VRIJE UNIVERSITEIT BRUSSEL

ISLANDS AND THE BLACK HOLE INFORMATION PARADOX

Emilie Grégoire

June 2021

Promotor: prof. dr. Ben Craps
Co-Promotor: dr. Surbhi Khetrapal
Supervisor: Philip Hacker

Faculty of Sciences and Bioengineering Sciences

ii

This master's thesis came about during the period in which higher education was subjected to a lockdown and protective measures to prevent the spread of the COVID-19 virus. The process of formatting, data collection, the research method and/or other scientific work the thesis involved could therefore not always be carried out in the usual manner. The reader should bear this context in mind when reading this master's thesis, and also in the event that some conclusions are taken on board.

Abstract

This thesis aims to provide a self-contained review of the black hole information paradox and recent developments concerning a partial resolution of the information problem. We start by introducing all the necessary semi-classical tools to formulate Hawking's original conclusions as a theorem. This should convince the reader that the assumptions and computations of Hawking indeed yield the conclusion that information is lost during the black hole evaporation process. Inspired by Page, the corresponding amount of information loss is then quantified by the entanglement entropy. This quantification highlights the partial information problem which is addressed in the second part of the thesis. Namely, we discuss recent developments in the derivation of an explicit expression for the fine-grained entropy of the Hawking radiation. Eventually, such a formula, which includes contributions from islands, could resolve the tension with unitary evolution. Although most of the available literature focuses on a derivation which relies on the holographic principle, this thesis is centred around a method including Euclidean gravitational path integrals and the replica trick. In order to understand the corresponding computations, all the necessary theories and formalisms are introduced. Thereafter, we use the acquired knowledge to formulate a version of the information paradox in eternal black holes. This problem is simpler but conceptually similar to the paradox encountered for black holes formed by gravitational collapse. To begin with, we apply the general formula of the fine-grained entropy to this version of the information paradox. After that, we will derive this entropy formula, without holography, with the use of the gravitational path integral. To see that this approach yields the same answers as obtained by holographic computations, we consider non-trivial saddles that arise in the replica method, referred to as replica wormholes.

Acknowledgements

First of all, I would like to thank professor Ben Craps for giving me the opportunity to study this very exciting subject. Based on only a few discussions, professor Craps was able to find a subject which corresponded perfectly to what I was looking for: conceptually challenging and very relevant in the context of quantum gravity. Of course, I also want to thank Surbhi and Philip for their support during the year. Not only did they provide the relevant references in literature when I had not clue where to look anymore, they also had the patience to deal with my large amount of questions every week. Thank you for all your patience, help and feedback.

I would also like to address several people which have contributed to this thesis in a way that is not necessarily related to physics. First of all, I would like to thank my parents, which have always believed in me, especially when I did not. Even though they have probably no idea what exactly I have been studying the last (5) year(s), they still supported me in any way they could. I am also very grateful to Mathilde, who did not want me to thank her in this thesis. Not only did she remind me to sleep and relax from time to time, she also allowed me to complain about difficulties for hours without running away. Someone who has played an enormous role during this 5-year long journey is Stijn. I think it is fair to say that we helped each other to learn so much about life that we could write another thesis about it. Therefore, I would like to thank him for being the most amazing friend one can have. Of course, I also want to thank all my other classmates, who have made these five years very enjoyable, despite the stress-full periods and insecurities. I really could not have imagined a better group to study physics with. Especially, I want to thank Aäron for all these moments where we could not stop laughing, mostly during totally inappropriate moments, and for all his incredible help during the years. Next, I also want to highlight that I will stop following Yarno, who has been my classmate for 11 years straight and never stopped talking a weird kind of French to me. Last but not least, Viktor, Floriano, Nordin, Max, Mitja, Roel and Sven, I am really thankful for all the help, support and friendship throughout these years.

Contents

Introduction

In the beginning of the 21st century, physics has been driven simultaneously towards the domain of the incredibly small *and* the exceptionally heavy in the universe [1]. On one side, quantum mechanics is able to describe the behaviour of matter and light in all details, especially, on an atomic scale[2]. While the original formalism, introduced in the 1920's by Werner Heisenberg, Paul Dirac and Erwin Schrödinger, mostly considers individual particles, more is needed to capture their interactions. Namely, one needs to combine quantum mechanics with the special theory of relativity, described by Albert Einstein in 1905 [3]. A framework which is able to do this is quantum field theory, which ultimately led to the development of the Standard Model during the second part of the 21st century. This paradigm of quantum field theory is able to describe the electromagnetic, strong and weak forces and all their (known) carrier particles.

Unfortunately, the most familiar force in our everyday lives, gravity, is not part of the Standard Model [4]. Namely, this force, which is particularly important when considering massive objects, is described by a totally different theory. This has been introduced by Albert Einstein in 1915 and is the very well known general theory of relativity. In contrast to Newton, Einstein's theory describes gravity in terms of the curvature of geometry, which is dynamically related to the matter content in the universe [5]. Furthermore, general relativity has an extremely wide range of implications, from the existence of black holes and gravitational waves up to gravitational lensing and cosmological models [6]. Nevertheless, it also leaves fundamental questions unanswered, one of them being how to reconcile gravity with the other three fundamental forces. Namely, when trying to quantize the gravitational theory with the same technique as used to e.g. transform classical electrodynamics into quantum electrodynamics, one encounters problems at high energies. More specifically, gravity turns out to be non-renormalizable. This observation led to a lot of pessimism about the possibility of finding a fundamental quantum field theory of gravity [7]. As a result, one often thinks about general relativity as an effective field theory, which only gives an accurate description of gravitational physics at low energies. This point of view allows us to define a regime in which one can treat a system in a semi-classical way i.e. describing matter with quantum mechanics while keeping gravity classical. Unfortunately, one can not use this approximation when quantum gravitational effects are expected to become important. An example of such a situation is when one tries to understand the earliest moments in the universe. Namely, general relativity breaks down in the extremely dense and hot regime an instant after the Big Bang. This, amongst other reasons, illustrates why finding a fundamental theory of quantum gravity remains an important mission for many physicists.

An interesting way of understanding quantum gravitational effects in the very early universe is by considering simpler problems. An example of such are black holes. Although these also contain a singularity and yield many complexities, the fact that we can study them from the outside simplifies our approach significantly. Namely, far from the black hole, one can ignore gravitational effects. As a result, one can probe the black hole from far away and still study some of its interesting features [8]. It is often believed that the understanding of these features may guide a new approach to the foundations of physics.

With these ideas in mind, this thesis aims to provide a self-contained guidance towards the understanding of a problem one encounters when studying black holes in the semi-classical approximation. This problem first appeared in the early 1970's, when Hawking and his colleagues studied the thermal behaviour of black holes. Remarkably, they found that black holes behave as objects with a certain temperature which emit, what is now called, Hawking radiation [9]. Furthermore, it was shown that they have an entropy proportional to the area of the horizon. From these observations, it suddenly seemed as if black holes could be viewed as ordinary quantum systems from the outside. Not only are these observations in contradiction with predictions from general relativity, e.g. the no-hair theorem,

they also result in a paradox. Namely, as the black holes radiate, they will lose mass and, presumably, eventually evaporate away. An important question which then arises is about the information that was initially trapped in the black hole [9]. From the unitary time evolution of quantum mechanics, one would expect such information to come out in a certain manner [10]. However, based on his calculations, Hawking concluded that such information is permanently lost from the universe when the black hole has completely evaporated [9]. As this contradicts important assumptions in quantum mechanics, and thus the semi-classical approximation from which his computations resulted, this conclusion is often referred to as *the information paradox*.

Although many physicists initially agreed with Hawking's very remarkable conclusion, Leonard Susskind and Don Page did not. They claimed that, with enough effort, one should always be able to reconstruct the information. An example which is often used to understand the intuition behind the information problem is the following. If one would put a match to the printed version of this thesis, one should be able to reconstruct all its content by studying the ashes, the smoke and photons. Among others, Susskind believed that this reasoning should also be applicable to black hole evaporation, although more complicated. This profound disagreement between Susskind and Hawking ultimately resulted in *the black hole war* which is described in Susskind's book [1].

In 1992, Don Page, who was a graduate student of Hawking in the seventies, was able to quantify the information loss in terms of entropies. Namely, he noted that *the entanglement entropy* of the radiation should follow the *Page curve* in order to be consistent with a unitary process. By doing so, he illustrated that Hawking's computations already yield non-unitary behaviour way before the black hole has evaporated completely. As a result, Hawking's approach seems to break down way before quantum gravitational effects are expected to appear. This enforced the idea that something went wrong with the original computations and not necessarily with quantum mechanics itself. Namely, Page illustrated that if one could find a way to compute an entanglement entropy which follows the Page curve, one may conclude that information *does* come out of the black hole [11].

Even though Page partially transformed the information paradox into a more concrete problem, it took physicists around decades to obtain more insight. Namely, in 2006, Ryu and Takayanagi proposed a holographic derivation of the entanglement entropy in quantum (conformal) theories. Specifically, they stated that the entanglement entropy in a $d + 1$ dimensional conformal field theory could be obtained from the area of a d-dimensional minimal surface in AdS_{d+2} [12]. This formula was proposed from the AdS/CFT duality, which was first introduced by Juan Maldacena [13]. Originally, this yielded a formula which was only applicable to static Anti-de-Sitter spacetimes [8], but it has been generalized to dynamical spacetimes in [14] and thereafter, quantum mechanical contributions were taken into account [16]. Later, a similar prescription was obtained for the entropy of the radiation, which is where the information problem arises [15]. This formula was derived and applied in a holographic theory. Surprisingly, this formula allows the inclusion of interior regions of the black hole, called *islands*. As argued in [26], this contribution leads to an entropy which follows the Page curve and is therefore consistent with unitarity. However, this was achieved by assuming the AdS/CFT duality, which is not manifest. Therefore, the next step will be to obtain the same result in a more general way [11].

In this thesis, we focus on a derivation of this formula which results from the Euclidean gravitational path integral, without the use of holography. This method mainly relies on the use of the so called *replica trick* which results in non-trivial saddles, called *replica wormholes*. Considering these contributions, one is able to find an entanglement entropy of the radiation which *is* consistent with unitarity, at least in a simplified setup. Due to the complexity of the method, we only discuss this for a black hole in AdS_2 that is glued to flat Minkowski spacetime. However, this is believed to serve as a good approximation to magnetically charged near extremal four-dimensional black holes[15].

Importantly, it must be noted that providing a correct entanglement entropy which follows the Page curve is only a part of the black hole information problem. Ultimately, one should be able to describe *how* information, that was thrown into the black hole, ends up escaping in the Hawking radiation. This can be addressed by using *entanglement wedge reconstruction* [13]. However, this will not be discussed in this thesis.

The purpose of this thesis is to provide a self contained review to the core ideas of the information paradox and recent developments towards a partial resolution. These chapters, should make anyone (familiar with general relativity and quantum mechanics) able to understand the paradox and the recent developments on the subject. Especially, the methodology used to derive the contributions coming from replica wormholes should become clear to the reader, without the need of going through a large amount of papers. In contrast to many lecture notes and papers about the information paradox

[e.g. [17],[18]], we do not focus on approaches including AdS/CFT but rather center the discussion on the gravitational method relying on the Euclidean path integral. Finally, note that all the necessary derivations, based on literature, and corresponding missing steps are written down in the appendices.

With the described purpose in mind, and in order to follow a historical route, the chapters are organized as follows. In chapter 1, we review some relevant classical aspects from general relativity and quantum mechanics. This enables us to formulate the framework of semi-classical physics, which is needed to discuss the process of Hawking radiation. In chapter 2, we study this process and show how this led Hawking to his original conclusion, which stated that information is lost when a black hole has evaporated completely. To highlight the paradoxical elements of this conclusion, we formulate the information problem as a theorem. Next, we discuss Page's approach to quantify the corresponding information loss, at the end of the second chapter. To do so, we first introduce the notions of fine- and coarse grained entropy. When this has been done, we formulate the so called Page-theorem, which highlights the partial problem we mainly consider in the second part of this thesis. A resolution of this problem consists of finding a fine-grained entropy for the Hawking radiation which is consistent with unitarity. To understand how one can find such an entropy with the help of gravitational path integrals, we introduce the necessary ingredients in chapter 3. Explicitly, we are going to introduce Euclidean path integrals, conformal field theories and JT gravity in two dimensions. These formalisms will be used to describe an AdS_2 black hole glued to flat spacetime in chapter 4. Namely, by using the connection between the Euclidean path integral and the thermal partition function, one is able to find an explicit expression for the gravitational fine-grained entropy. The obtained results are compared to the result derived from holography and it is shown that both approaches yield consistent answers, at least in the simplified setup we consider. To finish with, we discuss the implication of these results on the information paradox.

Chapter 1

Black holes and quantum mechanics

The Nobel Prize in Physics 2020 was divided among three people who were strongly involved with the discovery of the mysterious objects which are called black holes. One half was awarded to Roger Penrose *for the discovery that black hole formation is a robust prediction of the general theory of relativity* [19]. Namely, Penrose demonstrated how, according to general relativity, black holes can form under the right conditions. Previous researchers had demonstrated that such a surface, which traps light, could only form under non-physical conditions [20]. According to general relativity, one of the main properties of black holes is that nothing, not even light, can escape them. Additionally, an interesting theorem about black holes states that all solutions in general relativity should be fully characterized by three externally observable parameters. Namely, the *no-hair theorem* claims that these parameters are the mass, the momentum and the charge, while the other information "disappears" behind the event horizon [21]. However, when Bekenstein studied the dynamics of classical black holes, he started to doubt this idea. Namely, he saw that there is an astounding similarity between the laws obeyed by black holes and the classical laws of thermodynamics. From this observation, Bekenstein suggested that black holes may behave as thermal objects with a temperature and entropy. Furthermore, he postulated that this temperature is proportional to the surface gravity on the horizon, while the entropy is proportional to the area of the black hole [22]. This suggests that, if this entropy is thermodynamical and has an origin in statistical mechanics, the black hole can be viewed as an ordinary quantum state from the outside [8]. Such a state is typically characterized by a large amount of microstates for each possible set of external observable properties. Obviously, this contradicts the classical theorem which states that black holes have "no hair" and are fully characterized by only three observables. A couple of years later, Hawking studied matter around and in black holes with the use of quantum field theory. This led to the discovery that quantum mechanical effects cause black holes to create and emit particles and hence act like thermal objects [23]. Therefore, Hawking's computations confirmed the suggestions of Bekenstein and completely altered physicists view on black holes. The black hole evaporation process and its implications are the subject of the next chapter.

In this chapter, we acquire a detailed understanding of the ideas discussed above. The corresponding sections are organized as follows. In section 1.2, we start by reviewing some important notions of general relativity and classical black holes. Namely, we describe the Schwarzschild solution in detail and introduce the Penrose diagram for a physical black hole, which was formed by collapsing matter. At the end of that section, we discuss the similarities between the laws obeyed by classical black holes, and the laws of thermodynamics, studied by Bekenstein. Thereafter, we introduce an important object from quantum mechanics: the von Neumann entropy, in section 1.3. This notion will be extremely useful to describe the information loss described in chapter 2 and the rest of this thesis. Finally, we introduce an approximation which enables us to capture gravitational *and* quantum mechanical effects without knowing any details about quantum gravity theory. Namely, in section 1.4, we describe the *niceness conditions*, which must be satisfied when treating a system in the semi-classical approximation. Finally, we describe how one can describe quantum field theories on curved spacetime in this regime. Explicitly, we show that the vacuum is no longer uniquely defined and describe the *Bogoliubov transformations* which allows us to relate different definitions of empty space.

1.1 Conventions

Throughout this thesis, we will work in natural Planck units. This consists of setting the universal physical constants

$$\hbar = c = G = k_B = 1, \tag{1.1}$$

where c is the speed of light, G is the gravitational constant, k_B the Boltzmann constant and $\hbar$ is the reduced Planck constant [3].

Furthermore, we work with the metrics in the (-+++) signature, unless specified differently.

1.2 Classical black holes

Before including quantum mechanical effects around a black hole, it is useful to review the main classical properties of black holes in general relativity. First, we start by summarizing the main equations of this theory. Additionally we study some important solutions of these equations. We start with the Schwarzschild solution and then turn to Reissner-Nordstrom metric. Note that the following discussion is highly based on [5], [6] and [24].

As John Wheeler once stated, General relativity can be summarized as:

Spacetime tells matter how to move; matter tells space-time how to curve [5]. Namely, the first part of this sentence is encoded in the geodesic equations

$$\frac{d^2 x^\mu}{ds^2} + \Gamma^\mu_{\sigma\rho} \frac{dx^\sigma}{ds} \frac{dx^\rho}{ds} = 0, \tag{1.2}$$

where s is a scalar parameter of motion (for example proper time τ). Additionally, $\Gamma^\mu_{\sigma\rho}$ are Christoffel symbols, which can be written in terms of the metric components, as

$$\Gamma^\mu_{\rho\sigma} = \frac{1}{2} g^{\mu\nu} \left(\partial_\sigma g_{\nu\rho} + \partial_\rho g_{\nu\sigma} - \partial_\nu g_{\sigma\rho} \right) \qquad \text{with} \qquad ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu. \tag{1.3}$$

In contrast, the equations which connect the matter in the universe to the curvature of spacetime are the Einstein field equations (EFE)

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}, \tag{1.4}$$

where $G_{\mu\nu}$ is the Einstein tensor, $T_{\mu\nu}$ the stress-energy tensor which represents the matter content of the universe, κ the Einstein gravitational constant and Λ a cosmological constant. The Einstein tensor is defined as

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

where $R_{\mu\nu}$ is the Ricci curvature tensor and R the scalar curvature, which are defined as

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma} \qquad R_{\mu\nu} = R^\lambda_{\mu\lambda\nu} \qquad R = g^{\sigma\nu} R_{\sigma\nu}. \tag{1.5}$$

Equations (1.2) and (1.4) form the core of the mathematical formulation of general relativity and will be used frequently in our discussion. Interestingly, the Ricci scalar, or the *intrinsic* curvature, R tells us how global spacetime is curved and fundamentally has to do with the effect of parallel transportation on the manifold.

As one can expect from equation (1.4), finding a solution to the EFE is everything but simple. Fortunately, it helps to consider systems with a high amount of symmetry. We will start by looking at the *Schwarzschild* solution, which is the unique spherically symmetric solution[1] of the vacuum Einstein equations according to Birkhoff's theorem [5].

---

[1]A spacetime is said to be spherically symmetric if its isometry group contains an $SO(3)$ subgroup, whose orbits are 2-spheres [5]

1.2.1 The Schwarzschild metric

The unique spherically symmetric solution to the vacuum Einstein equations, i.e. in the absence of matter $T_{\mu\nu} = 0$ and without a cosmological constant $\Lambda = 0$, is the *Schwarzschild* solution. In Schwarzschild coordinates this solution is described by the metric

$$ds^2 = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin\theta^2 d\phi^2, \tag{1.6}$$

where $t \in (-\infty, \infty)$, M is interpreted as the mass of the gravitating object [6] and (θ, ϕ) spherical coordinates.

Equation (1.6) defines the Schwarzschild metric and describes the gravitational field outside a spherical mass (on the assumption that the electric charge and angular momentum related to the mass M are zero). Some of these spherically symmetric sources are stars and planets, which make the Schwarzschild geometry suitable to model the motion of planets in orbit around the sun [5]. Additionally, note that the Schwarzschild solution is also invariant under $t \to t + c$ and hence time independent.

By looking at (1.6), one sees that the metric coefficients blow up at $r = 2M$ and $r = 0$. In order to understand the Schwarzschild geometry, it is thus important to learn about the behaviour at these points. Although the first singularity, at $r = 2M$, is an artefact of a bad choice of coordinates, interesting things happen near this radius. We will discuss this *coordinate singularity* in more detail below. In contrast, the singularity at $r = 0$, is a *physical singularity*, as one can see by looking at the behaviour of a coordinate independent quantity at this radius. An example is the *Kretschmann invariant I*, which is clearly not well behaved at $r = 0$

$$R^{\mu\nu\rho\sigma} R_{\mu\nu\rho\sigma} = \frac{48M^2}{r^6}.$$

What exactly happens at such singularities in physics is a subject of many books an research papers [6]. However, we will only investigate the causal structure of this spacetime. As this is defined by the light cones, we consider radial null curves

$$ds^2 = 0 = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 \qquad \text{with} \qquad \phi = \theta = \text{constant}$$

from which we obtain

$$\frac{dt}{dr} = \pm\Big(1 - \frac{2M}{r}\Big)^{-1}. \tag{1.7}$$

This quantity measures the slope of the light cones on the $t - r$ plane of a space diagram. By taking the limit $r \to \infty$, one sees that the slope $\frac{dt}{dr} = \pm 1$ which is equal to the situation in flat spacetime. In contrast, when approaching $r \sim 2M$ one obtains $\frac{dt}{dr} \to \pm\infty$. This means that the light cones 'close up'. As a consequence, a light ray which approaches this radius never seems to get there. This is represented on Figure 1.1. As already anticipated by nothing that $r = 2M$ is a coordinate singularity, it turns out that this observation is an illusion. Although a light ray has no trouble reaching this point, an observer far away could never see this happening. In order to understand what really happens, it is appropriate to change coordinates. As the actual problem is that (1.7) diverges at $r \to 2M$, we look for a coordinate system where this does not happen. A convenient choice, which will make the slope run slower, is the *tortoise coordinate* r^*, defined as

$$r^* = r + 2M \ln\Big|\frac{r}{2M} - 1\Big|.$$

Furthermore, solving (1.7) explicitly, one finds

$$t = \pm r^* + \text{constant}. \tag{1.8}$$

The Schwarzschild metric in these coordinates is

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

where r is now a function of r^*. While the metric coefficients are now well behaved at the coordinate singularity, the surface of interest has been pushed to infinity [6]. Therefore, it is convenient to switch over to *light-cone* coordinates

$$u = t - r^* \qquad v = t + r^*, \tag{1.9}$$
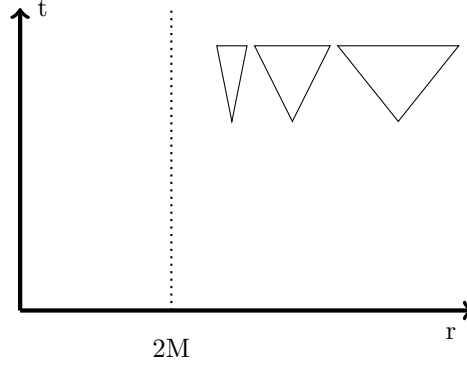
Figure 1.1: A representation of the light cones which seem to close up near the radius $r = 2M$. As a result, light signals which are sent out near this region seem to never reach $r = 2M$. Note that his Figure is highly inspired by Figure 5.7 in [6].
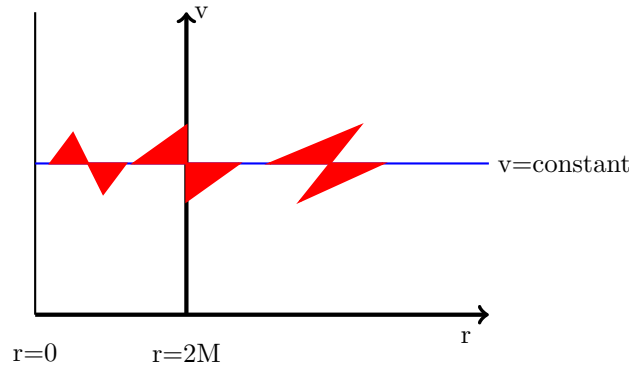
Figure 1.2: The light cones in (v, r) coordinates tilt over towards the physical singularity at $r = 0$ within the event horizon. This Figure is highly inspired by Figure 5.10 in [6]. Note that they should not change in size.

such that v =constant defines ingoing radial null geodesics while u =constant defines outgoing radial null geodesics. If one now considers the original coordinate r but replace the time coordinate t by u, one obtains the metric [6]

$$ds^2 = -\Big(1 - \frac{2M}{r}\Big)dv^2 + (dvdr + drdv) + r^2 d\Omega^2. \tag{1.10}$$

This choice of coordinates is referred to as *ingoing Eddington-Finkelstein coordinates*. This eliminates the coordinate singularity at $r = 2M$, while keeping it at finite distance. We can use these to analyse the slope of the light cones as we did in the original coordinates. This yields [6],

$$
\begin{aligned}
\text{infalling radial null geodesics:} \quad & \frac{dv}{dr} = 0 \\
\text{outgoing radial null geodesics:} \quad & \frac{dv}{dr} = 2(1 - \frac{2M}{r})^{-1}
\end{aligned} \tag{1.11}
$$

As a result, the light-cones are well behaved at $r = 2M$. However, it seems like they tilt over towards the physical singularity at $r = 0$, as shown on Figure 1.2. As a result, the surface defined at $r = 2M$ acts like a point of no return. This particularly means that no event at $r \leq 2M$ can influence another event at larger radii. Surfaces of this kind are called *event horizons*. Furthermore, a region which is separated from infinity by such a horizon is called a *black hole*, which is the central object in this thesis. Their name clearly captures the fact that, at least classically, nothing (not even light) can escape the event horizon.

In the following sections, we will see that this notion changed dramatically when Hawking discovered that quantum black holes create and emit quantum particles. Before we move on to that, it must

be noted that the Eddington-Finkelstein coordinates do not cover the whole manifold. Namely this can be extended to other directions with the use of *Kruskal coordinates*. To understand how, we will first discuss *Rindler coordinates*.

Rindler Patch

For later purposes, it will be interesting to discuss the so called Rindler coordinates. To do so, we closely follow the approach in [24].

This discussion starts in 2-dimensional Minkowski spacetime with the standard metric

$$ds^2 = -dT^2 + dX^2. \tag{1.12}$$

With the help of the relativistic velocity-addition formula

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}, \tag{1.13}$$

one can see that an uniformly accelerating observer follows a trajectory $X^\mu(\tau)$ along the X-direction, is given by

$$X(\tau) = \frac{1}{a}\cosh a\tau \qquad T(\tau) = \frac{1}{a}\sinh a\tau, \tag{1.14}$$

where a corresponds to the constant acceleration. From the fact that this trajectory obeys the relation

$$X^2 - T^2 = \frac{1}{a^2} \tag{1.15}$$

it is clear that the corresponding worldline is hyperbolic.

We can introduce a coordinate system which is naturally adapted to such an uniformly accelerated motion. These coordinates are referred to as Rindler coordinates (η, ρ) and are defined as

$$\begin{aligned} X = \rho\cosh\eta \qquad T = \rho\sinh\eta \\ \text{where} \qquad \rho \in [0,\infty[\qquad \eta \in]-\infty,\infty[. \end{aligned} \tag{1.16}$$

Note that these are related to the constant acceleration a and the proper time τ as

$$\rho = \frac{1}{a} \qquad \eta = a\tau. \tag{1.17}$$

It is again convenient to look at the light-cone coordinates

$$U = T - X = -\rho e^{-\eta} \qquad V = T + X = \rho e^{\eta}, \tag{1.18}$$

where we used the exponential definitions from the hyperbolic functions[2]. The metric in these coordinates becomes

$$\begin{aligned} ds^2 &= -dT^2 + dX^2 = -dU\,dV \\ &= -\rho^2 d\eta^2 + d\rho^2, \end{aligned} \tag{1.19}$$

which represents only a patch of Minkowski space, referred to as the *Rindler patch*. To see which patch, it is sufficient to look at the definitions of the light-cone coordinates (1.18). Namely, due to the definitions (1.16), these only describe the region which satisfies $U \leq 0$ and $V \geq 0$. this corresponds to the region enclosed by the right "triangle" on Figure 1.3. Physically, this means that we can send signals towards the region left from the future horizon $H^+ \equiv T = X$ but not receive any signals from this region. Similarly, we can only receive signals from the region on the left horizon $H^- \equiv T = -X$ but not send anything towards it. However, this is only an artefact of our coordinate choice, as was the case for the Schwarzschild event horizon. We can solve this by flipping the signs and X and T in (1.14). By considering all the possibilities, which are written down in the corresponding patches on Figure 1.3, we can describe the full Minkowski spacetime with the use of ρ and η. Interestingly, one can extend the Schwarzschild coordinates to describe the full spacetime of a black hole, with a similar method.

---

[2]When you express $\sinh\eta$ and $\cosh\eta$ in terms of exponentials, it is very easy to see that $\cosh\eta - \sinh\eta = e^{-\eta}$.
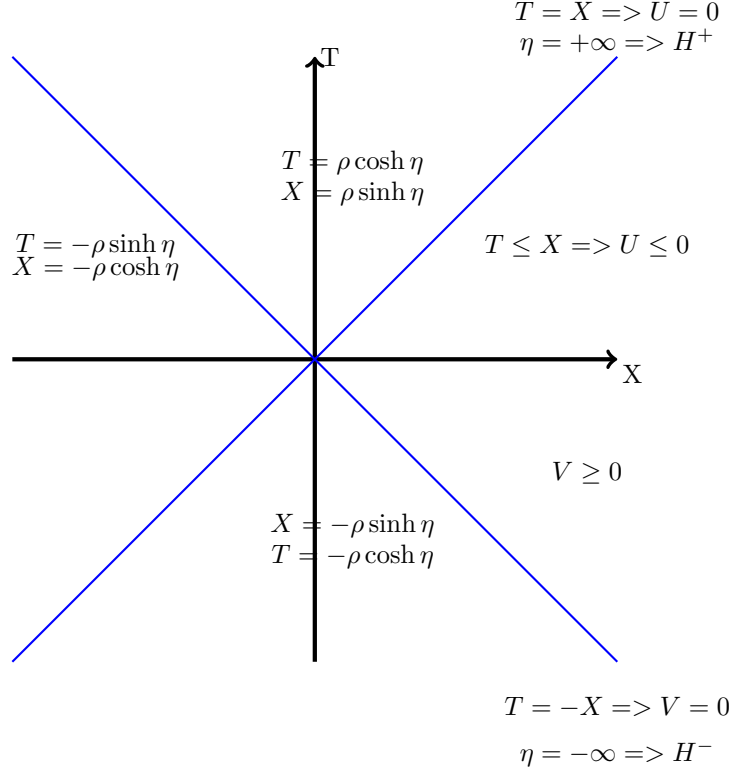
Figure 1.3: On this T-X plane, which represents the full Minkowski spacetime, we represented the future horizon H^+ and past horizon H^- by the blue lines. The right region enclosed by these two horizons is the patch covered by the Rindler coordinates defined in (1.14). Later, we will refer to this region as the *Rindler patch*. The remaining patches can be covered by switching signs in this definition. The explicit coordinates are written down in their corresponding patches. Figure inspired by diagram in section 1 of [24].

To end our discussion on Rindler coordinates, it is interesting to note that we can also define Rindler tortoise coordinates. Namely, we can define x and t as

$$\rho = \frac{1}{a}e^{ax} \qquad \eta = at. \tag{1.20}$$

These coordinates are related to the Minkowski lightcone coordinates as

$$U = -\frac{1}{a}e^{a(x-t)} \qquad V = \frac{1}{a}e^{a(x+t)}, \tag{1.21}$$

which can be seen by filling in (1.20) in (1.18). If we then define Rindler lightcone coordinates u en v in terms of these tortoise coordinates,

$$u = x - t \qquad v = x + t, \tag{1.22}$$

we find that there is a simple relationship between the Minkowski lightcone- and Rindler lightcone coordinates. Namely, we have that

$$U = -\frac{1}{a}e^{-au} \qquad V = \frac{1}{a}e^{au}. \tag{1.23}$$

Furthermore, we can see that we can use our new coordinates to conformally relate the Rindler metric (1.19) to the Minkowski metric

$$ds^2 = e^{2ax}(-dt^2 + dx^2) = e^{a(v-u)}dudv. \tag{1.24}$$

It turns out that their exists a coordinate system for which we can also do this for the radial part of Schwarzschild metric. This is not a surprise since one can show that every 2-dimensional metric can

be related to the flat metric (1.12) by a *Conformal transformation* [24]. Mathematically speaking,

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = e^{\Phi(x)}[\eta_{\mu\nu}dX^\mu dX^\nu]. \qquad (1.25)$$

In section 3.2, conformal transformations will be discussed in detail. For now, we only note that they change the scales but leave the angles invariant.

Kruskal Coordinates

We will now analyse the similarities between the Rindler patch (1.19) and the Schwarzschild metric (1.6). These similarities will then be used to motivate a coordinate choice which can extend the Schwarzschild coordinates to a system which covers the whole manifold.

Firstly, it is interesting to note that near the event horizon, a Schwarzschild black hole can be described by the Rindler metric (1.19). Starting with the metric (1.6), the proper distance from the event horizon up to $r > 2M = r_h$ is given by [24]

$$\rho = \int_{r_h}^r dr' \sqrt{g_{rr}} = \sqrt{r(r - r_h)} + r_h \sinh^{-1}\sqrt{\frac{r}{r_h} - 1}. \qquad (1.26)$$

Near the event horizon, this yields

$$\rho \sim 2\sqrt{r_h}\sqrt{r - r_h} \qquad a^\rho = a^r \frac{d\rho}{dr} \sim \frac{1}{\rho} \qquad (1.27)$$

such that, in this limit, the metric component g_{00} can be rewritten as

$$g_{00} = -(1 - \frac{r_h}{r}) \sim -(\frac{\rho}{2r_h})^2. \qquad (1.28)$$

By defining $\eta = \frac{t}{2r_h}$ and using the fact that $d\rho = \sqrt{g_{rr}}dr$, we thus find that near the horizon the metric (1.6) reduces to

$$ds^2 \approx -\rho^2 d\eta^2 + d\rho^2 + dx_\perp^2, \qquad (1.29)$$

where $x_\perp$ denote the transverse directions. As a result, the radial part simply corresponds to the Rindler metric (1.19). Interestingly, the fact that we recover the Rindler metric near the black hole is not surprising [24]. Namely, we can always define *normal coordinates* in the neighbourhood of a point on the manifold such that the corresponding metric is the flat metric [5].

As stated before, we can extend the Schwarzschild coordinates towards a system which covers all the spacetime patches in the same way as we did with the Rindler coordinates. Namely, as in (1.21), we expect that we can define lichtcone coordinates U, V which are related to the Schwarzschild coordinates (r, t) as

$$U = -\left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r-t}{4M}} \qquad V = \left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r+t}{4M}}. \qquad (1.30)$$

Additionally, as in (1.18), we can find coordinates T and R, referred to as *Kruskal coordinates*, which are related to U and V as

$$U = T - R \qquad V = T + R. \qquad (1.31)$$

By looking at (1.30) we find that these explicitly correspond to

$$R = \frac{1}{2}(V - U) = \left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r}{4M}} \cosh\left(\frac{t}{4M}\right)$$

$$T = \frac{1}{2}(V + U) = \left(\frac{r}{2M} - 1\right)^{1/2} e^{\frac{r}{4M}} \sinh\left(\frac{t}{4M}\right). \qquad (1.32)$$

These coordinates should be allowed to range over every value they can take without hitting the physical singularity $r = 0$. This region corresponds to

$$-\infty \leq R \leq \infty \qquad T^2 < R^2 + 1. \qquad (1.33)$$

With these coordinates, one can draw a spacetime diagram in the $T - R$ plane, referred to as a *Kruskal diagram*, where every point on the diagram is a two sphere [6]. To obtain this full solution one has to
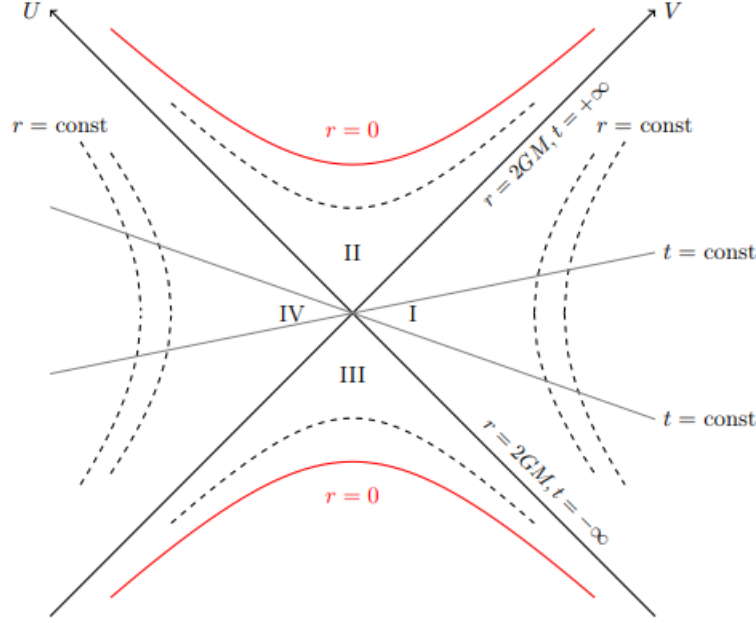
Figure 1.4: Kruskal diagram describing the maximally extended Schwarzschild solution. While region I is described by the metric (1.6) one can use the Kruskal coordinates to extend the range of U, V. Figure taken from the lecture notes [5]. Region II corresponds to the black hole, while region III corresponds to a white hole. Regions IV and I correspond to the outside regions of a spherical mass.

adapt the definitions of (1.32) for every patch, in an analogous way as in Figure (1.3). The resulting diagram is shown on Figure 1.4 and is referred to as *the eternal black hole*. This corresponds to a black hole that has always existed and contains a future and past singularity [5]. However, we do not expect this extended solution to be entirely physical.

Summarizing our approach, we have seen that the Schwarzschild metric covers only a patch of the full space-time in a same way as the Rindler metric does for Minkowski spacetime. Based on this analogy, we motivated why it is convenient to introduce lightcone coordinates. The result is that we have found an extension of the Schwarzschild solution (1.6) which is depicted on a Kruskal diagram.

It is often more interesting to constrain the Schwarzschild solution to a finite diagram by constructing its conformal diagram [6]. The corresponding diagram for Schwarzschild spacetime is depicted on Figure 1.5. This diagram also represents the full Schwarzschild solution, including the black hole, a white hole and asymptotically flat region which is connected to our universe through a wormhole. However, we do not expect all these regions to be physical. The fact that we theoretically find such features in the solution can be a result of the fact that we considered the vacuum Einstein equations and did not include matter [6].

Interestingly, one can partially use Figure 1.5 to obtain the conformal diagram to describe a physical black hole which was formed by a collapsing star. This diagram still has a singularity and a future horizon but does not contain a white hole (region II), past horizon or asymptotically regions (IV). This result depicted on Figure 1.6. Note that the green region is non-vacuum, representing the interior of the star, and thus does not correspond to a region depicted on Figure 1.5. This picture will be used throughout this thesis when we discuss particle creation near the horizon. In this subsection, we described the main aspects of the Schwarzschild spacetime which are relevant for our discussion in the remaining sections of this chapter. Furthermore, we introduced notions as the Rindler metric and uniformly accelerating observers in Minkowski. this will return in our discussion of the *Unruh effect*. Before we move on to that, we introduce another solution of the Einstein equation which will be of our interest when discussing black hole thermodynamics: the *Reissner-Nordstrom black hole*.
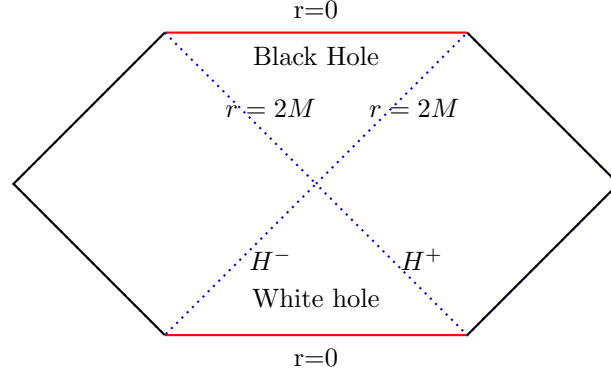
Figure 1.5: Conformal diagram for Schwarzschild solution. The red lines represent the physical singularity at $r = 0$. The blue lines represent the event horizon at $r = 2M$. Figure inspired by [6]
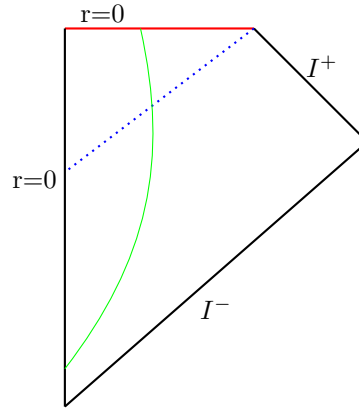
Figure 1.6: Conformal diagram for a black hole which was formed by a collapsing star. The blue dotted line represents the event horizon. The red line, represents the physical singularity. The region on the left of the green curve represents the interior of the star (i.e. a static spherical object). When the radius of the star reaches $r = 2M$, it collapses into a physical black hole and further into a singularity. When its radius is larger than $r = 2M$, the exterior of the star (white region) is described by Schwarzschild spacetime. Finally, note that I^- corresponds to the far past while I^+ corresponds to the far future. This Figure is inspired by Figure 5.17 in [6]

1.2.2   Reissner-Nordstrom black hole

This subsection is largely based on [24] and [46].

While the Schwarzschild metric is a solution of the vacuum Einstein equations, one can also find a similar solution which obeys

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu}$$
$$\nabla_\mu F^{\mu\nu} = 0,$$

(1.34)

where $T_{\mu\nu}$ is the Maxwell tensor

$$T_{\mu\nu} = \frac{1}{4\pi}\left(-\frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + F_{\mu\gamma}F_\nu^\gamma\right).$$

(1.35)

The solution of these equations is

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2d\theta^2 + r^2\sin\theta^2 d\phi^2$$
$$\text{with}\qquad f(r) = 1 - \frac{2M}{r} + \frac{Q^2}{4\pi r^2},$$

(1.36)

where M corresponds to the mass and Q to the charge. This describes a charged black hole in asymptotically flat space. The solution (1.36) is referred to as the Reissner-Nordstrom solution. Note that this corresponds to the Schwarzschild metric (1.6) where $Q = 0$. Furthermore this is again a static and spherically symmetric solution [46].

In contrast to the Schwarzschild black hole, a charged black hole has two horizons where $g_{00} = 0$:

$$r_\pm = M \pm \sqrt{M^2 - \frac{Q^2}{4\pi}}.$$

(1.37)

By looking at this expression, one can see that different situations can occur depending on the relation between the charge Q and the mass M. Namely, if $Q < 4\pi M^2$, then both horizons r_+ and r_- are real. In this case, the solution is referred to as a subextremal black hole. In contrast, if $Q^2 = 4\pi M^2$, both horizons coincide $r_+ = r_-$ and one has an *extremal* black hole with $r_\pm = M$ [24].

An interesting situation occurs when $Q > 4\pi M^2$, in that case, (1.37) becomes imaginary and the curvature singularity is not hidden behind a horizon. In that case, the black hole is said to have a naked singularity. However, it is often believed that "good" initial states do not lead to the creation of such singularities. This idea is referred to as the *cosmic censorship conjecture* [[24],[46]] but we are not going to discuss this.

Furthermore, note that the electric field generated by the charge Q can be found with the use of Gauss' law. The result is

$$E^r = \frac{Q}{4\pi r^2}$$

(1.38)

along the radial direction.

Kerr metric

Finally note that there is also a solution which represents a rotating black hole. This latter is the Kerr metric [46]

$$ds^2 = -\frac{\triangle(r)}{\rho^2}(dt - a\sin^2\theta d\phi)^2 + \frac{\rho^2}{\triangle(r)}dr^2 + \rho^2 d\theta^2 + \frac{1}{\rho^2}\sin^2\theta(adt - (r^2 + a^2)d\phi)^2$$

$$\text{with}\qquad \triangle(r) = r^2 + a^2 - 2Mr \qquad \rho^2 = r^2 + a^2\cos^2\theta$$

(1.39)

and $-M < a < M$. Again, M is the mass of the Black hole while the angular momentum is given by

$$J = aM.$$

(1.40)

Furthermore, this metric is generalized to the Kerr-Newman metric, which describes the spacetime surrounding an electrically charged, rotating mass [6]. Interestingly, it is believed that a gravitational collapse will form a physical black hole which settles down to a stationary equilibrium. The Kerr-Newman solution represents one family of black hole equilibrium states. As a result, one often ignores the gravitational collapse phase and describe a black hole fully by one of these equilibrium solutions [34].

The no-hair theorem

We will not analyse the spacetimes described by (1.36) and (1.39) as precisely as we did with Schwarzschild. However, these will be very interesting to analyse the so called *black hole laws*. Before we do so, note that classical black holes such as (1.6), (1.36) and (1.39) obey the *No-hair theorem*. this states that classical black holes have no microstates and are completely specified by the macro-variables M, J and Q [46]. Namely, if one looks at the Schwarzschild metric (1.6), one sees that this corresponds to a non-rotating, non-charged stable black hole which is simply parametrized by M. Additionally, we have that the Reissner-Nordstrom metric (1.36) is parametrized by M and Q, while the Kerr metric (1.39) depends on M and J [21].

Interestingly, in the next subsection, we are going to see that these classical black holes also obey laws which look very similar to the laws of thermodynamics. this inspired Bekenstein and Hawking to think of a black hole as a thermodynamical system with a temperature and entropy [[10],[22]]. However, this is not what one would expect from the no-hair theorem. In the remaining of this chapter, we will see that this is a result of quantum gravitational effects. Namely, it seems like black holes have many microstates in the UV completion of quantum gravity [46]. Before we are going to include quantum mechanics into our story, we start by looking at black hole thermodynamics.

1.2.3   Black hole thermodynamics

Before analysing the laws which are obeyed by classical black holes, it is useful to discuss the laws of classical thermodynamics. However, we are only going to present these very briefly. For those who are not at all familiar with thermodynamics or statistical mechanics, [27] is a very good reference. This subsection is largely based on [8],[24], [28], and [46].

The laws of thermodynamics

We start with the *zeroth law of thermodynamics* which states that *if two systems are each in thermal equilibrium with a third one, then they are also in thermal equilibrium with each other.*

This implies that the temperature throughout a system in thermal equilibrium is constant [[28],[25]]. The corresponding law can thus be summarized as

$$T = \text{constant} \qquad \text{throughout a system in thermal equilibrium.} \tag{1.41}$$

Secondly, the *first law of thermodynamics* is often thought of as the energy conservation law. Namely it states that small variations between equilibrium configurations of a system, the changes in the energy must obey the relation

$$dU = TdS + \sum_i \mu_i dN_i. \tag{1.42}$$

In this equation, U represents the internal energy, S is the entropy and μ_i are chemical potentials resulting from conserved charges [27]. These are interpreted as work terms [28].
Furthermore, the famous *second law of thermodynamics* states that, for any process, the entropy S of a closed system should obey

$$\delta S \geq 0. \tag{1.43}$$

In other words, *entropy never decreases, no matter which process is applied.*

To finish with, the *third law of thermodynamics* states that at absolute zero ($T = 0K$) the system must be in a state with the minimum possible energy. In other words, all thermodynamic processes stop at this temperature and the entropy reaches a constant value. This is also equivalent to the statement that *it is impossible to reduce the temperature of any closed system to zero temperature in a finite number of finite operations* or

$$T > 0. \tag{1.44}$$

We are now going to compare these four laws with theorems describing classical black hole interactions. Interestingly, we will see that these are formally identical to the laws of classical thermodynamics, if one identifies some of the black hole parameters with a temperature T and an entropy S.

The laws of black hole mechanics

The classical black holes that we discussed in section 1.2 also obey several laws. We will follow the same approach as above and go through every law in order to make the analogy with thermodynamics even stronger. Firstly, we introduce the notion of the *surface gravity* denoted as κ. This quantity is physically defined as the acceleration (due to gravity) near the horizon a^ρ times the redshift factor $\sqrt{-g_{00}}$

$$\kappa \equiv a^\rho \sqrt{-g_{00}}. \tag{1.45}$$

Intuitively, if a distant observer would throw an object, attached to a very long and strong massless rope, towards the black hole such that it hover near the horizon, the tension in your rope will corresponds to κM_{object}.

As an illustration, we can compute κ for the Reissner-Nordstrom black hole (1.36). Firstly, note that the proper gravitational acceleration is given by

$$\begin{aligned} a^\rho &= a^r \frac{d\rho}{dt} \\ &= \sqrt{g_{rr}} a^r \\ &= \frac{1}{\sqrt{g_{rr}}} \frac{\partial \ln \sqrt{-g_{00}}}{\partial r} \end{aligned} \tag{1.46}$$

where we used (1.27), which holds near the horizon, to obtain the second equality. Multiplying this with the redshift $\sqrt{-g_{00}}$ one obtains the surface gravity [47]

$$\kappa = \sqrt{-g_{00}} a^\rho \frac{1}{\sqrt{g_{rr}}\sqrt{-g_{00}}} \frac{1}{2} \frac{\partial g_{00}}{\partial r}. \tag{1.47}$$

Rewriting the zeroth component of the metric as

$$g_{00} = \frac{1}{r^2}(r - r_+)(r - r_-) \tag{1.48}$$

one obtains

$$\kappa = \frac{r_+ - r_-}{2r_+^2}. \tag{1.49}$$

As the solution (1.36) is static, this yields a constant κ. Interestingly, this is generally true due to the spherical symmetry of the metric [33]. This corresponds to the *zeroth black hole law*, namely *the surface gravity κ of a black hole is constant on the horizon*. This can be summarized as

$$\kappa = \text{constant}. \tag{1.50}$$

This has the same form as the zeroth law of thermodynamics (1.41) [25]. As a result, *if* we would think of the black hole as having a temperature proportional to κ, both statements are the same. By looking at the next law, we will see that this guess becomes even more probable.

The *first black hole law* can be formulated by looking at the response of a charged black hole to small changes in the black hole geometry [8]. To see this, we again work with the Reissner-Nordstrom black hole (1.36). Firstly, note that the area of the horizon (1.37) is

$$A = 4\pi r^2 = 4\pi \left(M + \sqrt{M^2 - \frac{Q^2}{4\pi}} \right)^2. \tag{1.51}$$

Consequently, a small variation of the area corresponds to

$$\delta A = 8\pi r_+ \delta r_+ = 8\pi r_+ \left(\frac{\partial r_+}{\partial M} \delta M + \frac{\partial r_+}{\partial Q} \delta Q \right). \tag{1.52}$$

Computing the derivatives, one obtains

$$
\begin{aligned}
\frac{\partial r_+}{\partial M} &= 1 + \frac{M}{\sqrt{M^2 - \frac{\pi Q}{4}}} \\
&= \frac{2r_+}{r_+ - r_-} \\
\frac{\partial r_+}{\partial Q} &= -\frac{1}{4\pi} \frac{Q}{\sqrt{M^2 - \frac{\pi Q^2}{4}}} \\
&= -\frac{Q}{4\pi} \frac{2}{r_+ - r_-},
\end{aligned}
\tag{1.53}
$$

where we used $r_+ - r_- = 2\sqrt{M^2 - \frac{Q^2}{4\pi}}$ to obtain the final forms. Plugging these results into (1.52) one obtains

$$
\delta A = \frac{16\pi r_+}{r_+ - r_-} \left[r_+ \delta M - \frac{Q}{4\pi} \delta Q \right],
\tag{1.54}
$$

which can be rewritten as

$$
\delta M = \frac{1}{4\pi} \left(\frac{r_+ - r_-}{r_+^2} \right) \delta(\frac{A}{4}) + \frac{Q}{4\pi r_+} \delta Q.
\tag{1.55}
$$

Furthermore, integrating the electric field (1.38), one sees that the potential due to the charge Q is

$$
V_Q = \frac{Q}{4\pi r_+}.
\tag{1.56}
$$

Additionally, if one identifies the second term with (1.49), one obtains that the small variation in mass can be written as

$$
\delta M = \frac{\kappa}{2\pi} \delta(\frac{A}{4}) + V_Q \delta Q.
\tag{1.57}
$$

This statement is the first black hole law. Clearly, this has a very similar form as (1.42). Again, this similarity inspires one to think that the black hole is a thermodynamical system with temperature T which is proportional to κ. Furthermore, one expects the proportionality constant to be $(2\pi)^{-1}$. Additionally, comparing (1.57) with (1.42) motivates the idea that the black hole also has an entropy. Namely, one expects that this is proportional to the area of the horizon as

$$
S_{BH} = \frac{A}{4}.
\tag{1.58}
$$

This idea looks even more attractive when looking at *the second black hole law*. Namely, this law states that *the area of the event horizon never decreases* [[24],[46]]. One can easily understand this for the Schwarzschild metric (1.6). Namely, as $r_H = 2M$ and because of the fact that no mass can escape this event horizon ($\delta M \geq 0$) one immediately has

$$
\delta A \geq 0.
\tag{1.59}
$$

This has also been proved for the Reissner-Nordstrom and Kerr black hole, for example with the help of differential geometry as in [25]. This law is formally the same as (1.43) and strengthens the speculation that the black hole has an entropy which is proportional to the area of the black hole horizon. Interestingly, this would contradict the no-hair theorem we introduced before. Namely, if the black hole has a thermodynamical entropy $S_{BH} \sim A$, which can have an origin in statistical mechanics, it must have a lot of microstates. Namely, such a thermodynamical entropy is highly related to the number of degrees of freedom of the system. Explicitly,

$$
S = k_b \ln(\Omega)
\tag{1.60}
$$

where Ω represents the number of microstates that can be formed with the degrees of freedom of this system [27]. If we let the Boltzmann entropy be proportional to the area of the black hole, one must

conclude that black holes have an enormous number of degrees of freedom. For example, if the area of Sagittarius A^* is expressed in Planck units $l_p^2 = \hbar G_N$ one calculates [3]

$$S_{sat} \approx 10^{82}.$$

This large number of degrees of freedom contradicts the no-hair theorem in classical general relativity[8].

Finally, the *the third black hole law* states that *it is impossible by any procedure, no matter how idealized, to reduce κ to zero by a finite sequence of operations* [28]. For the Reissner-Nordstrom metric this simply follows from $r_+ \geq r_-$, as one can see by looking at (1.49). As you may have suspected, this also looks very similar to the third law of thermodynamics (1.44) and reinforces our idea that κ has something to do with temperature.

To finish with, it is important to note that, at the moment, we do not have a good reason yet to call T the temperature and S the entropy of the black hole [46]. This will change when we couple general relativity to quantum fields theory by working in the semi-classical approximation [8]. this is exactly what Hawking has done to realize that black holes indeed behave as thermal objects, which radiate as blackbodies with temperature T [46]! Before we move on to that, we introduce an important concept from quantum mechanics: *the entanglement entropy*.

---

[3]This is easy to check. Consider the radius of the black hole $R_{sag} = 2.20 \times 10^7$ km. The conversion to Planck Units is $1km = 6.187 \times 10^{37} l_p$ and the area is then calculated as $\text{Area}_{sag} = 4\pi(R_{sag})^2 \sim 10^{82} l_p^2$.

1.3 Entanglement entropy

Up to now, we have been introducing important classical notions from general relativity. However, as the title of the title of this chapter suggest, we are interested in a quantum mechanical description of black holes. Specifically, we are interested in quantifying quantum information in and near a black hole. An object which is perfectly suited for that is *the entanglement entropy.* Before we introduce this , we review the definition of *pure and mixed quantum states.* Finally we give some examples which will be useful when computing the entanglement entropy of quantum states formed near the horizon of a black hole in the next section. Note that this section is largely based on [29], [18] and [30].

1.3.1 Quantum states

Before we start our discussion, it might be useful to ask ourselves what quantum information actually is. Firstly, note that one can think of classical information theory as the study and manipulation of bit strings [29], which are sequences of bits. These can represent sets or manipulate binary data [31]. In a similar way, one can think of quantum information theory as the study of *qubit strings*[4]. These encode the amount of information that is contained in a state of a quantum mechanical system. To see this, let us review the concepts of states and density matrices in quantum mechanics.

Remember that, in quantum mechanics, one is especially interested in the wave function $\Psi(x, t)$ of a particle. this is obtained by solving the Schrodinger's equation

$$i\frac{\partial \Psi}{\partial t} = -\frac{1}{2m}\frac{\partial^2 \Psi}{\partial x^2} + V\Psi,$$

which is the central equation of quantum mechanics.

In order to understand that such a wave function can be thought of describing the quantum state of the particle, it useful to remember *Born's statistical interpretation.* Namely, this states that $|\Psi(x, t)|^2 dx$ gives the probability of finding the particle between x and $x + dx$ at time t. Consequently, the wave-function has to be normalized.

Interestingly, the fundamental principles of quantum mechanics can be encoded in the language of linear algebra. Namely the state of a particle can be represented by a vector $|\Psi\rangle$, which is normalized. As this is possible only if it is square integrable, the quantum mechanical wave function should live in $L_2(-\infty, +\infty)$, which corresponds to the set of all square-integrable functions in the interval $]-\infty, +\infty[$ and is often denoted by L_2. From now on, we will just refer to this as *Hilbert space $\mathcal{H}$.* Furthermore, observables are represented by Hermitian operators $\mathcal{O}$, such that the expectation value of the observable Q in the state $|\Psi\rangle$ corresponds to

$$\langle \mathcal{O} \rangle = \langle \Psi | \mathcal{O} | \Psi \rangle. \tag{1.61}$$

Additionally, when measuring the observable, one is certain to get one of the eigenvalues λ_i of the corresponding operator $\mathcal{O}$, such that the probability of measuring a specific eigenvalue is $|\lambda|^2$ when expressed in the orthonormal basis of eigenvectors. However, this only holds because we assume that observables are only represented by operators which have eigenfunctions that constitute a complete set [32].

An operator which acts on $\mathcal{H}$, and will be of our main interest, is the *density matrix ρ.* If we denote some basis of the Hilbert space of system A as $|i\rangle$, the density matrix is often written as [30]

$$\rho = \sum_{ij} |i\rangle \rho_{ij} \langle j|. \tag{1.62}$$

Note that three essential properties of this density matrix are [29]:

- $\rho = \rho^\dagger$

- $\forall |\Psi\rangle \in \mathcal{H} : \langle \Psi | \rho | \Psi \rangle \geq 0$

- $\text{Tr } \rho = 1.$

---

[4]The difference between a qubit and a classical bit is that the state of a qubit is a linear superposition of states of a classical bit.[30]

Generally, the main use of ρ is to obtain expectation value of observables. Namely, it can be used to obtain (1.61) as [30]

$$\langle \mathcal{O} \rangle = \sum_{ij} \langle j | \mathcal{O} | i \rangle = Tr(\rho \mathcal{O}). \tag{1.63}$$

Pure and mixed states

States which can be described by a well defined single wavefunction Ψ are called *pure states*. However, not all states are pure. Namely, imagine that the system of interest is coupled to a reservoir. In that case, the full system is described in terms of the system and the reservoir. If the system and the reservoir do not interact, one can describe the corresponding state with a corresponding wavefunction $\Psi_{total}(x, x_R, t)$ which can be decomposed into a product of the two parts $\Psi_{system}(x, t)\Psi_{reservoir}(x_R, t)$. In contrast, when the system of our interest interacts with the reservoir, one can not longer separate a system wavefunction from a reservoir wavefunction. Furthermore, often we do not have access to the full state $\Psi_{total}(x, x_R, t)$. Specifically, when we deal with black holes later on, we will only have access either to the black hole or to the radiation degrees of freedom. As a result, we do not have a fully defined wavefunction in this case and our system is said to be in a *mixed state* [32]. For such states, we thus do not have all the information to compute the expectation values of observables in this state.

Interestingly, the density matrix provides a practical way of distinguishing pure from mixed states. Namely, start by noting that a pure state can be written as

$$\psi = \sum_i a_i | i \rangle , \tag{1.64}$$

such that the corresponding density matrix, generally defined as (1.62) is simply obtained by

$$\rho = \sum_{ij} | i \rangle a_i a_j^* \langle j | = | \psi \rangle \langle \psi | . \tag{1.65}$$

Furthermore, one can choose a special basis $| n \rangle$ such that the density matrix ρ is diagonal:

$$\rho = \sum_n | n \rangle \rho_{nn} \langle n | , \tag{1.66}$$

with $0 \leq \rho_{nn} \leq 1$. As a result, we can now interpret the diagonal elements as probabilities. Therefore, we set $\rho_{nn} \equiv p_n$ and interpret this as the probability that the system is in state n. By comparing the diagonal form (1.66) with (1.65), one sees that the density matrix of a pure state only has one eigenvalue $p_n = 1$ while $p_m = 0$ for all $m \neq n$. As a result, we have that $\rho^2 = \rho$ for pure states such that we can distinguish them from mixed states by comparing the traces of ρ^2 [30]

- Tr $\rho^2 = 1$ for pure states

- Tr $\rho^2 \leq 1$ for mixed states.

Bipartite systems

We will be especially interested in systems which consist of two subsystems. This is often referred to as *Bipartite systems*. Formally, the corresponding Hilbert space $\mathcal{H}_{AB}$ is defined by the direct product

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \tag{1.67}$$

The subspaces $\mathcal{H}_A$ and $\mathcal{H}_B$ are each spanned by their own basis which consists of orthonormal states denoted by $\{|i\rangle_A\}_{i=1}^{d_A}$ and $\{|j\rangle_B\}_{j=1}^{d_B}$ respectively [29]. The Hilbert space $\mathcal{H}_{AB}$ then has dimension $\dim(\mathcal{H}_{AB})=\dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$ and is spanned by the orthonormal basis $\{|i\rangle_A \otimes |j\rangle_B\}_{i=1,j=1}^{d_A,d_B}$.

Interestingly, one can expand every state $|\Psi\rangle_{AB} \in \mathcal{AB}$ as

$$|\Psi\rangle = \sum_i^{d_A} \sum_j^{d_B} c_{ij} |i\rangle_A \otimes |j\rangle_B . \tag{1.68}$$

An important object in our discussion is the partial trace. this enables us to reduce an operator, which is defined on the full Hilbert space $\mathcal{H}_{AB}$, to one acting only on a composite of the full system. Namely, imagine that we only have access to subsystem A, then we can construct a partial trace to capture the ignorance about B. To do so, note that the bra $\langle j|_B$ can actually be seen as an isometry $\langle k|_B : \mathcal{H}_{AB} \to \mathcal{H}_A$ which acts on an orthonormal basis as

$$\langle k|_B \left(|i\rangle_A \otimes |j\rangle_B\right) = \delta_{jk} |i\rangle_A .$$

With this in mind, the partial trace with respect to the reservoir B is defined as

$$Tr_B \mathcal{O}_{AB} = \sum_{k=1}^{d_B} \langle k|_B \, \mathcal{O}_{AB} \, |k\rangle_B .$$

The resulting operator, which is defined only on A is denoted as $\mathcal{O}_A$.

In order to obtain the entanglement entropy later, we use the definition of the partial trace to obtain a *reduced density matrix* ρ_A.

$$\rho_A \equiv Tr_B \rho, \tag{1.69}$$

where ρ corresponds to the density matrix of the total system AB [29]. If the corresponding state is pure, we can write ρ as in (1.65). Furthermore, we one may decompose $|\Psi\rangle$ as in (1.68) such that the reduced density is found as

$$
\begin{aligned}
\rho_A &= \sum_{k=1}^{d_B} \langle k|_B \, \rho \, |k\rangle_B \\
&= \sum_{k=1}^{d_B} \langle k|_B \left(\sum_{i,j} \sum_{l,m} \rho_{ijlm} |i\rangle_A |j\rangle_B \langle l|_A \langle m|_B \right) |k\rangle_B \\
&= \sum_{i,j} \sum_{l,m} \rho_{ijlm} |i\rangle_A \langle l|_A \left(\sum_{k=1}^{d_B} \delta_{jk} \delta_{km} \right) \\
&= \sum_{i,j} \sum_{l,m} \rho_{ijlm} |i\rangle_A \langle l|_A \, \delta_{jm} \\
&= \sum_{i,l} |i\rangle_A \langle l|_A \sum_j \rho_{ijlj}.
\end{aligned}
\tag{1.70}
$$

Interestingly, even if the complete bipartite system consisting of subsystems A and B is pure, it can be that reduced density matrices ρ_A and ρ_B are mixed. In other words, it is not because we have all the information of the full system that we have complete knowledge of the individual subsystems. In this case, we say that the subsystems A and B are *entangled* [18]. An object which is capable of measuring the amount of entanglement and thus, the lack of information, is the von Neumann entropy [29].

1.3.2 Von Neumann entropy

In section 1.2, we encountered, what we will now call, the thermodynamical entropy S. In this subsection, we will define another type of entropy, the von Neumann entropy S_{vN}, which will play an enormous role in our discussion about the information paradox. This has two main reasons. Firstly, it represents a measure for purity. Namely, the von Neumann entropy of a state vanishes when the state described by ρ is pure, which indicates full knowledge of the quantum state [29]. Additionally, S_{vN} will enable us to measure the quantum entanglement between two subsystems constituting a bipartite system [17]. For example, it is able to measure the entanglement between particles inside the black hole and the Hawking radiation.

Concretely, the von Neumann entropy of a system, with density matrix ρ, is defined as

$$S_{vN}(\rho) \equiv -Tr \rho \log \rho. \tag{1.71}$$

Interestingly, in the eigenbasis of ρ (see (1.66)), this yields [29]

$$S_{vN}(\rho) = -\sum_i p_i \log p_i. \tag{1.72}$$

Consequently, if ρ is pure, one obtains $S_{vN}(\rho) = 0$. Therefore, we may think of the von Neumann entropy as a measure of the purity of a state. Namely, if $dim(\mathcal{H}) = d < \infty$, the following statements hold

$$0 \leq S_{vN}(\rho) \leq \log d \qquad \text{maximal value obtained for} \qquad \rho = \frac{I}{d}$$
$$S_{vN}(\rho) = 0 \qquad \text{when } \rho \text{ corresponds to a pure state.} \tag{1.73}$$

Hence, this entropy measure how impure/mixed a state is or, equivalently, how much information of a state lacks. The maximal value of $S_{vN}(\rho) = \log d$ is encountered for a maximally mixed state.

Entanglement entropy

At this point, we would like to extend this notion towards bipartite systems. Based on (1.69) and (1.71), one could already expect that one can define the von Neumann entropy of a reduced density matrix as

$$S_{vN,A}(\rho) \equiv -Tr\rho_A \log \rho_A. \tag{1.74}$$

This entropy is referred to as the *entanglement entropy*. To see why let's consider a bipartite system again. It is shown in appendix A that the following statement is holds:

Lemma 1.3.1. *Consider $|\psi\rangle \in \mathcal{H}_{AB}$ and assume the corresponding full state is pure. One can show that in that case*

$$S_{vN,A} = S_{vN,B} \tag{1.75}$$

must hold. Furthermore, both entropy's are equal to zero if and only if $|\psi\rangle$ is unentangled across A and B. Otherwise stated

$$S_{vN,A} = S_{vN,B} = 0 \Leftrightarrow |\psi\rangle = |i\rangle_A \otimes |j\rangle_B \tag{1.76}$$

This property illustrates that the entropy defined in (1.74) measures the entanglement in a bipartite system. Namely, it is (non)-zero when the subsystems are (not) entangled [29]. Furthermore, the statement illustrates that even if the total state is pure, the composite states can be mixed, i.e. when they are entangled with each other.

The entanglement entropy will be used in section 2.2 to compute the amount of entanglement between the Hawking radiation and particles inside the black hole. Before we move on to that, note that some other important properties of the von Neumann entropy are [29]:

- Subadditivity:

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \tag{1.77}$$

- Invariancy under unitary time evolution:

$$S_{vN}(U\rho U^\dagger) = S_{vN}(\rho). \tag{1.78}$$

Note that the invariance under unitary time evolution implies that a pure state remains pure. We will see that this is one of the important elements of *the information paradox*.

In appendix A, one can find some examples of calculations which we shall encounter later. A full discussion about further properties can be found in [29]. In the next section, we are going to combine the knowledge about classical black holes from section 1.2 with some concepts from quantum mechanics introduced in this section.

1.4 Semi-classical physics

In section 1.2, we reviewed some notions of general relativity which describe gravity as a classical result of curvature in spacetime. Additionally, some concepts of quantum mechanics were introduced in section 1.3. Unfortunately, up to now, there does not exist a fundamental theory which is able to combine both theories. As a result, one would expect that we are currently unable to describe quantum fields in the neighbourhood of a black hole. However, from dimensional analysis, it is believed that quantum gravitational effects only become relevant at the *Planck scale* which is characterized by the *Planck length*: $l_{\text{Planck}} = \sqrt{\hbar G/c^3} \sim 10^{-33} cm$ [43]. This allows us to work in an intermediate semi-classical approximation, where matter fields are treated quantum mechanically while gravity is treated entirely classically. Explicitly, this yields working with quantum field theory on a fixed metric, as illustrated in the next chapter 1.4.2.

In this section, we sum up the conditions that define a regime for which this semi-classical limit is expected to be a good approximation. In literature, these conditions are often assumed (implicitly or explicitly), but we are going to list these *niceness conditions* explicitly. By doing so, we follow the approach in [10]. In the next chapter, these conditions will be checked explicitly for specific spacelike slices in a Schwarzschild spacetime. To finish with, we discuss how to describe quantum field theories on a curved spacetime in the semi-classical approximation. In the next chapter, this formalism will be used to show that black holes indeed evaporate.

1.4.1 The niceness conditions

To begin with, we sum up some explicit conditions which describe when one may work in the semi-classical approximation without missing important effects from quantum gravity. As explained before, the scale at which quantum gravitational effects are expected to become important is the Planck scale l_{Planck}. Therefore, the below of conditions explicitly describe the necessary relation between the geometrical properties (e.g. the intrinsic/extrinsic curvature) and this scale.

In order to apply quantum field theory on a general (curved) spacetime, the following *niceness conditions* should hold on a general spacelike slice under consideration:

(N1) the curvature R of the slice should obey $R << \frac{1}{l_p^2}$

(N2) the slice should be nicely embedded in the full spacetime; the extrinsic curvature K must obey $K << \frac{1}{l_p}$

(N3) the curvature R of the full spacetime, in the neighbourhood of the slice, should be small $R << \frac{1}{l_p^2}$

(N4) any quanta on the slice should have wavelength λ that obeys $\lambda >> l_p$.

(N5) the energy density U and the momentum density P should obey $U, P << l_p^{-4}$

(N6) the lapse and shift vectors needed to specify the evolution from one slice to the other should change smoothly with position $\frac{dN^i}{ds} << \frac{1}{l_p}$ and $\frac{dN}{ds} << \frac{1}{l_p}$.

When these conditions are satisfied, one may work in the semi-classical approximation on the slice under consideration, without missing important signatures from quantum gravity. In appendix C, it is shown that one can construct spacelike slices which obtain these conditions.

1.4.2 QFT in curved spacetime

When the niceness conditions are satisfied for the spacetime region of interest, one can actually work in the semi-classical approximation without worrying about quantum gravitational effects. Explicitly, we can treat matter as described by a quantum field theory on a curved spacetime which properties are defined by general relativity. To see how this works, we first review some general aspects from ordinary quantum field theory. Then we explain the important differences which arise when working on a curved spacetime. Especially, it tuns out that vacuum is not longer uniquely defined in that case. This is a result of the global covariancy of general relativity, as discussed below.

QFT on a flat spacetime

Before the case of a curved spacetime is considered, some important concepts of the canonical quantization formalism are reviewed. A full introduction to quantum field theory on a flat spacetime can be found for example in [3],[37] and [38].

Consider a massless scalar field ϕ in 2- dimensional Minkowski spacetime with metric (1.12). The corresponding action is

$$I = \int dX \, dT \left(\frac{1}{2} \partial_\mu \phi \partial^\mu \phi \right). \tag{1.79}$$

Varying this action and setting $\delta I = 0$ yields the equation of motion

$$\Box \phi = 0, \tag{1.80}$$

where $\Box$ is the d'Alembert operator $\Box = \frac{\partial^2}{\partial t^2} - \nabla^2$ [3]. As the scalar field ϕ is a quantum operator, it must satisfy the equal-time commutation relations [28]

$$[\phi(t, x^i), \phi(t, y^i)] = \delta^3(x^i - y^i). \tag{1.81}$$

One can choose a complete set of solutions of equation (1.80), and expand the general solution in this basis as [28]

$$\phi = \sum_k \left(a_k f_k + a_k^\dagger f_k^* \right) \tag{1.82}$$

or, after a Fourier transformation [24],

$$\phi = \int_{-\infty}^{\infty} dk \left(a_k f_k(X^\mu) + a_k^\dagger f_k^*(X^\mu) \right). \tag{1.83}$$

In Minkowski spacetime, the standard choice of basis functions is the set $\{f_k, f_k^*\}$ where

$$f_k(X^\mu) = \frac{e^{-i(\omega_k T - kX)}}{\sqrt{4\pi\omega_k}} \qquad \omega_k = |k|. \tag{1.84}$$

Additionally, the commutation relations of ϕ (1.81) imply that the Fourier coefficients a_k and $a_k^\dagger$ obey the following relations

$$[a_k, a_{k'}] = 0 \qquad [a_k^\dagger, a_{k'}^\dagger] = 0 \qquad [a_k, a_k^\dagger] = \delta_{k,k'}. \tag{1.85}$$

Consequently, we interpret a_k and $a_k^\dagger$ as the annihilator and creation operators of particles with momentum k. The Minkowski vacuum is then defined as the state which is annihilated by all the annihilation operators a_k

$$a_k \ket{0}_M = 0 \qquad \forall k. \tag{1.86}$$

Furthermore, the standard Fock space is constructed by applying arbitrary products of creation operators

$$\ket{k_1, k_2, \ldots, k_n} = a_{k_1}^\dagger a_{k_2}^\dagger \ldots a_{k_n}^\dagger \ket{0}_M. \tag{1.87}$$

The number of particles in Minkowski spacetime can then be formally obtained by defining the number operator N_k^M as

$$N_k^M = a_k^\dagger a_k \qquad \text{such that} \qquad \bra{0}_M (a_k^\dagger)^n N_k^M (a_k)^n \ket{0}_M = n. \tag{1.88}$$

These definitions can be extended to a curved spacetime with metric $g_{\mu\nu}$. However, one must be more careful when talking about states and numbers of particles. Namely, in curved spacetime, the definition of a particle and hence of the vacuum, is observer dependent. This is a result of the fact that general relativity is generally covariant. Explicitly, this means that there are many different legitimate choices for a time coordinate. As particles (e.g. frequency modes) are defined with respect to such a coordinate, the definition of particles is generally different for every choice. For example, consider Figure 1.6, which corresponds to the Penrose diagram of a black hole formed by gravitational collapse. In the flat asymptotic region, it is convenient to work with Minkowski coordinates (T, X) while, near the horizon one preferably works with Rindler or Kruskal coordinates. When a particle, described by the wave packet (1.84), starts out in a region far from the black hole it only consists of positive frequency modes. However, when the packet propagates near the horizon and turns back, it will consist of a mixture of positive and negative frequency modes. this correspond to quantum-particle creation [28], which was discovered by Hawking in [34].

First, we define a formalism which makes it possible to obtain the transformations between two different orthonormal bases corresponding to different coordinate systems. These kind of transformations is what we call the *Bogoliubov transformations*.

Bogoliubov transformations

Consider the case where are two appropriate time coordinates are available, for example t and t', one can define in a certain patch of a general spacetime. In that case, the general solution of the equation of motion (1.80) can be expanded in terms of different wave modes. Each contains positive frequency modes defined with respect to the corresponding time coordinate. Explicitly, say that in one basis we can expand the scalar field as

$$\phi = \sum_i (a_i f_i + a_i^\dagger f_i^*), \tag{1.89}$$

where the the modes f_i are defined as in (1.84). The operators a_i and $a_i^\dagger$ then annihilate/create A-particles, corresponding to the positive frequency modes with respect to the time coordinate t.

Secondly, one can expand the field with respect to the another complete set of solutions, defined with respect to another time coordinate t' as

$$\phi = \sum_i (b_i g_i + b_i^\dagger g_i^*). \tag{1.90}$$

The coefficients b and $b^\dagger$ are the corresponding annihilation and creation operators of the new "B" particles. The corresponding positive frequency modes are defined as

$$g_{k'}(X'^\mu) = \frac{e^{-i(\omega_{k'} T' - k' X')}}{\sqrt{4\pi\omega_{k'}}} \tag{1.91}$$

in analogy to (1.84). The vacuum with respect to the new time coordinate is then

$$b_i |0\rangle_g = 0 \qquad \forall i. \tag{1.92}$$

Interestingly, due to the completeness of both bases, one can expand each mode in terms of the other as,

$$g_i = \sum_j (\alpha_{ij} f_j + \beta_{ij} f_j^*) \qquad f_i = \sum_j (\alpha_{ij}^* g_j + \beta_{ij} g_j^*). \tag{1.93}$$

The coefficients α_{ij} and β_{ij} are referred to as the *Bogoliubov coefficients*, as they define the *Bogoliubov transformations* from one bases to another. In order to obtain these coefficients, one can use the orthonormality relations of the positive frequency modes f_i and g_i

$$(f_i, f_j) = \delta_{ij} \qquad (f_i^*, f_j^*) = -\delta_{ij} \qquad (f_i, f_j^*) = -(f_i^*, f_j), \tag{1.94}$$

where the conserved inner product is defined as

$$(f, h) = -i \int d^3x \sqrt{-g} \left(f \frac{\partial h^*}{\partial t} - \frac{\partial f}{\partial t} h^* \right). \tag{1.95}$$

With the use of these relations one can show that

$$\alpha_{ij} = (g_i, f_j) \qquad \beta_{ij} = -(g_i, f_j^*) \tag{1.96}$$

as illustrated in appendix B.1.

Interestingly, these coefficients allow us to obtain the creator/annihilator operator corresponding to one basis in terms of the other

$$a_i = \sum_j (\alpha_{ij} b_j + \beta_{ij}^* b_j^\dagger)$$
$$b_i = \sum_j (\alpha_{ij}^* a_j - \beta_{ij}^* a_j^\dagger). \tag{1.97}$$

Consequently, we can obtain the number of B particles in the f-vacuum with the use of

$$N_{B,i} = b_i^\dagger b_i. \tag{1.98}$$

As explained before, this yields a non-zero number of particles

$$
\begin{aligned}
\langle 0|_f \, N_{B,k} \, |0\rangle_f &= \langle 0|_f \, b_i^\dagger b_i \, |0\rangle_f \\
(\text{using (1.97)}) \qquad &= \langle 0|_f \sum_{j,k} (\alpha_{ij} a_j^\dagger - \beta_{ij} a_{ij})(\alpha_{ik}^* a_k - \beta_{ik}^* a_k^\dagger) \, |0\rangle_f \\
(\text{using} \qquad a_j \, |0\rangle_f = 0 \qquad \langle 0|_f \, a_i^\dagger = 0) \qquad &= \sum_{j,k} \beta_{ij} \beta_{ik}^* \, \langle 0|_f \, a_j a_k^\dagger \, |0\rangle_f \\
(\text{using 1.85}) \qquad &= \sum_j |\beta_{ij}|^2 .
\end{aligned}
\tag{1.99}
$$

As a result, the vacuum for one observer, which measures particles with respect to the time coordinate t, is not the vacuum for another observer which defines particles w.r.t t'. Consequently, the vacuum is not uniquely defined in curved spacetime.

In the next chapter, this formalism will be applied to understand the *Unruh effect*, which occurs for Rindler observers in Minkowski spacetime. Thereafter, this idea will be applied on black holes originating from gravitational collapse. This will confirm that black holes indeed behave as thermal objects as already suggested by the similarities with thermodynamical laws discussed in section 1.2.3.

Chapter 2

The black hole information paradox

In the previous chapter, we started by reviewing some important aspects of classical black holes. Importantly, we discussed the findings of Bekenstein and Hawking in the early seventies (originally published in [22] and [23]). Especially, in subsection 1.2.3, we saw that Bekenstein noticed interesting similarities between the laws of thermodynamics and the dynamics of classical black holes. These findings hint that a black hole behaves as a thermal object with a corresponding entropy and temperature. Remarkably, we explained that this suggests that a black hole could be viewed as an ordinary quantum as seen from the outside. However, this would contradict the no-hair theorem from general relativity, which states that a black hole is fully characterized by only three externally accessible observables. In this chapter, we will see that this is not the only problem with this thermal behaviour.

We build on the sections from the previous chapter and study their impressive implications in this chapter. To start with, we use the formalism of quantum field theory on curved spacetime, as defined in 1.4, to study quantum mechanical behaviour of black holes. Especially, we will see how the suggestions, presented by Bekenstein, gained credibility when Hawking used semi-classical physics to study the thermal behaviour of the black holes. Namely, by describing matter in a quantum mechanical way, while treating gravity classically, Hawking saw that black holes actually radiate particles as if they were thermal objects [23]. This is discussed in section 2.1.

The idea that black holes radiate clearly contradicts what one expects from a purely classical point of view, as this implied that nothing ever comes out of a black hole. Interestingly, this is not the only problem which Hawking encountered when studying black holes. Namely, a couple of years later, he concluded that considering black holes in this semi-classical limit results in information loss. Specifically, Hawking's computations suggested that the black hole should eventually evaporate (almost) completely, leaving only the evaporated particles behind. Assuming that black holes are formed from a pure state, this seemed like an enormous problem. Namely, as the in- and out-going particles are entangled, the final state, which only consists of the radiation, is in a mixed state. This is in strong contradiction with unitarity and therefore with quantum mechanics, from which this process initially resulted. Hawking therefore concluded that his semi-classical computation breaks down or that information disappears between the formation and disappearance of the black hole. Therefore, this conclusion has been referred to as the *information paradox*. This problem is formulated as a theorem and discussed in detail in section 2.2.

In the 1990's, Page, who was a student of Hawking, introduced a new way of thinking about the process of black hole evaporation. Namely, he suggested that one could quantify the amount of information, contained inside the black hole and the radiation, by computing the corresponding entanglement entropy. By doing so, he has shown that the computation of Hawking already yields a deviation from unitarity way before one expects the semi-classical approximation to break down [50]. Especially, he suggested that, if the black hole evaporation process is unitary, the entropy of its radiation should stop increasing at some moment [49]. The latest instant at which this should happen is now referred to as the *Page time*. These findings suggest that the semi-classical approximation is not necessarily breaking down, but Hawking's original computation *is*. This clearly gave a new perspective towards the paradox for many physicists and inspired them to find a new way of computing the entanglement entropy. This should yield a new formula with respect to the the ordinary von

Neumann entropy of the radiation, as discussed in section 2.3. Several years ago, such a formula has been derived, at least for simplified problems [50]. This resulted in an entropy which also includes disconnected interior regions of the black hole, referred to as *islands*. By doing so, one also considers gravitational effects while computing the entropy of the quantum mechanical radiation.

Summarizing our approach; in the first part of this chapter (section 2.1) we obtain a better understanding of the black hole radiation process. Then, in section 2.2, we formulate the so called information paradox. Starting with a discussion on the stretching of spacetime and the resulting process of particle creation in black hole geometries, we formulate the paradox as a theorem in subsection 2.2.3. After that, we discuss how the entropy of the radiation increases monotonically according to Hawking's computation. In the subsequent section 2.3, we introduce the notions of fine- and coarse grained entropy to formulate the "Page theorem" and describe the "Page curve". To finish with, we briefly present the recent developments concerning more general ways of deriving the the entanglement entropy formula for the radiation in section 2.3.7. However, this will only be discussed on a conceptual level for now. After a general discussion about the required theoretical formalism in chapter 3, this is treated it in detail in chapter 4.

2.1 Hawking radiation

In the first chapter, classical black holes were discussed and seemed to be objects from which, once the event horizon is crossed, nothing can escape. In this section, this classical view will be dramatically altered by an important prediction of Hawking which states that black holes evaporate. To do so, we will consider quantum field theory in curved spacetime. Note that this section is strongly based on [24],[28] and [34].

To start with, we study quantum field theory in Minkowski spacetime and consider a Rindler observer. In order to obtain the number of particles such an accelerating observer would measure in the Minkowski vacuum, we use the Bogoliubov transformations derived in section 1.4.2. Interestingly, will see that a Rindler observer perceives the vacuum, defined by an inertial observer, as a thermal bath. This observation is referred to as the *Unruh effect* and is derived in section 2.1.1. To finish with, a similar procedure is applied on the Schwarzschild black hole and will yield the result of Hawking's original paper: "quantum mechanical effects cause black holes to create and emit particles as if they were hot bodies" [34].

2.1.1 The Unruh effect

In section 1.2, we saw that so called Rindler coordinates (η, ρ) are able to describe a patch of Minkowski spacetime, denoted by R in Figure 1.3. In this section, we are going to use the general expressions of the Bogoliubov coefficients (1.96), to obtain the number of particles that a uniformly accelerating, e.g. a Rindler observer, measures in the Minkowski vacuum. In what follows, we discuss the approach used in [24].

In section 1.4.2, we started by expanding the scalar field in terms of modes, defined with respect to a certain time coordinate. Before one expands the scalar field in terms of Rindler modes, the correct equation of motion for a massless scalar field must be obtained. Furthermore, note that this discussion holds only in the Rindler-patch R (see Figure 1.3). We follow the discuss the derivation presented in [24] and work out the necessary steps below and in appendix B.2.

Equation of motion in Rindler coordinates

First, remember that one can define Rindler (tortoise) coordinates (t, x) which are related to the Minkowski lightcone coordinates (U, V) as

$$U = -\frac{1}{a}e^{-a(t-x)} = -\frac{1}{a}e^{-au} \qquad V = \frac{1}{a}e^{a(t+x)} = \frac{1}{a}e^{av}. \tag{2.1}$$

In these coordinates, the Rindler metric has the form

$$ds^2 = e^{2ax}(-dt^2 + dx^2) \tag{2.2}$$

or, in components,

$$g_{\mu\nu} = \begin{pmatrix} -e^{2ax} & 0 \\ 0 & e^{2ax} \end{pmatrix} \qquad g = \det(g_{\mu\nu}) = -e^{4ax} g^{\mu\nu} \sqrt{-g} = e^{2ax} g^{\mu\nu} = \eta^{\mu\nu}. \tag{2.3}$$

As a result, the action for the scalar field ϕ_R in Rindler coordinates is

$$
\begin{aligned}
I &= \int d^2x \frac{1}{2} \partial_\mu \phi_R \partial_\nu \phi_R g^{\mu\nu} \sqrt{-g} \\
&= \int d^2x \frac{1}{2} \partial_\mu \phi_R \partial_\nu \phi_R \eta^{\mu\nu} \\
&= \int d^2x \frac{1}{2} \partial_\mu \phi_R \partial^\mu \phi_R,
\end{aligned}
\tag{2.4}
$$

which is the same as the action in Minkowski coordinates (1.79). Consequently, the equation of motion is the same and we can expand the general solution in terms of positive frequency modes with the same form as (1.84). The only difference is the frequency, which is defined with respect to another time coordinate. Namely, the Rindler observer measures a time u and corresponding frequency Ω

while a Minkowski observer measures U and frequency ω. With these conventions, the Rindler scalar field is expanded in terms of $g_\Omega(u)$ as

$$\phi(u) = \int_0^\infty d\Omega \left(b_\Omega g_\Omega(u) + b_\Omega^\dagger g_\Omega^*(u) \right)$$

$$\text{where} \qquad g_\Omega(u) = \frac{1}{\sqrt{4\pi\Omega}} e^{-\Omega u}. \tag{2.5}$$

In contrast, for Minkowski coordinates, the general solution can be expanded in terms of $f_\omega(U)$. Note that, we must restrict the integral in (1.83) to the region described by the Rindler coordinates; $U \leq 0$. As a consequence, the solution is expanded as

$$\phi(U) = \theta(-U) \int_0^\infty d\omega \left(a_\omega f_\omega(U) + a_\omega^\dagger f_\omega^*(U) \right)$$

$$\text{where} \qquad f_\omega(U) = \frac{1}{\sqrt{4\pi\omega}} e^{-\omega U}, \tag{2.6}$$

where $\theta(-U)$ is the Heaviside function which yields zero for $U > 0$.

In order to obtain the number of particles detected by a Rindler observer, one needs to obtain the Bogoliubov coefficients as in (1.96). To do so, note that the inner product between two functions must be defined in a covariant way in order to obtain the correct (g, f). This can be done with the help of a timelike hypersurface Σ [24],

$$(\phi_1, \phi_2) = -i \oint_\Sigma \left(\phi_1 \partial_\mu \phi_2^* - \phi_2^* \partial_\mu \phi_1 \right) n^\mu \sqrt{-h} d^2 x, \tag{2.7}$$

where n^μ and $h_{\mu\nu}$ are, receptively, the normal vector and induced metric on this hypersurface. For more information on these subjects, see appendix C.

Furthermore, note that the covariant product obeys

$$(\phi_1, \phi_2)^* = (\phi_2, \phi_1) \qquad (\phi_1^*, \phi_2^*) = -(\phi_2, \phi_1). \tag{2.8}$$

With this covariant definition in mind, one can choose a convenient hypersurface to obtain the correct scalar product

$$\alpha_{\Omega\omega} = (g_\Omega, f_\omega). \tag{2.9}$$

For example, choose the hypersurface that corresponds to the slice $t = 0$, $-\infty < x < \infty$ or, in Minkowski coordinates, $T = 0, X \geq 0$. Then, the corresponding normal is $n_\mu = (\frac{\partial}{\partial t})_\mu$ and the scalar product yields

$$\alpha_{\Omega\omega} = -i \int_{-\infty}^\infty dx \left(g_\Omega^R(u) \frac{\partial}{\partial t} f_\omega^*(U) - f_\omega^*(U) \frac{\partial}{\partial t} g_\Omega^R(u) \right). \tag{2.10}$$

As the modes $f_\omega(U)$ do not explicitly depend on t, it is more convenient to write the derivative $\frac{\partial}{\partial t}$ in terms of a derivative with respect to U. To do so, note that on the hypersurface Σ:

$$u = t - x = -x \Longrightarrow \qquad du = -dx$$

$$n^\mu = (1, 0) = \frac{\partial}{\partial t} \Longrightarrow \qquad \frac{\partial}{\partial t} = \frac{\partial}{\partial u}. \tag{2.11}$$

Thus, by remembering the relation between u and U (see equation (2.1)), one can write

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial u} = e^{-au} \frac{\partial}{\partial U} \qquad \text{such that} \qquad \frac{\partial}{\partial t} f_\omega^*(U) = e^{-au} \frac{(i\omega)}{\sqrt{4\pi\omega}} e^{i\omega U}. \tag{2.12}$$

As a consequence of these remarks, the Bogoliubov coefficient can be obtained by computing the following integral

$$\alpha_{\Omega\omega} = \frac{i}{4\pi\sqrt{\omega\Omega}} \int_{-\infty}^\infty du\, e^{-i\Omega u} e^{i\omega U} \left(i\omega e^{-au} + i\Omega \right). \tag{2.13}$$

In Appendix B.2, it is shown that with the use of the Γ-function, this yields

$$\alpha_{\Omega\omega} = -\frac{1}{2\pi a} \sqrt{\frac{\Omega}{\omega}} e^{\frac{\pi\Omega}{2a}} e^{-\frac{i\Omega}{a} \ln \frac{\omega}{a}} \Gamma\left(\frac{i\Omega}{a}\right). \tag{2.14}$$

Furthermore, the second Bogoliubov coefficient is obtained as

$$\beta_{\Omega\omega} = -(g_\Omega, f_\omega^*). \tag{2.15}$$

It is shown in appendix B.2 that the explicit expression can be obtained from (2.14) and yields

$$\beta_{\Omega,\omega} = -e^{-\frac{\pi\Omega}{a}}\alpha_{\Omega\omega}. \tag{2.16}$$

Finally, with the use of (1.99), the number of particles detected by a a Rindler observer can obtained by

$$\langle 0|_M N_\Omega |0\rangle = \int_0^\infty dw |\beta_{\Omega\omega}|^2. \tag{2.17}$$

By noting that (1.94) implies that

$$\sum_k (\alpha_{ik}\alpha_{jk}^* - \beta_{ik}\beta_{jk}^*) = \delta_{ij}$$
$$\sum_k (\alpha_{ik}\beta_{jk} - \beta_{ik}\alpha_{jk}) = 0, \tag{2.18}$$

one obtains that

$$\int_0^\infty d\omega(|\alpha_{\Omega\omega}|^2 - |\beta_{\Omega\omega}|^2) = \delta(\Omega - \omega). \tag{2.19}$$

Combining this with (2.16), one obtains

(2.20)such that the number of B particle measured in the Minkowski vacuum is given by

$$\langle 0|_M N_\Omega^R |0\rangle = \frac{\delta(\Omega - \omega)}{e^{\frac{2\pi\Omega}{a}} - 1}. \tag{2.21}$$

If we interpret the delta as a 2-dimensional volume, we see that the number of B particles per volume is actually given by

$$N_{norm,\Omega} = \frac{(2\pi)^2}{V}N_\Omega, \tag{2.22}$$

such that the Rindler observer in the R-patch observes that the number of particles produced in each Rindler mode is

$$N_{norm,\Omega} = \frac{1}{e^{\frac{2\pi\Omega}{a}} - 1}. \tag{2.23}$$

This corresponds to a thermal spectrum with corresponding temperature

$$T_U = \frac{a}{2\pi}. \tag{2.24}$$

Hence, the Rindler observer sees the inertial vacuum state as a thermal bath, like blackbody radiation, with a temperature T_U. Corresponding B-particles are detected at a rate given by (2.23) [24]. This effect is referred to as the *Unruh effect*.

2.1.2 Black hole evaporation

Remember from section 1.2 that black holes in general relativity are regions from which nothing can escape. However, in [34], Hawking has shown that quantum mechanical effects make the black hole create and radiate particles. These particles are emitted at a rate one would expect from a hot body with temperature $T = \frac{\kappa}{2\pi}$. To show this, Hawking worked in the semi-classical approximation and considered quantum field theory in curved spacetime. Especially, he used a similar procedure as presented in section 2.1.1. As the computation involving black holes is a bit more involved, we will only briefly discuss the procedure and its consequences. The explicit derivation is presented in [28] and [34], on which this subsection is based.

First of of all, it is important to note that one needs to consider the time-dependent formation phase of the black hole to understand the process of particle creation. This is different from most

computations where only the stationary final phase of the black hole is taken into account, as discussed at the end of section 1.2.2. Especially, Hawking started by considering a spherically symmetric collapsing body which produces a black hole. The corresponding Penrose diagram was described in section 1.2, especially by Figure 1.6.

To obtain the emission spectrum of such a black hole, Hawking considered a wave packet which propagated back in time, from the far future I^+ to the far past I^-. To understand better why, one can think about the procedure used in section 2.1.1 and try to apply it on the gravitational collapse spacetime. Namely, as before, one needs to choose two bases of solutions which correspond to wavepackets, defined with respect to a convenient time coordinate on the spacetime patch under consideration. As this collapsing spacetime is nearly Minkowski in the far past, one can choose this slice to define positive frequency modes. More specifically, one can choose the usual particle states in flat space (2.6). A general solution of the equation of motion can then be expanded in positive frequency modes as in (1.89) and the vacuum is defined as in (1.86).

Finding a convenient other slice is more difficult. Namely, if one wants to study the evolution of these positive modes towards late times, one needs to define modes on both I^+ and the event horizon, denoted by the green line on Figure 1.6. As explained in [28], this is needed to obtain a complete set of solutions of the equations of motion. Furthermore, it is also more difficult to solve these equations for this geometry and global analytic solutions are not known. Therefore, instead of considering an exact solution of the equation of motion, Hawking pictured a wave packet that propagates from I^- towards the horizon of the black hole. At that point, he considered the case where the wave packet scatters in two parts. Namely, one back scatters off the geometry with a fraction $1 - \Gamma_\omega$, while the remaining part propagates along the horizon. This fraction is finally absorbed by the black hole and represents the part of the wavepacket which leads to particle production [28]. Concretely, Hawking used this picture to study a wavepacket which propagates back in time, i.e. from the black hole horizon to the far past I^-. By imposing boundary conditions such that the wave has only positive frequency modes at I^+, he dynamically obtained the behaviour of the wavepacket at I^- and decomposed this in positive and negative frequency parts.

With these ideas in mind and a similar procedure as used for the Unruh effect, Hawking obtained the following relation between the Bogoliubov coefficients

$$|\alpha_{\omega\Omega}| = e^{\frac{\pi\omega}{\kappa}} |\beta_{\omega\Omega}|, \tag{2.25}$$

where κ is the surface gravity. The spectrum of produced particles is then obtained by using the orthonormality relation (2.19) and yields

$$\langle N_\omega^{bh} \rangle = \int d\Omega |\beta_{\omega\Omega}|^2 = \frac{\Gamma_\omega}{e^{\frac{2\pi\omega}{\kappa}} - 1}. \tag{2.26}$$

In this context, Γ_ω is interpreted as the fraction of the wavepacket that is radiated to infinity [28]. This particle rate again corresponds to a thermal spectrum with temperature

$$T_{Haw} = \frac{\kappa}{2\pi} = \frac{1}{8\pi M}. \tag{2.27}$$

This result shows that the analogy between black hole mechanics and thermodynamics, discussed in section 1.2.3, is more than formal and $T \sim \kappa$ is more than a random guess [28]. From this prediction, it seems that the gravitational field of the black hole will create and radiate particles as if it was an ordinary body with temperature T_{Haw} [34]. We will refer to these emitted particles as *Hawking radiation*. Importantly, note that the particles radiated by a black hole have completely different states than particles which are usually created when a hot body radiates. Normally, the radiation is emitted from the constituents making up the hot body. In contrast, particles are pulled out of the vacuum in a region of the black hole [10].

Interestingly, such a thermal object also has a thermodynamical entropy. By filling in the expression of the temperature in (1.57), one sees that this should correspond to

$$S_{Bekenstein} = \frac{1}{4} Area. \tag{2.28}$$

Importantly, it may seem that the second law of thermodynamics is being violated. Namely, as the black hole evaporates, its area shrinks and, as a result, the entropy decreases. However, there is a

generalised second law which is obeyed by the generalised entropy

$$S_{gen} = \frac{1}{4} Area + S_{ext}.$$
(2.29)

where S_{ext} represents the entropy of the matter outside the black hole [34] and Area refers to the area of the event horizon. this will balance out the decrease in the area of the event horizon such that the full entropy still obeys the second law of thermodynamics [8].

An important consequence of the presented results is that this thermal emission leads to a mass decrease of the black hole and its eventual disappearance [34]. In the next section, we will see that this ultimately results in *the black hole information paradox.*

Interestingly, Hawking predicts that the effect of particle creation is non-negligible for *primordial black holes*[1]. Explicitly, he states that those with a mass less then $10^{15}g$ would have evaporated by now [34]. As it is expected that the evaporation yields a final explosive phase, this could possibly be detected at present. According to [42], it was believed that this final phase should be detectable now as photons with energy of order 100 MeV. However, cosmological and observational constrains suggest that there is a little chance of observing such an explosive phase, at least according to the Standard Model of particle physics [42].

Finally, note that the predicted temperature and entropy of the black hole can also be obtained in a classical approach, including path integrals. This has been done by Gibbons and Hawking in [44]. With this method, we will re-derive (2.27) and (2.28) in section 3.1.5.

---

[1]Primordial black holes are specific cases of black holes which have been generated by high densities and l inhomogeneities in the early universe. They are believed to be potential dark matter candidates, and they span a broad range of masses [41].

2.2   The information paradox

In section 2.1, it became clear that black holes radiate like thermal objects and have an entropy. In this section, we are going to study the evolution of this radiation process. To do so, we view particle creation as a process resulting from *spacetime stretching* and study how this evolves in time. Specifically, in section 2.2.2, we construct a set of slices in a Schwarzschild geometry which obey the niceness conditions that we introduced in section 1.4. These slices allow us to study the process of particle creation at the horizon and show that this eventually leads to *the black hole information paradox*. This conclusion is formulated as a theorem in subsection 2.2.3. To end with, the entanglement between the ingoing and outgoing quanta is discussed in subsection 2.2.4. Especially, we show that the corresponding entanglement entropy monotonically increases, at least according to Hawking's computation.

Note that this section is largely based on reference [10] and slightly on [[8],[18]].

2.2.1   The stretching of spacetime

To start with, we consider the process of particle creation as a result of stretching of a spacelike slice in a general geometry. Remember from our discussion on quantum field theory in subsection 1.4.2 that the vacuum is not longer uniquely defined in curved spacetime. Two implications of this result have been discussed in section 2.1. To do so, we considered an accelerating observer in subsection 2.1.1, while we studied the propagation of a wavepacket in subsection 2.1.2. In this section, we will look at the evolution of a spacelike slice and the corresponding wave modes defined on this region. This will yield an alternative way of looking at the particle creation process.

Consider a spacelike slice which is initially flat and on which one can define a vacuum state with respect to the Minkowski time coordinate. The corresponding complete basis $\{f_{i1}\}$ can be chosen to consist only of positive frequency modes, as in (1.89). However, if this specific region becomes curved after some time and thereafter returns to a flat geometry, the basis $\{f_{i3}\}$, which only consists of positive frequency modes, is not the same as the initial basis $\{f_{i1}\}$. Therefore, the initial vacuum state will not be the same as the final vacuum state, as discussed in section 1.4.2. This can be interpreted as particle creation resulting from the time dependency of the metric or the gravitational field [34].

We will consider such a spacelike slice, which must obey the niceness conditions[2] (defined in 1.4) and describe the evolution of its geometry as the *stretching* of a region on this slice. Furthermore, we describe the resulting particle creation process as a quantum fluctuation in the vacuum energy. Consequently, we picture the evolved slice as containing a virtual particle pair which is being created and annihilated over time [8], as depicted on Figure 2.1. In the next subsection we apply this view on a region near the event horizon of a Schwarzschild black hole.
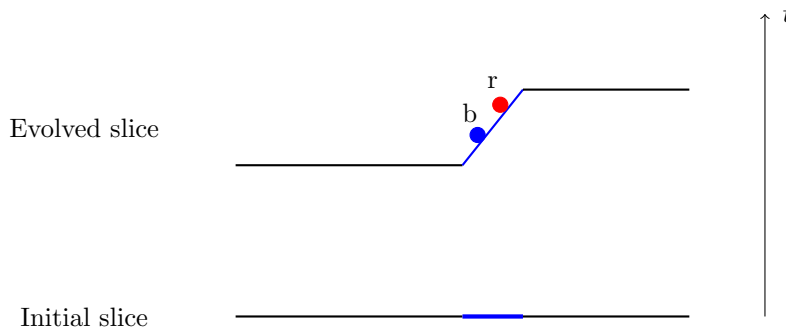
Figure 2.1: Evolution of a spacelike slice in a general geometry. While the initial slice is flat, the evolved slice contains a deformed region. As a consequence, the vacuum on this region is not the same as the initial vacuum in the same region. We interpret this as particle creation due to the time dependency of the metric. Figure inspired on Figure 1 in [10].

---

[2]This is necessary to use the results of quantum field theory in curved spacetime which are derived in the semi-classical approximation.

The state of the created particle pair on the evolved slice is

$$
\begin{aligned}
|\Psi\rangle_{pair} &= C b^{\dagger} r^{\dagger} |0\rangle_b |0\rangle_r \\
&= \frac{1}{\sqrt{2}} |0\rangle_b |0\rangle_r + \frac{1}{\sqrt{2}} |1\rangle_b |1\rangle_r .
\end{aligned}
\tag{2.30}
$$

Note that this is a simplified form of the actual state but our only goal is to capture the entanglement between the created particles, as will become clear later on.

As in [10], we assume that there is also some matter in a state $|\psi\rangle_M$ on the slice, but only very far $(L \sim 10^{77}$ light year) from where the particles are created. This assumption is referred to as *locality* and ensures that this matter only has a small effect on the particle creation. This enables us to write the complete state on the evolved slice as

$$
|\Psi\rangle \approx |\psi\rangle_M \otimes \frac{1}{\sqrt{2}} \Big(|0\rangle_r |0\rangle_b + |1\rangle_r |1\rangle_b \Big).
\tag{2.31}
$$

However, the assumption of locality still allows small departures from this state, which explains the approximation sign. To understand which departures are allowed and which are not, we computed the entanglement entropy of modifications in appendices A.2, A.3 and A.4. These computations show that one can make departures only in a sense that the created particles remain entangled with each other, but not with the matter far away. More precisely, we obtained that the entanglement entropy (defined in (1.74)) of a local state should yield

$$
S_{vN,A} = \ln 2.
\tag{2.32}
$$

In what follows, we are going to see that spacelike slices can be stretched in a black hole geometry, and discuss how this results in the information paradox.

2.2.2 Nice slices in a Schwarzschild geometry

In order to study the process of particle creation in a black hole geometry, we need to show that spacelike slices can be constructed and evolved in a way that the niceness conditions are continuously satisfied. This is needed such that we may use the semi-classical approximation, introduced in section 1.4.2. This discussion follows the approach in [10].

Especially, we are going to construct spacelike slices for the Schwarzschild black hole with metric (1.6),

$$
ds^2 = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2.
\tag{2.33}
$$

As this black hole has a physical singularity inside the horizon, we need to make sure that the spacelike slices constructed in this spacetime do not intersect this region. Namely, this would imply that the niceness conditions are violated. Therefore, we must explicitly construct slices S which do not suffer from this problem.

Importantly, remember that the notion *spacelike* refers to a spacetime interval for which $(\Delta s)^2 > 0$ [5]. By looking at the Schwarzschild metric (2.33), one observes that a spacelike slice corresponds to one where $dt^2 = 0$ when $r > 2M$ because $(1 - \frac{2M}{r}) > 0$. In contrast, when $r > 2M$, a spacelike slice corresponds to a slice which is constant in r as $(1 - \frac{2M}{r}) < 0$ in that regime.

Initial slice

We start with the construction of an initial slice $\mathcal{S}_1$, where it must be possible to define a quantum vacuum state. Assuming this is the case, it is shown in appendix C that one can construct a spacelike slice that obeys the niceness conditions. Namely, if one divides a spacelike slices in the following three parts:

(I) $r > 4M$ the slice is $t = t_1 =$ constant such that $dt = 0$ in this region

(II) $\frac{M}{2} < r < \frac{3M}{2}$ the slices are $r = r_1 =$ constant (as the spacelike and timelike directions switch when going inside the event horizon)

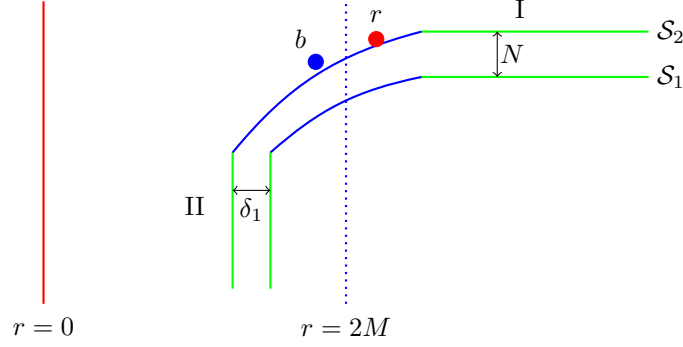(III) Join the parts (I) and (II) by a connector segment $\mathcal{C}$., the niceness conditions are satisfied along

Figure 2.2: Schematic representation of the slice construction $\mathcal{S}_1$ and $\mathcal{S}_2$. The green lines represent the $t = constant$ and $r = constant$ segments, while the connector segments are denoted by the blue lines. From this picture one sees that the condition that $\delta_1 \to 0$ implies that the connector segment gets stretched during the evolution from $\mathcal{S}_1 \to \mathcal{S}_2$. As a result, a particle pair is created, denoted by the red dots (radiated particles) and blue dots (ingoing quanta). Note that this does not correspond to the actual black hole geometry but only helps to understand the intuition behind the construction. The slice in the actual geometry is depicted on Figure 2.3. Note that this Figure is based on Figure 2 in [8].

the whole slice.

A schematic representation of such a construction can be found on Figure 2.2, where the initial slice is referred to as $\mathcal{S}_1$.

Later slices

We are ultimately interested in the evolution of the initial slice $\mathcal{S}_1$ in time. To study how it evolves, we need to define a second, future slice $\mathcal{S}_2$ such that it also obeys the niceness conditions:

(I) $r > 4M$ the slice is $t = t_1 + \Delta = constant$. As a result, the intrinsic geometry of this segment does not change, it just advances with a lapse function $N = (1 - \frac{2M}{r})^{\frac{1}{2}}$.

(II) The $r = constant$ part becomes $r = r_1 + \delta_1$ with $\delta_1 << M$ such that its geometry also remains unchanged. Especially, we consider the limit $\delta_1 \to 0$.

(III) the parts (I) and (II) are joined through a connector segment $\mathcal{C}$ in a way that leaves the two parts invariant. As a result, the connector segment has to stretch during the evolution $\mathcal{S}_1 \to \mathcal{S}_2$.

The schematic representation of the evolved slice is depicted on Figure 2.2, where it is referred to as $\mathcal{S}_2$. The stretching of the connector segment, with respect to that same region on slice $\mathcal{S}_1$, implies that a particle pair is created on the second slice. Namely, as explained in section 2.2.1, this can be understood as a result of quantum fluctuations and the time dependency of the metric[3].

If we consider this particle creation very near the event horizon of the black hole, these pairs are created such that one particle is trapped inside the black hole (blue dot on Figure 2.2) while the other escapes towards infinity (red dot on Figure 2.2). Consequently, the stretching of the connector segment results in Hawking radiation [10] that we discussed in section 2.1.2[4].

Of course, this particle creation process continues after one pair has been created. As long as the new slices, at later times, are constructed in a way that avoids breaking the niceness conditions, the connector segment keeps on stretching in the same way as described above. As a result, more and more particle pairs will be created in this region.

As we will be interested in the consequences of this process in physical black holes, one must note that we can extend this construction to a black hole formed by gravitational collapse (Figure 1.6). In such a geometry, the r =constant slice can be smoothly extended to $r = 0$ at early times. This

---

[3]Note that one could argue that our previous discussion is not valid as the metric (1.6) looks time-independent. However, this is only an illusion because, as we saw in section 1.2, the Schwarzschild coordinates break down at the event horizon. Namely, any coordinate system which really covers both regions must be time-independent [10].

[4]Note that we simply described the quantum behaviour of black holes in two different but equivalent ways.

is possible because there was no singularity before the hole was formed, i.e. at early times. The construction of the slices in this geometry is shown on Figure 2.3.
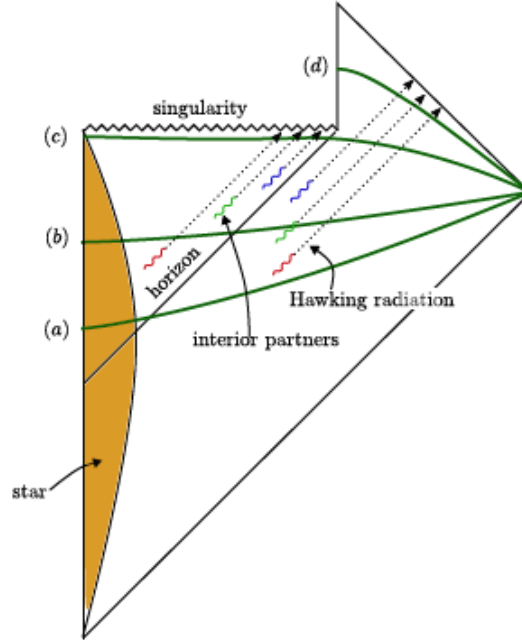
Figure 2.3: Nice slices in a black hole formed by gravitational collapse. By comparing this Figure to Figure 1.6, one sees that the orange region corresponds to the inside region of the star and the green lines correspond to the spacelike slices constructed before. The entangled quanta are denoted by the coloured short lines. As one can see, more and more quanta is trapped inside the horizon and at the black hole are shrinks due to the evaporation. This Figure is taken from [8]

2.2.3 The paradox

We are now going to combine the discussions in subsections 2.2.1 and 2.2.2 to formulate the black hole information paradox.

As stated before, the evolution of "nice" slices in the black hole geometry, leads to the creation of particle pairs. Importantly, one member of this pair escapes towards infinity, while the other one is trapped in the black hole, as depicted on Figure 2.2. The particles which float out to infinity are referred to as *Hawking radiation*. As explained in section 2.2.1, these particles are entangled with the particles trapped inside the event horizon. Remarkably, this entanglement gives rise to a problem when the black hole has radiated away almost all its mass. We will argue why this is always the case in a semi-classical approach. Namely, inspired by [10], we will formulate the essence of the black hole information paradox as a theorem and proof it. This will force us to Hawking's original conclusion: *Information is lost when the black hole has evaporated (almost) completely.*

We choose to formulate the theorem as follows:

Theorem 2.2.1. *Consider the evolution of spacelike slices, which obey the niceness conditions defined in section 1.4 and on which pure states are initially defined. If the spacetime stretching near the horizon creates particle pairs described by the leading order state 2.30, the evaporation of the black hole leads to the information paradox. Explicitly, this means that the radiation of such a black hole always leads to mixed states/remnants, which is paradoxical with the fact that this process results from working in the semi-classical approximation.*

Proof. To prove this statement, we consider the slices constructed in section 2.2.2. As these slices obey the niceness conditions, which is shown in appendix C, one may use the formalism of quantum field theory in curved spacetime. This formalism is described in section 1.4.2.

As explained in section 2.2.2, the connector segment of the slices are being stretched when time evolves, which results in the creation of particle pair (b_1, r_1). The resulting state of these particles is

defined as in equation (2.30). Additionally, the corresponding quantum state on slice S_2 is described by (2.31). Consequently, if one considers a second step in time, where the slice evolves exactly as S_1 to S_2, the state on the third slice S_3 is described by

$$|\Psi\rangle \approx |\psi\rangle_M \otimes \frac{1}{\sqrt{2}} \Big(\langle 0|_{r1} \langle 0|_{b1} + \langle 1|_{r1} \langle 1|_{b1} \Big) \otimes \frac{1}{\sqrt{2}} \Big(\langle 0|_{r2} \langle 0|_{b2} + \langle 1|_{r2} \langle 1|_{b2} \Big). \qquad (2.34)$$

Computing the entanglement of the radiated particles $\{r_1, r_2\}$ with $\{M, b_1, b_2\}$ thus yields the entropy

$$S_{vN,\{r_1,r_2\}}(t_3) = -\mathrm{Tr}\rho_{\{b_1,b_2,M\}} \ln \rho_{\{b_1,b_2,M\}} = 2\ln 2. \qquad (2.35)$$

With the same reasoning, we obtain the following state on slice S_n after N steps in the particle creation process,

$$|\Psi\rangle \approx |\psi\rangle_M \otimes \frac{1}{\sqrt{2}} \Big(\langle 0|_{b1} \langle 0|_{r1} + \langle 1|_{b1} \langle 1|_{r1} \Big) \otimes \frac{1}{\sqrt{2}} \Big(\langle 0|_{b2} \langle 0|_{r2} + \langle 1|_{b2} \langle 1|_{a2} \Big) ... \otimes \frac{1}{\sqrt{2}} \Big(\langle 0|_{bN} \langle 0|_{rN} + \langle 1|_{bN} \langle 1|_{rN} \Big),$$
$$(2.36)$$

with corresponding entanglement entropy of the radiation $\{r_1, ..., r_N\} = \{r_N\}$

$$S_{vN,\{r_N\}}(t_N) = N\ln 2. \qquad (2.37)$$

As the radiated quanta collects at infinity, the mass M of the black hole decreases. Consider the case where, after N steps, the mass reaches a value $M \sim m_{Planck}$. At that moment, the niceness conditions are not longer satisfied because the curvature $R \sim \frac{1}{M^2}$ does not obey $R << l_p^2$ anymore. Consequently, one can not longer work in the semi-classical approximation and we are therefore unable to describe the further evolution of the black hole evaporation process. However, one can only think of two possible outcomes at (very) late times [10]:

- Possibility I: The black hole evaporates completely. As the radiation is still entangled with the quanta inside the black hole (as quantified by (2.37)), this results in a final state which is mixed. As we started from a pure state containing only the matter

$$|\Psi\rangle = |\psi\rangle_M, \qquad (2.38)$$

  this means that unitarity has been broken.

- Possibility II: The evolution of the black hole stops at the moment that the mass reaches $M \sim m_{remnant}$. Consequently, the theory contains remnants. These correspond to objects with bounded mass and size, which allow arbitrarily high entanglement.

Possibility I is problematic because the violation with unitarity implies a breakdown of quantum mechanics. As we derived this process of black hole evaporation in the semi-classical approximation, which treats matter in a quantum mechanical way, this is a paradox. Namely, we implicitly assume that quantum mechanics is valid and prove that this implies that quantum mechanics is not valid.

The second possibility does not yield a direct violation of quantum mechanics but still contradicts principles of semi-classical physics. Especially, one expects only a finite number of possible states when an object's energy and spatial extend are bounded. However, this is not the case if remnants occur because they allow an arbitrary high amount of entanglement with systems far from the black hole. Furthermore, it has been argued that such remnants gives rise to loop divergences from the infinite number of possible objects in the loop [10].

The fact that our assumptions force us to accept possibility I or II is referred the *information paradox*. This name can be understood by remembering how a mixed state only captures a small amount of information about the black hole interior. Explicitly, one can not construct a quantum wavefunction, which entirely describes the final state, but one only has access to a density matrix, which captures only the information of the radiation.

In order to be complete, one must consider possible mistakes in the assumptions or approximations. We will not proof that these do not alter our result as this is perfectly described in [10]. Instead, we only summarize the possible mistakes and resolutions briefly.

One could possibly argue that:

- The set of niceness conditions presented in section 1.4 are not complete. For example, one could add a condition which avoids the segments to stretch "too much". However, it is argued in [10] that adding such a condition has consequences for physics in other situations, such as in Cosmology. This also seems to hold for other additional conditions.

- Deviations from the state (2.31) could alter the conclusion. Luckily, it is proofed in [10] that small corrections (as studied in appendix A.3) do not change our conclusion. Concretely it is shown that only modifications of order unity to our result could alter our theorem. However, it is stated that the semi-classical approximation breaks down in that case.

- There are small correlations between the infalling matter and the radiated pairs such that the *r*-quanta *do* carry information about the black hole interior. However, this is not happening, as discussed in [10].

As a result, we conclude that our assumptions indeed lead to the information paradox. □

Interestingly, the fact that this paradox arises is a result of all the important notions that we encountered in the previous subsections: black hole horizons, entanglement of quantum particles, QFT in curved spacetime and the thermal behaviour of a black hole. This can be understood by explaining the paradox in the following way, based on the discussion in [8]: the fact that one reaches a paradox can be seen as a consequence of quantum entanglement near the horizon. Namely, imagine that the black hole is formed by gravitational collapse from a pure state. As the vacuum in quantum field theory is an entangled state, one can divide space in two parts such that the whole state is pure but the substate in each subregion is mixed. As these two parts are separated by a horizon in a black hole, one particle of an entangled pair can escape towards infinity while the other one is trapped in the black hole interior. If one only considers the Hawking radiation, it is in a mixed state which looks like a thermal state with temperature T_{Haw}. As the black hole has evaporated completely, one can not consider the ingoing quanta anymore and one is only left with the thermal, mixed state. Consequently, unitarity has been broken as we started with a pure state and are only left with a mixed one.

The paradox we encountered leads to several possible conclusions. One of them is that new physics should be encountered when black holes evaporate, as quantum mechanics breaks down. Several other attempts are being made to solve this paradox, lots of them coming from string theory and especially the AdS/CFT correspondence. These results strongly suggest that information does come out but requires certain dualities to quantum systems where the geometry of spacetime is not manifest [8]. In the chapter 4, we will see that one can partially resolve the tension between black hole evaporation and quantum mechanics by considering non-trivial saddles of the gravitational path integral. This approach avoids holographic assumptions and seems consistent with other results. To understand which partial tension is resolved, we introduce the notion of the *Page curve* and *gravitational fine grained entropies* in the next section. Before we move on to that, we discuss the evolution of the entanglement entropy of the radiation, as originally computed by Hawking, in more detail.

2.2.4 Entropy of the radiation

Previously, we argued that the entanglement of the entropy increases with every time step in the evolution of black hole evaporation. Furthermore, it has been noted in the proof of theorem 2.2.1 that small deviations do not alter the conclusion that black hole evaporation yields the information paradox. However, it is interesting to explicitly show that small corrections do not alter the fact that the entanglement entropy increases. Namely, one could argue that the large number of quanta, after a large number of steps N, may allow overall complex entanglement which makes the entropy of the radiation decrease. We will show in this subsection this is not the case. Furthermore, in section 2.3, it will be discussed that this leads to a deviation from a unitary evolution already before the black hole has evaporated completely. Note that, in what follows, we present the approach used in [10].

In order to proof that small deformation from the state defined in equation (2.31) do not alter the conclusion that the entanglement entropy increases every step, it is convenient to write down the quantum state after n steps in terms of an orthonormal bases (as in section 1.3). Namely, consider the case where the full quantum state at time t_n is written down as $|\psi_{M,r}, \psi_r(t_n)\rangle$, where M refers to the matter, b to the ingoing particles and r to the outgoing particles. One can then choose a basis

of orthonormal states ϕ_i for the (M, b) quanta and an orthonormal basis χ_n for the radiation quanta, such that the full state is expanded as

$$|\psi_{M,b}, \psi_r(t_n)\rangle = \sum_{m,n} C_{mn} \phi_m \chi_n. \tag{2.39}$$

Remember that the Schmidt decomposition theorem (appendix A.1) states that one can make the convenient unitary transformation such that the pure state can be expressed as

$$|\psi_{M,b}, \psi_r(t_n)\rangle = \sum_i C_i \phi_i \chi_i. \tag{2.40}$$

In that case, C_i corresponds to square roots of the eigenvalues of the reduced density matrices $\rho_{M,b}$ and ρ_r simultaneously. As a result, the reduced density matrix corresponding to the hawking radiation r_i is simply given by

$$\rho_{ij,r} = |C_i|^2 \delta_{ij}. \tag{2.41}$$

The corresponding entanglement entropy at time t_n is then given by

$$S_{vN,rad}(t_n) = - \sum_i |C_i|^2 \ln |C_i|^2. \tag{2.42}$$

Now, consider the most general evolution from time step $t_n \to t_{n+1}$ denoted by the following change in the orthonormal states

$$\chi_i \to \chi_i \qquad \phi_i \to \phi_i^{(1)} S^{(1)} + \phi_i^{(2)} S^{(2)}. \tag{2.43}$$

Note that χ_i is assumed to remain unaltered because we assume that the emitted radiation at time t_n is not altered by the creation of the newest particle pair $\{b_{n+1}, r_{n+1}\}$. This corresponds to our assumption of locality, described in the previous subsection. In contrast, the state ψ_i of the quanta inside the black hole evolves into a tensor product of states consisting of S^i which describe the newest pair $\{b_{n+1}, r_{n+1}\}$. This was not considered in the proof of theorem 2.2.1 and yields corrections to the leading order result in which we are interested.

Including the changes (2.43), the full state (2.40) evolves to

$$|\psi_{M,b}, \psi_r(t_{n+1})\rangle = \sum_i C_i [\phi_i^{(1)} S^{(1)} + \phi_i^{(2)} S^{(2)}] \chi_i \equiv S^{(1)} \Lambda^{(1)} + S^{(2)} \Lambda^{(2)}$$

$$\text{with} \qquad \Lambda^{(1)} = \sum_i C_i \phi_i^{(1)} \chi_i, \qquad \Lambda^{(2)} = \sum_i C_i \phi_i^{(2)} \chi_i. \tag{2.44}$$

At this point, we can quantify what we meant with "small deformations from leading order" before. Namely, we only consider the deformations from (2.44) where

$$\left\| \Lambda^{(2)} \right\| < \epsilon << 1. \tag{2.45}$$

Furthermore, note that the orthonormality of $S^{(1)}$ and $S^{(2)}$ implies that

$$\left\| \Lambda^{(1)} \right\|^2 + \left\| \Lambda^{(2)} \right\|^2 = 1 \tag{2.46}$$

in order to have a correct normalization of $|\psi_{M,b}, \psi_r(t_{n+1})\rangle$.

At this point we would like to show that, when condition (2.45) is satisfied, the entanglement of all the radiation quanta emitted before and at t_{n+1} is larger the entanglement at t_n. Before we actually show this, we need to introduce three lemmas, which are proofed in appendix A.5 and where the notation is also discussed. With the use of these lemmas we proof the following theorem:

Theorem 2.2.2. *Suppose that at time t_n, the radiation $\{r\}_n$ has been emitted and the corresponding entanglement entropy with the inside of the black hole (consisting of matter M and ingoing quanta $\{b\}$) is denoted as $S_{\{r\}_n}$. Additionally, consider a newly created pair at t_{n+1} which has a quantum state which departs from state (2.30) by a small amount as defined in (2.45). In this case, the entanglement entropy of the radiation emitted at t_{n+1} satisfies*

$$S_{\{r\}_{n+1}} > S_{\{r\}_n} + \ln 2 - 2\epsilon, \tag{2.47}$$

e.g. the entanglement entropy of the radiation increases at each timestep, even if small deviations from leading order state (2.30) are considered.

Proof. First note that the strong subaddivity theorem [29] holds for the entropies of three systems:

$$S(A+B) + S(B+C) \geq S(A) + S(C). \tag{2.48}$$

For our system, one can take

$$A = \{r\}_n \qquad B = r_{n+1} \qquad C = b_{n+1}, \tag{2.49}$$

such that

$$\begin{aligned}
S_{\{r\}_n, r_{n+1}} + S_{r_{n+1}, b_{n+1}} &\geq S_{\{r\}_n} + S_{b_{n+1}} \\
&> S_{\{r\}_n} + \ln 2 - \epsilon,
\end{aligned} \tag{2.50}$$

where we used lemma A.5.3 in the last step. Furthermore using A.5.1 and noting that $S_{\{r\}_n + r_{n+1}} = S_{\{r\}_{n+1}}$, one finally obtains

$$S_{\{r\}_{n+1}} > S_{\{r\}_n} + \ln 2 - 2\epsilon, \tag{2.51}$$

which is what we needed to proof. $\qquad \square$

As a result of property (2.47), the entanglement entropy of the radiation, as originally computed by Hawking, always increases with time, unless there is an order unity change to the state (2.30). In the next section, we will see that this result has sever consequences for the unitarity of the black hole evolution even before $M \sim m_{planck}$. To understand that, new notions of entropies must be introduced first.

2.3   The Page curve

In section 2.1.2, we discussed how black holes behave as thermal objects, with a temperature T_{Haw} (2.27) and evaporate away their mass in the form of Hawking radiation. Furthermore, we saw that black holes also have an entropy $S_{gen} \sim Area$, from which it is seems that they could be viewed as ordinary quantum systems from the outside. However, in section 2.2, we discussed how the process of black hole evaporation results in the information paradox. This made Hawking conclude that information about the initial system is lost when the black hole has evaporated completely [8].

In this section, we are going to study the non-unitary behaviour of the black hole evaporation process by comparing the entropies of the black hole and the radiation. By doing so, we will see that a problem with unitarity already occurs way before the black hole has evaporated completely. In order to understand this, we will first introduce new concepts such as the *coarse- and fine grained entropies* in subsection 2.3.1. This will allow us to formulate the *Page theorem* and suggest that the way we computed the entropy of the radiation in section 2.2.4, may not be correct. Namely, during the past 15 years, a better understanding of the von Neumann entropy within gravitational systems was achieved [8]. This yielded a modified formula for the black hole entropy with respect to (2.29), which relies on the existence of *quantum extremal surfaces*. Furthermore, a new way to compute the entropy of the radiation was obtained more recently. We will see that these new way of computing the entropy results in an evolution which is consistent with unitarity and hence follows the *Page curve*.

Although we only present the recent developments and their consequences in this section, we will dive deeper in their origin in the next two chapters. Namely, in chapter 3, we introduce all the necessary ingredients which are needed to understand how one actually computes the correct entropies of the radiation. This will be done by considering *Replica wormholes* in chapter 4.

What follows is highly based on [8], which captures the ideas formulated by Page in [48] and [49].

2.3.1   Fine- and coarse- grained entropy

In order to understand how the entropy of the Hawking radiation illustrates non-unitary behaviour of the evaporation process, it is useful to distinguish two notions of entropy.

The first concept that is introduced is one we are already very familiar with from section 1.3: the von Neumann entropy

$$S_{vN} = -\text{Tr}\rho \log \rho \equiv S_{fine}. \tag{2.52}$$

To clearly distinguish with the coarse-grained entropy, defined below, we will refer to this entropy as the fine-grained entropy in this section. Remember from section 1.3 that this entropy quantifies the purity, and hence, the amount of information one has about the quantum state of the system. Furthermore, an important property of this entropy is that remains unchanged under unitary time evolution, as stated in (1.78). Finally, note that one can define the *fine-grained entropy* of quantum fields living only on a sub-region of space Σ. This is then written down as $S_{fine}(\Sigma)$. In such a case, the entropy can generally increase or decrease as the slice Σ moves forward in time.

The second notion of entropy, which plays a significant role in our discussion, is the *coarse-grained entropy*. In contrast to the fine-grained entropy, this entropy is obtained by considering only a subset A_i of all the observables of the system. To understand how this entropy is defined, consider a system with density matrix ρ. Secondly, construct a set of density matrices $\{\tilde{\rho}\}$ consisting of all the density matrices $\tilde{\rho}_i$ which obey

$$Tr(\tilde{\rho}_i A_i) = Tr(\rho A_i). \tag{2.53}$$

Explicitly, this means that one constructs a set of all density matrices which yield the same result for the observables under consideration A_i as the density matrix ρ, which describes the entire system. The coarse-grained entropy is then defined as the maximal fine grained entropy over all possible density matrices $\tilde{\rho}_i \in \{\tilde{\rho}\}$,

$$S_{coarse} = \max_{i=1,...,n} \Big[-\tilde{\rho}_i log(\tilde{\rho}_i) \Big], \tag{2.54}$$

where n denotes the number of $\tilde{\rho}_i$ that satisfy (2.53). The entropy thus corresponds to the fine-grained entropy of the density matrix $\tilde{\rho}_i$, which satisfies (2.53) and leads to the maximal entropy with respect to all other possible choices of $\tilde{\rho}_i$. Since ρ trivially obeys condition (2.53), it follows directly that

$$S_{fine} \leq S_{coarse} \tag{2.55}$$

always holds within a system.

A famous example of coarse-grained entropy is the entropy used in thermodynamics. Namely, this is obtained by choosing a few macro-observables, like the energy and volume, and then maximizing among all possible micro-states with that energy and volume. With the help of this example, one can naturally understand why (2.55) holds. Namely, the thermodynamical entropy provides a measure of the total degrees of freedom in a system. Therefore, it sets a limit on how much a system can be entangled with something else. As this is quantified by S_{fine}, the inequality between both entropy follows directly. Another important difference with the fine-grained entropy is that the coarse-grained entropy tends to increase under unitary time evolution. Consequently, it obeys the second law of thermodynamics.

Black hole entropies

If one assumes that the generalized entropy S_{gen} (2.29) corresponds to one of the two different entropies introduced above, it can only be the coarse-grained entropy [8]. Namely, by looking at its definition (2.29), one sees that, due to the proportionality to the area of the black hole, the entropy increases rapidly at early times. Of course, this is due to the formation of the black hole which happens in that period. Consequently, if one assumes that semi-classical physics holds and that evaporation processes evolve in an unitary way, S_{gen} can only correspond to the coarse-grained entropy which increases under unitary time evolution.

In contrast, the semi-classical entropy of the radiation computed in (2.37) is a fine-grained entropy of quantum fields, living only on a sub region of space Σ (e.g. the spacelike slices which obey the niceness conditions). In what follows, we will thus write

$$S_{gen} = S_{coarse,BH} \qquad S_{vN,rad} = S_{fine,rad}, \tag{2.56}$$

furthermore we will refer to the entropy of the radiation as computed in section 2.2 as the entropy computed by Hawking [8].

2.3.2 Information paradox revisited

We will now investigate and compare the evolution of the coarse- and fine-grained entropy related to the process of black hole evaporation.

Remember that, when a system is initially in a pure state, Lemma 1.3.1 holds. This implies that if we assume that the system consisting of the black hole degrees of freedom and the radiation produce a pure state, we must have that

$$S_{fine,BH} = S_{fine,rad}. \tag{2.57}$$

However, we saw in section 2.2 that $S_{fine,rad}$ rises during the process of black hole evaporation. Due to the above equality, this implies that the fine-grained entropy of the black hole should also increase. Importantly, this becomes problematic when the area of the black hole continuous to shrink. Namely, this implies that at a certain moment

$$S_{fine,BH} \geq S_{gen,BH} \tag{2.58}$$

holds, which directly contradicts property (2.55). Consequently, this means that the radiation is entangled with more degrees of freedom than present in the black hole. If one wants (2.57) to hold at all times, and hence obtain a unitairy process, the entropy of the radiation $S_{fine,rad}$ should start to decrease at the point where $S_{fine,rad} = S_{gen,BH}$ [8]. The explicit curve that $S_{fine,rad}$ should follow to yield consistency with unitarity has been suggested by Page in [48]. He formulated what is often referred to as the *Page theorem* and can be summarized as (inspired by [8], [18] and [49]):

Theorem 2.3.1. *If one assumes that: the black hole starts in a pure state, the Hawking evaporation is a unitary process and no large external sources interacts with the outgoing radiation, than the fine-grained entropy of the radiation may not exceed the coarse-grained entropy of the black hole at any time. Therefore, the fine-grained entropy of the radiation may monotonically increase only up to the Page time t_{page} where $S_{fine,rad} = S_{gen,BH}$. From that point, the entropy of the radiation should start to decrease back to zero as the black hole evaporates the remaining fraction of its mass.*
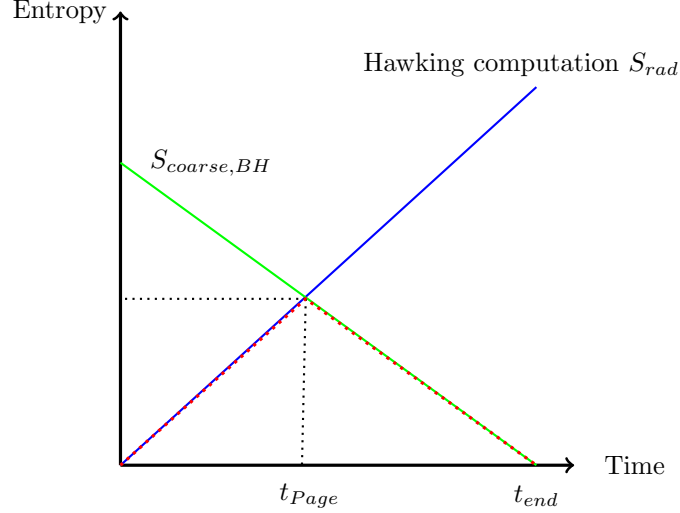
Figure 2.4: Schematic behaviour of the entropies of the black hole and the radiation. The exact curves depend on the details of the matter fields and the black holes and are presented in for example [49]. The green line represents the decrease in the generalized entropy of the black hole, which equals the coarse-grained entropy following our assumptions. The blue line represents the von Neumann entropy of the radiation, as computed by Hawking and presented in section 2.2. For the assumptions made in theorem 2.3.1, the correct fine-grained entropy should follow the Page curve and start decreasing at t_{Page}. The expected behaviour is schematically represented by the red dots. Finally, t_{end} represents the time at which the black hole has evaporated completely. Note that Figure is strongly inspired by Figure 7 in [8].

The schematic behaviour of the entropies are compared to the Page curve on Figure 2.4. From there, it is clear that the non-unitary behaviour of the radiation entropy already arises before quantum gravitational effects appear. In contrast to the information paradox, one can thus not simply state that the non-unitairy behaviour results from quantum gravity. Furthermore, this suggests that, if the assumptions made in (2.3.1) are correct, the fine-grained entropy of the radiation only corresponds to the increasing entropy as calculated by Hawking up to t_{Page}. After that the fine-grained entropy deviates from this computation and another radiation entropy should dominate. In chapter 4, we will see that large non-perturbative effects arise as non-trivial saddles of gravitational path integrals, yielding a fine-grained entropy for the radiation which is consistent with unitarity.

As stated in [8], theorem 2.3.1 relies only on the basic properties of fine-grained entropy, which makes it very robust. Furthermore, it is impossible to fix the discrepancy with unitarity by adding small deformation from the quantum state, as shown in section 2.2.4. Consequently, if there is a resolution, it should be a non-perturbative effect of order unity [8].

2.3.3   Gravitational fine-grained entropy

As suggested by the Page theorem, something is wrong with the von Neumann entropy of the radiation as computed in section 2.2.4. It turns out that the main problem is that Hawking ignored gravitational effects in his computation while the full quantum state of the initial system is obtained with the use of gravity [8]. Therefore, one should use a gravitational formula to obtain the correct fine-grained entropy of the radiation. Before we discuss such a formula, we introduce the notion of fine-grained entropies in general gravitational systems.

Firstly, note that the formula for the generalized entropy (2.29), and thus the coarse grained entropy of the black hole, can also be derived with the use of gravitational path integrals. This has been done by Gibbons and Hawking in [44] and will be discussed in section 3.1.4. Interestingly, one can also obtain a gravitational formula for the fine-grained entropy of the black hole by similar methods. The first description for such an entropy was discovered by Ryu and Takayanagi in a holographic context (i.e. in [51]) but has been generalized since then [8]. The form of the formula is similar to that of the generalized entropy (2.29), but differs from it in terms of the considered surface for which

the area is included. Namely, instead of considering the area of the horizon, one now considers the area of a *quantum extremal surface Q*. This surface is found by minimizing the generalized entropy

$$S_{gen} = \frac{\text{Area}(Q)}{4} + S_{vN}(\Sigma_Q) \tag{2.59}$$

in the spatial direction, while maximizing it in the time direction. Explicitly, the gravitational fine-grained entropy of the black hole is thus defined as

$$S_{grav,fine} = \min_Q \left\{ \text{ext}_Q \left[\frac{\text{Area}(Q)}{4} + S_{vN}(\Sigma_Q) \right] \right\}. \tag{2.60}$$

The surface Q is a co-dimension-2 object, which means that it has two dimensions less than the full spacetime, and is chosen such that it minimizes the generalized entropy (which is the quantity between the brackets). The second term $S_{vN}(\Sigma_Q)$, corresponds to the von Neumann entropy of the quantum matter fields in a region Σ_Q. This slice of the spacetime represents the region bounded by X and a cut-off surface (see Figure 2.5). Furthermore, if there are many quantum extremal surfaces, one should look for the global minimum. Additionally, note that we will refer to the quantity between brackets as the generalized entropy due to its similarities with (2.29). Finally, note that the location of the QES depends on the geometry of the interior of the black hole and time. How this is obtained specifically is discussed in section 4.1.

2.3.4 Fine-grained entropy of an evaporating black hole

In this subsection, the application of (2.60) on an evaporating black hole is shortly discussed. Although this discussion is rather conceptual, it is interesting to understand the influence of the evaporation process on the fine-grained entropy of the black hole. In the next subsection, this will be discussed for the entropy of the radiation, where the problematic discrepancy with the Page curve arises.

It is useful to consider two different types of contributions, e.g. the case where no area contribution is included and the case where it is. As the formula (2.60) dictates, one should always choose the contribution which minimizes the generalized entropy. This will be done at the end and yields the Page curve.

Vanishing area contribution

Firstly, we discuss the case where no extremal surface is encountered between the cut-off surface and the $r = 0$ boundary. This occurs at very early times (say t_0) when the black hole has formed but no Hawking radiation has escaped the horizon yet. Consequently, it does not matter where one puts Q, it will not catch more or less quanta and the matter contribution in $S_{vN}(\Sigma_Q)$ is constant. As a consequence, the QES lays at $r = 0$ such that its area shrinks to zero (see Figure 2.5) and thus minimizes the generalized entropy. The gravitational fine-grained entropy (2.60) at very early times is thus simply approximated as

$$S_{BH,gen,1} \approx S_{vN}(\Sigma_Q), \tag{2.61}$$

at early times [8]. After a while, the black hole starts evaporating and the entanglement entropy $S_{vN}(\Sigma_Q)$ starts to grow in a similar way as the Hawking entropy of the radiation grows (section 2.2.4). Namely, more and more ingoing quanta are included in the region Σ_Q, which are mixed states entangled with quanta outside Σ_Q. As a result, the vanishing QES contribution grows as

$$S_{BH,gen,1} \approx S_{vN}(\Sigma_Q) \sim t \ln 2. \tag{2.62}$$

The schematic representation of this vanishing area contribution is shown on Figure 2.5 while the growth of the corresponding entropy is depicted on Figure 2.7.

Non-vanishing area contribution

At early times (say $t_1 > t_0$), when the black hole has started evaporating, a second QES surface appears. As the corresponding entanglement entropy of the region Σ_Q depends on how much radiation is emitted, the position of this QES is time dependent. However, note that the location of Q turns out to be close to the horizon.
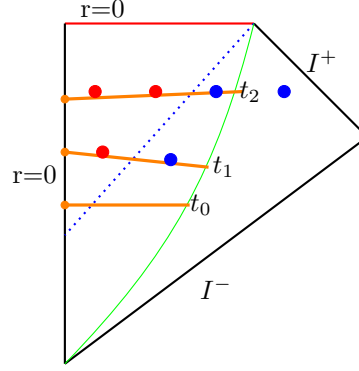
Figure 2.5: A schematic representation of the region Σ_Q, in a gravitational collapsing geometry. The captured number of quanta for different times are depicted by the red and blue dots and correspond to the case where QES is at $r = 0$. The green line represents the cutoff surface while the blue dotted line represent the event horizon. The orange lines represent the spacelike regions Σ_Q at different moments in time and the orange dots denote the locations of Q. As one can see, the number of ingoing quanta (red dots) included in the region Σ_Q grows with time. In contrast, the outgoing radiation quanta (blue dots) are less included as the region between the horizon and the cutoff surface becomes smaller with time. As a result the corresponding von Neumann entropy increases, since the ingoing states are less and less purified by the outgoing states. This Figure is based on Figures 8, 9 and 10 in [8]. Note that this diagram corresponds only to 1 spatial and 1 time direction such that a point represents a sphere.

The fine-grained entropy now depends on an area *and* a matter contribution. Nevertheless, since $S_{vN}(\Sigma_Q)$ is relatively small, the area contribution in (2.60) dominates. This can be understood because the QES lies very close to the horizon and not many quanta are captured, as seen on Figure 2.6. The corresponding fine-grained entropy is thus approximately given by

$$S_{BH,gen,2} \approx \frac{Area(Q(t))}{4}. \tag{2.63}$$

As the area of the black hole decreases, so does the area of Q. Consequently, the entropy decreases
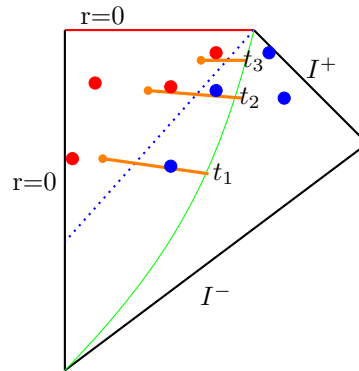
Figure 2.6: A schematic representation of the evolution of the region Σ_Q and the quanta it captures, in the case that the QES has a non-zero area. Note that the QES moves towards the horizon as time evolves. As a result, it does not capture many quanta. Therefore the "Area" contribution dominates over the entanglement entropy in (2.60). As the Area of the black hole shrinks with time, so does the Area of the QES. This contribution thus leads to a decreasing fine-grained entropy. This Figure is based on Figure 11 in [8].

with time and closely follows the evolution of the coarse-grained, thermodynamical entropy. The evolution of this entropy is depicted on Figure 2.7.

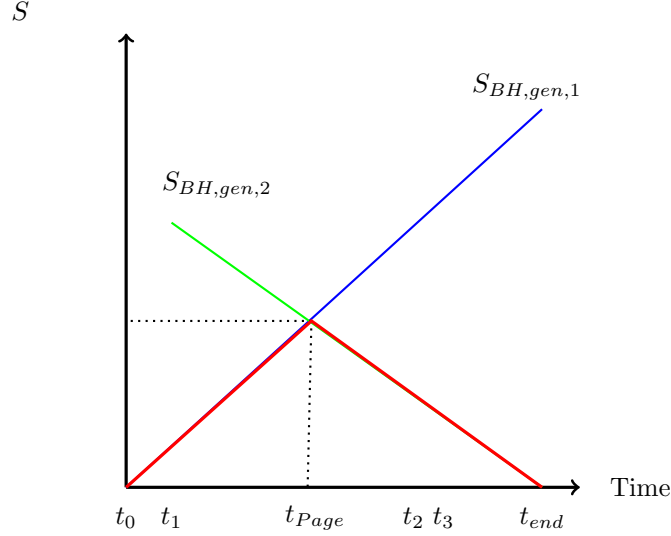Figure 2.7: The evolution of the generalized entropies, resulting from the non-vanishing and vanishing area contributions, are depicted in green and blue respectively. Additionally, the red curve corresponds to the Page curve which represents the minimum between both contributions at every moment in time. This Figure is based on Figure 13 in [8].

The page curve for the black hole entropy

As definition (2.60) tells us, one should consider the QES which minimizes the generalized entropy for every point in time t. The final gravitational fine-grained entropy thus corresponds to

$$S_{BH,fine}(t) = \min_Q\{S_{BH,gen,1}, S_{BH,gen2}\}. \tag{2.64}$$

One should thus compute which of the two contributions minimizes the entropy at which times. For the moment, we only discuss the result conceptually.

At very early times, only a vanishing area contribution exists. Consequently, one has no choice but consider this contribution. As this must start increasing from zero, it yields a smaller contribution than the non-vanishing QES contribution one up to t_{Page}. At that moment, the area of the QES has shrunk so much that $S_{BH,gen,1}(t_{Page}) = S_{BH,gen,2}(t_{Page})$. As the non-vanishing area contribution continues to decrease, this will be the relevant QES to consider up to t_3. As a result, one obtains the Page curve for the fine-grained entropy of the black hole, as depicted on Figure 2.7.

Importantly, this does not mean that the problem with unitarity is solved, as this arises due to the increasing entropy of the radiation, not the black hole. In what follows, we will see that a similar formula as 2.3.3 can be used to obtain a fine-grained entropy of the radiation which follows the Page curve as well.

2.3.5 Gravitational fine-grained entropy of Hawking radiation

In section 2.2.4, we have shown how the entanglement entropy of the Hawking radiation increases monotonically with time. However, the Page theorem 2.3.1 states that if the black hole evaporation process is unitary, the fine grained entropy of the radiation should start decreasing at t_{Page}. Consequently, the only manner to reach consistency with both theorems is by considering the possibility that the fine-grained entropy does not simply equal the von Neumann entropy calculated in section 2.2.4. Especially after t_{Page}, one should consider the case where the fine-grained entropy is calculated in a different way. Namely, after that point, the entropy should start decreasing in order to be consistent with Page's theorem.

Fortunately, there exists a generalization of (2.60) that can be used to obtain a correct fine-grained entropy of the Hawking radiation. Explicitly, it is given by

$$S_{fine,rad} = \min_Q\left\{\text{ext}_Q\left[\frac{\text{Area}(Q)}{4} + S_{vN}(\Sigma_{rad} \cup \Sigma_{island})\right]\right\}, \tag{2.65}$$

which has a very similar form as the fine-grained entropy of the black hole. The difference sits in the region enclosed by the cutoff surface and the quantum extremal surface. Namely, one could expect that computing the entropy of the radiation would only yield contributions from the outside of the black hole. However, as stated before, one should take gravitational effects into account as the original quantum state results from gravity. To do so, one can include regions *inside* the black hole in the computation of the radiation entropy, referred to as *islands*. Therefore, the surface Q in formula (2.65) refers to the boundary of such an island. Furthermore, the enclosed region on which the von Neumann entropy is computed is now disconnected. Namely, Σ_{rad} is the region between the cutoff surface all the way to up to infinity while $\Sigma_{islands}$ refers to the region from the boundary of the island up to $r = 0$. In analogy to the fine-grained entropy of the black hole, (2.65) is obtained by minimizing the expression between brackets (still referred to as the generalized entropy) with respect to the location and the number of islands.

We are now going to discuss the contributions which eventually lead to a gravitational fine-grained entropy consistent with unitarity.

No Island contribution

For similar reasons as discussed in the case of the black hole entropy, at very early times (t_0) there is no choice of an island which extremizes the entropy. Consequently, there exists a contribution where no-island is included and hence

$$S_{rad,gen,1} \approx S_{vN}(\Sigma_{rad}). \tag{2.66}$$

This means that the fine-grained entropy is simply the entanglement entropy of the radiation as computed in section 2.2.4. The corresponding region Σ_{rad} is depicted on Figure 2.8. Since this is simply the von Neumann entropy computed before, this contribution yields a monotonically increasing function $S_{rad,fine,1} \sim t \ln 2$, as depicted on Figure 2.10.
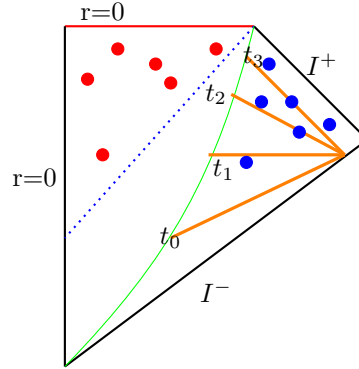

Figure 2.8: The schematic representation of the no-island contribution to the entropy. As there is no inside region included, the entanglement entropy only captures the outgoing radiation (blue dots). As a result, this grows with time, due to the increase of number of mixed states in the Σ_{rad} region, which is denoted by the orange lines. This Figure is based on Figure 15 in [8]

Island contribution

When the black hole has started evaporating, a non-vanishing island contribution, which minimizes the generalized entropy in (2.65), appears. This island lies very near the black hole horizon [8] and is centered around the origin. Consequently, the ingoing quanta is included in the computation of the von Neumann entropy as well. As a result, the outgoing radiation quanta states are purified by the quanta in the region $\Sigma_{islands}$. This is illustrated on Figure 2.9. Due to the smallness of the corresponding entanglement entropy, the area term in (2.65) will dominate. As this decreases with time, due to the shrinking of the area of the horizon, this contribution yields a decreasing fine-grained entropy for the radiation, as shown on Figure 2.10.
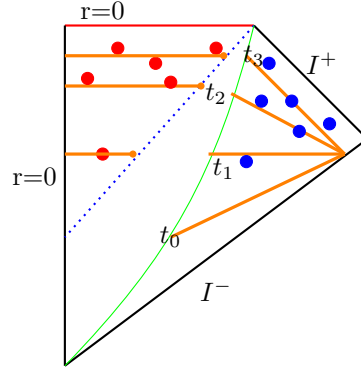
Figure 2.9: Schematic representation of the in and outer region which are included when a non-vanishing island appears. The orange lines on the left represent Σ_{island} while the ones on the right correspond to Σ_{rad}. The orange dots near the horizon refer to the boundary of the island at different moments in time. As one can see, due to the inclusion of the inside region, the quantum states of the outgoing particles are purified by the ingoing states. This results in a very small entanglement entropy such that the area term dominates in (2.65). This Figure is insipired on Figure 16 in [8].
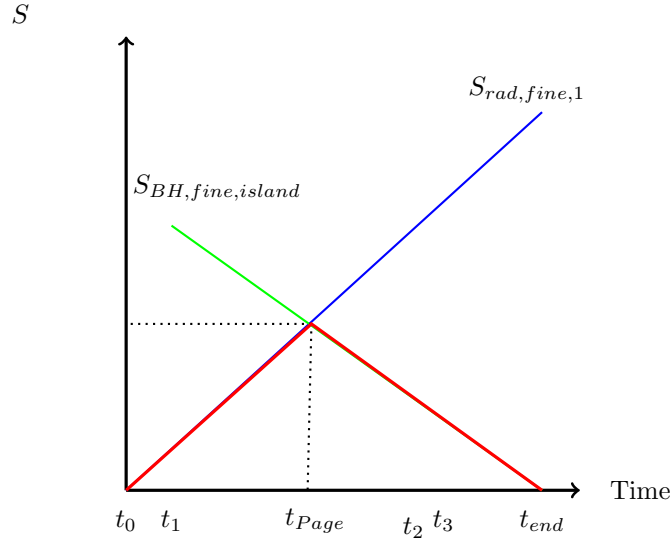

Figure 2.10: If one considers the contribution which extremizes the gravitational fine-grained entropy at each point in time, one obtains the Page curve. At early times $t < t_{Page}$, the entropy is increasing as originally computed by Hawking and in section 2.2.4. However, the island contribution minimizes the generalized entropy after t_{page} such that the resulting fine-grained entropy decreases. As a consequence, the entropy of the radiation is now consistent with unitarity. This Figure is based on Figure 18 in [8].

2.3.6 Page curve for the radiation

As one can already expect from our discussion on the fine grained entropy of black holes, minimizing the generalized entropy in (2.65) over the two discussed contributions yields the Page curve. Namely, at early times the no-island contribution minimizes the fine-grained entropy as it starts to increase from zero. This changes at t_{page} where $S_{rad,gen,1} = S_{rad,gen,2}$. At that point, the area of the event horizon has become so small that the island-contribution yields an entropy which equals the entropy computed by Hawking $S_{rad,gen,1}(t_{page})$. Consequently, the correct fine-grained entropy after t_{page} equals this contribution instead of the increasing entropy we considered up to now. Doing so yields an entropy which satisfies the Page theorem 2.3.1 and follows the corresponding Page curve depicted on Figure 2.10.

We thus conclude that the use of a gravitational formula for the fine-grained entropy of the radiation

resolves the tension with unitarity. Namely, we have found that $S_{BH,fine} = S_{rad,fine} \leq S_{BH,coarse}$ now holds at all times. This observation does not fully resolves the information paradox as we still do not know a way to obtain the explicit details of the quantum state of the radiation. However, the gravitational fine-grained entropy suggests that black hole evaporation is a unitary process [8] and thus, is not paradoxical with semi-classical physics.

2.3.7   Preview: replica wormholes

At this point, one could argue that it is not surprising that one finds the Page curve. Namely, one could think that formula (2.65) has simply been defined such that it computes the entropy of a pure state (consisting of the ingoing and outgoing quanta). However, formula (2.65) and the island contributions to the gravitational fine grained entropy have not been introduced "by hand". In contrast, the formula has been derived from methods including the gravitational path integral [8]. We will see how semi-classical gravity naturally yields the entropy consistent with unitarity in chapter 4. However, before on can understand these derivations, one must introduce a lot of new formalism's first. These are discussed in chapter 3. However, the reader already familiar with the discussed formalism can skip this and move on to the discussion about replica wormholes 4.2. This refers to a non-trivial saddle of the semi-classical gravitational path integral. As will be discussed later, the entanglement entropy can be obtained by computing the path integral over different copies of the original system. In gravity, this actually yields a sum over all possible topologies which connect the interiors of different copies (*replicas*). We will see that, there is a trivial manner of connecting such copies that leads to the monotonically increasing radiation entropy as computed by Hawking and in section 2.2. In contrast, one can also consider the topology where copies are connected through *wormholes*. This yields an entropy where island contributions are included and hence decreases. As this corresponds to a saddle of the gravitational path integral, which starts to dominate over the trivial contribution around t_{page}, the Page curve is naturally obtained with gravitational methods. A sketch of the different topologies is depicted on Figures 2.11 and 2.12. Note that these pictures result from taking only two copies of the entire system (consisting of the black hole and its radiation).
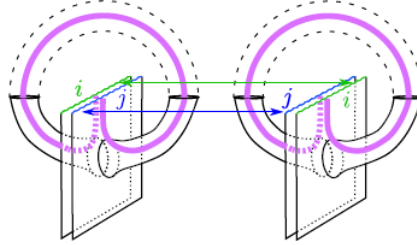
Figure 2.11: A trivial way of connecting two replicas of the black hole geometry is by gluing the outside regions to each other (blue and green lines). This corresponds to a saddle of the gravitational path integral that yields the Hawking radiation entropy. This Figure is taken from [8].
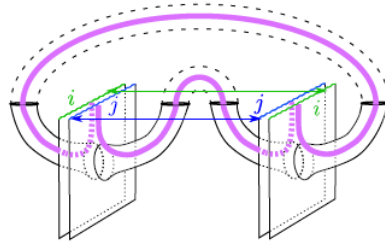
Figure 2.12: This sketch corresponds to the replica wormhole topology, where two copies of the original black hole are connected through their inside region. The connection "tube" is referred to as a wormhole. This saddle starts do dominate the gravitational path integral around Page time and yields an entropy consistent with unitarity. This Figure has been taken from [8].

Chapter 3

Two dimensional semi-classical gravity

In the previous chapter, we formulated the black hole information paradox as a theorem. Namely, we stated how the process of black hole evaporation, as predicted by Hawking, eventually results in remnants or mixed states when the black hole has radiated away (almost) completely. When the black hole was formed from a pure state, the second case is in conflict with unitarity, as explained in section 2.2.3. From that observation, Hawking concluded that information is lost within the process of black hole evaporation, which makes the semi-classical approximation break down.

Inspired by Page, we quantified the amount of information loss in terms of fine- and coarse grained entropies in section 2.3. This illustrated that Hawking's computations already yield non-unitary behaviour way before one expects quantum gravitational effects to play a role. This observation suggested that the semi-classical approximation should not necessarily yield problematic behaviour but Hawking's computations do. Namely, what was missing in the original computation of the radiation entropy is classical gravity. To solve this, one needs to introduce a gravitational formula for the fine-grained entropy which includes inside regions, referred to as islands. Recent developments have shown that this formula can be derived when assuming the AdS/CFT duality [15]. Interestingly, as briefly explained in section 2.3.7, one can obtain the same results with the help of semi-classical path integrals, without holography. The corresponding method will be of our main interest in chapter 4. However, to understand these developments, which mainly include the replica trick and Euclidean path integrals, we need to introduce several ingredients.

Firstly, we need to get familiar with gravitational path integrals and their connection to thermal partition functions. After the corresponding discussion in section 3.1, we introduce a very interesting subclass of quantum field theories in section 3.2: *conformal field theories*. These theories will turn out to be extremely useful to describe matter in the context of evaporating black holes. Namely, in section 3.4, we introduce the framework of JT gravity which enables us to couple a gravitational theory to a quantum mechanical matter. Explicitly, we will introduce a semi-classical approximation based on the same ideas as presented in section 1.4.2, but with the use of path integrals. Namely, one can treat matter quantum mechanically while treating gravity, described by the JT action, classically. This can be done by approximating the corresponding path integral by classical saddle-points. In order to find these saddles, one must find the equations of motion resulting from the variation of the JT action. Therefore, a better understanding of these equations and their classical solutions is discussed in section 3.3. Especially, we introduce AdS_2 and corresponding coordinate patches and symmetries. In chapter 4, all these ideas are finally applied to the problem of finding an explicit expression for the gravitational fine-grained entropy of Hawking radiation.

3.1 Gravitational path integrals

In the first chapter, the Bekenstein-Hawking entropy was derived by considering quantum field theory on a fixed, curved spacetime. In this subsection, the entropy of the black hole will be re-derived with the use of gravitational path integrals in a similar way as done by Gibbons and Hawking in [44]. This will yield the same result as in Chapter 1, although we now derive it in an entirely classical way. In order to understand this approach, we first discuss some important notions.

Firstly, the (Euclidean) path integral and its connection to the thermal partition function is discussed. Additionally, we consider the *saddle-point approximation* for gravitational path integrals. Additionally, the *Einstein-Hilbert action* will be introduced and the resulting equations of motion are going to be derived. Finally, all these concepts are used to derive the area law for the entropy of the Schwarzschild black hole, yielding the same result as obtained before (2.28).

Note that this section is highly inspired by [46] and [47]. Readers which are not familiar with path integrals in quantum theories are referred to [36] and [38] since this subsection will only contain a very brief discussion on this subject.

3.1.1 Path integrals in quantum mechanics

After the formulation of what is called the *operator formalism* of quantum mechanics around 1920 [52], Dirac speculated about the importance of the action $I[x(t)]$ in this theory. Namely, he suggested that the propagator could be written as the sum of all possible paths between initial and end position. Based on these ideas, Feynman developed a new formulation of quantum mechanics [52] which generalizes the action principle of classical mechanics [38]. In this formalism, one calculates the probability amplitude that a particle with initial position x_a at an initial time t_a propagates to x_b at a later time t_b, by performing a *path integral* over all possible paths which connects these two point in spacetime. Explicitly, for a one-dimensional theory this is written as,

$$\langle x_b | e^{-iHT} | x_a \rangle = \int Dx(t) e^{iI[x(t)]} \tag{3.1}$$

where $\langle x_b | e^{-iHT} | x_a \rangle$ is the probability amplitude in the canonical operator formalism, H is the Hamiltonian operator and $T = t_b - t_a$. Additionally, $I[x(t)]$ is the action from which one can obtain the equations of motions by the variational principle $\delta I = 0$. Altogether, the right hand side represents the weighted path integral over all possible paths $x(t)$.

By performing a wick rotation $t \to -i\tau$, one finds the Euclidean path integral, which will be of our interest. In contrast, (3.1) is referred to as the Lorentzian path integral.

Interestingly, this formalism can be extended to quantum field theories by taking the continuum limit [38]. In qft, these path integrals represent the transition amplitude corresponding to a field configuration ϕ_a at Euclidean time $\tau = 0$ to ϕ_b at $\tau = \beta$ and is given by

$$\langle \phi_b | e^{-\beta H} | \phi_a \rangle = \int_{\phi(\tau=0)=\phi_a}^{\phi(\tau=\beta)=\phi_b} D[\phi] e^{-I_E[\phi]}, \tag{3.2}$$

where I_E is the Euclidean action and we integrate over all possible field configurations ϕ. The values ϕ_a serve as boundary conditions and specify data at a specific moment of time. What performing the path integral exactly means depends highly on the topology of space [17].

There is an interesting link between path integrals and statistical mechanics. Namely, one can impose periodic boundary conditions, i.e. $\phi_a(\tau = 0) = \phi_b(\tau = \beta)$ such that all fields, which are periodic with period β in imaginary time, are integrated over. When doing so, one can show that the path integral corresponds to

$$\mathcal{Z} = Tr(e^{-\beta H}) = \int_{\phi(\tau=0)=\phi_a}^{\phi(\tau=\beta)=\phi_a} D[\phi] e^{-I_E[\phi]}. \tag{3.3}$$

$\mathcal{Z}$ denotes to the partition function for the canonical ensemble consisting of the field ϕ at temperature $T = \beta^{-1}$. This means that thermodynamical variables can be obtained by performing a path integral over periodic fields. This will be used to derive the area law of the black hole entropy.
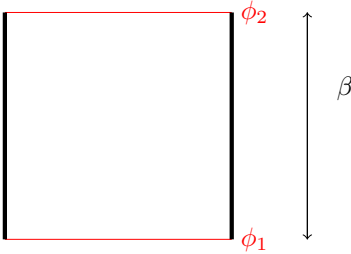
Before moving on to that, we dive deeper into the geometrical interpretation of the Euclidean path integral and establish the connection with the partition function.
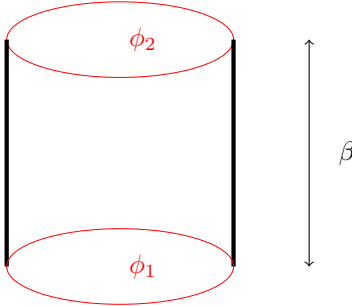
3.1.2 Thermal partition functions

In the remaining sections, we will be particularly interested in Euclidean geometries. These are obtained by performing a Wick rotation from real to imaginary time ($t \to \tau = it$). From now on, we will refer to Wick rotated path integrals as *Euclidean path integrals*. We already saw that these are closely related to the partition function of a thermal system. In this section, we use the relation between these Euclidean integrals and states in quantum field theories to establish this relationship.

We start our discussion, which is highly based on [46], with non-gravitational theories.

Identifying the transition amplitude with the Euclidean path integral, as in (3.2), allows us to interpret a theory in a more geometrical way. To do so, one has to specify the topology of the space where the theory lives. For example, if space is a one-dimensional line, the transition amplitude (3.2) is depicted by

where the red lines denote the boundary conditions that specify the data at $\tau = 0$ and $\tau = \beta$ and the path integral represents the area enclosed by these boundaries. Similarly, if space has the topology of a circle, the corresponding Euclidean path integral is an integral over a cylinder:

Density matrix

We can think of the red circles in the previous equation as spatial slices of the Euclidean manifold and refer to them as *cuts*. Namely, in order to define transition amplitudes we specified data on the cuts $\tau = 0$ and $\tau = \beta$. If instead, data is only specified on one of these cuts, the path integral can be interpreted as a quantum state. For example, the state

$$|\Psi\rangle = e^{-\beta H} |\phi_1\rangle \tag{3.4}$$

corresponds to the path integral

$$|\Psi\rangle = \int_{\phi_1}^{\phi(\beta)=?} D\phi \, e^{S[\phi]} \tag{3.5}$$

where $\phi(\beta) =?$ denotes that ϕ is not specified at $\tau = \beta$. With the same reasoning as before, this yields the following picture

$$|\Psi\rangle = \qquad\qquad\qquad\qquad\qquad \beta$$

,

where the green line represents the open cut. Formally, we can think of this state as a functional which can take a state $\langle\phi_2|$ as input and turn it into a complex number $\langle\phi_2|\Psi\rangle$.

This reasoning can be extended to the case where two cuts remain unspecified. In that case, we obtain an operator which turns two states into a complex number. As a consequence, any Euclidean path integral with two open cuts defines an unnormalized density matrix. For example, the density matrix $\rho \equiv e^{-\beta H}$, for a theory on a circle formally corresponds to

$$\rho \equiv e^{-\beta H} = \qquad\qquad\qquad\qquad\qquad \beta$$

$$(3.6)$$

where the green circles represent the two open cuts.

Remember that the thermal partition function $Z(\beta)$ of a thermal ensemble at temperature $T = \frac{1}{\beta}$ is given by

$$\mathcal{Z}(\beta) = \mathrm{Tr}\rho. \qquad\qquad (3.7)$$

As expected, this means that the computation of the partition function is closely related to the double cut Euclidean path integral depicted above. As a result, computing $\mathcal{Z}(\beta)$ corresponds to taking the trace of ρ which amounts to imposing periodic boundary conditions. This results in gluing together the two open cuts of the Euclidean cylinder and sum over all the possible periodic configurations ϕ_1. As a result, for a theory on a circle, one finds ,

$$\mathcal{Z}(\beta) = \sum_{\phi_1} \langle\phi_1| e^{-\beta H} |\phi_1\rangle = \sum_{\phi_1} \qquad\qquad \beta$$

The thermal partition function for a 2D theory on a circle is equal to the Euclidean path integral on a torus. Similarly, for a theory on a line, one computes the path integral on an infinitely long cylinder of period β.

We thus established the relationship between the thermal partition function and the Euclidean path integral in a geometrical way. It is good to keep these pictures in mind when the replica trick in two dimensional CFT's is introduced in section 3.2.4. Before this is discussed, the gravitational path integral and its semi-classical approximation are studied.

3.1.3 The Euclidean gravitational path integral

Remember that, in quantum field theories, one fixes the spacetime manifold M whereupon all fields ϕ are integrated over. In contrast, the spacetime M is not fixed in a gravitational theory. Therefore, one must integrate over the geometry as well as the quantum fields. The Euclidean gravitational path integral is

$$K = \int D[g]D[\phi]e^{-I_E[g,\phi]} \tag{3.8}$$

where I_E represents the Euclidean action functional corresponding to all the matter fields ϕ living on a spacetime with corresponding metric g.

Again, the interpretation of this path integral depends on the boundary conditions we impose. Inspired by our previous discussion, we define the thermal partition function $\mathcal{Z}(\beta)$ as the path integral on an Euclidean manifold. Specifically, we impose the periodic boundary condition that Euclidean time is periodic

$$\tau \sim \tau + \beta. \tag{3.9}$$

Consequently, Euclidean time corresponds to a circle of radius β and $\mathcal{Z}(\beta)$ can be obtained by integrating over the corresponding geometry [8].

Unfortunately, the measures of the path integral $D[g]$ and $D[\phi]$ are not well defined and we are thus unable to explicitly compute the path integral (3.8). Typically, one approximates this by assuming the integral is dominated by *saddle points* which correspond to classical solutions of the equations of motion. This method is known as the *saddle point approximation* and can be used for integrals of the form

$$Int = \int_{-\infty}^{\infty} e^{-f(x)}dx. \tag{3.10}$$

Namely, the negative exponential in this integral decreases so fast that it is enough to only consider the point where $f(x)$ is minimal [53]. As a result, if this function reaches its minimum value at x_0 one can expand it as Taylor series around x_0 and the integral is approximated as

$$Int \approx e^{f(x_0)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}f''(x_0)(x-x_0)^2+\cdots}dx. \tag{3.11}$$

Applying this on the Euclidean path integral (3.8) thus yields evaluating the exponential at a stationary point of $I_E[g,\phi]$. The corresponding metric $\tilde{g}$ and field $\tilde{\phi}$ are found by solving $\delta I_E[g,\phi] = 0$, i.e. the classical equations of motion. Hence, the thermal partition function, is obtained by considering a classical solution $\tilde{g}, \tilde{\phi}$ and the path integral around it like

$$\mathcal{Z}(\beta) \approx \exp\left\{-I_E[\tilde{g}, \tilde{\phi}] + I_E^{(1)} + ...\right\}. \tag{3.12}$$

The first term in the exponential denotes the Euclidean action evaluated at the *saddle point* of the path integral while the second term and the dots indicate quantum mechanical loop contributions [46]. Ignoring the quantum mechanical contributions and the matter fields, the classical gravitational path integral is simply approximated as

$$\mathcal{Z}(\beta)_{classical} \approx \exp\{-I_E[\tilde{g}]\}. \tag{3.13}$$

Finally, note that the thermal partition function $\mathcal{Z}(\beta)$ can be used to compute the entropy and energy as in standard thermodynamics. Namely,

$$S = (1 - \beta\partial_\beta)\log \mathcal{Z}(\beta) \qquad E = -\partial_\beta \log \mathcal{Z}(\beta). \tag{3.14}$$

We will now apply this on a gravitational theory to obtain the classical Bekenstein-Hawking entropy. In order to do this, we first introduce Einstein gravity using path integrals and derive the equations of motion. Secondly, the generalized entropy is obtained in a totally classical way with the use of (3.13) and (3.14).

3.1.4   Einstein-Hilbert Action

In chapter 1, we gave a very short review of general relativity. Interestingly, this can also be described with the use of path integrals. To see this, consider the 4-dimensional action

$$I_{EH} = \frac{1}{16\pi}\Big[\int_M d^4x\sqrt{-g}R\Big], \tag{3.15}$$

It is shown in appendix D.1 that varying this action with respect to the metric yields the Einstein equations in vacuum

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0. \tag{3.16}$$

However, this derivation assumed that the spacetime of interest is compact (i.e. $\partial M = \emptyset$). Unfortunately this is generally not the case [47]. As a result,

$$\delta g^{\mu\nu}|_{\partial M} = 0, \tag{3.17}$$

does not imply that the derivatives of the metric vanish, which means that $\delta I_{EH} \neq 0$.

   As shown in appendix D.1.3, Gibbons and Hawking (inspired by York [24]) proposed an additional boundary term I_{GH} which can be added to the Einstein-Hilbert action to solve the problem. Namely, this term cancels a boundary term resulting from the variation of R, even if the spacetime under consideration is not compact. The corrected action is

$$I_{\text{Einstein}} = I_{EH} + I_{GH} = \frac{1}{16\pi}\Big[\int_M d^4x\sqrt{-g}R + 2\int_{\partial M} d^3x\epsilon\sqrt{h}K\Big], \tag{3.18}$$

where K is the extrinsic curvature of the boundary and h is the determinant of the induced metric $h_{\mu\nu}$ (defined in (C.14)) on the boundary ∂M [46]. Furthermore, $\epsilon = \pm 1$ which is positive for timelike and negative for spacelike hypersurfaces.

3.1.5   Classical derivation of the Bekenstein-Hawking entropy

From section 3.1.3, we know that one can obtain the partition function of a thermal ensemble at temperature $T = \frac{1}{\beta}$ by computing an Euclidean path integral on an Euclidean cylinder with period β. In this section, this will be done in an entirely classical way with the use of (3.13). To do so, we use the Euclidean version of $I_{Einstein}$, derived in the previous section 3.18. As (3.13) tells us, one can obtain the partition function by evaluating this action for a classical solution $\bar{g}$ of the equations of motion resulting from $\delta I_{Einstein}=0$. In the previous subsection, we saw that this yields the Einstein vacuum equations (3.16) and hence we do not include matter fields ϕ. Luckily, we already know a classical solution of these equations: the Schwarzschild solution (see section 1.2.1). The corresponding metric (1.6) will be used to obtain the generalized entropy to compute the entropy of the Euclidean black hole with the use of path integrals. Notably, this yields the same result (e.g. (3.14)) as derived in chapter 1, although no quantum mechanical effects are now included. The derivation below is based on the approaches used in [[8],[47],[46]].

Euclidean black hole geometry

The geometry corresponding to the Euclidean black hole is obtained from the Schwarzschild metric (1.6) by performing a Wick rotation. This yields

$$ds_E^2 = \left(1 - \frac{2M}{r}\right)dt_E^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2d\Omega^2 \qquad t_E = t_E + \beta. \tag{3.19}$$

Importantly, this Euclidean version of (1.6) is only a solution of the Einstein equations for $r > r_h = 2M$ [24], therefore Euclidean Schwarzschild black holes do not have an interior [8].
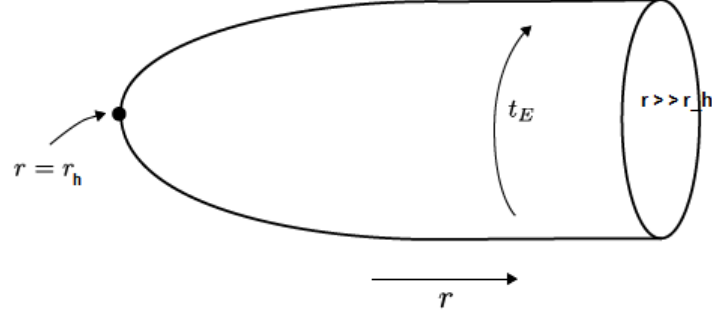
Figure 3.1: The geometry of the Euclidean time and radial directions of the Euclidean Schwarzschild black hole. For obvious reasons this geometry is often referred to as the cigar. Far from the horizon, $r \gg r_h$ one has a flat circle of circumference β which is the inverse temperature as seen by an observer far away. Figure taken from [8].

The Hawking temperature

Remember that, near the horizon $r \sim r_h$ the Schwarzschild metric (1.6) reduces to the Rindler metric (1.29). In a completely analogous way, the Euclidean Schwarzschild metric reduces to the Euclidean Rindler metric at $r \sim r_h$

$$ds_E^2 \approx (\frac{\rho}{2r_h})^2 dt_E^2 + d\rho^2 + r_H^2 d\Omega^2. \tag{3.20}$$

In order do zoom in on the event-horizon, a new coordinate $\tau_E \to 2r_h\tau_E$ can be introduced. With this definition and by ignoring the angular directions for a moment, one obtains

$$ds_E^2 = \rho^2 d\tau_E^2 + d^2\rho, \tag{3.21}$$

which is just the flat Euclidean metric in polar coordinates[1]. Note that, in order to fix the period of τ_E to 2π one needs to fix the inverse temperature to

$$\beta = 4\pi r_h, \tag{3.22}$$

which follows from $\tau_E \to 2r_h\tau_E$. This condition is needed in order to avoid a *conical singularity*. As a consequence, the temperature measured by a distant observer $r \gg r_h$ [8] is

$$T = \frac{1}{4\pi r_h} = \frac{1}{8\pi M} = T_{Haw} \tag{3.23}$$

which is the same temperature as we derived in (2.27). Interestingly, this temperature followed directly from classical arguments while (2.27) was found by considering quantum fields in a curved spacetime.

Note that the resulting geometry of the Euclidean Schwarzschild black hole is often referred to as *cigar* which can be understood by taking a look at Figure 3.1.

The Bekenstein-Hawking entropy

We will now use (3.13) to obtain the classical entropy of the black hole. To do so, one needs to evaluate the full Einstein-Hilbert action (3.18) for the classical Euclidean Schwarzschild solution (3.20).

First, note the Schwarzschild solution satisfies the Einstein vacuum equations and thus $R = 0$. As a result, the first term in (3.18) vanishes and in the only non-vanishing term is the Gibbons-Hawking boundary term

$$I_{\text{Einstein}}\big|_{R=} = \frac{1}{8\pi} \int d^3x \epsilon \sqrt{h} K. \tag{3.24}$$

In order to evaluate this term for the Schwarzschild solution, it is convenient to switch over to Gaussian normal coordinates. As explained in appendix C.1.2, this yields a metric of the form

$$ds_E^2 = d\rho^2 + \left(1 - \frac{2M}{r}\right) dt_E^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2), \tag{3.25}$$

---

[1] This can easily be seen by setting $x^0 = \rho \sinh \tau$ and $x^1 = \rho \cosh \tau$.

where ρ satisfies

$$d\rho = \frac{dr}{\sqrt{1 - \frac{2M}{r}}}. \tag{3.26}$$

This allows us to consider obtain the induced metric on the boundary ∂M, which we take to be $\rho = constant$. On this boundary, the components of the induced metric are

$$h_{ab} = \begin{bmatrix} 1 - \frac{2M}{r} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{bmatrix}. \tag{3.27}$$

Additionally, the corresponding determinant $\det h_{ab} = h$ is simply given by

$$h = r^4 \sin^2 \theta \left(1 - \frac{2M}{r}\right). \tag{3.28}$$

As demonstrated in appendix C.1.2, we can use this determinant to obtain the extrinsic curvature of the hypersurface embedded in M

$$K = -\frac{1}{\sqrt{h}} \frac{\partial}{\partial \rho} \sqrt{h}, \tag{3.29}$$

which corresponds to C.24 by changing the role of ρ by t.

Using the relation between r and ρ this yields the following expression for the extrinsic curvature

$$K = -\frac{1}{2h} \frac{\partial h}{\partial r} \frac{\partial r}{\partial \rho} = -\frac{1}{2r^2}(4r - 6M) \frac{1}{\sqrt{1 - \frac{2M}{r}}}. \tag{3.30}$$

Plugging this into $I_{Einstein}$ one obtains

$$I_{Einstein}[\bar{g}] = \frac{1}{8\pi} \int_{\partial M} d^3x \sin\theta (3M - 2r). \tag{3.31}$$

Unfortunately, this diverges when $r \to \infty$. To solve this, one needs to add a counterterm K_0 which is defined as the extrinsic curvature of the same boundary ∂M but embedded in flat Minkowski spacetime. This is explicitly evaluated by setting $M = 0$ in (3.30), resulting in

$$K_0 = -\frac{1}{2r}. \tag{3.32}$$

The regulated term we need to integrate is thus equal to

$$\sqrt{h}(K - K_0) = r^2 \sin\theta \sqrt{1 - \frac{2M}{r}} \left(-\frac{1}{2r^2}(4r - 6M) \frac{1}{\sqrt{1 - \frac{2M}{r}}} + \frac{2}{r}\right) = \sin\theta \left(-2r + 3M + 2r\sqrt{1 - \frac{2M}{r}}\right). \tag{3.33}$$

Especially, we are interested in the regime where $r >> 2M$, which corresponds to a region far away from the black hole horizon [47]. In such a region, we expect the classical approximation to be valid. In this limit, one can expand the square root by using $(1 - x)^n \approx 1 - nx + ...$, which yields

$$\sqrt{h}(K - K_0) \approx \sin\theta(3M - 2r + 2r(1 - \frac{M}{r})) \approx M\sin\theta. \tag{3.34}$$

The evaluation of the Einstein action for the Euclidean Schwarzschild solution then yields

$$I_{EH} = \frac{1}{8\pi} \int_{\partial M} d^3x \sqrt{h}(K - K_0) \approx \frac{1}{8\pi} \int_0^\beta d\tau \int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi M = \frac{\beta M}{2}. \tag{3.35}$$

Inserting this in (3.13) one obtains

$$\mathcal{Z}(\beta) = \exp\left\{-\frac{\beta M}{2}\right\} = \exp\left\{-\frac{\beta^2}{16\pi}\right\}, \tag{3.36}$$

where we used $\beta = 4\pi r_h = 8\pi M$ to obtain the second equality (see 3.22). Finally, the entropy is computed with the use of (3.14). With the obtained partition function (3.36), this yields

$$S = (1 - \beta\partial_\beta)\log\mathcal{Z}(\beta) = (\log\mathcal{Z}(\beta) + \frac{2\beta^2}{16\pi}) = \frac{\beta^2}{16\pi} = \frac{16\pi^2 r_h^2}{16\pi} = 4\pi M^2 \tag{3.37}$$

where the last equality is obtained by noting that $M^2 = \frac{r_h^2}{4}$. Remember that the area A of the black hole is given by

$$A = 4\pi r_h^2 = 16\pi M^2, \tag{3.38}$$

such that the classical entropy of the Schwarzschild metric can be written as

$$S = \frac{A}{4}. \tag{3.39}$$

This is exactly what we obtained in section 2.1.2, with arguments from quantum field theory [46]. The path integral derivation of the entropy was first performed by Gibbons and Hawking in [44] and a similar method will be used later on for the entropy of the radiation.

Note that we derived this result in an entirely classical way, e.g., we did not include any quantum mechanical contributions. However, to obtain the entropy of the radiation, where the information paradox arises, we will need to do this. Furthermore, one needs to include quantum mechanical effects in order to obtain a possible explanation for the large amount of black hole microstates one expects from the form (3.39). Several attempts have been made in the framework of string theory[24].

Before one actually includes quantum mechanical fields to this gravitational theory, it is important to get familiar with some important ideas from *conformal field theory* (CFT) in 2 dimensions. This will be introduced in the next section. After that, we can introduce the *semi-classical path integral* where we treat gravity classically (as in this section) but treat the coupled matter in a quantum mechanical way.

3.2   Conformal field theories in two dimensions

As our final goal is to describe Hawking radiation from a black hole into flat space, we need to find a way to describe quantum fields which exist on both spacetimes. As will become clear later, *conformal field theories* are perfectly suited for this problem. For this purpose, some relevant concepts of two dimensional CFT are introduced in this section. Additionally, the discussion ends with an illustration of how the entanglement entropy is computed in such a theory. Namely, *the replica trick* is introduced and applied on a single interval configuration. Importantly, we start with a discussion on a classical level (e.g. no quantum effects included) and extend the presented ideas to the quantum mechanical picture at the end.

Note that this section is based on [[46],[54]].

A *conformal transformation* is a change of coordinates $x^\mu \to x'^\mu(x)$ such that the corresponding change in the metric can be written as

$$g_{\alpha\beta}(x^\mu) \to \Omega^2(x^\mu) g_{\alpha\beta}(x^\mu), \tag{3.40}$$

i.e. a transformation which only changes the distances by a scale factor $\Omega^2(x)$.

A theory which is invariant under such a transformation, and thus looks the same at every scale, is called a *conformal field theory* (CFT). This type of theories arise in many different contexts such as string theory, AdS/CFT and statistical mechanics [46]. Furthermore, two dimensional conformal field theories form a special class of CFT's because the number of independent conformal transformations is infinite in this case, while being finite in higher dimensions [54]. In the following discussion, we will only be interested in this two dimensional case and discuss some properties. Furthermore, note that, in this case, any two dimensional gravitational theory which enjoys Weyl[2] and diffeomorphism invariance will reduce to a conformally invariant theory. Finally, we will assume that the background metric $g_{\alpha\beta}$ is fixed such that we can think of the conformal symmetry as a global one with corresponding Noether currents [54].

3.2.1   Conformal transformations

For simplicity, we will work on Euclidean signature worldsheets when analysing some properties of CFT's. However, everything could be done on Lorentzian worldsheets as well.

Consider a field theory on the Euclidean plane $\mathcal{R}^2$ with coordinates x_1 and x_2. It will soon become clear that it is convenient to introduce complex coordinates

$$z = x_1 + ix_2 \qquad \overline{z} = x_1 - ix_2. \tag{3.41}$$

In these coordinates, the Euclidean space has the metric

$$ds^2 = (dx_1)^2 + (dx_2)^2 = dz d\overline{z} \tag{3.42}$$

or

$$g_{\mu\nu} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}. \tag{3.43}$$

In 2 dimensions, conformal transformations in these coordinates have the nice property that they correspond simply to any holomorphic coordinate change. This can be understood by looking at the following form of (3.40)

$$ds^2 = dz d\overline{z} \to e^{\sigma(w,\overline{w})} dw d\overline{w}, \tag{3.44}$$

and investigate which type of transformations

$$z = f(w, \overline{w}) \qquad \overline{z} = \overline{f}(w, \overline{w}), \tag{3.45}$$

transform according to (3.44). Using the transformation properties of the metric for a general coordinate change $x \to x'$

$$g'_{\mu\nu}(x') = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} g_{\sigma\rho}(x) \tag{3.46}$$

---

[2]A theory with metric $g_{\alpha\beta}$ has a Weyl symmetry when it is invariant under the transformation $g_{\alpha\beta} \to e^{-2\omega}(x) g_{ab}$

the metric in w-coordinates looks like

$$ds^2 = \left(\frac{\partial f}{\partial w} \frac{\partial \overline{f}}{\partial w} dw dw + \frac{\partial f}{\partial w} \frac{\partial \overline{f}}{\partial \overline{w}} dw d\overline{w} + \frac{\partial f}{\partial \overline{w}} \frac{\partial \overline{f}}{\partial w} d\overline{w} dw + \frac{\partial f}{\partial \overline{w}} \frac{\partial \overline{f}}{\partial \overline{w}} d\overline{w} d\overline{w} \right). \tag{3.47}$$

As a result, these transformations lead to the form (3.44) with only non-diagonal terms, if

$$\frac{\partial f}{\partial w} \frac{\partial \overline{f}}{\partial w} = \frac{\partial f}{\partial \overline{w}} \frac{\partial \overline{f}}{\partial \overline{w}} = 0. \tag{3.48}$$

This is equivalent to the condition that f is a holomorphic function [3]

$$f = f(w) \qquad \overline{f} = \overline{f}(\overline{w}). \tag{3.49}$$

As a result, conformal transformations are equivalent to holomorphic functions in two dimensions [46]. Simple examples of two dimensional conformal transformations are

- Translations: $z \to z + a$
- Rotations: $z \to \lambda z$ with $|\lambda| = 1$
- Scale transformations (dilatations): $z \to \lambda z$ with $|\lambda| \neq 1$.

If a theory is classically invariant under one of these transformations, we can associate conserved currents to this symmetry (Noether's theorem). Interestingly, the conserved current corresponding to translational invariance is the energy-momentum tensor.

3.2.2 Energy-momentum tensor

Remember that conserved currents are often obtained by promoting a constant ϵ that appears in a symmetry to a function of spacetime coordinates (e.g. promoting a global symmetry to a local one). The corresponding change in the action is then given by

$$\delta I = \int d^2 x J^\alpha \partial_\alpha \epsilon, \tag{3.50}$$

where J^α is some function of the fields. If we want the theory to be invariant under the symmetry, the variation of the action must vanish. From (3.50), it is clear that this is the case when ϵ is constant. However, when the equations of motion are satisfied, we must have $\delta S = 0$ for every variation $\epsilon(x)$. Hence,

$$\partial_\alpha J^\alpha = 0 \tag{3.51}$$

must hold. As a result, J^α is the conserved current associated to the symmetry.

This idea can be applied on translational invariance. Consider a theory which is coupled to a dynamical metric $g_{\alpha\beta}(x)$ (e.g. coupled to gravity). Then the translation transformation

$$x^\alpha \to x^\alpha + \epsilon^\alpha, \tag{3.52}$$

can be viewed as a diffeomorphism. If one makes the corresponding change in the metric

$$\delta g_{\alpha\beta} = \partial_\alpha \epsilon_\beta + \partial_\beta \epsilon_\alpha, \tag{3.53}$$

then the action remains invariant if the change in the action is

$$\delta S = - \int d^2 x \frac{\delta S}{\delta g_{\alpha\beta}} \delta g_{\alpha\beta} = -2 \int d^2 x \frac{\delta S}{\delta g_{\alpha\beta}} \partial_\alpha \epsilon_\beta. \tag{3.54}$$

As a result, the conserved current arising from translational invariance is the energy-momentum tensor

$$T_{\alpha\beta} = -\frac{2}{\sqrt{g}} \frac{\delta S}{\delta g^{\alpha\beta}}. \tag{3.55}$$

---

[3]A complex function $f(z) = u(x,y) + iv(x,y)$ is said to be holomorphic if it is complex differentiable at every point of its domain. This holds if the Wirtinger derivative with respect to the complex conjugate of z vanishes. This corresponds to condition (3.48) [57]

On a flat worldsheet, this is evaluated on $g_{\alpha\beta} = \delta_{\alpha\beta}$ which obeys $\partial^\alpha T_{\alpha\beta} = 0$. Analogously, on a curved spacetime, the conservation equation (3.51) becomes

$$\nabla^\alpha T_{\alpha\beta} = 0. \tag{3.56}$$

In a flat background with metric (3.42), this equation has two components

$$\partial_{\overline{z}} T_{zz} = 0 \qquad \partial_z T_{\overline{zz}} = 0. \tag{3.57}$$

From these equations, we see that T_{zz} is a holomorphic function and $T_{\overline{zz}}$ is anti-holomorphic [46]. It is convenient to refer to them as

$$T_{zz} \equiv T(z) \qquad T_{\overline{zz}} \equiv \overline{T}(\overline{z}), \tag{3.58}$$

to simplify the notation.

T from Traceless

If a theory is also classically invariant under scale transformations, the trace of the stress tensor vanishes. Namely, when we vary the action with respect to a scale transformation

$$\delta g_{\alpha\beta} = \epsilon g_{\alpha\beta}, \tag{3.59}$$

we find

$$\delta S = \int d^2 x \frac{\partial S}{\partial g_{\alpha\beta}} \delta g_{\alpha\beta} = -\frac{1}{2} \int d^2 x \epsilon \sqrt{g} T^{\alpha\beta} g_{\alpha\beta} = -\frac{1}{2} \int d^2 x \epsilon \sqrt{g} T^\alpha_\alpha. \tag{3.60}$$

As this must vanish in a CFT (which is scale invariant), we have that

$$T^\alpha_\alpha = 0. \tag{3.61}$$

This is the key feature of a CFT in general dimensions [54]. However, the scale invariance is often anomalous[4] in Quantum Theory. In two dimensions, this conformal invariance survives but the vanishing trace only holds in flat space, as will be discussed below.

3.2.3 Quantum anomalies

Up to now, we have been introducing all the concepts classically. However, it is important to note that classical conformal invariance does not necessarily imply quantum conformal invariance. When this is not the case, this is referred to as a *quantum anomaly*. As a result, the properties of stress-energy tensor we presented above do not always hold on a quantum mechanical level. Before we dive deeper into this, we will briefly describe CFT as a subset of QFT. We will only quote the relevant results for our discussion but a full introduction of CFT can be found in [54]. From now on, a CFT refers to a conformal invariant QFT.

Quantum fields

Interestingly, a CFT is constructed around correlation functions and the behaviour of operators under conformal transformations. We start by looking at the general transformation rule of what is called a *primary operator*. Under a finite conformal transformation

$$z \rightarrow w(z) \qquad \text{and} \qquad \overline{z} \rightarrow \overline{w}(z) \tag{3.62}$$

these kind of operators transform with a rescaling

$$\mathcal{O}(z,\overline{z}) \rightarrow \tilde{\mathcal{O}}(w,\overline{w}) = \left(\frac{dw}{dz}\right)^{-h} \left(\frac{d\overline{w}}{d\overline{z}}\right)^{-\tilde{h}} \mathcal{O}(z,\overline{z}), \tag{3.63}$$

where h and $\tilde{h}$ are both real numbers referred to as *conformal weights*. Furthermore, their sum

$$\triangle = h + \tilde{h}, \tag{3.64}$$

---

[4]A symmetry is anomalous when it is a classical symmetry of a theory but not a symmetry of the quantum theory based on the same Lagragian.

is defined as the *scaling dimension* [46] of the operator. A consequence of conformal symmetry of the theory is that the correlation functions of these primary operators behave nicely under scale transformations $z \to \lambda z$:

$$\langle \mathcal{O}_1(z_1)\mathcal{O}_2(z_2)\mathcal{O}_3(z_3)...\mathcal{O}_n(z_n)\rangle = \lambda^{\triangle_1 + \triangle_2 + \triangle_3 + ... + \triangle_n}\langle \mathcal{O}_1(\lambda z_1)\mathcal{O}_2(\lambda z_2)\mathcal{O}_3(\lambda z_3)...\mathcal{O}_n(\lambda z_n)\rangle, \quad (3.65)$$

where the subscripts of the scaling dimensions denotes the operator they correspond to.

For a two point function, this implies the following behaviour

$$\langle \mathcal{O}(z_1)\mathcal{O}(z_2)\rangle \propto \frac{1}{|z_1 - z_2|^{2\triangle}}. \tag{3.66}$$

This can be understood from the conformal symmetry as follows. Due to translation invariance, the 2-point function can only be a function of $z_1 - z_2$. Additionally, rotational invariance implies that it can only depend on the modulus of such a vector. To finish with, it must transform like (3.65) under a scale transformation. These three conditions can only be mutually satisfied when the 2 point function has the form (3.66) [58].

Stress tensor transformation

An example of an operator which is not primary is the energy-momentum tensor. Explicitly, this transforms under a finite conformal transformation $z \to w(z)$ as

$$T(w) = \left(\frac{\partial w(z)}{\partial z}\right)^{-2}\left[T(z) + \frac{c}{12}\{w(z), z\}\right], \tag{3.67}$$

where $\{\tilde{z}, z\}$ is the *Schwarzian*, defined for a general transformation $z \to f(z)$ as

$$\text{Sch}(f(z), z) = \{f(z), z\} \equiv \frac{f(z)'''}{f(z)'} - \frac{3}{2}\left(\frac{f(z)''}{f(z)'}\right)^2, \tag{3.68}$$

and the primes denote derivatives with respect to z. This object will often return in our discussion. Furthermore, note that c in (3.67) is the *central charge* which characterizes the number of degrees of freedom of the CFT [54].

This transformation rule can be applied to an interesting example: the transformation from the Euclidean cylinder to the complex plane [56]. Start by parametrizing the cylinder with

$$w = \sigma + i\tau \qquad \sigma \in [0, 2\pi). \tag{3.69}$$

A conformal transformation from this cylinder towards the complex plane is then

$$w \to z = e^{-iw}. \tag{3.70}$$

Geometrically, this transformation corresponds to Figure 3.2. The fact that the cylinder and the plane are directly related by a conformal transformation implies that, if the properties of a CFT on the cylinder are known, they are also known on the plane. Namely, with the use of (3.67) we find that the stress tensor on the cylinder $T(w)$ is related to $T(z)$ by

$$T(w) = -z^2 T(z) + \frac{c}{24}, \tag{3.71}$$

where we used that $\{w(z), z\} = \frac{1}{2z^2}$ and $w'(z) = \frac{i}{z}$.

Weyl Anomaly

We finish the general discussion about two dimensional CFT's by noting that the condition (3.61) does not hold in curved spacetime.

Earlier, we saw that, in the classical case, the conformal symmetry implies that the trace of the stress tensor vanishes (3.61). Unfortunately, this does not hold in a general quantum theory. Namely, in flat space, $\langle T_\alpha^\alpha \rangle = 0$ holds, but in a curved background this is not longer true [54]. One can show that, on a 2-dimensional worldsheet with corresponding Ricci scalar R, the trace of the stress tensor becomes

$$\langle T_\alpha^\alpha \rangle = -\frac{c}{12}R. \tag{3.72}$$

This is a result of the fact that the conformal symmetry is anomalous in curved spacetime [55]. Equation (3.72) is often referred to as the *Weyl anomaly* [54].
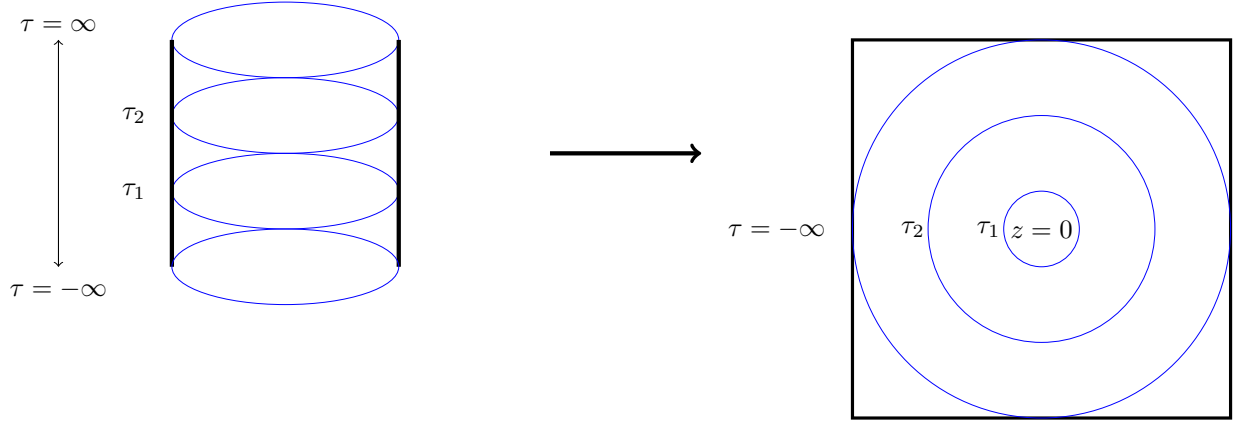
Figure 3.2: Conformal transformation from the cylinder towards the complex plane. Constant time τ slices on the cylinder are denoted in blue. These correspond to circles of fixed radius in the complex plane. The origin $z = 0$ corresponds to the distant past $\tau \to -\infty$. Figure based on Figure on page 19 in [56].

3.2.4   Entanglement entropy in 1+1-dimensional CFT

In the previous subsections we have introduced some basic concepts of conformal field theory and discussed the transformation properties of the energy-momentum tensor and primary operators. We will now apply this, together with the results of section (3.1), to compute the entanglement entropy in a conformal field theory. To do so, we use the *Replica Trick*, which is extended to a gravitational theory in section 4.2.1. Especially, we will focus on the computation of the entropy for a 2-dimensional CFT at zero and finite temperature. We discuss the results presented in [58] and [59].

The replica trick

Consider a lattice quantum field theory in one space and one time dimension. Initially, we consider an infinite line with lattice spacings a which are labelled by a discrete variable x. Additionally, time is considered to be continuous. A complete set of commuting observables (operators) is denoted as $\{\hat{\phi}(x)\}$ while the corresponding eigenvalues and eigenstates are respectively given by $\{\phi(x)\}$ and $|\{\phi(x)\}\rangle$. Remember from our discussion in section 3.1, that the density matrix, belonging to a thermal state at temperature $T = \frac{1}{\beta}$ can generally be expressed as an Euclidean path integral.

For an evolution from state $|\phi_a(x_1)\rangle$ at $\tau = 0$ to $|\phi_b(x_2)\rangle$ at $\tau = \beta$, described by the time-evolution operator $\hat{H}$, the density matrix can be represented by (as seen from equation (3.6))

$$\rho(\langle\phi_b(x_2)|\phi_a(x_1)\rangle) = Z(\beta)^{-1} \langle\phi_b(x_2)| e^{-\beta\hat{H}} |\phi_a(x_1)\rangle$$

$$= Z(\beta)^{-1} \times \quad \text{} \qquad (3.73)$$

The partition function $\mathcal{Z}(\beta)$ serves as a normalization factor, which ensures that $\mathrm{Tr}\rho = 1$, and is given by a cylinder of circumference β. This is obtained by gluing together the 'blue edge' in (3.73) which lies at $\tau = 0$ and $\tau = \beta$. This corresponds to setting $\phi_a(x) = \phi_b(x)$.

We will be interested in computing the reduced density matrix for a subset A which consists of points x in the disjunct intervals $(u_1, v_1), ..., (u_N, v_N)$. In equation formatting,

$$A = [x \in (u_j, v_j) \qquad \text{for} \qquad j \in [0, N]]. \qquad (3.74)$$

The reduced density matrix ρ_A is then computed by gluing together only the points which are not in A (denoted as $\overline{A}$),

$$\rho_A = \text{Tr}_{\overline{A}}\rho$$

$$= \frac{1}{\mathcal{Z}(\beta)} \times \text{Tr} \qquad\qquad\qquad\qquad\qquad (3.75)$$

As only the points $\overline{A}$ are sewed together (represented by the blue segments in (3.75)), the resulting cylinder has open cuts which correspond to the intervals of A on the $\tau = 0$ line[5]. The resulting Path integral representation is depicted on Figure 3.3.
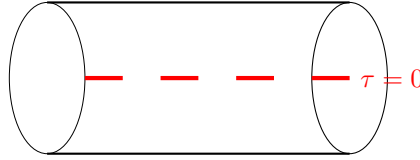
Figure 3.3: Cylinder of circumference β which represent the path integral ρ_A. This is obtained by gluing together the edges $\tau = 0$ and $\tau = \beta$ only at segments which are not part of A. This results in open cuts along $\tau = 0$ denoted by red dots.

Note that, computing the exact density matrix ρ_A for a quantum field theory is a very difficult operational challenge [58]. Therefore we aim to immediately calculate $\text{Tr}\rho_A$. For later purposes, e.g. section 4.2.1, we are more interested in $\text{Tr } \rho_A^n$ for $n \sim 1$. This is obtained by making n *replicas* (labelled by an integer $k \in [0, n]$) of our state (which is depicted on on Figure 3.3). Furthermore sewing these copies together along the open cuts, yields a n-sheeted Riemann surface $\mathcal{R}_n$ as in Figure 3.4. Note that these copies must be sewed together in a cyclic way such that

$$\phi_a(x)_k = \phi_b(x)_{k+1} \qquad \text{for} \qquad x \in A, \qquad\qquad (3.76)$$

---

[5]Generally we have n intervals belonging to A, however, due to practical reasons, we only drew $n = 3$ intervals in (3.75).

Figure 3.4: N-sheeted Riemann surface resulting from gluing the n replicas from our original state in a cyclic way. Note that this Figure has been taken from [59].

and especially $\phi_a(x)_1 = \phi_b(x)_n$ for all $x \in A$. Denoting the partition function on this n-sheeted structure by $Z_n(A)$, the trace of the reduced replica density matrix is given by
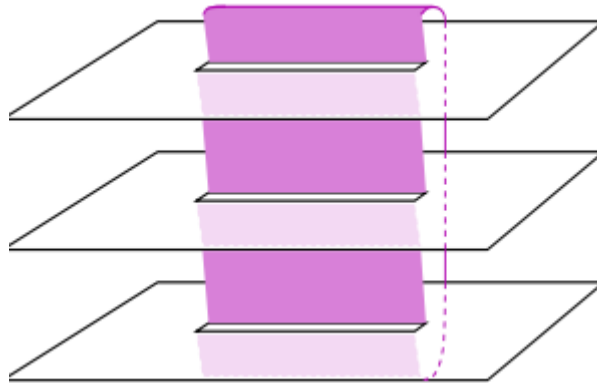
$$\mathrm{Tr}\rho_A^n = \frac{Z_n(A)}{Z^n}. \tag{3.77}$$

As $\mathrm{Tr}\rho_A^n = \sum_\lambda \lambda^n$, where λ are the eigenvalues of the reduced density matrix ρ_A ($\lambda \in [0, 1[$), the left hand is convergent and thus analytic for Re $n > 1$. We are thus allowed to take the derivative with respect to n [58]. As a result, the first derivative of (3.77) is well defined and this is exactly what we are looking for. Namely, the *von Neumann entanglement entropy* is found by evaluating the derivative at $n = 1$

$$S_{vN,A} = -\lim_{n \to 1} \frac{\partial}{\partial n} \frac{Z_n(A)}{Z^n}. \tag{3.78}$$

Remember that this is the entropy we used to obtain the entropy of the Hawking radiation in section 2.2. The computation of this entropy with the help of the replica trick will be of major importance when computing the entropy in the final section 4.2.1.

From n-sheeted surface to the complex plane

In most cases, calculating the partition function on a n-sheeted Riemann surface $\mathcal{R}_n$ is very difficult. Nevertheless, as explained by [59] in more details, the partition function on $\mathcal{R}_n$ can be expressed as a correlation function of another theory on a single sheet. This theory contains n copies of the CFT and lives on the complex plane $\mathbb{C}$, which makes its description much simpler.

As an example, consider the case where the region A, for which we want to calculate the entropy (3.78), only consists of a single interval $[u_1, v_1]$. In this case, $\mathcal{R}_n$ is constructed by sequentially joining the n sheets along the open cuts $x \in [u_1, v_1]$ at $\tau = 0$. The corresponding correlation function on the complex plane $z = x + i\tau$ can be expressed in terms of a type of local *twist fields* at the *branch points* $z = u_1$ and $z = v_1$. In general, such twist fields live in a QFT whenever there is a symmetry which acts the same way everywhere in space and that does not change the position of fields [59]. In our model, this symmetry will be referred to as the *replica symmetry*, e.g. a symmetry under the exchange of the n copies [15]. We will denote the twist fields by $\mathcal{T}_n$ and $\tilde{\mathcal{T}}_n$. These are associated to two opposite cyclic permutations σ and σ^{-1}:

$$\mathcal{T}_n \equiv \mathcal{T}_\sigma : \qquad \sigma : i \to i + 1 \qquad \mathbf{mod}\ n \tag{3.79}$$

$$\tilde{\mathcal{T}}_n \equiv \mathcal{T}_{\sigma^{-1}} : \qquad \sigma^{-1} : i \to i - 1 \qquad \mathbf{mod}\ n. \tag{3.80}$$

As a result, the n-copies of the CFT are still cyclically related as in (3.76), thanks to these twist fields.

In [59], it is shown in more detail that one can generally compute the partition function on $\mathcal{R}_n$ by considering a theory on the complex plane where n copies of the CFT live and twist fields $\mathcal{T}_n$ are present at the branch points $z = u_j$ and $z = v_j$. These fields generate cyclic permutations, defined in (3.80), between the n copies of the CFT. As a result, the condition (3.76) remains valid on the complex plane $\mathbb{C}$.

Concretely, the partition function $Z_{\mathcal{R}_n}$ on the n-sheeted surface, along A (which consists of N intervals), is related to the theory on the complex plane by

$$Z_{\mathcal{R}_n} \propto \langle \mathcal{T}_n(u_1, 0)\tilde{\mathcal{T}}_n(v_1, 0)...\mathcal{T}_n(u_N, 0)\tilde{\mathcal{T}}_n(v_N, 0)\rangle_\mathbb{C}. \tag{3.81}$$

More generally, the identification is extended to correlation functions of the theory on $\mathcal{R}_n$ as

$$\langle \mathcal{O}(x, \tau, \mathrm{sheet}k)...\rangle_{\mathcal{R}_n} = \frac{\langle \mathcal{T}_n(u_1, 0)\tilde{\mathcal{T}}_n(v_1, 0)\mathcal{O}_k(x, \tau)...\rangle_\mathbb{C}}{\langle \mathcal{T}_n(u_1, 0)\tilde{\mathcal{T}}_n(v_1, 0)...\mathcal{T}_n(u_N, 0)\tilde{\mathcal{T}}_n(v_N, 0)\rangle_\mathbb{C}}, \tag{3.82}$$

where $\mathcal{O}(x, \tau, \mathrm{sheet}k)$ denotes an operator of living on sheet k while $\mathcal{O}_k(x, \tau)$ is the corresponding operator on the single sheet theory. Note that the ratio in (3.82) takes into account all the proportionality constants [59].

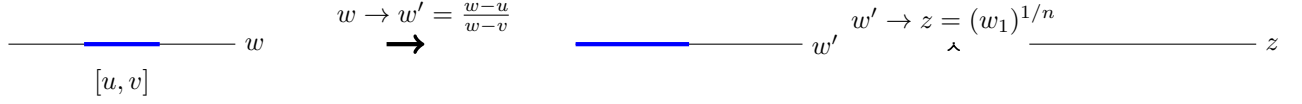$$w \to w' = \frac{w-u}{w-v} \qquad \qquad w' \to z = (w_1)^{1/n}$$

Figure 3.5: In this Figure, we show the effect of the uniformization transformation (3.83) schematically. The first transformation maps the branch points (u, v) to $(0, \infty)$ while the second transformation maps the n-sheeted structure to the complex plane.

Single interval at T=0

To illustrate how the replica trick works in practice, the entropy (3.78) will be computed for a CFT with a central charge c. Furthermore, we will consider A as consisting of a single interval $[u, v]$, setting $N = 1$ in (3.74). Additionally, it is assumed that the temperature is zero $T = 0$ and the field theory has no boundaries.

In order to obtain the entanglement entropy, one needs (3.77). Interestingly, this corresponds to the vacuum expectation value $\langle 0|0 \rangle_{\mathcal{R}_n}$ on $\mathcal{R}_n$. Furthermore, in a CFT, it is enough to know how this transforms under a general conformal transformation [58], e.g. the scaling dimension of the relevant fields. We will obtain this with the help of the stress tensor $T(w)$.

Start by mapping the complex coordinates $w = x + i\tau$ and $\overline{w} = w - i\tau$ of the original theory on $\mathcal{R}_n$ towards the complex plane $\mathbb{C}$. This is done with the *uniformization map* $w \to z(w)$ which is defined as

$$w \to w' = \left(\frac{w-u}{w-v} \right) \to z = \left(\frac{w-u}{w-v} \right)^{\frac{1}{n}}. \tag{3.83}$$

Note that w has been shifted to map the branch points u and v to 0 and ∞ respectively. The power $\frac{1}{n}$ then collects the n-sheets into one sheet corresponding to the full complex plane. This is shown schematically on Figure 3.5. As this coordinate transformation is holomorphic, and thus conformal, we know how the energy-momentum tensor $T(w)$ transforms under the change $w \to z(w)$. Taking the expectation value in (3.67), one finds

$$\langle T(w) \rangle_{\mathcal{R}_n} = \frac{c}{12} \{ z(w), w \} = \frac{c(n^2 - 1)}{24n^2} \frac{(v-u)^2}{(w-u)^2(w-v)^2}, \tag{3.84}$$

where we used that $\langle T(z) \rangle_{\mathbb{C}} = 0$ due to rotational and translation invariance [59].

From (3.82), it follows that the expectation value of $T(w)$ should also correspond to

$$\langle T(w) \rangle_{\mathcal{R}_n} = \frac{\langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) T_i(w) \rangle_{\mathbb{C}}}{\langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) \rangle_{\mathbb{C}}}, \tag{3.85}$$

for all i. Noting that $T(w) = \sum_i T_i(w)$ we can multiply the right hand side of (3.84) with n to obtain

$$\frac{\langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) T(w) \rangle_{\mathbb{C}}}{\langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) \rangle_{\mathbb{C}}} = \frac{c(n^2 - 1)}{24n} \frac{(v-u)^2}{(w-u)^2(w-v)^2}. \tag{3.86}$$

We would like to find an expression for the left hand side, in terms of the scaling dimension $\triangle$ of the primary twist fields. By doing so, we can use (3.86) to obtain $\triangle$ and hence the entanglement entropy. Interestingly, in [56] one finds that the *conformal Ward identity for primary fields* ϕ is given by

$$\langle T(w) \phi_1(z_1, \overline{z}_1)...\phi_N(z_n, \overline{z}_N) \rangle \sim \sum_i^N \left(\frac{h_i}{(w-z_i)^2} + \frac{1}{(w-z_i)} \frac{\partial}{\partial z_i} \right) \langle \phi_1(z_1, \overline{z}_1))...\phi_N(z_n, \overline{z}_N) \rangle,$$

where h_i corresponds to the conformal weights of the primary operators ϕ_i. Applying this to our problem yields

$$\langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) T(w) \rangle_{\mathbb{C}} = \left(\frac{h_{\mathcal{T}_n}}{(w-u)^2} + \frac{1}{(w-u)} \frac{\partial}{\partial u} + \frac{h_{\mathcal{T}_n}}{(w-v)^2} + \frac{1}{(w-v)} \frac{\partial}{\partial v} \right) \langle \mathcal{T}_n(u, 0) \tilde{\mathcal{T}}_n(v, 0) \rangle_{\mathbb{C}}, \tag{3.87}$$

where we assumed that $h_{\mathcal{T}_n} = h_{\overline{\mathcal{T}}_n}$.

Remember from the discussion about primary operators (equation (3.66)) that

$$\langle \mathcal{T}_n(u,0)\tilde{\mathcal{T}}_n(v,0)\rangle_{\mathbb{C}} \sim \frac{1}{|u-v|^{2\triangle}} = \frac{1}{(u-v)^{2\triangle_n}(\overline{u}-\overline{v})^{2\triangle_n}}, \tag{3.88}$$

with $\triangle = \sum h_i$. Using this in (3.87), one finds that

$$\langle \mathcal{T}_n(u,0)\tilde{\mathcal{T}}_n(v,0)T(w)\rangle_{\mathbb{C}} = \left(\frac{h_{\mathcal{T}_n}}{(w-u)^2} + \frac{1}{(w-u)}\frac{\partial}{\partial u} + \frac{h_{\mathcal{T}_n}}{(w-v)^2} + \frac{1}{(w-v)}\frac{\partial}{\partial v} \right) \frac{1}{(u-v)^{2\triangle_n}(\overline{u}-\overline{v})^{2\triangle_n}}$$

$$= \frac{2(u-v)^2 \triangle_n}{(u-v)^{2\triangle_n}(\overline{u}-\overline{v})^{2\triangle_n}(w-u)^2(w-v)^2}. \tag{3.89}$$

where we used $2\triangle_n = 4h_{\mathcal{T}_n}$. From this one easily finds that

$$\langle \mathcal{T}_n(u,0)\tilde{\mathcal{T}}_n(v,0)T(w)\rangle_{\mathbb{C}} = \frac{2(u-v)^2 \triangle_n}{(w-u)^2(w-v)^2}. \tag{3.90}$$

Finally, comparing this to (3.86) yields the scaling dimension

$$\triangle_n = \frac{c}{12}\left(n - \frac{1}{n} \right). \tag{3.91}$$

As stated before, these equations determine all the properties of the CFT under conformal transformations. Therefore, we conclude that $\mathrm{Tr}\rho_A^n$ behaves like a two point function of a primary operator with dimension $\triangle_n$ under these transformations [6]. As a result, one obtains

$$\mathrm{Tr}\rho_A^n = c_n \left(\frac{(v-u)}{a} \right)^{\frac{-c(1-1/n)}{6}}, \tag{3.92}$$

where a corresponds to the renormalization constant $Z(\beta)$ and makes the result dimensionless. Furthermore, the c_n's are proportionality constants which can not be obtained with this method. However $c_1 = 1$ must hold [59].

The entanglement entropy is obtained by inserting the obtained expression in equation (3.78). Doing so, one obtains

$$S_A = \left(-\frac{\partial}{\partial n}\left(c\frac{(v-u)}{a} \right)^{\frac{-c(1-1/n)}{6}} \right)\Big|_{n=1}$$

$$= \left[\left(c_n\frac{(v-u)}{a} \right)^{\frac{-c(1-1/n)}{6}} \ln\left(\frac{(v-u)}{a} \right)\left(\frac{-c(1-1/n)}{6} \right) \right]\Big|_{n=1} \tag{3.93}$$

$$= \frac{c}{3}\ln\left(\frac{(v-u)}{a} \right) + c'_1$$

where the non-universal constant $c'_1 \equiv \frac{\log c_n}{1-n}$ [59].

Single interval at finite temperature

This formula can be generalized to the case where the temperature is finite by using the transformation properties of $\mathrm{Tr}\rho_A^n$ under conformal transformations. Namely, we can use (3.91) to map our theory to a geometry which corresponds to a thermal state. To do so, note that two point functions are related by a conformal mapping $z \rightarrow w(z)$ as

$$\langle \mathcal{T}_n(z_1,\overline{z}_1)\tilde{\mathcal{T}}_n(z_2,\overline{z}_2)\rangle = |w'(z_1)w'(z_2)|^{\triangle_n}\langle \mathcal{T}_n(w_1,\overline{w}_1)\tilde{\mathcal{T}}_n(w_1,\overline{w}_1)\rangle. \tag{3.94}$$

---

[6]Note that this holds up to a possible constant factor. This is a result of the fact that the conformal ward identity contains proportionalities and no strict equality's although we assumed them in our calculations.

Remember that a thermal state with $T = \beta^{-1}$ can be represented by the Euclidean cylinder. Therefore, we map (3.92) to this geometry. This can be done by inverting (3.70), yielding the map

$$z = \frac{\beta}{2\pi} \ln w. \tag{3.95}$$

Using (3.92) and (3.94) one finds that

$$
\begin{aligned}
\text{Tr}\rho_{A,\text{cylinder}}^n &= |z'(w_1)z'(w_2)|^{-\triangle_n} \text{Tr}\rho_{A,\text{plane}}^n \\
&= \left(\frac{2\pi w_1}{\beta}\right)^{\triangle_n} \left(\frac{2\pi w_2}{\beta}\right)^{\triangle_n} c_n \left(\frac{|w_2 - w_1|}{a}\right)^{-2\triangle_n} \\
&= c_n \left[\frac{\beta}{\pi a} \sinh\left(\frac{\pi(z_2 - z_1)}{\beta}\right)^{-2\triangle_n}\right]
\end{aligned}
\tag{3.96}
$$

where we used (3.70) and $e^a e^b (e^a - e^b)^{-2} = \frac{1}{4}\cosh^2(\frac{a-b}{2})$ to obtain the last line.

As we now know the trace of the reduced density matrix in a thermal mixed state at temperature β^{-1} we can compute the entanglement entropy [58] by using (3.78)

$$
\begin{aligned}
S_{A,\text{finite T}} &= \frac{c}{3} \ln\left[\frac{\beta}{\pi a} \sinh\left(\frac{\pi(z_2 - z_1)}{\beta}\right)\right] + c_1' \\
&= \frac{c}{3} \ln\left[\frac{\beta}{\pi a} \sinh\left(\frac{\pi L}{\beta}\right)\right] + c_1'
\end{aligned}
\tag{3.97}
$$

where we used L to denote the length of the interval A[59].

In this section, a formula for the entanglement entropy of a single interval at finite temperature T in a two dimensional CFT has been obtained. This was achieved with the use of the replica method. This expression for the von Neumann entropy will be of major importance in our discussion about *quantum extremal surfaces* and *replica wormholes* in chapter 4. Before moving on to that, we need to briefly introduce *Anti-de sitter spacetime* and its different coordinate patches in section 3.3. After that, we finally study *JT-gravity* in section 3.4, which will give us a way to combine conformal matter and gravitational dynamics.

3.3   AdS$_2$ and its symmetries

The spacetime which will be of our main interest in the following sections is AdS_2. Therefore, the goal of this section is to get familiar with the different coordinate patches and geometric properties of this particular spacetime. Note that section is based on [46] and [62], [63].

3.3.1   Anti-de Sitter space in a general dimension

In section 3.1.4, we defined the Einstein Hilbert action and derived the resulting vacuum Einstein equations. Interestingly, when a negative cosmological constant is added to these Einstein equations, one finds that the maximally symmetric solution is Anti-de Sitter space [61]. This spacetime is characterized by a constant negative curvature [46]. In d dimensions, the corresponding AdS_d metric in global coordinates is given by

$$ds^2 = l^2(-\cosh^2\rho\, dt^2 + d\rho^2 + \sinh^2\rho\, d\Omega_{d-2}^2), \tag{3.98}$$

where l^2 is the curvature radius of the spacetime [64] and $t \in]-\infty,\infty[$ while ρ runs over $[0,\infty[$.

In order to find the corresponding Penrose diagram, one can simply extract a factor $\cosh^2\rho$ and define

$$\sigma = 2\tan^{-1}\tanh\left(\frac{\rho}{2}\right) \qquad d\sigma = \frac{d\rho}{\cosh\rho} \qquad \rho \in [0,\infty[\qquad => \qquad \sigma \in [0,\frac{\pi}{2}[. \tag{3.99}$$

By doing so, each value of t,σ yields a sphere S^{d-2}[63]. As a result, the Penrose diagram looks like a solid cylinder where σ is the radial coordinate and t,Ω describe its surface [46].

Other coordinates which are often used, cover only a patch of the full AdS space. For example, the metric of the *Poincaré patch* is

$$ds^2 = \frac{l^2}{z^2}(-dT^2 + dz^2 + d\vec{x}^2), \tag{3.100}$$

with $\vec{x} = (x^1,...,x^{d-1})$[46]. For example, in two dimensions, this is obtained by the coordinate transformation

$$\gamma = \cosh\rho\cos t + \sinh\rho \qquad z = \frac{1}{\gamma}$$

$$T = \frac{\cosh\rho\sin t}{\cosh^2\rho\sin^2 t - 1}(\sinh\rho - \cosh\rho\cos t). \tag{3.101}$$

Note that T covers only a finite range in t because $T \to \pm\infty$ is only achieved by setting the denominator in (3.101) equal to zero [65]. Next, note that the boundary in these coordinates is at $z = 0$. The patch covered by these Poincare coordinates (T,z) corresponds to the orange region on Figure 3.6.

3.3.2   AdS$_2$ coordinate systems

As was the case in our discussion of conformal field theories, we will only be interested in the 2-dimensional version of Anti-de Sitter space, denoted as AdS$_2$. We will also put the radius l to the unit value $l = 1$ in the metrics.

Interestingly, in two dimensions, the Poincaré patch (3.100), arises naturally as the near horizon limit of the extremal black holes. Explicitly, consider the 4 dimensional Reissner-Nordstrom black hole from section 1.2.2

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2 \qquad f(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2}. \tag{3.102}$$

In the *extremal limit* $M = Q$, we have that $f(r) = (1 - Q/r)^2$. As a result, the inner and outer horizons coincide. Defining

$$r = Q(1 + \lambda/z) \qquad t = \frac{QT}{\lambda} \tag{3.103}$$

and taking the limit $\lambda \to 0$ while keeping the other parameters fixed, one obtains

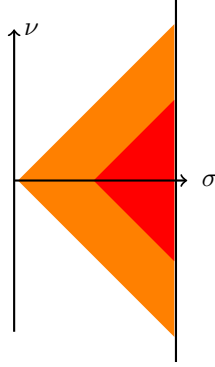$$ds^2 = \frac{Q^2}{z^2}(-dT^2 + dz^2) + Q^2 d\Omega_2^2 \tag{3.104}$$

Figure 3.6: Patches of AdS_2 in the Lorentzian signature. The Poincaré patch is denoted in orange while the black hole patch is shown in red. This Figure is based on Figure 1(b) in [63].

which is $AdS_2 \times S^2$. However, we will only be interested in the two dimensional metric

$$ds^2 = \frac{1}{z^2}(-dT^2 + dz^2), \tag{3.105}$$

where we have put $Q = 1$. As explained below equation (3.101), these coordinates covers only a patch of AdS₂ spacetime. Namely, the global spacetime consists of many Poincaré patches such that each patch gives a near-horizon region of a different patch of the global Reissner Nordstrom spacetime [46].

One can introduce lightcone coordinates to write the metric in another form

$$U = T + z \qquad V = T - z \tag{3.106}$$

$$ds^2 = -\frac{4dUdV}{(U-V)^2}. \tag{3.107}$$

This allows us to describe another interesting classical frame: *the Black Hole frame*. Explicitly, this is obtained by performing the coordinate change

$$U(u) = \frac{\beta}{\pi}\tanh\left(\frac{\pi}{\beta}u\right) \qquad V(v) = \frac{\beta}{\pi}\tanh\left(\frac{\beta}{\pi}v\right). \tag{3.108}$$

This yields the metric

$$ds^2 = -\frac{\pi^2}{\beta^2}\frac{4}{\sinh^2(\frac{\pi}{\beta}(u-v))}dudv \tag{3.109}$$

which describes the red coloured patch of global AdS₂ depicted on Figure 3.6. Similarly to (3.100), this patch is found in the near horizon regime of the *near*-extremal black hole [64].

Euclidean signature

Up to now, the metrics were introduced in the Lorentzian signature. However, we will mostly be interested in Euclidean AdS₂. After a Wick rotation of the Poincaré coordinates $T \to \tau = iT$ in (3.100) the Euclidean signature metric is

$$ds^2 = \frac{d\tau^2 + dz^2}{z^2}. \tag{3.110}$$

In this signature, AdS_2 is just the hyperbolic disk where orbits of τ are curves that touch the boundary at $\tau = \pm\infty$. Another set of natural coordinates in this signature are the coordinates of the Rindler patch [62], where the metric looks like

$$ds^2 = d\rho^2 + \sinh^2\rho d\hat{\tau}^2. \tag{3.111}$$

In these coordinates, the orbits of $\hat{\tau}$ translations look like circles on the hyperbolic disk. In this case, both coordinate choices cover all of hyperbolic space in contrary to the Lorentzian case where they cover different regions of the global space. This is shown on Figure 3.7.
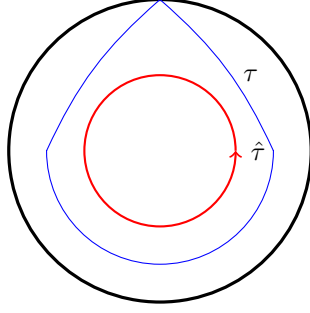
Figure 3.7: Hyperbolic space or Euclidean AdS_2. The red circle denotes an orbit in Rindler coordinate $\hat{\tau}$, while the blue teardrop represents τ. Figure based on Figure 1 in [62].
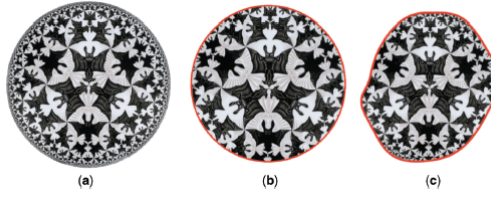

Figure 3.8: Figure a represents the full Euclidean AdS$_2$ space. On Figures b and c the full space has been cut out at the location of two different boundary curves (i.e. not related by a SL(2) transformation). If we rotate or displace such a cutout geometry by an SL(2), it remains unchanged. Figure taken from [62].

Interestingly, in two dimensions, the only curvature function that exist is the Ricci scalar [64]. As a result AdS$_2$ arises as a solution of JT gravity which we will define in the following section. Namely, the metric of this theory is fixed to AdS$_2$ because it is the only 2-dimensional solution with Ricci scalar which obeys $R = -2$ (i.e. has a negative constant curvature) [64]. This will be discussed in much more detail in the following section 3.4.

3.3.3 Symmetries and the Schwarzian

In order to study the symmetries of Euclidean AdS_2, we interpret this spacetime as a low energy limit of a well defined UV theory and follow the approach in [[62], [63]]. To do so, we fix Poincaré coordinates on the hyperbolic disk and imagine that the UV theory has a parameter u which corresponds to time on its boundary. Furthermore, we cut off the spacetime along a trajectory given by $(\tau(u), z(u))$ and fix the proper length of the boundary curve as

$$g_{\text{bdy}} = \frac{1}{\epsilon^2}. \tag{3.112}$$

We take ϵ to be small, which will be referred to as the *Schwarzian limit*. Taking (3.110) as our metric, (3.112) implies that the parametrization satisfies

$$g_{bdy} = \frac{1}{\epsilon^2} = \frac{\tau'^2 + z'^2}{z^2} \qquad \rightarrow z = \epsilon\tau' + O(\epsilon^3), \tag{3.113}$$

where the primes denote derivatives with respect to u.

Interestingly, one can cut off the spacetime in many different ways. Namely, for every arbitrary $\tau(u)$, we can find a $z(u)$ which obeys condition (3.113). This leads to a family of different solutions given by the $\tau(u)$ corresponding to different cutout shapes of AdS_2 as shown in Figure 3.8. Though the interior space is locally the same for different $\tau(u)$'s, the physics generally depends on the full cutout shape determined by $\tau(u)$. As an example, correlation functions of matter fields depend on these cut out shapes. Therefore, transformations of $\tau(u)$ generally alter the physics. However, overall translations and rotations of the whole shape in hyperbolic space leave the physics invariant. These global symmetry operations are generated by the SL(2) group and thus correspond to a symmetry

group of AdS$_2$. Explicitly, these act like

$$\tau(u) \rightarrow \tau'(u) = \frac{a\tau(u) + b}{c\tau(u) + d} \qquad \text{where} \qquad ad - cd = 1 \tag{3.114}$$

such that $\tau(u)$ and $\tau'(u)$ produce exactly the same cutout shape. This can be seen by considering the Poincaré metric (3.110) in complex coordinates $w = \tau + iz$ and notice that this is invariant under the Möbius transformation $w \rightarrow \frac{aw+b}{cw+d}$. However, it must be noted that the Einstein-Hilbert part of the gravitational action, is invariant under full reparametrizations of the boundary time $\tau \rightarrow f(\tau)$. This is an illustration of asymptotic full reparametrization symmetry of AdS$_2$. As we saw that, in the Schwarzian limit, AdS$_2$ only has a SL(2) symmetry, this means that the full reparametrization symmetry has been spontaneously broken. Explicitly, one can think of the $\tau(u)$'s as Goldstone bosons, which are called *boundary gravitons* in the Euclidean signature. These boundary gravitons break the full reparametrization symmetry down to SL(2) in the Schwarzian limit [62].

Schwarzian theory

Interestingly, this symmetry breaking has implications on the gravitational theory which will be of main interest in section 3.4. Namely, it will be shown that the AdS$_2$ metric is a classical solution of *Jackiw–Teitelboim* (JT) gravity. In this context, we will encounter a boundary contribution to the JT action of the form

$$I_{\text{bdy}} = \frac{1}{8\pi} \int_{\partial M} \phi_b K, \tag{3.115}$$

where K is the extrinsic curvature and ϕ_b is the boundary value of the dilaton, as explained later on. Using (3.113) and setting

$$\phi_b = \frac{\phi_r}{\epsilon}, \tag{3.116}$$

in order to let the dilaton diverge at the boundary, one obtains the action

$$I_{\text{bdy}} = \frac{1}{8\pi} \int_{\partial M} \frac{du}{\epsilon} \frac{\phi_r(u)}{\epsilon} K. \tag{3.117}$$

In appendix D.5, it is shown that this extrinsic curvature corresponds to

$$K = \frac{\tau'(\tau'^2 + z'^2 + z'z'') - zz'\tau''}{(\tau'^2 + z'^2)^{2/3}}. \tag{3.118}$$

With the use of equation (3.113), this can be rewritten in terms of the Schwarzian derivative, which was defined in equation (3.68). This yields,

$$K = 1 + \{\tau(u), u\}\epsilon^2 \tag{3.119}$$

such that the action becomes

$$I_{\text{bdy}} = \frac{1}{8\pi} \int du \phi_r \left(\{\tau(u), u\} + \frac{1}{\epsilon^2} \right). \tag{3.120}$$

As the other terms of the JT gravity action only yield topological contributions, this action should be responsible for breaking down the reparametrization symmetry down to $SL(2)$. Therefore, the form of (3.120) is not surprising. Namely, one expects the breaking to be local along the boundary and proportional to the dilaton ϕ_r. Furthermore, such a term has to involve the boundary graviton. An expression which is invariant under $SL(2)$ and satisfies all the mentioned requirements is the Schwarzian of the boundary graviton [63], hence (3.120) is not surprising.

To end with, note that Schwarzian theory described above makes an appearance in a lot of contexts, most notably in the SYK model and quantum mechanical models [55].

3.4   JT gravity

In section 3.1, we introduced Euclidean path integral and discussed how the saddle point approximation can be used to obtain the classical thermodynamical entropy for a Schwarzschild black hole. Additionally, the notion of conformal field theories was introduced in section 3.2. In this section, we would like to combine the results discussed in these sections and describe a gravitational theory which is coupled to a conformal field theory. Namely, the framework of *Jackiw-Teitelboim gravity*, which is able to describe the dynamics of a black hole in AdS_2, will be introduced. Interestingly, this model can incorporate quantum mechanical effects while treating gravity as in section 3.1. This approach is what we call the *semi-classical* approximation and will be discussed below.

Note that this section is largely based on [55] and [62] and we will always work in 2-dimensions from now on.

A logical starting point to describe gravitational dynamics with path integrals is the Einstein action (3.18). However, note that I_{Einstein} is purely topological in two dimensions. Indeed, (3.16) sets the matter stress-energy to zero $T_{\mu\nu} = 0$ (classically). Therefore, the theory described by this action can not be coupled to matter. To solve this, one must add another term to the action. A simple possibility is to introduce a scalar *dilaton* field ϕ and add the following action

$$I_{JT} = \frac{1}{16\pi} \int_M \sqrt{-g}\phi(R+2) + \frac{1}{8\pi}\int_{\partial M} \phi_b K. \tag{3.121}$$

The form of this action might not seem obvious, but it arises naturally when studying near-extremal (low temperature) black holes in higher dimensions [15]. The theory described by this action is referred to as *Jackiw–Teitelboim* (JT) gravity. In such theories, the Einstein term (3.18) is purely topological and proportional to the area of the event horizon of the black hole, while ϕ describes deviations from the extremal value of this area [62]. As mentioned before (i.e. in section 3.3.3), the last term including the extrinsic curvature has the effect of breaking a symmetry and encodes the dynamics. Furthermore, one can think of this last term as playing a similar role for the JT action as the Gibbons Hawking term introduced in section 3.1.4 for I.

3.4.1   Gravitational equations of motion

We start with examining classical solutions of this theory described by (3.121) and ignore the matter part I_{CFT} for a while.

Start by varying I_{JT} with respect to the dilaton. As this only depends linearly on this field, the resulting equation simply yields

$$R = -2. \tag{3.122}$$

The solutions of JT-gravity thus have negative constant curvature. Since the scalar curvature completely fixes the geometry in two dimensions (up to topology), this condition fixes the metric to be AdS_2 [64].

In addition to the condition (3.122), classical solutions of I_{JT}, also need to obey the condition

$$\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\nabla^2\phi + g_{\mu\nu}\phi = 0. \tag{3.123}$$

The derivation of this condition, which consists of varying the action I_{JT} with respect to the metric, is written down in appendix D.2.
As we know from (3.122) that the metric is fixed to AdS_2, we can solve this equation by using (3.110). This yields the general solution [62]

$$\phi = \frac{\alpha + \gamma\tau + \delta(\tau^2 + z^2)}{z^2}. \tag{3.124}$$

As in section 3.3.3, one can interpret this AdS_2 geometry as a low energy limit of a well defined UV theory. Furthermore, again imposing the boundary condition $\phi_b = \frac{\phi_r}{\epsilon}$ (3.116) and combining this with (3.123), one obtains the complete shape of the reparametrization $\tau(u)$ by inserting (3.113). This yields

$$\frac{\alpha + \gamma\tau(u) + \delta\tau(u)^2}{\tau'(u)} = \phi_r. \tag{3.125}$$

Interestingly, we could have found this condition in a much simpler way. Namely, if we plug in the condition (3.122) into the JT action (3.121), we are only left with the boundary term

$$I_{JT} = \frac{1}{8\pi} \int \frac{du}{\epsilon} \frac{\phi_r}{\epsilon} K, \tag{3.126}$$

where we used the boundary condition of the dilaton and $g_{bdy} = \frac{1}{\epsilon}$. Using the fact that the curvature is given by (3.119) in this setup, this yields the action

$$I_{JT} = \frac{1}{2\pi} \int du \phi_r \left(\{\tau(u), u\} + \frac{1}{\epsilon^2} \right). \tag{3.127}$$

Varying this action with respect to to the boundary graviton $\tau(u)$, one obtains

$$\left[\frac{1}{\tau'} \left(\frac{(\tau' \phi_r)'}{\tau'} \right) \right]' = 0, \tag{3.128}$$

which can be integrated to obtain exactly (3.125). As a result, the variation of the boundary JT action, with respect to the boundary graviton, captures the same information as varying the whole action with respect to the metric. From this observation, we conclude that in the JT gravity model, the Schwarzian $\{\tau(u), u\}$ captures the gravitational dynamics.

Note that this idea will be used in our discussion in section 4.2 about replica wormholes, where the boundary graviton arises in the context of gluing flat and curved space. In that setup, the geometry will be a bit more complicated yielding a slightly different action than (3.127), but the general idea remains the same.

3.4.2 Semi-classical approximation

Remember that in section 1.4, we described the "niceness conditions", which explain in which regime the semi-classical limit is expected to yield a good approximation. In this limit, we are allowed to define quantum field theories on a fixed curved spacetime. As a result, one can treat matter quantum mechanically while using general relativity. In this subsection, we would like to introduce a similar approximation with the use of path integrals. This will enable us to treat gravity in a dynamical but classical way while describing matter by a conformal field theory.

For later purposes (i.e. section 4.2), we would like to describe a CFT matter theory which lives on the spacetime described by the metric g (which is a classical solution of the theory described by $I_{\text{Einstein}} + I_{JT}$). Therefore, the total action we will study is given by

$$
\begin{aligned}
I_{\text{full}} &= I_{\text{Einstein}} + I_{\text{JT}} + I_{\text{CFT}} \\
&= \frac{\phi_0}{16\pi} \left[\int_M \sqrt{-g} R + 2 \int_{\partial M} \epsilon \sqrt{h} K \right] + \frac{1}{16\pi} \left[\int_M \sqrt{-g} \phi(R+2) + 2 \int_{\partial M} \phi_b K \right] + I_{CFT}[g, \chi],
\end{aligned}
\tag{3.129}
$$

where I_{CFT} represents the action of our matter theory and ϕ_0 is the area of the extremal black hole [62]. In this setup, the Einstein term is purely topological and only serves to give the extremal entropy, as will become clear later on [62]. Note that this matter does not couple directly to the dilaton such that I_{CFT} only depends on the quantum fields χ and the metric g.

Ideally, we would like to compute the path integral over the gravitational variables g, the dilaton ϕ and the matter fields X. Unfortunately, as already mentioned, doing so is very difficult. Instead, we look for saddle-points and evaluate the path integral as described in (3.12). However, these saddle-points consist of classical solutions of (3.129). As we want to include quantum mechanical effects, this is not sufficient. Luckily, this problem is solved by using an intermediate, *semi-classical* approach, where we treat matter quantum mechanically but avoid computing the full gravitational path integral by looking for classical saddle-points.

In practice, this consists of obtaining the *effective action* by integrating out the matter fields χ from I_{CFT}. In other words,

$$e^{-I_{EFT}[g]} = \int D\chi e^{-I_{CFT}[g,\chi]}, \tag{3.130}$$

which results in an action $I_{EFT}[g]$ that depends only on the metric. The next step in our semi-classical approach is to vary I_{EFT} and the gravitational action with respect to the metric g, yielding the general equation of motion

$$\delta I_{\text{Einstein}} + \delta I_{JT} + \delta I_{EFT} = 0. \tag{3.131}$$

While the variation of the two first terms yields (3.123), the variation of I_{EFT} with respect to the metric, is simply the stress tensor. This holds by definition (3.55). As a result, the complete semi-classical equation of motion is

$$\nabla_\mu \nabla_\nu \phi - g_{\mu\nu} \nabla^2 \phi + g_{\mu\nu}\phi + \langle T_{\mu\nu} \rangle = 0, \tag{3.132}$$

where the expectation value $\langle T_{\mu\nu} \rangle$ encodes the quantum mechanically treated matter.

Conservation equations

We take a closer look at the stress-energy tensor appearing in (3.132). To do so, we use a general two-dimensional metric in terms of lightcone coordinates u, v

$$ds^2 = -e^{2\omega(u,v)}dudv. \tag{3.133}$$

Note that, in two dimensions one can always write the metric in this way by a suitable choice of coordinates. This is equivalent to working in the *conformal gauge* [64].

In this framework, the Weyl anomaly (3.72) looks like

$$T_{uv} = -\frac{c}{12\pi}\partial_u \partial_v \omega(u,v), \tag{3.134}$$

which easily obtained by calculating the curvature explicitly. In appendix D.3, it is shown that the corresponding conservation equations (3.56) have the following solutions

$$T_{uu} = \frac{c}{12\pi}[\partial_u^2 \omega - (\partial_u \omega)^2] + F_u(u) \tag{3.135}$$

$$T_{vv} = \frac{c}{12\pi}[\partial_v^2 \omega - (\partial_v \omega)^2] + F_v(v), \tag{3.136}$$

where the $F_v(v)$ and $F_u(u)$ are functions of only u or v respectively. These functions can be seen as integration constants (as $\partial_u F_v(v) = 0$ these can be added without altering the conservation equations).

In AdS spacetime, the brackets vanish at large r. Consequently, the F's tell us the asymptotic value of the stress-energy. Namely, $F_u(u)$ corresponds to the asymptotic outgoing energy flux while $F_v(v)$ corresponds to the asymptotic incoming energy flux [55].

Finally, note that in quantum theory, one works with correlation functions and needs to take the expectation value of (3.136). The results obtained above are applicable on $\langle T_{ab} \rangle$ in any state [55].

3.4.3   Black hole evaporation

The semi-classical approximation of the model described by (3.129) can be used to describe the dynamics of a black hole in AdS$_2$. In order to enable the black hole under consideration to evaporate, it is coupled to an auxiliary system. This will function as a *reservoir system* which absorbs the radiation coming from the black hole. It is convenient to choose the reservoir such that it corresponds to a fixed flat spacetime where the same CFT lives. Namely, the reservoir can be described by half of Minkowski spacetime with metric

$$ds^2 = -dudv \qquad v > u, \tag{3.137}$$

which is then glued to the AdS$_2$ boundary along the boundary at $u = v = t$. The full system, consisting of the black hole and the reservoir, is thus described by the metric

$$ds^2 = -e^{2\omega(u,v)}dudv$$

$$\text{Reservoir:} \quad v > u \quad \text{static background} \quad \omega = 0 \tag{3.138}$$

$$\text{Black hole:} \quad u > v \quad \text{dynamical background} \quad \omega \neq 0.$$

However, this metric is not continuous at the boundary $u = v$. This is where the conformal invariance of the matter theory becomes essential. Thanks to this symmetry, we are free to define the matter theory on a smooth flat metric which is related to (3.138) by a conformal transformation. The expectation value of the energy-momentum tensor and other correlation functions in the physical metric (3.138) are then easily found by applying the correct transformation rules. Namely, in section 3.2, we saw that the energy-momentum tensor transforms like (3.67) under conformal transformations. Interestingly, we have already seen an example of a contribution resulting from such a transformations in (3.136). Namely, the inhomogeneous terms $F(u)(F(v))$ in the solution of the conservation equations (3.136), can be viewed as contributions from the conformal transformation from the reservoir to the black hole [55].

This framework allows the matter to pass freely between the black hole to flat spacetime, e.g. it allows the black hole to evaporate. In section 4.2, we will apply this on a more complicated geometry (i.e. the quotient manifold corresponding to a replica wormhole). In that case, gluing the AdS$_2$ geometry to a reservoir is not as trivial. This will result in the *conformal welding problem* described in section 4.2.

An expression for the stress tensor

To finish this section, we shortly illustrate how this evaporation is captured by the model. To do so, we will find an expression for $T_{\mu\nu}$ in terms of the Schwarzian.

As the condition $R = -2$ fixes the geometry to be AdS_2, the metric of the black hole region can be written as (see section 3.3)

$$ds^2 = \frac{4}{(U - V)^2} dU dV. \tag{3.139}$$

The asymptotic boundary is at $U = V = T$. As described above, we would like to find a transformations which connects this to the physical metric (3.138), which describes the black hole and the reservoir. Such a transformation F determines U and V in terms of the coordinates u, v such that $F(u) = U$ and $F(v) = V$. As a result, it is convenient to write (3.139) as

$$ds^2 = \frac{4}{(F(u) - F(v))^2} F'(u) F'(v) du dv, \tag{3.140}$$

recovering our general metric (3.138) by setting

$$2\omega = \ln \left(\frac{4}{(F(u) - F(v))^2} F'(u) F'(v) \right). \tag{3.141}$$

Using this in equation (3.136), one obtains

$$\langle T_{vv} \rangle = F_v(v) - \frac{c}{24\pi} (\{F(v), v\}) \tag{3.142}$$

where we used

$$\partial_v \omega(u, v) = \frac{1}{2} \left(\frac{F''(v)}{F'(v)} \right) + \frac{F'(v)}{F(u) - F(v)} \tag{3.143}$$

$$(\partial_v \omega(u, v))^2 = \frac{1}{4} \left(\frac{F''(v)}{F'(v)} \right)^2 + \frac{F''(v)}{F(u) - F(v)} + \left(\frac{F'(v)}{F(u) - F(v)} \right)^2 \tag{3.144}$$

$$\partial_v (\partial_v \omega(u, v)) = \frac{1}{2} \frac{F'''(v)}{F'(v)} - \frac{1}{2} \left(\frac{F''(v)}{F'(v)} \right)^2 + \frac{F''(v)}{(F(u) - F(v))} + \left(\frac{F'(v)}{F(u) - F(v)} \right)^2. \tag{3.145}$$

We thus see that the Schwarzian also arises in the context of the matter theory in our model. On the gravitational side, we have the Schwarzian corresponding to the boundary graviton $\tau(u)$ while on the CFT side we have the Schwarzian arising in the context of the transformation $F(v)$ which allows our quantum fields to pass freely trough the AdS_2 boundary. Solving the equation of motion is thus closely related to the symmetries of AdS_2 and finding the correct transformations from the black hole interior to the exterior.

In this chapter, we have acquired all the necessary knowledge to understand how one can use path integrals to describe JT gravity coupled to conformal matter. This will be applied in the next chapter to obtain a formula for the gravitational-fine grained entropy which was introduced in section 2.3.3.

Chapter 4

Replica wormholes

In chapter 2, we formulated the information paradox that arises for an evaporating black hole which was formed by gravitational collapse. Importantly, we have shown that this information problem can be quantified in terms of entropies. Especially, by considering the notions of fine- and coarse grained entropy, we saw that Hawking's computations already yield non-unitary results way before the black hole has evaporated completely. Namely, following Page's theorem 2.3.1, a unitary process which involves a pure state implies that $S_{fine,rad} = S_{fine,bh}$ always holds. However, we saw that, as a result of the increasing von Neumann entropy of the radiation and the shrinking horizon of the black hole, we encounter a regime where the fine-grained entropy of the radiation exceeds the coarse-grained entropy of the black hole. However, since the fine-grained entropy of the black hole is bounded by its course grained entropy, this yields a paradox. This observation clarified the information paradox and illustrated that, if one could show that the fine-grained entropy followed the Page curve, it could be concluded that all the information *does* come out of the black hole.

In this chapter we address the problem of deriving a formula for the gravitational-fine grained entropy and study how it evolves. However, we will discuss this for a simplified version of the information paradox, which can be formulated for *eternal black holes*. This setup is simpler than an evaporating black hole because it has no operator insertions in the definition of the state [68]. However, the resulting paradox is conceptually similar. Therefore, we hope that this already reveals some important keys about the resolution of the information paradox as originally formulated by Hawking.

In section 4.1, we formulate the information paradox for eternal black holes. Thereafter, we *use* the island-prescription, which has been derived in a holographic framework, to show that this new way of computing the entropy yields a result which *is* consistent with unitarity. Finally, in section 4.2, we *derive* this island-prescription for the gravitational-fine grained entropy of the radiation related to the eternal black hole. This derivation relies on the use of Euclidean path integrals and the replica trick, introduced in the previous chapter. Namely, we consider saddles with non-trivial topology which arise from the replicated Euclidean manifold consisting of an AdS_2 black hole glued to Minkowski space. Specifically, the saddles which yield contributions that make the fine-grained entropy consistent with unitarity are referred to as *replica wormholes*. At the end of this section, we discuss the implication of our result on the information paradox for eternal black holes and the original information problem introduced in chapter 2.

4.1   Quantum extremal surface

In this section, we will *use* the formula for the gravitational fine-grained entropy (2.65) to obtain an explicit expression for the entanglement entropy of quantum fields outside a black hole in AdS$_2$. To do so, we consider a simplified setup consisting of nearly-AdS$_2$ black holes in JT-gravity which are coupled to a 2d-CFT. This gravitational region is glued to flat Minkowski space, which serves as a reservoir and where the same 2d-CFT lives. As stated in [66], this model can be thought of as a toy model for asymptotically-flat near-extremal Reissner-Nordstrom black holes, where gravity is ignored in the asymptotically-flat region. Interestingly, in this setup, one can formulate a similar version of the Hawking information paradox presented in section 2.2. The goal of this section is to show that the entropy formula (2.65), resolves this simplified version of the information paradox. In section 4.2, we will then actually *derive* this formula with the use of the Euclidean gravitational path integral. For now, we simply *assume* the formula (2.65) which has been derived in holographic theories [15], and work in the Lorentzian picture [68].

In subsection 4.1.1, we start by describing the setup and explain how the formula for the fine-grained entropy should be applied on this model. To make the methodology more clear, we then discuss the case where the AdS$_2$ black hole is coupled to the flat reservoir at zero temperature in subsection 4.1.2 and compute the corresponding entropy. In section (4.1.3), we formulate a version of the information paradox which arises when considering *eternal black holes* coupled to Minkowski reservoirs at both sides. Thereafter, we obtain an explicit expression for the fine-grained entropy and discuss its implications on this version of the information problem.

Finally, note that, in what follows, we discuss some of the results obtained in [66] and partially follow the approach in [[15], [68]].

4.1.1   Setup

We start by describing the model that will be used to study the fine-grained entropy of quantum fields outside an AdS$_2$ black hole. As discussed in section 3.4, this setup, amongst other things, consists of coupling such a black hole to a reservoir, where gravitational effects are ignored. To do so, we consider two-dimensional JT gravity coupled to a CFT and glue it to a reservoir described by half of Minkowski space, where the same CFT lives.

Remember that the action of such a theory is given by

$$I_{\text{full}} = \frac{\phi_0}{16\pi} \left[\int_M \sqrt{-g}R + 2 \int_{\partial M} \epsilon \sqrt{h} K \right] + \frac{1}{16\pi} \left[\int_M \sqrt{-g}\phi(R+2) + 2 \int_{\partial M} \phi_b K \right] + I_{CFT}, \quad (4.1)$$

with boundary conditions $g_{bdy} = \frac{1}{\epsilon^2}$ and $\phi_b = \frac{\phi_r}{\epsilon}$. As explained in section 3.4, the equations of motion fix the metric to be locally AdS_2. Therefore, if we introduce the following coordinates

$$\begin{aligned} x^\pm &= T \pm z & \text{Gravitational region} \\ y^\pm &= t \pm \sigma & \text{Reservoir} \quad \sigma > 0, \end{aligned} \quad (4.2)$$

we obtain the following metrics

$$ds^2_{\text{grav}} = -\frac{4dx^+dx^-}{(x^- - x^+)^2} \qquad ds^2_{\text{reservoir}} = -\frac{1}{\epsilon^2}dy^+dy^-. \quad (4.3)$$

The coordinates from the gravitational region $x^\pm$ cover the Poincaré patch, as described in section 3.3 and the metric of the reservoir describes half of flat Minkowski space [66]. The corresponding Lorentzian geometry is presented on Figure 4.1.

Importantly, both regions need to be glued to each other near the AdS$_2$ boundary. This is done according to a map $x(t)$. Namely, this map identifies the curve $y^+ = y^- = t$, in the flat region, to the curve

$$x^+ = x(t) - \epsilon x'(t) \qquad x^- = x(t) + \epsilon x'(t) \quad (4.4)$$

in AdS$_2$ [68]. Furthermore, one obtains the dilaton profile in the gravitational region by solving the dilaton equation of motion (3.132). For the AdS$_2$ metric in Poincaré coordinates (4.3), this yields

$$\phi = -\phi_r \left[\frac{2}{(x^+ - x^-)} - \frac{x''(y^+)}{x'(y^+)} \right]. \quad (4.5)$$
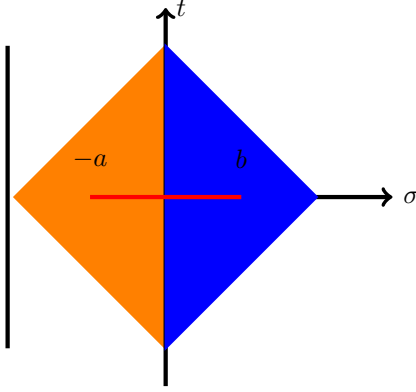
Figure 4.1: The Penrose diagram for the two-dimensional extremal black hole coupled to a reservoir at $T = 0$. The orange part represents the Poincaré patch of AdS_2 while the blue part corresponds to half of Minkowski space. This plays the role of the flat reservoir in our setup. We consider the situation where the same 2-dimensional CFT lives in both regions. The red line represents the interval $[-a, b]$ for which we compute the generalized entropy. Figure inspired by Figure 1 in [66].

as we assume $T_{++}^{(x)} = T_{--}^{(x)} = 0$ at the AdS_2 boundary [68]. This dilaton profile will be used to obtain an area term in the generalized entropy formula, as explained below.

Gravitational fine-grained entropy of the radiation

Of our main interest is the gravitational fine-grained entropy for radiation outside the AdS_2 black hole. Remember from section 2.3 that in a gravitational system, the original calculation of the von Neumann entropy (1.74) should be supplemented by the contribution from islands [15], according to the following rule:

$$S_{fine,rad} = \min_Q \left\{ \text{ext}_Q \left[\frac{\text{Area}(Q)}{4} + S_{vN}(\Sigma_{rad} \cup \Sigma_{island}) \right] \right\}. \tag{4.6}$$

In two dimensions, the boundary of an island, denoted by Q, corresponds to a point. Additionally, in 2-dimensional JT gravity, the corresponding $\text{Area}(Q)$ is given by the value of dilaton $\phi(a)$. Furthermore, the region $\Sigma_{rad} \cup \Sigma_{island}$, for which the ordinary entanglement entropy is computed, correspond to an interval that contains this island, as well as a region outside the black hole.

Note that, in what follows, we will first consider intervals $[-a, b]$ to lie on the $t = 0$ slice. In that case, we thus only extremize the generalized entropy with respect to the spatial direction. This yields the following expression for the fine grained entropy in our setup

$$S_{fine,rad}[-a, b]\Big|_{t=0} = \min_a \left[\frac{\phi_0}{4} + \frac{\phi(-a)}{4} + S_{vN}[-a, b] \right]. \tag{4.7}$$

Note that $S_{vN}[-a, b]$ is the entanglement entropy of an interval of length $a + b$ in a 2-dimensional CFT as derived in section 3.2, while $S_{fine,rad}[-a, b]$ is the gravitational fine-grained entropy. Furthermore, ϕ_0 corresponds to the extremal dilaton value and $\phi(-a)$ represents deviations from this maximum[64].

In what follows, we will minimize the expression between brackets, i.e. the generalized entropy S_{gen}, over all possible endpoints a. Explicitly, this corresponds to setting

$$\partial_a S_{fine,rad} = 0, \tag{4.8}$$

and solving the resulting equation for a. Plugging in this a into S_{gen} gives the gravitational fine grained entropy.

In the next subsection, we obtain the location of the QES for a setup where the gravitational region is in thermal equilibrium with the flat region at $T = 0$. Thereafter we apply it on the information paradox for eternal black holes coupled to a bath and check if it yields a unitary result. Finally, in the next section 4.2, we will derive the condition (4.8) and the expression (4.7) for these specific cases directly from the gravitational path integral in section.

4.1.2   Zero temperature configuration

We start by discussing the case where a black hole in JT gravity has temperature $T = 0$ and is coupled to and in equilibrium with a flat reservoir. This setup corresponds to Figure 4.1 and is referred to as a two-dimensional *extremal black hole*. In this subsection, we present the discussion in [66].

Firstly, note that we take the gravitational region to be in the Poincaré vacuum. This corresponds to setting the energy-momentum tensor in this region equal to zero $T_{++}^{(x)} = T_{--}^{(x)} = 0$. Additionally, as we take the $T = 0$ in the the gravitational region and this region is in thermal equilibrium with the reservoir, we should also have $T = 0$ in the flat region. Therefore, the energy-momentum stress tensor should also vanish in the flat region $T_{++}^{(y)} = T_{--}^{(y)} = 0$ [66].

We can now easily obtain the gluing map $x(t)$ with the help of the conformal anomaly (3.67). Namely, as explained in section 3.4, the gravitational and flat region should be related by a conformal transformation in order to have transparent boundary conditions. Therefore, we have that stress tensor of the black hole and the reservoir are related as

$$T_{++}^{(x)} = \left(\frac{\partial y^+}{\partial x^+} T_{++}^{(y)} - \frac{c}{24\pi} \{y^+, x^+\} \right). \tag{4.9}$$

As we both want T_{++}^{y} and T_{++}^{x} to vanish, this implies that the transformation $y \to x$ should be an SL(2,$\mathbb{R}$) transformation, because this yields a vanishing Schwarzian derivative. Furthermore, as stated in [66], the fact that we consider the full state, consisting of the black hole and the reservoir, is pure, implies that the transformation is simply

$$x^+ = y^+ \qquad x^- = y^-. \tag{4.10}$$

The generalized entropy

We will now obtain the location of the quantum extremal surface for the extremal black hole corresponding to the diagram on Figure 4.1. To start with, we need to write down the general expression of the generalized entropy. Specifically, this will be done for the interval $[-a, b]$ (in y coordinates), which includes a gravitational and a flat region. This interval is depicted by the red line on Figure 4.1. By plugging in the point $y^+ = -a$ (and hence $y^- = a$) into (4.10) and inserting the result in the general dilaton profile (4.5), one obtains

$$\phi(-a) = \frac{\phi_r}{a}, \tag{4.11}$$

The corresponding generalized entropy for this interval is thus given by [66]

$$S_{\text{gen}}(a) = \frac{\phi_0}{4} + \frac{\phi_r}{4a} + S_{\text{semi-cl}}[-a, b]. \tag{4.12}$$

Note that entropy $S_{\text{semi-cl}}[-a, b]$ actually corresponds to the entanglement entropy of an interval with two boundaries and length $a+b$. Hence, one can simply use (3.93), to obtain a more explicit expression. This yields the following form of the generalized entropy

$$S_{\text{gen}}(a) = \frac{\phi_0}{4} + \frac{\phi_r}{4a} + \frac{c}{6} \log \left[\frac{(a+b)^2}{a} \right], \tag{4.13}$$

where where the a in the denominator comes from a warp factor in the AdS$_2$ region of the metric [66]. This expression needs to be minimized over all possible spatial points a to obtain the fine-grained entropy (4.7)

$$\partial_a S_{\text{gen}}(a) = 0, \tag{4.14}$$

which yields the quadratic equation

$$-\frac{\phi_r}{4a^2} + \frac{c}{6} \frac{(a-b)}{a(a+b)} = 0$$

$$\Leftrightarrow \tilde{a}^2 - \tilde{a}(\tilde{b} + 1) - \tilde{b} = 0 \tag{4.15}$$

$$\text{with} \qquad \tilde{a} = a \frac{4c}{6\phi_r} \qquad \text{and} \qquad \tilde{b} = b \frac{4c}{6\phi_r}.$$
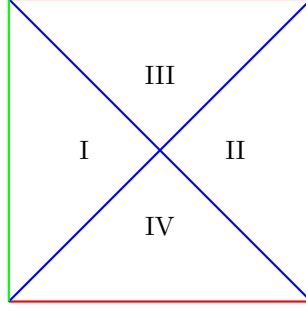
Figure 4.2: Double sided Penrose diagram representing an eternal black hole in AdS. Region III and IV contain a future/past singularity represented by the red lines. Region II covers the region that is outside the horizon from the point of view of an observer on the right boundary (green line) [67]. Region I is an exact copy of region II and contains a second boundary. Figure inspired by Figure 17.1 in [46].

Finally, we obtain the location of the QES as the solution of equation (4.15),

$$\tilde{a} = \frac{1}{2}\left((\tilde{b}+1) + \sqrt{\tilde{b}^2 + 6\tilde{b} + 1}\right).\tag{4.16}$$

Inserting this a in the entropy (4.13), one obtains the gravitational fine-grained entropy as defined in (2.65) for the interval $[-a, b]$.

In what follows, we will use the same reasoning to obtain the location of the QES corresponding to a configuration with finite temperature. In contrast to the $T = 0$ case we discussed in this subsection, the $T > 0$ can be applied on a simplified version of the information paradox.

4.1.3   Information paradox for the eternal black hole

The configuration where an AdS$_2$ black hole is coupled to a flat reservoir at finite temperature is more complicated. Namely, this setup can be thought of as consisting of two copies of the extremal black hole on Figure 4.1. In the Lorentzian picture, the resulting geometry is that of an *eternal black hole* in AdS$_2$, coupled to a flat reservoir on each side [68]. The corresponding diagram of the eternal black hole is depicted on Figure 4.2 while the diagram corresponding to the full system, where the eternal black hole is coupled to flat reservoirs at both sides, is depicted on Figure 4.3.

The idea behind this doubling of the previous setup relies on the thermofield double formalism, which is a trick to treat a thermal state as part of bigger system which is pure. We will not go through the details of this formalism here, but a discussion can be found in [46].

It is important to distinguish the eternal black hole from a black hole that forms by gravitational collapse. Namely, as one can see from Figure 4.2, the eternal black hole has a future *and* a past singularity. Additionally, it contains two asymptotic regions [46], in contrast to the black hole formed by gravitational collapse (see Penrose diagram 1.6). However, this geometry is still interesting because one can formulate a simpler version of the information paradox for eternal black holes [66]. In this subsection, we will formulate this version of the information paradox in the same way as is done in [66]. As this paradox is conceptually similar to the information problem discussed in chapter 2, we will not dive into details.

Hawking Entanglement entropy

Remember from section 2.2 that the information paradox was originally formulated for an evaporating black hole that is formed by gravitational collapse. Importantly, in AdS this evaporation only happens for small black holes. Unfortunately, these are not in thermal equilibrium and are difficult to study in holography [46]. However, we can consider an information paradox for an eternal black hole in AdS and try to resolve it [66]. As stated in [66], this setup is conceptually similar to that of a black hole formed by collapse. Therefore, one may hope that resolving this version of the paradox gives away some important keys towards the resolution of Hawking's information paradox.

To establish this version of the paradox, we will illustrate that the naive computation of the entanglement entropy yields a monotonically increasing contribution. As stated by the Page theorem 2.3.1, this is not consistent with unitarity.

Consider two intervals, $]-\infty, b[_L$ and $[b, \infty[_R$, in the reservoir regions on the $t = 0$ slice. These intervals correspond to the two black lines on Figure 4.3. First, one can note that the flat regions (V) and (VI) are related to the w-planes (III) and (IV) by the transformations

$$w^{\pm} = \pm \exp\left(\pm\frac{2\pi y_R^{\pm}}{\beta}\right) \qquad w^{\pm} = \mp \exp\left(\mp\frac{2\pi y_L^{\pm}}{\beta}\right). \tag{4.17}$$

Namely, one can view these y-regions as two *Rindler wedges* (section 1.2) of a single Minkowski spacetime with coordinates (w^+, w^-) [66]. This relation is useful because the state in this Minkowski spacetime is just the vacuum. As a result, we may compute the entropy with the use of the formula derived in section 3.2. For a metric of the form $ds^2 = -dw^+ dw^-$, the usual expression (3.93) becomes

$$S_{CFT} = \frac{c}{6} \log\left(\frac{(w_1^+ - w_2^+)(w_1^- - w_2^-)}{a}\right), \tag{4.18}$$

where a is a constant warp factor [68].

Inserting $t = 0$, $\sigma_R = b$ and $\sigma_L = -b$ in (4.2) and plugging in the result in (4.17), one obtains

$$S_{vN} = \frac{c}{3} \log \frac{\beta}{\pi} + c_1', \tag{4.19}$$

with the use of (4.18).

In order to study the evolution of this entropy, we move both intervals forward in time. This corresponds to setting

$$y_L^{\pm} = t \mp b \qquad y_R^{\pm} = t \pm b. \tag{4.20}$$

Inserting this into (4.17) and (4.18), one obtains the time dependent entanglement entropy

$$S_{gen}[R] = \frac{c}{3} \log\left[\frac{\pi}{\beta} \cosh\frac{2\pi t}{\beta}\right] + c_2'. \tag{4.21}$$

At late times $t >> \beta$, this linearly increases as

$$S_{vN} = \frac{2\pi c}{3} \frac{t}{\beta} + c_3'. \tag{4.22}$$

Just like the von Neumann entropy calculated in section 2.2.4 for a black hole formed by gravitational collapse, this entropy increases with time. Remember from chapter 2 that this growth becomes problematic when it exceeds the coarse-grained entropy of the black hole. Namely, this behaviour contradicts a basic property of the fine- and coarse grained entropy (2.55) $S_{fine} \geq S_{coarse}$. Furthermore, as stated in the Page theorem 2.3.1, this increase is not consistent with a unitary process and hence corresponds to information loss. As discussed in chapter 2.2, we can try resolve this tension by introducing a gravitational formula of the entropy. This is the island formula (4.6). In what follows, we will try to obtain an explicit expression for this gravitational entropy in the setup corresponding to Figure 4.3. Additionally, we would like to show that this does not exceed the maximal value of the coarse-grained entropy of the eternal black hole (which corresponds to the double of the Bekenstein entropy (2.28))

$$S_{max} = 2 S_{Bekenstein} = \frac{\phi_0}{2} + \frac{\pi \phi_r}{\beta}. \tag{4.23}$$

Of course, the factor 2 arises because we take into account the two black holes that form the eternal black hole.

4.1.4 Fine-grained entropy eternal black hole

Before we try to obtain an explicit expression for the gravitational-fine grained entropy (4.6) for radiation outside the eternal black hole, we start by considering a single interval $[-a, b]$. With a similar method as we used for the extremal black hole, we obtain the generalized entropy for this interval and obtain the location of the corresponding QES. Thereafter, we briefly explain why this result can simply be doubled to obtain the generalized entropy where we include an island contribution of the form $[-a, \infty]_L \cup [-\infty, a]_R$. The result will then be applied on the information paradox formulated in the previous subsection.
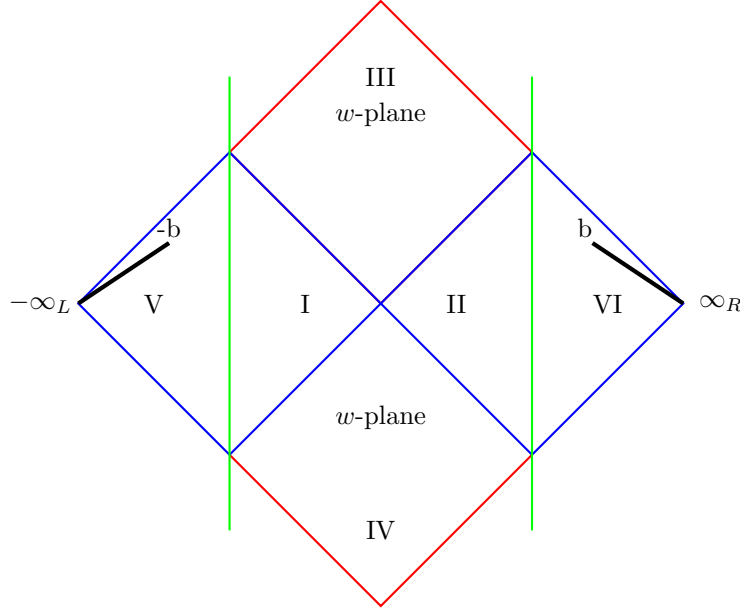
Figure 4.3: Eternal black hole coupled to two flat reservoirs on both sides. Regions I and II correspond to the region outside the horizon as in Figure 4.1. These regions are now coupled (at the boundary) to a reservoir (regions V en VI). Note that have two y-planes which are enclosed by the blue lines and where the combined state looks thermal. These two planes fit into a w-plane ($III+IV$) where the state is the Minkowski vacuum for the quantum fields. Figure based on Figure 8 in [66].

Single interval at finite temperature

We will use a similar methodology as used in section 4.1.2 to obtain the location of the QES arising in the setup on Figure 4.3. In this case, the gluing map $x(t)$ between the reservoir and black hole coordinates is not trivial. However, we know that the black hole and the reservoir should be in equilibrium, yielding a zero flux of energy between the regions. Inspired by the approach in [68], we therefore take a look at the Schwarzian equation of motion, obtained by the variation of the JT boundary term and I_{CFT} (3.132)

$$\partial_t \{x(\tau), \tau\} = T_{y^+ y^+} - T_{y^- y^-}. \tag{4.24}$$

Interestingly, the right hand side corresponds to the net flux of energy across the interface [68]. As the gravitational regions are in equilibrium, this flux should be zero, yielding

$$\partial_t \{x(\tau), \tau\} = 0. \tag{4.25}$$

A solution which corresponds to the configuration with inverse temperature β is then

$$x(t) = e^{\frac{2\pi t}{\beta}}. \tag{4.26}$$

As this map is only defined at the boundary, we need to find a way to obtain a coordinate which covers all the regions while satisfying (4.26) at the boundary. However, in the Lorentzian setup, we can simply extend (4.26) such that the $y^\pm$ coordinates also describe the gravitational region

$$x^+ = x(y^+) = e^{\frac{2\pi y^+}{\beta}} \qquad x^- = x(y^-) = e^{\frac{2\pi y^-}{\beta}}. \tag{4.27}$$

However, this is not always possible in a holomorphic way. As we will discuss in section 4.2, this can sometimes be solved by *conformal welding* .

In this case, we may simply consider (4.26) as a valid coordinate extension. Inserting this in the conformal anomaly yields

$$T_{y^+ y^+} = -\frac{c}{24\pi} \{x(y^\pm), y^\pm\} = \frac{\pi c}{12\beta^2}, \tag{4.28}$$

which is indeed consistent with a vanishing right hand side of the equation of motion (4.24).

Next, we would like to obtain the metric of the gravitational region in these coordinates. This metric is obtained by inserting (4.26) in the general form of the metric (4.3). By doing so, one obtains

$$ds^2_{grav} = -\frac{4\pi^2}{\beta^2}\frac{dy^+ dy^-}{\sinh^2(\frac{\pi}{\beta}(y^- - y^+))}. \tag{4.29}$$

This allows us to use following formula for the entanglement of the 2-dimensional CFT's

$$S_{vN} = \frac{c}{6}\log\left(\frac{(z_1^+ - z_2^+)(z_1^- - z_2^-)}{\Omega(z_1)\Omega(z_2)}\right), \tag{4.30}$$

for a metric of the form $ds^2 = -\Omega^{-2}dz^+ dz^+$. This is essentially the same formula as (3.93) but now the conformal factors Ω, arising in the metric (4.29), are taken into account.

Applying this formula to a single interval $[-a, b]_R$ at $t = 0$, as depicted in red on Figure 4.3, yields

$$S_{vN}[-a, b] = \frac{c}{6}\log\left(\frac{\sinh^2(\frac{\pi}{\beta}(a+b))}{\sinh\left(\frac{2\pi a}{\beta}\right)}\right) + \textbf{constant}, \tag{4.31}$$

where we inserted the conformal factors from (4.29) and (4.27).

Furthermore, by inserting (4.27) in (4.5), one finds that the dilaton profile in corresponds to

$$\phi = \frac{2\pi\phi_r}{\beta}\frac{1}{\tanh\frac{\pi}{\beta}(y^- - y^+)}. \tag{4.32}$$

Consequently, the generalized entropy is given by

$$S_{\text{gen}}[-a, b] = \frac{\phi_0}{4} + \frac{\pi\phi_r}{2\beta}\frac{1}{\tanh\frac{2\pi a}{\beta}} + \frac{c}{6}\log\left(\frac{\sinh^2(\frac{\pi}{\beta}(a+b))}{\sinh\left(\frac{2\pi a}{\beta}\right)}\right) + \textbf{constant}, \tag{4.33}$$

Finally, we find the location of the quantum extremal surface by minimizing this generalized entropy over a

$$\partial_a S_{gen} = 0 \qquad \Leftrightarrow \qquad \sinh\left(\frac{2\pi a}{\beta}\right) = \frac{3\pi\phi_r}{\beta c}\frac{\sinh\left(\frac{\pi}{\beta}(a+b)\right)}{\sinh\left(\frac{\pi}{\beta}(a-b)\right)}. \tag{4.34}$$

Writing this as exponentials, this results in a cubic equation for $e^{\frac{2\pi a}{\beta}}$ with solution

$$a \approx b + \frac{\beta}{2\pi}\log\left(\frac{6\pi\phi_r}{\beta c}\right). \tag{4.35}$$

Note that, we could have found equation (4.13) by taking the limit $\beta \to \infty$ in equation (4.33). Furthermore, note that the [66] uses the convention $4G_N = 1$ which explains why our results differ by a factor four from that paper.

Two disjunct intervals

In the previous part, we obtained the explicit expression for the generalized entropy of an interval $[-a, b]$ at $t = 0$. However, to address the information paradox in the eternal black holes, introduced in subsection 4.1.3, we need to take into account two disjoint intervals. Namely, we need to take into account a single interval at each side of the eternal black hole to make the entropy time-dependent.

Interestingly, in [66] and [68] it is argued that one can obtain the von-Neumann for two intervals factorizes as the double of the single interval entropy (4.31). Additionally, [66] suggests that the QES a is the same. To get some intuition, we present the arguments used in [66] and [68] and discuss the results obtained in [15]. Finally, we compare this with the maximal coarse-grained entropy of the eternal black hole to check if the tension with unitarity is resolved by the inclusion of islands.

Consider again the geometry of the eternal black hole. The two intervals disjoint intervals for which we compute the generalized entropy should include in the computation. The resulting Lorentzian geometry is depicted on Figure 4.4. Note that we set $\beta = 2\pi$ in what follows.

The end points of the intervals under consideration are

$$P_1 = (a, t_a) \qquad P_2 = (b, t_b) \qquad P_3 = (-a, -t_a + \pi) \qquad P_4 = (b, -t_b + \pi). \tag{4.36}$$

In this notation, the radiation region consists of

$$R = [P_4, \infty_L[\cup [P_2, \infty_R[, \tag{4.37}$$

while the island corresponds to

$$I = [P_3, P_1]. \tag{4.38}$$

As we take into account two intervals, we should modify the general formula (4.30) for the von Neumann entropy by

$$S_{vN}[2 \text{ intervals}] = \frac{c}{6} \log \left[\frac{|z_{21} z_{32} z_{43} z_{41}|^2}{|z_{31} z_{42}|^2 \Omega_1 \Omega_2 \Omega_3 \Omega_4} \right], \tag{4.39}$$

where $|z_{ij}|^2 = (z_i^+ - z_j^+)(z_i^- - z_j^-)$ and Ω again correspond to the conformal factor appearing in (4.29).

By plugging in the points (4.36) into (4.39) and some algebraic manipulations, one finds that the corresponding von Neumann entropy of the region $I \cup R$ is

$$S_{vN}(I \cup R) = \frac{c}{3} \log \left[\frac{2 \cosh t_a \cosh t_b |\cosh(t_a - t_b) - \cosh(a + b)|}{\sinh a \cosh\left(\frac{a+b-t_a-t_b}{2}\right) \cosh\left(\frac{a+b+t_a+t_b}{2}\right)} \right], \tag{4.40}$$

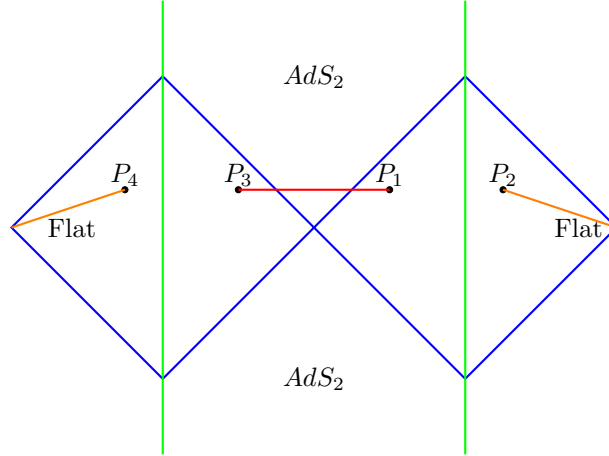where and a CFT with c free Dirac fermions is considered [15].

Figure 4.4: Lorentzian picture of two intervals in the eternal black hole geometry. The red line corresponds to the interval including the island contribution while the orange line refer to the intervals outside the black hole. This Figure based on Figure 15 in [68].

As in [15] we look at the generalized entropy at $t_a = t_b = 0$,

$$S_{gen}(I \cup R)_{t=0} = \frac{\phi_0}{4} + \frac{\phi_r}{4} \frac{1}{\tanh a} + \frac{c}{3} \log \left(\frac{4 \tanh^2 \frac{a+b}{2}}{\sinh a} \right). \tag{4.41}$$

The location of the boundary of the island $-a$, is thus found by solving

$$\partial_a S_{gen}(I \cup R)_{t=0} = 0$$
$$\Leftrightarrow -\frac{c}{3 \tanh a} - \frac{\phi_r}{8 \sinh^2 a} + \frac{2c}{3} \frac{1}{\sinh(a + b)} = 0 \tag{4.42}$$
$$\Leftrightarrow -\cosh a \sinh a \sinh(a + b) - \frac{3\phi_r}{8c} \sinh(a + b) + 2 \sinh^2(a + b) = 0$$

As stated in [15], it depends on the parameters b and $\frac{\phi_r}{c}$ if this has a real solution. However, at late times, the two-interval configuration sits in the "OPE limit". This allows us to factorize the von

Neumann entropy $S_{vN}[I \cup R]$ as two times the von Neumann entropy of the single interval $[-a, b]$ (4.31) [[66], [15],[68]] as computed in the previous subsection. Explicitly, we thus have

$$S_{vN}(I \cup R) \approx 2S_{vN}[-\sigma_a, b] = \frac{c}{3} \log\left(2\frac{\sinh^2(a+b)}{\sinh a}\right). \tag{4.43}$$

As the dilaton contribution of the two intervals just add up, we obtain the generalized entropy

$$S_{gen,island}(a) = \frac{\phi_0}{2} + \frac{\phi_r}{2\tanh a} + \frac{c}{3} \log\left(2\frac{\sinh^2(a+b)}{\sinh a}\right), \tag{4.44}$$

which is twice the single interval answer (4.33). Furthermore, note that we also used the island prescription sets $t_a = t_b$ [68].

As a result of (4.44), the quantum extremal surface in this case is the same as (4.35). Plugging this value into (4.44), we see that the fine-grained entropy of the radiation is dominated by the area term because $\frac{c}{3} \log\left(2\frac{\sinh^2(a+b)}{\sinh a}\right)$ is very small in the case where $\frac{\phi_r}{c} \gg 1$, in which we are interested [66]. Consequently we obtained,

$$S_{gen,island} = \frac{\phi_0}{2} + \frac{\phi_r}{2\tanh a} + \text{small} \approx S_{max}, \tag{4.45}$$

where S_{max} correspond to the maximal coarse-grained entropy of the black hole (4.23)[1]. Note that this result is valid when $t \gg \beta$ and $\frac{\phi_r}{c} \gg 1$. Remember that we are ultimately interested in the
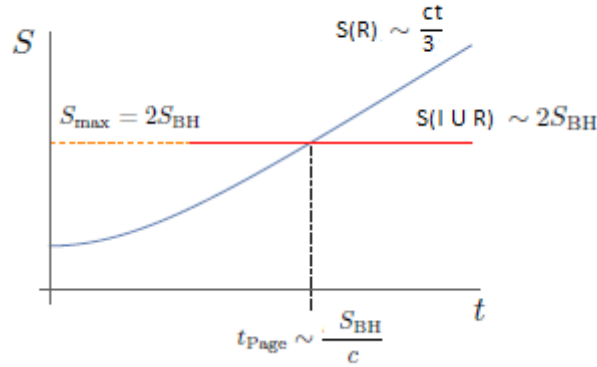

Figure 4.5: In order to illustrate the behaviour of the two types of contributions to the fine-grained entropy are plotted. As one can see, the entropy $S(R)$ that does not include contributions from the island, increases with time. This yields a fine-grained entropy of the radiation which becomes larger than the coarse-grained entropy of the black hole. As explained in section 2.3.1, this is not what one expects from a unitary process. In this section, we have shown that an island contribution should arise at late times $S(I \cup R) \approx S_{max}$. This Figure is entirely based on Figure 9 in [66].

correct fine-grained entropy (2.65), obtained by minimizing over the island and no-island contributions

$$S_{fine,rad} = \min\left\{S_{gen}(R), S_{gen}(I \cup R)\right\}. \tag{4.46}$$

As the no-island contribution $S_{gen}(R)$ (4.21) starts from a small value at $t = 0$, this will be the dominating contribution at early times. However at the moment where $S_{gen}(R) = S_{gen}(I \cup R)$ (at $t \gg \beta$), the island contribution minimizes the entropy, as this contribution is constant. As a result, the fine-grained entropy stops increasing and the problematic behaviour described in subsection 4.1.3 seems to be avoided. As an illustration, the behaviour of the contributions $S_{gen}(R)$ and $S_{gen}(I \cup R)$ is plotted on Figure 4.5.

As a result, the partial information problem, as described in 2.3 and 4.1.3 seems to be resolved for the eternal black hole [66]. However, to fully solve the information paradox in this simplified setup,

---

[1] Again note that we have set $\beta = 2\pi$ in this subsection

one should be able to explain *how* this information comes out of the black hole. In contrast, our findings only suggest that information comes out without explaining how information about matter inside the black is encoded in the Hawking radiation. Currently, physicists are trying to address this problem with *entanglement wedge reconstruction* [13].

Finally, note that, up to now, our computations of the fine-grained entropy relied on a formula (2.65), which was derived by holographic computations. Namely, in this section we simply *used* this formula and applied it on the Lorentzian geometry of an eternal black hole. In the next section, we are going to justify our computations by *deriving* the island-prescription for the fine-grained entropy of the radiation (2.65) with the use of the Euclidean path integral and the replica trick. Explicitly, we will show that (4.33) and (4.34) follow directly from gravitational methods without appealing to higher dimensional holography [15].

4.2   Replica wormholes

In the previous section, we naively used the formula (4.6), which has been obtained by assuming AdS/CFT duality in [26], to obtain an explicit expression for the entropy of the radiation. In this section, we will show that the entropy (4.33) and the condition (4.8) follow naturally from gravitational methods. Namely, we will use the replica method, discussed in section 3.2.4, on the gravitational system consisting of the black hole and the reservoir. As explained before, the von Neumann entropy can be obtained by evaluating the Euclidean path integral on this replicated system. To do so, we evaluate this path integral as a saddle point. Since the geometry is not fixed in the gravitational region, these saddles correspond to different geometries of the replicated system consisting of a black hole and a reservoir. Interestingly, two types of saddles arise. The obvious one leads to the standard Hawking entropy while another type corresponds to replica wormholes. These have a non-trivial topology which connects different copies of the black hole by a new geometry. These contributions start to dominate when the Hawking contributions begins to violate unitarity, yielding the entropy obtained with the island rule (4.33).

In what follows, we use the acquired knowledge of chapter 3 to understand the methodology and results presented in [15]. First, we reintroduce the replica trick to calculate the von Neumann entropy in a gravitational context (section 4.2.1). Similarly as in section 1.3, we consider a quotient manifold instead of computing the path integral over the n-sheeted replicated system. The physics on such a quotient manifold is described in detail in section 4.2.2. We then apply the replica method explicitly on the 2d JT gravity theory, coupled to a CFT, and derive the general form of the corresponding equations of motion in section 4.2.3. This will involve new aspects such as the *conformal welding problem* and rely on Schwarzian theory introduced earlier (sections 3.3.3 and 3.4.1). This technical discussion is then followed by the application of the general procedure on the single interval configuration at finite temperature in section 4.2.4. The results will be compared to the results of the previous section 4.1.3, and we will see that they are equivalent. To finish with, we discuss the implications on the configuration which is most relevant for the information paradox: the case where two intervals are included. Note that this section is largely based on [15] and partially on [68].

4.2.1   Replica method in a gravitational theory

Our goal is to show that the island rule (4.6) for the computation of the gravitational fine grained entropy can be obtained without assuming AdS/CFT, at least in the simplified setup presented in section 4.1. To do so, we will use the replica method to compute

$$S = -\partial_n \left(\frac{\log Z[\tilde{M}_n]}{n} \right) \Bigg|_{n=1} = -\partial_n (Tr\rho^n)\Big|_{n=1}. \tag{4.47}$$

As we work in a gravitational theory, obtaining $Tr\rho^n$ consists of computing the Euclidean gravitational path integral. As discussed in section 3.4.2, we work in the semi-classical approximation and assume this path integral is dominated by saddles. These correspond to different geometries connecting the copies of the system consisting of the black hole and the reservoir.

In order to understand what this replicated system looks like, we discuss the case of 2 replicas ($n = 2$). If we want to obtain $\text{Tr}\rho^2$, we need to evaluate the gravitational path integral on this replicated system, depicted on Figure 4.6. As discussed in more detail below, we choose to work in the Euclidean signature and analytically continue to obtain the Lorentzian result. This Euclidean point of view allows us to consider the coupled system of the black hole and flat space as the inner and outside of the standard hyperbolic plane. When we evaluate the gravitational path integral on this replicated system, we do not fix the geometry/topology of the gravitational (inside) region but only fix boundary conditions. As a result, the way two copies are sewed together is determined in a dynamical way by the gravitational theory. We will consider two saddles, i.e. two ways of sewing the replicas, corresponding to the gravitational path integral. These are depicted on the right of Figure 4.6. One of them simply glues together the replicas in a non-gravitational way, through branch cuts in the flat region. The non-trivial saddle is the replica wormhole where gravity glues together two copies of the original system. This topology starts to dominate when the Hawking computation fails to be consistent with unitarity.

In the remaining part of this subsection, we will explain how the replica method is used in a general gravitational theory. Next, we are going to apply that method on two dimensional JT gravity.
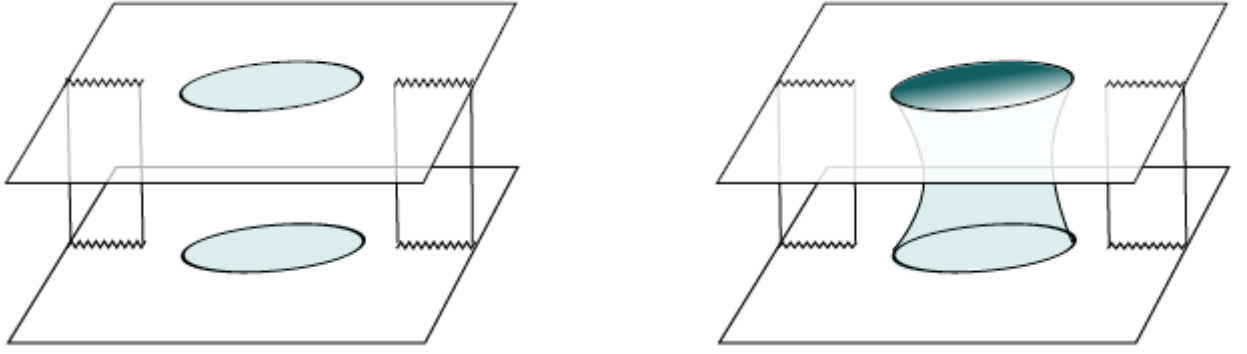
Figure 4.6: On the left, the two replicas of the system are sewed together through cuts in the flat region of space. As a result, this does not incorporate gravitational effects and thus yields the original Hawking result. The right picture corresponds to the geometry of a replica wormhole, where two replicas are glued together dynamically by gravity. This saddle is of our interest as it leads to an entropy which is consistent with unitarity and starts to dominate at Page time. This Figure has been taken from [15].

Replicated action

In section 3.2.4, we introduced the replica trick in order to compute the von Neumann entropy of a region A in a conformal field theory. Remember that this method is based on the observation that $\text{Tr}\rho_A^n$ can be viewed as an observable on a non-trivial topology $\tilde{M}_n$ corresponding to n copies of the original geometry sewed together as on Figure 3.4. As a result, one can compute the entropy by considering the Euclidean path integral over this n-sheeted structure and analytically continue to non-integer n. However; we saw that, when assuming replica symmetry, the replicated system can be seen as a single sheeted system containing n copies of the CFT, where twist fields are inserted at the branch points [59]. As a consequence, the unnormalized correlator of these twist fields can be seen as the partition function (3.81) [68]. A similar idea will be used to compute the entropy (4.47) in a gravitational theory coupled to a CFT.

We will consider a two-dimensional CFT which lives both on AdS_2 and on non-gravitational Minkowski space (which again serves as a reservoir). The full action corresponding to such a system on a non-trivial n-sheeted manifold $\tilde{M}_n$ has the following form:

$$I_{\text{full}}[\tilde{M}_n] = \frac{\log Z[\tilde{M}_n]}{n} = -\frac{1}{n}I_{\text{grav}}[\tilde{M}_n] + \frac{1}{n}\log Z_{\text{mat}}[\tilde{M}_n], \tag{4.48}$$

where $Z_{mat}[\tilde{M}_n]$ corresponds to the partition function of the two-dimensional conformal fields living on $\tilde{M}_n$.

While the manifold $\tilde{M}_n$ is fixed in the non-gravitational Minkowski region, one can consider a manifold with any geometry or topology in the gravitational region. This corresponds to the semi-classical approximation presented in (3.130). As a result, saddles of the full action (4.48) correspond to n-sheeted manifolds with a certain geometry and topology which are allowed by the boundary conditions we impose. One could try to look for these saddles by solving the equations of motions which arise from extremizing (4.48). However, like in the non-gravitational case, it is more easy to consider a single sheeted manifold

$$M_n = \frac{\tilde{M}_n}{Z_n}. \tag{4.49}$$

Under the assumption of replica symmetry, this can be seen as one where n copies of the CFT live and where twist fields are inserted at endpoints of the relevant intervals. In order to understand the difference between the manifold M_n and the replica manifold $\tilde{M}_n$, it is useful to look at Figure 4.7, where they are depicted for $n = 3$. Importantly, in the gravitational region, we must take into account the fact that the manifold M_n has conical singularities with opening angles $(2\pi/n)$, which result from the non-trivial way of sewing different copies. At these conical singularities, additional twist fields are inserted. Furthermore, it is convenient to enforce the conical singularities by adding cosmic branes
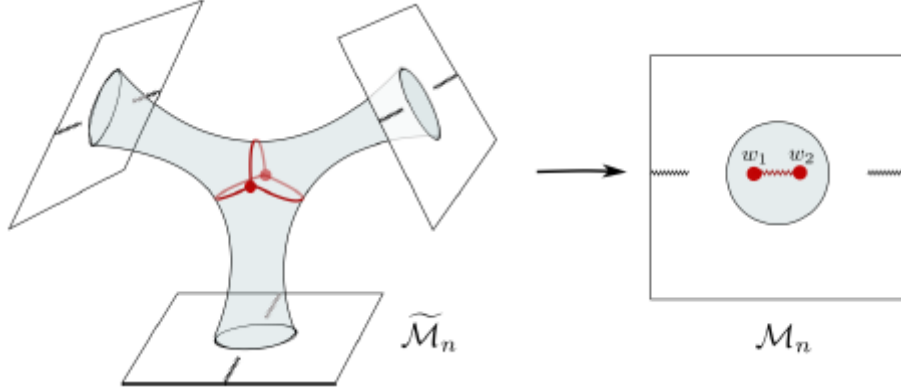
Figure 4.7: On this Figure, the replicated manifold $\tilde{M}_3$ consisting of 3 copies of the original system is depicted on the left. On the right, we have the simplified topology of a disk with conical singularities at w_1 and w_2. These singularities correspond to the red points on the left manifold. This Figure has been taken from [15].

with a certain tension T_n given by

$$4T_n = 1 - \frac{1}{n}. \tag{4.50}$$

We will not discuss what these cosmic branes are and what this tension refers to as this involves a discussion on String Theory [47]. Note however that these branes correspond to points in two dimensions and they add a new term to the gravitational action (4.48). Namely, the gravitational action on the single sheeted manifold is given by

$$I_{\mathrm{grav}}[M_n] = \frac{1}{n} I_{\mathrm{grav}}[\tilde{M}_n] + T_n \int_{\Sigma_{d-2}} \sqrt{g}. \tag{4.51}$$

Note however that this extra term will be canceled by a surface term coming from the Einstein term when integrated through the conical singularity [15].

We will thus be interested in classical solutions of the combined action

$$I_{\mathrm{full}}[M_n] = I_{\mathrm{grav}}[M_n] + \log Z_{\mathrm{mat}}[M_n]. \tag{4.52}$$

As stated before, classical solutions of the full action (4.48) will correspond to manifolds with a specific geometry and topology. Namely, different ways of sewing together different copies, result in different geometries of the quotient manifold. For example, the positions of the conical singularities and their corresponding twist fields can be obtained dynamically. As a result, for $n = 1$, this will simply yield the original manifold $M_1 = \tilde{M}_1$, because no copies have been made and hence no replica wormholes appear (such that the quotient manifold is just the original manifold without conical singularities). In contrast, we will be interested in the solution for $n \sim 1$. In this limit, we need to take into account the $O(n-1)$ effects of the cosmic branes and twist fields at the conical singularities. The action on the quotient manifold is then given by

$$\left(\frac{I^{\mathrm{tot}}}{n}\right)_{n \to 1} = I_1 + \delta\left(\frac{I}{n}\right), \tag{4.53}$$

where the second term contains the new effects. We can evaluate this action in a perturbative way by seeing these effects as modifications of the original geometry M_1 and hence consider the cosmic branes and twist fields as living on this geometry. Importantly, we will see that the explicit contributions $\delta\left(\frac{I}{n}\right)$ have the form of the generalized entropy [15]. From this observation, one can conclude that extremizing $\delta\left(\frac{I}{n}\right)$ should be equivalent to the condition (4.8). This will be shown explicitly later on. Before we move on to that, we dive deeper into the exact form of the gravitational action and how it is coupled to matter.

JT gravity

We will apply the replica method on the JT gravity theory coupled to a two dimensional CFT. We start by writing down the explicit action on the manifold $\tilde{M}_n$ as was done for an unspecified theory in (4.48). For JT gravity in two dimensions, this yields

$$-\frac{1}{n}I_{grav}[\tilde{M}_n] = I_{EH} + I_{GH} + I_{JT,bulk} + I_{JT,boundary}$$
$$= \frac{\phi_0}{16\pi}\Big[\int_{\Sigma_2} R + \int_{\partial\Sigma_2} 2K\Big] + \int_{\Sigma_2} \frac{\phi}{16\pi}(R+2) + \frac{\phi_b}{8\pi}\int_{\partial\Sigma_2} K. \tag{4.54}$$

This corresponds to the sum of the Einstein-Hilbert action (which is purely topological [15]), the JT action and their Gibbons-Hawking-like contributions which regulate them. As explained before, we instead consider the manifold M_n. The corresponding action on this manifold, which includes the contribution of the cosmic branes at the location of conical singularities w_i, is

$$-\frac{1}{n}I_{grav}[M_n] = I_{EH} + I_{GH} + I_{JT,bulk} + I_{JT,boundary} + I_{conic}$$
$$= \frac{\phi_0}{16\pi}\Big[\int_{\Sigma_2} R + \int_{\partial\Sigma_2} 2K\Big] + \int_{\Sigma_2} \frac{\phi}{16\pi}(R+2) + \frac{\phi_b}{8\pi}\int_{\Sigma_2} K - (1-\frac{1}{n})\Sigma_i[\phi_0 + \phi(w_i)].$$
$$\tag{4.55}$$

4.2.2 The quotient manifold

In the previous subsection, the replica trick was introduced in a gravitational context. Furthermore, we learned how to transform a gravitational action, defined on an n-sheeted manifold $\tilde{M}_n$, towards an action on the quotient manifold M_n. At this point, we would like to dive deeper into the geometry of the flat and gravitational regions on this manifold and see how these are *glued* together.

Original manifold M_1

We start by discussing how the original, single sheeted state, consisting of the black hole coupled to a reservoir, is constructed with the Euclidean path integral.

As in section 4.1, we start from the combined thermofield double state of the black hole and radiation as an initial state. This state is prepared by the Euclidean path integral, by imposing periodic boundary conditions in Euclidean time τ. This yields the Euclidean cylinder parametrized by coordinates y, depicted on Figure 4.8. If we cut this cylinder along the black dotted line, we obtain the thermofield initial double state. After Lorentzian analytic continuation, this yields the geometry on Figure 4.3, used in the previous section for the finite temperature configuration. We are most
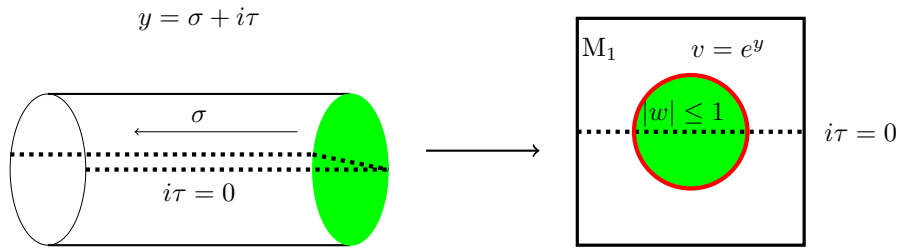
Figure 4.8: On the left we see the thermofield double state of the combined system consisting of the radiation and the black hole. This state is prepared by the Euclidean path integral (see section 3.1 and 4.1). The green regions correspond to the gravitational region (Euclidean AdS$_2$). On the right, this is represented as the inner disk of the hyperbolic plane parametrized by the w coordinates, referred to as M_1. The outside region (corresponding to the outside of the cylinder) similarly represents the non-gravitational region. In this region we use the $v = e^y$ coordinates. The initial state is then obtained by cutting along the dotted line, resulting in a geometry like Figure 4.3 after Lorentzian evolution. This Figure is inspired by Figure 2 in [15].

interested in the Euclidean picture on the right of Figure 4.8, where we represent our system by the

topology of a disk. With this picture, we can represent the gravitational region as the green inner region and the flat region as the outside of this disk. The metric which describes the flat outer region (with coordinates $v = e^y$) is

$$ds^2_{out} = dy d\overline{y} \qquad |v| \geq 1. \tag{4.56}$$

Furthermore, as the gravitational region is described by JT gravity, where $R = -2$ fixes the metric to be AdS$_2$, this yields the hyperbolic disk in the Euclidean signature with metric

$$ds^2_{grav} = \frac{4 dw d\overline{w}}{(1 - |w|^2)^2} \qquad |w| \leq 1. \tag{4.57}$$

The boundary between the gravitational and non-gravitational region is depicted by the red circle on Figure 4.8. This boundary is parametrized by θ from the inside,

$$w = e^{i\theta} \qquad |w| = 1 \tag{4.58}$$

and by the time τ along the boundary of the outside region

$$v = e^{i\tau} \qquad |v| = 1. \tag{4.59}$$

As a result, gluing both regions amounts to extending the coordinates from the inside to the outside region, in a way determined by the gluing map $\theta(\tau)$. Therefore, the gluing map can be thought of as the boundary graviton, introduced in section 3.3.3. On the original manifold M_1, this gluing map is simply $\theta(\tau) = \tau$ as the inner coordinates w can simply be extended to the outside region. In contrast, on the quotient manifold M_n, which generally contains conical singularities, the gluing is not trivial. This has important consequences on the CFT which describes the matter of the theory. This will be discussed below, after the gravitational region is analysed further.

Gravitational region on quotient manifold

The gravitational region on the original manifold M_1 of our coupled system consisting of the black hole and flat space, is described by the standard hyperbolic disk, described by the metric (4.57). However, this is generally not the case on the quotient manifold M_n. Namely, when extremizing the gravitational action (4.55), the equation of motion is not simply $R = -2$. Explicitly, it is shown in appendix D.4 that writing the general metric in the conformal gauge (3.133)

$$ds^2 = e^{2\rho} dw d\overline{w}, \tag{4.60}$$

yields the corrected equation of motion

$$-4 \partial_w \partial_{\overline{w}} \rho + e^{2\rho} = 16\pi(1 - \frac{1}{n}) \sum_i \delta^2(w - w_i) = 0. \tag{4.61}$$

Note that, imposing this equation makes the explicit cosmic brane action terms introduced in (4.51) vanish. This equation incorporates effects from the conical singularities. Consequently, the metric of the gravitational region will not correspond to the standard hyperbolic disk anymore and the correct metric has small deviations compared to (4.57). Explicitly, the following metric describes the gravitational region on the quotient manifold

$$ds^2_{\text{grav}, M_n} = \frac{4|dw|^2}{(1 - |w|^2)^2} e^{2\delta\rho}$$

$$\tag{4.62}$$

$$\text{with} \qquad \delta\rho \sim -\frac{(1 - |w|)^2}{3} U(\theta) \qquad U(\theta) = -\frac{1}{2}\left(1 - \frac{1}{n^2}\right) \frac{(1 - A^2)^2}{(e^{i\theta} - A^2)(e^{-i\theta} - A)^2}.$$

This metric describes the inner green disk of Figure 4.7.

Remember from our discussion about JT gravity (section 3.4.1) that the Schwarzian of the boundary graviton $\theta(\tau)$, encodes the gravitational dynamics of the theory. In this case, this equation of motion will get an additional contribution of $U(\theta)$ due to the form of the modified metric (4.62). Before we see this explicitly, we take a look at the matter side of the theory.
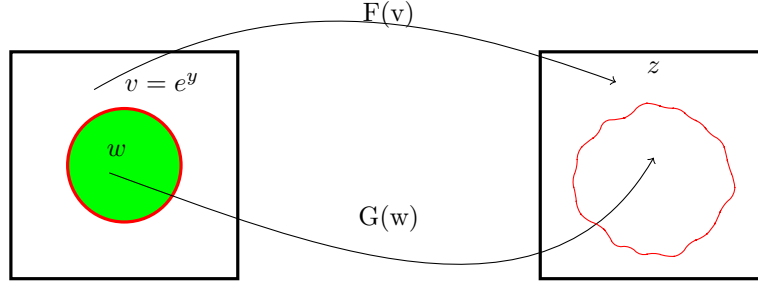
Figure 4.9: The conformal welding problem is depicted schematically on this Figure. The function $G(w)$ maps the gravitational region, represented by the green disk, on an inside region of the complex z plane. Similarly, $F(v)$ maps the flat outer region to another region on this complex disk in such a way that they are compatible at the boundary. Figure based on Figure 10 in [15].

The conformal welding problem

Up to now, we only discussed what the gravitational region of the quotient manifold should look like. However, this region must be joined to a flat Minkowski region described in section 4.2.2. Namely, we define a two dimensional CFT which is defined on a Minkowski *and* an AdS$_2$ background. Importantly, this implies that we should extend the interior coordinates to the flat coordinates in a holomorphic (e.g. conformal) way. This is necessary to compute the CFT stress tensor in the original physical coordinates through the transformation rule (3.67)) [68]. Naively, one could write down coordinate extensions of the form $(w(y, \overline{y}), \overline{w}(y, \overline{y}))$, as in the Lorentzian computation in section 4.1, which cover both regions on Figure 4.8. Unfortunately, it turns out that it is not always possible (e.g. in the Euclidean signature) to extend the gravitational w coordinates to an outside non-gravitational region in a holomorphic way [68]. To solve this, we look for a z-coordinate which covers both regions (instead of extending the w coordinates to the v coordinates). We can find such a coordinate by looking for holomorphic functions $F(w)$ and $G(v)^2$ which map each region of the disk to a region of the z complex plane such that

$$
\begin{aligned}
z &= F(e^{i\theta}) = G(e^{i\tau}) && \text{at the boundary} && |w| = |v| = 1 \\
z &= F(v) && \text{for} && |v| \geq 1 \\
z &= G(w) && \text{for} && |w| \leq 1.
\end{aligned}
\tag{4.63}
$$

We thus map the inner and outer disk to a plane where their boundaries are compatible. The corresponding mapping from the hyperbolic disk to the complex plane is represented schematically on Figure 4.9. The problem of finding such maps $F(w)$ and $G(v)$, given a gluing map $\theta(\tau)$, is referred to as the *conformal welding problem* [15]. Interestingly, we can always find such maps due to the Riemann mapping theorem, see appendix E and [72] for a more technical discussion on this subject.

The stress tensor

Once we have found a z-coordinate which entirely covers the (non)-gravitational regions, we can write the metric globally as [68]

$$
ds^2 = \Omega^{-2}(z, \overline{z})dzd\overline{z}.
\tag{4.64}
$$

This allows us to use (3.67) to obtain the physical stress-tensor

$$
T_{yy} = \left(\frac{dF(e^{iy})}{dy} \right)^2 T_{zz} - \frac{c}{24\pi} \{F(e^{iy}), e^{i\theta}\}.
\tag{4.65}
$$

As expected from our discussion in section 3.4.3, the stress tensor depends on the conformal map from the reservoir to the black hole. Note that, a similar expression can be found for T_{ww} and $T_{\overline{yy}}$ [15].

---

[2] The functions $F(w)$ and $G(v)$ only need to be holomorphic in the domain defined below [68]

4.2.3   Equations of motion

In the previous part of this section, the replicated action of the theory and the conformal welding problem were introduced. In what follows, these expressions are used to obtain a general form of the equations of motion. The classical solutions of this equation, correspond to the saddles in which we are ultimately interested to obtain an expression of the entropy.

  The construction of these equations has to be done in two steps. Firstly, one needs to vary the gravitational action with respect to the boundary graviton in order to describe the JT gravitational dynamics. Additionally, the action corresponding to the conformal field theory will be varied. At the end both contributions are combined and the general procedure to find a solution is presented.

Gravitational side

In section 2.4.1, we saw that, after filling in $R = -2$, the JT action only contains a boundary term. Furthermore we saw that this term evaluated to the Schwarzian $\{\tau(u), u\}$ and captures the gravitational dynamics through equation (3.128). In that context, $\tau(u)$ was the boundary graviton which represented ways of cutting the AdS$_2$ geometry. The resulting shapes of the resulting boundary were parametrized by u which corresponds to the boundary time of a well defined UV-theory.

  Interestingly, we can interpret the situation in our current model in the same way. Namely, the parameter τ corresponds to the time at the boundary of the flat region. Furthermore, $\theta(\tau)$ represents specific ways of gluing the flat to the gravitational region. As a result, $\theta(\tau)$ arises as the boundary graviton of our theory. Consequently, we expect that the equation of motion, which results from the variation of (4.55) with respect to the boundary graviton, is

$$\frac{\phi_r}{8\pi} \partial_\tau \{e^{i\theta(\tau)}, \tau\} = 0. \tag{4.66}$$

However, as discussed before, the geometry of our gravitational region does not exactly correspond to Euclidean AdS$_2$. This is a result of the induced conical singularities in this region after changing from $\tilde{M}_n \to M_n$. Explicitly, we show in appendix D.6 that the correct equation of motion resulting from the variation with respect to $\theta(\tau)$ is

$$\frac{\phi_r}{8\pi} \partial_r \left(\{e^{i\theta}, \tau\} + U(\theta)\theta' \right) = 0. \tag{4.67}$$

Later on, we will derive this for a specific configuration. First, we take a look at the variation of δI_{CFT} in order to obtain the full equations of motion corresponding to the variation of (4.52).

CFT side

Remember from previous sections that the variation of the action corresponding to the CFT with respect to the metric yields simply the energy-momentum tensor. However, to obtain the full equation of motion corresponding to the full action (4.52), we need to vary this action with respect to the boundary graviton. Note that to write the result in terms of the stress tensor, we should determine which change in the metric $\delta g_{\alpha\beta}$ corresponds to a variation of the boundary graviton. To obtain this, we will look at the change in the relationship between the coordinates in the gravitational and flat region, as is done in appendix of [15].

Remember that the coordinates at the boundary of the gravitational region are defined as

$$w = e^{i\theta}. \tag{4.68}$$

Additionally, the coordinates of the flat region are defined as

$$v = e^y \qquad \text{with} \qquad y = \sigma + i\tau \qquad \bar{y} = \sigma - i\tau. \tag{4.69}$$

As a result the following expressions

$$\log w = i\theta \qquad \log v = y = i\tau, \tag{4.70}$$

hold at the boundary. As the relation between τ and θ is encoded in the gluing map in this region, we can write the relation between w and v as

$$\log w = i\theta(-iy). \tag{4.71}$$

If one varies $\theta \to \delta\theta$, which is of our interest to derive the equation of motion, this relation changes to

$$\log w = i\theta(-iy) + i\delta\theta(-iy). \tag{4.72}$$

In order to avoid this, one can redefine the outside coordinates by an infinitesimal reparametrization

$$y \to \hat{y} = y + \xi^y. \tag{4.73}$$

We then choose ξ^y such that the relationship between the two coordinates is invariant under a variation with respect to $\theta(\tau)$. Explicitly, by filling in the reparametrization (4.73) in (4.71), one finds

$$\begin{aligned}\log w &= i\theta(-i\hat{y}) \\ &= i\theta(-iy) + i\theta'(-iy)\xi^2\end{aligned} \tag{4.74}$$

yields (4.72) when

$$\xi^y = i\frac{\delta\theta}{\theta'}. \tag{4.75}$$

Consequently, we can choose to define ξ^y as in (4.75) in order to keep the relation between the coordinates invariant. Importantly, we only want to change the outside coordinates (4.73) at the boundary. Therefore, we insert the Heaviside function $\Theta(\sigma)$[3] in the definition of ξ such that its contribution vanishes when $\sigma < 0$

$$\hat{\xi}^y = i\frac{\delta\theta(-iy)}{\theta'(-iy)}\Theta(\sigma) \qquad \hat{\xi}^{\overline{y}} = i\frac{\delta\theta(i\overline{y})}{\theta'(i\overline{y})}\Theta(\sigma) \tag{4.76}$$

where $\hat{\xi}^{\overline{y}}$ is the complex conjugate of $\hat{\xi}^y$.

With the use of (4.73) it is easy to see that this change in the coordinates is equivalent to the following change in the metric

$$ds^2 = dyd\overline{y} = d\hat{y}d\hat{\overline{y}} - 2\partial_{alpha}\xi^\beta d\hat{y}^\alpha d\hat{y}^\beta \qquad \text{such that} \qquad \delta g_{\alpha\beta} = -2\partial_{(\alpha}\xi_{\beta)}. \tag{4.77}$$

As the coordinate change defined in (4.73) holds only locally, we can interpret this $\delta g_{\alpha\beta}$ as the change in the metric due to a variation $\delta\theta$. As a result we can use the definition of the energy momentum tensor (3.55)

$$T_{\alpha\beta} = -\frac{2}{\sqrt{g}}\frac{\delta}{\delta g^{\alpha\beta}}\log Z_{\text{matter}}, \tag{4.78}$$

to find the explicit variation of the CFT action with respect to the boundary graviton. Furthermore, remember that the background metric of the outside region is flat, hence we know that the stress tensor is traceless (3.61). Using (4.77) we obtain

$$\begin{aligned}\delta\log Z_{\text{matter}} &= -\frac{1}{2}\int d\tau d\sigma(T_{yy}g^{yy} + T_{\overline{yy}}\delta g^{\overline{yy}}) \\ &= -2\int d\tau d\sigma(T_{yy}\partial_{\overline{y}}\xi^y + T_{\overline{yy}}\partial_y\xi^{\overline{y}}).\end{aligned} \tag{4.79}$$

The only thing that remains to do is evaluate the derivatives explicitly. The derivative only acts on the Heaviside function $\Theta(\theta)$ which gives a $\delta(\sigma)$ function. As a result, the integration over σ is simply

$$\delta\log Z_{\text{matter}} = i\int d\tau(T_{yy} - T_{\overline{yy}})\frac{\delta\theta(\tau)}{\theta(\tau)'}. \tag{4.80}$$

---

[3]We will plug this into an integral below.

Full equation of motion

We can combine the variation of the gravitational action with (4.80) to obtain the full equation of
motion corresponding to the variation of (4.52)

$$\frac{\phi_r}{8\pi}\frac{d}{d\tau}\left[\{e^{i\theta},\tau\} + U(\theta)\theta'^2\right] = i(T_{yy} - T_{\overline{yy}}). \tag{4.81}$$

As T_{yy} is given by (4.65), solving this equation consists of solving the conformal welding problem.

Before we move on to the next subsection, it might by useful to recapitulate the previous discussion. We started in section 4.2.1 by introducing the replica method in a gravitational context and derived the form of the full action on the quotient manifold $M_n = \frac{\tilde{M}_n}{Z_n}$. Additionally, we applied this method on the action corresponding to JT gravity. In section 4.2.2, we discussed the conformal welding problem and the geometry on this manifold M_n. This led to the expression of the physical stress tensor T_{yy} in terms of the welding maps from the disk topology towards the complex plane. To finish with, we derived the general equations of motion corresponding to $I_{full}[M_n]$ which yields 4.81.

   In the next subsection, we will combine all the obtained information to show that the condition (4.8) follows directly from our gravitational approach by looking at a specific configuration. Namely, we start by making n-copies of our system and explicitly transform this to the quotient geometry M_n. We will explicitly be interested in the case where we include a branch point " $-a$" in the gravitational region as this yields a n-fold replica. We will construct the corresponding equation of motion and see that the location of the quantum extremal surface can be obtained by solving this equation for $n \to 1$. This location will be compared to (4.35) to show that this approach is equivalent to the approach in section 4.1. Furthermore we show explicitly that (4.53) holds in our specific configuration.

4.2.4   Single interval at finite temperature

The first configuration that we consider is that of a single interval that contains one of the AdS$_2$ boundaries. In the Lorentzian signature, this corresponds to the configuration represented by Figure 4.3. In section 4.1.4, we obtained the location of the QES by naively extremizing the generalized entropy. Now, we are going to determine its location dynamically by solving the equation of motion in the Euclidean signature. In this setup, the interesting contribution to the path integral comes from a saddle point with a dynamical twist field at $y = -a$. This saddle point corresponds to the replica wormhole contribution discussed before.

   In what follows, we are going to use the replica method to construct the equation of motion for such n-folded replica geometry explicitly. This yields a complicated n-dependent equation with a similar form as (4.81), which is highly dependent on the welding maps F and G. However, we will be particularly interested in the limit $n \to 1$, where the equation of motion automatically yields the condition (4.34) and hence allows us to find the location of the QES dynamically. To end with, the explicit form of the gravitational action in this limit will be written down to show that (4.53) holds at finite temperature.

Equations of motion

Remember that our original system consists of an AdS$_2$ black hole glued to flat Minkowski space. We choose the simplest configuration of this system which corresponds to the topology of a disk as in the right picture of Figure 4.8. We make n-copies of this configuration and sew the replicas together along the branch points " $-a$" and "b", as depicted on the left of Figure 4.10 for $n = 2$. This corresponds to the insertion of twist fields at the conical singularities and " $-a$" on the quotient manifold on the right of the Figure 4.10.

   In order to obtain the equations of motion for an n-fold replica, we start by uniformizing the w-coordinate in a similar way as presented on Figure 3.5. Namely, we map the n-fold cover of the unit disk to the unit disk itself. Additionally, we set $\beta = 2\pi$ for a moment and restore it later

   Writing down the branch points as

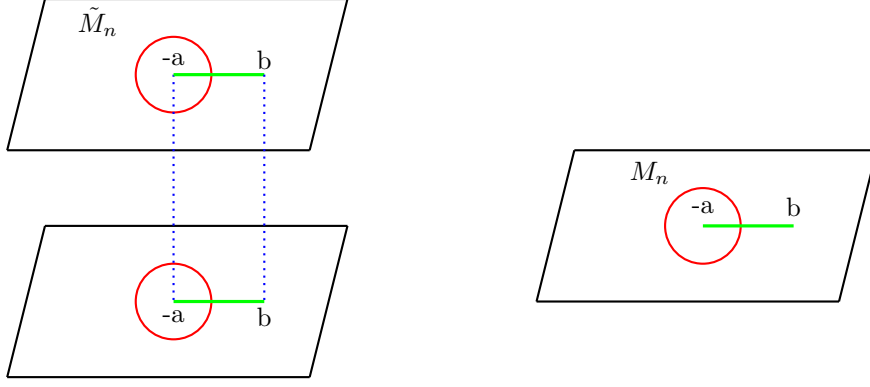$$w = A = e^{-a} \qquad v = B = e^b, \tag{4.82}$$

Figure 4.10: The left picture corresponds to $n = 2$ replicas, sewed together along the branch points $-a$ and b. The exact location of this point will be dynamically determined. This configuration corresponds to the replica wormhole structure of Figure 4.6. On the left we represent this configuration on the quotient manifold. Figure inspired by Figure 6 in [15].

one can uniformize the n copies by the following two transformations

$$w \rightarrow w' = \left(\frac{w - A}{1 - Aw} \right) \rightarrow \tilde{w} = (w')^{\frac{1}{n}}, \tag{4.83}$$

where the first transformation $w \rightarrow w'$ maps the branch point A to 0 and assures the point where the branch cut ends on the disk ($|w| = 1$), remains the same in the new coordinates.

By doing so, we can see our system as the standard unit disk without conical singularities, such that the geometry of the gravitational region is just AdS$_2$. As a result, the metric of this region is the hyperbolic disk

$$ds^2_{grav} = \frac{4|d\tilde{w}|^2}{(1 - |\tilde{w}|^2)^2}. \tag{4.84}$$

As discussed before, after filling in $R = -2$ in the JT gravity action, the dynamics are fully encoded in the variation of the boundary term. In analogy to our approach in the general case, we parametrize the boundary of the gravitational region by $\tilde{w} = e^{i\tilde{\theta}}$. Consequently, $\tilde{\theta}(\tau)$ arises as a boundary graviton and the relevant equation of motion is simply

$$\frac{\phi_r}{8\pi} \partial_\tau \{ e^{i\tilde{\theta}}, \tau \} = i(T_{yy}(i\tau) - T_{\overline{yy}}(-i\tau)), \tag{4.85}$$

where we also included the variation of the matter action (4.80).

Of course we need to go back to the w-plane. In contrary to approach in the previous subsection, the contribution of the conical singularities is included by using the Schwarzian decomposition identity [15]

$$\{ e^{i\tilde{\theta}}, \tau \} = \{ e^{i\theta}, \tau \} + \frac{1}{2} \left(1 - \frac{1}{n^2} \right) R(\theta)$$

$$R(\theta) = -\frac{(1 - A^2)^2 (\partial_\tau \theta)^2}{|1 - Ae^{i\theta}|^4}. \tag{4.86}$$

Plugging this into equation (4.85) one obtains

$$\frac{\phi_r}{8\pi} \partial_\tau \left(\{ e^{i\theta}, \tau \} + \frac{1}{2} \left(1 - \frac{1}{n^2} \right) R(\theta) \right) = i(T_{yy}(i\tau) - T_{\overline{yy}}(-i\tau)). \tag{4.87}$$

This puts the equation of motion in the same form of (4.81). In order to obtain the explicit form of the equations of motion in this configuration, we still need to find the stress tensor T_{yy} through the conformal welding problem (4.65). First note that, through the welding in Figure 4.9, we also map the twist points of the y-plane to the complex z-plane. However, we know that the stress tensor vanishes on the complex plane without any twist points. Therefore, we first remove these twist points

by the mapping $z \to \tilde{z} = z^{1/n}$. The stress tensor on the z plane is then obtained with the help of the conformal anomaly (4.65)

$$T_{zz}(z) = -\frac{c}{24\pi}\{z^{\frac{1}{n}}, z\} = -\frac{c}{48\pi}(1 - \frac{1}{n^2})\frac{1}{z^2} = -\frac{c}{48\pi}(1 - \frac{1}{n^2})\frac{1}{F(v)^2}. \qquad (4.88)$$

Finally, the welding map $(z \to v)$ needs to be inverted in order to obtain the physical stress tensor

$$T_{yy} = \left(\frac{dF(e^y)}{dy}\right)^2 T_{zz} - \frac{c}{24\pi}\{F(e^y), y\}. \qquad (4.89)$$

Or, using $v = e^y$

$$T_{yy} = e^{2y}\left[2(F'(v))^2 T_{zz} - \frac{c}{24\pi}\{F(v), v\}\right] - \frac{1}{2}. \qquad (4.90)$$

Inserting T_{zz} (i.e. 4.88) and using this in (4.87), we obtain the full equation of motion

$$\frac{6\pi\phi_r}{c\beta}\partial_\tau\left(\{e^{i\theta}, \tau\} + \frac{1}{2}\left(1 - \frac{1}{n^2}\right)R(\theta)\right) = ie^{2i\tau}\left[-\frac{1}{2}(1 - \frac{1}{n^2})\frac{F'(e^{i\tau})^2}{F(e^{i\tau})^2} - \{F, e^{i\tau}\}\right] + \text{complex conjugate} \qquad (4.91)$$

which is entirely written on the quotient manifold M_n and β-dependence is restored. Furthermore, n appears as a continuous parameter such that we may analytically continue this to non-integer n [15]. We will now show that taking the $n \to 1$ limit yields the condition (4.34).

Equation of motion at $n \sim 1$

Before we take a look at this limit, we discuss the situation $n = 1$. In this case, the manifold M_n is just the original geometry of the black hole glued to flat space. As a result, the gluing function $\theta(\tau)$ is trivial

$$\theta(\tau) = \tau, \qquad (4.92)$$

such that we can put $w = v$ everywhere and the conformal welding problem is trivial. In this case, it is convenient to set [15]

$$z = F(v) = \frac{v - A}{B - v} = G(w). \qquad (4.93)$$

As a result, we expect that in the limit $n \sim 1$, the gluing map $\theta(\tau)$ can be expanded around the trivial solution

$$\theta(\tau) = \tau + \delta\theta(\tau), \qquad (4.94)$$

where $\delta\theta$ is of order $n - 1$ [15].

At this point, we are interested in solving the equation (4.91) for the perturbation $\delta\theta$. To do so, we first need to write this equation in terms of this variable. As the right hand side depends on the welding map F, this also consists of solving the conformal welding problem in terms of $\delta\theta$.

Firstly, note that the maps F and G can be expressed as series expansions

$$G(w) = w + \sum_{l=0}^{\infty} g_l w^l \qquad F(v) = v + \sum_{-\infty}^{2} f_l v^l, \qquad (4.95)$$

due to the fact that they are holomorphic in their respective domain [68]. Furthermore, $e^{i\theta}$ can be expanded as a Fourier series, which we may do because there is a 2π periodicity in τ

$$\theta(\tau) = \tau + \sum_{m=-\infty}^{\infty} c_m e^{im\tau}. \qquad (4.96)$$

Plugging these expansions into the boundary matching condition

$$G(e^{i\theta(\tau)}) = F(e^{i\tau}), \qquad (4.97)$$

one can compare the coefficients of the series and write down the Schwarzian $\{F, v\}$ in terms of the perturbation $\delta\theta$. Namely, in appendix E.1 we show that

$$e^{2i\tau}\{F, v\} = -(\delta\theta''' + \delta\theta')_- = -\delta\{e^{i\theta}, \tau\}_-, \qquad (4.98)$$

where the minus subscript is used to indicate that we only used the negative frequency modes of $(\delta\theta''' + \delta\theta')$. This results from the series expansion of F, which only contains negative powers of v. This negative projection can be written with the help of the help of the Hilbert transform [4], defined as

$$H(e^{im\tau}) \equiv -sign(m)e^{im\tau} \tag{4.99}$$

and $H(1) = 0$. Consequently, we have that

$$e^{2i\tau}\{F, v\} = -\frac{1}{2}(1 + H)(\delta\theta''' + \delta\theta'). \tag{4.100}$$

This expression can be plugged into the equation of motion (4.91).

As the other terms containing F in the right hand side are being multiplied by $(n-1)$, inserting (4.95) in these terms would yield higher order contributions of $\delta\theta$. As a result, in these terms we simply use the zeroth order of the expansion (4.94) and use (4.93).

The equation of motion (4.91) is then fully written down in terms of the perturbation $\delta\theta$ as

$$\partial_\tau\left(\delta\theta''' + \delta\theta'\right) + \frac{ic}{3\phi_r}H(\delta\theta''' + \delta\theta') = (n-1)\left[\frac{c}{3\phi_r}\mathcal{F} - \partial_\tau R(\tau)\right] \tag{4.101}$$

where

$$\mathcal{F} = -i\frac{e^{2i\tau}(A - B)^2}{(e^{i\tau} - A)^2(e^{i\tau} - B)^2} + cc. \tag{4.102}$$

This equation can be solved by expanding both sides in Fourier series. By looking at (4.101), it is clear that the terms with the Fourier modes $e^{-ik\tau}$ with $k = -1, 0, 1$ will vanish (due to the $(k^2 - 1)k$ form of the coefficients which results from the derivatives). As a result, for these modes we have

$$\int_0^{2\pi} d\tau e^{-i\tau}\left(\frac{c}{3\phi_r}\mathcal{F} - \partial_\tau R(\tau)\right) = 0. \tag{4.103}$$

Solving the integrals as in appendix D.8, one can see that this *implies* the condition

$$\partial_a S_{gen} = 0 \qquad \sinh\left(\frac{2\pi a}{\beta}\right) = \frac{3\pi\phi_r}{\beta c}\frac{\sinh\left(\frac{\pi}{\beta}(a+b)\right)}{\sinh\left(\frac{\pi}{\beta}(a-b)\right)}, \tag{4.104}$$

which is exactly (4.34) which we found in section 2.5 by extremizing the generalized entropy. Note that once condition (4.104) is imposed, one is able to find a general solution of equation (4.91) in closed form, as shown in [68].

As mentioned in the introduction of this subsection, this is exactly what we wanted to show, i.e. by using the replica method for the gravitational JT theory, we can determine the location of the QES dynamically. This means that we obtain the same result as in section 4.1 without assuming AdS/CFT duality (in contrast to the derivation in [26], on which we relied in the previous section). However, we also want to show that this is not a coincidence. To do so, we will evaluate the full action (4.52) in the limit $n \sim 1$ and show that this corresponds to the generalized entropy (4.33). As a result, it is not a surprise that the equation of motion resulting from the variation of this action is equivalent to the condition (4.34). This will prove the argument made around equation (4.53), at least for the case where we include a single-interval.

Back to the generalized entropy

Remember that, we already argued that the variation of the gravitational action around $n \to 1$ should yield terms which correspond to the form of the generalized entropy. Therefore, one could already expect that the variation of the full action at $n \to 1$ should yield (4.34). Now, we can use our discussion in the finite temperature configuration to explicitly see this.

Filling in $R = -2$ into the gravitational action (4.55), one finds

$$-\frac{1}{n}I_{\text{grav}}[M_n] = \frac{\phi_r}{8\pi}\int_{\partial\Sigma_2} K, \tag{4.105}$$

---

[4]The appearance of this Hilbert transform may seem arbitrary but it has a natural interpretation in the Lorentzian signature, see [15] for more details.

where the first term results from I_{EH}. Based on our discussion in section 4.2.3, we know that (4.86) gives us the extrinsic curvature K. This yield the following explicit gravitational action for our configuration

$$-I_{\text{grav}}[M_n] = n\phi_0 + n\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\left[\{e^{i\theta},\tau\} + \frac{1}{2}\left(1 - \frac{1}{n^2}\right)R(\theta)\right].$$

(4.106)

We expect that evaluating this to leading order in $n - 1$ in the limit $n \approx 1$ yields the generalized entropy. To see this, first note that the contribution of I_{grav} in this limit is

$$-\log Tr(\rho_R)^n \approx -I_{\text{grav}}[M_n] + nI_{\text{grav}}[M_1].$$

(4.107)

For the term containing the Schwarzian, this becomes for $n \approx 1$

$$(n-1)\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\left(\{e^{i\theta},\tau\}\big|_n - \{e^{i\theta},\tau\}\big|_{n=1}\right)$$

$$\approx (n-1)\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\partial_n\{e^{i\theta},\tau\}.$$

(4.108)

For the term containing $R(\theta)$ we note that $I_{\text{grav}}(1) = 0$ and hence, the only contribution comes from $I_{\text{grav}}[M_n]$ expanded around $n \approx 1$

$$\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\frac{1}{2}n\left(1 - \frac{1}{n^2}\right)R(\theta)$$

$$\approx (n-1)\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau R(\theta)$$

$$\approx (n-1)\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau R(\tau)$$

(4.109)

where we used the fact that $R(\theta)$, defined in (4.86) can be expanded around θ. As $\partial\theta$ is already of order $n - 1$, we only kept $R(\tau)$ since we only keep the leading order in $n - 1$.

Putting all the gravitational contributions together, one obtains

$$-I_{\text{grav}}[M_n] + nI_{\text{grav}}[M_1] \approx (n-1)\left(\phi_0 + \frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\partial_n\{e^{i\theta},\tau\} + \frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau R(\tau)\right).$$

(4.110)

By evaluating the integral

$$\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau R(\tau) = -\frac{\phi_r}{4\tanh(a)}$$

(4.111)

which has been done explicitly in appendix D.7, this just the value of the dilaton evaluated at the branch point (a). These two terms thus already correspond to the first two terms of the generalized entropy (4.33). Note that we have set $\beta = 2\pi$ in our derivation.

We still need to take a look a the matter contribution to the action. By using the CFT Ward identity for $\partial_g \log Z_{matter}$, this turns out to be ([15],[68])

$$\log Z^{mat}[M_n] - n\log Z^{mat} = -(n-1)S_{vN}([-a,b]) + \delta_g \log Z_{matter},$$

(4.112)

where the second term results from an order (n-1) change in the metric. Evaluating the equation of motion (4.91) at $n = 1$ sets this contribution equal to $\frac{\phi_r}{8\pi}\int_0^{2\pi}d\tau\partial_n\{e^{i\theta},\tau\}$. Consequently, in the limit $n \approx 1$

$$-\log Tr(\rho_R)^n \approx -I_{\text{grav}}[M_n] + I_{\text{grav}}[M_1] + \log Z^{mat}[M_n] - n\log Z^{mat}$$

$$\approx (n-1)\left(\phi_0 + \frac{-\phi_r}{\tanh(a)} + S_{vN}([-a,b])\right)$$

$$\approx (1-n)S_{gen}[-a,b],$$

(4.113)

where we used (4.33) for $\beta = 2\pi$.

This is exactly what we obtained in (4.53) and explains why the condition (4.34) follows directly from the equations of motion. As a result, we have shown that one can obtain the explicit form of the generalized entropy with the help of gravitational methods without holography. Note that one can obtain the $T = 0$ zero results from section 4.1 by setting $\beta \to \infty$.

4.2.5 Eternal black hole and replica wormholes

We have shown that replica wormhole saddles yield the same contributions to the fine-grained entropy (2.65) as derived in section 4.1. However, this result was only achieved for the simplified case where only a single interval $[-a, b]$ is included. Remember from section 4.1.3 that we need to take into account two intervals to address the information paradox for the eternal black hole. In section 4.1.4, we have used the arguments in [66] and [68] to obtain a fine-grained entropy which is consistent with unitarity. These stated that at late times, the generalized entropy which includes two intervals simply corresponds to the double of the single-interval result. As stated in [68], this resulted from the fact that the twist operators are in the "OPE"-limit. However, we would like to know how and *if* these arguments can be justified by our gravitational method. If that is the case, we have found a more general resolution of the information paradox formulated for eternal black holes. Unfortunately, a full understanding of the two interval replica manifold has not been achieved yet [[15], [68]]. However, physical arguments are given in [15], which are worked out in [68]. We will briefly discuss the main idea behind the reasoning in [68].

The main complication that arises when considering two intervals instead of a single interval is that new equations of motion arise, beyond the Schwarzian equation. Namely, when branching the replicas along a single interval, the resulting topology of the gravitational region on the corresponding quotient manifold remains the same, i.e. it is a disk for all n. In contrast, the branching along two intervals results in a gravitational region with two branch points. As stated in [68], the resulting manifold has wormholes with higher topology, which introduce new equations of motion. Consequently, it is not longer enough to solve only the Schwarzian equation (4.81) and factorize the result. However, according to [68], these extra equations are irrelevant at late times. Therefore, it is shown in [68] that one can recover the QES derived in the previous section (4.44) by gluing together two copies of the single-interval solution in the Euclidean regime.

This would mean that the non-holographic approach is also able to resolve a partial information paradox for eternal black holes, hence generalizing the result of [26]. However, a detailed understanding of the two-interval computation with the gravitational replica method has not yet been obtained [68].

4.3 Resolution of the information paradox?

A question that possibly arises after reading this thesis is; what do these gravitational methods imply for the information paradox that was introduced in chapter 2? To partially answer this question, we briefly discuss some of the remarks that can be made about the results presented in this chapter.

First, remember that in section 4.1, we used a formula for the fine-grained entropy of the radiation which initially came from holographic computations [13]. We saw that this island-prescription for the fine-grained entropy resolves the tension with unitarity for the simplified setup of an eternal black hole [66]. However, there remains a debate about whether information escapes black holes in our universe. Namely, the Ryu-Takayanagi formula and its generalizations are especially well understood in the context of AdS/CFT [13] but it would be interesting to study black holes more generally, based on less assumptions about the true nature of the universe.

However, in section 4.2, we justified the computations used in section 4.1 by deriving the same results without the use of holography. Unfortunately, as mentioned before, a full understanding of the two-interval factorization property in the context of the replica method has not been obtained yet [68]. Additionally, the derivation in section 4.2 was entirely performed in only 2-dimensions, which makes many physicists very uncomfortable [11]. Furthermore, a lot of question marks arise when using gravitational Euclidean path integrals. For example, one could wonder how to know which saddles one must include in the computation or ask questions about the actual integration contour [8]. Additionally, we did not address the replica method on evaporating black holes which are formed by gravitational collapse, for which Hawking's original computations result in the information paradox. However, a better understanding of the application of the gravitational replica method on black holes formed by collapse was obtained in [68].

Unfortunately, even if we would be able to obtain a correct entanglement entropy of the radiation corresponding to realistic black holes, we would not have solved the Hawking information paradox introduced in 2.2. Namely, if the entanglement entropy follows the Page curve, and hence, is consistent with unitarity, we would only be able to conclude that *information does come out of the black hole.* In contrast, this result would not provide an explanation on *how* infalling matter escapesthe black hole into Hawking radiation [15]. Namely, finding the correct entropy only gives us the trace of the density matrix but not a formula for its individual matrix elements elements [8].

Interestingly, there also exist many arguments in the favor of information loss, many of them are described in [74]. In this paper, it is argued that the evolution from a pure into a mixed state does not violate any fundamental principle of physics. Furthermore, the presented results in this thesis do not address the case in which the black hole does not evaporate completely and results in remnants. Remember from section 2.2 that this is also a possibility.

Even though many other remarks could be made and even more questions are left unanswered, the developments concerning the fine-grained entropy have given us a lot of new insights towards the fundamental characteristics of nature. Next to providing interesting insights into the quantum mechanics of black holes, these studies involved the combination of many different fields of physics, e.g. quantum information, high energy physics, gravitation and many more [8]. As explicitly stated in [8], "*these developments have turned black holes into a light that illuminates many questions in many fields in physics.*

Conclusion

The aim of this thesis was to provide a self-contained review on the subject of the black hole information paradox and recent developments towards its resolution. Especially, the discussion was centered around a partial problem which was first highlighted by Page, consisting of an entanglement entropy which is not consistent with unitarity. To solve this, we discussed an approach that mainly relies on the gravitational path integral and the replica method.

In chapter 1, we started by reviewing some important notions from classical black holes. Above all, we discussed the similarities between the dynamical laws of black holes and the laws of thermodynamics. This already suggested that black holes had a temperature and an entropy, in contrast to what one expects from general relativity. Additionally, we summarized some basic concepts from quantum mechanics and introduced the von Neumann and entanglement entropy in section 1.3. These provided the necessary tools towards the formulation of the semi-classical approximation at the end of the first chapter. This framework allowed us to understand the black hole radiation process, which was analysed in sections 2.1 and 2.2. Especially, we have shown that the evolution of the black hole evaporation process ultimately results in the information paradox. In order to highlight the paradoxical behaviour, we formulated Hawking's conclusions as a theorem. This illustrated that, starting from Hawking's assumptions and semi-classical physics, one is forced to the conclusion that semi-classical physics breaks down when the black hole has evaporated (almost) completely. Namely, the pure state, from which one assumes it was formed, turns into a mixed state (or remnants). The first possibility seems to result in information loss and contradicts our belief that quantum mechanical processes are unitary in time.

In section 2.3, we saw that the amount of information loss can be quantified by the entanglement entropy of the radiation. This quantification led to the conclusion that Hawking's calculation already yields problematic, i.e. non-unitary behaviour, way before one expects quantum gravitational effects to become important. As a result, the entanglement entropy provided a concrete manner to check if information was indeed lost or not. Namely, if one could find a formula for the fine-grained entropy of the radiation which follows the Page curve, the amount of information is conserved and black hole evaporation is a unitary process.

The first approach towards the derivation of such a formula resulted from holographic computations [13]. However, this thesis focused on a method which, among other things, employs Euclidean path integrals and the replica trick, without the need of holographic assumptions. In order to understand how these gravitational methods are used to obtain a correct entropy, we had to introduce several "ingredients" in chapter 3. Explicitly, we started with a discussion about the relation between the Euclidean path integral and the thermal partition function, from which one can obtain thermodynamical quantities such as the entropy. At the end of section 3.1 we applied this relation to the Euclidean Schwarzschild black hole and derived the area law of the black hole entropy, yielding the same result as derived in section 2.1.2.

Next, in section 3.2, we considered the matter side of the story by introducing 2-dimensional conformal field theories. These form a special subset of ordinary quantum field theories and turn out to be extremely useful to describe matter fields in the setup consisting of black holes in AdS_2. Before we explicitly saw this, we introduced the replica trick to compute the von Neumann entropy in a two-dimensional CFT in section 3.2.4.

A conformal field theory was coupled to a gravitational theory for the first time in section 3.4. The specific gravitational action we considered to do this was the two-dimensional JT-gravity action. A classical solution of the equations of motion, which result from the variation of the JT action, is the AdS_2 spacetime. Some properties of this spacetime were discussed in more detail in section 3.3. Namely, we introduced the different coordinates and corresponding patches which describe AdS_2.

Furthermore, we have shown that the gravitational dynamics are fully encoded in the Schwarzian equation.

At the end of chapter 3, we discussed how one can work in the semi-classical approximation with gravitational path integrals. As explained in chapter 1, this yielded a formalism where one can treat the matter fields in a quantum mechanical way while considering classical gravity.

Finally, in chapter 4, we used all these elements to obtain explicit expressions for the fine-grained entropy corresponding to a simplified model of a black hole. This consists of an AdS$_2$ black hole coupled to a 2-dimensional conformal field theory. Additionally, the gravitational region is coupled to flat Minkowski space where the same CFT is defined. In section 4.1, we studied this model in the Lorentzian signature and illustrated how an information paradox can be formulated for eternal black holes. Although these must be distinguished from black holes formed by collapse, the corresponding information problem is conceptually similar to the paradox introduced in chapter 2. Next, we used the general formula for the gravitational fine-grained entropy to show that the inclusion of islands partially resolves the simplified version of the paradox.

While we simply assumed the general expression of the fine-grained entropy in section 4.1, we explicitly derived this formula in section 4.2. To do so, we considered the same model but in the Euclidean signature. This was necessary to make a connection with the thermal partition function, as described in section 3.1. In order to compute the gravitational path integral, we replicated the system consisting of AdS$_2$ glued to flat space. To simplify this computation, we followed the same reasoning as discussed in section 3.2.4 and considered the quotient manifold instead of the replicated manifold. The physics on this simpler manifold was discussed in section 4.2.2 and consists mainly of the conformal welding and adapted equations of motions.

As the geometry in the gravitational region is not fixed, one can evaluate the gravitational path integral as saddles which correspond to different geometries. In terms of replicated manifolds these saddles represent different ways of connecting the different copies of the original system to each other. On the level of the quotient manifold, this results in different locations of branch points and corresponding conical singularities.

We were especially interested in a non-trivial type of saddles. These correspond to geometries where the different replicas are glued together along the gravitational regions by so called replica wormholes. As illustrated in the last section, these saddles yield contributions which are equivalent to the island-contributions discussed in section 4.1.

In order to obtain the explicit form of the entropy resulting from replica wormholes, we had to solve the equation of motion, which is highly dependent on the boundary graviton and the welding maps, making the computation non-trivial. Luckily, we were especially interested in the limit where the number of replica's was $n \sim 1$. This allowed us to solve the equation in terms of perturbations and approximations around the trivial $n = 1$ case. For the finite temperature configuration, we saw that the condition to find the quantum extremal surface (studied in section 4.1) followed directly from the gravitational equations. As explained at the end of section 4.2.4, this is simply a result from the fact that the full action, for $n \sim 1$, contains the same terms as the generalized entropy proposed in section 4.1. As a result, it is not surprising that we found the QES condition by the variation of this action in that limit.

However, we only considered the case where a single-interval is taken into account. In order to address the information paradox formulated for an external black hole, one needs to take into account two-intervals. A full understanding of the corresponding computation which relies on the replica method has not yet been obtained.

Unfortunately, even if we knew how to extend the gravitational methods presented in this thesis towards the derivation of the entropy for black holes formed by collapse, this would not be enough to solve the complete information problem. Namely, from the consistency of the black hole evaporation process with unitarity, one can only conclude that information *does* come out. However, one can still not explain *how* it does. To understand this, one should be able to obtain the individual elements of the density matrix, describing the full state consisting of the black hole and the radiation, at all times. Even though many other remarks can be made and even more questions can be asked, the gravitational fine-grained entropy has given us many new insights about the quantum mechanics of black holes and other fields in physics.

Appendix A

Entanglement entropy

In this appendix, the entanglement entropy of bipartite systems is calculated explicitly. These examples come from [10] but the relevant steps are worked out below.

A.1   Proof of statement (1.76)

In order to prove equation (1.76) one should first introduce the *Schmidt decomposition theorem*. We present the approach in [29].

This theorem states that, if one considers a pure state $|\Psi\rangle \in \mathcal{H}_{\mathcal{AB}}$, one can always find orthonormal bases of $\mathcal{H}_A$ (denoted by $\{|u_i\rangle_A\}_{i=1}^{d_A}$) *and* $\mathcal{H}_B$ (denoted by $\{|v_j\rangle_B\}_{j=1}^{d_B}$) such that the full state Ψ_{AB} can be decomposed as

$$\Psi_{AB} = \sum_{i=1}^{n} \sqrt{p_i}\, |u_i\rangle_A \otimes |v_j\rangle_B\,, \tag{A.1}$$

where $n = \min\{d_A, d_B\}$ and $\sum_{i=1}^{n} p_i = 1$. The corresponding *Schmidt coefficients* are the square roots of the eigenvalues p_i of the reduced density matrices of ρ_A and ρ_B in their orthonormal bases. As both density matrices have the same eigenvalues, their entropy's $S_{vN,A}$ and $S_{vN,B}$ are clearly equal to each other [35]. The Schmidt decomposition theorem thus automatically implies statement (1.75). Furthermore, if the pure state has the following Schmidt decomposition

$$\Psi_{AB} = |u\rangle_A \otimes |v\rangle_B\,, \tag{A.2}$$

which is possible when there is only one non-zero Schmidt coefficient $p_i = 1$ while $p_j = 0\ \forall j \neq i$, then the substates are both pure. Namely, as the p_i's are the eigenvalues of the reduced density matrices, we have that

$$S_{vN,A} = S_{vN,B} = 0. \tag{A.3}$$

This proofs one implication of statement (1.76). Conversely, if A.3 holds, one states A and B are pure such that the reduced density matrices correspond to

$$\rho_A = |u\rangle_A \langle u|_A \qquad \rho_B = |v\rangle_B \langle v|_B\,. \tag{A.4}$$

As it is assumed that $|\Psi\rangle_{AB}$ is pure, the Schmidt decomposition of the total state must correspond to (A.2).

A.2   Entanglement entropy of state I

To capture which departures from the state (2.31) are allowed, we compute the entanglement entropy of

$$\langle\psi| = \langle\psi|_M \otimes \frac{1}{\sqrt{2}}(\langle 0|_r \langle 0|_b + \langle 1|_r \langle 1|_b) \tag{A.5}$$

and compare it to the entanglement entropy of two departures from this state in appendices A.3 and A.4. Although only the form of these states is written down in [15], we have worked out the entanglement entropy here.

First calculate the full density matrix

$$\rho = \frac{1}{2}(|\psi\rangle_M |0\rangle_r |0\rangle_b + |\psi\rangle_M |1\rangle_r |1\rangle_b)(\langle 0|_b \langle 0|_r \langle\psi|_M + \langle 1|_b \langle 1|_r \langle\psi|_M)$$

As

$$\langle\psi|_M = \frac{1}{\sqrt{2}}\Big(\langle\uparrow|_M + \langle\downarrow|_M \Big)$$

we see that

$$\rho_{c,b} = \mathrm{Tr}_M \rho = \frac{1}{2}\Big(|0\rangle_r |0\rangle_b \langle 0|_b \langle 0|_r + |0\rangle_r |0\rangle_b \langle 1|_b \langle 1|_r + |1\rangle_r |1\rangle_b \langle 0|_b \langle 0|_r + |1\rangle_r |1\rangle_b \langle 1|_b \langle 1|_r \Big)$$

where we used that

$$\langle\uparrow|\uparrow\rangle_M = \langle\downarrow|\downarrow\rangle_M = 1 \tag{A.6}$$
$$\langle\downarrow|\uparrow\rangle_M = \langle\uparrow|\downarrow\rangle_M = 0.$$

If we now take another partial trace, we find

$$\rho_b = \mathrm{Tr}_a \rho_{r,b} = \frac{1}{2}\Big(|0\rangle_b \langle 0|_b + |1\rangle_b \langle 1|_b \Big) \tag{A.7}$$

where we used the orthonormality of the c states. If we now represent the brackets by the following matrices

$$\langle 0|_b = (1 \quad 0) \qquad |0\rangle_b = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \langle 1|_b = (0 \quad 1) \qquad |1\rangle_b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

we find that $\rho_b = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The entanglement entropy is defined as

$$S_{vN,b} = -\mathrm{Tr}\rho_b \ln \rho_b \tag{A.8}$$

As ρ_b is a diagonal matrix we can write this as

$$S_{vN,b} = -\sum_{i=1}^{2} \lambda_i \log \lambda_i = -\frac{1}{2}\log\Big(\frac{1}{2}\Big) - \frac{1}{2}\log\Big(\frac{1}{2}\Big) = \log(2), \tag{A.9}$$

where λ_i are the eigenvalues, e.g. diagonal elements of the reduced density matrix. In the next subsections, we will compare this to the entanglement entropy's of two other states which represent small deviations from (2.31). In order for the state to remain local, only deviations which lead to the same entanglement entropy as (A.9) are allowed.

A.3   Entanglement entropy of state II

We are now going to compare the entanglement entropy of

$$|\psi\rangle = |\psi\rangle_M \otimes \left((\frac{1}{\sqrt{2}} + \epsilon) \langle 0|_r \langle 0|_b + (\frac{1}{\sqrt{2}} - \epsilon) \langle 1|_r \langle 1|_b \right) \tag{A.10}$$

with the entropy corresponding to a small deviation from that state.
The computation is exactly the same apart from the $(\frac{1}{\sqrt{2}} + \epsilon)$ coefficient. We now have

$$\rho_b = \begin{pmatrix} \frac{1}{2}(\frac{1}{\sqrt{2}} + \epsilon)^2 & 0 \\ 0 & \frac{1}{2}(\frac{1}{\sqrt{2}}) - \epsilon)^2 \end{pmatrix}$$

such that

$$S_{vN,b} = -\sum_i^{2} \lambda_i ln(\lambda_i) = -\frac{1}{2}(\frac{1}{\sqrt{2}} + \epsilon)^2 \log\Big(\frac{1}{2}(\frac{1}{\sqrt{2}} + \epsilon)^2\Big) - \frac{1}{2}(\frac{1}{\sqrt{2}} - \epsilon)^2 \log\Big(\frac{1}{2}(\frac{1}{\sqrt{2}} - \epsilon)^2\Big) \tag{A.11}$$

We can expand the logarithm's as $(1 + x) \approx x - \frac{x^2}{2} + O(x^3)$ for $x << 1$.

$$\log\left(\frac{1}{2}(\frac{1}{\sqrt{2}} + \epsilon)^2\right) = \log\left(\frac{1}{2}\right) + \log\left((\frac{1}{\sqrt{2}} + \epsilon)^2\right)$$

$$= \log\left(\frac{1}{2}\right) + 2\log\left(\frac{1}{\sqrt{2}}(1 + \sqrt{2}\epsilon)\right)$$

$$\approx -2\log(2) + 2\sqrt{2}\epsilon - 2\epsilon^2.$$

Similarly

$$\log\left(\frac{1}{2}(\frac{1}{\sqrt{2}} + \epsilon)^2\right) \approx -2\log(2) - 2\sqrt{2}\epsilon - 2\epsilon^2$$

Using this in we find

$$S_{vN,b} = \log(2) + \epsilon^2 - 4\epsilon^2 + 2\epsilon^2\log(2) = \log(2) - \epsilon^2(3 - 2\log(2)) \approx \log(2) \tag{A.12}$$

Which means that this state captures the same amount of entanglement as (2.31). Consequently, departures from the state 2.31) as the one discussed in this appendix are allowed.

A.4 Entanglement entropy of state III

Finally, we compute the entanglement entropy of

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_M |0\rangle_r + |\downarrow\rangle_M |1\rangle_r\right) \otimes \frac{1}{\sqrt{2}}(|0\rangle_b + |1\rangle_b) \tag{A.13}$$

which represents another type of departure from (2.31).
Note that we can write the partial density in a more general way

$$\rho_b = \sum_j^{d_c} \sum_i^{d_M} \langle j|_c \langle i|_M \left(\sum_{k,l}^{d_{c,M}} \sum_{a,b}^{d_b} |k\rangle_M |k\rangle_c |a\rangle_b \langle b|_b \langle l|_c \langle l|_M\right) |i\rangle_M |j\rangle_c \tag{A.14}$$

When using the orthonormality of the $|\psi\rangle_M$ states we see,

$$\rho_b = \sum_j^{d_c} \sum_{k,l}^{d_{c,M}} \sum_{a,b}^{d_b} \langle j|_c \left(\delta_k^i |k\rangle_c |a\rangle_b \langle b|_b \langle l|_c \delta_i^l\right) |j\rangle_c \tag{A.15}$$

and after using the orthonormality of the c states

$$\rho_b = \sum_{k,l}^{d_{c,M}} \sum_{a,b}^{d_b} \delta_k^i \delta_j^k |a\rangle_b \langle b|_b \delta_l^k \delta_i^l \tag{A.16}$$

or

$$\rho_b = \sum_{a,b}^{dim_B} |a\rangle_b \langle b|_b \tag{A.17}$$

This gives us

$$\rho_b = \frac{1}{2}\left(|0\rangle_b \langle 0|_b + |0\rangle_b \langle 1|_b + |1\rangle_b \langle 0|_b + |1\rangle_b \langle 1|_b\right)$$

Furthermore, calculating the outer products,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 1(0 & 1) \\ 0(0 & 1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Such that one obtains

$$\rho_b = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

after computing the outer product for the other components.

In order to use the same formula for the entanglement matrix, we need to change to the basis in which the matrix is diagonal. To do so we start with solving the characteristic equation of the partial density matrix

$$\det(\rho_b - \lambda I_{2x2}) = \lambda(\lambda - 1). \tag{A.18}$$

From this we know that the eigenvalues in the new basis are 1 and 0. This allows us to compute

$$S_b = -\sum_{i=1}^{2} \lambda_i ln(\lambda_i) = -1 ln(1) - 0 ln(0) = -ln(1) = 0. \tag{A.19}$$

The entanglement entropy for this state is zero. As a result, we conclude that states of this type do not capture the entanglement between the red and blue particles.

A.5   Entropy lemmas

In this section, we proof the lemmas used to proof theorem 2.2.2 about the monotonically increasing entropy of the radiation. To do so, we present the approach in [10] but work out the details not included in this paper. Sometimes a modification towards the approach in this paper was made. Before we move on to that, we introduce the following notation:

- $\{r\}_i$ refers to the emitted Hawking radiation at time t_i. Explicitly, it refers to all the outgoing quanta $r_1, r_2, ..., r_i$.

- $\{b\}_i$ refers to all the absorbed ingoing quanta at time t_i, e.g. $b_1, b_2, ..., b_i$.

- $\{b_{n+1}, r_{n+1}\}$ refers to the newly created particle pair created at t_{n+1}, where r_{n+1} is the outgoing and b_{n+1} the ingoing particle.

- ρ_{full_i} refers to the density matrix of the full system consisting over the matter of the black hole M, the ingoing quanta $\{b\}_i$, the radiation $\{r\}_i$ at time t_i.

- The entropy $S_{particles_i}$ refers to the entanglement entropy between a subsystem of particles and the remaining subsystems of the black hole, at time t_i. The corresponding reduced density matrix is $\rho_{particles_i}$.

Note that these lemma's are proofed in [10] but additional steps and sometimes other approaches are written down here.

Lemma A.5.1. *If condition (2.45) holds, then the entanglement of the newly created pair, (r_{n+1}, b_{n+1}) with the rest of the system is bounded as*

$$S_{\{r_{n+1}, b_{n+1}\}} \equiv -\operatorname{Tr} \rho_{\{r_{n+1}, b_{n+1}\}} \ln \rho_{\{r_{n+1}, b_{n+1}\}} < \epsilon \tag{A.20}$$

Proof. With the use of formula (1.70) and the orthonormality of the states, one finds that the reduced density matrix of the newly created particle pair (b_{n+1}, r_{n+1}) is

$$
\begin{aligned}
\rho_{\{r_{n+1}, b_{n+1}\}} &= \operatorname{Tr}_{\{b\},\{r\},M} \rho_{full_{n+1}} \\
&= \operatorname{Tr}_{\{b\},\{r\},M} |\psi_{full}(t_{n+1})\rangle \langle \psi_{full}(t_{n+1})| \\
&= \sum_{i,j}^{2} \langle S^i | \psi_{full}(t_{n+1})\rangle \langle \psi_{full}(t_{n+1}) | S^i \rangle \\
&= \begin{pmatrix} \langle \Lambda^{(1)} | \Lambda^{(1)} \rangle & \langle \Lambda^{(1)} | \Lambda^{(2)} \rangle \\ \langle \Lambda^{(2)} | \Lambda^{(1)} \rangle & \langle \Lambda^{(2)} | \Lambda^{(2)} \rangle \end{pmatrix},
\end{aligned} \tag{A.21}
$$

where $|\psi_{full}(t_{n+1})\rangle$ refers to the state (2.44). The second equality was obtained by noting that $|\psi_{M,b}, \psi_r(t_{n+1})\rangle$ is pure and the third equality is a result of the orthonormality of the S^i states. As condition (2.45) holds, we have that

$$\left\| \Lambda^{(2)} \right\|^2 = \langle \Lambda^{(2)} | \Lambda^{(2)} \rangle \equiv \epsilon_1 < \epsilon^2. \tag{A.22}$$

Furthermore, with the use of the Schwartz inequality

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle, \tag{A.23}$$

and the condition (2.46) one can show that

$$|\left\langle \Lambda^{(1)}\middle|\Lambda^{(2)}\right\rangle| \leq \left\langle \Lambda^{(1)}\middle|\Lambda^{(1)}\right\rangle\left\langle \Lambda^{(2)}\middle|\Lambda^{(2)}\right\rangle \equiv \epsilon_2 = (1-\epsilon_1)\epsilon_1 = \epsilon_1 - \epsilon_1^2 < \epsilon_1 < \epsilon. \tag{A.24}$$

As a result, we can write the density matrix as

$$\rho_{\{r_{n+1},b_{n+1}\}} = \begin{pmatrix} 1 - \epsilon_1^2 & \epsilon_2 \\ \epsilon_2 & \epsilon_1^2 \end{pmatrix}. \tag{A.25}$$

By ignoring $O(\epsilon^4)$ contributions, the corresponding eigenvalues are found to be

$$\lambda_1 = (\epsilon_1^2 - \epsilon_2^2) \qquad \lambda_2 = 1 - (\epsilon_1^2 - \epsilon_2^2) = 1 - \lambda_1. \tag{A.26}$$

The resulting entropy is thus given by

$$S_{\{r_{n+1},b_{n+1}\}} = -\sum \lambda_i \ln \lambda_i = -\lambda_1 \ln \lambda_1 - (1-\lambda_1)\ln(1-\lambda_1). \tag{A.27}$$

As $\lambda_1 = (\epsilon_1^2 - \epsilon_2^2) < \epsilon << 1$ one can expand this in a generalized Puiseux series leading to

$$\begin{aligned}
S_{\{r_{n+1},b_{n+1}\}} &\approx \lambda_1(1 - \ln \lambda_1) - \frac{\lambda_1^2}{2} - \frac{\lambda_1^3}{6} + O(\lambda^4 \\
&= (\epsilon_1^2 - \epsilon_2^2)(1 - \ln(\epsilon_1^2 - \epsilon_2^2) + O(\epsilon^3) \\
&= (\epsilon_1^2 - \epsilon_2^2)\ln \frac{e}{\epsilon_1^2 - \epsilon_2^2} + O(\epsilon^3) \\
&< \epsilon.
\end{aligned} \tag{A.28}$$

This proves the Lemma and shows that the newly created particle pair is only weakly entangled with the remainder of the system. $\qquad\square$

Lemma A.5.2. *The entanglement entropy, $S_{\{r\}_n,\{r_{n+1},b_{n+1}\}}$, corresponding to the radiation emitted at t_n and the newly created pair (b_{n+1}, r_{n+1}) obeys the following condition*

$$S_{\{r\}_n,\{b_{n+1},r_{n+1}\}} \geq S_{\{r\}_n} - \epsilon. \tag{A.29}$$

Proof. First note that the subaddivity property relating the entropy of 2 systems A and B yields

$$S(A+B) \geq |S(A) - S(B)|. \tag{A.30}$$

Setting

$$S_{\{r\}_n,\{r_{n+1},b_{n+1}\}} = S(A+B) \qquad S_{\{r\}_n} = S(A) \qquad S_{\{b_{n+1},r_{n+1}\}} = S(B) \tag{A.31}$$

this yields

$$\begin{aligned}
S_{\{r\}_n,\{r_{n+1},b_{n+1}\}} &\geq |S_{\{r\}_n} - S_{\{r_{n+1},b_{n+1}\}}| \\
&\geq |S_{\{r\}_n} - \epsilon| \\
&\geq S_{\{r\}_n} - \epsilon,
\end{aligned} \tag{A.32}$$

where we used Lemma A.5.1 to obtain the second inequality. To obtain the last inequality, we simply used the fact that

$$|a| \leq b \Leftrightarrow -b \leq a \leq b. \tag{A.33}$$

Setting $b = S_{\{r\}_n,\{r_{n+1},b_{n+1}\}}$, this simply yields the last line in (A.32), which proofs the second Lemma. $\qquad\square$

Lemma A.5.3. *The entanglement of the ingoing quanta created at t_{n+1}, denoted by b_{n+1}, obeys the following inequality*

$$S_{b_{n+1}} > \ln 2 - \epsilon. \tag{A.34}$$

Proof. Firstly, remember that the full quantum state at t_{n+1} can be written as (2.44). Explicitly writing down the orthonormal states of the newly created particle pair $\{r_{n+1}, b_{n+1}\}$ as

$$S^{(1)} = \frac{1}{\sqrt{2}} \left(|0\rangle_{r_{n+1}} |0\rangle_{b_{n+1}} + |1\rangle_{r_{n+1}} |1\rangle_{b_{n+1}} \right) \qquad S^{(2)} = \frac{1}{\sqrt{2}} \left(|0\rangle_{r_{n+1}} |0\rangle_{b_{n+1}} - |1\rangle_{r_{n+1}} |1\rangle_{b_{n+1}} \right), \tag{A.35}$$

allows us to write the full quantum state (2.44) as

$$\langle \psi_{full}(t_{n+1})| = \frac{1}{\sqrt{2}} \left[|0\rangle_{r_{n+1}} |0\rangle_{b_{n+1}} (\Lambda^{(1)} + \Lambda^{(2)}) + |1\rangle_{r_{n+1}} |1\rangle_{b_{n+1}} (\Lambda^{(1)} - \Lambda^{(2)}) \right]. \tag{A.36}$$

With the same technique as used in (A.5.1), one obtains the reduced density matrix

$$\rho_{r_{n+1}} = \begin{pmatrix} \frac{1}{2} \langle (\Lambda^{(1)} + \Lambda^{(2)}) | (\Lambda^{(1)} + \Lambda^{(2)}) \rangle & 0 \\ 0 & \frac{1}{2} \langle (\Lambda^{(1)} - \Lambda^{(2)}) | (\Lambda^{(1)} - \Lambda^{(2)}) \rangle \end{pmatrix}. \tag{A.37}$$

From equations (A.22) and (A.24) one can see that

$$\frac{1}{2} \left\langle (\Lambda^{(1)} + \Lambda^{(2)}) \middle| (\Lambda^{(1)} + \Lambda^{(2)}) \right\rangle = \frac{1}{2}(1 - \epsilon_1^2 + 2\left\langle \Lambda^{(1)} \middle| \Lambda^{(2)} \right\rangle + \epsilon_1^2) = \frac{1}{2} + \left\langle \Lambda^{(1)} \middle| \Lambda^{(2)} \right\rangle = \frac{1}{2} + \epsilon_2$$

$$\frac{1}{2} \left\langle (\Lambda^{(1)} - \Lambda^{(2)}) \middle| (\Lambda^{(1)} - \Lambda^{(2)}) \right\rangle = \frac{1}{2} - \left\langle \Lambda^{(1)} \middle| \Lambda^{(2)} \right\rangle = \frac{1}{2} - \epsilon_2 \tag{A.38}$$

such that

$$\rho_{b_{n+1}} = \begin{pmatrix} \frac{1}{2} + \epsilon_2 & 0 \\ 0 & \frac{1}{2} - \epsilon_2 \end{pmatrix} \tag{A.39}$$

This yields the entanglement entropy

$$\begin{aligned} S_{b_{n+1}} &= -\sum_i \lambda_i \ln \lambda_i \\ &= -(\frac{1}{2} + \epsilon_2) \ln \left(\frac{1}{2} + \epsilon_2 \right) - (\frac{1}{2} - \epsilon_2) \ln \left(\frac{1}{2} - \epsilon_2 \right) \\ &\approx -\frac{1}{2} \ln \left(\frac{1}{2} - \epsilon_2 \right)^2 - 4\epsilon_2^2 + O(\epsilon_2^4) \\ &\approx \ln 2 - 4\epsilon_2^2 + O(\epsilon_2^4) \\ &> \ln 2 - 2\epsilon^2 + O(\epsilon^4) \\ &> \ln 2 - \epsilon, \end{aligned} \tag{A.40}$$

where we have expanded the last term in Taylor series to obtain the third line. This proofs the third and last lemma.   □

Appendix B

Quantum field theory in curved spacetime

B.1   The general Bogoliubov coefficients

In this appendix, it is shown how (1.96) follows directly from the orthonormality of the f and g-modes, (1.94).

We start with α_{ij}. Use the form of the Bogoliubov transformations (1.93) and write

$$
\begin{aligned}
(g_i, f_j) &= (\sum_k \alpha_{ik} f_k + \beta_{ik} f_k^*, f_j) \\
&= \sum_k \alpha_{ik}(f_k, f_j) + \beta_{ik}(f_k^*, f_j).
\end{aligned}
\tag{B.1}
$$

Then one needs to use the orthonormality relations (1.94) to obtain

$$
(g_i, f_j) = \sum_k \alpha_{ik}\delta_{jk} + \beta_{ik}0 = \alpha_{ij},
\tag{B.2}
$$

which is what we wanted.

For the second Bogoliubov coefficient β_{ij}, one obtains, with the same reasoning,

$$
\begin{aligned}
-(g_i, f_j^*) &= -(\sum_k \alpha_{ik} f_k + \beta_{ik} f_k^*, f_j^*) \\
&= -\sum_k (\alpha_{ik}(f_k, f_j^*) + \beta_{ik}(f_k^*, f_j^*)) \\
&= -\sum_k \beta_{ik}(-\delta_{kj}) \\
&= \beta_{ij}.
\end{aligned}
\tag{B.3}
$$

We thus established that 1.96 holds.

B.2   Bogoliubov coefficients: Minkowski-Rindler

In this appendix, the Bogoliubov coefficients $\alpha_{\Omega\omega}$ and $\beta_{\Omega\omega}$, corresponding to the transformation between Minkowski and Rindler light coordinates, are computed. To do so, the approach in [24] is presented and supplemented by extra steps to make the proof completely clear.

B.2.1 $\alpha_{\Omega\omega}$

To obtain the first Bogoliubov coefficient $\alpha_{\Omega\omega}$, we need to compute the following integral

$$
\alpha_{\Omega\omega} = \frac{i}{4\pi\sqrt{\omega\Omega}} \int_{-\infty}^{\infty} du\, e^{-i\Omega u} e^{i\omega U} \left(i\omega e^{-au} + i\Omega \right).
\tag{B.4}
$$

To do so, we would like to write the integral in a form such that we can simply use the definition of the Γ-function ([40])

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \qquad \text{Integral converges for } \text{Re}(z) > 0 \qquad \text{(B.5)}$$

In order to obtain such a form, it is convenient to change variables such that the integration interval becomes $[0, \infty]$. A choice which enables this is

$$u \to z = e^{-au}. \qquad \text{(B.6)}$$

As a result, the following changes must be included

$$du = -\frac{1}{az} dz \qquad e^{-i\Omega u} = z^{\frac{i\Omega}{a}}$$
$$e^{i\omega U} = e^{-\frac{i\omega}{a} z}. \qquad \text{(B.7)}$$

The substitution thus yields the following form of the integral

$$\alpha_{\Omega\omega}^R = -\frac{1}{4\pi a} \frac{1}{\sqrt{\omega\Omega}} \int_0^\infty dz\, z^{\frac{i\Omega}{a}} e^{-\frac{i\omega z}{a}} \left(\omega + \frac{\Omega}{z}\right). \qquad \text{(B.8)}$$

By looking at the definition of the Γ-function (B.5), one sees that

$$\int_0^\infty dt\, t^n e^{-at} = \int_0^\infty \frac{du}{a} \left(\frac{u}{a}\right)^n e^{-u}$$
$$= \frac{\Gamma(n+1)}{a}, \qquad \text{(B.9)}$$

where we substituted $u = -at$. As a result, one can write (B.8) in terms of the Γ-function. However, one must note that the integral (B.8) is not convergent while this must hold to use the definition (B.5). To solve this, one can add an imaginary part to ω, in order to obtain a non-oscillating term. Explicitly we take

$$\omega \to \omega - i\epsilon \qquad \epsilon \to 0 \qquad \text{(B.10)}$$

such that the integral converges for $\text{Re}(z) > 0$.
With the following substitutions

$$x = -\frac{\omega}{a} \qquad s = \frac{i\Omega}{a} \qquad \text{(B.11)}$$

one can write the integral (B.8) in the convenient form

$$\alpha_{\Omega\omega} = -\frac{1}{4\pi a} \frac{1}{\sqrt{\omega\Omega}} \int_0^\infty dz \left[e^{ixz} z^s \omega + e^{ixz} z^{s-1} \Omega \right]. \qquad \text{(B.12)}$$

Applying (B.9), one obtains

$$\alpha_{\Omega\omega} = -\frac{1}{4\pi a} \frac{1}{\sqrt{\omega\Omega}} \left[\omega i^{(1+\frac{i\omega}{a})} (-\frac{\omega}{a})^{-(1+i\frac{\Omega}{a})} \Gamma(1 + i\frac{\Omega}{a}) + \Omega i^{\frac{i\Omega}{a}} (-\frac{\omega}{a})^{-\frac{i\Omega}{a}} \Gamma(\frac{i\Omega}{a}) \right]. \qquad \text{(B.13)}$$

Furthermore, one can note that

$$\omega(-\frac{\omega}{a})^{-(1+\frac{i\Omega}{a})} = \omega(-\frac{\omega}{a})^{-1}(-\frac{\omega}{a})^{-\frac{i\Omega}{a}} = -a(-\frac{\omega}{a})^{-\frac{i\Omega}{a}}, \qquad \text{(B.14)}$$

such that

$$\alpha_{\Omega\omega}^R = -\frac{1}{4\pi a} \frac{1}{\sqrt{\omega\Omega}} i^{\frac{i\Omega}{a}} (-\frac{\omega}{a})^{-\frac{i\Omega}{a}} \left[-ai\Gamma(1 + \frac{i\Omega}{a}) + \Omega\Gamma(\frac{i\Omega}{a}) \right]. \qquad \text{(B.15)}$$

With the use of a property of the Γ-function [40]

$$\Gamma(z+1) = z\Gamma(z), \qquad \text{(B.16)}$$

one can write

$$-ai\Gamma(1 + \frac{i\Omega}{a}) = -ai(\frac{i\Omega}{a})\Gamma(\frac{i\Omega}{a}) = \Omega\Gamma(\frac{i\Omega}{a}). \qquad \text{(B.17)}$$

This yields

$$
\begin{aligned}
\alpha_{\Omega\omega}^{R} &= \frac{i}{2\pi a}\frac{1}{\sqrt{\omega\Omega}}i^{\frac{i\Omega}{a}}\left(-\frac{\omega}{a}\right)^{-\frac{i\Omega}{a}}\Gamma\left(1+i\frac{\Omega}{a}\right) \\
&= -\frac{1}{2\pi a}\sqrt{\frac{\Omega}{\omega}}i^{\frac{i\Omega}{a}}\left(-\frac{\omega}{a}\right)^{\frac{-i\Omega}{a}}\Gamma\left(\frac{i\Omega}{a}\right).
\end{aligned}
\tag{B.18}
$$

Finally, rewriting terms as exponentials

$$
\begin{aligned}
i^{\frac{i\Omega}{a}} &= \left(e^{\frac{i\pi}{2}}\right)^{\frac{i\Omega}{a}} = e^{-\frac{\pi\Omega}{2a}} \qquad \text{as} \qquad e^{\frac{i\pi}{2}} = i \\
\left(-\frac{\omega}{a}\right)^{-\frac{i\Omega}{a}} &= \left(\frac{\omega}{a}\right)^{-\frac{i\Omega}{a}}(-1)^{-\frac{i\Omega}{a}} = e^{-\frac{i\Omega}{a}\ln\left(\frac{\omega}{a}\right)}(-1)^{-\frac{i\Omega}{a}} = e^{-\frac{i\Omega}{a}\ln\left(\frac{\omega}{a}\right)}e^{\frac{\Omega\pi}{a}}
\end{aligned}
\tag{B.19}
$$

one obtains the final expression for the Bogoliubov coefficient

$$
\alpha_{\Omega\omega}^{R} = -\frac{1}{2\pi a}\sqrt{\frac{\Omega}{\omega}}e^{\frac{\pi\Omega}{2a}}e^{-\frac{i\Omega}{a}\ln\left(\frac{\omega}{a}\right)}\Gamma\left(\frac{i\Omega}{a}\right).
\tag{B.20}
$$

B.2.2 $\beta_{\Omega\omega}$

We can avoid computing a complex integral by noting that the only difference between $\gamma_{\Omega\omega}$ is an overall minus sign and a change $\omega \to -\omega$. As a result, one can obtain $\beta_{\Omega\omega}$ by changing

$$
e^{-\frac{i\Omega}{a}\ln\left(\frac{\omega}{a}\right)} \to e^{-\frac{i\Omega}{a}\ln\left(\frac{-\omega}{a}\right)}
\tag{B.21}
$$

in (2.14).

Complex logarithms

However, one must pay attention with the $\log(-\omega)$ appearing in this term as we chose $\omega = \omega - i\epsilon$ to be complex. In this case, the logarithm is multi-valued and we must assure that

$$
\lim \epsilon \to 0 = \omega \qquad \text{and} \qquad \lim \epsilon \to 0 = -\omega,
\tag{B.22}
$$

e.g. the limit $\epsilon \to 0$ must bring us to the same ω (apart from the minus sign). To do so, one must choose a branch of the function [40]. In our case the limit $\epsilon \to 0$ therefore results in

$$
\ln\left(-\frac{\omega}{a}\right) = \ln\left(\frac{\omega}{a}\right) - i\pi.
\tag{B.23}
$$

As a result, keeping into the minus sign with respect to $\alpha_{\Omega\omega}$, one finally obtains

$$
\beta_{\Omega\omega} = -e^{-\frac{\pi\Omega}{a}}\alpha_{\Omega\omega}.
\tag{B.24}
$$

Appendix C

The niceness conditions for the Schwarzschild geometry

In this section, we prove that the spacelike slices in all regions of the Schwarzschild black hole obey niceness condition (N1)→(N3). Before we move on to that, the notions of hypersurfaces, normal vectors and extrinsic curvature are introduced.

C.1 Extrinsic curvature

Remember that the *intrinsic* curvature R tells us how global spacetime is curved and fundamentally has to do with the effect of parallel transportation on the manifold. In contrast, the *extrinsic* curvature K quantifies the curvature of a hypersurface Σ with respect to the higher-dimensional spacetime it is embedded in. The discussion below is based on [6], [24] and [69].

C.1.1 Definition

In order to get some intuition about the new concepts, one can picture a 2-dimensional global spacetime as a piece of paper or a plane as depicted on Figure C.1. If one defines a 1-dimensional hypersurface Σ (i.e. a curve) in this spacetime, it is possible to compute its extrinsic curvature by considering normal vectors n^μ defined in the points P and P' of the curve. Namely, the extrinsic curvature measures how much the direction of such a normal vector changes when it is parallel transported to another point on the curve. Explicitly, for Figure C.1, this yields

$$\frac{d\bar{n}}{ds} = -K\bar{e}_{tangent},\tag{C.1}$$

where $\bar{e}_{tangent}$ is a basis vector $\in T_P$.

The definition of the extrinsic curvature can be extended to cases where the global spacetime has a number of dimensions $d > 2$. In that case, one should first introduce the extrinsic curvature tensor K_{ab}, given by

$$\begin{aligned} K_{ab} &= -\bar{e}_b.\nabla_a\bar{n} \qquad \text{where} \qquad \bar{e}_b \in T_P \qquad a = 0, 1, 2, .., d-1 \\ &= -e_b^\mu e_a^\nu \nabla_\nu n_\mu,\end{aligned}\tag{C.2}$$

where the indices a are restricted to the directions tangential to Σ.

Note that, the normal 4-vector n_μ and the tangent 4-vector T^μ satisfy the following conditions

$$\begin{aligned} T^\mu n_\mu &= 0 \qquad \text{orthogonality between the tangent and the normal vector in a point} \\ n^\mu n_\mu &= \pm 1 \qquad \text{unit norm condition: + for timelike and - for spacelike surfaces,}\end{aligned}\tag{C.3}$$
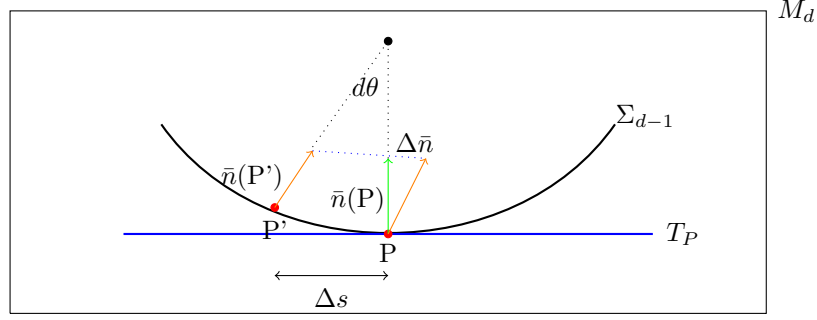
as one can expect from the sketch on Figure C.1.

Figure C.1: Schematic (and unrealistic) sketch of the definitions of normal vectors on a hypersurface. The orange vector correspond to the normal vector on the curve, defined in P'. This one has been parallel transported towards point P (as denoted by the blue dots). The green arrow represents the normal vector to the curve Σ in point P. The extrinsic curvature is a measure of the change in the normal vector when it is moved from one point on the hypersurface to another. Finally, note that T_P represents the tangent plane to the hypersurface at point P.

However, we will only be interested in the trace of this object, K, which we refer to as the extrinsic curvature. This is obtained by contracting the extrinsic curvature tensor, yielding

$$K = g^{ab}K_{ab} = g^{\mu\nu}\nabla_\nu n_\mu \qquad \mu = 0, 1, .., d. \tag{C.4}$$

This definition will be used to show that niceness condition (N2) is satisfied for the spacelike hypersurface in a Schwarzschild geometry (defined in section 2.2.2) in appendix C.2.2.

C.1.2   The induced metric

It is also possible to obtain the extrinsic curvature by defining an induced matrix h_{ab} on the tangent plane defined at a point of the hypersurface, yielding (3.29). To show this, we discuss the derivation in [47].

Firstly, one can derive a useful relationship between the scalar product of components of a basis and Christoffel symbols. To do so, consider a general basis $\bar{e}_\mu$ with $\mu = 0, 1, 2, 3$ in a 4-dimensional spacetime. Then, a general 4-vector can be expressed as linear combination of these basis vectors as

$$\bar{V} = V^\mu \bar{e}_\mu. \tag{C.5}$$

In the same way, one can also expand the covariant derivative of this vector in the same basis. This yields

$$\nabla_\mu \bar{V} = (\partial_\mu V^\lambda + \Gamma^\lambda_{\mu\nu} V^\nu)\bar{e}_\lambda = (D_\mu V^\lambda)\bar{e}_\lambda. \tag{C.6}$$

Interestingly, we can also write the basis vector themselves as a linear combination of the basis vectors. Of course, this simply yields

$$\bar{e}_\lambda = \delta^\nu_\lambda \bar{e}_\nu. \tag{C.7}$$

With the same reasoning as in (C.6), one can thus write down the covariant derivative in this basis as

$$\nabla_\mu \bar{e}_\nu = (\partial_\mu \delta^\alpha_\nu + \Gamma^\alpha_{\mu\nu})\bar{e}_\alpha = \Gamma^\alpha_{\mu\nu}\bar{e}_\alpha. \tag{C.8}$$

Furthermore, using the definition of a scalar 4-product, one has

$$\bar{e}_\alpha . \bar{e}_\beta = \delta^\mu_\alpha \delta^\nu_\beta g_{\mu\nu} = g_{\alpha\beta}. \tag{C.9}$$

With the use of these equality's, one obtains that

$$\bar{e}_\sigma . \nabla_\mu \bar{e}_\nu = \bar{e}_\sigma . (\Gamma^\alpha_{\mu\nu}\bar{e}_\alpha) = \Gamma^\alpha_{\mu\nu} g_{\sigma\alpha}. \tag{C.10}$$

This relation will be used to obtain a simplified expression for the extrinsic curvature K. Before we move on to that, we need to introduce the notion of an *induced metric* and define a convenient coordinate system for this purpose.
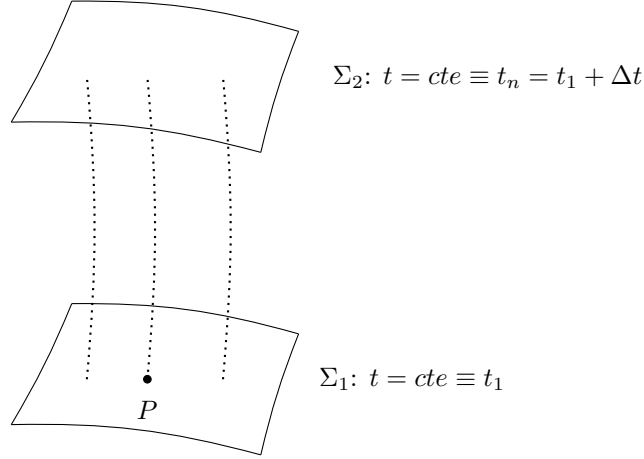
Figure C.2: 3-dimensional sketch of the construction of Gaussian normal coordinates based on Figure 27.1 in [69]. Two spatial hypersurfaces were constructed in a three dimensional spacetime. Along the dotted lines, which represent geodesics, the spatial coordinates are constant.

Gaussian normal coordinates

When working with space/time-like hypersurfaces, it is convenient to introduce *synchronous* coordinate system, also referred to as consisting of *Gaussian normal coordinates*. This is particularly interesting in cosmology [69] and we will introduce it here to discuss the notion of the induced metric.

Examine the t =constant hypersurface constructed in Figure C.2. If one considers the tangent plane T_P defined on a point P, one can construct a set of normal Gaussian coordinates. Namely, the tangent space in this point is spanned by the tangent basis vectors $\bar{e}^a$ with $a = 1, 2$. In contrast, the proper time coordinate corresponds to the unit normal to the hypersurface Σ_1: $x^0 = \bar{n}$. One can thus choose the basis vectors at the point P to have the form $\bar{e}^\mu = (\bar{n}, \bar{e}^a)$. Interestingly, one can use the orthogonality between the normal vector and the spatial coordinates together with equation (C.9) to see that the metric in Gaussian normal coordinates has the following components

$$
\begin{aligned}
g_{ta} &= \bar{e}_t.\bar{e}_a = \bar{n}.\bar{e}_a = 0 \\
g_{tt} &= \bar{e}_t.\bar{e}_t = -1 \qquad \text{due to condition C.3.}
\end{aligned}
\tag{C.11}
$$

As a result, the metric in these coordinates has the form

$$
ds^2 = -dt^2 + h_{ab}dx^a dx^b,
\tag{C.12}
$$

where h_{ab} is called the *induced metric* on the tangent plane at point P. Namely, a spacelike hypersurface $t = constant$ has a spatial (d-1) geometry defined by equation (C.12) with $dt = 0$. As a consequence, the metric defined on this surface is

$$
ds^2_{\text{hypersurface}} = h_{ab}dx^a dx^b \qquad a = 0, .., d-1.
\tag{C.13}
$$

Interestingly, one can also define the *transverse metric*, which isolates the part of the metric $g^{\mu\nu}$ that is transverse to the normal. This is defined as

$$
h_{\mu\nu} = g_{\mu\nu} - \epsilon n_\mu n_\nu \qquad \mu = 0, ..., d.
\tag{C.14}
$$

Induced metric and the extrinsic curvature

The use of normal coordinates allows us to find another expression of the extrinsic curvature K, which we defined earlier in equation (C.4). To start with, by looking at the definition of K_{ab} (C.2) and noting that $\bar{n}.\bar{e}_b = 0$, one finds that

$$
\nabla_a(-\bar{e}_b.\bar{n}) = -\bar{n}.(\nabla_a \bar{e}_b) - \bar{e}_b.(\nabla_a \bar{n}) = 0
\tag{C.15}
$$

such that

$$
K_{ab} = \bar{n}.\nabla_a \bar{e}_b.
\tag{C.16}
$$

As a result, using relation (C.10) and remembering that $\bar{n} = \bar{e}_t$, this yields

$$K_{ab} = g_{t\lambda}\Gamma^\lambda_{ab}. \tag{C.17}$$

Using the metric components in the normal coordinates (C.11), one finally obtains

$$\begin{aligned} K_{ab} &= g_{tt}\Gamma^t_{ab} \\ &= -\Gamma^t_{ab}. \end{aligned} \tag{C.18}$$

For a general hypersurface embedded in a spacetime of general dimension this is written as

$$K_{ab} = -\epsilon\Gamma^t_{ab} \qquad \text{where } a \text{ is restricted to tangential directions,} \tag{C.19}$$

where $\epsilon = +1$ for a spacelike and $\epsilon = -1$ for a timelike hypersurface.

Furthermore, one can compute the Christoffel symbols appearing in (C.18), in terms of the metric components

$$\Gamma^t_{ab} = \frac{1}{2}g^{t\lambda}\left(\partial_a g_{\lambda b} + \partial_b g_{\lambda a} - \partial_\lambda g_{ab}\right). \tag{C.20}$$

Explicitly inserting the components obtained in (C.11) yields

$$\Gamma^t_{ab} = -\frac{1}{2}g^{tt}\partial_t g_{ab} = \frac{1}{2}\partial_t h_{ab}. \tag{C.21}$$

Contracting K_{ab} with the induced metric,

$$K = -\frac{1}{2}(\partial_t h_{ab})h^{ab} \tag{C.22}$$

and noting that

$$\partial_t h = h h^{ab}\partial_t h_{ab} \tag{C.23}$$

the extrinsic curvature becomes

$$K = -\frac{1}{2}\partial_t\frac{1}{h} = -\frac{1}{\sqrt{h}}\partial_t\sqrt{h}, \tag{C.24}$$

which shows that (3.29) indeed holds.

C.2 Niceness conditions for spacelike slices in a Schwarzschild geometry

In what follows, we will show that the spacelike slices constructed in the Schwarzschild spacetime in a way described in section 2.2, obey the niceness conditions. These computations were not written out in [10], in which this construction of slices is presented, and are worked out in detail here. The Schwarzschild metric is given by

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \tag{C.25}$$

In order to check the niceness condition N1, we need to calculate the Ricci scalar R for every segment of the constructed spacelike slices in section 2.2.2.

C.2.1 Niceness condition N1

In this section we check if the $r < 2M$ and $r > 4M$ parts of the constructed slices obey $R << \frac{1}{l_p^2}$.

t=constant part

In the region $r > 4M$, a spacelike slice corresponds to a slice which is constant in time. On this segment, the metric on this slice is

$$ds'^2 = (1 - \frac{2M}{r})^{-1}dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{C.26}$$

where we have set $dt = 0$. Furthermore, this part of the slice obeys $r > 4M$.
In order to calculate the Christoffel symbol, we write down the corresponding action

$$S = \frac{1}{2}\int (1 - \frac{2M}{r})^{-1}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2.$$

This allows us to find the Euler Lagrange (EL) equations

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) - \frac{\partial L}{\partial x^\mu} = 0. \tag{C.27}$$

By solving every component of the EL-equation and by identifying it with the geodesic equation

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} = 0 \tag{C.28}$$

we can find the Christoffel symbols.

As there is no time dependence in the action we quickly see that $\Gamma^t_{\mu\nu} = 0$.
In contrast, the r component Euler Lagrange equation is

$$\frac{d}{d\tau}(\frac{\partial L}{\partial \dot{r}}) - \frac{\partial L}{\partial r} = 0$$

which gives

$$\ddot{r} + \dot{r}^2(1 - \frac{2M}{r})^{-1}\frac{2M}{r^2} - (1 - \frac{2M}{r})r\dot{\theta}^2 - 2r\sin^2\theta\dot{\phi}^2(1 - \frac{2M}{r}) = 0$$

After comparing this with the geodesic equation, one can easily read off the Christoffel symbols

$$\Gamma^r_{rr} = -\frac{M}{r(r-2M)} \qquad \Gamma^r_{\theta\theta} = -(r - 2M) \qquad \Gamma^r_{\phi\phi} = -\sin^2\theta(r - 2M) \tag{C.29}$$

The $\mu = \theta$ component of the EL equation reads

$$\ddot{\theta} + \frac{1}{r}\dot{r}\dot{\theta} + \frac{1}{r}\dot{\theta}\dot{r} - \sin\theta\cos\theta\dot{\phi}^2 = 0 \tag{C.30}$$

from which we can read off the following Christoffel symbols

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r} \qquad \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta. \tag{C.31}$$

Finally, the $\mu = \phi$ component of the EL equation reads

$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} + 2\cot\theta\dot{\phi}\dot{\theta} = 0$$

which gives us

$$\Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r} \qquad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta \tag{C.32}$$

These results are confirmed with the help of the Mathematica notebook [70]. Now we find the Ricci scalar by looking at the definition of the Riemann and Ricci tensors in terms of Christoffel symbols

$$R^\rho_{\mu\nu\sigma} = \partial_\mu\Gamma^\rho_{\nu\sigma} - \partial_\nu\Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \tag{C.33}$$

and

$$R_{\sigma\nu} \equiv R^\lambda_{\sigma\lambda\nu}. \tag{C.34}$$

such that

$$R = g^{\sigma\nu} R_{\sigma\nu} \tag{C.35}$$

As all non-diagonal and time components are zero, the Ricci scalar reduces to

$$R = g^{rr} R_{rr} + g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi}$$

We start with R_{rr} which is given by

$$R_{rr} = R^{\lambda}_{r\lambda r} = \partial_{\lambda}\Gamma^{\lambda}_{rr} - \partial_{r}\Gamma^{\lambda}_{\lambda r} + \Gamma^{\lambda}_{\lambda\gamma}\Gamma^{\gamma}_{rr} - \Gamma^{\lambda}_{r\gamma}\Gamma^{\gamma}_{\lambda r} = \frac{2M}{(2M - r)r^2} \tag{C.36}$$

Contracting this with g^{rr} gives us

$$g^{rr} R_{rr} = \frac{2Mr - 4M^2}{(2M - r)r^3} \sim \frac{1}{r^2},$$

where we used $r \sim M$ to obtain the proportionality.
Similarly the $R_{\theta\theta}$ component is given by

$$R_{\theta\theta} = 2 - \frac{M}{r} \qquad g^{\theta\theta} R_{\theta\theta} = \frac{M - 2r}{r^3} \sim \frac{1}{r^2}$$

Finally, the $R_{\phi\phi}$ component is given by

$$R_{\phi\phi} = -\frac{(M - 2r)\sin^2\theta}{r} \qquad g^{\phi\phi} R_{\phi\phi} = -\frac{M - 2r}{r^3} \sim \frac{1}{r^2}$$

We thus conclude that

$$R \sim \frac{1}{r^2} << \frac{1}{l_p^2} \tag{C.37}$$

which holds even for supermassive black holes with M between 10^5 and 10^{10} solar masses ($l_p \sim 10^{-35}$m). Finally, note that have compare these quantities because $r = 2M$ has the dimension of length.

r=constant part

We now do the same but for the $\frac{M}{2} < r = r_1 < \frac{3M}{2}$, in this region a spacelike slices corresponds to a surface which is constant in r. Consequently, the metric on this hypersurface is

$$ds''^2 = -(1 - \frac{2M}{r_1})dt^2 + r_1^2 d\theta^2 + r_1^2 \sin\theta^2 d\phi^2 \tag{C.38}$$

With the same procedure as for the first segment and by looking at the similarities with the previous case, we obtain the following Christoffel symbols

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta$$
$$\Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\theta.$$

With these new set of Christoffel symbols, one is able to determine the Ricci scalar in order to check niceness condition (N1). Note that the Ricci tensor has the following components

$$R_{\theta\theta} = 1 \qquad R_{\phi\phi} = \sin^2\theta.$$

Contracting these components with the metric (C.38), we obtain the Ricci scalar

$$R = \frac{1}{r_1^2} + \frac{\sin^2\theta}{r_1^2} \sim \frac{1}{r_1^2}$$

which indeed obeys $R << \frac{1}{l_p^2}$.

C.2.2   Niceness condition N2

In this section we check if the r=constant and $r > 4M$ parts of the constructed spacelike slices obey $K << \frac{1}{l_p^2}$.

t=constant

We start with the $r > 4M$ segment of the spacelike slice and ignore the angular directions (picture the hypersurface as constant in t and constant in the spatial directions θ and ϕ) yielding a curve as in the 2 dimensional diagram on Figure 1.6.

The tangent vector defined on a point on the $r > 4M$ segment is

$$T^\mu = (0, \frac{dr}{d\lambda} \equiv r') \tag{C.39}$$

where λ is the parameter on the curve Σ. The normal vector is then found by solving the set of equations resulting from the conditions in (C.3). Explicitly,

$$
\begin{aligned}
& T^\nu n_\nu = 0 \\
\Leftrightarrow \quad & r' n_r = 0 \qquad => n_r = 0
\end{aligned}
$$

$$
\begin{aligned}
& n^\mu n_\mu = -1 \\
\Leftrightarrow \quad & g^{\mu\nu} n_\nu n_\nu = 1 \\
\Leftrightarrow \quad & f(r)^{-1} n_t^2 + f(r) n_r^2 = -1 \\
\Leftrightarrow \quad & n_t = \sqrt{-f(r)} = \sqrt{-1 + \frac{2M}{r}}.
\end{aligned} \tag{C.40}
$$

Using the definition in (C.4) and the Christoffel symbols computed in the previous appendices (all of them are written down in (C.44), one obtains

$$
\begin{aligned}
K &= g^{\mu\nu} \nabla_\mu n_\nu \\
&= g^{\mu\nu} \left(\partial_\mu n_\nu - \Gamma^\alpha_{\mu\nu} n_\alpha \right) \\
&= f(r) \left[\frac{1}{2\sqrt{-1 + \frac{M}{r}}} (\frac{M}{r^2}) + \frac{M}{Mr - r^2} \sqrt{-1 + \frac{2M}{r}} \right]
\end{aligned} \tag{C.41}
$$

from which on sees that

$$K \sim \frac{1}{M}.$$

The $r > 4M$ segment thus obeys niceness condition (N2): $K << \frac{1}{l_p}$.

r=const

We follow the same procedure to check the (N2) niceness conditions for the $r < \frac{M}{2}$ segment of the spacelike slice. In this case, the tangent and normal vectors are

$$T^\mu = (t' \equiv \frac{dt}{d\lambda}, 0) \qquad n_\mu = (0, \sqrt{f(r)^{-1}}). \tag{C.42}$$

Consequently, the extrinsic curvature yields

$$
\begin{aligned}
K &= g^{\mu\nu} \nabla_\nu n_\mu \\
&= g^{\mu\nu} \left(\partial_r n_r - \Gamma^r_{rr} n_r \right) \\
&= g^{\mu\nu} \left(\frac{1}{\sqrt{\frac{r}{r-2M}}} \left(-\frac{M}{(r-2M)^2} \right) + \frac{M}{r^2} \sqrt{\frac{r}{r-2M}} \right).
\end{aligned} \tag{C.43}
$$

Again, by noting that $r \sim M$ one concludes that the external curvature scales as

$$K \sim \frac{1}{M}$$

which again obeys condition (N2).

C.2.3   Niceness condition N3

In this section we check if the curvature of the full spacetime, which metric is given by equation (1.6). Fortunately we already calculated a lot of Christoffel symbols and we will only have to make some changes. The resulting non-vanishing symbols are

$$\Gamma^t_{tr} = \Gamma^t_{rt} = -\frac{M}{2Mr - r^2}$$

$$\Gamma^r_{tt} = \frac{M(-2M + r)}{r^3} \qquad \Gamma^r_{rr} = \frac{M}{2Mr - r^2} \qquad \Gamma^r_{\theta\theta} = 2M - r \qquad \Gamma^r_{\phi\phi} = (2M - r)\sin^2\theta$$

$$\Gamma^\theta_{r\theta} = \Gamma\theta_{\theta r} = \frac{1}{r} \qquad \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta$$

$$\Gamma^\phi_{r\phi} = \Gamma^\phi_{\phi r} = \frac{1}{r} \qquad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.$$

(C.44)

With the use of the similarities with the previous cases, one can use the previous expressions for the Riemann tensor. One can see that this yields $R = 0$ for the full spacetime. In order to still say something about the curvature one may calculate the Kretschmann invariant

$$R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} = \frac{48M^2}{r^6}$$

(C.45)

from which we see that

$$R \sim \frac{1}{M^4},$$

which clearly satisfies $R << \frac{1}{l_p^4}$ as needed.

C.2.4   Niceness conditions of the connector segment

As we now know that the $\frac{M}{2} < r < \frac{3}{2}M$ and $r > 4M$ satisfy the niceness conditions, we can easily proof that the connector automatically satisfies the conditions as well.

Niceness condition N1

Firstly, one can proof that the fact that the segments $\frac{M}{2} < r < \frac{3}{2}M$ and $r > 4M$ satisfy niceness condition $N1$ implies the same holds for the connector segment.

Proof
The connector segment lies in the neigbourhoud of the r=const and t=const segments, which both satisfy condition N3. This condition states that the curvature of the full spacetime in the neighbourhood of the slice should be small everywhere $R << \frac{1}{l^2}$. As the connector is part of this neighbourhoud, it satisfies $R << \frac{1}{l^2}$ which is exactly (N1) for the connector segment.
We can resume this argument as

$$N3_{r<\frac{3M}{2}, r>4M} => N1_{connector}$$

Q.E.D.

Niceness condition N2

In this section we proof that the connector satisfies condition N2.

Proof

The t=const and r=const segments of the spacelike slices obey the condition N2, which states that the extrinsic curvature satisfies $K << l_p^{-2}$. As the connector segment is embedded in the same manifold and has the same extrinsic geometry, it automatically obeys N2 as well. We can resume this as

$$N2_{r<\frac{3M}{2}, r>4M} => N2_{connector}$$

Q.E.D.

Niceness condition N3

The connector segment satisfies condition N3.

Proof

As the neighbourhoud of the connector segment consists of the r=const and t=const segments which both obey $R << \frac{1}{l_p^2}$, condition N3 is automatically satisfied for the connector segment. We can resume the argument as

$$N1_{r<\frac{3M}{2},r>4M} => N3_{connector}$$

Q.E.D.

Appendix D

Equations of motion

D.1 Einstein Gravity

In this appendix the equations of motion, resulting from the variation of the Einstein-Hilbert action (3.15), is derived for d=4 dimensions. Importantly, it is going to become clear that one needs to add the Gibbons-Hawking term to make the variational principle well defined on this action. Note that this computation can be found in many places in literature, e.g. [24] but all the details are worked out here.

In 4 dimensions, the Einstein-Hilbert action is described by

$$I_{EH} = \frac{1}{16\pi} \Big[\int_M d^4x \sqrt{-g} R \Big] \tag{D.1}$$

which yields the Einstein equations in vacuum (3.16).

Varying this action with respect to the metric g, three terms need to be taken into account

$$\delta I_{EH} = \frac{1}{16\pi} \Big[\int_M d^4x (\delta \sqrt{-g}) R + \sqrt{-g} (\delta g^{\mu\nu}) R_{\mu\nu} + \sqrt{-g} g^{\mu\nu} (\delta R_{\mu\nu}) \Big]. \tag{D.2}$$

As $\delta \sqrt{-g} = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu}$, the previous equation can be rewritten as

$$\delta I_{EH} = \frac{1}{16\pi} \Big[\int_M d^4x \sqrt{-g} \Big[(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \delta g^{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu} \Big]. \tag{D.3}$$

D.1.1 Palatini identity

In order to calculate the variation of the Ricci curvature tensor, one can use the *Palatini identity*

$$\delta R_{\mu\nu} \equiv \delta R^\rho{}_{\mu\rho\nu} = \nabla_\rho \left(\delta \Gamma^\rho_{\mu\nu} \right) - \nabla_\nu \left(\delta \Gamma^\rho_{\mu\rho} \right). \tag{D.4}$$

This can be proofed by considering the explicit variation of the Riemann tensor (defined in C.33),

$$\begin{aligned} \delta R^\rho{}_{\mu\rho\nu} &= \delta \left(\partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\rho} - \Gamma^\rho_{\rho\lambda} \Gamma^\lambda_{\mu\nu} \right) \\ &= \partial_\rho \delta \Gamma^\rho_{\mu\nu} - \partial_\nu \delta \Gamma^\rho_{\mu\rho} + \delta(\Gamma^\rho_{\mu\lambda}) \Gamma^\lambda_{\nu\rho} + \Gamma^\rho_{\mu\lambda} \delta(\Gamma^\lambda_{\nu\rho}) - \delta(\Gamma^\rho_{\rho\lambda}) \Gamma^\lambda_{\mu\nu} - \Gamma^\rho_{\rho\lambda} \delta(\Gamma^\lambda_{\mu\nu}) \end{aligned} \tag{D.5}$$

and comparing it to

$$\nabla_\rho \left(\delta \Gamma^\rho_{\mu\nu} \right) - \nabla_\nu \left(\delta \Gamma^\rho_{\mu\rho} \right). \tag{D.6}$$

Namely, noting that the variations of the Christoffel symbols transform as tensors (in contrast to the symbols themselves), one finds that

$$\begin{aligned} \nabla_\rho \left(\delta \Gamma^\rho_{\mu\nu} \right) &= \partial_\rho (\delta \Gamma^\rho_{\mu\nu}) + \Gamma^\rho_{\rho\lambda} \delta \Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\mu\rho} \delta \Gamma^\rho_{\nu\lambda} - \Gamma^\lambda_{\nu\rho} \delta \Gamma^\rho_{\mu\lambda} \\ \nabla_\nu \left(\delta \Gamma^\rho_{\mu\rho} \right) &= \partial_\nu (\delta \Gamma^\rho_{\mu\rho}) + \Gamma^\rho_{\nu\lambda} \delta \Gamma^\lambda_{\mu\rho} - \Gamma^\lambda_{\mu\nu} \delta \Gamma^\rho_{\lambda\rho} - \Gamma^\lambda_{\nu\rho} \delta \Gamma^\rho_{\mu\lambda}. \end{aligned} \tag{D.7}$$

Taking the difference (D.6), one sees that the last terms cancel each other and such that one obtains (D.5).

D.1.2   The Einstein equation

By using the Palatini and the metric compatibility of the covariant derivative $\nabla_\rho g^{\mu\nu} = 0$ one finds that

$$g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\rho\left(g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu}\right) - \nabla_\nu\left(g^{\mu\nu}\delta\Gamma^\rho_{\mu\rho}\right). \tag{D.8}$$

One can write this as a total derivative by renaming the dummy indices in the second term ($\nu \leftrightarrow \rho$)

$$g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\rho(g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu} - g^{\mu\rho}\delta\Gamma^\nu_{\mu\nu}) \equiv \nabla_\rho W^\rho, \tag{D.9}$$

where W^ρ is defined as the quantity between the brackets.

With the help of the definition of the Christoffel symbols, in terms of the metric components, appearing in the covariant derivative, one can write the resulting term in the action as

$$\frac{1}{16\pi}\int_M d^4x\sqrt{-g}g^{\mu\nu}(\delta R_{\mu\nu}) = \frac{1}{16\pi}\int_M d^4x\partial_\rho(\sqrt{-g}W^\rho) \tag{D.10}$$

Consequently, one can use Stokes theorem in the following general form

$$\int_M d^4x\partial_\mu(A^\mu\sqrt{-g}) = \oint_{\partial M} A^\mu dS_\mu \tag{D.11}$$

to write (D.10) as a boundary term (setting $A^\mu = W^\rho$)

$$\frac{1}{16\pi}\int_M d^4x\partial_\rho(\sqrt{-g}W^\rho) = \frac{1}{16\pi}\oint_{\partial M} W^\rho dS_\rho. \tag{D.12}$$

As a result, one finds that the total variation is

$$\delta I_{EH} = \frac{1}{16\pi}\Big[\int_M d^4x\sqrt{-g}\Big[(R_{\mu\nu} - g_{\mu\nu}R)\delta g^{\mu\nu}\Big] + \frac{1}{16\pi}\oint_{\partial M} dS_\rho W^\rho. \tag{D.13}$$

On a compact manifold ($\partial M = \emptyset$), setting $\delta g^{\mu\nu} = 0$ at the boundary ∂M is enough to make this term vanish [24]. In that case, the corresponding equations of motion are simply the vacuum Einstein equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0. \tag{D.14}$$

However, as W^ρ contains variations of partial derivatives of the metric ($\delta\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\rho\lambda}\delta(\partial_\mu g_{\lambda\nu} + ...)$, the term containing the variation of R does not always vanish automatically. Namely, in general, asking $\delta g_{\mu\nu} = 0$ at the boundary is not enough to kill the variations $\delta\partial_\mu g_{\lambda\nu}$, which makes the variational principle ill-defined.

D.1.3   Gibbon-Hawking boundary term

The problem discussed above can be solved by adding the Gibbon-Hawking boundary term

$$I_{GH} = \frac{1}{8\pi}\int_{\partial M} d^3x\epsilon\sqrt{h}K. \tag{D.15}$$

To show that this term indeed cancels (D.12) we are going to write out the this term more explicitly with the use of the induced metric on the boundary ∂M.

The non-vanishing boundary term

Start by writing out the surface element of the non-vanishing boundary term (D.12) as

$$dS_\rho = \epsilon n_\rho dS = \epsilon n_\rho |h|^{1/2}d^3x^a \qquad a = 1, 2, 3, \tag{D.16}$$

where n_ρ is the normal vector and h is the determinant of the induced metric h_{ab} on the boundary ∂_M. For a discussion on this subject, see appendix C.1.2. This yields

$$\frac{1}{16\pi}\oint_{\partial M} dS_\rho(\sqrt{-g}W^\rho) = \frac{1}{16\pi}\oint_{\partial M} d^3x^a\epsilon n_\rho|h|^{1/2}(\sqrt{-g}W^\rho). \tag{D.17}$$

At this point, it is useful to rewrite $W^\rho = g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu} - g^{\mu\rho}\delta\Gamma^\nu_{\mu\nu}$ in a more convenient way. By noting that its individual terms can be written as

$$g^{\mu\nu}\delta\Gamma^\rho_{\mu\nu} = \frac{1}{2}g^{\mu\nu}g^{\rho\lambda}\left(\delta(\partial_\mu g_{\nu\lambda}) + \delta(\partial_\nu g_{\mu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})\right)$$

$$g^{\mu\rho}\delta\Gamma^\nu_{\mu\nu} = \frac{1}{2}g^{\mu\rho}g^{\nu\lambda}\left(\delta(\partial_\mu g_{\nu\lambda}) + \delta(\partial_\nu g_{\mu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})\right)$$

$$= \frac{1}{2}g^{\lambda\rho}g^{\mu\nu}\left(\delta(\partial_\lambda g_{\mu\nu}) + \delta(\partial_\nu g_{\mu\lambda}) - \delta(\partial_\mu g_{\nu\lambda})\right)$$

(D.18)

where the last equality was obtained by renaming the dummy indices $\lambda \leftrightarrow \mu$. Due to the cancellation of the second terms, W^ρ can be written as

$$W^\rho = g^{\mu\nu}g^{\rho\lambda}\left(\delta(\partial_\mu g_{\nu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})\right).$$

(D.19)

As a result, it is more easy to see that

$$n_\rho W^\rho|_{\partial M} = n_\rho g^{\mu\nu}g^{\rho\lambda}\left(\delta(\partial_\mu g_{\nu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})\right)$$

$$= n^\lambda g^{\mu\nu}\left(\delta(\partial_\mu g_{\nu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})\right).$$

(D.20)

Using (C.14) one is able to write

$$g^{\mu\nu} = \epsilon n^\mu n^\nu + h^{\mu\nu}$$

(D.21)

such that

$$n_\rho W^\rho|_{\partial M} = n^\lambda\left[(\epsilon n^\mu n^\nu + h^{\mu\nu})((\delta(\partial_\mu g_{\nu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})))\right].$$

(D.22)

By noting that $\epsilon n^\lambda n^\mu n^\nu$ is symmetric in indices λ and μ while $(\partial_\mu g_{\nu\lambda} - \partial_\lambda g_{\mu\nu})$ is asymmetric under their interchange, one obtains that

$$n_\rho W^\rho|_{\partial M} = n^\lambda h^{\mu\nu}(\delta(\partial_\mu g_{\nu\lambda}) - \delta(\partial_\lambda g_{\mu\nu})).$$

(D.23)

Although the condition $\delta g_{\mu\nu} = 0|_{\partial M}$ does not imply the variation of all the derivatives of $g_{\mu\nu}$ to vanish, it does imply that the tangential derivatives vanish: $\delta\partial_a g_{\mu\nu}|_{\partial M} = 0$ when a only runs over the tangential directions. Consequently

$$h^{\mu\nu}\delta(\partial_\mu g_{\nu\lambda}) = h^{ab}\delta(\partial_a g_{b\lambda}) = 0.$$

(D.24)

Finally, the remaining boundary term is thus

$$\frac{1}{16\pi}\oint_{\partial M} dS_\rho W^\rho = -\frac{1}{16\pi}\oint_{\partial M} d^3x^a \epsilon\sqrt{h}\, n^\lambda h^{\mu\nu}\delta(\partial_\lambda g_{\mu\nu})$$

(D.25)

Cancellation

Now it will be shown that (D.25) is cancelled by the Gibbons-Hawking term (D.15).

Remember that the extrinsic curvature is defined as in (C.4). Using (D.21), this yields

$$K = \nabla_\mu n^\mu = g^{\mu\nu}\nabla_\mu n_\nu = (\epsilon n^\mu n^\nu + h^{\mu\nu})\nabla_\mu n_\nu.$$

(D.26)

Noting that, with the use of the metric compatibility,

$$\nabla_\mu(n^\mu n_\mu) = \nabla_\mu(\pm 1) = 0$$

$$= n^\nu\nabla_\mu n_\nu + n_\nu\nabla_\mu n^\nu$$

$$= 2n^\nu\nabla_\mu n_\nu$$

(D.27)

the only non-vanishing part is

$$K = h^{\mu\nu}\nabla_\mu n_\nu = h^{\mu\nu}(\partial_\mu n_\nu - \Gamma^\alpha_{\mu\nu}n_\alpha).$$

(D.28)

Remembering that the induced metric is fixed due to the condition $\delta g_{\mu\nu}$ on the boundary, corresponding variation of the extrinsic curvature is

$$
\begin{aligned}
\delta K &= -h^{\mu\nu}\delta(\Gamma^{\alpha}_{\mu\nu})n_{\alpha} \\
&= -h^{\mu\nu}\frac{1}{2}g^{\alpha\lambda}\left(\delta(\partial_{\mu}g_{\nu\lambda}) + \delta(\partial_{\nu}g_{\mu\lambda}) - \delta(\partial_{\lambda}g_{\mu\nu})\right)n_{\alpha} \\
&= -h^{\mu\nu}\frac{1}{2}\left(\delta(\partial_{\mu}g_{\nu\lambda}) + \delta(\partial_{\nu}g_{\mu\lambda}) - \delta(\partial_{\lambda}g_{\mu\nu})\right)n^{\lambda}.
\end{aligned}
\tag{D.29}
$$

Again noting that the tangential derivatives of the metric vanish on the boundary, this finally yields

$$
\delta K = \frac{1}{2}h^{\mu\nu}\delta(\partial_{\lambda}g_{\mu\nu})n^{\lambda}
\tag{D.30}
$$

such that

$$
\delta I_{GH} = \frac{1}{16\pi}\oint d^3x^a\sqrt{h}h^{\mu\nu}\delta(\partial_{\lambda}g_{\mu\nu})n^{\lambda},
\tag{D.31}
$$

which indeed cancels (D.25).

This proofs that one indeed gets the vacuum Einstein equations (D.14) when varying the action $I_{EH} + I_{GH} = 0$.

D.2   JT dilaton

In this section we derive the condition (3.123). This derivation was not written down in [55], while the equation of motion was first encountered in this paper.

Look back at (3.129) before filling in $R = -2$ and vary this action with respect to the metric g

$$
\delta_g I_{JT} = \frac{1}{16\pi}\int_M d^2x(\delta\sqrt{-g})\phi(R+2) + \frac{1}{16\pi}\int_M d^2x\sqrt{-g}\phi\delta(R).
\tag{D.32}
$$

With the use of Palatini identity, we find that

$$
\delta R = \delta(g^{\mu\nu}R_{\mu\nu}) = R_{\mu\nu}\delta g^{\mu\nu} + \nabla_{\rho}\left(g^{\mu\nu}\delta\Gamma^{\rho}_{\nu\mu}\right) - \nabla_{\nu}\left(g^{\mu\nu}\delta\Gamma^{\rho}_{\rho\mu}\right).
\tag{D.33}
$$

Explicitly writing out the variation of the Christoffel symbols,

$$
\delta\Gamma^{\rho}_{\mu\nu} = \frac{1}{2}\delta(g^{\rho\gamma})\left(\partial_{\mu}g_{\gamma\nu} + \partial_{\nu}g_{\gamma\mu} - \partial_{\gamma}g_{\mu\nu}\right) + \frac{1}{2}g^{\rho\gamma}\left(\partial_{\mu}\delta g_{\gamma\nu} + \partial_{\nu}\delta g_{\gamma\mu} - \partial_{\gamma}\delta g_{\mu\nu}\right)
\tag{D.34}
$$

using $\delta g_{\mu\nu} = -g_{\mu\alpha}g_{\nu\beta}\delta g^{\alpha\beta}$ and the definition of the covariant derivative, we can write

$$
\delta\Gamma^{\rho}_{\mu\nu} = -\frac{1}{2}(g_{\nu\alpha}\nabla_{\mu}\delta g^{\alpha\rho} + g_{\mu\alpha}\nabla_{\nu}\delta g^{\alpha\rho} - g_{\mu\alpha}g_{\nu\beta}\nabla^{\rho}\delta g^{\alpha\beta})
\tag{D.35}
$$

such that

$$
\begin{aligned}
g^{\mu\nu}\delta\Gamma^{\rho}_{\mu\nu} &= -\nabla_{\alpha}\delta g^{\alpha\rho} + \frac{1}{2}g_{\alpha\beta}\nabla^{\rho}\delta g^{\alpha\beta} \\
g^{\mu\nu}\delta\Gamma^{\rho}_{\mu\rho} &= -\frac{1}{2}g_{\alpha\beta}\nabla^{\nu}\delta g^{\alpha\beta}.
\end{aligned}
\tag{D.36}
$$

As a result

$$
\delta R = (R_{\mu\nu} - \nabla_{\mu}\nabla_{\nu} + g_{\mu\nu}\nabla^2)\delta g^{\mu\nu}
\tag{D.37}
$$

such that the equation of motion reduces to

$$
\frac{\delta I_{JT}}{\delta g^{\mu\nu}} = 0 \qquad \Leftrightarrow \qquad R_{\mu\nu}\phi - \nabla_{\mu}\nabla_{\nu}\phi + g_{\mu\nu}\nabla^2\phi = 0.
\tag{D.38}
$$

When $R = -2$, the equation can be rewritten as

$$
\nabla_{\mu}\nabla_{\nu}\phi - g_{\mu\nu}\nabla^2\phi + g_{\mu\nu}\phi = 0.
\tag{D.39}
$$

This indeed corresponds to (3.123).

D.3 Conservation equations

This derivation was not written out explicitly in [55], where these conservation equations where encountered.

Start by observing that the only non-zero components of the corresponding Christoffel symbols are

$$\Gamma^u_{uu} = \partial_u \omega(u,v) \qquad \Gamma^v_{vv} = \partial_v \omega(u,v) \tag{D.40}$$

This yields a Ricci tensor

$$R_{uv} = \partial_\gamma \Gamma^\gamma_{uv} - \partial_u \Gamma^\gamma_{\gamma v} + \Gamma^\gamma_{\gamma\beta}\Gamma^\beta_{uv} - \Gamma^\gamma_{u\beta}\Gamma^\beta_{\gamma v} = -\partial_u \Gamma^v_{vv} = -\partial_u \partial_v \omega(u,v) \tag{D.41}$$

resulting in (3.134).
The conservation equations $\nabla_a T^a_b = 0$ then become

$$\nabla_a T^a_u = \nabla_u(g^{ud}T_{du}) + \nabla_v(g^{vd}T_{du}) = \nabla_u T_{vu} + \nabla_v T_{uu} = \partial_u T_{vu} - \Gamma^u_{uu}T_{vu} + \partial_v T_{uu} = 0 \tag{D.42}$$

$$\nabla_a T^a_v = \nabla_u(g^{ud}T_{dv}) + \nabla_v(g^{vd}T_{dv}) = \nabla_u T_{vv} + \nabla_v T_{uv} = \partial_u T_{vv} + \partial_v T_{uv} - \Gamma^v_{vv}T_{uv} = 0 \tag{D.43}$$

where we have used the metric compatibility of the connection.
From this, we see that

$$\partial_v T_{uu} = -\partial_u T_{vu} + T_{vu}\Gamma^u_{uu} = \frac{c}{12\pi}\partial_v\big(\partial^2_u \omega(u,v) - (\partial_u \omega(u,v))^2\big) \tag{D.44}$$

$$\partial_u T_{vv} = -\partial_v T_{vu} + T_{vu}\Gamma^v_{vv} = \frac{c}{12\pi}\partial_u\big(\partial^2_v \omega(u,v) - (\partial_v \omega(u,v))^2\big) \tag{D.45}$$

such that

$$T_{uu} = \frac{c}{12\pi}(\partial^2_u \omega - (\partial_u \omega)^2) + F_u(u) \tag{D.46}$$

$$T_{vv} = \frac{c}{12\pi}(\partial^2_v \omega - (\partial_v \omega)^2) + F_v(v) \tag{D.47}$$

where the inhomogeneous terms $F_v(v)$ and $F_u(u)$ are functions of only u or v respectively.

D.4 Dilaton equation of motion

By varying the gravitational action (4.55) with respect to the dilaton field ϕ we obtain the dilaton equation of motion. This derivation was not written out in [15], from which the resulting equation of motion was retrieved.

Remember that, in two dimensions, we can always find a coordinate w such that our metric looks like

$$ds^2 = e^{2\rho}dw d\overline{w} \tag{D.48}$$

corresponding to the green disk on Figure 4.7.
The corresponding non-vanishing Christoffel symbols are

$$\Gamma^w_{ww} = -2\partial_w \rho \qquad \text{and} \qquad \Gamma^{\overline{w}}_{\overline{w}\overline{w}} = -2\partial_{\overline{w}}\rho. \tag{D.49}$$

As a result the curvature R is given by

$$R = g^{\mu\nu}R_{\mu\nu} = 2g^{w\overline{w}}R_{w\overline{w}} = -4g^{w\overline{w}}\partial_w\partial_{\overline{w}}\rho = -8e^{-2\rho}\partial_w\partial_{\overline{w}}\rho \tag{D.50}$$

where we used $R_{w\overline{w}} = -2\partial_w\partial_{\overline{w}}\rho$. We can now use this expression to obtain the equation of motion by varying (4.55) with respect to the ϕ

$$\delta I_{grav,\phi}[M_n] = \int_{\Sigma_2} \frac{\delta\phi}{16\pi}(R+2) - (1 - \frac{1}{n})\sum_i \delta\phi(w_i) \tag{D.51}$$

such that, by using (D.50) and $\sqrt{g} = \frac{e^{2\rho}}{2}$ one obtains

$$\frac{\delta I_{grav}[M_n]}{\delta \phi} = 0$$

$$\Leftrightarrow \frac{1}{16\pi}\Big[\int \sqrt{g}(-8e^{-2\rho}\partial_w\partial_{\overline{w}}\rho)dwd\overline{w} + 2\int \sqrt{g}dwd\overline{w}\Big] - (1-\frac{1}{n})\sum_i \delta(w-w_i) = 0$$

$$\Leftrightarrow -\frac{1}{4\pi}\int \partial_w\partial_{\overline{w}}\rho\, dwd\overline{w} + \frac{1}{16\pi}\int e^{2\rho}dwd\overline{w} - (1-\frac{1}{n})\sum_i \int \delta^2(w-w_i)dwd\overline{w} = 0$$

$$\Leftrightarrow -4\partial_w\partial_{\overline{w}}\rho + e^{2\rho} = 16\pi(1-\frac{1}{n})\sum_i \delta^2(w-w_i) = 0,$$

(D.52)

which corresponds to (4.61). The first delta function arises because the variation of this term with respect to the scalar field is zero when $w \neq w_i$ as these only contains contributions from the conical singularities. The second arises when writing the term as an integral in order to set the integrand to zero.

D.5   Extrinsic curvature of the AdS$_2$ boundary

Remember that the extrinsic curvature is defined in (C.4) as

$$K = g^{\mu\nu}\nabla_\mu n_\nu.$$

(D.53)

In order to evaluate this on the AdS_2 boundary curve $(\tau(u),z(u))$ we have to find an expression for the tangent T^μ and normal vector n_μ at a point on the curve. Remember that the metric is

$$ds^2 = \frac{d\tau^2 + dz^2}{z^2}$$

(D.54)

and we fix the length of the boundary as

$$\frac{1}{\epsilon^2} = \frac{\tau^2 + z^2}{z'^2} \qquad z = \epsilon\tau' + O(\epsilon^3).$$

(D.55)

If one denotes the derivatives with respect to u as primes, tangent vector can be expressed as

$$T^\mu = (\tau'(u), z'(u)) = (\tau'(u), \epsilon\tau''(u)),$$

(D.56)

where we used (D.55) to obtain $z'(u)$ in terms of τ. With the use of conditions (C.3) and the metric (D.54) one finds that

$$T^\mu n_\mu = 0$$
$$= \tau' n_\tau + z' n_z$$
$$\Leftrightarrow n_\tau = -(\frac{z'}{\tau'})n_z$$

$$n^\mu n_\mu = 1$$

(D.57)

$$= z^2(n_\tau^2 + n_z^2)$$
$$= z^2[(\frac{z'}{\tau'})^2 + 1]n_z^2$$
$$\Leftrightarrow n_z^2 = \frac{1}{z^2[(\frac{z'}{\tau'})^2 + 1]} = \frac{\tau'^2}{z'^2[z'^2 + \tau'^2]}$$

such that

$$n_\mu = \frac{1}{z\sqrt{\tau'^2 + z'^2}}(-z', \tau').$$

(D.58)

Noting that, one can also write the extrinsic curvature as

$$K = (g^{\mu\nu} - n^{\mu\nu})\nabla_\mu n_\nu = \frac{T^\mu T^\nu}{T^2}\nabla_\mu n_\nu = \frac{T^\nu}{T^2}\nabla_T n_\nu \tag{D.59}$$

due to the second condition in (C.3) [71] and

$$\nabla_T n_\mu = \partial_u n_\mu - \Gamma^\alpha_{\mu\beta} n_\alpha T^\beta. \tag{D.60}$$

Noting that the only non-vanishing Christoffel symbols are $\Gamma^\tau_{\tau z} = \Gamma^z_{z\tau} = -\Gamma^z_{\tau\tau} = \Gamma^z_{zz} = -\frac{1}{z}$, one finds

$$T^\nu \nabla_T n_\nu = \frac{\tau'(\tau'^2 + z'^2 + z'z'') - zz'\tau''}{z^2\sqrt{\tau'^2 + z'^2}} \tag{D.61}$$

As

$$T^2 = g_{\mu\nu}T^\mu T^\nu = \frac{1}{z^2}(\tau'^2 + z'^2), \tag{D.62}$$

we find that the extrinsic curvature is

$$K = \frac{\tau'(\tau'^2 + z'^2 + z'z'') - zz'\tau''}{(\tau'^2 + z'^2)^{3/2}}. \tag{D.63}$$

Finally, one can use (D.55) to show that this can be written as

$$K = 1 + \epsilon^2\{\tau(u), u\} + O(\epsilon^4), \tag{D.64}$$

which indeed corresponds to (3.118). This is result was obtained in [15], [62] and [63] and but many steps where left out of these papers.

D.6 Extrinsic curvature on the inner circle of a quotient manifold

Based on the previous derivation and [15], we will show that (4.67) is the correct equation corresponding to the variation of the boundary term of I_{JT}. Namely, we need to show that K corresponds to

$$K = 1 + \epsilon^2\left[\{\theta, \tau\} + \left(\frac{1}{2} + U(\theta)\right)\theta'(\tau)^2\right]. \tag{D.65}$$

Firstly, we expand the exponential in the metric (4.62) and obtain

$$ds^2 = \frac{4}{(1 - |w|^2)^2}\left[1 - \frac{2}{3}(1 - |w|)^2 U(\theta) + ...\right]. \tag{D.66}$$

The cutout curve can be parametrized by τ such that $x^\mu(\tau) = (\theta(\tau), \gamma(\tau))$. To do so, write $w = e^{-\gamma}e^{-i\theta}$ and expand $e^{-\gamma}$ such that

$$ds^2 = \left(\frac{1}{\gamma^2} - \frac{2}{3}U(\theta)\right)d\theta^2 + \frac{d\gamma^2}{\gamma^2}. \tag{D.67}$$

If we fix the proper length of the boundary curve as

$$ds^2 = \frac{d\tau^2}{\epsilon^2}, \tag{D.68}$$

the parametrization satisfies

$$\gamma(\tau) = \epsilon\theta'\left[1 + \epsilon^2\left(\frac{1}{2}\frac{\theta''^2}{\theta'^2} - \frac{1}{3}U(\theta)\right)\theta'^2\right] \tag{D.69}$$

which is obtained by solving γ in a power series.

At this point, one can find an expression for the tangent T^a and normal n^a vectors to the boundary curve x^μ. In the same was as in appendix D.5, this yields

$$T^a = (\theta'(\tau), \gamma(\tau)') \qquad n_a = \frac{1}{\gamma(\tau)\sqrt{\theta'^2 + \gamma'^2}}(-\gamma'(\tau), \theta'(\tau)). \tag{D.70}$$

Computing the derivatives, the Christoffel symbols K and using (D.69) yields

$$K = \frac{T^\nu \nabla_T n_\nu}{T^2} = 1 + \epsilon^2 \left[\{\theta, \tau\} + \left(\frac{1}{2} + U(\theta) \right) \theta'(\tau)^2 \right]. \tag{D.71}$$

where we only kept $O(\epsilon^2)$ and threw away higher orders.

D.7   The curvature on the inner disk

The following computation was not worked out in paper [15], but fully discussed here.

The fist integral which is computed is (4.111):

$$I_{boun} = \frac{\phi_r}{2\pi} \int_0^{2\pi} d\tau R(\tau) = \frac{-\phi_r}{\tanh(a)} \tag{D.72}$$

where

$$R(\tau) = -\frac{(1-A^2)^2}{|1-Ae^{i\tau}|^4} = -\frac{(1-A^2)^2}{(1-Ae^{i\tau})^2(1-Ae^{-i\tau})^2}. \tag{D.73}$$

We want to use *Cauchy's integral theorem*

$$I_{boun} = \oint_C \frac{f(z)}{z-s} = 2\pi i Res(f(s)) \tag{D.74}$$

where f(z) is an analytic function. Therefore we start by writing the integral (D.72) in a convenient form

$$I_{boun} = \frac{\phi_r}{2\pi} \oint_C \frac{dz(i)(1-A^2)^2}{z(1-Az)^2(1-Az^{-1})^2} \tag{D.75}$$

$$\text{with} \qquad z = e^{i\tau} \qquad dz = ie^{i\tau}d\tau,$$

considering C as the boundary between the gravitational and flat region, ranging from $\tau = 0 \to \tau = 2\pi$ such that $A < 1$. Note that we are working on the quotient manifold such that branch cuts appearing in the circle. Noting that $z(1-Az^{-1})^2 = \frac{(z-A)^2}{z}$ and $(1-Az)^2 = A^2(z-A^{-1})^2$ we can rewrite the integral as

$$I_{boun} = \frac{\phi_r}{2\pi} \oint_C \frac{idz(1-A^2)^2}{A^2(z-A^{-1})^2(z-A)^2}. \tag{D.76}$$

As $z = A^{-1}$ is not enclosed by the circle C and hence only the $z = A$ singularity is included. The corresponding residue is

$$Res(A) = -i\frac{\phi_r}{2\pi}\frac{1+A^2}{-1+A^2}. \tag{D.77}$$

Using D.74, this yields

$$I_{boun} = \phi_r \frac{1+A^2}{-1+A^2} = \frac{\phi_r}{\tanh a}, \tag{D.78}$$

where the last expression was found by remembering that $A = e^{-a}$.

D.8   Obtaining the QES condition

The following computation was not worked out in paper [15], but fully discussed here.

In order to show that the QES condition (4.34) follows directly from the equation of motion, one needs to explicitly compute that the left hand side of

$$\int_0^{2\pi} d\tau e^{-i\tau} \left(\frac{c}{3\phi_r} \mathcal{F} - \partial_\tau R(\tau) \right) = 0. \tag{D.79}$$

When the computation is done, it should become clear that this condition is equivalent to the QES condition (4.34)

$$\partial_a S_{gen} = 0 \qquad \Leftrightarrow \qquad \sinh\left(\frac{2\pi a}{\beta}\right) = \frac{3\pi \phi_r}{\beta c} \frac{\sinh\left(\frac{\pi}{\beta}(a+b)\right)}{\sinh\left(\frac{\pi}{\beta}(a-b)\right)}. \tag{D.80}$$

Firstly note that

$$\mathcal{F} = -i\frac{e^{2i\tau}(A-B)^2}{(e^{i\tau}-A)^2(e^{i\tau}-B)^2} + i\frac{e^{-2i\tau}(A-B)^2}{(e^{-i\tau}-A)^2(e^{-i\tau}-B)^2} \tag{D.81}$$

$$\text{and} \qquad R(\tau) = -\frac{(1-A^2)^2}{(1-Ae^{i\tau})^2(1-Ae^{-i\tau})^2}.$$

The $\mathcal{F}$ part

We start by the computation of

$$\int_0^{2\pi} d\tau e^{-i\tau} \frac{c}{3\phi_r} \mathcal{F} = \int_0^{2\pi} d\tau \frac{c}{3\phi_r} \left[-i\frac{e^{i\tau}(A-B)^2}{(e^{i\tau}-A)^2(e^{i\tau}-B)^2} + i\frac{e^{-3i\tau}(A-B)^2}{(e^{-i\tau}-A)^2(e^{-i\tau}-B)^2} \right]. \tag{D.82}$$

Each term can be integrated individually with the help of the Cauchy integral theorem. First, note that we can rewrite the first term as

$$\int_0^{2\pi} d\tau \frac{c}{3\phi_r} - i\frac{e^{i\tau}(A-B)^2}{(e^{i\tau}-A)^2(e^{i\tau}-B)^2} = -\oint_C \frac{c}{3\phi_r} \frac{(A-B)^2}{(z-A)^2(z-B)^2} \tag{D.83}$$

where we have set $z = e^{i\tau}$. In this form, we can use the Cauchy theorem (D.74). Note that the Residue at A (which is the only singularity enclosed by C which corresponds to the boundary between the gravitational and flat space), yields

$$Res(i, A) = \frac{c}{3\phi_r} \frac{2}{A-B}. \tag{D.84}$$

With the same reasoning, we can set $z = e^{-i\tau}$ to rewrite the second term

$$\int_0^{2\pi} d\tau \frac{c}{3\phi_r} = i\frac{e^{-3i\tau}(A-B)^2}{(e^{-i\tau}-A)^2(e^{-i\tau}-B)^2} = -\oint_C \frac{c}{3\phi_r} \frac{z^2(A-B)^2}{(z-A)^2(z-B)^2} = ii. \tag{D.85}$$

The corresponding Residue is

$$Res(ii, A) = -\frac{c}{3\phi_r} \frac{2AB}{A-B}. \tag{D.86}$$

Using the exponential forms of $A = e^{-a}$ and $B = e^b$, the sum of the residues yields,

$$Res(i, A) + Res(ii, A) = -\frac{2c}{3\phi_r} \frac{\sinh\left(\frac{a-b}{2}\right)}{\sinh\left(\frac{a+b}{2}\right)}. \tag{D.87}$$

And hence the integral

$$\int_0^{2\pi} d\tau e^{-i\tau} \frac{c}{3\phi_r} \mathcal{F} = -4\pi i \frac{c}{3\phi_r} \frac{\sinh\left(\frac{a-b}{2}\right)}{\sinh\left(\frac{a+b}{2}\right)}. \tag{D.88}$$

The $\partial_\tau R(\tau)$ part

Secondly, we use Cauchy's integral theorem to compute

$$-\int_0^{2\pi} d\tau e^{-i\tau}\partial_\tau R(\tau) = -[R(\tau)e^{-i\tau}]_0^{2\pi} - i\int_0^{2\pi} d\tau e^{-i\tau} R(\tau) = -i\int_0^{2\pi} d\tau e^{-i\tau} R(\tau), \qquad \text{(D.89)}$$

which has been integrated by parts to obtain the last expression. Filling in $R(\tau)$ explicitly yields

$$-i\int_0^{2\pi} d\tau e^{-i\tau} R(\tau) = i\int_0^{2\pi} d\tau e^{-i\tau} \frac{(1-A^2)^2}{(1-Ae^{i\tau})^2(1-Ae^{-i\tau})^2}, \qquad \text{(D.90)}$$

which we rewrite by setting $z = e^{-i\tau}$

$$i\int_0^{2\pi} d\tau e^{-i\tau} \frac{(1-A^2)^2}{(1-Ae^{i\tau})^2(1-Ae^{-i\tau})^2} = -\oint_C \frac{(1-A^2)^2}{(1-z^{-1}A)^2(1-zA)^2} = iii. \qquad \text{(D.91)}$$

Again, A is the only singularity enclosed by C, the corresponding residue is

$$Res(iii, A) = -\frac{2}{-1+A^2} = \frac{1}{\sinh a}, \qquad \text{(D.92)}$$

where we used $A = e^{-a}$ again.

Putting everything together, we thus have that

$$\int_0^{2\pi} d\tau e^{-i\tau} \left(\frac{c}{3\phi_r}\mathcal{F} - \partial_\tau R(\tau)\right) = 0.$$

$$\Leftrightarrow -2\pi i\frac{2c}{3\phi_r}\frac{\sinh\left(\frac{a-b}{2}\right)}{\sinh\left(\frac{a+b}{2}\right)} + 2\pi i\frac{1}{\sinh a} = 0 \qquad \text{(D.93)}$$

$$\Leftrightarrow \frac{2c}{3\phi_r}\frac{\sinh\left(\frac{a-b}{2}\right)}{\sinh\left(\frac{a+b}{2}\right)} = \frac{1}{\sinh a}.$$

After inserting $\beta = 2\pi$, last condition is exactly the condition arising from $\partial_a S_{gen}(a) = 0$ (4.34), which is what we wanted to show.

Appendix E

Conformal welding problem

One of the main difficulties we encountered to solve the replicated equations of motion (4.91) is the conformal welding problem. In this section, we discuss the approach used in [15] to solve this by linearizing the welding maps $F(v)$ and $G(w)$ and using the matching condition at the boundary. We work out additional steps from the derivation in [15].

Interestingly, it is shown in [72] that general welding maps can always be found due to the Riemann mapping theorem. For a strictly mathematical discussion on the subject, the reader is referred to [72].

E.1 Linear solution to the conformal welding problem

In this appendix we show that (4.98) indeed holds. To do so, we follow approach used in [15]. en expansie fourier modes .

Expand $\theta(\tau)$ in Fourier series, which we may do because there is a 2π periodicity in τ

$$\theta(\tau) = \tau + \sum_{m=-\infty}^{\infty} c_m e^{im\tau} = \tau + \delta\theta, \tag{E.1}$$

where c_m coefficients are taken to be small [15].

Following the reasoning of we can write the holomorophic maps $G(w)$ and $F(v)$ as Laurent Polynomials

$$G(w) = w + \sum_{l=0}^{\infty} g_l w^l \qquad F(v) = v + \sum_{-\infty}^{2} f_l v^l, \tag{E.2}$$

where g_l and f_l are taken to be small.

In order to find the corresponding coefficients, we impose the boundary condition matching condition

$$G(e^{i\theta(\tau)}) = F(e^{i\tau}). \tag{E.3}$$

Comparing both sides of this equation for $l \leq -2$ and expanding $e^{i\theta}$ one finds

$$e^{i\tau} + ic_{-2}e^{-i\tau} + ic_{-3}e^{-2i\tau} + ic_{-4}e^{-3i\tau} + ... = e^{i\tau} + f_{-2}e^{-2i\tau} + f_{-3}e^{-3i\tau} + ... \tag{E.4}$$

from which we conclude that

$$f_{l+1} = ic_l \qquad \text{for} \qquad l \leq 2. \tag{E.5}$$

Similarly, for $l > 2$ one finds

$$g_{l+1} = -ic_l \qquad \text{for} \qquad l > 2. \tag{E.6}$$

Finally, for $l \in [0, 2]$, expanding $e^{i\theta}$, one finds that the boundary condition implies

$$e^{i\tau} + ic_{-1} + ic_0 e^{i\tau} + ic_1 e^{2i\tau} + g_0 + g_1 e^{i\tau} + g_2 e^{2i\tau} + ... = e^{i\tau} + f_0 + f_1 e^{i\tau} + f_2 e^{2i\tau} \tag{E.7}$$

from which we see

$$ic_{-1} = f_0 - g_0 \qquad ic_0 = f_1 - g_1 \qquad ic_1 = f_2 - g_2. \tag{E.8}$$

Note that there is an ambiguity by a small SL(2, C) action. We can remove it by setting $G(0) = 0$ and $F(v) = v+c$ for $v \to \infty$ [15]. This corresponds to setting (which can be seen from e.g. (E.7))

$$g_0 = f_2 = f_1 = 0. \tag{E.9}$$

As a result, we have an unique solution for the remaining coefficients

$$f_l = ic_{l-1} \quad \text{for} \quad l \leq 0 \quad \text{and} \quad g_l = -ic_{l-1} \quad \text{for} \quad l > 0. \tag{E.10}$$

We can now use this solution to compute the Schwarzian $\{F, v\}$. First, note that the derivatives are

$$F'(v) = 1 + \sum_{l=-\infty}^{2} f_l l v^{l-1}$$

$$F''(v) = \sum_{l=-\infty}^{0} f_l l(l-1)v^{l-2} + 2f_2 \tag{E.11}$$

$$F'''(v) = \sum_{-\infty}^{-1} f_l l(l-1)(l-2)v^{l-3} = \sum_{-\infty}^{-2} ic_l l(l-1)(l+1)v^{l-2}.$$

Using the smallness of the coefficients and, we find,

$$F'(v) \approx 1 \qquad . \tag{E.12}$$

Additionally, remember that we have set $F(v) = v + c$ to fix the SL(2,C) ambiguity and that $f_2 = 0$. Therefore, we have that the second derivative vanishes

$$F''(v) \approx 0. \tag{E.13}$$

Consequently, we find,

$$\{F, v\} \approx \sum_{-\infty}^{-2} ic_l l(l^2 - 1)v^{l-2} \tag{E.14}$$

Comparing $v^2\{F, v\}$ to $\delta\theta''' + \delta\theta'$, we see that

$$\delta\theta' = \sum_{-\infty}^{\infty} c_m i m e^{im\tau}$$

$$\delta\theta''' = -i \sum_{-\infty}^{\infty} c_m m^3 e^{im\tau} \tag{E.15}$$

$$\delta\theta''' + \delta\theta' = -i \sum_{-\infty}^{\infty} c_m m(m^2 - 1)e^{im\tau},$$

where we used (E.1). As a result, we obtain that

$$e^{2i\tau}\{F, v\} = -(\delta\theta''' + \delta\theta')_-, \tag{E.16}$$

where the minus subscript is used to indicate that we only used the negative frequency modes of the series in (E.15). Furthermore, noting that

$$\delta\{w, \tau\} = \delta\{e^{i\theta}, \tau\} \equiv \{e^{i\tau+i\theta}, \tau\} - \{e^{i\tau}, \tau\} = \delta\theta''' + \delta\theta' \tag{E.17}$$

such that

$$\delta\{w, \tau\}_- = (\delta\theta''' + \delta\theta')_- = -e^{2i\tau}\{F, v\} \tag{E.18}$$

We thus proofed (4.98).

Furthermore from (E.15) it is clear that this yields zero for $m = 0, +1, -1$ which ultimately results in (4.103).

Bibliography

[1] Susskind, L. (2009). The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics (Reprint ed.). Back Bay Books.

[2] Feynman, R. (1965b). The Feynman Lectures on Physics Vol. III Ch. 1: Quantum Behavior. The Feynman Lectures Website. `https://www.feynmanlectures.caltech.edu/III_01.html`

[3] Mandl, F., Shaw, G. (2010). Quantum Field Theory (2nd ed.). Wiley.

[4] The Standard Model. (2021, April 14). CERN. `https://home.cern/science/physics/standard-model`

[5] Blair, C. (2020, October). General Relativity. Lecture notes. `http://homepages.vub.ac.be/~cblair/GR20.html`

[6] Carroll, S. (2019). Spacetime and Geometry: An Introduction to General Relativity. Cambridge: Cambridge University Press. doi:10.1017/9781108770385

[7] Percacci, R. (2008). Asymptotic Safety. `https://arxiv.org/pdf/0709.3851.pdf`.

[8] Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., Tajdini, A. (2020). The entropy of Hawking radiation. ArXiv, 1–40. `https://arxiv.org/abs/2006.06872`

[9] Page, D. (2009). The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics. Leonard Susskind Little, Brown and Co, New York, 2008. . ISBN 978–0-316-01640-7. Physics Today, 62(5), 57–58. `https://doi.org/10.1063/1.3141946`

[10] Mathur, S. D. (2009). The information paradox: a pedagogical introduction. Classical and Quantum Gravity, 26(22), 1–30. `https://doi.org/10.1088/0264-9381/26/22/224001`.

[11] Musser, G. (2021b, February 4). The Black Hole Information Paradox Comes to an End. Quanta Magazine. `https://www.quantamagazine.org/the-black-hole-information-paradox-comes-to-an-end-20201029/`

[12] Ryu, S., Takayanagi, T. (2006). Holographic Derivation of Entanglement Entropy from the anti–de Sitter Space/Conformal Field Theory Correspondence. Physical Review Letters, 96(18). `https://arxiv.org/abs/hep-th/0603001`.

[13] Penington, G. (2020). Entanglement wedge reconstruction and the information paradox. Journal of High Energy Physics, 2020(9). `https://doi.org/10.1007/jhep09(2020)002`

[14] Hubeny, V. E., Rangamani, M., Takayanagi, T. (2007). A covariant holographic entanglement entropy proposal. Journal of High Energy Physics, 2007(07), 062. `https://doi.org/10.1088/1126-6708/2007/07/062`.

[15] Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., Tajdini, A. (2020). The entropy of Hawking radiation. ArXiv, 1–40. `https://arxiv.org/abs/2006.06872`

[16] Faulkner, T., Lewkowycz, A., Maldacena, J. (2013). Quantum corrections to holographic entanglement entropy. Journal of High Energy Physics, 2013(11). `https://arxiv.org/abs/1307.2892`

[17] Hartman, T. (2015). Lectures on Quantum Gravity and Black Holes. `http://www.hartmanhep.net/topics2015/gravity-lectures.pdf`

[18] Harlow, D. (2016, februari). Jerusalem lectures on black holes and quantum information. Reviews of Modern Physics. `https://doi.org/10.1103/revmodphys.88.015002`

[19] Nobel Foundation. (2020). Nobel Prizes 2020. NobelPrize.Org. `https://www.nobelprize.org/prizes/physics/2020/penrose/facts/`

[20] Nature Editorial. (2020, October 6). Physicists who unravelled mysteries of black holes win Nobel prize. Nature. `https://www.nature.com/articles/d41586-020-02764-w`.

[21] Hsu, R.-R. (1992). The no-hair theorem. `https://www.researchgate.net/publication/253684546_The_No_Hair_Theorem`

[22] Bekenstein, J. D. (1973). Black Holes and Entropy. Physical Review D, 7(8), 2333–2346. `https://doi.org/10.1103/physrevd.7.2333`

[23] Hawking, S. W. (1975). Particle creation by black holes. Communications In Mathematical Physics, 43(3), 199–220. `https://doi.org/10.1007/bf02345020`.

[24] Verschelde, H. (2021). Kwantum zwarte gaten en holografie. Lecture notes. UGent.

[25] Bardeen, J. M., Carter, B., Hawking, S. W. (1973). The four laws of black hole mechanics. Communications in Mathematical Physics, 31(2), 161–170. `https://doi.org/10.1007/bf01645742`

[26] Almheiri, A., Mahajan, R., Maldacena, J., Zhao, Y. (2020). The Page curve of Hawking radiation from semiclassical geometry. Journal of High Energy Physics, 2020(3). `https://doi.org/10.1007/jhep03(2020)149`

[27] Mandl, F. (1988). Statistical Physics, 2nd Edition (2nd ed.). Wiley.

[28] Traschen, J. (2000). An Introduction to Black Hole Evaporation. CDS, 29. `http://cds.cern.ch/record/468589/files/0010055.pdf`

[29] Chatwin-Davies, A. (2021). Modave Lectures on Quantum Information. ArXiv, 1–11. `https://arxiv.org/pdf/2102.02066.pdf`

[30] Verlinde, H. (2017). Introduction to Quantum theory: Entanglement and density matrix [Lecture notes]. Princeton University. `https://www.phy.princeton.edu//~verlinde/PHY305/`

[31] Hanson, C. (1996). MIT Scheme Reference - Bit Strings. `https://groups.csail.mit.edu/mac/ftpdir/scheme-7.4/doc-html/scheme_10.html`

[32] Griffiths, D. J. (2017). Introduction to Quantum Mechanics. Cambridge University Press.

[33] Natsuume, M. (2014). AdS/CFT Duality User Guide. ArXiv, 1–42. `https://arxiv.org/abs/1409.3575`.

[34] Hawking, S. W. (1975). Particle creation by black holes. Communications In Mathematical Physics, 43(3), 199–220. `https://doi.org/10.1007/bf02345020`.

[35] Bipartite states and Schmidt decomposition. Quantiki. (2019). Quantum Information Portal and Wiki. `https://www.quantiki.org/wiki/bipartite-states-and-schmidt-decomposition#:%7E:text=Schmidt%20Theorem%20(Schmidt%20Decomposition),1%7E%CF%87i%3D1`.

[36] Feynman, R. P., Hibbs, A. R., Styer, D. F. (2010). Quantum Mechanics and Path Integrals: Emended Edition (Dover Books on Physics) (Emended Editon ed.). Dover Publications. `http://www-f1.ijs.si/~ramsak/km1/FeynmanHibbs.pdf`

[37] Peskin, M. E., Schroeder, D. (2019). An Introduction To Quantum Field Theory. CRC Press.

[38] Zee, A. (2010). Quantum Field Theory in a Nutshell (2nd ed.). Princeton University Press.

[39] Figueroa-O'Farill, J. (1998). Chapter 2: Complex analysis. In Mathematical Techniques III (pp. 88–91). University of Eddington. `https://www.maths.ed.ac.uk/~jmf/Teaching/MT3/ComplexAnalysis.pdf`

[40] Orloff, J. (2020). Complex Variables with Applications. MIT OpenCourseWare. `https://math.libretexts.org/Bookshelves/Analysis/Book%3A_Complex_Variables_with_Applications_(Orloff)`

[41] Arbey, A., Auffinger, J., Silk, J. (2020). Primordial Kerr Black Holes. `https://arxiv.org/pdf/2012.14767.pdf`

[42] Carr, B. J., Khori, K., Sendouda, Y., Yokoyama, J. (2010). New cosmological constraints on primordial black holes. `https://arxiv.org/pdf/0912.5297.pdf`

[43] Rovelli, C. (2008). Quantum gravity. Scholarpedia. `http://www.scholarpedia.org/w/index.php?title=Quantum_gravity&action=cite&rev=170369`

[44] Gibbons, G. W., Hawking, S. W. (1977). Action integrals and partition functions in quantum gravity. Physical Review D, 15(10), 2752–2756. `https://doi.org/10.1103/physrevd.15.2752`

[45] Hawking, S. W. (1976). Breakdown of predictability in gravitational collapse. Physical Review D, 14(10), 2460–2473. `https://doi.org/10.1103/physrevd.14.2460`

[46] Hartman, T. (2015a). Course Website: Quantum Gravity and Black Holes. Tom Hartman's Physics Homepage. http://www.hartmanhep.net/topics2015/

[47] Verschelde, H. (2020). Kwantum Zwarte Gaten en Holografie. Lecture notes. UGent.

[48] Page, D. N. (1976). Particle emission rates from a black hole: Massless particles from an uncharged, nonrotating hole. Physical Review D, 13(2), 198–206. `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.13.198`

[49] Page, D. N. (2013b). Time dependence of Hawking radiation entropy. Journal of Cosmology and Astroparticle Physics, 2013(09), 028. `https://arxiv.org/pdf/1301.4995.pdf`

[50] Musser, G. (2021, February 4). The Black Hole Information Paradox Comes to an End. Quanta Magazine. `https://www.quantamagazine.org/the-black-hole-information-paradox-comes-to-an-end-20201029/`

[51] Ryu, S., Takayanagi, T. (2008). Holographic Derivation of Entanglement Entropy from AdS/CFT. `https://arxiv.org/pdf/hep-th/0603001.pdf`

[52] MacKenzie, R. (2000). Path Integral Methods and Applications. `https://arxiv.org/abs/quant-ph/0004090`

[53] Butler;, R. W. (2021). Saddlepoint Approximations with Applications (Cambridge Series in Statistical and Probabilistic Mathematics) by Ronald W. Butler (2007–08-27). Cambridge University Press; 1 edition (2007–08-27). `https://www.cambridge.org/core/books/saddlepoint-approximations-with-applications/2F7AB9244B8B82035C10BF0FF54EF3F8`

[54] Tong, D. University Of Cambridge. (2021). String Theory. `https://www.damtp.cam.ac.uk/user/tong/string/four.pdf`

[55] Maxfield, H. (2021, February 1). Lectures on black hole information and spacetime wormholes. Bhw Lectures 2021. `https://indico.cern.ch/event/970492/contributions/4156436/attachments/2183572/3689048/BHWHlecturesPart1.pdf`

[56] Stern, W. H. (2016). Conformal Field Theory. University of Hamburg. `https://www.math.uni-hamburg.de/home/stern/Notes/CFT/Notes_CFT.pdf`

[57] Bouboulis, P. (2010). Wirtinger's Calculus in general Hilbert Spaces. `https://arxiv.org/pdf/1005.5170.pdf`

[58] Calabrese, P., Cardy, J. (2004). Entanglement entropy and quantum field theory. Journal of Statistical Mechanics: Theory and Experiment, 2004(06), P06002.`https://doi.org/10.1088/1742-5468/2004/06/p06002`

[59] Calabrese, P., Cardy, J. (2009). Entanglement entropy and conformal field theory. `https://arxiv.org/pdf/0905.4013.pdf`

[60] Solvay Institutes. (2021, March 30). Black holes, information and wormholes [Lecture]. Solvay Colloquium, Brussels, Belgium. `https://www.youtube.com/watch?v=lqiWiudc2ZA`

[61] Freedman, D. Z., Van Proeyen, A. (2012). Supergravity (1st ed.). pg(489-451). Cambridge University Press.

[62] Maldacena, J., Stanford, D., Yang, Z. (2016). Conformal symmetry and its breaking in two-dimensional nearly anti-de Sitter space. Progress of Theoretical and Experimental Physics, 2016(12), 1–10.
`https://doi.org/10.1093/ptep/ptw124`

[63] Sarosi, G. (2019). AdS2 holography and the SYK model. Modave summer School.
`https://arxiv.org/abs/1711.08482`.

[64] Mertens, T. (2021). JT gravity - a review [Slides]. Math.Tecnico.Ulisboa. `https://math.tecnico.ulisboa.pt/seminars/download.php?fid=1199`

[65] Kaplan, J. (2016). Lectures on AdS/CFT from the Bottom Up [Slides]. Krieger School of Arts and Sciences. `https://sites.krieger.jhu.edu/jared-kaplan/files/2016/05/AdSCFTCourseNotesCurrentPublic.pdf`

[66] Almheiri, A., Mahajan, R., Maldacena, J. (2019). Islands outside the horizon. `https://arxiv.org/abs/1910.11077`

[67] Maldacena, J. (2003). Eternal black holes in anti-de Sitter. Journal of High Energy Physics, 2003(04), 021. `https://doi.org/10.1088/1126-6708/2003/04/021`

[68] Goto, K., Hartman, T., Tajdini, A. (2020). Replica Wormholes for an evaporating 2D black hole. `https://arxiv.org/pdf/2011.09043.pdf`

[69] Misner, C. W., Thorne, K. S., Wheeler, J. A., Kaiser, D. (2017). Gravitation. Princeton University Press. (p715-716). `http://xdel.ru/downloads/lgbooks/Misner%20C.W.%2C%20Thorne%20K.S.%2C%20Wheeler%20J.A.%20Gravitation%20%28Freeman%2C%201973%29%28K%29%28T%29%281304s%29_PGr_.pdf`.

[70] Hartman, T. (2019). General Relativity in Mathematica – Tom Hartman. Tom Hartman's Physics Homepage. `http://www.hartmanhep.net/greater2.php`

[71] Förste, S., Golla, I. (2017). Nearly AdS2 sugra and the super-Schwarzian. Physics Letters B, 771, 157–161. `https://arxiv.org/pdf/1703.10969.pdf`

[72] Sharon, E., Mumford, D. (2006). 2D shape analysis using conformal mapping. `https://link.springer.com/content/pdf/10.1007/s11263-006-6121-z.pdf`

[73] Hawking, S. W. (2005). Information loss in black holes. Physical Review D, 72(8). `https://doi.org/10.1103/physrevd.72.084013`

[74] Unruh, W. G., Wald, R. M. (2017). Information loss. `https://iopscience.iop.org/article/10.1088/1361-6633/aa778e/pdf`