PH-M24 Research Project - MPhys Dissertation

Black Holes and The Information Loss Paradox

Alistair Chopping, 951850

Department of Physics, Swansea University, Swansea, SA2 8PP, U.K.

E-mail: 951850@swansea.ac.uk

Abstract: We discuss the aspects of black hole physics that inexorably lead to the information loss paradox. Working with the Schwarzschild solution, a derivation of the Hawking temperature is given along with an overview of various numerical calculations performed to analyse an evaporating black hole's thermodynamics. In discussing the "näive" entropy of the radiation resulting from Hawking's calculation, and in giving an overview of the purity of quantum states, we highlight the issues leading to the information loss paradox. We describe the original intuition behind Page's conjecture regarding the "true" entropy of the radiation, and obtain the Page curve for an evaporating Schwarzschild black hole. We discuss the recent progress towards a solution of the information loss paradox, and give a qualitative overview of the island prescription. Finally, we use the island prescription to calculate the entropy dynamics of a pair of asymptotic-AdS_2 black holes in JT gravity, and find that one form of the information loss paradox is avoided.

Contents

1 Introduction **2**

1.1 Background 2

1.2 Structure of the Dissertation 5

2 Particle Production **6**

2.1 The Covariant Klein-Gordon Equation 6

2.2 The "$\boldsymbol{f}$" Fock Space 7

2.3 The "$\boldsymbol{p}$" Fock Space 8

2.4 Bogoliubov Transformations 9

3 The Unruh Effect **10**

3.1 Rindler Spacetime 10

4 Hawking Radiation & The Schwarzschild Black Hole **13**

4.1 The Radial Regge-Wheeler Coordinate 13

4.2 Ingoing Eddington-Finkelstein Coordinates 14

4.3 Kruskal-Szekeres Coordinates 14

4.4 The Temperature of Hawking Radiation 15

5 Purity of Quantum States **17**

5.1 Pure States 17

5.2 Mixed States & Density Matrices I - Ensembles 18

5.3 Mixed States & Density Matrices II - Composite Systems 21

5.4 Purity, Entanglement and Entropy 23

6 Mass, Temperature and Entropy Evolution of a Schwarzschild Black Hole 24

6.1 The Greybody Factor 25

6.2 Mass & Temperature Evolution 25

6.3 Bekenstein-Hawking Entropy 29

6.4 Radiation Entropy 30

6.5 The Page Curve of a Schwarzschild Black Hole 34

7 Information Loss & The Island Prescription 36

7.1 A Pre-Endpoint Paradox 36

7.2 Anti-de Sitter Space & Jackiw-Teitelboim Gravity 37

7.3 Generalised Entropy and the Island Prescription 41

7.4 The Island Prescription & The Page Curve 43

7.5 The Island Formula In Action 45

8 Discussion & Conclusions 54

8.1 Summary 54

8.2 Outlook & Discussion 55

A Definitions 58

B The Covariant Wave Equation 58

C Stress Tensor in $w^{\pm}$ coordinates 59

1 Introduction

1.1 Background

Einstein's theory of general relativity is one half (the other being quantum field theory) of the bedrock on which all of modern physics is built. Einstein's field equations can be derived from the simple Einstein-Hilbert action,

$$S_{\text{Hilbert}} = \frac{1}{16\pi G_N}\int_{\mathcal{M}} d^4x\sqrt{-g}\; R + \mathcal{L}_{\text{matter}} \quad \rightsquigarrow \quad R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G_N T_{\mu\nu}, \tag{1.1}$$

which thereby (in principle) encapsulates general relativity and all of its predictions in a single line. By solving the field equations, one can accurately predict how light bends around the sun, how the perihelion of Mercury's orbit precesses, and how an apple falls from a tree. Fighting in the trenches of the first world war, Karl Schwarzschild found the first exact solution of Einstein's field equations - a feat Einstein himself first thought impossible. A triumph no doubt, but in doing so Schwarzschild unwittingly revealed that Einstein's theory, despite describing a great number of physical phenomena with extraordinary accuracy, also predicts its own downfall.

The Schwarzschild metric as it was originally written features a coordinate singularity at a radial distance $r = \frac{2GM}{c^2}$ from the origin. This is commonly referred to as the "Schwarzschild radius", and is the radius that a spherically symmetric gravitating body needs to have in order for its escape velocity to equal the speed of light. The most familiar celestial bodies to us are those whose Schwarzschild radius is far smaller than the radius of the body itself - the earth's is roughly 8.85 millimetres, and the sun's is around 2.9 kilometres. A certain class of objects however are far *smaller* than their Schwarzschild radii, and these we call *black holes*. The classic tag line is that a black hole is "a region in spacetime from which not even light can escape", and to a good approximation this is true - the Schwarzschild radius in this case is the location of what is known as the *event horizon* or the "point of no return" in popular literature. Cross this seemingly innocuous point in space, and there is no hope of escape even if one could travel at light-speed. All world-lines of any particle thrown into the black hole point towards the *singularity*, a region of seemingly infinite spacetime curvature where our existing physical theories break down. A black hole is *so* inescapable in fact, that one can think of time and space as "swapping over" - just as we have no choice about whether or not the future comes, beyond the event horizon of a black hole travelling to the singularity becomes our future.

Hawking's 1975 paper "Particle creation by black holes"[1] represents a paradigm shift in our understanding of these objects. Therein, Hawking demonstrated that black holes are *not* simply regions of spacetime from which nothing could escape, but they do in fact emit radiation from their event horizons. This is in stark contrast to entirely classical Einstein gravity, where black holes only absorb, and do not emit, matter. A direct consequence is that since this so-called Hawking radiation carries energy away from the black hole, the black hole's mass must decrease over time - a phenomenon known as *black hole evaporation*.

The black hole information loss paradox is a direct consequence of Hawking's calculation, and has been a major problem in theoretical physics since Hawking himself proposed it in 1976 [2]. There are a number of ways to frame the problem, but one simple formulation is that since black holes evaporate, the information thrown into them appears to be destroyed at the evaporation time. This is in direct disagreement with quantum mechanics, which asserts that information cannot be destroyed without a fundamental principle of nature, *unitarity*, being violated. Hawking's calculation suggests that information is indeed lost and quantum mechanical unitarity *is* violated, a belief held by Hawking for a number of years. The famous Hawking-Thorne-Preskill bet [3] highlights the alternative position - many held the view that such a fundamental tenet of quantum mechanics could not possibly be wrong, and instead Hawking must have overlooked something in his original calculation. As time progressed, the general consensus became that indeed information should *not* be forever lost in black holes, and that there should be some as-yet-unknown mechanism by which information leaks out or can be recovered. In time, Hawking too came around to this viewpoint, conceding the bet in 2005 [4].

In the $\sim$45 years since its inception, swathes of theoretical physicists have worked on the information loss paradox, including Hawking himself [5], with none reaching a satisfactory solution. Many papers now considered to be classics have provided insight into the problem. Some have taken a quantum information theory approach such as the famous "Black holes as mirrors" paper by Hayden and Preskill [6]. They reach the conclusion that under certain assumptions regarding the black hole's internal dynamics, and assuming that any k-qubit quantum system thrown into the black hole is maximally entangled with a "reference system", one can decode the information stored in the (previously doomed) k-qubit system by collecting approximately $k + c$ qubits of Hawking radiation (with c small). They do not however address the issue of black hole evaporation and its role in the information loss paradox, nor do they treat the black hole as anything more than a highly efficient "thermaliser". Others have taken approaches

which let go of some fundamental postulate of physics, such as in the controversial "firewalls" papers by Almheiri *et al* and Susskind [7, 8]. Therein, the equivalence principle of General Relativity is explicitly violated by the presence of a so-called "firewall" - an infalling observer will be instantly "burned up" upon crossing the event horizon. All of these approaches however have provided *arguments*, not satisfactory *calculations*, to conclude that information is or is not lost.

Maldacena's landmark paper marking the birth of AdS/CFT [9] provided an entirely new toolkit for tackling a whole spectrum of problems across fields as diverse as quantum gravity, gauge theory and condensed matter physics. In particular, the AdS/CFT dictionary tells us that black holes in the bulk of AdS spacetimes[1] are dual (in a precise way) to certain manifestly unitary quantum field theories on the boundary. This indicates that black holes in AdS must evolve unitarily, and so information is indeed *not* lost in AdS black holes. Though this does not immediately imply that information is not lost for the kinds of black holes in our universe, it does provide a tantalising hint to at least the yes-or-no question "is information lost?". One would hope that a leap forward of the magnitude of AdS/CFT would eventually lead to a definitive solution of the problem, telling us exactly *how* information is not lost and where Hawking went wrong, but alas that calculation never materialised. Maldacena's conjecture has however provided insights that have directly inspired the apparent solution [10].

The apparent solution takes the form of the so-called *island prescription*, a new rule for calculating the entropy of Hawking radiation [11]. When we consider a bipartite quantum system $A + B$, obtaining the "state" of subsystem A involves "tracing out" subsystem B[2]. In the case at hand, we consider the black hole and the radiation as two halves of a bipartite quantum system, and as such calculating the state of the radiation involves "tracing out" the black hole. Doing this in a näive way leads to the information loss paradox. However, recent progress [11–13] suggests that there are regions of the black hole interior which actually contribute to the state of the radiation, and should *not* be traced out in this way. These regions are known as *islands*.

Many have tried, and ultimately failed, to satisfactorily address the black hole information loss paradox - in the words of polymath Piet Hein, *"Problems worthy of attack prove their worth by fighting back"* [14]. As of 2019, there is finally some promising progress towards the problem's solution, and the ultimate goal of this dissertation is to make contact with this recent progress.

[1] See chapter 7 for an overview of two-dimensional AdS spacetime.

[2] This idea is elucidated in chapter 5.

1.2 Structure of the Dissertation

The purpose of this dissertation is to formulate the theory which inexorably leads us to the black hole information loss paradox, and to give an overview of the recent developments regarding its apparent solution [11–13].

We begin in chapter 2 by showing that the notion of particle number is reference frame-dependent. This is done by considering Bogoliubov transformations between the annihilation and creation operators of two distinct solutions of the massless Klein-Gordon equation. This idea is carried forward to the Unruh effect in chapter 3, whereby an accelerated observer sees a thermal bath of particles in an inertial observer's vacuum, at a temperature proportional to their acceleration. Intuition would tell us that by the equivalence principle this effect should be intimately related to gravitational particle production. Thus, in chapter 4 we show how Hawking radiation is a direct implication of the Unruh effect. We derive the spectrum of Hawking radiation in the limit that there is no potential barrier outside the event horizon, and find that the Hawking temperature for a Schwarzschild black hole is $T = \frac{\hbar}{8\pi G_N M}$.

The necessary formalism regarding the purity of quantum states is introduced in chapter 5, wherein we show that "mixed states" can arise when we consider systems possessing an inherent classical uncertainty in the preparation of quantum states. More pertinent to that which follows, we also show that considerations of composite systems and their subsystems can also lead to mixed states. We prove that pure states cannot unitarily evolve into mixed states and vice-versa, before introducing the concept of entanglement entropy.

Chapter 6 begins with an exposition of various numerical calculations performed to elucidate the consequences of Hawking radiation. We obtain plots showing the mass, temperature and entropy evolution of a Schwarzschild black hole, as well as a plot showing how the entropy of the Hawking radiation changes as a function of time, according to Hawking's (now considered to be naïve) calculation. This leads us to the conclusion that Hawking's calculation must provide an incomplete picture, lest we give up on a fundamental tenet of quantum mechanics. Finally, the Page curve [16] for an evaporating Schwarzschild black hole is plotted, and we argue that it is reasonable to expect that a more sophisticated calculation of the radiation entropy should follow the Page curve.

The penultimate chapter of the dissertation (7) explains the presence of a paradox which arises immediately after the Page time, whereby the black hole and its Hawking

radiation appear to be in some sense "more entangled" than the black hole's degrees of freedom can accommodate. We then give a short overview of Anti-de Sitter (AdS) spacetime and Jackiw-Teitelboim (JT) gravity, such that the most recent developments regarding the island formula can be put into context. A qualitative overview of the island prescription for calculating the fine-grained entropy of Hawking radiation is given, and it is argued that the resulting entropy follows the Page curve. We then work through a calculation first performed in [12], using the island prescription to calculate the entropy dynamics of a pair of radiating asymptotic-AdS_2 black holes in JT gravity. Finally, we summarise the dissertation and detail some potential directions for future work in chapter 8. Throughout, we set the speed of light and Boltzmann's constant to 1.

2 Particle Production

The following is influenced by [17]. Throughout, we will employ the so-called "Heisenberg picture" whereby quantum states are fixed and all time dependence is encoded by the operators.

2.1 The Covariant Klein-Gordon Equation

We start with the Klein-Gordon equation for a massless scalar field ϕ in an arbitrary spacetime background characterised by a metric tensor g,

$$\nabla_\mu \nabla^\mu \phi = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} \partial^\mu \phi) = 0. \tag{2.1}$$

We can choose a complete basis of solutions $\{f_\omega, f_\omega^*\}$ to 2.1 which are orthonormal with respect to an inner product given by

$$(f, h) := -i \int_S d^3x \sqrt{-g} (f \dot{h}^* - \dot{f} h^*), \tag{2.2}$$

where we have used $g := \det(g_{\mu\nu})$ and where S is a Cauchy surface[3].

In order to perform the calculation of the Hawking temperature later on, we introduce an "Infra-red cut-off" to the system. This effectively gives the space a finite

[3]See A for a definition.

volume, and in doing so, the frequencies ω become quantised [18]. Therefore, the integral in 2.2 becomes a sum, and the orthonormality condition on $\{f_\omega, f_\omega^*\}$ becomes

$$(f_\omega, f_{\omega'}) \stackrel{!}{=} \delta_{\omega\omega'}. \tag{2.3}$$

The scalar field ϕ can be expressed in terms of the basis $\{f_\omega, f_\omega^*\}$ as

$$\phi = \sum_\omega (a_\omega f_\omega + a_\omega^\dagger f_\omega^*), \tag{2.4}$$

where the coefficients a_ω and $a_\omega^\dagger$ are to be identified with annihilation and creation operators, respectively. The annihilation operators are often associated with positive frequencies $\omega > 0$, and the creation operators with negative frequencies $\omega < 0$, though this is a free choice. In standard QFT in flat space, a scalar field takes the form

$$\phi = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} (a_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^\dagger\, e^{-i\vec{p}\cdot\vec{x}}), \tag{2.5}$$

and so we see that this is analogous to 2.4, where since we are considering massless modes in 1+1 dimensions the momentum is given by ω, and so f_ω is analogous to $\frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{i\vec{p}\cdot\vec{x}}$.

To properly associate the coefficients a and $a^\dagger$ with annihilation and creation operators, we can employ canonical quantisation. We impose equal-time commutation relations on ϕ,

$$[\phi(t, x_i), \phi(t, x_j)] = \delta_{ij}, \tag{2.6}$$

and commutation relations for the creation and annihilation operators

$$[a_\omega, a_{\omega'}^\dagger] = \delta_{\omega\omega'}, \tag{2.7}$$

$$[a_\omega, a_{\omega'}] = [a_\omega^\dagger, a_{\omega'}^\dagger] = 0. \tag{2.8}$$

where we have set $\hbar = 1$. We now construct a Fock space resulting from these creation and annihilation operators acting on a vacuum state.

2.2 The "f" Fock Space

To begin constructing a Fock space, we define a vacuum state $|0_f\rangle$ for the f basis by

$$a_\omega |0_f\rangle := 0 \quad \forall \omega > 0. \tag{2.9}$$

We can then construct any higher-energy state by acting on this vacuum state with any arbitrary number of creation operators. Namely,

$$(a_\omega^\dagger)^n \left|0_f\right\rangle, \tag{2.10}$$

is a state containing n particles all with energy ω. To that end, we define the so-called "number operator" used to count the number of particles in a particular state by

$$N_\omega^f := a_\omega^\dagger a_\omega, \tag{2.11}$$

such that we have

$$\left\langle 0_f\right| (a_\omega)^n N_\omega^f (a_\omega^\dagger)^n \left|0_f\right\rangle = n. \tag{2.12}$$

Note that we have

$$\left\langle 0_f\right| N_\omega^f \left|0_f\right\rangle = 0, \tag{2.13}$$

by definition of the number operator 2.11. This is simply the (rather trivial for now) statement that the vacuum state for the "f" basis is empty of "f" particles.

In order to see that the notion of particle number is observer-dependent, we must introduce another observer, or reference frame. This amounts to introducing a separate basis of solutions to 2.1, with its own Fock space and number operator.

2.3 The "*p*" Fock Space

We now define a second basis of solutions to 2.1, denoted $\{p_\omega, p_\omega^*\}$. We can express the field ϕ in this basis as

$$\phi = \sum_\omega (b_\omega p_\omega + b_\omega^\dagger p_\omega^*), \tag{2.14}$$

where again, b_ω and $b_\omega^\dagger$ are to be identified as annihilation and creation operators, which obey the expected commutation relations 2.7 and 2.8. A second Fock space can be constructed in exactly the same way as above, by defining a vacuum state $\left|0_p\right\rangle$ by

$$b_\omega \left|0_p\right\rangle := 0 \quad \forall \omega > 0, \tag{2.15}$$

and applying $b_\omega^\dagger$ to $\left|0_p\right\rangle$ n times to produce a state with n particles. We define a new number operator by

$$N_\omega^p := b_\omega^\dagger b_\omega, \tag{2.16}$$

such that

$$\left\langle 0_p\right| (b_\omega)^n N_\omega^p (b_\omega^\dagger)^n \left|0_p\right\rangle = n, \tag{2.17}$$

and again we have

$$\langle 0_p|\, N^p_\omega\, |0_p\rangle = 0. \tag{2.18}$$

Thus, we see that the vacuum state for the "p" basis is empty of "p" particles, as expected.

2.4 Bogoliubov Transformations

A non-trivial question to ask is "what is the f-vacuum expectation value of the p-number operator?". In other words, we are interested in objects like

$$\langle 0_f|\, N^p_\omega\, |0_f\rangle\,, \tag{2.19}$$

which will tell us whether or not the notion of a vacuum (i.e. the notion of particle number) is observer-independent. To answer questions like these, we need some mechanism by which we can compare the two bases of solutions. This is where Bogoliubov transformations come in. Since we have imposed that both bases are complete, it is guaranteed that expressing one basis in terms of the other is possible. We therefore have

$$p_\omega = \sum_{\omega'} (\alpha_{\omega\omega'} f_{\omega'} + \beta_{\omega\omega'} f^*_{\omega'}) \quad , \quad f_\omega = \sum_{\omega'} (\alpha^*_{\omega'\omega} p_{\omega'} - \beta_{\omega'\omega} p^*_{\omega'}), \tag{2.20}$$

where α and β are the so-called "Bogoliubov coefficients". We can derive a relation between the two sets of creation and annihilation operators using 2.4, 2.14 and 2.20, which gives

$$b_\omega = \sum_{\omega'} (\alpha^*_{\omega\omega'} a_{\omega'} - \beta^*_{\omega\omega'} a^\dagger_{\omega'}). \tag{2.21}$$

Thus, we see that an annihilation operator in the p frame is a combination of both annihilation *and* creation operators associated with the f frame. As a consequence of the canonical commutation relations imposed on the creation and annihilation operators, we have

$$\begin{aligned}
[b_\omega, b^\dagger_\omega] = \sum_{\omega'} & (\alpha^*_{\omega\omega'} a_{\omega'} - \beta^*_{\omega\omega'} a^\dagger_{\omega'})(\alpha_{\omega\omega'} a^\dagger_{\omega'} - \beta_{\omega\omega'} a_{\omega'}) \\
& - (\alpha_{\omega\omega'} a^\dagger_{\omega'} - \beta_{\omega\omega'} a_{\omega'})(\alpha^*_{\omega\omega'} a_{\omega'} - \beta^*_{\omega\omega'} a^\dagger_{\omega'}) \\
= \sum_{\omega'} & (a_{\omega'} a^\dagger_{\omega'} - a^\dagger_{\omega'} a_{\omega'})|\alpha_{\omega\omega'}|^2 + (a^\dagger_{\omega'} a_{\omega'} - a_{\omega'} a^\dagger_{\omega'})|\beta_{\omega\omega'}|^2 \stackrel{!}{=} 1,
\end{aligned}$$

and using 2.7 we have

$$\sum_{\omega'} (|\alpha_{\omega\omega'}|^2 - |\beta_{\omega\omega'}|^2) = 1. \tag{2.22}$$

Now, we can calculate the f-vacuum expectation value of the p-number operator. Using 2.21, we find

$$\boxed{\langle 0_f|\, N^p_\omega\, |0_f\rangle = \langle 0_f|\, b^\dagger_\omega b_\omega\, |0_f\rangle = \sum_{\omega'} |\beta_{\omega\omega'}|^2,} \tag{2.23}$$

which is in general non-zero. Thus, due to 2.21 involving creation operators, we see that the f-vacuum state $|0_f\rangle$ is in fact *not* empty of p-particles, and so the notion of particle number is truly observer-dependent.

3 The Unruh Effect

An important application of the ideas presented above is to the Unruh effect, whereby an accelerated observer sees a thermal bath of particles in an inertial observer's vacuum. In general relativity, inertial reference frames are those undergoing free-fall in a gravitational field, while the "accelerated" frame corresponds to one acted upon by some force. Thus, as we shall see, the Unruh effect directly implies the existence of Hawking radiation. The following is influenced by [18].

3.1 Rindler Spacetime

We begin by considering the Minkowski metric in coordinates (t, x), given by

$$ds^2 = -dt^2 + dx^2, \tag{3.1}$$

and define null (or "light cone") coordinates by

$$U := t - x,\ V := t + x, \tag{3.2}$$

such that the U and V axes lie at 45° on the (t, x) plane. Now, we introduce a uniformly accelerating "Rindler observer", with time and space coordinates (η, ξ) given in terms of the Minkowski coordinates by

$$t = a^{-1}e^{a\xi}\sinh(a\eta),\ x = a^{-1}e^{a\xi}\cosh(a\eta), \tag{3.3}$$

where a is the acceleration of the Rindler observer in their frame. We can see that (η, ξ) do indeed define an accelerating frame by noticing that

$$\frac{(ax)^2}{e^{2a\xi}} - \frac{(at)^2}{e^{2a\xi}} = 1, \tag{3.4}$$

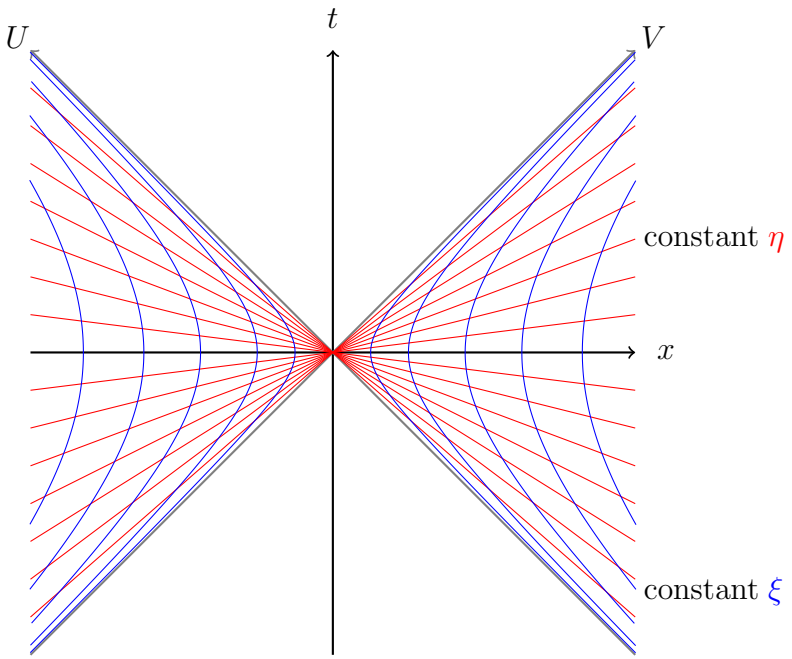

Fig. 1. A spacetime diagram showing the relationship between Minkowski coordinates (t, x), the Miknowski-null coordinates (U, V), and the Rindler coordinates (η, ξ). Lines of constant η are shown in red, with lines of constant ξ in blue.

namely 3.3 defines an infinite set of hyperbolae in the (t, x) plane for any given acceleration a. This is elucidated further in figure 1, where we parametrically plot 3.3 for varying values of a. We see that a line of constant space coordinate ξ is indeed a hyperbola asymptotic to the light cone described by (U, V). In other words, an observer following a line of constant ξ is uniformly accelerating closer and closer to the speed of light, without ever reaching it.

The coordinates (η, ξ) are often called Rindler coordinates, and they cover the right *Rindler wedge*; the portion of the spacetime diagram characterised by $U < 0$, $V > 0$. We will only work with the *right* Rindler wedge here, though this is a free choice and it is certainly possible to consider the left wedge (the region $U > 0$, $V < 0$). It is important to note that our accelerating observer has no access to the regions $U > 0$, $V > 0$ or $U < 0$, $V < 0$. Analogously to 3.2, we now define light-cone coordinates in the Rindler frame by

$$u = \eta - \xi, \; v = \eta + \xi. \tag{3.5}$$

We can relate these to the null coordinates in the inertial frame using 3.3 by

$$\begin{aligned} U &= t - x \\ &= a^{-1}e^{a\xi}(\sinh(a\eta) - \cosh(a\eta)) \\ &= a^{-1}e^{a\xi}(\frac{1}{2}(e^{a\eta} - e^{-a\eta}) - \frac{1}{2}(e^{a\eta} + e^{-a\eta})) \\ &= a^{-1}e^{a\xi}e^{-a\eta} \\ &= a^{-1}e^{-a(\eta-\xi)}, \end{aligned} \tag{3.6}$$

and similar for V. For the right Rindler wedge ($U < 0$, $V > 0$), we therefore have (using 3.5)

$$U = -a^{-1}e^{-au} \quad , \quad V = a^{-1}e^{av}. \tag{3.7}$$

For the remainder, we will focus on *right-moving* solutions to the Klein-Gordon equation 2.1 in this region of the spacetime diagram (again, this is a choice - left-moving solutions in the left Rindler wedge behave in exactly the same way). These are described by the U in 3.7 - namely the solutions we are looking for are of the form

$$\phi(u) \sim e^{-i\omega u(U)}, \tag{3.8}$$

where (rearranging 3.7)

$$u(U) = -\frac{1}{a}\ln(-aU), \tag{3.9}$$

with ω a single sharply-defined frequency. Each coordinate system (and therefore each reference frame) gives rise to a different basis of solutions, and the fact that the (U, V) and the (u, v) coordinate systems are related in this way means that we therefore must have a relation between these bases of the form 2.20. Therefore, a natural question to ask is one of the form 2.23, namely "do the non-inertial and the inertial observer agree on the vacuum state"?

Through a derivation that we will delay until the next chapter, it turns out that the accelerating observer (u, v) will see a thermal bath of particles in the inertial observer (U, V)'s vacuum, with a temperature given by

$$T = \frac{\hbar a}{2\pi}, \tag{3.10}$$

where a is the non-inertial observer's acceleration. It is important to note that this is an entirely general result when we have inertial and non-inertial frames related by an exponential map, as in 3.7. As mentioned, we would expect by the equivalence principle that there is also a particle production effect associated with the gravitational case with a temperature of the form 3.10. We now elucidate this, and give an expicit derivation of the Hawking temperature for a Schwarzschild black hole.

4 Hawking Radiation & The Schwarzschild Black Hole

In spherical polar coordinates (t, r, θ, ϕ) where $c = 1$, the Schwarzschild metric is given by

$$ds^2 = -\left(1 - \frac{2G_N M}{r}\right) dt^2 + \left(1 - \frac{2G_N M}{r}\right)^{-1} dr^2 + d\Omega_{S^2}^2, \tag{4.1}$$

where

$$d\Omega_{S^2}^2 = r^2(d\theta^2 + \sin^2(\theta) d\phi^2), \tag{4.2}$$

is the metric on the round 2-sphere. We will take the θ and ϕ coordinates to be constant throughout this discussion, allowing us to neglect the angular part of 4.1.

Note that 4.1 is cursed with a singularity at the so-called "Schwarzschild radius" $r = 2G_N M$ - the event horizon of the black hole. Since we are entirely concerned with physics around the event horizon, this is a disaster unless the singularity can be removed. Fortunately, $r = 2G_N M$ is not a curvature singularity and can indeed be removed by a sensible choice of coordinates. We now endeavour to remove this coordinate singularity. A more detailed exposition of the following coordinate systems can be found in [19].

4.1 The Radial Regge-Wheeler Coordinate

A "null geodesic" is the path in spacetime that a massless particle takes, and its interval is identically zero. Thus, for a null geodesic we find from equation 4.1 (again, where we ignore the angular part of the metric) that

$$dt^2 = \left(1 - \frac{2G_N M}{r}\right)^{-2} dr^2. \tag{4.3}$$

If we define the so-called "radial Regge-Wheeler" or "tortoise" coordinate r_* by

$$dr_*^2 := \left(1 - \frac{2G_N M}{r}\right)^{-2} dr^2, \tag{4.4}$$

equation 4.3 becomes $dt^2 = dr_*^2$. Solving 4.4, we find

$$r_* = r + 2G_N M \ln\left(\frac{r - 2G_N M}{2G_N M}\right). \tag{4.5}$$

We define two new coordinates:

$$u := t - r_* \quad , \quad v := t + r_*, \tag{4.6}$$

such that lines of constant v are ingoing null geodesics, and lines of constant u are outgoing null geodesics.

4.2 Ingoing Eddington-Finkelstein Coordinates

Here, we convert 4.1 to so-called "Ingoing Eddington-Finkelstein" (IEF) coordinates. This simply amounts to converting the t coordinate (which we will call the "Schwarzschild time" from now on) to the v coordinate given in 4.6. We start by calculating dt^2. Using $t = v - r - 2GM\ln(\frac{r-2G_NM}{2G_NM})$ from 4.6, we find

$$\begin{aligned} dt &= \frac{\partial t}{\partial v}dv + \frac{\partial t}{\partial r}dr \\ &= dv + \frac{\partial}{\partial r}\left(v - r - 2G_NM\ln\left(\frac{r-2G_NM}{2G_NM}\right)\right) \\ &= dv - \left(1 + \frac{2G_NM}{r-2G_NM}\right)dr, \\ \implies dt^2 &= dv^2 + \left(1 + \frac{2G_NM}{r-2G_NM}\right)^2 dr^2 - 2\left(1 + \frac{2G_NM}{r-2G_NM}\right)dvdr. \end{aligned} \tag{4.7}$$

Substituting this into equation 4.1, we obtain the Schwarzschild line element in ingoing Eddington-Finkelstein coordinates,

$$ds^2 = -\left(1 - \frac{2G_NM}{r}\right)dv^2 + 2dvdr. \tag{4.8}$$

Although in changing to IEF coordinates we have removed the divergence at $r = 2G_NM$, it is easy to see that the first term in 4.8 vanishes at this point, rendering the metric degenerate. A set of coordinates that extend smoothly through the event horizon are the so-called "Kruskal-Szekeres" coordinates [18].

4.3 Kruskal-Szekeres Coordinates

We first make a change of coordinates to the (u, v) system. Note that since $t = u + r_*$, we have

$$dt^2 = du^2 + dr_*^2 + 2dudr_*. \tag{4.9}$$

We also have

$$\begin{aligned} dr &= \left(\frac{\partial r_*}{\partial r}\right)^{-1} dr_* \\ &= \left(1 + \frac{2G_NM}{r-2G_NM}\right)^{-1} dr_*, \\ \iff dr &= \left(1 - \frac{2G_NM}{r}\right)dr_*. \end{aligned}$$

Plugging this into 4.1, we obtain

$$ds^2 = -\left(1 - \frac{2G_N M}{r}\right) dudv. \tag{4.10}$$

Note that the metric is still degenerate for $r = 2G_N M$, so 4.10 is only defined for $r > 2G_N M$. To solve this problem, we define another pair of coordinates in the region $r > 2G_N M$,

$$U := -\exp\left(-\frac{u}{4G_N M}\right), \; V := \exp\left(\frac{v}{4GM}\right). \tag{4.11}$$

The definitions of these coordinates imply that

$$du := 4G_N M \exp\left(-\frac{u}{4G_N M}\right) dU, \; dv := 4G_N M \exp\left(\frac{v}{4G_N M}\right) dV. \tag{4.12}$$

Plugging these into equation 4.10, we obtain the Schwarzschild metric in Kruskal-Szekeres coordinates,

$$ds^2 = -\frac{32(G_N M)^3}{r} \exp\left(-\frac{r}{2G_N M}\right) dUdV. \tag{4.13}$$

Note that to remove the coordinate singularity at $r = 2GM$ and obtain coordinates such that we can consider physics around the event horizon, we have had to introduce two coordinate systems (frames) which are related by an exponential map, manifest in equation 4.11. Therefore, we expect that there will be some kind of particle production effect. This effect is known as Hawking radiation, and has associated with it a temperature.

4.4 The Temperature of Hawking Radiation

Since the metric 4.13 is of the form $ds^2 \sim e^{-1} dUdV$, the Kruskal-Szekeres coordinates (U, V) in 4.11 define an inertial observer, freely falling through the event horizon and into the black hole [18]. We can therefore interpret the coordinates (u, v) as a non-inertial observer, perhaps in a rocket hovering at the event horizon. Then, a natural question to ask is one of the form 2.19, namely, do the two observers agree on the vacuum state of the system?

To answer this question, we consider right-moving solutions of 2.1 in the accelerating coordinates, ie $\phi \sim e^{-i\omega u(U)}$ where U is of the form 4.11. Using 2.20, we obtain

$$\frac{1}{\sqrt{\omega}} e^{-i\omega u(U)} = \sum_{\omega'} \frac{1}{\sqrt{\omega'}} (\alpha_{\omega\omega'} e^{-i\omega' U} + \beta_{\omega\omega'} e^{i\omega' U}), \tag{4.14}$$

where we have set $p_\omega = \frac{1}{\sqrt{\omega}} e^{-i\omega u(U)}$ and $f_{\omega'} = \frac{1}{\sqrt{\omega'}} e^{-i\omega' U}$. We can extract the Bogoliubov coefficients by employing inverse Fourier transforms, and thus we obtain

$$\alpha_{\omega\omega'} = \mathcal{C} \int_{-\infty}^{0} e^{i\omega' U} e^{-i\omega u(U)} dU, \;\; \beta_{\omega\omega'} = \mathcal{C} \int_{-\infty}^{0} e^{-i\omega' U} e^{-i\omega u(U)} dU, \tag{4.15}$$

where $\mathcal{C}$ is some normalisation factor. An important point to make is one about the choice of limits on the integrals in 4.15. Since $U < 0$ and $V > 0$ as indicated by 4.11, the non-inertial observer must be travelling in the *right* Rindler wedge. As mentioned before and as is evident from figure 1, an accelerated observer in the right Rindler wedge experiences a horizon at $U = 0$, and cannot access the region $U > 0$, $V > 0$. Therefore, we must impose that the solution to 2.1 we consider "has no support" in this region - namely, it must vanish for $U > 0$.

We first deal with the expression for $\alpha_{\omega\omega'}$. First, we make a change of variables and set

$$U = -e^{-i\frac{\pi}{2}} x, \implies dU = i dx. \tag{4.16}$$

Rearranging 4.11, we find $u = -4G_N M \ln(e^{-i\frac{\pi}{2}} x)$. Thus, the Bogoliubov alpha coefficient in 4.15 becomes

$$\begin{aligned} \alpha_{\omega\omega'} &= i\mathcal{C} \int_{-\infty}^{0} e^{-i\omega' e^{-i\frac{\pi}{2}} x} e^{i\omega 4G_N M \ln(e^{-i\frac{\pi}{2}} x)} dx \\ &= i\mathcal{C} e^{4G_N M\omega\pi} \int_{-\infty}^{0} e^{-\omega' x} e^{i\omega 4G_N M \ln(ix)} dx, \end{aligned} \tag{4.17}$$

where we have used

$$i\omega 4G_N M \ln(e^{-i\frac{\pi}{2}} x) = i\omega 4G_N M \ln(e^{i\frac{\pi}{2}} x) + 4G_N M\omega\pi,$$

and $e^{-i\frac{\pi}{2}} = -i$.

Now, consider the expression for $\beta_{\omega\omega'}$. Making the change of variables $U = -e^{i\frac{\pi}{2}} x$, we obtain

$$\beta_{\omega\omega'} = -i\mathcal{C} \int_{-\infty}^{0} e^{-\omega' x} e^{i\omega 4G_N M \ln(ix)} dx. \tag{4.18}$$

Comparing 4.17 and 4.18, we find that

$$\alpha_{\omega\omega'} = -e^{4G_N M\omega\pi} \beta_{\omega\omega'} \implies |\alpha_{\omega\omega'}|^2 = e^{8G_N M\omega\pi} |\beta_{\omega\omega'}|^2. \tag{4.19}$$

Using 2.22, we therefore have

$$\begin{aligned} &\sum_{\omega'} (e^{8G_N M\omega\pi} - 1) |\beta_{\omega\omega'}|^2 = 1, \\ &\implies \sum_{\omega'} |\beta_{\omega\omega'}|^2 = \frac{1}{e^{8G_N M\omega\pi} - 1}, \end{aligned} \tag{4.20}$$

and using 2.23, we find that the number of particles that the non-inertial observer (u, v) sees in the inertial observer (U, V)'s vacuum is given by

$$N = \frac{1}{e^{8G_N M\omega\pi} - 1}. \tag{4.21}$$

Comparing this to the Planck blackbody spectrum

$$N_{\text{Planck}} = \frac{1}{e^{\frac{\hbar\omega}{T}} - 1}, \tag{4.22}$$

we see that the non-inertial observer hovering at the event horizon sees a thermal bath of particles in the freely falling inertial observer's vacuum, with a temperature given by

$$\boxed{T = \frac{\hbar}{8\pi G_N M}.} \tag{4.23}$$

It turns out that 4.21 actually experiences a correction due to the presence of a potential barrier outside the event horizon - this is discussed in chapter 6.

5 Purity of Quantum States

In order to consider the consequences of Hawking radiation later on, we now take a detour away from black holes and into the nature of quantum states.

5.1 Pure States

The customary characterisation of the state of a quantum system is as a vector in some Hilbert space, namely $|\psi\rangle \in \mathcal{H}$ (or as an equivalence class of vectors $[|\psi\rangle] \subset \mathcal{H}$ called a "ray", but we will not dwell on this point). We have one system, with one state, described entirely by $|\psi\rangle$. Operators on $\mathcal{H}$ can be expressed in terms of their matrix elements $\mathcal{O}_{ab}$ as

$$\mathcal{O} = \sum_{a,b} \mathcal{O}_{ab} |a\rangle \langle b| , \tag{5.1}$$

where $|a\rangle$ and $|b\rangle$ are basis vectors. Expectation values of operators encode probabilities that a particular measurement outcome will occur, and are given by

$$\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle . \tag{5.2}$$

The state $|\psi\rangle$ evolves unitarily[4] via the Schrödinger equation, and since said equation is linear, it admits superposition solutions. For example,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{5.3}$$

is a state of a spin-$\frac{1}{2}$ particle described by a superposition of the states we call "spin up" ($|0\rangle$) and "spin down" ($|1\rangle$). It is important to note that superposition is an entirely quantum mechanical concept - there is some probability associated with the measurement of $|0\rangle$ or $|1\rangle$ (a probability of $\frac{1}{2}$ for each outcome in this case), but *before* the measurement, the system is in some sense in *both* states at the same time.

Measurement outcomes are extracted from the theory via the so-called Born rule. That is, if we have the previous state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and we want to find the probability that the particle is spin-down (ie, that the state is $|1\rangle$) upon measurement, we calculate this via

$$P(|1\rangle) = |\langle 1|\psi\rangle|^2. \tag{5.4}$$

In all of the above, the states $|\psi\rangle$, $|0\rangle$ and $|1\rangle$ are all called *pure* states.

5.2 Mixed States & Density Matrices I - Ensembles

The above formalism works perfectly fine when we are describing systems with well-defined states, for example a system consisting of a single spin-$\frac{1}{2}$ particle in a known state $|\psi\rangle$ (which may or may not be a superposition state). However, we are sometimes faced with the task of describing a system with some degree of uncertainty -perhaps our spin- $\frac{1}{2}$ particle is prepared in such a way that there are *classical* uncertainties associated with the state of the system[5]. To elucidate this point further, consider that there are two possible states of our system, $|\psi_1\rangle$ and $|\psi_2\rangle$, each with their own associated classical probabilities. These are shown in table 1.

Possible State	Classical Probability
$\|\psi_1\rangle = \frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)$	25%
$\|\psi_2\rangle = \frac{1}{\sqrt{3}}\|0\rangle + \sqrt{\frac{2}{3}}\|1\rangle$	75%

Table 1. The possible states of our toy system, with their classical probabilities shown.

[4]Namely, probabilities are conserved under time evolution.

[5]Perhaps we have a piece of laboratory equipment which prepares particles in state $|\psi_1\rangle$ or $|\psi_2\rangle$, according to the probabilities in table 1.

In table 1, we see that there is a 25% chance that the system is in the (pure) state $|\psi_1\rangle$ and a 75% chance that the system is in the (pure) state $|\psi_2\rangle$. Note that these are *classical* probabilities, entirely separate from the "quantum" probabilities associated with the fact that $|\psi_1\rangle$ and $|\psi_2\rangle$ are superposition states.

In the above situation we have what is known as an *ensemble*, whereby we have N copies of the system each in a (potentially) different state $|\psi_n\rangle$, where $n \in \{1, 2, ..., N\}$. How, then, do we characterise *the* state of the system? And how do we extract predictions from the theory? We do this by introducing a new formalism - the so-called *density matrix*.

A density matrix is an operator $\rho \in \text{End}(\mathcal{H})$ defined via the ensemble states $|\psi_n\rangle$ by [20]

$$\rho := \sum_{n=1}^{N} P_n \, |\psi_n\rangle \, \langle\psi_n| \,, \tag{5.5}$$

where P_n are known as the *weights* of the ensemble - the probability that $|\psi_n\rangle$ is the "true" state of the system[6]. As can be seen from 5.5, a general density matrix is in some sense built from a "mixture" of the pure states $|\psi_n\rangle$, so a ρ of this form is said to describe a "mixed" state[7].

How do we calculate expectation values of operators using the density matrix? We first write down the expression for the trace of an operator on $\mathcal{H}$,

$$\begin{aligned} \text{Tr}(\mathcal{O}) &:= \sum_a \mathcal{O}_{aa} \\ &= \sum_a \langle a| \, \mathcal{O} \, |a\rangle \,, \end{aligned} \tag{5.6}$$

where $\{|a\rangle\}$ is a set of basis vectors for $\mathcal{H}$ which are complete and orthonormal, namely

$$\sum_a |a\rangle \, \langle a| = 1 \quad , \quad \langle a|a\rangle = 1. \tag{5.7}$$

Expectation values in statistical mechanics are given by *ensemble averages*. In the case of our quantum ensemble, these are given by [20]

$$\langle \mathcal{O} \rangle = \sum_n P_n \, \langle\psi_n| \, \mathcal{O} \, |\psi_n\rangle \,, \tag{5.8}$$

[6] These are the "25%" and "75%" from table 1.

[7] ρ itself is sometimes colloquially referred to as "the state" of the system, though this is an imprecision.

analogous to 5.2. We can manipulate this expression in such a way that we can write it in terms of the density matrix, namely

$$\begin{aligned}
\langle \mathcal{O} \rangle &= \sum_n P_n \langle \psi_n | \mathcal{O} | \psi_n \rangle \\
&= \sum_n P_n \langle \psi_n | \mathcal{O} \cdot \mathbb{I} | \psi_n \rangle \\
&= \sum_{n,a} P_n \langle \psi_n | \mathcal{O} | a \rangle \langle a | \psi_n \rangle \\
&= \sum_{n,a} \langle a | \psi_n \rangle P_n \langle \psi_n | \mathcal{O} | a \rangle \\
&= \sum_a \langle a | \cdot \left(\sum_n P_n | \psi_n \rangle \langle \psi_n | \right) \cdot \mathcal{O} | a \rangle \\
&= \sum_a \langle a | \rho \mathcal{O} | a \rangle \\
&\equiv \mathrm{Tr}(\rho \mathcal{O}),
\end{aligned} \tag{5.9}$$

where $\{|a\rangle\}$ is an orthonormal basis for $\mathcal{H}$.

If there is only one state in the ensemble, the density matrix is of the form

$$\rho = |\psi\rangle \langle \psi| , \tag{5.10}$$

and therefore describes a pure state. Using this with 5.9, we find that

$$\begin{aligned}
\langle \mathcal{O} \rangle &= \mathrm{Tr}(\rho \mathcal{O}) \\
&= \sum_a \langle a | \rho \mathcal{O} | a \rangle \\
&= \sum_a \langle a | \psi \rangle \langle \psi | \mathcal{O} | a \rangle \\
&= \sum_a \langle \psi | \mathcal{O} | a \rangle \langle a | \psi \rangle \\
&= \langle \psi | \mathcal{O} \left(\sum_a | a \rangle \langle a | \right) | \psi \rangle \\
&= \langle \psi | \mathcal{O} | \psi \rangle ,
\end{aligned} \tag{5.11}$$

as expected for pure states. Thus, we see that the density matrix formalism is a powerful tool which can handle both mixed *and* pure states.

5.3 Mixed States & Density Matrices II - Composite Systems

It turns out that ensembles of the above type are not the only scenarios which give rise to mixed states. Mixed states also arise in the consideration of composite systems in which one is only interested in states of a subsystem. Let $\mathcal{H}$ (the composite system) be a Hilbert space which decomposes into a tensor product of two subspaces $\mathcal{H}_A$ and $\mathcal{H}_B$ (the subsystems), with $\dim(\mathcal{H}_A) \leq \dim(\mathcal{H}_B)$. Then, by the Schmidt decomposition theorem[8] we can always write a pure state on $\mathcal{H}$ as

$$\mathcal{H} \ni |\psi\rangle = \sum_{a=1}^{n} C_a \, |a\rangle_A \otimes |a\rangle_B \,, \tag{5.12}$$

where $\{|a\rangle_A\}$ and $\{|a\rangle_B\}$ are orthonormal states in $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, and $n = \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B))$. C_a here are complex coefficients satisfying

$$\sum_n |C_a|^2 = 1, \; 0 \leq |C_a|^2 \leq 1. \tag{5.13}$$

Note that these $\{|a\rangle_A\}$ also form a basis of $\mathcal{H}_A$[9], but it is not necessarily true that $\{|a\rangle_B\}$ form a basis of $\mathcal{H}_B$ (since $\dim(\mathcal{H}_A) \leq \dim(\mathcal{H}_B)$). From now on, we will denote $|a\rangle_A \otimes |b\rangle_B \equiv |a\rangle_A \, |a\rangle_B$.

Now, suppose we are only interested in the physics of subsystem A. We can use the density matrix formalism to calculate expectation values of operators $\mathcal{O}_A : \mathcal{H}_A \longrightarrow \mathcal{H}_A$ via

$$\begin{aligned} \langle \mathcal{O}_A \rangle &= \mathrm{Tr}_{\mathcal{H}_A \otimes \mathcal{H}_B}(\rho \mathcal{O}_A) \\ &= \mathrm{Tr}_{\mathcal{H}_A \otimes \mathcal{H}_B}(|\psi\rangle \langle\psi| \, \mathcal{O}_A) \\ &= \mathrm{Tr}_{\mathcal{H}_A}(\mathrm{Tr}_{\mathcal{H}_B}(|\psi\rangle \langle\psi|)\mathcal{O}_A). \end{aligned} \tag{5.14}$$

We now have an expression of the form 5.9, where we identify

$$\rho_{\mathcal{H}_A} := \mathrm{Tr}_{\mathcal{H}_B}(|\psi\rangle \langle\psi|), \tag{5.15}$$

to be the so-called *reduced density matrix* on $\mathcal{H}_A$ [21]. Using the expression for the trace of an operator, we find

$$\rho_{\mathcal{H}_A} = \sum_n \langle n|_B \, |\psi\rangle \langle\psi|n\rangle_B \,, \tag{5.16}$$

[8] This is a very well-known result, but see, for example, [22].

[9] Since the number of $|a\rangle_A$ states equals the dimension of $\mathcal{H}_A$.

where $\{|n\rangle_B\}$ is an orthonormal basis of $\mathcal{H}_B$. Plugging in 5.12 we obtain

$$\begin{aligned}
\rho_{\mathcal{H}_A} &= \sum_n \langle n|_B |\psi\rangle \langle\psi|n\rangle_B \\
&= \sum_n \langle n|_B \left(\sum_a C_a |a\rangle_A |a\rangle_B \sum_a C_a^* \langle a|_A \langle a|_B \right) |n\rangle_B \\
&= \sum_{a,n} C_a^* C_a \langle n|_B |a\rangle_B |a\rangle_A \langle a|_A \langle a|_B |n\rangle_B \\
&= \sum_{a,n} C_a^* C_a \, \delta_{na} (|a\rangle_A \langle a|_A) \delta_{an} \\
&= \sum_a |C_a|^2 |a\rangle_A \langle a|_A \,. \qquad (5.17)
\end{aligned}$$

Thus we see that $\rho_{\mathcal{H}_A}$ describes a mixed state of the form 5.5, with weights $P_n \leftrightarrow |C_a|^2$. The $|C_a|^2$ are the eigenvalues of ρ_A, with corresponding eigenstates $|a\rangle_A$. Similarly, we find

$$\rho_{\mathcal{H}_B} = \sum_a |C_a|^2 |a\rangle_B \langle a|_B \,. \qquad (5.18)$$

As intended, we have therefore shown that mixed states arise from the consideration of a composite system $(\mathcal{H}_A \otimes \mathcal{H}_B)$ when we are only concerned with the physics of a subsystem $(\mathcal{H}_\mathcal{A})$.

There is one more important question to answer in this section, namely, can pure states evolve unitarily into mixed states? We first define the *purity* of a density matrix,

$$P(\rho) := \mathrm{Tr}(\rho^2). \qquad (5.19)$$

It can be proved that [23]

$$\mathrm{Tr}(\rho^2) = \begin{cases} 1 & \text{if } \rho \text{ describes a pure state,} \\ < 1 & \text{if } \rho \text{ describes a mixed state,} \end{cases} \qquad (5.20)$$

and therefore the purity does indeed encode information about whether or not we have a pure or a mixed state.

For emphasis, we phrase the answer to the question posed above in terms of a theorem and proof. The proof is unsophisticated, and can be found in (for example) [24].

Theorem. *There exists no unitary transformation which maps a pure state to a mixed state, and vice-versa.*

Proof. Consider a unitary transformation U acting on a state as $|\psi\rangle \mapsto U|\psi\rangle$, $\langle\psi| \mapsto \langle\psi| U^\dagger$. Thus, a density matrix transforms as $\rho \mapsto U\rho U^\dagger$. We then calculate how the purity changes under U, namely

$$\begin{aligned}\mathrm{Tr}(\rho^2) &\mapsto \mathrm{Tr}((U\rho U^\dagger)^2)\\ &= \mathrm{Tr}(U\rho U^\dagger U\rho U^\dagger)\\ &= \mathrm{Tr}(U\rho\mathbb{I}\rho U^\dagger)\\ &= \mathrm{Tr}(\rho^2 U^\dagger U)\\ &= \mathrm{Tr}(\rho^2),\end{aligned}$$

where we have used the cyclicity of the trace and $U^\dagger U = \mathbb{I}$. Thus, a unitary transformation has no effect on the purity of a density matrix. □

The proof shows that pure states cannot evolve unitarily into mixed states (and vice-versa). For a state to change purity would therefore violate a central tenet upon which quantum mechanics is built, namely that states evolve unitarily via the Schrödinger equation. This is a very important point to remember.

5.4 Purity, Entanglement and Entropy

Entropy is a measure of the lack of knowledge one has about a particular system. One way of describing this relationship between entropy and information is

$$S \sim -I, \tag{5.21}$$

where we have denoted entropy by S and information by I - as our "amount of information" decreases, entropy increases. There are various measures of entropy, but the one we will find most useful is the so-called *von Neumann entropy*. Formulated in such a way as to immediately apply to mixed and pure quantum states, this is defined [21]

$$S := -\sum_n P_n \ln(P_n), \tag{5.22}$$

where the P_n are the classical probabilities from our mixed state definition 5.5. When the ensemble states $|\psi_n\rangle$ are orthonormal, the density matrix is diagonal, and the above

coincides with an equivalent definition of the von Neumann entropy [21],

$$S := -\mathrm{Tr}(\rho \ln \rho). \tag{5.23}$$

Note that when the ensemble has only a single element, i.e the system is in a pure state, the von Neumann entropy 5.22 vanishes since $P_n = 1$. This is expected - when the system is in a pure state there is no doubt which state the system is in; there is only one to choose from in the "ensemble" and hence there is no "lack of knowledge". On the other hand, when the ensemble contains more than one state there is indeed a lack of knowledge about the state of the system, and so the von Neumann entropy does not vanish.

However, what happens when we start considering composite systems and their subsystems as in the previous section? In this case, there is no ensemble to speak of, and so the von Neumann entropy cannot be a measure of the classical lack of knowledge about the system. In this case, we interpret the von Neumann entropy as a measure of the *entanglement* between the subsystems[10] [25]. For our system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, a pure state on the composite system $\mathcal{H}$ described by a density matrix ρ will have a vanishing von Neumann entropy, since the "total state" is pure. However, the reduced density matrices ρ_A or ρ_B will in general not describe pure states (as shown above), and therefore will have a non-vanishing von Neumann entropy. In this context, we interpret a non-vanishing von Neumann entropy as indicating entanglement between subsystems A and B.

One other important result is that $S(A) = S(B)$ when the overall state on $\mathcal{H}_A \otimes \mathcal{H}_B$ is pure. This is a consequence of the Schmidt decomposition theorem - using 5.17 and 5.18 with 5.22, we quickly see that

$$S(A) = -\sum_a |C_a|^2 \ln(|C_a|^2) = S(B). \tag{5.24}$$

6 Mass, Temperature and Entropy Evolution of a Schwarzschild Black Hole

In this chapter, we return to our discussion of Hawking radiation and detail the results of various numerical calculations performed in order to obtain plots showing the mass,

[10] It is valid to consider any function of a state satisfying a particular set of properties (such as the function being zero for separable states) as a measure of entanglement - see for example [26].

temperature and entropy evolution of a Schwarzschild black hole. Throughout the numerical calculations themselves, we use units such that $\hbar = c = k_B = G_N = 1$, and we also set the initial Schwarzschild radius r_s to 1.

6.1 The Greybody Factor

It turns out that although 4.21 gives the correct Hawking temperature, the spectrum given therein is in fact not quite correct, since there is a potential barrier outside the event horizon (at approximately $r \sim 3G_N M$). The Hawking modes must tunnel through said potential barrier in order to escape to infinity, and so there is a non-zero probability that these quanta will be reflected back into the black hole & thereby will not become Hawking radiation. Therefore, 4.21 must be modified by a so-called "greybody factor" to account for this. The aforementioned potential barrier is given by [18]

$$V(r) = \left(1 - \frac{r_s}{r}\right)\left(\frac{r_s}{r^3} + \frac{l(l+1)}{r^2}\right), \tag{6.1}$$

where r_s is the location of the event horizon, l is the angular momentum of the Hawking modes, and r is understood to be a function of the tortoise coordinate r_* defined by 4.5.

The presence of this potential barrier induces a correction to 4.21, which becomes

$$N_{l,\omega} = \frac{\Gamma_l(\omega)}{e^{\frac{\omega}{T}} - 1}, \tag{6.2}$$

where $\Gamma_l(\omega) \equiv |\mathcal{T}(\omega)|^2$, the probability that the Hawking modes will tunnel through the above potential barrier and escape to infinity (becoming Hawking radiation). We can therefore calculate $\Gamma_l(\omega)$ numerically for various values of l by solving the Schrödinger equation with 6.1 as the potential in the problem. Plotting these, we obtain 3.

6.2 Mass & Temperature Evolution

Since Hawking radiation carries energy away from the black hole, we would expect that an isolated black hole's mass will decrease as a function of time. In our $\hbar = 1$ units, the energy of a single Hawking mode is simply given by its frequency ω. Thus, the net energy flow out of the black hole due to all Hawking modes at a fixed frequency and angular momentum is given by

$$\frac{dE}{dt} = \omega N_{l,\omega} = \frac{\omega \Gamma_l(\omega)}{e^{\frac{\omega}{T}} - 1}. \tag{6.3}$$

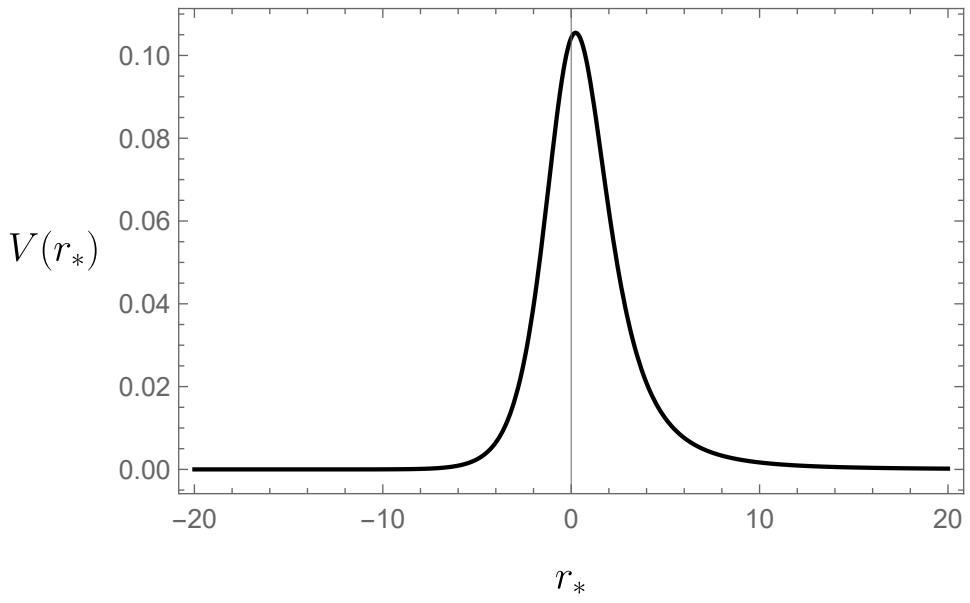

Fig. 2. The potential barrier present outside the event horizon of a Schwarzschild black hole, through which the Hawking modes must tunnel in order to escape to infinity. Here, the angular momentum of the Hawking modes was chosen to be $l = 0$.

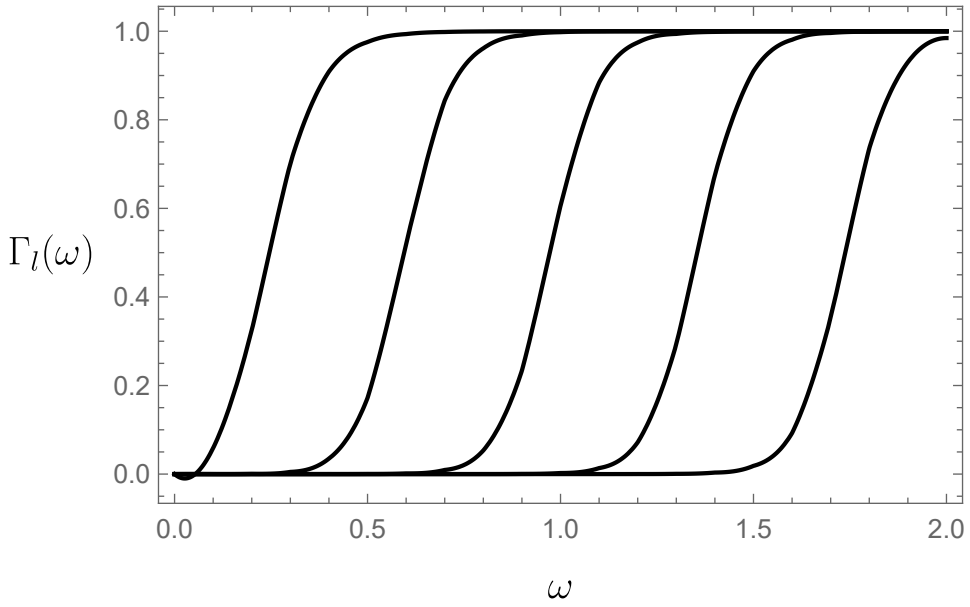

Fig. 3. A plot showing the greybody factor as a function of Hawking mode frequency, for angular momentum increasing left-to-right for $l = 0, 1, 2, 3$ and 4.

However, we are interested in the *net* flow of energy out of the black hole due to Hawking modes of *all* frequencies and angular momenta, and so to obtain this we must sum over the latter and integrate over the former. Using that the net energy flow out of the black hole must be equal to the net decrease in its mass ($\frac{dE_{\text{out}}}{dt} \stackrel{!}{=} -\frac{dM}{dt}$), we obtain

$$\frac{dM}{dt} = -\frac{dE_{\text{out}}}{dt} = -\sum_{l=0}^{\infty}(2l+1)\int_0^{\infty}\frac{d\omega}{2\pi}\cdot\frac{\omega\Gamma_l(\omega)}{e^{\frac{\omega}{T}}-1}, \tag{6.4}$$

where the $(2l+1)$ factor accounts for the degeneracy of the modes. We can scale the temperature out of the problem with the change of variables

$$x := \frac{\omega}{T} \implies d\omega = T dx, \tag{6.5}$$

and using 4.23 in natural units we obtain

$$\begin{aligned}\frac{dM}{dt} &= -T^2\sum_{l=0}^{\infty}\gamma(l) \\ &= -\sum_{l=0}^{\infty}\frac{\gamma(l)}{(8\pi M)^2},\end{aligned} \tag{6.6}$$

where we have defined

$$\gamma(l) := (2l+1)\int_0^{\infty}\frac{dx}{2\pi}\cdot\frac{x\Gamma_l(\frac{x}{4\pi})}{e^x-1}. \tag{6.7}$$

It turns out that $\gamma(l)$ is exceedingly small for $l > 1$, and so in practice one considers only the first two values of angular momentum (l = 0,1). Figure 4 shows this effect. Solving this differential equation, we can obtain a plot for the mass as a function of time, as shown in 5.

Since the mass of the black hole is intimately connected to its temperature (as given by 4.23), and (as we have found above) the mass is a function of time, in plotting 5 we can quickly find the temperature of the black hole as a function of time. Simply using

$$T = \frac{1}{8\pi M}, \tag{6.8}$$

we can plot the temperature evolution of the black hole, as shown in 6.

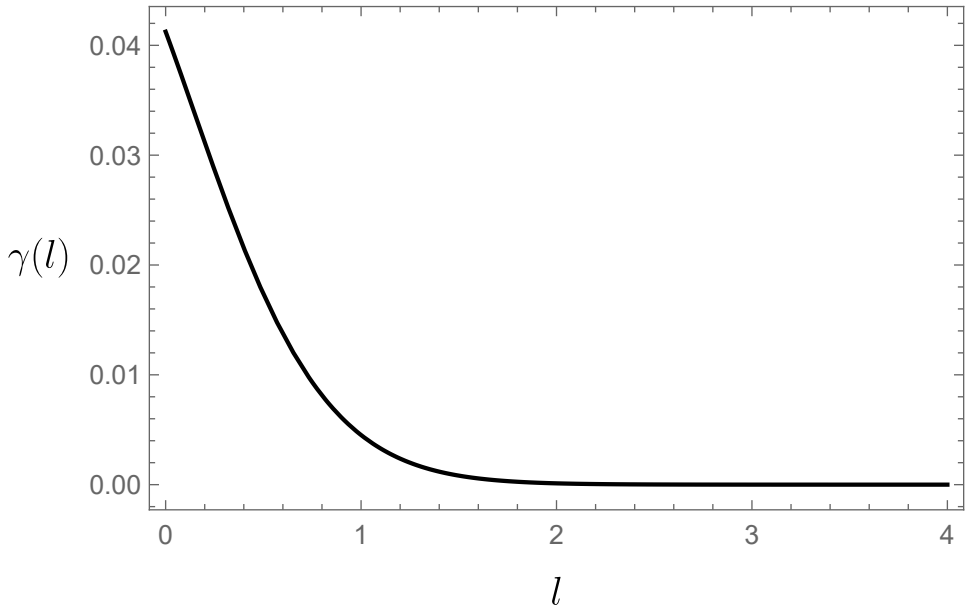

Fig. 4. A plot showing how $\gamma(l)$ falls off as a function of angular momentum. Since $\gamma(l)$ rapidly tends towards zero for $l > 0$, we choose to cut off the summations in our numerical calculations at $l = 1$.

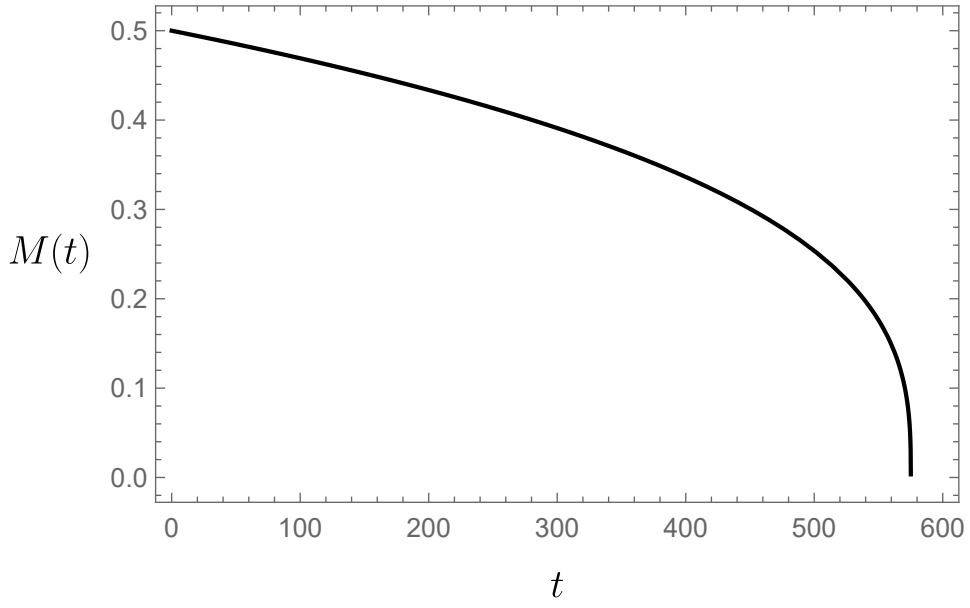

Fig. 5. A plot showing the mass of a Schwarzschild black hole as a function of time. Notice that the evaporation rate accelerates as the black hole approaches the end of its life.

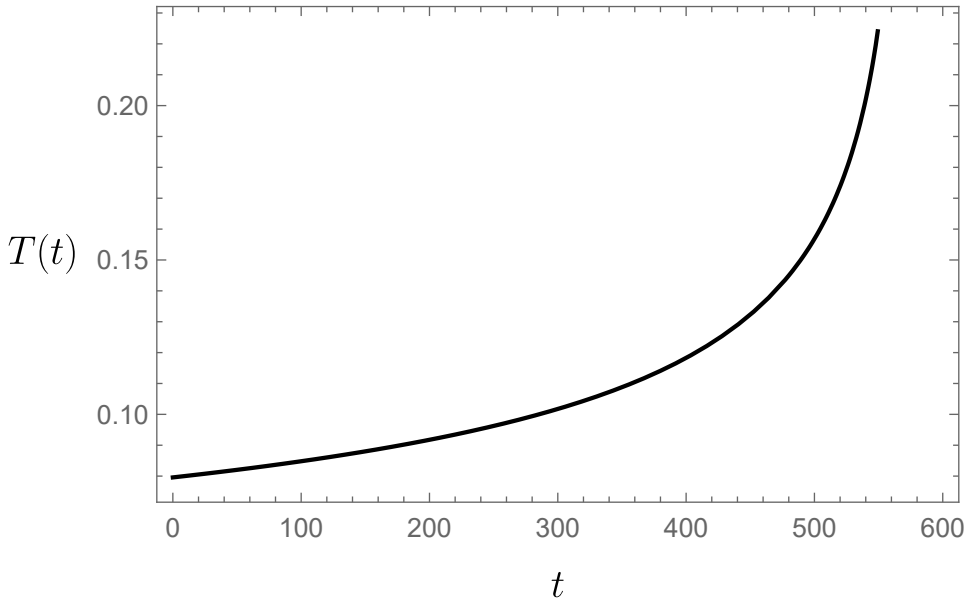

Fig. 6. A plot showing the temperature of a Schwarzschild black hole as a function of time.

6.3 Bekenstein-Hawking Entropy

Using the first law of thermodynamics, we can find the thermodynamic entropy of the black hole. Namely, we have

$$\begin{aligned} dM &= TdS, \\ \implies \frac{dS}{dt} &= \frac{1}{T}\frac{dM}{dt}. \end{aligned} \tag{6.9}$$

We can then find the black hole's entropy as a function of time simply by integrating the above. The Bekenstein-Hawking entropy of the black hole has the form [1] (where A is the area of the event horizon)

$$\begin{aligned} S &= \frac{A}{4} \\ &= \frac{4\pi r_s^2}{4} \\ &= \pi, \end{aligned} \tag{6.10}$$

at $t = 0$, where we have set the initial Schwarzschild radius to 1. We find that using $S(0) = \pi$ as an initial condition in our solving of 6.9 results in a final entropy of 0 once the black hole has evaporated completely. This is the expected result by intuition, and indicates that the plot obtained is indeed consistent with the Bekenstein-Hawking entropy. This plot is shown in 7.

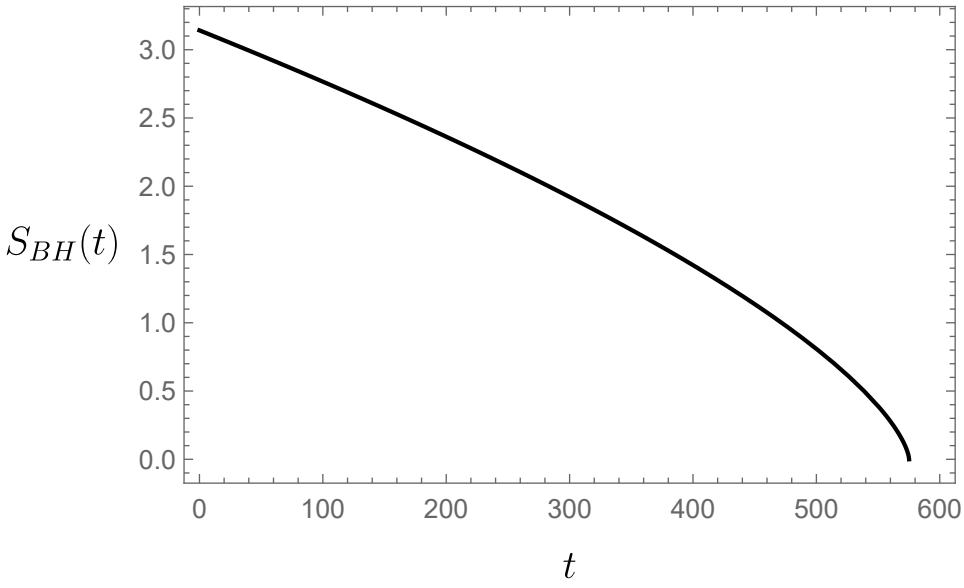

Fig. 7. A plot showing the Bekenstein-Hawking (ie, thermodynamic) entropy of a Schwarzschild black hole as a function of time.

6.4 Radiation Entropy

To find the (now considered to be naïve) entropy of the Hawking radiation, we model said radiation as a relativistic quantum gas of bosons (such as a photon gas for example) and ignore the presence of the potential barrier, which we will include by hand later on. We begin by considering a fixed number of particles N and writing down the canonical partition function,

$$\mathcal{Z}_N := \left(\sum_{i=1}^{\infty} e^{-\beta E_i} \right)^N , \tag{6.11}$$

where $\beta := \frac{1}{k_B T} \equiv \frac{1}{T}$ and $E_i = \hbar\omega_i \equiv \omega_i$. We will first consider only a single energy level with an occupation number labelled by n, such that the total number of particles N is given by

$$N = \sum_{n=0}^{\infty} n P_n, \tag{6.12}$$

where P_n is the probability that n particles occupy the energy level. The canonical partition function simply becomes

$$\mathcal{Z}_N = \left(e^{-\beta\omega} \right)^N , \tag{6.13}$$

in the limit that we have only a single energy level.

To check that this is a valid description of the Hawking radiation (at least in the limit that we ignore the potential barrier), we will allow N to vary by using the grand canonical ensemble and calculate the average number of particles in the radiation. The partition function here is

$$\begin{aligned}\mathcal{Z} &:= \sum_{N=0}^{\infty} z^N \mathcal{Z}_N \\ &= \sum_{N=0}^{\infty} z^N \left(e^{-\beta\omega}\right)^N \\ &= \sum_{N=0}^{\infty} \left(ze^{-\beta\omega}\right)^N \\ &= \frac{1}{1-ze^{-\beta\omega}},\end{aligned} \tag{6.14}$$

where $z = e^{\beta\mu}$ is the fugacity, with μ the chemical potential. The average number of particles in the grand canonical ensemble is given by

$$\langle N\rangle = z\frac{d}{dz}\ln(\mathcal{Z}), \tag{6.15}$$

so we therefore have

$$\begin{aligned}\langle N\rangle &= z\frac{d}{dz}\ln\left(\frac{1}{1-ze^{-\beta\omega}}\right) \\ &= \frac{ze^{-\beta\omega}}{1-ze^{-\beta\omega}} \\ &= \frac{1}{e^{\beta(\omega-\mu)}-1} \\ &= \frac{1}{e^{\frac{\omega}{T}}-1},\end{aligned} \tag{6.16}$$

where in the last line we have set the chemical potential to zero[11]. Notice that this coincides with 4.21, as we would hope.

To find the entropy of the radiation, we assume that the particles do not self-interact[12] and so follow Maxwell-Boltzmann statistics. As such, the occupation probability in 6.12 will be given by the Boltzmann factor

$$P_n = \mathcal{C}e^{-\beta n\omega}, \tag{6.17}$$

[11]Since we are describing a relativistic gas, the particles are massless. Hence, there is no energy barrier in creating a particle, which is equivalent to the statement $\mu = 0$.

[12]As would be the case for a gas of photons.

where $\mathcal{C}$ is a normalisation factor which can be found by requiring $\sum_{n=0}^{\infty} P_n \overset{!}{=} 1$. We find

$$\begin{aligned}\sum_{n=0}^{\infty} P_n &= \sum_{n=0}^{\infty} \mathcal{C} e^{-\beta n\omega} \\ &= \mathcal{C} \sum_{n=0}^{\infty} (e^{-\beta\omega})^n \\ &= \frac{\mathcal{C}}{1-e^{-\beta\omega}} \\ &\overset{!}{=} 1 \\ \implies \mathcal{C} &= 1 - e^{-\beta\omega}. \end{aligned} \tag{6.18}$$

We can then find the rate of change of the von Neumann entropy of the radiation using 5.22, namely

$$\begin{aligned}\frac{dS_{\text{vN}}}{dt} &= -\sum_{n=0}^{\infty} P_n \ln(P_n) \\ &= -\sum_{n=0}^{\infty} (1-e^{-\beta\omega}) e^{-\beta\omega} \ln((1-e^{-\beta\omega})e^{-\beta\omega}) \\ &= -\sum_{n=0}^{\infty} \frac{1}{\langle N\rangle + 1} \left(\frac{\langle N\rangle}{\langle N\rangle + 1}\right)^n \ln\left(\left(\frac{\langle N\rangle}{\langle N\rangle + 1}\right)^{n+1} \frac{1}{\langle N\rangle}\right), \end{aligned} \tag{6.19}$$

where we have used 6.16. Through the use of trivial laws of logarithms and the results

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \qquad \sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}, \tag{6.20}$$

we find that

$$\frac{dS_{\text{vN}}}{dt} = (\langle N\rangle + 1)\ln(\langle N\rangle + 1) - \langle N\rangle \ln(\langle N\rangle). \tag{6.21}$$

To include the effects of the potential barrier 6.1, we simply replace $\langle N\rangle$ in this expression by 6.2.

The above considers only Hawking modes at a fixed frequency and angular momentum. To consider contributions from *all* possible Hawking modes, we simply sum over angular momenta and integrate over the frequencies (just like in 6.2). Doing this, we finally obtain an expression which can be integrated to find the entropy of Hawking radiation as a function of time, namely

$$\frac{dS_{\text{rad}}}{dt} = \sum_{l=0}^{\infty} (2l+1) \int_0^{\infty} \frac{d\omega}{2\pi} ((N_{\omega,l}+1)\ln(N_{\omega,l}+1) - N_{\omega,l}\ln(N_{\omega,l})). \tag{6.22}$$

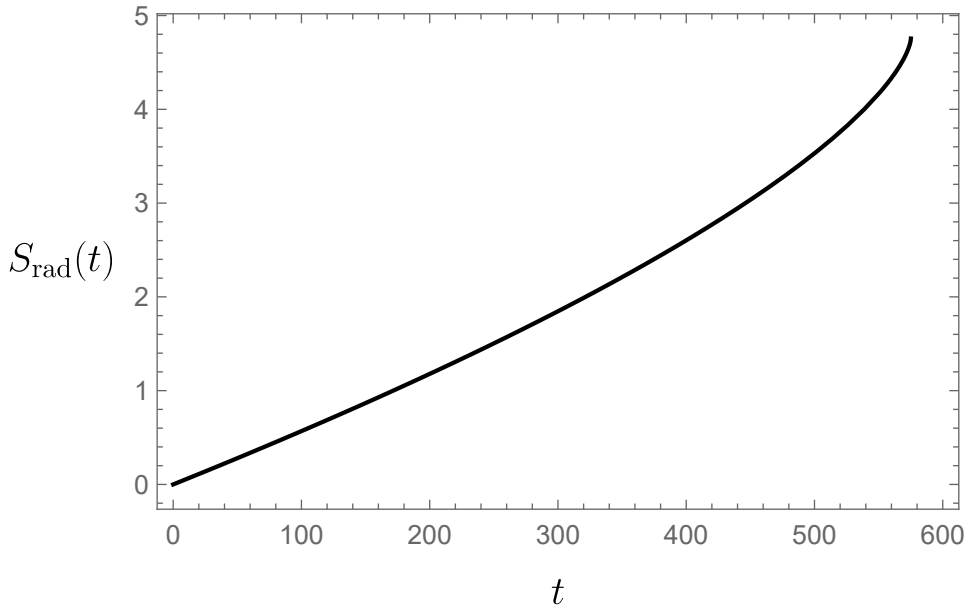

Fig. 8. A plot showing the (now considered to be näive) entropy of the Hawking radiation emitted by a Schwarzschild black hole as a function of time.

Performing said integration and plotting the result, we obtain 8. It is reasonable to consider the Hawking radiation as an entirely quantum mechanical system, and so it is reasonable to associate 8 with the *entanglement* entropy of the radiation. Notice that the radiation entropy increases monotonically as a function of time. This is a direct consequence of the fact that Hawking's calculation shows that the exterior Hawking modes which tunnel through the potential barrier and escape to infinity are in fact entangled with "interior modes" behind the event horizon [1]. To that end, each exterior Hawking mode has a "partner" behind the horizon. Thus, the more exterior Hawking modes escape the black hole, the higher the entanglement entropy of the radiation.

However, figure 8 poses a problem. If it is indeed reasonable to consider 8 to be the entropy described in chapter 5, we should be able to gain some insight about the purity of the quantum state of the radiation by looking at its entropy. We see from figure 8 that $S_{\text{rad}}(0) = 0$, and the entropy increases with time until it reaches some non-zero value at the end of the black hole's life, namely $S_{\text{rad}}(t_{\text{evap}}) \neq 0$. We therefore conclude that the quantum state of the radiation is pure at $t = 0$, and mixed at $t = t_{\text{evap}}$.

However, this cannot be a unitary process, since as we proved earlier there exists no unitary transformation which turns pure states into mixed states. Black hole evaporation therefore seems to somehow violate a central tenet of quantum mechanics. This

is the black hole information loss paradox.

We could try to argue the problem away by asserting that the black hole and the radiation are entangled and form a composite system (as is perhaps reasonable). This would lead us to the conclusion that the radiation is in fact in a mixed state for the entire evolution, and so figure 8 does not indicate a violation of unitarity. However, this would still pose a problem in that once the black hole evaporates away entirely, the radiation is no longer part of a composite system and so its entropy should vanish. As is clear from figure 8, this is not the case.

6.5 The Page Curve of a Schwarzschild Black Hole

In plotting 7 and 8 on the same axes, we can obtain the so-called "Page curve" given by $\min(S_{BH}, S_{\text{rad}})$. This was originally conjectured to be the result of a fully quantum mechanical treatment of the radiation entropy [16], which here we call the "true" entropy of the radiation, $S^{\text{true}}_{\text{rad}}(t)$. The intuition that leads to the Page curve is that the

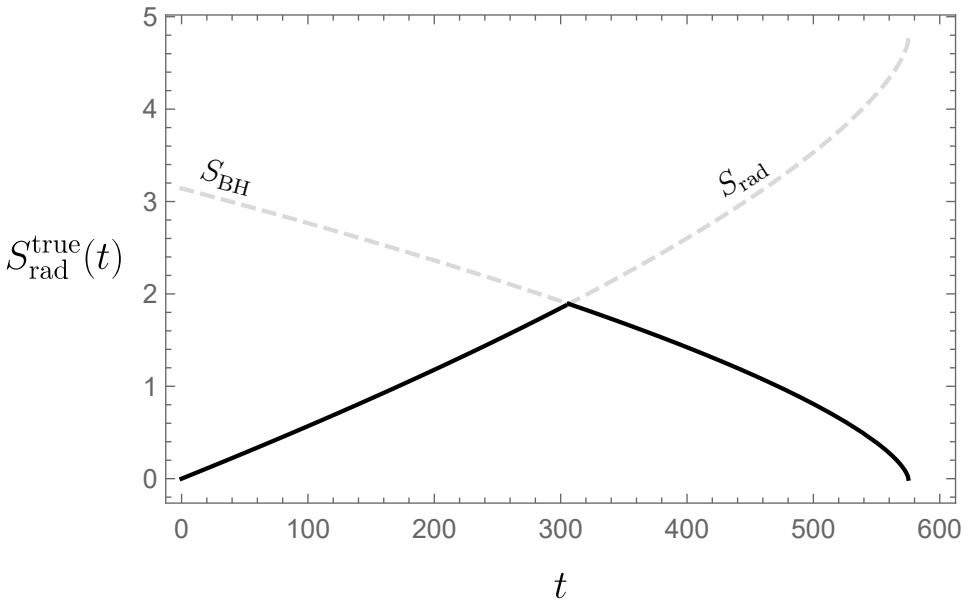

Fig. 9. The Page curve for a Schwarzschild black hole as a function of time. This is obtained by plotting the "näive" radiation entropy in figure 8 and the Bekenstein-Hawking entropy in figure 7 on the same axes. Note that the "true" radiation entropy in this plot begins and ends at $S = 0$, indicating that black hole evaporation is a unitary process.

radiation and the black hole should each be considered subsystems of a total composite

system [16]. If we assume that the overall state of the "radiation + black hole" system is pure (which is reasonable since the black hole formed from a pure state) and that the evolution of this state is unitary, then the radiation and the black hole must be entangled. Therefore, by looking at the reduced density matrix describing only the radiation, we will find that it describes a mixed state for the entire evolution of the system. We would therefore find no tension between the entropy dynamics and unitarity - the total state of the composite system remains pure from $t = 0$ to the evaporation time, and the state of the radiation remains mixed.

The Page curve certainly appears to avoid the issues associated with figure 8, but it does not constitute a solution of the information loss paradox, since until recently it was not known how to reproduce the Page curve by calculation. It was originally thought that one would need a full theory of quantum gravity to achieve this, but recent progress suggests that this is not the case - one can reproduce the Page curve using only semiclassical gravity [11, 15, 27]. We now endeavour to describe this progress.

7 Information Loss & The Island Prescription

As touched upon in the previous chapter, Hawking radiation and the resulting evaporation of black holes causes a number of problems that stem from the seemingly non-unitary evolution of the Hawking radiation system. The end-point of the evaporation is particularly troubling, and suggests that the information that should be present in the radiation is actually destroyed by the black hole when it evaporates away entirely. This is Hawking's original information loss paradox.

There is however another, more subtle, problem which appears before the endpoint of this process. To properly set up the problem, we apply the concepts discussed in chapter 5 to the black hole-radiation system.

7.1 A Pre-Endpoint Paradox

It is reasonable to describe the black hole as forming from a pure state. As the black hole begins to radiate, the system develops a bipartition. Namely, after some time we have a composite system consisting of the black hole and its emitted Hawking radiation. This is the same set-up as we had in 5, where we considered a composite quantum system described by a Hilbert space $\mathcal{H}$ which decomposed into a tensor product of Hilbert spaces describing each of its two subsystems, $\mathcal{H}_A \otimes \mathcal{H}_B$.

We assume that this composite system is formed via some unitary time evolution of the initial system (the young black hole which hasn't produced any Hawking radiation yet), and should therefore also be described by a pure state. Thus, one can calculate the entanglement entropy of the black hole and the radiation, and find that we have (via 5.24)

$$S_{\text{black hole}} = -\text{Tr}(\rho_{\text{black hole}} \ln(\rho_{\text{black hole}})) = -\text{Tr}(\rho_{\text{rad}} \ln(\rho_{\text{rad}})) = S_{\text{rad}}. \tag{7.1}$$

We now emphasise an important distinction between $S_{\text{black hole}}$ and $S_{\text{Bekenstein-Hawking}} \equiv S_{\text{BH}}$. That is, $S_{\text{black hole}} = S_{\text{rad}}$ is a measure of the *entanglement* between the black hole and the radiation, whereas S_{BH} is the *thermodynamic* entropy of the black hole. Thermodynamic entropies tell us about the microstates of a system. Namely, they tell us about the number of degrees of freedom the system has. The von Neumann (and therefore entanglement) entropy is often given the name "fine-grained", with the term "coarse-grained entropy" being associated with thermodynamic entropies such as S_{BH}. We will use these terms interchangeably.

The number of degrees of freedom a system has puts a limit on "how much entanglement" it can have with another system[13]. For example, if we have two spin-$\frac{1}{2}$ particles, each particle has only one degree of freedom - its spin - and therefore can only accommodate entanglement in one way. Since thermodynamic entropies quantify the number of degrees of freedom a system has, we therefore conclude that in a bipartite quantum system, the fine-grained (entanglement) entropy of a subsystem can never exceed the coarse-grained entropy of the subsystem it is entangled with.

Upon inspection of figure 9, we see that all is well at early times - $S_{\text{rad}}(t) < S_{\text{BH}}(t) \quad \forall\ t < t_{\text{Page}}$. However, after the Page time the "näive" fine-grained entropy of the radiation (S_{rad}) becomes *larger* than the coarse-grained entropy of the black hole (S_{BH}). This is clearly a problem - somehow the radiation and the black hole are "more entangled" than the black hole's degrees of freedom can accommodate. This is another form of the information loss paradox, and one whose solution we seek to show in section 7.5. These issues indicate that the radiation entropy shown in figure 8 cannot be correct. The reason we continually refer to 8 as the "näive" radiation entropy is that it does not follow the Page curve. A solution of the information loss paradox would take the form of an explicit calculation of the "true" fine-grained entropy of the radiation, which we expect should follow the Page curve. We now discuss some of the recent progress towards this solution.

7.2 Anti-de Sitter Space & Jackiw-Teitelboim Gravity

Here, we give a short overview of 2-dimensional Anti-de Sitter space and Jackiw-Teitelboim gravity such that we make clear the context in which the new insights into the information loss paradox are formulated.

7.2.1 2d Anti-de Sitter Space

$n-$dimensional Anti-de Sitter space (AdS_n) is a pseudo-Riemannian manifold with constant negative scalar curvature. AdS_n can be embedded in $n+1$ dimensional flat spacetime with two timelike coordinates and $n-1$ spacelike coordinates, denoted $\mathbb{R}^{2,n-1}$. AdS_2 is important for our purposes, which can be embedded in $\mathbb{R}^{2,1}$ with metric

$$ds^2 = -dX_0^2 - dX_1^2 + dX_2^2, \tag{7.2}$$

[13]This is not to be confused with *how entangled* those degrees of freedom are with eachother - this is quantified by the entanglement entropy.

subject to the hyperbolic embedding equation

$$-X_0^2 - X_1^2 + X_2^2 = \alpha^2, \tag{7.3}$$

where α is a length scale called the "AdS radius". One solution of the embedding equation is

$$\begin{aligned} X_0 &= \alpha \sin(\tau)\cosh(\rho), \\ X_1 &= \alpha \cos(\tau)\cosh(\rho), \\ X_2 &= \alpha \sinh(\rho), \end{aligned} \tag{7.4}$$

where we have introduced a coordinate system (τ, ρ). We then find

$$\begin{aligned} dX_0 &= \alpha \cos(\tau)\cosh(\rho)d\tau + \alpha \sin(\tau)\sinh(\rho)d\rho, \\ dX_1 &= -\alpha \sin(\tau)\cosh(\rho)d\tau + \alpha \cos(\tau)\sinh(\rho)d\rho, \\ dX_2 &= \alpha \cosh(\rho)d\rho, \end{aligned} \tag{7.5}$$

and substituting into 7.2 we find

$$ds^2 = \alpha^2(-\cosh^2(\rho)d\tau^2 + d\rho^2). \tag{7.6}$$

Defining so-called "global coordinates" via

$$r := \alpha \sinh(\rho),\ t := \alpha\tau, \tag{7.7}$$

we find a more familiar-looking form of the AdS_2 metric,

$$ds^2 = -\left(1 + \frac{r^2}{\alpha^2}\right)dt^2 + \frac{1}{1 + \frac{r^2}{\alpha^2}}dr^2. \tag{7.8}$$

One other coordinate system we will find most useful is so-called "Poincaré coordinates". We define new coordinates (t, z) by

$$t := \tau,\ z := \exp(\rho), \tag{7.9}$$

and plugging into 7.6 we find

$$ds^2 = \alpha^2(-\frac{1}{4}(z^2 + \frac{1}{z^2} + 2)dt^2 + \frac{1}{z^2}dz^2). \tag{7.10}$$

In the limit that z is small, and rescaling $t \longrightarrow 4t,\ z \longrightarrow \alpha z$, this reduces to the "Poincaré patch" metric,

$$ds^2 = \frac{-dt^2 + dz^2}{z^2}. \tag{7.11}$$

In light-cone coordinates $x^{\pm} = t \pm z$, this becomes

$$ds^2 = -\frac{4dx^+dx^-}{(x^+ - x^-)^2}. \tag{7.12}$$

7.2.2 Jackiw-Teitelboim Gravity

Jackiw-Teitelboim (JT) gravity is a $(1+1)$-dimensional theory of gravity where a scalar field called a dilaton is directly coupled to the Ricci scalar in the action [28]. One action describing JT gravity is of the form [29, 30]

$$S = \frac{1}{16\pi G_N}\int d^2x\sqrt{-g}\,[(\Phi R - V(\Phi)) + \mathcal{L}_{\text{matter}}], \tag{7.13}$$

where R is the Ricci scalar, Φ is the dilaton, $V(\Phi)$ is some potential for the dilaton, and g is some background spacetime metric. The dilaton plays a fundamental role in this model - its presence is required by the fact that the Einstein tensor in two dimensions is identically zero [28], and so the Einstein-Hilbert action 1.1 yields no dynamics. Additionally, the dilaton plays the role that the area of the horizon plays in higher-dimensional black holes. For example,

$$S_{BH} = \frac{A}{4G_N} \longrightarrow S_{BH} = \left.\frac{\Phi}{4G_N}\right|_{\text{horizon}} \tag{7.14}$$

in JT gravity, where $\Phi|_{\text{horizon}}$ is the value of the dilaton at the horizon [31].

Requiring the background to be AdS_2 (whose line element is given by 7.8) fixes the dilaton potential to be of the form

$$V(\Phi) = 2 - 2\Phi, \tag{7.15}$$

and also fixes the matter Lagrangian to be independent of the dilaton [29]. The resulting action is

$$S = \frac{1}{16\pi G_N}\int d^2x\sqrt{-g}\,[\Phi(R+2) - 2 + \mathcal{L}_{\text{matter}}], \tag{7.16}$$

and varying with respect to the dilaton gives

$$\begin{aligned} 0 \overset{!}{=} \delta S &= \frac{1}{16\pi G_N}\int d^2x \left(\frac{\delta(\sqrt{-g}\Phi(R+2))}{\delta\Phi} - \frac{\delta(2)}{\delta\Phi} + \frac{\delta(\mathcal{L}_{\text{matter}})}{\delta\Phi}\right)\delta\Phi \\ &= \frac{1}{16\pi G_N}\int d^2x \left(\sqrt{-g}(R+2)\frac{\delta\Phi}{\delta\Phi}\right)\delta\Phi, \end{aligned} \tag{7.17}$$

where we have used that $\sqrt{-g}(R+2)$ and $\mathcal{L}_{\text{matter}}$ are independent of the dilaton. We therefore find the constant curvature equation

$$R + 2 = 0, \tag{7.18}$$

showing that 7.15 does indeed fix the background to be AdS - the space has constant negative scalar curvature $R = -2$. Using $R = g^{\mu\nu}R_{\mu\nu}$ and $g_{\mu\nu}g^{\mu\nu} = 1$, we can see that 7.18 is equivalent to

$$R_{\mu\nu} = -2g_{\mu\nu}. \tag{7.19}$$

Importantly, this equation admits black hole solutions. In particular, the metric

$$ds^2 = -\left(1 - \frac{2G_N M}{r} + \frac{2r^2}{3}\right)dt^2 + \frac{1}{\left(1 - \frac{2G_N M}{r} + \frac{2r^2}{3}\right)}dr^2, \tag{7.20}$$

describes an asymptotic-AdS_2 black hole [32]. We can see this by looking at small radial coordinate, at which point the metric reduces to

$$ds^2 \approx -\left(1 - \frac{2G_N M}{r}\right)dt^2 + \frac{1}{\left(1 - \frac{2G_N M}{r}\right)}dr^2, \tag{7.21}$$

namely the Schwarzschild metric 4.1 with θ and ϕ constant. For large radial coordinate (namely, far from the black hole), 7.20 reduces to

$$ds^2 \approx -\left(1 + \frac{2r^2}{3}\right)dt^2 + \frac{1}{\left(1 + \frac{2r^2}{3}\right)}dr^2, \tag{7.22}$$

the AdS_2 metric 7.8 with $\alpha = \sqrt{\frac{3}{2}}$.

It turns out that AdS_2 is a natural space to consider when dealing with charged black holes in higher dimensions. Consider the metric of an extremal[14] Reissner-Nordström black hole in four dimensions,

$$ds^2 = -\left(1 - \frac{Q}{r}\right)^2 dt^2 + \frac{dr^2}{\left(1 - \frac{Q}{r}\right)^2} + r^2 d\Omega_{S^2}^2, \tag{7.23}$$

where $r^2 d\Omega_{S^2}^2$ is the metric on the round 2-sphere and Q is the charge of the black hole. Notice that the horizon for 7.23 is at $r = Q$. Now, define $\rho := r - Q$, and consider the limit that $\rho \approx 0$. This implies that $r \approx Q$, and so we are considering the near-horizon limit. We have

$$\begin{aligned} 1 - \frac{Q}{r} &= \frac{r - Q}{r} \\ &= \frac{\rho}{r} \\ &= \frac{\rho}{\rho + Q} \\ &\approx \frac{\rho}{Q}, \end{aligned} \tag{7.24}$$

[14] Namely, the charge is equal to the mass.

and

$$r = \rho + Q \implies dr = d\rho + dQ = d\rho, \tag{7.25}$$

since Q is a constant. 7.23 then becomes

$$ds^2 \approx -\left(\frac{\rho}{Q}\right)^2 dt^2 + \frac{Q^2 d\rho^2}{\rho^2} + Q^2 d\Omega_{S^2}^2, \tag{7.26}$$

which upon rescaling $\rho \to Q\rho$ and $t \to Qt$ gives

$$ds^2 \approx Q^2 \left(-\rho^2 dt^2 + \frac{d\rho^2}{\rho^2} + d\Omega_{S^2}^2\right). \tag{7.27}$$

It is easy to see that in the near-horizon limit, the geometry of an extremal Reissner-Nordström black hole is $\mathrm{AdS}_2 \times S^2$. As such, the asymptotic AdS_2 black hole in JT gravity is in fact a "toy model" for the spherically symmetric sector[15] of the near-horizon limit of Reissner-Nordström black holes in $(3+1)$ dimensions.

7.3 Generalised Entropy and the Island Prescription

Note that in chapter 6, we used only arguments from statistical mechanics and thermodynamics to derive the expression for the entropy of the black hole (6.9) and that of the Hawking radiation (6.22) - gravity appeared nowhere in the calculation. One would hope that if a *gravitational* formula were used we would gain a far more accurate picture of the behaviour of the radiation, since gravity itself is entirely responsible for its production. It turns out that insights from AdS/CFT [10] have surprisingly provided such gravitational formulae for the fine-grained entropy of both the black hole *and* its Hawking radiation [15], both of which follow the Page curve. The formulae and their arguments therefore seem to solve the information loss paradox, at least for asymptotic-AdS_2 black holes in JT gravity [11]. We now describe this progress.

First, we consider collecting the Hawking radiation from the black hole in some region R far away from the event horizon, near future null infinity ($\mathscr{I}^+$). In calculating the entropy of the quantum fields S_{QFT} on this region, we obtain figure 8, the result of Hawking's calculation. The new insight [11, 15, 27] is that there exist previously overlooked contributions to the von Neumann entropy of the Hawking radiation from so-called "islands" (denoted by I). These are regions of the black hole interior containing matter entangled with the radiation that must be taken into account in order to preserve unitarity and obtain entropy dynamics that follow the Page curve [11]. Figure 10 shows

[15] Namely, we take the S^2 factor in the near-horizon topology to be irrelevant, leaving only AdS_2.

a typical set-up, with a single radiation region R near $\mathscr{I}^+$ and a single island inside the event horizon. Many different island configurations are possible, and as proposed recently one can even obtain the Page curve by considering "islands in the stream" of the Hawking radiation itself [33].

The von Neumann entropy of the Hawking radiation is now understood to be given by the so-called "island formula" [11],

$$\boxed{S(R) = \min_{\partial I}\left\{\text{ext}_{\partial I}\left[S_{\text{QFT}}(R\cup I) + \sum_{\partial I}\frac{\text{Area}(\partial I)}{4G_N}\right]\right\}.} \tag{7.28}$$

We see that 7.28 contains a term that resembles the Bekenstein-Hawking entropy formula $S_{\text{BH}} = \frac{A}{4G_N}$, though here we are calculating the area of the boundary of an island. This, along with a term representing the standard QFT entanglement entropy of the quantum fields on both the radiation *and* the island regions, gives the *generalised entropy*,

$$S_{\text{gen}} := S_{\text{QFT}}(R\cup I) + \sum_{\partial I}\frac{\text{Area}(\partial I)}{4G_N}. \tag{7.29}$$

The island boundaries ∂I which extremise the generalised entropy are known as "Quantum Extremal Surfaces" (QESs), and are by definition surfaces of minimal generalised entropy in time and maximal generalised entropy in space[16] [27]. One can therefore think of the island boundaries as representing saddle points in the generalised entropy.

Another important point to make is that the term $S_{\text{QFT}}(R\cup I)$ is small throughout the evaporation process. This is because although the density matrix ρ_R on R describes a mixed state, the island region I contains (almost) all of the interior matter fields entangled with the radiation. R and I are two subsystems of the bipartite system $R\cup I$, and therefore the state on $R\cup I$ is (almost) pure[17].

In principle there could be an arbitrary number of possible islands, each with different boundary locations, and so we need to somehow take this into account when calculating the fine-grained entropy of the radiation. This is encoded into the island formula, and thereby into the "island prescription", which is given below [11].

[16] So-called "maximin" surfaces.

[17] Another way to word this is that the island "purifies" the state on R.

The Island Prescription:

1. Compute the generalised entropy for a chosen configuration of islands.
2. Extremise the generalised entropy with respect to the location of all island boundaries.
3. If there is more than one extremum in the generalised entropy, take the smallest of these to be the true entropy of the radiation.

Figure 10 shows just one island configuration, whereby we have one island I with two boundaries. Note that since the Penrose diagram suppresses two dimensions, each boundary ∂I in fact represents a 2-sphere, and so this particular island is a spherical shell. One can imagine other island configurations, where for example the island extends from some location inside the horizon all the way to the spatial origin. In this case the island is a solid sphere, rather than a shell. Figure 11 shows this island configuration.

7.4 The Island Prescription & The Page Curve

Here, we argue qualitatively that the island formula 7.28 does indeed produce entropy dynamics that follow the Page curve.

The black hole begins emitting Hawking radiation soon after it forms, but an island does not appear until some time after this [15]. Before an island forms, 7.28 simply reduces to Hawking's result (since $I = \partial I = \varnothing$)

$$S(R) = S_{\text{QFT}}(R), \tag{7.30}$$

and so the fine-grained radiation entropy rises as more and more Hawking modes (which are entangled with their "partners" behind the horizon) are collected at R. Namely, if no other mechanism interrupted this process we would obtain figure 8.

Now consider the island configuration shown in figure 11. In this case, it turns out that the QES is close to the event horizon of the black hole [15], and so one can approximate the area of the QES by the area of the event horizon. The area term in 7.29 therefore begins at some large value for early times, decreases as the black hole evaporates, and vanishes at the evaporation time. As mentioned earlier,

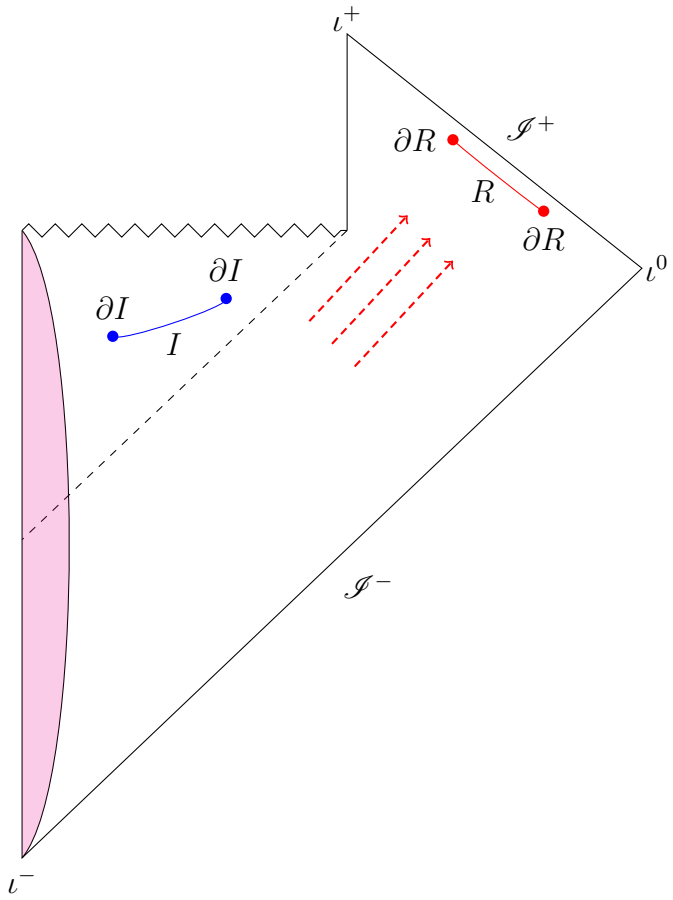

Fig. 10. The Penrose diagram for an evaporating black hole, showing the Hawking radiation in dashed red, the region R in which we collect the Hawking radiation, as well as an island I with its endpoints ∂I, the Quantum Extremal Surfaces. The portion filled in pink shows matter which collapses to form a black hole.

the $S_{\text{QFT}}(R \cup I)$ term is small for all times, and so the area term in the generalised entropy dominates after the island forms. The generalised entropy therefore mimics the thermodynamic Bekenstein-Hawking entropy of the black hole (shown in figure 7) after the island appears.

Finally, as we can see from 7.28, the true fine-grained entropy of the Hawking radiation at any particular time is given by the smallest of these two contributions - the early-time no-island contribution, and the later-time island contribution. We therefore have a "turn-over" in the fine-grained entropy at the Page time, and the island formula does indeed reproduce the Page curve, an example of which is shown in figure 9.

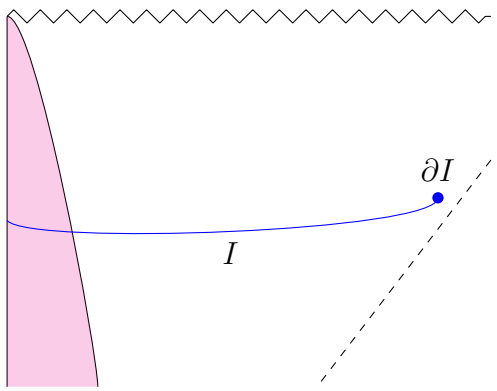

Fig. 11. A portion of a Penrose diagram for an evaporating black hole, showing a different island configuration to that depicted in figure 10. Here we see a single QES with an island that extends to the spatial origin.

Though the island prescription was initially considered for asymptotic-AdS$_2$ black holes in JT gravity, it is hoped that these ideas will give us some insight into the kinds of black holes present in *our* universe. Indeed, [34] investigates islands for Schwarzschild black holes in four dimensions.

7.5 The Island Formula In Action

We now work through an explicit example using the island formula. This calculation was first performed in [12], and the following is heavily influenced by the review of that calculation found in [31]. The calculation relies on a number of technical results from two dimensional conformal field theory[18] (CFT) which we will not discuss in detail here - the reader is instead directed to a number of references along the way.

The scenario we consider is a pair of asymptotic-AdS$_2$ black holes, with each of their exterior AdS regions "glued" to a copy of flat half-Minkowski space. Figure 12 shows the set-up. Each flat space region contains a bath - in this context a bath is some quantum mechanical system where we can neglect gravity altogether [12]. It is in the bath that the Hawking modes will be collected.

Defined across the entire spacetime is a CFT where the state is taken to be [31]

$$|\psi\rangle = \sum_n e^{-\frac{\beta E_n}{2}} |\psi_n\rangle_L \otimes |\psi_n\rangle_R \,, \tag{7.31}$$

[18] A CFT is a particularly special kind of quantum field theory which is invariant under so-called "conformal transformations" [35].

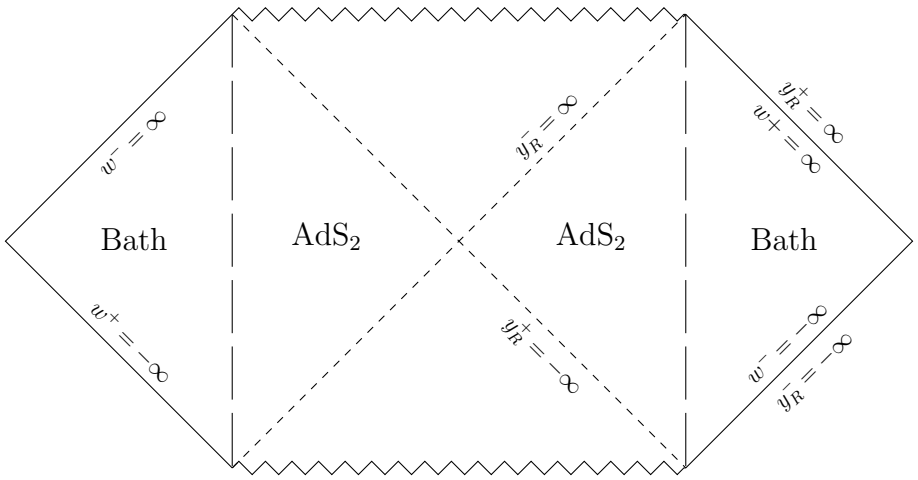

Fig. 12. A pair of asymptotic-AdS$_2$ black holes, each coupled to a bath living in flat Minkowski space. "Transparent" boundary conditions are imposed on the AdS-Minkowski interfaces, as indicated by the dashed vertical lines. $y_R^\pm$ covers the right-hand diamond (with $y_L^\pm$ covering the left), whereas the $w^\pm$ coordinates cover the entire spacetime.

where $\{|\psi_n\rangle_L\}$ and $\{|\psi_n\rangle_R\}$ define a basis of states on the left and right sides of the geometry respectively and the E_n are energy levels associated with these states. This is the so-called "thermofield double state". We consider this scenario in the context of JT gravity, with action [12]

$$S = \frac{1}{16\pi G_N}\int d^2x\sqrt{-g}\,[(\phi R + 2(\phi - \phi_0))] + S_{\rm CFT}, \tag{7.32}$$

where $S_{\rm CFT}$ is the action for the CFT. As mentioned earlier, this can be thought of as a toy model for the spherically symmetric sector of the near-horizon limit of charged black holes in higher dimensions - ϕ_0 is a constant that sets the extremal entropy of such a black hole. We impose "transparent" boundary conditions at the AdS-Minkowski interfaces such that any radiation leaving the black holes can propagate freely into the bath, and in turn any modes from the bath can propagate into the black hole. This ensures that the black holes and their respective baths are always in thermal equilibrium, and that the entire geometry is static. As such, the temperature of the black holes β^{-1} is equal to the temperature of the baths.

We take the metric in the AdS regions to be the Poincaré patch metric

$$ds^2 = -\frac{4dx^+dx^-}{(x^+ - x^-)^2}, \tag{7.33}$$

and the metric in the bath regions to be the standard 2d Minkowski metric

$$ds^2 = -dt^2 + d\sigma^2 = -dy^+ dy^-, \tag{7.34}$$

where we have defined $y^\pm := t \pm \sigma$. We introduce another set of coordinates $w^\pm$ which cover both the bath regions *and* the AdS regions, related to the Poincaré coordinates by

$$x^\pm = \pm \frac{\beta}{\pi} \frac{w^\pm \mp 1}{w^\pm \pm 1}, \tag{7.35}$$

and to the Minkowski coordinates by

$$w^\pm = \pm \exp\left(\pm \frac{2\pi y_R^\pm}{\beta}\right), \; w^\pm = \mp \exp\left(\mp \frac{2\pi y_L^\pm}{\beta}\right), \tag{7.36}$$

for the right-hand and left-hand black holes, respectively[19]. The right-hand black hole's event horizon is therefore at $w^- = 0$, shown by setting $y_R^- = \infty$ in the above. From now on, we will drop the R, L subscripts and the reader should assume we are considering the right-hand black hole unless otherwise stated. We can extend the $y^\pm$ coordinates into the AdS region by the relation

$$x^\pm = \frac{\beta}{\pi} \tanh\left(\frac{\pi y^\pm}{\beta}\right), \tag{7.37}$$

such that the $y^\pm$ coordinates cover the entire right diamond of the diagram 12.

In this set-up, we take the CFT to be in a thermal state from the point of view of the Minkowski coordinates. The stress-energy tensor is therefore [12]

$$T_{y^\pm y^\pm} = \frac{\pi c}{12\beta^2}, \tag{7.38}$$

where c is a constant known as the "conformal anomaly number" of the CFT [36], which we take to be large such that it is valid to keep the analysis entirely semiclassical [31]. We then want to find coordinates such that the CFT is in the vacuum state. This is analogous to the Unruh effect scenario in chapter 3, where we had an inertial observer (set of coordinates) with a vacuum state, and a non-inertial observer who measured this state to actually be thermal. It turns out that the coordinates to use are the Kruskal-like $w^\pm$ coordinates, where the stress tensor vanishes[20],

$$T_{w^\pm w^\pm} = 0. \tag{7.39}$$

[19] Notice that 7.36 are reminiscent of the Kruskal coordinates 4.11.

[20] See appendix C for a proof.

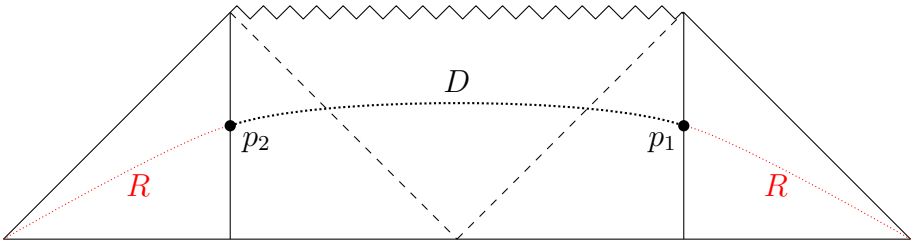

Fig. 13. We calculate the entanglement entropy of the interval D, a subset of the Cauchy slice $D \cup R$. Since we assume the overall state of the system to be pure, this yields the entanglement entropy of the radiation collected in the baths ($S(D) = S(R)$). p_1 and p_2 are the endpoints of the interval D.

Writing the Minkowski metric in these vacuum coordinates, we find

$$w^{\pm} = \pm \exp\left(\pm \frac{2\pi}{\beta} y^{\pm}\right) \implies y^{+} = \frac{\beta}{2\pi} \ln(w^{+}),\ y^{-} = -\frac{\beta}{2\pi} \ln(-w^{+}), \tag{7.40}$$

and so

$$ds^2 = -dy^{+}dy^{-} = \left(\frac{\beta}{2\pi}\right)^2 \frac{dw^{+}dw^{-}}{w^{+}w^{-}}. \tag{7.41}$$

We want to calculate the entanglement entropy of the radiation in the baths at some time t. To do so, we choose a Cauchy slice extending across the entire spacetime and calculate the entanglement entropy of a subset D of this slice which extends across the two AdS regions. This is equivalent to calculating the entropy of the radiation region R in the baths, since $S(D) = S(R)$ when the overall state of the system is pure. Figure 13 shows the set-up. The entanglement entropy of an interval in a 2-dimensional CFT is given by a 2-point correlation function of so-called "twist operators". For a metric of the form $ds^2 = -d\rho d\bar{\rho}$, this correlation function is given by [36]

$$\langle \mathcal{O}_{\Delta}(\rho, \bar{\rho}) \mathcal{O}_{\Delta}(\rho', \bar{\rho}') \rangle = \left[\frac{1}{(\rho - \rho')(\bar{\rho} - \bar{\rho}')}\right]^{\Delta}, \tag{7.42}$$

where Δ is the "conformal dimension" of the twist operator $\mathcal{O}$, with $(\rho, \bar{\rho})$ and $(\rho', \bar{\rho}')$ the coordinates of the end-points of the interval we are interested in. If the metric has a conformal factor after some coordinate transformation of the form

$$ds^2_{\rho,\bar{\rho}} \longrightarrow \Omega^{-2} ds^2_{\omega,\bar{\omega}}, \tag{7.43}$$

the correlation function transforms as [36]

$$\langle \mathcal{O}_{\Delta}(\rho, \bar{\rho}) \mathcal{O}_{\Delta}(\rho', \bar{\rho}') \rangle \longrightarrow \langle \mathcal{O}_{\Delta}(\omega, \bar{\omega}) \mathcal{O}_{\Delta}(\omega', \bar{\omega}') \rangle = \left[\frac{\Omega(\omega, \bar{\omega}) \Omega(\omega', \bar{\omega}')}{(\omega - \omega')(\bar{\omega} - \bar{\omega}')}\right]^{\Delta}. \tag{7.44}$$

Using that

$$x^n = \exp(n \ln(x)) \approx 1 + n \ln(x), \tag{7.45}$$

in the limit $n \to 0$, we obtain

$$\langle \mathcal{O}_\Delta(\omega, \bar{\omega}) \mathcal{O}_\Delta(\omega', \bar{\omega}') \rangle \approx 1 - \Delta \ln \left(\frac{(\omega - \omega')(\bar{\omega} - \bar{\omega}')}{\Omega(\omega, \bar{\omega}) \Omega(\omega', \bar{\omega}')} \right), \tag{7.46}$$

in the limit $\Delta \to 0$.

For our metric 7.41, we have

$$\Omega^{-2} = -\left(\frac{\beta}{2\pi} \right)^2 \frac{1}{w^+ w^-} \implies \Omega = \frac{2\pi i}{\beta} \sqrt{w^+ w^-}. \tag{7.47}$$

The end-points of the interval D (p_1 and p_2) are at $w_1^\pm = \pm \exp\left(\pm \frac{2\pi t}{\beta} \right)$ and $w_2^\pm = \mp \exp\left(\mp \frac{2\pi t}{\beta} \right)$ (respectively)[21]. Taking ω and $\bar{\omega}$ to be w_1^+ and w_1^- respectively, and taking ω' and $\bar{\omega}'$ to be w_2^+ and w_2^- respectively, we find

$$\langle \mathcal{O}_\Delta(w_1^+, w_1^-) \mathcal{O}_\Delta(w_2^+, w_2^-) \rangle \approx 1 - \Delta \ln \left(\frac{(w_1^+ - w_2^+)(w_1^- - w_2^-)}{\Omega(w_1^+, w_1^-) \Omega(w_2^+, w_2^-)} \right). \tag{7.48}$$

The entanglement entropy of the region D is then extracted from this as [31]

$$S(D) = \frac{c}{6} \ln \left(\frac{-(w_1^+ - w_2^+)(w_1^- - w_2^-)}{\Omega_1 \Omega_2} \right), \tag{7.49}$$

where we have used the notation

$$\Omega_{1,2} \equiv \left(\frac{2\pi}{\beta} \right) \sqrt{w_{1,2}^+ w_{1,2}^-}\,. \tag{7.50}$$

[21] We take the points to be inside the bath regions, but infinitesimally close to the AdS-Minkowski interface.

Using $w^+_{1,2}w^-_{1,2} = -1 \implies \sqrt{w^+_{1,2}w^-_{1,2}} = i$, we find

$$\begin{aligned} S(D) &= \frac{c}{6}\ln\left(\frac{-(w_1^+ - w_2^+)(w_1^- - w_2^-)}{\Omega_1\Omega_2}\right) \\ &= \frac{c}{6}\ln\left[\left(\frac{\beta}{2\pi}\right)^2\left(2\cosh\left(\frac{2\pi t}{\beta}\right)\right)\left(2\cosh\left(\frac{2\pi t}{\beta}\right)\right)\right] \\ &= \frac{c}{6}\ln\left(\left(\frac{\beta}{2\pi}\right)^2 4\cosh^2\left(\frac{2\pi t}{\beta}\right)\right), \end{aligned} \tag{7.51}$$

and so we finally obtain

$$\boxed{S_{\text{no-island}} \equiv S(D) = S(R) = \frac{c}{3}\ln\left(\frac{\beta}{\pi}\cosh\left(\frac{2\pi t}{\beta}\right)\right).} \tag{7.52}$$

Notice that for large times $t \gg \beta$, we have

$$\begin{aligned} S(R) &\approx \frac{c}{3}\ln\left(\frac{\beta}{\pi}\exp\left(\frac{2\pi t}{\beta}\right)\right) \\ &\sim \frac{2\pi c t}{3\beta}, \end{aligned} \tag{7.53}$$

and so the radiation entropy rises indefinitely as more and more Hawking modes are collected in the bath. This is expected, since the set-up used here (ie, that shown in figure 13) corresponds to the no-island scenario. However, this becomes a problem as soon as the radiation entropy exceeds the Bekenstein-Hawking entropy of the black hole, when the "pre-endpoint paradox" mentioned in section 7.1 is manifest. Without some mechanism for capping $S(R)$ at $2S_{\text{BH}}$ (the factor of 2 present because we have two black holes), the above "no-island" scenario exhibits an information loss paradox.

We can make an attempt to solve the paradox by implementing an island into the program via the set-up shown in figure 14. In doing so, we introduce two QESs at the points p_1 and p_3, with coordinates $w_1^\pm = w_3^\mp$. The points p_2 and p_4 are located at $w_2^\pm = w_4^\mp = \pm\exp\left(\pm\frac{2\pi t}{\beta}\right)$. As such, a version of the island formula 7.28 will need to be used, and the locations of the QESs will need to be extremised over. Since the overall state is pure, we can calculate the entropy on the region $R\cup I$ by computing the entropy on the two-interval region D. Calculating the CFT entanglement entropy on a two-interval region is in general complicated, but working in the large-time limit it

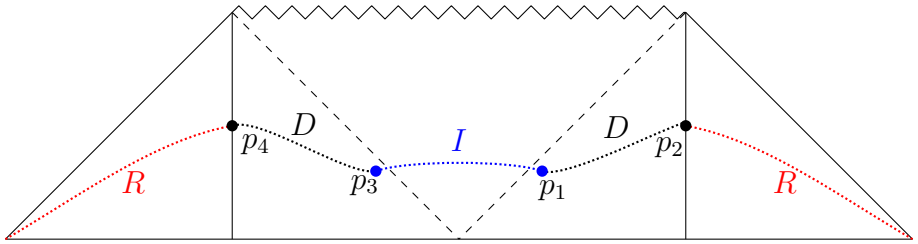

Fig. 14. A non-trivial set-up whereby we have a single island with two quantum extremal surfaces at p_1 and p_3. As such, we must calculate the entanglement entropy contributions from both D intervals on either side of the island.

is valid to simply calculate the entanglement entropy for the right-hand D region and double the result to account for the left-hand black hole[22].

The point p_1 is the QES, and according to the island prescription its location (namely $w_1^{\pm}$) must be extremised over. The island formula in this example is therefore

$$S(D) = 2 \times \min_{w_1^{\pm}} \left\{ \text{ext}_{w_1^{\pm}} \left[\frac{\phi(w_1^{\pm})}{4G_N} + S_{\text{QFT}}(D) \right] \right\}, \tag{7.54}$$

where the role of the area of the island boundary is played by the value of the dilaton at the QES. The point p_1 is considered to be in the AdS region (following [12]), and so we will need the conformal factor Ω_1 that comes from sending 7.33 to $w^{\pm}$ coordinates. We have

$$ds^2 = -\frac{4dx^+dx^-}{(x^+ - x^-)} \longrightarrow ds^2 = -\frac{16dw^+dw^-}{(1+w^+w^-)^2}, \tag{7.55}$$

and so we see that the conformal factor for p_1 is

$$\Omega_1^{-2} = \frac{4}{(1+w^+w^-)^2}. \tag{7.56}$$

The point p_2 is considered to be just inside the bath, and the conformal factor is therefore (as before)

$$\begin{aligned} \Omega^{-2} &= -\left(\frac{\beta}{2\pi}\right)^2 \frac{1}{w^+w^-} \\ \implies \Omega_2^{-2} &= -\left(\frac{\beta}{2\pi}\right)^2 \frac{1}{w_2^+w_2^-} \\ &= \left(\frac{\beta}{2\pi}\right)^2. \end{aligned} \tag{7.57}$$

[22]The entanglement entropy for a two-interval region is in general a function of the "cross-ratio" of the interval end-points [36, 37], but in the long-time limit this cross ratio goes to 1 [31].

We therefore find that the entropy contribution from the right-hand interval D is given by (following 7.49)

$$S_{\text{QFT}}(D) = \frac{c}{6}\ln\left(\frac{(w_2^+ - w_1^+)(w_1^- - w_2^-)}{\frac{1}{2}(1+w_1^+w_1^-)\frac{2\pi}{\beta}}\right)$$

$$= \frac{c}{6}\ln\left(\frac{\beta}{\pi}\right) + \frac{c}{6}\ln\left(\frac{(e^{\frac{2\pi t}{\beta}} - w_1^+)(w_1^- + e^{\frac{-2\pi t}{\beta}})}{(1+w_1^+w_1^-)}\right). \tag{7.58}$$

The dilaton in the $w^\pm$ coordinates in this model is given by [12]

$$\phi(w^\pm) = \phi_0 + \frac{2\pi\phi_r}{\beta}\frac{1-w^+w^-}{1+w^+w^-}, \tag{7.59}$$

where ϕ_r is a constant[23]. Using 7.54, we therefore find that the generalised entropy is

$$S_{\text{gen}} = \frac{\phi(w_1^\pm)}{4G_N} + S_{\text{QFT}}(D)$$
$$= \frac{\phi_0}{4G_N} + \frac{c}{6}\ln\left(\frac{\beta}{\pi}\right) + \frac{c}{6}f(w_1^\pm), \tag{7.60}$$

where we define

$$f(w_1^\pm) := \frac{\pi}{\beta k}\frac{1-w_1^+w_1^-}{1+w_1^+w_1^-} + \ln\left(\frac{(e^{\frac{2\pi t}{\beta}} - w_1^+)(w_1^- + e^{\frac{-2\pi t}{\beta}})}{(1+w_1^+w_1^-)}\right). \tag{7.61}$$

The constant

$$k := \frac{G_N c}{3\phi_r}, \tag{7.62}$$

is relevant for non-equilibrium (and therefore evaporating) black holes, where it sets the evaporation rate. We work in the "adiabatic approximation", where the evaporation is so slow that the black holes are effectively in equilibrium, and thus k is taken to be small [31]. We also work in the high temperature limit, where β is small.

Step 2 in the island prescription is to extremise the generalised entropy with respect to the island boundaries, which in our case amounts to extremising 7.61. We find

$$\frac{\partial f}{\partial w_1^+} \overset{!}{=} 0 \implies w_1^- = -\frac{\beta k}{\pi}e^{-\frac{2\pi t}{\beta}}, \tag{7.63}$$

$$\frac{\partial f}{\partial w_1^-} \overset{!}{=} 0 \implies w_1^+ = \frac{\beta k}{\pi}e^{\frac{2\pi t}{\beta}}, \tag{7.64}$$

[23] ϕ_r is a feature of the CFT on the boundary of AdS, dual to the gravitational theory according to the AdS/CFT correspondence [9]. This is not important for our discussion, and is not to be confused with the bulk CFT described by the thermofield double state.

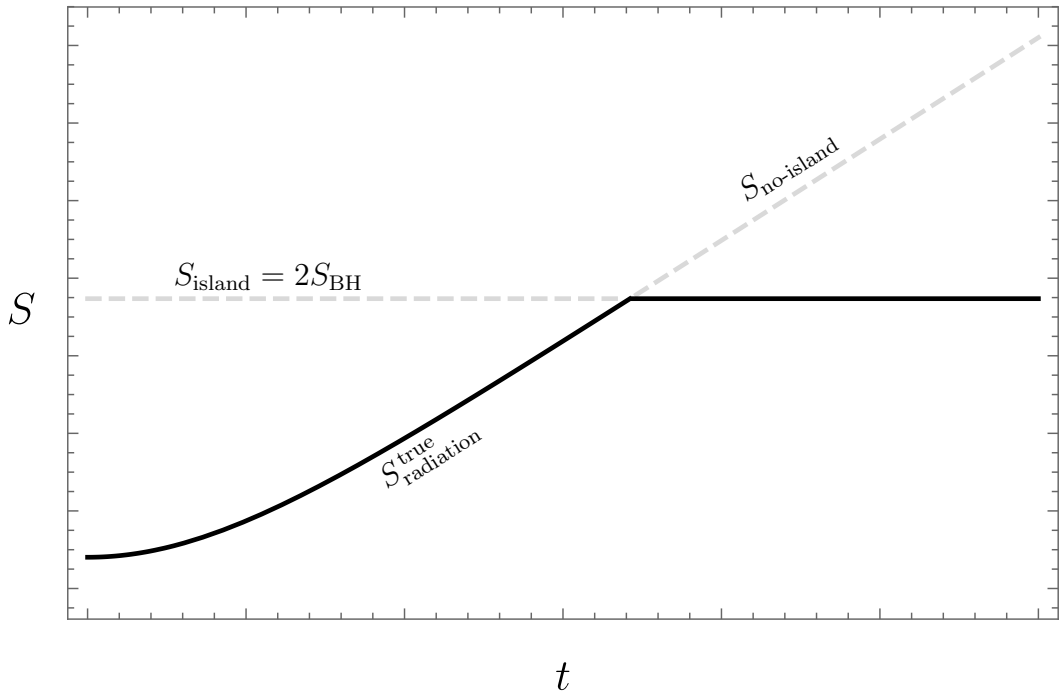

Fig. 15. The Page curve for a pair of asymptotic-AdS_2 black holes, each coupled to a bath living in flat Minkowski space, in the context of JT gravity.

in the limit that β and k are small. Substituting these values of $w_1^{\pm}$ back into 7.61, we find that the generalised entropy is given by

$$S_{\text{gen}} = \frac{\phi_0}{4G_N} + \frac{c}{6}\ln\left(\frac{\beta}{\pi}\right) + \frac{\pi c}{6\beta k}, \tag{7.65}$$

a constant. We identify the first term in this expression with the Bekenstein-Hawking entropy of the black hole, and identify the other two terms with quantum corrections to the Bekenstein-Hawking entropy arising from the CFT. We therefore find (using 7.54)

$$\boxed{S_{\text{island}} \equiv S(R \cup I) = S(D) = 2 \times S_{\text{gen}} = 2S_{\text{BH}},} \tag{7.66}$$

where we have doubled the result to account for both black holes. The full entropy of the radiation is then given by $\min(S_{\text{no-island}}, S_{\text{island}})$ in accordance with step 3 of the island prescription, which yields the Page curve shown in figure 15.

By including the contribution of an island in the calculation of the entanglement entropy of the baths, we therefore see that the entropy no longer increases indefinitely and instead saturates at $2S_{\text{BH}}$. As such, the fine-grained entropy of the radiation never exceeds the coarse-grained entropy of the black holes, and the information loss paradox present in the no-island scenario is avoided.

8 Discussion & Conclusions

8.1 Summary

In this dissertation, we formulated the theory required for understanding the issues that inexorably lead to the black hole information loss paradox. In chapters 2-4 we gave an overview of the key ideas that lead to Hawking radiation. We considered Bogoliubov transformations between annihilation and creation operators associated with two solutions of the Klein-Gordon equation, and we associated some of the resulting expressions with coordinate transformations of the Schwarzschild metric. We found that in the limit that we ignore the presence of any potential barrier outside the event horizon, these ideas lead to the conclusion that a Schwarzschild black hole emits a thermal bath of particles with temperature $T = \frac{\hbar}{8\pi G_N M}$.

In chapter 5, we formulated the elements necessary for a fully quantum mechanical treatment of entropy. We explained the concepts of pure and mixed states, and showed that mixed states arise when we have a classical uncertainty associated with quantum states, as well as when we consider subsystems of larger composite systems. We introduced the use of density matrices, and showed that they are central to a powerful formalism that allows us to deal with both pure *and* mixed states. Using the density matrix formalism we showed that a central tenet of quantum mechanics, unitarity, implies that pure states cannot evolve into mixed states and vice-versa. We then argued that a notion of entropy can be associated with quantum states, and that pure states have zero entropy while mixed states have non-zero entropy.

Chapter 6 saw an exposition of the thermodynamics of an evaporating Schwarzschild black hole. We began by introducing the greybody factor, which induces a correction to the spectrum of Hawking radiation and needs to be taken into account due to the presence of a potential barrier outside the event horizon. We obtained plots for the potential barrier for a single angular momentum mode, and the greybody factors for multiple values of angular momentum. Using these greybody factors, we were able to obtain plots for the mass, temperature and Bekenstein-Hawking entropy of the black hole as a function of time. Through a statistical mechanical calculation we were able to find a differential equation describing the evolution of the entropy of the Hawking radiation, and argued that this "näive" calculation results in entropy dynamics which violate quantum mechanics in a fundamental way. We then argued that Page's theorem provides a qualitative solution to the problem and argued that the Page curve should be the result of a more careful calculation.

Chapter 7 began with arguments for an entanglement paradox that occurs immediately after the Page time, significantly before the end-point of the evaporation process. We then gave a short description of Anti-de Sitter spaces and Jackiw-Teitelboim gravity, before providing an overview of the island prescription. We argued qualitatively that this reproduces the Page curve, thereby avoiding information loss. Finally, we put the island prescription to use and calculated the entropy dynamics for a pair of asymptotic-AdS_2 black holes in JT gravity, each coupled to a bath living in flat Minkowski space. In doing so, we found that the radiation entropy from a no-island contribution grows linearly (modulo early-time behaviour), but is capped at the Bekenstein-Hawking entropy of the black holes by a contribution from an island. As such, the fine-grained entropy of the radiation never exceeds the coarse-grained entropy of the black holes, and the form of information loss presented in 7.1 is avoided.

8.2 Outlook & Discussion

There are a number of interesting and important ideas we have neglected to present. We note that although chapters 2-4 covered only scalar particle production, the same ideas can be formulated using spinor fields to show that fermions are also produced. This is a non-trivial extension to the work presented here - in writing the Dirac equation in curved spacetime and coupling spinors to gravity for example, one must consider a so-called "spin connection" in place of the Christoffel symbols.

One somewhat unsatisfying point in our presentation is that we ignored the presence of the potential barrier outside the black hole horizon when calculating the spectrum of Hawking radiation in chapter 4. A direction the author wishes to look into in future work is how to derive the correct Hawking spectrum from the outset, such that the greybody factor does not need to included "by hand" as in chapter 6.

In chapters 4 and 6 we considered only the simple case of a Schwarzschild (ie, uncharged and non-rotating) black hole. An additional direction for future work could be to repeat the numerical calculations in chapter 6 for a Kerr (rotating) or a Reissner-Nordström (charged) black hole, in order to see how the Page curve for these differs from the Schwarzschild case. We expect that the resulting plots will indeed differ for a number of reasons. For instance, Hawking radiation in the Kerr case will carry not only energy, but also angular momentum away from the black hole [1]. We would therefore expect that the mass, temperature, black hole entropy and radiation entropy plots will all have some corrections dependent on the angular momentum of the black hole. A source of complication in the Reissner-Nordström case is that it has *two* horizons [38]

- the event horizon and an internal "Cauchy horizon" [39]. Moreover, since these black holes become "extremal" and there exists a lower bound on their mass, the black hole will not evaporate away entirely. We expect that this would drastically modify the plots given in 6 - in particular we would expect that the Page curve will not return to zero and will instead decay to some constant.

Though a short overview was given in chapter 7, JT gravity is an extremely rich model - the surface of which we only just scratched. One particularly attractive feature of JT gravity is that it appears to be fully quantisable [40]. The model has certainly proved a useful tool in probing information loss, and as such it may prove to be a vital stepping-stone towards the formulation of a full, four-dimensional theory of quantum gravity.

In section 7.5 we considered how the island prescription governs the entropy dynamics of asymptotically AdS black holes, and we imposed transparent boundary conditions at the AdS-Minkowski interfaces. It would be interesting to see how the analysis would change were we to impose "one-way" (instead of transparent) boundary conditions, such that modes from the baths are prevented from propagating into the AdS regions. Modes from the black hole however would still propagate into the baths, and the black holes would be allowed to evaporate. As such, the Bekenstein-Hawking entropy should have some time dependence, and we are led to predict that the Page curve in this case would likely resemble figure 9.

Computing the entropy of Hawking radiation by a standard QFT calculation on only the radiation region R is now considered to be "näive". As we showed in section 7.5, the correct thing to do is to include contributions from so-called "islands" I inside the black hole, and to that end calculating the true entropy of the radiation involves computing the entanglement entropy on $R \cup I$. As such, beyond the Page time the island prescription instructs us to think of the islands *not* as part of the black hole, but actually as part of the *radiation*. This is rather startling and so we restate for emphasis - to avoid information loss it appears as though we must consider parts of the black hole *interior* as in fact being *outside* of the black hole itself. Perhaps even more surprising is that it is gravity itself which instructs us to do this - it turns out that the island formula is not simply a postulate, but can in fact be derived from a semiclassical calculation involving so-called "replica wormholes" [13]. In future work the author would like to give some motivation for the island formula from the consideration of these replica wormholes. It would also be particularly interesting to consider the relevant aspects of holography that originally motivated the island formula, in particular the Ryu-Takayanagi conjecture [10] and its generalisation, the Engelhardt-Wall prescription [41].

Acknowledgments

I profusely thank Professors Tim Hollowood and Prem Kumar for their supervision. Their tutelage has provided an exciting first glimpse into the world of combining quantum mechanics with gravity, and I have learnt a great deal from our many discussions over the past year.

I would also like to thank Professors Adi Armoni and Carlos Núñez for the wealth of advice and support they have provided throughout the past four years. They have been excellent mentors, and have both consistently remained patient with my incessant questioning, for which I am extremely grateful.

A Definitions

Timelike, Spacelike and Null

• A tangent vector $X \in T_pM$ where M is some spacetime manifold is called "timelike", "spacelike", or "null" if $X^\mu g_{\mu\nu} X^\nu < 0$, $X^\mu g_{\mu\nu} X^\nu > 0$ or $X^\mu g_{\mu\nu} X^\nu = 0$ respectively, where g is the metric tensor
• A curve $\gamma : (a, b) \longrightarrow M$ where $(a, b) \subseteq \mathbb{R}$ and M is some spacetime manifold is called timelike, spacelike or null if all tangent vectors at every point of the curve are timelike, spacelike or null, respectively.

Cauchy Surface

• A Cauchy surface is any submanifold $S \subset M$ such that every timelike curve in M intersects S exactly once, where M is some spacetime manifold. An informal but intuitive picture is that a Cauchy surface is a "spacetime slice of constant time", namely the time coordinate does not vary from point to point on the surface.

B The Covariant Wave Equation

Here, we provide a derivation of the form of the covariant wave equation given in 2.1. Firstly, it is easy to see that $\nabla_\mu \nabla^\mu \phi$ reduces to the d'Alembertian operator $\Box = \partial_\mu \partial^\mu$ in flat space in Cartesian coordinates (where the Christoffel symbols vanish) and so $\nabla_\mu \nabla^\mu \phi = 0$ is indeed the massless Klein-Gordon equation. Using the definition of the Christoffel symbols, we have

$$\Gamma^i{}_{kl} = \frac{1}{2} g^{im} (\partial_l g_{mk} + \partial_k g_{ml} - \partial_m g_{kl}), \tag{B.1}$$

$$\implies \Gamma^i{}_{il} = \frac{1}{2} g^{ij} \partial_l g_{ji}. \tag{B.2}$$

It is well-known that this is equivalent to

$$\Gamma^i{}_{il} = \frac{1}{\sqrt{-g}} \partial_l \sqrt{-g}, \tag{B.3}$$

which can be proven using Jacobi's formula (see, for example, [42]). We then find

$$\begin{aligned} 0 = \nabla_\mu \nabla^\mu \phi &= \nabla_\mu \partial^\mu \phi \\ &= \partial_\mu \partial^\mu \phi + \Gamma^\mu_{\mu\nu} \partial^\nu \phi \\ &= \partial_\mu \partial^\mu \phi + \frac{1}{\sqrt{-g}} \partial_\nu \sqrt{-g} \partial^\nu \phi \\ &= \frac{1}{\sqrt{-g}} (\sqrt{-g} \partial_\mu \partial^\mu \phi + \partial_\nu \sqrt{-g} \partial^\nu \phi) \\ &= \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} \partial^\mu \phi), \end{aligned}$$

so finally we obtain 2.1,

$$\nabla_\mu \nabla^\mu \phi = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} \partial^\mu \phi) = 0. \tag{B.4}$$

C Stress Tensor in $w^\pm$ coordinates

Here, we prove that $T_{y^+y^+} = \frac{\pi c}{12\beta^2} \implies T_{w^+w^+} = 0$. The transformation rule for the stress-energy tensor from y^+ coordinates to w^+ coordinates is given by [12]

$$\left(\frac{\partial w^+}{\partial y^+}\right)^2 T_{w^+w^+} = T_{y^+y^+} + \frac{c}{24\pi} \{w^+, y^+\}, \tag{C.1}$$

where $\{w^+, y^+\}$ is the so-called "Schwarzian derivative" defined as

$$\{f, t\} := \frac{f'''(t)}{f'(t)} - \frac{3}{2} \left(\frac{f''(t)}{f(t)}\right)^2. \tag{C.2}$$

In the right-hand black hole, w^+ and y^+ are related as

$$w^+ = \exp\left(\frac{2\pi}{\beta} y^+\right), \tag{C.3}$$

and so we have

$$w^{+\prime} = \left(\frac{2\pi}{\beta}\right) \exp\left(\frac{2\pi}{\beta} y^+\right) \tag{C.4}$$

$$w^{+\prime\prime} = \left(\frac{2\pi}{\beta}\right)^2 \exp\left(\frac{2\pi}{\beta} y^+\right) \tag{C.5}$$

$$w^{+\prime\prime\prime} = \left(\frac{2\pi}{\beta}\right)^3 \exp\left(\frac{2\pi}{\beta} y^+\right). \tag{C.6}$$

Plugging these into C.2, we find

$$\begin{aligned}\{w^+, y^+\} &= \left(\frac{2\pi}{\beta}\right)^3 e^{\frac{2\pi}{\beta}y^+}\frac{\beta}{2\pi}e^{-\frac{2\pi}{\beta}y^+} - \frac{3}{2}\left(\left(\frac{2\pi}{\beta}\right)^2 e^{\frac{2\pi}{\beta}y^+}\frac{\beta}{2\pi}e^{-\frac{2\pi}{\beta}y^+}\right)^2 \\ &= \left(\frac{2\pi}{\beta}\right)^2 - \frac{3}{2}\left(\frac{2\pi}{\beta^2}\right) \\ &= -\frac{1}{2}\left(\frac{2\pi}{\beta^2}\right)^2 \end{aligned} \tag{C.7}$$

Finally, using C.1, we find

$$\begin{aligned}\left(\frac{2\pi}{\beta}e^{\frac{2\pi}{\beta}y^+}\right)^2 T_{w^+w^+} &= T_{y^+y^+} - \frac{c}{48\pi}\left(\frac{2\pi}{\beta^2}\right)^2 \\ &= \frac{\pi c}{12\beta^2} - \frac{4\pi^2 c}{48\pi\beta^2} \\ &= 0 \\ \implies T_{w^+w^+} &= 0, \end{aligned} \tag{C.8}$$

as we intended to show. The same analysis can be used to show that $T_{w^-w^-} = 0$.

References

[1] S.W. Hawking, "Particle creation by black holes", Communications in mathematical physics, Volume 43, p.199-220, (1975).

[2] S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse," Phys. Rev. D **14** (1976), 2460-2473 doi:10.1103/PhysRevD.14.2460

[3] "Black hole information bet", Available: http://theory.caltech.edu/ preskill/info_bet.html.

[4] S. W. Hawking, "Information loss in black holes," Phys. Rev. D **72** (2005), 084013 doi:10.1103/PhysRevD.72.084013 [arXiv:hep-th/0507171 [hep-th]].

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

[6] P. Hayden and J. Preskill, "Black holes as mirrors: Quantum information in random subsystems," JHEP **09** (2007), 120, doi:10.1088/1126-6708/2007/09/120, arXiv:0708.4025 [hep-th].

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

[8] L. Susskind, "Singularities, Firewalls, and Complementarity," [arXiv:1208.3445 [hep-th]].

[9] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2** (1998), 231-252, doi:10.1023/A:1026654312961, arXiv:hep-th/9711200 [hep-th].

[10] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT", Phys. Rev. Lett. **96** (2006), 181602 doi:10.1103/PhysRevLett.96.181602 [arXiv:hep-th/0603001 [hep-th]].

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

[12] A. Almheiri, R. Mahajan and J. Maldacena, "Islands outside the horizon," [arXiv:1910.11077 [hep-th]].

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

[14] P. Hein, "Grooks 1", Doubleday & Co, (1969).

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

[16] D. Page, "Information in Black Hole Radiation", Physical Review Letters. 71. 10.1103/PhysRevLett.71.3743, (1993).

[17] J. H. Traschen, "An Introduction to black hole evaporation," [arXiv:gr-qc/0010055 [gr-qc]].

[18] T.J. Hollowood, "Quantum Fields & Black Holes", Lecture Notes, (2020).

[19] F. Dowker, "Black Holes", Lecture notes, (2015).

[20] K. Blum, "Density Matrix Theory and Applications", 2nd ed, Plenum Press, (1996).

[21] M. Nielsen, I. Chuang, "Quantum Computation and Quantum Information", 10th Anniversary Edition, Cambridge University Press, (2010).

[22] F. Herbut, "Bipartite Entanglement Review of Subsystem-Basis Expansions and Correlation Operators in It", [arXiv:1410.4988 [math-ph]], (2014).

[23] J. S. Townsend, "A Modern Approach to Quantum Mechanics", 2nd ed, University Science Books, (2012).

[24] R. A. Bertlmann, "Theoretical Physics T2, Quantum Mechanics", Chapter 9, (2009). Lecture Notes - Available: https://homepage.univie.ac.at/reinhold.bertlmann/pdfs/T2_Skript_Ch_9corr.pdf

[25] C. H. Bennett, H. J. Bernstein, S. Popescu and B. Schumacher, "Concentrating partial entanglement by local operations", Phys. Rev. A **53** (1996), 2046-2052 doi:10.1103/PhysRevA.53.2046 [arXiv:quant-ph/9511030 [quant-ph]].

[26] M. B. Plenio and S. Virmani, "An Introduction to entanglement measures" Quant. Inf. Comput. **7** (2007), 1-51 [arXiv:quant-ph/0504163 [quant-ph]].

[27] G. Penington, "Entanglement Wedge Reconstruction and the Information Paradox," JHEP **09** (2020), 002 doi:10.1007/JHEP09(2020)002 [arXiv:1905.08255 [hep-th]].

[28] R. Jackiw, "Lower Dimensional Gravity," Nucl. Phys. B **252** (1985), 343-356 doi:10.1016/0550-3213(85)90448-1

[29] J. Engelsöy, T. G. Mertens and H. Verlinde, "An investigation of AdS_2 backreaction and holography," JHEP **07** (2016), 139 doi:10.1007/JHEP07(2016)139 [arXiv:1606.03438 [hep-th]].

[30] A. Almheiri and J. Polchinski, "Models of AdS_2 backreaction and holography," JHEP **11** (2015), 014 doi:10.1007/JHEP11(2015)014 [arXiv:1402.6334 [hep-th]].

[31] T. J. Hollowood and S. P. Kumar, "Islands and Page Curves for Evaporating Black Holes in JT Gravity," JHEP **08** (2020), 094 doi:10.1007/JHEP08(2020)094 [arXiv:2004.14944 [hep-th]].

[32] D. Tong, "General Relativity," Lecture Notes, (2019).

[33] T. J. Hollowood, S. Prem Kumar, A. Legramandi and N. Talwar, "Islands in the Stream of Hawking Radiation," [arXiv:2104.00052 [hep-th]], (2021).

[34] K. Hashimoto, N. Iizuka and Y. Matsuo, "Islands in Schwarzschild black holes," JHEP **06** (2020), 085 doi:10.1007/JHEP06(2020)085 [arXiv:2004.05863 [hep-th]].

[35] P. H. Ginsparg, "Applied Conformal Field Theory," [arXiv:hep-th/9108028 [hep-th]].

[36] P. Calabrese and J. L. Cardy, "Entanglement entropy and quantum field theory," J. Stat. Mech. **0406** (2004), P06002 doi:10.1088/1742-5468/2004/06/P06002 [arXiv:hep-th/0405152 [hep-th]].

[37] H. Casini and M. Huerta, "A Finite entanglement entropy and the c-theorem," Phys. Lett. B **600** (2004), 142-150 doi:10.1016/j.physletb.2004.08.072 [arXiv:hep-th/0405111 [hep-th]].

[38] S. M. Carroll, "Spacetime and Geometry", Addison Wesley, (2004).

[39] S. Chandrasekhar, "The mathematical theory of black holes", Clarendon Press, Reprinted Edition, (1998).

[40] L. V. Iliesiu, S. S. Pufu, H. Verlinde and Y. Wang, "An exact quantization of Jackiw-Teitelboim gravity," JHEP **11** (2019), 091 doi:10.1007/JHEP11(2019)091 [arXiv:1905.02726 [hep-th]].

[41] N. Engelhardt and A. C. Wall, "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime," JHEP **01** (2015), 073 doi:10.1007/JHEP01(2015)073 [arXiv:1408.3203 [hep-th]].

[42] S. Weinberg, "Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity", p.106-107, (1972).