Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF PHYSICS

Recovering the Page Curve from the AdS/CFT Correspondence

Author:
Johannes Røsok Eskilt

Supervisor:
Toby Wiseman

Submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London

September 2020

Abstract

In this dissertation, we discuss the black hole information paradox. We focus on the recent development were the minimal quantum extremal surface prescription can be applied to an evaporating black hole to recover the Page curve of the fine-grained entropy of the Hawking radiation. This suggests that the black hole evaporation process is unitary and conserves information. This procedure will also yield an entanglement wedge of the radiation that reaches into the interior of the black hole after the Page time. Moreover, we will show how one can explain these two results by applying the AdS/CFT conjecture twice to a 2d JT gravity system with a black hole, where the entanglement wedge of the radiation in the dual AdS_3 naturally accesses the interior of the black hole.

Acknowledgments

Leaving applied physics and starting this master's degree in theoretical physics might be one of the best choices I've taken in my life.

My year in London has been amazing, and that is mostly due to my friends in the QFFF programme and in Gradpad Wood Lane.

I have learned a lot this year, and I want to thank Toby Wiseman for accepting me as one of his many students and being a great supervisor.

Contents

Chapter 1

Introduction

The two pillars of fundamental physics are general relativity and quantum mechanics, but both theories, as they stand now, don't play well together. Thus, the holy grail of modern physics is to unify the two into a theory of quantum gravity. Nevertheless, the two theories tell us that our universe should be deterministic; if you know the current state of the universe, you should in principle be able to not only predict the entire future of the universe but also its past. We, therefore, run into a problem when we consider universes that contain a black hole.

Hawking shocked the world in 1973 when he showed that some things can escape a black hole, namely black body radiation. The black hole thus radiates as if it was a black piece of coal with a very low temperature. The fine-grained entropy of black body radiation is purely mixed; it doesn't carry any information. This is a big problem. Hawking's calculation showed that if you throw your diary into a black hole, the black hole will evaporate into radiation which contains no information. Thus a person born after your diary has been thermally absorbed by the black hole will never know that your diary ever existed. Thus, the law of determinism has been violated, contradicting both general relativity and quantum mechanics.

For decades, Hawking stood by his calculation, arguing that black holes, in fact, destroy information, and the problem has been a hot topic ever since. Today, it is known as the black hole information paradox. And it stands as one of the biggest mysteries in physics.

Entropy is closely linked to information; it tells us how much information we lack. If we know exactly what the quantum mechanical state of an object is, the so-called von Neumann entropy of that object is zero. This entropy is bounded above by the closely related thermodynamic entropy.

Another way of saying that black holes destroy information is that black hole evaporation breaks the unitarity of quantum mechanics. This property is important since it tells us that the time evolution of the entire state of the universe must preserve all information. Thus, if black holes destroy information, unitarity must be violated, breaking one of the axioms of quantum mechanics. Don Page showed in 1993 coarsely how the von Neumann entropy of the Hawking radiation of an initial pure black hole should behave, given that unitarity can not be violated. This entropy describes the entanglement between the radiation and the black hole. It starts at zero, rising linearly as more and more outgoing Hawking modes are created. But as the Hawking particles become more entangled with the black hole, there are no more degrees of freedom left in the black hole for the radiation to be entangled with. This happens at the so-called Page time. Thus, the new Hawking radiation, created after the Page time, is entangled with the early-time radiation, effectively lowering the von Neumann entropy. This is because two fully entangled particles have zero entropy when considered together. This goes on until the black hole has evaporated and the entropy is back to zero, as expected from unitarity. Thus, a unitary theory should give a radiation entropy that resembles this entropy curve, also called the Page curve, and such a theory came with the onset of holography.

In late 1997, the Argentinian physicist Juan Maldacena saw that there was a hidden relationship between a certain string theory in $\mathrm{AdS}_5 \times S^5$ and a supersymmetric Yang-Mills model in 4 dimensions. He conjectured that the two were exactly equivalent descriptions, two sides of the same coin. The conjecture was later generalized, and today it does not need string theory nor supersymmetry to be a consistent theory. Today, the so-called AdS/CFT correspondence conjectures that any anti-de Sitter (AdS) asymptotic spacetimes, no matter how complicated the dynamics of that universe is, is equivalent to a conformal field theory (CFT) living on the conformal boundary of the spacetime. Conformal field theories are quantum field theories were scale doesn't matter, and they are in most cases unitary.

The advent of the AdS/CFT conjecture shifted the general opinion on the black hole information paradox towards the belief that; no, black holes do not destroy information. If an evaporating black hole in an AdS asymptotic universe is equivalent to an unitary conformal field theory, then every process in the AdS universe must also adhere to unitarity. Given that the AdS/CFT conjecture is true, the information in an AdS asymptotic universe must be conserved, even when a black hole evaporates. Therefore, somehow the Hawking radiation must contain information about the interior of the black hole, namely everything the black hole has ever sucked in.

Black holes are essential for the AdS/CFT correspondence to hold true. Since the AdS bulk is allowed one extra spatial dimension, it can seemingly hold more degrees of freedom than the CFT dual, but black holes stop this from happening. This is due to the entropy of a black hole being proportional to its area and not its volume. The entropy is given by the Bekenstein formula

$$S = \frac{A}{4G_N}, \tag{1.1}$$

where A is the area of the event horizon. This entropy is thermodynamic, thus being an upper bound on the von Neumann entropy of the quantum fields inside it. When too much matter is stuck together, allowing many degrees of freedom in one region, gravity creates a black hole, which effectively sets an upper bound on the entropy and the degrees of freedom that can be contained by the event horizon. Thus, the theory can be equivalently described by a theory with one dimension less.

With the AdS/CFT came the RT conjecture, relating the entanglement entropy of a boundary region A with the minimal surface γ_A in the bulk which shares the same boundary as A. The formula is simply

$$S_A = \frac{\text{Area}(\gamma_A)}{4G_N}. \tag{1.2}$$

Remarkably, this is the same equation as the Bekenstein entropy of a black hole, except that γ_A is generally not the event horizon. If however the bulk only contains an eternal black hole, and we let A span the whole CFT, the two surfaces coincide. The surface also allows one to know which spatial region is encoded in a piece of the boundary theory. This is called the entanglement wedge. The entanglement wedge of A is simply the bulk spatial region encapsulated by γ_A and the conformal boundary itself A. All bulk operators in this region have well-defined boundary operators on A.

The RT conjecture above can be generalized to include quantum effects, changing the name of the surface to a quantum extremal surface. This new prescription can be applied to evaporating black holes successfully, which was done last year independently by Penington [39] and Almheiri, Engelhardt, Marolf, and Maxfield [6]. They discovered that before the Page time, the minimal extremal surface of the Hawking radiation entropy vanishes, but after this time, there is a phase-transition, creating a minimal extremal surface inside but close to the event horizon. This came as a surprise since this means that then entanglement wedge of the radiation is disconnected, with one part being most of the interior of the black hole. This means that information about most of the interior is encoded in the radiation after the Page time. The fine-grained entropy of the Hawking radiation was also shown to replicate the Page curve.

The astonishing fact of the development the past year is that the researchers were able to derive these formulas using the effective field theory of gravity. Namely, these results do not rely on any unknown details of the theory of quantum gravity, like the UV completion of the theory. The effective gravity theory, where quantum mechanics and gravity play well together, contains enough details to compute the fine-grained entropy of the radiation. Hawking's error was simply using the wrong formula.

There doesn't seem to be any intuitive reason for why the entanglement wedge of the radiation is partly inside the interior of the black hole, but it follows the spirit of the ER=EPR conjecture proposed by Maldacena and Susskind [34]. They linked entanglement of two particles to the wormhole connecting two black holes and conjectured that these two seemingly unrelated things are somewhat the same thing.

This dissertation will focus on a recent paper by Almheiri, Mahajan, Maldacena, and Zhao [7] which tried to show how the radiation reaches into the interior of the black hole after the Page time in a similar fashion to the ER=EPR conjecture. The authors proposed a system where there is a black hole in 2d JT gravity, which is a modified theory of gravity. The matter is a CFT, thus, it is taken to have an AdS_3 dual. This system shows that the entanglement wedge of the radiation naturally 'accesses' the interior of the black hole in the AdS_3 even though it seems mysterious to an inhabitant of the 2d universe. By applying the minimal extremal surface prescription we will also show that the entropy of the radiation also follows the Page curve.

This dissertation will start with preliminaries such as explaining different types of entropy and how the black hole paradox is linked to entanglement entropy. The third chapter goes into the details of the AdS/CFT correspondence and how you can apply it to a black hole in an asymptotic AdS spacetime. The fourth chapter then mentions how one can relate the entropy of the CFT to an extremal surface in the AdS dual, namely the RT surface and its generalizations. We will here also apply it to black holes to show that the entanglement entropy does indeed yield the Page curve. This all builds up to the fifth chapter where the AdS/CFT conjecture is applied in a so-called double-holography system. Here a 2d black hole made out of a CFT_2 matter has a AdS_3 dual which explains the Page curve by having an additional spatial dimension. In the last chapter, we summarize our results.

We will in addition set the physical constants $c = 1$, $\hbar = 1$, and $k_B = 1$, but Newton's constant G_N will be written explicitly. The sign convention for the flat metric is the mostly-plus, $\eta_{\mu\nu} = \text{diag}(-, +, ..., +)$.

Chapter 2

Preliminaries

2.1 Entropy

Our first step will be to learn the different notions of entropy. We start by looking at the density matrix. If we know the quantum state of a system $|\psi\rangle$ is pure at zero temperature, then the density matrix is given as the outer product [35]

$$\rho = |\psi\rangle\langle\psi|. \tag{2.1}$$

One quick way to check if a state is pure is to calculate ρ^2, since $\rho^2 = \rho$ if and only if the state is pure. In time-dependent systems, the density matrix will generally evolve over time and follow the von Neumann equation

$$i\frac{\partial\rho}{\partial t} = \left[\hat{H}, \rho\right], \tag{2.2}$$

where $\hat{H}$ is the Hamiltonian of the system. If we know the density matrix ρ of the quantum system, we can calculate the von Neumann entropy, also called the 'fine-grained' entropy, given as

$$S = -\mathrm{Tr}\left(\rho \ln \rho\right). \tag{2.3}$$

The von Neumann entropy will always obey $S \geq 0$, where the zero is ensured if and only if the quantum state is pure. The density matrix also allows us to find observables of operators

$$\langle A\rangle = \mathrm{Tr}\left(\hat{A}\rho\right). \tag{2.4}$$

It is important to distinguish von Neumann entropy from the type of entropy most undergraduates are familiar with, which is the thermodynamic entropy. It is practically impossible to know the quantum state of a gas in a box. Thus calculating the von Neumann entropy is not practically achievable, but you do have a set of observables of the system. You can acquire the temperature and the volume of the gas, and based on that you can find a 'coarse-grained' entropy. The idea is that you find a set of density matrices that yield the same set of macroscopic observables, and then you pick the density matrix that maximizes the von Neumann entropy. This coarse-grained entropy will then give you the thermodynamic entropy. This also ensures that the fine-grained entropy of the quantum system is always smaller than or equal to the coarse-grained entropy.

The coarse-grained entropy, or thermodynamic entropy, will always be in a mixed state when the temperature is non-zero. Thus, we can no longer write the density matrix as Eq. (2.1). This is because the temperature mixes the different possible eigenstates, and we can no longer assure that the system is in only one of these. The density matrix can now be written as [35]

$$\rho = \frac{e^{-\beta H}}{Z}, \tag{2.5}$$

where β is the inverse temperature, H is the Hamiltonian of the system, and $Z = \text{Tr}\ e^{-\beta H}$ is the partition function. In the limit $T \to 0$, $\beta \to \infty$, the system can only enter the ground state, and with the assumption that the ground state is non-degenerate, we get a pure state again. We also see that in the case $T \neq 0$, then $\rho^2 \neq \rho$ and so the state is mixed.

2.1.1 Entanglement Entropy

We now look at bipartite systems where the Hilbert space can be written as a tensor product of two subsystems A and B, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ [26, 29]. In a quantum field theory, this can be realized by dividing the whole space into two regions A and B. The example we are mostly going to be looking at later is where A is the black hole and B is everything outside the event horizon, mainly the Hawking radiation. In this case, the outside observer only have access to the density matrix in the region B which can be found from the total, pure density matrix, $\rho_{\text{tot}} = |\psi\rangle\langle\psi|$, by partially tracing over states in A to get the reduced density matrix,

$$\rho_B = \text{Tr}_A(\rho_{\text{tot}}). \tag{2.6}$$

Writing this explicitly in the language of bras and kets, this is equivalent to

$$\langle b|\rho_B|b'\rangle = \sum_a \left(\langle a| \otimes \langle b|\right) \rho_{\text{tot}} \left(|a\rangle \otimes |b'\rangle\right) \tag{2.7}$$

We now define the entanglement entropy of B as the von Neumann entropy of ρ_B,

$$S_B = -\text{Tr}_B \rho_B \ln \rho_B. \tag{2.8}$$

This entropy allows us to measure how entangled two subsystems are to each other. We will in this dissertation use the three terms, von Neumann entropy, fine-grained entropy, and entanglement entropy, interchangeably, but historically, entanglement entropy was used only when dividing a pure state into two sub-regions. This is because dividing a thermal state into two pieces will give you both an entanglement contribution and a thermal contribution to the entropy of either one of the subsystems. By using Schmidt decomposition [26] we get some interesting properties for the entanglement entropy.

- If the total system is pure, then the two entanglement entropies are the same, $S_A = S_B$. Remember that a pure state has zero von Neumann entropy, thus, the entanglement entropy is a result of only dividing the state into two parts. On the other hand, if the system has a non-zero temperature, the total system can't be in a pure state, and we generally don't get equal entanglement entropies, $S_A \neq S_B$.

- Entanglement entropy adheres to strong subadditivity. Dividing the system into three subsystems A, B, and C, we get the inequality

$$S_{ABC} + S_B \leq S_{AB} + S_{BC}. \tag{2.9}$$

  This yields a slightly weaker inequality

$$S_{AB} \leq S_A + S_B, \tag{2.10}$$

  which motivates a new definition: Mutual information,

$$I_{AB} = S_A + S_B - S_{AB} \geq 0. \tag{2.11}$$

  Being strictly non-negative, this quantity is a measure of how quantum entangled two systems are. If the two systems are unentangled, then simply $S_{AB} = S_A + S_B$.

As an example, we can look at a system where we can factorize the total density matrix as $\rho = \rho_A \otimes \rho_B$. Such a state is called a product state, and the reduced density matrix of A is, as the notation suggests, simply $\mathrm{Tr}_B(\rho) = \mathrm{Tr}_B\left(\rho_A \otimes \rho_B\right) = \rho_A \mathrm{Tr}_B\left(\rho_B\right) = \rho_A$. In a product state the total entropy is simply the sum of the entropy of the two reduced density matrices, $S(\rho) = S_A + S_B$, meaning there is no entanglement, and the mutual information is zero.

Another example is the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{2.12}$$

which is the spin-singlet (pure) state of two electrons. The density matrix is

$$\rho = \frac{1}{2}\left(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|\right). \tag{2.13}$$

And the reduced density matrix for one of the spins is $\rho_A = \frac{1}{2}\left(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\right)$. This yields $S_A = \ln 2$, thus, the mutual information is $I_{AB} = 2\ln 2$. Note that this reduced density matrix is simply the identity matrix divided by the number of elements. More generally, we can make a new definition: A pure system ρ divided into two subsystems is called maximally entangled if one of the reduced density matrices is the identity matrix divided by the number of elements, $\rho_A = \frac{I}{|A|}$. Thus, the spin-singlet state is maximally entangled. Note that in such a state, after taking the partial trace over B, we no longer have any information about what state A is in, unlike the previous example.

2.2 Euclidean Time Thermodynamics

In a quantum theory, all thermodynamic properties of a canonical ensemble can be found by calculating the partition function

$$Z[\beta] = \text{Tr}\left[e^{-\beta H}\right]. \tag{2.14}$$

There is a resemblance between this and the time evolution operator of quantum mechanics, $U = e^{-iHt}$. This connection is deeper than one might think, and we can cast thermodynamics in the language of regular quantum mechanics techniques such as path integrals.

An expectation value of an operator O in a thermal system is given as $\langle O\rangle = \frac{\text{Tr}\left[Oe^{-\beta H}\right]}{Z}$, and so we can find correlation functions

$$\langle T\phi(t_1, x_1)\phi(t_2, x_2)\rangle = \text{Tr}\left[\phi(t_1, x_1)\phi(t_2, x_2)e^{-\beta H}\right], \tag{2.15}$$

where T is the time-ordering operator. From the Heisenberg equation for field operators, $\phi(t, x) = e^{iHt}\phi(0, x)e^{-iHt}$, it isn't hard to show that the correlation function above is periodic under $t \to t + i\beta$. By Wick rotating time, $t = it_E$, the Euclidean time coordinate t_E becomes a compact coordinate with period β, $t_E \in [0, \beta)$.

As an example, let's use this to find the temperature of a Schwarzschild-generalized black hole. We can do this without doing it the conventional way of first finding the surface gravity κ, and then use the formula $T = \kappa/2\pi$. We start with a generalized Schwarzschild metric and introduce the Euclidean time coordinate,

$$ds^2 = f(r)dt_E^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2, \tag{2.16}$$

where $f(r)$ is a function such that there is a horizon at r_s, namely $f(r_s) = 0$. Then, we look at the metric close to the horizon by introducing the exterior coordinates $R^2 = r - r_s$ and set $R^2/r_s \ll 1$. The metric to first order in R^2 becomes

$$ds^2 = f(r_s + R^2)dt_E^2 + \frac{4R^2}{f(r_s + R^2)}dR^2 + (R^2 + r_s)^2 d\Omega^2 \tag{2.17}$$

$$\approx R^2 f'(r_s)dt_E^2 + \frac{4}{f'(r_s)}dR^2 + r_s^2 d\Omega^2, \tag{2.18}$$

where in the last line we series expanded, $f(r_s + R^2) \approx f(r_s) + R^2 f'(r_s) = R^2 f'(r_s)$. We are interested in how the metric behaves close to the horizon, $R^2 \to 0$, and we note the resemblance between the two first terms above and the metric of a 2d plane in polar coordinates, $ds^2 = dr^2 + r^2 d\theta^2$. The metric can be put in the more suggestive form

$$ds^2 = \frac{4}{f'(r_s)}\left(dR^2 + \frac{f'(r_s)^2}{4}R^2 dt_E^2\right) + r_s^2 d\Omega^2. \tag{2.19}$$

We will now require the Euclidean time to be periodic and lie in the interval $\frac{f'(r_s)t_E}{2} \in [0, 2\pi)$, otherwise, we will end up with a conical singularity. A conical singularity in itself is not a problem, but it would yield a divergent contribution to the gravitational path integral, which we will discuss below. Thus, we will stay clear of such singularities. As mentioned, the period of t_E is also the inverse temperature β, thus, to avoid a conical singularity at the horizon, we get the relation

$$\beta = \frac{4\pi}{f'(r_s)}. \tag{2.20}$$

For a Schwarzschild black hole in 4 dimensions, we simply have $f(r) = 1 - r_s/r$, which gives the temperature

$$T = \frac{1}{8\pi M G_N}, \tag{2.21}$$

where we used that $r_s = 2MG_N$, M being the mass of the black hole. Thus, by this very simple argument, we have shown that black holes exhibit a non-zero temperature.

We can also take advantage of path integrals technology to find the partition function. When we calculate transition amplitudes in quantum mechanics, going from an initial state $x(t_i)$ to a final state $x(t_f)$, we have [30, 35]

$$\langle x(t_f)|e^{-iH(t_f-t_i)}|x(t_i)\rangle = \int_{x(t_i)}^{x(t_f)} \mathcal{D}x \; e^{iS[x]}, \tag{2.22}$$

where S is the action. By Wick rotating and applying the path integral approach to the partition function, we can specialize this to the case $x(t_f) = x(t_i) = x$ and get

$$Z[\beta] = \sum_x \langle x|e^{-\beta H}|x\rangle = \int_{x(t_E+\beta)=x(t_E)} \mathcal{D}x \; e^{-S_E[x]}. \tag{2.23}$$

Here, S_E is the Euclidean action, and x is all periodic paths $x(t_E + \beta) = x(t_E)$. Due to the periodicity of t_E, this Euclidean geometry is equivalent to a cylinder where the Euclidean time is the periodic coordinate. Note that this approach is only valid in time-independent systems since we have replaced time with temperature. In quantum field theories, we are often more interested in fields and the equivalent partition function in Euclidean time becomes [35]

$$Z[\beta] = \int_{\phi(t_E+\beta,x)=\phi(t_E,x)} \mathcal{D}\phi \; e^{-S_E[\phi]}. \tag{2.24}$$

2.3 Hawking Radiation

We just saw that a black hole should have a temperature due to a simple argument stating that the Euclidean time shouldn't have a conical singularity at the horizon. This might not be convincing enough for most people, but Hawking showed more rigorously that black holes do in fact evaporate into seemingly thermal radiation as if it had a black body temperature of T that we found above.

We won't go into the details of Hawking's calculation, but we will look at how a massive scalar field is affected by the Schwarzschild metric. We assume that the coupling between the QFT and general relativity is a one-way street: The field is told by the metric how to behave, but the metric is not curved by the matter field. We, therefore, neglect any backreaction, which is equivalent to taking the limit where the mass of the black hole goes to infinity, $M \to \infty$, and Newton's constant goes to zero, $G_N \to 0$ so that the Schwarzschild radius $r_s = 2G_N M$ stays constant. Following Ref. [26], we can solve the Klein-Gordon equation in the Schwarzschild metric. Using the identity $\nabla^2\phi = \partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi)/\sqrt{-g}$ where g is the determinant of the metric, we get

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi) = m^2\phi. \tag{2.25}$$

The black hole has a spherical symmetry so we can look for solutions of the form

$$\phi_{\omega lm}(t, r, \Omega) = Y_{lm}(\Omega)e^{-i\omega t}\frac{\psi_{\omega l}(r)}{r}, \tag{2.26}$$

where (l, m) are a set of eigenvalues that describe the angular momentum of the particle, ω is an eigenvalue that gives the energy, and Ω is the set of angular coordinates. It is convenient to switch over to tortoise coordinates,

$$r^* = r + r_s \ln\left(\frac{r}{r_s} - 1\right), \tag{2.27}$$

since now we can write the solution as a particle moving in one spatial dimension in non-relativistic quantum mechanics,

$$\frac{\partial^2\psi_{\omega l}}{\partial(r^*)^2} + V(r)\psi_{\omega l} = \omega^2\psi_{\omega l}, \tag{2.28}$$

where the effective potential is

$$V(r) = \frac{r - r_s}{r^3}\left(\frac{r_s}{r} + l(l+1) + m^2r^2\right). \tag{2.29}$$

Here, r is implicitly a function of r^*. The Hawking radiation originates from a region very close to the event horizon at the Planck scale. And the potential gives a peak well outside the event horizon at $r_s < r = O(r_s)$ which the Hawking radiation has to climb out of, or tunnel through. A substantial amount of scattering is therefore happening due to this potential. Not only for outgoing Hawking modes but also incoming particles. As one also expects, the peak becomes harder to overcome for massive particles or particles with non-zero angular momentum $l > 0$. A massive particle needs to overcome its rest-mass to escape the black hole, $V(r \to \infty) = m^2$. Since we have shown that a black hole has a temperature of order $T \propto 1/r_s$, this is also the energy scale we should look at $\omega \approx 1/r_s$. A black hole with the same mass as our sun has a temperature of $T \approx 10^{-7}$ K which gives an energy 15 orders of magnitude less than the rest mass of an electron. Hence, a particle with an energy of order T will not have enough energy to attain its rest-mass energy, and thus be able to escape the effective potential of the black hole. It is, therefore, safe to only focus on massless particles.

For massless particles with no angular momentum, the peak of the effective potential sits at $r = 4r_s/3$. Even for these particles, there is a large amount of back-scattering which causes the black hole to favor high energy Hawking modes thus making the black hole not a perfect example of black body radiation. This is called a grey-body factor since the black hole absorbs most of its radiation, and it often complicates calculations.

2.4 Black Hole Information Problem

As we just saw, a black hole has a temperature, so one might hope to find other thermodynamic quantities. A very interesting one is the thermodynamic entropy which is proportional to the area of its event horizon A

$$S = \frac{A}{4G_N}, \tag{2.30}$$

A very simple derivation of this entropy can be made by using the thermodynamic identity $dE = TdS$ and the area,

$$A = 4\pi r_s^2 = 4\pi(2MG_N)^2. \tag{2.31}$$

The temperature T is found above and the energy is the same as the mass $E = M$, then a small change in entropy becomes

$$dS = 8\pi G_N M dM = 4\pi G_N dM^2 = \frac{dA}{4G_N}, \tag{2.32}$$

which gives the entropy in (2.30). A much harder derivation, but also a more convincing one, is to use gravitational path integral using Euclidean time as mentioned earlier. The latter derivation is a more rigors way of showing that a black hole does carry a thermodynamic entropy that is proportional to its area.

Since a black hole gives away radiation, its mass must decrease which also shrinks the event horizon. This happens extremely slowly, and a black hole the size of the sun will take around $t = 10^{64}$ years to completely evaporate, and all black holes formed right after the big bang with a mass less than $m < 10^{12}$ kg will be gone by now [8].

There is however a problem one faces with the black hole entropy. The black hole entropy is a gigantic number compared to the entropy of other astrophysical bodies. All these microstates are hidden behind the event horizon, but after the black hole evaporates into thermal radiation, all the vast information about the microstates is seemingly destroyed. This was Hawking's initial interpretation of black hole evaporation: Black holes destroy information. This is a huge problem since it violates the inherent determinism of both quantum mechanics and general relativity. In both theories, if you know the state of the universe exactly, then you should be able to predict not only the future but also the past. If you are born in a universe without any black holes, there is nothing you can do to figure out if a black hole ever existed. This problem is called the black hole information paradox, and it is arguably one of the biggest unsolved problems in physics. Since it combines the extremes of both quantum mechanics and general relativity one expects that the solution will give us profound insight into what properties a theory of quantum gravity would have.

Let's explain more in detail how the black hole paradox violates both quantum mechanics and general relativity. Quantum mechanics is notorious for being random. By measuring a system, the universe has to collapse the wavefunction, which before the measurement was a linear combination of states, to a seemingly randomly chosen single state. But this measurement hinges on the fact that we couple the system to an external environment, namely our measurement apparatus. The entire system, on the other hand, is still deterministic, and its time evolution is governed by a unitary transformation acting on the Hilbert space of states. Such transformations conserve information, and thus by accepting that black holes destroy information, one also has to throw away the unitary condition of time evolution. This is a hard pill to swallow since it is one of the axioms of quantum mechanics, and thus physicists have looked for other solutions.

General relativity is a classical theory, and its determinism should coincide more closely with our intuition. If you throw a tennis ball here on earth, and you know precisely its position and velocity, you can calculate its entire path before hitting the ground. The same goes for general relativity. If you know the position and velocity of all mass and energy at one point in time, then you should be able to not only determine the future but also the past of your universe. More accurately, if we know the state of the universe on a Cauchy surface, we can determine the entire future and past of the spacetime. The problem arises when we look at an evaporating black hole [18]. In Fig. 2.1 we have drawn then Penrose diagram of an evaporating black hole with two spacetime slices Σ_1 and Σ_2. They both originate at $r = 0$ and end at spatial infinity i^0, but Σ_1 starts at $r = 0$ before the black hole and its singularity have formed, while Σ_2 originates at $r = 0$ after the black hole has evaporated. Σ_1 is a proper Cauchy slice, if you know the state of the universe on this slice, you can determine the entire fate and past of the universe, including the black hole. In other words, the domain of dependence covers the entire spacetime. Σ_2, on the other hand is not a Cauchy slice. Its future domain of dependence is correct, but the inside of the black hole is not in its past domain of dependence. Thus, information about the spacetime has been lost going from Σ_1 to Σ_2.

We are going to look at an infalling shell of photons to better understand the information paradox, but firstly, we will learn a lesson about how a black hole is formed from such a shell and differences between horizons of black holes as discussed in Ref. [52]. The definition of a black hole is a region of space where you can not send a signal to null infinity, and the event horizon is the border of this region. Birkhoff's theorem gives us two solutions of the spacetime; the region outside the shell has the Schwarzschild solution, while the interior is the Minkowski vacuum. If you are positioned at the center of the collapsing shell, there are no experiments you can perform to foresee your unfortunate fate. If the radius of the shell at one point in time is slightly larger than its corresponding Schwarzschild radius, and you send out a signal with a flashlight, those photons will never reach null infinity since they

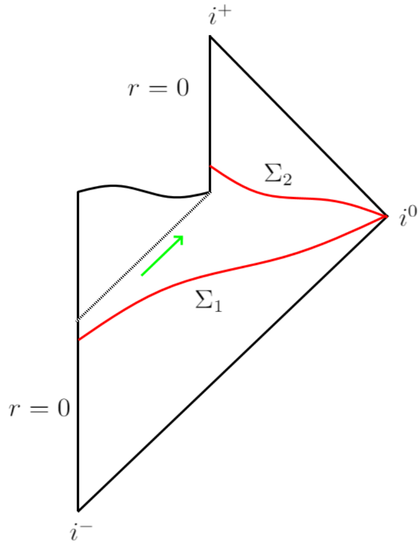

Figure 2.1: Penrose diagram of an evaporating black hole created from gravitational collapse. The green arrow shows the Hawking radiation which slowly evaporates the black hole. Knowing the state of the universe on the two Cauchy slices Σ_1 Σ_2 will give you two very different results. Σ_2 doesn't know that a black hole ever existed since the Hawking radiation is assumed to be purely thermal.

will be within the Schwarzschild radius by the time the shell has shrunk to be within its Schwarzschild radius. This is the grey area in Fig. 2.2. Therefore, you were already inside the black hole when you turned on the flashlight, and there was nothing you could do to know! This highlights the fact that event horizons (and black hole regions) are not possible to measure locally. Mathematically, you need to know the entire future of the universe to determine if you are right now inside a black hole or not. On the other hand, you have the apparent horizon which is defined as the surface where a null geodesic stays stationary. This horizon and the event horizon coincides in eternal, static black holes like the Schwarzschild solution, but in this example, the (outgoing) apparent horizon is formed once the shell reaches the Schwarzschild radius, and then it forever stays at this radius (assuming no evaporation). This horizon can also be determined by doing local experiments without knowing the future of the spacetime solution. Thus, these two types of horizons are fundamentally different.

2.5 The Page Curve

Assume the infalling shell of photons is in a pure state. Then, once the event horizon is formed, the fine-grained entropy of the black hole will initially be zero. But on the other hand, its coarse-grained entropy, proportional to its event horizon area, will increase until it reaches the Schwarzschild radius. Thus, it has a temperature, so one might ask where this thermal component comes from when starting with a pure state. The answer is that the vacuum in a quantum field theory is highly entangled

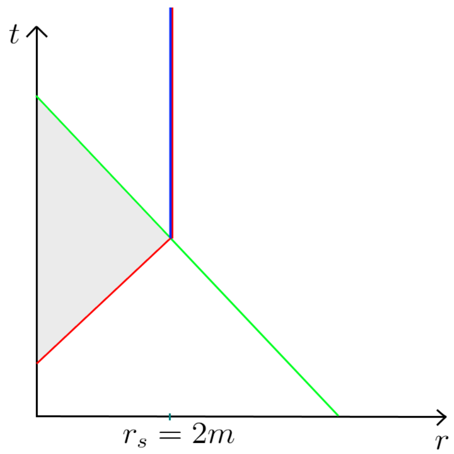

Figure 2.2: The infalling shell of photons (green) creates an event horizon (red) in an empty Minkowski space before an observer at the center knows it is inside it (grey area). An outgoing apparent horizon (blue) is created at the Schwarzschild radius once the shell has reached that radius. The figure is adapted from Ref. [52].

[8]. When a black hole is formed, we are creating a region of space that can never send a signal out to null infinity; this is the definition of a black hole. And thus, by separating the space into two regions, the interior and the exterior of the black hole, we naturally end up with an event horizon with a temperature, as seen from an observer in the exterior.

When one first learns about Hawking radiation, one is often told the story of the spontaneous creation and annihilation of matter and antimatter in the vacuum of a quantum field theory. These two particles are maximally entangled and normally destroyed after a very short time, but if this happens close to the event horizon, one of the particles can fall into the black hole while its partner, being outside the event horizon, can escape towards infinity. Since the outgoing particle has a positive energy to an outside observer, energy conservation tells us that the black hole must lose energy, and hence mass, in this process to the outside observer. Even though this explanation is popular, one needs to be wary of taking it literally. One of the problems is that the wavelength of the escaping particles is of the order of the size of the black hole [26], and scalar particles don't have such partners. In addition, there is a high effective potential barrier close to but outside of the event horizon which makes it easier for massless particles with zero angular momentum, both zero spin, and global angular momentum, to escape. Thus, as mentioned, the Hawking radiation consists of mostly photons. Even though we shouldn't take the above explanation literally, we will still think of outgoing particles as having an ingoing partner. This makes it clear that there is a finite part of the wave function propagating like a null geodesic (assuming its Hawking partner is massless) inside the black hole which is entangled to its outgoing partner.

When the black hole starts emitting Hawking radiation, the fine-grained entropy of the radiation will start increasing over time as more and more radiation is created. Most of the outgoing Hawking modes will fail to tunnel or climb their way out of the black hole potential and give rise to so-called grey body factors. If we want to quantitatively understand black hole evaporation, we can't discard them, but by neglect them, Hawking's calculation showed that the black hole radiation is purely thermal, and so the fine-grained entropy of the radiation seems to be strictly increasing until the black hole has completely evaporated. We started with a pure entropy, and seemingly we have created a highly mixed state of thermal radiation. Somehow, the unitarity of quantum mechanics has been violated, since one would expect the entire state to stay pure during the process.

If we assume that Hawking made an error somewhere, there are two options [26]. One is the remnant theory, which says that Hawking was correct up until the point where the black hole has evaporated to the size of the Planck length. Assuming unitarity can't be violated, the remnant needs an incredibly large entropy which surmounts the Bekenstein entropy. This theory is not taken seriously by most physicists since it is hard to imagine how such a vast amount of entropy can be captured by such a small object. Thus, we turn to the other option.

Assuming unitarity is preserved, many physicists are waging their careers on the assumption that the Hawking radiation is not purely thermal. Somehow the information about the black hole is encoded in radiation, and thus the entropy curve of the radiation will look very different. Assuming we start with a pure state, and if the evaporation process had respected unitarity, we would expect the fine-grained entropy of the radiation to eventually reach zero again when the black hole has fully evaporated. Secondly, the von Neumann entropy of the black hole and the radiation must be the same as we argued in the last section. If we have a state in a pure state, the von Neumann entropy of the subregion A (black hole) and the complement region B (rest of the universe, in this case, the radiation) must be equal.

The curve of the von Neumann entropy of the Hawking radiation is called the Page curve [37, 38]. And we can outline the rough shape of it without doing any math, see Fig. 2.3. Right after the pure, infalling shell settles, it starts evaporating Hawking radiation which increases its entropy linearly. Under the assumption of black holes evolving unitarily, the radiation is entangled with the interior of the black hole. As the fine-grained entropy of the radiation S_{rad} starts reaching the Bekenstein entropy S_{BH} of the black hole, there are no more available microstates in the black hole to be entangled with anymore. This happens when $S_{\text{rad}} = S_{\text{BH}}$ at the so-called Page time. After this point, the new Hawking radiation cannot be entangled to the black hole anymore, and must instead be entangled with early Hawking radiation sent out before the Page time. This makes the radiation state becoming purer and purer. Thus, the radiation entropy starts decreasing. When the black hole has completely evaporated, the system is entirely made out of a pure Hawking radiation and so the entropy is back to zero.

Whatever a theory of quantum gravity looks like, we should expect our way of understanding the universe to undergo several paradigm shifts, so why is it so hard to accept that unitarity might be violated? We can get clues from other theories, notably string theory. This theory is supposed to be the ultimate theory of everything, and it is inherently unitary. String theory has not passed the test of time yet, but it has yielded promising results. Strominger and Vafa were able to show in 1996 that counting the microscopic states of a particular type of black holes gives the expected

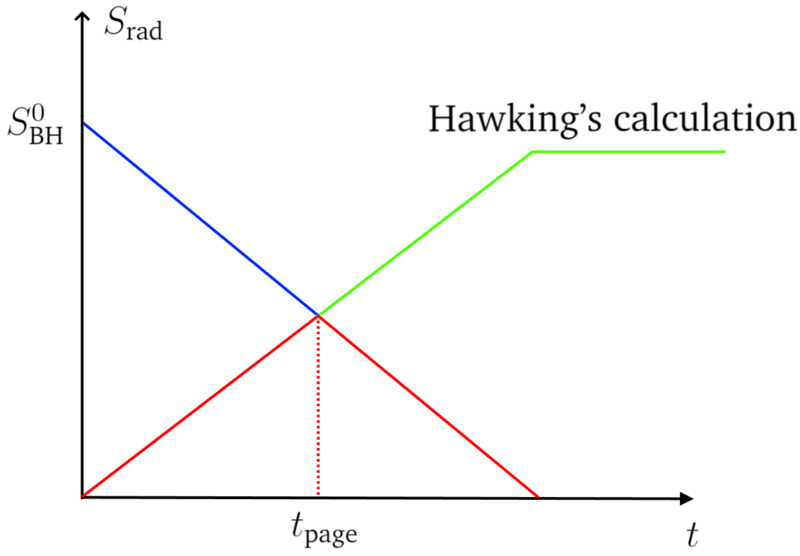

Figure 2.3: Qualitatively how we expect the Page curve (red) to look like. The fine-grained entropy of the radiation starts at zero and increases until it reaches the coarse-grained entropy of the black hole (blue), at the Page time. Then S_{rad} falls to zero along with the Bekenstein entropy, and we end up with a zero fine-grained entropy meaning the radiation ends up in a pure state. The green curve shows Hawking's original calculation where the entropy is expected to strictly increase until the black hole has completely evaporated.

Bekenstein entropy [47]. Subsequent discoveries have solidified the understanding of black holes in the context of string theory, but unfortunately, it doesn't help us understand how the information is stored in the Hawking radiation.

We can also look at less speculative theories, like the main ingredient of this dissertation, namely the AdS/CFT conjecture. This conjecture tells us that the unitarity of black holes in anti-de Sitter space is trivially satisfied since it has an equivalent description of living in a unitary conformal field theory on the boundary of the AdS spacetime. Since any quantum field theory adheres to unitarity, information must be conserved.

Chapter 3

AdS/CFT Correspondence

The idea of holography was created before Maldacena conjectured the AdS/CFT correspondence. The holographic principle, first proposed by Gerard 't Hooft [48] and later promoted by Leonard Susskind, was a set of proposed properties of quantum gravity where one would require one spatial dimension less to describe the same amount of information. The idea got its inspiration from the Bekenstein entropy of a black hole. The old thinking was that entropy in a spatial region was bounded by its volume, but the holographic principle suggests that it is instead bounded by the Bekenstein entropy with the area that encapsulates the region, an idea that is now known as the Bousso bound [13]. So if you try to put a huge amount of thermodynamic entropy in a box, that box will become a black hole before it can break the Bousso bound, effectively putting an upper bound on the allowed entropy in the box. There are however examples in classical general relativity called 'Wheeler's bag of gold' where there is a big region of space within an event horizon that can contain a lot more entropy than its event horizon area (divided my $4G_N$). Thus, one has to find a way to invalidate such solutions if holography is to hold true.

So if the holographic principle is true, it seems like gravity is a way for our 4-dimensional universe to put an upper bound on the entropy and the degrees of freedom in a 3-dimensional spatial region so that one still has enough degrees of freedom on a holographic 2d screen to describe the spatial region.

The vague ideas of holography were then in 1997 better realized by Maldacena in his gravity/gauge conjecture, later known better as the AdS/CFT conjecture. This conjecture states that an AdS asymptotic universe described by quantum gravity in d-dimensions can be equivalently described as a conformal field theory in $(d-1)$-dimensions living on the conformal boundary of the bulk universe.

Before we jump into his conjecture, we need to go through the two main ingredients: Anti-de Sitter (AdS) spacetime and conformal field theory (CFT).

3.1 Anti-de Sitter Spacetime

The anti-de Sitter spacetime is a vacuum solution of Einstein's equation that is maximally symmetric. This means that neither direction in space nor time is favored, and every single position at any given time gives the same metric. These two conditions are called isotropic and homogeneous, respectively. Due to these requirements, the spacetime must have a constant scalar curvature, and AdS spacetimes have negative curvature, while de Sitter spacetimes have positive curvature and Minkowski spacetimes have zero curvature. It can be shown that the Riemann tensor for a maximally symmetric $(d+1)$-dimensional spacetime to be [35]

$$\mathcal{R}_{\mu\nu\alpha\beta} = \frac{\mathcal{R}}{d(d+1)} \left(g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}\right), \tag{3.1}$$

where $\mathcal{R}$ is the Ricci scalar. Solving Einstein's equation with a cosmological constant Λ gives

$$\mathcal{R} = \frac{2(d+1)}{d-1} 8\pi G_N \Lambda, \tag{3.2}$$

valid for $d > 1$. Therefore, an anti-de Sitter spacetime describes a universe with a negative cosmological constant. This is the opposite of what we think of our universe, namely a positive cosmological constant which drives an acceleration of the expansion of our universe. AdS spacetimes, instead, make space shrink and two massive objects will always be pulled together no matter how fast they move apart.

To appreciate the symmetries of $(d+1)$-dimensional anti-de Sitter spacetime, we embed it in $\mathbb{R}^{(2,d)}$, namely a flat spacetime of d spatial dimensions and two timelike dimensions. The AdS spacetime is then the hypersurface

$$-T_0^2 - T_1^2 + \sum_{n=1}^{d} X_n^2 = -R^2. \tag{3.3}$$

Here R is the characteristic length scale of the spacetime. The metric of AdS spacetime is thus the induced metric on the hypersurface. The de Sitter spacetime can be defined similarly with the right-hand side being instead strictly positive. First notice that the symmetry of this spacetime is $SO(2,d)$, allowing us to, for example, rotate the d spatial dimensions, and the two time dimensions separately. We are now going to explore this hypersurface closer by only allowing one spatial dimension. Setting all spatial dimensions except X_1 to zero, after renaming $X_1 = X$ we get

$$T_0^2 + T_1^2 - X^2 = R^2, \tag{3.4}$$

which can be reparameterized to Rindler-like coordinates, $T_0 = R\cosh(\rho)\cos(\tau)$, $T_1 = R\cosh(\rho)\sin(\tau)$ and $X = R\sinh(\rho)$. This gives the metric

$$ds^2 = -dT_0^2 - dT_1^2 + dX^2 = R^2\left(-\cosh(\rho)^2 d\tau^2 + d\rho^2\right). \tag{3.5}$$

Note that the two new coordinates τ and ρ are dimensionless. The representation in Eq. (3.3) allows for closed timelike curves ($X_n = 0$) which generally is something one tries to avoid in relativity since it can make causality break down. Many authors, therefore, define the AdS spacetime as the universal cover of the above hypersurface, namely unwrapping the circle $\mathcal{S}^1$ ($X_n = 0$) and allow τ to be in the interval $[-\infty, \infty]$ without identifying two different τ [35].

At $\rho = 0$ the metric looks locally flat (with a conformal factor R^2), so to get a feel of the metric, we imagine that an observer is centered at $\rho = 0$ throws a tennis ball in the positive x-direction. The metric above admits a timelike Killing vector $k = \partial_\tau$ which has a corresponding conserved quantity $Q = \frac{d\tau}{d\lambda}\cosh(\rho)^2$ where λ is an affine parameter which we will use as the proper time for a massive particle. This equation can be interpreted as a conservation of energy. The tennis ball is massive, and so the geodesic will follow the path

$$1 = R^2\cosh(\rho)^2\left(\frac{d\tau}{d\lambda}\right)^2 - R^2\left(\frac{d\rho}{d\lambda}\right)^2, \tag{3.6}$$

which with the conservation of energy yields

$$\frac{d\rho(\lambda)}{d\lambda} = \pm\frac{1}{R}\sqrt{\frac{R^2Q^2}{\cosh(\rho)^2} - 1}. \tag{3.7}$$

The tennis ball has a positive start velocity, $\rho'(0) > 0$, which determines the energy Q. But we see that the observer positioned at $\rho = 0$ will observe the particle to stop moving at $\frac{R^2Q^2}{\cosh(\rho)^2} = 1$. At this point, the ball will start moving back to the observer. When it eventually reaches the observer it will have the same initial speed except it is moving in the opposite direction, which creates an oscillating motion. This supports the fact that the cosmological constant is negative in anti-de Sitter. It can generally be shown that the cosmological constant is

$$\Lambda = -\frac{(d+1)d}{16\pi G_N R^2}. \tag{3.8}$$

This example also shows an important symmetry of the anti-de Sitter space, not only is it symmetric in the spatial dimension, but also in the time dimension.

As we saw, no matter how far you throw the ball, it will eventually come back. Interestingly, this is not the case for massless objects. By shooting off a massless photon from $\rho = 0$, we see that the object reaches infinity at a finite time for the observer

$$\tau = \int_0^\infty \frac{d\tau}{d\rho}d\rho = \int_0^\infty \frac{d\rho}{\cosh(\rho)} = \pi/2. \tag{3.9}$$

What happens at infinity? That is a good question, and mathematically, it depends on what boundary condition we set here. An absorbing boundary condition would mean energy can leave our system, so instead, we reflect the massless particles back again. In the case of a black hole in an AdS asymptotic spacetime, we see that a big enough evaporating black hole can be in thermal equilibrium, where a certain amount of energy is Hawking radiation while the rest is the mass of the black hole. We will look at this in more detail later.

We can generalize the metric (3.5) to d spatial dimensions as

$$ds^2 = R^2\left(-\cosh(\rho)^2 d\tau^2 + d\rho^2 + \sinh(\rho)^2 d\Omega_{d-1}^2\right), \tag{3.10}$$

where $d\Omega_{d-1}^2$ is the solid angle of the unit $(d-1)$-sphere. These coordinates are called global coordinates, but another set of coordinates, that is especially useful when studying black holes, can be obtained by setting $r = R\sinh(\rho)$ and $t = R\tau$,

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_{d-1}^2, \tag{3.11}$$

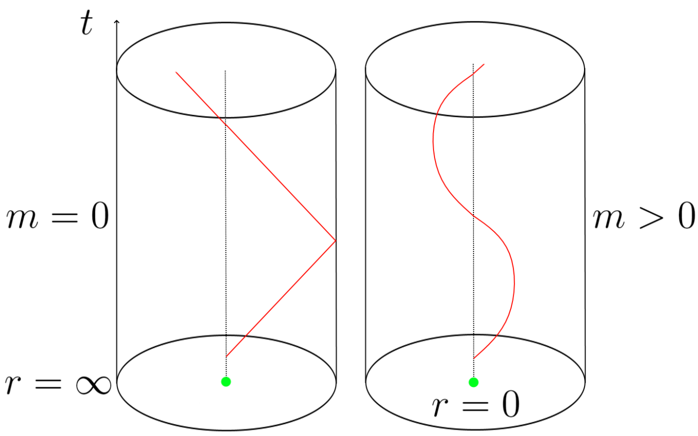

Figure 3.1: The AdS_3 spacetime for an observer centered at $r = 0$. From the observers perspective, the massless particle reaches infinity and bounces back, while a massive particle will always return back to the observer and never reach infinity.

where $f(r) = 1 + r^2/R^2$. This metric is intuitively easier to understand since in the limit $R \to \infty$ we simply have a Minkowski flat spacetime with spherical coordinates. This metric also gives a relationship between the AdS radius R and the time for a massless particle to reach $r \to \infty$. From (3.9) we get

$$t = \frac{\pi}{2} R. \tag{3.12}$$

So for massless particles, the AdS spacetime is similar to a box with a length of order R, while for massive particles, the universe is like an infinitely large box.

One of the coordinate systems that most physicists like to use is the Poincare patch, defined by the not so pleasant transformation

$$\begin{aligned} T_0 &= \frac{z}{2}\left(1 + \frac{1}{z^2}\left(R^2 + (x^i)^2 - t^2\right)\right) \\ T_1 &= \frac{Rt}{z} \\ X_i &= \frac{R}{z} x^i \\ X_d &= \frac{z}{2}\left(1 - \frac{1}{z^2}\left(R^2 - (x^i)^2 + t^2\right)\right), \end{aligned} \tag{3.13}$$

where $i \in \{1, ..., d-1\}$. Doing the algebra yields

$$ds^2 = \frac{R^2}{z^2}\left(-dt^2 + \left(dx^i\right)^2 + dz^2\right). \tag{3.14}$$

We now have the metric in a conformally invariant form. The transformation $(t, x^i, z) \to \lambda(t, x^i, z)$ leaves the metric invariant. It is important to note that the Poincare patch doesn't cover the entire manifold, unlike the global coordinates. This metric has a conformal boundary at infinity, $z \to 0$, which means we can remove the prefactor of the equation above, or more formally do a conformal transformation

$$g_{\mu\nu} \to g'_{\mu\nu} = \Omega^2 g_{\mu\nu}. \tag{3.15}$$

Here, (μ, ν) runs in the orthogonal directions, leaving $dz = 0$ in the metric in Eq. (3.14), and $\Omega = z/R$, which shows that the conformal boundary has a flat Minkowski metric. In the global coordinates, it is easy to check that the conformal boundary of the global coordinates has the topology $\mathbb{R} \times S^{d-1}$. This shows why these coordinates are helpful in the AdS/CFT framework; we can regularize the metric and get well-defined quantities on the AdS boundary where the CFT lives, even though it is infinitely far away.

3.2 CFT

A conformal field theory is a quantum field theory where scale doesn't matter. These are coordinate transformations $x^\mu \to x'^\mu(x)$ so that the new metric satisfies [51]

$$g_{\mu\nu}(x) \to \Omega^2(x) g_{\mu\nu}(x). \tag{3.16}$$

In Minkowski space, we can easily classify the group of transformations that satisfy the above requirement. The group of transformations that satisfy this requirement is called the conformal group, and the Poincare group is a subgroup of these transformations. The Poincare group already includes translation, boost, and rotation. But another type of transformation that obviously satisfies the requirement is scaling. We can scale all coordinates by the same factor, $x^\mu \to \lambda x^\mu$. This transformation is called dilation and doesn't change the angles between the coordinates. One not so obvious transformation is the special conformal transformation

$$x^\mu \to \frac{x^\mu + a^\mu x^2}{1 + 2x^\mu a_\mu + a^2 x^2}. \tag{3.17}$$

These different transformations have different generators. The Poincare generators are the usual generators for Lorentz transformations $M_{\mu\nu}$ and translation P_α

$$[M_{\mu\nu}, M_{\alpha\beta}] = i\left(\eta_{\nu\alpha} M_{\mu\beta} - \eta_{\mu\alpha} M_{\nu\beta} - \eta_{\nu\beta} M_{\mu\alpha} + \eta_{\mu\beta} M_{\nu\alpha}\right) \tag{3.18}$$
$$[M_{\mu\nu}, P_\alpha] = -i\left(\eta_{\mu\alpha} P_\nu - \eta_{\nu\alpha} P_\mu\right). \tag{3.19}$$

The invariance under scaling $x^\mu \to \lambda x^\mu$ generates the dilatation operator D which acts on a function as,

$$D f(x) = i x^\mu \partial_\mu f, \tag{3.20}$$

which gives the following algebra

$$[D, P_\alpha] = -i P_\alpha \tag{3.21}$$
$$[D, M_{\mu\nu}] = 0. \tag{3.22}$$

The last ingredient is the special conformal transformation which acts on a function f as,

$$K_\mu f(x) = i\left(x^2 \partial_\mu - 2x_\mu x^\nu \partial_\nu - 2x_\mu \Delta\right) f(x). \tag{3.23}$$

The commutator with other generators are

$$[M_{\mu\nu}, K_\rho] = -i(\eta_{\mu\rho} K_\nu - \eta_{\nu\rho} K_\mu) \tag{3.24}$$

$$[D, K_\mu] = iK_\mu \tag{3.25}$$
$$[P_\mu, K_\nu] = 2iM_{\mu\nu} - 2i\eta_{\mu\nu}D. \tag{3.26}$$

We can combine the generators of translation, scaling, and special conformal translation with the Lorentz generators to explicitly show that the algebra of the conformal group is isomorphic to $SO(d, 2)$, which is as we saw the symmetry group of pure AdS spacetime. We write put generators into a single matrix generator,

$$J_{ab} = \begin{bmatrix} M_{\mu\nu} & \frac{1}{2}(K_\mu - P_\mu) & \frac{1}{2}(K_\mu + P_\mu) \\ -\frac{1}{2}(K_\nu - P_\nu) & 0 & -D \\ -\frac{1}{2}(K_\nu + P_\nu) & D & 0 \end{bmatrix}, \tag{3.27}$$

where $a, b \in [0, d+1]$ and $\mu, \nu \in [0, d-1]$. Which gives us the following algebra analogous to the Lorentz symmetry group being isomorphic to $SO(1, d-1)$ [53],

$$[J_{ab}, J_{cd}] = i(G_{bc}J_{ad} - G_{ac}J_{bd} - G_{bd}J_{ac} + G_{ad}J_{bc}), \tag{3.28}$$

where $G_{ab} = \text{diag}(-, +, ..., +, -)$, which is exactly the same expression as we get for the Lorentz generators.

The conformal algebra when $d = 2$ is a special case and its algebra is infinite-dimensional, while the $d > 2$ case above is finite-dimensional. In other words, the number of independent conformal transformations we can do in 2d is infinite. This is because any 2d metric can, by a coordinate transformation, be written of the form $g_{\mu\nu} = \Omega^2\eta_{\mu\nu}$ globally, namely it can always be put on a form where it is conformally flat. In higher dimensions, it can only be written on this form locally. Historically, the term *conformal field theory* used to be exclusively understood as 2d conformal field theory, and it was mainly developed to better understand string theory in the 1980s. Due to the advent of AdS/CFT, the interest and development of higher dimensional CFT became more popular.

The way we describe CFTs is very different from the language of QFTs. In ordinary QFTs we are interested in S-matrices and scattering, but not in CFTs. The theory must be scale-invariant which doesn't allow massive excitations. In addition, we can't have length scales like Compton wavelength, and so scattering isn't of interest. Instead, we are mostly interested in mainly two aspects of a QFT: Correlation functions and operators.

The commutators in Eqs. (3.21) and (3.25) resemble ladder operators where P_μ is analogous to the creation operator and K_μ is the annihilation operator. We can also think of D as analogous to the Hamilton operator. We can therefore create a representation of the operators that are similar to something we already know from QFTs. In this representation, the operators are eigenfunctions of the dilation operator D, and it annihilates the vacuum state. If we choose the eigenvalue of D to be $-i\Delta$, then we get under a scaling transformation $x^\mu \to \lambda x^\mu$ that the fields transform as

$$\Phi(x) \to \Phi'(x) = \lambda^\Delta \Phi(\lambda x). \tag{3.29}$$

Δ is called the scaling dimension, and in this representation P_μ raises Δ while K_μ lowers it. The Δ must have a lower bound, and for unitary CFTs the bound is

$\Delta \geq (d-2)/2$. There must be operators with a scaling dimension at this lower bound which are annihilated by K_μ, analogous to the vacuum state being annihilated by the lowering operator in an ordinary QFT. Such operators are called primary operators. With this in mind, we build commutators relation for D and K_μ on top of an already existing Poincare representation of a field or operator $\Phi(x)$. This representation has the following commutator algebra for the conformal field group acting on an operator $\Phi(x)$ [1]

$$[K_\mu, \Phi(x)] = \left(i\left(x^2\partial_\mu - 2x_\mu x^\nu\partial_\nu - 2x_\mu\Delta\right) - 2x^\nu\Sigma_{\mu\nu}\right)\Phi(x), \tag{3.30}$$
$$[D, \Phi(x)] = i\left(-\Delta + x^\mu\partial_\mu\right)\Phi(x), \tag{3.31}$$
$$[P_\mu, \Phi(x)] = i\partial_\mu\Phi(x), \tag{3.32}$$
$$[M_{\mu\nu}, \Phi(x)] = (i(x_\mu\partial_\nu - x_\nu\partial_\mu) + \Sigma_{\mu\nu})\Phi(x), \tag{3.33}$$

where $\Sigma_{\mu\nu}$ acts on the intrinsic degrees of freedom of the field which for example gives rise to spin.

The conformal group severely restricts the correlation functions of an operator and using the algebra above, one can show that the 2-point correlation function of a scalar field $\phi(x)$ is [1]

$$\langle\phi(x)\phi(y)\rangle \propto \frac{1}{|x-y|^{2\Delta}}. \tag{3.34}$$

On the other hand, the correlation function between two primary scalar fields with different scaling dimension vanishes.

3.2.1 Properties of a CFT

Conformal field theory is a huge field in physics, and we don't have time to go through all details of it. Instead, we will in this section mention important properties we will need later.

The stress-energy tensor is an important quantity in a CFT, and it can easily be seen from the conformal invariance of the theory that the trace of this tensor must vanish. Consider conformally scaling the metric $g_{\mu\nu} \to e^{2\rho}g_{\mu\nu}$, then the infinitesimal variation is $\delta g_{\mu\nu} = 2\rho g_{\mu\nu}$, thus varying an arbitrary action with respect to the metric yields

$$\delta S = \int d^2x \frac{\delta S}{\delta g^{\mu\nu}}\delta g_{\mu\nu} = \int d^2x 2\rho\frac{\delta S}{\delta g^{\mu\nu}}g_{\mu\nu} = -\int d^2x\sqrt{-g}\rho T^{\mu\nu}g_{\mu\nu} \tag{3.35}$$
$$= -\int d^2x\sqrt{-g}\rho T^\mu{}_\mu, \tag{3.36}$$

where we have defined the stress-energy tensor as

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}}. \tag{3.37}$$

Since a conformal scaling shouldn't affect the action, we require the trace of the stress-energy tensor to be zero, $T^\mu{}_\mu = 0$. Note that this result is purely classical, and it can be shown that in a quantum theory of two dimensions the expectation value is [51]

$$\langle T^{\mu}{}_{\mu} \rangle = -\frac{c}{12}\mathcal{R}, \tag{3.38}$$

where c is the central charge which is roughly the number of degrees of freedom in the theory. $\mathcal{R}$ is the Ricci tensor, and this is often called the Weyl anomaly and is one of the reasons we need 26 dimensions for bosonic string theory to be a consistent theory. Later on, we will consider the stress-energy tensor of a 2d CFT in the semi-classical region. Namely apply standard quantum field theory techniques to a background that is not necessarily flat. Thus, we will simply always assume that we have quantum expectation value and don't explicitly write the brackets.

Often we want to go from one topology to another, for example, map a flat 2d plane to the surface of a cylinder as we will later in this chapter. First, it is often more convenient to work on 2d CFT in lightcone coordinates written as complex coordinates, $z = x + t = x + it_E$ and $\bar{z} = x - t = x - it_E$. t_E is the Euclidean time, and this makes the non-diagonal elements of the stress-energy tensor zero (given you have fixed the Weyl anomaly), while the two diagonal elements become only a function of their respective coordinates, $T_{zz}(z)$ and $T_{\bar{z}\bar{z}}(\bar{z})$. This is simply because the symmetric stress-energy tensor has three elements, and with energy conservation, we only have two independent elements T_{zz} and $T_{\bar{z}\bar{z}}$. For a good introduction to 2d CFTs, see Ref. [51]. We will just skip ahead and give the results. Going from a cylinder to a 2d plane is equivalent of the conformal transformation $z = e^{-iw}$, where $w = x + it_E$ and the spatial coordinate is the position on a unit circle $x \in [0, 2\pi)$. Undergoing such conformal transformations may yield a different stress-energy tensor and the transformation law is [51]

$$\left(\frac{\partial z}{\partial w}\right)^2 T_{zz}(z) = T_{ww}(w) - \frac{c}{12}S(z, w), \tag{3.39}$$

where

$$S(z, w) = \left(\frac{\partial^3 z}{\partial w^3}\right)\left(\frac{\partial z}{\partial w}\right)^{-1} - \frac{3}{2}\left(\frac{\partial^2 z}{\partial w^2}\right)^2\left(\frac{\partial z}{\partial w}\right)^{-2} \tag{3.40}$$

is called the Schwarzian. For the mapping from cylinder to 2d plane, we have $S(z, w) = 1/2$, so we get

$$T_{ww(\text{cylinder})}(w) = -z^2 T_{zz(\text{plane})}(z) + \frac{c}{24}. \tag{3.41}$$

The same result holds for the complex conjugate, $T_{\bar{z}\bar{z}}(\bar{z})$. If the plane has a vanishing energy in vacuum, $T_{\text{plane}}(z) = 0$, the cylinder will still have a non-zero vacuum energy, given the central charge is non-zero. We can then find the energy density E

$$E = \frac{\int dx T_{t_E t_E}}{2\pi} = -\frac{\int dx (T_{ww} + T_{\bar{w}\bar{w}})}{2\pi} = -\frac{c}{12}, \tag{3.42}$$

where we assumed that T_{ww} and $T_{\bar{w}\bar{w}}$ have the same central charge c. We will use this result when talking about black holes later in this chapter.

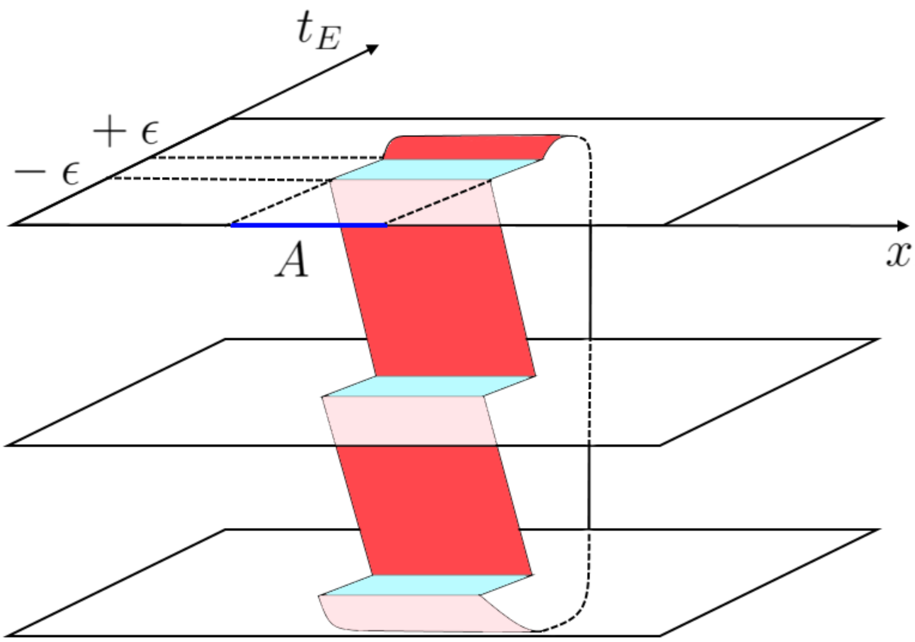

Figure 3.2: How the replica trick works. In this case, we have drawn a 3-sheeted Riemann surface $\mathcal{R}_3$. The blue line is the A-region we have to integrate over, and there is a small gap interval $[-\epsilon, +\epsilon]$ in Euclidean time that allows us to glue the sheets together. Adapted from Ref. [17].

3.2.2 Calculating the Entanglement Entropy in a QFT

A conformal field theory is a special class of quantum field theories, and so we need to use the tools from QFT to find the entanglement entropy of a CFT region. The formula for the entanglement entropy was given earlier as $S_A = -\text{Tr}_{\bar{A}}(\rho_A \ln(\rho_A))$ where A is a given subsystem and $\bar{A}$ is its complement. Earlier we looked at examples of two electron creating a singlet state, but now we have a field $\phi(x)$ that we trace over, only allowing x to be in the region A, $x \in A$. It is proved fruitful to consider the replica method [15, 35, 36], where we find $\text{Tr}_A(\rho_A^n)$ for a positive integer n and analytically vary it and differentiate to find the entropy,

$$S_A = -\frac{\partial}{\partial n}\text{Tr}_A(\rho_A^n)|_{n=1} = -\text{Tr}_A(\rho_A \ln(\rho_A)). \tag{3.43}$$

This can be generalized to adopt the so-called Rényi entropy,

$$S_A^{(n)} = \frac{1}{1-n} \ln \ \text{Tr}_A(\rho_A^n), \tag{3.44}$$

so that the entanglement entropy becomes $S_A = \lim_{n\to 1} S_A^{(n)}$, easily seen to be true by L'Hospital's rule. The diagonal elements in the density matrix can thus be found by using the path integral approach. For simplicity, we will consider a 1+1 dimensional theory, but our formulas can easily be generalized to multiple spatial dimensions. We choose a wavefunction in the ground state $|\Psi\rangle$

$$\langle\psi_0|\Psi\rangle = \Psi[\phi_0(x)] = \int_{t_E=-\infty}^{t_E=0} \mathcal{D}\phi \ e^{-S_E(\phi)}|_{\phi(t_E=0,x)=\phi_0(x)}. \tag{3.45}$$

Here $S_E = \int_0^\infty dt_E \mathcal{L}_E$ is the Euclidean action, and ϕ is a collection of all local fields. An element of the density matrix in the ground state becomes

$$\rho_{\phi_0\phi_0'} = \langle\psi_0|\Psi\rangle\langle\Psi|\psi_0'\rangle = \Psi[\phi_0(x)]\Psi^*[\phi_0'(x)]. \tag{3.46}$$

Here Ψ^* denotes the complex conjugate of Ψ. This can be achieved by replacing the limits of the integral in (3.45) to go from $t_E = \infty$ to $t_E = 0$ (since $t_E = it$ and everything else is real). To obtain the reduced density matrix ρ_A we take the trace over the complement region B by setting $\phi_0(x) = \phi_0'(x)$ and integrating over all fields $\phi_0(x)$ where $x \in B$. Putting this all together, the elements of the reduced density matrix becomes [35]

$$(\rho_A)_{\phi_0\phi_0'} = \frac{1}{Z}\int_{t_E=-\infty}^{t_E=\infty} \mathcal{D}\phi\, e^{-S_E(\phi)} \prod_{x\in A} \delta(\phi(0+\epsilon,x)-\phi_0(x))\delta(\phi(0-\epsilon,x)-\phi_0'(x)), \tag{3.47}$$

where Z is the partition function which ensures that $\mathrm{Tr}_A(\rho_A) = 1$. There is a cut at $t_E = 0$ in the region $x \in A$ with width 2ϵ for a small $\epsilon \ll 1$. $\phi_0(x)$ is attached at the upper part of the cut, and $\phi_0'(x)$ on the lower part. We now need to calculate the matrix products of n reduced density matrices and taking the trace, namely

$$\mathrm{Tr}_A(\rho_A^n) = (\rho_A)_{\phi_0\phi_1}(\rho_A)_{\phi_1\phi_2}\dots(\rho_A)_{\phi_{n-1}\phi_0} \tag{3.48}$$

where the sum over the indices is interpreted as path integrals. One can think of having a n number of independent fields $\phi_i(t_E, x)$ where we have the n number of boundary conditions

$$\phi_i(\epsilon, x) = \phi_{i+1}(-\epsilon, x), \tag{3.49}$$

for $i \in [1, n]$ where we identify $i + n \equiv i$. Figuratively, we glue the two sides of a cut to cuts of a new copy of the same reduced matrix. See Fig. 3.2. The topology we get from this gluing along the branch cuts is a n-sheeted Riemann surface $\mathcal{R}_n$. All these n boundary conditions collectively create one single field $\phi(t_E, x)$ living on $\mathcal{R}_n$. The trace then becomes

$$\mathrm{Tr}_A(\rho_A^n) = \frac{1}{Z^n}\int_{(t_E,x)\in\mathcal{R}_n} \mathcal{D}\phi_0\mathcal{D}\phi_1\dots\mathcal{D}\phi_{n-1}\, e^{-S_E[\phi_0]-S_E[\phi_1]-\dots-S_E[\phi_{n-1}]} = \frac{Z_{\mathcal{R}_n}}{Z^n}. \tag{3.50}$$

We are initially interested in a 2-dimensional CFT living on a plane, $\mathbb{R}^{1,1}$, but the Wick rotation to Euclidean time maps to the complex plane $\mathbb{C}$. The gluing is invariant under the cyclic permutation of the fields $\mathbb{Z}_N$, namely $i \to i + 1$ and $n + 1 \equiv 1$, and so the topology of the Riemann sheet is the quotient space $\frac{\mathbb{C}^n}{\mathbb{Z}_n}$. We are effectively no longer working with the original CFT, but instead with a cyclic product orbifold theory [43].

In an orbifold theory, one can introduce twist operators, satisfying the cyclic permutation of the fields and the twisted boundary conditions. We will not jump into twist field technology, but we will need some results from it: The partition function is given by the two-point function of two twist operators τ_n on a Riemann sheet and can be written as [43, 15]

$$Z_{\mathcal{R}_n} \propto \langle\tau_n(0, -l/2)\tau_{-n}(0, l/2)\rangle^n_{\mathcal{R}_n} \propto \frac{1}{|l|^{2n\Delta_n}} \tag{3.51}$$

Where we have used the 2-point correlation function from Eq. (3.34) by treating the fields as primary. Using the tools of 2d CFT, one can also show that the scaling dimension is

$$\Delta_n = \frac{c}{12}\left(1 - \frac{1}{n^2}\right). \tag{3.52}$$

Thus, we get from the formula for Renyi entropy

$$S_A^{(n)} = \frac{1}{1-n}\ln\left(\frac{Z_{\mathcal{R}_n}}{Z^n}\right) \tag{3.53}$$

$$= \frac{1}{1-n}\ln\left(\frac{(l/a)^{2n\Delta_n}}{(l/a)^{2n\Delta_1}}\right) \tag{3.54}$$

$$= \frac{c}{6}\frac{n+1}{n}\ln\left(\frac{l}{a}\right) \tag{3.55}$$

where we introduced a UV cutoff a. Taking the limit $n \to 1$, we get

$$S_A = \frac{c}{3}\ln\left(\frac{l}{a}\right). \tag{3.56}$$

It is rather remarkable that this equation is only a function of the length l and the central charge c. It seemingly doesn't care about any other details of the CFT_2, and it doesn't seem to tell us much about how the degrees of freedom are entangled. It does, however, show us that it is linear in the central charge, and so one often uses the equation to find the central charge of a CFT_2.

Similarly, a CFT_2 that lives on the boundary of a cylinder, $\mathbb{R} \times \mathbb{S}^1$, where the spatial dimension is the circle $\mathbb{S}^1$, has the entanglement entropy [15, 35]

$$S_A = \frac{c}{3}\ln\left(\frac{L}{\pi a}\sin\left(\frac{\pi l}{L}\right)\right), \tag{3.57}$$

where L is the circumference of the cylinder.

In a thermal system on a flat plane or a large cylinder, $L \to \infty$, we get

$$S_A = \frac{c}{3}\ln\left(\frac{\beta}{\pi a}\sinh\left(\frac{\pi l}{\beta}\right)\right), \tag{3.58}$$

which in the large temperature limit, $\beta \ll l$, simply becomes extensive, thermodynamic entropy as expected,

$$S_A \approx \frac{c\pi l T}{3}, \tag{3.59}$$

which signals that the entanglement is broken.

3.3 AdS/CFT Conjecture

In 1997 Maldacena found an intriguing relationship between asymptotic anti-de Sitter spacetimes and conformal field theories [32]. He initially conjectured this relationship by looking at a 4 dimensional, $\mathcal{N} = 4$ supersymmetric Yang-Mills model, a

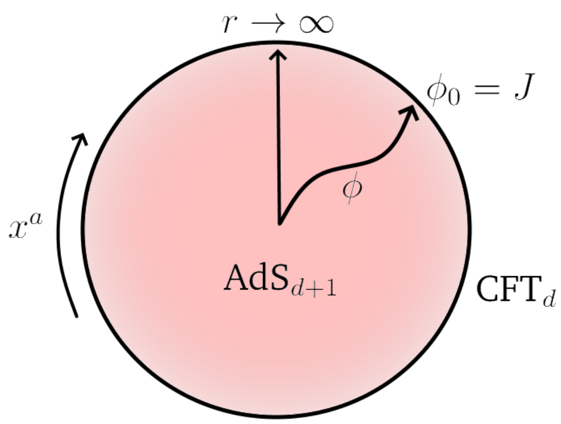

Figure 3.3: A spatial slice of AdS_3 that is equivalent to the CFT_2. More generally, an asymptotic AdS_{d+1} description is equivalent to a CFT_d description. The conformal field theory is placed infinitely far away, $r \to \infty$, which corresponds to $z \to 0$ in Poincare patch coordinates. The part of the bulk field ϕ that blows up or stays constant at the horizon ϕ_0 can be interpreted as the source J of the CFT.

supersymmetric generalization of the Yang-Mills theories that govern the weak and strong nuclear forces. He showed that this theory is equivalent to IIB string theory in an AdS_5 background so that the supersymmetric Yang-Mills theory lives on its conformal boundary.

Today we view the conjecture as being valid outside the realm of string theory. It states that a gravitational universe described by quantum gravity in d-dimensions which is AdS asymptotic can be equivalently described as a conformal field theory in $(d-1)$-dimensions living on the conformal boundary of the bulk universe. This is a powerful statement given that gravity is not well understood at high energies. One can describe gravity as an effective gravity theory at lower energies since the infinite set of parameters needed in the non-renormalizable theory are suppressed by very high energy scales. Thus, combining quantum mechanics and gravity at lower energies is not a problem, but at higher energies, we don't know what to do.

It might seem unclear how we apply it to our universe. Neither do we live in an AdS asymptotic universe, nor is our universe conformal. Had our universe been a quantum conformal field theory we wouldn't have any notion of distances. There would be no such thing as Compton wavelengths or black hole areas. Our best theory to describe gauge theories such as the strong nuclear force $SU(3)$ is classically conformal, but, as David Tong puts it, "the weight of the world relies on the fact that Yang-Mills fails to be conformal on a quantum level" [51]. The hope is nevertheless to have a solid understanding of how to relate the gauge theory with the gravity theory before we try to apply it to our seemingly de Sitter asymptotic and non-conformal quantum field theory universe.

3.3.1 Scalar Field

Let's look at an example where a field in pure AdS corresponds to a boundary field at the limit radial infinity, $r \to \infty$, where the conformal boundary lies. We will do the calculations in the Poincare patch where this limit is equivalent to $z \to 0$. The equation of motion for a massive scalar field ϕ in a fixed gravity background is

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi = m^2\phi, \tag{3.60}$$

which can be written as

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi\right) = m^2\phi. \tag{3.61}$$

We use the Poincare patch metric from Eq. (3.14) and write the flat coordinates as $x = (t, x^i)$ in d dimensions with the addition of z, which gives us the conformal boundary at $z \to 0$. Thus, the whole AdS spacetime is $(d+1)$-dimensional. The equation becomes

$$\left(\frac{z}{R}\right)^d \partial_z\left(\left(\frac{R}{z}\right)^{d-2}\partial_z\phi\right) + \left(\frac{R}{z}\right)^2\partial_a\partial_a\phi = m^2\phi, \tag{3.62}$$

where a runs over the interval $a \in [0, d-1]$. Assuming the coordinates orthogonal to z behave like a free wave, the solution will be of the form $\phi(x,z) = e^{ik_a x^a} f_k(z)$, giving us the equation

$$\partial_z^2 f_k - \frac{d-1}{z}\partial_z f_k = \left(\frac{m^2R^2}{z^2} + k^2\right) f_k. \tag{3.63}$$

We are only interested in how this equation evolves close to the boundary, so we look for solutions of the form $f_k \propto z^\Delta$ when $z \to 0$. That gives us the second-order equation

$$\Delta(\Delta-1) - (d-1)\Delta = m^2R^2, \tag{3.64}$$

which has two possible solutions

$$\Delta = \frac{d}{2} \pm \sqrt{\left(\frac{d}{2}\right)^2 + m^2R^2}. \tag{3.65}$$

As the notation suggests, Δ is the scaling dimension of the dual field living on the conformal boundary as we soon will see.

For simplicity, we assume we get two real solutions for Δ. Letting the largest one be written as Δ, we get the combined solution

$$\phi \approx \phi_0(x)z^{d-\Delta} + \phi_1(x)z^\Delta. \tag{3.66}$$

Here, $\phi_{0,1}$ are functions given by two independent boundary conditions. One can be found by looking at the past horizon $z \to \infty$, while the other can be found by looking at the conformal boundary $z \to 0$. Close to the boundary, $z = \epsilon \to 0$, the last term goes to zero, while the first term is either a constant $(m = 0)$ or it blows up $(m > 0)$. We can now identify the CFT source field φ by removing the divergent part of ϕ,

$$\varphi(x) = \lim_{z\to 0} z^{\Delta-d}\phi(x,z), \tag{3.67}$$

thus, we see that $\phi_1(x)$ acts as a normalizable boundary function.

Eq. (3.63) can be solved using Bessel functions [41], and requiring that the field doesn't increase exponentially in the bulk center at large z, we get the solution $f_k(z) = z^{d/2}K_{\Delta-\frac{d}{2}}(kz)$. Putting this into our equation for the field

$$\phi(x,z) = \int \frac{d^dk}{(2\pi)^d} e^{ik_a x^a} f_k(z), \tag{3.68}$$

where $f_k(z) = \phi_0(k)z^{d-\Delta} + \phi_1(k)z^{\Delta}$. Where $\phi_{0,1}$ has been Fourier transformed to momentum space. To evaluate this wave function on the conformal boundary, we rewrite the action by integration by parts,

$$S = -\frac{1}{2}\int dz d^dx\sqrt{-g}\left[g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + m^2\phi^2\right] \tag{3.69}$$

$$= S_{\text{boundary}} + \frac{1}{2}\int dz d^dx \phi\sqrt{-g}\left[\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi\right) - m^2\phi\right] \tag{3.70}$$

The second term is zero by the equation of motion, and S_{boundary} is the action on the conformal boundary $z = \epsilon \to 0$, which becomes

$$S_{\text{boundary}} = -\frac{1}{2}\int dz d^dx \partial_\mu\left(\sqrt{-g}\phi g^{\mu\nu}\partial_\nu\phi\right)|_{z=\epsilon} \tag{3.71}$$

$$= \frac{1}{2}\int_{\text{boundary}} d^dx\sqrt{-g}\phi g^{zz}\partial_z\phi \tag{3.72}$$

$$\propto \int d^dk\left(\epsilon^{-2(2\Delta-d)}(d-\Delta)\phi_0(-k)\phi_0(k) + d\phi_0(-k)\phi_1(k)\right), \tag{3.73}$$

where we have let $\epsilon \to 0$, only keeping the leading order terms. The first term diverges, and so we need to renormalize it by adding counter terms, similarly to how we renormalize in QFTs. But for simplicity, we will assume the mass is zero, making $\Delta = d$. Now, this term vanishes, and we only need to worry about the second term. Using the Bessel functions above, one can find the ratio [53, 41]

$$\frac{\phi_1(k)}{\phi_0(k)} \propto \left(\frac{k}{2}\right)^d \tag{3.74}$$

Now the action becomes

$$S_{\text{boundary}} \propto \int d^dk\phi_0(-k)\phi_0(k)k^d. \tag{3.75}$$

Fourier transforming back to position space $\phi_0(k) = \int dx e^{-ik_a x^a}\phi_0(x)$ and using the formula [41]

$$\int d^dk e^{ik_a x^a}k^n \propto \frac{1}{|x|^{d+n}}, \tag{3.76}$$

gives us

$$S_{\text{boundary}} \propto \int d^dx d^dy \frac{\phi_0(x)\phi_0(y)}{|x-y|^{2d}}. \tag{3.77}$$

Now, assuming the AdS/CFT correspondence, this should be the action for the conformal field theory living on the boundary. We can now find the generating functional for the gravity

$$Z_{\text{AdS}}[\phi_0] = \int_{\phi(z\to 0)=z^{d-\Delta}\phi_0} \mathcal{D}g\mathcal{D}\phi \; e^{iS[g,\phi]} \approx e^{iS_{\text{boundary}}[\phi_0]}, \tag{3.78}$$

where we used that in the semi-classical limit, we only need the classical solutions, and in that case, only the boundary term in non-zero. Next step is to calculate the correlation function for two operators using this generating functional,

$$\langle O(x)O(y)\rangle = -\frac{1}{Z_{\text{AdS}}}\frac{1}{\partial\phi_0(x)\partial\phi_0(y)}Z_{\text{AdS}} \propto \frac{1}{|x-y|^{2d}}. \tag{3.79}$$

Thus we get exactly what we would expect from the correlators of a two-point function in a CFT, also confirming that $\Delta = d$ is the scaling dimension. We also confirm that ϕ_0 can be interpreted as the source field on the boundary.

Now, we set $m^2 = 0$ to make the math easier, and we didn't run into many problems. Had we, however, allowed the mass to be positive, we would immediately run into infinities, but they could have been solved by adding counter terms to the action at the boundary, effectively removing the infinities. This is analogous to what we do in regular QFT when we run into infinities, we add counter terms infinite in size which effectively gives us an effective action without infinities. In addition, we would still get the same expression above, namely $\langle O(x)O(y)\rangle \propto |x-y|^{-2\Delta}$.

Our example was pure AdS geometry. Had we instead looked at an AdS asymptotic universe where the bulk geometry was far more complicated, then much of the math we already have done would still be true since one can think of the Poincare patch coordinates as being a small region close to the boundary. The big exception is the ratio in Eq. (3.74) since that ratio was determined by a boundary condition deep inside the spacetime. Thus, the boundary action can be drastically different, depending on what happens deep in the bulk. This again could give a very different correlation function for two boundary operators.

This calculation shows what the essential part of the AdS/CFT conjecture is. The conjecture can be precisely stated as [41]

$$Z_{\text{AdS}_{d+1}}[\phi \to \phi_0] = Z_{\text{CFT}_d}[\phi_0], \tag{3.80}$$

where $\phi \to \phi_0$ means we are removing the divergent prefactor of ϕ and setting it to be ϕ_0 at the boundary, acting as a source. Thus if the conjecture holds true, we see that quantum gravity with a negative cosmological constant is equal to something well-defined like a CFT. We don't really have an explicit equation of Z_{AdS}, and in the semi-classical limit the action in gravity with a scalar field would be

$$S_{\text{semi-cl}}[g,\phi] = \int d^{d+1}x\sqrt{-g}\left[\frac{R-\Lambda}{16\pi G_N} + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right]. \tag{3.81}$$

This action is not renormalizable so we shouldn't take it literally in high energy limit. Therefore, we are interested to know in what limit the CFT gives a semi-classical bulk dual.

When we talk about CFT, we often talk about the number of parameters N in the theory. If the CFT is a Yang-Mills system, the natural parameter to talk about is the size of the special unitary group $SU(N)$ which is a gauge symmetry of the action. 't Hooft was able to show that such systems simplified in the large N limit where the effective coupling of the system becomes $\lambda = g_{YM}^2 N$, g_{YM} being a normalization constant of the action. Leaving g_{YM}^2 constant while taking the high N limit, the Yang-Mills system becomes highly coupled, and it was shown explicitly in string theory that this was equivalent to a bulk AdS where the AdS radius was much larger than the Planck length. We therefore get the convenient result that semi-classical gravity in AdS$_{d+1}$ is dual to a strongly coupled CFT$_d$.

Our universe is seemingly not AdS asymptotic, so it doesn't seem like we can directly apply the conjecture to our universe, and whether or not we 'live' in a $(2+1)$ dimensional conformal field theory is not clear. There are however theories in cosmology, such as the Randall-Sundrum model [42] where we and all particles of the standard model live on a 4-dimensional membrane in a 5-dimensional AdS spacetime, and in that case, one can apply the conjecture to our entire universe. But AdS/CFT is nevertheless a handy tool that has been regularly used in theories far outside the realm of quantum gravity. The idea is that if you have a strongly coupled quantum field theory, the math might be easier to do if you jump into the dual gravity description where things aren't strongly coupled anymore. This technique has been implemented in vastly different branches of physics from condensed matter theory to nuclear physics [41].

3.4 Application to Black Holes

Let's look at an AdS universe with a black hole more in detail. The Schwarzschild-AdS metric in $d > 3$ dimensions is simply [26],

$$ds^2 = -f(r)dt^2 + \frac{1}{f(r)}dr^2 + r^2 d\Omega_{d-2}^2, \tag{3.82}$$

$$f(r) = 1 + \frac{r^2}{R^2} - \left(\frac{r_0}{r}\right)^{d-3}. \tag{3.83}$$

Here r_0 is the Schwarzschild radius if the universe was asymptotic flat. Now, the event horizon has a non-trivial equation $f(r_s) = 0$, but we can still find the temperature from Eq. (2.20),

$$T = \frac{d-3}{4\pi r_s} + \frac{(d-1)r_s}{4\pi R^2}. \tag{3.84}$$

In the asymptotic flat case, $R \to \infty$, we get the Hawking radiation for a Schwarzschild metric we found in the Preliminaries.

Since the metric is static, time is a Killing vector, preserving the overall energy of the spacetime. It also allows us to take use of the Tolman relation to find the temperature T_{obs} an observer at distance r feels [50],

$$T_{\text{obs}}(r)\sqrt{-g_{tt}(r)} = \text{constant} \tag{3.85}$$

And so, in the large r limit, $f(r \to \infty) = r^2/R^2$,

$$T_{\text{obs}}(r) \propto \frac{1}{r} \tag{3.86}$$

We see something interesting here. The temperature of any black body in an AdS asymptotic spacetime has no well-defined meaning as it has in an asymptotic Minkowski flat spacetime. In the latter case, the temperature is simply what an observer infinitely far away observes, but in AdS, the thermal radiation will be infinitely redshifted at spatial infinity. Thus the temperature will always be zero infinitely far away. The reason AdS is different is that the more distance there is between an object and an observer, the more the observer has to accelerate to keep the distance constant.

In the context of the AdS/CFT conjecture, we have to think about the AdS bulk in terms of quantum gravity, thus there will be Hawking radiation in the bulk, dominated by massless particles with no angular momentum. As we have seen, massless geodesics hit the boundary in finite time and are reflected back. Even though the temperature of the radiation is zero at infinity, it will be blueshifted back to a finite number again when it returns to the black hole, following Tolman's relation.

Therefore, it makes sense that black holes can be stable in an AdS asymptotic spacetime, unlike asymptotic flat spacetimes. This was shown more rigorously by Hawking and Page in 1983 using the Euclidean path integral approach [27]. They showed that there exists a minimal critical temperature T_c that gives a first-order phase transition between the unstable state $T < T_c$, and higher energies where the black hole can be in equilibrium with the Hawking radiation $T > T_c$. In this case, the heat capacity of the black hole is positive, namely, it becomes more massive when it gets warmer, unlike the negative heat capacity of Schwarzschild black holes in asymptotic flat spacetimes. This phase transition is called the Hawking-Page transition and happens when the radius of the black hole is on the order of the AdS radius [26].

As we earlier mentioned, black holes give an upper bound on entropy [13, 33]. The region contained by a shell of radius r_s can not have a higher entropy than the Bekenstein entropy of a black hole with Schwarzschild radius r_s. So in the case where you have a large number of fields in the bulk, it seems at first glance that the degrees of freedom could be larger in the bulk since they have an additional spatial dimension to propagate in, but fields carry energy which curves space into creating black holes if the energy is large enough. And so black holes bound the entropy and freedom of state so that we don't run into a contradiction. Thus, black holes are essential to understanding the AdS/CFT conjecture.

3.4.1 Example of $\text{AdS}_3/\text{CFT}_2$ Equivalence

In this section, we will look at an explicit example of AdS/CFT equivalence by looking at a black hole in AdS_3. Einstein's equation in 3d doesn't allow for black hole solutions in flat space, but luckily for us, there is a solution for eternal black holes in a universe with negative cosmological constant. This is called the BTZ black hole and was discovered surprisingly recently in 1992 [11]. One didn't expect to find a black hole solution in 3d gravity due to the limitations of three dimensions. In three dimensions, the Riemann tensor is fully determined by the Ricci tensor, which doesn't allow the Riemann tensor to enjoy any additional degrees of freedom causing physics like gravitational waves. In other words, 3d gravity has no local degrees of freedom. Therefore, we say that the spacetime must always be locally AdS even when a black hole is present. Thus, effectively, we have a bunch of coordinate charts, all locally AdS, being glued appropriately together so that a black hole emerges globally.

The black hole is rather remarkable since it doesn't have a curvature singularity, but it does have all other generic properties of Kerr and Schwarzschild black holes, such as an event horizon and surface temperature. For a non-rotating BTZ black hole, the metric is

$$ds^2 = -(r^2 - r_s^2)dt^2 + \frac{1}{r^2 - r_s^2}dr^2 + r^2 d\theta^2. \tag{3.87}$$

Here, the Schwarzschild radius is $r_s = M$, and we have set the AdS length scale to one, $R = 1$. Using Eq. (2.20), we find that the temperature is $T = r_s/2\pi$. As always the entropy is $S = A/4G_N$, but the area now is simply the circumference, thus

$$S = \frac{2\pi r_s}{4G_N} = \frac{2\pi^2 c}{3\beta}, \tag{3.88}$$

where c is the central charge of gravity $c = 3/2G_N$ as found by Brown and Henneaux in 1986 [14]. To find the entropy in the dual CFT$_2$, we are going to use a handy trick on can use in any conformal theory [51].

The partition function for any field theory is

$$Z[\beta] = \text{Tr}e^{-\beta H}. \tag{3.89}$$

The CFT dual of the bulk lives on a cylinder, so we are interested in the vacuum state on a cylinder, and as we saw in Eq. (3.42), this vacuum energy is non-zero, $H = -c/12$. Thus, in the low-temperature limit the partition function favors the vacuum state, so we approximately get

$$Z[\beta] = e^{\beta c/12}. \tag{3.90}$$

On the other hand, we know that large black holes in equilibrium dominate the canonical ensemble so we are interested in a high-temperature at the boundary. The trick is to use the conformal symmetry to our advantage, a trick first seen by Cardy. The conformal symmetry allows us to change the 2d coordinates by a prefactor. The conformal field theory is defined on a cylinder, but if we change our time to Euclidean time $t \to it_E$, the new topology of the CFT is a torus. The Euclidean time coordinate is valid in the region $t_E \in [0, \beta)$. Letting x be the position coordinate defined in the region $x \in [0, 2\pi)$. We now perform a scaling transformation, as allowed by the conformal symmetry,

$$t_E \to \frac{2\pi}{\beta} t_E$$

$$x \to \frac{2\pi}{\beta} x. \tag{3.91}$$

Now, the spatial coordinate is defined on the interval $x \in [0, 4\pi^2/\beta)$, while the Euclidean time coordinate is defined in the same region as the original spatial coordinate, namely $t_E \in [0, 2\pi)$. Now, physics doesn't care what we label our coordinates, so we label x to be our new Euclidean time coordinate, and we suddenly get the very nice expression

$$Z[\beta] = Z[4\pi^2/\beta] = e^{\pi^2 c/3\beta}, \tag{3.92}$$

where the last equality is true for high temperatures, $\beta \to 0$. Now that we have a partition function valid in high temperatures, we can find the energy and entropy

$$E = -\frac{\partial \ln Z}{\partial \beta} = \frac{\pi^2 c}{3\beta^2}, \tag{3.93}$$

$$S = \beta E + \ln Z = \frac{2\pi^2 c}{3\beta}. \tag{3.94}$$

The entropy in the CFT_2 is therefore the same as in the dual AdS_3 from Eq. (3.88). Often we don't have the luxury of being able to explicitly calculate whatever we want in arbitrary dimensions, but the case $\mathrm{AdS}_3/\mathrm{CFT}_2$ is often very popular since the math becomes much more tractable. One is often able to show a correspondence directly just like we did. This case will also be useful for us later in the dissertation.

Chapter 4

Entropy in Gravity Theories

We will in this chapter look at how one can find the entanglement entropy in a theory with gravity. And with AdS/CFT in our toolset, we can see how the entanglement entropy for a region in a pure CFT corresponds to the entanglement entropy for a corresponding region in the AdS bulk.

This chapter follows the chronological order of the development of the field. It was early on known that vacuum Minkowski space was highly entangled, and after the advent of AdS/CFT, Ryu and Takayanagi surprised the world with a conjectured relation between the entanglement in a CFT and the minimal surface in the dual bulk. Surprisingly, the formula is exactly the same as Bekenstein's entropy expression except for the area we are interested in isn't necessarily a black hole horizon.

Their conjecture was eventually proved and generalized, and after a decade, one was to apply its generalized formula to black holes. And as we will see, one was suddenly able to reproduce the Page curve of an evaporating black hole.

4.1 Area Law of Entanglement Entropy

A vacuum is not just empty space. The arrival of quantum field theories taught us that spontaneous creation and annihilation of particles happen all the time, and when one tries separate space into two halves, one separates the spontaneous processes of quantum mechanics. This makes the Minkowski vacuum highly entangled.

The ground state for a massless free scalar field in (3+1)d has been shown to have an entanglement entropy in a region to be proportional to its surface area and not its volume by Srednicki [45]. He gave a simple, but powerful argument why we should expect so. Let's say we have a pure universe, and we look at the entanglement entropy in a region A in a sphere with radius r, then the rest of the universe is its complement $\bar{A} = B$. When calculating the reduced density matrix of B we trace out the degrees of freedom inside the sphere, A, namely we integrate over the volume of A. So we expect the entanglement entropy of B, namely $S_B = -\text{Tr}\rho_B \ln \rho_B$, to scale with the volume of A, which is $O(r^3)$. But by the same argument, the entanglement entropy of A should scale with the volume of B which is not $O(r^3)$. This is a contradiction since a pure system should have $S_A = S_B$. On the other hand, the only thing the two regions share is the surface separating them. Therefore, it is natural that the entanglement entropy scales with the area of the entanglement surface.

Srednicki's calculated that the entanglement entropy becomes

$$S = \kappa \frac{\text{Area}(\partial A)}{a^2} \tag{4.1}$$

where a is a UV cutoff, ∂A is the boundary of the region A and κ is a numerical constant. As expected, it is proportional to the area of the surface separating the regions. In addition, the entropy is infinite and is regulated by a cutoff. This can be understood from equal time correlation functions between two fields that are spacelike separated. Taking the distance between them to zero gives a power-law behavior that diverges. Large correlation means large entanglement.

Note, however, that in a 2d CFT the area law fails. As we saw in Section 3.2.2 the entanglement entropy scales with the logarithm of the length of A and is not a constant, $O(r^0)$, as the area law would suggest.

4.2 Ryu-Takayanagi Formula

We now introduce the Ryu-Takayanagi formula which was presented in their 5-page long paper in 2006 [44]. They were able to relate the entanglement entropy of a conformal field theory living on the asymptotic boundary of a static AdS with a minimal surface in the AdS dual. The formula for the entanglement entropy of a sub-region A of a CFT in d dimensions states

$$S_A = \frac{\text{Area}(\gamma_A)}{4G_N}. \tag{4.2}$$

Here, G_N is Newton's constant in the AdS_{d+1} bulk, and γ_A is a $(d-1)$-dimensional spatial, static surface in the bulk. This surface must be extremal, and if there exist several such surfaces, we pick the one with the least area. This surface is always two dimensions smaller than the bulk, and its endpoints coincide with the endpoints of A at bulk infinity, $\partial\gamma_A = \partial A$. We also require the surface γ_A to be homologous to A, meaning we should be able to continuously deform the curve to coincide with A. In the case of a spacetime with a black hole, we would then not be able to deform the path through its curvature singularity. The Ryu and Takayanagi formulated also requires the spacetime to be static, and so an extremal surface would always lie on a constant time slice.

The first thing one might notice with this equation is that the equation is identical to the Bekenstein entropy of a black hole with the exception that we want the area of a different surface. But this isn't a coincidence, and it is what one would have excepted in an AdS spacetime with a stable black hole. If A is the entire CFT_2 region, then one can show that the RT surface wraps around the black hole, giving the same entropy as in the end of the last chapter in the case $\text{AdS}_3/\text{CFT}_2$. Therefore, the RT surface prescription can also describe thermal entropy and not just entanglement entropy.

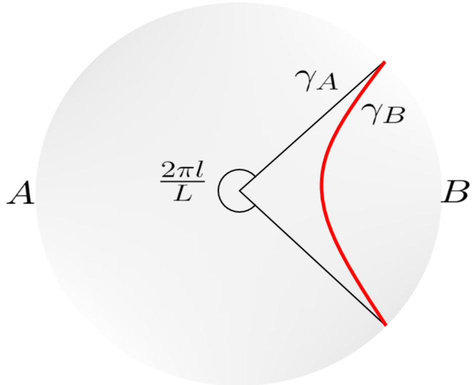

Figure 4.1: Constant time slice of a 3d AdS spacetime. The minimal surface γ_A connecting the boundary points of A is the same as the minimal surface γ_B for B. This is as expected since the entanglement entropy of regions A and its complement B should be the same when the total state is pure. We have also depicted the angle between the boundary points of A, namely $\Delta\theta = 2\pi l/L$ as used in our calculation.

To show the validity of this equation, we will look at an exactly solvable case and explicitly show that the minimal surface (divided by $4G_N$) in AdS gives the entanglement entropy of a region with the same boundary points. We will follow the procedure in [12] to find the minimal surface in AdS_3 and show that the RT formula gives the correct entanglement entropy for the corresponding CFT_2 region. Any surface area bounded at infinity will have a diverging area, thus, in the same way we need a UV cutoff a for the CFT_2, we need at cutoff radius ρ_0 in global coordinates (t, ρ, θ). We only calculate the area up to this radius, and later we will see that we can identify the length scale to the UV cutoff in the CFT as $e^{\rho_0} \propto L/a$.

The calculation is easiest to do if we embed the three dimensional AdS in a flat $(2+2)$-dimensional Minkowski-like spacetime $\mathbb{R}^{(2,2)}$. This is exactly the hypersurface defined in Eq. (3.3) with $d = 2$. In three dimensional AdS, the minimal surface will be a spatial geodesic which we know well how to calculate from any undergraduate class in General Relativity. But now, we constrain the geodesic to be on the hypersurface, so we add a Lagrange multiplier term to the action,

$$S = \frac{1}{2}\int d\lambda\ \left[g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu + \mu\left(g_{\mu\nu}x^\mu x^\nu + R^2\right)\right], \tag{4.3}$$

where μ is the Lagrange multiplier, and $\dot{x}^\mu = \frac{dx^\mu}{d\lambda}$. Varying this function with respect to x^μ we get $\mu x^\mu = \ddot{x}^\mu$. Contracting with x_μ yields $-R^2\mu = \ddot{x}^\mu x_\mu$. Noting that $\frac{d^2 x^\mu x_\mu}{d\lambda^2} = \frac{d^2\left(-R^2\right)}{d\lambda^2} = 0 = \ddot{x}^\mu x_\mu + \dot{x}^2$ gives $\mu = \dot{x}^2/R^2$, and we get

$$x^\mu \dot{x}^2 = \ddot{x}^\mu. \tag{4.4}$$

We have the freedom to reparameterize λ so that $\dot{x}^2 = 1$. This is because we want the distance to simply be $\Delta\lambda = \int d\lambda = \lambda_1 - \lambda_0$. Thus the general solution to the equation of motion for a spacelike geodesic becomes

$$x^\mu = m^\mu e^{\lambda/R} + n^\mu e^{-\lambda/R}, \tag{4.5}$$

where both m^μ and n^μ are constants and obey $m^2 = n^2 = 0$ and $2m^\mu n_\mu = -R^2$. The length of the geodesic is $\Delta\lambda = \lambda_1 - \lambda_0$, and we note

$$x^\mu(\lambda_1)x_\mu(\lambda_0) = m^\mu n_\mu \left(e^{(\lambda_1-\lambda_0)/R} + e^{-(\lambda_1-\lambda_0)/R}\right) = 2m^\mu n_\mu \cosh(\Delta\lambda/R) \tag{4.6}$$

$$= -R^2 \cosh(\Delta\lambda/R). \tag{4.7}$$

We switch over to global coordinates as we previously discussed, namely

$$x^0 = R\cosh(\rho)\cos(\tau), \tag{4.8}$$

$$x^1 = R\cosh(\rho)\sin(\tau), \tag{4.9}$$

$$x^2 = R\sinh(\rho)\cos(\theta), \tag{4.10}$$

$$x^3 = R\sinh(\rho)\sin(\theta). \tag{4.11}$$

Here, θ is the polar coordinate of the two spatial dimensions as a function λ. So the curve has endpoints at $\theta(\lambda_0) = 0$ and $\theta(\lambda_1) = 2\pi l/L$, where l is the length of the subsystem A, and L is the circumference, both at the cutoff radius $\rho = \rho_0$. The radial coordinate ρ is also a function of λ, and the endpoints are the same $\rho(\lambda_0) = \rho(\lambda_1) = \rho_0$. We have chosen a spatial slice, so τ remains constant. Entering the above coordinates into (4.6) finally gives us

$$\cosh(\Delta\lambda/R) = 1 + 2\sinh^2(\rho_0)\sin^2\left(\frac{\pi l}{L}\right). \tag{4.12}$$

We now have everything we need, and by entering in $\Delta\lambda = \text{Area}(\gamma_A)$ in the RT formula, we get by assuming $\rho_0 \gg 1$

$$S_A = \frac{\Delta\lambda}{4G_N} \simeq \frac{R}{4G_N}\ln\left(e^{2\rho_0}\sin^2\left(\frac{\pi l}{L}\right)\right) = \frac{c}{3}\ln\left(e^{\rho_0}\sin\left(\frac{\pi l}{L}\right)\right), \tag{4.13}$$

where the central charge is $c = 3R/2G_N$ as found by Brown and Henneaux in Ref. [14]. Comparing this to Eq. (3.57), this describes the relationship between the two cutoff lengths scales in the two theories, $e^{\rho_0} \propto L/a$.

If we now add matter to our AdS universe but still require it to be static, we get an eternal non-evaporating black hole in a AdS asymptotic universe. See Fig. 4.2. As seen in the figure, now the two minimal surfaces γ_A and γ_B are different which is because the black hole is in a thermal, mixed state. Thus, one should not expect the two CFT boundary regions A and B to have the same fine-grained entropy anymore. The two surfaces have to be different due to the requirement that the surfaces are homologous to its corresponding CFT region. In a constant time slice, the minimal surface can not cross the event horizon without blowing up, but by making A very small, the surface suddenly becomes disconnected into two pieces; one wraps around the event horizon and one is attached to the boundaries of A and is, therefore, closer to the conformal boundary [36]. This shows how one should think of entanglement entropy when the total state is mixed; a subregion gets a thermal contribution and an entanglement contribution to its fine-grained entropy. Note however that in this case, the connected surface that partially wraps around the black hole and ends on the boundary of A is still extremal, it is just not minimal anymore, and therefore, does not yield the correct fine-grained entropy.

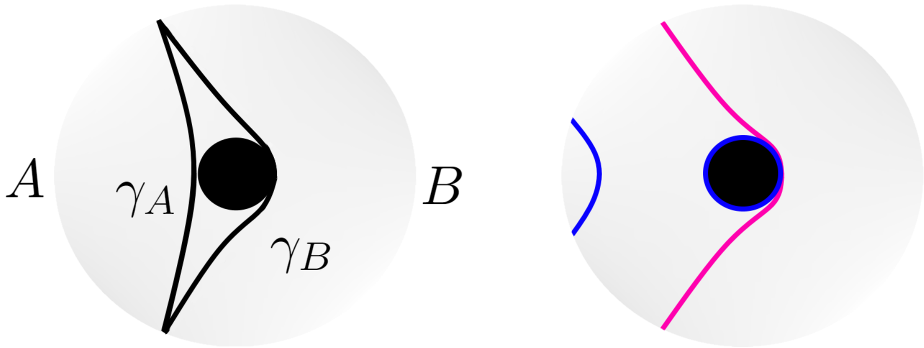

Figure 4.2: AdS asymptotic bulk with a non-evaporating black hole. The black hole is in a thermal state, so the AdS universe is mixed and hence the entropy of the two different regions of the boundary CFT is different (left picture), due to the homologous constraint which doesn't allow the path to enter the black hole in a constant time slice. In the right picture, we see that when the boundary A becomes very small, the minimal surface γ_B yields two contributions: the black hole event horizon area and the infinite radial part.

It might seem counter-intuitive that a part of the minimal extremal surface is the event horizon, but in a time-symmetric spacetime, the apparent horizon (the surface where an outgoing null geodesic doesn't expand) is simply the minimal surface. And for eternal non-evaporating black holes, the apparent horizon coincides with the event horizon. Therefore, as argued in Ref. [22], we can view the Bekenstein entropy of a black hole as an entanglement entropy as a result of dividing space into two parts; the interior of the black hole and the exterior as mentioned in Chapter 2.

In this paper, we are interested in finding a fine-grained entropy formula we can apply to a black hole to understand the entropy of its Hawking radiation. But realistic black holes are not static. They are formed from gravitational collapse, and they evaporate into radiation until they are gone. Therefore, we are not allowed to apply the RT formula to our realistic black hole.

The original Ryu-Takayanagi paper presented the formula simply as a conjecture. They noticed that for the calculations above and other examples, the formula gave the correct entanglement entropy. But seven years later, in 2013, Maldacena and Lewkowycz proved the conjecture by using gravitational path integrals [31], the same approach one can take to derive the Bekenstein entropy of a black hole.

One can understand the RT surface from the gravitational perspective. As mentioned, by dividing a pure vacuum into two regions, we get two mixed regions that have a fine-grained entropy proportional to the area of the dividing surface. And so the entanglement entropy on the boundary region A is equivalent to the entanglement entropy of the bulk spatial slice encapsulated by the RT surface γ_A and the boundary A. As we will later see, when we generalize the RT formula, everything on this spatial slice, like bulk operators and states, has a reconstructed representation on the CFT region A, which will be important later when asking where on the CFT boundary the bulk Hawking radiation has an equivalent description. In the case of a mixed bulk state (like a universe with a black hole), we get two different surfaces γ_A and γ_B like on the left in Fig. 4.2. The bulk region inside γ_A or γ_B is described in their respective boundary regions, while the bulk region outside both RT surfaces needs information about the entire boundary region in order to have a description. We will talk about this more in detail later when discussing entanglement wedges.

4.3 Generalized Entropy

The RT formula was generalized by Hubeny, Rangamani, and Takayanagi (HRT) [29] to include non-stationary spacetimes where the spacelike slice can now evolve in time. Thus in AdS_3, the extremal surface between two boundary points would generally not be the shortest spatial distance between them at a given time, but more generally, the shortest spacelike geodesic. For higher dimensions, the surface would be the extremal surface γ_A with the same boundaries as the sub-region A, $\partial\gamma_A = \partial A$. If there exists more than one such surface, you pick the one with the smallest surface area.

We are interested in finding the fine-grained entropy of the Hawking radiation of an evaporating black hole, but even with this generalization, we are not able to recover the Page curve. Let's say we have an evaporating black hole formed from gravitational collapse, then the only classical HRT surface that is homologous to the entire boundary is no longer the event horizon area, but instead the empty surface [39]. This is also an extremal surface, but then the entropy of the black hole will always be zero. Thus, we are not closer to understanding the fine-grained entropy of Hawking radiation using this framework.

The HRT surface gives the entanglement entropy up to order $O(1/\hbar)$, and thus can not describe the real world with quantum effects. So in 2013, the quantum correction was found by Faulkner, Lewkowycz, and Maldacena [25], and generalized further by Engelhardt and Wall [23]. They noted that the first quantum correction is, just like the RT surface, strikingly similar to the generalized, thermodynamic entropy of a black hole where we now introduce the quantum fields outside of its event horizon, namely

$$S_{\text{gen}}(X) = \frac{\text{Area}(X)}{4G_N^{(d+1)}} + S_{\text{semi-cl}}. \tag{4.14}$$

Here, X is the event horizon, and $S_{\text{semi-cl}}$ is the fine-grained entropy of the quantum fields outside the event horizon in a semi-classical geometry. If one only cares about the interesting physics of black holes, one should add a cut-off surface so that one does not include faraway galaxies. If one is interested in more realistic black holes, namely adding effects like grey-body factors, one should put the cut-off surface well outside of the peak of the effective potential barrier. Notice that this is simply the first quantum correction to the Bekenstein entropy. Engelhardt and Wall also showed that it obeys the second law of thermodynamics $\Delta S_{\text{gen}} \geq 0$ which supports the notion that we are dealing with proper thermodynamic entropy.

Engelhardt and Wall showed that the first quantum correction to the fine-grained entropy of the boundary region A is the surface that first extremalizes the generalized entropy and then picks out the minimal of these possible quantum extremal surfaces (QES) γ homologous to A,

$$S_A = \text{Min Ext}_\gamma \left[\frac{\text{Area}(\gamma)}{4G_N} + S_{\text{semi-cl}}(\Sigma_\gamma)\right] \tag{4.15}$$

Here, Σ_γ is a spatial slice that is bounded by the surface γ and its corresponding region A. Notice that we have the generalized entropy inside the brackets, except that the cut-off surface is the γ surface itself, namely the surface that extremalizes the generalized entropy. We emphasize that this is in no way the exact fine-grained entropy, and so we regard the system of being semi-classical, namely, we treat the quantum fields in the standard way, except that they are moving in a curved space-time.

Historically, the second term was considered a first-order correction to the RT prescription, namely, it could never compete with the area term. But the past 1-2 years this notion has changed as we will see when we apply it to black holes. There can be generalized entropies with vanishing area terms that has a huge contribution from $S_{\text{semi-cl}}$ that competes with a different non-vanishing quantum extremal surface that is dominated by an area term. Even more interestingly, Akers and Penington very recently showed that one has to be careful at how one apply this description to a system since one can run into paradoxes and contradictions [2]. They found a more refined QES formula which isn't as short and compact as the one above, but it gives the correct result. We don't have to worry about this, and the systems we will look at in this dissertation perfectly suits the naïve equation.

An important concept we have to discuss is entanglement wedge reconstruction. The AdS/CFT dictionary tells us that everything that happens in the AdS bulk can equivalently be described on the CFT boundary, but where on the boundary is a given bulk region described? Entanglement wedge reconstruction answers this question by allowing us to understand where on the boundary region a bulk operator is 'reconstructed' [39]. We define the entanglement wedge of the boundary A as a spacelike slice of the bulk that is encapsulated by the quantum extremal surface and the boundary A. More precisely, the entanglement wedge is the domain of dependence of this slice. So, if we have a boundary region A, we can reconstruct bulk operators as long as they live in the entanglement wedge of A. Therefore, the term $S_{\text{semi-cl}}(\Sigma_\gamma)$ is simply the semi-classical, fine-grained entropy of the quantum fields in the entanglement wedge of A. Since we know every operator in the entanglement wedge, it doesn't matter what spatial slice Σ_γ we choose as long as it ends on the QES and A. See Fig. 4.3 a).

When we later apply this formula to black holes, we are initially interested in what is happening in the gravity description and not the CFT on the boundary. So far, we have focused on the entropy on a boundary region in the CFT description, but what bulk region has the equivalent fine-grained entropy? The answer is, of course, the entanglement wedge. If we know the state on a spatial slice in the entanglement wedge, we know the entropy on this slice. When later discussing the information paradox, we are very interested in the entanglement wedge of the radiation, since all information once held by a black hole should eventually somehow be encoded in this wedge if unitarity is preserved.

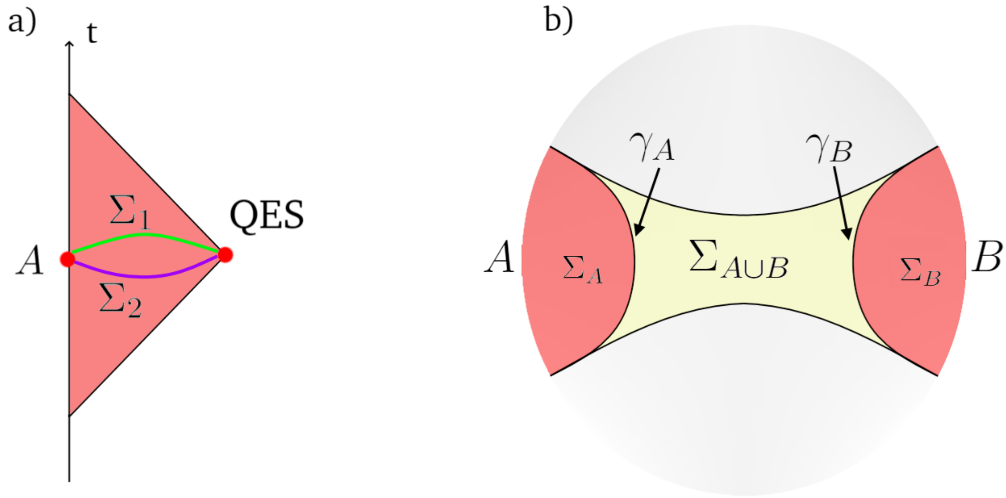

Figure 4.3: In a) we see the entanglement wedge of the boundary region A at a given boundary time t with its corresponding QES. Both Σ_1 and Σ_2 are valid spacelike slices that yield the same semi-classical bulk entropy since they both stretch from the QES to the boundary A at t, and they are both inside the entanglement wedge of A, namely they can not be influenced by anything outside this causal triangle. We also have b) which shows a constant time slice of AdS_3 with its conformal boundary. The red regions are the entanglement wedges of their corresponding boundary regions. The yellow plus red regions are the entanglement wedge of the union of the two boundary regions.

To understand entanglement wedges better, we take a look at Fig. 4.3 b). The figure depicts an AdS universe at a constant time slice. The conformal field boundary has been split into two sub-regions A and B with the corresponding quantum HRT surface $\gamma_{A/B}$. Their respective entanglement wedges have been colored in red showing where we would find the semi-classical fine-grained entropy given in the last term of (4.15). More fundamentally, any bulk operators on these spatial slices can be reconstructed as a boundary operator living on its corresponding CFT boundary region. If we look at the union of both regions $C = A \cup B$, we get that the entanglement wedge is the region colored in yellow plus the region in red. So the relevant fine-grained entropy would be $S_{\text{semi-cl}}(\Sigma_{A\cup B})$. Interestingly, a bulk operator living in the yellow region does not have a reconstructed boundary operator on either A or B, but the union of both does since neither of the two boundary regions have enough information about the operator [19].

From the classical HRT formula, it was known that if the bulk metric satisfied the null energy condition, the classical extremal surface γ_A can not intersect the causal wedge W_A of the boundary region A. See Fig. 4.4. The causal wedge is defined as $W_A = I^-(D_A) \cap I^+(D_A)$, where D_A is the domain of dependence of the boundary region A. And $I^{+(-)}(D_A)$ is the causal future (past) of D_A. Another way of saying this is that the classical extremal surface has to stay outside of the domain of influence.

Engelhardt and Wall also showed in their paper that the same is true for quantum extremal surfaces; it cannot intersect the causal wedge. In addition, the QES is

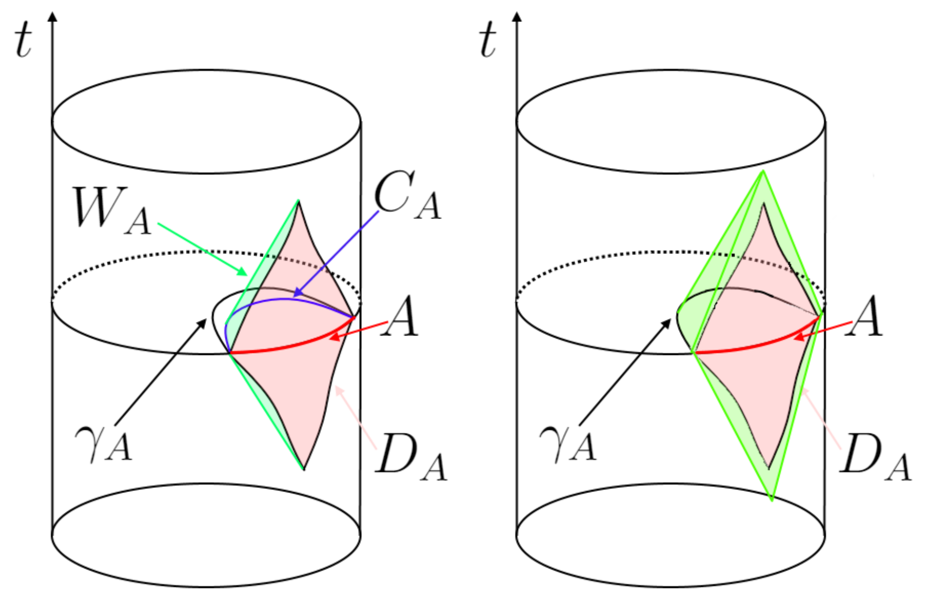

Figure 4.4: In the left figure, the boundary region A of an asymptotic AdS spacetime satisfying the null energy condition on a constant time slice. The domain of dependence is D_A, and the bulk causal wedge is W_A. The causal surface C_A is the spacelike boundary of W_A. γ_A can be either the classical or quantum extremal surface. Neither of them can intersect the bulk causal wedge W_A. In the right figure, the entanglement wedge of the quantum HRT surface γ_A is the green bulk region and is defined as the domain of dependence of the bulk spatial slice confined by γ_A and A. The figure is partially adapted from [23].

also spacelike or null separated from the causal surface C_A, see Fig. 4.4. It also has the strong subadditivity property that any fine-grained entropy needs.

We should also mention that there is an equivalent prescription to the minimal extremal surface formalism generalized by Engelhardt and Wall, and that is the maximin formalism. This prescription was known to be equivalent to classical surfaces in spacetime that don't violate the null energy condition but was very recently proven to hold true for quantum surfaces as well [3]. The equation becomes

$$S_A = \max_{\text{Cauchy}} \min_\gamma \left[\frac{\text{Area}(\gamma)}{4G_N} + S_{\text{semi-cl}}(\Sigma_\gamma) \right]. \tag{4.16}$$

You first find the surface that minimizes the generalized entropy within fixed Cauchy slices, then you pick out the one Cauchy slice that maximizes the generalized entropy. This will be your fine-grained entropy. These two prescriptions were shown to give the exact same quantum surface γ, which gives the same entropy. We will in the rest of this thesis only use the Engelhardt-Wall formalism.

4.4 Application to Black Holes

We are now interested in applying the quantum extremal surface formula to black holes to better understand Hawking radiation entropy. Our discussion will follow closely to Penington in Ref. [39] and Almheiri *et al* in Ref. [8].

We have a black hole formed from a collapsing shell of photons in an AdS asymptotic universe with an equivalent conformal field theory description on the boundary. We want to consider the entanglement entropy of the black hole and the radiation separately, but due to the spherical symmetry of the universe, we can't simply divide the boundary into a compact sub-region A and its complement B. Instead, we will turn the boundary into two systems coupled to each other; one being a conformal field theory with finite temperature T describing the dual, thermal black hole, and the other an auxiliary reservoir describing the Hawking radiation. As in Ref. [39], we assume we have an auxiliary Hilbert space reservoir $\mathcal{H}_{rad}$ where the Hawking radiation is absorbed when it reaches the conformal boundary, while a black hole with finite temperature T is described on the boundary as a CFT $\mathcal{H}_{CFT}$. In other words, we are in a controlled manner, extracting the Hawking radiation out of the AdS asymptotic universe. The CFT region is taken to be the entire conformal boundary of the black hole spacetime.

Now, the quantum extremal surface prescription tells us that entanglement entropy of the black hole is given as the minimal-extremal generalized entropy of the entire boundary region. Since a boundary has no boundary, the extremal surface is not anchored at infinity. Rather, due to spherical symmetry, the surface must be a $(d-1)$-sphere at a finite radius, where d is the number of bulk spatial dimensions. This also includes the empty surface, namely no area term.

Note that since we are allowing Hawking radiation to be absorbed on the boundary, the bulk entropy will depend on the boundary time. This is due to the spacelike slice between the quantum extremal surface and the boundary being dependent on what time the slice ends on the boundary. If we have two boundary times $t_2 > t_1$, then an outgoing Hawking mode can be on the spatial slice (or in the entanglement wedge of the boundary) at t_1, but before t_2 it hits the boundary and is absorbed, lowering the entropy of the bulk. Had we instead used reflecting boundary condition, the photon would still be on the spatial slice at t_2 and thus the entropy would have stayed the same.

To solve the information paradox, we can't simply find the entropy and entanglement wedge of the black hole and simply assume the entropy of the radiation to be equal (given the entire state is pure) and the entanglement wedge to be complementary to that of the black hole. We have to do the math for both, assuming none of the traits of unitarity. The black hole has a clear boundary region, so finding the quantum extremal surface and its corresponding entropy is standard procedure. On the other hand, the radiation absorbed into the radiation Hilbert space $\mathcal{H}_{\text{rad}}$ does not have a boundary in the normal sense, so we need to be wary of how to define its quantum extremal surface.

But the math checks out. Penington showed in Ref. [39] that the quantum extremal surface of the black hole and the radiation perfectly coincide, and that the two has entanglement wedges complementary to each other. Exactly as we expect from unitarity when we divide a pure region into two regions, creating two mixed states.

We will now explain the results of Penington's paper, but we will note one thing. In the review article Ref. [8], the authors didn't assume holography. They assumed that the universe was asymptotically flat thus having no conformal boundary. Instead, they introduced at cut-off shell surface at a couple of Schwarzschild radii away from the center of the black hole. This surface replaces the CFT boundary, so we define the black hole region as everything within the cut-off shell, while the radiation region is everything outside. AdS asymptotic universes simply has a negative cosmological constant. So, there is no reason to think that there are any fundamental differences between the degrees of freedom contained on a spatial slice originating from the quantum extremal surface extending to the CFT boundary, than the degrees of freedom in a flat asymptotic universe from the quantum extremal surface to the cut-off surface. The AdS slice is infinitely large since the CFT is infinitely far away from the black hole, but this infinite, empty space contribution has to be regularized to get a finite entropy. The Hawking radiation that escaped the cut-off shell is fundamentally the same as the Hawking modes we absorbed on the CFT boundary in the AdS asymptotic bulk case. One can think of it in terms of the origin of the quantum extremal surface formula. It was derived from the gravitational path integral which treats AdS asymptotic spacetimes on the same footing as asymptotic flat once. Thus, there are no reasons to think we need the AdS/CFT prescription for our results to hold. We can therefore apply our formulas to universes that resemble our own more.

4.4.1 Entropy of the Black Hole

We are going to present the result assuming we have a black hole formed from gravitational collapse of a pure state in an asymptotic flat spacetime. This spacetime has a cut-off surface far enough from the black hole that we can assume the outside spacetime to be approximated as semi-classical. This surface will be placed a couple of Schwarzschild radii outside of the black hole to make sure the peak of the effective potential of the black hole, experienced by the Hawking radiation, is well inside the shell. Previously, in the AdS/CFT prescription case, we saw that absorbing boundary conditions meant that the fine-grained entropy had to depend on the time on the CFT boundary. Now, we have the cut-off surface instead, and analogously, the time we calculate the entropies will be the time measured by an observer on this surface.

Immediately after the black hole has formed, no radiation has escaped the cut-off surface, and so no mass nor fields exist outside of the cut-off surface, and the entire system of fields and mass is the black hole. Therefore its fine-grained entropy is zero. We get the same result by applying the quantum extremal surface formula to the region. The extremal surface is simply zero, and there are no quantum fields that contribute. As the black hole starts radiating, it will still stay pure as the outgoing Hawking radiation and its interior partner are both encapsulated by the cut-off surface, making the entire state of the two entangled partners pure. But once the outgoing partner escapes the cut-off region, the entropy starts increasing since we only have information about one of the entangled particles within the cutoff region, thus increasing the semi-classical bulk entropy $S_{\text{semi-cl}}$. The area term, on the other hand, stays zero since the extremal surface is not minimized by adding a non-zero area term.

But there is also a different QES that initially is quite large, but eventually will be the minimal extremal generalized entropy and therefore be the fine-grained entropy of the black hole. As mentioned, the entanglement wedge should include the causal wedge, and we immediately conclude that a non-vanishing extremal surface must lie inside the black hole. Penington showed explicitly that there is, in fact, a QES lying close to but inside the event horizon. In ingoing Eddington-Finkelstein coordinates, its position at a given boundary time can be found by rewinding the surface time by the scrambling time,

$$v = -\frac{\beta}{2\pi} \ln S_{BH} \tag{4.17}$$

and then shooting off an ingoing null geodesic from the surface. Here, S_{BH} is the Bekenstein entropy, and v is the infalling time coordinate in Eddington-Finkelstein coordinates. The QES would lie on the infalling time of this geodesic.

One problem with the actual calculation is grey body factors, making it hard to find the exact radial position of the QES. As mentioned previously, there is an effective potential barrier right outside the event horizon, forcing most scalar field particles to be scattered back to the black hole. Penington was able to show that if you absorb the radiation into $\mathcal{H}_{\text{rad}}$ close to the horizon, well before the potential peak, the position of the QES is $r_{\text{QES}} = 2r_{\text{hor}} - r_s$ where r_s is the apparent horizon and r_{hor} is the event horizon.

For eternal black holes, these horizons coincide, but for evaporating black holes, the apparent horizon will have a larger radius than the event horizon. The explanation is simple; at a given time you create an outgoing null geodesic at the apparent horizon. When created, this geodesic will not move due to the definition of an apparent horizon, but a second after, the black hole has evaporated a bit, thus decreasing the event horizon and the apparent horizon. Now the massless particle is outside the

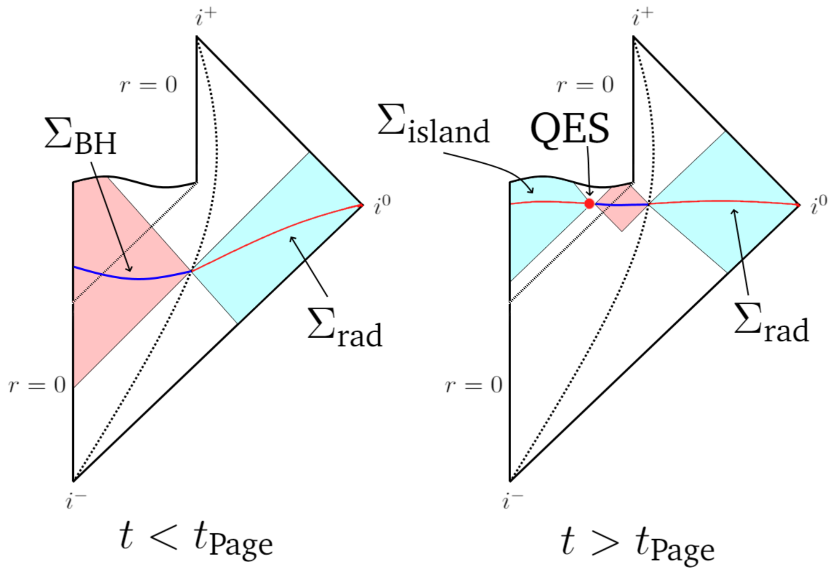

Figure 4.5: Penrose diagram of an evaporating black hole. Before the Page time, the black hole's entanglement wedge (red) is inside the cutoff surface (dashed) and the entanglement wedge of the radiation (blue) is outside it. Each has a corresponding spacelike slice where one can find the semi-classical bulk entropy. After the Page time, an island is created right behind the event horizon. The entanglement wedge of this island belongs to the entanglement wedge of the radiation and thus one has to consider this region as well.

apparent horizon and can propagate to null infinity. Thus, the apparent horizon can not be contained by the event horizon of an evaporating black hole. The position of the QES, $r = 2r_{\text{hor}} - r_s < r_{\text{hor}}$, is slightly inside the event horizon. See Fig. 4.5.

In addition, this non-vanishing QES generalized entropy gets an entropy contribution from the Hawking radiation between itself and the cut-off surface, but overall, the area term will dominate, causing the late-time entropy of the black hole to be very close to the Bekenstein entropy of a black hole. This QES entropy is therefore quite large, mostly due to its area term. Thus, the minimal QES will initially be the first one with a vanishing surface term. Over time, the vanishing QES generalized entropy will strictly increase, due to more and more outgoing Hawking radiation leaving the cut-off surface, leaving the black hole region more and more entangled with radiation outside of the cut-off surface. The second QES entropy, on the other hand, will strictly decrease as the area term becomes smaller and smaller.

Eventually, these two generalized entropy curves intersect and the minimal QES switches. See Fig. 4.6. This happens at the Page time, and this creates something that resembles a second-order phase transition due to the dent in the entropy curve. Now, the full quantum mechanical fine-grained entropy, given by the minimal-extremal generalized entropy, is strictly decreasing meaning it becomes less and less entangled with the Hawking radiation. Eventually, the black hole has completely evaporated, and no gravity or fields are contained within the cut-off surface, so the entropy vanishes.

The cutoff surface has replaced the CFT boundary, but the interpretation of the entanglement wedge on the surface is the same. We are observing the black hole and the radiation degrees of freedom from this surface, and the entanglement wedge is the domain of dependence of a spacelike slice going from the cut-off surface to the QES. Previously, the entanglement wedge was everything inside the cut-off surface, meaning all information about the black hole was inside this region, which intuitively makes perfect sense. But after the Page time, most of the degrees of freedom describing the black hole on the cut-off surface is encoded in the radiation. That means more rigorously that operators defined within the QES actually affect the entropy and state of the Hawking radiation outside the cutoff surface.

4.4.2 Entropy of the Radiation

We will now look at the same process but from the point of view of the Hawking radiation. This point of view perfectly complements the black hole both in terms of equal fine-grained entropy and an entanglement wedge covering the rest of the spatial slice of the universe, even the interior of the black hole after the Page time.

After the black hole has settled, it starts creating Hawking radiation that soon enough escapes the cut-off surface region. Outside this sphere, the radiation is far enough away from the black hole that we can assume the spacetime to be rigid and the effect of gravity to be weak. So one might be tempted to calculate the fine-grained entropy of this region by just using the second term in the generalized entropy, $S_{\text{semi-cl}}(\Sigma)$, where Σ is the spatial slice of everything outside of the cut-off shell. But this radiation originated from a black hole, and so we know it should be highly entangled with its origin, or rather, its interior partners. Thus the semi-classical entropy is not the total fine-grained entropy in the full quantum description. The formula tells us that the QES γ should bound the spacelike region Σ_γ, but everything outside of the cut-off surface is already bounded by itself and spatial infinity, and there is no reason to think one would get an extremal surface by placing a QES outside of the cut-off shell. The next question one might ask is, are we allowed to put a QES outside the region we are considering? The answer is yes. There is one possibility that we haven't considered, but which is analogous to the classical RT case mentioned earlier. We saw that in the case of a black hole in an AdS asymptotic universe, if the boundary region A is large, then the classical RT surface γ becomes disconnected, as in Fig. 4.2. Nothing is stopping us from doing something similar, namely allowing the spacelike slice Σ_γ to be disconnected. We will show in the next chapter that there might be a spatial connection between the two disconnected regions where an additional spatial dimension is allowed to 'access' the inside of the black hole. Penington nevertheless showed that QES lies right behind the horizon, and the region contained by the QES is often called an island. See Fig. 4.5. Now the entanglement wedge of the radiation contains the island plus everything outside the cut-off surface. Therefore, we need to add the bulk fields inside this region to correctly describe the mixed Hawking radiation. This motivates a redefinition of the entanglement entropy in Eq. (4.15) for the radiation [8],

$$S_{\text{rad}} = \text{Min Ext}_\gamma \left[\frac{\text{Area}(\gamma)}{4G_N} + S_{\text{semi-cl}}(\Sigma_{\text{rad}} \cup \Sigma_{\text{island}}) \right]. \tag{4.18}$$

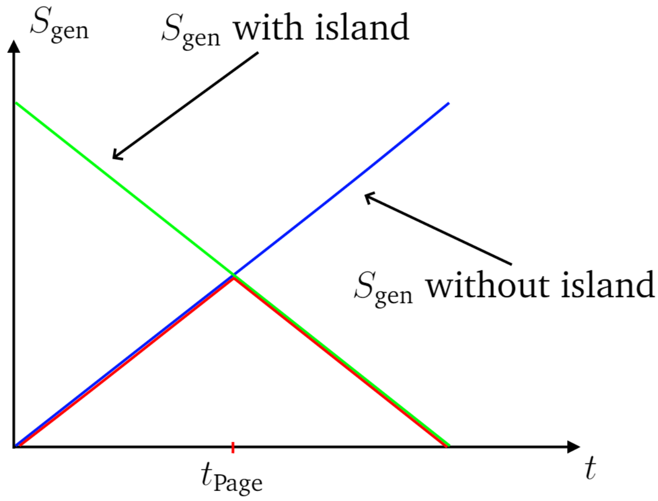

Figure 4.6: The generalized entropy of two extremal surfaces: One where the island vanishes, and one where there is an island. The true fine-grained entropy of the radiation is the minimal of these two curves, the red curve, which reproduces the Page curve.

In this new equation, we see explicitly that we need to add the island spatial region to the radiation region when calculating the semi-classical bulk entropy $S_{\text{semi-cl}}$.

It's important to see why this surface is extremal. The interior Hawking partners will propagate towards the singularity, and the modes closer to the singularity are entangled to the earlier Hawking radiation. This early radiation has already left the cut-off surface and by including its interior partner in the $S_{\text{semi-cl}}$-term we get no contribution to the fine-grained entropy since they combined are pure.

Imagine an ingoing Hawking mode that has an outgoing partner exactly at the cut-off surface. The position of this ingoing mode is where the position of the QES will be, and it is intuitive to see why. If you slightly move the QES further out, you will include interior modes that has partners inside the cut-off surface. The ingoing mode will be included in $S_{\text{semi-cl}}$, but not the outgoing partner, thus increasing the entropy.

Similarly, if you move the QES inwards, closer to the singularity, you will exclude interior modes which have an exterior mode in the entanglement wedge of the radiation. This also increases the entropy. Therefore, the QES is truly extremal. After the Page time, it is also minimal, therefore it is the true fine-grained entropy of the radiation.

Putting a piece of the entanglement wedge of the radiation inside the black hole seems like a leap of faith, but it was proven to be true by applying the replica method. The two papers [9, 40] showed that by using gravitational path integrals, they found a class of saddle points connecting black holes in the n different replicas through complex wormholes. Then, when taking the limit $n \to 1$ to get the full quantum mechanical fine-grained entropy, the presence of these wormholes leads to the new island rule in Eq. (4.18).

4.5 Wormholes and Firewalls

We will make a small digression and make a few comments about the infamous AMPS paradox presented in a paper made by Almheiri, Marolf, Polchinski, and Sully in 2012 [5]. The authors noted that after the Page time, the new, outgoing Hawking radiation seems to be strongly entangled to both the earlier radiation and the interior partner. This violates the strong sub-additivity identity of entanglement entropy [39], and thus we arrive at the paradox. To resolve it, the authors pointed out that one of three well-tested theories had to be given up: The equivalence principle, unitarity of quantum mechanics, or quantum field theory as we know it. The equivalence principle tells us that an infalling observer would never know when she crossed the event horizon of a black hole. Throwing out this principle, the authors imagined the event horizon being surrounded by a wall of fire only by an infalling observer when approaching the black hole. The energy of this firewall comes from the broken entanglement between the interior and the exterior particles of the radiation, which resolves the paradox.

The AMPS authors, however, made one extra assumption that we will not be making in this dissertation. They assumed the black hole degrees of freedom only describe the black hole interior [8]. In other words, they assumed that interior operators only act on the Hilbert space of the interior, and not the Hilbert space of radiation. Initially, this would seem like a valid assumption, but as we just saw, the Hawking radiation is actually partially described by the interior of the black hole after the Page time. And so there is no reason to think this additional hypothesis is true anymore. Thus, most physicists today believe the AMPS paradox to not hold true.

In 2013 Maldacena and Susskind wrote an article where they proposed a resolution to the AMPS firewall [34]. They conjectured a relation between quantum entanglement and something seemingly unrelated, namely wormholes. The conjecture was that an Einstein-Rosen (ER) bridge, a wormhole that connects the two sides of an extended, eternal Schwarzschild black hole, is equivalent to quantum entanglement first proposed by Einstein, Podolsky, and Rosen (EPR), giving rise to the amusing equation ER = EPR.

An eternal Schwarzschild black hole in an AdS background has two disconnected regions that have a CFT boundary. The only way the two disconnected bulk universes can 'communicate' is through a wormhole, while the dual CFTs can only 'communicate' through quantum entanglement. This lead the authors to the conjecture. They also noted more similarities between the two phenomenons, and they applied it to black hole evaporation. Hawking radiation emitted from a one-sided black hole must be entangled to each other, and by the conjecture, there must be a wormhole-like connection between the two, where the radiation replaces the other, disconnected black hole exterior.

This conjecture is not well-defined, but it does seem to hint that the universe creates a spatial connection between two entangled particles that we simply can't see. The next chapter will look at one realization of the ER=EPR conjecture where the spatial connection can be explained by a hidden spatial dimension.

Chapter 5

Double Holography

There are no seemingly good explanations for why the Hawking radiation seems to 'access' the interior of the black hole after Page time. If true, this interpretation would, to put it lightly, come with a paradigm shift, and would change the way we view quantum entanglement.

In this chapter, we will apply the principles of holography twice to show how one can explain the Hawking radiation 'reaching' into the interior of the black hole by accessing an additional spatial dimension. By applying the generalized entropy formula to such a system, we see how the late Hawking radiation and the black hole interior after Page time is naturally entangled. Then one might speculate if something similar, an additional spatial dimension, applies to our real world. This entire chapter is based on two recent papers. One by Almheiri, Mahajan, Maldacena, and Zhao [7], and a more recent paper by Chen, Fisher, Hernandez, Myers, and Ruan [20]. We will also in this section set the AdS length scale $R = 1$ for simplicity.

5.1 Preliminaries

Before jumping into our double holography system, there are a couple of preliminaries we will have to discuss.

5.1.1 2D Gravity

Einstein's equation in 4 dimensions yields an incredible large family of solutions containing rich physics, but the equation in 2 dimensions is nothing but dull [46]. One can show that the equation is satisfied trivially in vacuum, namely,

$$G_{\mu\nu} \equiv \mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu}, \tag{5.1}$$

is zero in 2d for all metrics. This can be seen by looking at the Riemann curvature tensor $\mathcal{R}_{\mu\nu\alpha\beta}$ which has a set of symmetries which only allows one independent element in 2d. Thus, one can easily show that Eq. (5.1) must vanish. One different way to show the triviality of Einstein's gravity in 2d is to show that the Einstein-Hilbert action

$$I_0 = \frac{\phi_0}{16\pi G_N} \int d^2x \sqrt{-g}\mathcal{R}, \tag{5.2}$$

is simply the integral of a total derivative, thus only depends on the topology of spacetime. Here, ϕ_0 is a constant, g is the determinant of the metric $g_{\mu\nu}$ and $\mathcal{R}$ is the Ricci scalar. To spice things up one can look at so-called $f(\mathcal{R})$-theories where you replace the scalar $\mathcal{R}$ in the Hilbert action with $f(\mathcal{R})$, allowing arbitrary functions f of the Ricci scalar. These theories were historically popular attempts at solving the mysteries of dark matter and dark energy by slightly altering Einstein's theory. These theories can be described more generally by keeping the Ricci scalar $\mathcal{R}$ in the Hilbert-Einstein action, but also adding an auxiliary field ϕ, which gives the same equation of motion for gravity. As an example, let's look at the case $f(\mathcal{R}) = \mathcal{R}^2$,

$$\frac{1}{16\pi G_N} \int d^2x \sqrt{-g}\mathcal{R}^2 \to \frac{1}{8\pi G_N} \int d^2x \sqrt{-g} \left(\mathcal{R}\phi - \frac{1}{2}\phi^2 \right). \tag{5.3}$$

Varying the above functional with respect to ϕ gives the classical equation of motion $\phi = \mathcal{R}$, and we end up with the functional on the left. One important example is JT gravity with matter field [6],

$$I = I_0(g, \phi_0) + \frac{1}{16\pi G_N} \left[\int d^2x \sqrt{-g}\phi \left(\mathcal{R} - \Lambda \right) + 2 \int_{\text{boundary}} \phi_b K \right] + I_M(g, \chi), \tag{5.4}$$

where Λ is the negative cosmological constant, and I_0 is the topological Einstein-Hilbert action in 2d from (5.2), and I_M is a matter term with a matter field χ. K is the extrinsic curvature at the boundary, while ϕ_b is the boundary value of ϕ. The equation of motion for ϕ in the second term forces the Ricci curvature to be equal to the cosmological constant, thus the metric is locally AdS. By varying the metric, however, we get a more complex equation of motion that gives a differential equation for the dilaton ϕ, which will describe the more complex geometry of the 2d universe. More accurately, the dilaton describes the backreaction of matter on this AdS_2 background. Note again that the first term in the action is trivially satisfied for all metrics, and so we have a AdS_2 spacetime with matter fields.

JT gravity has been well studied the past decade [4, 24], and it is well known that it exhibits black hole solutions. And since black holes in 2d gravity can't have surface terms, a co-dimension 2 surface is a point, the Bekenstein entropy is analogously shown to be $S_{BH} = \phi/4G_N$ where ϕ is evaluated on the event horizon position. This can be seen from the action itself since one simply replaces the prefactor, $\frac{1}{4G_N} \to \frac{\phi(x)}{4G_N}$.

5.1.2 JT Gravity

JT gravity forces the metric to be AdS_2, and thus, in the 2d gravity coordinates, we can write

$$ds^2_{\text{AdS}_2} = \frac{dt^2 - dx^2}{x^2} = \frac{-4dx^- dx^+}{(x^+ - x^-)^2}, \tag{5.5}$$

where $x^{\pm} = t \pm x$. By varying the JT action with respect to the metric, one can find that the vacuum solution for dilaton that admits a stable black hole [6, 4],

$$\phi = 2\bar{\phi}_r \frac{1 - \pi^2 T_0^2 x^+ x^-}{x^+ - x^-}, \tag{5.6}$$

where $\bar{\phi}_r = \epsilon\phi(x = \epsilon) = \epsilon\phi_b$ is the normalized boundary value of the dilaton, and T_0 is the temperature of the stable black hole.

One might be worried about the Weyl anomaly making the stress-energy tensor proportional to the Ricci curvature, thus non-zero, but it can simply be absorbed by the dilaton ϕ_0 in the Einstein-Hilbert action.

The dilaton above is obviously time-dependent and instead of a singularity positioned at a specific spatial coordinate (like $r = 0$ in Schwarzschild coordinates), the singularity is positioned at

$$x = \left(t^2 - \frac{\pi^2 T_0^2 - 1}{\pi^4 T_0^4}\right)^{1/2} - \frac{1}{\pi^2 T_0^2}, \tag{5.7}$$

which suggests that Poincare coordinates might not be our best choice of coordinates. Instead, we can switch over to coordinates where the dilaton is static, valid in the exterior region of the black hole, by setting $x^{\pm} = f(y^{\pm})$, where

$$f(u) = \frac{1}{\pi T_0}\tanh(\pi T_0 u). \tag{5.8}$$

Now, the dilaton becomes

$$\phi = 2\bar{\phi}_r \pi T_0 \coth\left(\pi T_0 \left[y^+ - y^-\right]\right) \tag{5.9}$$

and the metric

$$ds^2 = \frac{-4\left(\pi T_0\right)^2 dy^+ dy^-}{\sinh^2\left(\pi T_0 \left[y^+ - y^-\right]\right)}. \tag{5.10}$$

These coordinates don't just leave our JT gravity description static, but it is also the coordinates we get from considering the time coordinate t on the conformal boundary time u perspective, namely $u = f^{-1}(t)$.

By setting our dilaton from Eq. (5.9) to be our new spatial coordinate $\phi(\sigma) = r$, and plugging it into Eq. (5.10), we get a Schwarzschild-like metric, making it clear that we are dealing with a black hole [4],

$$ds^2 = -4r^2\left(1 - \left[\frac{2\bar{\phi}_r \pi T_0}{r}\right]^2\right) du^2 + \frac{r^2}{\left(2\bar{\phi}_r \pi T_0\right)^4 \left(1 - \left[\frac{2\bar{\phi}_r \pi T_0}{r}\right]^2\right)} dr^2. \tag{5.11}$$

Our goal later will be to couple our 2d gravity to a CFT_2 bath. And so we define the bath coordinates $y^{\pm} = u \pm \sigma$, where σ is the position coordinate in the bath CFT_2, valid for $\sigma > 0$. And the coupling happens at time $u = 0$.

Our stable black hole will have the Penrose diagram of an AdS-Schwarzschild black hole, as seen in Fig. 5.1, where we have two sides of the black hole. We are interested in the right side, and so we want to calculate the entanglement entropy of the right exterior part from the point of view of the right conformal boundary. Remember that both halves of this universe are causally disconnected, and we assume the whole universe (both halves) is pure. Due to symmetry, we expect them to have the same quantum extremal surface. The AdS boundary has reflective boundary conditions, thus keeping the evaporating black hole stable. This also makes the quantum fields in both halves pure.

The classical extremal surface can be found and becomes

$$S_{\text{cl-gen}} = \frac{A}{4G_N} = \frac{\phi(x^+, x^-)}{4G_N}, \tag{5.12}$$

and using Eq. (5.6), the requirement for an extremal surface is that $\frac{\partial S_{\text{gen}}}{\partial x^{\pm}} = 0$, which gives us

$$x^{\pm}_{\text{CES}} = \pm \frac{1}{\pi T_0}. \tag{5.13}$$

Now, this is just the bifurcation point. It is simply the point that divides the two halves of the universe. But due to the black hole being in equilibrium with the radiation, the bifurcation point will also be the quantum extremal surface of the right conformal boundary, since a spatial slice from this point to the boundary must have a pure semi-classical bulk entropy. This gives the following fine-grained entropy

$$S = \frac{\bar{\phi}_r \pi T_0}{2G_N} \tag{5.14}$$

which is constant in time and proportional to the initial temperature of the black hole. Keep in mind that we are assuming $\phi_0 \ll \phi$.

5.1.3 Branes

Branes are hypersurfaces which arose from string theory. They are arbitrary dimension surfaces that particles are confined on, living in a higher dimensional universe.

We will consider a Planck brane which was first introduced by Randall and Sundrum [42]. They considered a 4 dimensional (3+1)-brane embedded in a 5-dimensional AdS spacetime where all particles of the Standard Model were being confined on the brane, while the mediator of gravity, the graviton, was able to propagate in all the 4 spatial dimensions of the AdS. Their motivation was to explain the hierarchy problem; why is gravity so much weaker than all other forces of nature?

Our purpose is to confine a black hole on a (1+1)d Planck brane embedded in 3d Poincare patch coordinates. The gravity will be governed by the JT gravity explained above, and it will be coupled to a CFT matter. Given that the matter is CFT, we will take it to have an AdS asymptotic dual. But due do the dynamics of JT gravity coupled to matter, the metric isn't trivial, and the position of this brane isn't trivially the same as the conformal boundary of pure AdS_3, namely $z = \epsilon \to 0$ in Poincare patch coordinates. If however the region was perfectly flat, and there was a vanishing stress-energy tensor, the Planck brane would satisfy the requirements of a flat, conformal boundary being dual to a pure AdS_3. The position of the Planck brane from the 3d perspective is therefore not trivial.

5.1.4 Position of the Planck brane

Let's not make any assumptions about the metric on the Planck brane, nor the stress-energy tensor of the matter fields on the brane. This should not be confused with the stress-energy tensor due to the tension of the embedding of the Planck brane in AdS_3. It is well known that any 2d metric can be put of the conformally flat form, which in lightcone coordinates becomes $ds^2 = -\Omega^2(x)dx^+dx^-$. In lightcone coordinate, the non-diagonal elements of the stress-energy tensor vanish, so we need only to worry about $T_{x^+x^+}(x^+)$ and $T_{x^-x^-}(x^-)$ which, as we earlier mentioned, are functions of their respective coordinate. To determine the embedding of the Planck brane in the AdS spacetime, we need to change coordinates to a frame where the metric is flat. This will also make the stress-energy tensor vanish, with vanishing Weyl anomaly, since there is no energy to curve space. If there is no gravity present, the embedding is trivial since it is simply the conformal boundary positioned at $z = \epsilon \to 0$ in Poincare patch coordinates.

Following the procedure of [7], we switch over to new coordinates $w^\pm$ that we ensure has vanishing stress-energy components. We then find the new stress-energy tensor by using the transformation rule with the Schwarzian, as we previously discussed in Section 3.2.1. The transformation is

$$T_{w^+w^+} = \left(\frac{\partial w^+}{\partial x^+}\right)^{-2}\left(T_{x^+x^+} + \frac{c}{12}S(w^+, x^+)\right), \tag{5.15}$$

with an equivalent formula for $T_{w^-w^-}$. Note that we are not using complex coordinates, therefore we get a plus sign on the second term. We require the above stress-energy component (and $T_{w^-w^-}$) to be zero, thus we need w^+ to satisfy $T_{x^+x^+} + \frac{c}{12}\{w^+, x^+\} = 0$. By a Weyl transformation we can set the metric to be flat $ds^2 = -dw^+dw^-$, and thus we have avoided the Weyl anomaly, where the quantum expected value of the trace in a 2d CFT is proportional to the Ricci scalar [51]. Close to the Planck brane, we can assume all the complexity of the AdS asymptotic spacetime can be neglected, and we earlier mentioned, in 3d, any AdS asymptotic spacetime is always locally pure AdS. So close to the brane, the metric is

$$ds_3^2 = \frac{-dw^+dw^- + dz^2_{\text{Planck}}}{z^2_{\text{Planck}}}. \tag{5.16}$$

We require the induced metric Planck brane in the vanishing-energy coordinates $w^\pm$ to be the same as the gravity solution $x^\pm$, namely

$$-\frac{dw^+dw^-}{z^2_{\text{Planck}}} = -\frac{\Omega^2(x)dx^+dx^-}{\epsilon^2}, \tag{5.17}$$

which can be solved to give an explicit equation for z_{Planck}. In our AdS case, we have

$$\Omega^2(x) = \frac{4}{(x^+ - x^-)^2} \tag{5.18}$$

and thus, we know the exact position of the Planck brane in the bulk.

5.2 Setup

So far we have seen the advantage of applying the AdS/CFT conjecture once, but we will see how we can use it twice to get a better understanding of how an additional spatial dimension can 'access' the inside of the black hole, in the spirit of the ER=EPR conjecture. Since we will allow applying the conjecture twice, we have three equivalent descriptions of our system, see Fig. 5.3.

2D JT Gravity with a black hole - The most important description is the one with the 2d JT gravity. In this spacetime, there is a stable black hole of temperature T_0, which due to the topology of two dimensions, will have two disconnected, external regions. Both regions will have a conformal boundary point at infinity. The JT gravity will also be coupled to matter that is described as a CFT_2. At a certain boundary time, $u = 0$, the right exterior region will be coupled to a CFT bath of zero temperature. Deforming the operators to allow the two systems to be coupled releases energy into the JT gravity universe as an energy pulse. Such an event is called a quench, and the energy will eventually give energy to the stable black hole, making it expand and warmer, settling at a new horizon, but now the Hawking radiation of the right exterior is no longer confined to the AdS spacetime and will start leaking into the CFT bath.

3D Gravity - Since the matter of the JT gravity is a CFT_2, we take it to have an AdS_3 dual. In addition, the CFT bath will also have a dual which is pure AdS_3 before the quench. These two AdS_3 duals will be disconnected before the quench, but as will see, the brane that separated the systems will evolve and propagate deeper into the AdS bulk.

We are only going to be interested in the physics close to the conformal boundary where the geometry is close to purely AdS, and locally always AdS_3, even though it may not be globally AdS. Thus the exact geometry deep in this bulk description is not of importance.

Quantum Mechanics - Since the 2d JT gravity is AdS, we take it to have a 1d dual. A quantum field theory in zero spatial dimensions can be thought of as regular quantum mechanics (QM) where we consider a free scalar field evolving in time. Since the 2d universe has two conformal boundaries, we get two points of quantum mechanical description. The left point is called QM_L, and is entangled with the right point QM_R. They have all the degrees of freedom necessary to have the JT gravity dual between them. When the quench happens at boundary time $u = 0$, we couple the right QM point to the CFT bath which effectively supplies QM_R with energy.

Up until now, we have dealt with CFT theories with no boundaries, but not now. As seen in the 2d gravity, before the quench, we have a conformal field theory that stretches from QM_L to QM_R. These two points, when the 2d JT gravity is embedded in the 3d dual, give a conformal boundary with boundaries. After the quench, the only boundary will be QM_L.

To treat the CFTs properly we have to consider boundary CFT (BCFT) procedures, first studied by Cardy [16]. Later, Takayanagi showed how to treat the holographic dual of the BCFT [49]. He showed that if the CFT has a boundary, then we must consider a surface in the AdS asymptotic description which shares the boundary of the CFT. This is an idea similar to the branes considered in Randall and Sundrum's models [42]. In our asymptotic AdS_3, it tells us where the universe stops, so we will call them end-of-the-world (ETW) brane.

In what follows, we will consider the fine-grained entropy of two regions QM_L+Bath and QM_R. The former is the left exterior of the black hole in union with the CFT bath. This region will represent the Hawking radiation, while the right exterior of the black hole represents the black hole region. Remember that the left and the right side of the black hole are causally disconnected, and so their corresponding Hawking radiation will never leak into the other exterior, but the Hawking radiation of the right exterior will leak into the CFT bath after the quench.

The equation for the fine-grained entropy of the right black hole region is simply

$$S = \text{Min Ext}_x \left[\frac{\phi(x)}{4G_N^{(2)}} + S_{\text{semi-cl}}(\Sigma_x) \right], \tag{5.19}$$

where we explicitly write the dimension of the Newton's constant, and Σ_x is the spatial slice from the position of the QES, x, and to the boundary QM_R. The fine-grained bulk entropy is caused by the matter field χ in the matter action. The QES prescription then tells us to extremalize the entropy over $x = (x^+, x^-)$ and pick out the minimal entropy with a corresponding QES. This is what we just did above before the coupling.

However, the brilliance of the double holography system is that it allows us to find the von Neumann bulk entropy in the 2d gravity $S_{\text{semi-cl}}$ as a QES entropy in the AdS_3 bulk dual. So from a 3d perspective, by using Eq. (4.15), we get

$$S = \text{Min Ext}_x \left[\frac{\phi(x)}{4G_N^{(2)}} + S_{\text{semi-cl}}(\Sigma_x) \right] \tag{5.20}$$

$$= \text{Min Ext}_{x,\gamma} \left[\frac{\phi(x)}{4G_N^{(2)}} + \frac{\text{Area}(\gamma(x))}{4G_N^{(3)}} + S_{\text{semi-cl}}(\Sigma_\gamma) \right], \tag{5.21}$$

where Σ_x is the one-dimensional, spatial region from the QES to the conformal boundary QM_R, γ is the dual QES in the AdS_3 bulk, and Σ_γ is the spatial region encapsulated by γ and Σ_x. See Fig. 5.2. Note that γ must end on x and QM_R, so it is a function of the position of the QES, but we still have the freedom to extremalize the surface in the AdS_3 bulk. To the leading order, we can neglect the third term above, and then the matter entropy is simply the classical RT surface in the bulk. In this case, the 'length' of Σ_γ gives the biggest contribution to the matter bulk entropy.

Therefore, this system helps us understand how the entanglement wedge of radiation after the Page time is partly inside the black hole. It's due to the 'extra' bulk dimension reaching into it.

5.3 Before the Page Time

We will now look at what happens when we couple the 2d gravity theory with the CFT_2 bath. The system will go into a quench phase where there is a flow of energy,

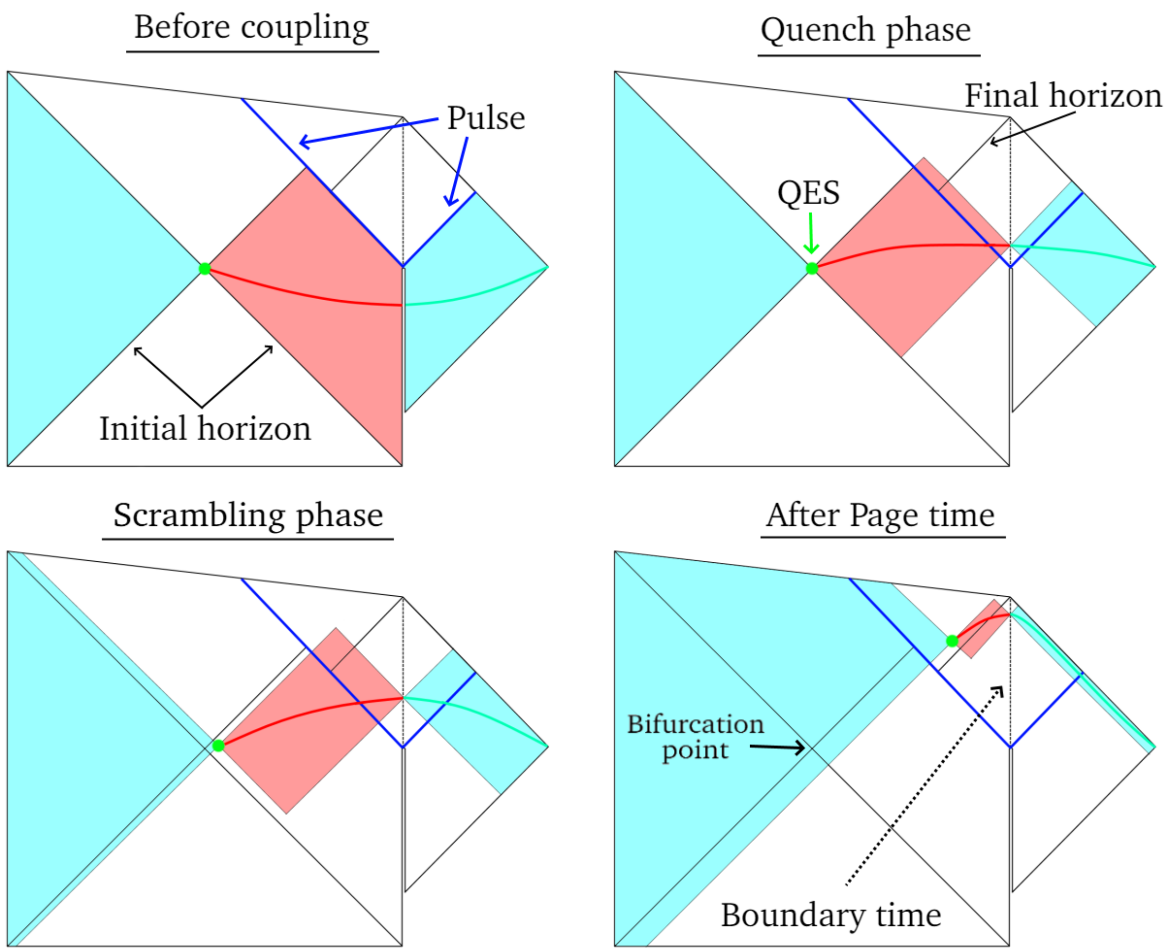

Figure 5.1: The two-dimensional spacetime at different time. The two entanglement wedges are shaded for QM_L+Bath (cyan) and for QM_R (red) which we will think of as the black hole region. The right 2d gravity at constant time slice is the red curve, while the dark cyan curve is the CFT_2 bath. The energy pulse caused by the coupling of the 2d gravity to the zero temperature bath is colored in blue. The minimal QES point is colored in green and is positioned at the bifurcation point before the coupling and immediately after in the quench phase. After the quench phase, it moves spacelike away from the black hole, and after Page time it jumps into the causal future of the quench but behind the final black hole horizon. Partially adapted from [20].

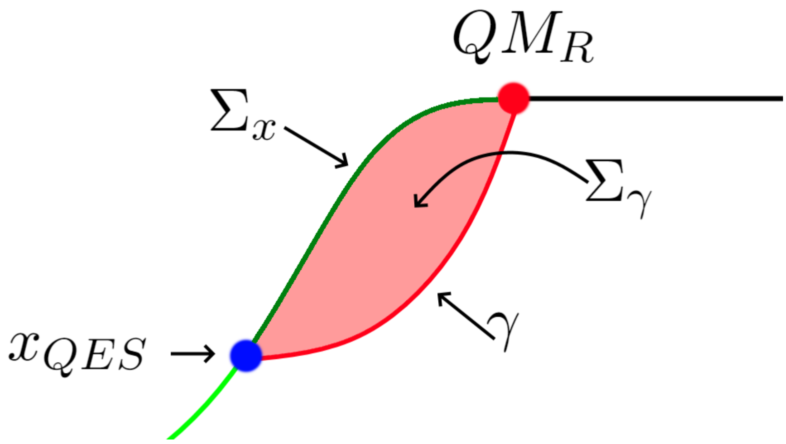

Figure 5.2: The fine-grained entropy of the CFT matter in the spatial region Σ_x is given by the minimal extremal surface prescription in the AdS_3 bulk. To first order, γ is simply the classical RT surface, and we can neglect the contribution from Σ_γ.

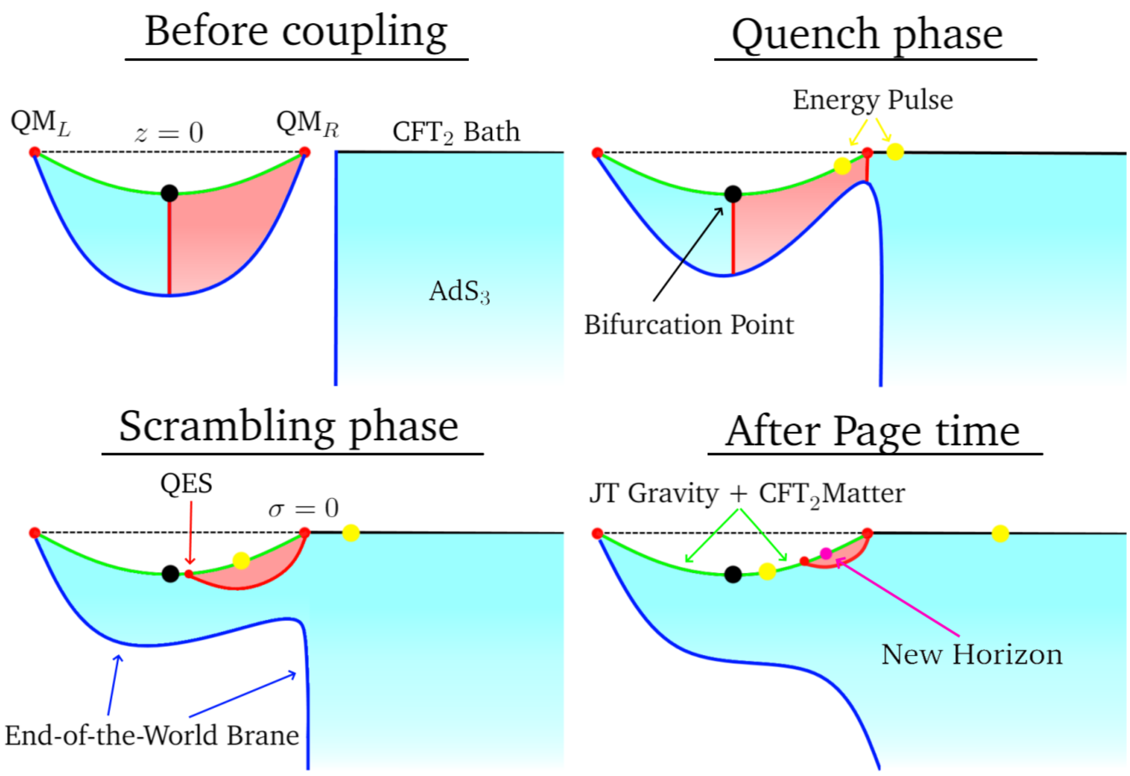

Figure 5.3: Black hole evaporation from the AdS_3 point of view. The entanglement wedge for the QM_L + Bath is colored in cyan and the entanglement wedge for the Hawking radiation is in red. This figure should be compared to Fig. 5.1 where the corresponding 2d entanglement wedges have the same colors.

moving as a null geodesic $x^- = 0$ into the 2d gravity theory. This gives a sharp increase in the entanglement entropy of QM_R which can be understood from the 3d dual as coming from the area term connecting the QM point, $\sigma = 0$, to the ETW brane which is now moving into the deeper parts of the bulk. From the 2d gravity side, it is explained by now having a large positive energy propagating in the 2d gravity, which changes the dilaton. Thus, from the point of view of the boundary point $\sigma = 0$, which is analogous to the observer on the cut-off surface in the last chapter, the increase in entropy can be viewed as a backreaction in the dilaton to the infalling shockwave. In terms of entanglement, the sharp increase can be explained in terms of the creation of short-range entanglement between the two energy pulses originating from $\sigma = 0$ and propagating into the bath and the right side of the gravity dual.

Immediately after the quench happens, a new extremal QES is created close to the original black hole horizon but outside it. This extremal surface moves away from the horizon in a spacelike way and does however not yield a minimal entropy until the quench phase is over. What happens at the bifurcation point is obviously causally disconnected from the boundary when the quench happens, so the dilaton shouldn't immediately change, but the surface is still allowed to move after the quench happens as Hawking modes, that are entangled to the initial black hole, are allowed to escape into the bath. When this new QES becomes minimal, it describes the fine-grained entropy and we call this the scrambling phase. And we see that the entanglement wedge of the radiation (QM_L + bath) is slightly inside the initial black hole, which comes from the Hawking radiation having escaped into the bath with enough information to encode this part of the interior.

Now, this case is different from what we were looking at in the last chapter. There, we had a vanishing minimal extremal surface before the Page time, but now there is no vanishing extremal surface. But we have many key differences. First of all, we are now starting with a stable black hole which has two causally disconnected regions that are entangled. In addition, we are not in Einstein gravity, we are using JT gravity. So we shouldn't be too surprised that we are getting these smaller qualitative differences.

When the system goes into the scrambling phase, the entanglement wedge of QM_L and the bath becomes connected. As the QES moves, the semi-classical bulk entropy of the matter is simply the classical RT surface in the bulk. So from Fig. 5.3, we see how the entanglement wedge of the radiation region $\sigma > 0$ 'stretches' into the black hole interior.

5.4 After the Page Time

After the Page time, a new quantum extremal surface has the lowest generalized entropy. As before, this QES was extremal before the Page time, but at the Page time, the QES close to the bifurcation point does no longer have a minimal generalized entropy. The late-time QES is positioned at the future of the energy pulse after it has created a new black hole horizon. The new surface lies just behind the event horizon and is creating an island inside the interior.

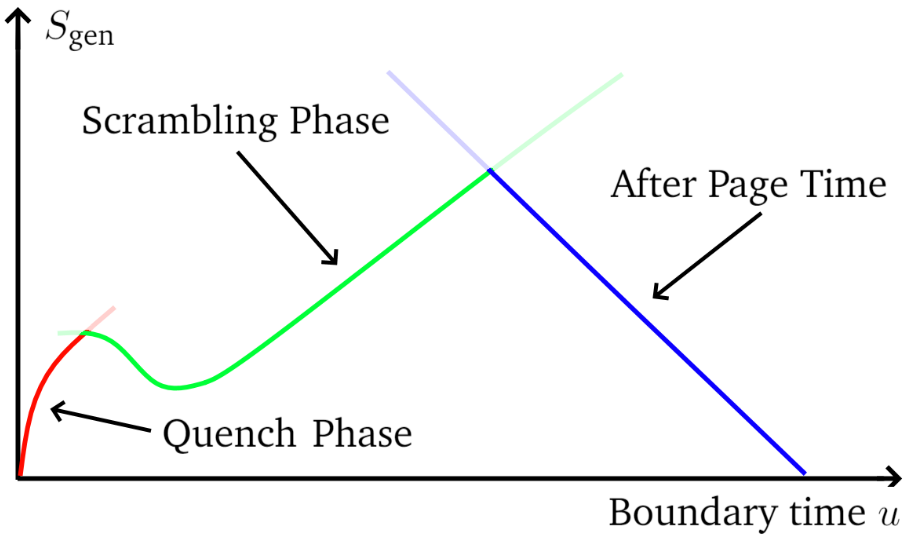

Figure 5.4: The reconstructed Page curve of the fine-grained entropy of the QM_R and QM_L+Bath over time. We assume the entropy before the quench from Eq. (5.14) can be neglected. At $u = 0$, the bath is connected to the JT gravity, which initially causes a sharp increase in entanglement entropy called the quench phase. Soon a new generalized entropy is minimal which has a QES not at the bifurcation point but close to it in the right exterior. This is the scrambling phase and lasts until the Page time. After the Page time, a new QES with minimal generalized entropy creates an island inside the interior of the new black hole. Eventually, the right black hole has evaporated and all its radiation has leaked into the bath causing a zero fine-grained entropy.

The onset of the late-time phase means that a large amount of Hawking radiation has escaped into the bath with enough information about the bulk to reconstruct most of the interior of the black hole, both the new and the original one. This was partially the case before the Page time as well during the scrambling phase, but then only a small part of the initial horizon was encoded in the Hawking radiation.

A creature on the 2d spacetime will be puzzled by why the entanglement wedge of the radiation has split into two disconnected parts after the Page time, but from the AdS_3 point of view, it is obvious why. The entanglement wedge of the black hole is simply the region encapsulated by the classical RT surface, to leading order, which stays close to the conformal boundary, but on the other hand, the entanglement wedge of QM_L + Bath must obviously fill the rest of the AdS_3 void, and so the disconnected entanglement wedges is in this theory simply explained by an additional spatial dimension which can 'stretch' the entanglement wedge into the interior of the black hole.

5.5 What Did We Learn?

As noted in Ref. [20], there are configurations where the scrambling phase never has the minimal generalized entropy, thus the quench phase will last until the late time phase. This is the result that more closely follows Ref. [6]. In that paper, the authors only considered an initial one-sided black hole that jumped from the quench phase to the late time phase after the Page time. But this approach misses a potential phase, the scrambling phase, which tells us that the bath plus QM_R contains information about a small region inside the initial black hole before the energy pulse.

We also considered dividing the system into QM_R and QM_L + Bath, but we could instead consider the entire 2d gravity region and compare it to the bath. In that case, the entanglement wedge of the bath would also 'reach' into the interior of the right side of the final black hole, but it would have no access to the left side of the black hole since no Hawking radiation from the left side would ever reach the bath. Thus, at late times, the entanglement wedge of the bath would reach into the interior of the right black hole, but would not be present in the left, disconnected side of the black hole.

Our system has a 2d black hole on a CFT boundary of a 3d AdS. Knowing that 2d CFTs are in many ways different from $d > 2$ CFTs and that JT gravity might not be representative of Einstein gravity, one might be skeptical if one can generalize our results to higher dimensions. This has already been done in Ref. [10] where the authors showed numerically that it does generalize to higher dimensions. They considered Einstein gravity and a two-sided black hole in 4 dimensions, where the matter was a CFT_4 which had a 5d AdS dual. They showed that there are two competing QES where the late-time entropy has a QES inside the event horizon, making the entanglement wedge of the radiation disconnected from the 4d perspective, but connected in the 5d dual.

Chapter 6

Discussion

Over a hundred years have passed since Karl Schwarzschild found the metric of eternal black holes in the trenches of the eastern front during the first world war. After that, it took roughly 60 years before Hawking was able to show apply both general relativity and quantum mechanics to show that black holes emit seemingly thermal radiation. Since then, progress on the matter has been exponential, especially after the tools of AdS/CFT conjecture were brought to light. A lot of progress has been made the past two decades, and it has been hard to condense everything down to a single dissertation.

Given the recent development the past year, we have in this dissertation argued that, no, black holes do not destroy information. This is mostly due to the strong evidence given by the AdS/CFT conjecture which trivially gives us the 'no' answer, but at the same time, it hides the answer to 'why?' too well. The development took off after Ryu and Takayanagi were able to find the first-order equation for boundary entanglement entropy, which is astonishingly similar to Bekenstein's black hole entropy. This was later generalized to include quantum effects, which made one able to apply it successfully to evaporating black holes. It was then shown to satisfy the unitarity required by quantum mechanics, and one was able to recover the Page curve. This is the fine-grained entropy curve of Hawking radiation where unitarity tells us that given the entire system being pure, the entropy should reach zero once the black hole has evaporated. The secret is that, in order to get this result after the radiation becomes maximally entangled with the black hole, the gravity is telling us to look at the interior of the black hole. To successfully describe the radiation region, we must include most of the interior in the entanglement wedge of the radiation, which violates a great deal of intuition. We have arrived at this result using nothing but the effective gravity theory. The fact that the gravitational path integral can be used to arrive at these results without us knowing what the real theory of quantum gravity is, means we don't have a good enough understanding of the gravitational path integral. The next step for researchers is, therefore, to understand it better and see what more the effective field theory knows.

If somehow the results we have presented here are proven wrong in the future, it is hard to see where the researchers might have gone wrong. Assuming that they haven't, this gives us a great deal of insight into how to unify quantum mechanics and general relativity. One of these insights could be something in the spirit of the ER=EPR conjecture, like the double-holography system considered, namely that there is an invisible but spatial connection between entangled particles.

We have left plenty out of the discussion. We mostly talked about the fine-grained entropy of a black hole and not how to retrieve information thrown into a black hole. A large part of the information paradox is figuring out how the information is encoded in the black hole, and the progress of this has been parallel to the development of entanglement wedges and entropy of black holes. For example, it was early known that unitary toy models of black holes yielded the so-called Hayden-Preskill decoding criterion [28]. Namely that small diaries, that don't contribute significantly to either entropy or backreaction, thrown into a black hole before the Page time can in principle be reconstructed at the Page time. While if they are thrown in after this time, you only need to wait a scrambling time before being able to reconstruct it. But the last couple of years, this has been verified to hold true for more realistic black holes [39, 6]. The development of information reconstruction is mostly due to the development of the fine-grained entropy of black holes and radiation. How the information is encoded in the black hole is still a mystery, but simpler examples such as 2d JT gravity, which are mathematically easier to solve, can maybe give us a clue. After the double holography paper was released, the same system has very recently been further studied to understand the reconstruction of the black hole interior [20, 21].

Our understanding of the entanglement between black holes and its Hawking radiation has expanded the past couple of years. It is a fast-moving field, and many talented string theorists seem to be moving to this new field. With so many eyes on the problem, and with the recent tools inspired by the AdS/CFT correspondence, one can only hope that the black hole information paradox will be solved once and for all in the not too distant future.

Bibliography

[1] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz. Large n field theories, string theory and gravity. *arXiv:hep-th/9905111*, 1999.

[2] C. Akers and G. Penington. Leading order corrections to the quantum extremal surface prescription. *arXiv:2008.03319 [hep-th]*, 2020.

[3] C. Akers, N. Engelhardt, G. Penington, and M. Usatyuk. Quantum maximin surfaces. *arXiv:1912.02799 [hep-th]*, 2020.

[4] A. Almheiri and J. Polchinski. Models of ads2 backreaction and holography. *arXiv:1402.6334 [hep-th]*, 2015.

[5] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully. Black holes: complementarity or firewalls? *JHEP*, 2013(62), 2013.

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

[7] A. Almheiri, R. Mahajan, J. Maldacena, and Y. Zhao. The page curve of hawking radiation from semiclassical geometry. *arXiv:1908.10996 [hep-th]*, 2019.

[8] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini. The entropy of hawking radiation. *arXiv:2006.06872 [hep-th]*, 2020.

[9] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini. Replica wormholes and the entropy of hawking radiation. *arXiv:1911.12333 [hep-th]*, 2020.

[10] A. Almheiri, R. Mahajan, and J. E. Santos. Entanglement islands in higher dimensions. *SciPost Phys.*, 9(001), 2020.

[11] M. Bañados, C. Teitelboim, and J. Zanelli. The black hole in three-dimensional space-time. *Phys. Rev. Lett.*, 69(13):1849–51, 1992.

[12] I. Bengtsson. Anti-de sitter space: Lecture notes. 1998.

[13] R. Bousso. A covariant entropy conjecture. *JHEP*, 1999(07), 1999.

[14] J. D. Brown and M. Henneaux. Central charges in the canonical realization of asymptotic symmetries. *Math. Phys.*, 104(207), 1986.

[15] P. Calabrese and J. L. Cardy. Entanglement entropy and quantum field theory. *J. Stat. Mech*, 0406(06002), 2004.

[16] J. Cardy. Boundary conditions, fusion rules and the verlinde formula. *Nucl. Phys. B*, 324(581–596), 1989.

[17] J. L. Cardy, O. A. Castro-Alvaredo, and B. Doyon. Form factors of branch-point twist fields in quantum integrable models and entanglement entropy. *J. Stat. Phys*, 130:129–168, 2008.

[18] S. Carroll. *Spacetime and Geometry, An Introduction to General Relativity*. Addison Wesley, 2004.

[19] C. F. Chen, G. Penington, and G. Salton. Entanglement wedge reconstruction using the petz map. *JHEP*, 2020(168), 2020.

[20] H. Z. Chen, Z. Fisher, J. Hernandez, R. C. Myersa, and S.-M. Ruan. Information flow in black hole evaporation. *arXiv:1911.03402 [hep-th]*, 2020.

[21] H. Z. Chen, Z. Fisher, J. Hernandez, R. C. Myersa, and S.-M. Ruan. Evaporating black holes coupled to a thermal bath. *arXiv:2007.11658 [hep-th]*, 2020.

[22] R. Emparan. Black hole entropy as entanglement entropy: a holographic derivation. *arXiv:hep-th/0603081*, 2006.

[23] N. Engelhardt and A. C. Wall. Quantum extremal surfaces: Holographic entanglement entropy beyond the classical regime. *JHEP*, 2015(73), 2015.

[24] J. Engelsöy, T. Mertens, and H. Verlinde. An investigation of ads2 backreaction and holography. *arXiv:1606.03438 [hep-th]*, 2016.

[25] T. Faulkner, A. Lewkowycz, and J. Maldacena. Quantum corrections to holographic entanglement entropy. *JHEP*, 2013(74), 2013.

[26] D. Harlow. Jerusalem lectures on black holes and quantum information. *Rev. Mod. Phys.*, 88(015002), 2016.

[27] S. Hawking and D. N. Page. Thermodynamics of black holes in anti-de sitter space. *Commun. Math. Phys.*, 87(577), 1983.

[28] P. Hayden and J. Preskill. Black holes as mirrors: quantum information in random subsystems. *JHEP*, 2007(09), 2007.

[29] V. Hubeny, M. Rangamani, and T. Takayanagi. A covariant holographic entanglement entropy proposal. *J. High Energy Phys.*, 07(062), 2007.

[30] M. Kachelriess. *Quantum Fields: From the Hubble to the Planck Scale*. Oxford University Press, 2017.

[31] A. Lewkowycz and J. Maldacena. Generalized gravitational entropy. *JHEP*, 2013(90), 2013.

[32] J. Maldacena. The large n limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, 2(4):231–252, 1998.

[33] J. Maldacena. The gauge/gravity duality. *arXiv:1106.6073v2 [hep-th]*, 2014.

[34] J. Maldacena and L. Susskind. Cool horizons for entangled black holes. *Fortsch. Phys.*, 61:781–811, 2013.

[35] H. Nastase. *Introduction to the AdS/CFT correspondence*. Cambridge University Press, 2015.

[36] T. Nishioka, S. Ryu, and T. Takayanagi. Holographic entanglement entropy: An overview. *arXiv:0905.0932 [hep-th]*, 2009.

[37] D. N. Page. Information in black hole radiation. *Phys. Rev. Lett.*, 71(3743), 1993.

[38] D. N. Page. Time dependence of hawking radiation entropy. *JCAP*, 1309(028), 2013.

[39] G. Penington. Entanglement wedge reconstruction and the information paradox. *arXiv:1905.08255 [hep-th]*, 2019.

[40] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang. Replica wormholes and the black hole interior. *arXiv:1911.11977 [hep-th]*, 2020.

[41] A. V. Ramallo. Introduction to the ads/cft correspondence. *arXiv:1310.4319 [hep-th]*, 2013.

[42] L. Randall and R. Sundrum. An alternative to compactification. *Phys. Rev. Lett.*, 83(4690), 1999.

[43] M. Rangamania and T. Takayanagi. Holographic entanglement entropy. *arXiv:1609.01287 [hep-th]*, 2016.

[44] S. Ryu and T. Takayanagi. Holographic derivation of entanglement entropy from ads/cft. *Phys. Rev. Lett.*, 96(181602), 2006.

[45] M. Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71(666), 1993.

[46] T. Strobl. Gravity in two spacetime dimensions. *arXiv:hep-th/0011240*, 2000.

[47] A. Strominger and C. Vafa. Microscopic origin of the bekenstein-hawking entropy. *arXiv:hep-th/9601029*, 1996.

[48] G. 't Hooft. Dimensional reduction in quantum gravity. *Conf. Proc. C*, 930308: 284–296, 1993.

[49] T. Takayanagi. Holographic dual of bcft. *Phys. Rev. Lett.*, 107(101602), 2011.

[50] R. C. Tolman. On the weight of heat and thermal equilibrium in general relativity. *Phys. Rev.*, 35:904–924, 1930.

[51] D. Tong. Lectures on string theory. *arXiv:0908.0333 [hep-th]*, 2012.

[52] M. Visser. Physical observability of horizons. *arXiv:1407.7295 [gr-qc]*, 2014.

[53] T. Wiseman. An introduction to ads-cft. 2020.