DISSERTAÇÃO DE MESTRADO IFT–D.001/2025

Black hole information loss issue and the global hyperbolicity of spacetimes

Juan Vitor Oliveira Pêgas

Orientador

George Emanuel Avraam Matsas

Fevereiro de 2025

P376b

Pêgas, Juan Vitor Oliveira

Black hole information loss issue and the global hyperbolicity of spacetimes / Juan Vitor Oliveira Pêgas. -- São Paulo, 2025

74 f: il. color.

Dissertação (mestrado) – Universidade Estadual Paulista (Unesp), Instituto de Física Teórica (IFT), São Paulo

Orientador: George Emanuel Avraam Matsas

1. Teoria quântica de campos. 2. Buracos negros (Astronomia). 3. Relatividade geral (Física) I. Título

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Física Teórica (IFT), São Paulo. Dados fornecidos pelo autor(a).

BLACK HOLE INFORMATION LOSS ISSUE AND THE GLOBAL HYPERBOLICITY OF SPACETIMES

Dissertação de Mestrado apresentada ao Instituto de Física Teórica do Câmpus de São Paulo, da Universidade Estadual Paulista "Júlio de Mesquita Filho", como parte dos requisitos para obtenção do título Mestre em Ciências, área: Física.

Comissão Examinadora:

Prof. Dr. GEORGE EMANUEL AVRAAM MATSAS (Orientador)
Instituto de Física Teórica/UNESP

Prof. Dr. GIANLUCA GREGORI
University of Oxford

Prof. Dr. JORMA LOUKO
University of Nottingham

Conceito: Aprovado

São Paulo, 25 de fevereiro de 2025.

Acknowledgments

To my advisor, Professor George Matsas, for all of the enthusiastic and enlightening discussions about how nature works. Thank you for taught me that “Semmi sem könnyű!”.

To my partner, Ana Vitória, for all of the countless hours discussing the curiosities and the running of the universe, from HTTPS to light cones. Thank you for always showing me that there is life outside physics.

To my parents, Elaine and Marcos, for all of the unrestricted love and support along the years.

To my friend and colleague, Gabriel Aguiar, for all of the lunches discussions about life, universe and the stock market.

To FAPESP for financial support under grant 2022/14028-3.

Resumo

Na ausência de uma teoria quântica de gravitação, a teoria quântica de campos em espaços-tempos curvos fornece o arcabouço mais sólido para estudarmos a interação entre gravidade e mecânica quântica. Uma das consequências mais marcantes dessa teoria é o efeito Hawking. Neste contexto, buracos negros iriam evaporar completamente devido à emissão de radiação térmica. Como consequência, um estado puro de um campo quântico preparado no passado assintótico que colapsasse em um buraco negro evoluiria para um estado misto no futuro assintótico, violando unitariedade. Este é o chamado "paradoxo da perda de informação em buracos negros". Nesta dissertação, argumentamos que, longe de ser paradoxal, a perda de informação é uma previsão da teoria quântica de campos em espaços-tempos curvos e não uma violação da mesma. Além disso, abordamos o problema de como conciliar a evolução unitária de um campo quântico com o efeito Hawking sem depender de um modelo específico de gravitação quântica.

Palavras Chaves: teoria quântica de campos em espaços-tempos curvos; perda de informação; buracos negros.

Áreas do conhecimento: física; teoria quântica de campos em espaços-tempos curvos; relatividade geral.

Abstract

In the absence of a quantum theory of gravity, quantum field theory in curved spacetimes provides the most solid framework to study the interplay between gravity and quantum mechanics. One of the most striking consequences of this theory is the Hawking effect. In this scenario, the fade of black holes is its complete evaporation due to the emission of thermal radiation. As a consequence, a pure state of a quantum field prepared in the asymptotic past that collapses in a black hole will evolve into a mixed state in the asymptotic future, violating unitarity. This is the so-called "black hole information loss paradox". In this dissertation, we argue that, far from paradoxical, the information loss is a prediction of quantum field theory in curved spacetimes and not a violation. In addition, we address the problem of how to conciliate the unitary evolution of a quantum field and the Hawking effect without relying on a specific quantum gravity model.

Key-words: Quantum field theory in curved spacetimes; Information loss; Black holes.

Field of knowledge: Physics; Quantum field theory in curved spacetimes; General relativity.

Contents

List of Figures

Chapter 1

Introduction

It is widely recognized that the main frameworks in the study of the fundamental interactions of nature are quantum field theory (QFT) and general relativity (GR). Over the past century, both theories have provided astonishing agreement with experimental tests in their regime of validity. However, the "big picture" is incomplete, lacking an entirely consistent framework encompassing these two successful physics branches in situations where the quantum nature of both fields and spacetime must be considered. In this scenario, one can argue that the most natural first step towards a quantum description of gravity is considering quantum effects in contexts where gravity cannot be neglected using the so-called quantum field theory in curved spacetimes (QFTCS). This approach considers the propagation of quantum fields in a classical curved spacetime background. Due to its classical approach to the metric, QFTCS can not be a fundamental theory of nature, although it is expected to provide an accurate description of situations where one can neglect the effects of quantum gravity. Since its development in the 60s, QFTCS has provided not only a description of quantum phenomena occurring in the early universe and near black holes, but it has also given a deeper understanding of QFT in Minkowski spacetime, e.g., the Unruh effect [1].

The most known effect predicted by QFTCS is undoubtedly Hawking radiation, which states that black holes could evaporate via the emission of thermal radiation [2]. Although using modern technology, it is impossible to detect this radiation, which, in the most favorable scenario, would have a temperature billions of times smaller than the cosmic microwave background, it was recently detected an analog effect using Bose-Einstein condensate [3]. Other analogs of the Unruh and Hawking effect have been reviewed in Ref. [4] (see also references [58-62]).

Despite its theoretical success, the Hawking effect has been raising doubts in the scientific community through the so-called "black hole information loss paradox", which can be stated as follows. Consider the evolution of a quantum state ϕ in a spacetime describing a collapsing star. Under certain conditions, the collapse will create a black hole, eventually evaporating due to the Hawking radiation. Assuming that the black hole evaporates completely, the final state of the field is characterized by a thermal distribution, which is a mixed state. Therefore, we have an evolution from a pure state to a mixed one, which is not allowed for closed quantum systems. As a consequence, the scientific community has been challenging Hawking's evaporation process due to this apparent violation of quantum mechanics, leading to the term "paradox". Past attempts to "resolve" this "paradox" by disregarding the role played by the black hole singularity have led to the introduction of all sorts of undesired features, such as violation of causality and of other basic principles in low curvature regimes (see Ref. [5] for a review).

Far from paradoxical, the loss of information is a natural consequence of the corresponding semiclassical spacetime not being globally hyperbolic. Enlightening this scenario and showing how one can conciliate the Hawking process with unitarity are the main goals of the present thesis. It should be stressed the possibility (but certainly not the obligation) of quantum gravity resolving the central spacelike singularity in the interior of black holes, recovering the unitary evolution throughout spacetime.

In this work, we proceed as follows. Chapter 2 is a review of basic concepts of general relativity relevant to the formulation of the Hawking process, focusing on the symmetries present in stationary spacetimes, the causal structure of general spacetimes, and the main features of black holes. Chapter 3 is concerned with the derivation of the Hawking process. We offer a brief review of the canonical quantization of scalar fields in curved spacetimes. The derivation of Hawking temperature and discussion of the resulting flux of radiation are presented. Finally, we discuss the formulation of the information loss "paradox", arguing that it is a prediction of the semiclassical theory. Chapter 4 is related to the conciliation of black hole evaporation via thermal radiation with the unitary evolution of quantum fields without relying on a specific quantum gravity model. We present our conclusions in Chapter 5.

We assume metric signature $(-+++)$ and natural units $\hbar = G = c = k_B = 1$, unless stated otherwise. Latin indexes (a, b,...) are used to denote tensorial quantities in arbitrary coordinate systems, and Greek indexes $(\mu, \nu, ...)$ are used to denote their value in a particular coordinate system.

Chapter 2

Basic concepts of general relativity

The fundamental ideas on which classical general relativity is based are that (i) the spacetime is a manifold M endowed with a Lorentzian metric (a rank-two tensor) g_{ab}[1] and (ii) the metric and matter fields are dynamical, evolving locally by the Einstein's field equation

$$G_{ab} = R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi T_{ab}. \tag{2.1}$$

This is a highly non-linear set of partial differential equations relating the metric degrees of freedom with the ones of the matter fields. There is a very non-trivial difference between Einstein's equation and other differential equations in physics. For instance, consider Maxwell's equation in the Lorenz gauge

$$\Box A^a = -4\pi j^a, \quad \nabla_a A^a = 0,$$

where we can fix the 4-current j^a and solve the above equation for A^a. However, in general, the stress-energy tensor T_{ab} depends on g_{ab}, so we can not, in analogy with the electromagnetic case, fix *ad hoc* the matter field and solve for the metric in Eq. (2.1). In this way, we have to solve both sides of Einstein's equation in a self-consistent way.

The above discussion makes it clear that it can be very challenging to solve Einstein's equation and extract the physical information from the solutions of Eq. (2.1). In this way, the study of symmetries and the causal structure of the spacetimes often help to unveil the physical content of the spacetime (M, g_{ab}). In this chapter, we review some concepts concerning symmetries and the causal structure of spacetimes, focusing on Minkowski and Schwarszchild solutions. The former helps to illustrate the rather mathematical definitions used in those contexts, while the latter is archetypal in studying black hole dynamics. We shall also review the mathematical definition and the main properties of black holes in

[1]Some additional, physically motivated, causal features are demanded, see Sec. 2.2.

classical general relativity, highlighting the main results of their structure.

2.1 Symmetries

The concept of symmetry is usually crucial in generating solutions of Eq. (2.1). Intuitively, a symmetry is a "motion" in spacetime under which all the physics is invariant. In this way, the protocol to search for symmetries is simply "comparing" the physical quantities defined in events before and after this "motion". If they are the same, we say that the theory has a symmetry.

Let us be more precise about what we mean physically by "motion" and "comparison". Consider a physical theory described by a manifold M and tensor fields $T^{(i)}$ defined on M. Let v^a be a vector field on M. Suppose we displace each point of M infinitesimally by an amount determined by the value of v^a at that point. Thus, v^a defines the "motion" of spacetime along the direction defined by the vector field. Now, we can have two sets of tensor fields: the original $T^{(i)}$ and the one that results from subjecting $T^{(i)}$ to the motion defined by v^a. We can, then, compare these two sets, by subtracting one from the other and dividing by the amount of sliding along v^a, and conclude if they are the same. In case affirmative, we say that v^a defines a *symmetry*.

Now that we have an intuitive picture of what symmetry is, let us make this notion mathematically precise. We follow here mainly the approach of Ref. [6]. Let M and N be manifolds (not necessarily of the same dimension). Let $\phi : M \to N$ be a C^∞ map and let $f : N \to \mathbb{R}$ be a function. We define the *pull back* of f to be the composed function $f \circ \phi : M \to \mathbb{R}$, see Fig. 2.1. Similarly, let $p \in M$. For $v_p \in V_p$ we define the *push forward* $\phi^* v \in V_{\phi(p)}$ by $(\phi^* v)(f) = v(f \circ \phi)$ for all smooth $f : N \to \mathbb{R}$, where V_p is the tangent space of p. In a similar manner, we can define the pull back $\phi_*(w_{\phi(p)})$ of a covector $w_{\phi(p)} \in V^*_{\phi(p)}$ by $\phi_*(w_{\phi(p)})v = w_{\phi(p)}(\phi^*(v))$. The same construction can be done for tensors of arbitrary ranks, where now we demand ϕ to be a diffeomorphism, i.e., bijective C^∞ map with inverse C^∞. By noticing that $(\phi^{-1})^*$ goes from $V_{\phi(p)}$ to V_p, we define the tensor $(\phi^* T)^{b1\ldots b_k}{}_{a_1\ldots a_l}$ at $\phi(p)$ by

$$(\phi^* T)^{b1\ldots b_k}{}_{a_1\ldots a_l}(\mu_1)_{b_1}\ldots(\mu_k)_{b_k}(\nu_1)^{a_1}\ldots(\nu_l)^{a_l} = T^{b1\ldots b_k}{}_{a_1\ldots a_l}(\phi_*\mu_1)_{b_1}\ldots([\phi^{-1}]^*\nu_l)^{a_l}. \tag{2.2}$$

Similarly, we could extend the action of the map ϕ_* to all tensors. However, since $\phi_* = (\phi^{-1})^*$ we only need to consider ϕ^* and $(\phi^{-1})^*$, see Ref. [6]. The

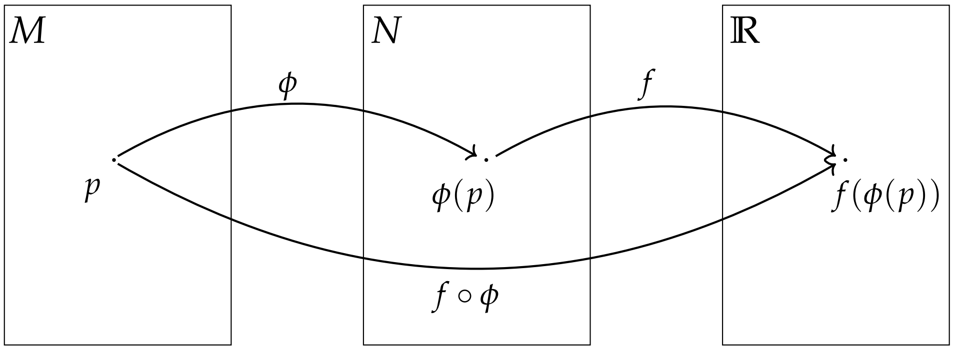


Figure 2.1: Ilustration of pull back.

notions of diffeomorphisms between manifolds and push forward of tensor fields define the concept of gauge freedom in general relativity. Roughly speaking, consider a physical theory described in terms of a manifold M and tensor fields $T^{(i)}$ defined on M. Let $\phi : M \to N$ be a diffeomorphism, which guarantees that M and N have the same manifold structure. Then, the solution $(N, \phi^* T^{(i)})$ is physically indistinguishable from the previous solution $(M, T^{(i)})$, which means that any physically well-posed statement about $(M, T^{(i)})$ will equally hold for $(N, \phi^* T^{(i)})$.

Now, we can use the diffeomorphism ϕ to construct the "motion vector" v^a described earlier as follows. Let M be a manifold. A *one-parameter group of diffeomorphisms* is a C^∞ map $\phi_t : \mathbb{R} \times M \to M$ such that for fixed t, $\phi_t : M \to M$ is a diffeomorphism. We can associate to ϕ_t the notion of a vector field v^a as follows. For fixed $p \in M$ and varying values of t, $\phi_t(p) : \mathbb{R} \to M$ is a curve, parameterized by t, called an orbit of ϕ_t, which passes through p at $t = 0$. Define v^a at p to be the tangent to this curve at $t = 0$. Hence, we have established the association between a one-parameter group of diffeomorphisms ϕ_t and a vector field v^a. In this way, we define the *Lie derivative* with respect to v^a, $\mathcal{L}_v$, of a tensor field $T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$ by

$$\mathcal{L}_v T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} = \lim_{t \to 0} \frac{\phi^*_{-t} T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} - T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}}{t}, \tag{2.3}$$

where all tensors in the equation above are evaluated at the same event p. The Lie derivative gives the notion of comparing $T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$ with $\phi^*_{-t} T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l}$ in a neighborhood of p for small values of t. In other words, the Lie derivative is the mathematical structure used to "compare" the original tensor field with the one "moved" along v^a. Therefore, in the light of the previous discussion, v^a defines a symmetry if $\mathcal{L}_v (T^{(i)}) = 0$. In particular, we have $\mathcal{L}_v(g_{ab}) = 0$ if v^a defines an isometry (i.e., spacetime symmetry). We can use this condition to study the special properties a vector field should satisfy to define a symmetry.

Firstly, from Eq. (2.3), it follows directly that $\mathcal{L}_v$ is a linear map from smooth tensor fields of rank (k, l) to smooth tensor fields of the same rank. Furthermore, for functions $f : M \to \mathbb{R}$ we have

$$\mathcal{L}_v(f) = v(f). \tag{2.4}$$

Secondly, the action of the Lie derivative in general tensor fields can be simplified if we introduce a coordinate system adapted to v^a in the sense that if x^0 is one of the coordinates, then $v^a = \left(\partial/\partial x^0\right)^a$. The action of ϕ_{-t} then corresponds to the transformation $x^\mu \to (x^0 + t, x^\nu)$, where $\nu = 1, 2, 3$. Then, it follows that

$$\phi^*_{-t} T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}(x^0, x^1, x^2, x^3) = T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}(x^0 + t, x^1, x^2, x^3). \tag{2.5}$$

As a consequence, the components of $\mathcal{L}_v T^{a_1 \dots a_k}{}_{b_1 \dots b_l}$ in this adapted coordinate system are given by

$$\mathcal{L}_v T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l} = \frac{\partial T^{\mu_1 \dots \mu_k}{}_{\nu_1 \dots \nu_l}}{\partial x^0}. \tag{2.6}$$

We can use this approach to find a coordinate-independent expression for the Lie derivative of tensor fields. From Eq. (2.6) we have

$$\mathcal{L}_v \mathrm{w}^\mu = \frac{\partial \mathrm{w}^\mu}{\partial x^0}. \tag{2.7}$$

Besides, defining the commutator $[v, \mathrm{w}]$ between the vector fields v^a and w^a to be

$$[v, \mathrm{w}] f \equiv v[\mathrm{w}(f)] - \mathrm{w}[v(f)]. \tag{2.8}$$

One can easily see that the commutator of any two vector fields is again a vector field. Using that $v(f) = v^a \nabla_a f$ we have

$$[v, \mathrm{w}] f = \left(v^a \nabla_a \mathrm{w}^b - \mathrm{w}^a \nabla_a v^b\right) \nabla_b f. \tag{2.9}$$

Thus, we can write the commutator in terms of the covariant derivative

$$[v, \mathrm{w}] = v^a \nabla_a \mathrm{w}^b - \mathrm{w}^a \nabla_a v^b. \tag{2.10}$$

It follows then, that in any coordinate system $\{x^\mu\}$, one can write the components of the commutator in this coordinate system as

$$[v, \mathrm{w}]^\mu = v^\nu \frac{\partial \mathrm{w}^\mu}{\partial x^\nu} - \mathrm{w}^\nu \frac{\partial v^\mu}{\partial x^\nu} \tag{2.11}$$

Thus, in the coordinate system adapted to v^a we have

$$[v, \mathrm{w}]^\mu = \frac{\partial \mathrm{w}^\mu}{\partial x^0}. \tag{2.12}$$

Therefore, the components of $\mathcal{L}_v \mathrm{w}^a$ and $[v, \mathrm{w}]^a$ are equal. However, since both the Lie derivative and the commutator are defined in a coordinate-independent manner, we obtain

$$\mathcal{L}_v \mathrm{w}^a = [v, \mathrm{w}]^a. \tag{2.13}$$

We can now derive the action of the Lie derivative in a dual vector w_a. From Eq. (2.4) we have

$$\begin{aligned} \mathcal{L}_v(\mathrm{w}_a u^a) &= v(\mathrm{w}_a u^a) \\ &= v^b \mathrm{w}_a \nabla_b u^a + v^b u^a \nabla_b \mathrm{w}_a \end{aligned}$$

Using now that $\mathcal{L}_v(\mathrm{w}_a u^a) = u^a \mathcal{L}_v \mathrm{w}_a + \mathrm{w}_a \mathcal{L}_v u^a$ with Eqs. (2.10) and (2.13) we get

$$\mathcal{L}_v \mathrm{w}_a = v^b \nabla_b \mathrm{w}_a + \mathrm{w}_b \nabla_a v^b. \tag{2.14}$$

We can generalize Eq. (2.14) for covectors by induction, arriving at

$$\mathcal{L}_v T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} = v^c \nabla_c T^{a_1 \ldots a_k}{}_{b_1 \ldots b_l} - \sum_i^{i=k} T^{a_1 \ldots c \ldots a_k}{}_{b_1 \ldots b_l} \nabla_c v^{a_i} + \sum_j^{j=l} T^{a_1 \ldots a_k}{}_{b_1 \ldots c \ldots b_l} \nabla_{b_j} v^c. \tag{2.15}$$

Finally, we can evaluate $\mathcal{L}_v(g_{ab}) = 0$ using Eq. (2.15)

$$v^c \nabla_c g_{ab} + g_{cb} \nabla_a v^c + g_{ac} \nabla_b v^c = 0. \tag{2.16}$$

where we have supposed that v^a defines an isometry. Now, using the fact that ∇_a is compatible with the metric, i.e., $\nabla_a g_{bc} = 0$, we arrive at

$$\nabla_a v_b + \nabla_b v_a = 0. \tag{2.17}$$

This equation is called *Killing's equation* and v^a is called a *Killing vector field*. Therefore, the necessary and sufficient condition to v^a be a generator of an isometry is that it satisfies Killing's equation (2.17).

From a more mathematical point of view, we say that ϕ is a *symmetry transformation* for a tensor field T if $\phi^* T = T$. In the case of the metric tensor, a symmetry transformation is called an *isometry*. It follows directly from Eq. (2.3) that if ϕ_t is a symmetry transformation, then $\mathcal{L}_v T = 0$. One can immediately see from Eq. (2.6) that if T is independent of one of the adopted coordinates, for example, x^0, we have a symmetry generated by $v^a = (\partial/\partial x^0)^a \equiv \partial^a/\partial x^0$. Therefore, if $g_{ab} = g_{ab}(x^i), i = 1, 2, 3$, then we have a Killing field $v^a = \partial^a/\partial x^0$. Conversely, choosing a coordinate system x^μ, Eq. (2.16) reads

$$v^\alpha \partial_\alpha g_{\mu\nu} + g_{\mu\alpha} \partial_\nu v^\alpha + g_{\nu\alpha} \partial_\mu v^\alpha = 0. \tag{2.18}$$

Now, let x^μ be a coordinate system and $v^a = \partial^a/\partial x^0$. Then, from Eq. (2.18) we get

$$\partial_{x^0} g_{\mu\nu} = 0 \Rightarrow g_{ab} = g_{ab}(x^i), \quad i = 1, 2, 3.$$

So far, we have established the concept of symmetry and its relation with Killing vector fields. We know from Noether's theorem that symmetries are associated with conserved quantities, so it is expected that Killing vectors lead to conserved quantities. These come in two ways. First, let ξ^a be a Killing vector field and γ a geodesic with tangent vector u^a. Then, $\xi_a u^a$ is constant along γ. To show this, we simply note that

$$\begin{aligned} u^a \nabla_a (\xi_b u^b) &= \xi_b u^a \nabla_a u^b + u^b u^a \nabla_a \xi_b \\ &= 0 \end{aligned}$$

since the first term vanishes due to the geodesic equation and the second one by Killing's equation.

The second type of conserved quantities comes from the stress-energy tensor T^{ab}. Again, assuming that ξ^a is a Killing vector field, we have the conserved four-current $J^a = T^{ab}\xi_b$:

$$\nabla_a J^a = \nabla_a (T^{ab}\xi_b) = (\nabla_a T^{ab})\xi_b + T^{ab} \nabla_a \xi_b = 0,$$

since the first term vanishes because T^{ab} is divergence-free and the second one due to Killing's equation. Now, suppose we have an isolated body. Let S and S' be three-dimensional surfaces that cross the world-tube of the body. It follows from

Gauss' theorem that

$$\int_{S'} T^{ab}\xi_b dS'_a - \int_S T^{ab}\xi_b dS_a = \int_V \nabla_a (T^{ab}\xi_b) dV = \int_V \nabla_a J^a dV = 0,$$

where V is the portion of the world-tube between S and S'. Therefore, $\int_S T^{ab}\xi_b dS_a$ is independent of the surface S and, thus, it is a conserved quantity.

Beyond conserved quantities, Killing vector fields can also characterize the spacetime itself. A spacetime is said to be *stationary* if it possesses an everywhere timelike Killing vector field ξ^a. It is said to be *static* if, in addition, there is a spacelike hypersurface Σ which is orthogonal to the orbits of the isometry, see Fig. 2.2. By Frobenius's theorem[2], this is equivalent to the requirement that the corresponding timelike Killing vector, ξ^a, satisfies $\xi_{[a}\nabla_b\xi_{c]} = 0$. Physically speaking, stationarity means invariance by "time" translation, where "time" means the parameter t of the isometry, while, in addition to time invariance, staticity means invariance by time reflection. For instance, if we consider the spacetime of the Earth (an axisymmetric sphere, in vacuum, rotating uniformly) it is clear that nothing happens in time, but in a time reversion, the Earth would rotate in the opposite direction. Therefore, in this simple example, the spacetime generated by Earth is stationary, but not static. In order to make it static, one should consider a non-rotating Earth.

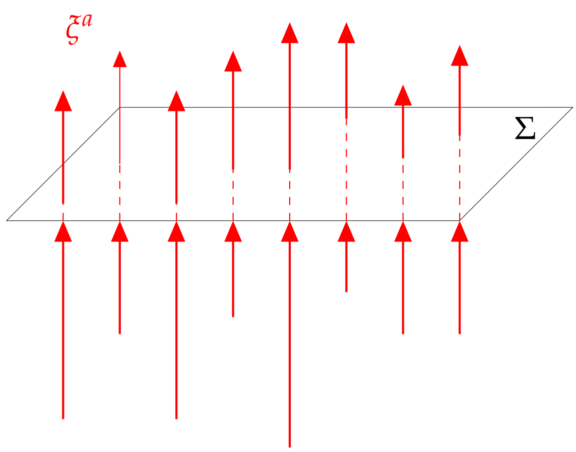


Figure 2.2: Representation of the timelike Killing vector ξ^a orthogonal to a space-like surface Σ.

[2]See Appendix B of Ref.[6] for the statement of Frobeniu's theorem.

One can give a general form of the metric tensor of a static spacetime using the above results as follows. Let ϕ_t be a one-parameter group of diffeomorphisms whose orbits are timelike. Let ξ^a be the corresponding timelike Killing vector field and Σ the spacelike surface to which ξ^a is orthogonal. Assuming that $\xi^a \neq 0$ on Σ, we can define a family of hypersurfaces, Σ_t, using the action of the isometry ϕ_t on Σ

$$\Sigma_t = \phi_t(\Sigma).$$

Since ϕ_t is an isometry, Σ_t is isometric to Σ and, thus, ξ^a is orthogonal to each Σ_t, see Fg. 2.3. Now, let $x^\mu = (t, x^i)$, $i = 1, 2, 3$, be a coordinate system in Σ. We can

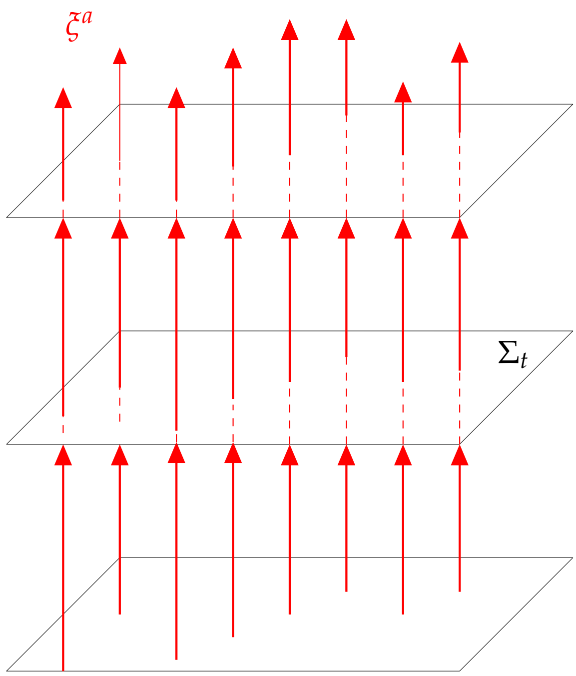


Figure 2.3: Representation of the family of hypersurfaces Σ_t.

use it to describe each Σ_t by means of the action of ϕ on each $p \in \Sigma$ demanding that

$$x^i(\phi_t(p)) = x^i(p), \tag{2.19}$$

i.e., the coordinates describing a point p in one of the hypersurfaces Σ_t are the same as the coordinates describing the image of this point under the action of ϕ_t, which lies in other Σ_t. Therefore, every event can be described by $\{t, x^i\}$. Using now the fact that $g_{\mu\nu} = g_{ab}(\partial^a_\mu \partial^b_\nu)$ along with the fact that ∂^a_t is orthogonal to ∂^a_i, the line element of such spacetime reads

$$ds^2 = g_{00}(x^i)dt^2 + g_{kj}(x^i)dx^k dx^j, \tag{2.20}$$

where we have used the fact that if ∂_t^a is a Killing vector, the components of the metric in this coordinate basis are independent of t. Eq. (2.20) is the general form of the metric components adapted to ξ^a in a static spacetime. Note that one can find an expression for $g_{00}(x^i)$ in terms of the Killing vector field by noticing that $\xi^a \xi_a = g_{00}$. A stationary spacetime will inevitably have cross terms of the kind $dtdx^i$ but, still, the components of the metric will be independent of t.

Finally, we mention the notion of conformal isometry which will be important in the study of causal structure. Let (M, g_{ab}) and (M', g'_{ab}) be spacetimes. A *conformal isometry*, ϕ, between these spacetimes is a diffeomorphism $\phi : M \to M'$ for which there is a function φ, called *conformal factor*, such that $\phi^* g_{ab} = \varphi^2 g'_{ab}$. Since ϕ is a diffeomorphism, φ is non-vanishing. Physically speaking, a conformal isometry is a transformation that leaves the causal structure of the spacetime unaltered, which is the same as saying that a null vector in (M, g_{ab}) is again a null vector in (M', g'_{ab})–this fact can readily be seen from the definition of conformal isometry.

2.2 Causal structure

As previously outlined, spacetime is a four-dimensional manifold endowed with a Lorentzian metric, dynamically coupled to the matter distribution via Einstein's equation. However, not all solutions of Eq. (2.1) are physically reasonable. For instance, Godel's universe, an exact solution of Einstein's equation describing a pressure-free perfect fluid, allows one to travel back in time, which is a non-expected property for a physical universe [7]. Hence, one should impose additional conditions on the spacetime to ensure that it really describes the physical world. In this manner, the study of the causal structure of spacetime has proved to be a powerful way to analyze scenarios that should be discarded as physically unreasonable. In fact, a causally well-behaved spacetime has an intrinsic relation with the unitary evolution of quantum fields, as we will show in a moment (see also Chap. 3). In this section, we discuss the most important of these conditions for general spacetimes.

2.2.1 Time-orientable spacetimes

The first condition that we should demand for our physical spacetime is a global notion of future and past. Let (M, g_{ab}) be a spacetime. The tangent vector space $\mathcal{V}_p$ of each $p \in M$ is isomorphic to Minkowski spacetime (this comes directly from the definition of manifolds). Thus, we can refer to the light cone of p as the light cone of the origin of $\mathcal{V}_p$. In this way, as in special relativity, we may designate half of the light cone as the future and the other half as the past; see Fig. 2.4. If a continuous definition of future and past for the light cone at each $p \in M$ can be made, then we say that (M, g_{ab}) is a *time-orientable* spacetime. It can be shown that every spacetime based on a simply connected manifold is time-orientable [8]. From now on, we will only consider time-orientable spacetimes.

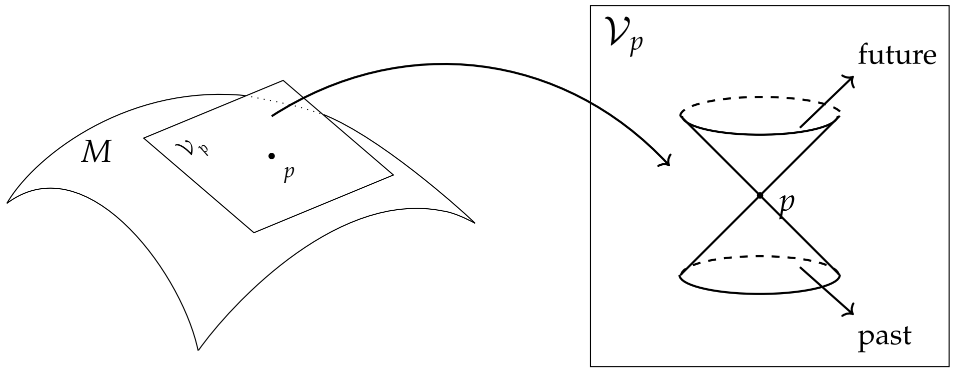


Figure 2.4: Lightcone of an event $p \in M$ as the lightcone of the origin of $\mathcal{V}_p$

2.2.2 Causality conditions

The notion of time-orientable spacetimes does not rule out situations that go against our intuitive thinking of how nature should behave. Consider, for instance, a simply connected spacetime (M, g_{ab}). From the last section, (M, g_{ab}) is time-orientable. Let γ be a closed timelike curve. Now, consider an observer $\mathcal{O}$ joining γ at some event $p \in M$ and leaving it at some $q \in M$, see Fig. 2.5. However, since the curve is everywhere timelike and it continues from q to p, the observer could prepare a rocket ship to travel from q to p, arriving at the instant of departure, and prevent the observer from setting out in the first place. It should be stressed that this paradoxical situation has nothing to do with the topology of spacetime, namely that γ is homotopic to the trivial closed curve at p, which prevents it from ruining the time orientation. Although this could appear an arbitrary example, Godel's universe, as commented above, has closed timelike curves but has the

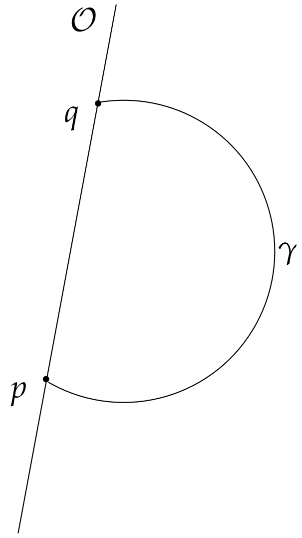


Figure 2.5: Example of an observer O crossing a closed timelike curve through p.

manifold structure of $\mathbb{R}^4$, i.e., it is time-orientable. This is an example of an exact solution of Einstein's equation that is time-orientable but has closed timelike curves [7]. In this way, other physically motivated conditions, named *causality conditions*, should be imposed on spacetimes to prevent the existence of closed timelike curves. In this section, we address the mathematical formulation of such conditions and their relation with unitarity. We follow here mainly the approach of Ref. [6].

Let $\gamma(t)$ be a curve parameterized by t and let u^a be its tangent. We say that $\gamma(t)$ is a *timelike* future directed (resp. past directed) curve if, at each $p \in \gamma(t)$, the vector field u^a is timelike and it is inside the future (resp. past) light cone of p. Besides, we say that $\gamma(t)$ is a *causal* future directed (resp. past directed) curve if, at each $p \in \gamma(t)$, the vector field u^a is either timelike or null and is in the future (resp. past) light cone of p. One can use these definitions to describe causal relations between events in M. Let $p, q \in M$. We say that q is in the chronological future of p, and denote this relation by $p \prec q$, if there exists everywhere a future-directed timelike curve $\gamma(t)$ such that $\gamma(0) = p$ and $\gamma(1) = q$. These causal relations of events in (M, g_{ab}) define causal sets in spacetime. We define the *chronological future* of $p \in M$, $I^+(p)$, as the set of events that are in the chronological future of p

$$I^+(p) = \{q \in M, p \prec q\}. \tag{2.21}$$

The same definition holds for the chronological past $I^-(p)$ of p by changing future with past in the above definition. We define the chronological future of a set $S \subset M$ by

$$I^+(S) = \cup_{p \in S} I^+(p). \tag{2.22}$$

An analogous definition holds for $I^-(S)$.

The notions of chronological future and past of a set can be used to define the general notion of horizons, ubiquitously present in the study of black holes. Let γ be a timelike curve. The *future horizon*, h^+, of γ is defined to be the boundary of $I^-(\gamma)$, written as $\dot{I}^-(\gamma)$ (see Fig. 2.6). Analogously, we can define the *past horizon*, h^-. Note that one can define h^+ and h^- for a family of observers. In addition, one can define Killing horizons. Let ξ^a be a Killing vector field. A *Killing horizon*, h, is a null surface having the property that ξ^a is normal to it.

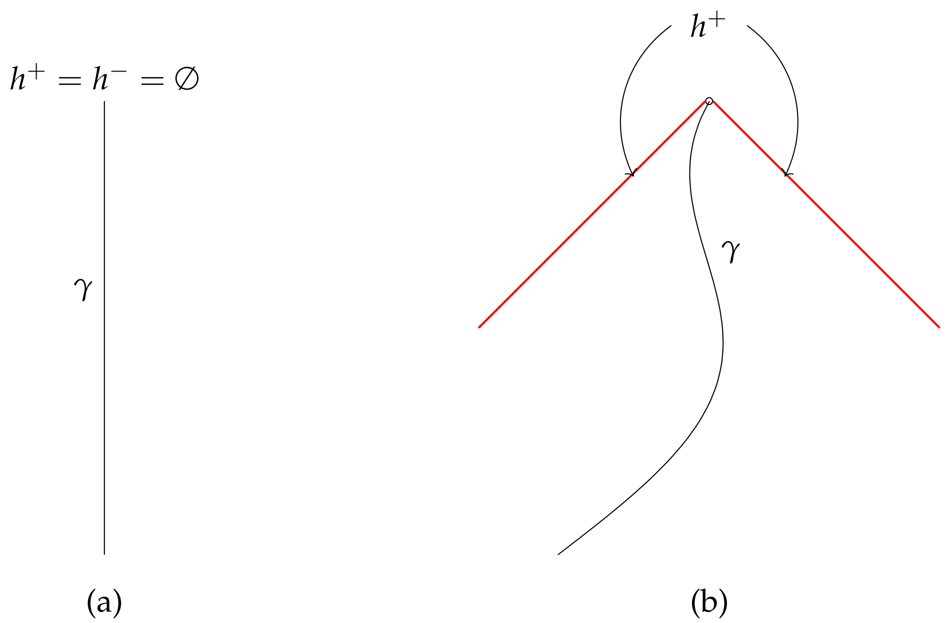


Figure 2.6: Example of past and future horizons for a inextendible timelike curve γ. In figure (a), we have a timelike curve in Minkowski spacetime with no past nor future horizons. In figure (b), we have a timelike curve in Minkowski spacetime with an event removed; in this case, γ has only a future horizon.

In addition, we say that q is in the causal future of p, and denote this relation by $p \ll q$, if there exists everywhere a future-directed causal curve $\gamma(t)$ such that $\gamma(0) = p$ and $\gamma(1) = q$. Analogously, we define the *causal future* of p, $J^+(p)$, as the set of events that are in the causal future of p

$$J^{+}(p) = \{q \in M, p \ll q\}. \tag{2.23}$$

Again, the causal past of p, $J^{-}(p)$, can be defined in the same manner just changing future by past in the above definition. We define the causal future of a set $S \subset M$ by

$$J^{+}(S) = \cup_{p \in S} J^{+}(p). \tag{2.24}$$

As before, an analogous definition holds for $J^{-}(S)$. See Fig. 2.7 illustrating the causal future of a closed set S. Physically, the set $I^{+}(S)$ consists of events that can be influenced by signals traveling slower than the speed of light coming from S, while $J^{+}(S)$ consists of events that can be influenced by signals traveling at a speed lower than or equal to the speed of light coming from S. Besides, the existence of a future horizon h^{+} of some timelike curve γ means that there are events that can not influence γ using signals traveling at a speed lower than the speed of light.

Now we can address the causality conditions. At the beginning of this section, we view that one should rule out the existence of closed timelike curves from the spacetime, this is called *chronological condition* [7]. However, even if the spacetime has no closed timelike curves, it could be "on the verge" of violating causality, i.e., there could be timelike curves that come arbitrarily close to intersecting themselves, although none of them actually do. One can, then, simply make an arbitrarily small perturbation on the metric of such spacetimes to generate closed timelike curves. In this manner, we should regard such spacetimes as physically unreasonable. One can rule out this kind of situation as follows. Let (M, g_{ab}) be a spacetime. If for all $p \in M$ and every neighborhood, O_p, of p there is a neighborhood V_p of p contained in O_p such that no causal curve intersects V_p more than once, then the previous situation can not occur. If (M, g_{ab}) satisfies this condition, then it is *strongly causal*. Thus, if a spacetime violates strong causality at p, there exist causal curves near p that come arbitrarily close to intersecting themselves.

However, one can still have a strongly causal spacetime but a modification of g_{ab} in an arbitrarily small neighborhood of two or more events produces closed causal curves. Thus, we need to impose some other condition to guarantee that the spacetime is not "on the verge" to produce closed timelike curves. Such conditions can be constructed as follows. Let t^a be a timelike vector at $p \in M$. Define $\tilde{g}_{ab}$ by

$$\tilde{g}_{ab} = g_{ab} - t_a t_b,$$

where g_{ab} is the spacetime metric. It is easy to see that $\tilde{g}_{ab}$ is also a Lorentzian signature metric at p: Let r^a be a spacelike vector orthogonal to t^a with respect to g_{ab} at p. Then, $\tilde{g}_{ab}r^a r^b = g_{ab}r^a r^b > 0$. Besides, $\tilde{g}_{ab}t^a t^b = g_{ab}t^a t^b - t_a t^a t_b t^b = t_a t^a(1 - t_b t^b) < 0$. Finally, we have $\tilde{g}_{ab}t^a r^b = g_{ab}t^a r^b = 0$. So $\tilde{g}_{ab}$ is a Lorentzian signature metric at p. Moreover, the light cone of $\tilde{g}_{ab}$ is larger than the one of g_{ab}, i.e., every timelike vector and null vector of g_{ab} is a timelike vector of $\tilde{g}_{ab}$. We already showed the first part of the claim. Now let k^a be a null vector of g_{ab}, thus, $\tilde{g}_{ab}k^a k^b = g_{ab}k^a k^b - t_a k^a t_b k^b = -t_a k^a t_b k^b < 0$. Therefore, if spacetime is "on the verge" of having closed timelike curves, then by "enlarging" the light cone at each point, we should be able to produce closed timelike curves. Thus, we define a spacetime (M, g_{ab}) to be *stably causal* if there exists a continuous non-vanishing timelike vector field t^a such that the spacetime $(M, \tilde{g}_{ab})$ possess no closed timelike curves. It follows that stable causality implies strong causality [6, 7]. Henceforth, every spacetime (M, g_{ab}) considered in this thesis is presumed to satisfy all the causality conditions discussed so far.

Now that we have described the physically reasonable conditions that a spacetime with no causal pathologies should possess, we can turn to the problem of what this causally well-behaved spacetime has to do with unitary evolution. In the previous definitions, attention was focused on the collections of events that could be influenced by a set S of events. Now we shall focus on the collection of events that can completely be determined by a set of events S.

First, we need the notion of extendible curves. Let $\gamma(t)$ be a future-directed causal curve. We say that $p \in M$ is a *future endpoint* of $\gamma(t)$ if for every neighborhood O_p of p, there exists a t_0 such that $\gamma(t) \in O_p$ for all $t > t_0$. The curve $\gamma(t)$ is said to be *future inextendible* if it has no future endpoints. Second, it is necessary to introduce the notion of achronal sets. A subset $S \subset M$ is said to be *achronal* if $I^+(S) \cap S = \emptyset$, i.e., no pair of points $p, q \in S$ can be joined by a timelike curve. A simple example of an achronal set is the surface of $t = const$ in Minkowski spacetime, where t is the usual Cartesian coordinate.

Now, let S be a closed, achronal set. We define the *future domain of dependence*

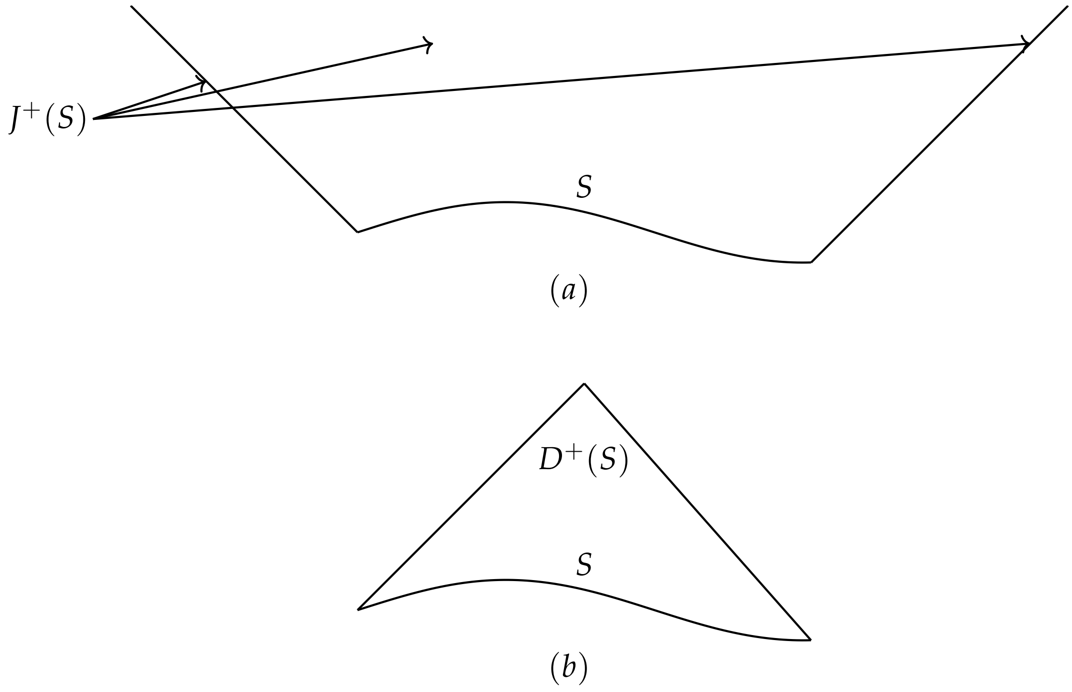


Figure 2.7: Example of the (a) causal future and (b) future domain of dependence of a closed set S.

of S, denoted $D^+(S)$, by

$$D^+(S) = \{q \in M, \text{ every past inextendible causal curve through } p \text{ intersects } S \text{ once and only once}\}.$$

(In fact, we could drop the imposition "once and only once" since we are only considering spacetimes with no closed causal curves). An analogous definition can be made for the past domain of dependence, $D^-(S)$ (see Fig. 2.7). The set $D^+(S)$ is fundamental in formulating unitary physical theories. Since information can not travel faster than light any signal sent to $p \in D^+(S)$ from the past of S must be registered on S. In this way, using the information in S as initial conditions, one should be able to predict what happens at $p \in D^+(S)$ (provided unitary evolution of the fields is given). Conversely, if $p \in I^+(S)$ but $p \notin D^+(S)$, then it would be possible to send a signal to p from the past of S without influencing S, and the complete knowledge of S will not be enough to determine conditions at p. A similar analysis holds for $D^-(S)$, where in this case, from the knowledge of conditions on S we expect to be able to retrodict conditions at all $q \in D^-(S)$. The *domain of dependence* of S, denoted $D(S)$, is defined simply by $D(S) = D^+(S) \cup D^-(S)$. In the light of the previous discussion, $D(S)$ represents the complete set of events for which all conditions should be determined by a knowledge of conditions on S.

A closed achronal set Σ for which $D(\Sigma) = M$ is called a *Cauchy surface*. Cauchy surfaces represent "instants of time" throughout the spacetime. A spacetime (M, g_{ab}) which possesses a Cauchy surface Σ is said to be *globally hyperbolic*. Therefore, in globally hyperbolic spacetimes the entire future and past of (M, g_{ab}) can be determined by data on Σ. An important property of a globally hyperbolic spacetime is that M can be foliated by Cauchy surfaces and the topology of M is $\mathbb{R} \times \Sigma$, where Σ denotes any Cauchy surface used in the foliation [9]. Therefore, from the previous discussion, it follows that global hyperbolicity is a much-desired condition for well-posed initial value formulation of physical field equations [6, 10]. It is challenging to conceive how information, and thus unitarity, could be conserved in non-globally hyperbolic spacetimes.

Finally, we comment on the relation between causality conditions and the well-posed unitarity in our spacetime. Let (M, g_{ab}) be a globally hyperbolic spacetime. Then (M, g_{ab}) is stably causal. Hence, (M, g_{ab}) is strongly causal. Therefore, the unitary evolution of physical systems is directly related to a causally well-behaved spacetime.

2.2.3 Carter-Penrose diagrams

Finally, we review the basic construction and properties of Cartan-Penrose diagrams, which are very useful in the study of the causal structure of spacetime since it has the nice feature of representing all null curves with a $\pi/4$-angle with the vertical. Consider the Minkowski spacetime $(\mathbb{R}^4, \eta_{ab})$ with the line-element cast in spherical coordinates (t, r, θ, ϕ)

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2, \tag{2.25}$$

where $d\Omega^2 \equiv d\theta^2 + \sin\theta^2 d\phi^2$ is the solid angle element. Consider the case of radial $(d\Omega^2 = 0)$ null curves in Minkowski spacetime

$$-dt^2 + dr^2 = 0,$$

whose solution is given by

$$t = \pm r + constant, \tag{2.26}$$

where the plus sign corresponds to *outgoing* (r increasing with t) null curves and the minus sign (r decreasing with t) to *ingoing* ones. One can show that solutions of the type (2.26) imposed to the coordinates of a null vector field k^a imply k^a satisfying the geodesic equation. Therefore, the curves given by Eq. (2.26) are radial null geodesics in the Minkowski spacetime.

Using these solutions, define the following coordinates transformation

$$\begin{aligned} u &= t - r \\ v &= t + r \end{aligned} \tag{2.27}$$

which gives

$$ds^2 = -\frac{1}{4}dvdu + \frac{1}{4}\left(u^2 - v^2\right)d\Omega^2, \tag{2.28}$$

The coordinates (u, v) are called *null coordinates* since $u = const$ and $v = const$ are solutions of the radial null geodesic equation. Note that since $t \in \mathbb{R}$ and $r \geq 0$, we have $u, v \in (-\infty, +\infty)$. Besides, it is important to stress that for any functions $f(u)$ and $g(v)$ it is still true that $f(u) = const$ (resp. $g(v) = const$) corresponds to outgoing (resp. ingoing) radial null geodesics in Minkowski spacetime. Now, define a new set of coordinates (U, V) by

$$\begin{aligned} u &= \tan U \\ v &= \tan V \end{aligned} \tag{2.29}$$

where now $U, V \in (\pi/2, \pi/2)$. The metric then reads

$$ds^2 = \frac{\sec^2 U \sec^2 V}{4}\left(-dUdV + \sin^2(U - V)d\Omega^2\right). \tag{2.30}$$

Considering now only the radial components of the metric, we can readily extract the conformal factor $\varphi^2 = \sec^2 U \sec^2 V/4$. The conformally related metric is clearly $ds'^2 = -dUdV$. We stress that ds'^2 should be regarded as the unphysical metric since it is not required to satisfy the field equations. However, the Carter-Penrose diagram will be formulated from ds'^2 instead of ds^2, which will allow us to use the unphysical metric (generally easier to treat) to study the causal structure of the physical metric (generally more complicated to treat).

Before moving to the diagram, it is convenient to define a new set of coordinates

(T, R) given by

$$\begin{aligned} U &= T - R, \\ V &= T + R, \end{aligned} \tag{2.31}$$

where $T \in (-\pi/2, \pi/2)$ and $R \in [0, \pi/2)$. Now, using (T, R) we can represent the whole spacetime, including its different types of infinities (see the comments in the next paragraph) in a finite diagram, shown in Fig. 2.8, which is the Carter-Penrose diagram of the Minkowski spacetime, where each point on it corresponds to a 2-sphere.

There are several regions of interest in this picture: (1) The *past timelike infinity* i^- given by $T = -\pi/2$ and $R = 0$. (2) The *spatial infinity* i^0 given by $T = 0$ and $R = \pi/2$. (3) The *future timelike infinity* i^+ given by $T = \pi/2$ and $R = 0$. (4) The *past null infinity* $\mathcal{J}^-$ given by the relation $T - R = -\pi/2$. (5) The *future null infinity* $\mathcal{J}^+$ given by the relation $T + R = \pi/2$. Note that all spacelike geodesics in Minkowski spacetime begin and end at i^0, all the timelike geodesics come from i^- and end at i^+ and all null geodesics come from $\mathcal{J}^-$ and end at $\mathcal{J}^+$. Furthermore, we have depicted a few additional structures to illustrate the previously made discussion: (i) Spacelike hypersurfaces Σ_1 and Σ_2. (ii) Curves of constant u and v. (iii) A curve γ. Firstly, note that because of the construction of the diagram, all null geodesics make a $\pi/4$-angle with the vertical, which illustrates the point made at the beginning of this discussion. Secondly, using this property we can immediately see that γ is in fact a timelike curve since one could easily draw the light cones along γ and verify that it is always inside of them. Lastly, if we pick any point of the diagram in the future (resp. past) of Σ_1 and trace all the past (resp. future) inextendible curves from it, we see that they all intersect Σ_1, so Σ_1 is a Cauchy surface of the Minkowski spacetime. The same argument holds for Σ_2.

Now that we have established the causal conditions one should have on a physical spacetime and illustrated it in the simple case of Minkowski spacetime, we can now move on to a case of major interest in general relativity: the Schwarzschild metric.

2.3 Schwarzschild spacetime

One of the most known (non-trivial) solution of Einstein's equation is the Schwarzschild metric. It is a spherically symmetric, static, and vacuum solution,

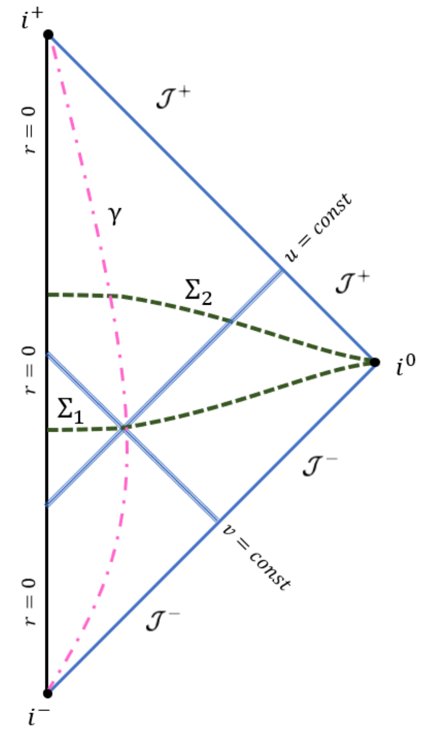


Figure 2.8: Carter-Penrose diagram of Minkowski spacetime

whose line element is most usually cast as

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \frac{dr^2}{(1 - \frac{2M}{r})} + r^2 d\Omega^2, \tag{2.32}$$

where $M = const$ is the single parameter characterizing the spacetime. A known result, named *Birkhoff's theorem*, states that the Schwarzschild spacetime is the only spherically symmetric solution of Einstein's vacuum equation.

To unveil the causal structure of Schwarzschild spacetime we proceed in the same manner as in the Minkowski case. We start by analyzing the equation of null radial curves

$$dt^2 = \frac{dr^2}{(1 - \frac{2M}{r})^2}$$

whose solution is given by

$$t = \pm \left(r + 2M \log\left|\frac{r}{2M} - 1\right|\right) + const. \tag{2.33}$$

Defining the "tortoise coordinate" r^* by

$$r^* = r + 2M \log \left| \frac{r}{2M} - 1 \right|, \tag{2.34}$$

we can write the equation for null radial curves in Schwarzschild metric

$$t = \pm r^* + const. \tag{2.35}$$

Following the procedure applied in the Minkowski case, we define the null coordinates

$$\begin{aligned} \bar{u} &= t - r^*, \\ \bar{v} &= t + r^*, \end{aligned} \tag{2.36}$$

and cast the line element as

$$ds^2 = -\left(1 - \frac{2M}{r}\right) d\bar{u} d\bar{v} + r^2 d\Omega^2, \tag{2.37}$$

where r should now be viewed as a function of $\bar{u}$ and $\bar{v}$ given implicitly by

$$r + 2M \log \left| \frac{r}{2M} - 1 \right| = \frac{\bar{v} - \bar{u}}{2}. \tag{2.38}$$

One can see that for finite values of t, the value of $r = 2M$ is equivalent to $(\bar{u} \to \infty,\ \bar{v})$ and $(\bar{u},\ \bar{v} \to -\infty)$.

Using Eq. (2.38) we can write

$$1 - \frac{2M}{r} = \pm \frac{2Me^{-\frac{r}{2M}}}{r} e^{\frac{\bar{v}-\bar{u}}{4M}},$$

where the plus sign refers to $r > 2M$ and the minus sign to $r < 2M$. Then, the line element reads

$$ds^2 = \mp \frac{2Me^{-\frac{r}{2M}}}{r} e^{\frac{\bar{v}-\bar{u}}{4M}} d\bar{u} d\bar{v} + r^2 d\Omega^2. \tag{2.39}$$

Defining the new set (u, v) of null coordinates

$$\begin{aligned} u &= \mp e^{-\frac{\bar{u}}{4M}}, \\ v &= e^{\frac{\bar{v}}{4M}}, \end{aligned} \tag{2.40}$$

where, again, the upper sign refers to $r > 2M$ and the lower one to $r < 2M$, the

Schwarzschild metric reads

$$ds^2 = -\frac{32M^3 e^{-\frac{r}{2M}}}{r} du dv + r^2 d\Omega^2. \tag{2.41}$$

Note that now the region $r = 2M$ is equivalent to $(u = 0, v)$ and $(u, v = 0)$. In this way, $r = 2M$ is a (bifurcating) *null surface*.

Finally, define the set (U, V) of null coordinates by

$$\begin{aligned} u &= \tan U, \\ v &= \tan V, \end{aligned} \tag{2.42}$$

with which the line element in the $U - V$ plane reads

$$ds^2 = -\frac{32M^3 e^{-\frac{r}{2M}}}{r} \sec^2 U \sec^2 V dU dV. \tag{2.43}$$

In this way, the conformal factor is now $\varphi^2 = \frac{32M^3 e^{-\frac{r}{2M}}}{r} \sec^2 U \sec^2 V$ and the unphysical metric is simply $ds'^2 = -dUdV$, where $U, V \in (-\pi/2, \pi/2)$.

In order to draw the Carter-Penrose diagram of Schwarzschild spacetime we should note first that from the definition of (u, v) it follows that

$$\begin{aligned} uv &= \mp e^{\frac{\tilde{v}-\tilde{u}}{4M}} \\ &= \mp e^{\frac{r^*}{2M}} \\ &= \mp e^{\frac{r}{2M}} \left| \frac{r}{2M} - 1 \right| \\ &= e^{\frac{r}{2M}} \left(1 - \frac{r}{2M} \right). \end{aligned}$$

Then, we have

$$uv \leq 1, \tag{2.44}$$

which gives

$$\tan U \tan V \leq 1, \tag{2.45}$$

where the equality holds at $r = 0$. Note, however, from Eq. (2.41) that ds^2 is ill-defined at $r = 0$. This is because $r = 0$ is a *true singularity* of the Schwarzschild metric. This fact can readily be seen by computing the *Krestchmann scalar* $K = R_{abcd} R^{abcd} = \frac{48M^2}{r^6}$ and noticing that K diverges as $r \to 0$. The term "singularity" means, in a loose sense, a breakdown in ordinary dynamical evolution. Next, define the coordinates (T, R) by

$$\begin{aligned} U &= T - R, \\ V &= T + R, \end{aligned} \tag{2.46}$$

where $T, R \in (-\pi/2, \pi/2)$. Now Eq. (2.45) reads

$$\tan(T + R)\tan(T - R) \leq 1. \tag{2.47}$$

We can use trigonometric relations in the above equation to conclude that the singularity $r = 0$ is located at $T = \pm\pi/4$, so we can only have $T \in (-\pi/4, \pi/4)$. We can use these new ranges to draw the Carter-Penrose diagram for Schwarzschild spacetime, depicted in Fig. 2.9. The diagram shows four regions: I, II, III, and IV. Region IV is given by $U, V < 0$ and it is causally connected with all other regions in the sense that it can influence (but not be influenced by) all the rest of spacetime. This region is called a *white hole*. Region I, given by $U < 0, V > 0$, is causally disconnected from region II (given by $U > 0, V < 0$) and vice-versa, so none of these regions can affect the other. Furthermore, region II, given by $U > 0, V > 0$, is causally connected with the rest of the spacetime, so that it can be influenced by the other regions. Since light travels with a $\pi/4$ angle, one can readily see that once light has entered region II it can never leave. In fact, since all information travels in timelike or lightlike trajectories, everything that enters region II can no longer leave it. This looks like the intuitive notion of a *black hole*. Indeed, as we will see in the next section, region II satisfies the topological definition of black holes. In addition, one can see that the future and past event horizons of the black hole are located respectively at $U = 0$ and $V = 0$, i.e., at $r = 2M$. Besides, once inside the black hole, one head inevitably to the singularity at $r = 0$, and it is destroyed by it[3]. In fact, since the singularity is completely inside Region II, no external observer can actually be influenced by it.

In order to clarify the above discussion, we have depicted in the diagram two timelike curves: γ_1, which goes safely to the asymptotic future, and γ_2, which enters the black hole and reaches the singularity in finite proper time. Finally, in this diagram, we have the same asymptotic (timelike, space-like and null) regions as the ones depicted in the Minkowski case, characterizing the asymptotic flatness feature of the Schwarzschild metric[4]. In addition, we have depicted a spacelike

[3]This is what physically characterizes a *spacelike singularity*.

[4]Asymptotically flat spacetimes has the property of "looking like" flat spacetime at space-like

hypersurface Σ that is also a Cauchy surface for the Schwarzschild spacetime, as can be seen from the same arguments made in the previous section.

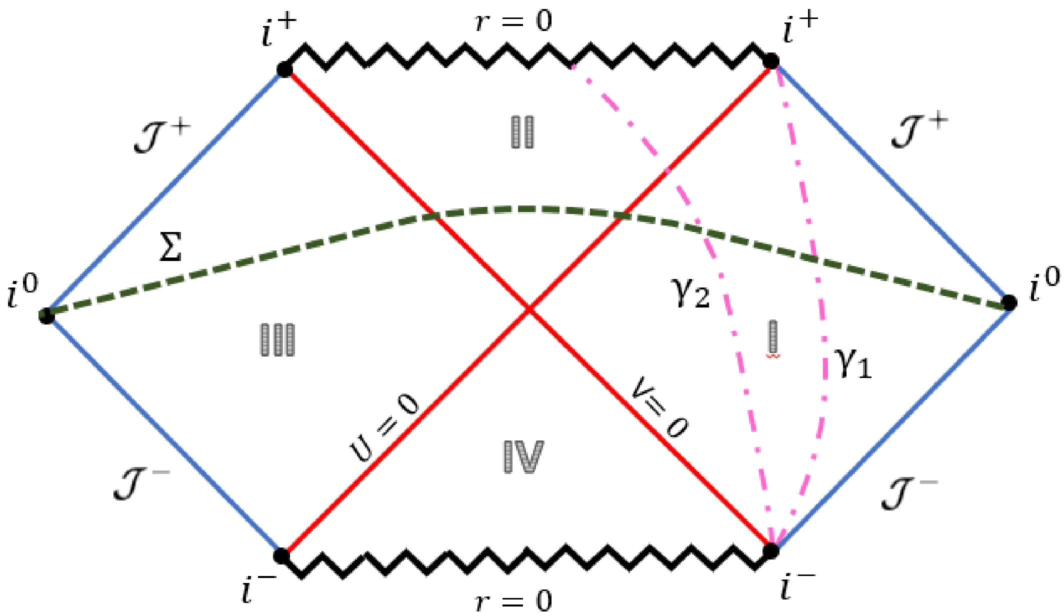


Figure 2.9: Carter-Penrose diagram of Schwarzschild spacetime.

Finally, we would like to address the characteristics of Killing vector fields in Schwarzschild spacetime, as they will play a fundamental role in quantum field theories. It is straightforward to see from Eq. (2.32) that $g_{\mu\nu}$, i.e., the metric components in the coordinate system adopted, is independent of t. From the results of Sec. 2.1, it follows that $\xi^a = \partial^a/\partial t$ is a Killing vector field. From the line element, we get

$$\xi^a \xi_a = -\left(1 - \frac{2M}{r}\right). \tag{2.48}$$

In this way, ξ^a is timelike for $r > 2M$, space-like for $r < 2M$ and null at $r = 2M$. But we already have seen that $r = 2M$ is a null hypersurface (the event horizon), so the event horizon is a Killing horizon. In the next section, we will see that this is the case for stationary spacetimes.

2.4 Black holes

The intuitive notion of a black hole is of a “region of no return", i.e., a region of spacetime where gravity is so strong that no particle nor light rays could ever

and null infinities, see Ref. [6] for a precise mathematical definition of it.

escape from it after entering. We have already seen that a black hole is present in the static, spherically symmetric, vacuum solution of Einstein's equation. The purpose of this section is to generalize the intuitive idea about black holes.

From the simple example we have treated, one would be tempted to define black holes in a spacetime (M, g_{ab}) to be simply a subset $A \subset M$ such that for any $p \in A$ we have $J^+(p) \subset A$. If this was the case, Region II of Fig. 2.9 could not be regarded as a black hole, since it contains the spacelike singularity at $r = 0$, which is not a subset of M. Moreover, with that definition, the causal future of any set of any spacetime would be a black hole. Therefore, this is not the definition that encodes the "no escaping" property of black holes. The striking feature of Region II in Fig. 2.9 is that it is confined to $r < 2M$, i.e., it does not extend "out to infinity". Hence, for asymptotically flat spacetimes, the impossibility of escaping to $\mathcal{J}^+$ seems to provide an appropriate characterization of black holes. Comparing Figs. 2.8 and 2.9 we see that $J^-(\mathcal{J}^+)$ in both cases is nonsingular, but for the Schwarzschild case, it does not include the entire physical spacetime, in contrast to the Minkowski case. Hence, the idea that $J^-(\mathcal{J}^+)$ is well-behaved but does not include the entire spacetime realizes the physical intuition of a black hole.

We can now use the above discussion to motivate the mathematical definition of black holes as follows. Let (M, g_{ab}) be an asymptotically flat spacetime. Let (M', g'_{ab}) be the associated unphysical spacetime. We say that (M, g_{ab}) is *strongly asymptotically predictable* if in (M', g'_{ab}) there is an open region $V' \subset M'$ with $\overline{M \cap J^-(\mathcal{J}^+)} \subset V'$ such that (V', g'_{ab}) is globally hyperbolic. In this way, a strongly asymptotically spacetime is said to contain a black hole if M is not contained in $J^-(\mathcal{J}^+)$. Thus, the *black hole region* B of such spacetime is defined to be

$$B = M - J^-(\mathcal{J}^+). \tag{2.49}$$

The boundary of B in M is called the *event horizon*. The same definition of a black hole could be given for asymptotically flat spacetimes that are not strongly predictable, see Ref. [6].

This definition of a black hole allows the existence of *naked singularities*, i.e., singularities present in the outside of the event horizon, which, therefore, could influence observers at large distances from the black hole. A more accurate definition of naked singularities can be given as follows. An inextendible spacetime (M, g) is said to be future nakedly singular if admits a future inextendible causal curve which lies entirely in the past of some point $x \in M$. From the definition of singularities, the existence of naked singularities results in a breakdown of predictability for external observers. Hence, a spacetime with naked singularities can not be globally hyperbolic. However, it is widely accepted that no such singularities occur in physically reasonable gravitational collapses. This conjecture is known as the *cosmic censor hypothesis*, which states that all physical spacetimes are globally hyperbolic. The validity of the cosmic censor would ensure that any outside observer could not be influenced by naked singularities. Thus, one would have a strongly asymptotically predictable spacetime. Therefore, if the cosmic censor conjecture holds, then every physical black hole is predictable despite the presence of the spacelike singularity in its interior (or any other type of singularity inside of more general black holes [6, 7]).

Let us now introduce the notion of a black hole at a given instant of time. Let (M, g_{ab}) be a strongly asymptotically predictable spacetime, with globally hyperbolic region V' in the unphysical spacetime. Let $B = M - J^{-}(\mathcal{J}^{+})$ be the black hole region of the spacetime. Let Σ be a Cauchy surface of V'. We shall refer to $\Sigma \cap B$ as the *black hole region at time* Σ. Thus, each connected component of $\Sigma \cap B$ is called *black hole at time* Σ. We then have the following theorem due to Hawking [11]. Let (M, g_{ab}) be a strongly predictable spacetime satisfying $R_{ab}k^a k^b \geq 0$ for all null k^a. Let Σ_1 and Σ_2 be spacelike Cauchy surfaces for the globally hyperbolic region V' with $\Sigma_2 \subset I^{+}(\Sigma_1)$. Let $h_1 = h \cap \Sigma_1$ and $h_2 = h \cap \Sigma_2$, where h is the event horizon. Then the area of h_2 is greater than or equal to the area of h_1.

To conclude this section, we shall present a striking property of event horizons of stationary spacetimes due to Hawking, known as the *strong rigidity theorem*. Let (M, g_{ab}) be a strongly asymptotically predictable stationary spacetime that satisfies Einstein's equation such that the metric and the matter fields are analytic. Then, the event horizon is a Killing horizon.

Once explored the key aspects of general relativity relevant to understanding the information loss issue, the next step is to discuss the framework of quantum field theory, where the principles and tools outlined in this chapter will find their application.

Chapter 3

Quantum field theory in curved spacetimes

In flat spacetime, it is usual to use the plane-wave decomposition of the field to construct the Fock space representation of the canonical commutation relations of our quantum field. However, even in this simple case, this construction is not unique, leading to different particle contents in the same state [1, 12]. In curved spacetimes, the situation is more dramatic, since the existence of solutions with time dependence $e^{\pm i\omega t}$ is only guaranteed if the spacetime is stationary. In the absence of timelike symmetry, there is no way to favor one Fock representation over any other and, thus, we have to abandon the particle concept. However, the theory must be well-posed even in this type of broader case. In this section, we review the formulation of the quantum field theory in curved spacetime without reference to the concept of particles.

3.1 Canonical quantization in curved spacetimes

Consider a globally hyperbolic spacetime (M, g_{ab}). Let $\mathcal{L}$ be the Lagrangian density describing the classical theory and $\{x^\mu\}$ some coordinate system for the spacetime. The equation of motion of the field $\phi(x^\mu)$ is given via the extremization of the action

$$S[g_{\mu\nu}, \phi] = \int_M d^4x \sqrt{-g}\mathcal{L}(\phi, \nabla_\mu \phi), \tag{3.1}$$

which gives

$$\frac{\partial \mathcal{L}}{\partial \phi} - \nabla_\mu \left(\frac{\partial \mathcal{L}}{\partial(\nabla_\mu \phi)} \right) = 0. \tag{3.2}$$

This is the well-known *Euler-Lagrange equation*. Now, consider the scalar field $\phi(x^\mu)$ described by the Lagrangian

$$\mathcal{L} = -\frac{1}{2} \left[\nabla^\mu \phi \nabla_\mu \phi + (m^2 + \xi R)\phi^2 \right], \tag{3.3}$$

where $\xi \in \mathbb{R}$ is a constant and $R = R_a^{\ a}$ is the Ricci scalar, which leads to the *Klein-Gordon equation*

$$-\Box\phi + (m^2 + \xi R)\phi = 0, \tag{3.4}$$

where $\Box \equiv \nabla_a \nabla^a$ is the d'Alambertian operator.

The well-posedness of Eq. (3.4) in globally hyperbolic spacetimes follows from the following theorem [10]: Let (M, g_{ab}) be a globally hyperbolic spacetime with smooth, spacelike Cauchy surface Σ. Then the Klein-Gordon equation has a well-posed initial value formulation in the following sense. Given any pair of C^∞ functions $(\phi_0, \dot{\phi}_0)$ on Σ, there exists a unique solution, ϕ, to Eq. (3.4), defined on all of M, such that on Σ we have $\phi = \phi_0$ and $n^a \nabla_a \phi = \dot{\phi}_0$, where n^a denotes the unit, future-directed, normal to Σ. Furthermore, for any closed subset $S \subset \Sigma$, the solution ϕ restricted to $D(S)$ depends only upon the initial data on S. In addition, ϕ is smooth and varies continuously with the initial data.

Now, Eq. (3.4) can be cast in a more suitable form. Consider, first, the term $\Box\phi = \nabla_\mu \nabla^\mu \phi$. Since ϕ is a scalar, $\nabla^\mu \phi = \partial^\mu \phi$. Next, let $v^\mu \equiv \partial^\mu \phi$. Then,

$$\begin{aligned} \Box\phi &= \nabla_\mu v^\mu \\ &= \partial_\mu v^\mu + \Gamma^\mu_{\mu\alpha} v^\alpha, \end{aligned}$$

where

$$\begin{aligned} \Gamma^\mu_{\mu\alpha} &= \frac{1}{2} g^{\mu\beta} \left(\partial_\mu g_{\beta\alpha} + \partial_\alpha g_{\mu\beta} - \partial_\beta g_{\mu\alpha} \right) \\ &= \frac{1}{2} \left(\partial^\beta g_{\beta\alpha} + g^{\mu\beta} \partial_\alpha g_{\mu\beta} - \partial^\mu g_{\mu\alpha} \right) \\ &= \frac{1}{2} g^{\mu\beta} \partial_\alpha g_{\mu\beta}. \end{aligned}$$

Now, note that if A is a matrix, we have

$$Tr(A^{-1} \partial_\alpha A) = \partial_\alpha \ln |A|,$$

where $|A|$ is the determinant of A. Using this result for the metric tensor, we have

$$g^{\mu\nu} \partial_\alpha g_{\mu\nu} = \partial_\alpha \ln(-g).$$

Applying this result in the expression for $\Gamma^{\mu}_{\mu\alpha}$

$$\begin{aligned}\Gamma^{\mu}_{\mu\alpha} &= \frac{1}{2}\partial_\alpha \ln(-g) \\ &= \partial_\alpha \ln\sqrt{-g} \\ &= \frac{1}{\sqrt{-g}}\partial_\alpha\sqrt{-g}.\end{aligned}$$

Thus, the d'Alambertian reads

$$\Box\phi = \partial_\mu\partial^\mu\phi + \frac{1}{\sqrt{-g}}(\partial_\mu\sqrt{-g})\partial^\mu\phi.$$

Finally, see that

$$\partial_\mu\left[\sqrt{-g}\partial^\mu\phi\right] = (\partial^\mu\phi)\partial_\mu\sqrt{-g} + \sqrt{-g}\partial_\mu\partial^\mu\phi$$

Then, we have

$$\Box\phi = \frac{1}{\sqrt{-g}}\partial_\mu\left[\sqrt{-g}\partial^\mu\phi\right].$$

Therefore, in general coordinates, the Klein-Gordon equation reads

$$-\frac{1}{\sqrt{-g}}\partial_\mu\left[\sqrt{-g}g^{\mu\nu}\partial_\nu\phi\right] + (m^2 + \xi R)\phi = 0. \tag{3.5}$$

Concerning the canonical quantization of the field ϕ, we need to define the momentum π associated with it. For this purpose, let Σ be a Cauchy surface for (M, g_{ab}). Let $t \in \mathbb{R}$ be the label of the different Cauchy surfaces used to foliate the spacetime, i.e., $\Sigma_t = \{t\} \times \Sigma$. Let n^a be the normal, unit vector field to each Σ_t and let $h_{ab} = g_{ab} + n_a n_b$ be the metric induced on each Σ_t. Thus, we define the canonical conjugate momentum π by

$$\pi = \frac{\partial\mathcal{L}}{\partial(n^\mu\nabla_\mu\phi)} = n^\mu\nabla_\mu\phi. \tag{3.6}$$

The canonical quantization follows from promoting ϕ and π to operators $\hat{\phi}$

and $\hat{\pi}$ and imposing the *canonical commutation relations*

$$\begin{aligned} \left[\hat{\phi}(p), \hat{\pi}(q)\right]\Big|_{\Sigma_t} &= \delta^{(3)}_{\Sigma_t}(p,q) i\hbar, \\ \left[\hat{\phi}(p), \hat{\phi}(q)\right]\Big|_{\Sigma_t} = \left[\hat{\pi}(p), \hat{\pi}(q)\right]\Big|_{\Sigma_t} &= 0, \end{aligned} \tag{3.7}$$

where $p, q \in \Sigma_t$ and $\delta^{(3)}$ is the Dirac-delta distribution on Σ_t

$$\delta^{(3)}_{\Sigma_t}(p,q) \equiv \frac{\delta(x^1(p) - x^1(q))\delta(x^2(p) - x^2(q))\delta(x^3(p) - x^3(q))}{\sqrt{h}}, \tag{3.8}$$

with x^μ being an arbitrary coordinate system for Σ_t and h the determinant of h_{ab} in these coordinates.

The construction of a representation of the canonical commutation in terms of the Fock space is given as follows. Let $\mathcal{S}^{\mathbb{C}}$ be the space of complex solutions of Eq. (3.4). Define, next, the following form on $\mathcal{S}^{\mathbb{C}}$

$$(\phi_1, \phi_2)_{\text{KG}} \equiv i \int_{\Sigma_t} d^3x \sqrt{h}\, n^\mu \left(\phi_1^* \nabla_\mu \phi_2 - \phi_2 \nabla_\mu \phi_1^*\right) = i \int_{\Sigma_t} d^3x \sqrt{h}\left(\phi_1^* \pi_2 - \phi_2 \pi_1^*\right), \tag{3.9}$$

called *Klein-Gordon form*. It follows directly from the above definition that

1. $(\phi_1 + \phi_2, \phi_3 + \phi_4)_{\text{KG}} = (\phi_1, \phi_3)_{\text{KG}} + (\phi_1, \phi_4)_{\text{KG}} + (\phi_2, \phi_3)_{\text{KG}} + (\phi_2, \phi_4)_{\text{KG}}$,
2. $(\alpha\phi_1, \beta\phi_2)_{\text{KG}} = \alpha^*\beta(\phi_1, \phi_2)$,
3. $(\phi_1, \phi_2)^*_{\text{KG}} = (\phi_2, \phi_1)_{\text{KG}} = -(\phi_1^*, \phi_2^*)_{\text{KG}}$,

where $\phi_n \in \mathcal{S}^{\mathbb{C}}$, $n = 1, 2, 3, 4$ and $\alpha, \beta \in \mathbb{C}$. Then, we say that $(\cdot, \cdot)_{\text{KG}}$ is a *sesquilinear map* on $\mathcal{S}^{\mathbb{C}}$. One can also show that Eq. (3.9) is independent on t, i.e., independent of the choice of the Cauchy surface. To see this, consider the quantity

$$J_\mu \equiv \phi_1^* \nabla_\mu \phi_2 - \phi_2 \nabla_\mu \phi_1^*. \tag{3.10}$$

Note that $\nabla_\mu J^\mu = 0$

$$\begin{aligned} \nabla_\mu J^\mu &= \nabla_\mu \phi_1^* \nabla^\mu \phi_2 + \phi_1^* \Box \phi_2 - \nabla_\mu \phi_2 \nabla^\mu \phi_1^* - \phi_2 \Box \phi_1^* \\ &= (m^2 + \xi R)\phi_1^* \phi_2 - (m^2 + \xi R)\phi_2 \phi_1^* \\ &= 0, \end{aligned}$$

where we have used Eq. (3.4) in the second line. Next, let $U \subset M$ be a region of the spacetime. Then, the Gauss's theorem reads

$$\int_{\partial U} d^3x\sqrt{h}\, n^\mu J_\mu = \int_U d^4x\sqrt{g}\nabla_\mu J^\mu = 0, \tag{3.11}$$

where n^μ is a normal unit vector to the boundary ∂U of U and h is the determinant of the induced metric on ∂U. Let ∂U be composed of two Cauchy surfaces Σ_1 and Σ_2 and a spatial volume V. Let $n_1^\mu = n^\mu$ and $n_2^\mu = -n^\mu$. Then, we have

$$\int_{\partial U} d^3x\sqrt{h}\, n^\mu J_\mu = \int_{\Sigma_1} d^3x\sqrt{h}\, n^\mu J_\mu - \int_{\Sigma_2} d^3x\sqrt{h}\, n^\mu J_\mu + \int_V d^3x\sqrt{h}\, n^\mu J_\mu = 0.$$

Assuming that J^μ decays rapidly near spatial infinity, we have

$$\int_{\Sigma_1} d^3x\sqrt{h}\, n^\mu J_\mu = \int_{\Sigma_2} d^3x\sqrt{h}\, n^\mu J_\mu$$

which is the same as

$$(\phi_1, \phi_2)_{\mathrm{KG}}\Big|_{\Sigma_1} = (\phi_1, \phi_2)_{\mathrm{KG}}\Big|_{\Sigma_2}. \tag{3.12}$$

Therefore, $(\cdot, \cdot)_{\mathrm{KG}}$ is independent of the choice of the Cauchy surface.

Next, we can use the canonical commutation relations to show that, for $\phi_1, \phi_2 \in \mathcal{S}^{\mathbb{C}}$, we have

$$[(\phi_1, \hat{\phi})_{\mathrm{KG}}, (\hat{\phi}, \phi_2)_{\mathrm{KG}}] = \hbar(\phi_1, \phi_2)_{\mathrm{KG}}. \tag{3.13}$$

To see this, we simply evaluate the left-hand side, using the fact that $\phi(x)$ is real along with Eq. (3.7)

$$\begin{aligned}
[(\phi_1, \hat{\phi})_{\mathrm{KG}}, (\hat{\phi}, \phi_2)_{\mathrm{KG}}] &= -\int d^3x d^3x'\, h\, [\phi_1^*(x)\hat{\pi}(x) - \hat{\phi}(x)\pi_1^*(x), \hat{\phi}(x')\pi_2(x') - \phi_2(x')\hat{\pi}(x')] \\
&= -\int d^3x d^3x'\, h\, (\phi_1^*(x)\pi_2(x')[\hat{\pi}(x), \hat{\phi}(x')] - \phi_1^*(x)\phi_2(x')[\hat{\pi}(x), \hat{\pi}(x')] \\
&\qquad - \pi_1^*(x)\pi_2(x')[\hat{\phi}(x), \hat{\phi}(x')] + \pi_1^*(x)\phi_2(x')[\hat{\phi}(x), \hat{\pi}(x')]) \\
&= i\hbar \int d^3x d^3x'\, \sqrt{h}\delta^{(3)}(x - x')(\phi_1^*(x)\pi_2(x') - \phi_2(x')\pi_1^*(x)) \\
&= i\hbar \int d^3x\, \sqrt{h}(\phi_1^*(x)\pi_2(x) - \phi_2(x)\pi_1^*(x)) \\
&= \hbar(\phi_1, \phi_2)_{\mathrm{KG}}.
\end{aligned}$$

This identity would be important in the establishment of the annihilation and creation operators soon.

The space $\mathcal{S}^{\mathbb{C}}$ can be decomposed as a direct sum of two subspaces,

$$\mathcal{S}^{\mathbb{C}} = \mathcal{S}^{\mathbb{C}}_{+} \oplus \mathcal{S}^{\mathbb{C}}_{-},$$

where on $\mathcal{S}^{\mathbb{C}}_{+}$ the Klein-Gordon form is positive definite and on $\mathcal{S}^{\mathbb{C}}_{-}$ is negative definite. Now, let $\{f_\alpha\}$ be complete set of solutions of $\mathcal{S}^{\mathbb{C}}_{+}$ and $\{f^*_\alpha\}$ of $\mathcal{S}^{\mathbb{C}}_{-}$ such that

$$\begin{aligned} (f_\alpha, f_\beta)_{\mathrm{KG}} &= -(f^*_\alpha, f^*_\beta)_{\mathrm{KG}} = \delta(\alpha - \beta), \\ (f_\alpha, f^*_\beta)_{\mathrm{KG}} &= 0. \end{aligned} \tag{3.14}$$

Solutions in $\mathcal{S}^{\mathbb{C}}_{+}$ are called *positive-norm solution* while the ones in $\mathcal{S}^{\mathbb{C}}_{-}$, *negative-norm solution*. The subspace $\mathcal{S}^{\mathbb{C}}_{+}$ can be made a Hilbert space, $\mathcal{H}$, where $(\cdot, \cdot)_{\mathrm{KG}}$ is the inner product on $\mathcal{H}$. Using this representation, we can decompose the quantum field by

$$\hat{\phi}(x) = \int d\alpha \left[\hat{a}_\alpha f_\alpha(x) + \hat{a}^\dagger_\alpha f_\alpha(x)^*\right], \tag{3.15}$$

where $\hat{a}_\alpha$ is an operator defined on $\mathcal{S}^{\mathbb{C}}_{+}$ and $\hat{a}^\dagger$ is its adjoint. It follows directly from the orthonormality conditions of (f_α, f^*_α) that

$$\hat{a}_\alpha = (f_\alpha, \hat{\phi})_{\mathrm{KG}}, \quad \hat{a}^\dagger_\alpha = (\hat{\phi}, f_\alpha)_{\mathrm{KG}}. \tag{3.16}$$

Besides, from the definition, the momentum operator reads

$$\hat{\pi} = \int d\alpha\, n^\mu \left[\hat{a}_\alpha \nabla_\mu f_\alpha + \hat{a}^\dagger_\alpha \nabla_\mu f^*_\alpha\right]. \tag{3.17}$$

Finally, from the canonical commutation relations (3.7) and Eq. (3.13) follow that

$$\begin{aligned} [\hat{a}_\alpha, \hat{a}_\beta] &= [\hat{a}^\dagger_\alpha, \hat{a}^\dagger_\beta] = 0, \\ [\hat{a}_\alpha, \hat{a}^\dagger_\beta] &= \delta(\alpha - \beta). \end{aligned}$$

To see this, we compute, using Eq. (3.13)

$$\begin{aligned} [\hat{a}_\alpha, \hat{a}^\dagger_\beta] &= [(f_\alpha, \hat{\phi})_{\mathrm{KG}}, (\hat{\phi}, f_\beta)_{\mathrm{KG}}] \\ &= (f_\alpha, f_\beta)_{\mathrm{KG}} \\ &= \delta(\alpha - \beta). \end{aligned}$$

The other relations can be derived similarly.

The operators $\hat{a}_\alpha$ and $\hat{a}^\dagger_\alpha$ are called *annihilation* and *creation operators*, respectively.

They define a *vacuum state* $|0\rangle$ characterized by

$$\hat{a}_\alpha |0\rangle = 0, \quad \forall \alpha. \tag{3.18}$$

Note that since the vacuum state is defined by the annihilation operator, which is given by the functions f_α via Eq. (3.16), the choice of f_α satisfying Eq. (3.14) determines the vacuum state. In this way, an important inquiry is to understand how to compare these different vacua for different choices of f_α.

Consider two complete sets of functions $\{f_\alpha, f_\alpha^*\}$ and $\{F'_\alpha, F'^*_\alpha\}$ satisfying conditions (3.14). Since the two sets are complete, we can write one in terms of the other

$$F_{\alpha'} = \int d\alpha \, [c_{\alpha'\alpha} f_\alpha + d_{\alpha'\alpha} f_\alpha^*] \,. \tag{3.19}$$

From the orthonormality conditions, we have

$$\begin{aligned} c_{\alpha'\alpha} &= (f_\alpha, F_{\alpha'})_{\text{KG}} = (F_{\alpha'}, f_\alpha)^*_{\text{KG}} \\ d_{\alpha'\alpha} &= -(f_\alpha^*, F_{\alpha'})_{\text{KG}} = (F^*_{\alpha'}, f_\alpha)_{\text{KG}}. \end{aligned} \tag{3.20}$$

Thus, we can invert Eq. (3.19), obtaining

$$f_\alpha = \int d\alpha' \, [c^*_{\alpha'\alpha} F_{\alpha'} - d_{\alpha'\alpha} F^*_{\alpha'}] \,. \tag{3.21}$$

Then, we can use the two sets to decompose the field $\hat{\phi}$

$$\hat{\phi} = \int d\alpha \left[\hat{a}_\alpha f_\alpha + \hat{a}^\dagger_\alpha f^*_\alpha\right] = \int d\alpha' \left[\hat{b}_{\alpha'} F_{\alpha'} + \hat{b}^\dagger_{\alpha'} F^*_{\alpha'}\right] . \tag{3.22}$$

Using Eq. (3.19) we can write

$$\int d\alpha \left[\hat{a}_\alpha f_\alpha + \hat{a}^\dagger_\alpha f^*_\alpha\right] = \int d\alpha \int d\alpha' \left[f_\alpha \left(c_{\alpha'\alpha} \hat{b}_{\alpha'} + d^*_{\alpha'\alpha} \hat{b}^\dagger_{\alpha'}\right) + f^*_\alpha \left(d_{\alpha'\alpha} \hat{b}_{\alpha'} + c^*_{\alpha'\alpha} \hat{b}^\dagger_{\alpha'}\right)\right] ,$$

from which we have

$$\hat{a}_\alpha = \int d\alpha' \left(c_{\alpha'\alpha} \hat{b}_{\alpha'} + d^*_{\alpha'\alpha} \hat{b}^\dagger_{\alpha'}\right) \tag{3.23}$$

and

$$\hat{a}^\dagger_\alpha = \int d\alpha' \left(c^*_{\alpha'\alpha} \hat{b}^\dagger_{\alpha'} + d_{\alpha'\alpha} \hat{b}_{\alpha'}\right) . \tag{3.24}$$

One can do the same procedure and extract the inverted relations

$$\hat{b}_{\alpha'} = \int d\alpha \left(c^*_{\alpha'\alpha} \hat{a}_\alpha - d^*_{\alpha'\alpha} \hat{a}^\dagger_\alpha\right) \tag{3.25}$$

and

$$\hat{b}^{\dagger}_{\alpha'} = \int d\alpha \left(c_{\alpha'\alpha} \hat{a}^{\dagger}_{\alpha} - d_{\alpha'\alpha} \hat{a}_{\alpha} \right). \tag{3.26}$$

These transformations mixing annihilation and creation operators relative to the two complete sets of modes are called *Bogolubov transformations*, with the coefficients $c_{\alpha'\alpha}$ and $d_{\alpha'\alpha}$ being the *Bogolubov coefficients*. One can show that the Bogolubov coefficients satisfy

$$\int d\alpha \; (c_{\alpha'\alpha} c^{*}_{\alpha''\alpha} - d_{\alpha'\alpha} d^{*}_{\alpha''\alpha}) = \delta(\alpha' - \alpha''), \tag{3.27}$$

$$\int d\alpha \; (c_{\alpha'\alpha} d_{\alpha''\alpha} - d_{\alpha'\alpha} c_{\alpha''\alpha}) = 0. \tag{3.28}$$

The vacuum states corresponding to the two sets of orthonormal basis are distinct if $d_{\alpha'\alpha}$ do not vanish for all α' and α. To see this, let $|0_{\alpha}\rangle$ and $|0_{\alpha'}\rangle$ be the vacua corresponding to f_{α} and $F_{\alpha'}$, respectively. Let $\hat{N}_{\alpha} = \hat{a}^{\dagger}_{\alpha}\hat{a}_{\alpha}$ and $\hat{N}_{\alpha'} = \hat{a}^{\dagger}_{\alpha'}\hat{a}_{\alpha'}$ be the *number operators*. Then, from Eqs. (3.23) and (3.24) it follows that

$$\langle 0_{\alpha'} | \, \hat{N}_{\alpha} \, | 0_{\alpha'} \rangle = \int d\alpha' |d_{\alpha'\alpha}|^2, \tag{3.29}$$

analogously for $\langle 0_{\alpha} | \, \hat{N}_{\alpha'} \, | 0_{\alpha} \rangle$. So we see that although the number of excitations of f_{α} in $|0_{\alpha}\rangle$ is zero, the excitations of $F_{\alpha'}$ in $|0_{\alpha}\rangle$ can be arbitrarily greater than zero, depending exclusively on the value of $d_{\alpha'\alpha}$.

From the previous discussion, we see that the choice of vacuum states depends on the choice of the functions f_{α} used to decompose the field. Since one could have infinite basis functions, one expects no natural choice of vacuum. However, there is a preferred vacuum state in static spacetimes, which can be constructed as follows. Let (M, g_{ab}) be a static spacetime. Let $\hat{\phi}$ be a quantum field decomposed in terms of a complete set $\{f_{\alpha}, f^{*}_{\alpha}\}$ as described above. First, we cast Eq. (2.20) as

$$ds^2 = -V^2(x^k)dt^2 + G_{ij}(x^k)dx^i dx^j \tag{3.30}$$

and note that $\sqrt{-g} = V\sqrt{G}$, where G is the determinant of g_{ij}. Then, the Klein-Gordon equation (3.5) reads

$$\partial^2_t f_{\alpha} = \frac{V}{\sqrt{G}} \partial_i \left[V\sqrt{G} g^{ij} \partial_j f_{\alpha} \right] - (m^2 + \xi R) V^2 f_{\alpha}. \tag{3.31}$$

Then, it is natural to let f_{α} have a t-dependence of the form $e^{-i\omega_{\alpha} t}$, where ω_{α} are positive constants related to the energy of the particle with respect to the Killing vector $\partial^a/\partial t$. Therefore, in a static spacetime, this is the natural choice of f_α, and thus of vacuum state, i.e., the one that preserves the time-translation symmetry. The state defined in this way is called *static vacuum*. Our next task is to show how these non-uniqueness of vacuum states can lead to different particle contents for different families of observers when gravitational collapse comes into play, which leads to the well-known *Hawking effect*.

3.2 Black hole evaporation: the Hawking effect

We are now in position to derive Hawking radiation [2]. We will concentrate on the case of a massless scalar field in the Schwarzschild spacetime, but the fundamental ideas can be applied to any quantum field in a general black hole spacetime. The physical setup is the following. We consider a black hole formed in the past by gravitational collapse as illustrated in Fig. 3.1. In addition, we consider a massless scalar field $\hat{\phi}$ prepared, in the asymptotic past, in the vacuum state $|0\rangle_{\mathcal{J}^-}$. The changing background geometry can lead to particle creation, as was the case in expanding universes [13]. In this way, our task is to determine how the in-vacuum is perceived by asymptotic observers in the future.

When no horizon is present, all the ingoing modes that left $\mathcal{J}^-$ propagate out to $\mathcal{J}^+$. However, when a horizon is present, part of the modes from the asymptotic past can propagate into the black hole, eventually reaching the spacelike singularity and, thus never reaching the future null infinity. In other words, all null geodesics have a past endpoint at $\mathcal{J}^-$, but their future endpoint can be the central singularity or $\mathcal{J}^+$. Hence, a complete description of the quantum state after the formation of the black hole is made in terms of modes that reach $\mathcal{J}^+$ and the ones that go into the black hole. In this manner, since all modes that fell into the black hole have crossed the event horizon, it is natural to describe these modes in terms of solutions of the Klein-Gordon equation at the event horizon. However, since the Killing vector field is null at the event horizon, no preferred set of orthonormal modes at the event horizon exists. Nonetheless, the quantum field can be decomposed in terms of two sets of solutions of the Klein-Gordon equation: the modes that represent "event–horizon particles states" according to "observers at the horizon" and modes that represent "asymptotic–observers particles states" according to static observers at $\mathcal{J}^+$, in a way similar to the case of the Unruh effect [1]. However, in the case of the Unruh effect, both sets of solutions have an unambiguous, physically well-defined, particle interpretation. On the other hand, the physical interpretation of the "event–horizon particles states" is not clear. First of all, as already discussed, since the Killing vector field at the horizon is null, there is no way to favor one Fock representation over any other. Besides, there is no such thing as "an observer at the horizon", since the horizon is a null surface and physical observers travel in timelike trajectories, so it is not even clear who would detect such particles and via what kind of measurements. Therefore, one is primarily interested in the results of measurements made by static observers in the asymptotic future [14, 15].

Now that the physical situation is established, we can translate it into the mathematical framework developed in the last section. Let $\{f_{\omega'}, f^*_{\omega'}\}$ be a complete set at $\mathcal{J}^-$ for the quantum field $\hat{\phi}$. From previous discussion, the in-modes, $f_{\omega'}$, that are purely positive frequency in $\mathcal{J}^-$ are of the form $\sim e^{-i\omega' t}$ at $\mathcal{J}^-$, where t is the proper time of static asymptotic observers in the past. Let $\hat{a}_{\omega'}$ be the anihilation operator associated with $f_{\omega'}$. Hence, the field can be decomposed as

$$\hat{\phi}(x) = \int d\omega' \left(\hat{a}_{\omega'} f_{\omega'}(x) + \hat{a}^\dagger_{\omega'} f^*_{\omega'}(x)\right). \tag{3.32}$$

In a similar manner, let $\{F_\omega, F^*_\omega\}$ be a orthonormal set at $\mathcal{J}^+$ for $\hat{\phi}$. The out-modes, F_ω, that are purely positive frequency in $\mathcal{J}^+$, are of the form $\sim e^{-i\omega t'}$, where t' is the proper time for static asymptotic observers in the future. We have seen that the complete decomposition of the field is made in terms of $\{F_\omega, F^*_\omega\}$ and the modes at the event horizon. Thus, let $\{u_\omega, u^*_\omega\}$ be a orthonormal set at the event horizon. In addition, let $\hat{b}_\omega$ and $\hat{c}_\omega$ be, respectively, the anihilation operator associated with F_ω and u_ω. Hence, the field can be decomposed as

$$\hat{\phi}(x) = \int d\omega \left(\hat{b}_\omega F_\omega(x) + \hat{b}^\dagger_\omega F^*_\omega(x) + \hat{c}_\omega u_\omega(x) + \hat{c}^\dagger_\omega u^*_\omega(x)\right). \tag{3.33}$$

As commented above, we are interested in measurements made by static observers in the asymptotic future, so we disregard any considerations from the set $\{u_\omega, u^*_\omega\}$.

Since the set $\{f_{\omega'}, f^*_{\omega'}\}$ is complete, we can use it to decompose the modes at $\mathcal{J}^+$ as

$$F_\omega = \int d\omega' \left(c_{\omega\omega'} f_{\omega'} + d_{\omega\omega'} f^*_{\omega'}\right). \tag{3.34}$$

Thus, the spectrum of created particles according to observers at $\mathcal{J}^+$ is computed from Eq. (3.29) to be

$$_{\mathcal{J}^-}\langle 0| \hat{b}^\dagger_\omega \hat{b}_\omega |0\rangle_{\mathcal{J}^-} = \int d\omega' \, |d_{\omega\omega'}|^2. \tag{3.35}$$

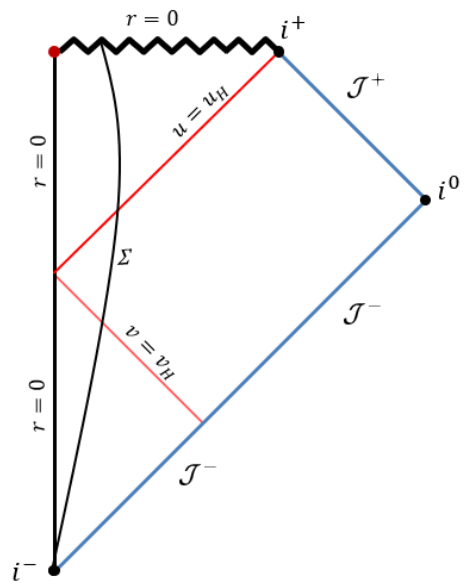


Figure 3.1: Carter-Penrose diagram for the spacetime of a collapsing spherical matter Σ. The line $u = u_H$ represents the event horizon.

Our next task is to compute the Bogolubov coefficients $c_{\omega\omega'}$ and $d_{\omega\omega'}$ to evaluate Eq. (3.35). The massless Klein-Gordon equation with $\xi = 0$ in Schwarzschild spacetime reads

$$\begin{aligned}\left(1-\frac{2M}{r}\right)^{-1}\frac{\partial^2\phi}{\partial t^2}-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\left(1-\frac{2M}{r}\right)\frac{\partial\phi}{\partial r}\right)&\\ -\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right)-\frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\phi^2}&=0,\end{aligned} \tag{3.36}$$

where we have used Eq. (3.5). Due to the time and spherical symmetry of the metric, we use as an *ansatz* for the normal modes, the functions

$$f_{\omega' lm} \sim \frac{1}{\sqrt{\omega'}} Y_{lm}(\theta,\phi) e^{-i\omega' t}\frac{g(r)}{r}, \tag{3.37}$$

where $Y_{lm}(\theta,\phi)$ are the usual spherical harmonics and $g(r)$ satisfies

$$\begin{aligned}\omega'^2\left(1-\frac{2M}{r}\right)^{-1}g(r)+\frac{2}{r}\left(1-\frac{2M}{r}\right)\left(\partial_r g(r)-\frac{g(r)}{r}\right)&\\ +\frac{2M}{r^2}\left(\partial_r g(r)-\frac{g(r)}{r}\right)+\left(1-\frac{2M}{r}\right)\left(\frac{2g(r)}{r^2}-\frac{2\partial_r g(r)}{r}+\partial_r^2 g(r)\right)&\\ -\frac{l(l+1)g(r)}{r^2}&=0.\end{aligned} \tag{3.38}$$

Note that here $M = \text{const}$ since we are in the asymptotic region.

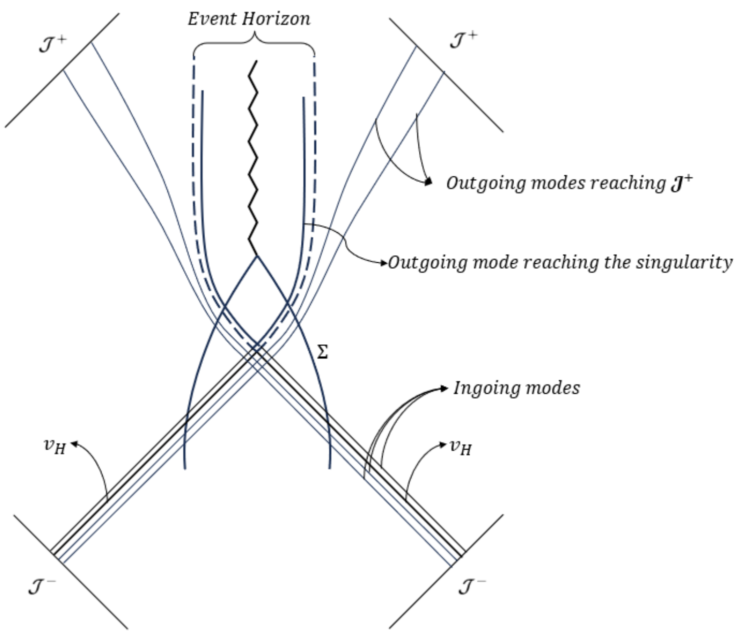


Figure 3.2: Illustration of the ingoing modes propagating to $\mathcal{J}^+$ or into the singularity in the spacetime of gravitational collapse of a body Σ. These modes are very high blue-shifted close to the event horizon. The limit ingoing ray is given by $v = v_H$.

The above equation has no closed form. However, as discussed above, we are only interested in three sets of solutions: the one at $\mathcal{J}^-$ (given by $t \to -\infty$ and $r \to \infty$), the one at $\mathcal{J}^+$ (given by $t \to \infty$ and $r \to \infty$), and the one at the event horizon ($r = 2M$). For static observers in the asymptotic past, we have

$$\partial_r^2 g(r) + \omega'^2 g(r) = 0. \tag{3.39}$$

We choose as solution $g(r) = e^{-i\omega' r}$ because we are interested in ingoing modes, i.e., the ones propagating from $r \gg 2M$ to $r = 0$. Therefore, the asymptotic form of the past modes is

$$f_{\omega' lm} \sim \frac{1}{\sqrt{\omega'}} Y_{lm}(\theta, \phi) \frac{e^{-i\omega' v}}{r}, \tag{3.40}$$

where $v = t + r$. The same reasoning with respect to the infinite future ($t \to +\infty$) gives us the asymptotic form of the future modes

$$F_{\omega lm} \sim \frac{1}{\sqrt{\omega}} Y_{lm}(\theta, \phi) \frac{e^{-i\omega u}}{r}, \tag{3.41}$$

where $u = t - r$.

Now, consider the out-mode propagating backward from $\mathcal{J}^+$. Part of the solution will be scattered and reach $\mathcal{J}^-$ with a frequency similar to the one it had when it was "emitted" in $\mathcal{J}^+$. The other part will pass through the center of the collapsing body and emerge as an incoming solution, reaching $\mathcal{J}^-$ with a frequency much greater than the one it had when left $\mathcal{J}^+$. This comes from the fact that these solutions pile up just outside the event horizon (see Fig. 3.2). In other words, $F_{\omega lm}$ suffers a very high blue shift near the event horizon. In this way, $F_{\omega lm}$ propagates by geometric optics back through the collapsing body and out to $\mathcal{J}^-$. However, we can see that there is a limiting value for v, namely, $v = v_H$ such that ingoing rays with $v > v_H$ do not escape the black hole and eventually reach the singularity. Thus, our out-mode $F_{\omega lm}$ is valid only for $v < v_H$, i.e.,

$$F_{\omega lm} \sim \frac{1}{\sqrt{\omega}} Y_{lm}(\theta, \phi) \frac{e^{-i\omega u(v)}}{r} \Theta(v_H - v), \tag{3.42}$$

where $\Theta(v_H - v)$ is the Heaviside function. All we need to know now is how an ingoing geodesic coming from $\mathcal{J}^-$ emerges as an outgoing one after passing through the center of the collapsing body.

Consider radial null geodesics of the Schwarzschild spacetime. The equation describing these curves is given by

$$\left(1 - \frac{2M}{r}\right) \frac{dt}{d\zeta} = \pm \frac{dr}{d\zeta}, \tag{3.43}$$

where ζ is the affine parameter along radial null geodesics and the plus (resp. minus) sign is along outgoing (resp. ingoing) ones. Let E be the conserved quantity associated with the Killing vector ∂_t^a

$$\left(1 - \frac{2M}{r}\right) \frac{dt}{d\zeta} = E. \tag{3.44}$$

Hence, Eq. (3.43) reads

$$\frac{dr}{d\zeta} = \pm E, \tag{3.45}$$

where again the plus sign is along outgoing null geodesics and the minus sign is along ingoing null ones. Therefore, we have

$$r(\zeta) = \pm E\zeta + a_{\pm}, \quad a_{\pm} \in \mathbb{R}. \tag{3.46}$$

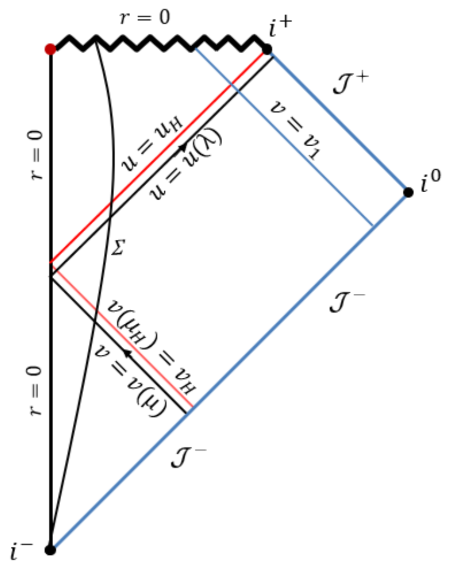


Figure 3.3: Carter-Penrose diagram for the spacetime of a collapsing spherical matter Σ. An ingoing null ray with $v < v_H$ from $\mathcal{J}^-$ passes through Σ, it is reflected at $r = 0$ and escapes to $\mathcal{J}^+$ as a $u = u(v)$ outgoing null ray. Ingoing rays with $v > v_H$ do not escape after reflecting at $r = 0$ and reach the spacelike singularity. The line $u = u_H$ represents the event horizon.

Consider now the ingoing ray $v = v_1 = \text{const}$ that crosses the event horizon, as in Fig. 3.3. We want to find $u(\lambda)$ along this geodesic, where $u \equiv t - r^*$ and we choose $\zeta = \lambda$ as the affine parameter. Choosing $r(\lambda = 0) = 2M$ and using Eq. (3.46), we have

$$r(\lambda) = -E\lambda + 2M. \tag{3.47}$$

Hence, from Eq.(3.44) and the fact that $dr^*/dr = (1 - 2M/r)^{-1}$ we have

$$\begin{aligned}\frac{du}{d\lambda} &= \frac{dt}{d\lambda} - \frac{dr^*}{d\lambda} \\ &= \frac{dt}{d\lambda} - \frac{dr^*}{dr}\frac{dr}{d\lambda} \\ &= \frac{2E}{1 - \frac{2M}{r(\lambda)}} \\ &= -\frac{4M}{\lambda} + 2E,\end{aligned} \tag{3.48}$$

where we have used Eq. (3.47). Therefore,

$$u(\lambda) = 2E\lambda - 4M \log \lambda/b, \quad b < 0. \tag{3.49}$$

The condition $b < 0$ comes from the fact that $\lambda < 0$ outside the black hole. Close to the event horizon, $\lambda \approx 0$, we have

$$u(\lambda) \approx -4M \log \lambda / b. \tag{3.50}$$

This is the equation describing all the outgoing geodesics close to the event horizon that intersect the ray $v = v_1$.

We can use Eq. (3.50) to find the function $u(v)$, where v is the ingoing ray that emerges when $u = \text{const}$ is reflected at $r = 0$, using the following reasoning. Consider a fixed value of λ along $v = v_1$, that is, choose any outgoing ray $u(\lambda)$ that crosses $v = v_1$. If we take the integration constant b to be the same along any ingoing geodesic that crosses $u(\lambda)$, then λ will describe the same outgoing ray along all such geodesics. In other words, fixing λ along any ingoing geodesic that crosses the event horizon also fixes the corresponding outgoing ray. Since this is valid for any outgoing geodesic, the same holds for the event horizon, which is given by $\lambda = 0$. Therefore, the affine distance between any outgoing ray $u(\lambda)$ and the event horizon is given simply by λ. Since λ is the same along any ingoing geodesic crossing $u(\lambda)$, this separation remains constant along the entire length of $u(\lambda)$ and the event horizon, extending all the way to $\mathcal{J}^+$.

Now, along $u(\lambda)$, the coordinate v is a function of the affine parameter $\zeta = \mu$, so we write $v = v(\mu)$. Next, let μ and μ_H be, respectively, the value of the affine parameter corresponding to the reflected ingoing ray $v(\mu)$ and the generator of the event horizon v_H, see Fig. 3.3. As with $u(\lambda)$, we can always choose the integration constant of $v(\mu)$ to be the same along any outgoing ray crossing $v(\mu)$ such that the value of μ is the same along all outgoing rays that cross $v(\mu)$. The same holds for v_H. Hence, the affine separation $\mu - \mu_H$ remains constant along $v(\mu)$ and v_H, extending all the way to $\mathcal{J}^-$. Let α be such value, i.e., $\mu - \mu_H = \alpha$. If we rescale the affine parameter μ by

$$\begin{aligned} \mu &\to a\mu + b, \\ \mu_H &\to a\mu_H + b, \end{aligned}$$

we end up with

$$\mu - \mu_H \to a\alpha.$$

Remembering that any affine function of an affine parameter is still an affine parameter, we can always choose a such that

$$\mu - \mu_H = \lambda, \tag{3.51}$$

i.e., the affine separation is preserved when ingoing geodesics are turned into outgoing ones. Using Eqs. (3.44) and (3.46) at $\mathcal{J}^-$ $(u \to -\infty)$ we have

$$\begin{aligned} t(\mu) &= E\mu + c_1 \\ r(\mu) &= E\mu + c_2, \end{aligned} \tag{3.52}$$

where c_1, c_2 are constants. Finally, using that $v \approx t + r$ at $\mathcal{J}^-$ we have

$$v(\mu) = 2E\mu + c_3, \quad c_3 \in \mathbb{R}. \tag{3.53}$$

Hence, $v - v_H = 2E(\mu - \mu_H)$, which gives

$$\lambda = \frac{1}{2E}(v - v_H). \tag{3.54}$$

We can now cast the above equation as $v_H - v = C_1\lambda$, where $C_1 < 0$ since $\lambda < 0$. Therefore, close to the horizon where the geometric optic limit holds, we have from Eq. (3.50) that an ingoing light ray with fixed v gives rise to an outgoing light ray with fixed u by

$$u(v) = -4M \log\left(\frac{v_H - v}{C}\right), \quad C \in \mathbb{R}^+, \tag{3.55}$$

where $C \equiv bC_1$. Therefore, we see from Eq. (3.42) that the mode in the asymptotic future, when traced back to $\mathcal{J}^-$ has the form

$$F_{\omega lm} \sim \frac{1}{\sqrt{\omega} r} Y_{lm}(\theta, \phi) e^{4M\omega i \log\left(\frac{v_H - v}{C}\right)} \Theta(v_H - v). \tag{3.56}$$

To obtain the Bogolubov coefficients we must expand $F_{\omega lm}$ in terms of $f_{\omega' lm}$. Due to spherical symmetry, only modes with the same $\{l, m\}$ are coupled. Therefore, our focus will be the ω part of the expansion. From Eq. (3.19) we have

$$F_{\omega lm} = \int_0^{+\infty} d\omega' (c_{\omega\omega' lm} f_{\omega' lm} + d_{\omega\omega' lm} f^*_{\omega' lm}) \tag{3.57}$$

Multiplying both sides by $e^{-i\omega'' v}/2\pi$ and integrating over v using Eq. (3.40) gives

$$\begin{aligned}\int_{-\infty}^{+\infty} dv\,\frac{e^{-i\omega'' v}}{2\pi}F_{\omega lm} &\propto \frac{1}{2\pi}\int_0^{\infty}\frac{d\omega'}{r\sqrt{\omega'}}\left(c_{\omega\omega' lm}\int_{-\infty}^{+\infty} dv\, e^{-iv(\omega''+\omega)} + d_{\omega\omega' lm}\int_{-\infty}^{+\infty} dv\, e^{iv(\omega'-\omega'')}\right)\\ &= \int_0^{\infty}\frac{d\omega'}{r\sqrt{\omega'}}\left(c_{\omega\omega' lm}\delta\left(\omega''+\omega'\right) + d_{\omega\omega' lm}\delta\left(\omega'-\omega''\right)\right)\\ &= \frac{d_{\omega\omega'' lm}}{r\sqrt{\omega''}},\end{aligned}$$

where we have discarded the first term since both ω' and ω'' are positive. Hence, using Eq. (3.56), we have

$$\begin{aligned}\frac{d_{\omega\omega'' lm}}{\sqrt{\omega''}} &\propto \int_{-\infty}^{+\infty} dv\,\frac{e^{-i\omega'' v}}{2\pi}\frac{1}{\sqrt{\omega}}e^{4M\omega i\log\left(\frac{v_H-v}{C}\right)}\Theta(v_H-v)\\ &= \int_{-\infty}^{v_H} dv\,\frac{e^{-i\omega'' v}}{2\pi}\frac{1}{\sqrt{\omega}}e^{4M\omega i\log\left(\frac{v_H-v}{C}\right)}\end{aligned}$$

Proceeding in the same manner, substituting $e^{-i\omega'' v}$ by $e^{i\omega'' v}$, we get $c_{\omega\omega'' lm}$. Therefore, the Bogolubov coefficients for the out-mode (3.57) are given by

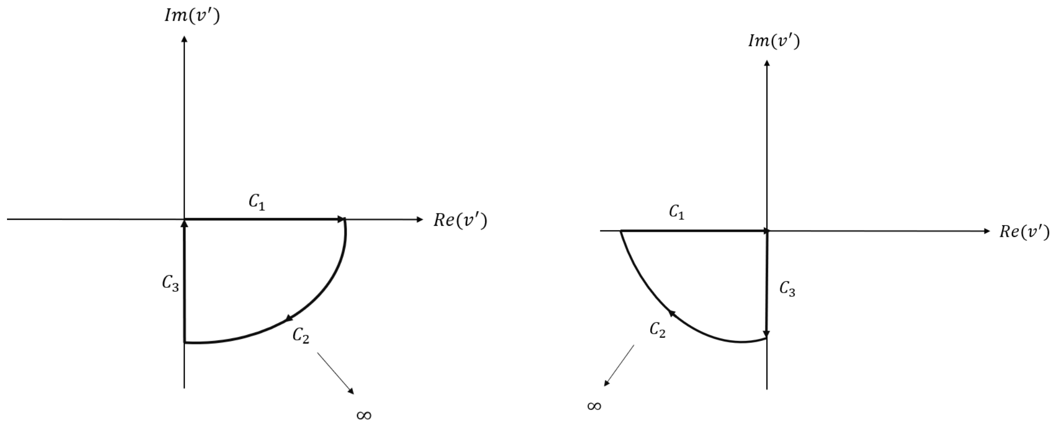
$$\begin{aligned}c_{\omega\omega' lm} &\propto \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-\infty}^{v_H} dv\, e^{i\omega' v}e^{4M\omega i\log\left(\frac{v_H-v}{C}\right)},\\ d_{\omega\omega' lm} &\propto \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-\infty}^{v_H} dv\, e^{-i\omega' v}e^{4M\omega i\log\left(\frac{v_H-v}{C}\right)}.\end{aligned} \tag{3.58}$$

Let us focus on the $c_{\omega\omega' lm}$ integral. Making a change of variables of the form $v' = v_H - v$ leads to

$$c_{\omega\omega' lm} \propto \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\int_0^{+\infty} dv'\, e^{i\omega'(v_H-v')}e^{4M\omega i\log\left(\frac{v'}{C}\right)}.$$

Since the integrand is analytic in the lower right quadrant of the complex plane, we can use a quarter-circle contour with radius R to evaluate the integral, see Fig. 3.4a. For $R\to\infty$, region C_2 does not contribute. Thus, we have

$$c_{\omega\omega' lm} \propto -\frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-i\infty}^{0} dv'\, e^{i\omega'(v_H-v')}e^{4M\omega i\log\left(\frac{v'}{C}\right)}.$$



(a) Contour for $c_{\omega\omega'}$ (b) Contour for $d_{\omega\omega'}$

Figure 3.4: Contour of integration on the complex plane for evaluating the Bogolubov coeficcients.

Making $v' = is$ gives

$$c_{\omega\omega' lm} \propto -\frac{i}{2\pi}\sqrt{\frac{\omega'}{\omega}}\int_{-\infty}^{0} ds\ e^{i\omega'(v_H - is)} e^{4M\omega i \log\left(\frac{is}{C}\right)}.$$

Besides, in the region of integration, we have

$$\log\left(\frac{is}{C}\right) = \log\left(\frac{|s|}{C}\right) - i\pi/2,$$

which makes the integral becomes

$$c_{\omega\omega' lm} \propto -\frac{i}{2\pi}\sqrt{\frac{\omega'}{\omega}} e^{2M\pi\omega} e^{i\omega' v_H}\int_{-\infty}^{0} ds\ e^{\omega' s} e^{4M\omega i \log\left(\frac{|s|}{C}\right)}. \tag{3.59}$$

A similar reasoning applied to $d_{\omega\omega' lm}$, where now we use the contour depicted in Fig. 3.4b, gives

$$d_{\omega\omega' lm} \propto \frac{i}{2\pi}\sqrt{\frac{\omega'}{\omega}} e^{-2M\pi\omega} e^{-i\omega' v_H}\int_{-\infty}^{0} ds\ e^{\omega' s} e^{4M\omega i \log\left(\frac{|s|}{C}\right)}. \tag{3.60}$$

Comparison of these results leads to

$$|c_{\omega\omega' lm}| = e^{4\pi M\omega} |d_{\omega\omega' lm}|. \tag{3.61}$$

We can use this result to evaluate the mean number of particles the observers in the asymptotic future attribute to the vacuum of the observers in the asymptotic past, $|0\rangle_{\mathcal{J}^-}$. Applying Eq. (3.61) in Eq. (3.27) we have

$$\begin{aligned}\lim_{T\to\infty}\lim_{\omega'\to\omega}\frac{1}{2\pi}\int_{-T/2}^{T/2}dt\,e^{it(\omega-\omega')} &= \int d\omega'\,\left(|c_{\omega\omega'}|^2-|d_{\omega\omega'}|^2\right)\\ &= \left(e^{8\pi M\omega}-1\right)\int d\omega'|d_{\omega\omega' lm}|^2.\end{aligned} \tag{3.62}$$

Hence, we have

$$\lim_{T\to\infty}\frac{T}{2\pi}=\left(e^{8\pi M\omega}-1\right)\int d\omega'|d_{\omega\omega' lm}|^2. \tag{3.63}$$

Therefore, from Eq. (3.29), the mean number of particles per unit of proper time for asymptotic observers in the future is given by

$$\langle N_\omega\rangle d\omega=\frac{1}{2\pi}\frac{d\omega}{e^{8\pi M\omega}-1}, \tag{3.64}$$

which is the thermal distribution of a blackbody radiating at a temperature given by

$$T_H=\frac{1}{8\pi M}. \tag{3.65}$$

Essentially, these results translate in the fact that, due to quantum effects, stationary black holes can create and emit particles as if they were hot bodies with temperature given by Eq. (3.65)[1]. This effective particle creation by black holes with thermal spectrum is known as the *Hawking effect*. This particle creation is related to thermal radiation, i.e., an energy flux of particles seen by static observers at future infinity, known as *Hawking radiation*. In fact, if we consider the (renormalized)[2] expectation value of the stress energy tensor, $\langle T_{\mu\nu}\rangle$, in the in-vacuum, i.e., the vacuum according to static observers at past infinity, for the

[1]We emphasize the fact that although black holes emit *as if they were* black bodies, a black hole could not be more different than an ordinary thermodynamical system. See Sec. 4.3 for more details on this issue.

[2]Here, renormalized stands for a renormalization of the cosmological constant, see Ref [16].

case of a spherically collapsing null shell, we end up with

$$\langle T_{uu} \rangle = \frac{1}{24\pi}\left(\frac{3M^2}{2r^4} - \frac{M}{r^3}\right), \tag{3.66}$$

$$\langle T_{vv} \rangle = \frac{1}{24\pi}\left(\frac{3M^2}{2r^4} - \frac{M}{r^3}\right), \tag{3.67}$$

$$\langle T_{uv} \rangle = \langle T_{vu} \rangle = \frac{1}{24\pi}\left(\frac{2M^2}{r^4} - \frac{M}{r^3}\right), \tag{3.68}$$

for times long before the collapse, and

$$\langle T_{uu} \rangle = \frac{1}{24\pi}\left(\frac{3M^2}{2r^4} - \frac{M}{r^3} + \frac{1}{32M^2}\right), \tag{3.69}$$

$$\langle T_{vv} \rangle = \frac{1}{24\pi}\left(\frac{3M^2}{2r^4} - \frac{M}{r^3}\right), \tag{3.70}$$

$$\langle T_{vu} \rangle = \langle T_{uv} \rangle = \frac{1}{24\pi}\left(\frac{2M^2}{r^4} - \frac{M}{r^3}\right). \tag{3.71}$$

for times long after the collapse [16]. Comparison between Eqs. (3.66) and (3.69) reveal that the effect of the collapse is to add a term proportional to M^{-2}, which is precisely the term that makes the energy flux, for long times, at infinity ($r \to \infty$) be non-vanishing and equal to

$$\langle T_{uu} \rangle = \frac{1}{768\pi M^2}, \tag{3.72}$$

Hence, asymptotic observers see a positive energy flux of radiation that they attribute as coming from the black hole. Furthermore, from both Eqs. (3.67) and (3.70), close to the black hole event horizon, the energy flux is given by

$$\langle T_{vv} \rangle = -\frac{1}{768\pi M^2}, \tag{3.73}$$

which is a negative energy flux of radiation going into the black hole.

This negative energy flux makes the area of the black hole shrink at a rate consistent with the energy flux observed at infinity, i.e., the black hole loses mass. One can then ask what the fate of the black hole is due to this mass loss. Due to the radiation being approximately thermal, we can use the Stefan-Boltzmann law to estimate the decrease of M for a chargeless black hole:

$$\frac{dM}{dt} = -\sigma A T_H^4, \tag{3.74}$$

where $\sigma \equiv \pi^2/60$ is the Stefan-Boltzmann constant (in natural units), $A = 16\pi M^2$ is the area of the EH, and t is the proper time of asymptotic static observers. From Eq. (3.65), we have

$$\frac{dM}{dt} = -\frac{M^{-2}}{15,360}, \tag{3.75}$$

which can be integrated to give

$$M^3(t) - M_0^3 = -\frac{\Delta t}{5,120}, \tag{3.76}$$

where $\Delta t \equiv t - t_0$ with M_0 and t_0 being, respectively, the initial mass of the black hole and the initial time of measurement. If we ignore the fact that the semiclassical picture discussed so far can change drastically in the final moments of the evaporation, it follows that the black hole will evaporate in a finite proper time given by

$$\Delta t = 5,120 M_0^3. \tag{3.77}$$

After this time, the final spacetime configuration will be, approximately, that of Minkowski space together with the thermal radiation emitted by the black hole. Since the spectrum of emitted particles is independent of the collapse details, the same conclusions hold for a general collapsing process resulting in a black hole. A spacetime diagram depicting the black hole formation and evaporation process is shown in Fig. 3.5.

3.3 Information loss

The idea that black holes evaporate gives rise to a striking consequence: the loss of quantum coherence. Consider the spacetime depicted in Fig. 3.5. Consider that a *pure* quantum state, $|\phi\rangle$, of a quantum field, ϕ, is prepared at time Σ_1, i.e., before the black hole has formed. Assuming that the laws used to evolve ϕ through the spacetime are unitary, the total state of the quantum field is still pure at some later time Σ_2 after the formation of the black hole. Now, consider the time Σ_3 after the total evaporation of the black hole. From the previous discussion, the Hawking radiation carries a positive-energy flux of particles to $\mathscr{I}^+$, according to

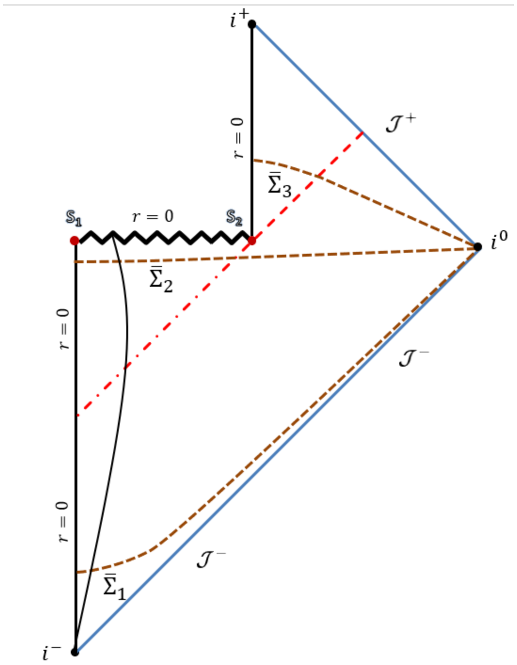


Figure 3.5: Carter-Penrose diagram illustrating the formation and evaporation of a black hole due to Hawking radiation. The surfaces $\overline{\Sigma}_i$, $i = 1, 2, 3$ are spacelike.

asymptotic static observers, to which there is a corresponding negative-energy flux of particles going into the black hole. Hence, the quantum state at $\overline{\Sigma}_3$ will remain entangled with the quantum field observables inside the (now) past black hole even though the black hole no longer exists. Therefore, although the system was initially in a pure state, the *entire* system at later times (after the complete evaporation) is described by a mixed state, so evolution from a pure to a mixed state has occurred. In other words, *information was lost*.

One can also see the information loss from the causal structure of the spacetime. Consider the spacelike surface $\overline{\Sigma}_3$ of Fig. 3.5. One can easily see that $\overline{\Sigma}_3 \not\subset D^+(\overline{\Sigma}_2)$, so $\overline{\Sigma}_3$ is not a Cauchy surface. One could expect that we could not predict what happens to ϕ after the evaporation of the black hole, say, at $\overline{\Sigma}_3$, from complete knowledge of the state of the field at $\overline{\Sigma}_2$. That this does not need to be the case can be illustrated as follows. Consider a massless scalar field ϕ propagating in Minkowski spacetime with an event p removed, see Fig. 3.6. In this case, the spacetime is no longer globally hyperbolic. Thus, apparently, the values of the field inside the future light cone of p would be undetermined. However, the value of the quantum field at p can be determined by continuity, allowing us to evolve ϕ inside the future light cone of p. Hence, even in this simple example of a non-globally hyperbolic spacetime, the field evolution is determined uniquely from the initial data set $\{\phi_0, \dot{\phi}_0\}$ on a spacelike surface ($t = const$, for instance). In

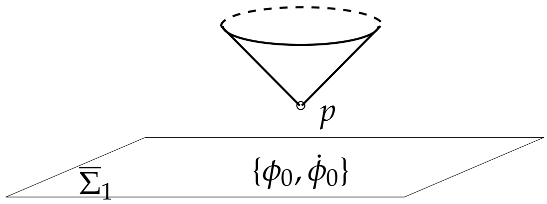


Figure 3.6: A initial data set $\{\phi, \bar{\phi}_0\}$ at a non-Cauchy surface $\overline{\Sigma}_1$ of $t = const$ for the quantum field ϕ in Minkowski spacetime with a point p removed.

our scenario, the event removed is the *naked singularity* S_2 [17]. The semiclassical theory assumes that this naked singularity *does not* influence the spacetime. As a result, from initial data on $\overline{\Sigma}_2$, we can evolve uniquely the quantum field both to the future and to the past. Now, consider the surface $\overline{\Sigma}_3$. Again, this is not a Cauchy surface since $D^-(\overline{\Sigma}_3)$ does not include any portion from inside the black hole. Therefore, the degrees of freedom from the field that fell into the black hole have been destroyed by the spacelike singularity, leaving the outside degrees of freedom in a mixed state. In other words, the spacelike singularity has opened the quantum system. Therefore, once again, we have the evolution from a pure state to a mixed state.

It is clear from the previous discussion that the pure-to-mixed-state evolution comes from the geometrical property of the spacetime of not being globally hyperbolic. However, the exact same sort of phenomenon occurs in a quantum field in flat spacetime if one considers its evolution from an initial Cauchy surface Σ_1 to a final surface $\overline{\Sigma}_2$ which fails to be a Cauchy surface for Minkowski spacetime (see Fig. 3.7). In this case, the quantum field inside $D(\overline{\Sigma}_2)$ will be correlated with radiation propagating out to infinity. Hence, the final state of the (initial pure) quantum state is mixed. Thus, again, there is a loss of quantum coherence, not because the spacetime fails to be globally hyperbolic, but because the final surface considered fails to be a Cauchy one. To physically illustrate this situation, consider that an atom in a laboratory emits a photon at some given instant of proper time $t = 0$. If that photon escapes out the window and is not absorbed (nor reflected) by any matter in the universe, then, as far as we are concerned, it is lost forever. Thus, the laboratory is forever correlated to degrees of freedom that it does not have access to anymore. Therefore, the state of the laboratory is mixed. Hence, any initial pure state in any accessible region around the laboratory will have been converted into a mixed state. In this way, information will be lost.

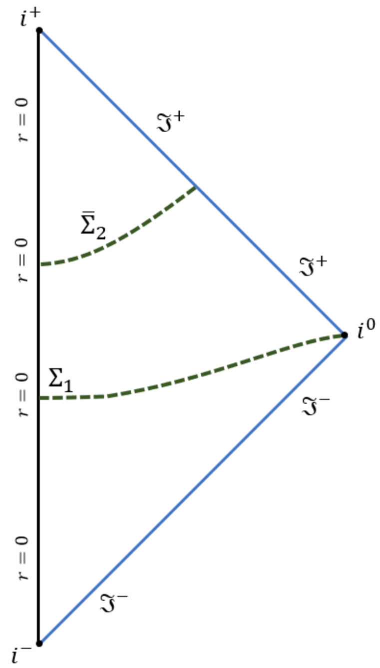


Figure 3.7: Carter-Penrose diagram of Minkowski spacetime depicting a Cauchy surface Σ_1 and a spacelike surface $\overline{\Sigma}_2$ that fails to be a Cauchy surface.

We emphasize that during the process of black hole formation and evaporation, the quantum field obeys well-posed (local unitary and causal) equations of motion. At no stage, except at the singularity, there is a breakdown of any known laws of physics. The information loss comes simply from the fact that the black hole acts as a "sink" for degrees of freedom of quantum fields. In this way, the pure-to-mixed-state evolution in the context of black hole evaporation is a *prediction* of the semiclassical theory and not a *paradox*.

Chapter 4

Globally hyperbolic evaporating black hole and the information loss issue

In the last section, we have discussed how the non-globally hyperbolicity of spacetimes can lead to loss of quantum coherence in the context of black hole evaporation. In this chapter, we discuss the possibility that the usual semiclassical spacetime of Figs. 4.3 and 4.4 is replaced by a globally hyperbolic one preserving all classical and semiclassical physics far from the Planck region. The simplest possibility would be given by a black hole that evaporates as usual until it reaches the Planck size, after which it would freeze out as shown in Fig. 4.2. Here, however, we look for a situation where the black hole completely evaporates in addition to being globally hyperbolic (see Figs. 4.5 and 4.6). To do so, instead of replacing the spacelike singularity of Figs. 4.3 and 4.4 with a super-curved spacetime region as was done in Fig. 4.1, we substitute the naked singularity with a null (thunderbolt) singularity. As a consequence, the corresponding spacetime preserves (i) the event horizon (EH) null surface, (ii) all the expected classical and semiclassical features (including Hawking radiation) far from the Planck region, and (iii) information over the Cauchy surfaces.

It must be stressed, however, that because parts of the late-time Cauchy surfaces are inside the EH (see Σ_2 in Figs. 4.5 and 4.6), late-time external observers will only have access to part of the information contained in the asymptotic past (see Σ_1 in Figs. 4.5 and 4.6), vindicating the mixed state of Hawking radiation. This is not any worse than what is found in the more familiar case where a black hole remnant is left at the end, Fig. 4.2, in which case observers outside the EH can always "blame" the black hole for the information they do not have access to. (Nevertheless, in our case, the black hole fully evaporates, as in the usual semiclassical scenario, in the sense that no massive remnant is left over at the end. Note that there are light rays from $\mathscr{I}^-$ that do not enter the horizon.) Whether Figs. 4.5 and 4.6 of a completely evaporating black hole are correct depends on what quantum gravity will say about the “last breath” of a Planck mass black hole.

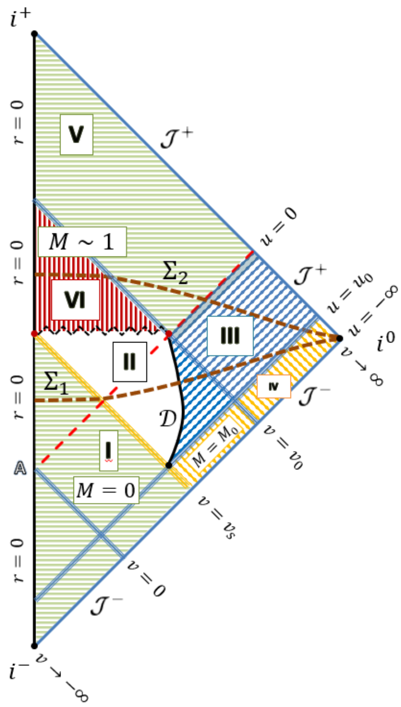


Figure 4.1: Possible conformal Carter-Penrose diagram associated with the total collapse of a spherically symmetric null shell ($v = v_s$) in a quantum gravity scenario. The spacelike singularity is replaced by a regular (yet super-curved spacetime) region. Regions I, II, III, IV, and V are similar to the corresponding ones in the semiclassical picture (although, *stricto-sensu*, there would be no EH). Nevertheless, the collapse would not give rise to a black hole in this scenario since information would leak to the full quantum gravity region VI. The dashed spacelike hypersurfaces Σ_1 and Σ_2 represent Cauchy surfaces of this globally hyperbolic spacetime. Information could be preserved along the whole spacetime depending on the equations that evolve the quantum gravity degrees of freedom.

Instead of considering specific quantum-gravity models (see, e.g., Refs. [18, 19, 20]), we focus on unveiling general conditions to allow for the appearance of the final thunderbolt singularity. (See, e.g., Ref. [21] for other spacetimes investigated irrespective of specific quantum-gravity theories.) We comply with Einstein and semiclassical equations everywhere in the low-curvature regions but not necessarily at the Planck scale, where we only assume the spacetime to possess a Lorentzian metric arbitrarily close to the singularity. This presumes that the forthcoming quantum-gravity theory will provide an effective description of the entire spacetime as a manifold endowed with metric, enabling the discrimination between spacetimes that only differ at the Planck scale.

We must acknowledge the possibility that quantum gravity changes the semi-

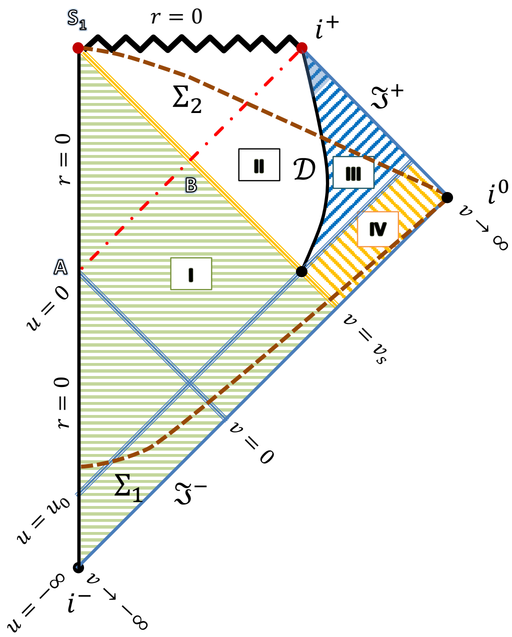


Figure 4.2: Conformal Carter-Penrose diagram associated with the total collapse of a spherically symmetric light null ($v = v_s$) with emission of Hawking radiation. This is a globally hyperbolic spacetime but in contrast to the usual semiclassical picture of Figs. 4.3 and 4.4, a black hole remnant is left out at the end.

classical picture by resolving the singularity, possibly altering the spacetime of Fig. 4.3 as, e.g., the one of Fig. 4.1. By replacing the singularity with a regular, yet super-curved, spacetime region, the quantum-gravity spacetime would gain a full quantum-gravity region, region VI, but would be deprived of the event horizon. Still, the quantum-gravity spacetime should preserve the classical and semiclassical (including Hawking radiation) features far from the Planck scale region. Importantly, the spacetime shown in Fig. 4.1 is supposed to be globally hyperbolic in contrast to the one of Figs. 4.3 and 4.4. We note, however, that even if quantum gravity turns the spacetime globally hyperbolic as in Fig. 4.1, it must comply with other nontrivial requirements to avoid loss of information. For instance, assuming that the black hole thermally evaporates until its mass reaches the Planck scale, quantum gravity should allow the storage of an arbitrarily large number of degrees of freedom in a Planck-mass object that would be entangled with the ones of Hawking radiation.

We start by elaborating on a simple model for the usual non-globally hyperbolic evaporating spacetime, which will be useful to describe some of the machinery used in the globally hyperbolic case. We then discuss how to change the naked singularity by a thunderbolt one and describe the model for the globally hyperbolic scenario.

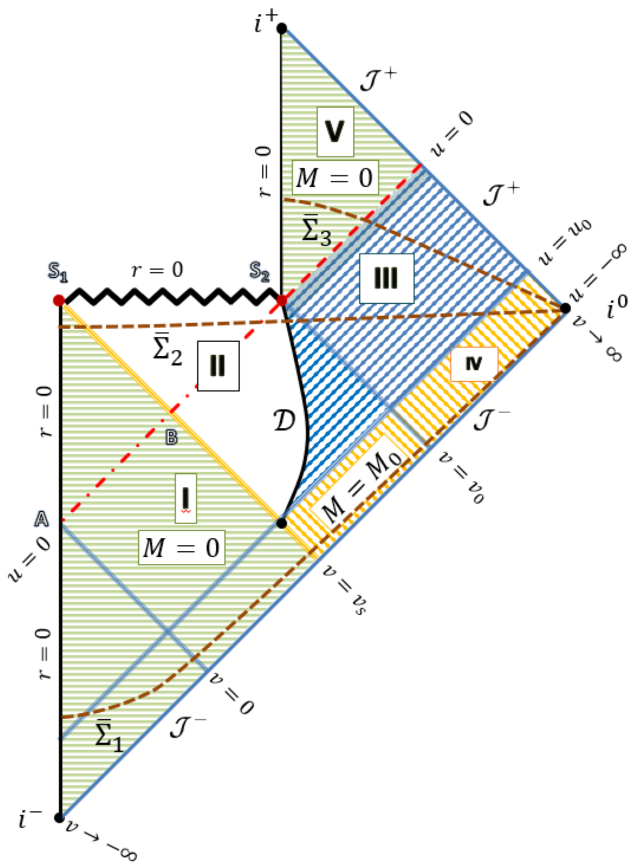


Figure 4.3: Conformal Carter-Penrose diagram associated with the total collapse of a spherically symmetric light shell ($v = v_s$). This leads to a black hole with mass M_0 that evaporates into Hawking radiation. The spacetime inside the shell, region I, is flat. Event $\mathbb{B}$ characterizes the crossing of the ingoing collapsing shell with the outgoing EH. The spacetime is not globally hyperbolic, and the dashed spacelike hypersurfaces $\overline{\Sigma}_i$, $i = 1, 2, 3$, are not Cauchy surfaces. The timelike curve $\mathcal{D}$, ending at the naked singularity $\mathbb{S}_2$, separates regions II and III approximated by metrics (4.3) and (4.4) of ingoing negative-energy and outgoing positive-energy Hawking radiation, respectively. Region IV represents a distant region where the spacetime is well approximated by a vacuum Schwarzschild metric. Hawking radiation is depicted by the translucent strip $u \lesssim 0$ at region III, and is assumed to carry out all black hole energy. Region V is a Minkowski spacetime in its own right, which emerges after the complete black hole evaporation.

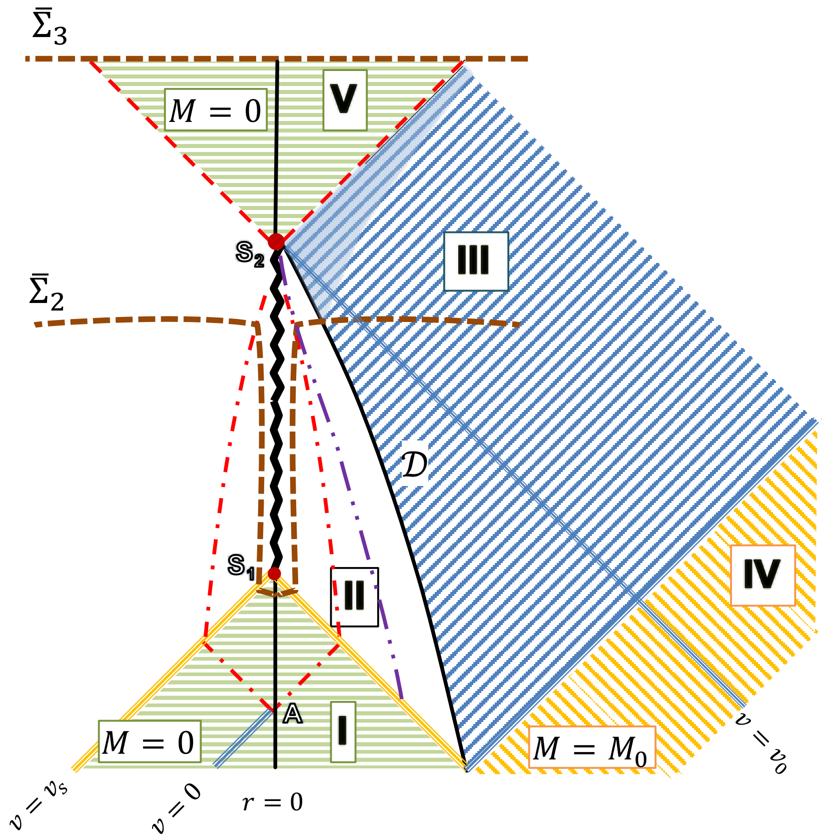


Figure 4.4: Sketch of the Finkelstein diagram corresponding to the Carter-Penrose one shown in Fig. 4.3. The diagram is axially symmetric around the $r = 0$ axis, but we have avoided repeating information in order not to overload the figure. One can see that $\bar{\Sigma}_2$ cannot be continuously deformed into $\bar{\Sigma}_3$ without crossing the singularity. There is nothing impeding information on $\bar{\Sigma}_2$ dropping in the singularity instead of reaching $\bar{\Sigma}_3$. We also show, here, the AH contained in region II depicted with a dashed-double-dotted line.

4.1 Nonglobally hyperbolic evaporating black hole

Let us consider a spherically symmetric thin shell with mass M_0 collapsing at the speed of light (with null coordinate $v = v_s$) to form a black hole, as shown in Figs. 4.3 and 4.4. We initially consider the usual semiclassical scenario where the black hole completely evaporates with the emission of Hawking radiation, leaving a Minkowski spacetime, region V. (We shall disregard possible massive particles that enter region V, making the metric deviate from flatness.) We shall broadly follow Ref. [22] in this section, which will serve to put in context our results of Sec. 4.2.

The spacetime inside the shell ($v < v_s$), region I, is flat and the metric can be cast as

$$ds_{\mathrm{I}}^2 = -dv^2 + 2dvdr + r^2 d\Omega^2, \tag{4.1}$$

where $r \geq 0$, $d\Omega^2 \equiv d\theta^2 + r^2 d\phi^2$, and $v =$ constant denotes ingoing null geodesics.

Concerning the metric outside the shell, let us assume that the evaporation of the black hole is negligible for $u < u_0$, where $u_0 =$ constant represents some outgoing null geodesic. As a consequence, Birkhoff's theorem guarantees that the spacetime in region IV, defined by $u < u_0 \cap v > v_s$, will be endowed with the usual Schwarzschild vacuum metric:

$$ds_{\mathrm{IV}}^2 = -\left(1 - \frac{2M_0}{r}\right) du^2 - 2dudr + r^2 d\Omega^2. \tag{4.2}$$

Then, the evaporation process starts (or becomes important) at $u = u_0$, where the original mass M_0 of the shell begins its gradual conversion into a positive-energy-density outgoing flux of radiation filling region III (associated with a negative-energy-density ingoing flux in region II flowing down into the black hole, as discussed in Sec. 3.2). Following Hiscock [22], regions II and III are assumed to share a common timelike hypersurface $\mathcal{D}$ (see Figs. 4.3 and 4.4) whose radius *monotonically* decreases from some $r = r_0$, at $u = u_0 \cap v = v_s$, to $r = 0$, at the naked singularity $\mathcal{S}_2$.

Hence, regions II and III should be well described by the ingoing and outgoing Vaidya metrics

$$ds_{\mathrm{II}}^2 = -\left(1 - \frac{2M_{\mathrm{II}}(v)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2 \tag{4.3}$$

and

$$ds_{\mathrm{III}}^2 = -\left(1 - \frac{2M_{\mathrm{III}}(u)}{r}\right) du^2 - 2dudr + r^2 d\Omega^2, \tag{4.4}$$

respectively, where $M_{\mathrm{II}}(v)$ and $M_{\mathrm{III}}(u)$ are monotonically decreasing mass functions:

$$dM_{\mathrm{II}}(v)/dv < 0 \quad (v_s < v < v_0), \tag{4.5}$$

$$dM_{\mathrm{III}}(u)/du < 0 \quad (u_0 < u < 0), \tag{4.6}$$

and the coordinate r must have the same value at every point on the common boundary $\mathcal{D}$, as defined in regions II and III, to ensure continuity of the 2-sphere areas. The dependence of the mass functions was chosen to reflect that Hawking radiation is dominated by the emission of massless particles, rendering $M_{\rm III}$ constant along the $u =$ constant null rays. An analogous reasoning for the negative-energy-density ingoing flux in region II drives $M_{\rm II}$ to be constant along the $v =$ constant null rays. Indeed, black holes are not expected to emit non-relativistic particles (say, neutrinos with energies of the order of 1 eV) before their Schwarzschild radius gets as minuscule as 10^{-8} m. Note, then, that

$$M_{\rm II}(v_s) = M_{\rm III}(u_0) = M_0, \quad M_{\rm II}(v_0) = M_{\rm III}(0) = 0. \tag{4.7}$$

We draw attention to the fact that we have set $u = 0$ at the final evaporation of the black hole. This is the same as setting the event horizon at $u = 0$, which is different from the divergent behavior ($u \to \infty$) of u that we have in the Schwarzschild case. This is a feature of the Vaidya metric, where now the coordinate u is not defined by $u = t - r^*$ with r^* given by Eq. (2.34), but rather by

$$du = \left(\frac{f_{\rm III}(t,r)}{1 - \frac{2M_{\rm III}(t,r)}{r}} \right)^{-1} dt - \left(1 - \frac{2M_{\rm III}(t,r)}{r} \right)^{-1} dr, \tag{4.8}$$

where $M_{\rm III}(t,r)$ is the mass function $M_{\rm III}(u)$ written in terms of $\{t,r\}$ and $f_{\rm III}(t,r)$ is the function used to write the Vaidya metric (4.4) as

$$ds^2 = -f_{\rm III}(t,r)dt^2 - \left(1 - \frac{2M_{\rm III}(t,r)}{r} \right)^{-1} dr^2 + r^2 d\Omega^2.$$

(The analytical form of $f_{\rm III}(t,r)$ is not known for non-linear mass functions). Since the spacetime is no longer stationary, the event horizon is not located at $r = 2M_{\rm III}(u)$, but rather inside the surface $r = 2M_{\rm III}(u)$ [26]. Hence, there is no divergence in Eq. (4.8) at the event horizon, as it is the case in the Schwarzschild metric. However, close to the junction between regions III and IV, we have $M_{\rm III}(u) \approx M_0$, which gives

$$du = dt - \left(1 - \frac{2M_0}{r} \right) dr,$$

where we have used the fact that near region IV, $f_{\rm III}(t,r) \approx 1 - 2M_0/r$. Thus, we arrive at

$$u = t - r - 2M \log \left(1 - \frac{2M_0}{r} \right),$$

which is the same definition of the Eddington-Finkelstein coordinate u. Therefore, we see that we can cover region III with a null coordinate u that is zero at the event horizon, but can be continuously matched with the usual Eddington-Finkelstein coordinate of Schwarzschild metric when approaching region IV. More details concerning the matching of the regions will follow next.

Now, in order to ensure the correct matching of the Vaidya metrics (4.3) and (4.4) along the boundary of regions II and III, we shall demand the first junction condition [23]

$$h^{\mathrm{II}}_{ab} = h^{\mathrm{III}}_{ab}, \tag{4.9}$$

where h_{ab} is the induced metric on the tridimensional hypersurface $\mathcal{D}$. (Because $\mathcal{D}$ is the region where positive- and negative- energy "particles" going to $\mathcal{J}^+$ and entering the black hole are assumed to pop up, respectively, there is no reason to expect the transition between the regions to be smooth on the stress-energy tensor and, thus, to impose the second junction condition $K^{\mathrm{II}}_{ab} = K^{\mathrm{III}}_{ab}$ for the extrinsic curvature K_{ab} on $\mathcal{D}$.) Let $\mathcal{D}$ be covered with coordinates $\xi^i = (\lambda, \theta, \phi)$, where ∂_λ is a future-pointing timelike vector. Thus, we can map the ξ^i coordinates to the ones defined in regions II and III on $\mathcal{D}$ as

$$(\lambda, \theta, \phi) \overset{\mathcal{D}}{\mapsto} (v(\lambda), r(\lambda), \theta, \phi) \tag{4.10}$$

and

$$(\lambda, \theta, \phi) \overset{\mathcal{D}}{\mapsto} (u(\lambda), r(\lambda), \theta, \phi), \tag{4.11}$$

respectively. Hence, the induced metrics are

$$ds^2_{\mathrm{II}} \overset{\mathcal{D}}{=} -\left[\left(1 - \frac{2M_{\mathrm{II}}(v)}{r}\right)\dot{v}^2 - 2\dot{v}\dot{r}\right] d\lambda^2 + r^2 d\Omega^2 \tag{4.12}$$

and

$$ds^2_{\mathrm{III}} \overset{\mathcal{D}}{=} -\left[\left(1 - \frac{2M_{\mathrm{III}}(u)}{r}\right)\dot{u}^2 + 2\dot{u}\dot{r}\right] d\lambda^2 + r^2 d\Omega^2, \tag{4.13}$$

where a dot represents derivative with respect to λ. Imposing the first junction condition (4.9) gives

$$\dot{r} \overset{\mathcal{D}}{=} \frac{1}{2(\dot{v} + \dot{u})}\left[\left(1 - \frac{2M_{\mathrm{II}}(v)}{r}\right)\dot{v}^2 - \left(1 - \frac{2M_{\mathrm{III}}(u)}{r}\right)\dot{u}^2\right]. \tag{4.14}$$

Using in the above equation the fact that $\dot{r} < 0$ [see discussion below Eq. (4.2)] with $\dot{v} > 0$ and $\dot{u} > 0$ (see Fig. 4.3), we get

$$\left(\frac{1-2M_{\rm III}(u)/r}{1-2M_{\rm II}(v)/r}\right)^{1/2} > \frac{dv}{du} > 0. \tag{4.15}$$

The $\mathcal{D}$ surface will be timelike if $ds^2 < 0$ for $\theta, \phi =$ constant. Imposing this in Eq. (4.13), we obtain

$$\left(1 - \frac{2M_{\rm III}(u)}{r}\right)\dot{u} > -2\dot{r} > 0, \tag{4.16}$$

because $\dot{r} < 0$. Recalling that $\dot{u} > 0$, the surface $\mathcal{D}$ will be timelike provided that

$$r > 2M_{\rm III}(u) \tag{4.17}$$

along $\mathcal{D}$. Applying condition (4.17) in Eq. (4.15), we get along $\mathcal{D}$

$$r > 2M_{\rm II}(v). \tag{4.18}$$

Equation (4.18) turns out particularly convenient when comparing the relative positions of $\mathcal{D}$ and the apparent horizon (AH). The latter is given by the geometric location where outgoing null geodesics have vanishing expansion θ. In the coordinates being used, this translates into the geometric location where outgoing null geodesics stem with $dr/dv = 0$. Let us make the ansatz (to be confirmed just ahead) that the AH is within region II. In this case, we consider the equation for outgoing null geodesics ($\theta, \phi =$ constant) given by metric (4.3):

$$\frac{dr}{dv} = \frac{1}{2}\left(1 - \frac{2M_{\rm II}(v)}{r}\right). \tag{4.19}$$

By imposing above $dr/dv = 0$, we obtain the equation for the AH:

$$r(v) \stackrel{AH}{=} 2M_{\rm II}(v), \quad v_s < v < v_0. \tag{4.20}$$

By comparing Eqs. (4.18) and (4.20), we see that the AH is inside the boundary $\mathcal{D}$ and, thus, dwells in region II, confirming the ansatz (see Fig. 4.4). Then, the AH pops up at the light shell with $r = 2M_0$ [see Eq. (4.7)] and diminishes monotonically as ruled by Eq. (4.5), eventually vanishing at $\mathcal{S}_2$.

Although the precise form of $M_{\rm II}(v)$ and $M_{\rm III}(u)$ depends on the evaporation details (that we do not address here), we know that the late stage of the evaporation process is dominated by Hawking radiation represented by the translucent strips in Figs. 4.3 and 4.4. Hence, we can give an explicit form for $M_{\rm III}(u)$ in this region based on Hawking's prediction. Assuming that the radiation reaching the asymptotic future is purely null outgoing [i.e., $M = M_{\rm III}(u)$], Eq. (3.75) reads

$$\frac{dM_{\rm III}(u)}{du} = -\frac{M_{\rm III}(u)^{-2}}{15,360}. \tag{4.21}$$

This can be integrated, leading to

$$M_{\rm III}(u) = \left(\frac{-u}{5120}\right)^{1/3}, \tag{4.22}$$

where we note that the end of the evaporation here is at $u = 0$. We emphasize that Eq. (4.22) is only a good approximation in the domain of region III where the optical geometric approximation used to derive Hawking's effect is valid.

4.2 Globally hyperbolic fully evaporating black hole

While the semiclassical theory provides a reliable description of the evaporation process for macroscopic black holes, its predictions are expected to break down when the black hole reaches a mass of the order of the Planck mass. After that, the black hole evaporation is governed by quantum gravity, which would supposedly describe the process in the so-called Planck scale. We shall look for a modification of the spacetime in this regime to make it globally hyperbolic (taking care of not jeopardizing any classical or semiclassical results). The simplest example would be a black hole that evaporates via Hawking radiation until it reaches the Planck size, after which it would freeze out (see Fig. 4.2). Here, however, we look for an instance where the black hole fully evaporates (in the sense that no massive remnant is left over at the end) in addition to being globally hyperbolic, as described by Figs. 4.5 and 4.6. (By construction any spacetimes that maintain unaltered classical and semiclassical physics can only be distinguished by their Planck scale different properties.)

In this token, we shall choose the metric in region V of Figs. 4.5 and 4.6 ($v > v_1 \cap r \lesssim 1$) to make the spacetime globally hyperbolic as depicted in the diagrams, while the evaporation model discussed in Sec. 4.1 will be valid everywhere outside region V. Note that P at $(u, v) = (u_1, v_1)$, representing a 2-sphere with an area of the Planck order, is the meeting point of regions II, III and V. It is interesting to note that observers in the far future visiting the asymptotically flat region $0 < u < u_1$

will be influenced (but not governed) by quantum gravity. Hence, we assume that we can also describe this region by metric (4.4) with a general mass function $M_{\rm III}(u)$ that should approach the Planck mass,

$$M_{\rm III}(u_1) \sim 1, \tag{4.23}$$

as $u \to u_1$.

Now, let us investigate the conditions that the metric in region V should satisfy to replace the naked singularity S_2 in Figs. 4.3 and 4.4 by a null (thunderbolt) singularity as in Figs. 4.5 and 4.6. Previous investigations, e.g. Ref. [18], have examined the Planck-scale region at the end of black hole evaporation and its connection to thunderbolt singularities based on specific quantum-gravity models. In contrast, we seek to uncover general conditions for the occurrence of a thunderbolt singularity and its prevailing features in a model-independent manner. In this way, let us begin considering a general ingoing Vaidya metric in region V but, now, with the mass function $M = M_{\rm V}(v,r)$ depending on v and r:

$$ds_V^2 = -\left(1 - \frac{2M_{\rm V}(v,r)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2, \tag{4.24}$$

where we use the same radial coordinate r to ensure the continuity of the 2-sphere areas. Let us cast $M_{\rm V}(v,r)$ as

$$M_{\rm V}(v,r) = \begin{cases} \mu_1(v,r), & v \geq v_0, \\ \mu_2(v,r), & v_0 \geq v > v_1, \end{cases} \tag{4.25}$$

with $\mu_i(v,r)$, $i = 1,2$, being monotonic positive functions:

$$\partial_v \mu_i(v,r) < 0, \quad \partial_r \mu_i(v,r) > 0. \tag{4.26}$$

Also, continuity of $M_{\rm V}(v,r)$ at $v = v_0$ is ensured by imposing

$$\mu_1(v_0,r) = \mu_2(v_0,r). \tag{4.27}$$

The rise of the null singularity depends on choosing Eq. (4.25) such that $r = 0$ is identified with an outgoing singular null ray, $u = 0$, for $v \geq v_0$. To achieve this, let us consider a one-parameter family of spherically symmetric timelike hypersurfaces S_i, defined by $\phi_i(x^\mu) =$ constant, namely, $r = r_i =$ constant for different values of r_i, and let us assume that $\nabla_\alpha \phi_i \nabla^\alpha \phi_i \xrightarrow{r_i \to 0} 0$. Then, in the $r_i \to 0$ limit, S_i degenerates into a null surface spanned by the null dual vector $\nabla_\alpha r$. Since $\nabla_\alpha r = (0, 1, 0, 0)$ and

$$g^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 - 2M_{\mathrm{V}}(v,r)/r & 0 & 0 \\ 0 & 0 & r^{-2} & 0 \\ 0 & 0 & 0 & (r\sin\theta)^{-2} \end{pmatrix} \tag{4.28}$$

for metric (4.24), we obtain

$$\nabla_\alpha r \nabla^\alpha r = 1 - 2M_{\mathrm{V}}(v,r)/r. \tag{4.29}$$

Therefore, a null outgoing ray (labeled by $u = 0$) stems in region V from the evaporation endpoint $\mathbb{S}_2$ at $r = 0$ provided

$$\lim_{r \to 0} \left[1 - 2M_{\mathrm{V}}(v,r)/r\right] = 0. \tag{4.30}$$

Combining Eqs. (4.30) and (4.25) for $v \geq v_0$ leads us to the conclusion that the existence of a zero-mass null singularity, $M_{\mathrm{V}}(v,r) \xrightarrow{r \to 0} 0$ (for $v \geq v_0$), requires

$$\lim_{r \to 0} \mu_1(v,r)/r = 1/2. \tag{4.31}$$

(Note that the above requirement is not satisfied, e.g., by linear mass models for evaporating black holes [22, 24, 25].)

The singular nature of the outgoing null ray departing from $\mathbb{S}_2$ at $u = 0$ ($r = 0$) in region V ($v \geq v_0$) can be investigated using the Kretschmann scalar $K = R_{abcd}R^{abcd}$:

$$K = 4r^{-6}\left[12\mu_1^2 - 16\, r\, \mu_1 \partial_r \mu_1 + 4r^2 \mu_1 \partial_r^2 \mu_1 + 8r^2(\partial_r \mu_1)^2 + r^4(\partial_r^2 \mu_1)^2 - 4r^3(\partial_r \mu_1)(\partial_r^2 \mu_1)\right]. \tag{4.32}$$

Equation (4.31) ensures the divergence of the term $\mu_1(v,r)^2/r^6$ for $r \to 0$ and, consequently, of K (provided the other terms are not fine-tuned to cancel out this divergence).

Finally, we must impose the mandatory continuity of the induced metrics, h_{ab}, on the boundaries of region V. (We do not impose the matching of the extrinsic curvatures, as there is no reason to expect the transition of the stress-energy tensor from semiclassical to quantum-gravity regions to be smooth.) The condition

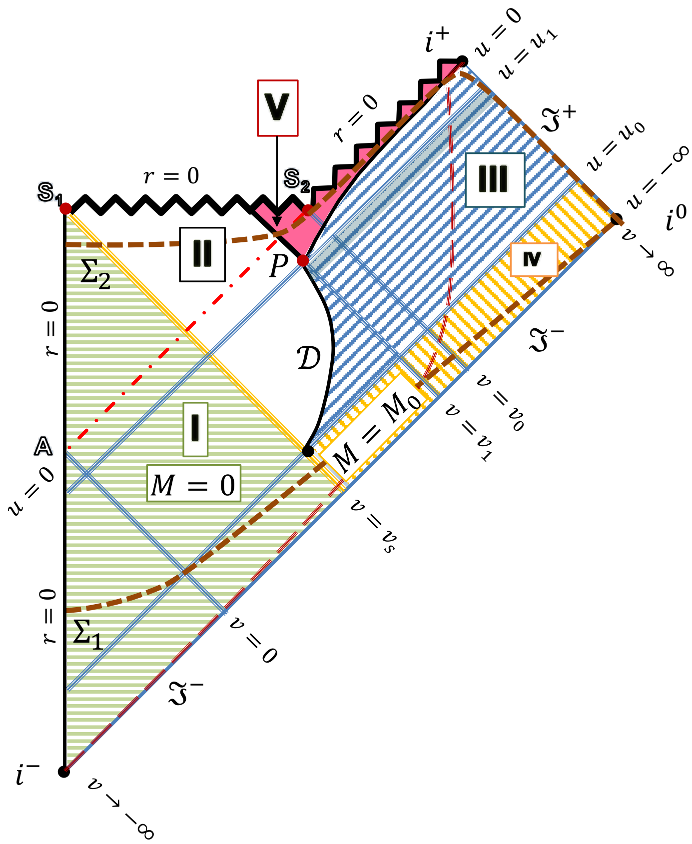


Figure 4.5: Conformal Carter-Penrose diagram associated with the total collapse of a spherically symmetric light shell into a black hole with mass M_0. This spacetime is similar to the usual semiclassical one except for the quantum-gravity naked singularity that is replaced by a null singularity. The timelike curve $\mathcal{D}$ separates regions II and III ruled by metrics (4.3) and (4.4) of ingoing negative-energy and outgoing positive-energy radiation, respectively. The mass of the black hole is of the order of the Planck mass, $M \sim 1$, at the event P. The Planck-scale region near the null singularity is represented by region V that is delimited by the null line $v = v_1$ joining P to the spacelike singularity at $r = 0$, and the timelike curve with Planck radius $r = 1$ joining P to i^+. The spacetime is globally hyperbolic, and the dashed spacelike hypersurfaces Σ_1 and Σ_2 are Cauchy surfaces. Information is preserved on each Cauchy hypersurface, although observers will not have access to the information that enters the black hole. It is worth noting that the black hole will still be held responsible for any information that observers outside the event horizon cannot access.

$h_{ab}^{\mathrm{II}} = h_{ab}^{\mathrm{V}}$ is automatically satisfied, since the corresponding induced metrics at

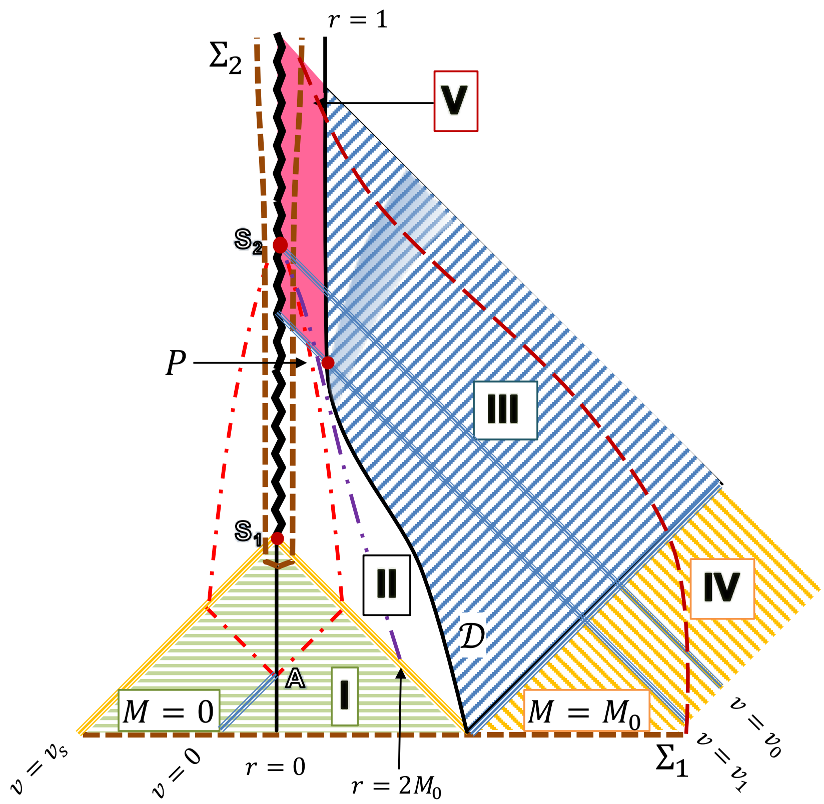


Figure 4.6: Sketch of the Finkelstein diagram corresponding to the Carter-Penrose one shown in Fig. 4.5. The diagram is axially symmetric around the $r = 0$ axis, but we have avoided repeating information to not overload it. An observer extending from i^- to i^+ (as depicted here and in Fig. 4.5 by long-dashed lines) would perceive no particles in regions I and IV but this would continuously change in region III building up into Hawking radiation at $u \lesssim u_1$, eventually. After that, the observer would accuse the presence of a black hole with Planck mass (decreasing as ruled by quantum gravity). It is particularly illuminating to see how the Cauchy surface Σ_2 curves itself, collecting information from past Cauchy surfaces.

$v = v_1$ are

$$ds^2_{\mathrm{II}} \overset{v=v_1}{=} r^2 d\Omega^2, \tag{4.33}$$

$$ds^2_{\mathrm{V}} \overset{v=v_1}{=} r^2 d\Omega^2. \tag{4.34}$$

Concerning the matching between regions III and V, we cover the corresponding boundary at $r = 1$ (in natural units) with coordinates $\xi^i = (\tau, \theta, \phi)$, where ∂_τ is a future-pointing timelike vector. Thus, we can map accordingly the ξ^i coordinates on the surface to the ones defined in regions III and V as

$$(\tau, \theta, \phi) \overset{r=1}{\mapsto} (v(\tau), 1, \theta, \phi), \tag{4.35}$$
$$(\tau, \theta, \phi) \overset{r=1}{\mapsto} (u(\tau), 1, \theta, \phi). \tag{4.36}$$

In this way, the induced metrics are

$$ds^2_{\mathrm{III}} \overset{r=1}{=} -[1-2M_{\mathrm{III}}(u)] \left(\frac{du}{d\tau}\right)^2 d\tau^2 + d\Omega^2 \tag{4.37}$$

and

$$ds^2_{\mathrm{V}} \overset{r=1}{=} -[1-2M_{\mathrm{V}}(v,1)] \left(\frac{dv}{d\tau}\right)^2 d\tau^2 + d\Omega^2. \tag{4.38}$$

Then, the junction condition, $h^{\mathrm{III}}_{ab} = h^{\mathrm{V}}_{ab}$, reads

$$\frac{du}{dv} \overset{r=1}{=} \left(\frac{1-2M_{\mathrm{V}}(v,1)}{1-2M_{\mathrm{III}}(u)}\right)^{1/2}, \tag{4.39}$$

where we impose $M_{\mathrm{III}}, M_{\mathrm{V}} \lesssim 1/2$ at the event P, $(u,v) = (u_1, v_1)$. This, combined with $\partial_v M_{\mathrm{V}}(v,1) < 0$, $\partial_u M_{\mathrm{III}}(u) < 0$, is enough to guarantee that Eq. (4.39) is well behaved on the whole boundary. A discussion about the behavior of the apparent and event horizon in this spacetime can be found in Ref. [26]

4.3 Black holes not as ordinary black boxes

The goal of this section is to clarify the distinction between the two prevailing issues concerning the black hole information loss:

A. the information loss observed at the final stages of the evaporation due to the black hole internal singularity expected to eventually erase information that enters the horizon, and

B. the paradox that arises (much earlier) in the semiclassical regime (after the Page time) if one imposes that black holes must behave as any other ordinary thermodynamical system (an imposition called "central dogma", see Ref. [27]).

Global hyperbolicity of the spacetime of Fig. 4.5 and the ensuing conservation of information at each Cauchy surface may not be enough to make supporters of the "central dogma" feel comfortable since they would like all this information to be accessible to external observers. In fact, it has been argued that if information that falls into the black hole does not begin to be accessible to external observers after some given (Page) time, a paradoxical situation would arise in the semiclassical regime when the black hole would still be macroscopic. The argument goes as follows (see, e.g., Ref. [27]).

Let Σ_ζ be a spacelike foliation of region $\mathrm{I} \cup \mathrm{II} \cup \mathrm{III} \cup \mathrm{IV}$ (of any of Figs. 4.3-4.6) with $\zeta \in \mathbb{R}$ being a labeling increasing to the future. Now, let Σ_ζ^{in} be the intersection of Σ_ζ with the black hole and $\Sigma_\zeta^{out} \equiv \Sigma_\zeta \backslash \Sigma_\zeta^{in}$. Then, tracing out the degrees of freedom of the quantum fields in the region Σ_ζ^{in} [resp., Σ_ζ^{out}] leads to the reduced fields' state $\hat{\rho}_{out}(\zeta)$ [resp., $\hat{\rho}_{in}(\zeta)$], whose corresponding von Neumann entropy we call $S_{out}^{vN}(\zeta)$ [resp., $S_{in}^{vN}(\zeta)$]. It is well known that if the state of the fields on Σ_ζ is pure (hence, with vanishing von Neumann entropy), then

$$S_{out}^{vN}(\zeta) = S_{in}^{vN}(\zeta), \tag{4.40}$$

after proper regularization and renormalization.[1]

Now, the assumption that Hawking radiation does *not* encode information that falls into the black hole implies that

$$S_{out}^{vN}(\zeta) \text{ is a } \textit{strictly increasing} \text{ function of } \zeta. \tag{4.41}$$

In particular, if we assume that Hawking radiation is truly a thermal state at Hawking temperature T_H, then $S_{out}^{vN}(\zeta)$ can be evaluated from the same (black hole) parameters that determine T_H.

On the other hand, *if* one *associates* the thermodynamical entropy of the black

[1]For free fields in d spacetime dimensions, the von Neumann entropy for general Hadamard states associated with a (spatial) region Σ with boundary $\partial\Sigma$ can usually be cast as

$$S_\Sigma^{vN} = \sum_{n<d-2} \frac{f_n(\partial\Sigma)}{\epsilon^{d-2-n}} + f(\partial\Sigma)\log\epsilon + S_0(\Sigma),$$

where ϵ is a short-distance cutoff, $n = 0, 1, \cdots$ are positive integers, f and f_n are state-independent constants (although the f_n's depend on the regularization process chosen) expressed as integrals on $\partial\Sigma$, and $S_0(\Sigma)$ is a finite state-dependent term. In $d = 4$, $f_0(\partial\Sigma) \propto A(\partial\Sigma)$, where $A(\partial\Sigma)$ is the area of $\partial\Sigma$, giving an area law for the leading term of von Neumann entropy divergences. If one is interested in variations of the entropy, one can renormalize S_Σ^{vN} by subtracting its value with respect to some reference (Hadamard) state. For more details (including interacting theories), see Refs. [28, 29, 30, 31] and references therein.

hole, S_{bh}^{therm}, as due to $\hat{\rho}_{in}(\zeta)$ at Σ_ζ^{in}, we would have

$$S_{in}^{vN}(\zeta) \leq S_{in}^{vN}[\tilde{\rho}]\Big|_{\text{sup}} \longleftrightarrow S_{bh}^{\text{therm}}(\zeta) = A(\zeta)/4, \tag{4.42}$$

where the first inequality is the trivial statement that S_{in}^{vN} is limited by the supremum of the von Neumann entropy among all possible states $\tilde{\rho}$ (satisfying the relevant coarse-grained constraints), and '$\longleftrightarrow$' stands for the *assumption* that $S_{in}^{vN}[\tilde{\rho}]\big|_{\text{sup}}$ equals the thermodynamical entropy of the black hole given by the Bekenstein-Hawking formula with $A(\zeta)$ being the area of the event horizon on Σ_ζ.

It is quite obvious that Eqs. (4.40), (4.41), and (4.42) are in tension: a strictly increasing function $S_{out}^{vN}(\zeta)$ bounded from above by the decreasing function $A(\zeta)/4$. This is the essence of the reasoning used by the followers of the "central dogma" to argue that Eq. (4.41) should be violated after some (Page) time; i.e., that Hawking radiation should eventually carry away information that fell into the black hole, "purifying" $\hat{\rho}_{out}(\zeta)$ for sufficiently large values of ζ. Yet, subtle as it may be, the identification in Eq. (4.42) depends on assuming that black holes are like any common ("black box") thermodynamical system, which is a quite nontrivial assumption.

While the walls of a "black box" do not have to "react" to attempts of external observers to "hide" entropy in its interior, the event horizon (being a *dynamical* entity) "denounces" any such attempt by changing its area. Therefore, while the entropy of a "black box" *must be* due to whatever it contains (to ensure a meaningful second law of thermodynamics), this is not so for black holes (being consistent with the *standard* view that modes falling into the black hole are inaccessible to external observers). According to the standard semiclassical description, the event horizon should be seen as a *causal* boundary separating degrees of freedom that (i) escape to the future null infinity and account for Hawking radiation from those that (ii) fall into the black hole (being both entangled with each other). In this way, the tension among Eqs. (4.40), (4.41), and (4.42) is avoided, not due to information "leaking" from the black hole [i.e., violation of Eq. (4.41)], but because the identification in Eq. (4.42) would not be valid for black holes.

Indeed, in the semiclassical realm, Eq. (4.41) is necessary to maintain the generalized second law true. This can be already seen from a simple calculation. Assuming that Hawking radiation as measured by asymptotic Killing observers is thermal with temperature T_H as given by Eq. (3.65), each electromagnetic radiation energy element dM_H carries an entropy element $dS_H = (4/3) \times dM_H/T_H$. This increase of external-to-the-horizon entropy is accompanied by a decrease of the black hole thermodynamical entropy of $dS_{bh}^{\text{therm}} = -8\pi M dM$, leading to

$$\left| \frac{dS_H}{dS_{bh}^{\text{therm}}} \right| = \frac{4}{3} > 1. \tag{4.43}$$

A more detailed analysis taking into account the full density matrix $\hat{\rho}_{out}$ describing Hawking radiation (including grey-body factors) leads to [32]

$$\left| \frac{dS_H}{dS_{bh}^{\text{therm}}} \right| \equiv R \gtrsim 1, \tag{4.44}$$

which ensures the validity of the generalized second law of thermodynamics.

In summary, if one adopts the "central dogma" and assumes that black holes must behave as any other ordinary thermal system, Hawking radiation should begin to deviate from thermality yet in the semiclassical domain (after some Page time), which is considered paradoxical for being at odds with the semiclassical derivation of the Hawking effect. On the other hand, if one does not adopt the "central dogma," the black-hole degrees of freedom (and thus its Hilbert space), do not need to be associated with the modes that fell into the black hole. In this case, black holes (with their horizons) must be seen as independent dynamical entities with their own Hilbert spaces, existing no contradiction between $A/4$ being the logarithm of the dimension of the black-hole Hilbert space (e.g., representing qubits encoded in each quantum of area) and the much larger Hilbert space associated with the swallowed matter/radiation.

Chapter 5

Conclusions

As discussed, quantum particle creation during the process of gravitational collapse and formation of black holes would lead to the emission of a thermal energy flux, as seen by asymptotic static observers, eventually resulting in the complete evaporation of the black hole. Semiclassical arguments strongly suggest that, in the process of black hole formation and complete evaporation, a pure quantum state will evolve to a mixed state, violating unitary evolution. In other words, information will be lost. Despite its theoretical robustness, this scenario is still posed as a paradox, even after fifty years of its discovery. Past attempts to solve this nonexistent paradox have led to all sorts of undesired features, such as a violation of causality and other basic principles in low curvature regimes.

As we have seen, far from paradoxical, the information loss is a natural consequence of the fact that the spacetime usually considered is not globally hyperbolic, i.e., the past domain of dependence of spacelike surfaces after the evaporation does not include the interior of the past black hole. In this way, the quantum state will remain entangled with the quantum field observables inside the (now) past black hole, even though the black hole no longer exists.

Nevertheless, quantum gravity may modify the semiclassical picture depending on its impact on the Planck scale physics. In this manner, we have also discussed the case where the final naked singularity of the usual semiclassical spacetime is replaced by a null (thunderbolt) singularity. All classical and semiclassical features are preserved in small curvature regions. In particular, late-time observers will experience Hawking radiation at $u \lesssim u_1$, after which they accuse the presence of a Planck-mass black hole (decreasing as ruled by quantum gravity). A pleasant feature of this spacetime is that it is globally hyperbolic. As a result, information is preserved on each Cauchy hypersurface, where the evolution is assumed to be unitary. Despite this, the fact that late-time Cauchy surfaces "invade" the EH prevents late-time observers from accessing complete early-time information. Thus, even after late-time observers (those ending at i^+) do not see Hawking radiation anymore, they can still "blame" the hole for the thermal state of Hawking radiation since its degrees of freedom will be still entangled with the ones inside the hole. Without a full quantum gravity theory, nothing can be said about the robustness of the conditions that lead to this scenario, neither in favor nor against it. Hence, the very fact that a fully evaporating black hole scenario exists, where information (and unitarity) can be preserved with "minimal" deviations from the semiclassical picture, is worthy of attention.

Bibliography

[1] W. G. Unruh, *Notes on black-hole evaporation, PhysRevD.* **14** (1976) 870.

[2] S. W. Hawking, *Particle creation by black hole, Commun. Math. Phys.* **43** (1975) 199.

[3] J. R. Muñoz de Nova, K. Golubkov, V . I. Kolobov, *Observation of thermal Hawking radiation and its temperature in an analogue black hole, Nature.* **569** (2019) 688.

[4] R. Schutzhold, *Ultra-cold atoms as quantum simulators for relativistic phenomena, arXiv:2501.03785* (2025).

[5] W. G. Unruh and R. M. Wald, *Information loss, Rep. Prog. Phys.* **80** (2017) 092002.

[6] R. M. Wald, *General Relativity,* Chicago: University of Chicago (1984).

[7] S. W. Hawking, G. F. R. Ellis, *The Large Scale Structure of Space-Time,* Cambridge University Press (1973).

[8] S. W. Hawking, W. Israel, *General relativity: An Einstein centenary survey,* Cambridge University Press (1979).

[9] R. P. Geroch, *Domain of dependence, J. Math. Phys.* **11** (1970) 437.

[10] R. M. Wald, *Quantum Field Theory in Curved Space-Time and Black Hole Thermodynamics,* Chicago: University of Chicago (1995).

[11] S.W. Hawking, *Black holes in general relativity, Commun.Math. Phys* **25** (1972) 152.

[12] S. A. Fulling, *Nonuniqueness of Canonical Field Quantization in Riemannian Space-Time,* Phys. Rev. D **7** (1973) 2850.

[13] L. Parker, *Particle Creation in Expanding Universes. Phys. Rev. Lett.* **21** (1968) 562

[14] S. W. Hawking, *Breakdown of predictability in gravitational collapse, Phys. Rev. D* **14** (1976) 2460.

[15] R. M. Wald, *On particle creation by black holes, Commun. Math. Phys.* **45** (1975) 9.

[16] P. C. W. Davies, S. A Fulling, and W. G. Unruh, *Energy-momentum tensor near an evaporating black hole, Phys. Rev. D* **13** *(1976) 2720.*

[17] R. M. Wald, Black holes, singularities, and predictability, in "Quantum Theory of Gravity – Essays in honor of the 60$^{\text{th}}$ *birthday of B. S. DeWitt, Bristol: Adam Hilger (1984).*

[18] S. W. Hawking and J. M. Stewart, Naked and thunderbolt singularities in black hole evaporation, Nuc. Phys. B **400** *(1993) 393.*

[19] D. A. Lowe, Semiclassical approach to black hole evaporation, Phys. Rev. D **47** *(1993) 2446.*

[20] A. Ashtekar, F. Pretorius, and F. M. Ramazanoğlu, Evaporation of two-dimensional black holes, Phys. Rev. D **83** *(2011) 044040.*

[21] S. Hossenfelder and L. Smolin, Conservative solutions to the black hole information problem, Phys. Rev. D **81** *(2010) 064009.*

[22] W. A. Hiscock, Models of evaporating black holes. II. Effects of the outgoing created radiation, Phys. Rev. D **23** *(1981) 2823.*

[23] E. Poisson, A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics, Cambridge: University of Cambridge (2004).

[24] D. W. Ring, A linear approximation to black hole evaporation, Classical and Quantum Gravity **23** *(2006) 5027.*

[25] W. A. Hiscock, Models of evaporating black holes. I, Phys. Rev. D **23** *(1981) 2813.*

[26] J. V. O. Pêgas, A. G. S. Landulfo, G. E. A. Matsas, D. A. T. Vanzella, Globally hyperbolic evaporating black hole and the information loss issue, Classical and Quantum Gravity. **42***, 065009 (2025).*

[27] A. Almheiri, T. Hartman, J, Maldacena, E. Shaghoulian, and A. Tajdini, The entropy of Hawking radiation, Rev. Mod. Phys. **93** *(2021) 035002.*

[28] H. Casini and M. Huerta, Lectures on entanglement in quantum field theory, PoS **TASI2021** *(2022) 2 (arXiv:2201.13310).*

[29] *S. N. Solodukhin, Entanglement entropy of black holes, Living Rev. Relativity* **14** *(2011) 8.*

[30] *R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall, A Quantum Focussing Conjecture, Phys. Rev. D* **93** *(2016) 064044.*

[31] *J. L. Miqueleto and A. G. S. Landulfo, Exact renormalization group, entanglement entropy, and black hole entropy, Phys. Rev. D* **103** *(2021) 045012.*

[32] *W. H. Zurek, Entropy Evaporated by a Black Hole, Phys. Rev. Lett.* **49** *(1982) 1683.*