Black Hole Information Paradox Notes

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`.

Last updated on January 20, 2021

Lecture 1: Introduction and Two-Point QFT Correlators

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(TODO ???). 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. $\mathcal{L} = 1/2(\partial_\mu \phi)^2$). (TODO how???)

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

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?)

Lecture 2: Entangled Modes accross Null Surfaces

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).

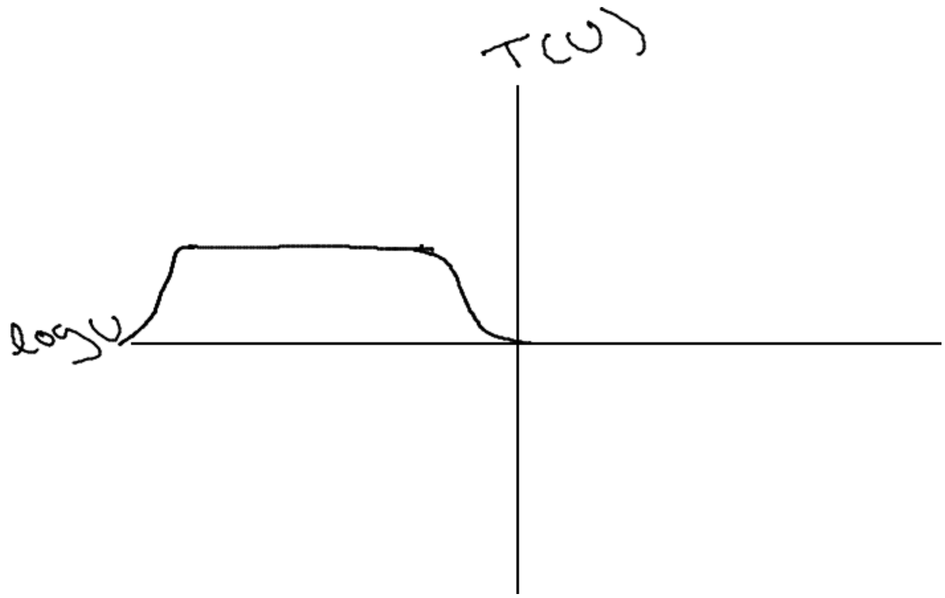
We'll start by introducing a smearing function (TODO what?) $T(U)$ with the following properties

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

(b) Support in interval $[U_l, U_r]$ where $\ell_{\text{UV}} \ll U_r, U_l \ll \ell_{\text{curv}}$ 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.

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

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.

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

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

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

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

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

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

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

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)

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

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

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

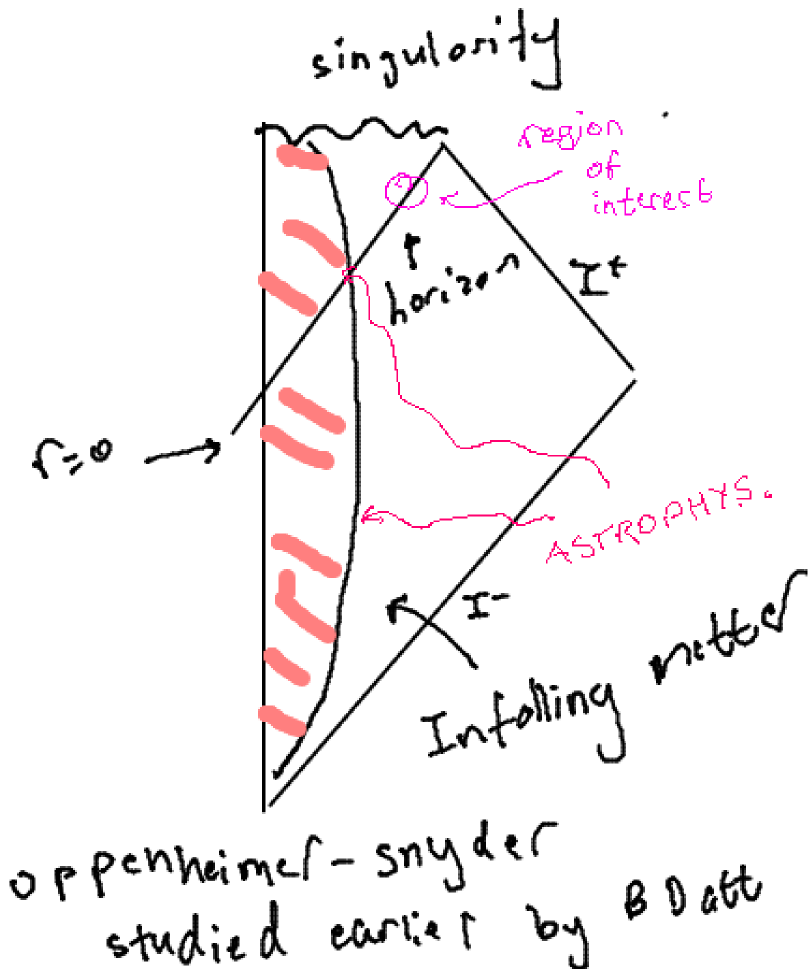
Similarly we can also show that

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

Next lecture we'll apply these results to black holes

Lecture 3:

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 is 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} are solutions which we choose as our basis) (where we choose f_{in} and f_{out} 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 $f_{\text{out}} \to b_{\omega,\ell}r^{(1-d)/2}e^{i\omega r_*}$.

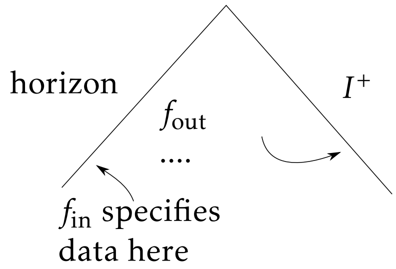


Figure 1: Description of how information in specified in Penrose diagram

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 then our expansion start breaking down. 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.