The Black Hole Information Paradox

Emma Winkels

Supervisor: Sebastian de Haro

May 16, 2012

CAPSTONE THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE

AMSTERDAM UNIVERSITY COLLEGE

Abstract

We conduct a study on the information paradox and its possible resolutions. We show a short argument for black hole evaporation by looking at the particles that constitute the black hole's thermal atmosphere. We find that the black hole indeed evaporates due to the well-known Hawking radiation. We move on to a concise argument of the paradox by showing that entangled particle pairs created near the event horizon lead to the evolution from pure to mixed states, which is in violation with the basic ideas of quantum mechanics. We examine several proposals for solving the information paradox. Focus is given to the resolution that attempts to solve the paradox by taking into account the gravitational back-reaction. We find that this back-reaction leads to a correction in the outgoing radiation state. This correction can potentially become very large, but its magnitude depends critically on the cut-off of the interaction region. The Rubik's cube model for black hole evaporation is described, together with a few other ideas on resolving the paradox. We note that the most promising current approach to solving the information paradox is by use of the AdS/CFT correspondence.

Contents

1 Introduction

Since its discovery by Hawking in 1976 [1], the Black Hole Information Paradox (from now on called I.P.) has been widely discussed by physicists all over the world. Even now, three and a half decades later, no definite resolution to the paradox has been found; hence it remains a topic of interest. The key concept in Hawking's paradox is the evaporation of black holes. He discovered that black holes do not have an infinite lifetime; they slowly evaporate until they completely disappear. The unexpected thing is, however, that the information about the matter that made up the black hole (i.e. the matter that created it and that fell into it afterwards) does not 'naturally' come out in this evaporation process. If Hawking's derivation is correct, this information would thus be lost forever. Information conservation is explicitly formulated in both classical and quantum mechanics, and a violation hereof would thus be in contradiction with the basic foundations of physics. In classical mechanics, Liouville's theorem describes the conservation of information by stating that the phase space of a system is always conserved [1]. The loss of information in quantum mechanics would mean that pure states evolve into mixed states, which means the theory is non-unitary [1]. Moreover, loss of information would also imply a violation of energy conservation, since information transmission requires energy.

Since the loss of information is a concept that is inconsistent with a broad range of physical principles, it is a very important goal to understand how the I.P. can be resolved. In order to do this, it is necessary to fully understand Hawking's derivation of the paradox and look very carefully at the assumptions he made, and the computations he carried out. One way to resolve the paradox is to identify which of Hawking's assumptions are wrong. Another way is to accept Hawking's derivation as correct, and thus a 'new physics' would need to arise in black holes that we do not yet know of. The first method is the least controversial and is used most often in order to deal with the paradox. This will also be the one focused on in this research.

Many solutions to the problem have been proposed, but none of them enjoys a widespread acceptance as the ultimate explanation. Some recent examples of possible solutions to the paradox are: black hole remnants [1], gravitational back-reaction [2], fuzzballs [3], large density of gravity states [4], and the final state boundary condition [5]. These are only a few of the large number of proposed solutions since the discovery of the paradox, and there is still no accordance amongst physicists which one is (moving towards) the final answer. It is thus clear that a lot more research on this topic needs to be done in order to resolve it.

This paper is built up of two parts; in section two, we first look at the evaporation of black holes, since this process is at the core of the I.P. Subsequently, a simple argument for Hawking's paradox is presented in order to understand what the paradox is really about. In this section, a general discussion of the paradox and its implications is also included. Some recent ideas on how to resolve the paradox are discussed in detail in the third section.

2 About the Information Paradox

Before we start our discussion of the I.P., it is useful to first have a look at the black hole geometry such that we get a feeling of the arena we will be working in. The standard representation of this geometry is given by the Schwarzschild metric which, in spherical coordinates, is given by [1]:

$$ds^2 = -\left(1 - \frac{2MG}{r}\right) dt^2 + \left(1 - \frac{2MG}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{2.1}$$

The Schwarzschild metric has an event horizon at $g_{00} = 0$; as we can see from (2.1), this happens when $r = 2MG$. The Schwarzschild metric is a solution of Einstein's field equations in empty space, in fact, it is the unique spherically symmetric vacuum solution [6]. M is the mass of the gravitating object, which is of course the mass of the black hole in our case. We will not proof the Schwarzschild metric here, such a derivation can be found in textbooks like [6].

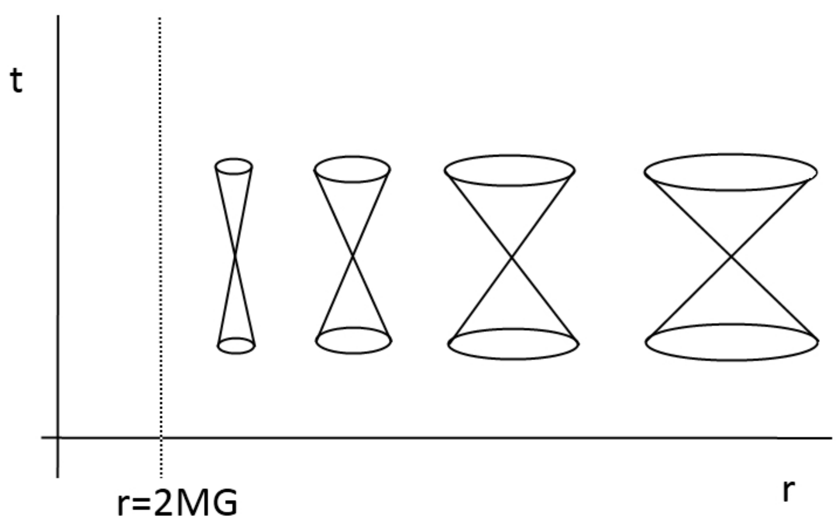

Figure 1: In the Schwarzschild geometry, light cones get narrower when approaching the event horizon at $r = 2MG$. At the horizon, the cones are folded up completely.

The behaviour of light cones in the Schwarzschild geometry is shown in Figure 1. As can be seen from the metric (2.1), radial null geodesics satisfy $\left(1 - \frac{2MG}{r}\right) dt^2 = \left(1 - \frac{2MG}{r}\right)^{-1} dr^2$, or $\frac{dt}{dr} = \pm\left(1 - \frac{2MG}{r}\right)^{-1}$. $\frac{dt}{dr}$ represents the slope of the light cones in Figure 1, since this depicts a (r, t)-diagram. As we come closer to the event horizon, $r \to 2MG$, and $\frac{dt}{dr} \to \pm\infty$. Hence, the slopes of the light cones get steeper as they approach the event horizon, which results in increasingly narrower cones (see Figure 1).

We can use Penrose diagrams to represent the causal structure of the Schwarzschild spacetime [1]. These diagrams are useful because they compactify the infinite spacetime such that it can be drawn entirely on the finite plane. A Penrose diagram for a Schwarzschild black hole is shown in Figure 2, where time runs upwards and space sideways. Lines tilted at angles of 45° correspond to light rays. In the diagram, remote distances are shrunk to aid the representation of distant spacetime. This results in the hyperbolic shape of constant-time and -space lines, as shown in Figure 2. Future and past time-like infinities are marked by $t = \pm\infty$, space-like infinity is marked by $r = \infty$, and future and past null-like infinity are labelled $\mathcal{I}^+$ and $\mathcal{I}^-$, respectively. The event horizon at $r = 2MG$ is also marked in the diagram in Figure 2. Once a particle or light-ray passes this horizon, it will inevitably hit the singularity, which is marked by a spacelike boundary in Figure 2.

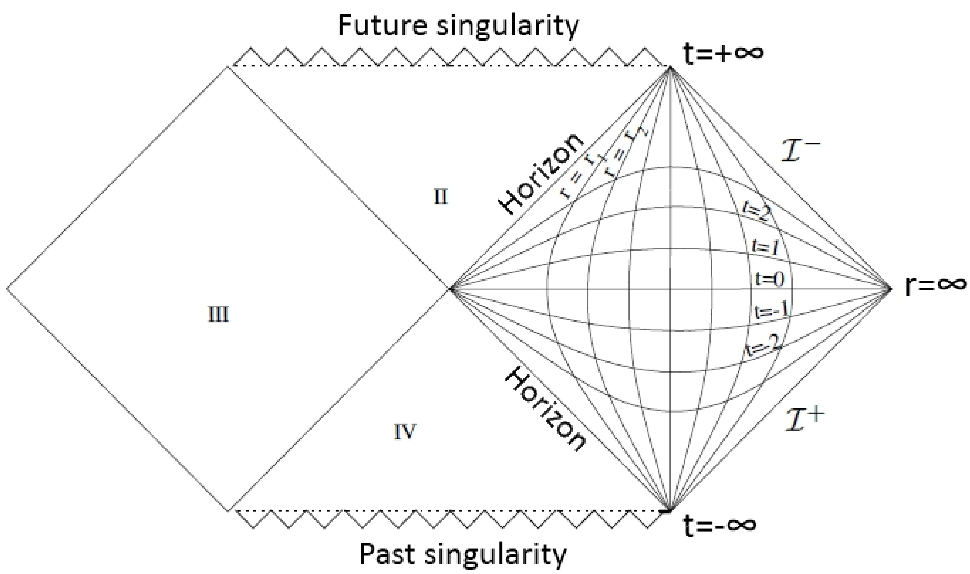

Figure 2: Taken from [1]. Penrose diagram for a Schwarzschild black hole. $\mathcal{I}^+$ is future null-like infinity, and $\mathcal{I}^-$ is past null-like infinity.

2.1 Hawking Radiation and Black Hole Evaporation

The I.P. originates from Hawking's discovery that black holes slowly evaporate, and thus do not have an infinite lifetime [1]. If all black holes indeed vanish at some point, the key question is what happens to the information that they contain. With 'information' we mean here the quantum states of all the matter that formed the black hole plus the matter that fell into it at later times. We will first show a short argument for black hole evaporation by looking at the particles that constitute the black hole's thermal atmosphere.

In this regime, we can use the Rindler metric to describe the black hole geometry, since we are close to the event horizon, i.e. $r \to 2MG$. We get the Rindler metric by taking this limit for r in the standard black hole metric given by (2.1). We will investigate the behaviour in the region close to the black hole's event horizon by replacing the standard Schwarzschild radial coordinate r in the Schwarzschild metric by the coordinate ρ, which measures the proper distance from the horizon [1]:

$$\begin{aligned} \rho &= \int_{2MG}^{r} \sqrt{g_{rr}\left(r'\right)} dr' = \int_{2MG}^{r} \left(1 - \frac{2MG}{r'}\right)^{-\frac{1}{2}} dr' \\ &= \sqrt{r\left(r - 2MG\right)} + 2MG \sinh^{-1}\left(\sqrt{\frac{r}{2MG} - 1}\right) \end{aligned} \tag{2.2}$$

Near the horizon we can use a simplified approximation of ρ by letting r go to $2MG$. In this case, the argument of the $arcsinh^{-1}$ is very small, and we know that $arcsinh^{-1}\left(\alpha\right) \approx \alpha$ for small α.

$$\begin{aligned} \rho &\approx \sqrt{r\left(r - 2MG\right)} + 2MG\sqrt{\frac{r}{2MG} - 1} \\ &= \sqrt{r\left(r - 2MG\right)} + \sqrt{r2MG - \left(2MG\right)^2} \end{aligned} \tag{2.3}$$

The first term on the r.h.s. of equation (2.3) can be written as:

$$\sqrt{r\left(r - 2MG\right)} = \sqrt{r}\sqrt{\left(r - 2MG\right)} \tag{2.4}$$

We can thus treat the $(r - 2MG)$ term as a small correction to $r = 2MG$ (since r approaches, but is not equal to, $2MG$). Expanding this term and neglecting all higher-order terms we get

$$\sqrt{r}\sqrt{\left(r - 2MG\right)} = \sqrt{2MG\left(r - 2MG\right)} + \mathcal{O}\left(\sqrt{r - 2MG}^2\right) \tag{2.5}$$

Which gives our approximate expression for ρ:

$$\rho \approx 2\sqrt{2MG\left(r - 2MG\right)} \tag{2.6}$$

We are interested in the part of the metric in (2.1) that describes the time evolution, i.e. the term:

$$-\left(1 - \frac{2MG}{r}\right) dt^2 \tag{2.7}$$

From the definition of ρ in (2.6) we see that we can write:

$$\rho^2 = 4\left(2MG\left(r - 2MG\right)\right) = 8MG\left(r - 2MG\right) = 8rMG - 16M^2G^2 \tag{2.8}$$

$$\frac{\rho^2}{\left(4MG\right)^2} = \frac{8rMG - 16M^2G^2}{16M^2G^2} = \frac{r}{2MG} - 1 \tag{2.9}$$

Now, we use the fact that r approaches $2MG$, but is not equal to $2MG$. Hence, we can write r in terms of a leading order term and a small correction: $r = 2MG + 2MGx$, where x is small. Now we see : $\frac{r}{2MG} - 1 = \frac{2MG(1+x)}{2MG} - 1 = 1 + x - 1 = x$. We want to compare this to the time evolution of the metric in equation (2.7), so we write: $\frac{2MG}{r} - 1 = \frac{2MG}{2MG+2MGx} - 1 = \frac{1}{1+x} - 1$. We know that $\frac{1}{1+x} \approx 1 - x$ for small x, so: $\frac{2MG}{r} - 1 \approx 1 - x - 1 = -x$. So we get:

$$\frac{\rho^2}{(4MG)^2} = \frac{r}{2MG} - 1 = -\left(\frac{2MG}{r} - 1\right) = \left(1 - \frac{2MG}{r}\right) \tag{2.10}$$

We can thus write the time evolution of the Schwarzschild metric as:

$$-\left(1 - \frac{2MG}{r}\right) dt^2 = -\frac{\rho^2}{(4MG)^2} dt^2 = -\rho^2 \left(\frac{dt}{4MG}\right)^2 \tag{2.11}$$

Since we substituted r with ρ, we see that we can define a new time coordinate as well: the Rindler time ω near the black hole is related to the regular Schwarzschild time t (for an outside observer) by [1]:

$$\omega = \frac{t}{4MG} \tag{2.12}$$

as follows from equation 2.11. Thus, a field (quantum state) near the horizon with frequency ν_R is seen as redshifted by the outside observer, since:

$$\begin{aligned} \nu &= 1/t \\ \nu_R &= \frac{1}{\omega} = \frac{4MG}{t} \\ \nu &= \frac{4MG}{t} \cdot \frac{1}{4MG} = \nu_R \frac{1}{4MG} \end{aligned} \tag{2.13}$$

This implies that the temperature T of the state seen by the outside observer is also redshifted compared to the Rindler temperature T_R of the state close to the horizon. Since $T_R = 1/2\pi$ [1], and $\nu \propto E \propto T$:

$$T = \frac{1}{2\pi} \cdot \frac{1}{4MG} = \frac{1}{8\pi MG} \tag{2.14}$$

The Rindler theory describes a potential barrier that specifies the amount of energy a particle needs to have to be able to escape from the black hole horizon at a certain point in space [1]. This potential is dependent on the angular momentum l of the particle and its distance from the singularity at the center of the black hole r [1]:

$$V_l = \frac{r - 2MG}{r} \left(\frac{l\,(l+1)}{r^2} + \frac{2MG}{r^3} \right) \tag{2.15}$$

If the angular momentum of the particle is zero and it is at a distance $r = 3GM$ from the singularity, the maximum height of this barrier is:

$$\begin{aligned} V_l &= \frac{3MG - 2MG}{3MG} \cdot \left(\frac{2MG}{(3MG)^3} \right) \\ &= \frac{6M^2G^2 - 4M^2G^2}{(3MG)^4} \\ &= \frac{2M^2G^2}{81M^4G^4} = \frac{1}{M^2G^2} \frac{2}{81} \end{aligned} \tag{2.16}$$

For any potential, a particle can pass over the barrier if its potential energy is larger than the potential energy of the barrier. We can model the particle as a harmonic oscillator, such that its potential energy is given by $V = \frac{1}{2}kx^2$. We also know that the spring constant k is proportional to ν^2, so we can roughly say that the particle will be able to pass over the potential barrier if its frequency squared is larger than the potential energy of the barrier. In the case of equation (2.16), the particle would thus need a frequency larger than

$$\nu = \left(\frac{1}{M^2G^2} \frac{2}{81} \right)^{\frac{1}{2}} = \frac{1}{MG} \sqrt{\frac{2}{81}} \tag{2.17}$$

to overcome the potential barrier. Since $\nu \propto E \propto T$, we can compare this value for the frequency to our average temperature calculated in equation (2.14). We see that $\frac{1}{8\pi} \approx 0.0398 < \sqrt{\frac{2}{81}} \approx 0.157$, but since this temperature is only an *average* temperature, and the difference between these two numbers is only a factor four; some particles will have a high enough temperature to escape the potential barrier. It is thus clear that some particles (with low angular momenta) of the black hole's thermal atmosphere can escape and the hole will hereby lose energy to its surroundings. This phenomenon is the well-known Hawking radiation.

We have described how a black hole loses energy and thus slowly evaporates. The problem is, however, that this radiation is strictly thermal, and thermal radiation does not carry information. If this thermal radiation is the only type of outgoing radiation from the hole, and there are no other processes we would need to take into account; all information inside the black hole would be lost. As described in the Introduction, this would have major consequences for many areas in physics, and it is thus important to investigate how this evaporation process exactly works.

2.2 A Simple Argument for the Information Paradox

Mathur gives a comprehensible argument leading to the I.P. in some of his papers of the last decade [3, 7]. The argument starts by naming a set of 'niceness conditions' which is an accepted set of conditions under which we must get so-called 'lab-physics'. The term 'lab-physics' refers to the exclusion of quantum gravity effects, i.e. the physics which is well-described by quantum fields on gently curved spacetime [3]. The complete set of 'niceness conditions' can be found in [3] and [7].

The second step in Mathur's argument consists of the construction of 'good slices' in the black hole geometry. This means that he looks at complete (extending from inside to outside the black hole) Cauchy surfaces on which to define quantum states. These 'good slices' satisfy all the 'niceness conditions' discussed above. Cauchy surfaces are space-like slices of the black hole spacetime geometry, i.e. these surfaces have a constant time outside the event horizon, and a constant r inside the black hole. The reason for this is that inside the black hole, time and space interchange roles. This happens because the first factor of the Schwarzschild metric (as seen in equation (2.1)) $\left(-1+\frac{2MG}{r}\right)$ changes sign as soon as the horizon is crossed (i.e. when we go from $r > 2MG$ to $r < 2MG$). These Cauchy surfaces arise from solving a differential equation in quantum field theory. The initial conditions on the surface determine the future and the past, of regions that are causally connected to it, in a unique manner [8]. At any point in the spacetime you can trace back along a null surface to the Cauchy surface: no point is disconnected. A Cauchy surface in a spacetime is thus intersected by each causal curve (causal curves are 'non-space-like', i.e. null or time-like) exactly once [8]. A quantum state in one instant in time is defined on one Cauchy surface. As we will do in section 3.1, Mathur looks at the evolution of quantum states from an initial to a final Cauchy surface, and shows that this evolution leads to a stretching of the initial slice, as we can see in Figure 3.

Figure 3 shows a schematic representation of the Swarzschild black hole. The left-most vertical line represents the singularity, whereas the right-most vertical line signifies the event horizon. Time runs upwards, and we thus see the time evolution of the initial infalling matter forming the black hole.

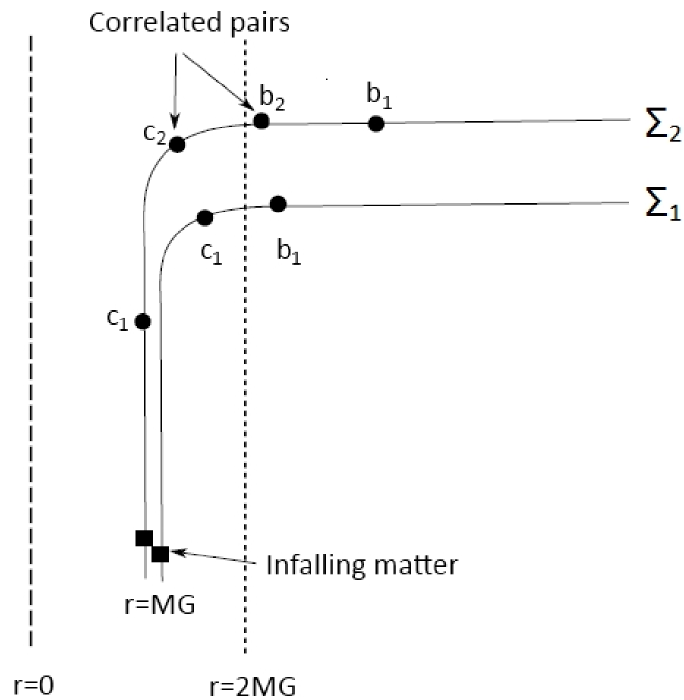

Figure 3: Adapted from [3]. Evolution of quantum states on an initial Cauchy surface, Σ_1, to a later Cauchy surface, Σ_2, leads to stretching of the initial spacelike slice. In this process, particle pairs (the c_i and b_i) are created.

Σ_1 and Σ_2 are two spacelike slices, i.e. Cauchy surfaces. It is clear from Figure 3 that Σ_2 is a later slice than Σ_1, since, outside the event horizon, Σ_2 ends further upwards (i.e. further in time) than Σ_1. Inside the black hole, however, the two slices have the same constant r, namely $r = M$. It is obvious that the slice Σ_2 is stretched compared to the slice Σ_1.

The next step is to look very closely at the evolution of a state from an earlier slice to a later slice, i.e. from Σ_1 to Σ_2 in Figure 3. The stretching of the Cauchy surface creates pairs of quanta on the slice [3], as is indicated in Figure 3. Some of these created particles can escape the black hole's gravitational pull and thus move away to infinity [7], i.e. this is the Hawking radiation discussed in section 2.1. As mentioned above, this radiation is crucial to the derivation of the I.P., thus the evolution of Cauchy surfaces is very important.

Let's call the state of the initial Cauchy surface (containing only the infalling matter M that creates the black hole) $|\psi\rangle_M$. As this surface evolves, correlated pairs are created as indicated in Figure 3. The quanta formed inside the event horizon are called c_i and the quanta outside of the horizon are called b_i. The complete state of the entire space-like slice is now given, to leading order, by [3, 7]:

$$|\Psi\rangle \approx |\psi\rangle_M \otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \tag{2.18}$$

where the numbers 0 and 1 represent the occupation number of the state and the term $\frac{1}{\sqrt{2}}$ is a normalization factor. We have taken a tensor product here between the initial matter state and the state of the created correlated pair of quanta because we are considering a leading order approximation. We assume that the particle pair is created far away form the infalling matter M such that the interaction between the two can be neglected. Small corrections to this state are of course possible, and an extensive discussion of these is given in [3, 7]. The total state in (2.18) is pure, but the constituent subsystems of the total state (i.e. the states of the created particles) are given by mixed states, i.e. we can describe them by a density matrix ρ [1]. We can compute the entanglement between the particles inside and outside of the event horizon by the Von Neumann entropy. This entropy is a measure of the departure from a pure state given by [1]:

$$S = -Tr\rho \ln \rho = \Sigma_j \rho_j \ln \rho_j \tag{2.19}$$

Our density matrix is given by $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$, where we represent the state of no particle creation $|0\rangle$ by the vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and the state $|1\rangle$ by a vector orthogonal to this one: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Thus we can write our density matrix as:

$$\rho = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.20}$$

We can easily compute the Von Neumann entropy using this expression for the density matrix:

$$\begin{aligned} S &= -Tr\rho \ln \rho \\ &= -Tr \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \ln\left(\frac{1}{2}\right) & 0 \\ 0 & \ln\left(\frac{1}{2}\right) \end{pmatrix} \\ &= -\left(\frac{1}{2}\ln\left(\frac{1}{2}\right) + \frac{1}{2}\ln\left(\frac{1}{2}\right)\right) \\ &= -\ln\left(\frac{1}{2}\right) \\ &= -\ln(1) + \ln(2) = \ln(2) \end{aligned} \tag{2.21}$$

At each step of the evolution of the Cauchy surface, a new particle pair is created, and the previously formed pairs move further away from each other (see Figure 3: b_2 and c_2 are the particles created in the following step). After N steps of evolution, our total state will look like [3, 7]:

$$
\begin{aligned}
|\Psi\rangle \approx |\psi\rangle_M &\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_1}|0\rangle_{b_1} + \frac{1}{\sqrt{2}}|1\rangle_{c_1}|1\rangle_{b_1}\right) \\
&\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_2}|0\rangle_{b_2} + \frac{1}{\sqrt{2}}|1\rangle_{c_2}|1\rangle_{b_2}\right) \\
&\dots \\
&\otimes \left(\frac{1}{\sqrt{2}}|0\rangle_{c_N}|0\rangle_{b_N} + \frac{1}{\sqrt{2}}|1\rangle_{c_N}|1\rangle_{b_N}\right)
\end{aligned} \tag{2.22}
$$

The entanglement entropy of this final state is thus $N \ln 2$, i.e. it increases with every new step in the evolution [3, 7]. This latter statement is what leads to Hawking's paradox. If the black hole indeed evaporates completely due to the Hawking radiation, the quanta b_i outside of the event horizon will be in an entangled state, but the states they are entangled with no longer exist. Thus, we can now only describe the final state in (2.22) by a density matrix [3, 7]. This means that the originally pure state $|\psi\rangle_M$ has evolved to a mixed state. As mentioned in the Introduction, the evolution from pure states to mixed states is in contradiction with basic quantum mechanics, since it implies that the theory is non-unitary.

The final step in Mathur's argument is one that he added to Hawking's original argument. This concluding step consists of the discussion of possible small corrections to the leading order evolution state (2.22) due to quantum gravity effects. Mathur concludes that these small corrections do not change the conclusion that the Von Neumann entropy increases with each step in the evolution (full details given in [3] and [7]).

3 Possible Solutions to the Information Paradox

3.1 Gravitational Back-Reaction

A solution to the I.P. was proposed in 1995 by three physicists, amongst which were the Verlinde twins and Y. Kiem. In their paper, *Black Hole Horizons and Complementarity*, they scrutinize Hawking's original derivation of the paradox, and add an additional factor: the gravitational back-reaction. This is the gravitational interaction between the matter falling into the black hole and the radiation that leaves the black hole. The main question is whether this gravitational back-reaction could indeed bring out the information about the matter that initially formed the black hole. In fact, this solution was based on previous studies in this area by other physicists, such as 't Hooft [2].

To investigate the effect of the back-reaction on the black hole evolution, we look at the propagation of a quantum state from one Cauchy surface to another. Cauchy surfaces are space-like slices of the black hole spacetime geometry, i.e. these surfaces have a constant time outside the event horizon, and a constant r inside the black hole. A quantum state in one instant in time (e.g. at time $t = 0$) is defined on one Cauchy surface. If the state evolves through time it is defined on a different Cauchy surface at time $t = 1$, and yet another Cauchy surface at time $t = 2$, etc (see Figure 4). $\Sigma_{initial}$ is the initial Cauchy surface at time $t = 0$ with a location somewhere outside the event horizon of the black hole. Σ_{final} is the final Cauchy surface at some later time, at a location closer to, but still outside of, the event horizon. As is clear from Figure 4, two types of gravitational interactions can occur in this scheme. The interaction indicated in red is between an outgoing wave (from the final to the initial surface) and an infalling wave (from the initial to the final surface), while the interaction indicated in blue is between two outgoing waves. We will show that both these types of interactions become increasingly more important in the evaporation process as time progresses on Σ_{final}.

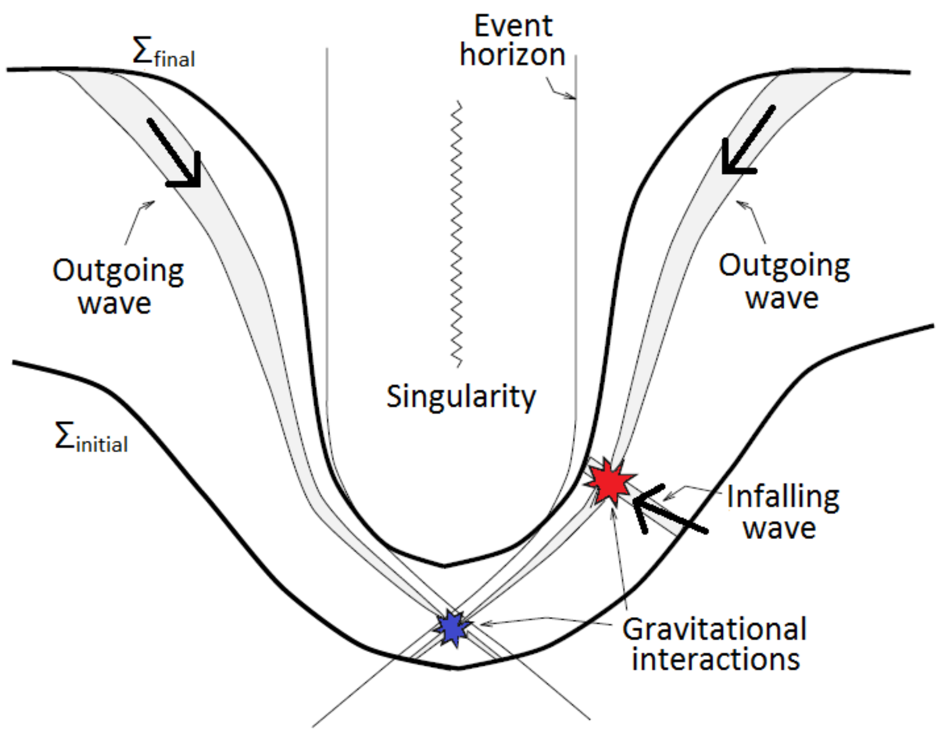

Figure 4: Adapted from [2]. Gravitational interactions occur between outgoing (from final to initial Cauchy surface) and infalling (from initial to final Cauchy surface) waves, and between outgoing waves themselves.

One of Hawking's assumptions in is derivation of the I.P. was that the radiation state (the infalling or outgoing wave) is well described by free field propagation on a fixed black hole background. Thus, we start with a concise review of free field propagation on the black hole geometry. Outside the black hole, this is given by the Schwarzschild metric [1]:

$$ds^2 = -\left(1 - \frac{2MG}{r}\right) dt^2 + \left(1 - \frac{2MG}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{3.1}$$

with $d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2$. It is convenient to write this metric using the Kruskal coordinates $x^+ = e^{v/4M}$ and $x^- = -e^{-u/4M}$, where $v = t + r^*$ is the advanced time coordinate, $u = t - r^*$ is the retarded time coordinate, and $r^* = r + 2M\ln\left(\frac{r}{2M} - 1\right) - 2M$ [2]. r^* is usually written as $r + 2M\ln\left(\frac{r}{2M} - 1\right)$ and is called the Tortoise coordinate [1]. It is chosen in such a way that the light cones no longer fold up as they approach the event horizon, like they do in the normal Schwarzschild metric as seen in Figure 1. We can rewrite the Schwarzschild metric in terms of the Tortoise coordinate by taking $\frac{dr^*}{dr} = \frac{-r}{2MG-1} = \frac{1}{1-\frac{2MG}{r}}$, such that we can rewrite dr^2 as $(dr^*)^2\left(1 - \frac{2MG}{r}\right)^2$. The metric in terms of the Tortoise coordinate then looks like $ds^2 = \left(1 - \frac{2MG}{r}\right)\left[-dt^2 + (dr^*)^2\right] + r^2 d\Omega^2$, thus radial null geodesics satisfy $dt^2 = (dr^*)^2$, or $\frac{dt}{dr^*} = \pm 1$. Therefore, in a (r^*, t) diagram, the light cones have constant slopes, i.e. they do not become narrower when approaching the event horizon. We can write the time coordinate in terms of the Tortoise coordinate and u, v to arrive at $t = v - r^* = u + r^*$, such that $\frac{dt}{dr^*} = v - 1 = u + 1$. This leads us to conclude that infalling radial null geodesics $\left(\frac{dt}{dr^*} = 1\right)$ are characterised by $v =$ constant and outgoing radial null geodesics by $u =$ constant. The behaviour of the light cones in a (v, r) diagram is shown in Figure 5.

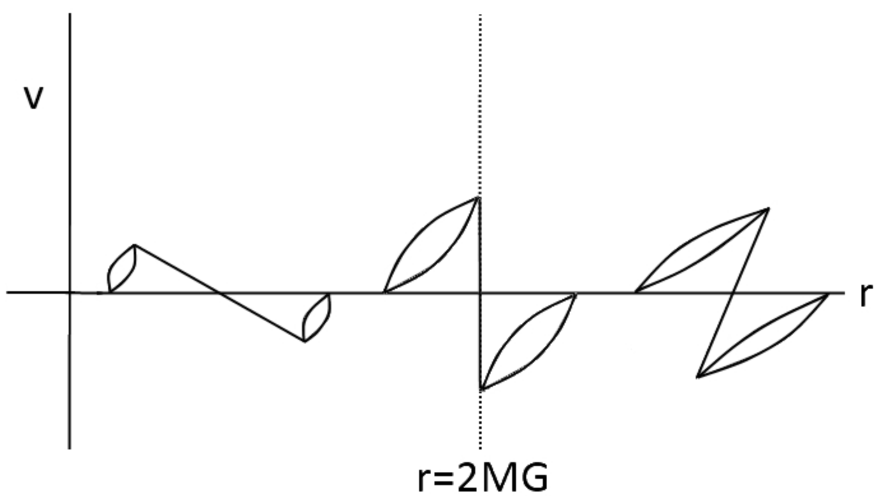

Figure 5: Behaviour of light cones in the u, v, r, Ω coordinate system. The light cones do not fold up at $r = 2MG$ but tilt over, leading to the fact that, behind the event horizon, only movement towards the singularity is allowed.

The new coordinate system x^+, x^-, r, Ω gives us the metric:

$$ds^2 = -\frac{32M^3}{r} e^{-r/2M} dx^+ dx^- + r^2 d\Omega^2 \tag{3.2}$$

We can see that this new metric is indeed the same as the original Schwarzschild metric in (3.1) by computing the derivative of x^+ with respect to v and of x^- with respect to u.

$$dx^+ = \frac{1}{4M} e^{v/4M} dv \tag{3.3}$$

$$dx^- = -\frac{-1}{4M} e^{-\frac{u}{4M}} du = \frac{1}{4M} e^{-\frac{u}{4M}} du \tag{3.4}$$

We then multiply equations (3.3) and (3.4) as is done in the metric (3.2):

$$dx^+ dx^- = \frac{1}{16M^2} e^{v-u/4M} dvdu \tag{3.5}$$

To work out this result, we need to fill in the definitions for u, v, dv, and du. First we calculate the exponential term:

$$v - u = t + r^* - t + r^* = 2r^* = 2r + 4M \ln\left(\frac{r}{2M} - 1\right) - 4M \tag{3.6}$$

$$\begin{aligned} e^{\frac{v-u}{4M}} &= e^{\frac{2r^*}{4M}} = e^{(2r+4M\ln(r/2M-1)-4M)/4M} = e^{\frac{r}{2M}+\ln\left(\frac{r}{2M}-1\right)-1} \\ &= e^{\frac{r}{2M}-1} e^{\ln\left(\frac{r}{2M}-1\right)} = \left(\frac{r}{2M} - 1\right) e^{\frac{r}{2M}-1} \end{aligned} \tag{3.7}$$

Subsequently, we calculate $dvdu$:

$$\begin{aligned} dv = dt + dr^*, du = dt - dr^* \\ dvdu = (dt)^2 - (dr^*)^2 \end{aligned} \tag{3.8}$$

Which leaves us with the following expression for dx^+dx^-:

$$dx^+ dx^- = \frac{1}{16M^2} \left(\frac{r}{2M} - 1\right) e^{\frac{r}{2M}-1} \left((dt)^2 - (dr^*)^2\right) \tag{3.9}$$

The next step is to compute dr^*, square it, and plug it into the definition in (3.9).

$$
\begin{aligned}
dr^* = \frac{dr^*}{dr}dr &= \left(1 + 2M\frac{1}{\frac{r}{2M}-1}\frac{1}{2M}\right)dr \\
&= dr + \frac{dr}{\frac{r}{2M}-1} = \frac{\left(\frac{r}{2m}-1+1\right)}{\frac{r}{2M}-1}dr \\
&= \frac{\frac{r}{2m}}{\frac{r}{2M}-1}dr
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
(dr^*)^2 &= \frac{\left(\frac{r}{2m}\right)^2}{\left(\frac{r}{2M}-1\right)^2}dr^2 = \frac{\left(\frac{r}{2m}\right)^2}{\frac{r^2}{4M^2}-\frac{2r}{2M}+1}dr^2 \\
&= \frac{1}{1-\frac{4M}{r}+\frac{4M^2}{r^2}}dr^2 = \frac{1}{\left(1-\frac{2M}{r}\right)^2}dr^2
\end{aligned}
\tag{3.11}
$$

This gives us our final definition for dx^+dx^-:

$$
dx^+dx^- = \frac{1}{16M^2}\left(\frac{r}{2M}-1\right)e^{\frac{r}{2M}-1}\left((dt)^2 - \frac{1}{\left(1-\frac{2M}{r}\right)^2}dr^2\right)
\tag{3.12}
$$

We can now plug equation (3.12) into the metric in (3.2) to get the original Schwarzschild metric back.

$$
\begin{aligned}
ds^2 &= -\frac{32M^3}{r}e^{-\frac{r}{2M}}\frac{1}{16M^2}\left(\frac{r}{2M}-1\right)e^{\frac{r}{2M}-1}\left((dt)^2 - \frac{1}{\left(1-\frac{2M}{r}\right)^2}dr^2\right) + r^2 d\Omega^2 \\
&= -\frac{2M}{r}\left(\frac{r}{2M}-1\right)\frac{1}{e}\left((dt)^2 - \frac{1}{\left(1-\frac{2M}{r}\right)^2}dr^2\right) + r^2 d\Omega^2 \\
&= -\left(1-\frac{2M}{r}\right)\frac{1}{e}\left((dt)^2 - \frac{1}{\left(1-\frac{2M}{r}\right)^2}dr^2\right) + r^2 d\Omega^2 \\
&= -\left(1-\frac{2M}{r}\right)\frac{1}{e}dt^2 + \left(1-\frac{2M}{r}\right)^{-1}\frac{1}{e}dr^2 + r^2 d\Omega^2
\end{aligned}
\tag{3.13}
$$

When comparing the metric in (3.13) to the Schwarzschild metric in (3.1), it is clear that the factor $1/e$ in (3.13) does not correspond to the metric in (3.1), and that the gravitational constant is missing from (3.13). The gravitational constant was left out because it is irrelevant to the computations done and thus it is set to 1, just like c is set to 1. We will continue to set G to 1 from now on. The $1/e$ factor would be solved if the factor $e^{v-u/4M}$ in equation (3.5) would be $e^{v-u+4M/4M}$. This could be done by changing the definition of x^+ and x^- to $x^+ = e^{v+2M/4M}$ and $x^- = -e^{-u+2M/4M}$. Another way in which this problem could be solved is by redefining r^* as follows: $r^* = r + 2M\ln\left(\frac{r}{2M}-1\right)$.

In fact, this latter definition of r^* is used in many other papers and books on black holes (see, for example, [1, 8, 9]).

Now that we have the appropriate metric for the region outside the black hole, we can start to look at free field propagation in this geometry. The Klein-Gordon equation for the field ϕ close to the horizon has the standard from [2]:

$$[\partial_u \partial_v - e^{\frac{v-u}{4M}} \left(-\frac{\Delta_\Omega}{r^2} + m^2 + \frac{2M}{r^3} \right)] r\phi \left(u, v, \Omega\right) = 0 \tag{3.14}$$

where $\Omega = (\theta, \phi)$. This equation takes on a very simple form in the region where r goes to $2M$, since the $e^{\frac{v-u}{4M}}$ term will go to zero in this region.

$$v - u = t + r^* - t + r^* = 2r^* = 2r + 4M \ln \left(\frac{r}{2M} - 1 \right) - 4M \tag{3.15}$$

As $r \to 2M, 2r \to 4M$, so in this limit the first and last term on the r.h.s. of the equation will cancel each other out. Also, $r/2M \to 1, r/2M - 1 \to 0$, and the natural logarithm of this will reach increasingly negative values as r comes closer to 2M. The $e^{\frac{v-u}{4M}}$ term will thus approach zero in this limit, and the Klein-Gordon equation can be written as:

$$\partial_u \partial_v \phi \left(u, v, \Omega\right) = 0 \tag{3.16}$$

where r is omitted since it's considered a constant. This property of the field in the region $r \to 2M$ means that the total field in this region can be decomposed in terms of the incoming and outgoing fields [2]. This can be shown by integrating (3.16) twice; first with respect to u, and then with respect to v:

$$\int du \frac{\partial}{\partial u} \frac{\partial \phi}{\partial v} = 0 \tag{3.17}$$

where $\frac{\partial \phi}{\partial v}$ is a constant with respect to u, but not with respect to v, thus we write this term in its most general form as $\frac{\partial \phi}{\partial v} = C_1 \left(v\right)$. Of course, we can do the same for the integral with respect to v to arrive at $\frac{\partial \phi}{\partial u} = C_2 \left(u\right)$. Thus:

$$\phi \left(u, v\right) = \int dv \cdot C_1 \left(v\right) + \int du \cdot C_2 \left(u\right) = f \left(v\right) + g \left(u\right) \tag{3.18}$$

Now, since $v = t + r$ is the advanced time coordinate and $u = t - r$ is the retarded time coordinate, we define the functions $f \left(v\right)$ and $g \left(u\right)$ to be the incoming and outgoing field, respectively:

$$\phi(u,v,\Omega) = \phi_{in}(v,\Omega) + \phi_{out}(u,\Omega) \tag{3.19}$$

In fact, the incoming field can be written in terms of the outgoing field as:

$$\phi_{in}(v,\Omega) = \phi_{out}\left(u(v), {}^{P}\Omega\right) \tag{3.20}$$

where $\Omega \equiv (\theta,\phi)$ and ${}^{P}\Omega$ is the antipodal point on the two-sphere, i.e.: ${}^{P}\Omega = (\pi-\theta, \pi+\phi)$. The antipodal point of a point on a sphere is the point that is diametrically opposite to it. In Cartesian coordinates, this would thus mean that $x \to -x, y \to -y$, and $z \to -z$. In spherical coordinates, we write this as $x \to -x = -r\cos(\phi)\sin(\theta), y \to -y = -r\sin(\phi)\sin(\theta)$, and $z \to -z = -r\cos(\theta)$. This transformation can be accomplished by letting $\phi \to \pi+\phi$ and $\theta \to \pi-\theta$, since $\cos(\pi+\phi) = \cos(\pi-\phi) = -\cos(\phi)$ and $\sin(\pi-\phi) = \sin(\phi)$, $\sin(\pi+\phi) = -\sin(\phi)$.

The incoming field can be written like (3.20) because the in- and out-signals have to match in the region where the horizon was formed in the beginning of gravitational collapse [2]. The reparametrization is given by [2]:

$$u(v) = v_0 - 4M\ln\left(\frac{v_0 - v}{4M}\right) + const. \tag{3.21}$$

The constant is of order M, and v_0 is the time at which the event horizon of the black hole is created.

The reparametrization in equation (3.21) comes from a reflection boundary condition on the surface of the black hole. To understand where this comes from, we can look at an analogous situation involving a moving mirror in a flat, two-dimensional spacetime (i.e. 2D Minkowski space). In 2D, this mirror is thus represented by a point. Let's consider our reflecting point to move along the path $x = z(t)$, where $z(t) = 0$ for $t < 0$ [9], as shown in Figure 6. A scalar field ϕ in this representation satisfies the Klein-Gordon equation:

$$\left(\Box + m^2\right)\phi = 0 \tag{3.22}$$

which has a standard set of solutions for a field ϕ dependent on x and t [9]:

$$u_k(t,x) \propto e^{ik\cdot x - i\omega t} \tag{3.23}$$

where

$$\omega \equiv \sqrt{(k^2 + m^2)} \tag{3.24}$$

$$k \equiv |k| = \sqrt{\left(\sum_{i=1}^{n-1} k_i^2\right)} \tag{3.25}$$

Here, n represents the number of dimensions of the Minkowski space, which equals 2 in our case. The orthonormal solution to equation (3.22) is [9]

$$u_k(t, x) = [2\omega (2\pi)^{n-1}]^{-\frac{1}{2}} e^{ik\cdot x - i\omega t} \tag{3.26}$$

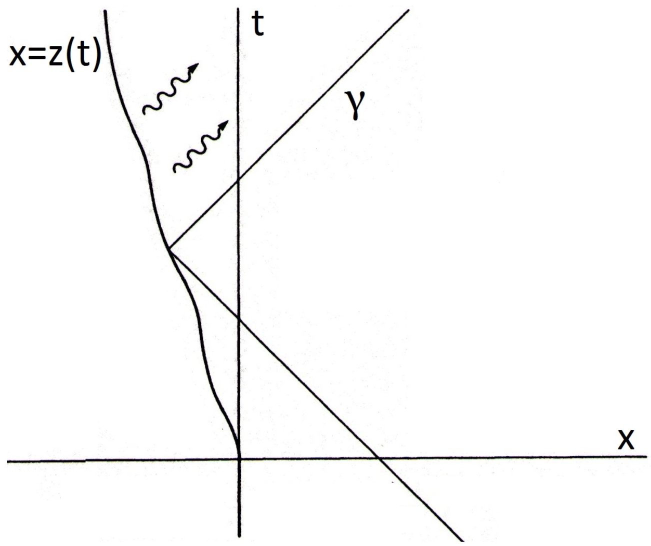

Figure 6: Taken from [9]. Reflection of a wave γ from a mirror at $x = z(t)$. The mirror radiates due to its movement.

In our case, we consider a massless scalar field, so the field equation in (3.22) reduces to

$$\Box\phi = \frac{\partial^2 \phi}{\partial u \partial v} = 0 \tag{3.27}$$

where the u and v coordinates are given, as before, by $u = t - x$ and $v = t + x$. We are in 2D, so all $(n - 1)$ terms in equation (3.26) equal 1 in our case. The definition of ω in equation (3.24) just reduces to $\omega = |k|$ since $m = 0$. Also the definition of k in (3.25) takes on a simpler form, since the sum is replaced by just the value k_1: $k \equiv |k| = \sqrt{k_1^2} = |k_1| = |k|$, so $\omega = |k| = |k|$. Thus, $e^{ik\cdot x - i\omega t} = e^{i|k|(x-t)} = e^{i\omega(x-t)}$. The most general solution to the field equation in (3.27) is then given by:

$$\phi = i\left(4\omega\pi\right)^{-\frac{1}{2}}\left(e^{-i\omega f(v)} + e^{i\omega g(u)} + e^{i\omega h(v)} + e^{-i\omega j(u)}\right) \tag{3.28}$$

where $f(v), g(u), h(v)$, and $j(u)$ are positive functions. Equation (3.28) is the analogue of equation (3.19), since we defined v to be the advanced time representing the incoming wave and u to be the retarded time representing the outgoing wave. It is clear that we can immediately throw away the last two terms of equation (3.28) because the incoming wave is only left-moving, and the outgoing wave is only right-moving. Since the first term on the r.h.s. of equation (3.28) should represent the incoming wave, and this wave has no further complications, $f(v)$ is linear in v, i.e. $f(v) = v$. The second term should represent the outgoing wave, which carries a Doppler shift due to the motion of the mirror. This term will thus have a more complicated u-dependence, therefore we need to establish the precise form of $g(u)$.

The massless field ϕ has the following reflection boundary condition:

$$\phi\left(t, z(t)\right) = 0 \tag{3.29}$$

We should make sure that the solutions in (3.28) indeed satisfy the boundary condition (3.29). We write (taking a minus sign for simplicity):

$$\begin{aligned}
&\phi = i\left(4\omega\pi\right)^{-\frac{1}{2}}\left(e^{-i\omega v} - e^{i\omega g(u)}\right) = 0 && (3.30)\\
&\Rightarrow \left(e^{-i\omega v} - e^{i\omega g(u)}\right) = 0 &&\\
&\Rightarrow e^{-i\omega v} = e^{i\omega g(u)} &&\\
&\Rightarrow v = -g(u) && (3.31)
\end{aligned}$$

The boundary condition tells us that the solution should be zero at time t when $x = z(t)$. Also, at the boundary $x = z(t)$, we have to match the in- and out-going waves, i.e. at the specific time and location of reflection (which we call $t_u$and $z(t_u)$, respectively), the in-going wave is equal to the out-going wave. This give us the following values for the u and v coordinates:

$$u = t_u - z(t_u), v = t_u + z(t_u) \tag{3.32}$$

Which leads us to the condition for $g(u)$:

$$v = -g(u) = t_u + z(t_u) \tag{3.33}$$

$$g(u) = -t_u - z(t_u) = -2t_u + u \tag{3.34}$$

By filling this definition for $g(u)$ into equation (3.30), we get our final set of solutions [9]:

$$u_k(u, v) = i(4\pi\omega)^{-\frac{1}{2}} \left(e^{-i\omega v} - e^{-i\omega(2t_u - u)}\right) \tag{3.35}$$

The $e^{-i\omega v}$ term characterizes the incoming wave (from $\mathcal{I}^-$), which is represented by the advanced time coordinate v and moves to the left (represented by the minus sign). The $e^{i\omega u}$ term characterizes the outgoing wave moving to the right (to $\mathcal{I}^+$), represented by the retarded time u. The factor t_u is defined as $u + z(t_u)$ and represents the Doppler shift of the reflection due to the motion of the mirror. We can check that this extra factor indeed represents a Doppler shift by writing the factor in the exponent of the second term on the r.h.s. of equation (3.35) as $-i\omega(2t_u - u) = -i\omega(2u + 2z(t_u) - u) = -i\omega(2z(t_u) + u)$. We take an example trajectory given by $z(t_u) = vt$, which gives us the term $-i\omega(2vt + u) = -i\omega(2vt + t - x) = -i\omega t(2v + 1) + i\omega x$. From this, we conclude that the frequency is indeed Doppler shifted, and is given by: $\omega' = \omega\left(1 + \frac{2v}{c}\right)$, which is exactly the standard Doppler shift formula (with $v_{source} - v_{receiver} = 2v$). Depending on the direction of v, ω is either redshifted or blueshifted.

Now that we understand the mirror analogy, we can translate what we have done there to the black hole scenario. We will look at the situation where a spherically symmetric ball of matter collapses and forms a black hole (see Figure 7). Again, we will do this in two dimensions since the full solutions of the wave equation (3.27) in this background metric are very complicated [9].

As we have seen above, the metric of the black hole geometry can be rewritten into terms of u and v. We will define our coordinates u and v slightly different this time by defining our spatial coordinate as $r^*_{new} = r^*_{old} - R^*_0$ [9]:

$$v = t + r^*_{old} - R^*_0, u = t - r^*_{old} + R^*_0 \tag{3.36}$$

where r^*_{old} is defined as above for the Schwarzschild metric ($r^*_{old} = r + 2M \ln\left(\frac{r}{2M} - 1\right)$, where we have left out the factor $-2M$ as discussed on page 16) and R^*_0 is a constant. Inside the ball of matter, we define our coordinates as [9]:

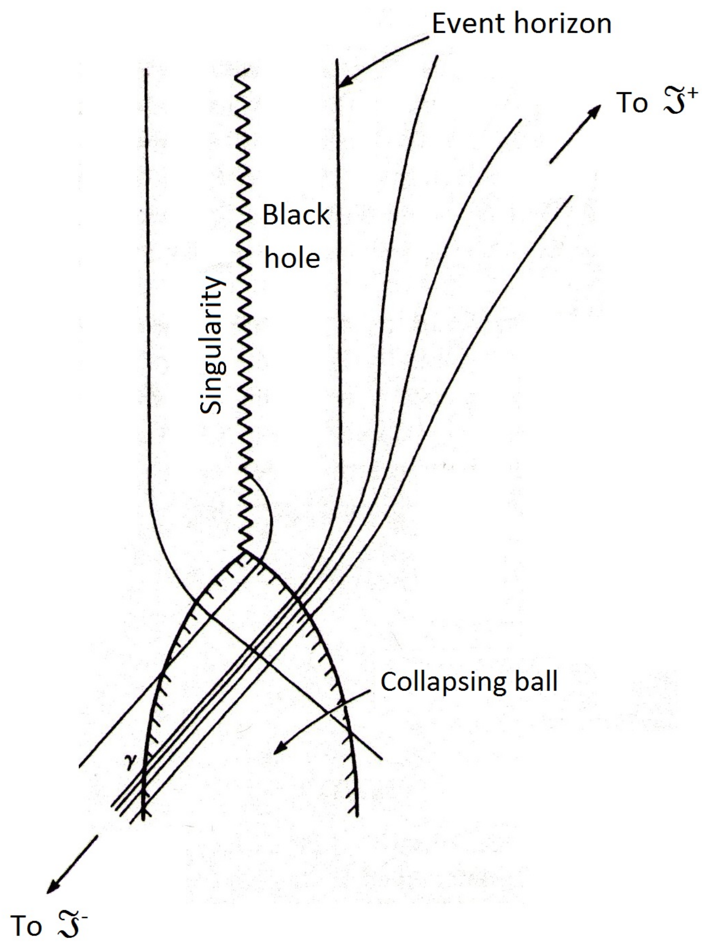

Figure 7: Taken from [9]. A collapsing ball of matter forms a black hole. Null rays coming from $\mathcal{I}^-$ are distorted. One such ray labelled γ forms the event horizon.

$$V = \tau + r - R_0, U = \tau - r + R_0 \tag{3.37}$$

where r is the normal Schwarzschild coordinate, and the relation between R_0 and R_0^* is the same as between r and r^*. R_0 is the radius of the ball of matter at time $\tau = 0$, i.e. before it starts to collapse. Like in equation (3.36), we have redefined our spatial coordinate to $r_{new} = r_{old} - R_0$. The surface shrinks along the trajectory $r = R(\tau)$. Notice that at $\tau = 0 = t$, the coordinates of the exterior and interior of the ball at its surface (i.e. $r = R_0$ and $r^* = R_0^*$) match: $u = U = v = V = 0$. To make the coming argument in the most general from, we will write the metric in the exterior region as:

$$ds^2 = C\left(r\right) dvdu \tag{3.38}$$

For the Schwarzschild metric, we have $dvdu = (dt)^2 - (dr^*)^2$, where $(dr^*)^2 = \left(1 - \frac{2M}{r}\right)^{-2} dr^2$, from equations (3.8) and (3.11). This leads us to the conclusion that $C(r) = -\left(1 - \frac{2M}{r}\right)$ for the Schwarzschild metric (the Ω term is left out because we are working in two dimensions). For the metric in the interior region we write a similar expression:

$$ds^2 = A(U, V)\, dU dV \tag{3.39}$$

In this general scenario, we must also generalize our definition of r^*:

$$r^* = \int C(r)^{-1}\, dr \tag{3.40}$$

To model the full four-dimensional symmetry of the situation at hand, we can restrict the treatment to the region $r \geq 0$, and reflect null rays at $r = 0$. This reproduces the effect of radially incoming rays propagating through the centre of the ball and out again [9]. Since we have just treated the moving mirror example, we know that we can impose such a reflection by demanding that the field ϕ is zero at $r = 0$.

We write the relation between the coordinates of the exterior and interior region as follows:

$$U = \alpha(u),\, v = \beta(V) \tag{3.41}$$

The exact shape of these functions is not important for understanding the reparametrization of equation (3.21). Analogous to equation (3.35), we get the following mode solutions:

$$u_k(u, v) = i\,(4\pi\omega)^{-\frac{1}{2}} \left(e^{-i\omega v} - e^{-i\omega\beta(\alpha(u) - 2R_0)}\right) \tag{3.42}$$

It is clear that the only difference between equations (3.35) and (3.42) is that the factor $2t_u - u$ is replaced by $\beta(\alpha(u) - 2R_0)$. The factor in equation (3.35) is the advanced time v evaluated at the reflecting boundary $x = z(t_u)$, just as the factor in equation (3.42) is the advanced time v evaluated at the reflecting boundary $r = 0$:

$$\begin{aligned} v &= \beta(V) = \beta(U + 2r - 2R_0) = \beta(\alpha(u) + 2r - 2R_0) \\ &= \beta(\alpha(u) - 2R_0) \end{aligned} \tag{3.43}$$

where we have set r to zero in the last step because we are looking at the boundary.

Along the line of collapse, $r = R(\tau)$, the two metrics given by equations (3.38) and (3.39) have to be equal. This leads us to the following expressions for $r = R(\tau)$ [9]:

$$\alpha'(u) = \frac{dU}{du} = \left(1 - \dot{R}\right) C\{[C\left(1 - \dot{R}^2\right) + \dot{R}^2]^{\frac{1}{2}} - \dot{R}\}^{-1} \tag{3.44}$$

$$\beta'(u) = \frac{dv}{dV} = C^{-1}\left(1 + \dot{R}\right)^{-1}\{[C\left(1 - \dot{R}^2\right) + \dot{R}^2]^{\frac{1}{2}} + \dot{R}\} \tag{3.45}$$

where $\dot{R} = \frac{dR}{d\tau}$. As the surface of the collapsing ball comes closer to the event horizon, i.e. C approaches 0, we can neglect the $C\left(1 - \dot{R}^2\right)$ term in the denominator of equation (3.44), so the expression simplifies to:

$$\begin{aligned}\alpha'(u) = \frac{dU}{du} &\approx \left(1 - \dot{R}\right) C\{[\dot{R}^2]^{\frac{1}{2}} - \dot{R}\}^{-1} \\ &= \left(1 - \dot{R}\right) C\left(-\dot{R} - \dot{R}\right)^{-1} \\ &= \frac{\left(1 - \dot{R}\right) C}{-2\dot{R}} = \frac{\left(\dot{R} - 1\right) C}{2\dot{R}}\end{aligned} \tag{3.46}$$

where we have used $[\dot{R}^2]^{\frac{1}{2}} = -\dot{R}$ because we have a collapsing surface. Since we are in the limit where C approaches 0, we may expand $R(\tau)$ as follows [9]:

$$R(\tau) = R_h + \nu(\tau_h - \tau) + \mathcal{O}\left((\tau_h - \tau)^2\right) + \ldots \tag{3.47}$$

Here, R_h is equal to R at the horizon, and τ is equal to τ_h when $R(\tau) = R_h$, and $\nu = -\dot{R}(\tau_h)$. We can then rewrite equation (3.46) to get, up to linear order since we are using the near-horizon limit [9]:

$$\begin{aligned}\frac{dU}{du} &\approx \frac{\left(\dot{R} - 1\right) C(R)}{2\dot{R}} = \frac{-\nu - 1}{-2\nu} C(R) \\ &= \frac{\nu + 1}{2\nu} C(R) = \frac{\nu + 1}{2\nu} \nu(\tau_h - \tau) \dot{C}|_{R_h} \\ &\approx \frac{(\tau_h - \tau)\dot{C}|_{R_h}}{2} = \kappa(\tau_h - \tau)\end{aligned} \tag{3.48}$$

where $\kappa = \left(\frac{1}{2}\frac{\partial C}{\partial r}\right)|_{R_h}$. In computing equation (3.48) we have used $C = 0 + (R - R_h)\dot{R}|_{R_h} = \nu(\tau_h - \tau)\dot{C}|_{R_h}$. By definition, $U = \tau - r + R_0 = \tau - R_h + R_0$, which means we can write $\tau_h - \tau = -(\tau - \tau_h) = -(U + R_h - R_0 - \tau_h)$, such that:

$$\frac{dU}{du} = -\kappa \left(U + R_h - R_0 - \tau_h\right) \tag{3.49}$$

Integrating equation (3.49) gives:

$$\begin{aligned}\int \kappa \cdot du &= \int \frac{-dU}{\left(U + R_h - R_0 - \tau_h\right)} \\ \kappa u &= -\ln\left|U + R_h - R_0 - \tau_h\right| + constant \end{aligned} \tag{3.50}$$

Equation (3.50) is the analogue of the reparametrization given in equation (3.21). We can see that, for our metric, we have $\kappa = 1/4M$, which is the surface gravity [9], and $U + R_h - R_0 - \tau_h = \frac{v_0 - v}{4M} = \frac{v_0 - t + r}{4M}$. As $U \to U_h = \tau_h + R_0 - R_h$, i.e. as we get closer to the event horizon and the time at which it was formed, $u \to \infty$. Indeed, in our metric, $u \to \infty$ as $\frac{v_0 - v}{4M} \to 0$, which happens when $v \to v_0$, i.e. when the advanced time v approaches the time of creation of the event horizon.

To describe the quantum physics near the event horizon, we can quantize the system according to the canonical quantization scheme by replacing the field variable ϕ by a quantized field operator [2, 9]. We impose the canonical commutation relations for the incoming and outgoing fields [2, 9]:

$$\begin{aligned}\left[\phi_{in}\left(v_1, \Omega_1\right), \partial_{v_2}\phi_{in}\left(v_2, \Omega_2\right)\right] &= -2\pi i \delta\left(v_{12}\right)\delta\left(\Omega_{12}\right) \\ \left[\phi_{out}\left(u_1, \Omega_1\right), \partial_{u_2}\phi_{out}\left(u_2, \Omega_2\right)\right] &= -2\pi i \delta\left(u_{12}\right)\delta\left(\Omega_{12}\right)\end{aligned} \tag{3.51}$$

where $v_{12} = v_1 - v_2$. Since the field variables ϕ_{in} and ϕ_{out}, together with their complex conjugates, form a complete orthonormal basis, we can expand the fields in terms of their modes as [9]:

$$\begin{aligned}\phi_{in}\left(v, \Omega\right) &= \sum_k \left[a_k u_k\left(v, \Omega\right) + a_k^\dagger u_k\left(v, \Omega\right)\right] \\ \phi_{out}\left(u, \Omega\right) &= \sum_k \left[b_k u_k\left(u, \Omega\right) + b_k^\dagger u_k\left(u, \Omega\right)\right]\end{aligned} \tag{3.52}$$

Here, the a_k and b_k are the annihilation operators that generate the Hilbert space on $\mathcal{I}^-$ and $\mathcal{I}^+$, respectively ($a_k^\dagger$ and $b_k^\dagger$ are the creation operators). In our case, equation (3.52) leads specifically to:

$$\phi_{in}(v,\Omega) = \sum_{l,m} \int_{-\infty}^{+\infty} \frac{d\omega}{\sqrt{2\pi\omega}} [a_{\omega lm} e^{-i\omega v} + a^{\dagger}_{\omega lm} e^{i\omega v}] Y_{lm}(\Omega)$$
$$\phi_{out}(u,\Omega) = \sum_{l,m} \int_{-\infty}^{+\infty} \frac{d\omega}{\sqrt{2\pi\omega}} [b_{\omega lm} e^{i\omega u} + b^{\dagger}_{\omega lm} e^{-i\omega u}] Y_{lm}(\Omega) \tag{3.53}$$

The spherical harmonics are necessary to form a complete basis for the Ω space, i.e. the expressions in (3.53) satisfy the Klein-Gordon equation. The factor $\frac{1}{\sqrt{2\pi\omega}}$ is a normalization factor due to integration over all of space. Since both $\phi_{in}(v,\Omega)$ and $\phi_{out}(u,\Omega)$ are complete sets, we can express their modes in terms of each other like [9]:

$$u_k(u,\Omega) = \sum_{k'} [\alpha_{kk'} u_{k'}(v,\Omega) + \beta_{kk'} u^*_{k'}(v,\Omega)] \tag{3.54}$$

This relation is known as the Bogolyubov transformation and the factors $\alpha_{kk'}$ and $\beta_{kk'}$ are called the Bogolyubov coefficients and are given by:

$$\alpha_{kk'} = (u_{k'}(v,\Omega), u_k(u,\Omega))$$
$$\beta_{kk'} = -(u_{k'}(v,\Omega), u^*_k(u,\Omega)) \tag{3.55}$$

The transformation given in equation (3.54) results from comparison between different observers in quantum field theory [10]. In our case, we are comparing a stationary observer at $\mathcal{I}^-$ carrying out measurements on the incoming field, and a stationary observer at $\mathcal{I}^+$ carrying out measurements on the outgoing field. The relation between their measurements is then given by the Bogolyubov transformations [2]:

$$b_{\omega lm} = \int_0^{\infty} d\omega' \left(\alpha_{\omega\omega'} a_{\omega' lm} + \beta_{\omega\omega'} a^{\dagger}_{\omega' lm}\right)$$
$$b^{\dagger}_{\omega lm} = \int_0^{\infty} d\omega' \left(\alpha_{\omega\omega'} a^{\dagger}_{\omega' lm} + \beta_{\omega\omega'} a_{\omega' lm}\right) \tag{3.56}$$

From the definitions in (3.56) it is clear that the Fock spaces based on the modes u_k and $u_{k'}$ are not, in general, equal. In the Fock representation, the basis vectors of the Hilbert space are constructed from the vector $|0\rangle$, which is the vacuum state (also called no-particle state) [9]. The vacuum state is annihilated by all annihilation operators of the corresponding field. The one-particle state is created by letting the creation operator act on the vacuum state $|0\rangle$. In our scenario, the incoming field has corresponding annihilation operators $a_{\omega lm}$, whereas the outgoing field has the annihilation operators $b_{\omega lm}$. Each field describes a different Fock space, and each field has its own vacuum state. For the incoming field we have: $a_{\omega lm}|0\rangle_{in} = 0, \forall\{\omega, l, m\}$, and for the outgoing field: $b_{\omega lm}|0\rangle_{out} = 0, \forall\{\omega, l, m\}$. However, since $a_{\omega lm} \neq b_{\omega lm}$ for general values of the operators, the annihilation operator of one field will not annihilate the vacuum state of the other field: $a_k|0\rangle_{out} \neq 0, b_k|0\rangle_{in} \neq 0$ [9]. This means that the vacuum of the u_k modes contains particles in the $u_{k'}$ mode, and vice versa. This causes an observer at $\mathcal{I}^-$ to see particles when looking at the outgoing field, even though an observer at $\mathcal{I}^+$ observes a vacuum state there (and vice versa).

We can calculate the Bogolyubov coefficients using equation (3.55) and the field modes given in equation (3.53) to get:

$$
\begin{aligned}
\alpha_{\omega\omega'} &= e^{-i\left(\omega'-\omega\right)v_0}\frac{e^{2\pi M\omega}\Gamma\left(1-i4M\omega\right)}{2\pi\sqrt{\omega\left(\omega'+i\epsilon\right)}} \\
\beta_{\omega\omega'} &= e^{+i\left(\omega'+\omega\right)v_0}\frac{e^{-2\pi M\omega}\Gamma\left(1-i4M\omega\right)}{2\pi\sqrt{\omega\left(\omega'-i\epsilon\right)}}
\end{aligned}
\tag{3.57}
$$

A full computation of these coefficients is beyond the scope of this paper, but we can make some plausibility arguments. One starts the computation by filling in $u(v) = v_0 - 4M\ln\left(\frac{v_0-v}{4M}\right)$ into the definition for $\phi_{out}(u,\Omega)$ in equation (3.53). Then, one equates $\phi_{out}(u(v),\Omega)$ with $\phi_{in}(v,\Omega)$. Since both expressions contain integrals, in order to compare individual modes one has to integrate. This is done by multiplying both sides with a factor $e^{-i\omega' v}$ and integrating over v. Now, when one looks at the definition of the Bogolyubov coefficients in equation (3.55), and at the fields written out in (3.53), the presence of some factors in equation (3.57) is clear.

The $\frac{1}{2\pi\sqrt{\omega\omega'}}$ term obviously comes from the normalization factor in the definition for the fields. The factor $e^{\pm i\left(\omega'\pm\omega\right)v_0}$ comes from the multiplication of the modes $u_k(u,\Omega), u_{k'}(v,\Omega)$ corresponding to the fields in (3.53), since u carries a factor of v_0 due to the reparametrization. The remaining factors in equation (3.57) all arise from an analytic continuation of the integral to the complex plane. This analytic continuation entails redefining v as $\frac{v}{i(\omega-\omega')}$. Analytic continuation means rotating the integration contour to the complex plane in such a way that the integral converges everywhere. For our integral to be convergent everywhere, one has to regularize it giving ω' an imaginary piece $i\epsilon$. This factor occurs in the denominator of both coefficients in (3.57).

The analytic continuation is done so that we can write the coefficients using the gamma function as in (3.57). The gamma function is defined as: $\Gamma(z) = \int e^{-t}t^{z-1}dt$. As can be seen in equation (3.57), our z is defined as $1 - i\omega 4M$. This leads to the conclusion that our t is identified with $\frac{v_0 - v}{4M}$, since $t^{z-1} = e^{(z-1)\ln(t)} = e^{-i\omega 4M \ln(t)}$, and we have a factor of $e^{-i\omega 4M \ln(\frac{v_0-v}{4M})}$ in our multiplication due to the definition of $u(v)$. Since $\Gamma(z) = \int e^{-t}t^{z-1}dt$, we still need to explain the factor of e^{-t}. This is exactly done by the redefinition of v stated above. This definition gets rid of the factor $i(\omega - \omega')$ in the exponent, such that we get the e^{-t} term.

The only factor of equation (3.57) that still needs explaining is the $e^{2\pi M\omega}$ term. Indeed, this term also comes from the redefinition of v. Since our t is dependent on v, the factor t^{z-1} in the gamma function also changes with this redefinition. We get an extra factor of $\left(\frac{i}{(\omega-\omega')}\right)^{-i\omega 4M} = e^{-i\omega 4M \ln\left(\frac{i}{(\omega-\omega')}\right)} = e^{-i\omega 4M\left(\ln(i) - \ln(\omega-\omega')\right)}$ due to this redefinition. One now uses the fact that $\ln(i) = \frac{\pi i}{2}$ to get $e^{-i\omega 4M\left(\frac{\pi i}{2} - \ln(\omega-\omega')\right)} = e^{2\pi\omega M}e^{i\omega 4M\ln(\omega-\omega')}$, where the second exponent goes to 1 as $(\omega - \omega')$ goes to 1. The Bogolyubov coefficients in equation (3.57) are thus in their asymptotic form.

The coefficients in (3.57) can be filled into equation (3.56) to get the exact relation between the creation- and annihilation-operators of the Hilbert spaces corresponding to the in- and out-going fields. Subsequently, these operators are filled into equation (3.53) to get a complete expression of the outgoing state on $\mathcal{I}^+$.

Equation (3.21) was the result for the reparametrization when gravitational back-reaction is neglected. The next step is to rewrite this equation by taking these interactions into account. Thus, we want to know what the effect is of some infalling wave on an outgoing wave that interacts with it (see Figure 4). When an infalling wave carries some amount of energy δM, the black hole's event horizon will become a factor $2\delta M$ larger, since $r = 2M$ [1]. This change in radius has an effect on the time v_0 at which the event horizon is created, as can be seen from Figure 8. This new, slightly larger, event horizon is created at the earlier time $v_0 + \delta v_0$ (where δv_0 is a negative number; see Figure 8). This small change δv_0 can be calculated from the reparametrization (3.21) and is given by:

$$\delta v_0 = -4\delta M e^{\frac{v_0 - v}{4M}} \tag{3.58}$$

We thus need to know the precise form of δM to know what δv_0 looks like. In fact, this term can be computed explicitly within a linearised approximation of the Einstein field equations [2].

The computation of δM begins with the assumption that the back-reaction effects are small, such that we can represent them by a small perturbation to the classical metric: $d\tilde{s}^2 = ds^2 + h_{\mu\nu}dx^\mu dx^\nu$ [2, 6]. Because of our assumption that $h_{\mu\nu}$ is small, we can neglect any terms that are higher than first order in this quantity, which means that $h_{\mu\nu}$ satisfies the linearised Einstein field equations [2, 6].

The usual Einstein equations are given by $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi T_{\mu\nu}$, where $R_{\mu\nu}$ is the Ricci tensor, R is the Ricci scalar (the contracted Ricci tensor), and $T_{\mu\nu}$ is the stress-energy tensor [6]. The linearised version of these equations is an approximation in which the non-linear terms in the metric are ignored (for a derivation see e.g. [6]). The source of our perturbation $h_{\mu\nu}$ is equal to the stress-energy tensor of the field, so we may write [2]:

$$h_{\mu\nu}(x) = 8\int d^4y D_{\mu\nu}{}^{\lambda\sigma}(x,y)\, T_{\lambda\sigma}(y) \tag{3.59}$$

where $D_{\mu\nu}{}^{\lambda\sigma}$ is the propagator for the field on the black hole background [2]. With this perturbation, we want to find an approximate solution to the linearised Einstein field equations. We will derive the solution in the region close to the event horizon of the black hole, where we assume the metric looks like the one in (3.2), but now with the perturbation incorporated into the x^+ and x^- coordinates:

$$ds^2 = -\frac{32M^3}{r}e^{-r/2M}\left(dx^+ + h_{--}\left(x^-,\Omega\right)dx^-\right)\left(dx^- + h_{++}\left(x^+,\Omega\right)dx^+\right) + r^2 d\Omega^2 \tag{3.60}$$

The linearised Einstein equations for the metric in (3.60) are given by [2]:

$$\frac{2^9 M^4}{e}\left(\Delta_\Omega - 1\right)h_{\pm\pm} = T_{\pm\pm} \tag{3.61}$$

We can integrate this equation to get the perturbation $h_{\pm\pm}$ back [2]:

$$h_{\pm\pm}\left(x^\pm,\Omega\right) = \frac{e}{2^9M^4}\int d\tilde{\Omega} f\left(\Omega,\tilde{\Omega}\right)T_{\pm\pm}\left(x^\pm,\tilde{\Omega}\right) \tag{3.62}$$

where f is the finite shift in the metric due to the introduction of the matter δM.

Now that we know what δM looks like, we can fill this definition into equation (3.58) to get our final expression for the shift in the advanced time v:

$$\delta v_0\left(\Omega_1\right) = 8\int d^2\Omega_2 f\left(\Omega_1,\Omega_2\right)P_{in}\left(\Omega\right) \tag{3.63}$$

where the factor $\frac{e}{2^9M^4}$ is incorporated into f, and $f\left(\Omega_1,\Omega_2\right)$ satisfies:

$$\left(\Delta_\Omega - 1\right)f\left(\Omega_1,\Omega_2\right) = -2\pi\delta^{(2)}\left(\Omega_{12}\right) \tag{3.64}$$

where Δ_Ω is the Laplacian over the Ω-space. The momentum P_{in} is given, as usual, by $P_i \equiv \int_t T_{ti}d^{n-1}x$

[9]; in our case, this results in:

$$P_{in}(\Omega) = \int_{v_0}^{\infty} dv \cdot e^{\frac{v_0 - v}{4M}} T_{vv}(v, \Omega) \tag{3.65}$$

since T_{vv} is the only non-zero component of the stress-energy tensor.

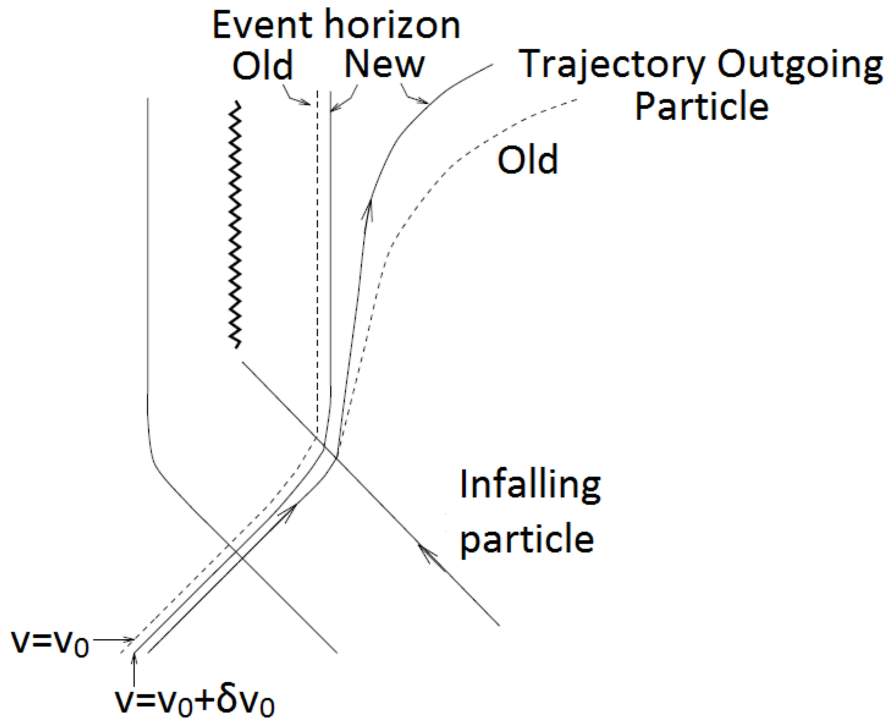

Figure 8: Taken from [2]. An infalling shell of matter carrying energy δM changes the position of the horizon by an amount $2\delta M$. The time of creation of this horizon is shifted an amount δv_0. Out-going light-rays are also affected by this change.

This small change δv_0 in the origination time of the event horizon leads to a change in the retarded time u of an outgoing wave. The time u will be delayed by a factor δu, which can be explicitly computed. The outgoing wave will thus reach an outside observer at a later time than it would have done without the interaction with the wave of energy δM (see Figure 8). From this reasoning, it is clear that the gravitational interaction between an infalling and outgoing wave as shown in Figure 4 can indeed have an effect on the evaporation process of the black hole. This effect can be included in the original reparametrization given by equation (3.21) simply by adding the factor $\delta v_0$to the original v:

$$\phi_{out}(v, \Omega) = \phi_{in}\left(v(u) + \delta v_0(\Omega), {}^{P}\Omega\right) \tag{3.66}$$

Here, $\delta v_0(\Omega)$ is given by equation (3.63), and $v(u)$ can be found from equation (3.21):

$$
\begin{aligned}
u(v) &= v_0 - 4M \ln\left(\frac{v_0 - v}{4M}\right) \\
\frac{u(v) - v_0}{-4M} &= \ln\left(\frac{v_0 - v}{4M}\right) \\
e^{\frac{v_0 - u(v)}{4M}} &= \left(\frac{v_0 - v}{4M}\right) \\
v(u) &= v_0 - 4Me^{\frac{v_0 - u}{4M}}
\end{aligned}
\tag{3.67}
$$

These corrections are incorporated by substituting equation (3.66) for the original (equation (3.20)) in the derivation of the Hawking radiation.

Ultimately, this gravitational back-reaction may have a large enough effect on the Hawking radiation process such that the information locked up inside the black hole actually comes out along with the outgoing radiation. The idea is that, due to the newly defined outgoing field given by (3.66), the commutator between the in- and out-going fields will not vanish for $v > v_0$, where v_0 is the time at which the event horizon of the black hole is created [2]. However, the magnitude of the corrections to the final state is a complicated and subtle issue on its own. It could, in principle, be computed exactly, but one finds that this calculation depends critically on the boundary of integration over the retarded and advanced time coordinates [2]. In fact, the cut-off of the interaction region will make the corrections either very large or increasingly shrinking with time for the final state on $\mathcal{I}^+$. In [2], a short discussion on this matter is worked out qualitatively, but exact details on this manner were left for a later publication.

3.2 Other Proposed Solutions

As we discussed in the Introduction, numerous possible explanations for the I.P. have been proposed over the years. It is not the intention to discuss the majority of them here, but some interesting, intuitive, or promising ideas will be discussed shortly. To start with, we will review some ideas of S.D. Mathur as discussed in [3] and [7]. Mathur has published dozens of papers on black holes and the paradox over the last decade (e.g. [3, 7]). He has written about the nature of the paradox, about other people's solutions to it and his own criticism on these, and of course about his own proposed way to fix the problem - the fuzzball solution.

Mathur distinguishes between three main types of fallacies made in the search for a solution to the I.P. [3]. The first error lies in the expression of evolution from an initial Cauchy surface to a later Cauchy surface. Mathur states that the error in this step is the replacement of the correct evolution by a new evolution, which is not proven to exist in the theory of gravity [3, 7], i.e. the third step in the argument discussed in section 2.2 is violated. The second type of error comes from wrongly changing the state on the initial slice, i.e. in a way that is not permitted by the theory of gravity. Lastly, people make mistakes in describing the consequences and effects of particular evolutions from initial to later Cauchy surfaces.

An example of the first type of error is present in the so-called Rubik's cube model of black hole evaporation [3]. This model was proposed by Czech, Larjo, and Rozali in 2011 and represents the black hole as a large collection of Rubik's cubes. Each cube can be in any of a finite number of possible states, which we shall call $C_1, ..., C_m$. The state of each cube can be altered by applying one of four possible operations: left ($\vec{L}$), right ($\vec{R}$), up ($\vec{U}$), and no-change ($\vec{N}$) [3]. We know that the evolution of Cauchy slices leads to particle creation. We will name the states of the particles after the operation of evolution in which they were created, i.e. a particle created in an *up* move is in the state u, a particle created in a *left* move is in the state l, etc. Now, the evolution of quanta on Cauchy slices is modelled by a sum of these operators with equal amplitude (since there is no favoured operation), multiplied by the corresponding state of the particle created:

$$\begin{aligned}\{C_1, ..., C_m\} \to &\frac{1}{2}(\{\vec{N}C_1, ..., \vec{N}C_m\}|n\rangle + \{\vec{L}C_1, ..., \vec{L}C_m\}|l\rangle \\ &+ \{\vec{R}C_1, ..., \vec{R}C_m\}|r\rangle + \{\vec{U}C_1, ..., \vec{U}C_{m-1}\}|qu\rangle)\end{aligned} \tag{3.68}$$

The extra q in the last term of equation (3.68) represents a quantum in state q that is emitted as soon as the Rubik's cube is in its 'solved' state. In fact, as soon as it reaches this state, the Rubik's cube is taken out of the list of cubes that make up the black hole, and thus the quantum q, carrying a certain energy, is needed to account for the deletion of one cube [3]. We have made the arbitrary choice here that the operation $\vec{U}$ brought the last cube C_m to its 'solved' state.

Czech, Larjo, and Rozali proof in their paper that this evolution indeed leads to complete evaporation of the black hole, with all the energy stored in the emitted quanta q. Their final state is a linear combination of general sequences of quanta in the allowed states l, u, r, n, q [3]. This is a pure state, i.e. it is not entangled with anything. Mathur's explanation of this result is that the Rubik's cube model is in fact not a good model for an evaporating black hole, but is analogous to that of a burning paper. In the case of a paper burning away, there is indeed no question of a paradox since the emitted quanta form a pure state. The crucial difference between a black hole and the burning paper model is that, in a black hole, each step in the evolution leads to the *same* entangled state as given in (2.18) (up to small corrections), whereas a burning paper emits radiation that depends on the state of the paper at that moment [3]. Each step in the evolution of the latter situation thus takes on a different form.

The problem with the Rubik's cube model is that the evolution given in (3.68) is not the one given by the Schwarzschild metric [3], which describes the geometry of a black hole. This metric would give an evolution like the one studied in section 2.2, which has an analogous form in the Rubik's cube regime [3]:

$$\begin{aligned}\{C_1, ..., C_m\} \to \frac{1}{2}\{C_1, ..., C_m\} &\otimes \frac{1}{2}(n'_1 \otimes n_1 + l'_1 \otimes l_1 + r'_1 \otimes r_1 + u'_1 \otimes u_1) \\ &... \\ &\otimes \frac{1}{2}(n'_N \otimes n_N + l'_N \otimes l_N + r'_N \otimes r_N + u'_N \otimes u_N) \end{aligned} \tag{3.69}$$

where the quanta denoted with an accent $'$ are the members of the created pairs that reside inside the event horizon. This description captures the correct form of the evolution as discussed in section 2.2. Indeed, the evolution in (3.69) leads again to the paradox through the entanglement between the quanta $\{n_i, l_i, r_i, u_i\}$ and $\{n'_i, l'_i, r'_i, u'_i\}$. Hence, we conclude that we really need to look at the entangled state of the pair of particles that is created, and not merely focus on the particles that are seen by an outside observer, as is done in (3.68).

Mathur mentions briefly the possibility of black hole remnants in [3, 7]. In this idea, one simply puts a bound on the evaporation of the black hole. This means that, when the black hole reaches the Planck size, it stops emitting quanta and a small remnant of the black hole lives on forever. This entails the postulation of new particles called remnants, which contain all the information of the black hole and live on forever, which subsequently leads to the conclusion that there is no paradox. However, we computed the entanglement entropy of the final state to be $N \ln 2$ in section 2.2. This means that the black hole remnant would need at least N number of possible states to reach this amount of entanglement with the radiated quanta. The argument in section 2.2 that led us to the entropy $N \ln 2$ could have started with a black hole of any size, meaning that N is unbounded, i.e. the degeneracy of the black hole remnant is unbounded. Hence, we would have a system with an unbounded number of possible states (i.e. an infinite number of remnant particle species), but with bounded size and energy, which is still in contradiction with the expected behaviour of quantum mechanical systems [3, 7]. Mathur groups the notion of remnants and information loss together, whereas other authors, such as Susskind [1], reject the remnant idea altogether because of its highly pathological nature. The idea of black hole remnants is also mentioned in [4], where it is approached with the same scepticism as in Susskind's book. The authors of [4] note that the problem of unbounded degeneracy could be circumvented, but this usually leads to a sea of other problems with, for example, energy conservation, locality, or interactions between remnants and ordinary fields.

Another, very intuitive, solution to the I.P. is to let the event horizon copy each piece of information that falls into the black hole, and send one copy to the world outside the black hole [1]. However, the quantum Xerox principle (or no-cloning principle) forbids such a scenario. This principle states that quantum information can never faithfully be copied [1]. If one could perform such a cloning process, the Heisenberg uncertainty principle could be violated, since we could perform a set of measurements on the system and all its clones. Thus, this idea does not provide us with a solution to the paradox either.

A more recent approach to studying the I.P. and the general nature of black holes (see e.g. [4]) is a concept which emerged from string theory called Anti de Sitter-Conformal Field Theory (AdS/CFT) correspondence. The idea that is central to this correspondence is the Holographic Principle. As the name suggests, this principle states that it is possible to describe a three (or N) dimensional space by a two (or N-1) dimensional hologram at its boundary [1]. AdS space has a negative cosmological constant, which makes it naturally ball-like. The holographic principle then states that we can describe everything in the bulk (usually taken to be $AdS(5)$ [1]) by a theory whose degrees of freedom are identified with the boundary (the corresponding 4-dimensional boundary of AdS space). An additional restriction is that the boundary can have no more than one degree of freedom per Planck area. When one considers all the symmetries of AdS space, one inevitably comes to the conclusion that quantum gravity in $AdS(5)$ (actually in $AdS(5) \otimes S(5)$, where $S(5)$ is the 5-sphere) should be described by a conformal, Lorentz invariant, quantum field theory on its boundary [1]. In fact, there is only one class of symmetries that are appropriate for this description: the $SU(N)$ supersymmteric Yang-Mills theories [1]. The AdS/CFT correspondence states that the superstring theory in the bulk of AdS space in five dimensions and the four dimensional $SU(N)$ on the AdS boundary are completely equivalent.

A naive use of AdS/CFT correspondence as discussed in [3, 7] leads to the conclusion that there is no paradox at all. Since the Yang-Mills CFT is manifestly unitary, and the CFT is equivalent to the gravity theory in the bulk, this latter theory should also be unitary and thus no information is ever lost. However, this is obviously a circular argument. Mathur goes on stating that AdS/CFT might very well be a tool to solve the paradox, but that it cannot simply be evoked as a ready-made solution to the paradox in its current state. Before any such claim can be made, the formation and evaporation processes of black holes first have to be studied in detail in the Yang-Mills CFT on the boundary or in the gravity theory in the bulk. In fact, this discussion is similar to the one in a recent paper on quantitative approaches to solving the I.P. [4]. The authors of [4] first state that the paradox is indeed circumvented when AdS/CFT is evoked. However, they note that this argument does not explain why and where Hawking's derivation fails. Such an explanation is sought after in the rest of the article.

AdS/CFT correspondence is also discussed in [1]. We will not discuss the details of the argument here, but Susskind comes to the conclusion that "any phenomenon which crucially depends on the finiteness of horizon entropy will be gotten wrong by the approximation of quantum field theory in a fixed background" (p.146). Indeed, the study of information-loss and -recovery from black holes falls into this category. This thus suggests that the AdS/CFT correspondence is a crucial factor that needs to be taken into account when discussing the I.P.

4 Conclusions

The black hole information paradox has played a large and interesting role in physics over the last 30 years or so. In light of the discussions laid out in this paper, it is clear that this will keep on being a prominent and fascinating subject for a long time to come. As we have seen in section 2, the paradox originates in Hawking's discovery of black hole evaporation: black holes lose energy to their surrounding through the so-called Hawking radiation. In the second half of section 2, we saw how this outgoing radiation leads to a paradox regarding the information inside the black hole. Entangled particle pairs are created near the black hole, and it is precisely this entanglement that leads to the information paradox. When one member of the pair falls into the black hole, whereas the other one escapes, the latter particle would be entangled with nothing once the black hole has completely evaporated. This means that an initially pure state has evolved into a mixed state, which is a violation of the unitary nature of quantum mechanics.

We have discussed a method for computing the corrections to the outgoing radiation of an evaporating black hole due to gravitational back-reaction. This entailed looking at the evolution of quanta from initial to later Cauchy surfaces. We calculated the Bogolyubov transformations, which allows the exact computation of the density matrix that describes the outgoing Hawking radiation. The corrections due to the interaction between outgoing and infalling waves can potentially become very large, but this depends critically on the cut-off of the interaction region. Our discussion was indeed limited, since later papers on the gravitational back-reaction were not included. Neither did we discuss the quantum mechanical details of the back-reaction. These factors should definitely be taken into account if one wishes to accomplish a complete discussion of the possibility of these corrections to solve the information paradox.

We discovered that there is an enormous diversity of opinions on the ultimate solution to the information paradox A selection of proposed resolutions was discussed briefly, but a detailed account of these is beyond the scope of this paper. The most promising current view on the information paradox is the widely applicable AdS/CFT correspondence. Studying the black hole geometry using this correspondence, and thus understanding the black hole's holographic properties, will certainly shed light on the information paradox and its final resolution.

References

[1] Susskind L, Lindesay J. 2005. *Black holes, information, and the string theory revolution.* Singapore: World Scientific Publishing Co. Pte. Ltd.

[2] Kiem Y, Verlinde H, Verlinde E. 1995. *Black hole horizons and complementarity.* Physical Review D. 52(12): 70537065. arXiv:hep-th/9502074v2

[3] Mathur SD. 2011. *What the information paradox is not.* arXiv:1108.0302v2 [hep-th]

[4] Balasubramanian V, Czech B. 2011. *Quantitative approaches to information recovery from black holes.* Classical and Quantum Gravity 28:163001. arXiv:1102.3566v1 [hep-th]

[5] Horowitz GT, Maldacena J. 2003. *The black hole final state.* JHEP 0402:008. arXiv:hep-th/0310281v

[6] Carroll S. 2004. *Spacetime and geometry: An introduction to general relativity.* San Francisco: Pearson Education Inc.

[7] Mathur SD. 2009. *The information paradox: a pedagogical introduction.* Classical and Quantum Gravity 26:224001. arXiv:0909.1038v2 [hep-th]

[8] Townsend PK. 1997. *Black Holes: Lecture notes for a 'Part III' course 'Black Holes' given in DAMTP, Cambridge.* arXiv:gr-qc/9707012

[9] Birrell ND, Davies PCW. 1982. *Quantum fields in curved space: Cambridge monographs on mathematical physics.* Cambridge: Cambridge University Press.

[10] Giddings SB, Nelson WM. 1992. *Quantum emission from two-dimensional black holes.* Physical Review D. 46: 2486-2496. arXiv:hep-th/9204072v3