The Black Hole Firewall Paradox

Dissertation by

Tycho Sikkenk

Supervised by

Professor Fay Dokwer

Department of Physics, Imperial College London

In Partial Fulfillment of the Requirements
for the Degree of
Master of Science (MSc)

Imperial College
London

June-September 2013

(Submitted September 20, 2013)

"Nobody has the slightest idea what is going on."

Joe Polchinski on the Firewall Paradox, KITP Fuzz or Fire Conference, August 2013

Abstract

Famously, the quantum treatment of black holes predicts the appearance of Hawking radiation emanating from the black hole. Detailed analysis of the behaviour of the wave modes in the proximity of the horizon in the Schwarzschild background shows that the Hawking modes will be entangled with partner modes at the interior of the black hole, so that an observer far removed from the black hole will detect a mixed state. Thus, it seems that the collapse and subsequent evaporation of the black hole caused by emission the Hawking radiation is not unitary. Information is lost during the process and the state of the black hole is independent of that of the initial configuration of the spacetime. The solution to this paradox was found in the notion of black hole complementarity, which postulates that the observer at infinity will measure a different reality than an infalling spectator. Further support dispelling the conclusions of the information paradox came with the holographic principle and the gauge/gravity duality of the AdS/CFT correspondence, that conclusively proved that even the life cycle of an AdS black hole is governed by a unitary matrix of the conformal field theory on the boundary of the spacetime.

However, it was recently shown that the basic tenets of black hole complementarity are internally incompatible. At least for black holes in an advanced stage of evaporation, an infalling observer is able to determine the state of the Hawking radiation both while it is still close to the horizon and far away at infinity, whilst these should be related by a semi-classical evolution. As a consequence, even before crossing the horizon such an observer will be incinerated by high energy quanta: a firewall.

Tycho Sikkenk
Imperial College London
September 20, 2013

Contents

1 Hawking Radiation

1.1 Black Hole Entropy

Consider a small amount of matter falling into an ordinary Schwarzschild black hole of mass M with radius R_s and area A given by

$$R_s = 2M \qquad\qquad A = 4\pi R_s^2 = 16\pi M^2. \tag{1.1.1}$$

Here we have normalised the universal constants $G = c = \hbar = 1$ so we can use Planck units. Of course, the gas must have some entropy, so that after disappearing behind the horizon of the black hole the entropy of the rest of the Universe has decreased. This is in clear violation of the original second law of thermodynamics, which states that the entropy S of an isolated system can never decrease: $dS_{tot} \geq 0$.

Assume that instead of letting it fall into the black hole, we put the matter in a small black box for safekeeping. Instead of vanishing behind the horizon and into the singularity, the matter's entropy would simply be inside of the box. We could even reopen the lid, making the entropy reappear. Hence, the contradiction can easily be resolved by ending the separation between the black hole and the rest of the Universe, effectively absorbing the former into the definition of the latter. In order for this to be possible, however, the black hole must have some feature that allows external observers to determine its entropy [2].

Suppose the matter consists of a photon carrying just a single entropy quantum, i.e. one bit of information, when it crosses the horizon. Note that in general, a photon will convey to the black hole much more information than the single qubit specifying its spin. Using the fact that the momentum of a photon is given by $p_{ph} = k = h/\lambda$, Heisenberg's Uncertainty Principle gives

$$\Delta(x)\Delta(p) = h\Delta(x)\Delta(\lambda)^{-1} \geq 1/2. \tag{1.1.2}$$

We want the energy of the photon to be maximally specified so that the mass increase of the black hole is known. Then, in order for the photon to carry the minimum amount of entropy when entering the black hole region, it should have as high an uncertainty as possible. Hence, we should leave it maximally unlocalised and its wavelength maximally specified. Such a photon would have to be spread out over the entire horizon of the black hole. In that case, it will have a wavelength approximating the Schwarzschild radius, i.e. $\lambda \sim R_s$. When the single bit photon falls into the black hole, it incrementally increases the mass and thus the area of the horizon,

$$\delta M = E_{ph} \sim \frac{1}{\lambda} \sim \frac{1}{R_s} \quad \leftrightarrow \quad \delta R_s = 2\delta M \sim \frac{1}{R_s} \quad \leftrightarrow \quad \delta A_s = 8\pi R_s \delta R_s \sim 1. \tag{1.1.3}$$

Hence, by dimensional analysis we must conclude that

adding one bit of information to a black hole increases the area of the horizon by approximately one Planck area ℓ_p^2.

Furthermore, assuming a smooth transition between normal spacetime and a black hole, a point-like black hole with zero mass and area must have zero entropy. Integrating and imposing this boundary condition, we find the Bekenstein-Hawking entropy

$$S_{BH} = \int dS = \int \frac{\delta S}{\delta A} dA \sim A. \tag{1.1.4}$$

In other words, the entropy measured in bits is proportional to the horizon area measured in Plank units [41]. It is as if the entropy of matter becomes associated with the boundary of the black hole when it crosses the horizon. The horizon of a black hole can be divided into Planck areas ℓ_p^2 that each harbour approximately one bit of information (see Fig.(2)).

Notice that a Planck area ℓ_p^2 is of order 10^{-70} m^2, so that the entropy of a black hole is very large. This reflects the fact that black holes are by definition the most dense objects in the universe. From statistical mechanics, we know that entropy is related to the number of configurations $\mathcal{N}$ the system can be in as $S = \ln \mathcal{N}$. For a solar mass black hole with a radius of ~ 3 km, this implies that

$$\mathcal{N} \sim 10^{77},$$

a gigantic number of microstates. This raises the question how the black hole can give rise to such a huge number of internal configurations. Because the Bekenstein-Hawking entropy is proportional to the

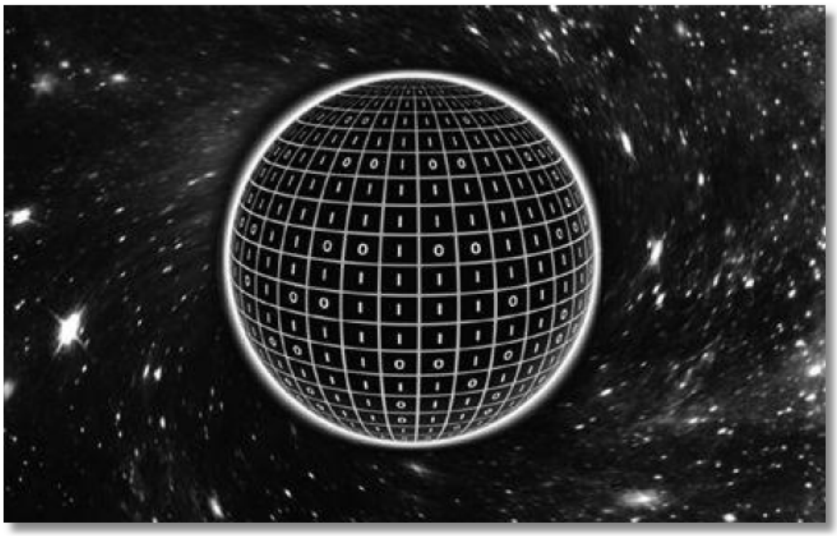


Figure 2: The horizon of a black hole horizon can be divided into Planck areas l_p^2, that each harbour one bit of information.

horizon surface area, it is only logical to search for 'deformation modes' of the horizon that would allow for classical differentiation between the states. However, through a series of proofs, collectively known as the *no-hair theorems*, it has been established that classically black holes can have no other properties besides mass, angular momentum and charge. Any deformation of the horizon would either float off to spatial infinity or inwards to the singularity, restoring the horizon to its spherically symmetric form. It is interesting to note that these theorems depend crucially on the Cosmic Censorship Conjecture, that postulates that the physical universe does not tolerate singularities without horizon [30]. Also, if the microstates of the black hole were present at the horizon, any infalling observer would be able to detect them there and violate the equivalence principle that is at the core of general relativity. Hence, the horizon can hold no information, corresponding to an entropy of $S = \ln 1 = 0$. Another option is to differentiate the configurations by quantum mechanical properties of the interior. However, incorporating quantum effects into the theory of black holes has lead to even more fundamental problems, dubbed the information paradox [25].

We have seen that a black hole of energy $E = M$ has a certain entropy S_{BH}, depending on its area. Then, according to the fundamental relation of thermodynamics,

$$TdS_{SH} = dE, \tag{1.1.5}$$

it must also have a corresponding temperature T that is given by

$$T = \left(\frac{dS_{BH}}{dE}\right)^{-1} \sim \left(\frac{dA_s}{dM}\right)^{-1} \sim \left(\frac{d}{dM}(M^2)\right)^{-1} \sim \frac{1}{M}. \tag{1.1.6}$$

Classically, the horizon acts as a one-way membrane, so that black holes are perfectly absorbing bodies that do not radiate. Temperature by itself does not automatically mean that a body radiates. However, by the law of detailed balance for equilibrated systems, a body with an absorptance $\sigma(k)$ for matter quanta with wavenumber k must have a rate of emission $\Gamma(\varepsilon(k))$ for an emissivity $\varepsilon(k)$ of the same quanta. By Kirchoff's law, this rate of emission can then be related to the black body spectrum as

$$\Gamma(\varepsilon(k)) = \int \frac{d^3k}{(2\pi)^3}\alpha(k)\frac{1}{e^{2\pi k/T} - 1}. \tag{1.1.7}$$

In the case of black holes, $\alpha(k)$ approximates unity at least for those quanta with $\lambda \leq R_s$, so that they are nigh perfect black body emitters. Hence, our semi-classical derivation predicts that black holes must radiate, whereas the classical geometry redirects all outgoing physical worldines back into the interior. Including quantum processes in the description of black holes is thus expected to lead to particle creation. In 1975, Hawking described such an effect in a seminal paper on the quantum fluctuations of the vacuum in the presence of a gravitational potential [15].

According to the time-energy uncertainty relation of quantum mechanics

$$\Delta(T)\Delta(E) \geq 1/2 \tag{1.1.8}$$

relatively large energy fluctuations $\Delta(E)$ are permitted as long as they have a lifetime smaller than the short interval $\Delta(T)$. In a pure vacuum virtual particle-antiparticle pairs are constantly created and annihilated, temporarily violating energy conservation. However, for all other quantum numbers such as angular momentum, charge or lepton number, conservation is guaranteed. Consequently, applying an electric field to a pure vacuum will separate charged particles from their oppositely charged antiparticles, generating a detectable current through the vacuum. This is known as the *Schwinger Effect*. The associated physical constant ϵ_0 is now even used as a baseline to convert between the SI units of electric charge and force.

For a canonical Schwarzschild black hole, the region of spacetime near the horizon is unexceptional in every aspect. Notably, the curvature of spacetime near the horizon is small, so that the vacuum state approximates that of ordinary Minkowski spacetime. Particle-antiparticle pairs will be created from the vacuum. If this process occurs across the horizon, the gravitational force of the black hole is sufficiently high to confine one member of the pair to the interior, while enabling the other to escape to infinity. While falling in, the gravitational energy of the particle at the interior diminishes, so that it has a negative total energy with respect to an observer at infinity and energy is conserved. Notice that unlike the Schwinger Effect, this process is not restricted to particles that are electrically charged. A gravitational field is a general feature of curved spacetimes that affects all forms of matter. Furthermore, this effect is not biased towards matter or antimatter, so that the radiation produced by black holes should consist of all kinds of particles. Nevertheless, vacuum fluctuations with smaller energy are more likely to occur than those with higher energies, so one would expect infrared photons to feature most prominently in the black hole radiation [25].

1.2 Quantum Field Theory on Curved Spacetimes

In order to be able to prove the emission of Hawking radiation by black holes, we will first revise some general aspects of quantum field theories on curved backgrounds. In order to do quantum field theory, we need to be able to the define what we mean by 'particle'. Given some particle state, we can then use technical instruments such as propagators and Feynman diagrams to calculate the outcome of scattering experiments. Only for a specific subset of physical spacetimes it is possible to define a positive-definite inner product between wavefunction operators that leads to a consistent definition of the vacuum. Furthermore, on curved spacetimes, the notion of a vacuum state is no longer valid globally, so that two observers generally find different numbers of excitations while describing the same region of spacetime.

The classical action of a real scalar field $\phi(x)$ in curved spacetime is

$$S = -\frac{1}{2}\int d^4x\,\sqrt{-g}\,(\nabla_\mu\phi\nabla^\mu\phi + m^2\phi^2). \tag{1.2.1}$$

The equation of motion for $\phi(x)$ derived from this action is the Klein-Gordon equation

$$\nabla^\mu\nabla_\mu\phi - m^2\phi = 0, \tag{1.2.2}$$

where ∇_μ is the covariant derivative operator compatible with metric $g_{\mu\nu}$ on manifold $\mathcal{M}$. Assume there is some form of causal structure attributed to the spacetime $(\mathcal{M}, g_{\mu\nu})$. Then, for every achronal set $\Sigma \in \mathcal{M}$ consisting of pairs of points (p, q) connected by timelike curves there is a domain of dependence $D(\Sigma)$ comprising the entire causal past and future of Σ, defined by

$$D(\Sigma) = \{p \in \mathcal{M}\,|\,\text{every (inextendible) causal curve through } p \text{ intersects } \Sigma\ \}. \tag{1.2.3}$$

If all points are causally related to Σ so that $D(\Sigma) = \mathcal{M}$, then Σ is said to be a Cauchy surface of the spacetime. It follows that Cauchy surfaces are 3-dimensional spacelike hypersurfaces. Any spacetime in which a Cauchy surface exist is called globally hyperbolic.

Globally hyperbolic spacetimes possess some interesting structural characteristics that admit a classical picture of their evolution from some initial conditions on a Cauchy surface. Two theorems that are very relevant to our present discussion are listed here [50]:

Theorem 1.2.1 (Dieckmann[8], Geroch[12]): , If spacetime $(\mathcal{M}, g_{\mu\nu})$ is globally hyperbolic with Cauchy surface Σ, then $\mathcal{M}$ has topology $\mathbb{R} \times \Sigma$. Manifold $\mathcal{M}$ can be foliated by a one-parameter family of smooth Cauchy surfaces Σ_t, where t acts as a type of 'time coordinate'. All hypersurfaces of coordinate $t = const.$ form Cauchy surfaces.

Theorem 1.2.2 (Hawking and Ellis[13]): For any globally hyperbolic spacetime $(\mathcal{M}, g_{\mu\nu})$ with Cauchy surface Σ, there exist a global solution ϕ to the Klein-Gordon equation (1.2.2) valid on all of $\mathcal{M}$ for which $\phi = \phi_0$ and $n^\mu \nabla_\mu \phi = \tilde{\phi}_0$ on Σ, where n^μ is the 'future' unit normal to Σ. In this sense the function ϕ is the classical evolution of the pair of functions $(\phi_0, \tilde{\phi}_0)$ on Σ. Even stronger, for every closed subspace $S \in \Sigma$ the solution ϕ_S restricted to the domain of dependence $D(S)$ depends only on the initial configuration of the field on S.

Essentially, the Cauchy surface of a spacetime provides the framework to build a quantum field theory. By singling out a single time-like coordinate t, we can consistently define the momentum density π conjugate to the field ϕ so that we can come to a phase space formulation of the evolution of the scalar field.

Suppose $(\mathcal{M}, g_{\mu\nu})$ is a globally hyperbolic spacetime foliated by spacelike Cauchy surfaces of constant time-like parameter t, as proposed by theorem (1.2.1). We can then introduce a 'time evolution' vector field t^μ tangential to the congruence of causal curves through surfaces Σ_t and define it to satisfy $t^\mu \nabla_\mu t = 1$ everywhere on $\mathcal{M}$. t^μ is easily decomposable over the complete basis provided by the unit normal n^μ to Σ_t and the shift vector N^μ as

$$t^\mu = \alpha n^\mu + N^\mu, \tag{1.2.4}$$

where α is a constant called the lapse factor. On Σ_t, we can introduce local coordinates (t, x^i), for which $t^\mu \nabla_\mu x^i = 0$ for all $x = 1, 2, 3$ so that we can reduce $t^\mu \nabla_\mu \phi = \dot{\phi}$. In terms of these local coordinates, the Klein-Gordon equation (1.2.2) is rewritten in terms of a Lagrangian density with respect to time-like coordinate t

$$S = \int \mathcal{L}\ dt. \tag{1.2.5}$$

The spatial Lagrangian density $\mathcal{L}$ is expressed as an integral over Σ_t and its coordinate basis as

$$\mathcal{L} = -\frac{1}{2} \int_{\Sigma_t} \Big[-\left(n^\mu \nabla_\mu \phi\right)^2 + h^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi + m^2 \phi^2 \Big]\, \alpha \sqrt{h}\ d^3x, \tag{1.2.6}$$

where $h_{\mu\nu}$ is the induced metric on Cauchy surface Σ_t. Inverting Eq.(1.2.4), we find the relation

$$n^\mu \nabla_\mu \phi = \frac{1}{\alpha}(t^\mu - N^\mu) \nabla_\mu \phi = \frac{1}{\alpha} \left(\dot{\phi} - N^\mu \nabla_\mu \phi \right),$$

whence we must conclude that on Σ_t the conjugate momentum density π to ϕ is

$$\pi \equiv \frac{\delta S}{\delta \dot{\phi}} = (n^\mu \nabla_\mu \phi) \sqrt{h}. \tag{1.2.7}$$

Collectively, all the values of the variables $\phi(x)$ and its conjugate $\pi(x)$ form the phase space M to a curved spacetime equipped with a Klein-Gordon action. By theorem (1.2.2), the global solution ϕ can then be regarded as the classical evolution of the functions (ϕ_0, π_0) on some Cauchy surface Σ_0.

This situation is similar to the large n limit of the quantum mechanical case where phase space M_{qm} is composed of the set of n positions $q_1, \ldots, q_n$ and their conjugate momenta $p_1, \ldots, p_n$. These variables undergo a classical evolution with the Hamiltonian $H_{qm} = H(q_1, \ldots, q_n; p_1, \ldots, p_n)$ according to Hamilton's equations of motion

$$\frac{dq_a}{dt} = \frac{\partial H}{\partial p_a}, \qquad \frac{dp_a}{dt} = -\frac{\partial H}{\partial q_a}, \qquad a = 1, \ldots, n. \tag{1.2.8}$$

By collecting together all points of the phase space in the variable $y = (q_1, \ldots, q_n; p_1, \ldots, p_n)$, these equations can be summarised in terms of an antisymmetric $2n \times 2n$ matrix $\Omega^{\alpha\beta}$, which satisfies

$$\Omega^{\alpha\beta} = \begin{cases} 1 & \text{if } |\alpha - \beta| = n \\ 0 & \text{if } |\alpha - \beta| \neq n. \end{cases} \tag{1.2.9}$$

We obtain, for $\alpha, \beta = 1, \ldots, 2n$,

$$\frac{dy^\alpha}{dt} = \sum_{\alpha=1}^{2n} \Omega^{\alpha\beta} \frac{\partial H}{\partial y^\beta}. \tag{1.2.10}$$

Thus, the matrix $\Omega^{\alpha\beta}$ allows us to derive the evolution of the quantum mechanical system from a differential of the Hamiltonian [50]. It follows more generally that any system with some phase space M should similarly permit a *symplectic form* ω that defines a vector field V_H describing the system's flow from the differential dH. Hamilton's equations of motion then correspond to the integral curves of V_H on M that determine the dynamically possible motions of the system [26].

One way to elucidate the relationship between the two formulations of Hamiltonian systems is to look at the cotangent bundle $T^*(\mathcal{Q})$ of the n-dimensional configuration manifold $\mathcal{Q}$. $\mathcal{Q}$ is composed of the positional data $(q_1, \ldots q_n)$ of the original phase space manifold M. By definition, the cotangent bundle $T^*\mathcal{N}$ of m-dimensional manifold $\mathcal{N}$ is the $2m$-dimensional manifold formed by all the m points on $\mathcal{N}$ together with the cotangent vectors dual to every point. We can look at the local coordinates $(q_1, \ldots q_n)$ on $\mathcal{Q}$ as supporting the conjugate momenta $(p_1, \ldots p_n)$ as the components of the cotangent vectors on $T^*(\mathcal{Q})$. The phase space M used in the more abstract picture then corresponds to the complete tangent bundle itself, i.e. $M = T^*(\mathcal{Q})$. Furthermore, configuration space $\mathcal{Q}$ has the natural structure of a vector space, so we can locally choose a basis. The basis components of the vectors in $\mathcal{Q}$ then globally define 'linear coordinates' $q_1, \ldots, q_n$ on $\mathcal{Q}$. By identifying phase space M with the cotangent bundle of Q these coordinates then extrapolate to globally well-defined coordinates $q_1, \ldots, q_n; p_1, \ldots, p_n$ on M.

Because of the vector space nature of M, we can identify its tangent space $T_y(M)$ at some point $y \in M$ with M itself. The symplectic form ω can then be used to specify a corresponding bilinear function on M as

$$\begin{aligned} \Omega \quad : \quad & M \times M \mapsto \mathbb{R} \\ & (y_1, y_2) \mapsto x \in \mathbb{R}. \end{aligned} \tag{1.2.11}$$

Function Ω is called the *symplectic structure* on M. Any manifold (M, ω) with such a symplectic structure defined by symplectic form ω is called a *symplectic manifold*. Since elements of the cotangent space are linear functionals acting on $T(M)$, having a symplectic structure is equivalent to stating that ω must be a member of $T^*(M) \times T^*(M)$. As the map between the Hamiltonian H and the differential equations of motion is unique, we demand the symplectic form to be equally non-degenerate. Hence, every differential dH corresponds to a unique vector field $V_H \in M$. ω must act on V_H to give dH,

$$dH = \omega(V_H \, , \, \cdot). \tag{1.2.12}$$

Moreover, we require the Hamiltonian to be constant along flow lines over the integral curves of V_H, which corresponds to

$$dH(V_H) = \omega(V_H \, , \, V_H) = 0. \tag{1.2.13}$$

This proves that ω satisfies the definition of an alternating bilinear form. Every alternating bilinear form is antisymmetric, because

$$\begin{aligned} 0 &= \omega(V_{H_1} + V_{H_2} \, , \, V_{H_1} + V_{H_2}) \\ &= \omega(V_{H_1} \, , \, V_{H_1}) + \omega(V_{H_1} \, , \, V_{H_2}) + \omega(V_{H_2} \, , \, V_{H_1}) + \omega(V_{H_2} \, , \, V_{H_2}) \\ &= \omega(V_{H_1} \, , \, V_{H_2}) + \omega(V_{H_2} \, , \, V_{H_1}) \\ & \qquad \therefore \quad \omega(V_{H_1} \, , \, V_{H_2}) = -\omega(V_{H_2} \, , \, V_{H_1}). \end{aligned} \tag{1.2.14}$$

Since ω is antisymmetric and bilinear, it must be a 2-form formed by the wedge product of two elements of phase space M. Finally, we also demand that symplectic form ω itself should not change under flow lines of the Hamiltonian, which in the less abstract quantum mechanical case simplifies to the requirement that Hamilton's equations of motion maintain the same form throughout the phase space. More technically, this means that the Lie derivative L of ω along integral curves of V_H must vanish. By using the momentum map identity $i_{V_H}\omega = dH$ and the nilpotency of the exterior derivative d, we obtain

$$\begin{aligned} 0 \quad = \quad & L_{V_H}(\omega) = d(i_{V_H}\omega) + i_{V_H} d\omega = \underbrace{d(dH)}_{0} + d\omega(V_H) \\ & \therefore \quad d\omega = 0, \end{aligned} \tag{1.2.15}$$

so that we find that ω should be a closed 2-form [26].

By definition, differential forms are naturally integrated over their (sub)manifold. Stokes' theorem states that if ρ is an $(n-1)$-form on manifold R and ∂R is the boundary of R, then

$$\int_R d\rho = \oint_{\partial R} \rho. \tag{1.2.16}$$

Suppose then we integrate ρ over two m-dimensional hypersurfaces Σ and Σ' that coincide everywhere except for a well-defined region where they enclose an $(m+1)$-dimensional hypervolume S, as in Fig.(3). The difference between the two calculations is then given by an integral of n-form $d\rho$ over hypervolume S,

$$\left(\int_\Sigma \rho\right) - \left(\int_{\Sigma'} \rho\right) = \int_S d\rho. \tag{1.2.17}$$

If the differential form ρ is closed, then $d\rho = 0$ and any integral over ρ will be independent of the specific hypersurface over which is integrated. The same holds true for the symplectic form ω.

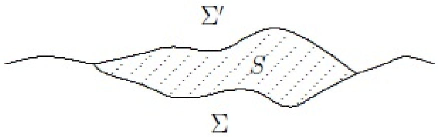


Figure 3: m-dimensional hypersurfaces Σ and Σ' that overlap almost completely. Where they do not coincide, they form a $(m+1)$-dimensional hypersurface between them.

This provides a very attractive candidate for the symplectic structure function Ω. For the case of the Klein-Gordon curved spacetime, it must be the integral of the symplectic form over any Cauchy surface Σ_0,

$$\begin{aligned}
\Omega(y_1\,,\,y_2) &= \Omega([\phi_1,\pi_1]\,,\,[\phi_2,\pi_2]) \\
&= \int_{\Sigma_0} (\pi_1\phi_2 - \pi_2\phi_1)\, d^3x \\
&= \int_{\Sigma_0} [\phi_2 n^\mu \nabla_\mu \phi_1 - \phi_1 n_\mu \nabla_\mu \phi_2]\sqrt{h}\, d^3x \\
&= \int_{\Sigma_0} dS^\mu [\phi_2 \nabla_\mu \phi_1 - \phi_1 \nabla_\mu \phi_2]
\end{aligned} \tag{1.2.18}$$

Because of the closed nature of the symplectic form these integrals do not depend on the Cauchy surface over which they are calculated. We say Klein-Gordon spacetime has a *natural* symplectic structure.

By Darboux's theorem, any symplectic manifold $(\mathcal{M},\omega)$ of dimension $2n$ can be locally made to look like the standard symplectic space $\mathbb{R}^{2n}$. Extrapolating this result to the limit where $n \to \infty$, it is directly applicable to the Klein-Gordon symplectic manifold. In the standard symplectic space, the symplectic structure Ω is given by a non-singular antisymmetric matrix that is conventionally taken to be

$$\Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \tag{1.2.19}$$

where I_n is the $n \times n$ identity matrix. However, by a change of basis, we can change into block diagonal form

$$\Omega = i \begin{pmatrix} \sigma_2 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_2 \end{pmatrix} \qquad \text{where} \qquad i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.2.20}$$

In this basis, real solutions of the Klein-Gordon equation are separated into n independent pairs of two solutions $\{\phi,\phi'\}$ with suppressed index $i = 1, 2, \dots n$. These solutions and their conjugate momenta have only $\Omega([\phi,\pi]\,,\,[\phi',\pi']) = 1$ as non-vanishing symplectic function. For every pair, we subsequently rotate to the complex solution $\psi = (\phi - i\phi')/\sqrt{2}$ to diagonalise the 2×2 matrix. We define the wedge product norm $||\psi||$ by

$$\begin{aligned} ||\psi||^2 &= i \int dS_\mu \psi^* \overleftrightarrow{\nabla}^\mu \psi \\ &= \frac{i}{2} \int dS_\mu (\phi + i\phi') \overleftrightarrow{\nabla}^\mu (\phi - i\phi') \\ &= \frac{1}{2} \int dS_\mu \left[\left(\phi \overleftrightarrow{\nabla}^\mu \phi' \right) - \left(\phi' \overleftrightarrow{\nabla}^\mu \phi \right) \right] \\ &= 1, \end{aligned} \tag{1.2.21}$$

so that ψ is of unit norm. Since this holds for all of the n different pair of solutions, there is a complex basis $\{\psi_i\}$ of solutions to the Klein-Gordon equation for which we can use the symplectic structure of the manifold to define the hermitian *Klein-Gordon inner product*

$$\begin{aligned} (\psi_i, \psi_j) = i \int dS_\mu \psi_i^* \overleftrightarrow{\nabla}^\mu \psi_j &= \begin{cases} ||\psi_i||^2 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \\ &= \delta_{ij}. \end{aligned} \tag{1.2.22}$$

Note, however, that there is no basis that allows for diagonalisation of the 2-form into the $2n \times 2n$ identity matrix I_{2n} since the inner product is not positive definite. Because of the antisymmetry of the symplectic structure, we have

$$\begin{aligned} (\psi_i^*, \psi_j^*) = i \int dS_\mu \psi_i \overleftrightarrow{\nabla}^\mu \psi_j^* &= \begin{cases} ||\psi_i^*||^2 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \\ &= -\delta_{ij}. \end{aligned} \tag{1.2.23}$$

Separating the complex solutions from their complex conjugates by a permutation of basis vectors, the symplectic function can be described by the diagonal matrix

$$\tilde{\Omega} = \begin{pmatrix} (\psi_i, \psi_j) & (\psi_i, \psi_j^*) \\ (\psi_i^*, \psi_j) & (\psi_i^*, \psi_j^*) \end{pmatrix} = \begin{pmatrix} \delta_{ij} & 0 \\ 0 & -\delta_{ij} \end{pmatrix}. \tag{1.2.24}$$

The fact that the Klein-Gordon inner product admits negative values makes it impossible to quantise the solutions. Consider, however, a massless scalar field in ordinary flat Minkowski spacetime with metric $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$. In this spacetime, the Cauchy surfaces are foliated by a global inertial time t. As such, transition between different Cauchy surfaces gains the meaningful content of time translation. Incorporating the Minkowski metric, the Klein-Gordon equation reduces to

$$\partial_t^2 \phi(x) - \vec{\nabla}^2 \phi(x) = 0, \tag{1.2.25}$$

which has obvious exponential solutions of the form $\phi = e^{i\omega t - i\mathbf{k}\mathbf{x}}$ for all frequencies $\omega \in \mathbb{R}$ with respect to t that satisfy $\omega^2 = |\mathbf{k}|^2$. If we restrict to the subspace of positive frequency solutions with a time Fourier transform

$$\psi(\omega, \mathbf{x}) = \int dt \, e^{i\omega t} \, \psi(t, \mathbf{x}) \tag{1.2.26}$$

that vanish for $\omega < 0$, then the KG inner product is positive-definite by Eq.(1.2.24). We can use the positive frequency solutions to build up a Hilbert space $\mathcal{H}$, with binary operation defined by Eq. (1.2.22). The negative frequency solutions, whose time Fourier transform vanishes for $\omega > 0$, can be put into linear correspondence with vectors of the complex conjugate Hilbert space $\mathcal{H}^*$ that is dual to $\mathcal{H}$, i.e. $\mathcal{H}^* = \bar{\mathcal{H}}$. As vectors of $\mathcal{H}^*$ then, the negative frequency solutions adhere to the positive-definiteness condition necessary to build a Hilbert space[49].

By Noether's theorem, there is a conserved charge associated with time translations between the Cauchy surfaces, notably the Hamiltonian

$$H = \int T^{00} \, d^3x \tag{1.2.27}$$

where $T_{\mu\nu}$ is the stress- energy tensor given by

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L} \, \delta^\mu_\nu \tag{1.2.28}$$

for the Klein-Gordon Lagrangian $\mathcal{L}$ from Eq.(1.2.1). A generic quantised field solution ϕ can then be described as a linear combination of positive and negative frequency solutions,

$$\phi(x) = \sum_i \left[a_i \psi_i(x) + a_i^\dagger \psi_i^*(x) \right] \tag{1.2.29}$$

where a_i and their hermitian conjugates $a_i^\dagger$ are operators satisfying the commutation relations

$$[a_i, a_j] = 0 = [a_i^\dagger, a_j^\dagger], \qquad [a_i, a_j^\dagger] = \delta_{ij}. \tag{1.2.30}$$

The creation and annihilation operators $\{a_i\}$ describe a (bosonic) Fock space of particles as excitations of a unique vacuum state $|0\rangle$ that is the ground state of the Hamiltonian in Eq.(1.2.27). As usual, the vacuum state is defined as the state that is destroyed by annihilation operators a_i for all corresponding frequencies ω_i. Furthermore, it is conventionally normalised to unity, i.e.

$$a_i|0\rangle = 0 \quad \forall i, \qquad \langle 0|0\rangle = 1. \tag{1.2.31}$$

It is the uniformity of Minkowski spacetime that allows for a globally valid quantum field theory with a uniquely defined vacuum. Its global inertial time t singles out a preferred orthonormal basis of KG solutions $\{\psi_i\}$ for which we can consistently and meaningfully define some conjugate momenta. As a consequence, the creation and annihilation operators used to define particle states in a Fock space are uniquely specified. For general globally hyperbolic spacetimes, the Cauchy surfaces are not as nicely ordered. Although they are indeed spatial slices of a 4-dimensional spacetime, it is impossible to define a universal concept of time on them. As a consequence, there is no conserved current that can function as Hamiltonian to evolve the spacetime between Cauchy surfaces. Without such a Hamiltonian, there is no single quantum field theory that is valid everywhere in the spacetime.

In the special case that the spacetime is stationary, there is a Killing vector k with respect to which the Lie derivative $\mathrm{L}_k(g_{\mu\nu})$ of the metric disappears. More graphically, the metric does not change under transformations along the integral curves of k. In the space of solutions $\mathcal{S}$ the solutions u_i are described by functions, so that the Lie derivative locally is a eigenvalue equation of the form

$$i\,\mathrm{L}_k u_i = i\,k^\mu \partial_\mu u_i = -i\,\omega_i u_i \tag{1.2.32}$$

Since both the Lie derivative and the Klein-Gordon operator $(\nabla^\mu \nabla_\mu - m^2)$ produce some sort of scalar when applied to solution ψ_i, they must commute. Hence, we can find simultaneous eigenmodes of the operators. The Lie derivative can then be regarded to map solutions of the KG equation to solutions, i.e. $\mathrm{L}_k : \mathcal{S} \mapsto \mathcal{S}$. L_k is anti-hermitian in $\mathcal{S}$, since $(\mathrm{L}_k\psi_i, \psi_j) = -\omega_i\delta_{ij} = -(\psi_i, \mathrm{L}_k\psi_j)$. It is a well-known result in matrix algebra that such operators can be diagonalised with purely imaginary eigenvalues, so that the eigenvalues ω_i are restricted to the reals, $\omega_i \in \mathbb{R}$. By definition, the Killing vector k is timelike asymptotically, so that its corresponding Lie derivative approaches ∂_t towards infinity. The eigenvalue equation Eq.(1.2.32) takes the form of a Schrödinger-like equation where we can identify the eigenvalues ω_i with a frequency. Since the eigenvalues of L_k are purely imaginary, complex conjugate solutions will have negative eigenvalues, i.e.

$$\begin{aligned} i\partial_t u_i &= -i\omega_i u_i \\ i\partial_t u_i^* &= i\bar{\omega}_i u_i^* \end{aligned} \tag{1.2.33}$$

where $\omega_i \geq 0$ and $\bar{\omega}_i \leq 0$. We can characterise the solutions to the Klein-Gordon equation as positive frequency solutions u_i and negative frequency solutions u_i^* with respect to some Killing time specified by vector k. Two eigenfunctions whose corresponding eigenvalues are distinct are by definition orthogonal, so that

$$(u_i, u_j^*) = 0.$$

Moreover, if we exclude the zero frequency mode with $\omega = 0$, we can choose to normalise the basis $\{u_i\}$ in such a way that $(u_i, u_j) = \delta_{ij}$, so that the positive frequency solutions indeed satisfy Eq.(1.2.24).

In stationary spacetimes then, we can restrict to the unique basis $\{u_i\}$ of 'positive frequency' eigenfunctions with respect to the Killing vector k, i.e.

$$k^\mu \partial_\mu u_i = -i\omega_i u_i, \qquad \omega_i > 0, \tag{1.2.34}$$

that are solutions of the Klein-Gordon equation with a purely positive-definite inner product $(u_i, u_j) = \delta_{ij}$ in order to define a legitimate Hilbert space $\mathcal{H}$ in which to perform quantum field theory. Similar to the Minkowski flat spacetime case, the negative frequency solutions have a linear correspondence with vectors of the complex conjugate Hilbert space $\mathcal{H}^* = \bar{\mathcal{H}}$ and can be incorporated as such. Any quantum scalar field $\phi(x)$ on a stationary spacetime can then be described by a linear combination of positive and negative frequency modes,

$$\phi(x) = \sum_i \left[a_i u_i(x) + a_i^\dagger u_i^*(x)\right], \tag{1.2.35}$$

similar to Eq.(1.2.29) valid in Minkowski spacetime. This may come as no surprise, as Minkowski spacetime is merely a special example of stationarity where the Killing vector $k = \partial_t$ is constant over the spacetime. The operators $\{a_i\}$ and their hermitian conjugates are annihilation and creation operators respectively, with which particles can be created from a unique vacuum $|0\rangle$. In a frame of reference that follows the integral curves of Killing vector k, there is time translation invariance so that this vacuum will remain an empty state. Physically, the vacuum is empty in the sense that it can be interpreted as the lowest energy state of the Hamiltonian

$$H = \int_{\Sigma_\tau} dS^\mu \, T_{\mu\nu} k^\nu, \tag{1.2.36}$$

where Σ_τ is some spacelike Cauchy surface of constant time $\tau = const.$ and $\hat{T}_{\mu\nu}$ is the stress-energy tensor of the spacetime. This Hamiltonian then dictates time evolution of all operators O to other Cauchy surfaces of constant time $\Sigma_{\tau'}$ in the past and future by the Heisenberg equation of motion,

$$\frac{dO}{d\tau} = i\,[H, O] \tag{1.2.37}$$

Excitations of the vacuum created by creation operators $a_i^\dagger$ can be interpreted as particles. The spectrum of $\mathcal{H}$ can thus be summarised as $\{|0\rangle, a_i^\dagger|0\rangle, a_i^\dagger a_j^\dagger|0\rangle, \ldots\}$ [45].

1.3 Sandwich Spacetime

Consider an evolving 'sandwich' spacetime M that is split up in three regions as in Fig.(4), so that $M = M_- \cup M_o \cup M_+$. If there is a scalar field ϕ on M, the Klein-Gordon equation must hold throughout all of M. For times $t < t_1$, the spacetime is stationary and admits an asymptotically timelike Killing vector k_-, so that there we can quantise the field by choosing positive frequency modes $\{f_i\}$ with respect to k_- satisfying $(k_-)^\mu \partial_\mu f_i = -i\omega_i f_i$ to define a unique local 'in' vacuum $|0\rangle_{in}$. For intermediate times $t_1 < t < t_2$ the spacetime has a time-dependent metric. In M_o, there are many bases $\{\psi\}$ that can be used for quantisation, so that there does not exist a single notion of particles. On M_+ for $t > t_2$ the spacetime is again stationary and has some asymptotically timelike Killing vector $k_+ \neq k_-$ different from that of M_-. On M_+, we can then quantise the scalar field using the unique basis of positive frequency modes $\{g_i\}$ with respect to k_+, that satisfy $(k_+)^\mu \partial_\mu g_i = -i\omega_i g_i$ [45].

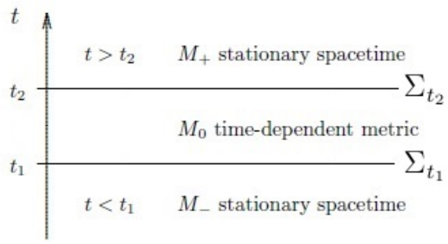


Figure 4: Sandwich Spacetime. For $t < t_1$, region M_- admits a Killing vector k_-. Between t_1 and t_2 spacetime region M_0 is evolving with a corresponding a time-dependent metric. For later times $t > t_2$ spacetime M_+ again settles down, returning to stationarity, with a Killing vector k_+.

On initial subregion M_- a general solution to the Klein-Gordon equation can be expressed as a linear combination of positive frequency solutions f_i, and negative frequency solutions f_i^*,

$$\phi(x) = \sum_i [a_i f_i + a_i^\dagger f_i^*]. \tag{1.3.1}$$

Instead, we can also choose the basis of solutions $\{g_i\}$ of the final stationary region and perform the alternative expansion

$$\phi(x) = \sum_i [b_i g_i + b_i^\dagger g_i^*]. \tag{1.3.2}$$

Since both sets of modes are complete, we can also express one set in terms of the other by a *Bogoliubov transformation.* Note that in general, it does not hold automatically that positive frequency solutions of one set can be expanded purely in terms of positive frequency solutions of the other. We have, for example,

$$g_i = \sum_j (A_{ij} f_j + B_{ij} f_j^*). \tag{1.3.3}$$

Applying the orthonormality conditions (1.2.24), the *Bogoliubov coefficients* A and B are easily extracted by taking the Klein-Gordon inner product between ingoing and outgoing modes,

$$A_{ij} = (g_i, f_j), \qquad B_{ij} = -(g_i, f_j^*). \tag{1.3.4}$$

Furthermore, since both the sets of modes have to satisfy the orthonormality conditions (1.2.24), the *Bogoliubov coefficients* A and B are faced with the restrictions

$$\begin{aligned} AA^\dagger - BB^\dagger &= \mathbb{I}, \\ AB^T - BA^T &= 0. \end{aligned} \tag{1.3.5}$$

Alternatively, we can also express the early solutions in terms of the later ones, leading to an inversion of Eq.(1.3.3), as

$$f_i = \sum_j (A'_{ij} g_j + B'_{ij} g_j^*). \tag{1.3.6}$$

Using this relation, we then find

$$\begin{aligned} g &= A(A'g + B'g^*) + B(A'^*g^* + B'^*g) \\ &= (AA' + BB'^*)g + (AB' + BA'^*)g^* \\ &= (AA^\dagger - BB^\dagger)g + (AB^T - BA^T)g^* = g \end{aligned}$$

only if $A' = A^\dagger$ and $B' = -B^T$ [45]. Expressed in terms of original Bogoliubov coefficients, we have

$$f_i = \sum_j (A_{ji}^* g_j + B_{ji} g_j^*). \tag{1.3.7}$$

Moreover, the alternative Bogoliubov coefficients A' and B' must also satisfy the conditions (1.3.5), so that we find the additional requirements

$$\begin{aligned} A^\dagger A - B^T B^* &= \mathbb{I}, \\ A^\dagger B - B^T A^* &= 0. \end{aligned} \tag{1.3.8}$$

Combining Eqs.(1.3.3) and (1.3.7) with Eqs.(1.3.1) and (1.3.2) respectively, we can express early annihilation operators a_i in terms of the late operators $b_i, b_i^\dagger$ and vice versa [10]. Remembering that the KG inner product is anti-linear in its second argument, we find

$$\begin{aligned} a_i &= (\phi, f_i) = \sum_j (A_{ji} b_j + B_{ji}^* b_j^\dagger), \\ b_i &= (\phi, g_i) = \sum_j (A_{ij}^* a_j - B_{ij}^* a_j^\dagger) \end{aligned} \tag{1.3.9}$$

If $B = 0$ then by the conditions (1.3.5) and (1.3.8) we find $AA^\dagger = A^\dagger A = \mathbb{I}$. The change of basis from $\{f_i\}$ to $\{g_i\}$ is nothing but a unitary transformation permuting the annihilation operators but leaving the vacuum state unchanged. If the coefficients B do not vanish, the in vacuum state $|0_{in}\rangle$ and the out vacuum state $|0_{out}\rangle$, defined by

$$\begin{aligned} a_i|0_{in}\rangle &= 0 \quad \forall i, \\ b_i|0_{out}\rangle &= 0 \quad \forall i, \end{aligned} \tag{1.3.10}$$

will not coincide. This discrepancy is even more obvious when inspecting the number operators N_i^{in} and N_i^{out},

$$\begin{aligned} N_i^{in} &= a_i^\dagger a_i \qquad \text{on } M_- \\ N_i^{out} &= b_i^\dagger b_i \qquad \text{on } M_+. \end{aligned} \tag{1.3.11}$$

For the i-th mode of k, the amount of particles in the i-th mode seen by an observer in final region M_+ in the initial vacuum $|0_{in}\rangle$ is

$$\begin{aligned} \langle 0_{in}|N_i^{out}|0_{in}\rangle &= \langle 0_{in}|b_i b_i^\dagger|0_{in}\rangle \\ &= \langle 0_{in}|\sum_j (A_{ji}a_j^\dagger - B_{ji}a_j)\sum_k (A_{ik}^* a_k - B_{ik}^* a_k^\dagger)|0_{in}\rangle \\ &= \langle 0_{in}|\sum_j (-B_{ji}a_j)\sum_k (-B_{ik}^* a_k^\dagger)|0_{in}\rangle \\ &= \sum_j B_{ij}B_{ij}^* = (BB^\dagger)_{ii}. \end{aligned} \tag{1.3.12}$$

The total number of particles detected by the observer is then $\sum_i \langle N_i^{out}\rangle = \mathrm{Tr}\,(BB^\dagger)$. Hence, if the matrix B is nonzero, the in vacuum is different from the out vacuum and there is geometrical particle production caused by the non-stationarity of the metric in the intermediate region M_0 of sandwich spacetime [10,45].

To gain more insight into the particle content that the in vacuum state $|0_{in}\rangle$ has for an observer in M_+, we rewrite its definition (1.3.10) by means of Eq.(1.3.9),

$$\begin{aligned} 0 &= a_i|0_{in}\rangle = \sum_j (A_{ji}b_j + B_{ji}^* b_j^\dagger)\,|0_{in}\rangle \\ &= (b_k + \sum_{ij} B_{ji}^* A_{ik}^{-1} b_j^\dagger)\,|0_{in}\rangle \\ &= (b_k + \sum_j V_{jk} b_j^\dagger)\,|0_{in}\rangle, \end{aligned} \tag{1.3.13}$$

where we used right multiplication by A_{ik}^{-1} to obtain the second line and have defined $V_{jk} \equiv \sum_i B_{ji}^* A_{ik}^{-1}$. As was shown by Wald, this definition is completely legitimate. The inverse matrix A^{-1} used in the definition of the symmetric matrix V is guaranteed to exist by the relations (1.3.5) and (1.3.8) [48]. For simplicity, first reduce Eq.(1.3.13) to the case of a single mode in k, i.e.

$$0 = (b + Vb^\dagger)|0_{in}\rangle. \tag{1.3.14}$$

We can see the in vacuum $|0_{in}\rangle$ as a function $\mathfrak{f}$ in b and $b^\dagger$ of the out vacuum $|0_{out}\rangle$. We then find

$$\begin{aligned} 0 &= (b + Vb^\dagger)\,|0_{in}\rangle \\ &= (b + Vb^\dagger)\,\mathfrak{f}\,|0_{out}\rangle \\ &= (\mathfrak{f}\,b + [b\,,\,\mathfrak{f}]\,) + Vb^\dagger\mathfrak{f}\,|0_{out}\rangle. \end{aligned}$$

Since operator b by definition annihilates the out vacuum, we have

$$[b, \mathfrak{f}] = -Vb^\dagger\mathfrak{f}. \tag{1.3.15}$$

The commutation relations (1.2.30) establish that $[b, b^\dagger] = 1$ is the only non-vanishing commutator between the available operators on which $\mathfrak{f}$ can depend. Eq.(1.3.15) then shows that $\mathfrak{f}$ must be some kind of polynomial in the creation operator $b^\dagger$, i.e. $\mathfrak{f} = \mathfrak{f}(b^\dagger)$. Furthermore, since the commutator $[b, b^\dagger]$ gives just a C-number, $\mathfrak{f}$ must be an exponential in this polynomial. We therefore try something of the form

$$|0_{in}\rangle = \mathfrak{f}\,|0_{out}\rangle = Ce^{\mu b^\dagger b^\dagger}|0_{out}\rangle, \tag{1.3.16}$$

as a solution to Eq.(1.3.14)[25]. In the proposed solution, C is a constant incorporated to normalise the vacuum to unity and μ a variable that is determined by Eq.(1.3.15),

$$\begin{aligned} -Vb^\dagger\mathfrak{f} &= [b,\mathfrak{f}] = C[b, e^{\mu b^\dagger b^\dagger}] \\ &= C\sum_n \frac{\mu^n}{n!}[b,(b^\dagger b^\dagger)^n] \\ &= C\sum_n \frac{\mu^n}{n!}[2nb^\dagger(b^\dagger b^\dagger)^{n-1}] \\ &= 2\mu C\, b^\dagger \sum_n \frac{\mu^{n-1}}{(n-1)!}(b^\dagger b^\dagger)^{n-1} \\ &= 2\mu\, b^\dagger \mathfrak{f}, \end{aligned} \tag{1.3.17}$$

where we have used the inductive relation $b(b^\dagger b^\dagger)^n = (b^\dagger b^\dagger)^n b + 2nb^\dagger(b^\dagger b^\dagger)^{n-1}$. Looking at the preceding equation, we see that we should set $\mu = -\frac{1}{2}V$, and we get

$$|0_{in}\rangle = Ce^{-\frac{1}{2}Vb^\dagger b^\dagger}|0_{out}\rangle, \tag{1.3.18}$$

Returning to the full equation (1.3.13), we compose the solution by summing over all the modes of k,

$$|0_{in}\rangle = C\exp\left(-\frac{1}{2}\sum_{ij} V_{ij}b_i^\dagger b_j^\dagger\right)|0_{out}\rangle. \tag{1.3.19}$$

An immediate consequence of the relation between the in and out vacua is that the amplitude for creation of an odd number of particles vanishes[25],

$$C\,\langle 0_{out}|(b_{k_1}b_{k_2}\dots b_{k_n})\exp\left(-\frac{1}{2}\sum_{ij} V_{ij}b_i^\dagger b_j^\dagger\right)|0_{out}\rangle = 0 \qquad n \text{ odd}. \tag{1.3.20}$$

Particles are produced in even numbers. In particular, the amplitude for pair production is just the matrix V_{ij} with normalisation C,

$$C\,\langle 0_{out}|b_{k_1}b_{k_2}\exp\left(-\frac{1}{2}\sum_{ij} V_{ij}b_i^\dagger b_j^\dagger\right)|0_{out}\rangle = C\,V_{k_1k_2}. \tag{1.3.21}$$

Moreover, for the state (1.3.19) to be well-defined in the out Hilbert space, it needs to be normalisable by adapting constant C. This is equivalent to the requirement

$$\sum_i \langle 0_{in}|N_i^{out}|0_{in}\rangle = \sum_{ij} B_{ij}B_{ij}^* < +\infty. \tag{1.3.22}$$

which implies that the total quantity of particles created during the non-stationary period in region M_0 is some finite number, irrespective of the amount of time that passes between t_1 and t_2[48].

In conclusion, there will be geometrical particle creation in spacetimes of sandwich form, in which a volatile, evolving period is asymptotically flanked by two stationary periods. Although we have focused on scalar fields, this results holds for all types of matter. Furthermore, these particles are always generated in particle-antiparticle pairs, because the creation amplitude vanishes for odd particle numbers.

1.4 Hawking Radiation

Consider the spacetime describing the gravitational collapse of a spherically symmetric massive star into a Schwarzschild black hole. Outside of the matter of the star this spacetime will have the Schwarzschild metric

$$ds^2 = -\left(1-\frac{2M}{r}\right)dt^2 + \left(1-\frac{2M}{r}\right)^{-1}dr^2 + r^2 d\Omega_2^2, \tag{1.4.1}$$

where $d\Omega_2^2 = d\theta^2 + \sin^2\theta d\varphi^2$ is the metric of the 2-sphere S^2. For null geodesics it holds that $ds^2 = 0$, so that

$$dt^2 = \left(1-\frac{2M}{r}\right)^{-2}dr^2. \tag{1.4.2}$$

Defining the *tortoise coordinate* r_* in such a way that null geodesics satisfy $dt^2 = dr_*^2$ and requiring r_* to be real and increasing with r, we obtain

$$dr_*^2 \equiv \left(1 - \frac{2M}{r}\right)^{-2} dr^2, \qquad r_* = r + 2M \ln\left(\frac{r - 2M}{2M}\right) \tag{1.4.3}$$

In terms of coordinates $(t, r_*, \theta, \varphi)$ the Schwarzschild metric becomes

$$ds^2 = \left(1 - \frac{2M}{r}\right)\left(-dt^2 + dr_*^2\right) + r^2 d\Omega_2^2, \tag{1.4.4}$$

Solutions of the form $t \pm r_* = const.$ to the defining equation $dt^2 = dr_*^2$ correspond to null geodesics going in and coming out of the black hole, respectively. This defines the *lightcone coordinates*

$$\begin{aligned} u &= t - r_* \qquad \text{with } u = const. \quad \leftrightarrow \quad \text{outgoing null geodesics} \\ v &= t + r_* \qquad \text{with } v = const. \quad \leftrightarrow \quad \text{ingoing null geodesics} \end{aligned} \tag{1.4.5}$$

As can be easily deduced from the Penrose diagram of the collapsing star spacetime in Fig.(5), there are numerous spatial hypersurfaces that intersect all causal curves. The collapsing star diagram is globally hyperbolic. Most notable are the two Cauchy surfaces

1. - $\Sigma_- = \mathfrak{I}^-$ at past infinity, from which all causal curves originate.
2. - $\Sigma_+ = \mathfrak{I}^+ \cup H^+$, registering causal curves that end up at the singularity as well as those that terminate at future infinity.

Although the Schwarzschild black hole solution to Einstein's equation is static by the existence of an asymptotically timelike Killing vector $k = \partial_t$ and invariance of the metric under time reversal $t \to -t$, the spacetime of the collapsing star is not. In general, the formation of a black hole involves some complicated and highly intricate dynamical processes that disturb the time translation invariance of the metric. As the spacetime is approximately Minkowskian far away from the black hole, stationarity is retrieved in the far past near $\mathfrak{I}^-$ and far future near $\mathfrak{I}^+$. In analogy with sandwich spacetime, modes from the far past pass through a non-stationary, evolving part of spacetime before ending up at a stationary far future. This will lead to geometrical particle production by the black hole [9]. While the collapse is a transient phenomenon, the infinite time dilation at the horizon as perceived by an observer at infinity means that created particles can take an arbitrary amount of time to escape the black hole region and travel to future infinity.

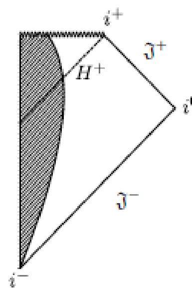


Figure 5: The Penrose diagram for the spacetime of a spherically symmetric star that undergoes gravitational collapse into a Schwarzschild black hole.

In the far past, all modes of the scalar field are specified by initial data on early Cauchy surface $\mathfrak{I}^-$ in a stationary part of spacetime. Hence, there is a preferred notion of time to define positive frequency modes with, so that the field expansion of ϕ uniquely defines the in vacuum. At late times, wave packets have moved to the future infinity $\mathfrak{I}^+$ or fallen into the black hole through H^+, so modes are determined by final data on $\mathfrak{I}^+ \cup H^+$. In summary, there will be three sets of modes:

$\{f_i\}$: positive frequency on $\mathfrak{I}^-$
$\{g_i\}$: positive frequency on $\mathfrak{I}^+$ and zero on H^+
$\{h_i\}$: 'positive frequency' on H^+ and zero on $\mathfrak{I}^+$

Since spacetime is not stationary near $H+$, there will not be a timelike Killing vector that can define a conserved Hamiltonian and thus no consistent way of doing quantum field theory. However, it turns out it is possible to draw physical conclusions that avoid this problem. The Hilbert space $\mathcal{H}_{\Sigma_+} = \mathcal{H}_{\mathfrak{J}^+} + \mathcal{H}_{H^+}$ of states on the composite Cauchy surface $\Sigma_+ = \mathfrak{J}^+ \cup H^+$ is naturally isomorphic to the tensor product space $\mathcal{H}_{\mathfrak{J}^+} \times \mathcal{H}_{H^+}$ and can thus be thought of as a joint system. The system describing the particles propagating out to infinity is completely separable from that of the particles falling into the black hole. Now, in general, for any state $|\Psi\rangle$ in a tensor product system $\mathcal{H}_A \times \mathcal{H}_B$ there is a *reduced density matrix* $\rho_A \in \mathcal{H}_A \times \mathcal{H}_A^*$ formed by taking the trace of $|\Psi\rangle \otimes |\Psi^*\rangle \in (\mathcal{H}_A \times \mathcal{H}_B) \times (\mathcal{H}_A^* \times \mathcal{H}_B^*) \cong (\mathcal{H}_A \times \mathcal{H}_A^*) \times (\mathcal{H}_B \times \mathcal{H}_B^*)$ with respect to a basis of $\mathcal{H}_B$. It may be defined as

$$\rho_A = \sum_i \langle e_{i_B}| \Big(|\Psi\rangle\langle\Psi| \Big) |e_{i_B}\rangle, \tag{1.4.6}$$

where $\{e_{i_B}\}$ is an orthonormal basis on $\mathcal{H}_B$ and the inner products on the right hand side are taken in $\mathcal{H}_A \times \mathcal{H}_B$. Now, a state like $|\Psi\rangle$ that is spread out over two separable Hilbert spaces has a Schmidt decomposition $|\Psi\rangle = \sum_i \lambda_i |e_{i_A}\rangle \otimes |e_{i_B}\rangle$ onto orthonormal bases $\{e_{i_A}\}$ and $\{e_{i_B}\}$ of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. We can rewrite definition (1.4.6) into

$$\rho_A = \sum_i \lambda_i^2 \, |e_{i_A}\rangle\langle e_{i_A}|. \tag{1.4.7}$$

It follows that the expectation value of any observable O_A for the subsystem $\mathcal{H}_A$ is given by

$$\begin{aligned}
\langle O_A \rangle &= \langle \Psi | O_A | \Psi \rangle \\
&= \sum_{ij} \lambda_i \lambda_j \Big(\langle e_{i_A}| \otimes \langle e_{i_B}| \Big) O_A \Big(|e_{j_A}\rangle \otimes |e_{j_B}\rangle \Big) \\
&= \sum_{ij} \lambda_i \lambda_j \Big(\langle e_{i_A}| O_A | e_{j_A}\rangle \Big) \Big(\langle e_{i_B} | e_{j_B}\rangle \Big) \\
&= \sum_i \lambda_i^2 \, \langle e_{i_A} | O_A | e_{j_A} \rangle \\
&= \mathrm{Tr} \Big(\sum_i \lambda_i^2 \, |e_{j_A}\rangle\langle e_{i_A}| O_A \Big) = \mathrm{Tr}\,(\rho_A O_A).
\end{aligned} \tag{1.4.8}$$

Since the probabilities of the possible results of any experiment on subsystem $\mathcal{H}_A$ can be expressed in terms of expectation values of such observables, all information about $\mathcal{H}_A$ can be retrieved from reduced density matrix ρ_A. The behaviour of system $\mathcal{H}_A$ is described completely by ρ_A, while leaving undetermined the properties of the composite system $\mathcal{H}_A \times \mathcal{H}_B$. In our case, we can 'trace out' all information belonging to particles falling into the black hole to find a reduced density matrix that describes only behaviour of particles propagating towards $\mathfrak{J}^+$. As such, the appearance of ρ_A instead of a regular state vector reflects our incapability to recover all information about the whole system once the particle states have traversed the horizon.

Now, the creation and annihilation operators corresponding to infalling particles will be subjected to a Bogoliubov transformation, whereas the particles propagating to $\mathfrak{J}^+$ never enter the intermediate non-stationary region of spacetime and their associated vacuum does not undergo a change of definition. This will induce a transformation of the state vector $|\Psi\rangle \in \mathcal{H}_{\Sigma_+}$ of the form $|\Psi\rangle \to \Psi' = \mathfrak{f}\Psi$, where $\mathfrak{f}$ is an endomorphism on joint system $\mathcal{H}_{\Sigma_+}$ of the form

$$\mathfrak{f} = I_{\mathcal{H}_{\mathfrak{J}^+}} \times U_{\mathcal{H}_{H^+}}, \tag{1.4.9}$$

where I is the identity operator on $\mathcal{H}_{\mathfrak{J}^+}$ and U an unitary operator that acts on $\mathcal{H}_{H^+}$. Since a unitary transformation preserves the inner product, we have $\langle \Psi | O_A | \Psi \rangle = \langle \Psi' | O_A | \Psi' \rangle$. We see by Eq.(1.4.8) that the density matrix $\rho'_{\mathcal{H}_{\mathfrak{J}^+}}$ corresponding to $|\Psi'\rangle$ will equal the original matrix $\rho_{\mathcal{H}_{\mathfrak{J}^+}}$. The density matrix remains unchanged by transformations (1.4.9) and does not depend on the definition of 'positive frequency' modes on H^+ [49]. Therefore, we can arbitrarily appoint some modes as positive frequency and use those to define a complete set on composite Cauchy surface $\Sigma_+ = \mathfrak{J}^+ \cup H^+$.

Like the case of sandwich spacetime, we can expand the scalar field in terms of both the complete sets of modes

$$\phi = \sum_i [a_i f_i + a_i^\dagger f_i^*] = \sum_i [(b_i g_i + b_i^\dagger g_i^*) + (c_i h_i + c_i^\dagger h_i^*)]. \tag{1.4.10}$$

The future basis $\{g_i, g_i^*, h_i, h_i^*\}$ is related to the past basis $\{f_i, f_i^*\}$ by Bogoliubov transformations of the form

$$g_i = \sum_j [A_{ij} f_j + B_{ij} f_j*], \qquad h_i = \sum [C_{ij} f_j + D_{ij} f_j*], \tag{1.4.11}$$

because the Hilbert spaces on $\mathfrak{I}^+$ and H^+ are completely separable[48]. Late time observers at future infinity $\mathfrak{I}^+$ inspecting the in vacuum state defined by $a_i|0_{in}\rangle = 0 \quad \forall\, i$ will detect $\langle N_i^{out}\rangle = (BB^\dagger)_{ii}$ particles in the i-th mode[9,31].

The exact form of the Klein-Gordon equation (1.2.2) in Schwarzschild spacetime with metric (1.4.4) is more conveniently found by the using the formula

$$\Box = \nabla^\mu \nabla_\mu = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} g^{\mu\nu} \partial_\nu). \tag{1.4.12}$$

In addition expanding the scalar field solution in spherical harmonics as $\phi(t, r_*, \theta, \varphi) = \chi_l(r_*, t) Y_{lm}(\theta, \varphi)$, the equation of motion for the scalar field becomes

$$\partial_t^2 \chi_l - \partial_{r_*}^2 \chi_l + \left(1 - \frac{2M}{r}\right) \left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3} + m^2\right] \chi_l = 0. \tag{1.4.13}$$

This equation is the Klein-Gordon equation (1.2.25) for a massless scalar field in a 2-dimensional Minkowski spacetime (with Cartesian coordinates t, r_*) with a scalar potential

$$V_l(r_*) = \left(1 - \frac{2M}{r}\right) \left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3} + m^2\right]. \tag{1.4.14}$$

We can depict the evolution of the spacetime by propagating wave packets. Classically, the black hole does not allow escape once the horizon is passed. Waves propagating into the future from past infinity can literally disappear from the theory at the singularity. It is thus more appropriate to describe the spacetime by waves moving back in time from future infinity.

As $r_* \to -\infty$ (i.e. as $r \to 2M$ by Eq.(1.4.3)) the potential $V_l(r_*)$ behaves as $(1 - 2M/r) \sim \exp(r_*/2m)$, falling off exponentially with r_*. For $r_* \to +\infty$ (i.e. for $r \to +\infty$) there are two possible outcomes, depending on the mass m of the scalar field. If $m \neq 0$, $V_l(r_*)$ decreases as $(m^2 - 2Mm^2/r) \sim (m^2 - 2Mm^2/r_*)$, whereas for a massless scalar field $V_l(r_*) \sim l(l+1)/r^2 \sim (l(l+1)/r_*^2$ in this limit. In the case of the massive scalar field, every wave packet should behave like a free ($V_l = 0$) massless solution in (t, r_*) space propagating in from ($r_* \to -\infty$) combined with a massive solution coming in from ($r_* \to +\infty$) in the far past. In the asymptotic future, the situation is reversed, with every wave packet being a combination of a free massless solution travelling to $r_* \to -\infty$ and a massive wave propagating to $r_* \to +\infty$, as in Fig.(6a)[49]. For an asymptotically flat Schwarzschild black hole, the Hawking radiation is massively dominated by low angular momentum modes with low l[27]. For a typical $l = 0$ mode of a massless scalar field, the potential tends to zero both near the horizon and close to spatial infinity, obtaining the form of a potential barrier surrounding the black hole region, as in Fig.(6b). Hence, any solution in (t, r_*) space evolving back in time will partly be transmitted and partly scatter off the potential[9].

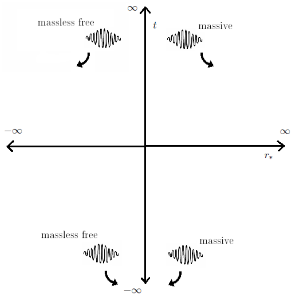
Near to $\mathfrak{I}^\pm$, where the potential tends to zero, massless solutions to Eq.(1.4.13) behave as free waves of the form $[a \exp(-i\omega u) + b \exp(-i\omega v)]$ in (t, r_*) spacetime, where u, v are the lightcone coordinates given by Eq. (1.4.5). We define 'early modes'

$$\begin{aligned} f_{i\,+} &= f_{(lm\omega)\,+} \sim f_{\omega\,+} = e^{-i\omega u} && \text{(outgoing at the horizon)} \\ f_{i\,-} &= f_{(lm\omega)\,-} \sim f_{\omega\,-} = e^{-i\omega v} && \text{(ingoing at the horizon)} \end{aligned} \tag{1.4.15}$$

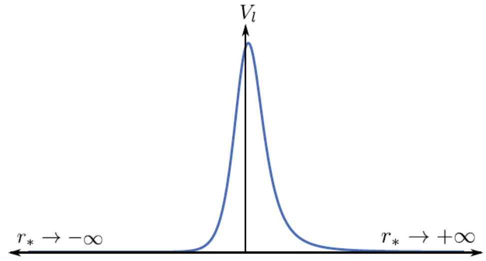
on $\mathfrak{I}^-$ and 'late modes'

$$\begin{aligned} g_{i\,+} &= g_{(lm\omega)\,+} \sim g_{\omega\,+} = e^{-i\omega u} && \text{(outgoing at the horizon)} \\ g_{i\,-} &= g_{(lm\omega)\,-} \sim g_{\omega\,-} = e^{-i\omega v} && \text{(ingoing at the horizon)} \end{aligned} \tag{1.4.16}$$

on $\mathfrak{I}^+$. As we trace back time from $\mathfrak{I}^+$, outgoing solutions will travel towards decreasing r_* and encounter the potential barrier. Part of the wave, say $g_{i\,+}^{(r)}$, is reflected by the potential, whereas the remaining part



(a) The behaviour of the massive scalar field propagating in two-dimensional (t, r_*) space under potential $V_l(r_*)$.



(b) The potential surrounding the black hole horizon for a low angular momentum mode of a massless scalar field.

Figure 6: The potential faced by a scalar field in the spacetime of a star collapsing into a Schwarzschild black hole takes on a different form depending on the mass of the particle. This highly influences the behaviour of the wave packets.

$g^{(t)}_{i+}$ is transmitted (Fig.7). Since $g^{(r)}_{i+}$ never enters the non-stationary part of spacetime, it will arrive at $\mathfrak{I}^-$ undisturbed. This will correspond to a term $B = 0$, $A_{\omega\omega'} \propto \delta(\omega - \omega')$ in Eq.(1.4.11). The transmitted part $g^{(t)}_{i+}$ of the plane wave passes through the barrier and will enter the collapsing matter, where it will be distorted before emerging and reaching $\mathfrak{I}^-$ [9].

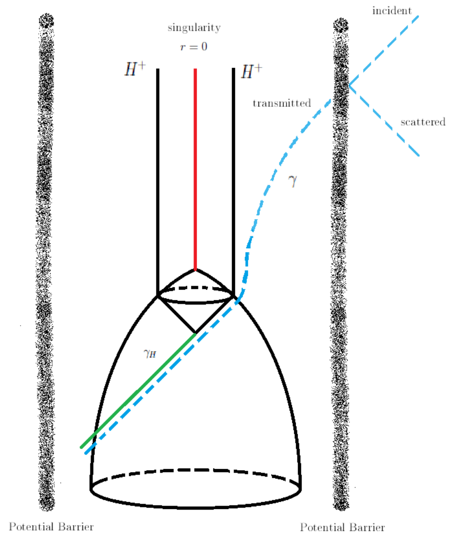


Figure 7: An outgoing plane wave g_{i+} propagating back in time from $\mathfrak{I}^+$ is scattered off or transmitted through the potential barrier surrounding the black hole, with both parts eventually reaching $\mathfrak{I}^-$.

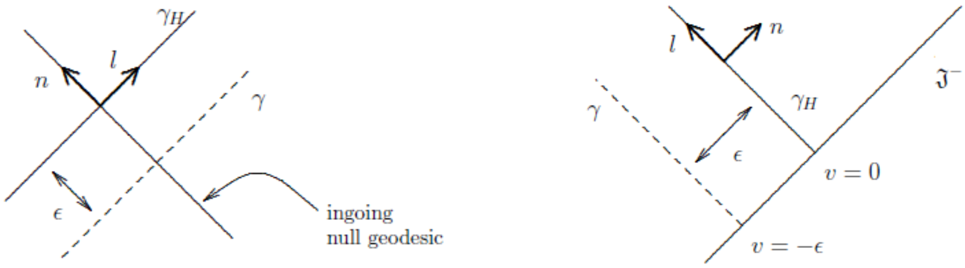
Consider the nature of a purely outgoing mode g_{i+} on $\mathfrak{I}^+$. Such a plane wave mode has some definite frequency ω, but is completely delocalised in (t, r_*) space (in accordance with the uncertainty principle). By superposing outgoing modes with different frequencies we may construct wave packets centered around some frequency ω_0 and coordinate u_0 that are localised on $\mathfrak{I}^+$. These wave packets vanish upon approaching the horizon as $u \to \infty$. Nonetheless, the frequency of individual modes diverges to infinity on the horizon as a consequence of the gravitational blueshift caused by the black hole. In terms of the Kruskal coordinate $U = -e^{-\kappa u}$, where $\kappa = 1/4M$ is the surface gravity of the black hole, these solutions are of the form e^{iS} with $S = \omega\kappa^{-1} \ln(-U)$. Let $\mu(\lambda)$ be any geodesic crossing the horizon with affine parameter λ chosen to have its zero value at the intersection with the horizon. Since λ depends smoothly on U and satisfies $dU/d\lambda \neq 0$ on the horizon where $U = 0$, in this region U has to depend linearly on λ. It follows that near $\lambda = 0$ the outgoing modes oscillate as $\exp[i\omega\kappa^{-1} \ln(-\alpha\lambda)]$, where $\alpha = dU/d\lambda|_{\lambda=0}$. Because $\alpha \neq 0$, S diverges at $U = 0$. Single modes observed by a freely falling observer crossing the horizon at $\lambda = 0$ have an 'infinite oscillation' singularity [49].

Let γ_H be a null geodesic generator of the horizon that is continued into the past until it intersects with $\mathfrak{I}^-$ at v_0. Without loss of generality, we may set $v_0 = 0$, as the spacetime is invariant under translations $v \rightarrow v + const.$[9]. Since the locally measured frequency of the plane wave diverges to infinity at γ_H, the *geometrical optics approximation* is valid: if $\Phi = Ae^{iS}$ and $|S| \gg 1$ and S is varying rapidly compared to $\ln A$, then the Klein-Gordon equation $\nabla^\mu \Phi \nabla_\mu \Phi = 0$ implies that

$$\nabla^\mu S \nabla_\mu S \cong 0. \tag{1.4.17}$$

Thus, to an approximation that becomes more and more exact as we get closer to γ_H, the outgoing plane wave g_{i+}, will have the form $G_i\, e^{iS}$, where G_i is a constant. Furthermore, by Eq.(1.4.17) the surfaces of constant phase S are null surfaces near γ_H since their normal vectors are null.

Because S diverges to infinity at γ_H, let l instead be a tangent null vector to inverse surfaces $S^{-1} = const.$, so that $l = hdS^{-1}$ for some $h \neq 0$. For such a tangent vector l, the autoparallel equation holds for transport along γ_H, i.e. $l \cdot \nabla l^\mu = 0$. Also, let n be a future-directed null vector that is parallelly transported along γ_H like l. Then, $l \cdot \nabla n^\mu = 0$, with $l \cdot n \neq 0$. For every $p \in \gamma_H$, consider a point p' an infinitesimal distance $-\epsilon$ away from γ_H along the null geodesic to which n is tangent. The collection of these points form a null curve $\gamma = \gamma(\epsilon)$ (see Fig.(8a)).



(a) The connecting vector n that is parallelly transported along null integral curve γ_H of l defines a new null curve γ, a distance $-\epsilon$ away.

(b) At $\mathfrak{I}^-$, the ingoing lightcone coordinate v is an affine parameter on integral curves of connecting vector n and γ crosses $\mathfrak{I}^-$ at $v = -\epsilon$.

Figure 8: Tangent vector l and future-directed null vector n, which is parallelly transported along γ_H, near H^+ and $\mathfrak{I}^-$.

On this curve γ, we obtain

$$S^{-1}\big|_\gamma = -\epsilon\, n \cdot \partial\, S^{-1}\big|_{\gamma_H}, \tag{1.4.18}$$

as in the limit $\epsilon \rightarrow 0$ we have $S^{-1}\big|_{\gamma_H} = 0$. Furthermore, we can show that

$$l \cdot \nabla \left(l \cdot \partial S^{-1}\right)\big|_{\gamma_H} = 0. \tag{1.4.19}$$

It follows that $l \cdot \partial S^{-1}\big|_{\gamma_H}$ is parallelly transported along γ_H. By Eq.(1.4.18), this means that S^{-1}, and by extension, S, are constant along γ[31]. Hence, the shape of the wave near $v_0 = 0$ at $\mathfrak{I}^-$ can be obtained by continuing the null geodesic generators of the surfaces of constant S back to $\mathfrak{I}^-$[49]. At $\mathfrak{I}^-$, the ingoing lightcone coordinate v is an affine parameter on integral curves of connecting vector n and γ crosses $\mathfrak{I}^-$ at $v = -\epsilon$ as in Fig.(8b). In contrast, near the horizon lightcone coordinate u diverges to infinity so that Kruskal coordinate $U = -e^{-\kappa u}$ affinely defines distance $-\epsilon$ along connecting vector n. Since the distance $-\epsilon$ is taken to be infinitesimal, U is linear in ϵ in close approximation. Coordinate U is semi-negative definite by definition and has a positive slope $\alpha = dU/d\lambda|_{H^+}$ along integral curves of n near the horizon, such that $U = -\alpha\epsilon$ on curve γ (Fig.(9)).

In conclusion, the outgoing plane wave mode may approximately be represented as $G_i\, e^{iS}$ near the horizon H^+, which has a null generator γ_H. Surfaces of constant phase S are null by the geometrical optics approximation. Then, let γ be the null geodesic generator of one such surface of constant phase S_γ. On γ, the Kruskal coordinate is given by $U = -\alpha\epsilon$, so that the phase of the wave mode can be described by $S_\gamma = \omega\kappa^{-1} \ln(-U) = \omega\kappa^{-1} \ln(\alpha\epsilon)$. Continuing γ back to $\mathfrak{I}^-$, we find that it intersects with $\mathfrak{I}^-$ at $v = -\epsilon$. On $\mathfrak{I}^-$, the phase of the wave is $S_\gamma = \omega\kappa^{-1} \ln(-\alpha v)$. Furthermore, only the reflected part $g_{i+}^{(t)}$ can reach $\mathfrak{I}^-$ for $v > 0$. The transmitted part $g_{i+}^{(t)}$ of the wave will be of the form

$$g_{i+}^{(t)} = \begin{cases} 0 & v > 0 \\ G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)} & v < 0,\ ||v||\ \text{small}, \end{cases} \tag{1.4.20}$$

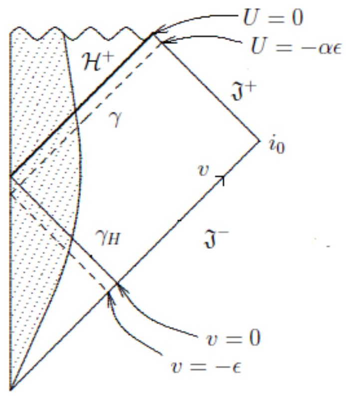


Figure 9: Penrose diagram of the scattering of an outgoing plane wave mode g_{i+}

once it has arrived at $\mathfrak{J}^-$ [31].

We wish to express $g^{(t)}_{i+}$ in terms of the complete basis $\{f_j \sim e^{-i\omega' v}, f^*_j \sim e^{i\omega' v}\}$ provided by the outgoing early modes, so that we may find Bogoliubov coefficient B in the continuous limit

$$g^{(t)}_{i+}(\omega) = \int_0^\infty d\omega' \, (A_{\omega\omega'} f_{\omega'} + B_{\omega\omega'} f^*_j), \tag{1.4.21}$$

of Eq.(1.4.11) that will allow us to compute the number of particles a late observer sees in the early in vacuum. By comparison with the frequency Fourier transform

$$\begin{aligned} g^{(t)}_{i+}(\omega) &= \int_{-\infty}^\infty d\omega' \, e^{-i\omega' u} \tilde{g}^{(t)}_{i+}(\omega') \\ &= \int_0^\infty d\omega' \, e^{-i\omega' u} \tilde{g}^{(t)}_{i+}(\omega') + \int_0^\infty d\omega' \, e^{i\omega' u} \tilde{g}^{(t)}_{i+}(-\omega'), \end{aligned} \tag{1.4.22}$$

we find

$$A_{\omega\omega'} = \tilde{g}^{(t)}_{i+}(\omega') \qquad \text{and} \qquad B_{\omega\omega'} = \tilde{g}^{(t)}_{i+}(-\omega'). \tag{1.4.23}$$

If we then manage to encounter a way to relate $\tilde{g}^{(t)}_{i+}(\omega')$ to $\tilde{g}^{(t)}_{i+}(-\omega')$, the Bogoliubov coefficients remain determined completely by the condition $AA^\dagger - BB^\dagger = \mathbb{I}$ from Eq.(1.3.5) [9].

To find such a relation, we use another Fourier transform to write

$$\begin{aligned} \tilde{g}^{(t)}_{i+}(\omega') &= \int_{-\infty}^\infty dv \, e^{i\omega' v} g^{(t)}_{i+}(v) \\ &\approx \int_{-\infty}^0 dv \, e^{i\omega' v} G_i \, e^{i\kappa^{-1}\omega \ln(-\alpha v)}. \end{aligned} \tag{1.4.24}$$

This approximation is only valid for large $|\omega'|$, for which $\exp(i\omega' v)$ oscillates rapidly and there are cancellations for the large v modes not described by solution (1.4.20). However, remember that we are interested in evolving back 'late time' wave packets on $\mathfrak{J}^+$ that are localised around some large value $u = u_0$ and have a small spread Δu. Once these wave packets have arrived on $\mathfrak{J}^-$ they are centered around some $v = v_0$ with width Δv. Here,

$$-u_0 = \kappa^{-1} \ln -\alpha v_0 \qquad \text{and thus } -\alpha v_0 = e^{-\kappa u_0}, \tag{1.4.25}$$

so that v_0 is small. Moreover, for a deviation Δu, we find

$$\Delta u = -\kappa^{-1} \ln[-\alpha(v_0 + \Delta v)] - u_0 = \kappa^{-1} \ln\left(\frac{v_0}{v_0 + \Delta v}\right) = -\kappa^{-1} \ln\left(1 + \frac{\Delta v}{v_0}\right) \simeq \frac{\Delta v}{\kappa |v_0|}, \tag{1.4.26}$$

by a Taylor expansion of $\ln(1+x)$ for small x. Then, we see $\Delta v \simeq \kappa |v_0| \Delta u \ll \Delta u$ and we conclude that the large u_0, small Δu wave packets from $\mathfrak{J}^+$ have morphed into a form with very narrow spread Δv around small v_0 at $\mathfrak{J}^-$. Thus, the physically interesting wave packets investigated here are mainly high frequency modes on $\mathfrak{J}^-$ [31].

By simply applying Eq.(1.4.24), we obtain

$$\tilde{g}^{(t)}_{i+}(-\omega') = \int_{-\infty}^{\infty} dv\, e^{-i\omega' v} g^{(t)}_{i+}(v) \approx \int_{-\infty}^{0} dv\, e^{-i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)}. \quad (1.4.27)$$

The complex logarithmic function $\ln(-z)$ in this equation is multiple-valued. If a complex number is written in polar form as $z = re^{i\theta}$, then its logarithm $\ln z = \ln r + i\theta$ is invariant under transformations $\theta \to \theta + 2\pi n$ for $n \in \mathbb{N}$. The logarithmic function has its branch point at the origin and branches for all values of r. Hence, we may take a branch cut to lay on the positive real axis. This requires us to deform the integration path to avoid the branch cut. This can be done by shifting the contour an infinitesimal distance $i\epsilon$ upwards in the complex plane, as in Fig.(10), and taking the limit $\epsilon \to 0$ at the end of the calculation. Because the integral has no further poles inside of the closed contour, it defines a region

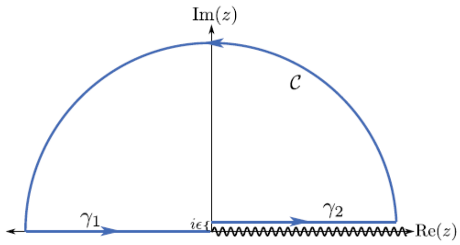


Figure 10: The contour integral used to evaluate $\tilde{g}^{(t)}_{i+}(-\omega')$ with a branch cut of $\ln(-z)$ along the positive real axis.

of the complex v-plane where the function is holomorphic. This means we can use Cauchy's integral theorem:

$$0 = \oint dv\, e^{-i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)} = \left\{ \int_{\gamma_1} + \int_{\gamma_2} + \int_{\mathcal{C}} \right\} dv\, e^{-i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)} \quad (1.4.28)$$

Without loss of generality, we may make the assumption that $\omega' > 0$, so that the integral decays exponentially for $\mathrm{Im}\, v > 0$ and the integration over arc segment $\mathcal{C}$ vanishes correspondingly as we take the arc radius to infinity. It follows from Eq.(1.4.28) that $\int_{\gamma_1} = -\int_{\gamma_2}$. As a consequence, we can rewrite Eq.(1.4.27) to

$$\begin{aligned}
\tilde{g}^{(t)}_{i+}(-\omega') &\approx \int_{-\infty}^{0} dv\, e^{-i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)} \\
&= -\int_{i\epsilon}^{\infty + i\epsilon} dv\, e^{-i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(-\alpha v)} && \left(\int_{\gamma_1} = -\int_{\gamma_2} \right) \\
&= -\int_{-\infty - i\epsilon}^{-i\epsilon} dv\, e^{i\omega' v} G_i\, e^{i\kappa^{-1}\omega \ln(\alpha v)} && (v \to -v) \\
&= -\int_{-\infty}^{0} dv\, e^{i\omega' v} G_i\, e^{i\kappa^{-1}\omega[\ln(-\alpha v) + i\pi]} && (\epsilon \to 0) \\
&= -e^{-\pi\omega/\kappa}\, \tilde{g}^{(t)}_{i+}(\omega') && (1.4.29)
\end{aligned}$$

Feeding this result into Eq.(1.4.23), we relate the Bogoliubov coefficients as

$$B_{ij} = -e^{-\frac{\pi\omega_i}{\kappa}} A_{ij}. \quad (1.4.30)$$

Now, by the condition $AA^\dagger - BB^\dagger = \mathbb{I}$ we obtain

$$\delta_{ij} = \sum_k (A_{ik}A^*_{jk} - B_{ik}B^*_{jk}) = \sum_k \left(e^{\frac{\pi(\omega_i + \omega_j)}{\kappa}} - 1 \right), \quad (1.4.31)$$

whence we conclude that the number of particles of the i-th wave mode detected by an observer at $\mathfrak{J}^+$ in the in vacuum as valid at $\mathfrak{J}^-$ is

$$\langle N^{out}_i \rangle = (BB^\dagger)_{ii} = \frac{1}{e^{\frac{2\pi\omega_i}{\kappa}} - 1}. \quad (1.4.32)$$

This is the Planck spectrum for black body radiation at the Hawking temperature [9,31]

$$T_H = \frac{\kappa}{2\pi} = \frac{1}{8\pi M}. \tag{1.4.33}$$

Notice that this coincides with the black hole temperature predicted from entropy principles in Eq.(1.1.6), fixing the constant of proportionality. As expected, including quantum effects in the model of a star collapsing into a Schwarzschild black hole evokes the pair production of scalar particles that propagate to infinity in the form of *Hawking radiation*. More significantly, it is possible to generalise this result to stars that are not spherically symmetric when collapsing and any other kind of resultant black hole [9,31]. Furthermore, it can be shown that the same effect occurs for other types of fields, for which the out vacuum will be a state containing particle-antiparticle excitations of the early in vacuum [34].

1.5 Entanglement of Hawking Pairs

In order to study more closely the nature of the Hawking pairs making up the radiation emitted by a Schwarzschild black hole, we take a more intuitive approach based on their wavelike characteristics. Consider the evolution of ingoing positive frequency modes starting at past null infinity $\mathfrak{J}^-$. Suppose these waves reach $r = 0$ during the period the black hole is formed, so that they scatter back towards $\mathfrak{J}^+$ as outgoing in vacuum modes. From subsection (1.3), we know that there is particle production when positive frequency modes evolve into a mix of positive and negative frequency modes and leave a non-vanishinig Bogoliubov coefficient B. Hence, particle creation is possible only if the wave modes are distorted. We can classify these outgoing modes by the way they stretch between Cauchy surfaces Σ. The geometrical optics approximation, which was justified by the infinite blueshift of Hawking radiation at the horizon, then stipulates that the phase S of these wave modes near the horizon stays constant along null hypersurfaces, which can be depicted as the trajectory followed by light rays (Fig.(11)) [25].

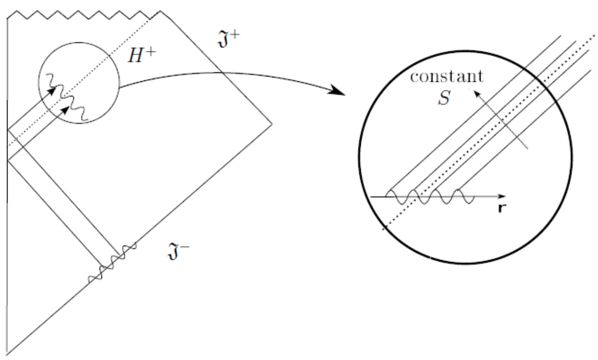


Figure 11: Penrose diagram detailing the evolution of a positive frequency mode starting at $\mathfrak{J}^-$, scattering back at $r = 0$ and propagating towards $\mathfrak{J}^+$ with constant phase S along null hypersurfaces.

Consider the immediate vicinity of the horizon closely after the black hole has formed, as sketched in the close-up of Fig.(11). For black holes that are sufficiently large, this region of spacetime has very small curvature. Wave modes in this region will still be in the in vacuum state as defined on $\mathfrak{J}^-$. Furthermore, on the portions of the 'initial' Cauchy surfaces that are straddling the horizon the null Kruskal coordinates U ,V may be used as an affine parameter. On late Cauchy surfaces near $\mathfrak{J}^+$, however, spacetime is approximately Minkowskian and the canonical null coordinates

$$X^+ = t + r, \qquad X^- = t - r \tag{1.5.1}$$

can be used to describe the form of wave modes. Because the initial Cauchy surfaces are connected to the late Cauchy surfaces by outgoing null rays, we can evolve the wave modes from early to later surfaces by keeping the phase S constant. Then, the shape of the wave modes at the later Cauchy surfaces is determined by the function

$$X^- = X^-(U), \tag{1.5.2}$$

i.e. the relation between the coordinates X^- and U. Since U is defined as $U = -\exp(-u/4M) = [(r - 2M)/2M]^{1/2} \exp(-X^-/4M)$, the function (1.5.2) is logarithmic, $X^- M \sim \ln -U$. Notice, however, that the Kruskal coordinate U is dimensionless, whereas we would prefer to work with a coordinate that has units of length like Minkowskian coordinate X^-. Considering that the only natural length scale in the Schwarzschild geometry is given by M, we can define a new null coordinate y^- by

$$X^- = -M \ln\left(-\frac{y^-}{M}\right) \qquad \text{for } -M \lesssim y^- < 0. \tag{1.5.3}$$

The boundaries of validity $|y^-| \ll M$ of this equation can be justified by setting $y^- = 0$ on the horizon. Then, $y^- < 0$ for spacetime points outside the horizon and $y^- > 0$ for events inside the horizon. As such, events for which $y^- \lesssim M$ will no longer be in close proximity to the horizon, so that U is no longer an affine parameter and the meaning of y^- becomes obscure. Here, spacetime is approximately flat, so the relation (1.5.3) will now be of linear form instead,

$$X^- = -y^- \qquad \text{for } y^- \lesssim -M. \tag{1.5.4}$$

The shape of wave modes starting out on a part of the initial Cauchy surface that is far from the horizon is thus determined by an evolution along outgoing null geodesics that are, in ingoing Eddington-Finkelstein (EF) coordinates where $t_* = v - r$, infinitesimally close to the straight lines of Minkowski space. Such wave modes are undisturbed by the gravitational forces of the black hole, and no particles are produced (Fig. (12a)).

Now consider the spacetime events in the range $y^- > 0$ that are on the inside of the black hole horizon. Here, there are no natural coordinates that are preferred in the physical sense. However, as we have shown in subsection (1.4), the final result of the calculation will not depend on the choice of wave modes inside the horizon, so we are free to allocate them arbitrary coordinates. This allows us to define a coordinate Y on the slice of a Cauchy surface inside the horizon that linearly measures distances along hypersurface. We then assign to every null ray a null coordinate Y^- that equals Y at the intersection with the hypersurface slice. Y^- is on equal footing with coordinate X^- outside the horizon, with the exception that X^- is a physically justified choice because it can be used to define real particles at $\mathfrak{I}^+$ with the correct energy-momentum tensor. The goal is now to find out how coordinate Y^- evolves from its value $Y^- = Y$ where it crosses the Cauchy surface back to the vicinity of the horizon, where U is an affine parameter. Choosing coordinate Y^- in this way, we should find a relation similar to that of X^- with U, i.e. $Y^- \sim \ln U$. Then, by the way y^- is defined, we find the dimensionful equation

$$Y^- = -M \ln\left(\frac{y^-}{M}\right) \qquad \text{for } y^- > 0. \tag{1.5.5}$$

Hence, any distortion of wave modes that depart closely outside the horizon is mirrored by deformations of wave modes on the inside[25].

To study the behaviour of wave modes in the range $-M \lesssim y^- < 0$, we make the assumption that their initial wavelength λ_1 is much smaller than the scale M over which the black hole metric is curved,

$$\lambda_1 = \epsilon M, \qquad \epsilon \ll 1, \tag{1.5.6}$$

so that their oscillations look like those on a piece of Minkowski spacetime. As the wavelength of these modes will increase as they propagate away from the black hole towards $\mathfrak{I}^+$, they will eventually become of order M near infinity. Although the assumption (1.5.6) may thus lead of order unity imprecisions in the overall radiation rate, it does not interfere with a qualitative description of their evolution. Such small λ wave modes will have a large number of oscillations in their range,

$$\mathcal{N}_{oscillations} = \frac{\Delta y^-}{\lambda_1} = \frac{M}{\epsilon M} \gg 1 \tag{1.5.7}$$

and so the width of one oscillation is of order ϵ. Consider such a narrow oscillation ς_1 extending over the range $-\alpha < y^- < -\alpha + \epsilon$ with $\alpha > 0$. In the frame of reference associated with coordinate X^-, the wavelength λ_{ς_1} will then be

$$\lambda_{\varsigma_1} = |\delta X^-| = \left|\left(\frac{dX^-}{dy^-}\right)\delta y^-\right| = \left[M\left(\frac{M}{|y^-|}\right)\frac{1}{M}\right]\epsilon = M\frac{\epsilon}{\alpha}. \tag{1.5.8}$$

An adjacent oscillation ς_2 closer to the horizon will span the range $-\alpha - \epsilon < y^- < -\alpha$, so that it will have wavelength

$$\lambda_{\varsigma_2} = \left[M\left(\frac{M}{|y^-|}\right)\frac{1}{M}\right]\epsilon = M\frac{\epsilon}{\alpha - \epsilon} \tag{1.5.9}$$

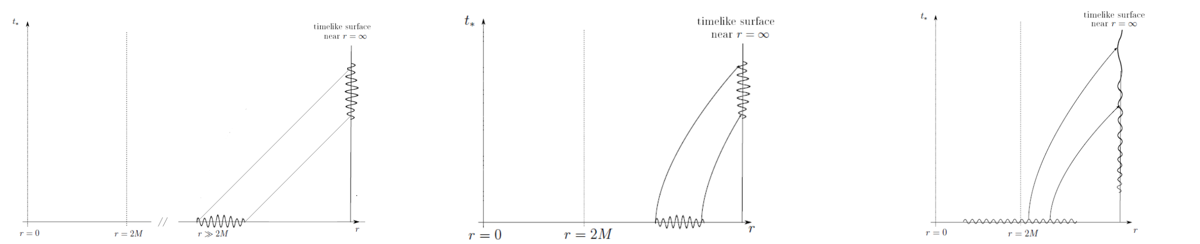
on the later Cauchy surface where lengths are measured by X^-. If $\alpha \gg \epsilon$, the pair of adjacent oscillations was still some distance away from the horizon on the initial Cauchy surfaces, albeit within range to be influenced by the gravitational effects of the black hole. Despite the logarithmic map between the frames of reference, these oscillations will still be of comparable wavelength after the evolution to the later Cauchy surface,

$$\frac{\lambda_{\varsigma_2}}{\lambda_{\varsigma_1}} = \frac{\alpha}{\alpha - \epsilon} \approx 1 \qquad \text{for } \alpha \gg \epsilon. \tag{1.5.10}$$

The stretching they undergo is of almost uniform nature and resembles an ordinary rescaling of coordinate y^-. Graphically, this is due to the fact that outgoing null geodesics in ingoing EF coordinates that begin at such a distance $\alpha \gg \epsilon$ from the horizon do not diverge much, despite being logarithmically curved (Fig.12b). There is very little change in the stretching factor of spacetime over the oscillation period of the wave modes, so that the adiabatic theorem predicts there should not be geometrical particle creation. Indeed, let $f \sim e^{iky^-}$ be a positive frequency mode on the initial Cauchy surface that partially evolves to negative frequency mode $g^* \sim e^{-ik\gamma y^-}$ under rescaling of y^- by constant $\gamma > 0$. Then, by Eq.(1.3.4) we see that the Bogoliubov coefficient B for this wave mode vanishes:

$$\begin{aligned} B = -(g, f^*) &= -i \int dS_\mu g^* \overleftrightarrow{\nabla}^\mu f^* \\ &\sim -i \int dy^- \; e^{-ik\mu y^-} \overleftrightarrow{\partial_y} \, e^{-iky^-} \to 0. \end{aligned} \tag{1.5.11}$$

Consider, however, the few oscillations straddling the horizon for which $\alpha \sim \epsilon$. In this case, the wavelength of oscillations becomes longer and longer as their associated departing distance α from the horizon decreases (Fig.(12c)). The deformation of these oscillations in close proximity of the horizon is completely non-uniform. A positive frequency mode $f \sim e^{iky^-}$ will partially evolve into a negative frequency modes $g \sim e^{iKX^-}$, where in this case the change of coordinates cannot be approximated by rescaling by a constant. Thus, there will be particle creation [25].



(a) Wave modes far from the horizon are not subjected to the gravitational forces exerted by the black hole and are not distorted.

(b) Wave modes close to the horizon are stretched uniformly in case the null geodesics remain approximately parallel.

(c) Wave modes straddling the horizon will undergo a very non-uniform stretching that will lead to particle production.

Figure 12: The distortion of wave modes starting from various degrees of separation from the horizon depicted in ingoing EF coordinates. To facilitate comparison of stretching rates all modes are evolved to a timelike surface near infinity. Only for those modes that initially are very close to the horizon the stretching factor of spacetime changes sufficiently to cause significant particle creation.

In order to properly visualise these effects we need a consistent way of establishing Cauchy surfaces with which to foliate Schwarzschild spacetime. Far away from the black hole, we take it to approximate a spacelike slace in flat spacetime by choosing the hypersurfaces $t = const.$. We will denominate this part of the Cauchy surface by Σ_{out}. As can be seen from the metric (1.4.1), the radial coordinate r becomes timelike for $r < 2M$ inside the horizon. Hence, for this region of spacetime we use surfaces $r = const.$ (but not $r = 0$ or $r = 2M$) to form the inner parts Σ_{in} of the Cauchy surface. Furthermore, we connect these two parts by a smooth, spacelike connector surface Σ_{con} spanning the horizon as in Fig.(13). Not shown in Fig.(13) is a final part of the Cauchy surface where we extend the surface Σ_{in} of constant r back in time before the black hole formed and smoothly bend it to reach $r = 0$. Cauchy surfaces of this form provide a way of slicing the spacetime. Graphically, 'later' surfaces in the foliation can be created by a shift to the upper left corner combined with an enlargement of Σ_{con} (Fig.(13)) [25].

Using the fact that the phase S of the wave modes stays constant along null geodesics, we then evolve the wave modes between different Cauchy surfaces in the foliation. During the early stages of evolution

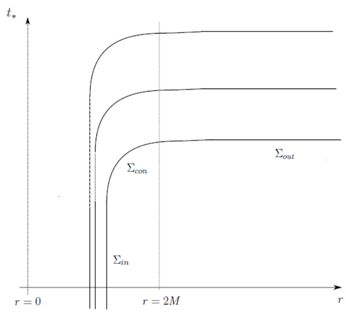


Figure 13: A type of Cauchy surfaces in ingoing EF coordinates, where $t_* = t + 2M \ln[(r-2M/)2M] \to t$ as $r \to \infty$, that is composed of a horizontal part Σ_{out} on the outside, vertical part Σ_{in} on the inside and a bridging connector part Σ_{con} over the horizon. 'Later' surfaces can be constructed by a shift to the upper left corner combined with an enlargement of Σ_{con}.

the wavelength of the mode increases uniformly. Since the stretching factor of spacetime is still small, there is no significant distortion of its general shape. Further evolution along the null hypersurfaces of constant phase S leads to a non-uniform distortion of the wave mode on both sides of the horizon (Fig.(14)), so that particles are created. Since this happens at distances of order M from the horizon, these particles will have wavelength $\lambda \sim M$. The particles on the outside propagate away to infinity as Hawking radiation with a temperature $T \sim \lambda^{-1} \sim 1/M$, which indeed agrees with relation (1.4.33). Following the wave mode even more until it has moved to distances $\gtrsim M$ from the horizon will see the null hypersurfaces return to nigh parallel trajectories, causing particle production to cease.

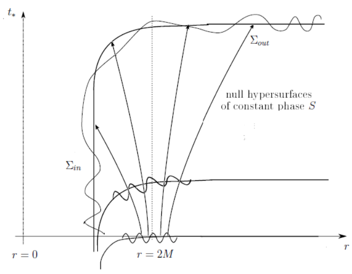


Figure 14: A wave mode on the initial Cauchy surface is evolved to later ones. At first, its distortion is mainly uniform. Only when the stretching factor of spacetime grows larger, the stretching of the wave mode becomes highly non-uniform and particles are created. Evolving the mode even more will stop particle production in finite period of time Δt_*.

Hence, during their evolution the wave modes will attain a wavelength of order M. This implies that wave modes that initially started out with a shorter wavelength will take a longer time before entering the period of non-uniform stretching and will thus begin emitting particles at a later moment in time. Consider two wave modes straddling the horizon with different initial wavelengths beginning from the same early Cauchy surface as sketched in Fig.(15). While the wave mode with the shorter wavelength is

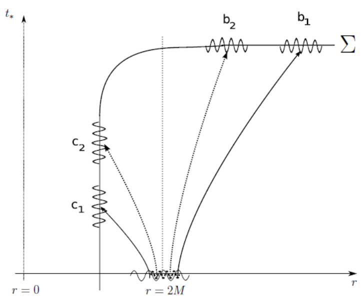


Figure 15: The creation of separate pairs b_1, c_1 and b_2, c_2 of entangled particles caused by the evolution of two Fourier modes with different wavelengths.

still in a phase of uniform stretching, the other is already distorted sufficiently to warrant the production of quanta b_1 inside and c_1 outside of the horizon on the later Cauchy surface. Exactly because the formations that lead to particle production are evenly distributed over both sides of the horizon, one particle of the created pair will fall towards the singularity inside of the black hole whereas the other will propagate to infinity. The state of this pair can then be expressed in terms of creation operators acting on the local 'out' vacua at the respective positions of the particles as

$$|\psi_1\rangle = Ce^{\mu b_1^\dagger c_1^\dagger}|0_{out,b_1}\rangle|0_{out,c_1}\rangle, \tag{1.5.12}$$

where operator $b_1^\dagger$ creates quantum b_1 and operator $c_1^\dagger$ creates quantum c_1. Proceeding the evolution of the shorter wavelength wave mode eventually causes it to enter a non-uniform stretching period, during which the inside quantum b_2 and outside quantum c_2 are produced. Assuming that the wave modes were well-defined localised wave packets on the initial Cauchy surface, the quanta are formed in isolated regions of spacetime. Hence, the state of a pair is given by

$$|\psi_k\rangle = Ce^{\mu b_k^\dagger c_k^\dagger}|0_{out,b_k}\rangle|0_{out,c_k}\rangle. \tag{1.5.13}$$

The Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \dots$ of these pair states should be regarded as separate systems lying in an overarching tensor product space $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2 \times \dots$ comprised of all separate pair state vectors. By an adaptation of Eq.(1.3.18), the overall in vacuum $|0_{in}\rangle$ can then be described by the state in $\mathcal{H}$ that is the direct product of the states $|\psi_k\rangle$,

$$|0_{in}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle \otimes \dots \tag{1.5.14}$$

Becaue the constant μ was found to be of order unity in Eq.(1.3.17), we can expand the exponential in the state of the pairs in powers of μ. Terms with higher powers of μ will have to be of ever smaller relevance, so let us restrict ourselves to the first two terms. Then, the state of a pair as given by Eq.(1.5.13) can be expanded as

$$\begin{aligned} |\psi_k\rangle &= C\left[|0_{out,b_k}\rangle \otimes |0_{out,c_k}\rangle + \mu\left(b_k^\dagger|0_{out,b_k}\rangle \otimes c_k^\dagger|0_{out,c_k}\rangle\right) + \dots\right] \\ &= C\left[|0_{out,b_k}\rangle \otimes |0_{out,c_k}\rangle + \mu\Big(|1_{out,b_k}\rangle \otimes |1_{out,c_k}\rangle\Big) + \dots\right] \end{aligned} \tag{1.5.15}$$

Quintessentially, it now becomes clear that the pair states $|\psi_k\rangle$ are entangled states. Such states are inseparable in the sense that they cannot be written as a mere direct product of states of their subsystems. As a result, it is no longer possible to attribute definite states to the subsystems individually. Instead, the separate subsystems are described only by a reduced density matrix as defined in Eq.(1.4.6). Hence, we cannot restrict attention to only the quanta b_k in the Hawking radiation and still describe them by pure quantum states because the particles b_k are completely entangled with the quanta c_k on the inside of the horizon.

2 The Black Hole Information Paradox

2.1 The Paradox

Since the black hole radiates Hawking quanta at a temperature T_H inversely proportional to its mass M, it will have a negative heat capacity

$$C_{BH} = \left(\frac{dT_H}{dM}\right)^{-1} = -8\pi M^2 < 0. \tag{2.1.1}$$

Suppose then that the temperature of the universe as determined by the CMB is lower than the Hawking temperature of the black hole, creating a temperature gradient with the net effect of the black hole losing mass by particle emission. Lower mass black holes have a higher Hawking temperature, raising the magnitude of the temperature gradient and causing the black hole to radiate more strongly. It is predicted that this runaway process eventually leads to the complete evaporation of the black hole [14]. By Stefan's law, the energy outflow P of the black hole is given by

$$P = \frac{dM}{d\tau} \simeq -\frac{\pi^2}{60} A T_H^4 = \frac{1}{15360\pi M^2}. \tag{2.1.2}$$

This differential equation is separable, so we can write $15360\pi M^2 dM = d\tau$. The black hole mass is now a decreasing function $M(\tau)$ of proper time τ. The lifespan τ_{ev} of the black hole may then be calculated by integrating the right hand side over τ from proper time zero when the black hole forms to τ_{ev} when it has completely evaporated. This will be of the same order as an integration of the left hand side of the equation from initial mass M_0 to zero,

$$\tau_{ev} = \int_0^{\tau_{ev}} d\tau = 15360\pi \int_{M_0}^{0} M^2 dM = 5120\pi M^3. \tag{2.1.3}$$

in Planck units [45].

The Penrose diagram of the gravitational collapse of a star that incorporates the notion that the resultant black hole decreases in size and eventually evaporates completely is outlined in Fig.(16). The

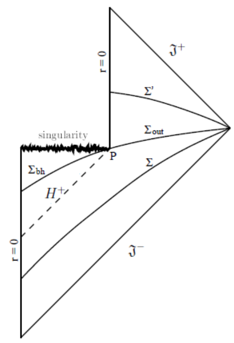


Figure 16: Penrose diagram for a star undergoing gravitational collapse into a Schwarzschild black hole that eventually dissipates through Hawking radiation. Point P partitions the intermediate Cauchy surface Σ_P into two disjoint spatial hypersurfaces Σ_{bh} and Σ_{out} that respectively lie to the interior and the exterior of the horizon. Near P, a theory of quantum gravity is needed to describe the evolution of the black hole.

initial hypersurface Σ is a Cauchy surface. On this slice of spacetime, the scalar field function ϕ satisfies the boundary conditions ($\phi = \phi_1, \dot{\phi} = \dot{\phi}_1$) that determine the state of the field on any other Cauchy surface of the spacetime. This includes Cauchy surfaces like $\Sigma_P = (\Sigma_{bh} \cup \Sigma_{out})$ that extend over the horizon H^+ of the black hole. On such hypersurfaces, the state of ϕ should include the entangled states of the Hawking radiation quanta b_k and the quanta c_k on the inside of the horizon. As the black hole radiates it decreases in size and finally evaporates completely, the entanglement pair particles c_k of the Hawking quanta disappear with it. In some sense information is lost.

Indeed, there is a von Neumann 'entanglement' entropy $S_{vN_{(rad)}}$ associated with the fact that the Hawking radiation is not a definite, pure quantum state but rather a mixed state described only by a reduced density matrix $\rho_{\mathfrak{I}^+}$. It is given by

$$S_{vN}^{(rad)} = -\operatorname{Tr}(\rho_{\mathfrak{I}^+} \ln \rho_{\mathfrak{I}^+}). \tag{2.1.4}$$

For the state (1.5.15) of one Hawking pair, the reduced density matrix (as defined in Eq.(1.4.6)) for the outside quantum b that radiated to infinity is

$$\rho_b = C\left(|0_b\rangle\langle 0_b| + \mu^2|1_b\rangle\langle 1_b| + \dots\right), \tag{2.1.5}$$

resulting in an entanglement entropy $S_{vN}^{(b)} = -\left[C\ln C + C\mu^2\ln\left(C\mu^2\right)\right]$ of order unity. The energy of an average emitted Hawking quantum is $\sim T_H \sim 1/M$, so that during its evaporation the black hole will radiate a number

$$\mathcal{N}_{quanta} \sim \frac{M}{(M^{-1})} = M^2 \tag{2.1.6}$$

of Hawking quanta. Since the entropy for each pair (b_k, c_k) is of order unity, we see that the total entanglement entropy of the Hawking radiation is $S_{vN}^{(rad)} \sim M^2 \sim S_{BH}$, of the same order as the semi-classical Bekenstein-Hawking entropy of the black hole [25].

Because of the information loss that encompasses the total evaporation of the black hole, the evolution from Cauchy surfaces like Σ and Σ_P to later spacelike hypersurfaces as Σ' is *non-unitary.* Although they slice spacetime, such hypersurfaces Σ' are not Cauchy surfaces. Their past domains of dependence $D^-(\Sigma')$ do not include the black hole region surface Σ_{bh}. Hence, it seems that Hawking radiation puts QFT in curved spacetime in contradiction with one of the fundamental principles of quantum mechanics [31,45].

2.2 Black Hole Complementarity

Hitherto, we have studied the behaviour of a quantised field ϕ on an asymptotically stationary spacetime that is curved by the gravitational effects of a forming black hole. However, we have not treated gravity itself as a quantum phenomenon. Instead, we have maintained a classical description, which is predicted to break down at least at the last stages of evaporation. This treatise of the collapsing star process is called the *semi-classical approximation.* One might then ask oneself whether including quantum gravity could effect a change in the state (1.5.15) of one set of pair produced particles so that the interior particle c_k and the exterior particle b_k are no longer entangled and the evaporation of the black hole again becomes a unitary process. Such a non-entangled state would need to be of the form

$$|\psi_k\rangle = \Big(C_0|0_{b_k}\rangle + C_1|1_{b_1}\rangle + \dots\Big) \otimes \Big(D_0|0_{c_k}\rangle + D_1|1_{c_1}\rangle + \dots\Big). \tag{2.2.1}$$

Notice that the state (2.2.1) is not just a small perturbation of the entangled state (1.5.15). In order to cause such a order unity change, quantum gravity must completely change the dynamics of the wave modes [25].

In a 1993 paper, Susskind, Thorlacius and Uglum proposed that the correct description of the evaporation of a black hole indeed contains such changes for an outside observer [37]. In that case, the Hawking radiation detected by a distant observer far removed from the gravitational effects of the black hole is in a pure state. Information can no longer be lost in the black hole and its evaporation process may be described completely by a unitary transformation. In this light, consider again the Penrose diagram of the formation and evaporation of a black hole sketched in Fig.(16). The state of the field $|\phi(\Sigma')\rangle$ on late hypersurface Σ' must be a pure state that is related to the original incoming state $|\phi(\Sigma)\rangle$ on Cauchy surface Σ by some unitary operator S, the S-matrix. The domain of dependence of Σ' contains only Σ_{out} and not Σ_{bh}, which implies $|\phi(\Sigma_{out})\rangle$ must also be a pure state that has evolved unitarily from initial state $|\phi(\Sigma)\rangle$. As such, $|\phi(\Sigma_P)\rangle$ becomes a product state of states on Σ_{bh} and Σ_{out},

$$|\phi(\Sigma_P)\rangle = |\phi(\Sigma_{bh})\rangle \otimes |\phi(\Sigma_{out})\rangle, \tag{2.2.2}$$

where $|\phi(\Sigma_{out})\rangle \in \mathcal{H}_{\Sigma_{out}}$ and $|\phi(\Sigma_{bh})\rangle \in \mathcal{H}_{\Sigma_{bh}}$ is independent of the initial state. This implies that states of infalling matter are scrambled and randomised at the moment of crossing, which clearly violates the equivalence principle. Even assuming changes of order unity due to quantum gravity effects, there appears to be a strong contradiction between some widely-held principles from general relativity on the one hand and quantum mechanics on the other.

However, it was suggested that this contradiction is only apparent. The existence of a state $|\phi(\Sigma_P)\rangle$ with which both the interior and exterior of the black hole can be described simultaneously seems unphysical. Such a state contains information about correlations without operational meaning, since communication between the inside and the outside of the black hole is impossible. Only some kind of 'superobserver' outside the universe who can concurrently perform experiments in both regions can describe a state $|\phi(\Sigma_P)\rangle$ in the tensor product Hilbert space $\mathcal{H}_{\Sigma_{out}} \times \mathcal{H}_{\Sigma_{bh}}$. The inconsistencies in the derivation can be removed by giving up the global validity of the governing principles in a retreat to an operationalistic viewpoint of nature . A distant observer outside the horizon will observe Hawking radiation in a pure state and will think that the equivalence principle is violated whenever matter enters the black hole region. An infalling observer will detect no sudden scrambling upon crossing the horizon, but will predict that the Hawking radiation is perceived as a mixed state by the outside observer. Note that there is no conflict with the *no-cloning theorem* of quantum information theory, because the Hilbert spaces of the interior and exterior states do not overlap.

These proposals were enshrined in three postulates:

Postulate 1: **Purity of Hawking radiation**

The process of formation and evaporation of a black hole, as viewed by a distant observer, can be described entirely within the context of standard quantum theory. In particular, there exists a unitary S-matrix which describes the evolution from infalling matter to outgoing Hawking-like radiation.

Postulate 2: **Semi-classical behaviour outside the stretched horizon**

Outside the stretched horizon of a massive black hole, physics can be described to good approximation by a set of semi-classical field equations.

Postulate 3: **Absence of infalling drama**

The global event horizon of a very massive black hole does not have large curvature or energy density. Therefore, an infalling observer will see nothing extraordinary when crossing the horizon, in accordance with the equivalence principle of general relativity.

Together, these three postulates form a conjecture by the name of *black hole complementarity* (BHC)[37]. The observations of an outside observer versus those made by an infalling observer are complementary in the sense of Bohr: neither can be said to be the true picture of reality. Instead, the two pictures are complementary to each other.

Let us have a closer look at the formation and subsequent evaporation of the black hole from the viewpoint of the distant observer. Far away from the horizon, an observer will detect Hawking radiation with a temperature $T_H = 1/(8\pi M)$. However, because of the black hole's gravity it will have experienced an infinite redshift since its emission, so that the distant observer will think the temperature diverges as one approximates the horizon. Now, the entropy density $s(T)$ associated with a field in a general field theory can always be parametrised by an equation of the form

$$s(T) = \gamma(T)T^3, \tag{2.2.3}$$

where $\gamma(T)$ represents the 'effective' degrees of freedom of the field at T and is in most canonical cases accepted to be a monotomically increasing function. Thus, $s(T)$ will always increase with temperature and the entropy density will diverge at the horizon, in clear contradiction with the Bekenstein-Hawking entropy $S_{BH} = A/4$ of a black hole. The free quantum field theory applied to the exterior of the black hole will attribute too much entropy in modes in close proximity to the horizon. In order to recover the Bekenstein-Hawking result, we will need a new theory that has γ decreasing with T closer to the horizon. Therefore, we introduce some cutoff distance ϵ from the horizon beyond which the ordinary free quantum field theory is a valid approximation[40].

The region near the horizon can be examined by replacing the radial coordinate r in the Schwarzschild metric (1.4.1) by a *Rindler coordinate* ρ, defined by

$$\begin{aligned} \rho &\equiv \int_{2M}^{r} \sqrt{g_{rr}(r')}dr' \\ &= \int_{2M}^{r} \left(1 - \frac{2M}{r'}\right)^{-\frac{1}{2}} dr' \\ &= \sqrt{r(r-2M)} + 2M \sinh^{-1}\left(\sqrt{\frac{r}{2M} - 1}\right), \end{aligned} \tag{2.2.4}$$

that measures the proper distance from the horizon. In terms of this new coordinate, the metric takes the form

$$ds^2 = -\left(1 - \frac{2M}{r(\rho)}\right)dt^2 + d\rho^2 + r^2(\rho)d\Omega^2. \tag{2.2.5}$$

Near the horizon, Eq.(2.2.4) is approximated by $\rho \approx 2\sqrt{2M(r-2M)}$, so the metric reduces to

$$ds^2 \cong -\left(\frac{\rho}{4M}\right)^2 dt^2 + d\rho^2 + r^2(\rho)d\Omega^2. \tag{2.2.6}$$

Since the spacetime is stationary, the gravitational redshift experienced by a mode in the vicinity of the horizon before it arrives at infinity is simply given by the ratio of the proper time τ to Minkowski time t valid at infinity. Hence, the frequency ω_∞ detected at infinity is related to the frequency $\omega_{\sim 2M}$ near the horizon as

$$\frac{\omega_\infty}{\omega_{\sim 2M}} = \frac{d\tau}{dt} = \frac{\rho}{4M}. \tag{2.2.7}$$

Since all frequencies in the distribution undergo the same redshift, the overall distribution will remain unchanged. Hence, the proper temperature $T_{\sim 2M}(\rho)$ near the horizon as predicted by an observer at infinity is given by

$$T_{\sim 2M}(\rho) = \frac{\omega_{\sim 2M}}{\omega_\infty} T_H = \frac{1}{2\pi\rho}. \tag{2.2.8}$$

Now, the entropy density of a free scalar field ϕ in $(3+1)$-dimensional quantum field theory is given by

$$s(T) = \frac{2}{\pi^2}\zeta(4)\, T^3 = \frac{2\pi^2}{45} T^3. \tag{2.2.9}$$

Then, the entropy stored in an area segment B^2 of a thin layer $\delta\rho$ at a distance ρ from the horizon is

$$\begin{aligned} \delta S(\rho) &= \frac{2\pi^2}{45} T^3_{\sim 2M}(\rho)\, \delta\rho B^2 \\ &= \frac{2\pi^2}{45}\frac{1}{(2\pi\rho)^3}\, \delta\rho\, B^2. \end{aligned} \tag{2.2.10}$$

As we have argued the quantum field theory that is key to the derivation of Eq.(2.2.9) loses its validity at distances closer than ϵ from the horizon. The full entropy may be found by integrating $\delta S(\rho)$ from ϵ to infinity:

$$\begin{aligned} S &= \frac{B^2}{(2\pi)^3}\frac{2\pi^2}{45}\int_\epsilon^\infty \frac{d\rho}{\rho^3} \\ &= \frac{B^2}{360\pi\, \epsilon^2}. \end{aligned} \tag{2.2.11}$$

If we demand that the entropy at distances greater than ϵ does not exceed the value predicted by Bekenstein and Hawking, it follows that we must have $\epsilon^2 \lesssim 1$ in Planck units. At distances less than a Planck length ℓ_p away from the horizon the numbers of degrees of freedom as measured by the function $\gamma(T)$ must be negligible. This is a reason to argue that for an external observer the mathematical horizon at $r = 2M$ may be substituted by an effective membrane called the *stretched horizon*, which is a surface $\rho = \rho_0 \sim \ell_p$ one Planck length further outwards. Instead of being a null hypersurface like the mathematical horizon, the stretched horizon is a timelike membrane. This means that separate regions of the stretched horizon can be casually connected, enabling dynamical properties like electrical conductivity and viscosity to develop on its surface [40].

Derivation of some of the dynamical characteristics of the stretched horizon will be easier in a different coordinate system, in which the metric is identified with that of Minkowski spacetime. To this end, we return to the approximation (2.2.6) of the Schwarzschild metric near the horizon. If we concentrate on a small angular region, which we may center at $\theta = 0$ without loss of generality, the angular coordinates (θ, φ) can be parametrised by Cartesian coordinates

$$\begin{aligned} x &= 2M\theta\cos\varphi \\ y &= 2M\theta\sin\varphi. \end{aligned} \tag{2.2.12}$$

Furthermore, we can also replace t with the dimensionless *Rindler time* $\omega = t/4M$, so that the metric becomes

$$ds^2 = -\rho^2 d\omega^2 + d\rho^2 + dx^2 + dy^2. \tag{2.2.13}$$

Rindler coordinates ρ and ω are radial and hyperbolic angle variables for ordinary Minkowski spacetime that can be recovered by defining

$$\begin{aligned} T &= \rho \sinh \omega \\ z &= \rho \cosh \omega, \end{aligned} \tag{2.2.14}$$

so that the metric is reparametrised as

$$ds^2 = -dT^2 + dx^2 + dy^2 + dz^2. \tag{2.2.15}$$

The approximation of the region of Schwarzschild spacetime near the horizon, also called *Rindler spacetime*, with ordinary flat spacetime is named the *Rindler approximation*. An observer at rest in Rindler coordinates experiences a uniform proper acceleration. This situation corresponds to an observer hovering at a constant distance from the event horizon of a black hole in Schwarzschild coordinates by the equivalence of acceleration and gravitation [40].

In Rindler spacetime, the Lagrangian for an electromagnetic field is

$$\mathcal{L} = -\frac{\sqrt{-g}}{16\pi} g^{\mu\nu} g^{\sigma\tau} F_{\mu\sigma} F_{\nu\tau} + j^\mu A_\mu. \tag{2.2.16}$$

As usual, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength, which is based upon the electromagnetic potential $A_\mu = (-\xi, \vec{A})$. Also, we have included the conserved current density j, for which $\partial_\mu j^\mu = 0$. By the principle of least action, the equations of motion are

$$\partial_\mu(\rho F^{\mu\nu}) = -4\pi j^\nu. \tag{2.2.17}$$

Furthermore, since the field strength is an antisymmetric rank-2 tensor, it identically satisfies the Bianchi identity

$$\partial_{[\lambda} F_{\mu\nu]} = \partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. \tag{2.2.18}$$

Substituting in the Rindler metric (2.2.13) into the electromagnetism Lagrangian, we can rewrite

$$\begin{aligned} \mathcal{L} &= -\frac{\rho}{16\pi}\left(-\frac{2\,\delta^{ij}}{\rho^2} F_{0i}F_{0j} + \delta^{ij}\delta^{kl} F_{ik}F_{jl}\right) + j^\mu A_\mu \\ &= \frac{1}{8\pi}\left[\frac{\delta^{ij}}{\rho}(\partial_0 A_i - \partial_i A_0)(\partial_0 A_j - \partial_j A_0) - \rho(\epsilon^{kij}\partial_i A_j)(\epsilon_{kab}\partial^a A^b)\right] + j^\mu A_\mu \\ &= \frac{1}{8\pi}\left[\frac{(\dot{\vec{A}} + \vec{\nabla}\xi)^2}{\rho} - \rho(\vec{\nabla} \times \vec{A})^2\right] + j \cdot A, \end{aligned} \tag{2.2.19}$$

where we have kept to the custom of designating (Rindler) time derivation by dotting, i.e. $\dot{\vec{A}} = \partial_0 \vec{A}$. This form of the Lagrangian suggests we should define an electric field $\vec{E}$ and magnetic field $\vec{B}$ given by

$$\vec{E} = -\vec{\nabla}\xi - \dot{\vec{A}} \qquad \vec{B} = \vec{\nabla} \times \vec{A}, \tag{2.2.20}$$

so that

$$\mathcal{L} = \frac{1}{8\pi}\left(\frac{|\vec{E}|^2}{\rho} - \rho|\vec{B}|^2\right) + j \cdot A. \tag{2.2.21}$$

In close analogy to ordinary electromagnetism in flat spacetime, in terms of the electric and magnetic fields Eqs.(2.2.17) and (2.2.18) take on a Maxwellian form:

$$\begin{aligned} \frac{1}{\rho}\dot{\vec{E}} - \vec{\nabla} \times (\rho\vec{B}) = -4\pi\vec{j}, \qquad & \vec{\nabla} \cdot \left(\frac{1}{\rho}\vec{E}\right) = 4\pi j^0, \\ \dot{\vec{B}} + \vec{\nabla} \times \vec{E} = 0, \qquad & \vec{\nabla} \cdot \vec{B} = 0. \end{aligned} \tag{2.2.22}$$

Together, this set of equations governs the effects of charges and currents on the magnetic and electric fields, as well as their interactions [40].

We start by considering the electrostatic properties of the stretched horizon. This implies we study the effects of fields associated with charges that are stationary in Rindler coordinates. In ordinary

Schwarzschild spacetime, this corresponds to charges that are at a constant distance from the stretched horizon. We take this distance to be macroscopically large, so that all length scales associated with the charges are much larger than $\rho_0 \sim \ell_p$.

By the second of the Maxwellian equations (2.2.22), the electric charge confined in a region of space is proportional to the strength of the electric field through that region. Hence, the surface charge density on the stretched horizon is proportional to the component of $\vec{E}$ perpendicular to the horizon,

$$\begin{aligned} j^\omega &= \frac{1}{4\pi\rho} E_\rho \bigg|_{\rho=\rho_0} \\ &= -\frac{1}{4\pi\rho} \partial_\rho \xi \bigg|_{\rho=\rho_0} . \end{aligned} \tag{2.2.23}$$

Choosing the Coulomb gauge sets $\vec{\nabla} \cdot \vec{A} = 0$, so that the second Maxwellian equation changes to

$$\vec{\nabla} \cdot \left(\frac{1}{\rho} \vec{E} \right) = -\vec{\nabla} \cdot \left(\frac{1}{\rho} \vec{\nabla} \xi \right) = \left(-\frac{1}{\rho} \partial_\rho \xi \bigg|_{\rho=\rho_0} \right) = 0 \tag{2.2.24}$$

for $\rho \neq \rho_0$ near but not on the horizon. Therefore,

$$0 = -\vec{\nabla} \cdot \left(\frac{1}{\rho} \vec{\nabla} \xi \right) = \frac{1}{\rho^2} \partial_\rho \xi - \frac{1}{\rho} \vec{\nabla}^2 \xi \qquad \leftrightarrow \qquad \partial_\rho^2 \xi - \frac{1}{\rho} \partial_\rho \xi = -\nabla_\perp^2 \xi, \tag{2.2.25}$$

where $\nabla_\perp = (\partial_x, \partial_y)$. Assuming that ξ is analytic, we may expand in powers of ρ. Since the RHS is two powers of ρ smaller than the LHS, it can be discarded if we limit ourselves to second order. We then obtain a second order differential equation without a first order term, so that we may assume

$$\xi(\rho, x, y) = F(x, y) + \rho^2 G(x, y) + \mathcal{O}(\rho^3). \tag{2.2.26}$$

Substituting this approximate solution into Eq.(2.2.25), we obtain

$$\nabla_\perp^2 F + \rho^2 \nabla_\perp^2 \approx 0. \tag{2.2.27}$$

Since we have made the ansatz that ρ_0 is much smaller then all other length scales, evaluating this expression at the horizon causes it to simplify to

$$\nabla_\perp^2 F(x, y) = 0. \tag{2.2.28}$$

$F(x, y)$ and, by extension, ξ are harmonic functions on the compact black hole horizon manifold. A famous result from DeRham-cohomology then stipulates that $\phi = const.$, so that the stretched horizon is a 2-dimensional equipotential surface in space. This causes it to act as a charged electrical conductor in equilibrium [40].

Exactly because the stretched horizon acts as an electrical conductor to a distant observer, it should also have some associated non-vanishing current density $\vec{j}$. Evidently, the ρ component of the current density will vanish because the stretched horizon is the 2-dimensional (x, y)-plane. The other components can be found by using the Maxwellian equations on the Rindler time derivative of Eq.(2.2.23),

$$4\pi \partial_\omega j^\omega = \frac{1}{\rho_0} \dot{E}_\rho = \left(\vec{\nabla} \times \rho \vec{B} \right)_\rho = \rho (\nabla_\perp \cdot B_\perp). \tag{2.2.29}$$

Clearly, this has the correct form to be a continuity equation with the definitions

$$4\pi j_x = -\rho B_y, \qquad\qquad 4\pi j_y = \rho B_x, \tag{2.2.30}$$

that determine how the electric current behaves on the stretched horizon.

Now, suppose there is an electromagnetic wave propagating towards the stretched horizon along the ρ axis, so that it intersects perpendicularly. Outside the horizon, $j^\mu = 0$, so that the first and third of the Maxwellian equations (2.2.22) give

$$\begin{aligned} \dot{B}_x &= \partial_\rho E_y, & \dot{B}_y &= -\partial_\rho E_x, \\ \frac{1}{\rho} \dot{E}_x &= -\partial_\rho (\rho B_y), & \frac{1}{\rho} \dot{E}_y &= \partial_\rho (\rho B_x). \end{aligned} \tag{2.2.31}$$

These equations can be cast into more recognisable form with a redefinition $\vec{\beta} = \rho\vec{B}$ of the magnetic field and the introduction of a new tortoise coordinate $\varpi = \ln\rho$, for which $\partial_\varpi = \rho\partial_\rho$ by the chain rule. We find

$$\begin{aligned} \dot{\beta}_x &= \partial_\varpi E_y, & \dot{\beta}_y &= -\partial_\varpi E_x, \\ \dot{E}_x &= -\partial_\varpi \beta_y, & \dot{E}_y &= \partial_\varpi \beta_x. \end{aligned} \tag{2.2.32}$$

Then, by Rindler time derivation of the LHS of these equations and back substitution, we find

$$\begin{aligned} \ddot{\beta}_x &= \partial_\varpi \dot{E}_y = \partial_\varpi^2 \beta_x, & \ddot{\beta}_y &= -\partial_\varpi \dot{E}_x = \partial_\varpi^2 \beta_y, \\ \ddot{E}_x &= -\partial_\varpi \dot{\beta}_y = \partial_\varpi^2 E_x, & \ddot{E}_y &= \partial_\varpi \dot{\beta}_x = \partial_\varpi^2 E_y, \end{aligned}$$

so that to all relevant intents and purposes, $\partial_\omega = \pm\partial_\varpi$. Hence, the mathematical equations admit waves propagating in either direction along the u-axis with Rindler time ω. Physically, however, only waves propagating towards the black hole from the outside are allowed, for which we must pick the + sign. For these ingoing waves, we then have

$$\beta_x = E_y \qquad \beta_y = -E_x. \tag{2.2.33}$$

When these waves cross the stretched horizon, they will generate a current whose components are given by ohmic relations

$$j_x = \frac{1}{4\pi}E_x, \qquad j_y = \frac{1}{4\pi}E_y, \tag{2.2.34}$$

according to Eq.(2.2.30). The stretched horizon of a black hole can be interpreted as an ohmic conductor with a conductivity of $1/4\pi$. For a 2-dimensional resistor like the stretched horizon, the surface resistance is a topological quantity independent of scale. With a square as the fundamental shape, a conductivity of 4π corresponds to a surface resistance of $377\Omega/square$. If we were to connect a battery to the horizon in a circuit as in Fig.(17), an external observer in possession of an ammeter and knowledge of Ohm's law would be able to determine this resistance just as if the stretched horizon were a physical conducting membrane [40].

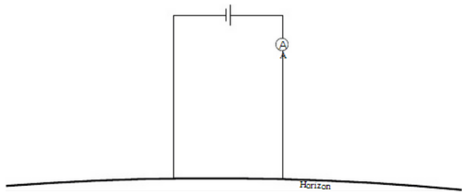


Figure 17: If we create an electrical circuit by connecting a battery to the stretched horizon, an external observer will be able to confirm, for example by repeatedly measuring the amperage for different voltages, that it behaves as a ohmic resistor with a resistance of $377\Omega/square$.

As a last example of the properties of the stretched horizon, consider a particle with charge e that is falling freely into the black hole. Without loss of generality, we can choose the charge to be positioned at z_0 along the z-axis in the Minkowksi coordinates defined by Eq.(2.2.14). As stationary observers in Rindler spacetime correspond to uniformly accelerating observers in Minkowski spacetime, the coordinates of the two spacetimes are simply related by a Lorentz boost along the z-axis. Then, because the electric field is boost invariant along the boost direction, $E_\rho = E_z$. Interpreting the electric field as an ordinary Coulomb field due to charge e, we can then write

$$\begin{aligned} E_\rho &= E_z \\ &= \frac{e(z - z_0)}{[(z - z_0)^2 + x_\perp^2]^{3/2}} \\ &= \frac{e(\rho\cosh\omega - z_0)}{[(\rho\cosh\omega - z_0)^2 + x_\perp^2]^{3/2}}. \end{aligned} \tag{2.2.35}$$

Using the surface density equation (2.2.23), we obtain

$$j^\omega = \frac{e}{4\pi\rho_0}\frac{\rho_0\cosh\omega - z_0}{[(\rho_0\cosh\omega - z_0)^2 + x_\perp^2]^{3/2}}. \tag{2.2.36}$$

Now, we have $\cosh x = (e^x + e^{-x})/2$, so that for large Rindler time ω the hyperbolic cosine is well approximated by an exponential and we have

$$j^\omega = \frac{e}{8\pi\rho_0} \frac{\rho_0 e^\omega}{\left[\frac{\rho_0^2}{4} e^{2\omega} + x_\perp^2\right]^{3/2}}. \tag{2.2.37}$$

To better appreciate the meaning of this equation, it proves fruitful to apply a rescaling $x_\perp = e^\omega y_\perp/2$, after which it becomes

$$j^\omega - \frac{e}{4\pi} \frac{e^{-2\omega}}{(\rho^2 + y_\perp^2)^{3/2}}. \tag{2.2.38}$$

Evidently, the charge spreads out exponentially with Rindler time ω on its course towards the stretched horizon. This is in complete contrast with what one would expect intuitively, as spacetime near the horizon is a low curvature, approximately Minkowskian region. Based on solely this principle one would assume a point charge to approach the horizon asymptotically, with the transverse charge density remaining in a delta function-like form [40].

Because the global horizon of a black hole in 4-dimensional Schwarzschild spacetime is a null hypersurface, it is constricted in the degrees of freedom it can display by the no-hair theorems. The timelike surface of the stretched horizon faces no such inhibitions and is capable of carrying the plethora of microstates that lie at the basis of the black hole entropy. To determine the nature of these microstates, we have a closer look at the definition of the stretched horizon.

For a canonical Schwarzschild black hole, let us consider radial incoming null geodesics. These geodesics cross the event horizon of a certain area A. Approximately one Planck length ℓ_p backwards along the null rays, they intersect with a surface with an area of one Planck unit larger, which we can define to be the stretched horizon. This construct maps all the points of the event horizon onto points of the stretched horizon that lie one Planck length more radially outwards. Now, the global horizon consists of light rays which can be interpreted as a fluid moving with time on the 2-dimensional spatial surface of the horizon [44]. By the null geodesic map between the event horizon and the stretched horizon, it follows that there is also a fluid flow on the last surface. The stretched horizon can then be thought of as a dynamical system of some continuous, viscous fluid. However, the fluid description of a flow is often an emergent property of statistical physics. Therefore, it was speculated that the continuous classical fluid should be replaced by a fluid of discrete 'horizon atoms' that collectively determine the microphysics of the stretched horizon [37]. As demonstrated, the thermodynamic properties of the stretched horizon are invariant quantities independent of the black hole's size and mass. It follows that the surface density of the proposed horizon atoms should remain constant throughout the lifetime of the black hole, from formation to evaporation. Hence, when incoming matter or outgoing Hawking quanta increase respectively decrease the area of the stretched horizon, there must be atoms spontaneously appearing and disappearing [37].

To an external observer the dynamics of a black hole can be fully described by a hot conducting membrane just outside the global event horizon. In addition to possessing thermodynamic properties like temperature, entropy, energy and viscosity, this stretched horizon also behaves as a conventional ohmic conductor in a circuit with a definite measure of electrical resistivity. Next to undergoing time dilational effects, any object thrown into the black hole will be heated up divergently and have its charges smeared out over the horizon. Furthermore, the bits containing the information associated with the object will be scrambled and randomly returned to infinity in the Hawking radiation. However, the reality of the existence of the stretched horizon is not invariant to the choice of observer. To an infalling observer, this description of the black hole seems completely ungrounded as any attempt to detect the stretched horizon and determine its properties is doomed to fail. These observations seem to be in contradiction, but are really complementary to one another if we choose the Hilbert spaces of the observers to only include those events that are physically able to influence their experiments [37,40].

2.3 The Holographic Principle

As we have seen, quantum field theory attributes too many degrees of freedom to its fields to provide an appropriate description of the complete Schwarzschild spacetime outside of the black hole. In particular, the problem arises from field modes very close to the horizon, which cause the the entropy density to diverge as one approximates the event horizon. Hence, it has been argued that a small distance cutoff may be required to understand situations involving gravitation. From this point of view, quantum field theory should be regulated to a three dimensional lattice field theory with some very small spacing a. Let us consider a 3-dimensional lattice of discrete spin-like degrees of freedom on a large spherical region Γ of space. Each lattice site is equipped with a spin carrying one bit of information. If the volume of region Γ is V, the number of spins is $n = V/a^3$ and Γ can support

$$\mathcal{N}_{states}(V) = 2^n \tag{2.3.1}$$

orthogonal quantum states in its interior. The maximum entropy admitted by the system is the natural logarithm of the total number of states and thus satisfies

$$S_{max} = n \ln 2 = \frac{V}{a^3} \ln 2. \tag{2.3.2}$$

Hence, we see that the maximum entropy is proportional to the volume of the system. This principle holds more generally. Whenever the energy density of a system is bounded from above so that there is a notion of maximum energy density, it is proportional to the volume of the space.

However, this result does not take into account the effects of gravity on the calculation. Gravity gives rise to black holes, and black holes subsequently give rise to more stringent entropy bounds [22]. Consider again a spherical region of space Γ with radius R and boundary $\partial\Gamma$ of area $A = 4\pi R^2$. Contained inside of Γ there can be a thermodynamic system with an entropy S_Γ corresponding to a mass M. Because the horizon area of a black hole increases with the mass, M cannot exceed the mass of a black hole of area A. Now imagine the system is surrounded by a collapsing spherically symmetric light-like matter shell with precisely the right amount of additional energy $E_g = E_{crit} - E_M$ to create a black hole together with the original mass M of the system, as in Fig.(18). The resultant black hole will have an event horizon with

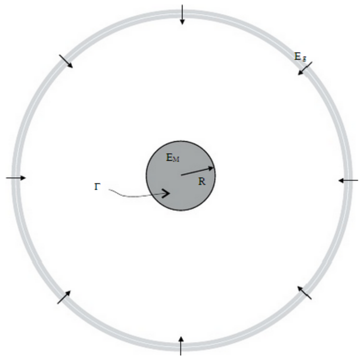


Figure 18: Cross-section of the collapse of a spherically symmetric shell surrounding some matter with mass E_M in a sphere with surface area A. If the mass of the shell is $E_g = E_{crit} - E_M$, a black hole with area A and entropy $S_{BH} = A/4$ will form.

an area A and a corresponding Bekenstein-Hawking entropy $S_{BH} = A/4$. By the generalised second law, the entropy of the universe cannot decrease, so that the original entropy S_Γ inside Γ had to be less or equal then S_{BH}. Hence, we find that the maximum entropy that can be contained in a region of space is the Bekenstein-Hawking entropy, which is proportional to the area of the boundary of the space instead of its volume [3,38,40].

The fact that the amount of information in a 3-dimensional space is limited by a 2-dimensional bound has had some far-reaching consequences that were pioneered by 't Hooft. In a 1994 paper, he inferred that it should mean that all events occurring in a volume V may be described by degrees of freedom on the boundary of the volume [19]. In the limit of a very large volume, the bounding surface approximates a flat plane at infinity. The universe at any point in time may be interpreted as a 3-dimensional projection of information stored on a 2-dimensional 'projector screen', as if it were a hologram. In order to give rise to a 4-dimensional spacetime filled with events the hologram should be cinematic, with the bounded degrees of freedom changing over time. This idea has since been called the *holographic principle* [19,38].

Let us see how to construct such a screen in a spacetime that is asymptotically flat. At asymptotic distances close to infinity, we can use a coordinate system $(X^+, X^-, x_\perp)$, where X^+, X^- are the Minkowski lightcone coordinates defined in Eq.(1.5.1) and $x_\perp = (x, y)$ are the coordinates specifying points on a the 2-dimensional screen. As per custom, we take X^+ to function as the lightcone time variable. Consider then the set of null rays residing on surfaces $X^+ = const.$ in the limit where $X^- \to \infty$ as depicted in Fig.(19). In Minkowski space, these light rays form a congruence defining a three dimensional lightlike

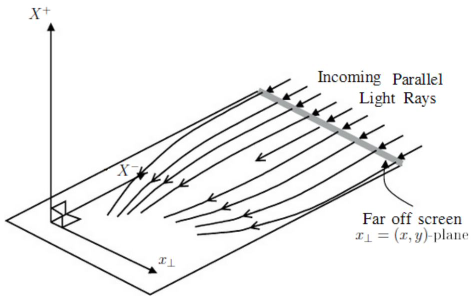


Figure 19: Null rays propagating from $X^- \to \infty$ on a lightlike surface of $X^+ = const.$ eventually find a caustic point.

surface. More generally, they define a lightlike surface named the *light sheet*. Under some justifiable and generally accepted assumptions regarding the nature of the matter filling spacetime, the light rays will converge to some caustic. Although the light sheets typically contain these singular caustic lines, they can be uniquely specified. The collection of light sheets obtained by varying the planes $X^+ = const.$ contain all events in spacetime, with the exception of those points inside of a black hole. Any spacetime point p can be assigned a ligthcone coordinate $X^+ = X_0^+$ corresponding to the light sheet it is on and a planar coordinate $x_\perp = x_{\perp_0}$ determined by the asymptotically transverse coordinate of the light ray it is on. It has proven difficult to allocate a X^- value, but there are still some conclusions that can be drawn.

Suppose a black hole passes through the light street with lightcone time variable X_0^+ as sketched in Fig.(20). The horizon of the black hole is a 2-dimensional null surface in the 3-dimensional light sheet. Points on the horizon have a unique set of coordinates $(X^+, x_\perp)$ and can thus be mapped onto the screen in a one-to-one correspondence. Because the entropy of the black hole is given by $S_{BH} = A/4$,

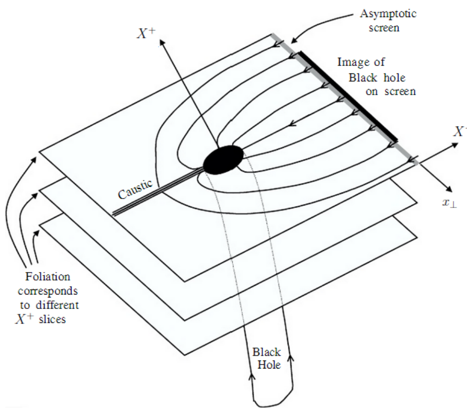


Figure 20: A black hole passing through a X^+ slices of spacetime. The stretched horizon has an one-to-one correspondence with some image on the asymptotically defined 2-dimensional screen.

the horizon will have an entropy density of $s_{BH} = 1/4$ in Planck units. The holographic mapping then also defines an entropy density $s_\perp$ on the image of the black hole on the screen. Consider then the congruence of light rays with cross sectional area α that connect the stretched horizon with the screen at $X^- = \infty$. On these light rays, we can measure distances by some affine parameter λ. In the asymptotic region near the screen the spacetime approximates Minkowski spacetime, so that null rays are parallel and thus

$$\frac{d\alpha}{d\lambda} \to 0 \qquad \text{as } X^- \to \infty. \tag{2.3.3}$$

Assuming the *Weak Energy Condition (WEC)* holds, the Raychaudhuri equation states that any expansion of the area α of the cross section of the congruence cannot but decrease,

$$\frac{d^2\alpha}{d\lambda^2} \leq 0. \tag{2.3.4}$$

Hence, if we trace the parallel light rays back from the screen to the stretched horizon they must eventually converge to some caustic whilst the the cross sectional area of their congruence diminishes. By the holographic principle, a patch of the stretched horizon of the black hole and its image on the asymptotic screen will carry the same entropy (Fig.(21)). However, as we see a bigger patch is needed on the projector screen to register the same amount of information, so that the *holographic bound*

$$s_{\perp} \leq \frac{1}{4} \tag{2.3.5}$$

follows automatically [38,40].

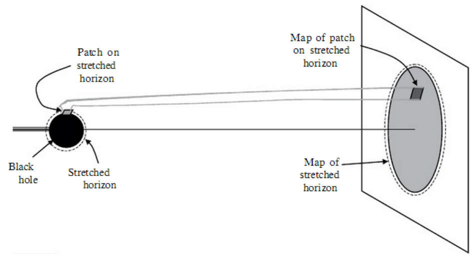


Figure 21: The image of a patch of the stretched horizon on the screen has a bigger surface area.

Notice that this argument hinges crucially on the asymptotic flatness of spacetime to specify the projector screen in such a way that condition (2.3.3) holds. If we were to consider screens for which the light rays are diverging as we move outwards from the stretched horizon, i.e. with $d\alpha/d\lambda \to> 0$, the area of a patch in the bulk would still be smaller than that of its image on the screen and the holographic bound (2.3.5) would still be valid. However, any attempt to formulate a similar argument for screens with asymptotically converging congruences for which $d\alpha/d\lambda \to< 0$ will surely fail. This puts stringent limits on the extent to which the holographic bound can be generalised to more arbitrary geometries [40].

In order to study the maximum entropy bound of more general spacetimes, let us consider some region Γ confined by the spacelike boundary $\partial\Gamma$. It has been shown that at any point on such boundaries there are four perpendicular light rays called *branches* [4], which are two by two past and future directed. For both temporal directions there is one branch composed of ingoing rays, whereas the other consists of outgoing null rays. Collectively, the four branches of all points of $\partial\Gamma$ form as many congruences that define the light sheets corresponding to the boundary. On any of the light sheets the congruences will undergo a positive or negative expansion as we move outwards from the boundary. In ordinary Minkowski spacetime, for a convex boundary $\partial\Gamma$ the ingoing congruences will converge, whereas the outgoing ones will diverge. However, this is not necessarily so in non-static universes, where other combinations are possible [40]. As argued, only congruences on negatively expanding light sheets will give rise to some maximum entropy. We can straightforwardly generalise the holographic bound encountered for lightlike screens by stating **Bousso's rule**: *The entropy passing through the negative expansion light sheets constructed from some spacelike boundary $\partial\Gamma$ is less than $A/4$, where A is the area of $\partial\Gamma$* [4].

The formalism of Bousso's rule is set up explicitly to enable simple visual interpretation in the form of Penrose diagrams. Remember that for spherically symmetric spacetimes the angular coordinates are suppressed so that every point in a diagram represents a $(d-1)$-sphere. Hence, any point p in a Penrose diagram can be taken to be some boundary, with the corresponding light sheets represented by two straight lines intersecting at p at right angles. In general spacetimes, the direction of the negative expansion light sheets depends on the position of the boundary in the spacetime. Therefore, in *Penrose-Bousso (PB)* diagrams the spacetime is divided into several regions of constant negative expansion combination of branches. The typical direction of the negative branches of a region are indicated at some point.

To better understand the implications of Bousso's rule in this graphical representation, consider the PB diagram for a positive curvature, matter dominated universe sketched in Fig.(22) which can be separated into four regions $I-IV$. In region I, the universe is expanding so that the cross sections of the future-directed congruences is increasing. This implies that the past branches, which terminate on the $t=0$ initial 'big bang' boundary, have negative expansion. In this case, applying Bousso's rule means that the entropy passing through either of the light sheets cannot surpass one fourth of the area of the 2-sphere p. However, we can place no bounds on the entropy on spacelike surfaces bound by p.

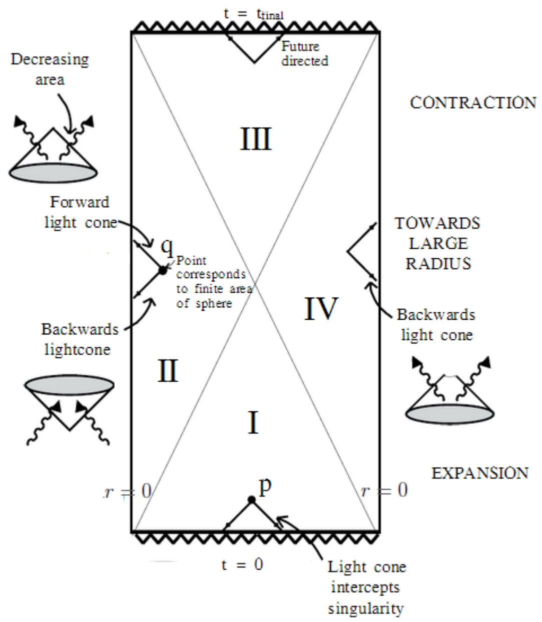
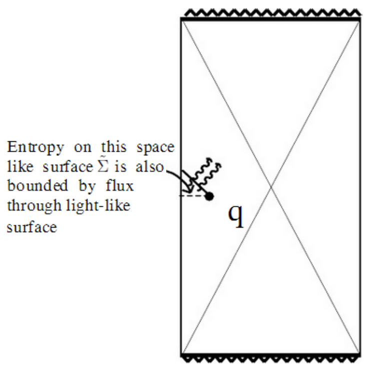
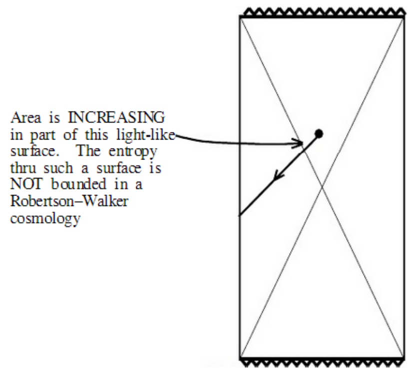


Figure 22: Penrose-Bousso diagram for a matter dominated universe that starts with a big bang at $t = 0$, grows, recollapses and ends in a big crunch at $t = t_{final}$.

The expansion and subsequent contraction have a bigger effect on points of the spacetime that are further away from the spatial origin. Hence, a different picture arises in region II, where the negative expansion light sheets are defined by the branches beginning at 2-sphere q and terminating at $r = 0$ as in Fig.(23a). On both of the light sheets, the entropy is bounded by the surface area of q. Now let us



(a) In region II, there can be spacelike surfaces like $\tilde{\Sigma}$ with an entropic bound determined by a light sheet to its future.



(b) In region III, there are future negative light sheets, but none of them is the complete future of some spacelike surface.

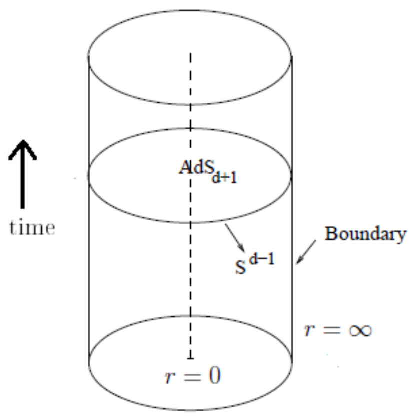
Figure 23: Region II and III of the matter dominated universe.

consider a spacelike surface $\tilde{\Sigma}$ bounded by q and connected to the spatial origin. Clearly, any entropy passing through surfaces must subsequently traverse the future directed negative expansion light sheet. Since we move from $\tilde{\Sigma}$ to this light sheet along with the direction of the entropic time arrow by the second law of thermodynamics the holographically bounded entropy on the light sheet must then be larger then that on $\tilde{\Sigma}$. Hence, one of the conditions for an entropy bound on spacelike surfaces bound by some 2-sphere is that there is a corresponding negative expansion light sheet in its future that complete comprises its entropy. Indeed, such a future light sheet is absent for spacelike surfaces through p in region I. Notice furthermore that the diagram is essentially folded into the surface of a cylinder, where crossing the spatial origin involves moving to the exact antipodal point on the 2-sphere. Hence, the situation is completely equivalent in region IV and the same holographic bound holds for spacelike surfaces with a negative expansion light sheet in their future.

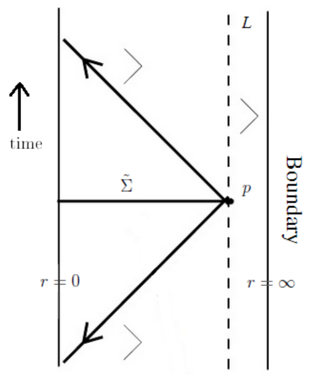
In region III, the relevant light sheets are both future directed because of the contraction of the universe as sketched in Fig.(23b). Notwithstanding, none of them are completely to the future of any spacelike surface, so there are no spacelike entropy bounds [40].

In this fashion, we can study the entropy bounds on arbitrary lightlike and spacelike surfaces in general spacetimes, on the condition that they must be symmetric. Nevertheless, none of these considerations make much physical sense as the surfaces $\tilde{\Sigma}$ are generally not Cauchy surfaces. Also, most spacetimes will not be stationary, so that there is no Killing vector with respect to which there is some notion of time so that the Cauchy surfaces can be evolved by some measurable Hamiltonian of the form (1.2.36).

Suppose, however, we consider symmetric solutions to Einstein's equations with negative cosmological constant Λ. These so-called *Anti-de Sitter (AdS)* spacetimes are the Lorentzian analogues of hyperbolic spaces and as such have an innate boundary. For the $(d+1)$-dimensional spacetime AdS_{d+1}, the boundary contains the time direction and a sphere S^{d-1}. If $d = 3$, the boundary reduces to a simple circle evolving with time, i.e. $\mathbb{R}_{time} \times S^1$, so that the entire spacetime is given by a solid cylinder. This picture can be extended to form the Penrose diagram of d-dimensional AdS spacetimes, where each point is given by a S^{d-1} sphere (Fig.(24a)). In the case $d = 4$, we can suppress the angular coordinate on every line



(a) The Penrose diagram of AdS_{d+1} spacetime may be represented as a solid cylinder, where the vertical direction is time. The boundary contains the time axis and a sphere S^{d-1}, represented as a circle [22].



(b) Penrose-Bousso diagram of AdS_4 spacetime. Any spacelike surface $\tilde{\Sigma}$ in AdS will have an entropy that is holographically bounded by $A/4$, where A is the area of 2-sphere p.

Figure 24: Two Penrose diagrams for Anti-de Sitter spacetime.

of latitude of the cylinder to recover the more familiar notion of the 2-dimensional Penrose diagram, where every point corresponds to a 2-sphere as usual. The PB diagram then consists of an infinite strip bounded on the left by the spatial origin and on the right by a boundary at spatial infinity. For any 2-sphere p on surfaces L that are static with time, the negative expansion light sheets are pointed inwards towards the spatial origin. The entropy of any spacelike surface $\tilde{\Sigma}$ bounded by the spatial horizon and such a point p must later pass through the future directed light sheet and is thus holographically limited by the area A of the 2-sphere at p. The entire spacetime can be foliated by Cauchy-like surfaces in such a way that the entropy on each surface maximally equals $A/4$ [40]. The non-global hyperbolicity of AdS is due to the behaviour of the spacetime near the boundary. However, under appropriate regularity and boundary conditions on the fields it is still possible to define a Hamiltonian description [17].

We have found that Anti-de Sitter spacetime is almost globally hyperbolic, with a holographic bound on the entropy of every Cauchy-like surface. Moreover, it is also a static cosmology, so that there is a timelike Killing vector k with respect to which we can define the time on the vertical axis and a Hamiltonian (1.2.36) with which we can evolve between Cauchy-like surfaces. These three properties imply that all physics taking place in the *bulk* of the spacetime may be described by a Hamiltonian acting on a Hilbert space containing a much-reduced number of states. In particular, the dimension of the Hilbert space is [40]

$$\mathcal{N}_{states} = e^{S_{max}} = e^{S_{BH}} = e^{A/4}. \tag{2.3.6}$$

2.4 The Ads/CFT Correspondence

In the preceding subsection, we have shown that general AdS spacetimes are static and can be foliated by Cauchy-like surfaces, as well as having a natural boundary that holographically limits the entropy of the bulk. Motivated by string theoretical reasons, one of these is of particular interest. $\text{AdS}_5 \times S^5$ is a 10-dimensional product space consisting of a (4+1)-dimensional AdS spacetime and an internal 5-sphere S^5. Generally, the supergravity theories derived from superstring theories are static models, without a cosmological constant Λ. However, by Kaluza-Klein decomposing the additional compactified dimensions of the 5-sphere it is possible obtain an effective field theory on a dimensionally reduced spacetime that does have some cosmological constant [40]. In this product space, the 5-dimensional AdS spacetime can be viewed as a solid ball of four spatial dimensions times the infinite time axis with a metric

$$ds^2 = \frac{R^2}{(1-r^2)^2}\left[-(1+r^2)^2 dt^2 + 4dr^2 + 4r^2 d\Omega_3^2\right]. \tag{2.4.1}$$

Here, t is time, r is a radial coordinate that admits values $0 \leq r < 1$ and Ω_3 parametrises the angular coordinates of this unit 3-sphere. Furthermore, the geometry has a uniform hyperbolic curvature $-1/R^2$, for a radius of curvature denoted by R. Near $r = 0$ the metric looks like that of flat space. With increasing r, the time-time component g_{00} and the radius of the 3-sphere will increase. The growth of g_{00} can be interpreted as a type of gravitational potential that increases from $r = 0$ towards the boundary. For example, a slowly moving scalar particle satisfying the Klein-Gordon equation (1.2.2) experiences a gravitational potential $V \sim \sqrt{-g_{00}}$ that pulls it back to $r = 0$. If such a particle is put at rest at a large value of r, it will start an oscillatory motion in the r-direction, effectively confining particles to the origin [22]. For some constant β, there is a proper distance

$$L = \int_0^L dL' = \int_q^1 dr \, \frac{2}{1-r^2}\left[\frac{1}{R^2} + \beta^2\left(\frac{1-r^2}{1+r^2}\right)^2\right]^{-1/2} \to \infty \tag{2.4.2}$$

from any point p in the bulk to the boundary, so massive particles on timelike geodesics can never reach it (Fig.(25)) but instead are bounced back to the origin by the gravitational potential. However, light

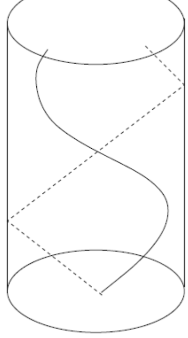


Figure 25: Timelike (solid line) and null (dashed line) radial goedesics in AdS spacetime.

travelling along radial null geodesics can travel in a round trip from $r = 0$ to the boundary and back in a finite coordinate time

$$t = \int_0^t dt' = \left(\int_0^1 - \int_1^0\right) dr \frac{2}{1+r^2} = 4\left[\arctan{(r)}\right]_0^1 = \pi. \tag{2.4.3}$$

Hence, the picture is sketched that Anti-de Sitter space should be regarded as a mirror box with perfectly reflecting inner walls. The size of the space inside of the box is of order R, the curvature scale. Generally, for non-diverging spacetimes R will be a macroscopic quantity that is much larger than any microscopic scale such as the Planck length ℓ_p or the string theoretic string scale ℓ_s [40].

AdS is a spacetime with many symmetries. For example, from the metric (2.4.1) it is obvious that the spatial 3-sphere will have an associated $O(4)$ symmetry group that permutes its points. Other symmetries become more apparent if we study the metric in the neighbourhood of a point at the boundary at $r = 1$. There, the 3-sphere space is well approximated by the geometry of its 3-dimensional flat tangent space parametrised by coordinates $\vec{x} = (x^1, x^2, x^3)$. We can then apply a coordinate transformation $y = (1-r)$ to find

$$ds^2 = \frac{R^2}{y^2}\left(-dt^2 + d\vec{x}^2 + dy^2\right) \qquad \text{for } r \to 1. \tag{2.4.4}$$

From a different viewpoint, Eq.(2.4.4) can be regarded not as an approximation but rather as the exact metric of an incomplete patch of AdS_5 spacetime. A massive particle on a timelike geodesic trajectory can reach $y = 0$ from any point q in the bulk to the boundary in finite proper time

$$\tau = \int_0^\tau d\tau' = \int_q^\infty dy \, \frac{1}{y\sqrt{\varepsilon^2 y^2 - \frac{1}{R^2}}} \nrightarrow \infty, \tag{2.4.5}$$

where ε is some constant. The spacetime described by metric (2.4.4) is geodesically incomplete. It has a horizon at $y = \infty$. In this interpretation, the time coordinate differs from that of AdS spacetime as described by the metric (2.4.1)[40]. From the AdS metric in the form (2.4.4), it is obvious that the geometry of an incomplete patch is invariant under the canonical Poincaré transformations of a 4-dimensional Minkowski spacetime with coordinates $(t, \vec{x})$. These coordinates also display a 'dilatation' isometry

$$(t, \vec{x}, y) \mapsto \lambda(t, \vec{x}, y) \tag{2.4.6}$$

for some constant λ. The spacetime of an incomplete patch is scale invariant. In short, the full symmetry group, which includes the dilatations and the Poincaré group, of AdS_5 is $O(4,2)$[40]. The isometries of AdS spacetime must also act on its edge, where they will have the endomorphic action of permuting boundary points. On the boundary, the action of the complete symmetry group can be identified with that of the conformal group in four dimensions. This is the group of all transformations that preserve all angles on the edge of AdS. Note that because of the conformal symmetry we can specify an arbitrary radius for the AdS spacetime, so that picking the boundary to be a unit 4-sphere times the time axis by limiting r to values between zero and one was entirely legitimate[22].

To extend the spacetime to the full $\text{AdS}_5 \times S^5$ only requires defining five more angular coordinates ω_5 that parametrise the compact directions of the unit 5-sphere. Since the hypersphere should be curved over the same radius R as the AdS part of the product space, it is a surface of curvature $1/R^2$ and we should add a term

$$ds_5^2 = R^2 d\omega_5^2 \tag{2.4.7}$$

to the metric (2.4.1). Evidently, the internal 5-sphere will have an additional $O(6)$ symmetry group. However, by the conformal symmetry of the boundary we may Weyl rescale the metric (2.4.1) by a factor $\frac{R^2}{(1-r^2)^2}$ to

$$d\tilde{s}^2 = \left[-(1+r^2)^2 dt^2 + 4dr^2 + 4r^2 d\Omega_3^2\right] + \left[(1-r^2)^2 d\omega_5^2\right]. \tag{2.4.8}$$

This produces a new rescaled metric $d\tilde{s}^2$ that is in the same conformal group but is finite at the boundary where $r = 1$. Because the radius $(1 - r^2)$ of the 5-sphere vanishes approaching the boundary, the 10-dimensional $\text{AdS}_5 \times S^5$ spacetime has a negligibly small number of internal degrees of freedom in asymptotic regions. The boundary of the full $\text{AdS}_5 \times S^5$ geometry is well approximated by a reduced $(3 + 1)$ spacetime[40].

According to the holographic principle, the entropy of the bulk of $\text{AdS}_5 \times S^5$ is bound by the area of the boundary of AdS_5 at $y = 0$. As mentioned earlier, 't Hooft has suggested that this should be interpreted more broadly as proof of the existence of a theory that has its degrees of freedom on the boundary where they determine completely the evolution of the bulk[19]. However, from Eq.(2.4.1) we infer that that metric diverges on approaching the boundary at $r = 0$. If we refrain from imposing some cutoff scale on the theory, this will mean that the numbers of degrees of freedom per unit coordinate area grows to infinity. At first glance this is a strong indication that the boundary theory has the nature of a continuum quantum field theory. Like the metric of spacetime at the boundary, this QFT must respect the conformal symmetry[40].

Motivated by reasons from string theory, of all AdS spacetimes the 10-dimensional $\text{AdS}_5 \times S^5$ is of special interest. On $\text{AdS}_5 \times S^5$ gravitational phenomena can be described by low energy string theory on the condition that the theory realises a sufficient degree of supersymmetry. Naturally, this so-called $\mathcal{N} = 4$ supersymmetry must also carry over to the holographic theory on the 4-dimensional boundary of AdS_5. All the symmetries considerably narrow down the number of candidate boundary theories that can holographically describe the evolution of the bulk. It turns out that here is only one known class of superconformal Lorentz invariant quantum field theories that also has the appropriate amount of supersymmetry. These are the $SU(N)$ *supersymmetric Yang-Mills (SYM)* theories [40]. Indeed, it has been shown to be possible to describe an exact 'recipe' to convert correlation functions of the SYM boundary theory to string theory calculations performed in the bulk of AdS [51]. First proposed by Maldacena in 1997, this duality between Type IIB superstring theories in an AdS background and supersymmetric Yang-Mills field theories on its boundary has been named the *AdS/CFT correspondence* [20].

Because SYM is conformally invariant, it does not have a characteristic scale and therefore its parameters must be dimensionless. It has proven convenient to define the theory in such a way that the AdS boundary is parametrised by dimensionless coordinates $(t, \vec{x})$. For the sake of uniformity, we also adhere to the convention that all SYM quantities, such as momentum and energy, are dimensionless. However, in the AdS interior the spacetime is not invariant under conformal transformations. Quantities appearing in the bulk theory will have their usual dimensions. Hence, there will be a conversion factor, given by the radius of curvature R, relating SYM variables to their quantum gravity equivalents. For example, a time interval Δt_{SYM} measured on the boundary is given by an interval

$$\Delta t_{bulk} = R\,\Delta t_{SYM} \tag{2.4.9}$$

in the quantum gravity theory in the bulk. Likewise, a mass M in the bulk translates to a mass of M/R in the Yang-Mills boundary theory. There are, however, two natural dimensionless constants in the bulk string theory. These are

1) The radius of curvature of the AdS spacetime measured in Planck units ℓ_p or string units ℓ_s. Because both R and ℓ_p or ℓ_s are dimensionful length scales characteristic to the bulk theory, their quotients R/ℓ_p and R/ℓ_s are dimensionless. This is also reason to deviate from the unit normalisation convention and reinstate the 10-dimensional Newton's constant G.
2) The relation between string length and Planck length is given by

$$g_s^2 \ell_s^8 = \ell_p^8 = G$$

where g_s is the dimensionless string coupling constant.

The SYM theory defined on the boundary is a gauge theory. There are two constants associated with gauge group $SU(N)$, namely

1) The rank N of the gauge group.
2) The gauge coupling constant g_{ym} that determines the strength of the gauge field self-interaction.

Evidently, the two bulk parameters R and g_s are holographically related to the SYM parameters N and g_{ym}. As was first derived by Maldacena in his original paper [20], these relations can be taken to be

$$\begin{aligned} \left(\frac{R}{\ell_s}\right)^4 &= N g_{ym}^2 \\ g_s &= g_{ym}^2. \end{aligned} \tag{2.4.10}$$

Combining these relations, we find for the 10-dimensional Newton's constant that [40]

$$G = \frac{R^8}{N^2}. \tag{2.4.11}$$

In order to learn something more about the reduction in the number of degrees of freedom caused by the holographic principle, consider describing the bulk by a ordinary quantum field theory with a cutoff of microscopic scale, such as the Planck length. This can for instance be done by constructing a random spatial lattice with an average spacing of ℓ_p in the 9 dimensions of $\text{AdS}_5 \times S^5$. Each of the lattice sites will contain a single simple degree of freedom, like the spin state of a photon. As we have seen, the area of the boundary is infinite. To regulate this divergence, we also introduce a surface L with a finite area at $y = \delta_{IR}$ like in Fig.(24b). The volume of the 5-sphere is simply given by

$$V_{\omega_5} = \frac{8\pi^2}{15} R^5. \tag{2.4.12}$$

Furthermore, using the metric (2.4.4) the volume element of AdS_5 is calculated as

$$dV_{AdS_5} = \left(\frac{R}{y}\right)^4 d\vec{x}\, dy. \tag{2.4.13}$$

Integrating this volume element up to the boundary at δ_{IR}, we find that the 9-dimensional volume is

$$V(\delta) \sim \frac{R^9}{{\delta_{IR}}^3}. \tag{2.4.14}$$

Therefore, the number of bulk lattice sites of a 9-dimensional Planck volume, each holding one bit, is

$$\mathcal{N}_{sites} = \frac{V}{\ell_p^9} \sim \frac{1}{{\delta_{IR}}^3} \frac{R^9}{\ell_p^9}. \tag{2.4.15}$$

Not taking into account gravitational effects, the maximum entropy comprised in the bulk will be of the same order. If, however, we factor in gravity this measure will be a great overestimation. By the holographic principle, the entropy and number of degrees of freedom are bounded to an order

$$S_{max} \sim \mathcal{N}_{dof} \sim \frac{A}{l_p^9}, \tag{2.4.16}$$

where A is the 8-dimensional area of boundary cutoff L. Since the area element of the complete product spacetime is

$$dA_{AdS_5} = \left(\frac{R}{y}\right)^3 d\vec{x}, \tag{2.4.17}$$

we can easily compute that, evaluated at $y = \delta_{IR}$,

$$A \sim \frac{R^8}{{\delta_{IR}}^3} \tag{2.4.18}$$

and hence

$$S_{max} \sim \mathcal{N}_{dof} \sim \frac{1}{{\delta_{IR}}^3} \frac{R^8}{\ell_p^8}. \tag{2.4.19}$$

We see that the holographic principle dictates a reduction in the degrees of freedom of the bulk theory by a factor ℓ_p/R. This implies that a complete description of all the processes in the bulk requires a density of degrees of freedom of

$$\frac{\mathcal{N}_{dof}}{V} \sim \frac{\ell_p}{R} \frac{1}{\ell_p^9}, \tag{2.4.20}$$

or merely ℓ_p/R degrees of freedom per spatial Planck volume. Nevertheless, the theory must still be able to completely characterise all physics in the bulk, even when the AdS spacetime becomes very large. Hence, we will be most interested in the case where the microscopic parameters for the bulk theory remain fixed but R increases, i.e. the limit

$$\begin{aligned} g_s &= constant, \\ R/\ell_s &\to \infty. \end{aligned} \tag{2.4.21}$$

By equation (2.4.10), this corresponds on the SYM boundary side to [40]

$$\begin{aligned} g_{ym} &= constant, \\ N &\to \infty. \end{aligned} \tag{2.4.22}$$

As we have argued, the proper area of a finite coordinate patch tends to infinity as the surface is taken closer to the boundary. Hence, one would expect the density of degrees of freedom on the boundary to diverge. This is internally consistent with the idea that SYM, in essence a continuum field theory, will have an infinite amount of degrees of freedom in any spatial volume. This suggest that the infrared (IR) cutoff L of the bulk theory used to regulate down the diverging boundary area is related in some way to an ultraviolet (UV) cutoff in the boundary SYM theory that can act as an average lattice spacing length.

To come to a better understanding of this relationship, we need to address some of the properties of D-branes [40].

D-branes are mathematical constructs appearing in the vacuum state of supersymmetric string theories. Since open strings may end on D-branes, the arrangement of the D-branes determines the types of string states in a system. Any string may either connect two separate branes or end on the same brane with both its endpoints. Hence, they can be classified by two coordinates i and j, corresponding to the branes on which they end. Strings are oriented objects, so that the classification cannot be symmetric, i.e. a $[ij]$ string is not equal to a $[ji]$ string.

Because all strings have some characteristic tension, it requires energy to span a string between branes. The separation of D-branes thus gives certain masses to the connecting strings. Furthermore, the spatial orientation of the D-branes highly constricts the way strings can move and vibrate. Now, since string excitations can be interpreted as different particle states, the orientation of the D-branes in spacetime controls the particle spectrum of the string theory. The particle sector originating from strings $[ii]$ beginning and ending on the same brane, for example, will contain several massless particles. If a D_p-brane, which fills p spatial dimensions and the time axis, is embedded in a $(d+1)$-dimensional spacetime, there will be $(d-p)$ massless scalars living on its world volume. In fact, these massless scalars are Goldstone bosons corresponding to the broken symmetries of the vacuum in the $(d-p)$ directions perpendicular to the brane [52].

Now consider strings $[ij]$ with $i \neq j$ that have their endpoints on different branes. These strings overlap in their domains and might interact with one another. For example, one of the mechanisms by which such an interaction could occur is the simple joining of endpoints, so that a string in the $[ij]$ sector may combine with a $[jk]$ string to form a $[ik]$ string. As previously argued, these strings will have a mass corresponding to the seperation of the branes in the stack. Moving the branes closer and closer until they coincide then preserves the particle sectors of the theory, but sends the masses of these particles to zero. The total particle spectrum originating from a system of N of these D-branes at the same coordinate comprises many different interacting quantum fields, depending in nature on the open strings on the branes. These fields can then be described by a corresponding $SU(N)$ gauge theory [52].

Consider then the situation of some D3-brane embedded in a 10-dimensional spacetime. This brane will fill the time axis, as well as three spatial coordinates which we group together by denoting them by $\vec{x}$. The six remaining coordinates will be named $\tilde{z}^m$, and we define overarching variable $\tilde{z}$ by $\tilde{z}^2 = \sum_m (\tilde{z}^m)^2$. Then, the scalar fields arising as Goldstone excitations by vacuum symmetry breaking can be identified with the location $\tilde{z}$ of the D3-brane. Juxtaposing N of these D3-branes gives rise to different fields. Manipulating the open string spectrum of the branes, we can create a situation in which the fluctuations of the stack are given by a $\mathcal{N}=4$ supersymmetric Yang-Mills gauge theory. All of the fields in this theory are in a single supermultiplet of the adjoint representation of $SU(N)$, so that they are given by $N \times N$ matrices. By the supergravity equations of motion, a stack of D-branes initially placed at the origin $\tilde{z}=0$ will source a geometry of the form

$$ds^2 = F(\tilde{z})(-dt^2 + d\vec{x}^2) + F(\tilde{z})^{-1} d\tilde{z}^2, \tag{2.4.23}$$

where

$$F(\tilde{z}) = \left(1 + \frac{a g_s N}{\tilde{z}^4}\right)^{-1/2} \tag{2.4.24}$$

for some numerical constant a. Now, if we impose an IR cutoff L at $y = \delta_{IR}$ and assume there must also be some microscopic cutoff scale for the bulk theory, we find the limits (2.4.22) for a very large AdS spacetime with the microscopic parameters fixed. Then, we may take the limit in which $\frac{a g_s N}{\tilde{z}^2} \gg 1$, so that by approximation

$$F(\tilde{z}) \cong \frac{\tilde{z}^2}{(a g_s N)^{1/2}}. \tag{2.4.25}$$

Then, the D-brane metric becomes

$$\begin{aligned} ds^2 &= \frac{\tilde{z}^2}{(a g_s N)^{\frac{1}{2}}}(-dt^2 + d\vec{x}^2) + \frac{(a g_s N)^{\frac{1}{2}}}{\tilde{z}^2} d\tilde{z}^2 \\ &= \frac{\sqrt{a}R^2}{\ell_s^2}\frac{1}{\tilde{z}^2}(-dt^2 + d\vec{x}^2) + \frac{\ell_s^2}{\sqrt{a}R^2}\tilde{z}^2 d\tilde{z}^2. \end{aligned}$$

Simply rescaling the coordinates as

$$\begin{aligned} \tilde{z} &\rightarrow R\frac{a^{1/8}}{\sqrt{\ell_s}}\,\tilde{z} \\ (t,\vec{x}) &\rightarrow R\frac{\sqrt{\ell_s}}{a^{1/8}}\,(t,\vec{x}), \end{aligned} \tag{2.4.26}$$

means that this line element becomes

$$ds^2 = R^2\left[\tilde{z}^2(-dt^2 + d\vec{x}^2) + \frac{1}{\tilde{z}^2}d\tilde{z}^2\right], \tag{2.4.27}$$

which may be identified with Eq.(2.4.4) after it is expressed in terms of coordinate $y = 1/\tilde{z}$. Hence, the geometry sourced by the stack of D3-branes is equal to the metric on the boundary of $\mathrm{AdS}_5 \times S^5$. In the form (2.4.27) the boundary of AdS is at $\tilde{z} = \infty$ and there must be a horizon at $\tilde{z} = 0$, because the time-time component vanishes at that point. The six coordinates $\tilde{z}^m$ under $\tilde{z}$ that indicate the position of the branes can now be recognised as the Cartesian versions of the $\mathrm{AdS}_5 \times S^5$ coordinates z, ω_5, with

$$z^2 \equiv \tilde{z}^2 = (z^1)^2 + \cdots + (z^6)^2 \tag{2.4.28}$$

as radial coordinate. In the SYM theory, these coordinates can be thought of as six scalar fields ϕ^m living on the brane world volume. The relation between the coordinates and the fields is then given by

$$z^m \leftrightarrow \frac{g_{ym}\ell_s^2}{R^2}\,\phi^m. \tag{2.4.29}$$

However, this cannot be a direct relation as the ϕ^m are $N \times N$ matrices in the adjoint representation of $SU(N)$. Instead, Eq.(2.4.29) should be interpreted as stating that the N eigenvalues of the six fields ϕ^m are identified as the six coordinates z^m of the N branes[39]. In this spirit, it is possible to write down an exact expression only for the radial component z,

$$z^2 = \left(\frac{g_{ym}\ell_s^2}{R^2}\right)^2 \frac{1}{N}\,\mathrm{Tr}\,(\phi^2) \tag{2.4.30}$$

where $\phi^2 = \sum_m (\phi^m)^2$ [40].

Now, the ϕ^m are quantum fields subjected to quantum fluctuations, which will reflect on the position of the coincident D3-branes. It is known that the fields scale with a dimension of inverse length, so that the squared mean of any of their N^2 components satisfies

$$\langle \phi_{ab}^2 \rangle \sim {\delta_{UV}}^{-2}, \tag{2.4.31}$$

where δ_{UV} is the ultraviolet cutoff imposed to regulate the diverging number of degrees of freedom in a continuum SYM quantum field theory. By the relation (2.4.30) and Eq.(2.4.10), this translates to an uncertainty in the average value of z of

$$\begin{aligned} \langle z \rangle^2 &\sim \left(\frac{g_{ym}\ell_s^2}{R^2}\right)^2 \frac{N}{{\delta_{UV}}^2} \\ &\sim {\delta_{UV}}^{-2}. \end{aligned} \tag{2.4.32}$$

From here, it is obvious that the location of the brane is influenced by the extent to which the field theory on the boundary is cut off on the ultraviolet end. Evidently, examining the branes with low frequency up to only a relatively large cutoff δ_{UV} will center them very close to the origin at $z = 0$. But as the frequency of the probe increases and δ_{UV} decreases, the brane appears to move toward the boundary at $z = \infty$. Here the proper area of any finite coordinate patch will diverge. This also implies that the number of degrees of freedom of the fields that is associated with such a path will tend to infinity towards the boundary. Thus, there will necessarily be a lower limit to the UV cutoff scale δ_{UV} that is given by the IR cutoff scale δ_{IR}, so that

$$\delta \equiv \delta_{UV} \sim \delta_{IR}. \tag{2.4.33}$$

This relation between the UV and IR cutoffs was dubbed the *UV/IR connection*[39].

As a corollary of the UV/IR connection, the region $y = \delta$ inside of the cutoff L can be described by a supersymmetric Yang-Mills boundary theory, as long as it is ultravioletly regulated with a cutoff length that is of the same order as δ. The boundary, a 3-dimensional spatial surface evolving in time, will be divided into cutoff cells with a volume of δ^3. As such, a boundary patch of unit coordinate area will contain $1/\delta^3$ cells. In a consistent cutoff theory, the fields will be constant inside every such patch. There will thus be density of N^2 degrees of freedom, corresponding to the $N \times N$ components of the field matrices, per unit coordinate area. As a result, there will approximately be a total number of degrees of freedom

$$\mathcal{N}_{dof} \approx \frac{N^2}{\delta^3}. \tag{2.4.34}$$

Then, as calculated in Eq.(2.4.18), the 8-dimensional area of the unit coordinate patch is given by $A \sim R^8/\delta^3$. The number of degrees per unit area becomes

$$\frac{\mathcal{N}_{dof}}{A} \sim \frac{N^2}{R^8} = \frac{1}{G}, \tag{2.4.35}$$

where we have used Eq.(2.4.11) to express the 10-dimensional Newton's constant in terms of rank N and radius of curvature R. With the knowledge that the degrees of freedom can be taken to be simple, spinlike states, we find that the entropy density on the boundary is of the same order,

$$s \sim \frac{\mathcal{N}_{dof}}{A} \sim \frac{1}{G}, \tag{2.4.36}$$

so that we have retrieved the bound required by the holography from string theory principles [39,40]. Hence, we can conclude that the AdS/CFT correspondence provides a self-consistent way of correlating effects in quantum gravity with proceedings in a purely field theoretic theory that does not take gravitational effects into account.

2.5 The AdS-Schwarzschild Spacetime: a Black Hole in a Box

As we have seen, an AdS geometry can be viewed as a spherically symmetric spacetime with an outwardly increasing gravitational potential. Under influence of this potential, any massive particle will perform an oscillatory motion around the coordinate origin and will eventually be drawn in to that point. It is then completely imaginable that one could keep introducing mass into the spacetime, until eventually the Chandrasekhar limit is exceeded and the lump will collapse into a black hole under the influence of gravity [41]. The resultant spacetimes are solutions to the 5-dimensional Einstein equations with a negative cosmological constant. Stable black holes will have a Schwarzschild radius that is at least as large as the radius of curvature R. Furthermore, if we extend the spacetime to full $\text{AdS}_5 \times S^5$ they will homogenously fill the compact dimensions of the 5-sphere [40].

By a coordinate transformation of Eq.(2.4.1), we find that a more generalisable way of writing the metric of AdS_5 spacetime is

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_3^2 \tag{2.5.1}$$

where $f_{AdS}(r) = \left(1 + \frac{r^2}{R^2}\right)$. Adding mass to the geometry leads to a modification of function $f(r)$ to $f_{A-S}(r) = \left(1 + \frac{r^2}{R^2} - \frac{\mu G}{R^5 r^2}\right)$ for the metric of AdS-Schwarzschild spacetime. The horizon of the black hole sits on a Schwarzschild radius that is the largest solution of

$$f_{A-S}(r) = \left(1 + \frac{r^2}{R^2} - \frac{\mu G}{R^5 r^2}\right) = 0, \tag{2.5.2}$$

so it depends on a parameter μ that is proportional to the black hole mass M. The thermodynamic properties of the black hole can be established by calculating the area element associated with AdS-Schwarzschild spacetime, integrating it to find the area of the horizon and applying the Bekenstein-Hawking formula. The entropy of the AdS black hole will depend on its mass M as

$$S \sim \left(\frac{M^3 R^{11}}{G}\right)^{1/4}. \tag{2.5.3}$$

By the fundamental relation of thermodynamics (1.1.5), the temperature at which the black hole radiates its Hawking radiation is

$$T = \left(\frac{dS}{dM}\right)^{-1} \sim \frac{M^2 R^{11}}{G\left(\frac{M^3 R^{11}}{G}\right)^{3/4}}. \tag{2.5.4}$$

Thus, the black hole mass is related to the temperature as

$$M \sim \frac{R^{11}T^4}{G}. \tag{2.5.5}$$

The AdS/CFT correspondence promises that it must be possible to describe the Hawking radiation process of the black hole occuring in the bulk of the AdS spacetime by a conformal field theory on its lower-dimensional boundary. Translated into dimensionless quantities in the SYM theory on the boundary of AdS that are related as in Eq.(2.4.9), this becomes

$$\begin{aligned} E_{sym} &\sim \frac{R^8}{G} T_{sym}^4 \\ &\sim N^2 T_{sym}^4. \end{aligned} \tag{2.5.6}$$

By integrating Planck's law, the energy density ε in a volume V is given by

$$\varepsilon = \frac{E}{V} \sim T^4. \tag{2.5.7}$$

Consider then a patch of the boundary of unit coordinate 3-area. Instigating a cutoff in the SYM theory, there will be N^2 degrees of freedom in such a patch. Up to a constant of proportionality, Eq.(2.5.5) can then be interpreted as the energy of a thermal gas of N^2 different quanta in the volume of a boundary sphere. This leads to the conclusion that the holographic image of the black hole is formed by the simple quantum mechanical system of N^2 particle species propagating on the boundary. Evidently, the evolution of the black hole, from formation to evaporation, must be completely describable in terms of some unitary transformation [40].

When applied to the black hole case, the AdS/CFT correspondence offers the perfect tool to study the details of Hawking radiation in an environment where it is categorically impossible to have order unity changes to entanglement state (1.5.15) and see where bulk quantum theory fails. To this end, we will consider a minimally coupled 10-dimensional scalar dilaton field ϕ. If we take ϕ to be constant over the 5-sphere, it will have a 5-dimensional action in AdS spacetime of the form

$$S = \int d^5x \, \sqrt{-g} \, g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi. \tag{2.5.8}$$

In order for a Hamiltonian description to exist, we must also demand that the field ϕ satisfies a vanishing boundary condition at the AdS boundary, i.e. $\phi \to 0$ as $r \to \infty$. Belonging to this action is the massless Klein-Gordon equation

$$\Box \phi = \nabla^\mu \nabla_\mu \phi = 0. \tag{2.5.9}$$

Substituting the generalised AdS metric (2.5.1), by Eq.(1.4.12) we have partwise

$$\begin{aligned} \frac{1}{\sqrt{-g}} \partial_t (g^{tt} \sqrt{-g} \partial_t) &= -\frac{1}{f} \partial_t^2 \\ \frac{1}{\sqrt{-g}} \partial_r (g^{rr} \sqrt{-g} \partial_r) &= \frac{1}{r^3} \partial_r (r^3 f \partial_r) \\ \frac{1}{\sqrt{-g}} \partial_{\sigma_i} (g^{\sigma_i \sigma_j} \sqrt{-g} \partial_{\sigma_j}) &= \frac{1}{r^2} \Delta_{S^3} \end{aligned} \tag{2.5.10}$$

where we have chosen coordinates $\sigma_i = (\sigma_1, \sigma_2, \sigma_3)$ on S^3 and Δ_{S^3} is the 3-dimensional spherical Laplacian [11]. The complete d'Alembertian will be the sum of these factors,

$$\Box = -\frac{1}{f} \partial_t^2 + \frac{1}{r^3} \partial_r (r^3 f \partial_r) + \frac{1}{r^2} \Delta_{S^3} \, . \tag{2.5.11}$$

From this form of the operator it is clear that solutions to the 5-dimensional Klein-Gordon equation in general AdS spacetimes will generally be periodic functions of coordinate time t. However, in the black hole metric in turns out that field decays exponentially,

$$\phi \to \exp(-\gamma t) \tag{2.5.12}$$

where γ is some dimensionless monotonically increasing function of the black hole mass M [17,40]. This behaviour of the field reflects also on the quantum correlation functions in the AdS-Schwarzschild bulk. Because of Eq.(2.5.12), correlators must have the form

$$\langle \phi(t)\phi(t')\rangle \sim \exp\left(-\gamma|t-t'|\right) \tag{2.5.13}$$

for some large interval $\Delta t = |t-t'|$. Since any correlator will decrease exponentially, the bulk spacetime displays a phenomenon called *asymptotic stability.* Any perturbation of the field ϕ at initial time t will dissipate away and ultimately become negligibly small. Such behaviour is mostly seen in infinite thermal systems, where γ is some sort of dissipation coefficient. However, it stands perpendicular to what one would expect to observe for any system with finite entropy such as the AdS black hole [40].

The key insight comes from the realisation that the entropy of a system is an emergent quantity derived from its number of microstates. It follows that any system with finite entropy must have a discrete spectrum of states. In general, a closed system with a discrete spectrum of states and corresponding finite entropy may be described by a thermal density matrix and a thermal correlation function

$$\begin{aligned} F(t) = \langle A(0)A(t)\rangle &= \frac{1}{Z}\,\mathrm{Tr}\left(e^{\beta H}A(0)e^{iHt}A(0)e^{-iHt}\right) \\ &= \frac{1}{Z}\sum_{ij} e^{-\beta E_i}e^{i(E_j-E_i)t}|A_{ij}|^2, \end{aligned} \tag{2.5.14}$$

where Z is the partition function and we have inserted two complete sets of energy eigenstates $\sum_i |\psi_i\rangle\langle\psi_i|$. Let us then consider the long term average of $F(t)F^y*(t)$ that is given by

$$L = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{+T} dt\; F(t)F^*(t). \tag{2.5.15}$$

Inserting expression (2.5.14) for the thermal correlation function, this becomes

$$L = \frac{1}{Z^2}\sum_{ijkl} e^{-\beta(E_i+E_k)}|A_{ij}|^2||A_{kl}|^2\;\delta(E_i-E_j+E_l-E_k), \tag{2.5.16}$$

which obviously has a non-zero, positive value. Therefore, the correlator $F(t)$ of a general sytem with finite entropy cannot possibly vanish in the infinite time limit. Hence, perturbations of such finite quantum system will not be found to vanish. Since the proceedings in the bulk of AdS are completely determined by calculations of a conventional quantum system on its boundary, its correlators should not tend to zero. By estimation, the value of the long time average L of finite systems like the AdS black hole is of the order e^S, where S is the finite entropy. However, without a cutoff the entropy attributed to the fields is UV divergent in the neighbourhood of the boundary, so that correlation functions in the approximation QFT in the bulk vanish exponentially as in Eq.(2.5.13). This leads to the remarkable conclusion that the miscounting of short wavelength modes near the horizon in the approximation QFT is responsible for the erroneous description of various phenomena, such as those apparently causing information loss [40].

According to Hawking's claim, mixed-state Hawking radiation follows as a general property of event horizons and should hold in any geometrical background. By the duality of gravity theories in the bulk of AdS and conformal gauge theories on its boundary, Hawking radiation in the AdS-Schwarzschild geometry can be completely described by some kind of unitary process on the boundary. This implies that information does not disappear into non-existence with the evaporation of the AdS black hole. Even though neither the string theory in the bulk of the AdS spacetime nor the SYM theory on the boundary have the right properties to describe the physical world, they offer a mathematically consistent theory of some world and thus invalidate the conclusion of non-unitary evolution of the information paradox. Notice, however, that the AdS/CFT correspondence only conclusively proves that Hawking radiation must in fact be pure, and not whether the original Hawking calculation leading to black hole radiation and evaporation is inaccurate or equivocal.

In conclusion, the holographic principle and the gauge/gravity correspondence of AdS/CFT decisively validate postulates 1 and 2 of black hole complementarity. Assuming the validity of the third postulate, the case seems strong that observations made by infalling observers differ significantly from those by stationary observers far removed from the black hole and its gravitational effects as predicted by BHC.

3 The Black Hole Firewall Paradox

3.1 The Paradox: Internal Contradiction of the Postulates of Black Hole Complementarity

Up to very recently, black hole complementarity, which is backed up by the AdS/CFT correspondence, was widely believed to have resolved the information paradox conclusively. However, it has since been discovered that BHC, as defined by its three axioms that are restated here for convenience,

Postulate 1: *Purity of Hawking radiation*
Postulate 2: *Semi-classical behaviour outside the stretched horizon*
Postulate 3: *Absence of infalling drama*

is internally inconsistent [1].

Consider a black hole that was formed through the gravitational collapse of matter in some pure state and is now evaporating by emitting Hawking radiation. Assume that the black hole has grown 'old' in the sense that it has surpassed its *Page time*, the point when it has only half its initial entropy remaining. Arbitrarily setting the Page time at Schwarzschild time $t = 0$, we may divide the Hawking radiation into an early part R_E emitted while $t < 0$ and a late part R_L that is emitted for $t > 0$. By postulate 1, the Hawking radiation is in a pure state, so that R_E should be entangled with R_L and

$$|\Psi\rangle = \sum_i |\psi_{i_E}\rangle \otimes |e_{i_L}\rangle, \tag{3.1.1}$$

where the states $|e_{i_L}\rangle$ form some orthonormal basis for the late radiation. Because of the exponential relationship between the number of states of a system and the corresponding entropy, the early subspace containing the states $|\psi_{i_E}\rangle$ will be much larger than the late subspace of states $|e_{i_L}\rangle$,

$$(\mathcal{N}_{states})_L < e^{S_{Page}} = e^{S/2} \ll e^S = (\mathcal{N}_{states})_E. \tag{3.1.2}$$

As a consequence, the reduced density matrix describing the late-time radiation can be taken to be approximately equal to the identity for 'old' black holes.

As was argued leading up to Fig.(7), the decomposition of the quantum fields outside the black hole horizon will give rise to a situation where the outgoing wave modes are scattered off or transmitted through an effective potential depending on their angular momentum j. In order not to complicate the argument, we will treat these black hole *gray-body factors* by assuming some low momentum waves are transmitted completely, whereas the higher partial waves will be fully reflected. Although this massive simplification is unjustified, we will come to the same qualitative conclusion if we include more arbitrary gray-body factors. Because the total energy radiated away from the black hole is finite, we can think of the Hawking radiation as living in a finite-dimensional Hilbert space. Now, we can assume the state $|\Psi\rangle$ is randomly distributed in this space, so that it has uniform measure along some orthonormal basis like in the microcanonical ensemble [1].

Consider the late projection operator $P^i = |e_{i_L}\rangle\langle e_{i_L}|$ that projects onto the state $|e_{i_L}\rangle$. Note that we are free to specify index i, so that we can project onto any state of the late radiation. We then propose that an observer measuring the early radiation will be able to predict the action of P^i by using the operator $\hat{P}^i = L|\psi_{i_E}\rangle\langle\psi_{i_E}|$ acting on the early radiation, or

$$L \sum_j |\psi_{i_E}\rangle\langle\psi_{i_E}|\psi_{j_E}\rangle \otimes |e_{j_L}\rangle = \hat{P}^i|\Psi\rangle \approx P^i|\Psi\rangle = |\psi_{i_E}\rangle \otimes |e_{i_L}\rangle, \tag{3.1.3}$$

for $E \gg L$. Here, E and L measure the finite number of dimensions of the early and late Hilbert spaces, so that $1 \leq i, j, \cdots \leq L$ and $1 \leq a, b, \cdots \leq E$ for some arbitrary basis of R_E. Note that if the states $|\psi_{i_E}\rangle$ formed an orthonormal basis by themselves this approximation would become an equality. The relative error of this assumption is

$$\begin{aligned}\mathcal{E} &= \frac{||(P^i - \hat{P}^i)|\Psi\rangle||^2}{||P^i|\Psi\rangle||^2} = \frac{\langle\Psi|P^{i\dagger}P^i - P^{i\dagger}\hat{P}^i - \hat{P}^{i\dagger}P^i + \hat{P}^{i\dagger}\hat{P}^i|\Psi\rangle}{\langle\Psi|P^{i\dagger}P^i|\Psi\rangle} \\ &= \frac{\langle\psi_{i_E}|\psi_{i_E}\rangle\left(1 - 2L\langle\psi_{i_E}|\psi_{i_E}\rangle + L^2\sum_j\langle\psi_{j_E}|\psi_{i_E}\rangle\langle\psi_{i_E}|\psi_{j_E}\rangle\right)}{\langle\psi_{i_E}|\psi_{i_E}\rangle} \\ &= (1 - L\langle\psi_{i_E}|\psi_{i_E}\rangle)^2 + L^2\sum_{i\neq j}|\langle\psi_{i_E}|\psi_{j_E}\rangle|^2. \end{aligned} \tag{3.1.4}$$

Now, if we expand the early states $|\psi_{i_E}\rangle$ in some arbitrary basis $|e_{a_E}\rangle$ with measures c_{ia}, we will find

$$\overline{c_{ia}^* c_{jb}} = \frac{1}{LE}\delta_{ij}\delta_{ab}, \qquad \overline{c_{ia}^* c_{jb} c_{kc}^* c_{ld}} = \frac{1}{L^2E^2}(\delta_{ij}\delta_{kl}\delta_{ab}\delta_{cd} + \delta_{il}\delta_{jk}\delta_{ad}\delta_{bc}), \tag{3.1.5}$$

since we have chosen to average over $|\Psi\rangle$ with uniform measure. Then, we see that

$$\begin{aligned}
\overline{\langle\psi_{i_E}|\psi_{j_E}\rangle} &= \sum_{a,b} \overline{c_{ia}^* c_{jb}}\,\langle e_{a_E}|e_{b_E}\rangle \\
&= \frac{1}{LE}\sum_a \langle e_{a_E}|e_{a_E}\rangle\,\delta_{ij} \\
&= \frac{1}{L}\delta_{ij},
\end{aligned} \tag{3.1.6}$$

and

$$\begin{aligned}
\overline{\langle\psi_{i_E}|\psi_{j_E}\rangle\langle\psi_{k_E}|\psi_{l_E}\rangle} &= \sum_{a,b,c,d} \overline{c_{ia}^* c_{jb} c_{kc}^* c_{ld}}\,\langle e_{a_E}|e_{b_E}\rangle\langle e_{c_E}|e_{d_E}\rangle \\
&= \frac{1}{L^2E^2}\left(\sum_{a,c}\langle e_{a_E}|e_{a_E}\rangle\langle e_{c_E}|e_{c_E}\rangle\,\delta_{ij}\delta_{kl}\right) \\
&\quad + \frac{1}{L^2E^2}\left(\sum_{a}\langle e_{a_E}|e_{a_E}\rangle\langle e_{a_E}|e_{a_E}\rangle\,\delta_{il}\delta_{jk}\right) \\
&= \frac{1}{L^2}\delta_{ij}\delta_{kl} + \frac{1}{L^2E}\delta_{il}\delta_{jk}.
\end{aligned} \tag{3.1.7}$$

Thus, the average of the relative error (3.1.4) is given by

$$\begin{aligned}
\overline{\mathcal{E}} &= (1 - L\frac{1}{L}\delta_{ii})^2 + L^2\sum_{i\neq j}\left(\frac{1}{L^2}\delta_{ij}\delta_{ji} + \frac{1}{L^2E}\delta_{ii}\delta_{jj}\right) \qquad (no\, traces) \\
&= 0 - L^2\frac{1}{L^2E}(L-1) \\
&\simeq \frac{L}{E},
\end{aligned} \tag{3.1.8}$$

for $E \gg L \gg 1$. Note that the exponential decrease of $\overline{\mathcal{E}}$ after the Page time of the black hole implies that the approximation in Eq.(3.1.3) will become more and more precise. Hence, by superposing operators $\hat{P}^i$ we can create early radiation operators that act on the full state $|\Psi\rangle$ like a projection operator projecting onto any subspace of the late radiation [1].

Now, we apply this knowledge to an outgoing late-time Hawking mode, which we take to be a localised wave packet with width of order the Schwarzschild radius R_s and corresponding frequencies of order R_s^{-1}. By postulate 2, there will be a unique set of raising and lowering operators $(b^\dagger, b)$ associated with this mode, so that we can take a sum of projection operators P^i to project onto the eigenspaces of the number operator $N_{out} = b^\dagger b$. Thus, an observer far from the horizon measuring the early radiation is able to determine the number of particles in a given mode of the late-time Hawking radiation R_L as would be measured by another observer in region III of Fig.(26). Imagine now that after measuring R_E, the observer starts diving into the black hole and encounters the late radiation in region I, at some distance from the stretched horizon. As the observer could have deduced, the late-time radiation would have been in an eigenstate of the number operator N_{out} far from the black hole. By postulate 2, the modes making up this radiation would have undergone an essentially free evolution back in time through region II towards the horizon, were it not for the measurements of the observer. However, it is impossible for the radiation to simultaneously be in the in vacuum state defined by $a|0_{in}\rangle = 0$ associated with the infalling observer and an N_{out} eigenstate, which can be easily demonstrated by considering the Bogoliubov transformation relating the 'distant' operators $(b^\dagger, b)$ to the infalling operators $(a^\dagger, a)$. Assuming that the late radiation is in a state Ψ_L which can be identified with the a-vacuum, we see

$$\begin{aligned}
N_{out}|\Psi_L\rangle &= b^\dagger b|\Psi_L \\
&= \left[\int_0^\infty d\omega\,(B^* a_\omega^\dagger + C^* a_\omega)(B a_\omega + C a_\omega^\dagger)\right]|\Psi_L\rangle \\
&= \left[\int_0^\infty d\omega\, B^* C a_\omega^\dagger a_\omega^\dagger + C^* C a_\omega a_\omega^\dagger\right]|\Psi_L\rangle,
\end{aligned} \tag{3.1.9}$$

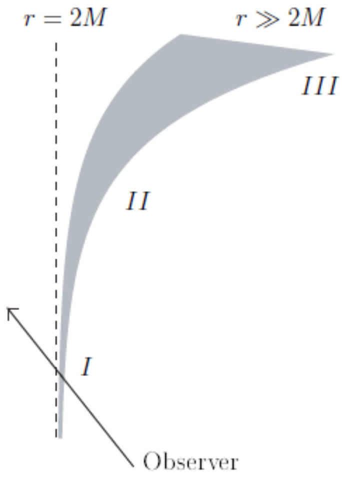


Figure 26: EF-diagram that shows an observer detecting the late Hawking radiation R_L. If, as projection correspondence suggest, her measurement is given by an eigenstate of $N^{out} = b^\dagger b$, postulate 3 is violated. If they are given by an eigenstate of $N^{in} = a^\dagger a$ such as the ground state required by postulate 3, postulate 1 cannot be satisfied. Furthermore, if the outcome depends on the time the observer falls in, we find a contradiction with postulate 2.

which obviously can never return an eigenvalue of $|\Psi_L\rangle$. The only conclusion possible is that the observer will encounter excitations of its a-vacuum. Even worse, these particles will be highly blue-shifted to barely sub-Planckian frequencies $\omega_* \gg R_s^{-1}$, so that the infalling observer will be incinerated by a *firewall* of high-energy modes. This is in stark contrast with third postulate of BHC. The problem lies in the fact that the observer can know what will be the state of the late radiation R_L far from the black hole while also measuring it in the vicinity of the horizon while falling in, thus spoiling the argument of black hole complementarity [1].

Thus, we have arrived at a fresh contradiction, aptly called the *firewall paradox*. In short, it is a contradiction between the entanglement requirements of the postulates. Whereas the purity of the Hawking radiation implies that the late radiation is fully entangled with the early radiation, the absence of drama for an infalling observer demands that every Hawking mode is entangled with its partner mode behind the horizon as in Fig.(15). As we have shown, the claim that either one of the entanglements should be given up depending on the location of the observer can be falsified when the black hole has surpassed its Page time. Then, both entanglement statements should hold simultaneously, which is equivalent to quantum cloning. For example, consider the entanglement entropies of early Hawking modes A, an outgoing Hawking mode B and its partner mode C on the inside of the black hole. The von Neumann entropies of these three systems should satisfy the strong subadditivity constraint

$$S_{AB} + S_{BC} \geq S_B + S_{ABC}. \tag{3.1.10}$$

If the formation and evaporation of the black hole is a unitary process as required by postulate 1, the von Neumann entropy of the Hawking radiation will have a plot as in Figure (27) [28]. Initially, the emitted

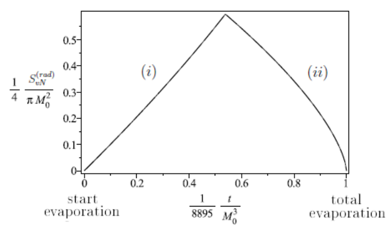


Figure 27: Plot of the von Neumann entanglement entropy of all Hawking radiation emitted up to a moment t in Schwarzschild time reparametrised to measure the evaporation of the black hole, as a fraction of the initial black hole entropy $S_{BH}(t = 0)$.

modes will be entangled with the black hole, so that the system of all Hawking radiation emitted so far can only be described by a reduced density matrix. Correspondingly, the von Neumann entropy associated with the system will increase (region (i)). The von Neumann entropy will peak when the black hole entropy $S_{BH}(t)$ equals the semi-classical entropy $S^{(rad)}(t)$ in the radiation, which occurs after the black hole has lost about 40.25% of its original entropy $S_{BH}(t = 0)$ at about 53.81% of the evaporation time [28]. An 'old' black hole, however, will be in a further stage of evaporation, so that the entanglement entropy of the system will decrease with every extra particle emitted (region (ii)) and so $S_A > S_{AB}$. Furthermore, it is well-known that entanglement can cause the von Neumann entropy of a joint system to be lower than the sum of that of its components. The systems A and BC must meet the subadditivity condition on their respective entropies,

$$S_A + S_{(BC)} \geq S_{A(BC)} \geq |S_A - S_{(BC)}|. \tag{3.1.11}$$

By postulate 3, infalling observers should not experience anything extraordinary while passing the horizon so that the modes B and C are fully entangled and $S_{BC} = 0$. As a result, Eq.(3.1.11) gives us that $S_{ABC} = S_A$. Strong subbadditivity (3.1.10) then implies that $S_A \geq S_B + S_A$, which must evidently fail because mode B individually is described by a thermal density matrix [1].

The Hawking radiation emitted by a Schwarzschild black hole consists mainly of low angular momentum modes [27]. When evolved back, these modes face a gravitational potential surrounding the black hole region as in Fig.(7), which either reflects or transmits them. This is a direct consequence of the fact that a Schwarzschild black hole has a temperature $T_H \sim 1/M$ and a radius $R_s \sim M$, so that $T_H R_s \sim 1$. High angular momentum modes with an energy of the order T_H are effectively imprisoned by the gravitational potential (1.4.14) with escape being very improbable. Therefore, it might be expected that the preceding firewall argument only holds for low angular momentum modes, which form but a small fraction of all partial waves and postulate 3 might survive in a weakened form, where the probability of an infalling observer to observe quanta is suppressed by a power law. This would only require moderately small modifications of ordinary local quantum field theory of this observer. However, using a method known as *black hole mining* we can prove that our conclusions will be valid for high angular momentum modes as well [1].

The aim of the mining process is to lower a particle detector attached to a long string through the potential barrier to absorb the higher partial waves and then lift it back up [46]. Because the high angular momentum modes generally lie close to the horizon, we will make a detector drop from a static point at radius $r_0 \gg R_s$ far from the black hole to within a proper distance $L \ll R_s$ of the stretched horizon, where the locally measured temperature is $T_{loc} \sim 1/L$. It will then probe modes of angular momentum of the order $R_s/L \gg 1$. We will take the detector to have initial mass $m_{det} \sim \epsilon^{-1}L^{-1} \gg T_{loc}$ for some small constant ϵ, so that its stability will not be substantially effected by the heat bath of particles. Furthermore, we may also specify that the size of the detector is of the order ϵL, so that its small absorption cross-section will make it likely that it will first detect particles at target height L. The detector is to be left at this height long enough for it to completely absorb a Hawking quantum, which will effect a marginal mass increase of $T_{loc} \sim 1/L \sim \epsilon m_{det}$ before it is reeled back up to height r_0. Raising and lowering the detector only very slowly will ensure the mining process is adiabatic and no entropy is generated, so that the net amount of energy extracted from the black hole is a positive quantity that has been shown to be of order T_H [46].

Hence, we find that it is possible to mine the high angular momentum modes of a black hole, resulting in a lower mass for the black hole and an increase in energy for the matter far from the black hole. By postulate 2 the detector with the absorbed mode can be evolved to future null infinity, where there is a well defined out-state. At future null infinity these modes, as well as the mining equipment itself, must be seen a subsystem of of the final state. According to postulate 1 the out-state at null infinity is in a pure state, so that we find that late-time high partial modes must be entangled with earlier radiation. Now, it is only natural to consider the detector as part of the Hawking radiation, instead of arbitrarily adding some artificial mass to the spacetime to make up its mass. In this case, the sole act of measuring the modes will not have changed the overall state. We find that regardless of whether the mining actually takes place, the high partial modes emitted by an old black hole are entangled with the early time radiation. We find that postulate 3 is violated for modes of all kinds of angular momenta, suggesting that an observer diving in after the Page time encounters a firewall of Planck density, Planck energy radiation and will thus surely be burned up, violating black hole complementarity [1,6].

So far, the argument for the occurrence of a firewall that will be encountered by an observer falling into a black hole has relied crucially on the entanglement of the Hawking radiation it emits. In order to create a sharply defined paradox, we had to restrict to old black holes that have passed their Page time.

However, whether or not there are firewalls should be an intrinsic property of a black hole, independent of the degree to which it is entangled with some external system such as the Hawking radiation. The *scrambling time* of a black hole measures how fast it can thermalise the matter by which it is formed and randomise the bits that encode the corresponding information. After the scrambling time, almost any small subsystem of a black hole is in a completely inseparable state and the degrees of freedom encountered by an infalling observer are typical for the remainder of the lifespan of the black hole. It has been suggested that black holes are so-called fast scramblers [32], putting their scrambling time at $R_s \ln(R_s/\ell_p)$. Then, the time after which we might predict to see firewalls or expect new physics outside the black hole is remarkably smaller than the Page time [1].

In conclusion, we have found that if black hole complementarity breaks down, it breaks down massively. If there is indeed a firewall, it is everywhere surrounding the black hole and is likely to materialise quickly after the black hole has formed, necessitating a major rethinking of the common perception of black hole evaporation process. By effectively reinstating the information paradox to its former glory, it has sharply divided the physics community between those who prefer to value postulate 3 and thus rather give up unitarity and argue against the AdS/CFT correspondence and those with a more quantum physical background, who are adamant that unitary evolution is a key aspect of modern day physics. Various proposals have sought to resolve the paradoxical firewall appearance, but so far none have been able to do so conclusively. Suggestions include:

- *EP=EPR*
A wormhole or Einstein-Rosen (ER) bridge between two black holes can be seen as the geometrical manifestation of the Einstein-Poldolski-Rosen (EPR) entanglement of their microstates. This was called the *EP=EPR relation*. It is then tempting to speculate that there is a more general mechanism at hand, so that other systems might also be susceptible to such a description. Suppose a black hole is connected to the Hawking radiation by a wormhole with many different outlets, so that all the quanta are entangled with the black hole. It is then possible that the actual measurement of an early Hawking mode causes the creation of a high energy particle that travels through the wormhole to the near horizon region, where the observer, who dove in after the measurement, will encounter it. Such radically non-local action effectively makes the occurrence of the firewall dependent on the early measurements of an observer at infinity, which are limited by practical bounds [23]. The quantum computation necessary to determine the degree of entanglement of the early radiation requires a time span that is larger than the black hole lifetime to carry out [16,42]. However, it is recognised that the states of two entangled black holes is very special. It was shown that small perturbations in the state of such black holes have large consequences, what has been beautifully named the *black hole butterfly effect* [33]. Indeed, in a recent article Marolf and Polchinski, two of the original creators of the paradox, have argued that generic entangled states cannot generally be interpreted as a wormhole that connects them [24].

- *Strong Complementarity*
It has been proposed by various sources [5,16] that a 'strong' version of black hole complementarity might still hold. Like ordinary BHC, strong complementarity proposes that a distant observer would predict a firewall and the infalling observer sees nothing extraordinary, but abandons the requirement that there is a global effective field theory outside the horizon. Instead, it assumes an even more operationalistic view by demanding only that no single observer will be able to detect violations. This would imply that the the infalling observer is unable to compare a mode of the late radiation encountered at the near-horizon region with its state at infinity, since they are no longer related by an evolution governed by a global effective field theory. However, Marolf and Polchinski reason that one can imagine a continuous string of observers diving into the black hole at successive points in time, each with a valid local effective field theory. However, two observers in causal contact are able to communicate, so that the relaxed form of the second postulate is still violated [24].

- *State-dependence*

If the AdS/CFT correspondence were more complete, we could use the dictionary between bulk and boundary operators to translate a field theory operator $\hat{T}_{\mu\nu}$ on the boundary to find its dual energy momentum theory $T_{\mu\nu}$ of the interior of the black hole. By calculating its expectation value in the infalling vacuum state, it is then possible to investigate the existence of a firewall for an infalling observer. However, so far the gauge/gravity duality has not yet been sufficiently developed to enable such a method. Moreover, it seems likely that a perfect conversion will not be possible [1]. The bulk energy-momentum tensor $T_{\mu\nu}$ involves the Hawking partner modes c, which have negative energy with respect to an asymptotic observer. There should then be a bulk field theory operator with the properties

$$[c, c^\dagger] = 1, \qquad [H, c^\dagger] = -\omega c^\dagger, \tag{3.1.12}$$

which, however, cannot exist [24]. Now, typical states in a field theory will have thermal spectra for the partner modes b and c. Any particular example of these states can then be elected to construct a representation of the commutators (3.1.12) [29,47]. Nevertheless, this does require a stark modification of quantum mechanics, where interior observables would become non-linear maps $\mathcal{H} \times \mathcal{H} \mapsto \mathcal{H}$, depending on the base state chosen. In order to have an interior interpretation of a given dual state, one must first specify which of the modes b was used to define the Fock space ladder operators. This has inspired an adaption of the famous phrase by Einstein: "God not only plays dice with the world, She also plays pin the tail on the quantum donkey" [47]. It is not yet clear whether it would be possible to construct a valid quantum theory based on this proposal. Things are complicated by the fact that every instance of interaction of the black hole with external spacetime changes its microstate and thus invalidates the original representation associated with the base chosen. Hence, the base state must be a dynamical quantity evolving along with the system itself. As of yet, there is no theory available that is capable of describing such twofold evolution [24].

3.2 Should the Calculation of the Hawking Flux Still Hold?

Another question that has been raised is whether the usual Hawking calculation is still valid in the light of the firewall paradox. Since the occurrence of Hawking radiation is due essentially to a Bogoliubov transformation like (1.3.9) of the in vacuum, the fact that an infalling observer sees a firewall instead of the in vacuum seems to indicate an internal inconsistency. If this were indeed found to be the case, it would signal a fundamental failure of black hole complementarity that is beyond repair. Therefore, we will examine this area of research in somewhat more detail.

In summary, the argument is as follows. By the technique of thermo-field doubling, a thermal state on a Hilbert space $\mathcal{H}$ described by a density matrix of the form

$$\rho = \frac{1}{Z} \sum_E e^{-\beta E} |E\rangle\langle E|, \tag{3.2.1}$$

where Z is the partition function of the system, can be artificially duplicated in order to obtain a pure state

$$|\Psi\rangle = \frac{1}{\sqrt{Z}} \sum_E e^{-\beta E/2} |E\rangle \otimes |\tilde{E}\rangle, \tag{3.2.2}$$

called the the *thermo-field vacuum*, that lives on two copies of the original Hilbert space $\mathcal{H} \times \tilde{\mathcal{H}}$. Of course, tracing over the second copy returns the thermal density matrix (3.2.1). Any observable of the original system can then be calculated more easily by relation (1.4.8) [43],

$$\mathrm{Tr}\,(\rho\, O) = \langle \Psi | O | \Psi \rangle. \tag{3.2.3}$$

We can think of the doubled system as a gas of free thermal bosons. Such a doubled system will have two sets of ladder operators, with the annihilation operators destroying the thermo-field vacuum. The creation operators, however, will have the meaning of creating a boson in the one copy or annihilating one in the other copy, as in Figure (28) [7].

Originally, the second copy of the Hilbert space was interpreted solely as a mathematical construct without physical meaning, useful only to simplify calculations [43]. However, it was found that observers

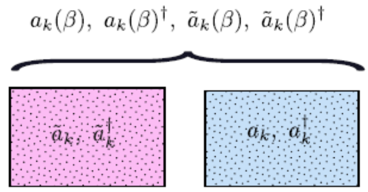


Figure 28: The boson gas can be seen as a system of two boxes. The left box can be interpreted as the thermo-field doubled copy of a thermal state in the right box. The two sets of creation operators associated with the thermo-field vacuum $|\Psi\rangle$ then have the meaning of creating bosons in their respective boxes or annihilating them in the other. Both a and $\tilde{a}$ annihilate the thermo-field vacuum state $|\Psi\rangle$.

on the outside of an eternal Schwarzschild black hole can interpret the interior as a thermo-field doubled copy of the exterior, and vice versa [18]. This idea was strengthened further by the discovery that the same will hold for accelerating observers in Minkowski space, who can describe the state in other Rindler wedge as the themo-field duplicate of their own thermal state. With the advent of the AdS/CFT correspondence, these findings could be put in a clearer light. Maldacena himself proposed that the *Hartle-Hawking vacuum* state associated with the modes of the the thermal radiation that maintains the equilibrium of the eternal AdS black hole has a natural dual in the thermo-field vacuum state of the CFTs on the two asymptotic boundaries of the extended AdS-Schwarzschild geometry [21]. As we will see, this is a very accommodating result: by definition, eternal black holes are in contact with a thermal heat bath of radiation at exactly the right temperature to maintain equilibrium. As a consequence, they form definite systems with well-formulated energy eigenstates. The thermo-field vacuum state is a product state of two entangled systems dual to the interior and exterior of the black hole, which will each be described by a thermal density matrix. On the other side of the correspondence, the outgoing b-modes of the radiation will have a thermal density matrix ρ_b, at least as seen from the Hartle-Hawking vacuum [24]. Then, according to the *eigenstate thermalisation hypothesis (ETH)* the same conclusion will hold for any generic energy eigenstate, such as the a-vacuum of an infalling observer. There can truly be said to be a flux of Hawking quanta at temperature T_H emanating from the black hole in the form of the b modes [24].

The eigenstate thermalisation hypothesis is a statement concerning the energy eigenstates of *chaotic* quantum system. To come to a better understanding of its meaning, consider first the classical situation of a box filled with a gas of N particles modeled by hard spheres. The Hamiltonian corresponding to such a system is simply

$$H = \frac{1}{2m}\sum_{i=1}^{N}\mathbf{p}_i^2 = \frac{1}{2m}\mathbf{P}^2, \tag{3.2.4}$$

where m is a mass that is taken to be uniform over all the particles and $\mathbf{P}$ is a vector composed of N momentum 3-vectors $\mathbf{p} = (p_x, p_y, p_z)$. The Hamiltonian must take on a value equal to the total energy U given by the ideal-gas relation $U = \frac{3}{2}NT$, so that there is a $\mathbf{P}$-sphere of possible solutions. By Sinai's theorem a box of hard spheres will be a chaotic system, implying that each location of the $\mathbf{P}$-sphere is as likely as any other to be a solution to the ideal-gas relation [35]. The probability that the solution is within an angular section between θ and $\theta + d\theta$ with respect to some axis is given by

$$\begin{aligned} f(\theta)d\theta &\sim (\sin\theta)^{(3N-2)}d\theta \\ &\sim (\sin\theta)^{(3N-3)}d\cos\theta \\ &\sim (1-\cos^2\theta)^{(3N-3)}d\cos\theta. \end{aligned} \tag{3.2.5}$$

Suppose we know define the momentum of the first particle along an arbitrary z-axis by the kinetic enegy relationship, so that $p_{1z} = (2mU)^{(1/2)}\cos\theta$. Then, we find

$$\begin{aligned} f(p_{1z})dp_{1z} &\sim (1-\frac{p_{1z}^2}{2mU})^{(3N-3)/2}dp_{1z} \\ &\sim \exp\left(\frac{-p_{1z}^2}{2mT}\right)dp_{1z}, \end{aligned} \tag{3.2.6}$$

where we have taken the large-N limit to arrive at the exponential relation of the second line. Hence, we have found the appropriate thermal Maxwell-Boltzmann distribution for the single component p_{1z}.

Consider then a second component p_{1y} when p_{1z} is fixed. Its probability distribution is then that of p_{1z} given by Eq.(3.2.6), with the exception of the replacements

$$\begin{aligned} 3N &\to 3N-1 \\ 2mU &\to 2mU - p_{1z}^2 \simeq 2mU \quad \text{as } N \to \infty, \end{aligned} \tag{3.2.7}$$

so that we see that we obtain the same thermal distribution for any n individual components of the $3N$-vector $\mathbf{P}$, as long as $n \ll N$[35].

Now, consider the same system from a quantum perspective. The evolution of its N constituent particles is governed by a Schrödinger equation that can be solved in the energy basis, where it has the form

$$H|E_\alpha\rangle = U_\alpha |E_\alpha\rangle. \tag{3.2.8}$$

Here, the Hamiltonian H is a quantised version of that given in Eq.(3.2.4). Furthermore, we specify boundary conditions on the energy eigenfunctions $\psi_\alpha(\mathbf{X})$ such that they vanish at the walls of the box and when the spheres come into contact. The momentum space wavefunction can then be expanded in terms of the energy eigenstates as

$$\tilde{\psi}(\mathbf{P}, t) = \sum_\alpha C_\alpha \exp{(-iU_\alpha t)} \tilde{\psi}_\alpha(\mathbf{P}). \tag{3.2.9}$$

By means of the wavefunction we can calculate the probability distribution of momenta for the first particle at a particular time t,

$$\begin{aligned} f(\mathbf{p}_1, t) &= \int d^3p_2 \dots d^3p_N |\tilde{\psi}(\mathbf{P}, t)|^2 \\ &= \sum_{\alpha,\beta} C_\alpha^* C_\beta e^{i(U_\alpha - U_\beta)t} \int d^3p_2 \dots d^3p_N \tilde{\psi}_\alpha^*(\mathbf{P}) \tilde{\psi}_\beta(\mathbf{P}). \end{aligned} \tag{3.2.10}$$

Here, we can define a quantity

$$\Phi_{\alpha\beta}(\mathbf{p}_1) \equiv \int d^3p_2 \dots d^3p_N \tilde{\psi}_\alpha^*(\mathbf{P}) \tilde{\psi}_\beta(\mathbf{P}) \tag{3.2.11}$$

that loosely characterises how the momentum of the first particle is distributed over the different energy eigenfunctions. It must satisfy a normalisation condition

$$\int d^3p_1 \Phi_{\alpha\beta}(\mathbf{p}_1) = \delta_{\alpha\beta}. \tag{3.2.12}$$

On physical grounds, we expect $f(\mathbf{p}_1, t)$ to approximate a thermal distribution in the classical large-N limit after system has settled. However, how this might come to be is unclear in this form. To this end, consider the long-time average of $f(\mathbf{p}, t)$ that is given by

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, f(\mathbf{p}, t) = \sum_\alpha |C_\alpha|^2 \Phi_{\alpha\alpha}(\mathbf{p}_1). \tag{3.2.13}$$

Clearly, the initial perturbations caused by the anisotropic starting conditions of the system have become completely negligible by taking the long-time average, so that this will surely be a thermal distribution. It is obvious that such a distribution can only be realised if each $\Phi_{\alpha\alpha}(\mathbf{p}_1)$ is thermal by itself. Furthermore, if the overlaps of separate energy eigenfunctions are small, i.e. small $\Phi_{\alpha\beta}(\mathbf{p}_1)$ for $\alpha \neq \beta$, then the $\alpha \neq \beta$ terms in the distribution (3.2.10) will usually make a negligible contribution and $f(\mathbf{p}_1, t)$ will be of thermal nature at most times t, not just on average[35].

These two assumptions, that together form the eigenstate thermalisation hypothesis, are supported by a result in quantum statistical mechanics known as *Berry's conjecture*. Basically, it speculates that the energy eigenfunctions of a bounded, isolated quantum system with a classically chaotic counterpart behave as if they are gaussian random variables, in the meaning that it satisfies some sort of Wick's theorem:

$$\lim_{\alpha\to\infty} \int d\mathbf{X}\, \psi_\alpha(\mathbf{X} + \mathbf{X}_1) \dots \psi_\alpha(\mathbf{X} + \mathbf{X}_n) = \sum_{pairs} J(\mathbf{X}_{i_1} - \mathbf{X}_{i_2}) \dots J(\mathbf{X}_{i_{n-1}} - \mathbf{X}_{i_n}). \tag{3.2.14}$$

Here, we utilise a normalised integration measure so that $\int d\mathbf{X} = 1$ and the sum on the other hand of the equality is over all possible pairings over the $\mathbf{X}_i$. The $J(\mathbf{X})$ are correlation functions given by

$$J(\mathbf{X}) \sim \int d\mathbf{P} \; \exp\left(i\mathbf{P}\cdot\mathbf{X}\right) \delta(H(\mathbf{P},\mathbf{X}) - U_\alpha), \tag{3.2.15}$$

with $J(\mathbf{0}) = 1$ [35]. Although Berry's conjecture remains unproven, there is ample numerical evidence to argue in its favour [36]. It is this argument that is used to justify the extrapolation of the thermal nature of the radiation surrounding the eternal black hole as seen from the Harte-Hawking state to the vacuum of an observer falling into an eternal black hole and reason for the existence of Hawking radiation, even considering this observer will be incinerated by a firewall [24].

Although Marolf and Polchinski seem fully convinced by this line of reasoning supporting the firewall paradox [24], there are several things that have remained unclear to us and might point at a fallacy in the argument. These will be outlined here for the benefit of the reader.

- First of all, if the argument does stand up to scrutiny, it will only argue for the case of the eternal AdS black hole. This does not automatically imply that the same thing will hold for other classes of black holes.

- Secondly, the eigenstate thermalisation hypothesis was first derived for the classical system of a box filled with gas and subsequently extended to include the quantum treatment of this system. It is believed that Berry's conjecture will hold only in those quantum systems that will similarly have a chaotic classical counterpart. It remains to be seen whether the eternal AdS black hole is indeed such a system. If so, how should we identify the chaotic degrees of freedom? The classical treatment of black holes does not envision the appearance of Hawking radiation, so that the source must either lay with the black hole or with the thermal heat bath in which it is immersed.

- Furthermore, it is stated in in the original ETH paper that "another important unsettled issue is how high the energy needs to be before Eq.(3.2.14) [of Berry's conjecture] is sufficiently accurate" [35]. This is insignificant for the case of the gas-filled box, but all the more relevant for the black hole situation. By the inverse mass relation of the Hawking temperature $T_H \sim 1/M$, the temperature of a black hole decreases as it grows in size. Hence, it might be concluded that Berry's conjecture will not be accurate for sufficiently large black holes. This would invalidate the ETH, and with that the argument in favour of the occurrence of Hawking radiation in the light of the firewall paradox, in these specific cases. However, the consensus in the physics community is that of uniformity: if a class of black holes is found to radiate, it will do so regardless of how it was formed.

4 Conclusion

It is a universally accepted fact that black holes will emit Hawking quanta at a temperature inversely proportional to their mass, which causes their eventual evaporation. This inevitably led to the information paradox, which appeared to have been solved by the black hole complementarity proposal. Although BHC may seem like a reasonable suggestion inspired on the idea to preserve a maximal of known physics, at least for a black hole past its Page time its three postulates cannot all hold: an infalling observer will encounter a firewall of radiation. There is something radically different going on. In trying to resolve the firewall paradox, the physics community is resorting to remedies that to me at least seem more and more far-fetched, such as the EP=EPR wormhole proposal. It seems unlikely that the paradox will be solved any time soon, and its resolution will have some profound consequences. As Marolf and Polchinski argue, the AdS/CFT correspondence has proven of great significance but has not conclusively put the lid on the question raised originally by the information paradox. "While gauge/gravity duality is a powerful tool, there is a gap in the current understanding of quantum gravity, on that must be filled in order to move on to quantum cosmology" [24].

What we know for certain are the measurements an asymptotic observer can make, since at infinity the effects of the black hole are negligible. I think that we should at least entertain the idea that what we should give up are in fact the no-hair theorems. As a result from quantum gravity, the black hole horizon must be special. I would like to speculate that there really is 'nothing' behind the horizon. I propose that when matter falls into a black hole, instead of crossing the horizon its characteristics are transferred to the horizon. I think we should look at a black hole more as any other kind of physical object, albeit with a negative heat capacity. Then, when an observer tries to cross the horizon into the black hole it is only natural to encounter these degrees of freedom in the form of a firewall. Notice that this violates postulate 3 of black hole complementarity but not the equivalence principle that underlies it, since now the horizon is a special place. In this way, it would be possible to circumvent the disastrous conclusions of the firewall paradox while maintaining what are first principles of our current understanding of physics: the equivalence principle of general relativity, semi-classical field behaviour outside event horizons and unitarity of all processes, including the formation and evaporation of a black hole.

References

[1] A. Almeheiri, D. Marolf, J. Polchinski and J. Sully, "Black Holes: Complementarity or Firewalls?", (2012). arXiv: 1207.3123v4 [hep-th]

[2] J. Bekenstein, "Black Holes and Entropy", Phys. Rev. D **7**, 2333 (1973).

[3] J. Bekenstein, "Entropy bounds and black hole remnants", Phys. Rev. D **49**, 1912 (1994).

[4] R. Bousso, "The Holographic Principle", Rev. Mod. Phys. **74**, 825 (2002). arXiv: 0203101 [hep-th]

[5] R. Bousso, "Complementarity Is Not Enough", (2012). arXiv:1207.5192v2 [hep-th]

[6] R. Bousso, "Firewalls From Double Purity", (2013). arXiv:1308.2665v2 [hep-th]

[7] B. Chowdhury, "Black holes vs. firewalls and thermo-field dynamics", (2013). arXiv:1305.6343v2 [gr-qc]

[8] J. Dieckmann, "Cauchy Surfaces in Globally Hyperbolic Spacetimes", J. Math. Phys. **29**, 578 (1988).

[9] F. Dowker, "Black Holes Lecture Notes QFFF Imperial College", (2013).

[10] A. Fabbri and J. Navarro-Salas, "Modelling Black Hole Evaporation", Imperial College Press (2005).

[11] O. Gannot, "Quasinormal Modes for Schwarschild-AdS Black Holes: Exponential Convergence to the Real Axis", (2013). arXiv: 1212.1907v4 [math.SP]

[12] R. Geroch, "Domain of Dependence", J. Math. Phys. **11**, 437 (1970).

[13] S. Hawking and G. Ellis, "The Large Scale Structure of Space-Time", Cambridge Uni. Press (1973).

[14] S. Hawking, "Black hole explosions?", Nature **248**, 5443 (1974).

[15] S. Hawking, "Particle Creation by Black Holes", Commun. Math. Phys. **43**, 199 (1975).

[16] D. Harlow and P. Hayden, "Quantum Computation vs. Firewalls", (2013). arXiv:1301.4504v4 [hep-th]

[17] G. Holzegel and J. Smulevici, "Stability of Schwarzschild-AdS for the Spherically Symmetric Einstein-Klein-Gordon System", Commun. Math. Phys. **317**, 205 (2013).

[18] W. Israel, "Thermo-field dynamics of black holes", Phys. Lett. A **57**, 107 (1976).

[19] G. 't Hooft, "Dimensional Reduction in Quantum Gravity", Class. Quant. Grav. **11**, 621 (1994).

[20] J. Maldacena, "The Large N Limit of Superconformal Field Theories and Supergravity", Adv. Theor. Math. Phys. **2**, 231 (1998). arXiv: 9711200v3 [hep-th]

[21] J. Maldacena, "Eternal black holes in Anti-de-Sitter", (2001). arXiv: 0106112v6 [hep-th]

[22] J. Maldacena, "The gauge/gravity duality", (2011). arXiv: 1106.6073v1 [hep-th]

[23] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes", (2013). arXiv:1306.0533v2 [hep-th]

[24] D. Marolf and J. Polchinski, "Gauge/Gravity Duality and the Black Hole Interior", (2013). arXiv: 1307.4706v2 [hep-th]

[25] S. Mathur, "What Exactly is the Information Paradox?", (2008). arXiv: 0803.2030v1 [hep-th]

[26] D. McDuff and D. Salamon, "Introduction to Symplectic Topology", Oxford Clarendon Press (1995).

[27] D. Page, "Particle Emission Rates from a Black Hole: Massless Particles from an Uncharged, Non-rotating Hole", Phys. Rev. D **13**, 198 (1976).

[28] D. Page, "Time Dependence of Hawking Radiation Entropy", (2013). arXiv:1301.4995v3 [hep-th]

[29] K. Papadodimas and S. Raju, "An Infalling Observer in AdS/CFT", (2012). arXiv:1211.6767v1 [hep-th]

[30] R. Penrose, "Gravitational Collapse: The Role of General Relativity", Rivista del Nuovo Cimento **1**, 252 (1969).

[31] H. Reall, "Black Holes Lecture Notes Part III Cambridge", (2008).

[32] Y. Sekino and L. Susskind, "Fast Scramblers", JHEP **0810**, 65 (2008). arXiv: 0808.2096 [hep-th]

[33] S. Shenker and D. Stanford, "Black holes and the butterfly effect", (2013). arXiv:1306.0622v1 [hep-th]

[34] M. Soffel, B. Muller and W. Greiner, "Dirac particles in Rindler space, Phys. Rev. D, **22**, 1935 (1980).

[35] M. Srednicki, "Quantum Chaos and Statistical Mechanics", (1994). arXiv: 9406056v1 [cond-mat]

[36] M. Srednicki, "Chaos and Quantum Thermalization", (1994). arXiv: 9403051v2 [cond-mat]

[37] L. Susskind, L. Thorlacius and J. Uglum, "The Stretched Horizon and Black Hole Comoplementarity", (1993). arXiv: 9306069v2 [hep-th]

[38] L. Susskind, "The World as a Hologram", (1994). arXiv: 9409089v2 [hep-th]

[39] L. Susskind and E. Witten, "The Holographic Bound in Anti-de Sitter Space (1998). arXiv: 9805114v1 [hep-th]

[40] L. Susskind and J. Lindesay, "An Introduction to Black Holes, Information and the String Theory Revolution - The Holographic Universe", World Scientific (2005).

[41] L. Susskind, "The Black Hole War: My Battle with Stephen Hawking to Make the World Safe for Quantum Mechanics", Back Bay Books (2008).

[42] "Singularties, Firewalls, and Complementarity", (2012). arXiv:1208.3445v1 [hep-th]

[43] Y. Takahashi and H. Umezawa, "Thermo-field Dynamics", Intern. Journal of Mod. Phys. B **10**, 1755 (1996).

[44] K. Thorne, R. Price and D. MacDonald, "Black Holes: The Membrane Paradigm", Yale Uni. Press (1986).

[45] P. Townsend, "Black Holes Lecture Notes Part III Cambridge", (1997). arXiv: 9707012v1 [gr-qc]

[46] W. Unruh and R. Wald, "Acceleration Radiation and Generalized Second Law of Thermodynamics", Phys. Rev. D **25**, 942 (1982).

[47] E. Verlinde and H. Verlinde, "Black Hole Entanglement and Quantum Error Correction", (2012). arXiv:1211.6913v1 [hep-th]

[48] R. Wald, "On Particle Creation by Black Holes", Commun. Math. Phys. **45**, 9 (1975).

[49] R. Wald, "General Relativity", Uni. Chicago Press (1984).

[50] R. Wald, "Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics", Uni. Chicago Press (1994).

[51] E. Witten, "Anti de Sitter Space and Holography", (1998). arXiv: 9802150v2 [hep-th]

[52] B. Zwiebach, "A First Course in String Theory", Cambridge Uni. Press (2004).