde{a}^\dagger\rangle - 1 = \langle\tilde{a}^\dagger\tilde{a}\rangle = \frac{e^{-2\pi\omega_0}}{1-e^{-2\pi\omega_0}}$$

Now with this we have (after $|\dots|^2$ both sides)

$$\frac{e^{-2\pi\omega_0}}{1-e^{-2\pi\omega_0}} = |c_2|^2\langle\psi|aa^\dagger|\psi\rangle + \langle\chi|\chi\rangle + 0 \Rightarrow \langle\chi|\chi\rangle = 0 \Rightarrow |\chi\rangle = 0$$

Where the last term vanishes because we assume χ is orthogonal to $a^\dagger|\psi\rangle$ and $a|\psi\rangle$. After all this we get

$$\boxed{\tilde{a}|\psi\rangle = e^{-\pi\omega_0}a^\dagger|\psi\rangle} \tag{3}$$

Similarly we can also show that

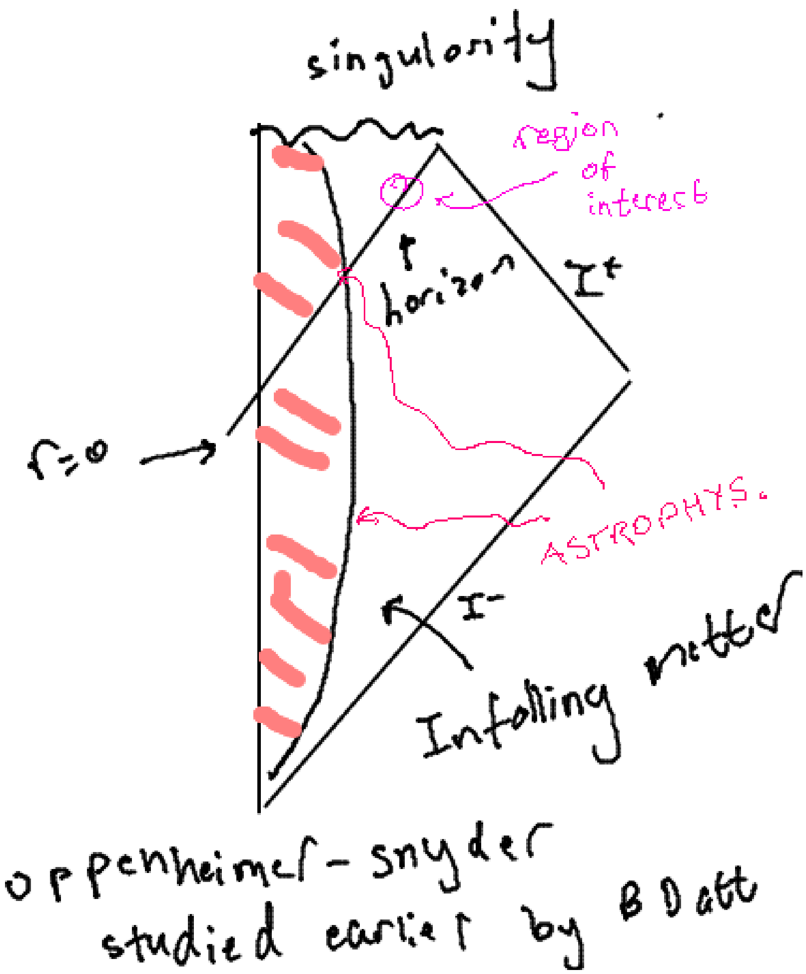
$$\boxed{\tilde{a}^\dagger|\psi\rangle = e^{\pi\omega_0}a|\psi\rangle} \tag{4}$$

Next lecture we'll apply these results to black holes

Lecture 3: Quantum Fields in a Black Hole Background

January 20,2021

Lets start by review Black Holes in flat space.



In the late time limit the metric becomes very simple

$$ds^2 \to -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_{d-1}^2 \qquad f(r) = 1 - \frac{\mu}{r^{d-2}}$$

Where μ is the mass parameter is related to the mass by

$$\mu = 8\pi^{(2-d)/2}\Gamma(d/2)GM_{\text{real}}/(d-1)$$

The horiizon is when $f(r_h) = 0$ and thus $r_h = \mu^{d-2}$. Lets talk about what we mean by as $t \to \infty$. We mean that $t \gg r_h$ after collapse but $t \ll t_{\text{evap}}$. This is because the collapsing black hole is not valid for when the black hole is evaporating. Also note that the region of interest hast the property that there is a lot of time in that region. It's also useful to go to Tortoise coordinates in order to examine propagating fields

$$dr_* = \frac{dr}{f(r)} \qquad \text{Near } r \to \infty,\ f(r) \to 1 \Rightarrow r_* \to \infty \qquad r \to r_h,\ f(r) \to 2k(r - r_h) \Rightarrow r_* \to \frac{1}{2k}\log[(r-rh)\underline{2k}]$$

Where $k = f'(r_h)/2 =$ surface gravity. The underlined term is a choice of constant. So now we have

$$ds^2 = f(r)[-dt^2 + dr_*^2] + r(r_*)^2 d\Omega^2$$

Something else we should note is that the horizon is not as special as we think. Lets go to Kruskal coordinates

$$U = -\frac{1}{k}e^{k(r_*-t)} \Rightarrow dU = (dt - dr_*)e^{k(r_*-t)} \qquad V = \frac{1}{k}e^{k(r_*+t)} \Rightarrow dV = (dr_* + dt)e^{k(r_*+t)}$$

$U < 0$ outside the horizon. This means we have $dUdV = (dr_*^2 - dt^2)e^{2kr_*}$. However near the horizon we found that the exponential becomes $2k(r - r_h)$. The metric in Kruskal coordinates becomes

$$ds^2 \to -dUdV + r^2 d\Omega_{d-1}^2 \text{ near } r \to r_h$$

Horizon is at $U = 0$ while V remains finite so $t \propto \log(V/U) \to \infty$. The coordinates are basically flat near the horizon. Behind the horizon U becomes positive. For $r < r_h$ we find that $f(r)$ changes negative so t is a spacelike coordinate and r_* is a time coordinate.

We're done reviewing classical black holes so now lets consider the propagation of fields. Consider the field that is minimally coupled

$$\left(\frac{1}{\sqrt{-g}}\partial_\mu g^{\mu\nu}\sqrt{-g}\partial_\nu - m^2\right)\phi = 0$$

In tortoise coordinates $\sqrt{-g} = f(r)r^{d-1}$(sphericalcontribution) and $g^{**} = -g^{tt} = 1/f(r)$. The wave equation becomes

$$\frac{1}{f(r)r^{d-1}}\partial_* r^{d-1}\partial_*\phi = \frac{1}{f(r)}\partial_t^2\phi + \frac{1}{r^2}\Box_\Omega\phi - m^2\phi$$

We can solve the above by noting near the horizion $f(r) \to 0$ so the equation becomes

$$\frac{1}{f(r)}(\partial_*^2\phi - \partial_t^2\phi) = 0$$

This is independent of the angular part, mass, and additional interactions. We can then write

$$\phi \to \int d\omega e^{-i\omega t}\left[A_\omega(\Omega)e^{-i\omega r_*} + B\omega(\Omega)e^{i\omega r_*}\right] + \text{hermitian conjugate}$$

This however isn't the most convenient thing we could do. First let $Y_\ell(\Omega)$ as a spherical harmonics where ℓ is a collective symbol for all the angular quantum numbers. We can choose spherical harmonics as our basis of solutions and have (where we choose f_{in} and f_{out}. These are solutions which we choose as our basis)

(a) $f_{\text{in}}(\omega, \ell, r_*)e^{-i\omega t}Y_l(\Omega)$ where as $r \to r_h$ $f_{\text{in}} \to h_{\omega,\ell}e^{-i\omega r_*}$

(b) $f_{\text{out}}(\omega, \ell, r_*)e^{-i\omega t}Y_l(\Omega)$ where as $r \to r_h$ $f_{\text{in}} \to e^{i\omega r_*} + g_{\omega,\ell}e^{-i\omega r_*}$.

f_{in} has the property that as t increase r_* must decrease. On the other hand f_{out} has the property that as t increase r_* must increase. Both of these happen to keep phase constant. These solutions above are chosen so that the they're orthogonal in the Klein-Gordon norm. This isn't enough to fix those however. We also choose $g_{\omega,\ell}$ so that as $r_* \to \infty$ we have $f_{\text{out}} \to b_{\omega,\ell}r^{(1-d)/2}e^{i\omega r_*}$.

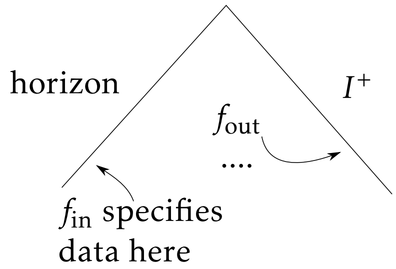


Figure 1: Description of how information in specified in Penrose diagram

f_{in} as $r \to r_h$ looks like $e^{-i\omega(r_*+t)}$. Also remember that $r_* \to -\infty$ at horizon but $t \to \infty$ so $(r_* + t)$ remains finite. The point of this whole song and dance is for

$$\phi = \sum_\ell \int d\omega \left[A_{\omega,\ell} f^{\text{out}}(\omega,\ell,r_*) + B_{\omega,\ell} f^{\text{in}}(\omega,\ell,r_*)\right] e^{-i\omega\bar{t}} Y_\ell(\Omega) + \text{hermitian conjugate}$$

A, B are not normalized but are anhilation operators and hermitian conjugate are the creation operator. What happens when we cross the horizion. Behind the horizion we can write a similar expansion

$$\phi = \sum_\ell \int d\omega \left[\tilde{A}_{\omega,\ell} e^{i\omega t} Y_\ell^*(\Omega) + C_{\omega,\ell} e^{-i\omega t} Y_\ell(\Omega)\right] \tilde{f}^{\text{out}}_{\omega,\ell}(r_*) + \text{hermitian conjugate}$$

Where as $r \to r_h$ from inside we have $\tilde{f}^{\text{oout}}_{\omega,\ell}(r_*) \to e^{-i\omega r_*}$. Note that we do not go deep inside the horizon. This is because our expansion starts breaking down if we go too far in. By continuity the $C_{\omega,\ell}$ modes have an expansion

$$C_{\omega,\ell} = A_{\omega,\ell} h_{\omega,\ell} + B_{\omega,\ell} g_{\omega,\ell}$$

Also we know that $\tilde{A}_{\omega,\ell}$ are new modes that aren't related to modes outside of the horizon

Lecture 4: Hawking Radiation

January 21,2021

Lets first recap. We took a field in KG equation in BH ST and said this field has some form that look like

$$\phi = \sum_\ell \int d\omega \left[A_{\omega,\ell} f^{\omega t}(\omega,\ell,r_*) Y_\ell(\Omega) + B_{\omega,\ell} f^{\text{in}}(\omega,\ell,r_*) Y_\ell(\Omega)\right] e^{-i\omega t} + \text{hermitian conjugate}$$

We could in principle solve for $f^{\text{in,out}}$ but we could just look at the near horizion behavior

$$f^{\text{out}} \to e^{i\omega r_*} + g_{\omega,\ell} e^{-i\omega r_*} \qquad f^{\text{in}} \to h_{\omega,\ell} e^{-i\omega r_*}$$

We can also look at the field inside the black hole and find

$$\phi = \sum d\omega \ldots$$

And again looking at the near horizion behavior

$$\tilde{f}_{\text{out}}(\omega,\ell,r_*) \to e^{-i\omega r_*}$$

And asserting continuity we get

$$C_{\omega,\ell} = A_{\omega,\ell}h_{\omega,\ell} + B_{\omega,\ell}g_{\omega,\ell}$$

Near the horizion the metric looks like

$$ds^2 \to -dUdV + r^2 d\Omega_{d-1}^2$$

This should remind us of near horizion modes that we saw in the first two lectures. In our older notes we have

$$a_{nh} = \frac{r_0^{d-1}}{\sqrt{\pi\omega_0}} \int \partial_U \phi(U, V = 0, \Omega)(-U)^{-i\omega_0} T(-U) Y_l^*(\Omega) dU d\Omega$$

$$\tilde{a}_{nh} = \frac{r_0^{d-1}}{\sqrt{\pi\omega_0}} \int \partial_U \phi(U, V = -\epsilon, \Omega) U^{i\omega_0} T(U) Y_\ell(\Omega) d\Omega$$

Since we have the field expansion we could in principle plug in ϕ and evaluate the integral very carefully. But in fact we don't have to do the integral becuase near the horizion what does the field look like?

$$\phi \approx \sum_{\omega,\ell} e^{-i\omega t} A_{\omega,\ell}(e^{i\omega r_*} + g_{\omega,\ell}e^{-i\omega r_*}) + (\text{terms with } B_{\omega,\ell}) \approx \sum_{\omega,\ell} A_{\omega,\ell}(U^{i\omega/k} + g_{\omega,\ell}V^{-i\omega/k})Y_\ell(\Omega)$$

Where the above is up to constants. This is because $U \propto e^{k(r_*-t)}$. The $B_{\omega,\ell}$ multiply a different set of terms $e^{-i\omega t}e^{-i\omega r_*}Y_\ell(\Omega)$ which is approximately $B_{\omega,\ell}V^{-i\omega/k}$.

Now looking at a_{nh} and $\tilde{a}_{nh}$ which are basically fourier transforms in $\log U$ we see that the integral picks up the modes with $\omega = \omega_0$ and $-\omega_0$ respectivley. So we can say that

$$a_{\omega,\ell} \approx \int A_{\omega',\ell} q(\omega',\omega) d\omega' +$$

In fact we will find that

$$a_{nh} = a_{\omega_0 k,\ell}$$

Since $q(\omega',\omega)$ picks up modes from near $A_{\omega_0 k}$. We only need that

(a) $q(\omega',\omega)$ is sharply peaked around $\omega' = \omega$.

(b) after smearing $[a_{\omega,\ell}, a^\dagger_{\omega,\ell}] = 1$.

The summary is that the near horizion modes become slightly smeared global modes centered around frequency $\omega_0 k$. We can now immediatley compute the two point function of these modes. Using $\omega_0 = \omega/k$ lets define $\beta = 2\pi/k$.

$$\boxed{\langle a_{\omega,\ell} a^\dagger_{\omega,\ell} \rangle = \frac{1}{1 - e^{-2\pi\omega_0}} = \frac{1}{1 - e^{-\beta\omega}}}$$

And look at that! We have *Hawking radiation*. Lets look at bit more at the $\tilde{a}$ modes

$$\tilde{a}_{nh} \approx \tilde{a}_{\omega_0 k,\ell} = \text{ smeared version of } \tilde{A}_{\omega,\ell}$$

We're actually going to put this aside for now if we're only talking about the exterior of the BH. To get thermal occupancy for $a_{\omega,\ell}$ modes we assumed

(a) Horizion was smooth. this is what fixed the occupancy of the near horizion modes

(b) There was a late time emergent $t \to t + \delta t$ isometry. this is what allows us to relate the near horizion (nh) modes to global modes.

The short distance properties correspond to the smoothness of horizion and the long distance property corresponds to the late time emergent isometry. Something to note is that $a_{\omega,\ell}$ is proporitional to $A_{\omega,\ell}$ which is outgoing to I^+. The stress tensor far away also depends on the properties of f_{out}. Also our derivation does not constrain $B_{\omega,ell}$ since it doesn't appear in the near horiizon mode. So we can choose some things for $B_{\omega,\ell}$

(a) We could put $B_{\omega,\ell}$ in a vacuum $\to$ Unruh state

(b) populate $B_{\omega,\ell}$ thermally $\to$ Kruskal state

(c) $\ldots$

The advantage of this derivation is that it is clear how we can correct this theory by finding the higher order terms. This is an advantage to the rindler anlogy or the ray tracing arguemnt where finding higher order terms is less well defined.

We can also consider AdS Black Holes. In asympotitically global AdS

$$ds^2 \xrightarrow{r\to\infty} -(1+r^2)dt^2 + \frac{dr^2}{1+r^2} + r^2 d\Omega_{d-1}^2$$

The AdS radius is 1. So now we want to consider black holes that form from the collapse of matter.

$$ds^2 \xrightarrow{t\gg 1} -f(r)dt^2 + \frac{dr^2}{f(r)^2} + r^2 d\Omega_{d-1}^2$$

Which is identitical with what we had above except $f(r) = 1 + r^2 - \mu/r^{d-1}$ and μ has the same relationship with M. The horizion once again is at $f(r_h) = 0$ meaning that we are interested in a late infalling observer "take enough" that effect of infalling matter have died out.

Behind the horizion U becomes positive for $r < r_h$ $f(r)$ changes sign so for $r < r_h$ t is a psace-likecoordinate r_* is a time coordinate

$$T = V + U \qquad x = V - U$$

We are considering fields which are normalised

$$\phi \to 0 \qquad \text{as} \qquad r \to \infty$$

System undisturbed by external sources evolves autonomously. For fields of mass m

$$\phi \to \frac{1}{r^D}$$

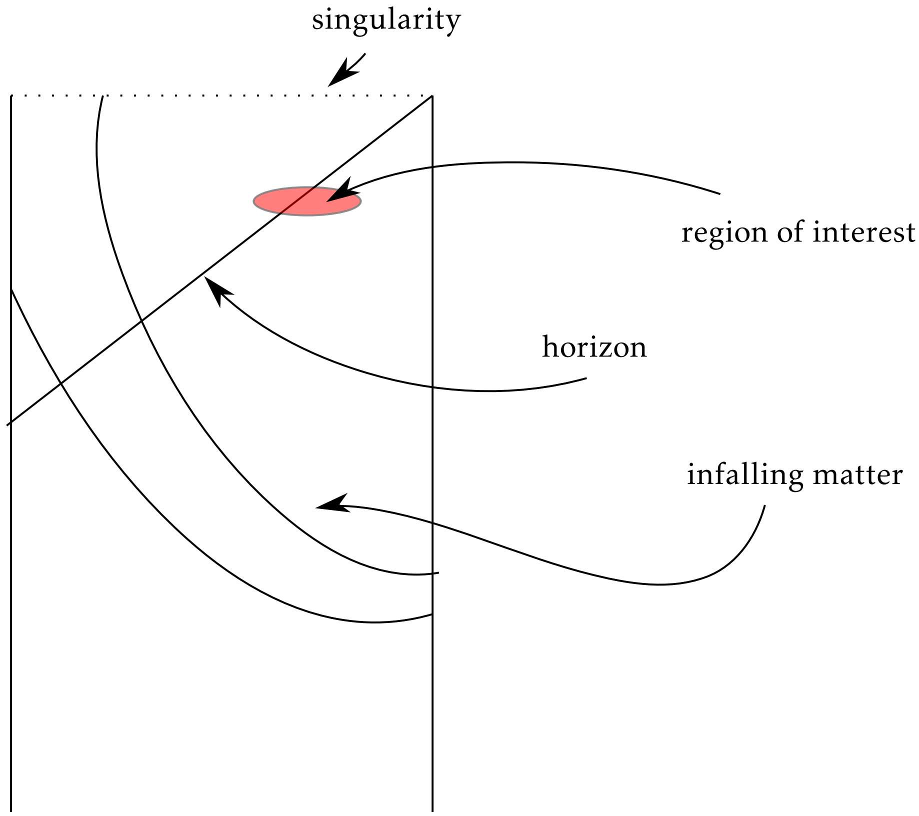


Figure 2: Penrose diagram for AdS black hole

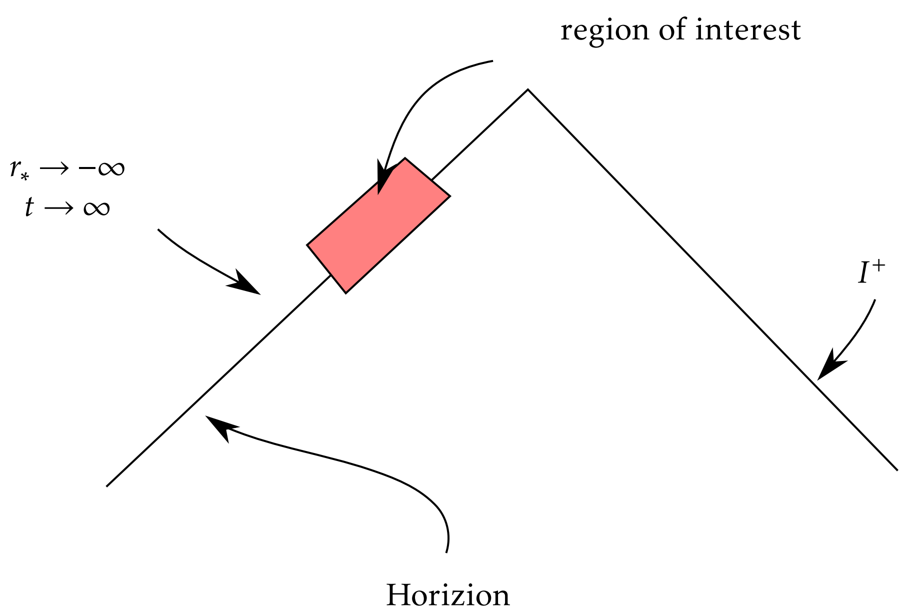


Figure 3: Zooming in on penrose diagram for late time ignore infalling matter

Where $D = d/2 + \sqrt{d^2/U} + m^2$ is also dimension of the dual operator to ϕ. Like before we'll consider tortoise coordinates

$$dr_* = \frac{dr}{f(r)}$$

Near the horizion once again $r_* \to -\infty$ and near infinity $r_* \to \text{const}$. We can once again write down a expansion

$$\phi = \sum_\ell \int d\omega A_{\omega,\ell} f^{\text{st}}(\omega,\ell,r_*)e^{-i\omega t}Y_\ell(\Omega) + \text{hermitian conjugate}$$

Where f^{st} is a standing wave solution. We will also demand near $r_* \to -\infty$

$$f_{\text{st}} \to e^{i\omega r_*} + \underline{e^{-i\delta_{\omega,\ell}}} e^{-i\omega r_*}$$

Where the underlined term is a phase. We need to satisfy

$$r^{d-1}[\phi,\dot{\phi}] = i\delta(r_* - r_*')\hat{\delta}(\Omega,\Omega')$$

Near the horizion

$$[\phi,\dot{\phi}] \approx \int d\omega\ [A_{\omega,\ell}, A^\dagger_{\omega',\ell}]\left(e^{i\omega r_* - \omega' r_*'} + g_{\omega,\ell} g^*_{\omega',\ell'} e^{-i\omega r_* + \omega' r_*'}\right) + (\text{linear in } g_{\omega,\ell}) e^{i\omega r_* + \omega' r_*'}$$

(TODO what?). He says if you didn't follow you can work it out yourself. In the same way behind the horizion we can write down a expansion again

$$\phi = \int d\omega[A_{\omega,\ell}e^{-i\delta\omega\ell}e^{-\omega t}Y_\ell(\Omega) + \tilde{A}_{\omega,\ell}e^{i\omega t}Y^*_\ell(\Omega)]\tilde{f}_{\text{st}}(\omega,\ell,r_*) + \text{hermition conjugate}$$

Where $\tilde{f}_{\text{st}}(\omega,\ell,r_*) \xrightarrow{r\to r_h^-} e^{-i\omega r_*}$. As in flat space we introduce near horizion modes and once again

$$a_{nh} = a_{\omega_0 k,\ell} \quad \tilde{a}_{nh} = \tilde{a}_{\omega_0 k,\ell} \quad k = f'(r_h)/2 \quad a_{\omega,\ell} = \int A_{\omega,\ell} a_\ell(\omega,\omega')d\omega' \quad \tilde{a}_{\omega,\ell} = \int \tilde{A}_{\omega,\ell}\tilde{q}_\ell(\omega,\omega')d\omega'$$

And this leads us

$$\langle a_{\omega,\ell}a^\dagger_{\omega,\ell}\rangle = \frac{1}{1-e^{-\beta\omega}}$$

This is the same as before but here we don't have any flux at infinity. Even ghough $A_{\omega,\ell}$ smeared is thermally occupied there is no flux at inifnity. We obtain a black holes with a thermal atmosphere around it.

$$f^{\text{st}} \xrightarrow{r\to\infty} \frac{\sqrt{G_{\omega,\ell}}}{r^D}$$

And the dual operator $O_{\omega,\ell} = \sqrt{G_{\omega,\ell}}A_{\omega,\ell}$.

Lecture 5: Hawking's Original Paradox

January 27, 2021

Today we're going to formulate the paradox as Hawking stated it. Last time we derived a formula that gives the Hawking temperature

$$T = \frac{k}{2\pi}$$

Before this was derived people knew that

$$dM = \frac{kdA}{8\pi} + \Omega dJ + \phi dQ \qquad dA \geq 0$$

Where A is the area, Ω is the angular velocity of the horizion, J is angular momentum, ϕ is potential of the horizion and Q is chrage. This comes from analyzing classical processes around the particle. E.g. you throw some particle of charge and mass M into a Black Holes and then see what happens. We also know that the area of a black hole is alwyas increasing. This in fact looks a lot like thermodynamics

$$dU = TdS + \text{work terms} \qquad dS \geq 0$$

Feynman likes to say that the same equations have the same solutions. For example LC circuits and springs. They're different physical systems but they can solved in the same way. The finding of the black hole temperature elavated the connection between thermodynamics and blackhole from a formal one to a physcial one. The entropy of the black hole is

$$\boxed{S = \frac{A}{4}}$$

Something we should note is that black holes have a lot of entropy compared to other objects in the universe.

Lets look at Hawking original paper. In it he argued that if you start with some inital state, the final state is a mixed state because you end up losing some information. The first point that Hawking made is that Hawking had argued in a previous paper that black holes steadily create and emit particle with a thermal spectrum. This radiation should take away energy so the black hole must lose mass and eventually evaporate. The details can seen here. We worked out

$$\frac{dM}{dt} = -cAT^{d+1}$$

To work out the constant c you need to work out the greybody factors. This almost follows from dimensional analysis. In flat space we found that

$$M \propto r_h^{d-2} \qquad T \propto \frac{1}{r_h} \qquad A \propto r_h^{d-1}$$

Putting these proportionalies we get

$$\frac{dr_h}{dt} \propto -\frac{1}{r_h^{d-1}}$$

The interesting part of this is that in a time $\propto r_h^d \propto \frac{A}{T}$ we have that $r_h \to 0$. What this tells us is if there's a black hole in a universe where nothing is trying to fight against hawking radiation the black hole will evaporate in a time propotional to the power of the inital radius.

Hawking's next point is that if you take a black hole you need data on the horizion and I^+ to determine I^-. This is different from minkowski where you only need information of I^+ to know the inital state of I^-. What this is saying is that you lose information. From here hawking notes that this is also true quantum mechanically. He does this by computing the occupancy of something. In the first we lectures we had fields as

$$\phi = \int \left[A_{\omega,\ell} f^{\text{out}}(\omega,\ell,r_*) + B_{\omega,\ell} f^{\text{in}}\right] Y_\ell(\omega) e^{-i\omega t} + \text{h.c.}$$

We then define $a_{\omega,\ell}$ which are smeared versions of $A_{\omega,\ell}$ and we got from these

$$\left\langle a_{\omega,\ell} a^\dagger_{\omega,\ell}\right\rangle = \frac{1}{1-e^{\beta\omega}}$$

We then found that there was some flux at infinity. Hawking's point is that this flux you compute does not depend on the state of the $B_{\omega,\ell}$ modes. Not only does it not depend on the B modes but you can compute the state of the $A_{\omega,\ell}$ modes. First we have

$$N_\omega = a^\dagger_\omega a_\omega \qquad \langle N_\omega\rangle = \frac{1}{e^{\beta\omega}-1} \qquad \left\langle N^2_\omega\right\rangle = \frac{1+e^{\beta\omega}}{(e^{\beta\omega}-1)^2}$$

The derive this you can of the following. First we know that

$$(\tilde{a}_{\omega,\ell} - a^\dagger_{\omega,\ell} e^{-\beta\frac{\omega}{2}})\,|\psi\rangle = 0$$

We find that the kind of state that satisfies this equation is

$$|\psi\rangle = e^{e^{-\beta\omega/2} a^\dagger_{\omega,\ell}\tilde{a}^\dagger_{\omega,\ell}}\,|N_{\omega,\ell}=0, \tilde{N}_{\omega,\ell}=0\rangle$$

Then we can find correlators of any polynomial can be worked by writing out a density matrix

$$\left\langle N^q_{\omega,\ell}\right\rangle = \text{tr}(\rho_{\omega,\ell} N^q_{\omega,\ell}) \qquad \rho_{\omega,\ell} = \frac{1}{1-e^{-\beta\omega}} e^{-\beta\omega N_{\omega,\ell}}$$

(something about therma distribution here.) The important part about of all of this is the fact that this is not a pure state. And so we find is that whatever state we start with we have on I^- we end up with a thermal state on I^+ (characterized by the thermal distribtuin.) This is in paradox with the unitarity of quantum mechanics.

Quantum mechniacal evolution takes a state and evolves ith with a unitary operator. So lets say we have some density matrix. AFter time evolution we have

$$|\psi\rangle\langle\psi| \to U\,|\psi\rangle\langle\psi|\,U^\dagger$$

So the fact that you can start in a pure state and go to a mixed state is a sign of the fact that we've lost information along the way. This is because we can think of a mixed state as starting with a pure state and throwing away some information. This is our first paradox! Hawking then gives some intuiion for why this might happen. There is some hidden surface behind the horizion so we should adopt a principle of ignorance. The point of this is say we have some state on I^- as a vector in some hilbert space H_1

$$\psi_A \in H_1$$

And then we have $\psi_B^H \in H_2$ which is a state at the horizion or interor hiddens urface. And finally we have $\psi^+ \in H_3$ which is a state on I^+. Now we assume locality. So we have some S matrix

$$S_{ABC}\psi_A^-\psi_B^H\psi_C^+$$

Is the amplitude to go from $\psi^- \to \psi^H \otimes \psi^+$. So if you sum over all possible states of the hidden surface

$$\sum_{\psi^H} S_{ABC}\psi_A^-\psi_B^H\psi_c^+ S^*_{\tilde{A}\tilde{B}\tilde{C}}\psi_{\tilde{A}}^{-*}\psi_{\tilde{B}}^{H*}\psi_{\tilde{C}}^{+*} = \underbrace{\sum_B S_{ABC}S^*_{\tilde{A}B\tilde{C}}\psi_A^-\psi_{\tilde{A}}^{-*}}_{\rho_{c\tilde{c}}}\psi_C^+\psi_{\tilde{C}}^{+*}$$

So we get a superscattering operator (some generalzation of the unitary evolution operator) that takes pure states to mixed states. There is a nice diagrammitic way to summarize the argument which is in Figure 4. Information requires energy to be stored. Some system has a lot

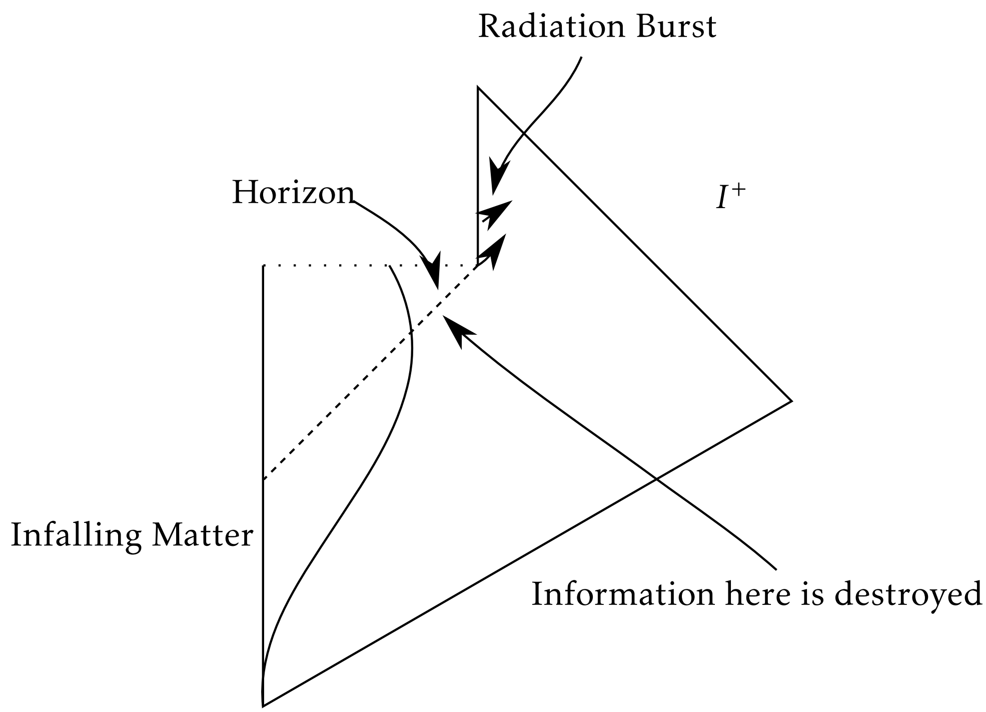


Figure 4: Diagrammitic summary of Hawking's argument of information paradox

of microstates and the microstate gives the information. This is how hard-disks store information. This extends to quantum information. And so the fact taht you have a large number of microstates means you need a lot of energy. The final radiation burst when the black hole evaporates doesn't have enough energy to encode the information so that radiation burst shouldn't have the lost information. So remnants are not viable candiate to resolve information paradox.

Elaboration on energy needing energy: you can store information in microstates. The number of microstates is e^S. The fact that information requires energy is the fact that S at some energy that is bounded.

So there are two elements of Hawking's argument that are crucial

(a) Computation of the thermal density matrix. This is the key computation

(b) Intution from the casual structure of space time.

What we'll ask tomorrow is it really true that computation shows mixed state? And also we'll ask is that is the intuiton about the casual strucutre valid at the accuracy required.

Lecture 6: Mixed and Pure States

January 28, 2021

We'll start by asking a question that seems unrelated to the information paradox: how close are pure states to mixed states? Lets start by discussing the difference between pure states and mixed states. In quantum mechanics our observations are inherantly probailistic. Consider schrodinger's cat. The state of a cat is

$$|\psi\rangle = a_1 |\text{dead}\rangle + a_2 |\text{alive}\rangle$$

And the probability of finding the cat dead is $|a_1|^2$ and alive $|a_2|^2$. This is a pure state. We could similarly have a classical mixture where we have the same probability of having a dead or alive cat. These two states (quatnum state and clasical mixture) can be distinguished. Well the EV of an operator A can be found with

$$\langle\psi|A|\psi\rangle = |a_1|^2 \langle 1|A|1\rangle + |a_2|^2 \langle 2|A|2\rangle + a_2^* a_1 \langle 2|A|1\rangle + a_1^* a_2 \langle 1|A|2\rangle$$

Wheras for the classical mixture we have

$$\langle A\rangle = |a_1|^2 \langle 1|A|1\rangle + |a_2|^2 \langle 2|A|2\rangle$$

So the difference is in the cross terms. The classical mixture is usually represetned by a density matrix

$$p = |a_1|^2 |1\rangle\langle 1| + |a_2|^2 |2\rangle\langle 2|$$

And the way you find an expectation value is compute the $\mathrm{tr}(\rho A)$. Of course you could also compute a density matrix for a pure state

$$p_{\text{pure}} = |\psi\rangle\langle\psi|$$

And notice that this pure state density matrix $p_{\text{pure}}^2 = p_{\text{pure}}$. This property differentiates pure states from mixed states. Now we want to ask how close are pure and mixed states. Lets add a little bit of physics to our situation. Say we have a system with discrete energy level. The mean energy is E_0. We want to consider energy in the region

$$E_0 - \Delta \le E_i \le E_0 + \Delta$$

Where Δ is the spread. We could consider a state

$$|\psi\rangle = \sum_{i=1}^{w} a_i |E_i\rangle$$

Where w is the number of eignetstates in the interval $[E_0 - \Delta, E_0 + \Delta]$. This w is nothing but e^S, the exponential of the entropy. However in stat mech we don't usually consider pure states like this but mixed states with a density matrix

$$\rho_{\text{micro}} = \frac{1}{w} \sum_i |E_i\rangle\langle E_i|$$

This density matrix is the microcanonical density matrix and what this is doing is saying is that you have equal probability to be in one of the corresponding eigenstate. Now the question we wnat to ask is clear. How close is

$$\sum a_i |E_i\rangle \qquad \text{To} \qquad \rho_{\text{micro}}$$

Well in a Hilbert space these two things are orthogonal(there exists a observable where the EV of the observable with $|\psi\rangle$ is 0 and with the density is 1). But if we take a "typical state" $|\psi\rangle$ we find that these are "extremley close" to ρ_{micro}. Lets define typical state an extremley close. We want a physical notion of cloesness. Lets say we have some observer which can make physical observations. From the POV of physical observations how easy or difficult is it to distinguish the two? Physical observations always have to do with probabilities of different outcomes. And these probabilities come from p = projector and trying to find

$$\langle\psi|p|\psi\rangle$$

The physical significance of this statement is consider some observable $\mathcal{O}$ with spectracl composition $\mathcal{O} = \lambda p + \sum_i \lambda_i p_i$. Upone measuring $\mathcal{O}$ what is the probability of getting λ. The expectation value has information on the average value but also λ. Lets say we take

$$\mathcal{O} \to 1000\mathcal{O} \Rightarrow \lambda \to 1000\lambda$$

In this rescaling the probabilities are invariant. So if we have a typical pure state and some projector p how does $\langle\psi|p|\psi\rangle$ compare with $\text{tr}(\rho P)$ where p is some micro mixed state. This is a more physical way of defining similarlity. Now we'll finally define a typically state. Recall some properties of the hilbert space

$$\sum a_i |E_i\rangle \qquad \sum |a_i|^2 = 1$$

One way to think about this state is to think about it as a sphere. The hilbert space is a sphere in very high dimensions and we're just picking points on the surface of the sphere. Note we still have som ambiguity in phase ($|\psi\rangle \to E^{i\phi}|\psi\rangle$) The question now is lets say we have the surface of the sphere if we pick a random point on the sphere and ask how close is

$$\langle\psi|p|\psi\rangle \qquad \text{to} \qquad \text{tr}(\rho P)$$

We'll do this by defining a volume measure on the sphere. What is a natural probability to shove onto a sphere. Well first we define a measure on a hilbert space

$$d\mu_\psi = \frac{1}{V}\delta\left(\sum |a_i|^2 - 1\right)\pi d^2 a_i \Rightarrow \int d\mu_\psi = 1$$

The d^2a_i is for complex numbers, δ function is for normalizaiton and $\frac{1}{V}$ is another normalziation factor. Now lets define

$$\delta = \langle\psi|p|\psi\rangle - \mathrm{tr}(\rho p)$$

We want to compute $\langle\delta\rangle_M$ and $\left\langle\delta^2\right\rangle_M$. Now lets do some computations

$$\langle\psi|p|\psi\rangle = \sum_i |a_i|^2 \langle E_i|p|E_i\rangle + \sum_{i\neq j} a_i a_j^* \left\langle E_j|p|E_i\right\rangle$$

We need to average this over all possible pure states using the measure we defined above $\int\langle\psi|p|\psi\rangle d\mu_\psi$. To compute this we need $\int |a_i|^2 d\mu_\psi$ and $\int a_i a_j^* d\mu_\psi$. This is easy to work out because $\left\langle |a_i|^2\right\rangle$ cannot depend on i. This means we can just as easily compute

$$\frac{1}{w}\left\langle\sum |a_i|^2\right\rangle_\mu = \left\langle |a_i|^2\right\rangle_\mu = \frac{1}{w}$$

Similarly

$$\left\langle a_i a_j^*\right\rangle_\mu = \delta_{ij}$$

There is no correlcation betwen a_i and a_j because we're picking things randomly. All of this gives us

$$\boxed{\int\langle\psi|p|\psi\rangle d\mu_\psi = \frac{1}{w}\sum_i \langle E_i|p|E_i\rangle = \mathrm{tr}(\rho p) \Rightarrow \langle\delta\rangle = 0}$$

This does not end the story. We need to now compute $\left\langle\delta^2\right\rangle$. It's a bit more involved but we can do it.

$$\begin{aligned}\left\langle\delta^2\right\rangle &= \int [\langle\psi|p|psi\rangle - \mathrm{tr}(p\rho)]^2 d\mu \\ &= \int \sum_{i,j,k,l} \left\langle E_j|p|E_i\right\rangle (a_i a_j^* - \delta ij/2) \langle E_l|p|E_k\rangle (a_k a_l^* - \delta_{kl}/w) d\mu\end{aligned}$$

We need $\int a_i a_j^* a_k a_l^* d\mu = \dfrac{\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk}}{w(w+1)}$

If we're careful we can notice that

$$\delta^2 \leq \sum_{i,j} \frac{1}{w(w+1)} \left\langle E_i|p|E_j\right\rangle\left\langle E_j|p|E_i\right\rangle \leq \sum_i \frac{1}{w(w+1)} \left\langle E_i|p^2|E_i\right\rangle \leq \frac{1}{w+1}$$

The second relation comes from notion that E_j is almost but not a complete basis. The above result tells us

$$\left\langle\delta^2\right\rangle_\mu \leq \frac{1}{w+1}$$

This is significant since $w \propto e^S$ meaning that average deviation are of size $e^{-\frac{s}{2}}$. So pure states not only on average look like mixed states but also the deviation is small. To summarize

(a) For a given observable for most pure states look exponentially to the maxiamlly mixed state

(b) Volume of "atypical state"(a state where δ is large) is exponentially small. (this results comes from $\left\langle\delta^2\right\rangle_\mu \propto \frac{1}{e^S}$

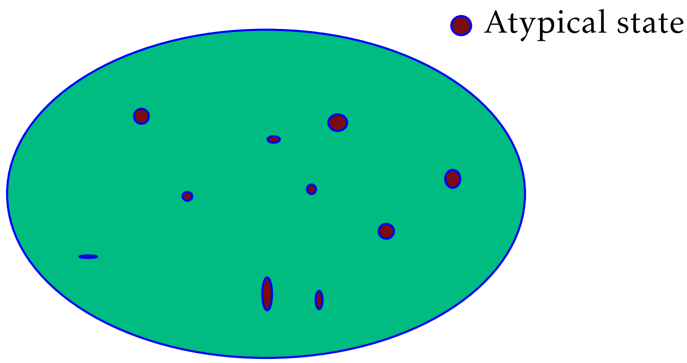


Figure 5: Illustration of the Hilbert state and atypical state

Now lets talk about things that coudl cause confusion

(a) basis vs typical states: We have n qubits which can be spanned by basis

$$|00\ldots0\rangle$$
$$|10\ldots0\rangle$$
$$|01\ldots0\rangle$$

We have σ_3^i that has definite values in all these states. Basis states occupy 0 volume.

(b) ETH = Eigenstate Thermlization Hypothesis: Our result is kinematical. We never used ETH.

(c) Entanglement entropy. $S = 0$ for pure state and $S = \ln W$ for mixed states. You might say that engatnelment entropy differentiates pure state and mixed state. But the entanglement entropy isa fine grained observable and the expecation value of observavble. $S \neq \langle A\rangle$. The reason for this simple. One can always find a basis of states for which $S = 0$. If S was the expectation value for an operator then it's the same for all basis. Thus S is not an expectation value of any operator. However S is a function of correlation functions.

So how does this all apply to hawking original paradox.

$$\left\langle a_{\omega,\ell}a^\dagger_{\omega,\ell}\right\rangle = \frac{1}{1-e^{-\beta\omega}} \qquad \left\langle a_{\omega,\ell}a^\dagger_{\omega',\ell}\right\rangle = 0 \text{ for } \omega \neq \omega'$$

This lead to the conclusion that final state was mixed. However if the had found

$$\left\langle a_{\omega,\ell}a^\dagger_{\omega,\ell}\right\rangle = \frac{1}{1-e^{-\beta\omega}} + O\left(e^{-\frac{S}{2}}\right) \qquad \left\langle a_{\omega,\ell}a^\dagger_{\omega',\ell}\right\rangle = O\left(e^{-\frac{s}{2}}\right) \text{ for } \omega \neq \omega'$$

(The results by hawking aren't exact) then we would have found that perfectly consistent with pure states. What we're saying is that Hawking's original paradox is not a paradox. In fact it would have been weird if we hadn't found a mixed state. In fact the accuracy of our computation $\ll e^{-S/2}$. Consider the following. What if we formed a b.h. in a micro ensemble or in canonical ensemble. Observables in these two are different by an amount we can compute $\frac{1}{\sqrt{S}} \gg e^{-\frac{S}{2}}$. The fact that simple correlators look thermal does not imply that the final state is thermal. As we have just shown pure states and thermal states can look expoentially close to eachother. The punchline is this

Hawking's original coputation is not precise enough to be a paradox

There are some things we have not answered here. We can refine the paradox. Also what about the "principle of ignorance" that came from looking at the causal structure of spacetime.

Lecture 7: Trying to Localize Operators behind the Horizon

February 03, 2021

There were two parts to Hawking's original paradox

(a) concrete computation

$$\left\langle a_{\omega,\ell}, a^{\dagger}_{\omega,\ell} \right\rangle = \frac{1}{1 - e^{-\beta\omega}}$$

(b) Some intuitional: Hawking looked at the casual structure of the black hole geometry. This is Figure 4. So because the causal structure is different we clearly don't have information on the insdide. This is what gives us the "principle of ignorance" where the observer doesn't know what goes on inside the black hole.

Last time we talked about (a) with stat mech. If $|\psi\rangle$ is a typical state in a large hilbert space from a given energy band

$$\left| \langle \psi | A | \psi \rangle - \frac{1}{e^S} \mathrm{tr}(A) \right| \propto e^{-\frac{s}{2}}$$

Basically mixed a pure states are very close. One thing Suvrat wanted to emphesize last time is that black hole has has a natural perturbative parameter

$$T \propto \frac{1}{r_H} \qquad G_N T^{d-1} = \text{natural perturbative parameter} \propto \frac{1}{S}$$

No one has worked out final state entropy non-perturbativley and this is the only way we could transform hawking's original paradox to an actual paradox.

Someone could ask what about the "principle of ignorance." First lets try to frame this more precisley. An observer outside the blackhole doesn't have "no information" about a black hole. There is a "no hair" theorem. A blackhole is characterized by mass, angular momentum, and charge. A observer can observe these things. However this information isn't enough to get all the information about the inital state. We should emphesize that the no hair theorem is a *classical result*. There is no analagous quantum result. This is not some technical point about lack of proof. This is a substanstive point about the difference between quantum and classical. We will claim that we shall never expect a quantum no-hair theorem.

In classical physics, why do we always know the mass of a black hole? It's because classically, gravity obeys a gauss's law (aparently this guy learned this in high school???) We have an expresssion

$$H = \frac{1}{16}\int (\partial^i \underbrace{h_{ij}}_{\text{deviation from flat metric}} - \partial_j h_{ii}) n^j d^2 s$$

We'd also like to expect something like this to will work in QM as well. Lets try to see this

$$|\psi\rangle = \sum_i a_i |E_i\rangle$$

Note that a blackhole formed by collappse and evaporates it cannot be an energy eigenstate since energy eigenstates don't evolve. Thus it must be a superposition of distinct energies(these energy eigenstates are wrt Hamiltonian of some imaginary theory of quantum gravity.) From here we can measure

$$\langle \psi | H | \psi \rangle = \sum |a_i|^2 E_i$$

Quantum mechanically we can measure more than the value of $\langle H \rangle$ but can also measure other observables like

$$\left\langle \psi | H^2 | \psi \right\rangle = \sum_i |a_i|^2 E_i^2 \neq \langle \psi | H | \psi \rangle^2$$

We do expect spreads in energy. We also have information about

$$p(E, E+\Delta) = \sum_{E<E_i<E+\Delta} |a_i|^2$$

There is no classical analouge of this. Quantum mechanically we have more information. But this is not the end of the story. Lets say $\mathcal{O}$ is someother observable without a conserved charge. Like $\mathcal{O} \to \phi$ as $r \to \infty$. We can ask about correlators

$$\langle H\mathcal{O} \rangle \neq \langle H \rangle \langle \mathcal{O} \rangle$$

Lets say someone was trying to prove a theorem(a quantum version of no-hairtheroem)

$$\langle H\mathcal{O} \rangle \to \text{universal value at late time}$$

Is this accurate up to $e^{-\frac{S}{2}}$ corrections. These exponentially small corrections is where the information is. So quantum mechanically the obstacles to no-hair theorem are much more numerous than in classical mechanics. This is why there is no quantum no hair theorem. The spirit of the no-hair theroem is that the geometry settles down. So proving $\langle H\mathcal{O} \rangle$ would be proving no-hair in spirit.

Lets look at this another way. Recall Figure 4 Lets say we draw some slice through the horizon. In local QFT it's possible to change something inside the horizon without changing anything outside the horizion. Now lets say we want to prove there is a principle of ingnraocne. Then the person had to say "hey there's two kinds of states in the black hole" and that no observer outside the horizon can know if the field is excted or not inside the horizon. For the principle of ignorance too hold there has to exists a unitary U such that U commutes with all operators outside but changes the inside.

$$\left\langle\psi|U^{\dagger}\mathcal{O}(\text{out})U|\psi\right\rangle = \langle\psi|\mathcal{O}(\text{out})|\psi\rangle$$

So it must be the case that $U^{\dagger}HU = H$. So U must commute with the hamiltonian. But know we are in trouble because if U commutes with the Hamiltonian, U cannot be localized to the black hole interior. And if it is not purely localized insie the interior then it must change some observable outside the horizion. This sometimes goes under the name "there are no gauge invariant quantum operators in quantum gravity." To summarize. If someone wanted to make hawking's princile of ignoracne precise then there would have to be at least one unitary U so that this unitary U commmutes with all observables outside but change something inside. But this would mean that U commutes with the Hamiltonian. In this case then U would have zero energy and thus could not be localized. If it is not localized in the interior then it must fail to commute with some operator on the outside.

Lets look at yet another argument. In any gauge theory lets say we wanted to put a charge somewhere. An operator with charge isn't gauge invaraint. We neeed to dress this operator up with a wilson line to infinity (or wilson loops) to make this gauge invariant. In graivty the "Wilson line" run to infinity. There's another way to say this. Lets say we want to act with some operator and create an excitation insie the black hole.

$$\phi(x)|\psi\rangle$$

You might have thought that ϕ is a scalar field and thus is gauge invarint. However if you consider

$$x^{\mu} \to x^{\mu} + \epsilon^{\mu} \Rightarrow \phi \to \phi + \partial_{\mu}\epsilon^{\mu}$$

How do you make this difeomrophism invariant? well we could define the observable relationally to infinity. So you can define a notion of "dressing" in gravity by starting at the boundry of the metric, drop some geodesic, and them measure the operator at the end of the geodesic. So in gravity there is no such thing as a localized gauge invariant operator. This is actually a well known result that there is no such thing as local gauge-invariant operator in quantum-gravity. Mathematically we could also just say "we fix a gauge."

So in this lecture we have gone over three intuitional arguments for why we can't change the interior without knowing outside the horizon. These refute the "prinicple of ignorance" that Hawking argued for. The error in Hawking's argument we found are

(a) Computation of low-pt correlatiors are insuffiecnt to conlcude that final states are thermal

(b) Intuition of principle of ignorance is very sublte.

So both of these tell you that we have information loss have been refuted. Next time we'll talk about a more formidable paradox. Some paradoxes by Mathur and AMPS firewall/fuzzball paradox

Lecture 8: Some Results from Quantum Information

February 04, 2021

Today we're going to go on a detour. We're going to talk about some simple techniques and questions from quantum informtion.

The argument of Mathur/AMPS:

(a) if you have a smooth horizon then the modes inside and outside the horizon must be entangled if the horizon is smooth. We encountered this result in some sense when we looked at correlators accross the null surface.

(b) Typicality requires entangleemnt between modes that are close to the horizon (near-horizon modes) and far-away modes at late times.

(c) These two requirements are inconsistent because entanglement in monogamous.

At this point we're not in a place to make this paradox precise. Lets first talk about entanglement. He wants to emphasize that these arguments look at the inside (makes quantitative reference to the interior.)

So what are the original paradoxes used entanglement entropy. We'll use Bell corrleators and their refinement CHSH correlators to phrase these paradoxes, monogomy, and entanglement. These correlators are well defined in a theory of gravity.

Definition 1: (CHSH Correlators) Lets say we have two distinct systems A_1, A_2, B_1, B_2. Now lets say each observable takes values $\in [-1,1]$. First lets look at the classical case. Consider the situtaiton where A and B two coins.

$$A_1 = \begin{cases} +1 & \text{heads} \\ -1 & \text{tails} \end{cases} \qquad A_2 = \begin{cases} +1 & \text{dime} \\ -1 & \text{quarter} \end{cases}$$

And same idea for B. We're not going to put any constraints on the probabilty distribution. The point is to define a joint observable, the CSSH observable

$$C_{AB} = A_1(B_1 + B_2) + A_2(B_1 - B_2)$$

In the classical case it should be clear that $|C_{AB}| < 2 \Rightarrow |\langle C_{AB}\rangle| \le 2$. In the quantum mechanical case there is something case Tsirelson's bound which says

$$|\langle C_{AB}\rangle| \le 2\sqrt{2}$$

In Q.M we again have

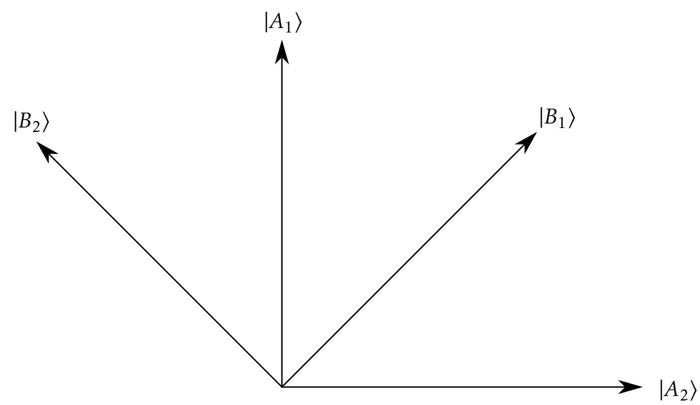
$$C_{AB} = A_1(B_1 + B_2) + A_2(B_1 - B_2) \Rightarrow |A_i|, |B_i| \le 1$$

The statement that the operators probe different systems is that $[A_i, B_j] = 0$. So how does this Tsirelson's bound happen? Lets look at a configuration where this happens. The flaw in our CM system we assigned all values simulatenously. But now if we have some spinor and A_1 was measuring the z component and A_2 was measuring the x component then we could violate the correlator.

$$\text{If } 2 \le |\langle C_{AB}\rangle| \le 2\sqrt{2}$$

Then the system is said to be entangled(correlators between systems exceed maximum allowed classical values.)

Say we have a joint state $|\psi\rangle$. Lets see how we'd saturate Tsirelson's bound. Lets use the notation $|A_i\rangle = A_i|\psi\rangle$ and $|B_i\rangle = B_i|\psi\rangle$. And now we have a very simple way to sataurate the bound. Consider



$|B_1\rangle + |B_2\rangle = \sqrt{2}|A_1\rangle$ $|B_1\rangle - |B_2\rangle = \sqrt{2}|A_2\rangle \Rightarrow \langle C_{AB}\rangle = \langle A_1|(|B_1\rangle + |B_2\rangle) + \langle A_2|(|B_1\rangle - |B_2\rangle) = 2\sqrt{2}$
Say $A_i^2 = B_i^2 = 1$. Then

$$C_{AB}^2 = 4 - [A_1, A_2][B_1, B_2]$$

Now lets talk about the monogamy of entanglement. Contuining from our definition above now lets say we have a system C where

$$[C_i, A_j] = [C_i, B_j] = 0$$

From here we can define an observable

$$C_{AC} = A_1(C_1 + C_2) + A_2(C_1 - C_2)$$

From our notion of entanglement defined above we can again see how entangled things are. Something else to note is that there is a remarkable inequality

$$\langle C_{AB}\rangle^2 + \langle C_{AC}\rangle^2 \leq 8 \Rightarrow \{|\langle C_{AB}\rangle| < 2 \Rightarrow |\langle C_{AC}\rangle| < 2\} \text{ and converse}$$

"This inequality shows that classical correlators can be shared but quantum correlators cannot." This also shows that correlators are monogomous[1].

We can now talk about average entanglement between subsystesms. We have two systems $e^{S'}$ and e^S in a big system with hilbert space dimension $e^{S+S'}$. We'll also assert that $e^{S'} \ll e^S$. This turns out to not be that strong a statement and really is here to make our life easier. The reason this isn't a very stron statement is because black holes have lots of entropy. Say one system has 10^{10} qubits and the other system has $(1-10^{-6}) \times 10^{10}$. That's a small difference in qubits but a huge difference in e^S and $e^{S'}$. Now we want to ask in some precise sense "if we take a typical state of the large system, how are the subsystesms entangled." There are two answers to this

[1] cute

question. The answer relies on considering the *density matrix*. Recall that we defined before as taking the Haar measure on the big system. We'll make the following claims

(a) Given operators with simple properties in smaller subsystem then we can find some operators in the larger subsystem to saturate Tsirelson's bound

(b) The entanglement entropy between the subsystem obeyes a "Page curve."

Lets try to prove property (a). We can write a typical state as

$$|\psi\rangle = \sum_{n=1}^{e^S}\sum_{m=1}^{e^{S'}} a_{mn}|m,n\rangle \qquad d\mu_\psi = \pi d^2 a_{mn}\delta\left(\Sigma|a_{mn}|^2 - 1\right)$$

The density matrix of the smaller subsystem is

$$\rho_{mn'} = \sum_{n=1}^{e^S} a_{mn}a^*_{m'n}$$

It should be emphasized that the eigenvalues of the larger density matrix are **the same**. One can always choose a basis to "diagonalize" a_{mn} (actually since this is rectangular it's more precise to do a singular value decompositiion) meaning that you can always write (after some change of basis in smaller and larger system)

$$|\psi\rangle = \sum_{\alpha=1}^{e^{s'}} \sqrt{\rho_\alpha}\,|\alpha,\alpha\rangle$$

This tells us that the rank of the larger density matrix at most have $e^{S'}$. So the first question we can ask is what is $\langle\rho_{mm'}\rangle$ (wrt to the Haar measure). We can fidn this with

$$\langle\rho_{mm'}\rangle = \int\sum_n a_{mn}a^*_{m'n}d\mu_\psi = \frac{1}{e^{S+S'}}\sum_n \underbrace{\delta_{mm'}\,\delta_{nn}}_{e^S} = \frac{\delta_{mm'}}{e^{S'}}$$

There is a more useful thing we could ask

$$\left\langle\sum_{mm'}\left|\rho_{mm'} - \frac{1}{e^{s'}}\delta_{mm'}\right|^2\right\rangle = \frac{e^S + e^{S'}}{e^{S+S'}+1} - \frac{1}{e^{S'}} = \delta e^{S'}$$

This computation is done in the lecture value. To get a good notion of the average, we should divide everything by $1/e^{S'}$. At large S and S' we have

$$\delta \approx \frac{1}{e^{S+S'}}$$

This δ is telling us about the average deviation squared and $\sqrt{\delta}$ is the size of the average deviation of eigenvalues.

$$\sqrt{\delta} = \frac{1}{e^{\frac{S+S'}{2}}} \ll \frac{1}{e^{S'}}$$

This tells us that the average density matrix is close to the identity [normalized]. Notice by the way that this does not work if we trtyto use this argument for the larger system. We can now choose a orhtonormal (schmidt?) basis such that

$$|\psi\rangle = \frac{1}{e^{S'/2}} = \sum_{m=1}^{e^{S'}} |m, \tilde{m}\rangle$$

Note that not all states can be written in this way but for a typical state we should be fine. Say we have a pseudospin operator in the smaller subsystem. What we mean is that these operators should behave

$$A_1^2 = A_2^2 = 1 \Rightarrow (A_1 + A_2)^2 = (A - 1 - A_2)^2 = 2$$

This is our way of saying we're looking at a qubit. We can find such operators in QFT as well, it's not only qubit operators. The result we claimed in the beginning is that we can find some operator in the larger subsytem to saturate Tsirelson's bound. Lets show that. Say we have some state. Lets define the matrix elements. Let $|m\rangle$ be the basis that appears inthe schmidt decomposition

$$A_1 |m\rangle = \sum_q (A_1)_{mq} |q\rangle \quad A_2 |m\rangle = \sum_q (A_2)_{mq} |q\rangle$$

The point about finding operators taht are entangled with these operators is that given these operators we can define another operator $\tilde{A}_1 |\tilde{m}\rangle$ that actsin the larger subsystem.

$$\tilde{A}_1 |\tilde{m}\rangle = \sum_q (A_1)_{qm} |\tilde{q}\rangle \leftarrow \text{transpose of above}$$

The matrix elements are the transpose of the smaller subsystem. And we have something similar for $\tilde{A}_2$. It should be clear that

$$|\tilde{A}_1| = |\tilde{A}_2| = 1$$

$$\tilde{A}_1 |\psi\rangle = \frac{1}{e^{S'/2}} = \sum_{m=1}^{S'} \tilde{A}_1 |m, \tilde{m}\rangle = \frac{1}{e^{S'/2}} = \sum_{mq} (A_1)_{qm} |m, \tilde{q}\rangle$$

Renaming $q \leftrightarrow m$ we have

$$\tilde{A}_1 |\psi\rangle = \frac{1}{e^{S'/2}} = \sum_{m,q} (A_1)_{mq} |q, \tilde{m}\rangle = A_1 |\psi\rangle$$

And similarly we can find $\tilde{A}_2 |\psi\rangle = A_2 |\psi\rangle$. Now define

$$B_1 = \frac{1}{\sqrt{2}}\left(\tilde{A}_1 + \tilde{A}_2\right) \quad B_2 = \frac{1}{\sqrt{2}}\left(\tilde{A}_1 - \tilde{A}_2\right) \quad B_1^2 = B_2^2 \quad C_{AB} = \text{same as above}$$

This gives us

$$\langle\psi|C_{ab}|\psi\rangle = \sqrt{2}\left\langle\psi|A_1^2|\psi\right\rangle + \sqrt{2}\left\langle\psi|A_2^2|\psi\right\rangle = 2\sqrt{2}$$

To summarize: what we hae argued here is that we had some large system that we devided into two parts (where the dimension of smaller much smaller than alrger.) We then showed that the density matrix of the smaller system is nearly diagonal. Now focusing on an arbitrary operator we showed that we construct a operator in the larger system that is maximally entangled.

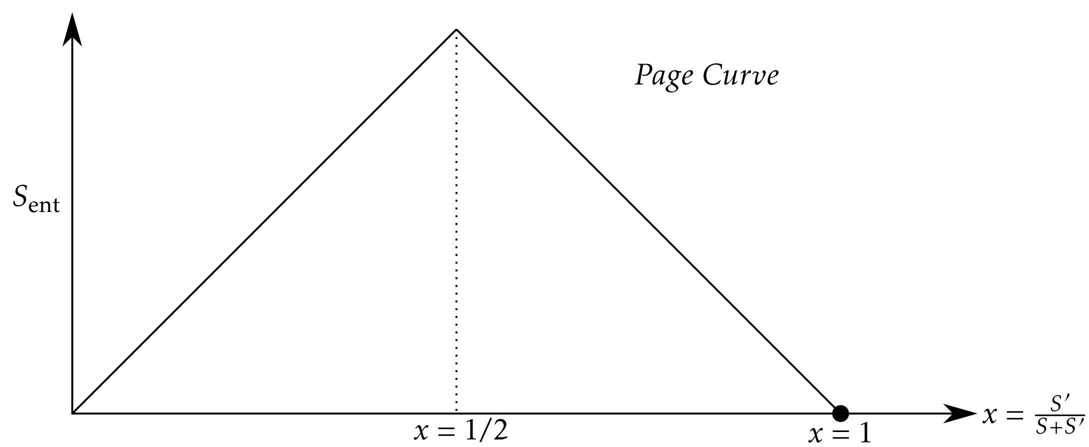
We'll now talk about the page curve. Consider the entanglement entropy of two systems. Lets try to find the exepctation value of this wrt the haar measure

$$\langle -\mathrm{tr}(\rho \ln \rho) \rangle_{\mu_\psi} \approx S'$$

This is true because the density matrix is almost identity. This is also true for the larger system. So we have an assymetry. In general

$$-\mathrm{tr}(\rho \ln \rho) = \min(S', S) \text{ provided } |S' - S| \gg 1$$

The statement is that theentaglement entropy is the dimension of the smaller ofthe two systems. We can now plot the entanglement entropy as a function of the fraction of system size.



The peak is very sharp in a thermosynamic ssystem. We'll apply this to Black Holes soon.

Lecture 9: Marthur/AMPS paradox

February 10, 2021

We're going to start talking about the next paradox of the course: the monogamy paradox. This was formulated by Marthur and then refined by AMPS. This paradox is actually relaly simple to describe

(a) Modes accross horizon are entangled

(b) For old black hole near horizion modes are entangled with far away modes.

(c) This contradicts the monogamy of enganglement

So this is pretty simple to write down. Lets try to make all of this precise. First we'll make (a) precise and then make (b) precise and then we'll see (c). The reason to do this is in more detail than required is because we want to tie up as many loose end as possible. It should also be noted that much of the literature that formulated this paradox was in terms of Von Neumann entropy. However VN entropy is a UV divergent quantity which is a disadvantage. We'll instead focus on CSHS correlators and use an EFT. Lets start by talking about entanglement accross the horizon.

We've defined modes accross the horizon before. We start with some field which has modes $a_{\omega,\ell}$ and $\tilde{a}_{\omega,\ell}$ are right moving modes that live in the near horizon region. These are local degrees of freedom. These aren't mdoes that come from integrating the quantum field over all space. This gives us a way of describing local degrees of freedom. We wrote down some equations describing these modes

$$(a_{\omega,\ell} - e^{-\beta\frac{\omega}{2}}\tilde{a}_{\omega,\ell})|\psi\rangle = 0 + O(\epsilon)$$

$$(a^\dagger_{\omega,\ell} - e^{\beta\omega/2}\tilde{a}_{\omega,\ell})|\psi\rangle = 0 + O(\epsilon)$$

The reason we don't care about $O(\epsilon)$ is because we'll find a $O(1)$ paradox. Now we also found that there is a correlation function

$$\langle\psi|a_{\omega,\ell}\tilde{a}_{\omega,\ell}|\psi\rangle = \frac{e^{-\beta\omega/2}}{1-e^{-\beta\omega}}$$

We can't put this into a bell inequality just yet because these aren't bounded operators. They don't have norm 1. We now want to extract some operators that are operators of norm 1 that then we can show exceeds the classical entanglemnt entropy. Another point: thes a and $\tilde{a}$ are gotten from integrating a quantum field with a smear function. However they are still simple harmonic degrees of freedom. Suvrat wants to emphasize that what we're doing is simple. So given these operators we can define the number operator

$$N_{\omega,\ell} = a^\dagger_{\omega,\ell}a_{\omega,\ell} \qquad \tilde{N}_{\omega,\ell} = \tilde{a}^\dagger_{\omega,\ell}\tilde{a}_{\omega,\ell}$$

Given this number operator we can write a projector onto the eigenstates of the number operator

$$|n\rangle\langle n| = \int_0^{2\pi} e^{i\theta(N_{\omega,\ell}-n)}\frac{d\theta}{2\pi}$$

This integral is only 1 is $N_{\omega,\ell} = n$ and 0 otherwise. The operator on the RHS is well defined in an effective field theory. the LHS is just notation since the hilbert space for a black hole is hard to describe. Also we should point out that in the prescence of interaction, $N_{\omega,\ell}$ is not exactly quantized. However the projector will still get close to 1. There might be seom $O(\epsilon)$ corrections but it won't actually affect the $O(1)$ final paradox. We can also define

$$|n\rangle\langle n+1| = \frac{1}{\sqrt{n+1}}|n\rangle\langle n|a_{\omega,\ell}$$

$$|n+1\rangle\langle n| = \frac{1}{\sqrt{n+1}}a^\dagger_{\omega,\ell}|n\rangle\langle n|$$

Now just like we defined these operators we can define corresponding operators for modes on the other side of the horizon

$$|\tilde{n}\rangle\langle\tilde{n}| = \int e^{i\theta(\tilde{N}_{\omega,\ell}-n)}\frac{d\theta}{2\pi} \quad |\tilde{n}\rangle\,\tilde{n}+1 = \frac{1}{\sqrt{n+1}}|\tilde{n}\rangle\langle\tilde{n}|\tilde{a}_{\omega,\ell} \quad |\tilde{n}+1\rangle\langle\tilde{n}| = \frac{1}{\sqrt{n+1}}\tilde{a}^\dagger_{\omega,\ell}|\tilde{n}\rangle\langle\tilde{n}|$$

So now we can define A_1 (which will go into the CHSH inequlatiy)

$$A_1 = \sum_{n=0}^{\infty}(|2n+1\rangle\langle 2n+1| - |2n\rangle\langle 2n|) \quad A_2 = \sum_{n=0}^{\infty}(|2n+1\rangle\langle 2n| + |2n\rangle\langle 2n+1|)$$

The way to think about these oscialltors. Think about a single SHO. In this SHO we have a lot of number levels with odd and even levels. These operators plug these operators into two. What A_1 and A_2 are doing behave like σ_z and σ_x on each group of two levels. We'll also see that the higher terms don't really matter. Okay. Within an EFT we can square A_1

$$A_1^2 = \sum_{n,m}(|2n+1\rangle\langle 2n+1| - |2n\rangle\langle 2n|)(|2m+1\rangle\langle 2m+1| - |2m\rangle\langle 2m|)$$

The only thing we actually need here is the fact that if you multiply two projectors thjen

$$|n\rangle\langle n||m\rangle\langle m| = \delta_{nm}|n\rangle\langle n|$$

Thus our sum above can turn into

$$A_1^2 = \sum_n |2n+1\rangle\langle 2n+1| + |2n\rangle\langle 2n| = 1$$

And a similar results follows for A_2^2. One more thing for future reference. We can find that

$$A_1A_2 = \sum_n |2n+1\rangle\langle 2n| - |2n\rangle\langle 2n+1| \quad A_2A_1 = \sum_n |2n\rangle\langle 2n+1| - |2n+1\rangle\langle 2n| = -A_1A_2$$

This means that A_1 and A_2 anti-commute. Thus

$$(A_1 + A_2)^2 = 2$$

Thus we've found that A_1 and A_2 are pseudo-spin operators like we defined last lecture. Now we want to compute some correlators for these operators.

$$\langle\psi|A_1\tilde{A}_1|\psi\rangle$$

We could just compute these in EFT. But we can take a shortcut. To do this we define a density matrix and trace out all the other DOF in the black hole. So lets go back to

$$\tilde{N}_{\omega,\ell}|\psi\rangle = \tilde{a}^\dagger_{\omega,\ell}\tilde{a}_{\omega,\ell}\psi = \tilde{a}^\dagger_{\omega,\ell}a^\dagger_{\omega,\ell}|\psi\rangle e^{-\beta\omega/2}$$

These modes also commute with eachother since they're on either side of the horiozn meaning that

$$= a^\dagger_{\omega,\ell}\tilde{a}^\dagger_{\omega,\ell}|\psi\rangle e^{-\beta\omega,2} = a^\dagger_{\omega,\ell}a_{\omega,\ell}|\psi\rangle = N_{\omega,\ell}|\psi\rangle$$

This actually holds for higher powers of $N_{\omega,\ell}$ as well. $\tilde{N}^q_{\omega,\ell}|\psi\rangle = N^q_{\omega,\ell}|\psi\rangle$. Therefore it's convient to think about this state $|\psi\rangle$ as

$$|\psi\rangle \approx \sum a_n|n,\tilde{n}\rangle\otimes\psi_n^{\text{rest}}$$

Furtheromore from

$$(\tilde{a}_{\omega,\ell} - e^{-\beta\omega/2}a^\dagger_{\omega,\ell})|\psi\rangle = 0$$

We can intuit that

$$|\psi\rangle \approx \left(\sum_n e^{-\beta\omega n/2}|n,\tilde{n}\rangle\right)\underbrace{\left(1-e^{-\beta\omega}\right)^{1/2}}_{\text{normalization}}\otimes\psi^{\text{rest}}$$

Thus we can define an effective density matrix

$$\rho_{\text{eff}} = (1-e^{-\beta\omega})\sum_n |n,\tilde{n}\rangle\langle n,\tilde{n}|\, e^{-\beta\omega n}$$

The above is an approximate statment which is useful for computing simple correlators. The the exact statement of what an approximate factorization is the followign. Consider $\mathcal{O}$ which is a simple oeprator of $a_{\omega,\ell}\tilde{a}_{\omega,\ell}$ with h.c. Then

$$\langle\psi|\mathcal{O}|\psi\rangle = \text{tr}(\rho_{\text{eff}}\mathcal{O}) + O(\epsilon)$$

Lets compute some simple correlators of these $\mathcal{O}$

$$\langle\psi|A_1\tilde{A}_1|\psi\rangle = \langle\psi|(|2n+1\rangle\langle 2n+1| - |2n\rangle\langle 2n|)\otimes(|2\tilde{n}+1\rangle\langle 2\tilde{n}+1| - |2\tilde{n}\rangle\langle 2\tilde{n}|\,|\psi\rangle = 1$$

Where the only non-vanishing comes from the first terms in each parenthesis click and the second terms in each parenthisis clikcing. Similalrly we can find

$$\begin{aligned}
\langle\psi|A_2\tilde{A}_2|\psi\rangle &= \sum_n \langle\psi|\left(|2n+1\rangle\langle 2n| + |2n\rangle\langle 2n+1|\right)\otimes\left(|2\tilde{n}+1\rangle\langle 2\tilde{n}| + |2\tilde{n}\rangle\langle 2\tilde{n}+1|\right)|\psi\rangle \\
&= \sum_n \langle\psi|\left(|2n+1,2\tilde{n}+1\rangle\langle 2n,2\tilde{n}| + |2n,2\tilde{n}\rangle\langle 2n+1,2\tilde{n}+1|\right)|\psi\rangle \\
&= \sum_n 2e^{-\beta(2n+1)\omega/2}e^{-\beta(2n\omega)/2}(1-e^{-\beta\omega}) \\
&= \frac{2e^{-\beta\omega/2}}{1+e^{-\beta\omega}}
\end{aligned}$$

Where the second to last equality comes from the density matrix. When we insert the operator into the density matrix we pick upthe coeff of the odd and even term. Now lets define

$$B_1 = \cos\theta\tilde{A}_1 + \sin\theta\tilde{A}_1 \quad B_2 = \cos\theta\tilde{A}_1 - \sin\theta\tilde{A}_2 \Rightarrow B_1^2 = B_2^2 = 1 \text{ because } \left\{\tilde{A}_1,\tilde{A}_2\right\} = 0 \text{ and } \tilde{A}_1^2 + \tilde{A}_2^2 = 1$$

So now we can finally get our CSSH operator

$$C_{AB} = A_1(B_1+B_2) + A_2(B_1-B_2)$$

So we can compute

$$\boxed{\langle\psi|C_{AB}|\psi\rangle = 2\cos\theta\,\langle\psi|A_1\tilde{A}_1|\psi\rangle + 2\sin\theta\,\langle\psi|A_2\tilde{A}_2|\psi\rangle = 2\cos\theta + \frac{4e^{-\beta\omega/2}}{1+e^{-\beta\omega}}\sin\theta}$$

We can check that the boxed values is max at $\tan\theta = \frac{2e^{-\beta\omega/2}}{1+e^{-\beta\omega}}$. Using we can find that the maximum value is

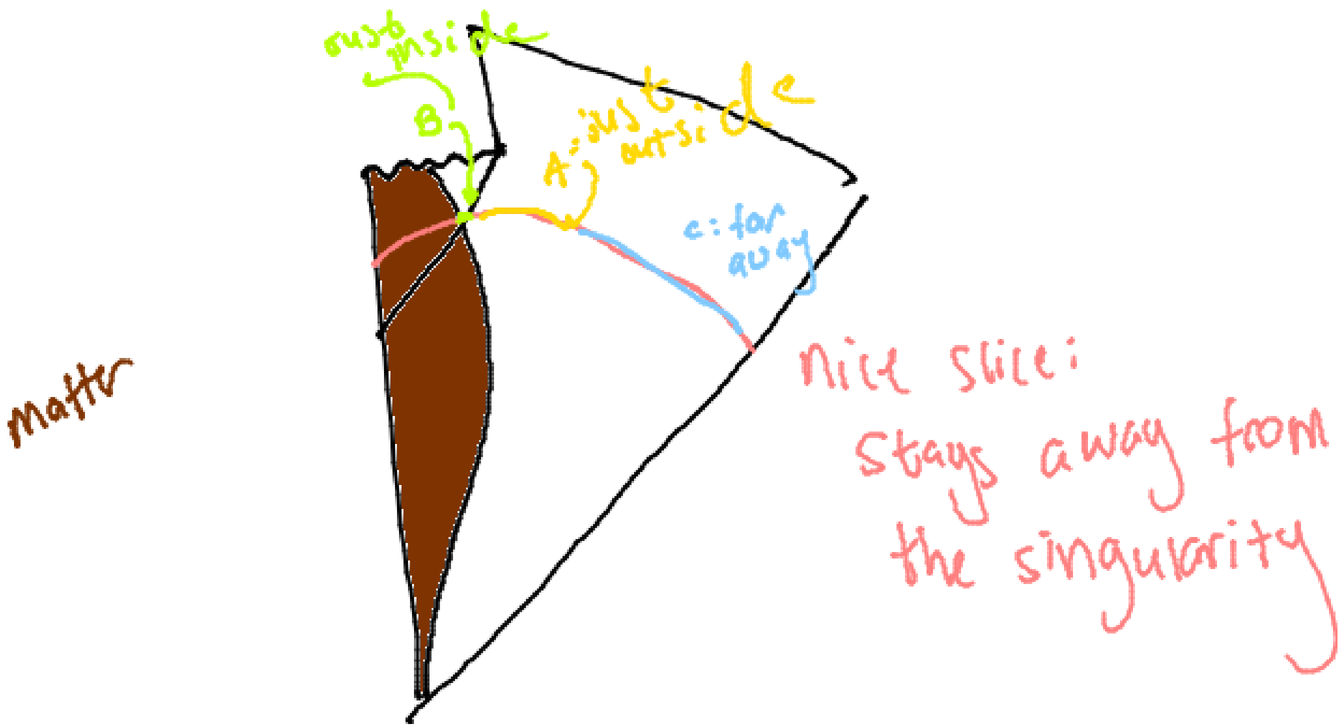
$$\max(\langle\psi|C_{AB}|\psi\rangle) = \frac{2}{1+e^{-\beta\omega}}(1+6e^{-\beta\omega}+e^{-2\beta\omega})^{\frac{1}{2}} > 2$$

This expression is always larger than 2. To see that this is true just square the red terms

$$(1+e^{-\beta\omega})^2 = 1+e^{-2\beta\omega}+2e^{-\beta\omega} < 1+6e^{-\beta\omega}+e^{-2\beta\omega}$$

So we did a long thing to prove a very simple result. We can define CHSH operators accross the horizon that violate the bell inequalities(e.g. they display entanglemnt.) This result is very robust. In some sense this is intuitive. However we've shown this int terms of simple correlators.

Now we need to prove part (b) of our parado



Now lets wave our hands **VERY** fast. We can start this hand-waving with a few assumption

(a) The dynamics can be formulated in an effective hilbert space $e^{S_{\text{bh}}}$.

(b) In this Hilbert space we can find

$$\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \subset \mathcal{H}$$

Namely the Hilbert space effectivley factorized (where A, B, and C) is defined in the picture above.

(c) The modes $a_{\omega,\ell}$ and $a^\dagger_{\omega,\ell}$ act within $\mathcal{H}_A$ and the $\tilde{a}$ operators act within $\mathcal{H}_B$.

(d) The early hawking radition is captured by $\mathcal{H}_C$.

Consider an old black hole. What old means is that the BH has more than half evaporated(entropy is half of S_{BH}.) What we'd like to say is that the effective dimension of $\mathcal{H}_C$ is somehow larger than the complement

$$\dim(\mathcal{H}_C) \gg e^{S_{\text{bh}}/2} \gg \dim(\overline{\mathcal{H}_C})$$

Now if we do this then this is the setup we had at the last lecture. We have some space that is much bigger than the other space. Thus A_1 and A_2 act on the complement o $\overline{\mathcal{H}_C}$. Thus we can find some C_1 and C_2 which has the property that if we define

$$C_{AC} = A_1(C_1 + C_2) + A_2(C_1 - C_2)$$

Then we'll be close to violating silresons?ss bound

$$\langle\psi|C_{AC}|\psi\rangle = 2\sqrt{2} - O(\epsilon)$$

But earlier we found that

$$\langle\psi|C_{AB}|\psi\rangle > 2$$

And we had a beautiful inequality

$$\langle\psi|C_{AB}|\psi\rangle^2 + \langle\psi|C_{AC}|\psi\rangle^2 \leq 8 \text{ violated :3c}$$

Lets emphasize a few points.

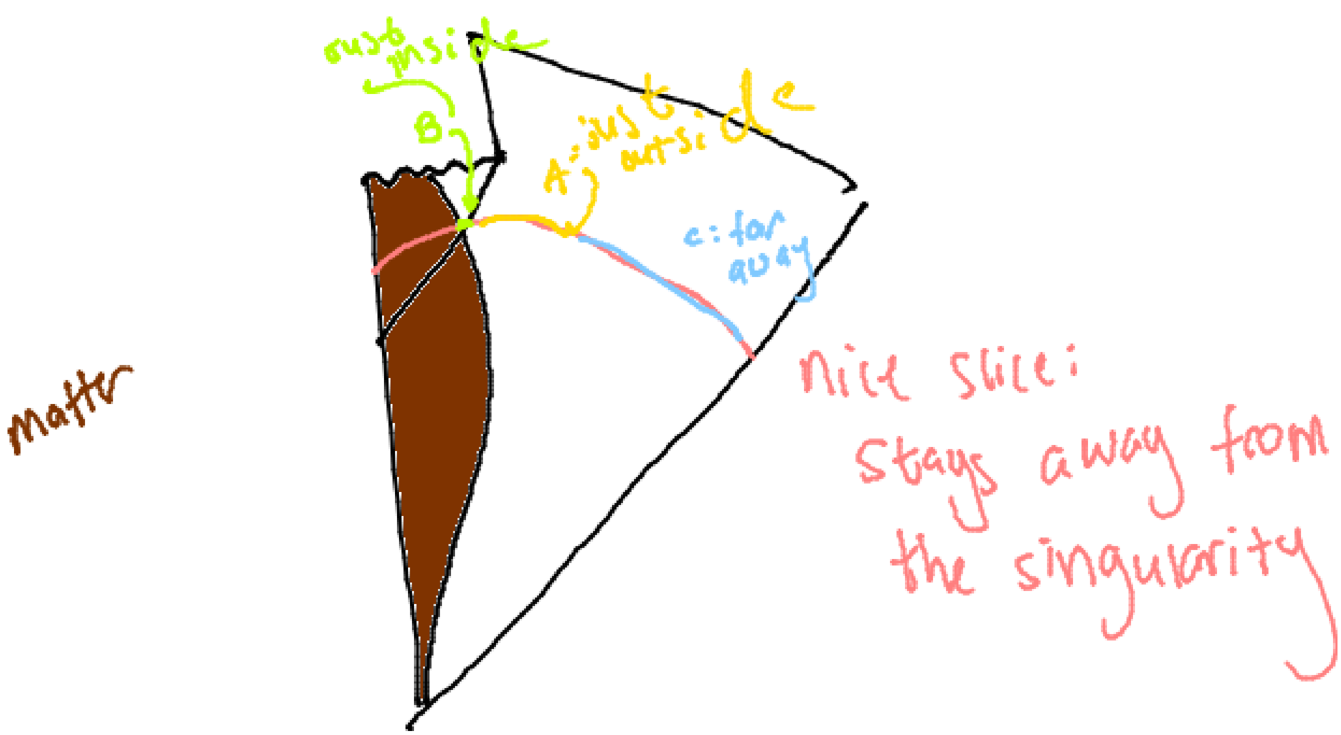
(a) This paradox is different to the hawking paradox because it makes reference ot the interior and not just to the exterior.

(b) This is an $O(1)$ violation.

Lecture 10:

February 11, 2021

Today we'll talk about the original formulation of the monogamy paradox by Marthur using the subadditivity of Von Neumann entropy. Recall the drawing we had last lecture



Also a quick note on notation. S_{IJ} is the entropy of the region $I \cup J$. The reason we didn't start with von neumann entropy to formulate this paradox is because in a QFT we still don't know how to define these quantities well and these quantities are UV divergent. So what we need to do is assume that our intuition holds for situations where we don't know how to define these quantities precisley. Lets try to show the paradox now

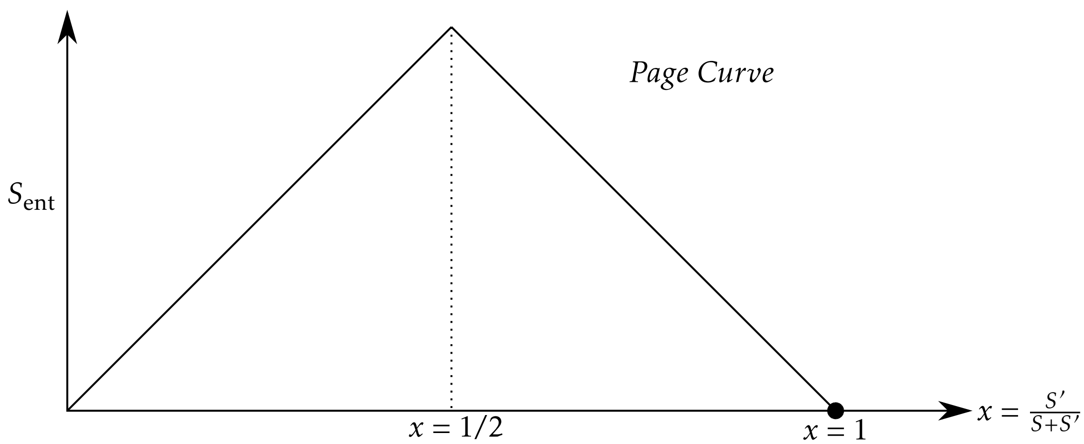
We'll first consider $A \cup B$. We want to show that

$$S_{AB} < S_A \quad S_{AB} < S_B$$

What this is saying is that S_{AB} is let mixed than either A or B by itself. We can restate this as asaying that the DOF in A and B are maximally engangled

$$S_{AB} \approx 0 \text{ and } S_A \approx S_B \Rightarrow A, B \text{ are maximally entangled}$$

It turns out to set up the paradox we only need the first assertion instead of the second strong assertion. The next thing we need to consider is the entropy of the region C. Uner our assumption of factorization of the Hilbert space, we expect that S_C will obey a page curve(see figure from lecture 8 below.)



Physically the page curve is saying that DOF from A is moving to C. This is actually really schematic, for on the x axis becomes times and $S'/S + S'$ isn't linear with time. The important aspect is that there's a rise then a fall. We'll call $x = 1/2$ the "page time." We'll consider time after the page time. At $t + \delta t$ some additional DOF move from region A to region C. So we could see that

$$C(t + \delta t) \approx AC(t)$$

$$S_C(t + \delta t) < S_C(t) \text{ for } t > t_{\text{page}} \Rightarrow S_{AC} < S_C$$

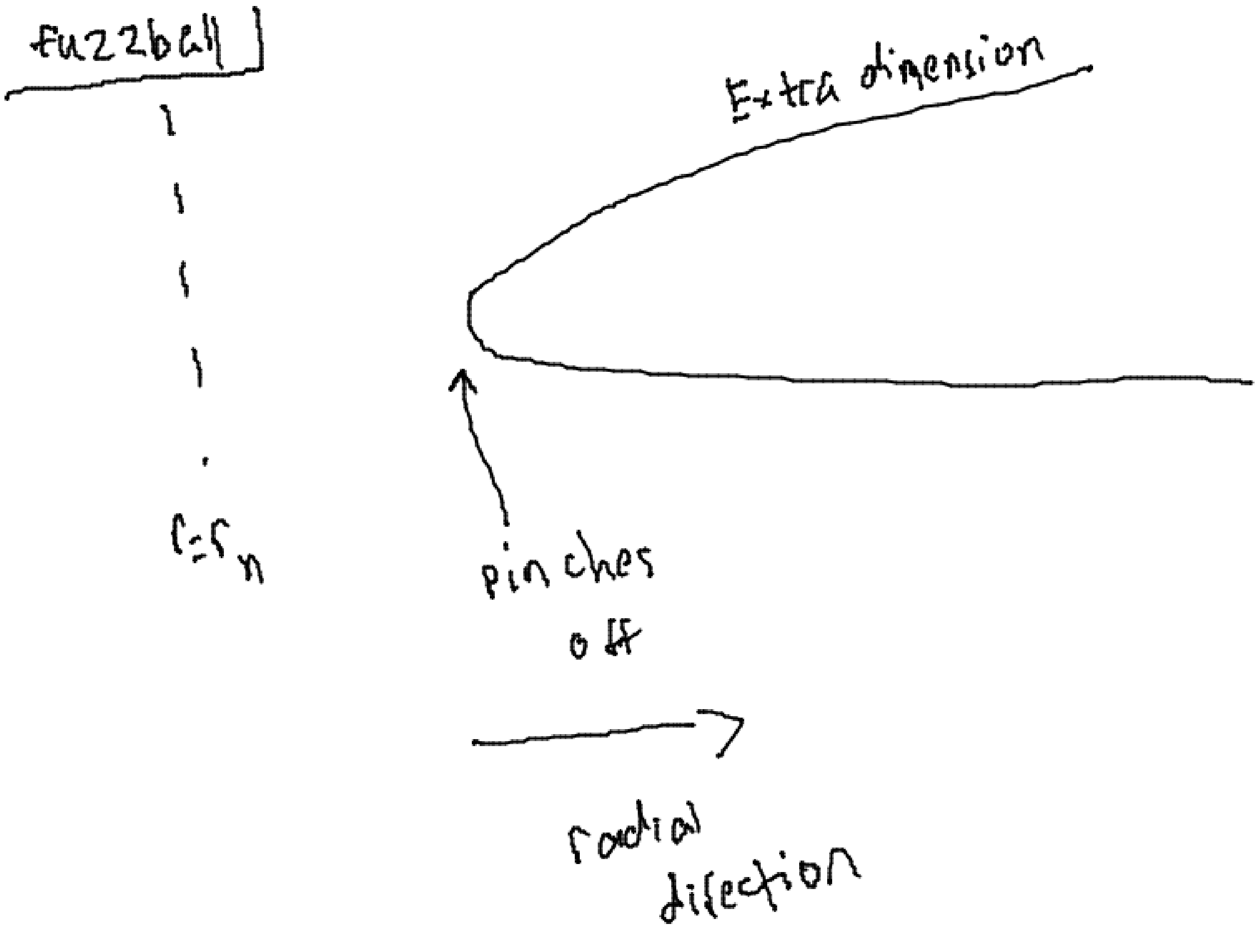
The Von Neumann entropy satisfies some inequalities

$$S_{AB} + S_{AC} \geq S_B + S_C$$

This is the strong subadditivity of Von Neumann entropy. What this is saying is that if AB are entangled then S_{AB} is small then S_{AC} cannot be too small. But here we're in trouble since

$S_{AB} < S_B$ and $S_AC < S_C$ after the page time. We'll go on to say a few more things about this paradox. Before we continue we should note that this was formulated in arxiv:0909:1038. The original paradox wasn't talking about regions but was talking about bit models. We've slightly generalized that here by talking about regions. So what we're going to do for the next hour is we're going to talk about how this paradox might be resolved.

The proposal of fuzzballs is that bhs evaporate like ordinary objects. But how does BH evaporate different from this computer in front of us? tHe difference is tthat ordinary obejcts have no interior partners for emitted radiation. If we were talking about a BHs this means BH has no interior. This anhilates the region B and thus gets rid of $S_{AB} < S_B$ and $S_{AB} < S_A$. So at late times A is only entangled with C.



Topological obstruction to collapse (huh?). Fuzzball complentanity

$$\text{observer with } E \gg T_H \text{ sees smooth interior}$$

$$\text{observer with } E \ll T_H \text{ see the fuzzball}$$

Now lets consider some metric operator $g_{\mu\nu}$ somewhere in the exterior. We proved a while ago

$$\langle \psi | g_{\mu\nu} | \psi \rangle = g_{\mu\nu}^{\text{universal}} + O\left(e^{-S/2}\right)$$

We could also argue that if we considered Euclidean quantum gravity then we could let

$$g_{\mu\nu}^{\text{universal}}(x) = g_{\mu\nu}^{\text{convential}}(x) \text{ outside}$$

Solutions by themselves are not enough by themselves to state that BH don't exsist.

I have no idea what's going on here

Next we'll talk about firewalls. AMPS paradox is the marthur paradox. We should distinguish this from AMPSSS paradoxes as well as the Marolf/Polchinski papers. The latter two apply to large (non evapoarating) AdS black holes. anywyas back to firewalls. They propose that

(a) Geometry all the way to the horizion (from theo utside) resembles the convential geometry.

(b) There is no interior

Lets say we try to compute $\langle T_{\mu\nu}\rangle$ at the horizon. To do this we need to compute

$$\lim_{x_1\to x_2}\langle \partial_\mu\phi(x_1)\partial_\gamma\phi(x_2)\rangle - \text{countertes}$$

Where do the counter terms come from? They come from (what?)

$$\lim_{x_1\to x_2}\langle\phi(x_1)\phi(x_2)\rangle = \frac{N}{|x_1-x_2|^{(d-1)/2}}$$

And if we have no interior then $T_{\mu\nu}$ diverges at the horizon. A person heading to a horizon will hit a literal firewall(ouch) just behind the horizon. What are the weaknesses of this argument? First there is no mechanism for producing or for stabilizing the firewall. Secondly this is a large violation of effective field theory because it says quantum effects reach all the way to the horizon.
The last resolution Suvrat talked about is the holography of information.

tl;dr: Fuzzball/fireweall says interior doesn't exists. holography of inromation: for a black hole, the fact that the hilbert space factorizes is not true and the commutators are large. This will resolve paradox?

Lecture 11: Hilbert Space in AdS

February 17, 2021

Today we'll talk about Hilbert space in AdS. First recall the metric of AdS

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

In global AdS we have a consistent IR regulator which we don't have in Poincare AdS. Here we've set the length scale $\ll_{=1}$. We'll talk about gravity and quantum gravity. When we talk about gravity the above will become an asymptotic definition. e.g.

$$ds^2 \xrightarrow{r\to\infty} -(1+r^2)dt^2 + \frac{dr^2}{1+r^2} + r^2 d\Omega^2$$

We'll also say that

$$N = \frac{\ell_{\text{ads}}}{\ell_{\text{plank}}} \gg 1 \text{ and also assume that N is larger than all other numbers}$$

This is the standard convention used in AdS/CFT. So this is some simple stuff about the geometry of AdS. First lets talk about simple scalar fields in AdS and then talk about how this generalizes. A minimally couled scalar field is goverend by the klein gordon equation

$$\left(\nabla^2 - m^2\right)\phi(t, r, \Omega) = 0$$

We're going to impose the boundary condition that $\phi \to 0$ as $r \to \infty$. This is a normalizable boundary condition. With these boundary conditions we can solve the klein gordon equation(we won't go through this exercise, it's just separation of variables.) The solution turns out to be

$$\phi = \sum_{n \geq 0} a_{n,\ell} c_{n,\ell} e^{-i(2n+\ell+\Delta)t} Y_\ell(\Omega) \chi_{n,\ell}(r) + \text{h.c.}$$

$$\text{where } \chi_{n,\ell}(r) = r^\ell (r^2+1)^{-(\Delta+2n+\ell)/2} \times 2F_1(-n, -\Delta - n + d/2, d/2 + \ell, -r^2)$$

Where $c_{n,\ell}$ and $a_{n,\ell}$ are normalization. Something to note is that as $r \to \infty$ we have

$$\phi = \frac{1}{r^\Delta}$$

Where $\Delta = d/2 + \sqrt{(d/2)^2 + m^2}$. There's a second interesting aspect to these solutions oscillate in some way and are a quantized set of solutions. The reason they're quantized is because AdS is like a box. We also require that the the field dies at the box. And thus there are a fixed number of wavelengths that can oscillate in this boundary. Therefore things are quantized. We'll also find that the $a_{n,\ell}$ obey the canonical commutation relations. It's usually convinient to talk about the asymptotic value of the field

$$O(t, \Omega) = \lim_{r \to \infty} r^\Delta \phi(r, t, \Omega)$$

In AdS/CFT this operator is the operator that's dual to the bulk field ϕ. There are some interesting relations between the modes of this operator and the modes of the (something?). The way we'd write down the modes of this operator is by doing a fourier transform

$$O_{n,\ell} = \int d\Omega \int dt\ O(t, \Omega) e^{i(2n+\ell+\Delta)t} Y_\ell^*(\Omega) = \sqrt{G_{n,\ell}} a_{n,\ell}$$

$\sqrt{G_{n,\ell}}$ is just some factor. It's interesting because it's telling us the commutator of $O_{n,\ell}$ and $O^\dagger_{n,\ell}$ (which we can find with $[a_{n,\ell}, a^\dagger_{n',\ell'}] = \delta_{nn'}\delta_{\ell\ell'}$. So we have these mdoes, these creation and anhilation operators. Lets construct the Hilbert space in perturbatiuon theory.

$$a_{n,\ell}|0\rangle = 0 \forall n, \ell$$

We can now define excitations to create a very simple fock space

$$\ldots a^\dagger_{n_2,\ell_2} a^\dagger_{n_1,\ell_1}|0\rangle$$

Okay now lets try to generalize this Hilbert space. So these modes are were generated by considering the a and a daggers. But we could have done this in another way. We could've taken our boundary field and taken out the boundary operator

$$O^\dagger_{n,\ell} = \int O(t, \Omega) e^{-i(2n+\ell+D)t} Y_\ell^*(\Omega) dt d\Omega^{d-1}|0\rangle$$

This generates (in a free field theory) a state that is unnormalized and parallel to $a^\dagger_{n,\ell}$. We can do this for other operators

$$O^\dagger_{n_i,\ell_i} = \int O(t_i,\Omega_i)e^{-i(2n_i+\ell_i+D)t_i}Y^*_\ell(\Omega_i)dt_i d\Omega_i^{d-1}|0\rangle$$

And in this way we can create a fock space. So what we've found is that we can describe this hilbert space and describe them in two different ways. The first was with creation and anhilation and the second is using boundary operators

$$\text{span}\{O(f_1)\dots O(f_n)|0\rangle\} \text{ where } O(f) = \int O(t,\Omega)f(t,\Omega)dtd^{d-1}\Omega$$

So what's the big deal? The reason is the following. This generalizes nicely to the interacting theory and gravity.

Interacting Theory: We have some vacuum that's defined by $H|0\rangle = 0$. Now consider the set of operator that are $O(t_1)\dots O(t_n)|0\rangle$. Notice here we dropped the Ω. They're not important and it's really annoying to write it out so we're just not going to anymore. Anyways think of an operator of this kind. If we wanted a state that was normalizbale we'd need to smear things out a little bit. Anyways what happens to this state whne we evolve it in time? What does this state go to? If we evovlve this state in time we'll find this remarkable fact

$$e^{-iH\tau}O(t_1)\dots O(t_n)|0\rangle = e^{-iH\tau}O(t_1)e^{iH\tau}e^{-iH\tau}\dots O(t_n)e^{iH\tau}e^{-iH\tau}|0\rangle$$

We've just inserted a bunch of ones. But we've just sandwhiched operators inside unitary transforms so we're evolving each operator by $-\tau$. This means all together we have

$$e^{-iH\tau}O(t_1)\dots O(t_n)|0\rangle = O(t_1-\tau)\dots O(t_n-\tau)|0\rangle = O(t_1')\dots O(t_n')|0\rangle$$

The final exponential just does an arbitaray phase shift on the vacuum which we don't care about. This turns out to also be true in a theory of gravity. We have a notion of asympotic states that are of the same form under time evolution. Namely

$$\text{span}\{O(f_1)\dots O(f_n)|0\rangle\}$$

forms a superselection sector[2]. they time evolve states into themselves. Something we should be worried about is why this is a good description in gravity. Well

(a) asympottic operators make sense

(b) asympototic time-translations make sense

But where did all those branes, D-branes, and other things that sound too big brain for me go? Well the point is in a non-perturbative theory of gravity all these weird branes and things could be formed by our very simple span of boundary operators. By the way, I don't know if I wrote this down but the physical picture we should have of the $O(f)$ is that we're throwing things into the space at the boundary. Something we should say is that the Hilbert space we've written down is an overcomplete hilbert space. Specifically if we consider

$$O(f_1)\dots O(f_n)|0\rangle \text{ where } f_1 = e^{-i(2n+\ell+\Delta)t},\ t\in[0,\pi] \quad f_2 = \text{same thing with} t\in[\pi,2\pi]$$

[2] sounds like where all the superhero aparatments are in a homebrew mutant and monsters game

We see that $O(f_1)|0\rangle = O(f_2)|0\rangle$. Now lets describe how we can squeeze this Hilbert space. So far we've talked about smearing. Suvrat claims that we can create the same space with

$$O(\tilde{f}_1)\ldots O(\tilde{f}_n)|0\rangle$$

Where $\tilde{f}$ have support only in $t \in [0,\epsilon]$. We claim that this space is dense in the original Hilbert space. Earlier we said that the physical interpretation we where throwing in some stuff and we got some set of states. Now we're saying that we can get the same set of states by throwing in all of our stuff in some small interval ϵ. We will now prove that this result is true in QFT nd in gravity. Something to note is that this is not the principle of holography of information[3]. Lets start by thinking about single particle state

$$|\psi\rangle = O(f)|0\rangle \text{ where } f \text{ has arbitrary support(as opposed to } \tilde{f} \text{ in some small time)}$$

Now supose that this is orthogonal to all states $O(\tilde{f})|0\rangle$. Then $|\psi\rangle$ has the property

$$\langle\psi|O(\tau)|0\rangle = 0,\ \forall\tau\in[0,\epsilon]$$

We can insert a complete set of eigenstates above to get

$$\langle\psi|O(\tau)|0\rangle = \sum_E \langle\psi|E\rangle\langle E|O(\tau)|0\rangle = \sum_E \langle\psi|E\rangle\langle E|O(0)|0\rangle e^{iE\tau}$$

notice that this expression has a nice property. It's analytically when τ is extended in the upper half plane. This is because energy is positive and when we extend τ to the upper hlaf plane something happens? τ doesn't have any singularities in the upper half plane[4]. So we have an analytic function in the upper half plane and vanishes in some small region $[0,\epsilon]$. Since this function vanishes we know that in the real plane everything vanishes. What this tells us is

$$\langle\psi|O(t)|0\rangle \text{ vanishes for all } \tau$$

However this is a contradicition since

$$|\psi\rangle = O(f)|0\rangle \Rightarrow \int \langle\psi|O(f)|0\rangle f(t)dt = \langle\psi|\psi\rangle$$

So any state that can created from smearing over all time can be creatd by smearing over some small region. What we've said is suppose there was an element of span $\{O(f)|0\rangle\}$ which could not be approximated arbitrarily well by $O(\tilde{f})|0\rangle$ we could find some g such that $O(g)|0\rangle$ is orthogonal to $O(\tilde{f})|0\rangle$. Lets try to prove the multiparticle version of this. We wnat

$$O(f_1)\ldots O(f_n)|0\rangle$$

We can get the "same space" through $O(\tilde{f}_1)\ldots O(\tilde{f}_n)|0\rangle$. Lets do this by contradidction again. Suppose theres a $|\psi\rangle$ that orthogonal to the span of $\tilde{f}$ stuff. This means that

$$\langle\psi|O(\tau_1)\ldots O(\tau_n)|0\rangle = 0\ \forall\tau_1,\ldots,\tau_n \in [0,\epsilon$$

[3]oh wait was I supposed to think that it was?

[4]TODO: revisit

Now lets insert a bunch of energy eignetates like we did before

$$\langle\psi|E_1\rangle\langle E_1|O(\tau_1)|E_2\rangle\langle E_2|O(\tau_2)|E_3\rangle\ldots\langle E_n|O(\tau_n)|0\rangle = 0$$

But now we can do the same trick

$$= \sum_E \langle\psi|E_1\rangle\langle E_1|O(0)|E_2\rangle\ldots\langle E_n|O(0)|0\rangle e^{iE_1\tau_1}e^{iE_2(\tau_2-\tau_2)}\ldots e^{iE_n(\tau_n-\tau_{n-1})}$$

And if we do a change of variables to $z_i = \tau_i - \tau_{i-1}$ where $\tau_0 = 0$. And this is analytical when z_i are extended in the upper half plane. Now here we again invoke a theorem in complex analysis "edge of the wedge" theorem. We use this theorem to say that we now have

$$\langle\psi|O(\tau_1)\ldots O(\tau_n)|0\rangle = 0,\ \tau\in[0,\epsilon] \to \langle\psi|O(t)\ldots O(t)|0\rangle = 0$$

The exact same contradiction we had before!

Something to note is that this is an trivial example of the "Reeh Schlieder" theorem. It's "trivial" because we consider the entire boundary between time $0 \to \epsilon$. In general we use the Reeh Schlieder theorem for an arbitrary set. What we've proven is also true for gravity (which isn't true in general for Reeh Schlier theorems) since we only used the positivy of energy. In general Reeh Schlieder uses the psecturm something.

Lecture 12: Holography of Information

February 22, 2021

Okay lets get started[5].

$$\text{Let } A_\epsilon = \text{"algebra" of operators in } [0,\epsilon]$$

When we say algebra we basically means a vector space but with operators $O(f_i)$. We proved in the last lecture is that

$$A_\epsilon|0\rangle \text{ is dense in the full Hilbert space}$$

Note that when we say full hilbert space we're talking about a dynamically closed superselection sector. Lets get some perspective

(a) free-field theory: what we proved last time holds for free-field theory(as well as many more theories.) Consider some state

$$|n,\ell\rangle = a^\dagger_{n,\ell}|0\rangle$$

Can we createt his operator by acting on a smaller time band in a vacuum? We can actually check this numericaly. Take $\tilde{f}_q$ to be a basis of functions with support in $[0,\epsilon]$(a basis for A_ϵ?). Also note that we can take this as a countable basis. Now we want

$$\text{Minimize } \left| |n,\ell\rangle - \sum_{q=1}^{q_{\max}} O(\tilde{f}_q)C_q|0\rangle \right|^2$$

Given some $\tilde{f}_q$ we can numerically minimize wrt C_q. We can call this minimum (the residue) $r_{q_{\max}}$. We can then show that $r_{q_{\max}}$ is monotically decreasing with $q_{\max}$. Suvrat uploaded some mathematica code to see this.

[5]watching this lecture a bit later than the actual lecture date 2021-02-18 since Texas was frozen during the normal lecture time :(

(b) AdS/CFT: Consider an operator $\mathcal{O}$ that is a light primiary operator. Now lets consider

$$A_\epsilon |0\rangle \approx \{\mathcal{O}(\tau_1)\ldots\mathcal{O}(\tau_n)|0\rangle\}$$

Products of $\mathcal{O}$ generate higher primariers and descendants. Action of higher primariers and descendants generates the hilbert space(something about Polchinski's string theory books here? *sweating*)

Something to note about $A_\epsilon|0\rangle$ generating H is that it can be physcially interpreted as telling us about entanglement in the vacuum. Lets go away from all this CFt and AdS and go back to quantum mechanics. Consider an EPR pair

$$|\mathrm{EPR}\rangle = a|00\rangle + b|11\rangle$$

Here we have $\dim(H) = 4$. But notice if we act on the first qubit with an arbitrary hermmitian operator then we can generate any state in the Hilbert space. Lets prove this.

$$(1+\sigma^3)|\mathrm{EPR}\rangle = 2a|00\rangle \quad (\sigma^2+i\sigma^y)|\mathrm{EPR}\rangle = 2a|10\rangle \quad (1-\sigma^3)|\mathrm{EPR}\rangle = 2b|11\rangle \quad (\sigma^x-i\sigma^y)|\mathrm{EPR}\rangle = 2b|01\rangle$$

Where the σ only act on the first qubit. Since we've managed to get a basis we've found that we can generate any state in the Hilbert space. Notice that this doesn't mean we can send messages bewteen Earth and Mars since there involves no loss of locality since we cannot "act" with Hermitian operators. How would you act on a state wit h $(1+\sigma^z)$? It's not a physcially allowed transformation of state. The most generally allowed allowed transformations of states are acting with unitary operators. Furthermore we cannot use this result to measure the second qubit. E.g. consider x_1 some hermitiation on the first qubit and U_2 a unitary on the second qubit

$$\langle\mathrm{EPR}|U_2 x_1 U_2^\dagger|\mathrm{EPR}\rangle = \langle\mathrm{EPR}|x_1|\mathrm{EPR}\rangle$$

Now what we'll prove the *holography of information* which is only for gravity. A copy of all information on the bulk of a Cauchy slice is available near the boundary. We'll now make this precise. First what does *near the boundary* mean? It is illustrated in Figure 7. More formally say there are two states $|\psi\rangle \neq |\phi\rangle$ we know that $\exists x \in A_\epsilon$ such that $\langle\psi|x|\psi\rangle \neq |\phi|x|\phi\rangle$. Equivalently all operators $H \to H$ can be approximated arbitrarily well by operators in A_ϵ. This result is trivial from the perspective of AdS/CFT but we are not invoking AdS/CFT so this works also for black holes in flat spacetime. So what is the thing that is different about gravity as opposed to QFT? The main physical input will be kinda trivial: energy in gravity can be measured from infinity(asymptotic region. Lets get some perspective

(a) Gauss's law. We have an equivalent statement in AdS. First we need boundary conditions

$$g_{\mu\nu} = g_{\mu\nu}^{\mathrm{AdS}} + h_{\mu\nu}$$

We choose a Fefferman Graham gauge

$$h_{r\mu} = 0 \text{ and } h_{ij} \text{ where } i,j \neq r \xrightarrow{r\to\infty} \frac{1}{r^{d-2}}$$

There's a version of the Gauss law in AdS with these kind of boundary conditions is

$$H = \frac{d}{16\pi G}\lim_{r\to\infty}\int d^{d-1}\Omega\ r^{d-2}h_t t$$

This is an particular gauge. To see this in different gauges look at hep-th/9902121 and gr-gc/9501014

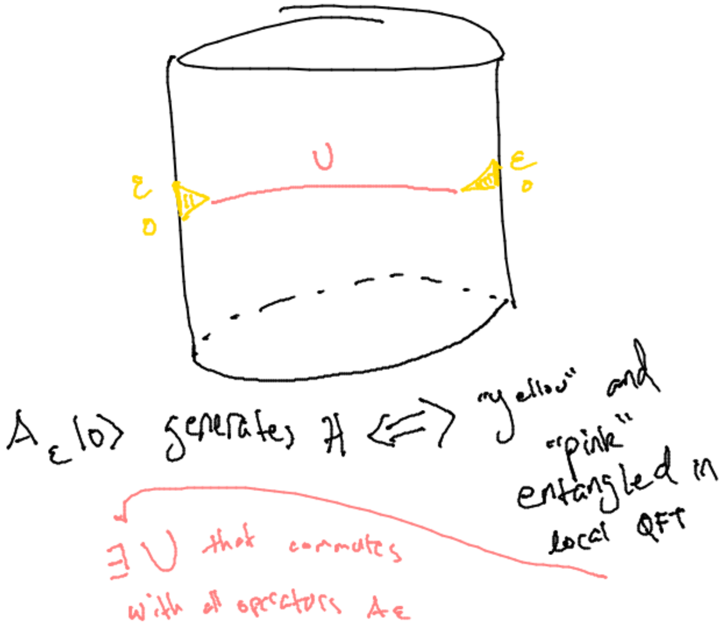


Figure 6: Intepretation of $A_\epsilon |0\rangle$ generating Hilbert space in AdS/CFG

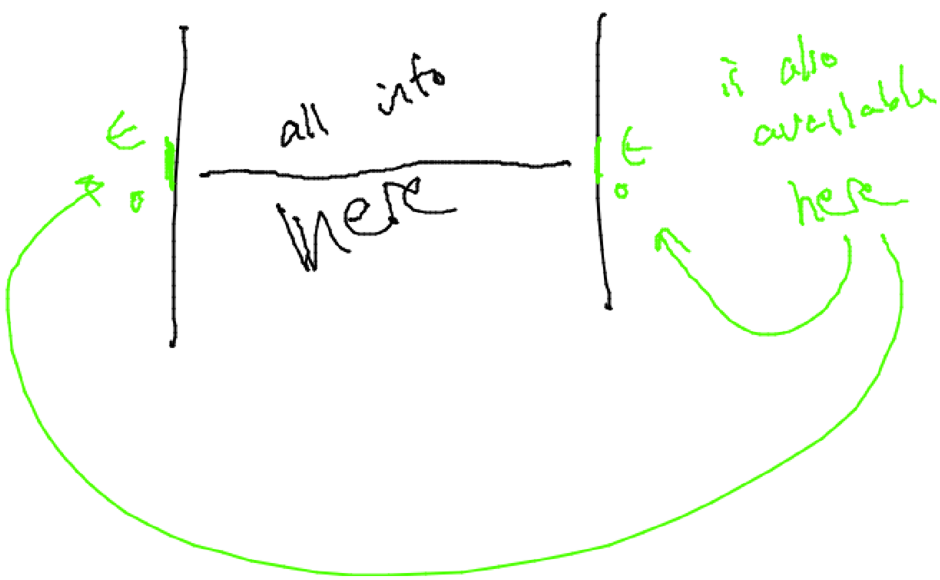


Figure 7: Definition of "near the boundary" for holography of information

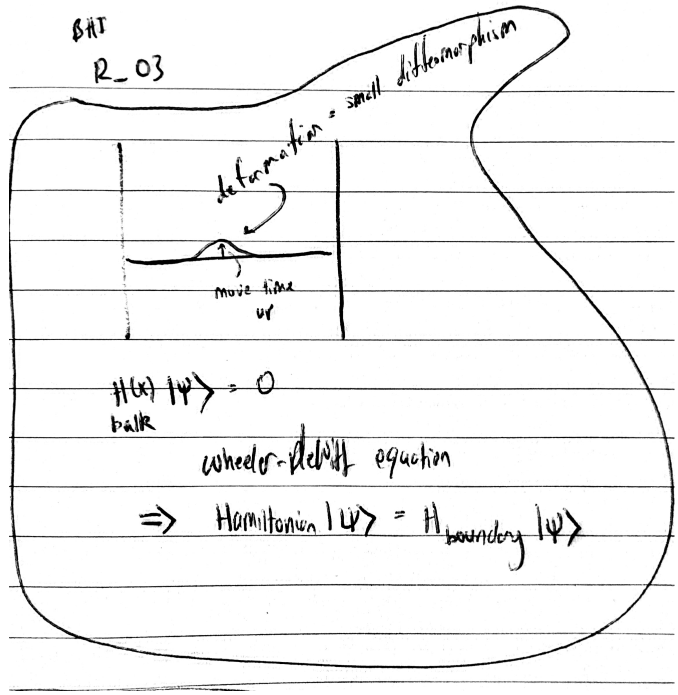


Figure 8: A perspective on the boundary version of gauss' law in AdS

(b) See Figure 8

(c) The third way is how we'd look at things in AdS/CFT. We have an extrapolate dictionary boundary limits of bulk operators are boundary operators.

$$\lim_{r\to\infty} r^{d-2} h_0 0 \propto T_{tt} \text{ the boundary stress tensor}$$

This means that

$$\lim_{r\to\infty} \int h_{00} r^{d-2} d^{d-1}\Omega = H$$

Energy from infinity in gravity.

Now lets say one more thing. Say we've measured the energy at infinity. Now consider Figure 9. Putting everything together. Say we have $|\psi_1\rangle \neq |\psi_2\rangle$. There must be some x so that

$$\langle\psi|x|\psi\rangle \neq \langle\phi|x|\phi\rangle$$

Choose some basis for the hilbert space $\{|n\rangle\}$. This x can be expanded

$$x = \sum_{n,m} c_{n,m} |n\rangle\langle m|$$

Now invoke Rich-Lieder theorem so that $x_n, x_m \in A_\epsilon$ so that $|n\rangle = x_n|0\rangle$ and $|m\rangle = x_m|0\rangle$. Thus we can write

$$x = \sum_{n,m} c_{n,m} x_n |0\rangle\langle 0| x_m^\dagger$$

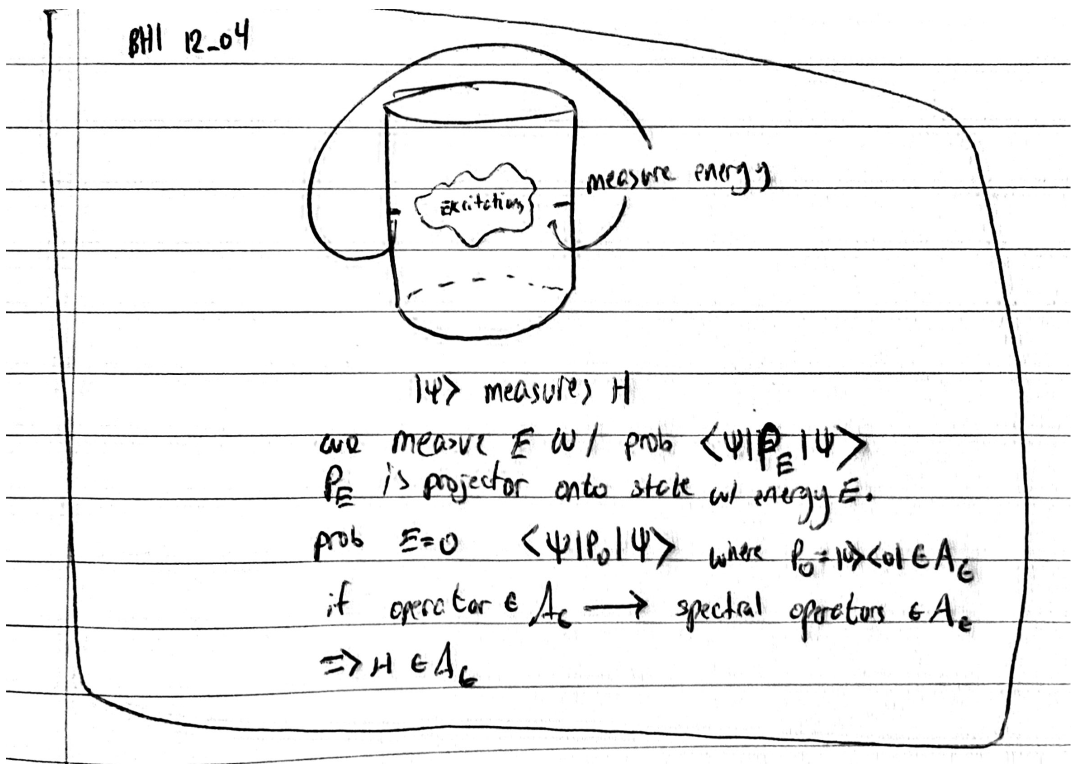


Figure 9: A proof that the projector operator onto $E = 0$ is a asympotic operator. Formally $P_0 \in A_\epsilon$

So far everything here works in a general QFT. But from here we'll assert a special property of gravity. In gravity $P_0 \in A_\epsilon$ meaning that

$$x = \sum_{n,m} c_{n,m} x_n P_0 x_m^\dagger \text{ is a linear combination of products of elements in } A_\epsilon$$

Since this is an algebra (closed under linear combinations and products) we see that only in gravity $x \in A_\epsilon$. *Thus we have just proved that there exists $x \in A_\epsilon$ that allows us to distinguish states in the bulk.* We'll discuss this as well next time.

Lecture 13: Low-energy AdS tests of Holography of Information

March 08, 2021

Lets start with a review. Last lecture we proved that in a theory of gravity in a state is available in some small time band on the boundary. This is not true in a QFT. From the point of view from AdS/CFT the result should be obvious. However here proved the result without ever invoking AdS/CFT. What we will do today is bring things down to Earth.

Lets go to global AdS and we have some observers (a group of very strong astrophysicist) on the boundary who have detectors which only work in $t \in [0, \epsilon]$. What we want to ask is that can the astrophyscist determine the state in the bulk? First we'll make some simplifications. The state is a low-energy excitations (e.g. it's not a black holes.) Instead it's maybe some gravitons or some other low-energy excitations. To make this precise we need to demand that our state $|g\rangle$ = global state that the boys want to characterize satisfies

$$1 - |P_{E<\Lambda}|g\rangle|^2 \ll 1$$

The Λ is some uv cutoff and P is a projector of energy less than Λ. Lets also consider what the physicist have access to. First we assume that they have textbook QM abilities. Consider the stern-gerlach experiment. Rotating particles by turning on a magnetic field corresponds to a unitary transformations. Furthermore we have measurements.

(a) if x is a low-energy opertaor near the boundary (at boundary in $t \in [0,\epsilon]$) then we allow the observers to act with $U = e^{iJx}$ a unitary. In other word we allow $|g\rangle \to e^{iJx}|g\rangle$.

(b) We'll allow these poor souls to measure the energy using gravitational effects. They can't go into the bulk but they can measure of the energy at the boundary.

$$\int h_{00} d^{d-1}\Omega$$

Lets start with a warmup task. Can the physicist determine if $|g\rangle = |0\rangle$ or not. The first thing to recognize is that in a LQFT this simple task is not possible. Consider a local QFT. All correlators on the border are unchanged in a LQFT

$$\langle 0|U^\dagger x U|0\rangle = \langle 0|x|0\rangle \text{ in a LQFT}$$

Now lets explain how this is simple to do in gravity Measure the energy and determine probability that $E = 0$.

$$\langle g|P_{E=0}|g\rangle \leftarrow \text{ Born rule}$$

But remember in global AdS the vacuum projector is unique $P_0 = |0\rangle\langle 0|$.

$$\Rightarrow \langle g|P_0|g\rangle = |\langle 0|g\rangle|^2 = \begin{cases} 1 & |g\rangle = 0 \\ \text{otherwise} & |g\rangle \neq 0 \end{cases}$$

Using this we can distinugish between $|0\rangle$ and $U|0\rangle$.

Consider another warmup task. Let x be a low-energy hermitian operator near the boundary[6]. we'll denote $|x\rangle = x|0\rangle$. What we want to determine if $|g\rangle = |x\rangle$ or not. This is different from the previous example since $|x\rangle$ is not characterized by a conserved charge. In a LQFT this is impossible. In garvity we can do this in two steps

(a) Act with e^{iJx}

(b) Measure energy and dtermine probability of getting 0.

[6]near boundary in this lecture means at boundary in $t \in [0,\epsilon]$. Might be good to reiterate that here.

Why does this work? Well in the first step we have

$$|g\rangle \to e^{iJx}|g\rangle$$

And then in step two measuring the energy gives us

$$\langle g|e^{-iJx}P_0 e^{iJx}|g\rangle$$

To make a sipmlification we assume that $\langle g|0\rangle = 0$ (e.g. it's orthogonal the environment.) From here it's some simple algebra.

$$\langle g|e^{-iJx}P_0 e^{iJx}|g\rangle = \langle g|\left(1 - iJx + \frac{J^2x^2}{2}\right)|0\rangle\langle 0|\left(1 + iJx + \frac{J^2x^2}{2}\right)|g\rangle + O(J^3)$$

Now what happens to each of these terms. First we have $\langle g|0\rangle = 0$ and thus the $O(1)$ term vanishes. Similarly $O(J)$ terms vanish on examination

$$O(J) : -i\langle g|Jx|0\rangle\langle 0|g\rangle + \text{h.c.} = 0$$

So the first non vanishing term is the $O(J^2)$ term. on examination we have

$$O(J^2) : \langle g|Jx|0\rangle\langle 0|Jx|g\rangle = J^2|\langle g|x\rangle|^2 + O(J^3)$$

Now after this warmup we're almost done(just like my workouts!)

some complications discussed here

Right now we have measure

$$|\langle g|x\rangle|^2$$

From here we need to determine the phase. Now we need to pcik our favorite operator x_r and declare that $\langle g|x_r\rangle$ is real. And now that we have done this we note that $x + x_r$ is also hermitian so using the protocol we can also determine

$$|\langle g|x\rangle + \langle g|x_r\rangle|^2 = |\langle g|x\rangle|^2 + \langle g|x_r\rangle^2 + 2\langle g|x_r\rangle \operatorname{Re}(\langle g|x\rangle)$$

Thus we can determine the real part of $\langle g|x\rangle$. Using this as well as $|\langle g|x\rangle|^2$ we can almost get the imaginary part of $\operatorname{Im}\langle g|x\rangle$ up to a sign. This sign can actually be found in 2008.01740.

We are now going to shfit gear and talk about flat space. First we'll specialize to $D = 4$ and consider

$$ds^2 \xrightarrow{r\to\infty} -du^2 - 2dudr + r^2\gamma_{AB}d\Omega^A d\Omega^B \text{ where } \gamma_{AB} \text{ is a round metric on 2-sphere}$$

In flat space to make specifying boundary conditions easier we will use the gauge choice

$$g_{rr} = 0 = g_{rA} \quad \partial_r\left(\frac{\det(g_{AB})}{r^2}\right) \text{ this gauge is called Bondi gauge}$$

In this gauge the boundary we will choose are

$$g_{uu} \to -1 + O\left(\frac{1}{r}\right) \quad g_{ur} \to -1 + O\left(\frac{1}{r^2}\right) \quad g_{uA} \to O(1) \quad g_{AB} = r^2\gamma_{AB} + O(\gamma)$$

Meaning that we will not fluctuations that are so large that they change the leading term meaning that the asymptoic structure is not destroyed.

What are spacetimes taht are consistent with these conditions? One example is

$$ds^2 = -du^2 - 2dudr + r^2\gamma_{AB}d\Omega^A d\Omega^B + \frac{2m}{r}du^2 + \gamma C_{AB}d\Omega^a d\Omega^B + D^B C_{AB} du d\Omega^A$$

Lets talk about the components of this metric

(a) C_{AB} is called the "sheare" and must be symmetric and traceless $\gamma^{AB}C_{AB} = 0$ and has 2 independent componenets. These two componenets are very physical and contain information about physical components of the graviton, specifically the 2-polarizations of outgoing radiation. It might be more convinient to talk about $N_{AB} = \partial_u C_{AB}$ which is the Bondi News. Both of these are function of u and Ω.

(b) m is called the Bondi mass aspect which can be integrated

$$(\text{Bondi Mass}) = M(u) = \int \sqrt{\gamma}\, m(u,\Omega) d^2\Omega$$

What the Bondi mass is telling us is that how much mass remains to be radiated or how much mass remains inside. We find that

$$\lim_{u\to-\infty} M(u) = H \text{ in the ADM Hamiltonian}$$

There are similar boundary condiitons you can impose on other fields. Suppose you have some bulk scalar field

$$\phi(r,u,\Omega) \xrightarrow{r\to\infty} \frac{1}{r} \underbrace{O(u,\Omega)}_{boundaryvalue}$$

In a more relaistic theory we'll also have massive fields but we will not be able to discuss massive fields since massive fields die off faster than any polynomial power of r as we approach i^+. Since we're only consideirg black holes this doesn't matter too much. Consider a very large BH then most of its radiation is in massless particles $T \propto \frac{1}{r}$ and thus a BH basically does not emit massive particles at all.

Lecture 14: Holography of Information in Flat Space

March 17, 2021

Today we have some good news and bad news about holography of information in asymptotically flat space. Good news: operator algebra at scri+ is simpler than in AdS

$$[N_{AB}(u,\Omega), N_{CD}(u',\Omega')] = i16\pi G\partial_\mu\delta(u-u')\frac{1}{\sqrt{\gamma}}\delta(\Omega-\Omega')[\gamma_{A(C}\gamma_{D)B} - 1/2\gamma_{AB}\gamma_{CD}]$$

This is called "extremley simple". I guess in relation to other operator relations. We can fourier transform these operators to get

$$N^{\pm}_{AB}(\omega,\Omega) = \int e^{i\omega u} N_{AB}(u,\Omega) du$$

Now given the commutation relations we have previously what are the commutations of the momentum space operators

$$[N^+, N^-] = i\frac{16\pi G}{\sqrt{\gamma}}\delta(\Omega-\Omega')[\gamma_{A(C}\gamma_{D)B} - \gamma_{AB}\gamma_{CD}/2]\times\int e^{i\omega u}e^{-i\omega' u'}\partial_u\delta(u-u')dudu'$$

We can integrate by parts to evaluate the above integral. It should be clear that the above evaluates to

$$[N^+, N^-] = \frac{16\pi G}{\sqrt{\gamma}}\delta(\Omega-\Omega')[\gamma_{A(C}\gamma_{D)B} - \gamma_{AB}\gamma_{CD}/2]\times\omega 2\pi\delta(\omega-\omega')$$

Now with the operator algebra we can try to construct the hilbert space of the theory. The moment we get commutation relations we can construct the hilbert space. The first thing is that there should exists a $|0\rangle$ vaccum such that

$$N^+_{AB}(\omega,\Omega)|0\rangle = 0 \text{ where } \omega > 0$$

There will be some problems here but

$$N^-_{AB}(\omega,\Omega)|0\rangle \in (\text{excited state})$$

We can also define the smeared version of the N operator with

$$N(f) = \int \sqrt{\gamma}N_{AB}(u,\Omega)f^{AB}(u,\Omega)dud^2\Omega$$

From this we can generate a fock space

$$\text{span}\{N(f_1)\dots N(f_n)|0\rangle\}$$

Last time we shoed that we can also have scalar fields in this theory. We can also smear these in the same way and then define the fock space as

$$\text{span}\{O(h_1)\dots O(h_n)N(f_1)\dots N(f_m)|0\rangle\}$$

And this is all massless excitations that can be built on top of $|0\rangle$. In the context of the S what we've found is the out space and the S matrix is the map between "in" and "out" space. Since we're worring about information theoretic aspects we dont need to worry about S.

Now lets talk about the "bad news." the vacuum we talked about above $|0\rangle$ is a infinitely degenerate vacuum. The distinct vacua are described by "supertranslation charges."

$$Q_{\ell,m} = \frac{1}{4\pi G}\int \sqrt{\gamma}m(u\to-\infty,\Omega)Y_{\ell m}(\Omega)d^2\Omega$$

Why is this called the "supertranslation charges?" Well first we know that Q_{00} is the hamiltonian which generates translations in u. $Q_{\ell,m}$ are in a sense "angle dependent" translations of u.

In the context of this supertranslation we should specify vacuums by their supertranslation chage

$$Q_{\ell,m}|\{s\}\rangle = S_{\ell,m}|\{s\}\rangle$$

We can decompose the supertranslation charge into

$$Q_{\ell m} = Q^{\text{hard}}_{\ell,m} + Q^{\text{soft}}_{\ell,m}$$

Where hard charges are carried by excitations and soft charges are cairred by vacuums. However in this theory we have a special feature that ther are no hard charges. We can normalize these vacua

$$\langle\{s\}|\{s'\}\rangle = \pi\delta(s-s')$$

We can then define a fock space for each vacuum

$$H_s = \text{span}\{O(h_1)\dots O(h_m)N(f_1)\dots N(f_n)\}|\{s\}\rangle$$

And then we can define a general hilbert space

$$H = \bigotimes_{\{s\}} H_s$$

This is the "bad hews": the hilbert space is more complicated.

Now in this hilbert space any two distinct states in H can be distinugisted by observations in $u \in (-\infty, -1/\epsilon)$[7]. Lets prove this in two steps. First we define algebra of operators in $(-\infty, -1/\epsilon)$ as $A_{-\infty}$.

(a) Vacuum operators

$$T_{s,s'} = |\{s\}\rangle\langle\{s'\}| \in A_{-\infty}$$

The above operator takes on vacuum to another. *This is unique to gravity*. First we note that

$$H \in A_{-\infty} \quad Q_{\ell,m} \in A_{-\infty}$$

Now if a operator is in a algebra then the spectral projectors of these operators are in the algebra. This means that

$$P_0 = \int |\{s\}\rangle\langle\{s\}|\pi dS_{\ell,m} \in A_{-\infty}$$

We're being a bit schematic since we're ignore IR cutoffs. Similarly we can write a spectral decomposion of the chrages

$$Q_{\ell,m} = \int ds\; P_{\ell,m}[s]s \text{ and thus } P_{\ell,m} \in A_{-\infty}$$

Since algebras are closed we can then consider

$$P_0\pi_{\ell,m}P_{\ell,m}[S_{\ell,m}] = |\{s\}\rangle\langle\{s\}| \in A_{-\infty}$$

Now lets write down a scary looking operator

$$P_o\exp\left\{-\frac{i}{2}\int_{-\infty}^{-1/\epsilon} \sqrt{\gamma}C_{AB}(u,\Omega)G^{AB}(u,\Omega)d^2\Omega\right\}|\{s\}\rangle = |\{s'\}\rangle$$

The above operator is also in $A_{-\infty}$. Physically this corresponds to being able to change charges at infinity (wha?) And with this we have $T_{s,s'} \in A_{-\infty}$.

[7]This is diferent than information in a local qFt. However in a LQFT we'd need to look at all future null finity (all of sciplus).

(b) Now we want to show that all opertaros that map $H \to H$ are in the algbera $A_{-\infty}$. To do this we need the riech-lieder arguemnt. First we smear

$$\int_{-\infty}^{\infty} N_{AB}(u) f^{AB}(u) du \int_{-\infty}^{\infty} N_{A'B'}(u') f^{A'B'}(u') du' |\{s\}\rangle$$

We want to prove

$$\text{span}[N(f_1) \dots N(f_n)] |\{s\}\rangle = \text{span}[N(\tilde{f}_1) \dots N(\tilde{f}_n)] |\{s\}\rangle$$

Where $\tilde{f}_i$ supports in $(-\infty, -1/\epsilon)$. First let

$$|\psi\rangle = N(f) |\{s\}\rangle$$

E.g. a excitations of a single graviton on the vacuum. Assume

$$\langle\psi| N(\tilde{f}) |\{s\}\rangle = 0 \text{ for all } \tilde{f}$$

(we're doing a proof by contradiction). The above would imply

$$\langle\psi| N(u) |\{s\}\rangle = 0 \forall\ u \in (-\infty, -1/\epsilon)$$

The above is equal to

$$= \sum_E \langle\psi|E\rangle \langle E| N(u) |\{s\}\rangle = \sum_E \langle\psi|E\rangle \langle E| N(0) |\{s\}\rangle e^{iEu}$$

The second equality follows from time evolution. But the above is analytic when u is extended in upper half plane. And thus this vanishes for all real u.

$$\langle\psi| N(u) |\{s\}\rangle = 0 \forall u \in (-\infty, \infty)$$

If we smear the above with $f(u)$ we should get $\langle\psi|\psi\rangle$. And thus we've reached a contradiction. Now let $A : H \to H$. This operator must have expansion in this hilbert space

$$A = \sum_{s,s',n,m} c(n,m,s,s') |n_{\{s\}}\rangle \langle m_{\{s'\}}|$$

But now we just proved that $|n_{\{s\}}\rangle = x_n |\{s\}\rangle$ for some $x_n \in A_{-\infty}$ and similarly for $|m_{\{s'\}}\rangle$ so we can write

$$A = \sum_{n,m,s,s'} x_n |\{s\}\rangle \langle\{s'\}| x_m^\dagger = \sum_{n,m,s,s'} x_n T_{s,s'} x_m^\dagger$$

This is a linear combination of three elements of the algebra and due to closure we find that for an arbitarry $A : H \to H$ we have $A \in A_{-\infty}$.

What were some assumptions we made to extend results to UV complete theory.

(a) Vacuum operators were part of the algebra $A_{-\infty}$. This is the same as saying that the full thoery shares low-energy properties of low-energy theories.

(b) Positivitiy of energy.

Lecture 15: Low energy Tests of Black Hole Information

March 19, 2021

Suppose we're in flat space and $|0\rangle$ is some vacuum in flat space. Let $O(u,\Omega)$ be the boundary value of a massless scalar field.

$$O(u,\Omega) = \lim_{r\to\infty} r\phi(r,u,\Omega)$$

We'll excite the vacuum with a unitary operator

$$|f\rangle = e^{-i\lambda\int f(u,\Omega)O(u,\Omega)dud\Omega}|0\rangle$$

Suppose $f(u,\Omega)$ has support in $u\in[0,1]$. One way to think about the state is as some shell that is hitting $u=0$ and null infinity. Now consider some astrophysicist that can only observer things very far away. They can only observe on $u\in[-\infty,-1/\epsilon]$. The questions we want to ask is if the astrophyscists can deterine the shape of the shell before it hits $u\to-\infty$. We're lalowed to probe the state perturbtivley (in orders of λ). First note that this is impossible in a local qft. How will we do this magic in gravity? What we need to is measure a two point function. In particular we measure the correlcaiton function of the metirc and some dynamiocal field

$$\lim_{\tilde{u}\to-\infty}\langle f|M(\tilde{u})O(u,\Omega)|f\rangle$$

Where $M(\tilde{u})$ is the Bondi which turns into the hamiltonian as $\tilde{u}\to-\infty$. In the above limit we get

$$= 4\pi G\langle f|HO(u,\Omega)|f\rangle$$

Lets write down the correlcation function

$$C\langle 0|e^{-i\lambda\int f(u',\Omega')O(u',\Omega')}HO(u,\Omega)e^{i\lambda\int O(u'',\Omega'')f(u'',\Omega'')}|0\rangle$$

Working perturbativley we expand this to first order in λ.

$$= \int -i\lambda\left[\langle 0|O(u',\Omega')HO(u,\Omega)|0\rangle - \langle 0|HO(u,\Omega)O(u',\Omega')|0\rangle\right]f(u',\Omega')du'd\Omega'$$

Since the vacuum has no energy only the first term contributes. WE can compute the first term by noting

$$O(u',\Omega')H = HO(u',\Omega') + i\frac{\partial}{\partial u'}O(u',\Omega')$$

The second term comes from the Heisnebrg equations of motion and could also come from the commutation relations(?) Plugging everything in we get

$$\int \lambda\langle 0|\partial_{u'}O(u',\Omega')O(u,\Omega)|0\rangle f(u',\Omega')$$

Apprently this two point function isn't too hard to evaluate and it evaluates to

$$\langle\ldots\rangle = -\frac{1}{4\pi}\frac{1}{u'-u-i0}\delta(\Omega,\Omega')$$

Plugging this in gives us

$$C(u,\Omega) = \langle f|M(-\infty)O(u,\Omega)|f\rangle = G\lambda\int_0^1 \frac{f(x,\Omega)}{x-u-i0}dx + O(\lambda^2)$$

Where x bounds come from the support of $f(u,\Omega)$. Now we claim is if we know $C(u,\Omega)$ for $u \in (-\infty, -1/\epsilon)$ we can reconstruct $f(x,\Omega)$. The formal argument is by contradiction. Suppose $f_1 \neq f_2$ gave the same $C(u,\Omega)$ then means that

$$\int_0^1 \frac{f_1 - f_2}{x-u-i0}dx = 0 \forall u \in (-\infty, -1/\epsilon)$$

This is analytic when u is extended in the upper half plane and this would imply $f_1(x,\Omega) = f_2(x,\Omega)$ and this tells us that if we know $C(u,\Omega)$ then we know $f(u,\Omega)$. We'll also now look at a hands on argument. We can do a series expansion since x/u is small

$$C(u,\Omega) = -G\lambda\sum_{n=0}^{\infty}\int_0^1 f(x,\Omega)\frac{x^n}{u^{n+1}}dx$$

Where we can extract terms by extracting the moments

$$\int_0^1 f(x,\Omega)x^n dx$$

Some discussions: the price of being far away. Higher moments of f are supprssed by (x/u) where x and u are some measure. What this means is that the futher away you are the harder you have to work harder to extract higher order moments.

Lets now talk about global symmetires. First ocnsider our original state $|f\rangle$. Now consider so $\tilde{O}$ that is related to O by a global symmetry. Then we have

$$|\tilde{f}\rangle = e^{i\int f(u,\Omega)\tilde{O}(u,\Omega)}|0\rangle$$

How would the correlator be affected?

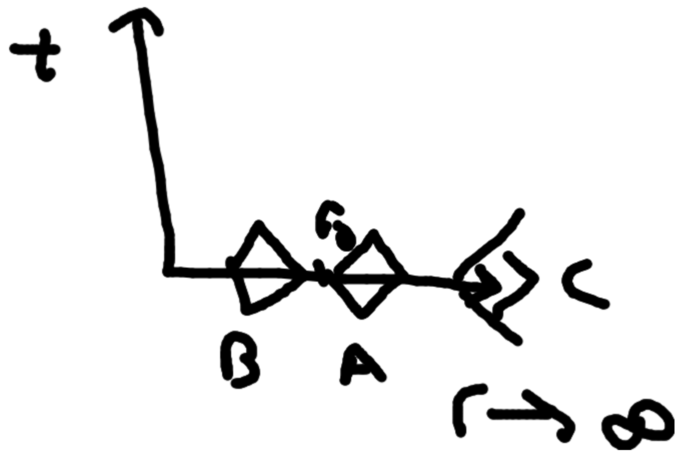
$$\textit{punchilne is } \tilde{f} \textit{ and } f \textit{ can be distinguished}$$

What are the applications of this to black hole informaiton. The questions we want to ask is "how does infomraion come out when black holes evaporate." The answer of this is that information is always outside and accessible even before evaporation.

To prove this lets try to set up an illuminating paradox. First we go back to global AdS

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

Now lets define the foloiwng geometry



Where $a, a^\dagger \in A$ and $\tilde{a}, \tilde{a}^\dagger \in B$ with fequency ω_0. We define $N = a^\dagger a$ and $\tilde{N} = \tilde{a}^\dagger \tilde{a}$. We can then set up CSSH operators

$$A_1 = \sum_{n=0}^{\infty} (|2n+1\rangle\langle 2n+1| - |2n\rangle\langle 2n|)$$

$$A_2 = \sum (|2n+1\rangle\langle 2n| + |2n\rangle\langle 2n+1|)$$

And similar ones for $\tilde{A}_{1,2}$. Then we could write

$$B_1 = \cos\theta \tilde{A}_1 + \sin\theta \tilde{A}_2 \quad B_2 = \cos\theta \tilde{A}_1 - \sin\theta \tilde{A}_2 \quad \tan\theta = \frac{2e^{-\pi\omega_0}}{1 + e^{-2\pi\omega_0}}$$

We can then compute

$$\langle 0|C_{AB}|0\rangle = \langle 0|A_1(B_1 + B_2) + A_2(B_1 - B_2)|0\rangle = \frac{2}{1 + e^{-\pi\omega_0}}(1 + 6e^{-\pi\omega_0} + e^{-2\pi\omega_0})^{1/2} > 2$$

We have that operators in A and B are gauge fixed quase local field operators. If we go in to LQFT there exists Q_1, Q_2 where $Q_i |0\rangle = B_i |0\rangle$. In gravity we have $P_0 \in C$ and combining everything we can consturct operators such that $|C_i| < 1$ (bounded operator) where $C_i |0\rangle = B_i |0\rangle$. So we can construct

$$|B_i\rangle\langle 0| = Q_i P_0 = Q_i \langle 0||0\rangle \quad |0\rangle\langle B_i| = P_0 Q_i^\dagger \quad |B_i\rangle\langle B_i| = Q_i P_0 Q_i^\dagger$$

Also if we grind through a little algebra we can find the eexplicity form $C_i |0\rangle$

$$C_i = \frac{1}{\langle B_i^2\rangle - \langle B_i\rangle^2} (|B_i\rangle\langle 0| + |0\rangle\langle B_i| - \langle B_i\rangle (|0\rangle\langle 0| - |B_i\rangle\langle B_i|)) \in C$$

This is so complex so to assert

$$|C_i| < 0 \quad C_i |0\rangle = |B_i\rangle$$

Now we can construct C_{AC} and C_{AB} so that we get another monogomy of infomration paradox. And this suggests that there is a paradox in black holes.

Lecture 16: Von Neumann Entropy as $\mathcal{I}^+$

Lets start by reviewing some stuff

(a) Stuff we've proven

  (a) We assumed that the full UV theory shares some low-energy properties and hamiltonian of the full theory has a real spectrum bounded below.

  (b) Using the above assumption we showed that all information about massess particles is at $\mathcal{I}^+_-$

(b) Stuff we didn't prove

  (a) Full UV theory is well defined.

  (b) Information at $\mathcal{I}^+_-$ is the same as $\mathcal{I}^-_+$.

Now we go back to the monogamy paradox we set up last lecture. We managed to find operators so that

$$\langle 0|C_{AB}|0\rangle^2 + \langle 0|C_{AC}|0\rangle^2 > 8$$

We will generalize this to black holes now. Suppose a black hole microstate is $|\psi\rangle$. Then we expect there exists some CSSH operators where

$$\langle \psi|C_{AB}|\psi\rangle > 2$$

To prove this we'll assume there exists

$$Q\,|0\rangle = |\psi\rangle,\ Q_{B_i}\,|0\rangle = B_i\,|\psi\rangle = |B_i\rangle$$

Using these operators we can write tehe following operator

$$|B_i\rangle\langle\psi| = Q_{B_i} P_o Q^\dagger \in \text{region } C \quad \text{hermitian conjugate}$$

$$|B_i\rangle\langle B_i| = Q_{B_i} P_o Q^\dagger_{B_i} \quad |\psi\rangle\langle\psi| = Q P_o Q^\dagger$$

Using this we can define a opertor

$$C_i = \frac{1}{\langle B_i^2\rangle - \langle B_i\rangle^2}[|B_i\rangle\langle\psi| + |\psi\rangle\langle B_i| - \langle B_i\rangle(|\psi\rangle\langle\psi| - |B_i\rangle\langle B_i|]$$

And after a little algebra we can define

$$\langle \psi|C_{AC}|\psi\rangle^2 + \langle \psi|C_{AB}|\psi\rangle^2 > 8$$

Another monogamy paradox!

we'll now talk about enganglement entropy at $\mathcal{I}^+$. In the context of quantum gravity we'll only look at a segment of $\mathcal{I}^+$. First we will define the entanglement entropy. Consider algebra of operators in $(-\infty, u_0)$ which we call A_{u_0}. Now suppose a theory is in some state $|\psi\rangle$. We now look for a $\rho \in A_{u_0}$ such that $\text{tr}(\rho O) = \langle\psi|O|\psi\rangle$ for all $O \in A_{u_0}$. Why is this the same as the partial trace definition? Suppose we have a factoraizable hilbert space H with $|\psi\rangle \in H$.

$$H = H_{\text{sys}} \otimes \tilde{H}_{\text{sys}} \Rightarrow \rho = \text{tr}_{\tilde{H}_{\text{sys}}} |\psi\rangle\langle\psi| \Rightarrow \rho \in \text{algebra of operator in } H_{\text{sys}}$$

And all of this tells us that $\text{tr}(\rho O) = \langle\psi|O|\psi\rangle$ for all O in H_{sys}. Now once we have the ρ_{u_0} operator we can define the entropy

$$S(u_0) = -\text{tr}\left(\rho_{u_0} \ln \rho_{u_0}\right)$$

Provided we regulate the entropy correctly. So we can approximate arbitrarily well ρ_{u_0} by an operator in $A_{-\infty}$. If we compute everything what we'll find is that $S(u_0)$ is constant(this formalizes the fact that the information is always outside.) Something to note is that this is note the conventional page curve.

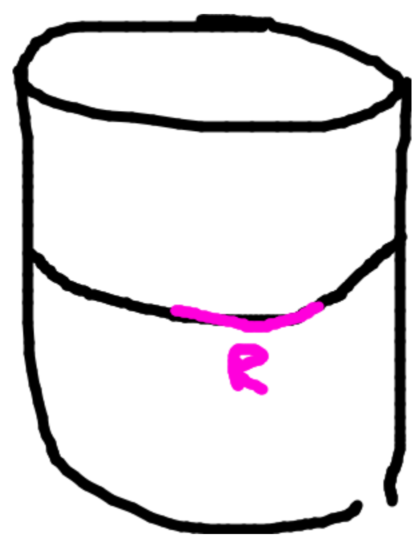
some stuff not typed out here.

Now we'll talk about a perspective on page curves from islands. The first thing to note is that the questions asked is different so answers are different. The basic setup is that we have gravity coupled to a CFT_d in asymptotic AdS. Now we take this entire system of gravity and CFT andcouple it to a CFT_d with no gravity. The boundary of these two systems has a AdS boundary that is transparent. Now taking this picture we could reduce the picture with a holographi duality to a CFT_d with no gravity on a half plane that ends at a "defect" which supports CFT_{d-1}. Now in this non-gravitional system we choose an arbitrary line and define points to the right of this line as the radiation region and to the left as the BH region. Now we have a facotirzed hilbert space and and now have a page curve. If we restore the original graivtional system then the page curve can also be interpeted as informtaion transfer from near-boundary region to nongravtaional bath.

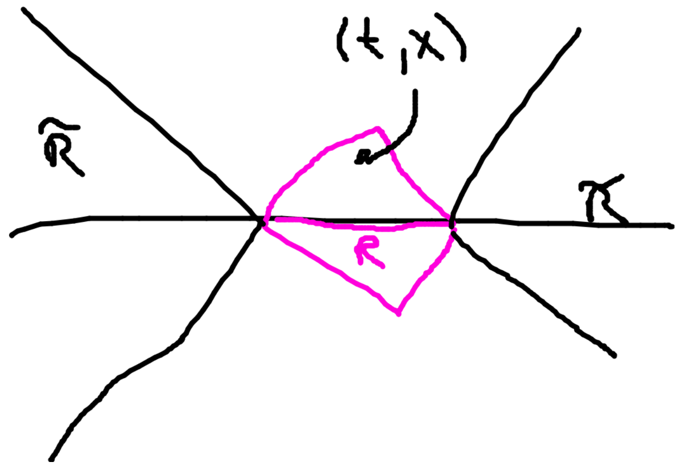
Lecture 17: Entanglement Entropy in Holographic CFT

March 21, 2021

Before we talk about islands next lecture we'll take a little side tour. First we'll talk about ads.



What is the entanglement entropy of R with it's complement $\tilde{R}$ and the von neumann entropy of R in some state. The first thing to note is that this question is asked independent of gravity. The next is to note that the e.e. depends only the causal diamond.



This is because the operator algebra of the causal diamond is isomorphic to the operator algebra of the base of the diamond. In a CFT we have from the fact that operators are heisenberg operators

$$O(t,x) = e^{iHt}O(0,x)e^{-iHt} \quad H = \int H(x)dx$$

It turns out that things outside of R don't affect things meaning that

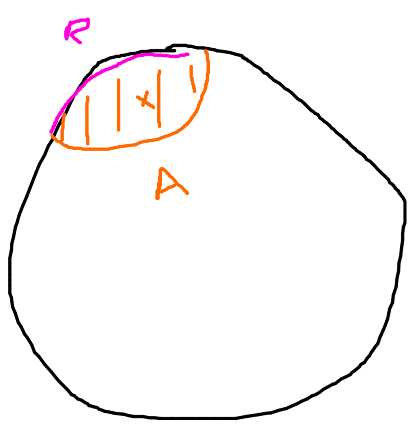
$$O(t,x) = e^{i\int_R H(x)dx\times t}O(0,x)e^{-i\int_R H(x)dx\times t} \in A_R$$

Where A_R is the operator algebra in R. The reason this works is that the rest of the terms in H commute with $O(0,x)$ by microcausality(?)

In CFTs with a holographic dual there is a proposal for how to compute the entanglement entropy of $S(R)$. The idea is to use bulk as an auxiliary device to compute $S(R)$.

$$S(R) = \min\left[\text{ext}\left[\frac{A}{4G} + S_{\text{bulk}}(x)\right]\right]$$

A is the rea of the bulk surface anchored on R and x is the region between R and the surface. S_{bulk} is the free field entropy of fields in x.



Extremize the inner bracket ends up corresponding to "maximizing" in time and minimizing in space. We then take the minimum of all possible extremal surfaces. We'll now do an example for how this is done. Sometimes there is also a homology constraint: x has no boundries except for the extremal surface and R. Furthermore there are UV-divergences so we need to match UV divergences. We can still extract cut-off independt quantities. This is because of subregion duality: CFT region R has informaion to the bulk region x.

Example 1: (Computation in AdS$_3$)

Lets first focus on the $A/4G$ term. First recall the metric of AdS_3

$$ds^2 = -(1+r^2)dt^2 + \frac{dr^2}{1+r^2} + r^2 d\theta^2$$

We expect the extremal surface will not move in time at all. So we're looking for an r and θ curve. Namely what we're trying to extremize

$$\int \sqrt{\left(\frac{dr}{d\tau}\right)^2 \frac{1}{1+r^2} + r^2 \left(\frac{d\theta}{d\tau}\right)^2} \, d\tau$$

In higher dimensions we're not always extremizing a geodesic length. We could just as well consider

$$\int \sqrt{\frac{1}{1+r^2} + r^2 \left(\frac{d\theta}{dr}\right)^2} \, dr$$

First notice there is a symmetry $\theta \to \theta + c$. This integral of motion gives us

$$\frac{d\theta}{dr} \equiv \dot{\theta} = \pm \frac{c/r}{\sqrt{(r^2-c^2)(1+r^2)}}$$

we have a choice in the constant. After evaluating the integral we get

$$\theta = \theta_0 + \cot^{-1}\left[c\sqrt{\frac{1+r^2}{r^2-c^2}}\right] \leftarrow \text{ ingoing branch}$$

$$\theta = \theta_0 - \cot^{-1}\left[c\sqrt{\frac{1+r^2}{r^2-c^2}}\right] \leftarrow \text{ return branch.}$$

Now we will evaluate the integral. Notice that it's diverent so we have a cutoff

$$A = 2\int_c^{1/2\epsilon} \frac{r dr}{\sqrt{(1+r^2)(r^2-c^2)}} = -\log(1+c^2) + 2\log 1/\epsilon$$

We can elimnate c by choosing how much the surface has moved in θ. If you look at our solutions of θ

$$c = \cot\frac{\delta\theta}{2}$$

All together we get

$$A = 2\log\left(\sin\frac{\delta\theta}{2}\times\frac{1}{\epsilon}\right)$$

Using this we can compute the entropy

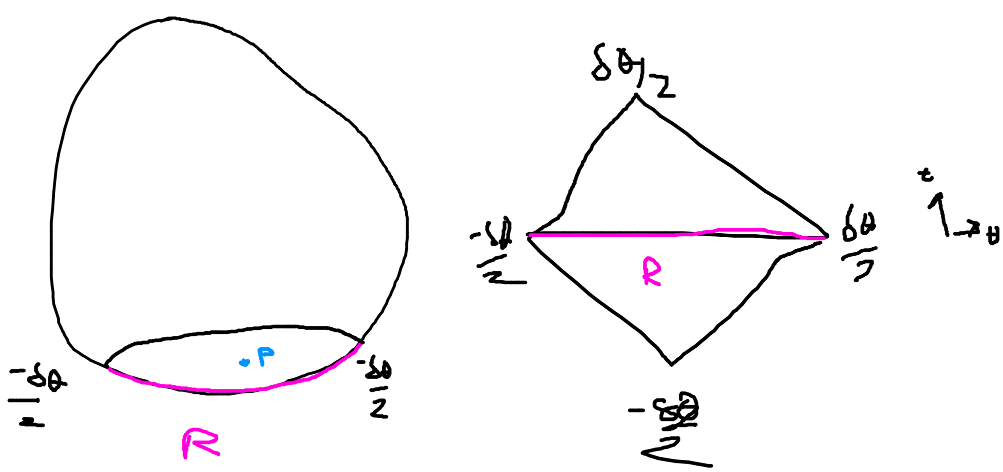
$$\frac{A}{4G} = \frac{1}{2G}\log\left(\frac{1}{\epsilon}\sin\frac{\delta\theta}{2}\right)$$

The central crhage of a CFT $C_N = 3/2G$ and therefore we predict

$$S = \frac{C_N}{3}\log 1/\epsilon \sin\frac{\delta\theta}{2}$$

This matches exactly with an equivalent calculation with only CFT techniques.

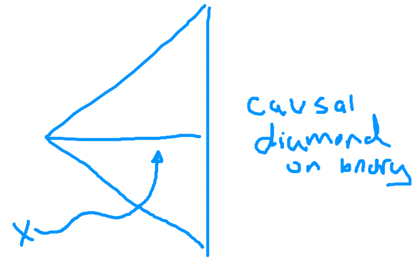
We also have a staement about the entaglement wedge.



For every point in the bulk there exists a timelike trajectory that takes that point to the causal diamond on the boundary. We have

$$r = \cot\frac{\delta\theta}{2} \quad \theta = 0 \quad t = 0 \quad dt = \pm\frac{dr}{1+r^2}$$

To reach from r_0 to ∞ takes time $\cot^{-1} r_0$. So the light ray from p reaches the boundary at $\theta = 0$ and $t = \delta\theta/2$. So the geometry is that we have some region in the boundary that is connected to all point on the boundary.



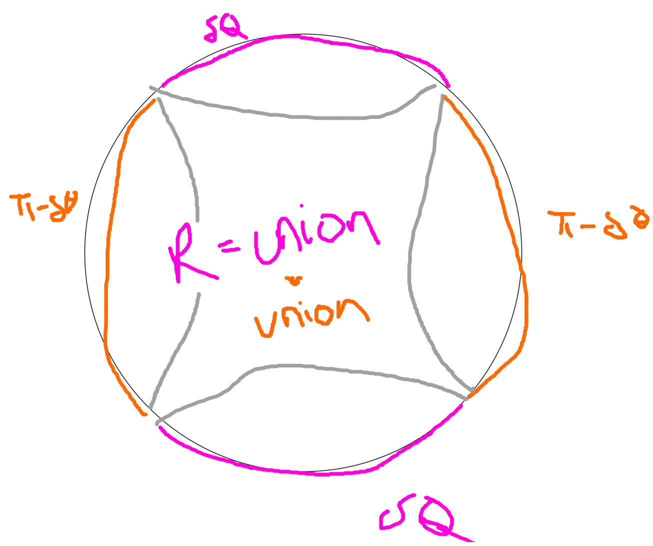
Say we have some bulk minimally coupled to a field ϕ which obeys

$$(\partial^2 - m^2)\phi = 0 \quad \lim_{r\to\infty} r^\Delta \phi(r,t,\theta) = O(t,\theta) \text{ on diamond}$$

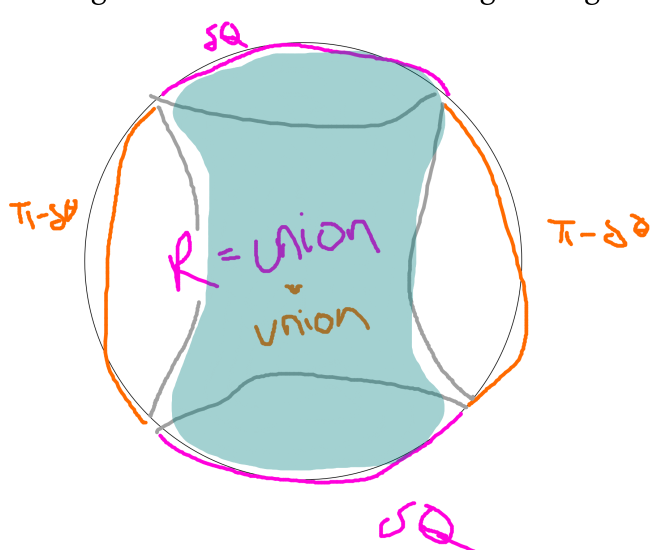
So we can write

$$\phi(r',t',\theta') = \int O(\theta,t)K(\theta,\theta',t,t',r)d\theta dt$$

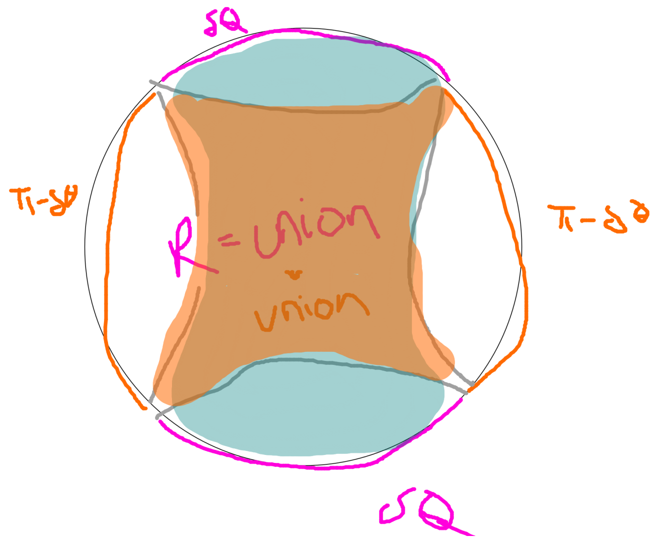
Where K is some kernel which can eb written down explicitly(in momentum space). Now lets consider a nontrivial entanglement wedge.



For $\delta\theta < \pi/2$ then the entropy is given by the orange surface. And for $\delta\theta > \pi/2$ the white surface win. In this second regimed we have the following entanglemnt wedge.



The organe region below are out of causal contact with boundary caused diamonds



This means that in this non-trivial entanglement wedge

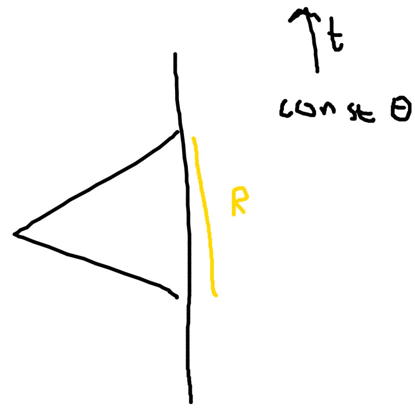
$$\phi(r', t', \theta') = \int O(\theta, t) K(\theta, \theta', t, t', r) d\theta dt \textbf{ does not work}$$

Lecture 18: Introduction to Islands

March 22, 2021

Lets start with the causal wedge proposal. This is only true in gravity

This region R has information about evertyhing between the dashed line and R. AT this point we might wonder why gravity is important. Even without gravity the following is still true.



The diamond built on R has information about the bulk causal wedge. But there is difference between the two figures. At any time we have the information on some boundary.

From a bulk perspective it's suprising that R on the boundary can be completed to the causal diamond on R. So why can't bulk info enter this causal diamond? The answer is because in gravity it cannot because the hamiltonain is a boundary term.

At $t = 0$ there is no formula

$$\phi(t,r,\Omega) \neq \int O(t',\Omega')K(t,r,\Omega,\Omega')d\Omega' = 0$$

We'll go over the replica trick to derive entanglement entropy. We'll think of the state as a functional of fields $\psi[\phi(x)]$ at $t = 0$. With this we can take the action

$$\int_{\phi(x,t_E=-\infty)=0}^{\phi(x,0)=\phi(x)} e^{-S_E[\phi]} D\phi = \psi[\phi(x)]$$

This is an integral on the lower half plane. What is this function physcially? It's the (unnomralized) vacuum state. To see this note that the path integral is

$$\langle\phi_1(x)|E^{-H\tau}|\phi_2(x)\rangle = \text{path integral of euclidean time } \tau$$

In this we can resolve the identity

$$= \sum_E \langle\phi_1(x)|E\rangle\langle E|\phi_2(x)\rangle e^{-E\zeta}$$

For a time parmaeter η. If we think of this as a fun of $\phi_1(x)$ as as $\zeta \to \infty$ then we can see that the above integral is the vacuum state.

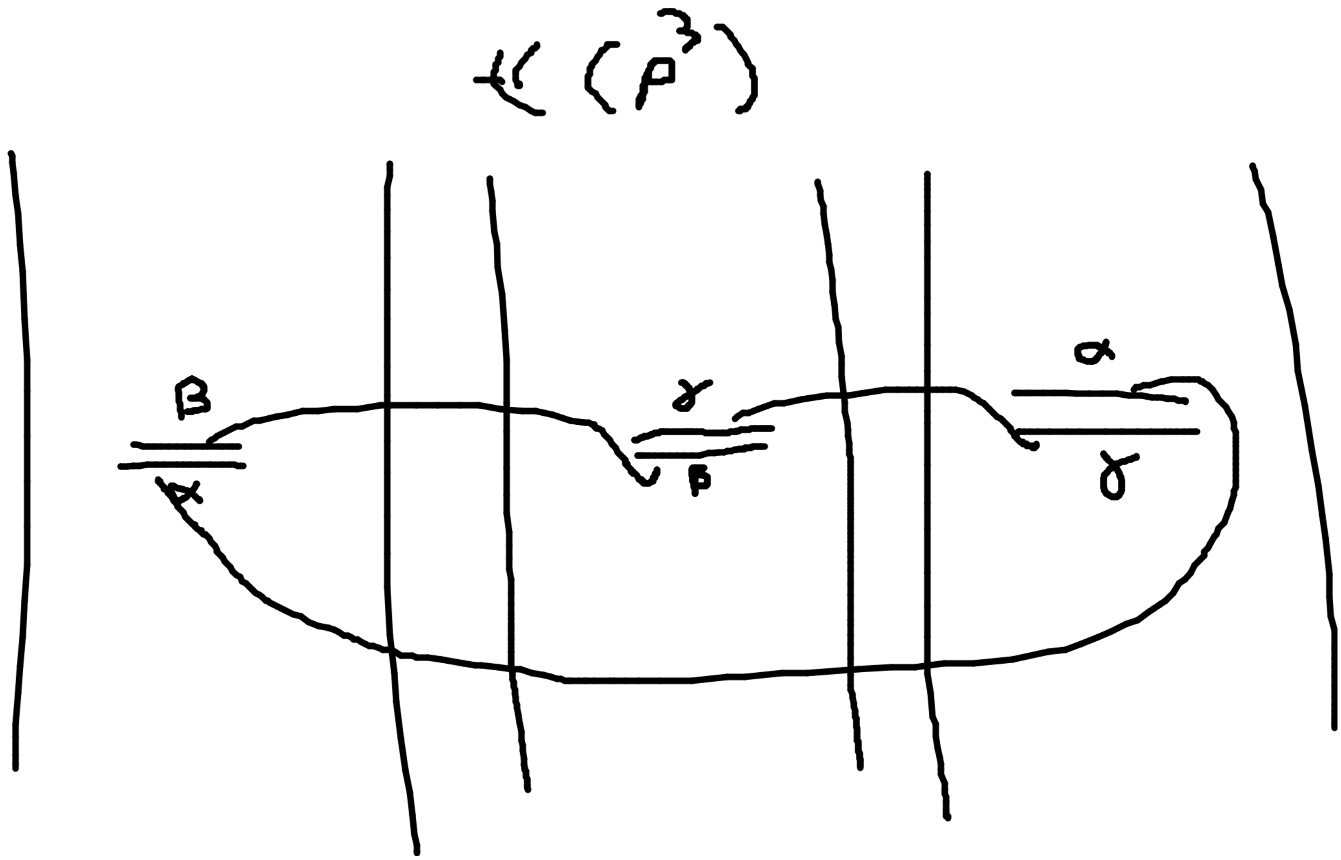
Now how could we do this for a density matrix? If we have some region R and its complement $\tilde{R}$ how do we find ρ_R the density matrix when the state is in a vacuum. Lets try

$$\psi[\phi(x)]\psi^*[\tilde{\phi}(x)] = \rho[\phi,\tilde{\phi}]$$

Let $\alpha(x)$ and $\beta(x)$ have domain R . We can write the density matrix path integral in teh vaccuuum state (which is unnormalized) as

$$\rho(\alpha,\beta) = \int_{\phi(x,0^-)=\alpha(x)}^{\phi(x,0^+)=\beta(x)} e^{-S} D\Phi$$

Now we want to compute $\text{tr}(\rho^n)$. A diagrammtic representation could be



We'll write

$$\text{tr}(\rho^n) = \frac{Z(n)}{Z(1)^n} = \frac{\text{(path integral on surface)}}{\text{(path integral on first surface only)}^n}$$

Now we want to compute the von neumann entropy $-\text{tr}(\rho \log \rho)$

$$-\text{tr}(\rho \log \rho) = \lim_{n \to 1} \frac{1}{1-n} \log \text{tr} \rho^n = -\partial_n [\log Z(n) - n \log Z(1)] \Big|_{n=1}$$

Now the idea in holographics entanglement entropy we replicate the bounadry and find a bulk saddle which has the right (replicated) boundary conditions. We the nevaluate the bulk action on this saddle. Maldecena and Yu(kowisk?) gave us a proof for how to compute this which we'll now spell out

(a) For $n = 1$ bulk metric is AdS

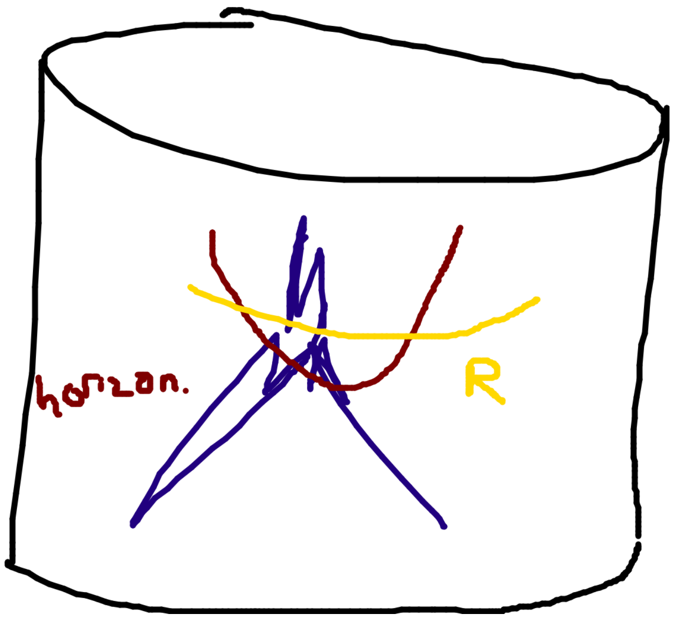
(b) We need metric "near" $n = 1$

(c) So we analytically continue the metric to one which has conical singularities

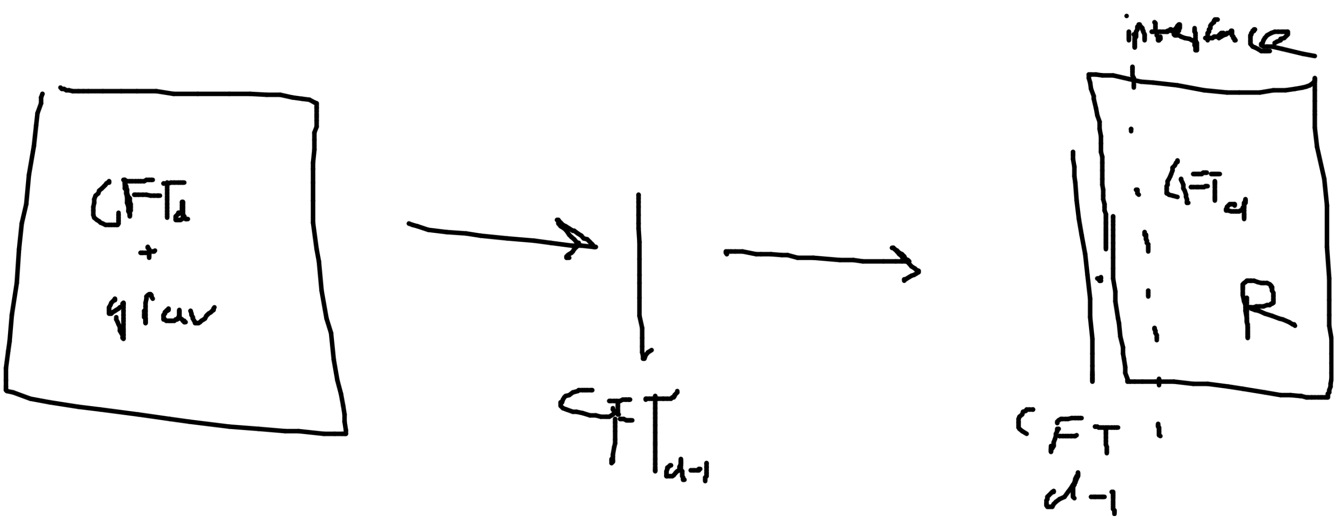
(d) On analytically continued metric we compute the bulk action

(e) We knew we were going to get Area

And now we'll couple CFTs coupled to bath.



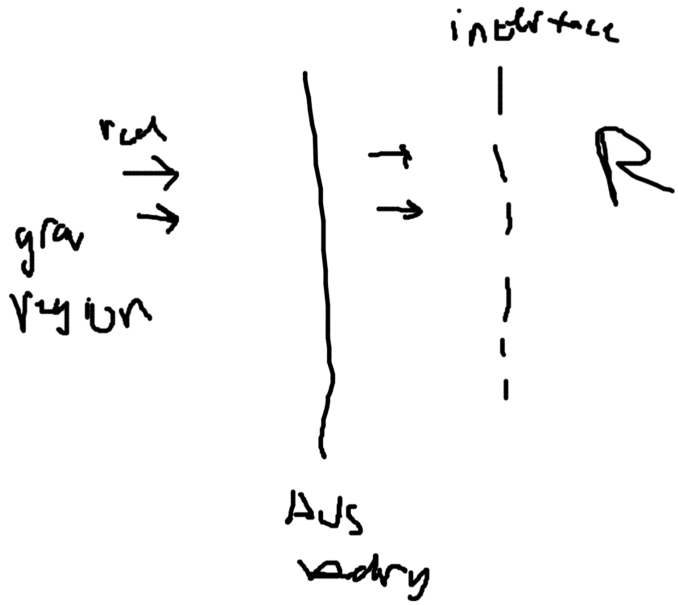
The entropy of the entire boundary remains 0 from the R-T formula. This is also consisten with the princitple of holography of info. Lets talk in the bulk a theory of gravity plus a CFT_d from AdS/CFT is dual to a CFT_{d-1}. We can now couple this dual with the same CFT_d that lives in the bulk but without gravity. When we do this coupling the CFT_d has more degrees of freedom so if we had any energy in the CFT_{d-1} would flow to CFT_d. If we choose some interface and the the right we define as R we would now ask what is the entropy of region R.



In the third picture we can replace CFT_{d-1} with the original dual CFT_d+gravity. We can then turn on the coupling with transparent boundary conditions. This is the setup for the island proposal. We will now compute the entropy of R. First we extend a Cauchy slice accross the entire geometry. Then if we want to compute $S(R)$

$$S(R) = \min[\text{ext}[A(\text{boundary of islands})/4G + S_{\text{semi-cl}}(\text{R}\cup\text{ island})]]$$

Where $S_{\text{semi-classical}}$ is the entropy as computed using the rules of QFT in curved spacetime.



As radiation enters the region then the entropy $S(R)$ just grows (from the semicalssical entropy.) What do we expect from the semicalssical entropy? Well it hsould increase monotonically until BH evaporates. We also expect semicallsical entropy for the entire cauchy slice is zero. Now if we go to late times the $S(R)$ might become very large and an island might form in the interior and the moment this happens we have to pay a cost in area. AT late enough time the radiation in the island might purify the radiation in R and thus even though the island is expensive it resuses the semicallsical entropy. It turns out after the page time the island wins because the reducition in the semicallsical entorpys compensates for the area.