The information loss problem

Kavir Asswarnicumar

Supervised by Professor Robert De Mello Koch

A dissertation presented for the degree of
Master of Science
Department of Physics
University of the Witwatersrand

11 March 2022

Declaration

I declare that this dissertation is my own work. It is being submitted in requirement for the Degree of Master of Science, Research by Dissertation at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Kavir Asswarnicumar

11 March 2022

Date

Acknowledgments

This dissertation would not be possible without the support from so many individuals. I would like to start by thanking my parents Sunita and Ashok. Thank you mum for persuading me to study what Einstein studied. You always showed an interest in my work and I will forever be grateful. But most of all thank you creating this amazing bond between us. I will always be your baby. Dad, thank you for supporting me and ensuring that I achieve my dreams. Everything you do is for our family, and will do my best to follow the example to set for us.

To my sister Alisha: at times when I needed you most, you played the role as protector, teacher and mother. I will always be grateful. To my brother Jitesh, thank you for always going the extra mile for our family. Your role as a bigger brother and sometimes as a father will stay with me forever.

To my wife Kuvashnee, you're one of the most intelligent and resilient individuals I know. I am inspired by you daily. Thank you for your unconditional support and making sure that I have a life outside academia. I cannot wait to start the rest of our lives together.

To Basil and Rani Naidoo, thank you for mentoring me and supporting me when I needed you most. I am truly blessed to become a part of your family.

I would also like to thank all of my friends and colleagues for all the great conversations that help me get through difficulties in everyday life. To Edivaldo, we have been friends since we were kids and got to experience many great memories together. Our soccer sessions in the back yard were the best. I am truly proud of the relationship we have built and the type of men we have become. I would also like to thank a very dear friend Kevin. Thank you for always being there for me and my family. Not many things can compare to our fifa sessions. We have grown together and I am proud to say that you are more like a brother. To Ntokozo, our friendship started with a simple question you asked me. I always think about what you told me, because it has shaped the way I think. You are a visionary.

My pursue in theoretical physics is inspired by one of the greatest teachers I know, my supervisor Professor Robert De Mello Koch. Prof, I would like to thank you for teaching me the tricks of the trade. Thank you for the wonderful and inspiring conversations and your continuous support. None of this would be possible without you. You know, when I attended your quantum field theory course, I knew this is where I belong. You reignited my love for physics. Apart from academia, I will take your lessons with me everywhere I go.

Finally, I would like to thank the NRF for supporting this research. The views and opinions presented here are of the authors only.

For Radica,

Abstract

The paradox of information loss within black holes is a major problem in theoretical physics. Stephen Hawking showed that pure states evolve into mixed states and entanglement entropy for the Page curve increases monotonically. These two results violate the unitarity of quantum mechanics and imply that gravitational effects at the quantum scale is not well understood. This paradox has lead to the idea's of holography and subsequently the AdS/CFT correspondence. Much progress towards the paradox relies on holography, which includes a formula for computing entanglement entropy for gravitational systems. This formula consists of two terms: 1) A quantum extremal surface and 2) Entropy for semi-classical gravity. We present the computation on how to locate quantum extremal surfaces using JT gravity in AdS_2 and subsequently show the Island surface on a Penrose diagram. We carry out the saddle point approximation, showing where the semi-classical expansion comes from and how to compute higher order corrections using Feynman diagrams. We find that the formula for entanglement entropy in gravitational systems within the Island proposal does indeed produce a unitary Page curve. However, the Island proposal is inconsistent because it doesn't address massive gravitons which are needed in dimensions higher than 2. It also fails to address that the stress-energy tensor $T_{\mu\nu}$ is not conserved, because energy and momentum leaks into the non-gravitational bath.

Contents

List of Figures

1 Introduction

The year 2020 has been exciting for theoretical physicists because many new interesting ideas have been developed towards solving one of the most mysterious problems in physics. This problem is called the black hole information loss paradox.

The Physics Nobel Prize in 2020 was awarded one half to Roger Penrose for the theoretical discovery that black hole formation is a prediction within Einstein's theory of general relativity. The other half of the prize was jointly awarded to Andrea Ghez and Reinhard Genzel for discovering experimentally, a super-massive compact object at the center of our galaxy. Secretary General Goran Hansson from the Royal Swedish Academy of Science remarked: "*...this years prize is about the darkest secrets of the universe...*".

Black holes are the results of studying singularities inherent within general relativity. Before 2020, black holes seemed merely as dark entities that engulf whatever crosses it's horizon. The equations strongly support the idea of black holes, regardless of how absurd it seems. It took the experimental physics community many years to finally confirm signals that describe a super-massive compact object, which we think[1] is a black hole. One could argue that this discovery, together with it's mathematical frame work, is much like Dirac's relativistic quantum theory predicting anti-particles, which were subsequently discovered.

Even though theoretical physicists have a lot to celebrate, the celebrations are done with some sympathy because Stephen Hawking was not alive to be part of the celebration. Stephen Hawking explored deep questions about general relativity and quantum mechanics. These questions were explored by probing theoretical aspects of black holes. It began in 1965 when Hawking attended a lecture by Roger Penrose [1]. This iconic lecture inspired Stephen Hawking, and lead him to the idea that singularities could be at the beginning of the expansion of our universe. Hawking further postulated that black holes are thermodynamic systems with non-zero temperature. Up until 2019, many scientists did not believe in the existence of black holes. So Hawking's conjecture was way ahead of it's time. More than 50 years later experimental evidence finds a black hole at the center of our galaxy! To make matters even more interesting, Stephen Hawking subsequently developed the idea that black holes evaporate [2]. Black holes now possessed a quantum aspect. The evaporation is due to quantum fluctuations. Evaporation is possible when entangled particles across the horizon were considered. In this case the negative energy particle crosses the horizon and falls into the singularity. The positive energy particle propagates away from the black hole, thus creating the overall effect of the black hole loosing energy. The information coded into the quantum states that are entangled have an entropy. Stephen Hawking finally argued that the black hole evaporation process is not quantum mechanically unitary [3]. It seemed like initial data was lost and that the entanglement entropy of the fields increased monotonically. Together with Don Page, the Information Loss Paradox was coined and this problem was

[1]The signals all point to one direction, unique properties shared with those of a black hole.

formally stated in [3, 4]. Three years before Stephen Hawking's death, a talk was given on the 28th August 2015 at a conference in Stockholm's KTH Royal Institute of Technology. At this conference, Stephen Hawking proposed a solution to the black hole information loss paradox [5]. This idea was inspired by a lecture given by Andy Strominger, where supertranslations for stationery black hole horizons were considered. Supertranslations are extra conserved charges, which form the BMS^+ transformations [6]. Hawking argued that the 'lost' data was encoded in the values of these charges. Unfortunately this paradox was not solved by Hawking, and has not yet been solved. Currently, there are no existing solutions that can explain the black hole information loss paradox without referring to conjectures.

In this dissertation, we examine a recent proposal [7] that claims to produce a unitary Page curve. We need to refer to conjectures by which many solutions are based on. One of the first conjectures is based on the idea of holography. Gerhard t'Hooft started studying this paradox and realized that surfaces on lower dimensions are enough to describe the physics of higher dimensions [8]. This realization was further developed by Leonard Susskind and was named holography [9]. One of the most crucial conjectures within the solution we study is called the AdS/CFT correspondence principle [10]. This conjecture is influential as we can see from the number of citations it has and the pathway it has carved for theoretical physicists interested in string theory[2]. Without discussing the details, this correspondence postulates a duality between gauge theory and gravity. In particular, it says that a theory of quantum gravity defined on a specific spacetime is equivalent to a quantum field theory that is defined on the boundary of that spacetime. One example is an exact equivalence between string theory on the Ad$S_5 \times S^5$ spacetime (a spacetime with constant negative curvature) and the $\mathcal{N} = 4$ super Yang Mills theory. Further important refinements of the conjecture is given by Edward Witten [11] and Steven Gubser, Igor Klebanov and Alexander Polyakov [12].

Within the black hole information loss paradox, it is important to discuss and distinguish the different types of entropy. The entropy of radiation for the entangled particles that escape the black hole is called the entanglement entropy. This is the fine-grained entropy. The black hole itself has an entropy and the result for computing it was achieved in a classical context by Bekenstein [13]. This is the coarse-grained entropy. This result needs to be modified for radiating black holes. Due to the quantum fluctuations that make black hole evaporation possible, a different formalism for computing the entropy was needed. An important milestone for the black hole information loss paradox comes from computing entanglement entropy. It was done in the context of AdS/CFT, by Ryu and Takayanagi [14]. Even though gravity effects are very small outside a black hole, theories of gravity encode information in a very different way from theories without gravity. The is one of the first detours taken compared to Hawking's computation. Hawking was essentially using an incorrect formula for the entropy. The Ryu-Takayanagi formula for gravitational entropy is

[2]String theories are theories of quantum gravity.

essentially the search for a quantum extremal surface. The subsequently important result is the description of a surface called the "Island" [15]. The authors in this paper used the ideas of holography to access higher dimensions. The higher dimensions allow for a geometric connection of the internal modes of the black hole with it's radiation outside. This could be viewed as an application of the ER=EPR relation [16] which is a theory that attempts to explain a version of the AMPS paradox [17]. The following papers have computed the Hawking entropy at late times and shown a unitary Page curve [7, 18, 19, 20, 21].

The rest of this dissertation is structured as follows: in Section II we review aspects about general relativity that are essential to the black hole information loss paradox and it's possible solutions. We show why gravity in 2 dimensions is interesting. Other aspects about general relativity includes Penrose diagrams, black holes and the Page curve. We describe how to represent the physics of black holes within a diagram. Thereafter we explain the equations that describe black holes. Another aspect that is essential to this paradox involves studying quantum states and observable's. We explain the unitarity of quantum mechanics and how quantum states are used to compute entropy. The measurement of entropy in relation to time is shown by the Page curve. We explain this curve and what is expected from a unitary theory against what has been computed by Stephen Hawking. In Section III we first compute the saddle point approximation and higher order corrections. Before presenting an application of the saddle point approximation, we show a relation between thermodynamics and the Euclidean path integral. Thereafter we show how applying this approximation scheme to the gravitational path integral results in semi-classical gravity and thermodynamics of a black hole. In Section IV we explore two solutions to the paradox. We first explain entanglement and it's implications for the black hole information loss paradox. We then discuss the fuzzball proposal and how it solves the paradox. Thereafter we discuss a modern version of the black hole information loss paradox, called the AMPS paradox. Finally we explain the holography of information and how it solves the paradox without conjectures. In Section V we present the bulk reconstruction scheme and the entanglement wedge. We show how the equations of motion are calculated with it's specific boundary conditions. We conclude the section by describing what the entanglement wedge contains and how to picture it on Penrose diagrams. In Section VI we present the Island proposal. We first explain the formula needed to compute entanglement entropy for gravitational systems. Thereafter we show how island surfaces arise and represent them on Penrose diagrams. We conclude the section by showing how island surfaces affect the generalized entropy and how this proposal achieves a unitary Page curve. We end with Section VII and present conclusions. We discuss the importance of the literature presented and how this literature is central to understanding this paradox and it's possible solutions. We list some shortfalls of the possible solutions and provide recommendations.

2 A Review of special cases in General Relativity

2.1 Gravity in 2 dimensions - The Jackiw-Teitelboim Action

In the first section of this dissertation, we review one of the most interesting action's in gravity[3]. This action was proposed to study quantum gravity and take a different path to that taken by string theorists. This action is interesting because one of the proposed solutions to the black hole information loss paradox is computed in 2 dimensions. Studying lower dimensional models can help us understand the properties and dynamical characteristics of the physical world. This is why this action was proposed. For example in solid state physics, a 1 dimensional model is sufficient to describe dynamical properties of phonon's. In quantum mechanics, tunneling phenomena is calculated and described in 1 dimension. The harmonic oscillator is also computed in 1 dimension. It's generalization to 3 dimensions is straightforward. However in this case the 3 dimensional harmonic oscillator will have levels that cause degeneracy (energy with same quantum number). The 1 dimensional harmonic oscillator displays no degeneracy at all and this implies that even though lower dimensional models can be useful, it's generalization to higher dimensions needs careful treatment. There are many other fascinating examples of how lower dimensional "toy models" can be successful for describing the properties and dynamics of processes that occur in the classical and quantum world.

In the early 1980's, two physicists Roman Jackiw and Claudio Teitelboim published articles and consequently developed what we call today *Jackiw-Teitelboim Gravity* (JT gravity). Roman Jackiw [22] considered studying gravity on a line (1 temporal and 1 spatial dimension = 2 Dimensional space-time) and on a plane (1 temporal and 2 spatial dimension = 3 Dimensional space-time). At the time, many physicists had been using string theory to study quantum gravity in dimensions higher than 4. Jackiw's intent to study quantum gravity in lower dimensions was not for unification purposes, but instead hopefully to illuminate the dynamics of our physical 4 and perhaps higher dimensional quantum gravity theories. On the other hand, Teitelboim's [23] motivation was the same except that he wanted the lower dimension models to also describe dynamics of super-gravity. At any rate, both physicists shared the same sentiments when it came to the usefulness of lower dimensional gravity.

Studying gravity in 2 dimensions is interesting because it has peculiar and non-generic features. For instance, Einstein's field equation is given by

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \tag{1}$$

where Λ is the cosmological constant, G_N is Newton's gravitational constant, $T_{\mu\nu}$ is the

[3]Gravity in this context refers to Einstein's theory of general relativity.

stress-energy tensor, $g_{\mu\nu}$ is a function called the metric tensor and $G_{\mu\nu}$ is Einstein's tensor defined by

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R. \tag{2}$$

In the absence of matter where $T_{\mu\nu} = 0$, the field equation (1) becomes

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\Lambda g_{\mu\nu}. \tag{3}$$

If we multiply (3) by $g_{\mu\nu}$, we can solve for the cosmological constant

$$g^{\mu\nu} R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} g^{\mu\nu} R = -\Lambda g_{\mu\nu} g^{\mu\nu}$$

The result is

$$\Lambda = R \left(\frac{1}{2} - \frac{1}{D} \right) \tag{4}$$

Note that $g_{\mu\nu} g^{\mu\nu} = D$, which denotes the dimension of the spacetime in which the theory is defined. So when $D = 2$, we are constrained to set the cosmological constant to zero $\Lambda = 0$ and perhaps the curvature scalar to zero $R = 0$. The dynamics of gravity in two dimensions is highly constrained, so we would have liked a model with the freedom to keep both the curvature scalar and the cosmological constant non-zero.

To remedy this, building on earlier work [23], Jackiw proposed [22] a constant curvature equation

$$R + 2\Lambda = 0 \tag{5}$$

The action whose variation gives (5) is

$$S = \int d^2x N \sqrt{-g} \, (R + 2\Lambda) \tag{6}$$

where N is a Lagrange multiplier (a scalar field), $g = \det(g_{\mu\nu})$ and the factor $\sqrt{-g}$ is needed to ensure that the volume integration measure is invariant under general coordinate transformations $x^\mu \to x'^\mu$. This together with the facts that R is a scalar, N is a scalar and Λ is also a scalar, ensure that the action itself is a scalar as it must be. This demonstrates that Jackiw's equation (5) can be recovered using the principle of least action. In the rest of this dissertation we refer to the theory defined by the action (6) as JT gravity.

2.2 Penrose Diagrams

General relativity is a dynamical theory of spacetime. One key idea used to formulate the theory, is that even for curved spacetimes, if we zoom in close enough the geometry is qualitatively flat. In general relativity, the distance between two nearby points called the line element is given by

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{7}$$

where $g_{\mu\nu}(x)$ is the metric tensor. The metric may depend on the spacetime coordinate x^μ. Spacetime's are structures that may be infinite. These structures are characterized by their topological and causal properties. A common tool that is used to study these properties is the Penrose diagram. This is because with well chosen coordinate transformations, Penrose diagrams "compactify" the geometry of a given spacetime and represent it on a finite plane. Penrose diagrams are well suited for spacetime's that have spherical symmetry. Therefore, Penrose diagrams can also be applied to black holes. In order to preserve the causal structure of a given spacetime, we consider light rays which propagate along $ds^2 = 0$. This implies in 2D, that $dt \pm dr = 0$. The next step is to transform original coordinates to light-cone coordinates

$$F^+ = f(t+r), \qquad F^- = f(t-r)$$

The function f is a transformation that makes the new coordinates finite. The range of the original coordinates t and r always needs to be specified. Typically $0 \leq r \leq \infty$ and $-\infty \leq t \leq +\infty$. Although t and r have infinite range, F^+ and F^- have finite range

$$F^+ = \tanh(t+r), \qquad -1 \leq F^+ \leq 1$$
$$F^- = \tanh(t-r), \qquad -1 \leq F^- \leq 1$$

Let's review at an example from [24], for a 2 dimensional Minkowski spacetime. The goal here is to calculate and draw it's Penrose diagram. The metric in this case is

$$ds^2 = -dt^2 + dx^2 \tag{8}$$

where $-\infty < t < \infty$ and $-\infty < x < +\infty$. The transformation to light-cone coordinates is

$$a_\pm = t \pm x$$

$$a_+ + a_- = 2t, \qquad a_+ - a_- = 2x$$

Taking the derivative on each side and squaring gives

$$-(da_+^2 + da_-^2 + 2da_+da_-) = -4dt^2, \qquad da_+^2 + da_-^2 - 2da_+da_- = 4dx^2$$

Therefore,

$$ds^2 = -da_+da_- \tag{9}$$

We now need to compactify the spacetime which entails mapping the light-cone coordinates to a finite interval,

$$a_\pm = \tan(\bar{a}_\pm), \qquad \bar{a}_\pm = \frac{\tau \pm \theta}{2}$$

Using the above transformation in (9),

$$da_+ = \frac{\partial}{\partial\tau}\tan(\frac{\tau+\theta}{2})d\tau + \frac{\partial}{\partial\theta}\tan(\frac{\tau+\theta}{2})d\theta$$
$$da_- = \frac{\partial}{\partial\tau}\tan(\frac{\tau-\theta}{2})d\tau + \frac{\partial}{\partial\theta}\tan(\frac{\tau-\theta}{2})d\theta$$

Doing the derivatives and applying the chain rule gives

$$da_+ = \frac{1}{2}\sec^2(\frac{\tau+\theta}{2})d\tau + \frac{1}{2}\sec^2(\frac{\tau+\theta}{2})d\theta$$
$$da_- = \frac{1}{2}\sec^2(\frac{\tau-\theta}{2})d\tau + \frac{1}{2}\sec^2(\frac{\tau-\theta}{2})d\theta$$

For the Minkowski metric we get

$$\begin{aligned} ds^2 &= -da_+da_- \\ &= -\frac{1}{4}\frac{1}{\cos^2\bar{a}_+}\frac{1}{\cos^2\bar{a}_-}d\tau^2 + \frac{1}{4}\frac{1}{\cos^2\bar{a}_+}\frac{1}{\cos^2\bar{a}_-}d\theta^2 \\ &= \frac{1}{4}\frac{1}{\cos^2\bar{a}_+\cos^2\bar{a}_-}(-d\tau^2 + d\theta^2) \end{aligned} \tag{10}$$

The factor in front of (10) is a conformal factor which performs a scaling of the line element. Compactification is achieved by dropping the conformal factor, which gives a new spacetime with the same causal and topological structure as the original spacetime. The Penrose diagram is a geometrical visualization of the new spacetime. Where initially we had t and x, we now have τ and θ. The boundaries of our new coordinates satisfy $|\frac{\tau+\theta}{2}| \leq \frac{\pi}{2}$ because of the asymptotic properties of the tangent function. This choice of function is one of the main ingredients responsible for mapping the infinite spacetime to a finite plane. The Penrose diagram resulting from our calculation is Figure 1, a diamond,

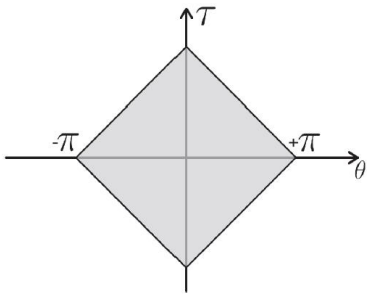


Figure 1: Penrose diagram for a 2D Minkowski space-time [24].

In order to study the qualitative features of a Penrose diagram, we show the finer details. These features will give a full description of what information the Penrose diagram is able to relay. The diagram we consider is a half diamond, implying that we now would have the limits $-\infty < t < +\infty$ and $0 < r < +\infty$. A detailed diagram is given in Figure 2

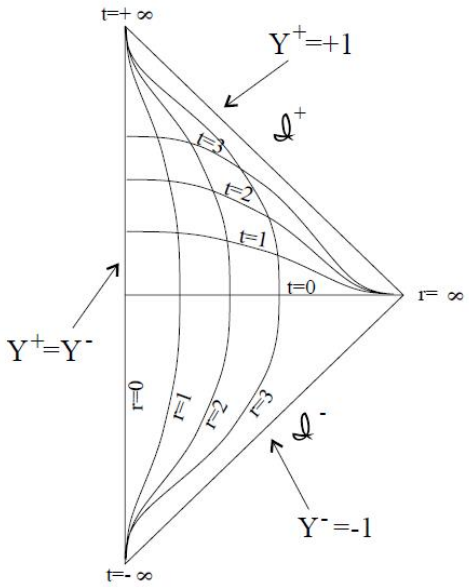


Figure 2: Detailed Penrose diagram for a 2D Minkowski space-time with constant t and r curves [25].

The boundaries of space-time are represented by the 45 degree lines and their endpoints. The curves shown in the above diagram represent lines of constant t and constant r. They taper off at infinity, so they curve even though the lines are straight and parallel in the original spacetime. The boundaries mentioned include the past time-like infinity $t = -\infty$ and the future time-like infinity $t = +\infty$. Space-like infinity, $r = \infty$ becomes a single point in the diagram. In this diagram $Y^+ = 1$ and $Y^- = -1$ represent null or light like infinity. These are the starting and endpoints of light rays, represented by light-like infinities, denoted by $\mathcal{I}^{\pm}$. This is particularly important because we can model photons that enter or escape our black hole horizon. Later we will use them to model incoming matter during the formation of black holes.

2.3 Black Holes

In 1916, Karl Schwarzschild [26] discovered what we know today as the"*Schwarzschild metric*" given by

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{11}$$

where $r_s = 2G_N M$ and M is the mass of the source that creates the gravitational field and G_N is Newton's gravitational constant. The radius r_s is called the Schwarzschild radius [27].

The metric (11) is spherically symmetric and time-independent. Hence the Schwarzschild metric is able to describe the spacetime in empty space which surrounds a spherical static object. Research led by Stephen Hawking [2] suggests there exists an object whose surface is at $r_s = 2G_N M$. Such objects are called static **black holes**. Equation (11) can be written as

$$ds^2 = -\left(1 - \frac{2G_N M}{r}\right) dt^2 + \left(1 - \frac{2G_N M}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{12}$$

Classically the black hole spacetime might not be defined beyond $2G_N M$. Certainly the interior of this region can't classically communicate with the outside region. Notice that when $r = 2G_N M$, a clock at rest will register no time ($dr = d\theta = d\phi = 0$). The surface $r = 2G_N M$ describes a one way membrane since if a spaceship is inside $r < 2G_N M$, then the radial coordinate r acts like the time coordinate t due to the change of signs in (12). Therefore a spaceship would continue to move inward as time progresses (moving inward means to move towards $r = 0$). Such a membrane is called the **event horizon**. This implies that all matter that enters a black hole will eventually fall towards $r = 0$ and so there is an infinitely dense point of mass at $r = 0$ which is surrounded by a curved spherically symmetric space-time given by the Schwarzschild metric.

In the 1970's, quantum mechanical features of black holes were discovered, primarily motivated by results obtained by Stephen Hawking [2]. Black holes are able to radiate photons as if they are a black body with temperature T. The black hole should also have an entropy since they have a temperature and $T = \left(\frac{\partial S}{\partial E}\right)^{-1}$. This is paradoxical since once matter is within $2G_N M$, it will eventually end up at $r = 0$. Classically once you have passed beyond the event horizon, there is no way to return. How can a black hole then radiate energy? Quantum fluctuations create particle-antiparticle pairs which exist for a short time before recombining. Consider such a pair. For a very brief moment the pair can be split by falling on different sides of the black hole's event horizon. Suppose the particle with negative energy is inside the event horizon whilst the particle with positive energy is outside the event horizon. Instead of the usual process whereby the pair recombines, the particle with negative energy now needs to conform to black hole physics and move towards $r = 0$. In the case of the positive energy particle, it will keep traveling towards infinity. The net result is that the black hole gained a particle with negative energy and "lost" a particle with positive energy. In this way, the black hole is said to radiate energy! Such type of radiation is called *Hawking Radiation* [27].

Another important result, due to Bekenstein [13], shows that the entropy S of a black hole is proportional to the area A of its event horizon.

$$S = \frac{k_\beta}{4G_N \hbar} A \tag{13}$$

where k_β is the usual Boltzmann's constant. Formulas that use the holographic principle have been used to give geometrical methods for the computation of the entanglement entropy[14], as well as formulas for the fine grained entropy of an evaporating black hole [19, 20].

One may wonder if the black hole described above is simply a property of the mathematical equations and has nothing to do with physics. Perhaps the black hole is nothing but a bad coordinate choice. Therefore we will discuss other coordinate choices and examine the singularities and event horizon. In order to do this, we will make different coordinate choices so that we can study different patches of spacetime. In some cases one coordinate choice is preferred over another depending on the application.

We know that the event horizon is described by the radial coordinate r. Suppose we consider transformation of r such that the event horizon describes the region $-\infty$ of spacetime. In that case the system of coordinates will only allow us to study the region $r > 2GM$. In this system there is no such thing as a membrane that changes the sign of coordinates and the event horizon is not captured by these new coordinates. This is accomplished by considering a new radial coordinate r^* such that

$$\frac{1}{1-\frac{2GM}{r}}dr^2 = \left(1-\frac{2GM}{r}\right)dr^{*2} \tag{14}$$

The metric now takes the form

$$ds^2 = \left(1-\frac{2GM}{r}\right)(dt^2 - dr^{*2}) - r^2 d\Omega^2 \tag{15}$$

In (15), we see that at $r = 0$ the singularity is still present after the coordinate change. One anticipates tremendous gravitational forces around this point and the geometry still describes a black hole. The only difference here is that we do not know exactly how far we could travel near a black hole and measure signals before we fall into a black hole. The radial coordinate is given as

$$r^* = r + 2GM \log\left(\frac{r-2GM}{2GM}\right) \tag{16}$$

As $r \to 2GM$, $r^* \to -\infty$. This coordinate system is called the tortoise coordinate system. The coordinate r^* forms a mapping so that in tortoise coordinates the new r^* is defined over $-\infty < r^* < \infty$. These coordinates are used to describe light rays in the radial direction.

It is worth reviewing one of the most useful coordinate systems that represent trajectories of photons. To describe this coordinate system, we will follow the steps in [27]. We first need to consider the following transformations

$$\left(\frac{r}{2GM}-1\right)e^{\frac{r}{2GM}} = u^2 - v^2, \qquad t = 2GM \ln|\frac{u+v}{u-v}| \tag{17}$$

By taking derivatives of (17), we are able to get expressions for dt and dr so that we can analyze the resulting metric. Doing the derivatives gives us

$$\begin{aligned} 2udu - 2vdv &= \left(\frac{1}{2GM}e^{\frac{r}{2GM}} + (\frac{r}{2GM} - 1)e^{\frac{r}{2GM}} \times \frac{1}{2GM}\right) dr \\ &= e^{\frac{r}{2GM}}\left(\frac{2GM}{(2GM)^2} + \frac{r}{(2GM)^2} - \frac{2GM}{(2GM)^2}\right) dr \\ &= e^{\frac{r}{2GM}}\frac{r}{(2GM)^2}dr \end{aligned} \tag{18}$$

Therefore

$$dr = e^{-\frac{r}{2GM}}\frac{8(GM)^2}{r}(udu - vdv) \tag{19}$$

Taking the derivative of the time t coordinate gives the following

$$\begin{aligned} dt &= 2GM\left(\frac{-1}{-(u+v)}du + \frac{-1}{-(u+v)dv}\right) - 2GM\left(\frac{-1}{-(u-v)}du + \frac{1}{-(u-v)}dv\right) \\ &= 2GM\left(\frac{du+dv}{u+v} - \frac{du-dv}{u-v}\right) \\ &= \left(\frac{2GM}{u+v} - \frac{2GM}{u-v}\right)du + \left(\frac{2GM}{u+v} + \frac{2GM}{u-v}\right)dv \\ &= \frac{4GM}{u^2-v^2}\left(udv - vdu\right) \\ &= \frac{4GMe^{\frac{-r}{2GM}}}{\left(\frac{r}{2GM} - 1\right)}\left(udv - vdu\right) \end{aligned} \tag{20}$$

where in the last line we used the transformation (17). Lastly, all that is left is to square both (19) and (20) and finally obtain the metric

$$ds^2 = \frac{-32(GM)^3}{r}e^{\frac{r}{2GM}}(dv^2 - du^2) + r^2d\Omega^2 \tag{21}$$

Similarly to the tortoise coordinates, this metric contains finite terms as $r \to 2GM$. However, the singularity at $r = 0$ remains providing further evidence that the singularity is geometrical. In this system v is a time-like coordinate, while u is space-like. Furthermore, radial light worldline's move along trajectories so that $du = \pm dv$. Particles like photons are constrained to follow trajectories so that $dv > |du|$. This choice of coordinates is called Kruskal-Szekeres Coordinates. Figure 3 shows it's features diagrammatically.

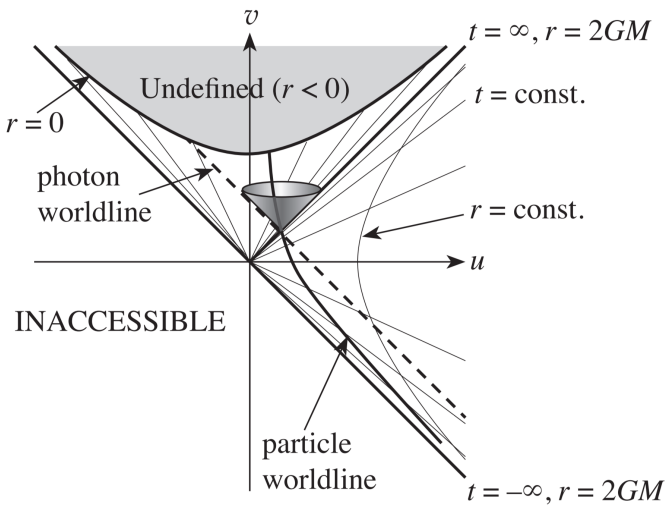


Figure 3: The Kruskal-Szekeres coordinate system defined by coordinates u and v [27].

One is free to explore and invent many possible coordinate systems. The choice of coordinates allows us to focus on patches of spacetime and in some instances, gives a global description. Therefore, the choice of coordinates is considered by the application - which sections of space-time are relevant for the analysis. In some sense, no coordinate choice is a preferred choice without asking what it is that is being studied. The Kruskal-Szekeres system is extensively used in studies of gravity. Therefore when incoming matter and Penrose diagrams are used to model black hole formation, the Kruskal-Szekeres coordinate system quantifies the locations in spacetime.

2.4 Black hole formation

The concept of black hole singularities began with Oppenheimer and Snyder in 1939. They considered one of the first physical processes of black hole formation. Subsequently, Penrose and Hawking developed a set of axioms that describe singularities in the context of general relativity and cosmology. These axioms culminated in the singularity theorems [28], and ultimately one could argue that black hole singularities are a feature of general relativity. One physical process that displays geometric properties like black holes are stars. The lifespan of a star is quite interesting because stars collapse much like the implosion of a sky scraper. One model for the formation of a black hole is by considering the collapse of a thin spherical shell of massless matter. We start by sketching out important geometric properties of a black hole using empty space Penrose diagrams. The Penrose diagram needs to cover the singularity and at most, the event horizon. Figure 4 shows exactly this, although some comments are needed for clarification. The diagram has two singularities, a past and future singularity. The future singularity is found inside a black hole. The past singularity is found

within a white hole. The space on the left is the region needed to correctly extend our Penrose diagram.

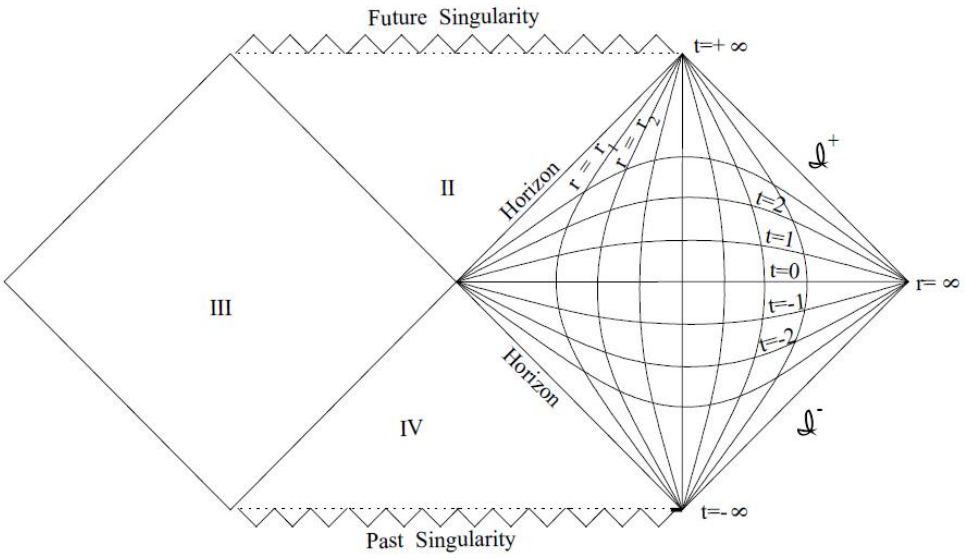


Figure 4: Extended Penrose diagram for a blackhole indicating it's singularities and event horizon [25].

In order to model collapsing gravitational matter, we first consider a light-like infalling ray, with energy E. This starts out in flat Minkowski spacetime, far away from the black hole's event horizon. Notice that the ray is at a 45 degree angle to display light-like infalling matter in Figure 5.

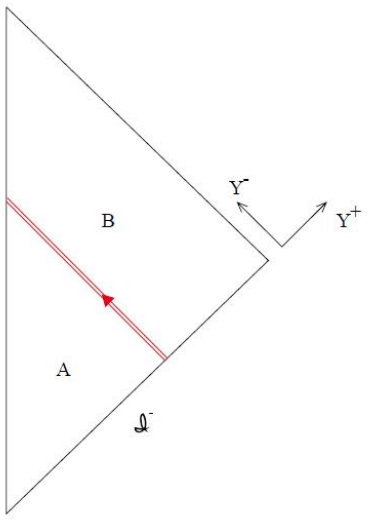


Figure 5: Half of a Penrose diagram in Minkowski space-time with an infalling light-like ray [25].

The ray eventually reaches the boundary of Minkowski spacetime at $r = 0$. To fully describe black hole formation, the ray needs to travel towards an event horizon and subsequently into the singularity. Now we can consider that same ray, but use the Penrose diagram for the black hole - Figure 6 displays the process.

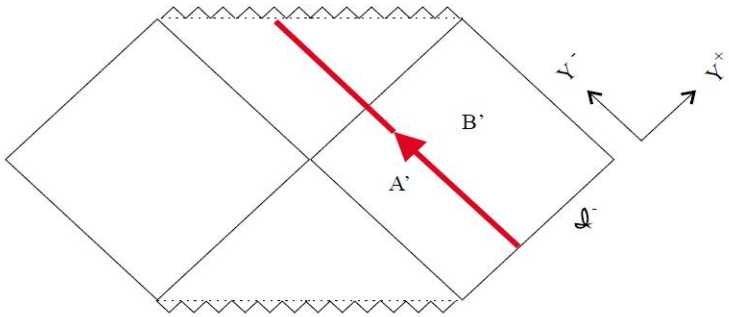


Figure 6: Naive representation of a Penrose diagram of a blackhole with an infalling light-like ray [25].

In order to produce a consistent picture of black hole formation, we need to merge diagrams in Figures 5 and 6. In the first diagram where the ray begins, Figure 5 has region A that is physical and we require all the data in that space. Similarly, we require all the data in space B' in the diagram in Figure 6. Both A and B' are physically significant and we "glue" them in such a way that there is continuity of the radial coordinate r [25]. The resulting diagram is Figure 7, which is the most widely used Penrose diagram for studies of black hole formation and evaporation.

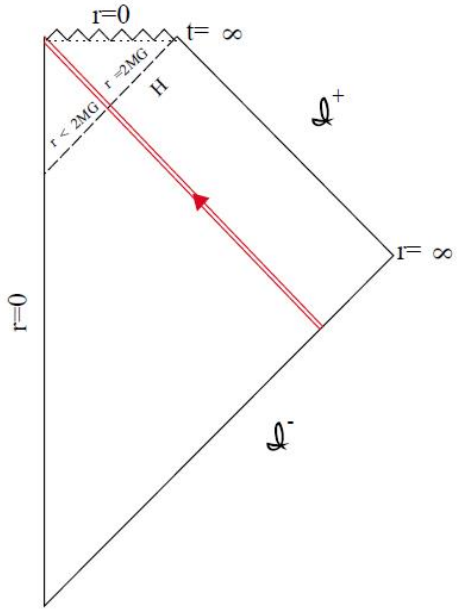


Figure 7: Correct Penrose diagram of a blackhole with an infalling light-like ray [25].

Notice how the infalling matter crosses the horizon at $r = 2GM$ and falls into the singularity at $r = 0$. This model describes the formation of black holes using Penrose diagrams. This diagram is robust because it is able to depict incoming matter that results in black hole formation and can also represent Hawking radiation that escapes the black hole. This subsection concludes the basic aspects of general relativity that is crucial to understanding the black hole information loss paradox. In the next section, we review important aspects about quantum mechanics and how to compute the Page curve.

2.5 Information Loss and the Page Curve

In this section, we review the Page curve. The Page curve gives a precise formulation of the information paradox. Hawking's Page curve showed that the entropy of the radiation increases monotonically. Unitarity of quantum mechanics seems to be violated. In this section we first review some of the basic concepts of quantum states in order to understand where unitarity comes from. We also discuss the density matrix and show how it is used to compute Von Neumann entropy (entanglement entropy).

Properties of a quantum system are described by specific mathematical objects. These objects include vector states $|\psi\rangle$, when acted upon by a position dual vector $\langle x|$ gives the wave function $\psi(x) = \langle x|\psi\rangle$ in position space. The system is also described by the Hamiltonian H, an operator which determines the dynamical evolution of the system, and is equal to the sum of the kinetic and potential energy. Even though matrix mechanics was simultaneously developed to understand quantum theory, the wave function is famous due to it's statistical interpretation by Born [29]. This interpretation implies that our fundamental understanding of nature at the atomic scale is purely probabilistic. One very important concept for the wave function is that a particle in question has to exist somewhere, should we want to study it's dynamics. This statement can be framed mathematically as

$$\int_{-\infty}^{\infty} \psi(x)\psi^*(x)dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1 \tag{22}$$

The equation above implies that the sum of all probabilities for the particle is unity. Therefore in an experiment, a particle has 100% chance of being found. To actually locate the particle's position, we need to perform a measurement. The mathematical quantity we expect to obtain by performing a measurement is called an expectation value and is calculated as follows

$$\langle x\rangle = \langle\psi|x|\psi\rangle = \int_{-\infty}^{\infty} \psi(x)x\psi^*(x)dx \tag{23}$$

Similarly, a measurement of it's momentum is expected to yield

$$\langle p \rangle = \langle \psi | p | \psi \rangle = \int_{-\infty}^{\infty} \psi(x) - i\hbar \frac{d}{dx} \psi^*(x) dx \tag{24}$$

The position x and momentum p have mathematical structure. They are operators (matrices) that act on state vectors $|\psi\rangle$. The structure follows the laws of linear algebra, and here it means that an eigenvector of an operator gives us a state with a precise value of the corresponding observable. The eigenvalues of the operators give us the possible outcomes of a measurement. The statement we follow comes from Sakurai [30]. This idea originated from Dirac, in which we use state vectors. For instance, we can describe the state of a system to be in

$$|\alpha\rangle = \sum_a c_a |a\rangle \tag{25}$$

This is because we don't know for sure which state it's occupying, so we can place it in a linear combination of all possible eigenstates $|a\rangle$. To perform a measurement is to take the system and randomly allocate one of the possible eigenstates to it. The probability for the system to be measured in any one of the eigenstates is

$$P_a = |\langle a | \alpha \rangle|^2 \tag{26}$$

This argument is reminiscent of (22) in that the probability density $\rho(x)$ is given by

$$\rho(x) = |\langle x_a | \psi \rangle|^2 = |\psi(x_a)|^2 \tag{27}$$

The conservation of probability is important because it has no classical analogue and has huge implications for the black hole information loss paradox. The paradox arises because initial pure state data on past null infinity $\mathcal{I}^-$ is supposed to unitary transform to pure state data on future null infinity $\mathcal{I}^+$. As was shown by Stephen Hawking, this does not occur and the data on $\mathcal{I}^+$ is a mixed state. Lately, it's been argued that we don't know for sure if the states on $\mathcal{I}^+$ are indeed mixed. The argument shows that we can only confirm if a state is mixed by calculating corrections to the correlation functions to order $e^{-S/2}$, where S is the entropy for the black hole [31, 32]. The equation of conservation of probability is

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j}(\vec{x}) = 0 \tag{28}$$

where $\vec{j}(\vec{x})$ is the probability current.

One of the direct consequences of the conservation of probability is the concept of unitarity. Unitarity implies that pure states will always evolve in time into other pure states. Quantum field theory is built by constructing dynamics that incorporates unitarity. Given

state vectors in an orthonormal basis $|b_i\rangle$ and $|a_i\rangle$, a unitary operator U acts on the states as follows:

$$\begin{aligned} |b_1\rangle &= U|a_1\rangle \\ |b_2\rangle &= U|a_2\rangle \\ &\cdots \\ |b_N\rangle &= U|a_N\rangle \end{aligned}$$

The operator U is unitary in the sense that we require $UU^\dagger = U^\dagger U = 1$. We build this operator as follows

$$U = \sum_n |b_n\rangle\langle a_n| \tag{29}$$

Lets act with this operator on $|a_m\rangle$:

$$U|a_m\rangle = \sum_n |b_n\rangle\langle a_n|a_m\rangle = \sum_n |b_n\rangle\delta_{nm} = |b_m\rangle$$

If we are able to compute $U^\dagger$, then we could check if the requirement of unitarity has been preserved. Taking the adjoint of U gives

$$\begin{aligned} U^\dagger &= \sum_n (|b_n\rangle\langle a_n|)^\dagger \\ &= \sum_n |a_n\rangle\langle b_n| \end{aligned} \tag{30}$$

Therefore, we have

$$\begin{aligned} U^\dagger U &= \left(\sum_m |a_m\rangle\langle b_m|\right)\left(\sum_n |b_n\rangle\langle a_n|\right) \\ &= \sum_m \sum_n |a_m\rangle\langle b_m|b_n\rangle\langle a_n| \\ &= \sum_n |a_n\rangle\langle a_n| \\ &= 1 \end{aligned} \tag{31}$$

as we hoped to show. The result is not the number 1, but rather the identity operator. Unitary evolution of the state is given by

$$\psi(x,t) = U(t)\psi(x,0)$$

Inserting the above transformation into the Schrodinger equation gives

$$\frac{\partial}{\partial t}\left(U(t)\psi(x,0)\right) = \frac{-i}{\hbar}HU(t)\psi(x,0)$$
$$\frac{\partial U(t)}{\partial t}\psi(x,0) = \frac{-i}{\hbar}HU(t)\psi(x,0)$$
$$\frac{\partial U(t)}{\partial t} = \frac{-i}{\hbar}HU(t)$$

For all closed systems, the Hamiltonian is independent of time, so that we can generally solve the equation above. Since the unitary operator depends on t, we must integrate:

$$\int \frac{dU}{U} = -\frac{i}{\hbar}Hdt$$

$$U(t) = e^{(-i/\hbar)Ht}$$

Our state vector at some time t is described by

$$|\psi(t)\rangle = e^{(-i/\hbar)Ht}|\psi(0)\rangle$$

When the Hamiltonian is also time dependent, we have

$$U(t) = e^{-i/\hbar\int_0^t dtH(t)}$$

and it's adjoint is

$$U(t)^\dagger = e^{i/\hbar\int_0^t dtH^\dagger(t)}$$

What does unitary evolution imply about the Hamiltonian? Unitarity requires that

$$\begin{aligned} U(t)^\dagger U(t) &= e^{i/\hbar\int dt(H^\dagger(t)-H(t))} \\ &= 1 \end{aligned} \tag{32}$$

This will hold only when $H(t) = H^\dagger(t)$. Hence unitary evolution in quantum theory implies that the Hamiltonian needs to be a hermitian operator.

A wave function always describes a pure state, so that a classical mixture (for example) can't be described using a wave function. In 1927, Jon Von Neumann [33] developed a density operator formalism that is able to encode a complete description of physical processes that may correspond to pure or mixed ensembles. In terms of the usual description of the wave function as a state in a Hilbert space, a pure state $|\alpha\rangle$ describes a superposition of the possible states of the physical system [30]. We interpret this state as a usual quantum superposition: the system is in all of these different states at the same time. There is no analog of this in classical mechanics: particles are always in definite states. In contrast to

this, a mixed state is one for which the system is in a definite state - we just don't know which state it is in. Such a mixture can be present in both a classical or a quantum system, and the fact that we don't know for sure which state the system is in reflects our ignorance and not some fundamental uncertainty of nature. The density operator can describe the most general situation for which the system is both a mixture of states, and each state in the mixture can be some superposition of states.

The density operator for a pure state is given by

$$\rho = |\psi\rangle\langle\psi| \tag{33}$$

The density operator has trace 1, corresponding to the fact that the state vectors are correctly normalized. A more crucial property of ρ for any pure state is that it is indempotent,

$$\begin{aligned} \rho^2 &= (|\psi\rangle\langle\psi|)\,(|\psi\rangle\langle\psi|) \\ &= |\psi\rangle\,(\langle\psi|\psi\rangle)\,\langle\psi| \\ &= |\psi\rangle\langle\psi| \\ &= \rho \end{aligned} \tag{34}$$

The time evolution of the density operator can be constructed from the time dependent state vectors $|\psi\rangle$ and the unitary operator $U(t)$:

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle$$

$$\langle\psi(t)| = \langle\psi(0)|U^\dagger(t)$$

Therefore, the density operator is evolved in the following way.

$$\rho(t) = U(t)|\psi(0)\rangle\langle\psi(0)|U^\dagger(t) \tag{35}$$

Suppose a bipartite quantum system consists of two subsystems A and B. A possible wave-function describing the quantum system is

$$|\Psi\rangle = \sum_{\alpha,\beta} c_{\alpha\beta}|\alpha\rangle \otimes |\beta\rangle \tag{36}$$

where α and β are sets of commuting variables that are required to represent subsystems A and B respectively. In order to study a specific subsystem, say subsystem A, the density operator needed to describe subsystem A (also known as the reduced density matrix) is given by

$$\rho_A(\alpha, \alpha') = \sum_{\beta} c_{\alpha\beta}c^*_{\alpha'\beta}|\alpha\rangle\langle\alpha'| \tag{37}$$

Furthermore, the expectation value of an A-operator is defined by

$$\langle a \rangle = \mathrm{tr}(a\rho_A) \tag{38}$$

where tr denotes the trace operation, and $\mathrm{tr}(\rho) = 1$ which is equivalent to saying that the sum of all probabilities is equal to 1. The density matrix is hermitian $\rho = \rho^\dagger$ and non-negative. These requirements indicate that the eigenvalues of ρ_A give the probability that system A is in the corresponding state. Due to entanglement between subsystems A and B, the reduced density matrix ρ_A describes a mixed state even if the full system is in a pure state.

The density operator for a mixed state can also be written as a weighted sum

$$\rho_A^{\mathbf{mixed}} = \sum_i P_i \rho_A^{\mathbf{pure}} = \sum_i P_i |\Psi_i\rangle\langle\Psi_i| \tag{39}$$

with $\sum_i P_i = 1$. The dynamics of the density operator is given by the Liouville - Neumann equation

$$i\frac{\partial \rho}{\partial t} = \frac{1}{\hbar}\left[H, \rho\right] \tag{40}$$

where H is the Hamiltonian.

Now that pure and mixed states have been reviewed, we can introduce the Von Neumann entropy, S. The Von Neumann formula computes the entropy when given the density matrix. It is also called the entanglement entropy S_E when it is evaluated using the reduced density matrix, since it gives a measure of the entanglement between two subsystems. The Von Neumann entropy is given by

$$S = -\mathrm{tr}(\rho \ln \rho) \tag{41}$$

For a pure state, $S = 0$ and for a mixed state $S > 0$. The maximum entanglement entropy is obtained for a totally incoherent mixed density matrix with all eigenvalues equal to $1/D$ with D the dimension of the Hilbert space. This maximum entropy is given by

$$S_{\mathbf{max}} = \ln D \tag{42}$$

For a pure state $\rho^2 = \rho$ so that ρ can be viewed as a projection operator, and it's entropy vanishes. Equation (41) represents entropy that measures the degree of entanglement between subsystems A and B, hence the name entanglement entropy [30]. For a purely mixed state, the Von Neumann entropy reproduces the usual definition of thermodynamic entropy from statistical mechanics, which is given by

$$S_T = -\mathrm{tr}(\rho_{MB} \mathsf{ln} \rho_{MB}) \tag{43}$$

where ρ_{MB} is the Maxwell-Boltzmann probability distribution given by

$$\rho_{MB} = \frac{1}{Z}\mathsf{exp}(-\beta H) \tag{44}$$

with H the Hamiltonian. β is defined using the temperature as $T = \frac{1}{\beta}$ and $Z = tr(\mathsf{exp}(-\beta H))$ is the partition function.

The important concepts in quantum mechanics have been discussed and now can describe the Page curve. As reviewed in the introduction, this is related to the question of whether or not information about the matter that formed a black hole is lost as the black hole evaporates. According to quantum mechanics, the Schrodinger equation takes pure initial states and evolves them into pure final states

$$|\Psi_n^{\mathsf{final}}\rangle = \sum_m S_{nm}|\Psi_m^{\mathsf{initial}}\rangle \tag{45}$$

where S is a unitary operator acting on the Hilbert space of the theory, known as the scattering matrix. The clearest signal of information loss is exhibited in the evolution of pure initial states (before the black hole has formed) given by

$$\rho^{initial} = |\Psi^{initial}\rangle\langle\Psi^{initial}| \tag{46}$$

into mixed final states containing Hawking radiation, given by the density operator

$$\rho^{final} = \sum_i P_i|\Psi_i^{final}\rangle\langle\Psi_i^{final}| \tag{47}$$

The evolution of a density operator is not given by a unitary time evolution operator, but rather by a super operator which maps density matrices into density matrices. Under evolution by the super operator the density matrix remains hermitian and it's trace remains 1 [34]. Super operators are usually written as a dollar ($) as follows

$$\rho_{mm'}^{final} = \$_{mm',nn'}\rho_{nn'}^{initial} \tag{48}$$

Super operators evolve pure states into mixed states, which signals information loss.

The Page curve [4] plots the entanglement entropy as a function of time. As mentioned above, the initial black hole state is pure so that the initial entropy of the system is zero. After the black hole has formed Hawking radiation begins to leave the black hole. As a result, the entanglement entropy between the Hawking radiation and the black hole starts increasing, the entropy of the black hole decreases (it shrinks and its horizon area decreases) and the entropy of the radiation increases. At the retention time, also called Page time, the two entropies are equal. At this time the entanglement entropy reaches its maximum and starts decreasing. The information about the initial state of the system starts coming out in the Hawking radiation. At the end of the evaporation the final state is a pure state of the radiation, which again has vanishing entropy. Thus, the Page curve increases from zero to a maximum and then decreases back down to zero [34]. Figure 8 shows the Page Curve with Hawking's famous result [3], and what is expected from unitary requirements of quantum mechanics.

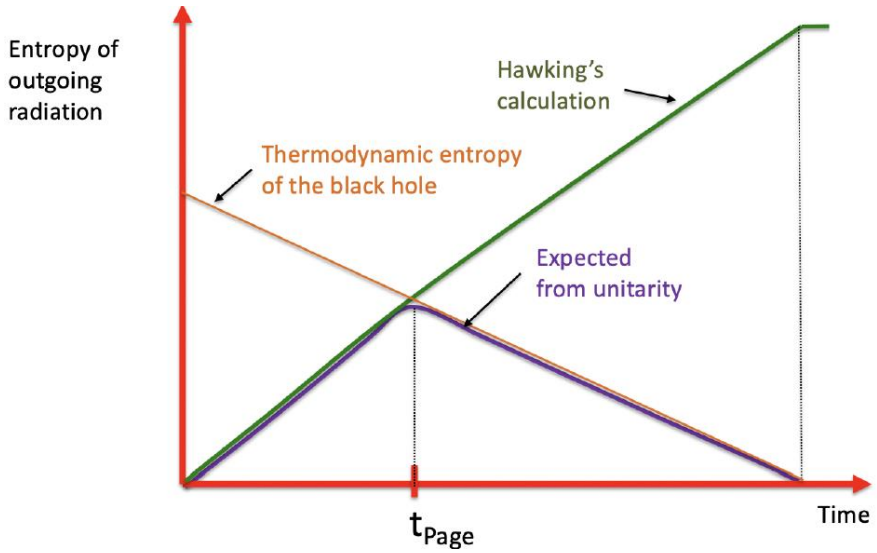


Figure 8: Page Curve indicating the evolution of entropy S with time t, for thermodynamic entropy of a black hole, Hawking's result and what we expect from a unitary quantum theory [35].

Therefore, the Page Curve is a great way to track our progress and provides a goal post in order to solve the black hole information loss paradox. The goal is to produce a Page curve consistent with unitarity. This means that the entropy of the Hawking radiation needs to eventually drop down to zero after the Page time. The recent results provide a calculation that results in a curve that reproduces what is expected from unitarity requirements. Many believe that these recent papers [7, 19, 20] have brought us very close to the end of the paradox, a milestone in theoretical physics. Sophisticated methods have been used and a lot has been computed. The method we focus on is the semi-classical expansion. One of the most important implications of [7] was that they calculated the Page curve entirely from the saddle point of the gravitational and quantum field theory path integral [36]. The saddle point analysis is one of the most important tools we will need to understand. In the next session we discuss and show how to compute the saddle point approximation.

3 Saddle point analysis

In this section we will develop the saddle point analysis by considering how to compute integrals of the form $\int e^{Mf(x)}$ in situations when the constant M is large. We would like to understand the saddle point analysis because it is used to develop the semi-classical expansion of the path integral which is central to our study. As mentioned above, [7] uses this semi-classical expansion to produce results consistent with unitarity. This is important because now we are witnessing old idea's that seem to have far reaching applications. In order to properly understand the saddle point analysis, we begin with the method to evaluate integrals which are of a similar nature to the gravitational path integral.

3.1 Method of saddle point approximation for a generic function

Let's start with a function $f(x)$, Taylor expanded around some point x_0

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots \tag{49}$$

Suppose that this point x_0 is the point at which $f(x)$ has a maximum such that $f(x_0) > f(x)$ $\forall\ x \neq x_0$. In that case $f'(x_0) = 0$ and (49) becomes

$$f(x) = f(x_0) - \frac{|f''(x_0)|}{2!}(x - x_0)^2 + \cdots \tag{50}$$

The saddle point technique will help us evaluate integrals of the form

$$\int_a^b e^{Mf(x)}dx \tag{51}$$

where M is a large constant and $f(x)$ has a maximum at x_0 in the interval $[a, b]$. Inserting (50) into (51) gives

$$\begin{aligned}
\int_a^b e^{Mf(x)}dx &= \int_a^b e^{M(f(x_0) - \frac{|f''(x_0)|}{2!}(x-x_0)^2 + \cdots))}dx \\
&\approx \int_a^b e^{M(f(x_0) - \frac{|f''(x_0)|}{2!}(x-x_0)^2))}dx \\
&= e^{M(f(x_0)} \int_a^b e^{-M\left(\frac{|f''(x_0)|}{2}(x-x_0)^2\right)}dx \\
&\approx e^{M(f(x_0)} \int_{-\infty}^{\infty} e^{-M\left(\frac{|f''(x_0)|}{2}(x-x_0)^2\right)}dx \\
&\approx e^{M(f(x_0)} \sqrt{\frac{2\pi}{M|f''(x_0)|}}
\end{aligned} \tag{52}$$

This approximation is only valid when M is large and where x_0 is the point that $f(x)$ achieves it's maximum. In (52), line 2 comes from the fact that at large M it is accurate to restrict

the expansion up to quadratic terms because the major contribution to the integral comes from the vicinity of x_0. In line 4, we change the limits from $[a, b]$ to $(-\infty, \infty)$ because the integrand is very small outside the interval $[a, b]$ when M is large.

3.2 Applying the saddle point approximation to the integral representation of the Gamma function

In this section we apply the saddle point approximation to evaluate the Gamma function $\Gamma(n)$, for large n, starting from it's integral representation. This is done to give a simple setting in which the saddle point approximation can be developed.

The Gamma function is defined as

$$\Gamma(n) = (n-1)! \tag{53}$$

We can modify (53) by setting $n = N + 1$. Then we have

$$\Gamma(N+1) = N! \tag{54}$$

The integral representation of the Gamma function is given by

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx \tag{55}$$

The integral representation of the Gamma function can then be written as

$$\Gamma(N+1) = \int_0^\infty x^N e^{-x} dx = N! \tag{56}$$

Since $e^{\ln x} = x$, the integrand of (56) can be written as an exponential function, which is useful if we want to evaluate it using the saddle point approximation

$$N! = \int_0^\infty e^{-x + N \ln x} dx \tag{57}$$

Some comments about (57): N is a large constant. The exponential does not take the form $e^{Nf(x)}$, which is required for the saddle point evaluation. To remedy this, we can make a variable substitution which must involve N. Let's try $x = Ny$, so that $dx = Ndy$, and $\ln(x) = \ln(Ny) = \ln N + \ln y$. Then (57) becomes

$$
\begin{aligned}
N! &= \int_0^\infty e^{-Ny+N\ln N+N\ln y} N dy \\
&= \int_0^\infty dy e^{-Ny+N\ln y} N e^{N\ln N} \\
&= \int_0^\infty dy e^{-Ny+N\ln y} N(N^N) \\
&= N^{N+1}\int_0^\infty dy e^{-Ny+N\ln y} \\
&= N^{N+1}\int_0^\infty dy e^{N(\ln y - y)}
\end{aligned}
\tag{58}
$$

All we have done so far, is modifying (57) so that we can use the saddle point approximation when N is large. We can consider a maximum point y_0 within $f(y) = \ln y - y$

$$
\frac{\partial f(y)}{\partial y}|_{y=y_0} = \frac{1}{y} - 1 \tag{59}
$$

so that we find $y_0 = 1$. Furthermore $f''(y) = -\frac{1}{y^2}$, so that $|f''(1)| = 1$ and $f(1) = -1$.

The integration limits in (58) can be changed as in (52) so that we can simplify the integral we need to $(-\infty, \infty)$, and use the Gaussian integral. The final answer becomes

$$
\begin{aligned}
N! &\approx N^{N+1}\sqrt{\frac{2\pi}{N|f''(1)|}} e^{Nf(1)} \\
&\approx N^{N+1}\sqrt{\frac{2\pi}{N}} e^{-N} \\
&\approx N^{N+1/2}\sqrt{2\pi} e^{-N}
\end{aligned}
\tag{60}
$$

We can interpret the result in (60) since we know something about $\ln N!$ for large N from Sterling's approximation. Taking the logarithm of (60) gives

$$
\begin{aligned}
\ln N! &\approx (N + 1/2)\ln(N) + \frac{1}{2}\ln(2\pi) - N \\
&\approx N\ln N - N
\end{aligned}
\tag{61}
$$

In line 1 of (61) we neglect all the smaller terms since N is large. This is Sterling's formula for large N.

3.3 Higher order corrections

In this subsection we compute the size of the corrections that we dropped in equation (52). Recall the first line in (52) with integration limits $(-\infty, \infty)$ is

$$\begin{aligned}\int_{-\infty}^{\infty} e^{Mf(x)}dx &= \int_{\infty}^{\infty} exp[M(f(x_0) - \frac{|f''(x_0)|}{2!}(x-x_0)^2 + a(x-x_0)^3 + b(x-x_0)^4\cdots)]dx \\ &= e^{Mf(x_0)}\int_{-\infty}^{\infty} e^{-\frac{|f''(x_0)|}{2}(x-x_0)^2 + Ma(x-x_0)^3 + Mb(x-x_0)^4\cdots}dx \\ &= e^{Mf(x_0)}\int_{-\infty}^{\infty} e^{-M\frac{|f''(x_0)|}{2}x^2 + Ma\,x^3 + Mb\,x^4 + \cdots}dx\end{aligned} \tag{62}$$

The constants a and b represent derivatives of f evaluated at the saddle point x_0. They come from the Taylor expansion (50)

$$a = \frac{f'''(x_0)}{3!}, \qquad b = \frac{f''''(x_0)}{4!}$$

One could go as far as desired depending on the accuracy that is needed. It should be noted that the higher the power in $(x - x_0)$, the smaller the correction. In this computation we stop at order $(x - x_0)^4$.

The integrand above resembles a quantum field theory in zero dimensions [37]. Zero dimensions physically represent a universe with a single event. The generating functional depends on only one variable x. One way to evaluate this integral is to write it as

$$\begin{aligned} & e^{Mf(x_0)}\int_{-\infty}^{\infty} e^{-M\frac{|f''(x_0)|}{2}x^2 + Ma\,x^3 + Mb\,x^4 + \cdots)}dx \\ = \; & e^{Mf(x_0)}\int_{-\infty}^{\infty} e^{-M\frac{|f''(x_0)|}{2}x^2 + Ma\,x^3 + Mb\,x^4 + \cdots)}dx \\ = \; & e^{Mf(x_0)}\int_{-\infty}^{\infty} e^{-M\frac{|f''(x_0)|}{2}x^2}\left(1 + Ma\,x^3 + Mb\,x^4 + \frac{M^2a^2\,x^6}{2!}\cdots\right)dx \end{aligned} \tag{63}$$

where we have expanded the exponential to get the last line.

The corrections of order b^2, a^4 and a^2b are computed by multiplying out the terms that come from expanding the exponential

$$\begin{aligned}e^{aMx^3+bMx^4+\cdots} = 1+(aMx^3+bMx^4+\cdots)+\frac{1}{2!}(aMx^3+bMx^4+\cdots)^2+\frac{1}{3!}(aMx^3+bMx^4+\cdots)^3+ \\ \frac{1}{4!}(aMx^3+bMx^4+\cdots)^4+\cdots\end{aligned} \tag{64}$$

In this expansion we are able to select any order of M that would be required. The higher order corrections that we are interested in are

$$\frac{b^2M^2x^8}{2}, \qquad \frac{a^4M^4x^{12}}{4!}, \qquad \frac{3a^2bM^3x^{10}}{3!} \tag{65}$$

By manually carrying out the multiplications, we note that the relation of a, b and M is

$$a^n \times b^m \to M^{n+m}$$

The right arrow in the above is used to say that for each power of a (n) and each power of b (m), the constant M will come with a power of $n+m$. Depending on the accuracy we want to work with, we will end up summing a specific number of terms that appear in the expansion. All odd terms can be neglected since they end up giving us zero in the integral (63). We can use this to specify n to be an even integer since a is coupled to x^3. This implies that for every a^n when n is odd, we can ignore those terms for our integral. This can be extended for all the higher terms in the Taylor expansion.

The integrals in (63) are moments of a Gaussian, so they are straightforward to perform. The result of the first four terms are

$$I = e^{M(f(x_0)}\sqrt{\frac{2\pi}{M|f''(x_0)|}}\left(1 + 3bM\frac{1}{(M|f''(x_0)|)^2} + \frac{15}{2}(aM)^2\frac{1}{(M|f''(x_0)|)^3}\right) \tag{66}$$

These corrections are obtained by taking moments of a Gaussian. We are faced with solving integrals of the form

$$\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}x^c dx$$

where c is a number. One could try out integration by parts or solve numerically. Taking moments of a Gaussian requires us to take derivatives with respect to α on both sides. For example,

$$\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}dx = \sqrt{\frac{2\pi}{\alpha}}$$

$$\int_{-\infty}^{\infty} \frac{d}{d\alpha}\left(e^{-\frac{\alpha}{2}x^2}\right)dx = \frac{d}{d\alpha}\sqrt{\frac{2\pi}{\alpha}}$$

$$\int_{-\infty}^{\infty} e^{-\frac{\alpha}{2}x^2}x^2 dx = \sqrt{\frac{2\pi}{\alpha}}\frac{1}{\alpha}$$

We can apply this method of integration for the integral (63) with the corrections defined in (65). This computation results in the following

$$\frac{105}{2}\sqrt{\frac{2\pi}{M|f''(x_0)|}}\frac{M^2b^2}{(M|f''(x_0)|)^4}, \tag{67}$$

$$\frac{15\times 7\times 9\times 11}{4!}\sqrt{\frac{2\pi}{M|f''(x_0)|}}\frac{(aM)^4}{(M|f''(x_0)|)^6}, \frac{15\times 7\times 9\times 3}{3!}\sqrt{\frac{2\pi}{M|f''(x_0)|}}\frac{a^2bM^3}{(M|f''(x_0)|)^5} \quad (68)$$

It turns out that we can reproduce the above results by using the Feynman rules for the 0-dimensional quantum field theory. Our goal is to show that the semi-classical expansion is the usual loop expansion of perturbative quantum field theory.

The "free" part of the "action" is $M\frac{|f''(x_0)|}{2}x^2$ so that we have the following rule for the propagator, i.e. for the line in the Feynman diagram

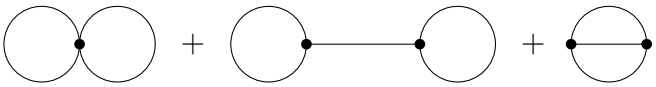
$\frac{1}{M|f''(x_0)|} \longleftrightarrow$ 

There is a cubic interaction and a quartic interaction

$Ma \rightarrow$ 

$Mb \rightarrow$ 

To reproduce the result (66) we need to sum the following vacuum graphs

$1 +$ 

The symmetry factor is a product that combines the following rules:

1. For each closed loop, we have a factor of 2
2. A set of n lines are called equivalent only if they have the same endpoints. However a permutation of the lines results in the same diagram. Therefore they are equivalent. For every set of a equivalent lines, we have a factor of $a!$.
3. For every set of a equivalent vertices, we have a factor $a!$. Vertices connected to distinct external points are not considered.
4. For every set of a equivalent diagrams, we have a factor of $a!$

Using the above rules, we calculate the symmetry factor of the first term. The first term is just 1, which matches with the numerical factor in first term in (66). For the second term, the symmetry factor is 8 because there are two closed lines and a pair of equivalent lines, so this term comes with a coefficient $4!/8 = 3$. This matches with the numerical factor in the second expression in (66). For the third term the symmetry factor is 8 because there are two closed lines and a pair of equivalent vertices, so this term comes with a coefficient $3!3!/8 = 9/2$. For the fourth term the symmetry factor is 12 because there are 3 equivalent lines and a pair of equivalent vertices, so this term comes with a coefficient of $3!3!/12 = 3$. Thus, the third and fourth terms sum to give

$$\frac{9}{2} + 3 = \frac{15}{2}$$

which matches with coefficient of $15/2$ in the third expression in (66). Notice that the correction we found comes with a $1/M$ relative to the leading term, so it is indeed small when M is large. Since M is playing the role of $\hbar^{-1}$ (i.e. M appears in our formula where $\hbar^{-1}$ would have appeared), our loop expansion will now become an expansion in powers of M^{-1}.

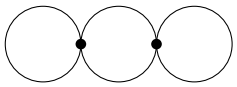
We can find the first correction by summing the following diagrams:

The $(bM)^2$ are for the two vertices, and the $\frac{1}{(M|f''(x_0)|)^4}$ are for the four lines found in the diagrams. What we need to check are the symmetry factors. The first symmetry factor is


$= \frac{(4!)^2}{(2\times2\times2\times2)\times(2\times2)\times2} = \frac{9}{2}$

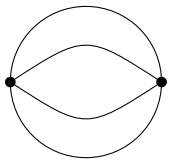
This is because we have four closed loops, 2 sets of 2 equivalent lines and 2 equivalent diagrams.

The next diagram gives us


$= \frac{(4!)^2}{(2\times2)\times2\times2} = 36$

This symmetry factor comes about because there are 2 closed loops, 2 equivalent lines and vertices.

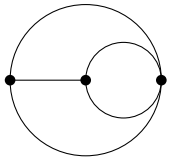
The last diagram gives

 $= \frac{(4!)^2}{(4!)\times 2} = 12$

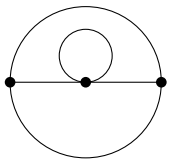
Four equivalent lines and two equivalent vertices. The total symmetry factor for these three diagrams gives

$$\frac{9}{2} + 36 + 12 = \frac{105}{2}$$

which is precisely what we have for the coefficient in (67). We can continue with these rules for the symmetry factor and compute all the higher order corrections. It becomes tedious but doable and we will show another correction which is the second term in (68). The diagrams to sum are the following:

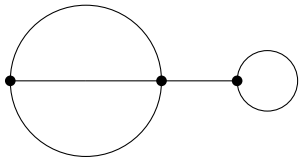
 $= \frac{(3!)^2(4!)}{2\times 2\times 2} = 108$

The factor comes from two 3-leg vertices coupled with 3! and one vertex coupled with 4! (all the following diagrams will come with this factor in the numerator so we won't keep explaining it). There are 2 sets of equivalent lines and two equivalent vertices. The next diagram is

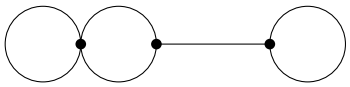
 $= \frac{(3!)^2(4!)}{2\times 2\times 2} = 108$

Here, we have only one closed loop, two equivalent lines and vertices.

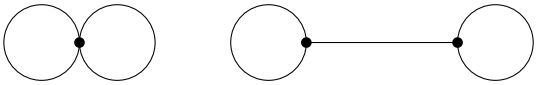
The next diagram could be

 $= \frac{(3!)^2(4!)}{2\times 2\times 3!} = 36$

This diagram has one closed loop, three equivalent lines and two equivalent vertices. The next diagram we have is

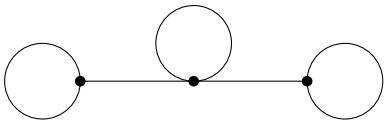
 $= \frac{(3!)^2(4!)}{2\times 2\times 2} = 108$

This diagram has two closed loops and 2 equivalent lines.
Next up is

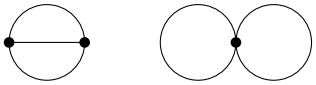
 $= \frac{(3!)^2(4!)}{(2\times 2)\times 2\times (2\times 2)\times 2} = \frac{27}{2}$

The first diagram has 2 closed loops and equivalent vertices, the second diagram also has two closed loops and equivalent lines.

Then we get

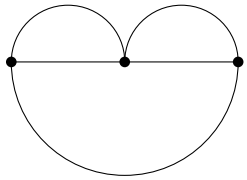
 $= \frac{(3!)^2(4!)}{(2\times 2\times 2)\times 2} = 54$

This diagram has three closed loops and two equivalent vertices. The second last diagram we have is

 $= \frac{(3!)^2(4!)}{(2\times 2)\times 2\times 3!\times 2} = 9$

On the left diagram there are three equivalent lines with two equivalent vertices, the diagram on the right has two closed loops and equivalent lines.

The last[4] diagram we have is



$= \frac{(3!)^2(4!)}{(2\times 2)\times 3!} = 36$

since there are two equivalent lines and three equivalent vertices.

Summing all the symmetry factors for these diagrams gives us

$$108 + 108 + 36 + 108 + \frac{27}{2} + 54 + 9 + 36 = \frac{945}{2}$$

which corresponds to the numerical factor in the second term of (68). We have shown that the corrections can be computed using Feynman diagrams in zero dimensions.

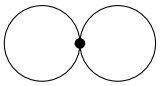
We have learned that the semi-classical expansion is the same as doing perturbation theory in quantum field theory. The only difference is that we expanded about the point $x_0 \neq 0$. In other quantum field theory notes, the expansion is usually done around $x = 0$. The fact that $x_0 \neq 0$ means the field has a classical field strength that is non-zero. We have verified that our methods that worked when the classical field vanished and continue to work when the classical field is non-vanishing. We have more physical insight into the semi-classical expansion.

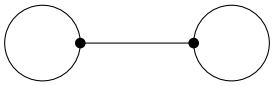
3.4 The semi-classical expansion is the loop expansion: Counting loops and $\hbar$'s

In the previous subsection, we described the semi-classical expansion by computing the saddle point approximation. Each Feynman diagram comes with a definite power of $\hbar$ related to the number of loops in the diagram. There are $\frac{1}{\hbar}$'s for each vertex and $\hbar$ for each propagator. Each vacuum diagram will come with a factor $\hbar^{I-V}$ where I is the number of internal propagators and V are the number of vertices in the diagram. The connected diagrams with loops have a relation $L = I - V + 1$. Therefore our connected diagrams come with $\hbar^{L-1}$ where L is the number of loops within a diagram. In our computation of the saddle point, $\frac{1}{M}$ plays the role of $\hbar$. Therefore counting loops in connected diagrams are equivalent to counting powers of $\frac{1}{M}$ minus 1.

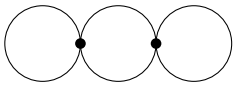
[4]There is no specific order in which is needed to sum the diagram.

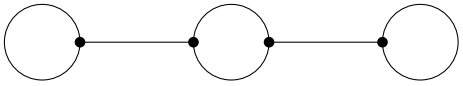
This can be demonstrated with the diagrams we have already summed. Lets study the following diagrams


$= \frac{Mb}{(M|f''(x_0)|)^2} = \frac{1}{M}\frac{b}{(|f''(x_0)|)^2}$


$= \frac{(Ma)^2}{(M|f''(x_0)|)^3} = \frac{1}{M}\frac{a^2}{(|f''(x_0)|)^3}$

Counting the number of loops for both diagrams result in $L = 2$. Therefore we expect to have $\frac{1}{M}$. For the first diagram we see that this is consistent because the first diagram contains one vertex of Mb and two propagators $\frac{1}{M|f''(x_0)|}$. The second diagram is also consistent because it contains two vertices that contribute $(Ma)^2$ and three propagators $\frac{1}{M|f''(x_0)|}$ to connect and close the diagram. Lets examine the following diagrams


$= \frac{(Mb)^2}{(M|f''(x_0)|)^4} = \frac{1}{M^2}\frac{b^2}{(|f''(x_0)|)^4}$


$= \frac{(Ma)^4}{(M|f''(x_0)|)^6} = \frac{1}{M^2}\frac{a^4}{(|f''(x_0)|)^6}$

These two diagrams have three loops, $L = 3$ and we expect the factor $\frac{1}{M^2}$. The first diagram contains two vertices that contribute $(Mb)^2$ and four $\frac{1}{M^4}$ propagators that close and connect the diagram. In the second diagram we find that there are four vertices that contribute $(Ma)^4$ and six $\frac{1}{M^6}$ propagators that connect and close the diagram. Therefore we have shown that both diagrams have the factors $\frac{1}{M^2}$. In conclusion we have demonstrated that the semi-classical expansion is the loop expansion.

3.5 Application: Simple 2D model using the saddle point method

In this subsection we apply the saddle point approximation scheme to a model in 2 dimensions. This model is described by the following action

$$S = \int_{t_1}^{t_2} dt \left(\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + V \right) \tag{69}$$

Consider a general expansion for $x(t)$, namely that

$$x(t) = \bar{x}(t) + \sum_n c_n x_n(t) \tag{70}$$

where $x_n(t)$ is a complete set of orthonormal functions and $\bar{x}(t)$ satisfy the boundary conditions. Furthermore, the functions x_n vanish at those boundary conditions. The first step is to substitute (70) into (69)

$$S[x(t)] = S[\bar{x}(t) + \sum_n c_n x_n(t)] \tag{71}$$

We can immediately recognize that the above equation is in a form that can be Taylor expanded. Using the following for the expansion

$$S[\bar{x}(t) + \sum_n c_n x_n(t)] = \sum_{i=0}^{\infty} \frac{1}{i!} \left(\int dt \sum_n c_n x_n(t) \frac{\delta}{\delta x(t)} \right)^i S[x(t)]|_{x(t)=\bar{x}(t)}$$

The expansion gives the following equation

$$\begin{aligned} S[\bar{x}(t) + \sum_n c_n x_n(t)] = S[\bar{x}(t)] + \sum_n c_n \int dt x_n(t) \frac{\delta S[x]}{\delta x(t)}|_{x=\bar{x}} \\ + \frac{1}{2} \sum_{n_1,n_2} c_{n_1} c_{n_2} \int dt_1 x_{n_1}(t_1) \int dt_2 x_{n_2}(t_2) \frac{\delta^2 S[x]}{\delta x(t_1) \delta x(t_2)}|_{x=\bar{x}} \\ + \frac{1}{3!} \sum_{n_1,n_2,n_3} c_{n_1} c_{n_2} c_{n_3} \int dt_1 x_{n_1}(t_1) \int dt_2 x_{n_2}(t_2) \int dt_3 x_{n_3}(t_3) \frac{\delta^3 S[x]}{\delta x(t_1) \delta x(t_2) \delta x(t_3)}|_{x=\bar{x}} + \\ \frac{1}{4!} \sum_{n_1,n_2,n_3,n_3} c_{n_1} c_{n_2} c_{n_3} c_{n_4} \int dt_1 x_{n_1} \int dt_2 x_{n_2} \int dt_3 x_{n_3} \int dt_4 x_{n_4} \frac{\delta^4 S[x]}{\delta x(t_1) \delta x(t_2) \delta x(t_3) \delta x(t_4)} + \cdots \end{aligned} \tag{72}$$

Since $S[x(t)]$ is a functional, we will consider functional derivatives as seen in the expansion above. We assume that $\hbar$ is small, the integral depends mainly on values in the vicinity of the saddle point (in this case we denote one such saddle as $\bar{x}$). This requires that $\frac{\delta S[\bar{x}]}{\delta x(t)}|_{x=\bar{x}} = 0$ since $S[x(t)]$ will have a local maximum value at $\bar{x}$. In order to compute (72), we first compute the first and higher order functional derivatives. The first functional derivative is

$$\begin{aligned}
\frac{\delta S[x]}{\delta x(t)} &= \int_{t_1'}^{t_2'} dt' \frac{\delta}{\delta x}\left(\frac{1}{2}\left(\frac{dx}{dt'}\right)^2 + V\right) \\
&= \int_{t_1'}^{t_2'} dt' \left(\frac{\delta}{\delta x}\left(\frac{1}{2}\left(\frac{dx}{dt'}\right)^2\right) + \frac{\delta}{\delta x}(V)\right) \\
&= \int_{t_1'}^{t_2'} dt' \left(\frac{dx}{dt'}\frac{\delta}{\delta x}\left(\frac{dx}{dt'}\right) + \frac{\partial V}{\partial x}\frac{\delta x(t')}{\delta x(t)}\right) \\
&= \int_{t_1'}^{t_2'} dt' \left(\frac{dx}{dt'}\frac{d}{dt'}\frac{\delta x(t')}{\delta x(t)} + \frac{\partial V}{\partial x}\frac{\delta x(t')}{\delta x(t)}\right) \\
&= \int_{t_1'}^{t_2'} dt' \left(\frac{dx}{dt'}\frac{d}{dt'}\delta(t'-t) + \frac{\partial V}{\partial x}\delta(t'-t)\right)
\end{aligned} \tag{73}$$

At this point, there is nothing wrong with the last line above. However the sequence of derivatives will look nicer by doing standard integration by parts. This will leave us with an equation that is interesting to interpret. The formula for integration by parts is

$$\int_{-\infty}^{\infty} f\frac{dg}{dt}dt = fg|_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{df}{dt}gdt$$

Considering $\int_{t_1'}^{t_2'} dt' \frac{dx}{dt'}\frac{d}{dt'}\delta(t'-t)$, then $f \equiv \frac{dx}{dt'} \to \frac{df}{dt'} \equiv \frac{d^2x}{dt'^2}$ and $\frac{dg}{dt'} \equiv \frac{d}{dt}\delta(t'-t) \to g \equiv \delta(t'-t)$. Therefore

$$\int_{t_1'}^{t_2'} dt' \frac{dx}{dt'}\frac{d}{dt'} = \frac{dx}{dt'}\delta(t'-t)|_{t_1'}^{t_2'} - \int_{t_1'}^{t_2'} dt' \frac{d^2x}{dt'^2}\delta(t'-t)$$

We then get a more meaningful equation for (73)

$$\frac{\delta S[x]}{\delta x(t)} = \int_{t_1'}^{t_2'} dt' \left(-\frac{d^2x}{dt'^2} + \frac{\partial V}{\partial x}\right)$$

At the saddle point $\bar{x}$

$$\begin{aligned}
\frac{\delta S[\bar{x}]}{\delta x(t)} &= \int_{t_1'}^{t_2'} dt' \left(-\frac{d^2\bar{x}}{dt'^2}\delta(t'-t) + \frac{\partial V(\bar{x})}{\partial x}\delta(t'-t)\right) = 0 \\
&= -\frac{d^2\bar{x}}{dt^2} + \frac{\partial V(\bar{x})}{\partial x}
\end{aligned} \tag{74}$$

The above gives us the first functional derivative which describes the equation of motion for a particle with unit mass that experiences a potential $-V$. The second functional derivative using (74)

$$\begin{aligned}\frac{\delta^2 S[x]}{\delta x(t_2)\delta x(t_2)} &= \frac{\delta}{\delta x(t_2)}\left(-\frac{d^2x(t_1)}{dt^2}+\frac{\partial V(x(t_1))}{\partial x}\right)\\ &= -\frac{d^2}{dt^2}\frac{\delta x(t_1)}{\delta x(t_2)}+V''(x(t_1))\frac{\delta x(t_1)}{\delta x(t_2)}\\ &= -\frac{d^2}{dt^2}\delta(t_1-t_2)+V''(x(t_1))\delta(t_1-t_2)\end{aligned} \tag{75}$$

where V'' indicates a double partial derivative with respect to x. The third functional derivative gives

$$\begin{aligned}\frac{\delta^3 S[x]}{\delta x(t_3)\delta x(t_2)\delta x(t_1)} &= \frac{\delta}{\delta x(t_3)}\left(-\frac{d^2}{dt^2}\delta(t_1-t_2)+V''(x(t_1))\delta(t_1-t_2)\right)\\ &= -\frac{d^2}{dt^2}\frac{\delta}{\delta x(t_3)}(\delta(t_1-t_2))+V'''(x(t_1))\delta(t_1-t_2)\delta(t_1-t_3)\\ &= V'''(x(t_1))\delta(t_1-t_2)\delta(t_1-t_3)\end{aligned} \tag{76}$$

Subsequently the fourth functional derivative is

$$\frac{\delta^4 S[x]}{\delta x(t_4)\delta x(t_3)\delta x(t_2)\delta x(t_1)} = V''''(x(t_1))\delta(t_1-t_2)\delta(t_1-t_3)\delta(t_1-t_4) \tag{77}$$

Now we have the terms that will be substituted in (72). The expansion simplifies due to the orthogonality of the functions x_n. The first and second term in (72) are trivial. The first term is $S[\bar{x}]$ and the second term is zero because $\bar{x}(t)$ satisfies the classical equations of motion. The third term gives us

$$\frac{1}{2!}\sum_{n_1,n_2} c_{n_1}c_{n_2}\int dt_1\int dt_2 x_{n_1}(t_1)x_{n_2}(t_2)\frac{\delta^2 S[x]}{\delta x(t_2)\delta x(t_1)}|_{x=\bar{x}}$$

$$= \frac{1}{2!}\sum_{n_1,n_2} c_{n_1}c_{n_2}\int dt_2\int dt_1 x_{n_1}(t_1)x_{n_2}(t_2)\left(-\frac{d^2}{dt^2}\delta(t_1-t_2)+V''(\bar{x}(t_1))\delta(t_1-t_2)\right)$$

$$= \frac{1}{2!}\sum_{n_1,n_2} c_{n_1}c_{n_2}\int dt_2\int dt_1 x_{n_1}(t_1)\left(-\frac{d^2}{dt^2}\delta(t_1-t_2)+V''(\bar{x}(t_1))\delta(t_1-t_2)\right)x_{n_2}(t_2)$$

$$= \frac{1}{2!}\sum_{n_1,n_2} c_{n_1}c_{n_2}\int dt_2 x_{n_1}(t_2)\left(-\frac{d^2}{dt^2}+V''(\bar{x}(t_2))\right)x_{n_2}(t_2)$$

$$= \frac{1}{2!}\sum_{n_1,n_2} c_{n_1}c_{n_2}\delta_{n_1,n_2}\left(-\frac{d^2}{dt^2}+V''(\bar{x}(t_2))\right)$$

$$= \frac{1}{2!}\sum_n c_n^2 \lambda_n$$

where λ_n is defined by the following eigenvalue problem

$$-\frac{d^2}{dt^2}x_n + V''(\bar{x}(t))x_n = \lambda_n x_n, \qquad x_n(0) = 0 = x_n(t)$$

All of the information we will need about the semi-classical evaluation has been shown. The following terms are correction terms, which give the quantum correction to the leading result. The fourth term is

$$\frac{1}{3!}\sum_{n_1,n_2,n_3} c_{n_1}c_{n_2}c_{n_3}\int dt_1\int dt_2\int dt_3 x_{n_1}(t_1)x_{n_2}(t_2)x_{n_3}(t_3)\frac{\delta^3 S[x]}{\delta x(t_1)\delta x(t_2)\delta x(t_3)}|_{x=\bar{x}}$$

$$= \frac{1}{3!}\sum_{n_1,n_2,n_3} c_{n_1}c_{n_2}c_{n_3}\int dt_1\int dt_2\int dt_3 x_{n_1}(t_1)x_{n_2}(t_2)x_{n_3}(t_3)\left\{V'''(\bar{x}(t_1))\delta(t_1-t_2)\delta(t_1-t_3)\right\}$$

$$= \frac{1}{3!}\sum_{n_1,n_2,n_3} c_{n_1}c_{n_2}c_{n_3}\int dt_1\int dt_2\left\{\int dt_3 x_{n_3}(t_3)\delta(t_1-t_3)\right\}x_{n_1}(t_1)x_{n_2}(t_2)V'''(\bar{x}(t_1))\delta(t_1-t_2)$$

$$= \frac{1}{3!}\sum_{n_1,n_2,n_3} c_{n_1}c_{n_2}c_{n_3}\int dt_2\left\{\int dt_1 x_{n_1}(t_1)x_{n_3}(t_1)V'''(x(t_1))\delta(t_1-t_2)\right\}x_{n_2}(t_2)$$

$$= \frac{1}{3!}\sum_{n_1,n_2,n_3} c_{n_1}c_{n_2}c_{n_3}\int dt_2 x_{n_1}(t_2)x_{n_3}(t_2)V'''(\bar{x}(t_2))x_{n_2}(t_2)$$

The fifth term is (where we dropped the factor $\frac{1}{4!}$)

$$\sum_{n_1,n_2,n_3,n_4} c_n^4\int dt_1\int dt_2\int dt_3\int dt_4 x_{n_1}(t_1)x_{n_2}(t_2)x_{n_3}(t_3)x_{n_4}(t_4)\frac{\delta^4 S[x]}{\delta x(t_1)\delta x(t_2)\delta x(t_3)\delta x(t_4)}|_{x=\bar{x}}$$

$$= \sum_{n_1,\ldots} c_n^4\int dt_1\int dt_2\int dt_3\int dt_4 x_{n_1}(t_1)x_{n_2}(t_2)x_{n_3}(t_3)x_{n_4}(t_4)\left\{V''''(\bar{x}(t_1))\delta(t_1-t_2)\delta(t_1-t_3)\delta(t_1-t_4)\right\}$$

$$= \ldots\int dt_2\int dt_1 x_{n_1}(t_1)x_{n_2}(t_2)\left\{\int dt_3 x_{n_3}(t_3)\delta(t_1-t_3)\right\}\left\{\int dt_4 x_{n_4}(t_4)\delta(t_1-t_4)\right\}V''''(\bar{x}(t_1))\delta(t_1-t_2)$$

$$= \ldots\int dt_2\left\{\int dt_1 x_{n_1}(t_1)x_{n_3}(t_1)x_{n_4}(t_1)V''''(\bar{x}(t_1))\delta(t_1-t_2)\right\}x_{n_2}(t_2)$$

$$= \sum_{n_1,n_2,n_3,n_4} c_n^4 \int dt_2 x_{n_1}(t_2) x_{n_3}(t_2) x_{n_4}(t_2) V''''(\bar{x}(t_2)) x_{n_2}(t_2)$$

Our goal in this chapter is to compute the path integral when $\hbar$ is small. The saddle point approximation requires us to use the first term $S[\bar{x}]$ and the third term.

$$\begin{aligned} Z &= \mathcal{N} \int [dx] e^{-\frac{S}{\hbar}} \\ &= \mathcal{N} \int \prod_n (2\pi\hbar)^{-1/2} e^{-\left\{ \frac{S[\bar{x}]}{\hbar} - \frac{\lambda_n c_n^2}{2\hbar} \right\}} dc_n \\ &= \mathcal{N} e^{-\frac{S[\bar{x}]}{\hbar}} (2\pi\hbar)^{-1/2} \int \prod_n e^{-\frac{\lambda_n c_n^2}{2\hbar}} dc_n \\ &= \mathcal{N} e^{-\frac{S[\bar{x}]}{\hbar}} (2\pi\hbar)^{-1/2} \prod_n \sqrt{\frac{2\pi\hbar}{\lambda_n}} \end{aligned} \tag{78}$$

We have freely used the very well known identity that the Dirac delta function (for a constant a) obeys

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a)$$

In conclusion, we have demonstrated that the corrections to the semi-classical expansion can be reproduced by using Feynman diagram techniques. Consequently, the semi-classical expansion is well understood when we keep the sub-leading terms in our expansion. Hence we now understand the finer details needed to evaluate the gravitational path integral.

3.6 Thermodynamics, the Partition function and it's relation to the Euclidean path integral

In this section we will review the connection between the Euclidean path integral and thermodynamics. This will serve as useful background for the following section where we study the classical saddle point and black hole thermodynamic properties. These properties include thermodynamic entropy S and the average energy of a system E. For this section, any factor $\frac{1}{k_B T}$ will be written as β. We have in mind a thermodynamic system with n microstates that is in thermal contact with a heat reservoir at temperature T. Such systems are treated using the canonical formalism in statistical mechanics. This system is given by the famous Boltzmann's distribution which comes with the Boltzmann's factor $e^{-E_n\beta}$. The sum over all microstates of the system gives the partition function Z.

$$\begin{aligned} Z &= \sum_n e^{-E_n\beta} \\ &= \sum_n \langle n|e^{-E_n\beta}|n\rangle \\ &= \sum_n \langle n|e^{-H\beta}|n\rangle \\ &= tr(e^{-H\beta}) \end{aligned} \tag{79}$$

The thermal partition function describes how the energy is distributed over the accessible states of the system.

There is a useful connection between statistical mechanics and the path integral which we will now review. The path integral is given by Feynman's sum over histories which can be written as

$$\langle \phi_f|e^{\frac{-tH}{\hbar}}|\phi_i\rangle = \int_{\phi_i}^{\phi_f}[d\phi]e^{\frac{iS}{\hbar}} \tag{80}$$

where H is the Hamiltonian and S is the action of the system we want to study. We can obtain the Euclidean path integral by considering imaginary time τ and the Euclidean action S_E such that $iS \to -S_E$. This gives us the following equation for the Euclidean path integral

$$\langle \phi_f|e^{\frac{-\tau H}{\hbar}}|\phi_i\rangle = \int_{\phi_i}^{\phi_f}[d\phi]e^{\frac{-S_E}{\hbar}} \tag{81}$$

If we take an integral of the above equation with respect to some basis of states with label q and choose the start and endpoints of the path integral to be state q, then we have the following identity

$$\begin{aligned} \int dq \left(\int_q^q [d\phi]e^{\frac{-S_E}{\hbar}}\right) &= \int dq \left(\langle q|e^{\frac{-\tau H}{\hbar}}|q\rangle\right) \\ &= tr\left(e^{\frac{-\tau H}{\hbar}}\right) \end{aligned} \tag{82}$$

This implies that (82) is a partition function with $\tau = \hbar\beta$, so that

$$Z(\beta) = tr(e^{-H\beta}) = \int dq \left(\int_q^q [d\phi]e^{\frac{-S_E}{\hbar}}\right) \tag{83}$$

Such an integral is integrating over periodic paths with period given by $\tau = \hbar\beta$. Hence the relation between the Euclidean path integral and thermodynamics is given by

$$Z(\beta) = \int_{\phi(0)=\phi(\hbar\beta)} [d\phi]e^{\frac{-S_E}{\hbar}} \tag{84}$$

The uncertainties in statistical mechanics and quantum theory arise from fluctuations. By moving to Euclidean space we have shown a relationship between the thermal fluctuations in statistical mechanics and the quantum fluctuations in quantum theory. This relation is important because in the next subsection we will show how black hole thermodynamics arise from the black hole saddle point.

3.7 Black hole Saddle

In the last two subsections we studied the semi-classical expansion and discovered ways to evaluate the higher order corrections to the leading semi-classical result. We also showed how statistical mechanics is related to quantum theory. This was accomplished by showing a relation between the path integral and the partition function. In this subsection we will discover how a **gravitational path integral** can lead to black hole thermodynamics [36].

The procedure followed when quantizing classical fields consists of the canonical quantization scheme, or Feynman's path integral. The vacuum expectation value of time ordered fields is given by a path integral as follows

$$\langle 0|T(\phi(x_1)...\phi(x_n))|0\rangle = \int [d\phi] e^{iS} \phi(x_1)...\phi(x_n) \tag{85}$$

The T represents the complete set of time ordered correlation functions. The measure is $[d\phi]$ is the sum over all fields and S is the action. The gravitational path integral includes the matter field $\phi(x)$ and metric $g_{\mu\nu}$ [36]

$$\langle g_2, \phi(x_2), s_2 | g_1, \phi(x_1), s_1\rangle = \int D[g,\phi] e^{iS_E[g,\phi]} \tag{86}$$

where $D[g,\phi]$ is the measure for field configurations g and ϕ, $S_E[g,\phi]$ is the action, and the integral is taken over all matter fields ϕ and gravitational fields g. Equation (86) is the amplitude for describing the transition from a state with metric g_1 and matter fields $\phi(x_1)$ on a surface s_1, to a state 2 specified by its metric g_2 and matter field $\phi(x_2)$ on surface s_2 [36]. This equation sums all possible configurations of g and ϕ in between surfaces s_2 and s_1. It is also assumed that s_1, s_2 and the region between them is compact, or alternatively that g and ϕ go to zero sufficiently fast at spatial infinity. This assumption is met by joining both surfaces via a time-like tube at large radius so that the region contained in the boundary is compact [36].

The relation between thermodynamics and the path integral implies that

$$Z(\beta) = \langle g_2, \phi(x_2), s_2 | g_1, \phi(x_1), s_1\rangle = \int D[g,\phi] e^{iS_E[g,\phi]} \tag{87}$$

where $Z(\beta)$ is the thermal partition function, which is now described by a path integral with boundary conditions such that Euclidean time is a circle of proper size β [38] where $\beta = \frac{1}{T}$. The action S_E is given by the Einstein-Hilbert action

$$S_E[g,\phi] = \frac{-1}{16\pi G_N}\int \sqrt{g}(R+...) + \text{boundary terms} \tag{88}$$

Our goal is to evaluate the path integral. This cannot be computed exactly so we will have to use approximation methods developed in section 3, which is the saddle point approximation

$$Z(\beta) \approx exp\left(-S_E[\bar{g},\bar{\phi}] + S^{(1)} + ...\right) \tag{89}$$

where $\bar{g}$ and $\bar{\phi}$ are solutions to the classical equations of motion and $S^{(n)}$ is the n'th correction. This is the semi-classical approximation to the path integral. To leading order in (89) we have

$$\log Z(\beta) \approx -S_E[\bar{g}] \tag{90}$$

where the matter fields are dropped from the background.

The Gibbons-Hawking-York term is one of the boundary terms mentioned in (88) and is given by [36]

$$S_E[g,\phi] = \frac{-1}{16\pi G_N}\int_M \sqrt{g}R - \frac{1}{8\pi G_N}\int_{\partial M}\sqrt{h}K \tag{91}$$

The integrand of the path integral is specified by (91). The starting point is to specify a spatial boundary at some finite distance, say $r = r'$. As usual $g = det(g_{\mu\nu})$, R is the curvature scalar, G_N is Newtons gravitational constant. The boundary term is chosen so that h is the metric of a hypersuface Σ on boundary ∂M and K is the measure of extrinsic curvature given by $K = h^{ij}K_{ij}$, where the repeated indices represent the usual Einstein summation [36, 39]. The expression K_{ij} is given by

$$K_{ij} = \frac{1}{2}\mathcal{L}_n h_{ij} = \frac{1}{2}U^n\partial_n h_{ij} \tag{92}$$

where U^n is the unit normal vector to the hypersurface and the partial derivative is applied in the normal direction at specified data. This follows from the fact that the i and j components of the metric are scalar functions. This is how we compute the Lie derivative of an ordinary function [39].

In (91) we set $R = 0$ for the vacuum Einstein equations and we are left with

$$S_E[g,\phi] = -\frac{1}{8\pi G_N}\int_{\partial M}\sqrt{h}K \tag{93}$$

The above equation can be evaluated for various metrics. All we are required to do at this moment is to calculate $\sqrt{h}$ and K. Consider a Schwarzschild Euclidean black hole described by metric [27]

$$ds^2 = \left(1 - \frac{2GM}{r}\right) d\tau^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{94}$$

This is different from the usual Schwarzschild metric because it is the Euclidean version, so have replaced $t \to i\tau$. In order to evaluate (91) we need to specify a surface, for example $r = r'$ so that we have the following induced metric on the hypersurface

$$h_{ij} = \begin{pmatrix} (1 - \frac{2GM}{r'}) & 0 & 0 \\ 0 & r'^2 & 0 \\ 0 & 0 & r'^2 \sin^2\theta \end{pmatrix} \tag{95}$$

The above matrix is a diagonal 3 by 3 matrix and the determinant gives us

$$det(h_{ij}) = (1 - \frac{2GM}{r'}) r'^4 \sin^2\theta \qquad , \sqrt{h} = \sqrt{det(h_{ij})} = \sqrt{1 - \frac{2GM}{r'}} r'^2 \sin\theta \tag{96}$$

We can calculate K by computing the inverse metric components of h_{ij}. Since h_{ij} is diagonal, it is straightforward.

$$h^{\tau\tau} = \left(1 - \frac{2GM}{r'}\right)^{-1}, h^{\theta\theta} = \frac{1}{r'^2}, h^{\phi\phi} = \frac{1}{r'^2 \sin^2\theta} \tag{97}$$

The components of K_{ij} require more detail since all of the components depend on the unit normal vector to the hypersurface. In general, the unit normal vector is given by [39]

$$U_n = \frac{\partial_n f}{\sqrt{|g^{\alpha\beta}\partial_\alpha f \partial_\beta f|}} \tag{98}$$

where $U^n U_n = \pm 1$ depending on whether the hypersurface specified is time-like or space-like. In our chosen hypersurface, we set $f = r$, so that the normal vector gives

$$\begin{aligned} U_n &= \frac{\partial_n r}{\sqrt{|g^{\alpha\beta}\partial_\alpha r \partial_\beta r|}} \\ &= \frac{\partial_n r}{\sqrt{g^{rr}}} \end{aligned} \tag{99}$$

We can raise indices with the metric

$$U^n = U_\alpha g^{\alpha n} = \frac{\partial_\alpha r}{\sqrt{g^{rr}}} g^{\alpha n} \qquad , U^r = \frac{\partial_\alpha r}{\sqrt{g^{rr}}} g^{rr} = \partial_\alpha r \sqrt{g^{rr}} = \partial_\alpha r \sqrt{1 - \frac{2GM}{r'}} \tag{100}$$

Calculation for the components of K are

$$K_{\tau\tau} = \frac{1}{2} U^r \times \partial_r h_{\tau\tau} = \frac{1}{2} \sqrt{1 - \frac{2GM}{r'}} \times \frac{2GM}{r'^2} = \frac{GM}{r'^2} \sqrt{1 - \frac{2GM}{r'}}$$

$$K_{\theta\theta} = \frac{1}{2} U^r \times \partial_r h_{\theta\theta} = \frac{1}{2} \sqrt{1 - \frac{2GM}{r'}} \times 2r' = r' \sqrt{1 - \frac{2GM}{r'}}$$

$$K_{\phi\phi} = \frac{1}{2}U^r \times \partial_r h_{\phi\phi} = \frac{1}{2}\sqrt{1-\frac{2GM}{r'}} \times 2r' \sin^2\theta = r' \sin^2\theta \sqrt{1-\frac{2GM}{r'}}$$

Therefore

$$K_{\tau\tau} = \frac{GM}{r'^2}\sqrt{1-\frac{2GM}{r'}}, K_{\theta\theta} = r'\sqrt{1-\frac{2GM}{r'}}, K_{\phi\phi} = r' \sin^2\theta \sqrt{1-\frac{2GM}{r'}} \tag{101}$$

With the components given by (97) and (101), the extrinsic curvature K is

$$\begin{aligned} K &= h^{\tau\tau}K_{\tau\tau} + h^{\theta\theta}K_{\theta\theta} + h^{\phi\phi}K_{\phi\phi} \\ &= \left(\left(1-\frac{2GM}{r'}\right)^{-1}\frac{GM}{r'^2}\sqrt{1-\frac{2GM}{r'}}\right) + \left(\frac{1}{r'^2}r'\sqrt{1-\frac{2GM}{r'}}\right) + \left(\frac{1}{r'^2\sin^2\theta}r'\sin^2\theta\sqrt{1-\frac{2GM}{r'}}\right) \\ &= \frac{GM}{r'^2}\left(1-\frac{2GM}{r'}\right)^{\frac{-1}{2}} + 2r'^{-1}\left(1-\frac{2GM}{r'}\right)^{\frac{1}{2}} \end{aligned} \tag{102}$$

In order to calculate (93), it's simpler to first compute

$$\begin{aligned} \sqrt{h}K &= \sqrt{1-\frac{2GM}{r'}}r'^2\sin\theta \times \left(\frac{GM}{r'^2}\left(1-\frac{2GM}{r'}\right)^{\frac{-1}{2}} + 2r'^{-1}\left(1-\frac{2GM}{r'}\right)^{\frac{1}{2}}\right) \\ &= \sin\theta(-3GM+2r') \end{aligned} \tag{103}$$

The action is now

$$\begin{aligned} S_E[g,\phi] &= -\frac{1}{8\pi G}\int_{\partial M}\sqrt{h}K \\ &= -\frac{1}{8\pi G}\int_0^\beta d\tau\int_0^{2\pi}d\phi\int_0^\pi d\theta\sin\theta(-3GM+2r') \\ &= \frac{3\beta M}{2} - \frac{r'\beta}{G} \end{aligned} \tag{104}$$

The boundary conditions have been chosen on the geometry where $r \to \infty$. Hence the action above is divergent. We can regulate this divergence by adding another counter term to the action [36, 38]

$$S_{counter} = \frac{1}{8\pi G}\int_{\partial M}\sqrt{h}K^0 \tag{105}$$

Here, K^0 is the extrinsic curvature of the hypersurface on the same manifold ∂M, embedded in flat space. In order to calculate K^0, the metric we use is a flat space metric defined by [36, 38]

$$ds^2_{\text{flat space}} = \left(1 - \frac{2GM}{r'}\right) d\tau^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{106}$$

Denote the metric of flat space by m_{ij}. The components are

$$m_{\tau\tau} = \left(1 - \frac{2GM}{r'}\right), m_{\theta\theta} = r^2, m_{\phi\phi} = r^2 \sin^2\theta$$

The inverse components are

$$m^{\tau\tau} = \left(1 - \frac{2GM}{r'}\right)^{-1}, m^{\theta\theta} = r^{-2}, m^{\phi\phi} = \frac{1}{r^2 \sin^2\theta} \tag{107}$$

The normal component is simply $U^n = \partial_n r$ since $g^{rr} = 1$. The components for K^0 are as follows

$$K^0_{\tau\tau} = \frac{1}{2} U^r \times \partial_r m_{\tau\tau} = 0$$
$$K^0_{\theta\theta} = \frac{1}{2} U^r \times \partial_r m_{\theta\theta} = r$$
$$K^0_{\phi\phi} = \frac{1}{2} U^r \times \partial_r m_{\phi\phi} = r \sin^2\theta$$

Therefore the components are

$$K^0_{\tau\tau} = 0, K^0_{\theta\theta} = r, K^0_{\phi\phi} = r \sin^2\theta \tag{108}$$

As done above for the extrinsic curvature, K^0 gives

$$\begin{aligned} K^0 &= m^{\tau\tau} K^0_{\tau\tau} + m^{\theta\theta} K^0_{\theta\theta} + m^{\phi\phi} K^0_{\phi\phi} \\ &= \left(1 - \frac{2GM}{r'}\right)^{-1} (0) + (r)(r^{-2}) + r \sin^2\theta \left(\frac{1}{r^2 \sin^2\theta}\right) \\ &= \frac{2}{r'} \end{aligned} \tag{109}$$

The counter term is therefore

$$\begin{aligned} S_{counter} &= \frac{1}{8\pi G} \int_{\partial M} \sqrt{h} K^0 \\ &= \frac{1}{8\pi G} \int_0^\beta d\tau \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta \sqrt{1 - \frac{2GM}{r'}} r'^2 \sin\theta \frac{2}{r'} \\ &= \frac{1}{8\pi G} 8\pi\beta r' \sqrt{1 - \frac{2GM}{r'}} \\ &= \frac{r'\beta}{G} \left(1 - \frac{GM}{r'} + \mathrm{O}(\frac{1}{r'})\right) \end{aligned} \tag{110}$$

where in the last line we used the binomial approximation since r' is chosen at a finite distance where $r' > GM$.

The final action, including the counter term, is

$$\begin{aligned} S_E[g,\phi] &= \frac{3\beta M}{2} - \frac{r'\beta}{G} + \frac{r'\beta}{G}\left(1 - \frac{GM}{r'} + \mathrm{O}(\frac{1}{r'})\right) \\ &= \frac{\beta M}{2} = \frac{M}{2T} = \frac{M}{2} 8\pi GM = 4\pi GM^2 \end{aligned} \tag{111}$$

where in the last line we recall $\beta = \frac{1}{T}$, and Hawking's result for the temperature of a black hole $T = \frac{1}{8\pi GM}$.

So far we have assumed $\kappa_B = \hbar = 1$. From now on, lets also set $G = 1$. From (90) we have that

$$\log Z(\beta) = -S_E = 4\pi M^2 \tag{112}$$

Recall for a thermal system, the average energy E of the system is given by [40]

$$E = -\partial_\beta \log Z \tag{113}$$

We can modify (112) to make all β dependence explicit by using Hawking's result for the temperature of the black hole $\beta = 8\pi M$. Therefore the average energy is given by

$$E = -\partial_\beta \log Z = -\partial_\beta \left(\frac{-\beta^2}{16\pi}\right) = M \tag{114}$$

We can also use the fundamental relation for a thermal system, to compute its entropy using the following formula [40]

$$\kappa_B T \log Z \approx T S_{entropy}(E) - E \tag{115}$$

The entropy is therefore

$$S_{entropy} = \log Z + \beta E = -4\pi M^2 + 8\pi M^2 = 4\pi M^2 \tag{116}$$

This result is in perfect agreement with the entropy of a black hole [27]

$$S_{entropy} = \frac{\kappa_B}{4G\hbar} A \tag{117}$$

where $A = 4\pi(2GM)^2$. This calculation shows that the classical saddle point with Euclidean time circle of size β corresponds to the solution of a Euclidean Schwarzschild black hole. This concludes our discussion on the semi-classical expansion.

4 A possible resolution

4.1 Entanglement puzzles in black hole information loss

The black hole information loss paradox requires us to consider entangled particles split across the event horizon. The entanglement between the two particles has implications when understanding the paradox. In this section we study entanglement by computing CHSH operators [41]. The CHSH operator is given by

$$\mathcal{O}_{AB} = A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2) \tag{118}$$

where we require that A_1 and A_2 are observables for a **local** system A and B_1 and B_2 for a **local** system B. We will assume that these observables are hermitian and live in different Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Therefore we expect the hermitian observables to commute $[A_i, B_j] = 0$ and we choose them so that their eigenvalues are in the range of $[-1, 1]$. Suppose we calculate the expectation value of the CHSH operator on a state $|\psi\rangle$. In this case, classical systems obey $\langle\psi|\mathcal{O}_{AB}|\psi\rangle \leq 2$ while quantum systems obey $\langle\psi|\mathcal{O}_{AB}|\psi\rangle \leq 2\sqrt{2}$ [42]. These systems have different bounds because in a quantum system, the local observable A is entangled with local observable B. Entanglement is a purely quantum mechanical phenomena.

Let's consider a pair of harmonic oscillators. The state $|\psi\rangle$ is given by

$$|\psi\rangle = \mathcal{N} e^{x a_A^\dagger a_B^\dagger} |0\rangle \tag{119}$$

where, a_A, $a_A^\dagger$ and a_B,$a_B^\dagger$ are a pair of creation and annihilation operators. They act on the vacuum state $|0\rangle$ in the following way:

1. The action of $a_A^\dagger a_B^\dagger |0\rangle = |1\rangle_A \otimes |1\rangle_B$
2. ${}_A\langle n|m\rangle_A =_B \langle n|m\rangle_B = \delta_{n,m}$
3. $\left(a_A^\dagger a_B^\dagger\right)^2 |0\rangle = \left(a_A^\dagger a_B^\dagger\right) |1\rangle_A \otimes |1\rangle_B = \sqrt{2}|2\rangle_A \otimes \sqrt{2}|2\rangle_B = 2|2\rangle_A \otimes |2\rangle_B$
4. $\left(a_A^\dagger a_B^\dagger\right)^3 |0\rangle = 6|3\rangle_A \otimes |3\rangle_B$

where in the last result one needs to merely repeat the action of the first point above. These results will be used below

$$
\begin{aligned}
|\psi\rangle &= \mathcal{N} e^{x a_A^\dagger a_B^\dagger}|0\rangle \\
&= \mathcal{N}\left(1 + x a_A^\dagger a_B^\dagger + \frac{x^2}{2}\left(a_A^\dagger a_B^\dagger\right)^2 + \frac{x^3}{3!}\left(a_A^\dagger a_B^\dagger\right)^3 + \cdots\right)|0\rangle \\
&= \mathcal{N}\left(|0\rangle + x a_A^\dagger a_B^\dagger|0\rangle + \frac{x^2}{2}\left(a_A^\dagger a_B^\dagger\right)^2|0\rangle + \frac{x^3}{3!}\left(a_A^\dagger a_B^\dagger\right)^3|0\rangle + \cdots\right) \\
&= \mathcal{N}\left(|0\rangle + x|1\rangle_A \otimes |1\rangle_B + x^2|2\rangle_A \otimes |2\rangle_B + x^3|3\rangle_A \otimes |3\rangle_B + \cdots\right)
\end{aligned}
\tag{120}
$$

The adjoint gives

$$
\langle\psi| = \mathcal{N}^*\left(\langle 0| +_B \langle 1| \otimes_A \langle 1|x +_B \langle 2| \otimes_A \langle 2|x^2 +_B \langle 3| \otimes_A \langle 3|x^3 + \cdots\right) \tag{121}
$$

We therefore get

$$
\begin{aligned}
\langle\psi|\psi\rangle &= \mathcal{N}^2\left(1 +_B \langle 1| \otimes_A \langle 1|x^2|1\rangle_A \otimes |1\rangle_B +_B \langle 2| \otimes_A \langle 2|x^4|2\rangle_A \otimes |2\rangle_B +_B \langle 3| \otimes_A \langle 3|x^6|3\rangle_A \otimes |3\rangle_B + \cdots\right) \\
&= \mathcal{N}^2\left(\langle 0|0\rangle + x_B^2\langle 1| \otimes_A \langle 1|1\rangle_A \otimes |1\rangle_B + x_B^4\langle 2| \otimes_A \langle 2|2\rangle_A \otimes |2\rangle_B + x_B^6\langle 3| \otimes_A \langle 3|3\rangle_A \otimes |3\rangle_B + \cdots\right) \\
&= \mathcal{N}^2\left(1 + x^2 + x^4 + x^6 + \cdots\right) \\
&= \mathcal{N}^2(1 - x^2)^{-1}
\end{aligned}
\tag{122}
$$

In order to have a normalized state $|\psi\rangle$, we set the above equal to one and solve for the normalization constant $\mathcal{N}$. In this case $\mathcal{N} = \sqrt{1 - x^2}$.

Now that we have a well defined state for $|\psi\rangle$, we can compute the CHSH operator. The operator is given by

$$
\mathcal{O}_{AB} = A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2)
$$

where the observables for A_1, A_2, B_1 and B_2 are given by

$$
\begin{aligned}
A_1 &= |0\rangle_{AA}\langle 0| - a_A^\dagger|0\rangle_{AA}\langle 0|a_A \\
A_2 &= a_A^\dagger|0\rangle_{AA}\langle 0| + |0\rangle_{AA}\langle 0|a_A \\
B_1 &= \frac{1}{\sqrt{2}}\left(|0\rangle_{BB}\langle 0| - a_B^\dagger|0\rangle_{BB}\langle 0|a_B + a_B^\dagger|0\rangle_{BB}\langle 0| + |0\rangle_{BB}\langle 0|a_B\right) \\
B_2 &= \frac{1}{\sqrt{2}}\left(|0\rangle_{BB}\langle 0| - a_B^\dagger|0\rangle_{BB}\langle 0|a_B - a_B^\dagger|0\rangle_{BB}\langle 0| - |0\rangle_{BB}\langle 0|a_B\right)
\end{aligned}
$$

In order to compute the CHSH operator, we need the following linear combinations

$$B_1 + B_2 = \sqrt{2}\left(|0\rangle_{BB}\langle 0| - a_B^\dagger|0\rangle_{BB}\langle 0|a_B\right)$$
$$B_1 - B_2 = \sqrt{2}\left(a_B^\dagger|0\rangle_{BB}\langle 0| + |0\rangle_{BB}\langle 0|a_B\right)$$

The ladder operators act as follows

$$a_A^\dagger|0\rangle_A = |1\rangle_A$$

The adjoint gives us

$$\left(a_A^\dagger|0\rangle_A\right)^\dagger =_A \langle 0|a_A$$

Hence

$$_A\langle 0|a_A =_A \langle 1|$$

Therefore, the CHSH operator becomes[5]

$$\mathcal{O}_{AB} = \sqrt{2}(|0\rangle_A|0\rangle_{BB}\langle 0|_A\langle 0| - |0\rangle_A|1\rangle_{BB}\langle 1|_A\langle 0| - |1\rangle_A|0\rangle_{BB}\langle 0|_A\langle 1| + |1\rangle_A|1\rangle_{BB}\langle 1|_A\langle 1| + |1\rangle_A|1\rangle_{BB}\langle 0|_A\langle 0| + |1\rangle_A|0\rangle_{BB}\langle 1|_A\langle 0| + |0\rangle_A|1\rangle_{BB}\langle 0|_A\langle 1| + |0\rangle_A|0\rangle_{BB}\langle 1|_A\langle 1|) \quad (123)$$

Since

$$\langle\mathcal{O}_{AB}\rangle = Tr\left(\rho\mathcal{O}_{AB}\right)$$

we know

$$\begin{aligned} Tr\left(\rho\mathcal{O}_{AB}\right) &= \sum_b \langle b|\psi\rangle\langle\psi|\mathcal{O}_{AB}|b\rangle \\ &= \sum_b \langle b|\psi\rangle\langle\tilde{\psi}|b\rangle \\ &= \langle\psi|\tilde{\psi}\rangle \end{aligned} \quad (124)$$

where $\sum_b |b\rangle\langle b| = \mathbf{1}$ and $|\tilde{\psi}\rangle = \mathcal{O}_{AB}^\dagger|\psi\rangle$. It is a simple exercise to compute $\mathcal{O}_{AB}^\dagger|\psi\rangle$ given the above result for the CHSH operator. The first step is finding the adjoint of $\mathcal{O}_{AB}$. Since it contains terms with two bra's and two ket's each, we need to use the result

$$\begin{aligned} (|a\rangle|b\rangle\langle c|\langle d|)^\dagger &= \left(\langle c|\langle d|\right)^\dagger\left(|a\rangle|b\rangle\right)^\dagger \\ &= |d\rangle|c\rangle\langle b|\langle a| \end{aligned} \quad (125)$$

This tells us that all we need to do for the adjoint of $\mathcal{O}_{AB}$ is to insert the states in the reverse order. This gives

[5]Here have dropped the $\otimes$ symbol for simplicity.

$$\mathcal{O}_{AB}^{\dagger} = \sqrt{2}(|0\rangle_A|0\rangle_{BB}\langle 0|_A\langle 0| - |0\rangle_A|1\rangle_{BB}\langle 1|_A\langle 0| - |1\rangle_A|0\rangle_{BB}\langle 0|_A\langle 1| + |1\rangle_A|1\rangle_{BB}\langle 1|_A\langle 1| + |0\rangle_A|0\rangle_{BB}\langle 1|_A\langle 1| + |0\rangle_A|1\rangle_{BB}\langle 0|_A\langle 1| + |1\rangle_A|0\rangle_{BB}\langle 1|_A\langle 0| + |1\rangle_A|1\rangle_{BB}\langle 0|_A\langle 0|) \quad (126)$$

Recall that the state is given by

$$|\psi\rangle = \sqrt{1-x^2}\,(|0\rangle + x|1\rangle_A|1\rangle_B + \cdots)$$

We now obtain the following result for $|\tilde{\psi}\rangle$;

$$\mathcal{O}_{AB}^{\dagger}|\psi\rangle = \sqrt{2}\sqrt{1-x^2}(|0\rangle_A|0\rangle_{BB}\langle 1|_A\langle 1|x|1\rangle_A|1\rangle_B + |1\rangle_A|1\rangle_{BB}\langle 0|_A\langle 0|0\rangle + |0\rangle_A|0\rangle_{BB}\langle 0|_A\langle 0|0\rangle + |1\rangle_A|1\rangle_{BB}\langle 1|_A\langle 1|x|1\rangle_A|1\rangle_B) \quad (127)$$

Using the orthogonality of states, we get

$$\mathcal{O}_{AB}^{\dagger}|\psi\rangle = \sqrt{2}\sqrt{1-x^2}(x|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B + |0\rangle_A|0\rangle_B + x|1\rangle_A|1\rangle_B) \quad (128)$$

Finally we obtain the result

$$\begin{aligned}\langle\psi|\tilde{\psi}\rangle &= \sqrt{2}(1-x^2)\left(\langle 0|x|0\rangle_A|0\rangle_B + \langle 0|0\rangle_A|0\rangle_B +_B \langle 1|_A\langle 1|x|1\rangle_A|1\rangle_B +_B \langle 1|_A\langle 1|x^2|1\rangle_A|1\rangle_B\right)\\ &= \sqrt{2}(1-x^2)(x+1+x+x^2)\\ &= \sqrt{2}(1-x^2)(1+x)^2\end{aligned} \quad (129)$$

where $0 < x < 1$. Hence we observe entanglement between local observables A and B.

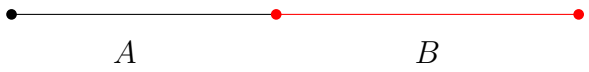


Figure 9: Observable A_1 and A_2 defined in the black region A, entangled with B_1 and B_2 defined in red region B.

Examples of the computation for entanglement using CHSH operators can also be found in [31].

Imagine we add another region to (9), the region C, in green denoted below.

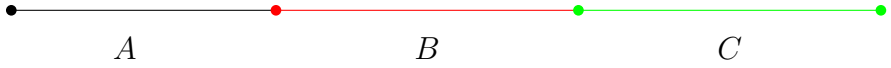


Figure 10: Observables B_1 and B_2 defined in the red region B, entangled with C_1 and C_2 defined in green region C.

Adding the region C which we assume is entangled with B, is problematic because B is already entangled with A. Since region A is entangled with B, we have the corresponding CHSH operator

$$\begin{aligned} \langle\psi|\mathcal{O}_{AB}|\psi\rangle &= \langle\mathcal{O}_{AB}\rangle \\ &= C_{AB} \end{aligned} \tag{130}$$

If region B is also entangled with region C, then we should also have

$$\begin{aligned} \langle\psi|\mathcal{O}_{BC}|\psi\rangle &= \langle\mathcal{O}_{BC}\rangle \\ &= C_{BC} \end{aligned} \tag{131}$$

The papers [43, 44], demonstrate that entanglement is monogamous. It was shown that there is a bound on the entanglement between three regions,

$$C_{AB}^2 + C_{BC}^2 \leq 8 \tag{132}$$

Since both C_{AB} and C_{BC} are entangled, we expect

$$C_{AB}^2 + C_{BC}^2 = (2\sqrt{2})^2 + (2\sqrt{2})^2 = 16 \tag{133}$$

This clearly breaks the bound and establishes the monogamy of entanglement. It is crucial to the black hole information loss paradox. Within the paradox, we could assume states behind the horizon are in region A. The states just outside the horizon would be defined in region B. Therefore due to the monogamy of entanglement, the states far away from the horizon in region C cannot be entangled with the states in region B. This describes a modern version of the black hole information loss paradox. Many of the recent solutions try and address the monogamy of entanglement. In the next subsection we review these solutions and how they attempt to solve the paradox.

4.2 Fuzzballs, Firewalls and the Holography of Information

4.2.1 Issues with the horizon

In the last subsection, we explored the concept of monogamy of entanglement. This concept poses a simple question: How is it possible that the radiation just outside the horizon can

never be entangled with the radiation far away from the horizon? In this subsection we review theories that try to answer this question.

In string theory, we can study black holes by constructing black hole microstates. This can be done by analyzing extremal 2-charge holes in D1-D5 systems at strong coupling. In such systems, it was shown that the D1-D5 system has a geometry with no horizon (and even no singularity) [45, 46, 47, 48]. It was further shown in more generic states that the D1-D5 system has no horizon in [49, 50, 51]. Therefore in studying black holes using string theory, the black holes don't have a horizon or singularity. These kind of black holes are compelling because their surface area satisfies a Bekenstein entropy relation [52]

$$\frac{A}{G} \sim \sqrt{n_1 n_5} \sim S \tag{134}$$

where S is the Bekenstein-Wald entropy constructed by n_1 D1 and n_5 D5 branes. If these black holes don't have a horizon then surely they cannot evaporate. They do not evaporate in the conventional way, however they release energy via eroregion emission [53, 54]. These black hole solutions in string theory have no horizon or singularity. However they have Planck scale fluctuations of supergravity fields which approach a smooth geometry outside of where the classical horizon would be located. Such solutions create the overall picture of a something like a "fuzzball" [55]. Fuzzballs don't have the horizon structure which is essential to the black hole information loss paradox. Hence there is no monogamy of entanglement because there is no horizon structure to distinguish the states that are behind the horizon and outside the horizon. In conclusion, fuzzballs present a possible solution to the black hole information loss paradox.

The idea of using string theoretic computations for the black hole information loss paradox was first considered by [55]. The intuition behind this argument asks a simple but clever question. What is so special about the horizon? Why is there no paradox for a piece of coal? Another approach was taken by [17], where they used the idea of subaditivity of entropy in [55] to draw a different conclusion. The authors in [17], developed what is known as the AMPS paradox, where the abbreviation represents the surnames of the authors: A. **A**lmheiri, D. **M**arolf, J. **P**olchinski and J. **S**ully. The AMPS paradox attempts to make sense of formal statements regarding the black hole information loss paradox. Simultaneously, these statements are inconsistent and result in a unique conclusion. The statements are:

1. The process of black hole evaporation is unitary and quantum theory is enough to analyze this system.

2. The physics outside a black hole can be well approximated by semi-classical field equations.

3. Asymptotic observers can model the black hole as a quantum system where the dimension of the Hilbert space for these states is exponential in the Bekenstein entropy $S(M)$, for a black hole of mass M.

4. A free falling observer experiences nothing unusual when traversing the horizon.

For the statements 1 and 2 to be consistent, [17] leads to the conclusion that infalling observers will experience high-energy modes, a "firewall". If this is indeed the case, then the horizon is not smooth and the correlation functions computed in Hawking's result needs to be revisited. Recently much debates within the black hole information loss paradox has placed the fuzzball proposal against the firewall proposal. This is because both proposals attempt to resolve the paradox uniquely. As we mentioned, the fuzzball proposal says that there is no horizon. The AMPS paradox says that there is a horizon, but it is not smooth. Therefore, each proposal indirectly implies that the other proposal must be wrong.

4.2.2 Gravity stores information differently

Another proposal to the black hole information loss paradox is called the Holography of Information. Due to [56, 57, 58], we have learned that there are no local gauge invariant operators in gravity. This implies that we cannot factorize the Hilbert space separately for $\mathcal{H}_A$, $\mathcal{H}_B$, $\mathcal{H}_C$. It's not possible to act with observables in region A without affecting regions B and C. The degrees of freedom in region C is actually contained in both A and B. This is the main idea behind the Principle of holography [25, 32, 59].

Statement of holography of information: A copy of all information on the bulk of a Cauchy slice is available near the boundary [60].

In order to explain holography of information, suppose we have two distinct states such that

$$|\psi_1\rangle \neq |\psi_2\rangle \tag{135}$$

There is an observable X which will satisfy

$$\langle\psi_1|X|\psi_1\rangle \neq \langle\psi_2|X|\psi_2\rangle \tag{136}$$

which distinguishes between the two states. We choose a basis for the full Hilbert space with elements $|n\rangle$, so that any observable X is constructed in the following way

$$X = \sum_{n,m} c_{n,m}|n\rangle\langle m| \tag{137}$$

For every state $|n\rangle$ in the full Hilbert space, there is a corresponding state x_n in a time band $[0, \epsilon]$ [60]

$$|n\rangle = x_n|0\rangle, \qquad |m\rangle = x_m|0\rangle$$

Therefore, we learn that

$$X = \sum_{n,m} c_{n,m} x_n |0\rangle\langle 0| x_m^\dagger \tag{138}$$

We can measure energy of gravity from infinity by computing the ADM mass which essentially follows from a Gauss Law in AdS. To measure the energy of a state, we need to integrate the subleading terms of the metric in the Fefferman-Graham gauge [60]

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

The boundary conditions imply that the component $h_{r\mu} = 0$ near the boundary, and that $h_{ij} \to \frac{1}{r^{d-2}}$ as $r \to \infty$. We will label the quantity that's being integrated by H, which is the Hamiltonian of the quantum gravity. It's given by

$$H = \frac{d}{16\pi G} \lim_{r\to\infty} \int d^{d-1}\Omega h_{tt} r^{d-2} \tag{139}$$

H is also given by

$$H = \sum_E E P_E \tag{140}$$

where P_E is the projection onto a subspace of states where H has eigenvalues E. H and P_E are elements of a boundary algebra. The factor $P_0 = |0\rangle\langle 0|$ represents the projection onto the vacuum, which is defined in $[0, \epsilon]$.

This therefore gives us

$$X = \sum_{n,m} c_{n,m} x_n P_0 x_m^\dagger \tag{141}$$

In the equation above, all the operators exist within $[0, \epsilon]$. Therefore in gravity[6] the operator X exists in $[0, \epsilon]$. This implies that since X is defined in the small time band and it is able to give us distinct observables. We conclude that all information in the bulk (distinct observables) are equivalent to the information on the boundary in the small time band $[0, \epsilon]$. This proposal uniquely solves the paradox by stating that the information behind the horizon, is also outside the horizon.

[6]In this context, we are referring to quantum gravity.

5 Bulk Reconstruction - how to compute operators using the Gauge/Gravity duality

In the last section we presented the possible solutions to the black hole information loss paradox and discussed how they solve the paradox. In this section, we will discuss a famous strategy that is used to evaluate fields for interesting boundary conditions. It is called bulk reconstruction. Bulk reconstruction is sometimes referred to as the HKLL reconstruction, named after Alex **H**amilton, Daniel **K**abat, Gilad **L**ifschytz and David **L**owe. This is because these authors gave a first account [61] for constructing bulk[7] observables and representing them with boundary CFT observables. Since this topic has become somewhat standard, much of what we discuss can also be found in [62, 63, 64].

5.1 Free scalar ϕ in AdS

The "extrapolate dictionary" can be represented as axioms that describe details about the gauge/gravity duality. The first two axioms are:

1. All bulk quantum states $\mathcal{O}_{bulk}|\psi\rangle$ in the Hilbert space $\mathcal{H}_{AdS}$ correspond to boundary quantum states $\mathcal{O}_{boundary}|\psi\rangle$ in Hilbert space $\mathcal{H}_{CFT}$. This statement can be better phrased as an isomorphic mapping between the two Hilbert spaces [62, 63]

$$\psi : \mathcal{H}_{AdS} \to \mathcal{H}_{CFT}$$

2. The boundary value of a bulk operator in AdS is well approximated by a boundary operator in the CFT, on a Cauchy slice Σ [62, 63],

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

The boundary operator $\mathcal{O}(t, \Omega)$ is well understood, since the conformal field theory has a precise definition. A conformal field theory is a quantum field theory that enjoys conformal invariance. The bulk field is still mysterious, since it's supposed to be defined by a nonperturbative quantum gravity. Our understanding of bulk operators in quantum gravity is not exact, but it is well approximated by a semiclassical path integral over local bulk fields [62]. Given a Cauchy slice Σ, the bulk field ϕ is well approximated by the boundary CFT operator.

In what follows, we begin our analysis by constructing the bulk field operator $\phi(r, t, \Omega)$. This is usually done by solving the Klein-Gordon equation

$$(\Box - m^2)\phi = 0$$

[7]The spacetimes we encounter have boundaries. Some of these boundaries are described by asymptotic boundary conditions. Everything within the boundary is called the bulk.

with boundary conditions

$$\lim_{r\to\infty} r^{\triangle}\phi(t,r,\Omega) = \mathcal{O}(t,\Omega) \tag{142}$$

where $\triangle$ is the scaling dimension of the CFT operator. The Klein-Gordon equation is for the free scalar in AdS.

The next step is to solve the equation, namely the equations of motion for a free scalar field in AdS. The procedure is standard and appears in many applications [65]. To quote a few examples, this problem is very similar to solving the one dimensional harmonic oscillator, the cylindrical wave guide in three dimensions, the vibrating membrane described using the wave equation, and the central force problem in quantum mechanics. The procedure goes as follows:

1. Write down the background metric you are working on, in order to find the appropriate differential operator.
2. Use separation of variables ansatz.
3. Transform the equation into familiar/standard forms of differential equations (like the hypergeometric differential equation for example).
4. Often a change of variables is needed to complete the above step.
5. Then solve the equation using the standard solutions of differential equations.

The background metric is AdS. In global coordinates the metric is

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

Following [63], we make an ansatz for the field ϕ, expanding in terms of creation and annihilation operators

$$\phi(t,r,\Omega) = \sum_{nlm}\left(f_{nlm}a_{nlm} + f^*_{nlm}a^\dagger_{nlm}\right)$$

The separation of variables ansatz is

$$f_{nlm}(t,r,\Omega) = e^{-i\omega_{nl}t}X_{nl}(r)Y_{lm}(\Omega)$$

The d'Alembertian operator is

$$\Box = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\right)$$

When this operator is the AdS metric, the Ω component has the spherical Laplacian on a $d-1$ dimensional sphere, which has eigenfunctions

$$\nabla^2_{S^{d-1}} Y_{lm}(\Omega) = -l(l+d-2)Y_{lm}(\Omega)$$

Then the t and r components produce the following differential equation

$$\frac{1}{r^{d-1}}\left(\partial_r\left(r^{d-1}(1+r^2)\partial_r(X_{nl})\right)\right) + \left(\frac{\omega^2}{1+r^2} - \frac{l(l+d-2)}{r^2} - m^2\right)X_{nl} = 0 \tag{143}$$

This is the equation we need to solve. It can be done in two ways: separate this equation and solve for $r \to \infty$ and $r \to 0$ [63]. These conditions simplify the differential equation and note that $r \to \infty$ is in fact the boundary condition. We then patch these solutions together to obtain the full solution. The second way makes a clever choice of changing variables, transforming (143) into the form of a hypergeometric differential equation. After making a suitable change of variables, we get the solution

$$\phi(r,t,\Omega) = \sum_{nlm} a_{nlm}c_{nlm}e^{-i(2n+l+\triangle)t}X_{nl}(r)Y_{lm}(\Omega) + h.c \tag{144}$$

where h.c is the hermitian conjugate of the first term. The function X_{nl} is related to a hypergeometric function and is given by

$$X_{nl} = r^l(1+r^2)^{\frac{\triangle+2n+l}{2}}F_1\left(-n, -\triangle - n + d/2, d/2 + l, -r^2\right)$$

c_{nlm} is a normalization constant. This solution gives a quantized set of energies for a scalar field in AdS because $r \to \infty$ and ω has the form

$$\omega_{nl} = 2n + l + \triangle, \qquad \triangle = d/2 + \sqrt{(d/2)^2 + m^2}$$

A different approach can be found in [66], where global coordinates were used by the authors to make a convenient variable substitution. This approach, that $r = \tan\rho$ is convenient because it transforms the t and r components of the metric into a similar form (differing by a sign). It is then much easier to make a new variable substitution in order to bring the differential equation into the form of the hypergeometric differential equation.

We have solved for the free scalar in AdS. Now we will apply the boundary conditions. We consider the case that $r \to \infty$. As shown in [64] for large r we have two sets of modes. We need normalizable modes which have a $r^{-\triangle}$ fall off. The normalization for these modes N_{nl} can be obtained using the Klein-Gordon inner product. Therefore, applying the extrapolate dictionary (142) to our solved field for large r, we get [63]

$$\mathcal{O}(t,\Omega) = \sum_{nlm}\left(N_{nl}e^{-i\omega_{nl}t}Y_{lm}(\Omega)a_{nlm} + h.c\right) \tag{145}$$

We can explicitly define the creation and annihilation operators in terms of operators of the boundary CFT and some known functions. Since (145) contains both creation and annihilation operators, it is convenient to introduce positive and negative frequency modes [63]

$$\begin{aligned}\mathcal{O}(t,\Omega) &= \mathcal{O}_+(t,\Omega) + \mathcal{O}_-(t,\Omega) \\ &= \sum_{nlm}\left(N_{nl}e^{-i\omega_{nl}t}Y_{lm}(\Omega)a_{nlm} + N_{nl}e^{i\omega_{nl}t}Y^*_{lm}(\Omega)a^\dagger_{nlm}\right)\end{aligned} \tag{146}$$

This will allow us to focus on each term separately. Let's choose to work with the positive frequency modes

$$\mathcal{O}_+(t,\Omega) = \sum_{nlm} N_{nl}e^{-i\omega_{nl}t}Y_{lm}(\Omega)a_{nlm}$$

By choosing appropriate orthonormal and complete functions, we can get an explicit form of the annihilation operator a_{nlm},

$$a_{nlm} = N_{nl}^{-1}\int_{-\pi}^{\pi} dt e^{i\omega_{nl}t}\int d\Omega Y^*_{lm}(\Omega)\mathcal{O}_+(t,\Omega) \tag{147}$$

For this boundary value problem, our solution is periodic in time and we limit the integral from $-\pi$ to π [63, 64]. Substituting this expression into our field gives

$$\begin{aligned}\phi(r,t,\Omega) &= \sum_{nlm}\left(f_{nlm}N_{nl}^{-1}\int_{-\pi}^{\pi} dt e^{i\omega_{nl}t}\int d\Omega Y^*_{lm}(\Omega)\mathcal{O}_+(t,\Omega) + h.c\right) \\ &= \int_{-\pi}^{\pi} dt\int d\Omega\left(\sum_{nlm} f_{nlm}N_{nl}^{-1}e^{i\omega_{nl}t}Y^*_{lm}(\Omega)\right)\mathcal{O}_+(t,\Omega) + h.c \\ &= \int_{-\pi}^{\pi} dt'\int d\Omega' K_+(r,t,\Omega;t',\Omega')\mathcal{O}_+(t',\Omega') + h.c\end{aligned} \tag{148}$$

We can demand that K be real. Since K is not real in general, we can make it real and therefore combine the positive and negative modes. This is possible because we can add any function $e^{i(\triangle-m)t'}$ to K_+ if it integrates to zero against $\mathcal{O}_+$. We can choose any integer m, place the bulk field at $r = t = 0$ and make K real [63]. Now that K can be made real, the positive and negative modes can be combined and the bulk point at $r = t = 0$ can be shifted to any point in the bulk using isometrics of AdS. By taking advantage of this peculiar feature of K, we can make it real and obtain $K(r,t,\Omega;t',\Omega')$ which is commonly called the *smearing function*. It is the key ingredient for representing bulk fields with approximate boundary CFT fields on a Cauchy slice Σ. The authors [61] in the HKLL reconstruction studied this function in different coordinates and discovered the following

1. For global coordinates, the smearing function has support on boundary points which are space-like separated from a point in the bulk.

2. Smearing functions in Poincare coordinates cover a subspace of AdS space called the Poincare patch.

3. When the AdS dimension is even, the smearing function has support at space-like separation. With odd dimension, the smearing function has support over the entire subspace boundary.

4. The most interesting case is in Rindler coordinates, the smearing function has support on a compact region of the complexified geometry. In this case, an analytic continuation is needed taking the spatial coordinates on the boundary to imaginary values.

5. Two-point functions of boundary CFT operators produce bulk two-point functions.

We can write a simple expression for (148) since K can be made real

$$\phi(z) = \int d^d X \sqrt{-\gamma} K(z; X) \mathcal{O}(X) \tag{149}$$

where the bulk point is given by $z = (r, t, \Omega)$ with the corresponding boundary point at $X = (t', \Omega')$. The integration measure is over the boundary that includes the Lorentz invariant factor $\sqrt{-\gamma}$ where $\gamma_{\mu\nu}$ is the metric on the boundary CFT. The integration is over boundary points which are space-like separated from z. The function K is not a function, but in fact a distribution. It's only well defined when integrated against a function, similar to the delta function. The procedure is also based on solving equations of motion, as was shown for the free scalar. We can write the action for a free scalar as

$$\begin{aligned} \phi(x) &= \int dx' \sqrt{-g} \phi(x') \delta^{d+1}(x - x') \\ &= \int dx' \phi(x') \left(\Box' - m^2 \right) G(x, x') \end{aligned} \tag{150}$$

Using the delta distribution, it is convenient to describe a Green's function $G(x, x')$ which is space like separated [67].

$$(\Box' - m^2) G(x, x') = \frac{1}{\sqrt{-g}} \delta^{d+1}(x - x') \tag{151}$$

The Green's function behaves as follows for large values of r [67]

$$G(x, x') \backsim \frac{1}{2 \triangle - d} \left(r'^{\triangle} L(x, X') + r'^{-(d-\triangle)} K(x, X') \right)$$

We can re-derive (149) by doing integration by parts on the second line of (150).

$$\begin{aligned}\phi(x) &= \int dx'\phi(x')(\Box' - m^2)G(x,x') \\ &= \int dX' r^\mu \left(\phi(x')\partial'_\mu G(x,x') - G(x,x')\partial'_\mu\phi(x')\right) + \int dx' G(x,x')(\Box' - m^2)\phi(x') \\ &= \int dX' K(x,X')\mathcal{O}(X') + \int dx' G(x,x')(\Box' - m^2)\phi(x')\end{aligned} \tag{152}$$

where the asymptotic behavior of the Green's function and the extrapolate dictionary are applied in the first term on the second line. Here, we have adopted the notation in [63], where the bulk point is $x = (r, t, \Omega)$ and the boundary point in the CFT is $X' = (t', \Omega')$. To deal with bulk interactions use (152) to produce a perturbation series for $\phi(x)$ written in terms of $\mathcal{O}(X')$ [63, 67]. For example, consider the following interaction with some coupling constant g,

$$(\Box - m^2)\phi = g\phi^2 \tag{153}$$

The bulk interaction term on the RHS can be written in terms of boundary CFT's in the following way

$$g\phi^2 = g\int dX' dX'' K(x,X')K(x,X'')\mathcal{O}(X')\mathcal{O}(X'') \tag{154}$$

Therefore, plugging (154) into (153) gives

$$\phi(x) = \int dX' K(x,X')\mathcal{O}(X') + g\int dX' dX'' dx' K(x,X')K(x,X'')G(x,x')\mathcal{O}(X')\mathcal{O}(X'') + \mathcal{O}(g^2) \tag{155}$$

Therefore we have demonstrated the bulk reconstruction. We have shown how to evaluate observables in the bulk by using boundary CFT's. This is interesting because it indicates that we could (in principle) evaluate the observables behind the horizon, by using boundary conditions of CFT observables far away from the horizon.

5.2 The Causal and Entanglement Wedge

By making a reparamererization of coordinates [63, 64], we can easily bring the AdS metric to the following form

$$ds^2 = -(\rho^2 - 1)d\tau^2 + \frac{d\rho^2}{\rho^2 - 1} + \rho^2 dH^2_{d-1} \tag{156}$$

where $dH^2_{d-1} = d\chi^2 + \sinh^2\chi d^2\Omega_{d-1}$. The new coordinates have range $\rho > 1$ and $-\infty < \tau < \infty$ so that the reparameterization does not make the AdS metric singular. Due to the

range of the new coordinates, the metric no longer describes the full AdS space. Instead, it will cover only a subregion because $\rho > 1$. This subregion in AdS is shown in the Figure below.

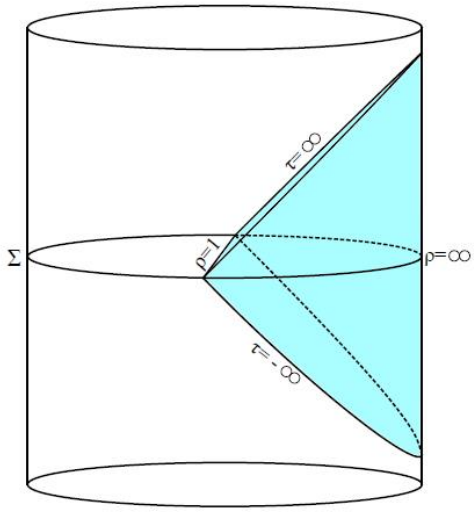


Figure 11: Spacetime in AdS - the bulk physics is defined in a subregion (shaded in blue) due to a change of coordinates [63].

The Figure above is called the AdS-Rindler Wedge. This 'wedge' shaded in blue, contains all the bulk physics for this metric. When we apply the bulk reconstruction, we choose a Cauchy slice Σ. On this slice bulk operators are well approximated by boundary CFT's. Since the bulk is defined within the wedge, we can make a cut along the τ axis at the angle $\theta = 0$. This cut will produce something like a 'blank sheet', with a blue diamond in the center, shown in the figure below.

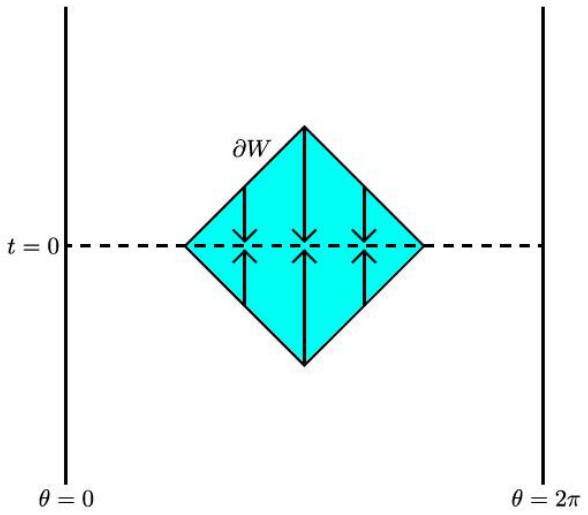


Figure 12: AdS-Rindler Wedge at $d = 2$ [63].

In the Figure above, the 'blank sheet' represents the boundary. The blue diamond extends out of the 'blank sheet', so that ∂W lies on the boundary. The blue diamond in Figure 12 represents a *Causal Wedge*. This is because the blue diamond represents the intersection of all bulk points (past and future) with ∂W. In this setting, ∂W represents the domain of dependence for a boundary spatial subregion X_R. The set of points within ∂W must intersect X_R through any boundary causal curve. Given a boundary spatial subregion X_R of an asymptotic geometry, there is a causal wedge $C[X_R]$ which is the intersection of bulk both future and past of some $D[X_R]$ (domain of dependence) [28, 63].

The boundary spatial subregion X_R contains information about the bulk. It also contains information about the entanglement entropy S for CFT's with holographic dual theories. The entanglement entropy computation in gravity is due to [14]. The Ryu-Takayanagi formalism states that the entanglement entropy of a subregion R of the CFT is proportional to the area of an extremal bulk surface $\mathcal{Q}_R$ homologous to R. The extremal surface and boundary spatial subregion is related by the following

$$\partial H_R = \mathcal{Q}_R \cup R \tag{157}$$

where H_R is a region defined in the bulk. Then, the domain of dependence of H_R is called the entanglement wedge of R

$$D[H_R] = W[R] \tag{158}$$

In conclusion, we have shown how to compute bulk observables by using boundary CFT observables. We have also shown the bulk and boundary of an AdS spacetime on a Penrose diagram. The bulk and boundary sections can be used to define the causal wedge. We finally described where the extremal surface comes from and how the entanglement wedge arises. These are the crucial idea's needed to understand the island proposal in the next section.

6 The Island proposal

The reason why many people believe that the black hole information paradox is close to its solution is because during the last 3 years, we have learned important concepts which allow us to compute the gravitational fine grained entropy. The formula for the gravitational fine grained entropy was modified and work on, which have produced many promising results. In this section we present how entanglement entropy is calculated for gravitational systems. Such calculations involve quantum extremal surfaces, therefore we show how to locate quantum extremal surfaces in JT gravity. Finally, we show how the quantum extremal surfaces and islands produce a unitary Page curve. This is called the Island proposal.

6.1 Entanglement entropy for gravitational systems

The entropy of a black hole S_{BH} is proportional to the area of the event horizon A. The generalized entropy S_{gen} describes the entropy of the black hole as well as the entropy for radiation outside the black hole S_{out},

$$S_{\text{gen}} = \frac{\text{Area of horizon}}{4\hbar G_N} + S_{\text{out}} \tag{159}$$

Using the holographic principle, Ryu and Takayanagi developed a formula for entanglement entropy in gravitational systems [14]. This required that we minimize S_{gen}. This minimization procedure equates to finding a surface that minimizes S_{gen} in the spatial direction and maximizes in the time direction [35]. Therefore the formula for entanglement entropy in gravitational systems searches for an extremal surface X,

$$S_{\text{RT}} = \min_X \left[\text{ext}_X \left(\frac{\text{Area(X)}}{4\hbar G_N} + S_{\text{out}}(X) \right) \right] \tag{160}$$

Refinements to the extremal surface is given by Engelhardt and Wall [18], and the modified version of the entanglement entropy in gravitational systems is given by

$$S_E = \min_\chi \left[\text{ext}_\chi \left(\frac{\text{Area}(\chi)}{4G_N} + S_{\text{semi-classical}}(\Sigma_\chi) \right) \right] \tag{161}$$

where χ is a codimension-2 surface called the quantum extremal surface and Σ_χ is the region bounded by χ. $S_{\text{semi-classical}}(\Sigma_\chi)$ is the entropy calculated from semi-classical gravity. This is the entropy **outside** the black hole. Therefore (161) requires us to find χ by extremizing. We compute the global minimum if there are multiple extremal surfaces χ.

6.2 How to find the quantum extremal surface using AdS/CFT

The formula for the gravitational entanglement entropy [14, 18] allows us to study anti-de Sitter spacetime on a spatial slice. This slice (yellow surface) contains all the physics we

want to study and we can see this diagrammatically in Figure 13

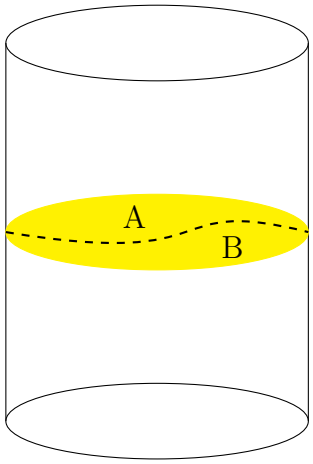


Figure 13: AdS space-time studied on a spatial slice

We split the slice into sections A and B. On A, the surface is of co-dimension two in AdS gravity. The region B is a sub-region of the spatial slice. To describe this region we consider a quantum theory on the boundary, and we use the AdS/CFT correspondence principle [68]. The papers [14, 18] try to find what piece of the bulk is dual to surface B and what is the von Neumann entropy on B. What we have learned so far is that for each region A, there exists the following functional

$$S_{\text{gen}}[A] = \frac{Area(A)}{4G_N} + S_{bulk}[B] \tag{162}$$

where S_{gen} is the generalized entropy and S_{bulk} is the entropy of bulk matter fields in region B. In order to compute the entanglement entropy in region B, we consider

$$S_{EE}[B] = \min_A \left[S_{\text{gen}}[A]\right] \tag{163}$$

where S_{EE} denotes the entanglement entropy. The minimizing surface A^* is called the quantum extremal surface for B.

In this section we consider an example of these definitions in JT gravity with AdS_2 background. The example considers gravity coupled to a dilaton action

$$I_{JT}[\phi, g] = \frac{1}{16\pi G_N}\int d^2x\sqrt{-g}\phi(R+2) + \frac{1}{8\pi G_N}\int dt\sqrt{-\gamma}(K-1)\frac{\phi_r}{\epsilon} \tag{164}$$

The first term in the above equation is the usual JT gravity action coupled to a dilaton ϕ. This is a scalar field and contributes all the dynamics for this system. Here R is the Ricci scalar, the metric is given by $g = \det(g_{\mu\nu})$ and Newtons constant G_N. The constant is dimensionless, whilst ϕ has dimensions of length. The second term in (164) is a boundary

term called the Gibbons-Hawking boundary term. Here, ϕ_r is the boundary value of the dilaton also in dimensions of length, K is the extrinsic curvature defined by the induced metric γ and ϵ is a very small constant representing some ultraviolet cutoff.

By varying (164) with respect to the dilaton ϕ we get

$$\begin{aligned}\frac{\delta}{\delta\phi}I_{JT}[\phi,g] &= \frac{1}{16\pi G_N}\int d^2x\sqrt{-g}\left(\frac{\delta}{\delta\phi}(\phi R+2\phi)\right)+\frac{1}{8\pi G_N}\int dt\sqrt{-\gamma}(K-1)\frac{1}{\epsilon}\frac{\delta\phi_r}{\delta\phi}\\ &= \frac{1}{16\pi G_N}\sqrt{-g}(R+2)\\ &= 0\end{aligned} \tag{165}$$

After varying (164) we set it equal to zero. The result is that $R=-2$, implying that we are in Anti-de Sitter spacetime.

In this problem the boundary conditions are:

$$\gamma_{tt}=\frac{1}{\epsilon^2}\qquad \phi|_{BD}=\frac{\phi_r}{\epsilon}$$

The quantum extremal surface is simply a point[8] in 2 dimensions and so the Area(A) is given by the value of the field at A, $\phi(A)$.

The metric is given by

$$ds^2=\frac{-dt^2+dx^2}{x^2}$$

and we can make an educated guess for the dilaton which gives

$$\phi(t,x)=\frac{-\phi_r}{x}$$

In the above, x was chosen negative $x<0$ for simplicity. We then can introduce lightcone coordinates

$$x^{\pm}=t\pm x$$

The boundary is therefore at $x=-\epsilon$ (denoted in Figure (14) below). The setup is given as follows

Our gravitational system in AdS is dual to some quantum theory on the boundary. One possible candidate could be SYK fermions or a quantum dot that is evolving in time. The dual is shown simply because we compute the entropy in the dual system using holography.

[8]A co-dimension 2 surface implies that it's dimensions are determined by the dimension of the spacetime minus 2. In this case, the JT gravity model is in 2 dimensions. Therefore the quantum extremal surface as zero dimensions and is represented by a single point.

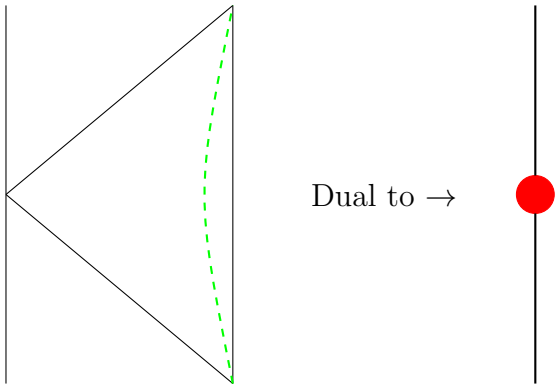


Figure 14: The set up for our gravity theory in AdS space-time and it's holographic dual theory [68].

In order to actually consider black hole evaporation and the resulting matter fields outside the horizon, the author's in [68] considered adding matter fields χ. The action with these matter fields has the following structure

$$I[\phi, g, \chi] = I_{JT}[\phi, g] + I_{CFT}[\chi, g] \tag{166}$$

The entanglement entropy in a CFT_2 for a space-like line segment $A \to B$ is

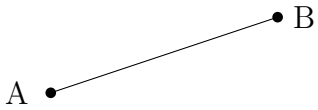


Figure 15: Entanglement entropy calculated on a line segment using holography

The entanglement entropy for this line segment is given by

$$S(AB) = \frac{c}{3} \log\left(\frac{l}{\epsilon_{uv}}\right) \tag{167}$$

The author's in [68] then consider lightcone coordinates and implement an ultraviolet cutoff on proper distance. They get the following equation for entanglement entropy in the CFT

$$S(AB) = \frac{c}{6} \log(\Delta x^+) + \frac{c}{6} \log(\Delta x^-) - \frac{c}{6} \log(\Omega(A)) - \frac{c}{6} \log(\Omega(B)) \tag{168}$$

Here Ω is defined by the metric of the CFT

$$ds^2 = \frac{-1}{\Omega(x^+, x^-)^2} dx^+ dx^- \tag{169}$$

We are however interested in computing the entanglement entropy for the following set up

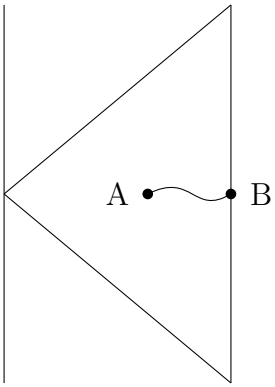


Figure 16: Region to calculate entanglement entropy for the gravity system.

It is natural to suspect that the entanglement wedge covers the entire Poincare patch. So physically we set some arbitrary point A and consider a slice to some point B in the boundary. The points are defined by the coordinates $B(0,0)$ and $A(x,-x)$. They find that

$$\begin{aligned} S_{gen}[x] &= \frac{-\phi_r}{x} + \frac{c}{6}\log(-x) - \frac{c}{6}\log(\Omega(A)) \\ &= \frac{-\phi_r}{x} + \frac{c}{6}\log(-x) - \frac{c}{6}\log(-x) \\ &= \frac{-\phi_r}{x} \end{aligned} \tag{170}$$

$min_A S_{gen}[x]$ is then given by $x \to -\infty$. This calculation gives the point A^*

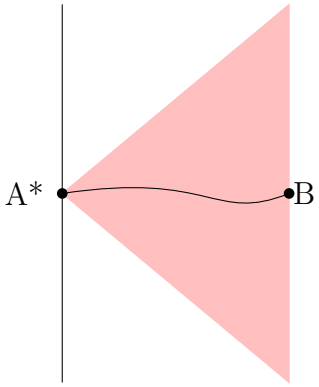


Figure 17: The resulting entanglement wedge for our gravitational system.

This indeed verifies that the entanglement wedge of B covers the entire Poincare patch. Now we may couple the quantum system to a "wire" (the wire could contain a CFT on the boundary). We do this because it can be realized as an attempt to make an AdS black hole evaporate. The system we coupled to can be SYK Fermions on a wire describing a CFT_2.

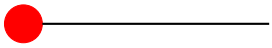


Figure 18: Holographic dual theory which contains the CFT on a wire.

The Penrose diagram for the gravity system with its dual is given in Figure 19

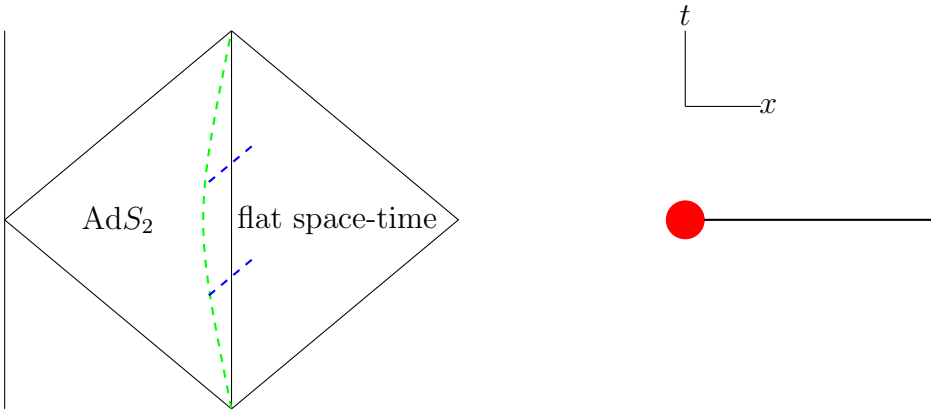


Figure 19: Penrose diagram of the gravity system coupled to a flat space CFT with no gravitons.

On the left hand side diamond we have the usual AdS spacetime. We then "glue" another half diamond which has no dynamical gravity. The boundary represented by the green dotted line is now transparent and will allow for radiation modes (in blue) to cross the boundary and into the flat space. Flat space and allows for transparent boundary conditions between the two spacetimes. The entanglement entropy calculation for this dual system is as follows

$$\begin{aligned} S_{gen}[x] &= \frac{-\phi_r}{x} + 2\frac{c}{6}\log(-x) - \frac{c}{6}\log(\Omega(A)) \\ &= \frac{-\phi_r}{x} + \frac{c}{6}\log(-x) \end{aligned} \tag{171}$$

According to the methods explained at the beginning of this chapter, we can now take the derivative of this quantity with respect to x and set it equal to zero in order to find the quantum extremal surface A^*.

$$\frac{d}{dx}S_{gen}[x] = \frac{\phi_r}{x^2} + \frac{c}{6x} = 0$$

Hence A^* is located at $A^* = -\frac{\phi_r 6}{c}$. The entanglement wedge for this system is therefore as shown in Figure 20

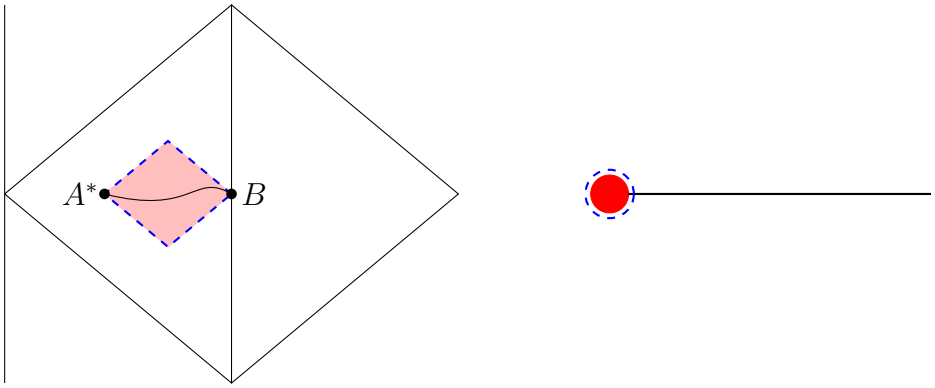


Figure 20: Entanglement wedge for the gravity and dual quantum system (surrounded by blue dots), the entanglement wedge does not cover the entire AdS space-time.

Notice that here only the quantum theory of the dual system was considered. This indicates that once the dual system is coupled to a wire, it can no longer create all of the AdS space. In this case the entanglement wedge covers only a subspace of the Poincare patch. If we intend to focus on the wire CFT of the dual system (the compliment system), we get the following entanglement wedge

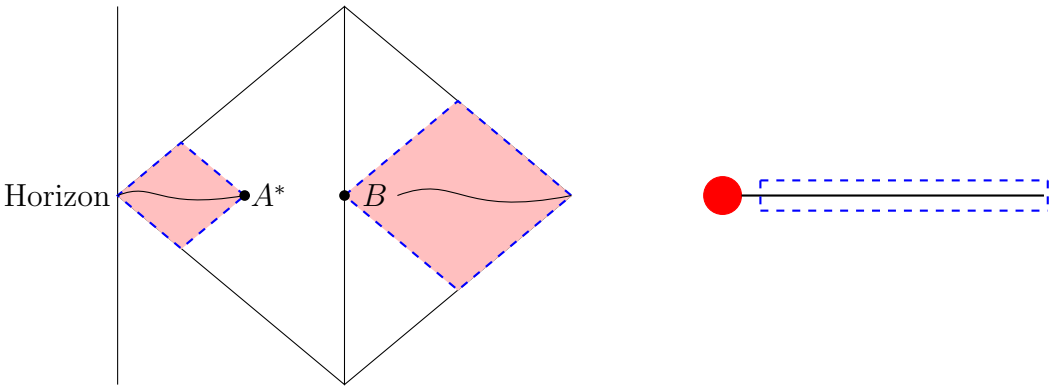


Figure 21: Entanglement wedge for the gravity and dual CFT system, the entanglement wedge contains an isolated region called an Island.

For this case, the entanglement wedge of the wire is disconnected, so we have the usual causal wedge and a disconnected patch near the Poincare horizon. The regions are $[B, \infty) \bigcup [H, A^*]$. The disconnected piece is what is known as the **entanglement island**.

In conclusion, we have seen how to locate the quantum extremal surface in JT gravity with AdS background. This has lead to a region called the island surface. One resolution of the information paradox is due to this island surface and subsequently showing that unitarity of the Page curve is achieved. In the next subsection we show how this Page curve is

achieved by applying the generalized entropy to the process of black hole evaporation.

6.3 A unitary Page curve

In order to understand how the Island proposal is applied to the black hole information loss paradox, we need to describe the black hole evaporation process in stages.

1. Early times: this is when the formation of the black hole begins, and no evaporation has taken place. The black hole is in a pure state and there is no Hawking radiation. At this stage there is also no quantum extremal surface present. So the generalized entropy follows

$$S_{\text{gen}} \sim S_{\text{semi-classical}}(\Sigma_\chi)$$

where Σ_χ is a spatial region chosen for a fixed time slice. This spatial region at a specific time slice defines it's causal wedge. As time evolves, different time slices define different causal wedges. This is why Σ_χ grows with time. Therefore at early times after the black hole has formed, the entropy will start from zero and then increase as the Hawking radiation increases outside the horizon. The Hawking radiation then saturates where $S_{\text{gen}} > 0$. Figure 22 depicts this stage of black hole evaporation in terms of the generalized entropy

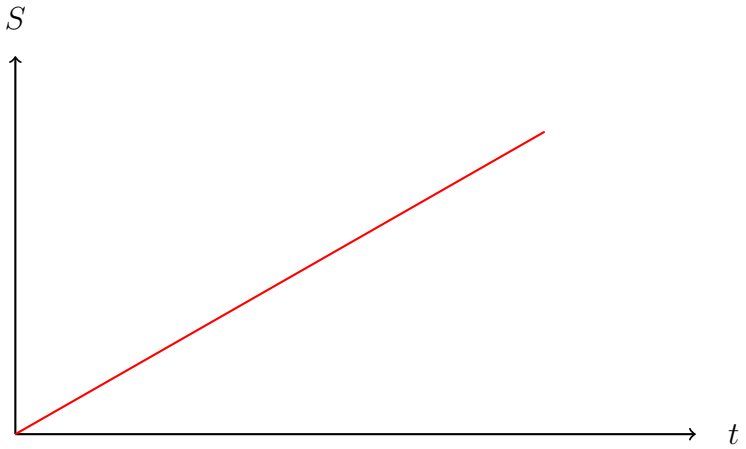


Figure 22: Page curve at early times in the evaporation process of black holes with no quantum extremal surface.

2. Later times: since the black hole has developed and emitted Hawking radiation, a quantum extremal surface will contribute to the generalized entropy. When the black hole radiates half it's mass at some time t', the generalized entropy will be dominated

by the quantum extremal surface. This value is large[9] in comparison with the Hawking radiation outside the black hole, therefore the generalized entropy is

$$S_{\text{gen}} \approx \frac{\text{Horizon Area(t')}}{4G}$$

As the black hole evaporates, the area of it's horizon also evaporates. Therefore the generalized entropy at late times behaves as follows

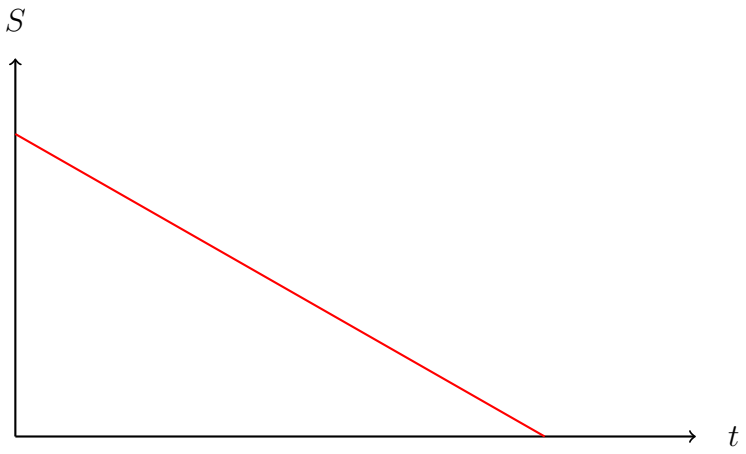


Figure 23: Page curve at late times in the evaporation process of black holes with quantum extremal surface(s).

This indicates that the quantum extremal surface causes the generalized entropy to decrease. The generalized entropy decreases to zero because the black hole completely evaporates.

The full picture of the black hole evaporation process within the Island proposal, tells us that we need to begin with the black hole in a pure state. At the beginning there is no entropy, but as the black hole forms, evaporation starts to take place and the Hawking radiation increases. During this process, the thermodynamic entropy of the black hole decreases as the horizon area decreases. After the formation of the black hole, at late times the quantum extremal surface will start to dominate the generalized entropy. The generalized entropy will then take on the value approximately equal to current thermodynamic entropy of the black hole. The thermodynamic entropy of the black hole will decrease all the way down to zero, producing what looks like a unitary Page curve.

[9]The thermodynamic entropy of the black hole is proportional to it's area, which will be much larger then the entropy of the Hawking radiation emitted at time t'.

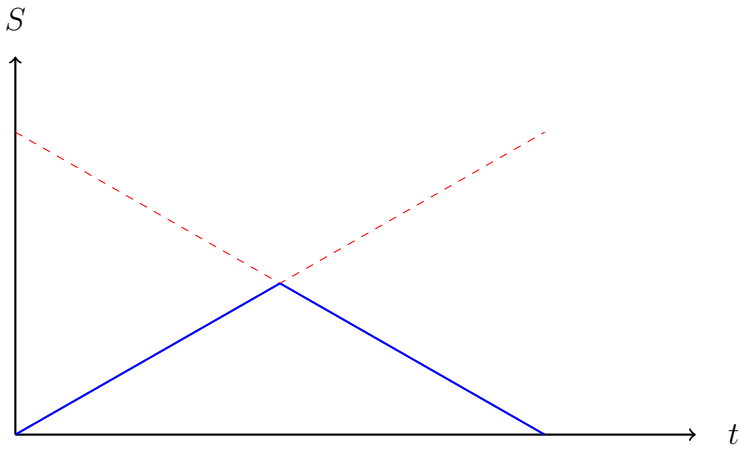


Figure 24: Unitary Page curve achieved by the island proposal.

where the blue curve in Figure 24 represents the unitary Page curve produced by the island proposal. This is the crucial aspect of the island proposal and what makes many researches believe that the black hole information loss paradox is near it's end.

To conclude this section, we discuss briefly the idea of forced evaporation. Large black holes in AdS do not evaporate, the radiation "bounces off" from the boundary and returns to the black hole [15]. The island proposal forces the black hole to radiate by coupling the system to a non-gravitational bath with transparent boundary conditions. Therefore radiation is allowed to leak into the bath. This is problematic because such a system does not conserve the stress energy tensor $T_{\mu\nu}$ [69]. Therefore even though it may seem that the paradox is reaching it's end, there are still questions that the Island proposal needs to address.

7 Conclusion

In this research, we set out to understand the details of the information loss paradox, and discover it's possible solutions. Recall that the crucial points about the paradox is that it's inconsistent with quantum mechanics; initial pure states evolve into mixed states. The information on the initial states gets 'destroyed', because the final states are always mixed. The Page curve calculated by Hawking increases and doesn't follow what we expect for a unitary process. The Page curve we expect should start from zero before the black hole forms, increase almost linearly and then achieve it's maximum at the Page time. Thereafter it should drop down to zero.

When studying gravity in 2 dimensions, we need to use the JT gravity action. This is because gravity in 2 dimensions has constraints on the curvature tensor R and cosmological constant Λ. Gravity is the curvature of spacetime. Curvature is quantified by the curvature tensor R. Therefore restricting a model to 2 dimensions directly places restrictions on the curvature tensor. Defining points on a curved line will need more than 1 spatial dimension. Therefore gravity on a line is problematic. The degrees of freedom are highly restricted since a particle traveling on a line may only move forwards or backwards. Curvature on a plane is not straightforward to imagine. We could use cylindrical coordinates. Still, the dynamics for a particle is limited to traveling on a plane. A plane is flat so we still experience a problem with curvature. We show that JT gravity is important because we locate extremal surfaces in a JT gravity model.

In the process of reviewing Penrose diagrams, we have learned how compatification works. Compactification is possible because we map coordinates that have infinite range, to coordinates that have finite range. This mapping implies a transformation of the metric. The resulting metric has a similar form to the original one. The main idea here is that we do not take an infinite physical spacetime and compactify it into a finite physical spacetime. All we are doing is mapping the coordinates of a infinite physical spacetime to a finite physical spacetime. The finite spacetime is able to capture the same physics of the original spacetime (because of it's causal structure). The physics is invariant because we are only performing a mapping of coordinates. Penrose diagrams are important because we can depict the singularity and event horizon of a black hole. We can also depict the causal wedge, entanglement wedge, extremal surface and island surface.

The literature presented in subsections 2.3 to 2.5 reviews black holes, which have event horizons described by the Schwarzchild metric. The physical process that models a the formation of a black hole is a thin spherical shell. In these subsections we also describe the black hole information loss paradox and how the density matrix is used to compute entropy. This is important because it lays the foundations for advanced computations of entropy. We also explain the Page curve and what is expected from a unitary theory compared to Hawking's

result.

In Section 3 we presented the computation of a saddle point approximation for a general integral. We learned that the semi-classical expansion is based on this approximation. We find the semi-classical expansion through the leading order in the saddle point approximation. We demonstrated that one can use Feynman diagram techniques to compute all the higher order terms in a saddle point approximation. We reviewed the relation between statistical mechanics and quantum mechanics. This relation is demonstrated by computing the classical black hole saddle by which thermodynamics for black holes are recovered. This demonstrated what the semi-classical expansion can be used for. It also gives deep insight as to where semi-classical gravity comes from. This is important because the gravitational entanglement entropy formula has an entropy term that is calculated from semi-classical gravity. The higher order terms are the quantum corrections. Perturbative quantum field theory requires that we expand $\hbar$, we then showed an expansion in $\hbar \rightarrow$ loop expansion. This is a useful result and indicates how effective the saddle point approximation is.

In Section 4 we discovered that entanglement is monogamous. This discovery has implications on the radiation outside a black hole. The radiation far away and near the horizon cannot be entangled. We learned how some proposals try and solve the paradox. The fuzzball proposal claims there there is no horizon. The firewall proposal claims that the horizon is not smooth. Holography of information says that the radiation inside the black hole is also outside the black hole. These are the main competing solutions to the black hole information loss paradox.

The literature presented in Section 5 shows how to solve equations of motion that invoked quantum gravity manually, and used CFT boundary conditions. It is not made explicit why the equations that were being solved, were equations for quantum gravity. However this section is important because we give formal definitions of the entanglement wedge and extremal surface needed to understand aspects about Section 6. This is where many conjectures were employed: the ads/cft conjecture and quantum extremal surfaces. This is a weak area of the literature because conjectures cannot be used solely to solve paradoxes. In some cases solutions which contain conjectures develop even further paradoxes. Ideally we want to gain an understanding of the paradox through equations that is well formulated mathematically and have been experimentally tested.

In Section 6 we presented the Island proposal. This proposal has received a lot of attention due to the unitary Page curve it claims to produce. Therefore we stated the modified equation for gravitational entanglement entropy and presented a way to calculate the location of a quantum extremal surface. Subsequently we plotted the entanglement wedge and Island on a Penrose diagram. This confirmed the importance of what was presented in section 2 of the literature. We have learned that the quantum extremal and island surface creates a

decreasing generalized entropy. This brings the Page curve back down all the way to zero. If the black hole evaporates fast enough, then we might not even see a rise in the Page curve. Why does the generalized entropy take on a value approximately equal to the black hole entropy? This implies that at late times, the quantum extremal surface is the surface of the black hole. This is not clearly discussed the in the proposal. However, it is evident that the island and quantum extremal surface that develops is large enough, and that computing the area of such a surface will give a generalized entropy that could be much larger than the entropy from semi-classical gravity. Even though the black hole entropy dominates the generalized entropy, there should still be some contribution of the entropy from semi-classical gravity. Therefore the Page curve would not go exactly down to zero. Towards the end section 6, we stated some recent criticisms of the island proposal. It seems that the conservation of the stress-energy tensor is not conserved. Due to the arrangement of this proposal, it also needs to address the fact that gravitons become massive for models higher than 2 dimensions.

The black hole information loss paradox is one of the most studied cases because of it's implications. If any mathematically consistent proposal solves this paradox, it will have to be formulated with 'new' physics. Such new physics will indeed be a description of quantum gravity, the unification of general relativity and quantum theory. Understanding gravitational forces at the quantum scale is the final big breakthrough in theoretical and experimental physics. It is therefore imperative that this research continues. Many of the solutions might not solve the paradox at the moment, but we have learned a great deal as to what works and what needs improvement. Any working solution will have to address the problems stated in this research. This research serves as a central location for new researchers trying to understand this paradox as well as a checklist for possible solutions to the black hole information loss paradox.

References

[1] R. Penrose, "Gravitational Collapse and Space-time Singularities," Phys. Rev. Lett.**14**, 57 (1965).

[2] S. Hawking, "Particle Creation by Black Holes," Commun. Math. Phys. **43**, 199-220 (1975).

[3] S. Hawking, "Breakdown of Predictability in Gravitational Collapse," Phys. Rev. D **14**, 2460-2473 (1976).

[4] D. N. Page, "Information in black hole radiation," Phys. Rev. Lett. **71**, 3743-3746 (1993) [arXiv:hep-th/9306083 [hep-th]].

[5] S. Hawking, "The Information Paradox for Black Holes," (2015) [arXiv:1509.01147 [hep-th]].

[6] A. Strominger, "On BMS invariance of gravitational scattering," (2014) [arXiv:1312.2229 [hep-th]].

[7] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, "Replica Wormholes and the Entropy of Hawking Radiation," JHEP **05**, 013 (2020) [arXiv:1911.12333 [hep-th]].

[8] G. 't Hooft, "Dimensional reduction in quantum gravity," Conf. Proc. C **930308**, 284-296 (1993) [arXiv:gr-qc/9310026 [gr-qc]].

[9] L. Susskind, "The World as a hologram," J. Math. Phys. **36**, 6377-6396 (1995) [arXiv:hep-th/9409089 [hep-th]].

[10] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Int. J. Theor. Phys. **38**, 1113-1133 (1999) [arXiv:hep-th/9711200 [hep-th]].

[11] E. Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253-291 (1998) [arXiv:hep-th/9802150 [hep-th]].

[12] S. Gubser, I. R. Klebanov and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," Phys. Lett. B **428**, 105-114 (1998) [arXiv:hep-th/9802109 [hep-th]].

[13] J. D. Bekenstein, "Black holes and entropy," Phys. Rev. D **7**, 2333-2346 (1973).

[14] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," Phys. Rev. Lett. **96**, 181602 (2006) [arXiv:hep-th/0603001 [hep-th]].

[15] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, "The Page curve of Hawking radiation from semiclassical geometry," JHEP **03**, 149 (2020) doi:10.1007/JHEP03(2020)149 [arXiv:1908.10996 [hep-th]].

[16] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," (2013) [arXiv:1306.0533v2 [hep-th]].

[17] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, "Black Holes: Complementarity or Firewalls?," JHEP **02**, 062 (2013) doi:10.1007/JHEP02(2013)062 [arXiv:1207.3123 [hep-th]].

[18] N. Engelhardt, A.C. Wall, "Quantum extremal surfaces: Holographic entanglement entropy beyond the classical regime," (2015) [arXiv:1408.3203 [hep-th]].

[19] G. Penington, "Entanglement Wedge Reconstruction and the Information Paradox," [arXiv:1905.08255 [hep-th]].

[20] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, "The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole," JHEP **12**, 063 (2019) [arXiv:1905.08762 [hep-th]].

[21] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, "Replica wormholes and the black hole interior," [arXiv:1911.11977 [hep-th]].

[22] R. Jackiw, "Lower Dimensional Gravity," Nucl. Phys. B **252**, 343-356 (1985).

[23] C. Teitelboim, "Gravitation and Hamiltonian Structure in Two Space-Time Dimensions," Phys. Lett. B **126**, 41-45 (1983).

[24] H. Nastase, "Introduction to the AdS/CFT Correspondence." ISBN: 978-1-107-08585-5. Cambridge University Press. United Kingdom, United Kingdom by TJ International Ltd. Padstow Cornwall, 2015 (2015).

[25] L. Susskind, J. Lindesay "AN INTRODUCTION TO BLACK HOLES, INFORMATION, AND THE STRING THEORY REVOLUTION." ISBN: 981-256-083-1. World Scientific Publishing Co. Pte. Ltd., 2005 (2005).

[26] K. Schwarzschild, "On the gravitational field of a mass point according to Einstein's theory," Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1916**, 189-196 (1916) [arXiv:physics/9905030 [physics]].

[27] T.A. Moore, "A general relativity workbook." (pp. 352-365), 2013. Mill Valley: University Science Books.

[28] S. Hawking, G. Ellis, "The large scale structure of space-time" , (1973) Cambridge University Press.

[29] M. Born, " The statistical interpretation of quantum mechanics," www.nobelprize.org (1954)

[30] J.J. Sakurai and J. Napolitano, "Modern quantum mechanics." Harlow: Pearson (2014).

[31] S. Raju, "Lessons from the Information Paradox," (2021) [arXiv:2012.05770v2 [hep-th]].

[32] K. Papadodimas, S. Raju and P. Shrivastava, "A simple quantum test for smooth horizons," (2019) [arXiv:1910.02992V1 [hep-th]].

[33] J. Von Neumann, "Wahrscheinlichkeitstheoretischer aufbau der quantenmechanik," Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse (1927): 245-272.

[34] B. Ydri, "Lectures on General Relativity, Cosmology and Quantum Black Holes." ISBN: 978-0-7503-1478-7. IOP ebooks. Bristol, UK: IOP Publishing, 2017 (2017).

[35] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, "The entropy of Hawking radiation," (2020) [arXiv:2006.06872 [hep-th]].

[36] S.W. Hawking, "General Relativity, An Einstein centenary survey." (pp. 746-750), 1979. Cambridge University Press.

[37] R. de Mello Koch and S. Ramgoolam, "Strings from Feynman Graph Counting:without large N." Phys.Rev.D **85**,026007 (2012) [arxiv:1110.4858 [hep-th]].

[38] T. Hartman, "Lectures on Quantum Gravity and Black Holes." (pp. 68-69), 2015. Cornell.

[39] S.M. Carroll, "An Introduction to Spacetime and Geometry." (pp. 429-452), 2004. Addison Wesley.

[40] F.A.M. Frescura, "Statistical Mechanics Notes: Honours." (pp. 79-82), 2009. School of Physics, University of Witwatersrand.

[41] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, "Proposed experiment to test local hidden variable theories," Phys. Rev. Lett. **23**, 880-884 (1969) doi:10.1103/PhysRevLett.23.880.

[42] B. S. Cirelson, "QUANTUM GENERALIZATIONS OF BELL'S INEQUALITY," Lett. Math. Phys. **4**, 93-100 (1980) doi:10.1007/BF00417500.

[43] B. Toner and F. Verstraete, "Monogamy of bell correlations and Tsirelson's bound," (2006) [quant-ph/0611001v1] .

[44] B. Toner, "Monogamy of non-local quantum correlations," Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 465 (2009) 59-69, [quant-ph/0601172].

[45] M. Cvetic and D. Youm, "General Rotating Five Dimensional Black Holes of Toroidally Compactified Heterotic String" Nucl. Phys. B 476, 118 (1996) [arXiv:9603100 [hep-th]].

[46] V. Balasubramanian, J. de Boer, E. Keski-Vakkuri and S. F. Ross, "Supersymmetric Conical Defects: Towards a string theoretic description of black hole formation" Phys. Rev. D 64, 064011 (2001) [arxiv: 0011217 [hep-th]].

[47] J. M. Maldacena and L. Maoz, "De-singularization by rotation" JHEP 0212, 055 (2002) [arXiv: 0012025 [hep-th]].

[48] O. Lunin and S. D. Mathur, "Metric of the multiply wound rotating string" Nucl. Phys. B 610, 49 (2001), [arXiv:0105136 [hep-th]].

[49] O. Lunin, J. Maldacena and L Maoz, "Gravity solutions for the D1-D5 system with angular momentum" (2002) [arXiv:0212210 [hep-th]].

[50] O. Lunin and S. D. Mathur, "AdS/CFT duality and the black hole information paradox" Nucl. Phys. B 623, 342 (2002) [arXiv:0109154 [hep-th]].

[51] I. Kanitscheider, K. Skenderis and M. Taylor, "Fuzzballs with internal excitations" (2007) [arXiv:0704.0690 [hep-th]].

[52] O. Lunin and S. D. Mathur, "Statistical interpretation of Bekenstein entropy for systems with a stretched horizon" Phys. Rev. Lett. 88, 211303 (2002) [arXiv:0202072 [hep-th]].

[53] V. Cardoso, O. J. C. Dias, J. L. Hovdebo and R. C. Myers, "Instability of non-supersymmetric smooth geometries" Phys. Rev. D 73, 064031 (2006) [arXiv:0512277 [hep-th]].

[54] B. D. Chowdhury and S. D. Mathur, "Radiation from the non-extremal fuzzball" (2007) [arXiv:0711.4817 [hep-th]].

[55] S. D. Mathur, "The Information paradox: A Pedagogical introduction," (2009) Class. Quant.Grav. 26 224001, [arxiv: 0909.1038 [hep-th]].

[56] B. S. DeWitt, "The quantization of geometry, in Gravitation: an introduction to current research" (L. Witten, ed.). John Wiley and Sons, (1963).

[57] K. Kuchar, "The problem of time in canonical quantization, in Conceptual Problems of Quantum Gravity" (A. Ashtekar and J. Stachel, eds.). Birkhauser, (1991).

[58] S. B. Giddings, D. Marolf and J. B. Hartle, "Observables in effective gravity", Phys. Rev. D 74 (2006) 064018, [arXiv:0512200 [hep-th]].

[59] R. Bousso, "Complementarity Is Not Enough", Phys. Rev. D 87 (2013) 124023, [arXiv:1207.5192 [hep-th]].

[60] S. Banerjee, J. W. Bryan, K. Papadodimas and S. Raju, "A toy model of black hole complementarity", JHEP 05 (2016) 004, [arXiv:1603.02812 [hep-th]].

[61] A. Hamilton, D. Kabat, G. Lifschytz and D.A. Lowe, "Holographic representation of local bulk operators," (2006) [arXiv:0606141v1 [hep-th]].

[62] X. Dong, D. Harlow, A.C. Wall "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality," (2016) [arXiv:1601.05416v3 [hep-th]].

[63] D. Harlow, "TASI Lectures on the Emergence of Bulk Physics in AdS/CFT," (2018) [arXiv:1802.01040v3 [hep-th]].

[64] N. Kajuri, "Lectures on Bulk Reconstruction," (2020) [arXiv:2003.00587v8 [hep-th]].

[65] J.B. Seaborn, "Hypergeometric Functions and Their Applications." ISBN 978-1-4419-3097-2, Texts in Applied Mathematics. Springer Science Business Media, 1991.

[66] V. Balasubramanian, P. Kraus and A. Lawrence, "Bulk vs. Boundary Dynamics in Anti-de Sitter Spacetime," (1998) [arXiv:9805171v4 [hep-th]].

[67] I. Heemskerk, D. Marolf, J. Polchinski and J. Sully, "Bulk and Transhorizon Measurements in AdS/CFT," (2012) [arXiv:1201.3664v2 [hep-th]].

[68] A. Almheiri, R. Mahajan, J. Maldacena, "Islands outside the horizon," (2019) [arXiv:1910.11077 [hep-th]].

[69] H. Geng, A. Karch, C.P. Pardavila, S. Raju, L. Randall, M. Riojas, S. Shashi "Inconsistency of Islands in Theories with Long-Range Gravity," (2021) [arXiv:2101.03390 [hep-th]].