Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Mathematics

Black Hole Entropy and the Information Loss Paradox

Pierce Comerford
Supervised by Prof. Manya Sahni O'Hara

April 6, 2025

Submitted in partial fulfilment of the requirements for the degree of
B.A. Theoretical Physics

Declaration

I hereby declare that this Final Year Project is entirely my own work and that it has not been submitted as an exercise for a degree at this or any other university.

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at `http://www.tcd.ie/calendar`.

I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at `http://tcd-ie.libguides.com/plagiarism/ready-steady-write`.

Signed: Pierce Comerford

Date: Sunday, 6 April, 2025

Abstract

Since their characterization as one-way membranes in 1958 [1], the ostensibly time-asymmetric properties of black holes has been characteristic of all conceptual issues that arise in their study. The initial contention by Wheeler was that entropy simply vanished in a black hole [2], which motivated the establishment of black hole thermodynamics [3, 4] and eventually lead to the discovery that black holes radiate with a certain temperature [5] and have entropy proportional to their surface area. The exact thermal evaporation of a black hole, however, violates the unitarity of quantum mechanics, prohibiting retrodiction [6]. This is known as the information loss paradox. In this paper we explore the development of time-asymmetric properties of black holes in various stages of the development of black hole mechanics, and explore some current literature in the debate – namely proposed resolutions to the information paradox, the issues surrounding them, and the corrections they require for the Bekenstein-Hawking entropy. Wald's formulation of black hole entropy as a Nöther charge of black hole spacetimes [7] is used to explore how corrections may be computed.

Acknowledgements

I first ought to thank Manya for her extremely helpful advice, direction, and encouragement throughout the course of this project, and for helping to make it fun.

Also, thanks to Chaolun for his helpful questions during the talk, and physics advice in general over the last however-so-many years.

I also should thank my fellow classmates – especially Ronan and Chloe, for keeping me grounded and reminding me to eat. I could name more, but the appreciation extends to everyone in TP, anyone who passed through the maths library as this was being done, and anyone who could be considered 'in the same boat'.

Thanks also to everyone outside the boat: Ellen (of course), David, Tom and Alex (for the magic), Mic and Thomas (for the doughnut), Dexter, and Mark.

And of course: to my parents, for their unwavering support, and my brother, Jim, for reminding me of the importance of a simple explanation.

Contents

List of Figures

Conventions

Here, and in the following, we set $G = c = \hbar = k_{\mathrm{B}} = 4\pi\varepsilon_0 = 1$ unless otherwise noted.

We use the $(-, +, +, +)$ metric convention.

Upright (Roman) characters are sometimes used to denote differential forms, to distinguish, for example, the Lagrangian L from the Lagrangian n-form L.

1 | Introduction

Black holes are intriguing and unusual objects, commonly regarded as one of the most surprising consequences of Einstein's theory of general relativity. All black holes are characterised by a surrounding surface known as the *event horizon* [3, 9, 10, 11], sometimes called a 'one-way membrane' [1, 12] to stress the idea that, once a particle passes the horizon, it is impossible for the particle to come back out. This property is what leads black holes to be thought of as regions of space from which "nothing, not even light, can escape" [10].

The 'no-escape' description misses some intricacies of modern black hole physics. Rotating (Kerr and Kerr–Newman) black holes are surrounded by an additional region, outside the event horizon, known as the ergosphere, in which no object can be seen as stationary by distant observers, an effect known as *frame-dragging*. As Roger Penrose pointed out [8], this effect can be used to extract rotational energy from the black hole, up to the point that the black hole degenerates to a static (Schwarzschild) solution with *irreducible mass* m_{ir} [12], so called because it cannot decrease under any physical process. As Zel'dovich [13] points out, this process also allows classical rotating black holes to radiate.

The surface area A of the event horizon is proportional to m_{ir}, and the consequent rule $\mathrm{d}A \geq 0$ is reminiscent of the second law of thermodynamics. Indeed it is possible to construct four laws governing the behaviour of black holes which exactly resemble the laws of thermodynamics, such as Bekenstein's thermodynamic equation

$$\mathrm{d}M = \frac{\kappa}{2\pi}\frac{\mathrm{d}A}{4} + \mathbf{\Omega} \cdot \mathrm{d}\mathbf{J} + \Phi \mathrm{d}Q, \tag{1.1}$$

with the quantity

$$S = \frac{A}{4} \tag{1.2}$$

occupying the position of entropy, and the quantity

$$T = \frac{\kappa}{2\pi} \tag{1.3}$$

(where κ is the *surface gravity*) occupying that of temperature [2, 3, 4]. The other quantities are energy terms coming from angular momentum and electric charge.

Historically it has been considered absurd for a black hole to have any temperature other than zero, as that would require the black hole to equilibrate with its surroundings by radiating [4], although some authors differ [14]. But a semiclassical calculation by Stephen Hawking showed that the nonzero vacuum expectation value at the event horizon means that, on average, energy is slowly sapped away from a black hole in a process called *Hawking radiation* [5], with a temperature equal to (1.3). Hawking radiation leads to the eventual evaporation of the black hole.

At the start of evaporation, the state just outside the black hole is described by a superposition of ingoing and outgoing modes. After evaporation, it is described only by the outgoing modes. This violates the unitarity principle of quantum mechanics (that if the state at t_0 is known, it can be known in principle at $t_0 \pm \Delta t$), because there appears to be more information about the state of the object which collapses to form the black hole than about the state of the system after evaporation [15]. The breakdown of this principle is known as the 'black hole information loss paradox', and is one of the key questions that a full theory of quantum gravity is expected to be able to answer [6, 16].

It is generally considered that the resolution to the paradox is that information is not lost, but is somehow encoded in Hawking radiation [16, 17]. String-theoretic results provide a statistical-mechanical account of black hole entropy in terms of microstates. From this, one can reproduce the Bekenstein-Hawking entropy at leading order with corrections. Higher-order gravity theories [18] (*i.e.*, theories which generalise the Einstein–Hilbert action to include higher-order derivatives of the metric tensor) also reproduce these results as they may be considered more accurate effective field theories for an underlying quantum gravity. Wald [7] provides a toolkit for computing black hole entropy in general gravity theories by formulating the entropy as a Nöther charge in the classical case.

In this project, a review of black-hole thermodynamics will be given, and the laws including (1.1) re-derived and explained. The semi-classical treatment of black holes, especially quantum events at the event horizon in curved spacetime, will be discussed, with attention given to Hawking radiation and its derivation. The implications for unitary black hole time evolution is discussed, following [19] by modelling the black hole as a scattering operator. Wald's formulation of black hole entropy as a Nöther charge will be examined [7] as a toolkit for finding corrections to (1.2), with an envoi to how this contends with statistical-mechanical or effective-QFT results.

2 | Classical Black Holes

2.1 The Schwarzschild Metric

In 1915, Albert Einstein completed the monumental task [20], started by him in 1911 [**?**], of uniting the special theory of relativity with Newtonian gravitation by introducing his general theory of relativity. In particular, the theory relates the curvature of spacetime to the distribution of mass–energy within it, given by the Einstein field equations [21],

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{2.1}$$

where $R_{\mu\nu}$ is the Riemannian curvature of spacetime, $R = R^{\lambda}{}_{\lambda}$ is the Ricci (or scalar) curvature, $g_{\mu\nu}$ is the metric of spacetime, and $T_{\mu\nu}$ is the stress–energy–momentum tensor. Without a mass–energy distribution ($T_{\mu\nu} = 0$), this reduces to the vacuum field equations,

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0. \tag{2.2}$$

One useful aspect of the vacuum Einstein equations of motion is found by taking the trace, in which case we obtain

$$\begin{aligned} R - \frac{1}{2}R \times 4 &= 0 \\ \implies R - 2R &= 0 \\ \implies R &= 0, \end{aligned} \tag{2.3}$$

i.e., the Ricci scalar vanishes in vacuum.

The first spherically symmetrical solution to (2.2) was found by Karl Schwarzschild in the same year, given by the Schwarzschild metric [22],

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2 d\Omega^2. \tag{2.4}$$

Here, $d\Omega$ is the spherical volume element, $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. The metric approximates the gravitational field around a spherically symmetrical celestial object such as the Sun, and in this case M is the mass of the object at the origin. It turns out, due to Birkhoff's theorem, that it is the *only* spherically symmetric solution of the Einstein field equations.

When one considers this as such an approximation, it is normally supposed that the radius of the spherical object described is larger than its Schwarzschild radius $r_s = 2M$, and thus that no test mass placed in the system would approach the region $r < r_s$. However, objects whose radius are or approach the Schwarzschild radius are in fact physically reasonable, as shown theoretically by Oppenheimer, Tolman, *et. al.* [23, 24], and eventually observed by astronomers in 1971 [25]. Such an object is known as a black hole.

The Schwarzschild black hole has unusual properties. The hypersurface $r = 2M$ is a singularity in the standard spherical coordinates given in (2.4), but by switching to a different coordinate system, it can be considered a one-way system: anything which passes the event horizon of the black hole, up to and including photons, are drawn inexorably toward $r = 0$.

In particular, one can switch to Kruskal–Szekeres coordinates, given by

$$U = -4M \exp\left(-\frac{u}{4M}\right); \quad V = 4M \exp\left(\frac{v}{4M}\right), \tag{2.5}$$

where we can identify $4M$ as the inverse of the surface gravity κ, and the u and v are the Eddington-Finkelstein coordinates,

$$u = t - r_* \text{ and } v = t + r_* \quad \Longleftrightarrow \quad t = \frac{1}{2}(v + u) \text{ and } r_* = \frac{1}{2}(v - u). \tag{2.6}$$

The 'tortoise coordinate' r_* above is

$$r_* = r + 2M \ln\left(\frac{r}{2M} - 1\right), \tag{2.7}$$

which approaches $-\infty$ as r approaches the Schwarzschild radius $2M$. To an observer in this system, nothing ever reaches $r = 2M$, and the black hole is always just about to form. This is the frame of a stat static observer far from the black hole.

We can rewrite the Schwarzschild metric with the tortoise coordinate by considering its derivative

$$\begin{aligned}\frac{\mathrm{d}r_*}{\mathrm{d}r} &= 1 + \frac{1}{\frac{r}{2M} - 1} \\ \implies \mathrm{d}r_* &= \left(1 + \frac{1}{\frac{r}{2M} - 1}\right)\mathrm{d}r \\ \implies \mathrm{d}r &= \left(1 - \frac{2M}{r}\right)\mathrm{d}r_*.\end{aligned}$$

Then the Schwarzschild metric reduces to

$$\mathrm{d}s^2 = \left(1 - \frac{2M}{r}\right)\left(\mathrm{d}t^2 - \mathrm{d}r_*^2\right) - r^2\mathrm{d}\Omega^2. \tag{2.8}$$

To put this in terms of the Eddington-Finkelstein coordinates, then, we note that

$$\mathrm{d}t = \frac{1}{2}(\mathrm{d}v + \mathrm{d}u) \text{ and } \mathrm{d}r_* = \frac{1}{2}(\mathrm{d}v - \mathrm{d}u), \tag{2.9}$$

and thus

$$\mathrm{d}s^2 = \left(1 - \frac{2M}{r}\right)(\mathrm{d}u\,\mathrm{d}v) - r^2\mathrm{d}\Omega^2. \tag{2.10}$$

Then we can also obtain the metric in Kruskal-Szekeres coorinates by noting

$$u = -4M\log\left(-\frac{U}{4M}\right); \quad v = 4M\log\left(\frac{V}{4M}\right), \tag{2.11}$$

and thus

$$\mathrm{d}u = -\frac{4M}{U}\mathrm{d}U; \quad \mathrm{d}v = \frac{4M}{V}\mathrm{d}V. \tag{2.12}$$

Inserting these into (2.10) we find

$$\begin{aligned}\mathrm{d}s^2 &= \frac{1}{\exp\left(\frac{v-u}{4M}\right)}\left(1 - \frac{2M}{r}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2 \\ &= \exp\left(\frac{-r_*}{2M}\right)\left(1 - \frac{2M}{r}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2 \\ &= \exp\left(\frac{-r}{2M} - \log\left(\frac{r}{2M} - 1\right)\right)\left(1 - \frac{2M}{r}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2 \\ &= \left(\frac{r}{2M} - 1\right)^{-1}\exp\left(\frac{-r}{2M}\right)\left(1 - \frac{2M}{r}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2 \\ &= \frac{2M}{r}\exp\left(-\frac{r}{2M}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2.\end{aligned}$$

The space- and time-like coordinates for Kruskal–Szekeres spacetime are

$$T = \frac{1}{2}(V + U), \quad X = \frac{1}{2}(V - U), \tag{2.13}$$

so the metric is then

$$\mathrm{d}s^2 = \frac{2M}{r} \exp\left(-\frac{r}{2M}\right)\left(-\mathrm{d}T^2 + \mathrm{d}X^2\right) - r^2\mathrm{d}\Omega^2. \tag{2.14}$$

Notice that Kruskal–Szekeres coordinates are similiar to Rindler coordinates in special relativity, in the sense that an observer in Kruskal coordinates is in freefall towards the black hole.

Having the metric in the form (2.14) now allows us to define a final useful set of coordinates: *conformal coordinates* $\bar{T}$ and $\bar{X}$. A conformal mapping is a transformation

$$g_{\mu\nu} \mapsto (c(x))^2 g_{\mu\nu}, \tag{2.15}$$

such that $c(x)$ is a smooth function of the coordinates. Under a conformal transformation, the space-/time-/light-like property of a given interval remains invariant. The conformal coordinates are implicitly defined by

$$U = \tan\tilde{U}; \quad V = \tan\tilde{V}, \tag{2.16}$$

such that

$$\mathrm{d}U = \frac{\mathrm{d}\tilde{U}}{\cos^2\tilde{U}}; \quad \mathrm{d}V = \frac{\mathrm{d}\tilde{V}}{\cos^2\tilde{V}} \tag{2.17}$$

and

$$\mathrm{d}s^2 = \frac{1}{\cos^2\tilde{U}\cos^2\tilde{V}}\frac{2M}{r}\exp\left(\frac{-r}{2M}\right)\mathrm{d}U\mathrm{d}V - r^2\mathrm{d}\Omega^2. \tag{2.18}$$

Then we set $c(\tilde{U},\tilde{V}) = \cos\tilde{U}\cos\tilde{V}$ in (2.15) to obtain a spacetime which is causally equivalent to the Schwarzschild one but contained in a finite box

$$\tilde{U} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad \tilde{V} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \tag{2.19}$$

as portrayed in Fig. 5.2.

This kind of representation of spacetime is called a Penrose diagram, and the timelike and spacelike coordinates for such a diagram are given by

$$\tilde{T} = \frac{1}{2}(\tilde{V} + \tilde{U}), \quad \tilde{X} = \frac{1}{2}(\tilde{V} - \tilde{U}) \tag{2.20}$$

which are also regular at the Schwarzschild radius.

2.2 The Kerr Metric

The Kerr solution is a non-spherically symmetrical solution to (2.2) describing a rotating black hole. It is given by

$$\begin{aligned} \mathrm{d}s^2 = &-\left(1-\frac{2Mr}{\rho^2}\right)\mathrm{d}t^2 - \frac{2Mar\sin^2\theta}{\rho^2}\mathrm{d}t\mathrm{d}\phi + \frac{\rho^2}{\Delta}\mathrm{d}r^2 \\ &+ \frac{\sin^2\theta}{\rho^2}\left[\left(r^2+a^2\right)^2 - \Delta a^2\sin^2\theta\right]\mathrm{d}\phi^2 + \rho^2\mathrm{d}\theta^2, \end{aligned} \tag{2.21}$$

where $\Delta = r^2 - 2Mr + a^2$ and $\rho^2 = r^2 + a^2\cos^2\theta$. The parameter a can be shown to correspond to the angular momentum per unit mass, that is, $a = \frac{J}{M}$.

2.3 The Reissner–Nordström Metric

When electromagnetic interactions are relevant, one instead looks at the Einstein-Maxwell equation, where the energy–momentum distribution is given by the electromagnetic field due to the point charge, described by

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi\left(F_{\mu}^{\rho}F_{\nu\rho} - \frac{1}{4}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma}\right), \tag{2.22}$$

where $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$ is the Faraday tensor.

The analog of the Schwarzschild metric for the Einstein-Maxwell equation is the Reissner-Nordström metric, given by

$$\mathrm{d}s^2 = \left(1-\frac{2M}{r}+\frac{Q^2}{r^2}\right)\mathrm{d}t^2 - \left(1-\frac{2M}{r}+\frac{Q^2}{r^2}\right)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2, \tag{2.23}$$

where Q is the charge of the black hole.

2.4 The Kerr–Newman Metric

The metric for a black hole that is both charged and rotating is given by generalizing (2.24) to include charge, given by

$$\begin{aligned} \mathrm{d}s^2 = &-\left(1-\frac{2Mr-Q^2}{\rho^2}\right)\mathrm{d}t^2 - \frac{2Mar\sin^2\theta}{\rho^2}\mathrm{d}t\mathrm{d}\phi + \frac{\rho^2}{\Delta}\mathrm{d}r^2 \\ &+ \frac{\sin^2\theta}{\rho^2}\left[\rho^2\left(r^2+a^2\right) - (2Mr-Q^2)a^2\sin^2\theta\right]\mathrm{d}\phi^2 + \rho^2\mathrm{d}\theta^2, \end{aligned} \tag{2.24}$$

where we have set the discriminant to $\Delta \mapsto \Delta = r^2 - 2Mr + a^2 + Q^2$.

The roots of the discriminant are given by $\Delta = (r - r_+)(r - r_-)$, where

$$r_\pm = M \pm \sqrt{M^2 - a^2 - Q^2}, \tag{2.25}$$

which are as singularities in the coordinates used in (2.24), and are event horizons analogous to r_s in the Schwarzschild case.

The Kerr–Newman metric is the most general black hole metric we have discussed. While Birkhoff's theorem states that the Schwarzschild metric is the unique spherically symmetric solution to the vacuum field equations, there is a similar conjecture for the Einstein-Maxwell equations: the so-called 'no-hair' conjecture states that classical black holes have only three degrees of freedom: mass M, angular momentum J, and charge Q [26, 27].

A star has many more degrees of freedom that this, which are apparently lost when the star collapses into a black hole: this means the initial state of a collapsing system cannot be retrodicted from the final state. This already represents a classical toy model of the information loss paradox, which will be explored later.

3 | Black Holes as Thermodynamical Objects

3.1 Killing Vectors and Conserved Quantities

To properly make sense of quantities like energy, and to look at the properties of a surface like an event horizon, in a Schwarzschild spacetime, it is useful to examine the spacetime's Killing vectors and Killing *horizons*. A vector ξ_μ is a *Killing vector* if the Lie derivative of the metric with respect to ξ_μ vanishes, *i.e.*

$$\mathcal{L}_\xi g_{\mu\nu} = 0, \tag{3.1}$$

where the Lie derivative along a vector field ξ acts on a tensor η of rank (k, l) as

$$\mathcal{L}_\xi \eta^{i_1,\ldots,i_k}_{j_1,\ldots j_l}(p) = \left[\frac{\mathrm{d}}{\mathrm{d}\lambda}(F_\lambda \eta)^{i_1,\ldots,i_k}_{j_1,\ldots j_l}\right]_{\lambda=0}. \tag{3.2}$$

Here, F_λ is a one-parameter family of diffeomorphisms generated by ξ. The Lie derivative is thus a generalization of the directional derivative, in the sense that it measures the rate of change of a tensor η along the flow of a vector field ξ. The equation (3.1) implies F_λ is an isometry of the spacetime defined by the metric g. Then the Killing vectors correspond to symmetries of the spacetime: we can thus construct quantities

$$\mathcal{Q} = \xi_\mu p^\mu \tag{3.3}$$

such that $\mathcal{Q}$ is conserved along geodesics, where p^μ is the four-momentum.

The quantities $\mathcal{Q}$ are those found by varying the action functional of a free particle along a geodesic, given by

$$S^{\text{geo}}[x^\mu] = m \int \sqrt{-g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}}, \tag{3.4}$$

where λ is an affine parameter. Notably, these are *not* the conserved quantities associated with the

Einstein-Hilbert action (see 6.1), but it will be shown later that there *are* conserved quantities associated with (6.1) which relate to the Killing vectors in a special way.

A Killing horizon is a hypersurface Σ to which a Killing vector ξ is normal at the surface. Then on an event horizon we have

$$\xi^\mu \nabla_\mu \xi^\nu = \kappa \xi^\nu, \tag{3.5}$$

where κ is the surface gravity, which is the acceleration needed for a observer to stay at the horizon as measured by an observer at infinity.

In particular, (3.1) implies

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0, \tag{3.6}$$

which is called the Killing equation.

We can compute the surface gravity using the above two equations by computing the derivative of the magnitude squared of the Killing vector ξ^μ as

$$\begin{aligned}
\nabla_\rho(-\xi_\mu \xi^\mu) &= -2\xi^\mu \nabla_\rho \xi_\mu \\
&= 2\xi^\mu \nabla_\mu \xi_\rho \quad \text{(by (3.6))} \\
&= 2\kappa \xi_\rho.
\end{aligned} \tag{3.7}$$

The Kerr(–Newman) and metric has in particular a timelike killing vector

$$\xi_t^\mu = (\partial_t)^\mu = (1, 0, 0, 0), \tag{3.8}$$

and rotational Killing vector

$$\xi_\phi^\mu = (\partial_\phi)^\mu = (0, 0, 0, 1). \tag{3.9}$$

At the event horizon, the timelike Killing vector is null. Its magnitude-squared is thus

$$\begin{aligned}
(\partial_t)_\mu (\partial_t)^\mu &= g_{\mu\nu} (\partial_t)^\mu (\partial_t)^\nu \\
&= g_{tt}((\partial_t)^t)^2 \\
&= -\left(1 - \frac{1}{2M}\right) \times 1 \times 1 \\
&= -\left(1 - \frac{1}{2M}\right).
\end{aligned}$$

Then the surface gravity is obtained by first switching to Eddington-Finkelstein coordinates, where it is well-defined at infinity, in which case we have

$$\partial_t = \frac{\partial r}{\partial t}\partial_r + \frac{\partial v}{\partial t}\partial_v = \partial_v, \tag{3.10}$$

and

$$\xi_r = g_{rv}\xi^v = 1, \tag{3.11}$$

and thus, using (3.7), we have

$$\begin{aligned} 2\kappa(\partial_t)_r &= \nabla_r\left(1 - \frac{2M}{r}\right) \\ \implies 2\kappa|_{r=2M} &= \frac{2M}{(2M)^2} \\ \implies \kappa &= \frac{1}{2M}. \end{aligned}$$

In SI units,

$$r_s = \frac{2GM}{c^2}, \tag{3.12}$$

and thus the surface gravity in Schwarzschild spacetime is

$$\kappa = \frac{c^4}{4GM}. \tag{3.13}$$

3.2 The Penrose Process

While black holes are commonly thought of as regions where "nothing, not even light, can escape" [10], this is strictly meant to illustrate the key point that, past the event horizon at $r = 2M$, spacelike geodesics become timelike, and all bodies are compelled toward the centre of the black hole: energy is not precluded from extraction. In particular, Penrose [8, 28] describes a mechanism for extracting rotational energy from a rotating black hole by taking advantage of the fact that objects in a certain region outside the event horizon – called the ergosphere – can be given negative energy [2].

The ergosphere is a region in Kerr(–Newman) spacetime in which the magnitude-squared of the (asymptotically) timelike Killing vector $\partial_t^2 = -g_{tt}$ is negative, *i.e.*

$$g_{tt} = -\left(1 - \frac{2Mr}{\rho^2}\right) > 0 \iff \rho^2 - 2Mr < 0, \tag{3.14}$$

which defines the region $M + \sqrt{M^2 - a^2} < r <= M + \sqrt{M^2 - a^2\cos^2\theta}$.

In this region the magnitude-squared of the tangent vector u^μ to a curve x^μ given by

$$u^2 = g_{\mu\nu} u^\mu u^\nu, \tag{3.15}$$

which is negative (thus timelike) only for $u^\phi \neq 0$. Thus the ergosphere defines a region in which it is impossible to remain stationary to observers at infinity. This effect is known as frame-dragging. Note from the section on Killing vectors that the conserved quantity associated with the timelike Killing vector is the asymptotic energy E – to identify it as such, notice that at the Minowski limit $r \to \infty$, we should have $E = p^0$ and $g_{tt} = -1, g_{rr} = 0$, which implies

$$\lim_{r \to \infty} E = p^0 \tag{3.16}$$

where p^μ is the four-momentum.

The Penrose process is then rather simple: outside a black hole we may throw a particle of energy E in such a way that it decays into two smaller particles of energies E_+ and E_- during its time in the ergosphere. By (3.16) it is possible for E_- to be negative, and thus, by conservation of four-momentum, we can have $E_+ > E$, *i.e.* we have extracted some energy from the black hole by first sending a smaller amount in.

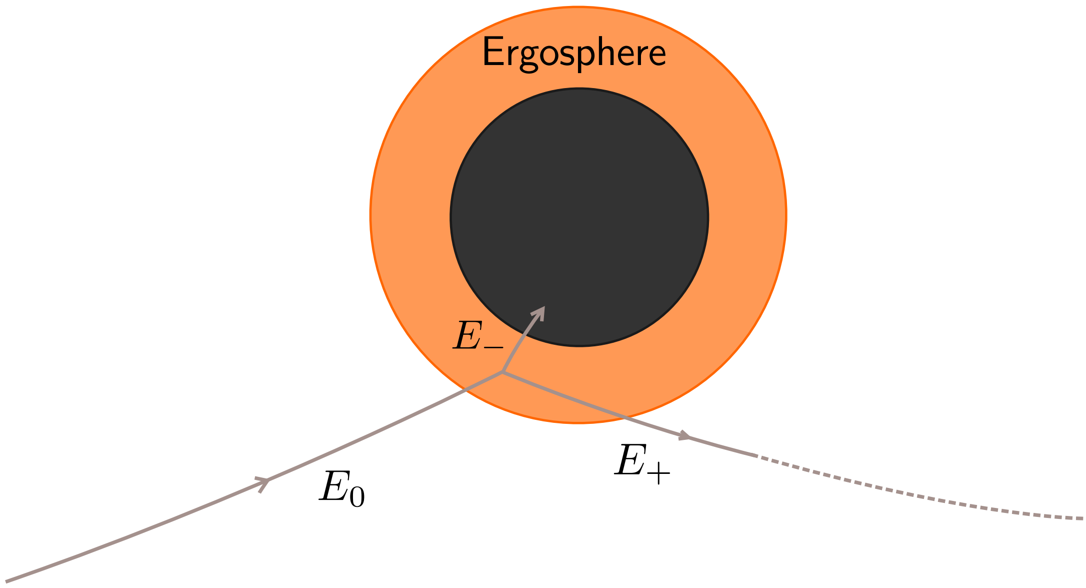

Figure 3.1: The Penrose process. A particle of energy E_0 enters the ergosphere and decays into particles of energy E_+ and E_-. The latter particle has negative energy in the ergosphere, so by conservation of energy, the former particle leaves with more energy than the initial E_0. Partially adapted from [8]

3.3 The Area Theorem and Irreversible Processes for Classical Black Holes

In this example, however, rotational energy extraction can be performed only up to a certain point, at which the Kerr(–Newman) black hole ceases to rotate and degenerates to a Schwarzschild black hole with mass m_{ir}, a quantity known as the *irreducible mass* of the original black hole. The irreducible mass m_{ir} is defined in relation to the surface area A of the hole's event horizon by

$$A = 16\pi m_{\mathrm{ir}}^2, \tag{3.17}$$

since we can then show that A (and thus m_{ir}) cannot be reduced in any process, and stays the same only for a select set of processes called *reversible processes* [12]. This is known as the area theorem. It is easy to see that m_{ir} is a mass quantity corresponding to that of a Schwarzschild black hole by recognising that, for a Schwarzschild black hole,

$$\begin{aligned} A_s &= 4\pi r_s^2 \\ &= 4\pi(2M)^2 \\ &= 16\pi M^2. \end{aligned}$$

This quantity m_{ir} and the notion of reversible and irreversible processes in the context of black hole mechanics was first introduced by Christodoulou [12]. First note that a Killing horizon is a surface on which a Killing vector is normal to it. A Kerr(–Newman) has timelike killing vector ∂_t and azimuthal Killing vector ∂_ϕ, and the Killing horizon χ^μ can then easily be given by the surface

$$\chi^\mu = (\partial_t)^\mu + \Omega(\partial_\phi)^\mu, \tag{3.18}$$

where $\Omega = \mathrm{d}\phi/\mathrm{d}t|_{r_+}$ is the angular velocity of the event horizon. Properly, we say that (3.18) *generates* the event horizon, by which we mean it is null at the event horizon.

To define Ω properly we recall the notion of frame-dragging defined above and consider a particle following along the curve $\theta = 0$ on the event horizon. The velocity four-vector is then zero in the r and θ components and thus

$$u^2 = g_{\mu\nu}u^\mu u^\nu = 0 \tag{3.19}$$

is satisfied when

$$g_{tt}\mathrm{d}t^2 + 2g_{t\phi}\,(\mathrm{d}t\mathrm{d}\phi) + g_{\phi\phi}\mathrm{d}\phi^2 = 0, \tag{3.20}$$

which for the general Kerr–Newman case is given by

$$
\begin{aligned}
\mathrm{d}s^2 &= -\left(1 - \frac{2Mr}{\rho^2}\right)\mathrm{d}t^2 - \frac{2Mar\sin^2\theta}{\rho^2}\mathrm{d}t\mathrm{d}\phi \\
&\quad + \frac{\sin^2\theta}{\rho^2}\left[\left(r^2+a^2\right)^2 - \Delta a^2\sin^2\theta\right]\mathrm{d}\phi^2 \\
&= -\left(\frac{\Delta - a^2\sin^2\theta}{\rho^2}\right)\mathrm{d}t^2 - \frac{2a\sin^2\theta}{\rho^2}\left(r^2+a^2-\Delta\right)\mathrm{d}t\mathrm{d}\phi \\
&\quad + \frac{\sin^2\theta}{\rho^2}\left[\left(r^2+a^2\right)^2 - \Delta a^2\sin^2\theta\right]\mathrm{d}\phi^2 = 0.
\end{aligned}
$$

Substituting this into 3.20, rearranging, and evaluating at r_+ we have (remembering that r_+ is a root of the discriminant Δ)

$$
\begin{aligned}
\Omega &= \frac{\mathrm{d}\phi}{\mathrm{d}t}\Big|_{r=r_+} \\
&= -\frac{g_{t\phi}}{g_{\phi\phi}} \pm \sqrt{\left(\frac{g_{t\phi}}{g_{\phi\phi}}\right)^2 - \frac{g_{tt}}{g_{\phi\phi}}} \\
&= \frac{a\left(r_+^2+a^2\right)}{\left(r_+^2+a^2\right)^2} \pm \sqrt{\left(\frac{a\left(r_+^2+a^2\right)}{\left(r_+^2+a^2\right)^2}\right)^2 - \frac{a^2\left(r_+^2+a^2\right)^2}{\left(r_+^2+a^2\right)^4}} \\
&= \frac{a}{r_+^2+a^2}.
\end{aligned}
$$

We can use this to examine how the negative-energy (E_-) particle 2 'takes' energy from the black hole and 'gives' it to the positive-energy (E_+) particle 1. When particle 2 passes the event horizon, the particle's direction with respect to the Killing horizon is timelike, that is

$$p_2^\mu\left(\partial_t\right)_\mu + \Omega p_2^\mu\left(\partial_\phi\right)_\mu > 0. \tag{3.21}$$

But $\left(\partial_t\right)_\mu p^\mu$ and $\left(\partial_\phi\right)_\mu p^\mu$ are just the conserved quantities associated with their respective Killing vectors, and can be straightforwardly identified with the classical analogues of energy E and angular momenum J, so we effectively have

$$-E_- + \Omega J_- > 0, \tag{3.22}$$

where J_- is the angular momentum of the particle. This is intuitive: the particle robs the black hole of a certain amount of angular momentum, proportional to the event horizon's angular velocity.

This is the first hint that black holes can radiate. In fact, [13] considers the scattering of cylindrical electromagnetic waves of sufficiently low energy $E = \omega$ and angular momentum J by a rotating black hole within the ergosphere. The waves are redshifted by an added frequency $J\Omega$. To conserve the total energy flux, the waves are thus amplified. A small electromagnetic perturbation in the ergosphere should then amplify into electromagnetic radiation which carries away the black hole's angular momentum.

Then (at least classically) energy cannot then be robbed from a nonrotating black hole. We can also show that, correspondingly, no amount of energy loss will decrease the area of any black hole. To see this, we first explicitly compute the area of the outer horizon r_+. The induced metric on the horizon is (2.24) evaluated at the hypersurface Σ_+ defined by $r = r_+$, $\mathrm{d}t = \mathrm{d}r = 0$, whose only two nonzero elements are

$$g_{\theta\theta} = \rho^2|_{r_+}, \quad g_{\phi\phi} = \frac{\sin^2\theta}{\rho^2|_{r_+}}\left(r_+^2 + a^2\right)^2. \tag{3.23}$$

The volume element of an arbitrary n-dimensional space is $\sqrt{g}\,\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n$. On Σ_+ this is just $\sqrt{g_{\theta\theta}g_{\phi\phi}}\mathrm{d}\theta\mathrm{d}\phi$, so the area is just

$$A = \int \sin^2\theta\left(r_+^2 + a^2\right)\mathrm{d}\theta\mathrm{d}\phi = 4\pi\left(r_+^2 + a^2\right). \tag{3.24}$$

We rewrite the reduced mass in terms of r_+ and a as

$$\begin{aligned} m_{\mathrm{ir}}^2 &= \frac{A}{16\pi} \\ \implies m_{\mathrm{ir}} &= \frac{1}{2}\left(r_+^2 + a^2\right)^{1/2} \\ &= \frac{1}{2}\sqrt{\alpha} \end{aligned}$$

where

$$\alpha = A/4\pi = r_+^2 + a^2 \tag{3.25}$$

is the so-called "rationalized area". A change in energy for the black hole is equivalent to a change in mass M, so to see the effect of a small change δE on the black hole we can differentiate with respect to M to obtain

$$\frac{\mathrm{d}m_{\mathrm{ir}}}{\mathrm{d}M} = \frac{1}{4} \times \frac{1}{\sqrt{\alpha}} \times \left(\frac{\mathrm{d}\alpha}{\mathrm{d}M}\right).$$

Then we have

$$\frac{\mathrm{d}\alpha}{\mathrm{d}M} = 2\left(M\frac{\mathrm{d}r_+}{\mathrm{d}M} + r_+\frac{\mathrm{d}M}{\mathrm{d}M}\right) - 2Q\frac{\mathrm{d}Q}{\mathrm{d}M}. \tag{3.26}$$

The term $\mathrm{d}M/\mathrm{d}M = 1$ trivially. The derivative of the outer horizon r_+ with respect to M is

$$\begin{aligned} \frac{\mathrm{d}r_+}{\mathrm{d}M} &= \frac{\mathrm{d}}{\mathrm{d}M}\left(M + \left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}\right) \\ &= 1 + \frac{2M - 2Q\frac{\mathrm{d}Q}{\mathrm{d}M} - 2\mathbf{a}\cdot\frac{\mathrm{d}\mathbf{a}}{\mathrm{d}M}}{2\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}, \end{aligned} \tag{3.27}$$

and the derivative of the angular momentum per unit mass $\mathbf{a}$ is

$$\frac{\mathrm{d}\mathbf{a}}{\mathrm{d}M} = \frac{\mathrm{d}\left(\mathbf{J}/M\right)}{\mathrm{d}M} = \mathbf{J}\frac{\mathrm{d}\left(1/M\right)}{\mathrm{d}M} - \frac{1}{M}\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M} = -\frac{\mathbf{J}}{M^2} + \frac{1}{M}\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M}, \tag{3.28}$$

so the full derivative of the rationalized area is

$$\begin{aligned}\frac{\mathrm{d}\alpha}{\mathrm{d}M} &= 2M + \frac{2M^2 - 2MQ\frac{\mathrm{d}Q}{\mathrm{d}M} - 2\mathbf{a}\cdot\left(\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M} - \frac{\mathbf{J}}{M}\right)}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}} + 2r_+ - 2Q\frac{\mathrm{d}Q}{\mathrm{d}M}\\
&= \left[2M + 2r_+ + \frac{2M^2 + 2a^2}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\right] - \left[\frac{2MQ}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}} + 2Q\right]\frac{\mathrm{d}Q}{\mathrm{d}M}\\
&- \frac{2\mathbf{a}}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\cdot\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M}.\end{aligned} \tag{3.29}$$

Rewriting more neatly, we have

$$\begin{aligned}\frac{\mathrm{d}\alpha}{\mathrm{d}M} = &\left[\frac{\left(M + r_+\right)\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}} + 2M^2 + 2a^2}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\right]\\
&- \left[\frac{2MQ + 2Q\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\right]\frac{\mathrm{d}Q}{\mathrm{d}M}\\
&- \frac{2\mathbf{a}}{\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M}.\end{aligned}$$

Finally we obtain

$$\begin{aligned}\frac{\mathrm{d}m_{\mathrm{ir}}}{\mathrm{d}M} = &\left[\frac{\left(M + r_+\right)\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}} + 2M^2 + 2a^2}{4\alpha^{1/2}\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\right]\\
&- \left[\frac{2MQ + 2Q\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}{4\alpha^{1/2}\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\right]\frac{\mathrm{d}Q}{\mathrm{d}M}\\
&- \frac{2\mathbf{a}}{4\alpha^{1/2}\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}M}.\end{aligned} \tag{3.30}$$

For illustrative purposes we can assume that the black hole does not gain any charge during this process, so $\mathrm{d}Q/\mathrm{d}M = 0$. Then multiplying through by $\mathrm{d}M$ we obtain

$$\mathrm{d}m_{\mathrm{ir}} = \frac{2\mathbf{a}}{4\alpha^{1/2}\left(M^2 - Q^2 - a^2\right)^{1/2}}\left(-\mathrm{d}\mathbf{J} + \frac{r_+^2 + a^2}{a}\mathrm{d}M\right). \tag{3.31}$$

Using the definition of Ω and eq. 3.22, we see that this implies $\mathrm{d}m_{\mathrm{ir}} \geq 0$, which in turn implies $\mathrm{d}A \geq 0$. This defines the notion of irreversibility for a black hole.

3.4 The Black Hole 'Thermodynamic Equation'

The rule $\Delta A \geq 0$ bears resemblance to the second law of thermodynamics, $\Delta S \geq 0$, with S being entropy. Bekenstein formalized this analogy [3] by providing the expression

$$\mathrm{d}M = \Theta \mathrm{d}\alpha + \mathbf{\Omega} \cdot \mathrm{d}\mathbf{L} + \Phi \mathrm{d}Q, \tag{3.32}$$

which resembles the fundamental equation of thermodynamics, where

$$\alpha = \frac{A}{4\pi} \tag{3.33}$$

is again the rationalized area, and

$$\Theta = \frac{1}{4}\frac{(r_+ - r_-)}{\alpha}, \tag{3.34}$$

$$\mathbf{\Omega} = \frac{\mathbf{a}}{\alpha}, \tag{3.35}$$

$$\Phi = \frac{Qr_+}{\alpha}. \tag{3.36}$$

In particular, Θ and α take on the roles of temperature and entropy, respectively. This can be seen by looking at (3.30), and multiplying through by $\mathrm{d}M$, so that we have

$$\begin{aligned} \mathrm{d}\alpha = &\left[\frac{(M + r_+)\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}} + 2M^2 + 2a^2}{(M^2 - Q^2 - a^2)^{\frac{1}{2}}}\right]\mathrm{d}M \\ &- \left[\frac{2MQ + 2Q\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}}{(M^2 - Q^2 - a^2)^{\frac{1}{2}}}\right]\mathrm{d}Q \\ &- \frac{2\mathbf{a}}{(M^2 - Q^2 - a^2)^{\frac{1}{2}}}\mathrm{d}\mathbf{L}. \end{aligned} \tag{3.37}$$

Then, noting that $r_+ - r_- = 2\left(M^2 - Q^2 - a^2\right)^{1/2}$, we obtain

$$\begin{aligned} \frac{1}{4}(r_+ - r_-)\,\mathrm{d}\alpha &= \Big(M\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}} + M\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}} \\ &\quad + \left(M^2 - Q^2 - a^2\right) + M^2 + a^2\Big)\mathrm{d}M \\ &\quad - \left(MQ + 2Q\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}\right)\mathrm{d}Q - \mathbf{a}\cdot\mathrm{d}\mathbf{L} \\ &= \left(2Mr_+ - Q^2\right)\mathrm{d}M - (Qr_+)\,\mathrm{d}Q - \mathbf{a}\cdot\mathrm{d}\mathbf{L}. \end{aligned} \tag{3.38}$$

Identifying $2Mr_+ - Q^2$ as α, we recover (3.32).

3.5 The Four Laws

Bardeen, Carter, *et. al.*, rederived (3.32), identifying the quantity $\Theta = \kappa/2$, where κ is the surface gravity of the black hole [4].

To see how this is true for the Schwarzschild case (which we will be dealing with henceforth), take $a \to 0$ and $Q \to 0$, and note that $r_+ - r_- = 2\left(M^2 - Q^2 - a^2\right)^{\frac{1}{2}}$. Then

$$\Theta = \frac{1}{4}\frac{2\sqrt{M^2}}{A/(4\pi)} = \frac{2M\pi}{4\pi(2M)^2} = \frac{1}{2}\kappa. \tag{3.39}$$

Now we can absorb a factor of π in the $T\mathrm{d}S$-like term in (3.32) so that it becomes

$$\Theta \mathrm{d}\alpha = \frac{1}{2}\kappa \mathrm{d}\frac{A}{4\pi} = \frac{\kappa}{2\pi}\mathrm{d}\frac{A}{4}, \tag{3.40}$$

thus recovering the form (1.1).

They [4] also derived a set of four laws for the mechanics of black holes, analogous to the laws of thermodynamics. The zeroth law holds that the surface gravity is constant across the event horizon. This follows from the definition. The first law is just (3.32), and the second law is the area theorem.

All of these laws are either rigorously proven or true by definition, except the third law. It states that there is no finite sequence of operations which can be used to reduce the surface gravity κ to zero. This is important to keep in mind when considering the extent to which a black hole is a thermodynamical object.

4 | Hawking Radiation

4.1 The 'Temperature' of a Black Hole?

The laws described above were originally considered as simply a neat analogy, and the authors of [4] stressed that neither temperature nor entropy could actually be well-defined quantities for a black hole, based on the argument that, in order for a black hole to come into thermal equilibrium with its surroundings, it would need to be able to radiate. While it is mentioned above that an electromagnetically perturbed rotating black hole will radiate out all of its rotational energy, this does not apply to nonrotating classical black holes.

Bekenstein's paper [3] makes an information-theoretic argument for why entropy for nonrotating black holes might still be well-defined, however, suggesting that the surface area the black hole gains when a particle is dropped past the event horizon corresponds to an increase in entropy that reflects the loss of information outside observers have as to the state of the particle – namely whether it still exists, or is destroyed.

Eventually, Hawking proved that black holes actually *do* radiate in an approximate quantum treatment of black holes [5].

4.2 Quantum Field Theory in Curved Spacetime

Naïve attempts to insert gravity into quantum field theory by simply treating the metric as an ordinary tensor field and quantizing it fail, since gravitation is not renormalizable within QFT. But if we take from general relativity only the conclusion that space may have a curvature, and disregard for a moment the relationship of this curvature to the force of gravity, it is possible to formulate quantum field theory with a curved spacetime background. This formalism is by no means a theory of quantum gravity – it should break down at the scale where both highly quantum and highly general-relativistic effects should be taken into account (the *Planck scale*) – but at the event horizon of a black hole it is a good approximation.

Recall that in flat spacetime the Lagrangian density for a free scalar field ϕ is given by the Klein-Gordon action

$$\mathcal{L}^{\text{KG}} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi - \frac{1}{2}m^2\phi^2, \tag{4.1}$$

whose equations of motion are the Klein-Gordon equation

$$\left(\partial^\mu\partial_\mu + m^2\right)\phi = 0. \tag{4.2}$$

Normally this is quantized by imposing the constraints

$$[\phi(\tau,x),\phi(\tau',x')] = [\pi(\tau,x),\pi(\tau',x')] = 0, \quad [\phi(\tau,x),\pi(\tau',x')] = i\delta(x-x'), \tag{4.3}$$

and expanding the field in terms of creation and annihilation operators, yielding

$$\phi(\mathbf{x},t) = \int \frac{\mathrm{d}^d p}{(2\pi)^d}\frac{1}{\sqrt{2E_\mathbf{p}}}\left(a_\mathbf{p}\,\mathrm{e}^{ip_\mu x^\mu} + a^\dagger_\mathbf{p}\,\mathrm{e}^{-ip_\mu x^\mu}\right)|_{p^0=E_\mathbf{p}}, \tag{4.4}$$

where d is the number of spatial coordinates of the system.

We can similarly write down the Lagrangian density of a scalar field ϕ of mass m in curved spacetime of dimension $n = d+1$ by factoring in the square root of the metric (to ensure that the volume form is manifestly covariant) and replacing the partial derivative with the covariant derivative [29], obtaining

$$\mathcal{L}^{\text{CST}} = \sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\phi\,\nabla_\nu\phi - \frac{1}{2}m^2\phi^2 - \xi R\phi^2\right), \tag{4.5}$$

where $g = \det g_{\mu\nu}$ and where we have added ξ as an extra coupling constant to the usual Ricci curvature R.

In what follows, we need only look at massless particles like photons. In addition, we assume that we are not going to be dealing with quantum corrections to gravity, so we would not expect to have to use the coupling of the scalar field to the Ricci curvature. We can thus look at the simple case of a massless ($m = 0$), minimally coupled ($\xi = 0$) scalar field in $1+1$ dimensions, where the Schwarzschild metric (2.4) reduces to

$$\mathrm{d}s^2 = \left(1-\frac{2M}{r}\right)\mathrm{d}t^2 - \left(1-\frac{2M}{r}\right)^{-1}\mathrm{d}r^2, \tag{4.6}$$

i.e. the solid-angle terms disappear and our definitions (2.5) of Kruskal coordinates and (2.7) of the 'tortoise' coordinate remain unchanged. Since the Schwarzschild metric is spherically symmetric it is easy to generalise to higher dimensions after the fact. The Lagrangian (4.5) reduces to the familiar Klein-Gordon Lagrangian (multiplied by the volume element $\sqrt{-g}$) whose equations of motion are simply a covariant version of the Klein-Gordon equation for a massless particle,

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi = 0. \tag{4.7}$$

This is quantized in the usual way by introducing the conjugate momentum $\pi = \partial\mathcal{L}/\partial\dot{\phi}$ and asserting the commutation relations (4.3) for timelike coordinate τ and spacelike coordinate x. The Dirac delta function is extended to be defined as

$$\int \delta(x - x')d\Sigma = 1, \tag{4.8}$$

where Σ is a spacelike hypersurface in the given coordinate system.

So far this is almost identical to the standard quantization of a scalar field in flat spacetime [30, 31]. Unlike the flat case, however, the expansion of the scalar field ϕ in terms of creation/annihilation operators $a_p, a_p^\dagger$ for momentum p satisfying the relations

$$[a_p^\dagger, a_{p'}] = \delta(p - p') \tag{4.9}$$

is non-unique. In particular, we can expand ϕ in the space-/time-like Kruskal coordinates T, X given by (2.13) to obtain

$$\phi(T, X) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi} \frac{1}{\sqrt{2E_p}} \left(a_p \, \mathrm{e}^{i(E_pT - pX)} + a_p^\dagger \, \mathrm{e}^{-i(E_pT - pX)} \right), \tag{4.10}$$

and similarly in tortoise coordinates t, r^* to obtain

$$\phi(t, r^*) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}p}{2\pi} \frac{1}{\sqrt{2E_p}} \left(b_p \, \mathrm{e}^{i(E_pt - pr^*)} + b_p^\dagger \, \mathrm{e}^{-i(E_pt - pr^*)} \right), \tag{4.11}$$

with $E_p = |p|$ for massless particles [31]. The expansion (4.10) can be seen to represent the mode expansion for an infalling observer, while (4.11) can be considered the mode expansion for a static observer away from the event horizon.

There is no *a priori* reason that a_p would be equal to b_p, and indeed it will be shown that this is not the case.

4.3 The Bogoliubov Transformation

The general relation between operators a_p and b_p satisfying the relations (4.9) is given by a *Bogoliubov transformation* [32],

$$\begin{aligned} b_q &= \int_0^\infty \mathrm{d}p(\alpha_{pq}a_p + \beta_{pq}a_p^\dagger) \\ b_q^\dagger &= \int_0^\infty \mathrm{d}p(\alpha_{pq}^* a_p^\dagger + \beta_{pq}^* a_p), \end{aligned} \tag{4.12}$$

where $\alpha_{pq}, \alpha_{pq}^\dagger$ and $\beta_{pq}, \beta_{pq}^\dagger$ are the Bogoliubov coefficients, which satisfy the relations

$$\begin{aligned} &\int_{-\infty}^\infty \mathrm{d}p(\alpha_{pq}\alpha_{pq'}^* - \beta_{pq}\beta_{pq'}^*) = \delta(q-q'), \\ &\int_{-\infty}^\infty \mathrm{d}p(\alpha_{pq}\beta_{pq'} - \alpha_{pq}\beta_{pq'}) = 0. \end{aligned} \tag{4.13}$$

The coefficients can be read off in integral form by examining equations (4.10) and (4.11) and inserting the expressions (4.12). First we transform (4.10) into the light-cone Kruskal coordinates U, V given by (2.5) and (4.11) into Eddington-Finkelstein coordinates u, v given by (2.7) and use the fact that $E_p = |p|$ so that we can use

$$E_p t - pr^* = \begin{cases} |p|t + |p|r* = E_p v \text{ for } p < 0 \\ |p|t - |p|r^* = E_p u \text{ for } p > 0, \end{cases} \tag{4.14}$$

and similarly

$$E_p T - pX = \begin{cases} |p|T + |p|X = E_p V \text{ for } p < 0 \\ |p|T - |p|X = E_p U \text{ for } p > 0. \end{cases} \tag{4.15}$$

Then in lightcone coordinates, we have

$$\phi(U,V) = \int_0^\infty \frac{\mathrm{d}E_p}{\sqrt{2\pi}} \frac{1}{\sqrt{2E_p}} \left(\mathrm{e}^{iE_pU}\, a_{E_p} + \mathrm{e}^{-iE_pU}\, a_{E_p}^\dagger + \mathrm{e}^{iE_pV}\, a_{-E_p} + \mathrm{e}^{-iE_pV}\, a_{-E_p}^\dagger \right), \tag{4.16}$$

for the Kruskal-Szekeres observer, and

$$\phi(u,v) = \int_0^\infty \frac{\mathrm{d}E_q}{\sqrt{2\pi}} \frac{1}{\sqrt{2E_q}} \left(\mathrm{e}^{iE_qu}\, b_{E_q} + \mathrm{e}^{-iE_qu}\, b_{E_q}^\dagger + \mathrm{e}^{iE_qv}\, b_{-E_q} + \mathrm{e}^{-iE_qv}\, b_{-E_q}^\dagger \right), \tag{4.17}$$

for Eddington-Finkelstein coordinates. Equating the two expressions and using (4.12), we have

$$\int_0^\infty \frac{\mathrm{d}E_p}{\sqrt{E_p}}\left(\mathrm{e}^{iE_pU}\,a_{E_p}+\mathrm{e}^{-iE_pU}\,a^\dagger_{E_p}+\mathrm{e}^{iE_pV}\,a_{-E_p}+\mathrm{e}^{-iE_pV}\,a^\dagger_{-E_p}\right)=$$
$$\int_0^\infty \frac{\mathrm{d}E_q}{\sqrt{E_q}}\left(\int_0^\infty \mathrm{d}E_p\left(\alpha_{E_pE_q}a_{E_p}+\beta_{E_pE_q}a^\dagger_{E_p}\right)\mathrm{e}^{iE_qu}+\right.$$
$$\int_0^\infty \mathrm{d}E_p\left(\alpha^*_{E_pE_q}a^\dagger_{E_p}+\beta^*_{E_pE_q}a_{E_p}\right)\mathrm{e}^{-iE_qu}+$$
$$\int_0^\infty \mathrm{d}E_p\left(\alpha_{-E_p-E_q}a_{-E_p}+\beta_{-E_p-E_q}a^\dagger_{-E_p}\right)\mathrm{e}^{iE_qv}+$$
$$\left.\int_0^\infty \mathrm{d}E_p\left(\alpha^*_{-E_p-E_q}a^\dagger_{-E_p}+\beta^*_{-E_p-E_q}a_{-E_p}\right)\mathrm{e}^{-iE_qv}\right).$$

Then matching coefficients of a_{E_p} we have

$$\begin{aligned}\int_0^\infty \frac{\mathrm{d}E_p}{\sqrt{E_p}}\,\mathrm{e}^{iE_pU}&=\int_0^\infty\int_0^\infty \frac{\mathrm{d}E_q\mathrm{d}E_p}{\sqrt{E_q}}\left(\alpha_{E_pE_q}\,\mathrm{e}^{iE_qu}+\beta^*_{E_pE_q}\,\mathrm{e}^{-iE_qu}\right)\\ \implies \frac{1}{\sqrt{E_p}}\,\mathrm{e}^{iE_pU}&=\int_0^\infty \frac{\mathrm{d}E_q}{\sqrt{E_q}}\left(\alpha_{E_pE_q}\,\mathrm{e}^{iE_qu}+\beta^*_{E_pE_q}\,\mathrm{e}^{-iE_qu}\right).\end{aligned} \tag{4.18}$$

We can rewrite the coefficients as

$$\tilde{\beta}_{E_pE_q}=\begin{cases}\frac{\alpha_{E_pE_q}}{\sqrt{|E_q|}} \text{ if } E_q\le 0\\ \frac{\beta^*_{E_pE_q}}{\sqrt{E_q}} \text{ if } E_q>0,\end{cases} \tag{4.19}$$

so that (4.18) takes the form

$$\frac{\mathrm{e}^{iE_pU}}{\sqrt{E_p}}=\int_{-\infty}^\infty \mathrm{d}E_q\left(\tilde{\beta}_{E_pE_q}\,\mathrm{e}^{-iE_qu}\right), \tag{4.20}$$

where we recognise the left-hand side as the Fourier transform $\mathcal{F}(u)$ of the function $\tilde{\beta}_{E_pE_q}(E_q)$. Taking the inverse Fourier transform, we obtain

$$\begin{aligned}\tilde{\beta}_{E_pE_q}&=\frac{1}{2\pi}\int_{-\infty}^\infty \mathrm{d}u\frac{\mathrm{e}^{iE_pU}}{\sqrt{E_p}}\,\mathrm{e}^{iE_qu}\\ &=\begin{cases}\frac{1}{2\pi}\int_{-\infty}^\infty \mathrm{d}u\frac{\mathrm{e}^{iE_pU}}{\sqrt{E_p}}\,\mathrm{e}^{-iE_qu} \text{ if } E_q\le 0\\ \frac{1}{2\pi}\int_{-\infty}^\infty \mathrm{d}u\frac{\mathrm{e}^{iE_pU}}{\sqrt{E_p}}\,\mathrm{e}^{+iE_qu} \text{ if } E_q>0,\end{cases}\end{aligned}$$

that is,

$$\alpha_{E_pE_q}=\frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^\infty \mathrm{d}u\,\mathrm{e}^{i(E_pU-E_qu)}, \tag{4.21}$$

$$\beta^*_{E_pE_q}=\frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^\infty \mathrm{d}u\,\mathrm{e}^{i(E_pU+E_qu)}. \tag{4.22}$$

To solve these integrals, note the definition of the Kruskal coordinates in terms of the Eddington-

Finkelstein coordinates given by (2.5), so that

$$
\begin{aligned}
\alpha_{E_pE_q} &= \frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{i\left(-E_q u - 4ME_p \exp\left(-\frac{u}{4M}\right)\right)} \\
&= \frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{i\left(-E_q u - \frac{E_p}{\kappa} \exp\left(-\kappa u\right)\right)}, \\
\beta^*_{E_pE_q} &= \frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{i\left(E_q u - 4ME_p \exp\left(-\frac{u}{4M}\right)\right)} \\
&= \frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{i\left(E_q u - \frac{E_p}{\kappa} \exp\left(-\kappa u\right)\right)}.
\end{aligned}
$$

Making the substitution $v = \exp(-\kappa u)$, we have

$$
u = -\frac{\ln v}{\kappa}; \quad \mathrm{d}u = -\frac{\mathrm{d}v}{\kappa v}, \tag{4.23}
$$

and the integrals become

$$
\begin{aligned}
\alpha_{E_pE_q} &= -\frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{\infty}^{0} \frac{\mathrm{d}v}{\kappa v}\, v^{i\frac{E_q}{\kappa}}\, \mathrm{e}^{-i\frac{E_p}{\kappa}v} \\
&= \frac{1}{2\pi\kappa}\sqrt{\frac{E_q}{E_p}}\int_{0}^{\infty} \mathrm{d}v\, v^{i\frac{E_q}{\kappa}-1}\, \mathrm{e}^{-i\frac{E_p}{\kappa}v} \\
&= \frac{1}{2\pi\kappa}\sqrt{\frac{E_q}{E_p}}\frac{\Gamma\left(i\frac{E_q}{\kappa}\right)}{\left(i\frac{E_p}{\kappa}\right)^{i\frac{E_q}{\kappa}}}, \\
\beta^*_{E_pE_q} &= -\frac{1}{2\pi}\sqrt{\frac{E_q}{E_p}}\int_{\infty}^{0} \frac{\mathrm{d}v}{\kappa v}\, v^{-i\frac{E_q}{\kappa}}\, \mathrm{e}^{-i\frac{E_p}{\kappa}v} \\
&= \frac{1}{2\pi\kappa}\sqrt{\frac{E_q}{E_p}}\int_{0}^{\infty} \mathrm{d}v\, v^{-i\frac{E_q}{\kappa}-1}\, \mathrm{e}^{-i\frac{E_p}{\kappa}v} \\
&= \frac{1}{2\pi\kappa}\sqrt{\frac{E_q}{E_p}}\frac{\Gamma\left(-i\frac{E_q}{\kappa}\right)}{\left(i\frac{E_p}{\kappa}\right)^{-i\frac{E_q}{\kappa}}} = \frac{1}{2\pi\kappa}\sqrt{\frac{E_q}{E_p}}\Gamma\left(-i\frac{E_q}{\kappa}\right)\left(i\frac{E_p}{\kappa}\right)^{i\frac{E_q}{\kappa}},
\end{aligned}
$$

where we have used the definition of the Γ function. The norm-squared of the coefficients are

$$
\begin{aligned}
|\alpha_{E_pE_q}|^2 &= \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\frac{\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2}{\left(i\frac{E_p}{\kappa}\right)^{i\frac{E_q}{\kappa}}\cdot\left(-i\frac{E_p}{\kappa}\right)^{-i\frac{E_q}{\kappa}}} \\
&= \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\frac{\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2}{(-1)^{i\frac{E_q}{\kappa}}} = \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\frac{\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2}{\mathrm{e}^{-\pi\frac{E_q}{\kappa}}} = \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2\mathrm{e}^{\pi\frac{E_q}{\kappa}}, \\
|\beta_{E_pE_q}|^2 &= \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2\left(i\frac{E_p}{\kappa}\right)^{i\frac{E_q}{\kappa}}\cdot\left(-i\frac{E_p}{\kappa}\right)^{i\frac{E_q}{\kappa}} \\
&= \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2(-1)^{i\frac{E_q}{\kappa}} = \frac{1}{4\pi^2\kappa^2}\frac{E_q}{E_p}\left|\Gamma\left(i\frac{E_q}{\kappa}\right)\right|^2\mathrm{e}^{-\pi\frac{E_q}{\kappa}},
\end{aligned}
$$

and we thus have $|\alpha_{E_pE_q}|^2 = \mathrm{e}^{2\pi E_q/\kappa}\,|\beta_{E_pE_q}|^2$. The ratio $\mathrm{e}^{-2\pi E_q/\kappa}$ is the Boltzmann factor, representing the probability of a particle having energy E_q.

Clearly the Bogoliubov coefficients are nonzero, and so the definition of 'vacuum' depends on one's reference frame. Let $|0_a\rangle$ denote the Kruskal–Szekeres vacuum and $|0_b\rangle$ denote the Eddington-Finkelstein vacuum. To see what $|0_a\rangle$ looks like to an observer in the Eddington-Finkelstein system, we want to know the expectation value of the number operator $N_{E_q} = b_{E_q}b^\dagger_{E_q}$, which can be expressed in terms of the coefficient $\beta_{E_pE_q}$ using

$$
\begin{aligned}
\langle N_{E_q}\rangle &= \langle 0_a|b_{E_q}b^\dagger_{E_q}|0_a\rangle \\
&= \langle 0_a|\int_0^\infty\int_0^\infty \mathrm{d}E_p\,\mathrm{d}E'_p\Big(\alpha_{E_pE_q}a_{E_p} + \beta_{E_pE_q}a^\dagger_{E_p}\Big)\Big(\alpha^*_{E'_pE_q}a^\dagger_{E'_p} + \beta^*_{E'_pE_q}a_{E'_p}\Big)|0_a\rangle \\
&= \langle 0_a|\int_0^\infty\int_0^\infty \mathrm{d}E_p\,\mathrm{d}E'_p\Big(\alpha_{E_pE_q}\alpha^*_{E'_pE_q}a_{E_p}a^\dagger_{E'_p} + \alpha_{E_pE_q}\beta^*_{E'_pE_q}a_{E_p}a_{E'_p} \\
&\qquad + \beta_{E_pE_q}\alpha^*_{E'_pE_q}a^\dagger_{E_p}a^\dagger_{E'_p} + \beta_{E_pE_q}\beta^*_{E'_pE_q}a^\dagger_{E_p}a_{E'_p}\Big)|0_a\rangle \\
&= \int_0^\infty\int_0^\infty \mathrm{d}E_p\,\mathrm{d}E'_p\Big(\alpha_{E_pE_q}\alpha^*_{E'_pE_q}\langle 0_a|a_{E_p}a^\dagger_{E'_p}|0_a\rangle + \beta_{E_pE_q}\beta^*_{E'_pE_q}\langle 0_a|a_{E_p}a^\dagger_{E'_p}|0_a\rangle \\
&\qquad + \beta_{E_pE_q}\beta^*_{E'_pE_q}\langle 0_a|[a^\dagger_{E_p}, a_{E'_p}]|0_a\rangle\Big).
\end{aligned}
$$

Now using (4.3), and the fact that the vacuum expectation value for a-particles in the a-vacuum is by definition zero,

$$
\begin{aligned}
\langle N_{E_q}\rangle &= \int_0^\infty\int_0^\infty \mathrm{d}E_p\,\mathrm{d}E'_p\;\beta_{E_pE_q}\beta^*_{E'_pE_q}\delta(p-p') \\
&= \int_0^\infty \mathrm{d}E_p\;\beta_{E_pE_q}\beta^*_{E_pE_q}.
\end{aligned}
$$

Using the completeness relations (4.13), we have

$$\begin{aligned}
&\int_0^\infty \mathrm{d}E_p(\alpha_{pq}\alpha^*_{pq'} - \beta_{pq}\beta^*_{pq'}) = \delta(q-q'),\\
&\implies \int_0^\infty \mathrm{d}E_p(\mathrm{e}^{2\pi E_q/\kappa}\,|\beta_{E_pE_q}|^2 - |\beta_{E_pE_q}|^2) = \delta(0),\\
&\implies \left(\mathrm{e}^{\frac{2\pi E_q}{\kappa}} - 1\right)\int_0^\infty \mathrm{d}E_p|\beta_{E_pE_q}|^2 = \delta(0),\\
&\implies \int_0^\infty \mathrm{d}E_p|\beta_{E_pE_q}|^2 = \langle 0_a|b_{E_q}b^\dagger_{E_q}|0_a\rangle = \frac{\delta(0)}{\mathrm{e}^{\frac{2\pi E_q}{\kappa}} - 1}
\end{aligned}$$

over all space. Then finally the mean density of particles in the Kruskal–Szekeres as measured by an observer in the Eddington-Finkelstein vacuum is

$$n_{E_q} = \frac{\langle N_{E_q}\rangle}{\delta(0)} = \frac{1}{\mathrm{e}^{\frac{2\pi E_q}{\kappa}} - 1}, \tag{4.24}$$

which is just the blackbody distribution for photons of frequency $\omega = E_q$ and temperature

$$T = \frac{\kappa}{2\pi}. \tag{4.25}$$

Somewhat miraculously[1], then, to observers far from the black hole, it emits particles as a perfect blackbody with temperature identical to the one given in (3.32) when one considers entropy $S = A/4$.

4.4 Observing Hawking Radiation

In SI Units the surface gravity of a Schwarzschild black hole is given by

$$\kappa = \frac{c^4}{4GM}, \tag{4.26}$$

which combined with eq. (1.3) yields

$$T = \frac{\hbar c^3}{8\pi G M k_B}. \tag{4.27}$$

Let us look at a concrete example. The nearest known black hole, Gaia BH1 [33], is also one of the smallest known, with a mass of $M_{\text{BH1}} \approx 9.78 M_\odot$. We would expect that, since $T \propto 1/M$, this system has the best chance of radiating at a measurable temperature, but a simple computation only gives us $T_{\text{BH1}} \approx 7 \times 10^{-7}$ K. The cosmic microwave background, recall, has a temperature of $T_{\text{CMB}} = 2.725$ K, so our signal-to-noise ratio is very poor.

[1] To see why this appears miraculous, consider classical gravitation in geometric units $G = c = k_\mathrm{B} = 1$. Here temperature T has dimension [mass] = [energy], while surface gravity κ has units [mass^{-1}]. In the semi-classical case, this is fine, as the quantities can be related by $\hbar$. No such constant seems to be available in classical gravity, however, so the question appears to be: how is (4.25), an apparently purely quantum formula, anticipated by (3.32), an apparently purely classical one? An in-depth analysis is given by [14], in which it is suggested that the black hole thermodynamical equation is indeed 'true', independent of quantum effects.

Another test might be to wait till the estimated time that a black hole might evaporate and see whether this indeed occurs. This would also give us some insight into whether effects due to quantum gravity become relevant when the black hole has shrunk sufficiently. Let us estimate how long it would take for Gaia BH1 to evaporate. The power radiated is given by the Stefan-Boltzmann law

$$P = \frac{\mathrm{d}M}{\mathrm{d}t}c^2 = -4\pi\sigma R^2 T^4, \tag{4.28}$$

where $\sigma \approx 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$ is the Stefan-Boltzmann constant. The power emitted by Gaia BH1, with Schwarzschild radius

$$R_{\mathrm{BH1}} = \frac{2G \times 9.78 M_{\odot}}{c^2} \approx 28889\,\mathrm{m}, \tag{4.29}$$

is approximately $P = 7 \times 10^{-23}\,\mathrm{W}$. The rate at which the mass of the black hole is lost is then

$$\frac{P}{c^2} = 7.8 \times 10^{-40}\,\mathrm{kg\,s^{-1}}, \tag{4.30}$$

which is small enough that we can consider it constant. The black hole would then not be expected to evaporate for $2.5 \times 10^{70}\,\mathrm{s} \approx 8 \times 10^{62}$ years.

It seems in practice that Hawking radiation is not a measurable effect, which poses questions for the testability of quantum field theory in curved spacetime. However, it is possible to construct so-called 'analog systems' to black holes – for instance, 'dumb holes' [34]. Consider a perfect fluid governed by the equations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{4.31}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{\nabla p}{\rho} - \nabla \Phi \tag{4.32}$$

where, assuming the curl of the velocity vanishes, the velocity can be expressed as $\mathbf{v} = \nabla \psi$ for some potential ψ. The pressure in the fluid p is dependent on the density ρ.

The equations (4.32) can be rewritten using the variables

$$h(\rho) = \int^{\rho} \frac{\mathrm{d}p(\bar{\rho})}{\bar{\rho}}, \quad \xi(\rho) = \log \rho, \tag{4.33}$$

to yield

$$
\begin{aligned}
\frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{1}{\rho}\nabla \cdot (\rho \mathbf{v}) = 0 \\
\implies \frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{\mathbf{v}}{\rho} \cdot \nabla \rho + \nabla \cdot \mathbf{v} = 0 \\
\implies \frac{\partial \xi}{\partial t} + \mathbf{v} \cdot \nabla \xi + \nabla \cdot \mathbf{v} = 0
\end{aligned}
\tag{4.34}
$$

and

$$
\frac{\partial \nabla \psi}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + \frac{\nabla p}{\rho} = 0 \tag{4.35}
$$

which we can then integrate to obtain

$$
\implies \frac{\partial \psi}{\partial t} + \frac{1}{2}\mathbf{v}^2 + h = 0. \tag{4.36}
$$

To examine the behaviour of perturbations to the system, let us write the variables in terms of a zeroth order term and a perturbation denoted by a prime, that is,

$$
\begin{aligned}
\xi &= \xi_0 + \xi', \\
\psi &= \psi_0 + \psi'.
\end{aligned}
\tag{4.37}
$$

Then the linearised version of (4.32) for the perturbation ξ', we have

$$
\begin{aligned}
\frac{\partial \xi'}{\partial t} + \nabla \psi_0 \cdot \nabla \xi' + \nabla(\nabla \psi') = 0 \\
\implies \frac{1}{\rho_0}\left(\frac{\partial \rho_0 \xi'}{\partial t} + \nabla(\rho \mathbf{v} \cdot \xi')\right) + \nabla(\rho \nabla \psi') = 0.
\end{aligned}
\tag{4.38}
$$

For the perturbations ψ', we have

$$
\begin{aligned}
\frac{\partial(\cancel{\psi_0} + \psi')}{\partial t} + \frac{1}{2}\left(\cancel{(\nabla \psi_0)^2} + 2\nabla \psi_0 \cdot \nabla \psi' + \cancel{(\nabla \psi')^2}\right) + \cancel{h} + h' = 0 \\
\implies \frac{\partial \psi'}{\partial t} + \nabla \psi_0 \cdot \nabla \psi' + h'(\exp(\xi_0))\xi' = 0 \\
\implies \frac{\partial \psi'}{\partial t} + \nabla \psi_0 \cdot \nabla \psi' + c^2 \xi' = 0
\end{aligned}
\tag{4.39}
$$

where in the last term we have used the fact that $c^2 = \mathrm{d}p/\mathrm{d}\rho$ and the definition of h.

Then we obtain an expression for ξ' given by

$$
\xi' = -\frac{1}{c^2}\frac{\partial \psi'}{\partial t} - \frac{1}{c^2}\mathbf{v_0} \cdot \nabla \psi', \tag{4.40}
$$

where we define $\mathbf{v}_0 = \nabla\psi_0$.

Substituting this expression into (4.38), we obtain

$$\frac{1}{\rho_0}\left(\frac{\partial}{\partial t}\frac{\rho}{c^2}\frac{\partial\psi'}{\partial t} + \frac{\partial}{\partial t}\frac{\rho_0\mathbf{v}_0}{c^2}\cdot\nabla\psi' + \nabla\cdot\left(\frac{\rho_0\mathbf{v}}{c^2}\frac{\partial\psi'}{\partial t}\right) - \nabla\cdot\rho_0\nabla\psi' + \nabla\cdot\left(\mathbf{v}\frac{\rho_0}{c^2}(\mathbf{v}\cdot\nabla\psi')\right)\right) = 0 \tag{4.41}$$

which can be rewritten as

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\psi') = 0 \tag{4.42}$$

where $g^{\mu\nu}$ is equivalent to the metric described by

$$\mathrm{d}s^2 = \frac{\rho_0}{c}\left((c^2 - v_0^2)\mathrm{d}t^2 + 2\mathbf{v}_0\,dt\,\mathrm{d}\mathbf{x} - \mathrm{d}\mathbf{x}^2\right), \tag{4.43}$$

where $\mathbf{v}_0 = \nabla\psi_0$. To recognise this, note that $\det g_{\mu\nu} = \rho_0^2$.

Then (4.42) is just the Klein-Gordon equation for a massless scalar field ψ' in a curved metric.

In this way, one can create fluid systems which behave like black holes and should also radiate like them [34]. This is claimed to have been done by *e.g.* [35], possibly strengthening the validity of Hawking's original calculation. The extent to which experimental data on analog systems can tell us anything about real black holes is, however, debated [2].

5 | The Information Loss Paradox

5.1 The Breakdown of Physics?

Bekenstein's original paper made use of a formal similarity—between the expression for Shannon entropy as a measure of information, and thermodynamic entropy—to postulate that an increase in the area of a black hole, and thereby its entropy, correspond to an increase in "inaccessibility of information about its internal configuration" [3]. Once Hawking showed that Bekenstein's result, and the four laws of black hole mechanics in general [4], were more than a formal analogy [15], it became immediately apparent that information about projectiles sent into a black hole is lost in a very real sense.

Hawking showed shortly thereafter that the evolution of a system from a gravitationally collapsed stellar object to an evaporated or evaporating black hole is non-unitary [15]. The paper was originally titled *The Breakdown of Physics in Gravitational Collapse* [16]. It is worth looking at why Hawking felt his result had such disastrous implications for physics as a study.

One way of interpreting the resulting thermal spectrum in eq. 4.24 is as a vacuum fluctuation resulting in the production of entangled infalling and outfalling modes, with the negative-energy modes $|b_{E_{-q}}\rangle$ falling into the black hole and the positive energy modes $|b_{E_q}\rangle$ emerging from it, in analogy with the Penrose process.

To simplify the argument, we approximate the energy modes of the particles as discrete (we relabel $E_q \to E_n$ and $|b_{E_q}\rangle \to |n_{\text{out}}\rangle$ and $|b_{-E_q}\rangle \to |n_{\text{in}}\rangle$), and model the black hole as two quantum systems I (internal information) whose basis is $|n_{\text{in}}\rangle$ and H (Hawking radiation) whose basis is $|n_{\text{out}}\rangle$. Because the ingoing and outgoing states are due to pair production, they begin maximally entangled – that is, the state of the field near the event horizon as radiation begins is given by the (see Appendix B) bipartite state

$$|\phi\rangle = \sum_{n}^{\dim H} p_n |n_{\text{in}}\rangle \otimes |n_{\text{out}}\rangle. \tag{5.1}$$

By definition (Appendix B), the purity of the initial state is $\mathcal{P}(\rho_{\text{in}}) = 1$. After some evaporation has occurred, our description of the state of the system just outside the event horizon can only be given by the outgoing modes, since information about the ingoing ones is inaccessible. Then the density matrix ρ_{out} after some evaporation has occurred is given by

$$\begin{aligned}
\rho_{\text{out}} &= \mathrm{Tr}_{\text{in}}(\rho_{\text{in}}) \\
&= \sum_{n,m} p_n p_m^* \langle m_{\text{in}}|n_{\text{in}}\rangle |n_{\text{out}}\rangle\langle n_{\text{out}}| \\
&= \sum_{n,m} p_n p_m^* \delta_{nm} |n_{\text{out}}\rangle\langle n_{\text{out}}| \\
&= \sum_{n} |p_n|^2 |n_{\text{out}}\rangle\langle n_{\text{out}}|, \qquad (5.2)
\end{aligned}$$

where $p_n = \mathrm{e}^{-2\pi E_n/\kappa}$ is the probability of the particle having discrete energy E_n. Clearly $\mathcal{P}(\rho_{\text{out}}) < 1$ by completeness, since we do not sum over the entire basis of the original system. Since eventually the black hole completely evaporates, we are left only with the mixed (impure) outside states, and thus evaporation represents a non-unitary pure-to-mixed transformation. Hawking represented this transformation by a *superscattering operator* $\$$, which, unlike the usual S-matrix, is non-invertible [15, 36], meaning we cannot retrodict the initial state $|\phi\rangle$ from the final density matrix ρ_{out}. To Hawking, loss of retrodictability, which is effectively predictability, meant a loss of the predictive power of physics itself – but the paper was finally published under the more mundane *Breakdown of Predictability in Gravitational Collapse* [2, 15].

One interpretation of this result aligns with Bekenstein's original interpretation of black hole entropy: it is equivalent to saying that we lose information about the system because of its exactly thermal evaporation: by the no-hair conjecture, the thermal radiation only carries away radiation pertaining to an equivalence class of black holes with three degrees of freedom whereas the initial stellar object had many, many more than that [6, 37, 38, 39]. In particular, the von Neumann entropy of a quantum system in a state $|\psi\rangle$ is given by

$$S(\rho) = -\,\mathrm{Tr}(\rho \log \rho) = \sum_i \lambda_i \log \lambda_i \qquad (5.3)$$

where $\rho = |\psi\rangle\langle\psi|$ is the density operator with eigenvalues λ_i. For a pure state, ρ only has eigenvalues 1 and 0, so von Neumann entropy must be zero outside the black hole at the start of evaporation and increase as the state becomes increasingly mixed and we begin to lose information.

5.2 Unitary or not?

That the unitarity violation constitutes a *paradox* is contingent on how important it is that unitarity be conserved. At the very least, throwing out unitarity would mean a redefinition of quantum mechanics

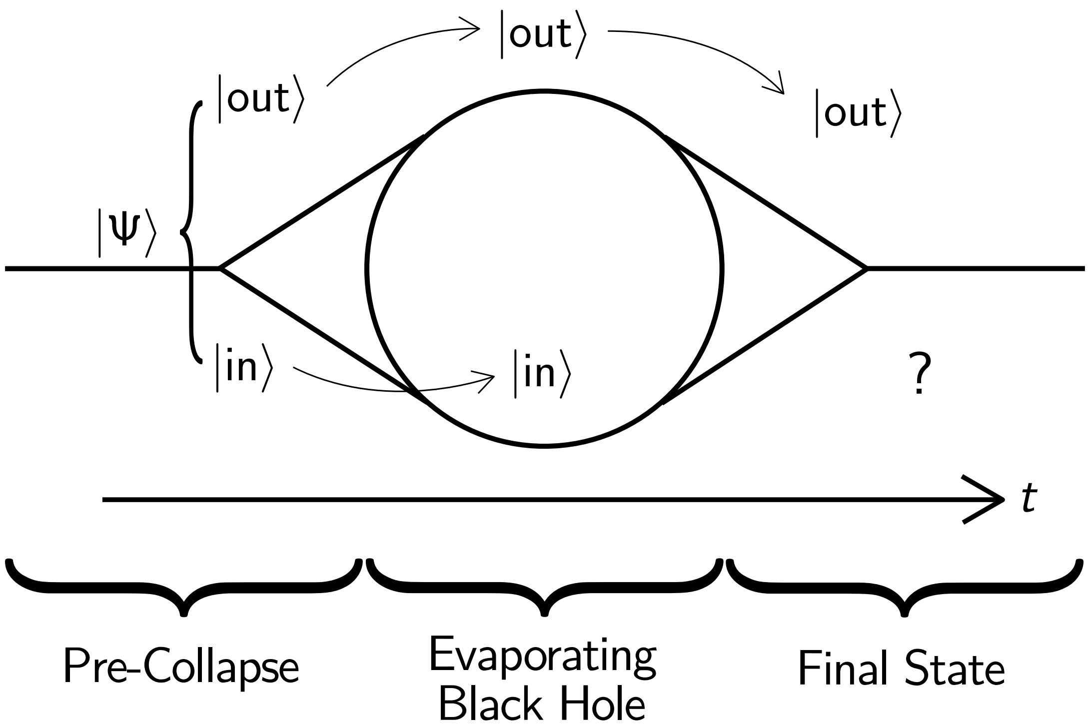

Figure 5.1: At the moment of collapse, the black hole begins to Hawking radiate. Immediately the outgoing modes leave and the ingoing modes disappear beneath the event horizon. This is true right up until the moment of evaporation. The state outside the black hole thus has to have ingoing modes traced out – and once the black hole evaporates we are left only with this outgoing mixed state. This is a nonunitary transformation.

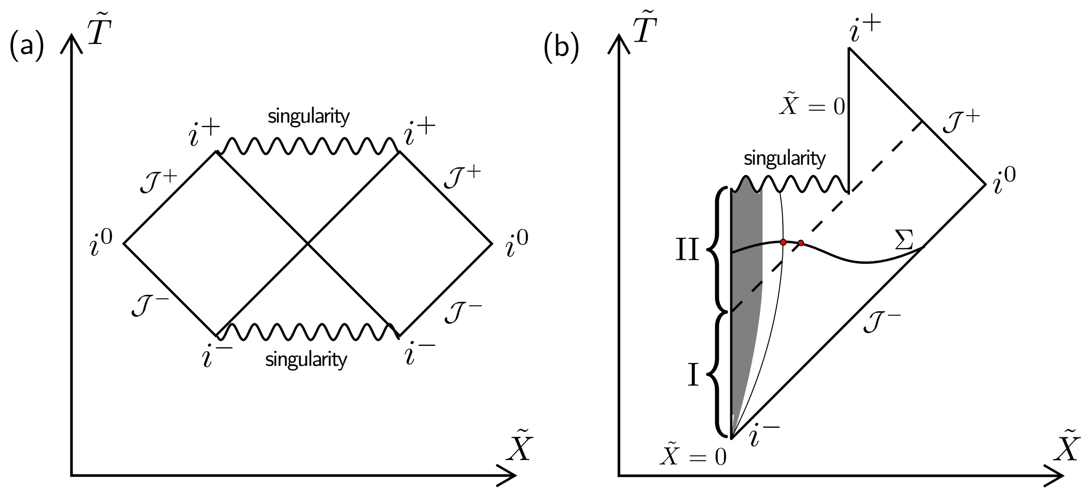

Figure 5.2: (a) Conformal diagram of an eternal (maximally-extended Kruskal–Szekeres) black hole. The past singularity is a so-called 'white hole', the time-reversed counterpart to a black hole. Note that (even if practically unreasonable) the white hole is a physically valid time-reversed solution to the Einstein field equations, so black holes do not *a priori* break CPT invariance. (b) Conformal diagram of an evaporating black hole (region II) formed from the collapse of a stellar object (region I). If the process is unitary, the Hawking radiation just outside the event horizon (dashed line) carries out all information about ingoing particles (dotted line) and thus there are spacelike hypersurfaces Σ where the same state $|\Psi\rangle$ exists at two points simultaneously as ingoing information (green) and outgoing information (red).

[6, 40]. It seems obvious to many (apparently including Hawking) that unitarity is sacrosanct, and says something fundamental about the extent to which physical laws should be able to specify the evolution of a system. Quantum mechanics, after all, deterministically specifies the evolution of states, and admits deterministic interpretations in terms of trajectories [41, 42] – but the end state of an evaporated black hole system does not deterministically retrodict anything.

If the thermal distribution derived above is exact, unitarity may only be preserved if, by the time the black hole evaporates to Planck scale, quantum gravity effects prevent its further evaporation, and we are left with some eternal black hole which has all of the initial object's degrees of freedom inside it. But this manifestly violates CPT invariance, since if we reverse the picture of gravitationally-collapsing object we find that there is no black hole at the end of the process, which was apparently impossible in the other direction [15]. We must correct at least one of: unitarity, CPT symmetry, or our expressions for temperature and/or entropy. (Although the CPT theorem assumes unitarity and may be violated by non-unitarity in any case [43].)

5.3 Unitarity *vs* Linearity

Arguments for and against black hole evaporation being a unitary process will be discussed in the next section. A simple explication of the fate and accessibility of the quantum information in such a scenario will first be discussed, partially following [19, 40].

It is illuminating to consider the example of a black hole which has been evaporating for some time. As in the non-unitary case, we can model the internal information and Hawking radiation as systems I and H which start out maximally entangled.

By assumption, the black hole should be represented by a (mixing) unitary operator U on the system $H \otimes I$, and the outgoing radiation $|n_{\text{out}}\rangle$ should after some time be described by a pure state. Then after a short time T, the state $|\phi\rangle$ at the event horizon should evolve into a tensor product of states $|\phi(T)\rangle = |n_{\text{in}}\rangle \otimes |n_{\text{out}}\rangle$.

But if this is true for two different states $|\psi_1\rangle$ and $|\psi_2\rangle$, then it must also be true for their superposition (since there is no special basis in which it couldn't be true). In particular, the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|\psi_1\rangle + |\psi_2\rangle \right) \tag{5.4}$$

should after the time interval T evolve to

$$|\Psi(T)\rangle = \frac{1}{\sqrt{2}} \left(|\psi_1^{\text{in}}\rangle \otimes |\psi_1^{\text{out}}\rangle + |\psi_2^{\text{in}}\rangle \otimes |\psi_2^{\text{out}}\rangle \right) \equiv |\Psi^{\text{in}}\rangle \otimes |\Psi^{\text{out}}\rangle \tag{5.5}$$

where the last equality makes sense if and only if $|\psi_1^{\text{in}}\rangle = |\psi_2^{\text{in}}\rangle = |\Psi^{\text{in}}\rangle$ on some spacelike hypersurface Σ (see Fig. 5.2). That is, the operator U is unitary if and only if it clones the state $|\psi_1^{\text{in}}\rangle$ onto the state $|\psi_2^{\text{in}}\rangle$. But this violates the no-clone theorem, meaning in turn that it violates the linearity of quantum mechanics [44].

5.4 Preserving Linearity

Several ideas have been proposed to deal with black hole information in a way which preserves unitarity as well as linearity. One such idea is that of black hole complimentarity, which suggests that the fate of a quantum state entering a black hole is observer-dependent. Under this assumption, infalling observers will observe that the state will fall into the black hole uninterrupted; observers at infinity will see the black hole as being a perfect mirror. In this way, the above inequalities can hold without linearity being violated for any observer.

There are some issues with this resolution. The most obvious is that it still allows the existence of a spacelike hypersurface (5.2) where a quantum state is cloned, and merely abolishes *observation* of this fact. By treating black holes as a unitary scattering process, [19] observe that the complimentary hypothesis is empirically compatible with a nonlinear black hole evolution, due to the thermalisation time of the black hole being just above the upper bound necessary for two instances of cloned information to reach one observer.

A related attempt is to prevent (5.5) by proposing that entanglement is broken at the event horizon. This would involve releasing large amounts of radiation at the event horizon, creating a so-called 'firewall' [45]. While preserving unitarity and linearity, the equivalence principle is now broken, since it is now always possible to distinguish when one is crossing the event horizon.

Much of the work done now deals with the statistical-mechanical origin of black hole entropy, and its (in-)compatibility with semiclassical calculations.

6 | Corrections to the Entropy

6.1 Very General Relativity

Let us turn then to the question of correcting Bekenstein's thermodynamical description. While a fully statistical-mechanical account of black holes is beyond the scope of this paper, results common to (for example) the string-theoretic derivation of black hole entropy in terms of microstates may be reproduced by considering corrections to general relativity as an effective field theory.

In n dimensions, the vacuum Einstein field equations are the Euler-Langrange equations derived from the Einstein-Hilbert action, given by

$$S = \frac{1}{16\pi} \int \mathrm{d}^n x \sqrt{-g} R. \tag{6.1}$$

This is the simplest action integral that can be expressed only in terms of scalars which depend only on the metric and (at most) its first-order derivatives. In general, however, a gravity theory need not be so simple: it may involve a nonzero torsion, or relax metric compatibility constraints, or contain higher-order derivatives of the metric tensor.

It should be of particular note that our derivation of Hawking radiation does not make any special use of general relativity – the only relevant conditions are that spacetime be curved and that some set of coordinates has no special importance. Therefore, as long as we look at a spacetime which retains the concept of surface gravity κ, we may keep (1.3) and look for a more general derivation of the entropy S in gravity theories with some reasonable constraints. In the spirit of general relativity, we will keep the theory generally covariant by imposing invariance of the Lagrangian under arbitrary diffeomorphism.

The language of differential forms (Appendix A) is useful to describe general theories. We start with a

Lagrangian n-form L with components

$$\mathrm{L}_{a_1 \dots a_n} = \mathrm{L}_{a_1 \dots a_n}[\phi, \nabla_a \phi], \tag{6.2}$$

where ϕ are the fields including the metric tensor g_{ab} and ∇_a is a (not necessarily Levi-Civita) connection. The action is then given by

$$\mathrm{S}^{\mathrm{grav}} = \int \mathrm{L}. \tag{6.3}$$

Our aim is to derive a general version of (3.32) with (1.3) occupying the usual status of Hawking temperature and to discuss corrections yielded in higher-derivative theories.

6.2 The Hamiltonian Formulation of General Relativity

General relativity (by design) treats time on the same footing as the other spatial coordinates. It is (often) possible, however, to foliate spacetime into 1 time coordinate and $n-1$ spacelike hypersurfaces Σ_i, and decompose the metric into the form [46]

$$\mathrm{d}s^2 = -N^2 \mathrm{d}t^2 + h_{ij}(\mathrm{d}x^i + N^i \mathrm{d}t)(\mathrm{d}x^j + N^j \mathrm{d}t), \tag{6.4}$$

where h_{ij} is the induced spatial metric, N encodes time evolution, and N_i encodes spatial evolution (between hypersurfaces).

For a theory of gravity, the Hamiltonian n-form is then given by [46, 47]

$$\mathrm{H} = \epsilon \left(N\mathcal{H} + N^i \mathcal{H}_i \right) \tag{6.5}$$

where $\mathcal{H}$ and $\mathcal{H}_i$ are Lagrange multipliers for enforcing temporal and spatial constraints.

The Hamiltonian on a spacelike hypersurface Σ_k is

$$H = \int_{\Sigma_k} \mathrm{H} \wedge \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^{n-1} \tag{6.6}$$

For asymptotially flat spacetimes, this has to be augmented [47] by a surface integral like

$$H^{\mathrm{AFS}} = H + \oint_{\partial\Sigma} \mathrm{d}^{n-2}x \left(\partial_i h_{kj} - \partial_k h_{ij} \right) \tag{6.7}$$

in order to be compatible with the Einstein field equations.

The surface terms (*i.e.*, the terms under the integral sign in (6.7)) are in fact readily associated [47] with the conserved quantities given by the Killing vectors (3.3).

6.3 Wald's Formalism

Wald [7] showed that, for a general theory of gravity, it is possible to relate quantities conserved when varying the action (3.4) to quantities conserved when varying (6.3). In particular, in a black hole spacetime with reasonable constraints, one can find a conserved charge Q and relate it to the conserved quantities of a black hole horizon, rederiving (3.32) with Q taking on the role of entropy.

We consider a diffeomorphism generated by the flow of a vector field χ^a and the variation of the Lagrangian n-form L along it, given by

$$\begin{aligned}\hat{\delta}\mathrm{L} &= E^{ab}\hat{\delta}g_{ab} + E^{(i)}\hat{\delta}\phi_{(i)} + \mathrm{d}\Theta[\phi, \hat{\delta}\phi] \\ &= \mathcal{L}_\chi \mathrm{L} = \{\iota_\chi, \mathrm{d}\}\mathrm{L} \\ &= \mathrm{d}\iota_\chi \mathrm{L},\end{aligned}$$

where on the last line we have used the fact that L is closed, since $L \propto \epsilon$, where ϵ is the Levi-Civita tensor, because there is only one linearly independent n-form in an n-dimensional space. The terms $E^{(i)}$ and $E^{\mu\nu}$ are the equations of motion for the fields $\phi_{(i)}$ and g_{ab} (the metric) respectively. The term containing the symplectic current Θ is a total derivative, so the action only changes by a boundary term when varied along χ^a. Thus the vector field generates a family of discrete symmetries to which we can associate a Nöther current

$$J = \Theta - \iota_\chi \mathrm{L}. \tag{6.8}$$

The exterior derivative of the current is

$$\begin{aligned}\mathrm{d}J &= \mathrm{d}\Theta - \mathrm{d}\iota_\chi \mathrm{L} \\ &= \hat{\delta}L - E^{ab}\hat{\delta}g_{ab} - E^{(i)}\hat{\delta}\phi_{(i)} - \hat{\delta}\mathrm{L} \\ &= -E^{ab}\hat{\delta}g_{ab} - E^{(i)}\hat{\delta}\phi_{(i)} \\ &= 0 \quad \text{(on-shell)}\end{aligned}$$

where by *on-shell* we mean that the equations of motions are satisfied for this particular variation along χ^a. J is then closed, so we can use the general identity $\mathrm{d}^2 = 0$ to define the conserved Noether charge Q by

$$\mathrm{d}J = Q. \tag{6.9}$$

Varying the Nöther current we have

$$\begin{aligned}
\delta J &= \delta[\Theta] - \iota\chi\delta \mathrm{L} \\
&= \delta[\Theta] - \iota_\chi \left[E^{ab}\hat{\delta} g_{ab} + E^{(i)}\hat{\delta}\phi_{(i)} + \mathrm{d}\Theta \right] \\
&= \delta[\Theta] - \mathcal{L}_\chi \Theta - \mathrm{d}\iota_\chi\Theta \\
&\equiv \delta_1[\Theta] - \delta_2[\Theta] - \mathrm{d}\iota_\chi\Theta,
\end{aligned} \tag{6.10}$$

where the final definition of the two variations δ_1 and δ_2 give rise to the symplectic form

$$\Omega[\phi, \delta_1\phi, \delta_2\phi] = \delta_1[\Theta(\phi, \nabla\phi)] - \delta_2[\Theta(\phi, \nabla\phi)], \tag{6.11}$$

which measures the phase-space area spanned by the two independent variations and is the density of a variation of the Hamiltonian which generates the χ^a-evolution. We thus have

$$\delta H = \delta \int_\Sigma J - \int_\Sigma \mathrm{d}\iota_\chi\Theta, \tag{6.12}$$

where Σ is a closed surface. Now recall from the previous section that the surface terms (6.7) are associated with the quantities $\mathcal{Q}$ from the spacetime's Killing vectors. Then equation (6.12) can be rewritten as

$$\mathrm{d}\delta\mathcal{Q} = \int_\Sigma \delta \mathrm{d}Q - \int_\Sigma \mathrm{d}\iota_\chi\Theta, \tag{6.13}$$

which implies (using Stokes' theorem)

$$\delta\mathcal{Q} = \int_{\partial\Sigma} \delta Q - \int_{\partial\Sigma} \iota_\chi\Theta, \tag{6.14}$$

i.e. we have exactly related the charges of (3.4) to the charges of (6.1)[1]

In particular we may look at the case where the variation is along a linear combination of the timelike Killing vector $(\partial_t)^a$ or rotational Killing vector $(\partial_\phi)^a$, in which case the equations of motion are satisfied. In particular, consider the case that the vector field χ is the Killing horizon $\chi^a = (\partial_t)^a + \Omega(\partial_\phi)^a$ of a black hole. Notice that at the horizon the vector field vanishes by definition and so the last term containing the interior product can be disregarded. Then we easily recover a thermodynamical equation like (3.32) by varying the asymptotic energy M associated with $(\partial_t)^a$ and the asymptotic angular momenta $\mathbf{L}$ associated with (a family of) rotational Killing vector(s) $(\partial_\phi)^a$ as

$$\delta M + \mathbf{\Omega} \cdot \delta \mathbf{L} = \int_{\partial\Sigma} \delta Q[\chi]. \tag{6.15}$$

[1] The interior product term $\iota_\chi\Theta$ is a numerical factor which is irrelevant to the physical interpretation of the quantities $\mathcal{Q}$ [7].

To explicitly include the surface gravity in this equation, it is enough to re-express the Nöther charge in terms of χ^a and $\nabla_a \chi_b$ and replace the first derivative with the binormal ε_{ab} which is always true at the event horizon. The resulting quantity $\tilde{Q}$ is analogous to the quantity Q but for a Killing horizon $\tilde{\chi}$ defined by $\tilde{\chi} = \kappa\chi$. One can then construct a diffeomorphism from the hypersurface defined by $\tilde{\chi}$ to the one defined by χ. By identifying χ as the Killing horizon for a black hole with surface gravity set to unity, we recover the quantity

$$Q = \kappa\tilde{Q}, \tag{6.16}$$

and plugging into (6.15) we obtain

$$\frac{\kappa}{2\pi}\delta S[\chi] = \delta M + \mathbf{\Omega} \cdot \delta \mathbf{L}. \tag{6.17}$$

where

$$S = 2\pi \int_{\partial\Sigma} \tilde{Q}[\chi]. \tag{6.18}$$

The result is two-fold: firstly, it reveals that, for general relativity specifically, the Nöther charge of the Einstein-Hilbert action when varied along a black hole horizon is (up to an integral) the entropy of the black hole. Secondly, it provides an algorithm to derive the entropy of a black hole in a more general theory of gravity.

6.4 Bekenstein's Formula Revisited

Let us now use the above prescription to re-derive the Bekenstein-Hawking entropy in n-dimensional general relativity. The components of the Lagrangian n-form are (see Appendix A) given by

$$\mathrm{L}^{\mathrm{GR}}_{a_1 \dots a_n} = \frac{R\epsilon_{a_1 \dots a_n}}{16\pi}. \tag{6.19}$$

The action is

$$\mathrm{S}^{\mathrm{GR}} = \int \mathrm{L}^{\mathrm{GR}}. \tag{6.20}$$

We vary the Lagrangian in the usual way [9], with

$$\begin{aligned} \delta \mathrm{L}^{\mathrm{GR}} &= \frac{\epsilon_{a_1 \dots a_n}}{16\pi} \left[G_{ab}\delta g^{ab} + \nabla_d \left(g^{ab}\nabla^d \delta g_{ab} - g^{da}\nabla^b \delta g_{ab} \right) \right] \\ &= \frac{\epsilon_{a_1 \dots a_n}}{16\pi} \left[G_{ab}\delta g^{ab} + \mathrm{d} \left(\epsilon_{da_1 \dots a_{n-1}} g^{de} g^{fh} (\nabla_f \delta g_{eh} - \nabla_e \delta g_{fh}) \right) \right], \end{aligned} \tag{6.21}$$

noting that the total derivative term, which we normally disregard, can be written as an exterior derivative when we use the Levi-Civita symbol. Normally we would disregard this total derivative, but here can extract from it the symplectic current and consider a variation along a vector field χ using the Lie derivative $\mathcal{L}_\chi L$ to obtain

$$
\begin{aligned}
\Theta_{a_1\ldots a_{n-1}} &= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} g^{de} g^{fh} (\nabla_f \delta g_{eh} - \nabla_e \delta g_{fh}) \\
&= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} g^{de} g^{fh} (\nabla_f \mathcal{L}_\chi g_{eh} - \nabla_e \mathcal{L}_\chi g_{fh}) \\
&= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} g^{de} g^{fh} (\nabla_f (\nabla_e \chi_h + \nabla_h \chi_e) - \nabla_e (\nabla_f \chi_h + \nabla_h \chi_f)) \\
&= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} g^{de} g^{fh} (\nabla_{[f} \nabla_{e]} \chi_h + \nabla_f \nabla_h \chi_e - \nabla_e \nabla_h \chi_f) \\
&= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} \left(g^{de} g^{fh} (R_{feh}{}^{g} \chi_g) + g^{de} g^{fh} (\nabla_f \nabla_h \chi_e - \nabla_e \nabla_h \chi_f) \right) \\
&= \frac{\epsilon_{da_1\ldots a_{n-1}}}{16\pi} \left(g^{de} g^{fh} (R_{feh}{}^{g} \chi_g) + g^{de} g^{fh} (\nabla_f \nabla_h \chi_e) - g^{df} g^{eh} (\nabla_f \nabla_h \chi_e) \right) \qquad (6.22)
\end{aligned}
$$

where in the sixth line we have used the compatibility of the metric with the connection $\nabla_k g_{ij} = 0$. Since, by the reasoning above, we can neglect the interior derivative

$$
\begin{aligned}
(\iota_\chi \mathrm{L})_{a_1\ldots a_{n-1}} &= \frac{1}{16\pi} R (\iota_\chi \epsilon)_{a_1\ldots a_{n-1}}) \\
&= R \chi^d \epsilon_{da_1\ldots a_{n-1}}.
\end{aligned}
$$

Thus we can also drop the term containing the Riemann curvature by recognising

$$
\begin{aligned}
g^{de} g^{fh} R_{feh}{}^{g} \chi_g &= g^{de} g^{fh} R_{fehi} g^{gi} g_{gj} \chi^j \\
&= g^{de} g^{fh} R_{fehi} \delta^i_j \chi^j \\
&= g^{de} R^h_{ehi} \chi^i = g^{de} R_{ei} \chi^i. \qquad (6.23)
\end{aligned}
$$

Using the equations of motion, we have

$$
\begin{aligned}
g^{de} g^{fh} R_{feh}{}^{g} \chi_g &= \frac{1}{2} g^{de} R g_{ei} \chi^i \\
&= \frac{1}{2} \delta^d_i R \chi^i \\
&= \frac{1}{2} R \chi^d. \qquad (6.24)
\end{aligned}
$$

When we multiply this by the Levi-Civita tensor, we get a scalar multiple of the interior product, which we can ignore. Now

$$\begin{aligned}
J_{a_1\dots a_{n-1}} &= \frac{\epsilon_{da_1\dots a_{n-1}}}{16\pi}\left(g^{de}g^{fh}(\nabla_f\nabla_h\chi_e) - g^{df}g^{eh}(\nabla_f\nabla_h\chi_e)\right) \\
&= \frac{\epsilon_{da_1\dots a_{n-1}}}{16\pi}\left((\nabla_f\nabla^f\chi^d) - g^{df}g^{eh}(\nabla_h\nabla_f - [\nabla_h,\nabla_f])\chi_e\right) \\
&= \frac{\epsilon_{da_1\dots a_{n-1}}}{16\pi}\left(\nabla_f\nabla^f\chi^d - \nabla_h\nabla^d\chi^h - g^{df}g^{eh}R_{hfe}{}^{g}\chi_g\right) \\
&= \frac{\epsilon_{da_1\dots a_{n-1}}}{16\pi}\left(\nabla_e\nabla^e\chi^d - \nabla_e\nabla^d\chi^e\right) \\
&= \frac{\epsilon_{da_1\dots a_{n-1}}}{16\pi}\times 2\nabla_e\left(\nabla^{[e}\chi^{d]}\right),
\end{aligned} \tag{6.25}$$

where we again can get rid of the curvature term as above. Then

$$J_{a_1\dots a_{n-1}} = \frac{\epsilon_{da_1\dots a_{n-1}}}{8\pi}\nabla_e(\nabla^{[e}\chi^{d]}) \tag{6.26}$$

We can find Q by noting the definition $J_{a_1\dots a_{n-1}} = (\mathrm{d}Q)_{a_1\dots a_{n-1}} = \nabla_{[a_1}Q_{a_2\dots a_{n-1}]}$. Noticing that the Levi-Civita tensor already antisymmetrises the expression, this is simply given by

$$Q_{a_1\dots a_{n-2}} = -\frac{\epsilon_{dea_1\dots a_{n-2}}}{16\pi}\nabla^d\chi^e. \tag{6.27}$$

Finally, to find the Nöther charge in this formalism, we define

$$\tilde{Q}_{a_1\dots a_{n-2}} = \frac{Q_{a_1\dots a_{n-2}}}{\kappa} = -\frac{\epsilon_{dea_1\dots a_{n-2}}}{16\pi\kappa}\nabla^d\chi^e. \tag{6.28}$$

To integrate this expression, we notice [10] that the Komar mass of a black hole in general relativity is given by

$$M = -\frac{1}{8\pi}\int \epsilon_{a_1\dots a_n}\nabla^{a_{n-1}}\chi^{a_n}, \tag{6.29}$$

and that the surface gravity $\kappa = (4M)^{-1}$. Then

$$S = 2\pi\int\tilde{Q} = 2\pi\times\frac{1}{2}\frac{M}{\kappa} = 4\pi M^2 = \frac{16\pi M^2}{4} = \frac{A}{4}. \tag{6.30}$$

6.5 Higher-Derivative Gravity

A natural generalisation of general relativity with higher-order curvature terms is Lovelock gravity [18, 48], whose Lagrangian n-form can explicitly be written, to second order in curvature, as

$$\mathrm{L}^{\mathrm{LL}} = \frac{\epsilon}{16\pi}\left(-2\Lambda + \alpha_1 R + \alpha_2(\mathcal{R}^2)\right), \tag{6.31}$$

where $\mathcal{R}^2$ is the Gauss-Bonnet term, which appears in effective action functionals of some string theories like heterotic string theory [49]. It is given by

$$\begin{aligned}\mathcal{R}^2 &= R^2 + R_{abcd}R^{abcd} - 4R_{ab}R^{ab} \\ &= \alpha R^2 + \beta R_{ab}R^{ab} + \gamma R. \end{aligned} \tag{6.32}$$

Lovelock gravity can be considered as an effective action for heterotic string theory. It is straightforward to show using Wald's formalism that the Gauss-Bonnet term adds an extra scalar term to Bekenstein's entropy which is dependent on the dimensionality of the system and of order A^{-1}[50].

Indeed, for extremal black holes, [51] derives an order A^{-1} correction, as well as a logarithmic one. The logarithmic correction is matched by an effective-QFT derivation by [52], and appears to contradict recent results by Page [17] which suggest that black hole entropy should first increase and then decrease[6].

7 | Conclusions and Further Work

Here we have examined the evolution of the development of black hole mechanics from their initial study as time-asymmetric regions up until the foundations for today, where they are looked at in a statistical mechanical light.

Charged, rotating classical black holes mirror many of the aspects of (semi-)quantum ones, including the ability to radiate electromagnetically, delete information about the pre-collapsed system, and obey a thermodynamical equation with a temperature proportional to their own surface gravity.

That many of these properties re-emerge when black holes are treated in conjunction with quantum effects is not in itself surprising, but the culmination of all of these properties to produce a decidedly un-quantum effect (unitarity violation) is a cause for concern and demands further attention.

We have looked at some corrections to the thermodynamical description of classical black holes by using general relativity as an effective action and matched certain results from quantum gravity.

The method does not, however, reproduce results from a more standard, statistical-mechanical, string-theoretic treatment [51] of black hole entropy as a sum of microstates of a collection of branes – in particular, such a method should yield corrections of order $\mathcal{O}(\log A)$.

It is possible, however, to use Wald's formalism to compute *quantum* corrections to the black hole entropy using the path integral formalism of quantum mechanics in combination with Wald's method, reproducing the $\mathcal{O}(\log A)$ results [52]. In this case quantum field theory is used as an effective action, instead of gravity. Examining this method and its relation to current statistical-mechanical approaches to the information loss paradox will be discussed in a future work.

Bibliography

[1] David Finkelstein. Past-future asymmetry of the gravitational field of a point particle. *Phys. Rev.*, 110:965–967, May 1958. doi: 10.1103/PhysRev.110.965. URL `https://link.aps.org/doi/10.1103/PhysRev.110.965`.

[2] Carla Rodrigues Almeida. The thermodynamics of black holes: from penrose process to hawking radiation. *Eur. Phys. J*, 2021. URL `https://link.springer.com/article/10.1140/epjh/s13129-021-00022-9`.

[3] Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev.*, 7, 1973. URL `https://link.springer.com/article/10.1140/epjh/s13129-021-00022-9`.

[4] J.M. Bardeen, B. Carter, and S.W. Hawking. The four laws of black hole mechanics. *Commun. Math. Phys*, 31:161–170, 1973. URL `https://www.semanticscholar.org/paper/The-four-laws-of-black-hole-mechanics-Bardeen-Carter/8ce8d2e2d30f51c42d80ff22fe2b6030493aba65`.

[5] Stephen W. Hawking. Particle creation by black holes. *Commun. Math. Phys*, 43:199–220, 1975. URL `https://www.semanticscholar.org/paper/Particle-creation-by-black-holes-Hawking/6b9d83c086ea962b6e7ef5caaaf2c241a4d47ff8`.

[6] David Wallace. Why black hole information loss is paradoxical. 2018. URL `https://arxiv.org/abs/1710.03783`.

[7] Robert M. Wald. Black hole entropy is noether charge. *Phys. Rev. D.*, 48, 1993. URL `https://journals.aps.org/prd/abstract/10.1103/PhysRevD.48.R3427`.

[8] Roger Penrose. Black holes. *Scientific American*, 226:38–47, 1972. URL `https://www.scientificamerican.com/article/black-holes/`.

[9] Sean Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. 2004.

[10] Robert M. Wald. *General Relativity*. The University of Chicago Press, 1984.

[11] Subrahmanyan Chandrasekhar. *The Mathematical Theory of Black Holes*. 1983.

[12] Demetrios Christodoulou. Reversible and irreversible transformations in black-hole physics. *Phys. Rev. Lett.*, 25, 1970. URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.25.1596.

[13] Ya.B. Zel'dovich. Amplification of cylindrical electromagnetic waves reflected from a rotating body. *Soviet Physics JETP*, 35(6), 1972.

[14] Erik Curiel. Classical black holes are hot. 2023.

[15] Stephen W. Hawking. Breakdown of predictability in gravitational collapse. *Phys. Rev. D.*, 14, 1976. URL https://journals.aps.org/prd/abstract/10.1103/PhysRevD.48.R3427.

[16] Gordon Belot, John Earman, and Laura Ruetsche. The hawking information loss paradox: The anatomy of a controversy. *Brit. J. Phil. Sci*, 50:189–229, 1999. URL https://www.journals.uchicago.edu/doi/10.1093/bjps/50.2.189.

[17] Don Page. Hawking radiation and black hole thermodynamics. 2004. URL https://arxiv.org/abs/hep-th/0409024.

[18] C. Charmousis. *Higher Order Gravity Theories and Their Black Hole Solutions*, pages 299–346. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-88460-6. doi: 10.1007/978-3-540-88460-6_8. URL https://doi.org/10.1007/978-3-540-88460-6_8.

[19] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. 2007. URL https://arxiv.org/abs/0708.4025.

[20] Albert Einstein. Die feldgleichungen der gravitation. *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, pages 844–847, 1915. URL https://link.springer.com/chapter/10.1007/978-3-662-57411-9_11.

[21] Albert Einstein. Die grundlage der allgemeinen relativitätstheorie. *Annalen der Physik*, 49, 1916. URL https://onlinelibrary.wiley.com/doi/10.1002/andp.19163540702.

[22] Karl Schwarzschild. Über das gravitationsfeld eines massenpunktes nach der einsteinschen theorie. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, pages 189–196, 1916. URL https://ui.adsabs.harvard.edu/abs/1916SPAW.......189S/abstract.

[23] J. Robert Oppenheimer and George Michael Volkoff. On massive neutron cores. *Physical Review*, 55:374–381, 1939. URL https://link.aps.org/doi/10.1103/PhysRev.55.374.

[24] J. Robert Oppenheimer and Hartland Snyder. On continued gravitational contraction. *Physical Review*, 56:455–459, 1939. URL `https://journals.aps.org/pr/abstract/10.1103/PhysRev.56.455`.

[25] Paul Murdin and Betty Louise Webster. Optical identification of cygnus X-1. 1971.

[26] Werner Israel. Event horizons in static electrovac space-times. *Commun. Math. Phys*, 8:245–260, 1968. URL `https://link.springer.com/article/10.1007/BF01645859`.

[27] Brandon Carter. Axisymmetric black hole has only two degrees of freedom. *Phys. Rev. Lett*, 26:331–333, 1971. URL `https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.26.331`.

[28] Roger Penrose. "golden oldie": Gravitational collapse: The role of general relativity. *General Relativity and Gravitation*, 34:1141–1165, 2002.

[29] L.H. Ford. Quantum field theory in curved spacetime. 1997.

[30] Michael Peskin and Daniel Schroeder. *An Introduction to Quantum Field Theory*. CRC Press, 1995.

[31] Viatcheslav Mukhanov and Sergei Winitzki. *Introduction to quantum effects in gravity*. Cambridge University Press, 2007.

[32] Dorothea Deeg. *Quantum Aspects of Black Holes*. PhD thesis, Ludwig-Maximilians-Universität München, 2006.

[33] Kareem El-Badry, Hans-Walter Rix, Eliot Quataert, Andrew W Howard, Howard Isaacson, Jim Fuller, Keith Hawkins, Katelyn Breivik, Kaze W K Wong, Antonio C Rodriguez, Charlie Conroy, Sahar Shahaf, Tsevi Mazeh, Frédéric Arenou, Kevin B Burdge, Dolev Bashi, Simchon Faigler, Daniel R Weisz, Rhys Seeburger, Silvia Almada Monter, and Jennifer Wojno. A sun-like star orbiting a black hole. *Monthly Notices of the Royal Astronomical Society*, 518(1):1057–1085, 11 2022. URL `https://doi.org/10.1093/mnras/stac3140`.

[34] W.G. Unruh. Experimental black hole evaporation? *Physical Review Letters*, 46(21), 1981.

[35] Jeff Steinhauer. Observation of quantum hawking radiation and its entanglement in an analogue black hole. *Nature Physics*, 12, 2016.

[36] Stephen W. Hawking. The superscattering matrix for two-dimensional black holes. *Physical Review Letters*, 69(25):4069–4072, 1992. doi: 10.1103/PhysRevLett.69.4069.

[37] S. W. Hawking. Superscattering matrix for two-dimensional black holes. *Physical Review D*, 50(6):3982–3986, September 1994. ISSN 0556-2821. doi: 10.1103/physrevd.50.3982. URL `http://dx.doi.org/10.1103/PhysRevD.50.3982`.

[38] Stephen W. Hawking. Information loss in black holes. *Physical Review D*, 72:084013, 2005. doi: 10.1103/PhysRevD.72.084013.

[39] Valeri P. Frolov and Igor D. Novikov. *Black Hole Physics: Basic Concepts and New Developments*. Springer, 1998. ISBN 978-0792341166. doi: 10.1007/978-94-011-5139-9.

[40] Steffen Gielen. Does black hole evaporation imply that physics is nonunitary, and if so, what must the laws of physics look like? *International Journal of Quantum Information*, 07(05):847–878, August 2009. ISSN 1793-6918. doi: 10.1142/s0219749909005596. URL `http://dx.doi.org/10.1142/S0219749909005596`.

[41] Hugh Everett. "relative state" formulation of quantum mechanics. *Reviews of Modern Physics*, 29: 454–462, 1957. doi: 10.1103/RevModPhys.29.454.

[42] Detlef Dürr, Sheldon Goldstein, and Nino Zanghì. *Quantum Mechanics Without Quantum Philosophy*. Theoretical and Mathematical Physics. Springer, 2012. ISBN 978-3-642-23867-3. doi: 10.1007/978-3-642-23869-7.

[43] Ralf Lehnert. CPT symmetry and its violation. 8(11), 2016.

[44] W.K. Wootters and W.H. Zurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982. URL `https://doi.org/10.1038/299802a0`.

[45] Ahmed Almheiri, Donald Marolf, Joseph Polchinski, Douglas Stanford, and James Sully. An apologia for firewalls. *Journal of High Energy Physics*, 2013(9), 2013. URL `http://dx.doi.org/10.1007/JHEP09(2013)018`.

[46] R. Arnowitt, S. Deser, and C.W. Misner. *The Dynamics of General Relativity*, chapter 7. Wiley, 1962.

[47] Tullio Regge. Role of surface integrals in the hamiltonian formulation of general relativity. *Annals of Physics*, 88:286–316, 1974.

[48] David Lovelock. The einstein tensor and its generalizations. *Journal of Mathematical Physics*, 12(3), 1971.

[49] Barton Zwiebach. Curvature squared terms and string theories. *Physics Letters B*, 156(5):315–317, 1985. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(85)91616-8.

[50] Vivek Iyer and Robert M. Wald. Some properties of the noether charge and a proposal for dynamical black hole entropy. *Physical Review D*, 50(2):846–864, July 1994. ISSN 0556-2821. doi: 10.1103/physrevd.50.846. URL `http://dx.doi.org/10.1103/PhysRevD.50.846`.

[51] Andrew Strominger and Cumrun Vafa. Microscopic origin of the bekenstein-hawking entropy. *Physics Letters B*, 379(1–4):99–104, June 1996. ISSN 0370-2693. doi: 10.1016/0370-2693(96)00345-0. URL `http://dx.doi.org/10.1016/0370-2693(96)00345-0`.

[52] Ashoke Sen. Logarithmic corrections to schwarzschild and other non-extremal black hole entropy in different dimensions. *Journal of High Energy Physics*, 2013(4), April 2013. ISSN 1029-8479.

[53] Michael Nielsen and Isaac L Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[54] John Goold. *Notes on Quantum Information Theory*. 2021.

[55] Mikio Nakahara. *Geometry, Topology and Physics*. Institute of Physics Publishing, 2003.

[56] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2013.

A1 | Some Quantum Information Theory Concepts

Throughout the text, a semiclassical black hole is modelled as a pair of maximally entangled systems I representing internal information and H representing Hawking information. Here we will clarify some of those terms [53, 54].

A quantum system $\mathcal{H}$ is a Hilbert space, *i.e.* an infinite dimensional vector space with elements

$$|\Psi\rangle = \sum_i c_i |\phi_i\rangle,$$

where $|\psi_i\rangle$ are the basis vectors of the space, with an inner product defined by

$$\langle\psi_i|\psi_j\rangle = \int \psi_i^*(x)\psi_j(x)\mathrm{d}x, \quad \langle\psi_i|\psi_j\rangle \geq 0 \tag{A1.1}$$

where $\psi(x) = \langle x|\psi\rangle$ is the wavefunction associated with the vector $|\psi\rangle$, that is, the coefficients of the vector in position space, spanned by the vectors $|x\rangle$ which satisfy

$$\hat{x}|x\rangle = x|x\rangle, \tag{A1.2}$$

that is, they are eigenstates of the position operator $\hat{x}$.

The *dimension* of $\mathcal{H}$ is the span of the basis vectors of the system, denoted by

$$\dim \mathcal{H} = |\mathcal{H}|.$$

Two quantum systems $\mathcal{H}_1$ with basis $\{|\psi_i^1\rangle\}$ and $\mathcal{H}_2$ with basis $\{|\psi_i^2\rangle\}$ can be composed into a system $\mathcal{H}$ denoted by

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2.$$

where $\otimes$ here denotes the tensor product of the spaces. The space $\mathcal{H}$ is then spanned by the basis $\{|\phi_i^1\rangle\} \otimes \{|\phi_j^2\rangle\}$, where $|\psi_i^1\rangle \otimes |\psi_j^2\rangle$ is a product state of $|\psi_i^1\rangle$ and $|\psi_j^2\rangle$. A quantum system $\mathcal{H}$ composed of two such subsystems is called *bipartite*, and a state in such a system is a *bipartite state*.

A superposition of vectors in $\mathcal{H}$ describes the state of a quantum mechanical system. A state is *pure* if it is described by a single state vector $|\Psi\rangle$. The density operator of a quantum state is given by

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|,$$

where p_i is the probabilities that the state is described by $|\psi_i\rangle$. For a pure state $|\Psi\rangle$,

$$\rho = |\Psi\rangle\langle\Psi|,$$

and we describe a state as *mixed* otherwise. The purity operator

$$\mathcal{P}(\rho) = \mathrm{Tr}(\rho^2) \tag{A1.3}$$

measures the extent to which a state is pure, in which case $\mathcal{P} = 1$, or *mixed*, in which case

$$\mathcal{P} = \frac{1}{\dim \mathcal{H}}.$$

A state $|\Psi\rangle$ is *entangled* if it cannot be described by a single product

$$|\Psi\rangle \neq |\psi_i^1\rangle \otimes |\phi_j^2\rangle,$$

which means the systems are correlated in some way – the evolution of one depends nonlocally on the evolution of another. A *maximally entangled* state is given by

$$|\Psi\rangle = \frac{1}{\sqrt{|\mathcal{H}|}} \sum_{i=0}^{|\mathcal{H}|-1} |\psi_i^1\rangle \otimes |\psi_j^2\rangle. \tag{A1.4}$$

This is a state where the two systems are as strongly correlated as possible. We define the *reduced density matrix* $\rho_{\mathcal{H}_1}$ of the subsystem $\mathcal{H}_1$ by

$$\rho_{\mathcal{H}_1} = \mathrm{Tr}_{\mathcal{H}_2}(\rho) = \sum_i \langle\psi_i^2|\rho|\psi_i^2\rangle.$$

When we take the partial trace, we refer to it as *tracing out* the subsystem $\mathcal{H}_2$. The state described by the reduced density matrix is then necessarily mixed.

A2 | Gravity in Terms of Differential Forms

In order to discuss a more general theory of gravity it is useful to switch to a non-coordinate basis $(e^{a_1}, \ldots, e^{a_n})$ and consider the Lagrangian n-form

$$\mathrm{L} = \mathcal{L}\omega,$$

where ω is the volume form on the given manifold.

In general, the only linearly independent n-form in n-dimensional space is the Levi-Civita symbol ϵ with components $\epsilon_{a_1,\ldots a_n}$, so the Lagrangian n-form will always be given by

$$\mathrm{L} = \mathcal{L}\epsilon.$$

We will make some use of the exterior derivative, interior product, and Cartan's formula which relates the two for differential forms. The exterior derivative d maps a p-form ω to a $(p+1)$-form $\mathrm{d}\omega$ by [55]

$$(\mathrm{d}\omega)_{\mu_1\ldots\mu_{p+1}} = (p+1)\partial_{[\mu_1}\omega_{\mu_2\ldots\mu_{p+1}]}.$$

The interior derivative can be written in terms of the Hodge dual

Likewise, the interior product maps p-forms to $(p-1)$-forms, although they are defined with respect to a specific vector field X. If ω is a p-form which acts on the fields $Y_1, \ldots, Y_p$ by $\omega(Y_1, \ldots, Y_p)$ then the interior product $\iota_X\omega$ of ω with respect to X acts on fields $X_1, \ldots X_{p-1}$ by

$$\iota_X\omega(X_1, \ldots, X_{p-1}) = \omega(X, X_1, \ldots, X_{p-1}).$$

Cartan's formulates is an expression for the Lie derivative of a p-form ω in terms of the exterior derivative and interior product, given by [56]

$$\mathcal{L}_X \omega = \{d, \iota_X\}\omega \tag{A2.1}$$

where curly brackets denote the anticommutator.