Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Derivation of Hawking radiation by analogy with the Unruh effect

Relatore:
Prof. Roberto Casadio

Presentata da:
Damiano Scevola

Anno Accademico 2022/2023

Alla mia famiglia

Abstract

L'effetto Hawking è un fenomeno riguardante i buchi neri secondo cui essi emettono una radiazione termica composta da particelle che vengono prodotte a causa dell'interazione tra il campo quantistico che le definisce e l'orizzonte degli eventi. In questa tesi ci si propone di ricavare la temperatura della radiazione e l'entropia di un buco nero di Schwarzschild utilizzando il principio di equivalenza tra un ipotetico osservatore stazionario in prossimità dell'orizzonte degli eventi e un altro osservatore in moto uniformemente accelerato nello spaziotempo piatto, che osserva una radiazione termica per effetto Unruh. Infine si procede a studiare il fenomeno dell'evaporazione dei buchi neri e il paradosso dell'informazione.

The Hawking effect is a phenomenon regarding black holes according to which they emit thermal radiation composed of particles that are produced due to the interaction of the quantum field defining them and the event horizon. In this dissertation, we propose to derive the temperature of the radiation and the entropy of a Schwarzschild black hole using the equivalence principle between a hypothetical stationary observer in proximity of the event horizon and another observer in uniform accelerated motion in flat spacetime, who observes a thermal radiation due to Unruh effect. Finally, we proceed to study the phenomenon of black hole evaporation and the information loss paradox.

Contents

Introduction

Black holes are one of the most exotic physical systems that exist in our universe. A black hole is a region of spacetime where the curvature becomes so intense that all possible paths of any physical object, including light, do not escape such region, but instead all point towards the singularity, a point of infinite density which is located at the center of the region. The border of a black hole is called the event horizon, above which some photons can still escape the gravitational pull. The discovery of black holes was led by theory, in fact, their existence was predicted for the first time in 1916 when Karl Schwarzshild solved Einstein's equations of general relativity for a spherical gravitational source. However, until the 1960s they were considered just a mathematical curiosity, like many other exotic predictions that newborn theories make. In 1967 the first neutron star was observed, and the possibility of a gravitational collapse so strong to form a black hole became acceptable. In the following years, the observations of the trajectories of astronomical objects hinted at the existence of black holes in order to explain their motion. In 2017, scientists were able to reconstruct the image of M87, the black hole at the center of the Virgo A galaxy, from data collected by the Event Horizon Telescope, and that constituted the first direct observation of a black hole. Due to their strong gravity and because the Schwarzschild radius increases with mass, until 1974 it was thought that black holes could only get larger by absorbing more and more matter. In that year, however, Stephen Hawking elaborated a theory that opened the possibility that quantum effects can make black holes radiate matter, shrink in dimension and completely disappear (after an enormous amount of time). This goes by the name of the Hawking effect, and it is the phenomenon that we want to study in this dissertation.

To derive the expression for the temperature of Hawking thermal radiation, we need a quantity called *surface gravity* which characterizes *Killing horizons*, and in chapter 1 we are going to define both such concepts and understand what they represent in an actual Schwarzschild black hole. After that, in chapter 2 we move to discussing quantum field theory for a real scalar field, pointing out in particular how the concepts of *vacuum* and *particles* arise from the quantization procedure, and how they behave in curved spacetime. Then, in chapter 3 we are going to analyze the Unruh effect, which consists in the production of particles in accelerated reference frames in flat spacetime, using the theory from the previous chapter. Finally, in chapter 4 we will realize that stationary observers just above the event horizon of a Schwarzschild black hole are analogous to accelerated observers in flat spacetime, and use the quantities derived for the Unruh effect to understand what Hawking radiation actually consists of and what some of its consequences are.

The main reference that has been used to develop this dissertation is chapter 9 of the book "Spacetime and Geometry: An Introduction to General Relativity" by Sean Carrol [1], and the other references are cited throughout the discussion. Most of the topics that we cover can also be found in [2]. A prerequisite to being able to understand what follows, other than undergraduate Physics knowledge, is a familiarity with basic concepts of differential geometry and general relativity such as differentiable manifolds, metric and tensor fields, hypersurfaces, covariant derivatives, Killing vector fields, and the Schwarzschild solution to Einstein field equations. In addition to Carrol's book, these topics are also extensively discussed in [3, 4, 5].

Let us now dive into the details and take our steps towards a basic understanding of Hawking radiation.

Chapter 1

Killing horizons and surface gravity

In this chapter, we are going to derive some results regarding null hypersurfaces, Killing vector fields, redshift factors, and Killing horizons which will be useful to understand what follows.

1.1 Null hypersurfaces

Let us first recall what a hypersurface is.

Def. 1.1 (Hypersurface). *Given a scalar function f defined on a differentiable manifold $\mathcal{M}$, a hypersurface $\Sigma \subset \mathcal{M}$ is the set of all points $P \in \mathbb{M}$ where the function has a fixed value $C \in \mathbb{R}$:*

$$\Sigma = \big\{P \in \mathcal{M} | f(P) = C\big\}.$$

A way to define a hypersurface given a chart on a manifold is by constraining one of the coordinates to be a fixed constant. We can define a vector field that is orthogonal to a given hypersurface Σ at each point, in the sense that it is orthogonal to all vectors in the tangent space of the hypersurface (which is a submanifold). If the hypersurface is defined by a function f being constant, the orthogonal vector field $\vec{\zeta}$ will have the following expression for components:

$$\zeta^\mu = g^{\mu\nu}\nabla_\nu f. \tag{1.1}$$

Proof. *Given an arbitrary point $P \in \Sigma$ and a vector $\vec{V} \in T_P\Sigma$, their scalar product is*

$$g_{\mu\nu}\zeta^\mu V^\nu = V^\nu g_{\mu\nu} g^{\mu\alpha}\nabla_\alpha f = V^\nu \delta^\alpha_\nu \nabla_\alpha f = V^\alpha \nabla_\alpha f = \nabla_{\vec{V}} f = 0,$$

where the last equality comes from the fact that $\vec{V}$ is tangent to Σ, therefore the *derivative of f along $\vec{V}$ is 0 because f stays constant by definition on the hypersurface.*

An hypersurface Σ is said *timelike* if its orthogonal vector field $\vec{\zeta}$ is spacelike, *spacelike* if $\vec{\zeta}$ is timelike and *null* if $\vec{\zeta}$ is null everywhere on Σ. In addition, the following identity for the orthogonal vector field $\vec{\zeta}$ holds:

$$\zeta_{[\alpha}\nabla_\beta\zeta_{\gamma]} = 0, \tag{1.2}$$

where braces over the indices stand for anti-symmetrization.

Proof. *By* (1.1) *we can write*

$$\begin{aligned}
6\zeta_{[\alpha}\nabla_\beta\zeta_{\gamma]} &= +\zeta_\alpha\nabla_\beta\zeta_\gamma + \zeta_\beta\nabla_\gamma\zeta_\alpha + \zeta_\gamma\nabla_\alpha\zeta_\beta \\
&\quad - \zeta_\alpha\nabla_\gamma\zeta_\beta - \zeta_\beta\nabla_\alpha\zeta_\gamma - \zeta_\gamma\nabla_\beta\zeta_\alpha \\
&= +\nabla_\alpha f\nabla_\beta\nabla_\gamma f + \nabla_\beta f\nabla_\gamma\nabla_\alpha f + \nabla_\gamma f\nabla_\alpha\nabla_\beta f \\
&\quad - \nabla_\alpha f\nabla_\gamma\nabla_\beta f - \nabla_\beta f\nabla_\alpha\nabla_\gamma f - \nabla_\gamma f\nabla_\beta\nabla_\alpha f = 0,
\end{aligned}$$

where we used the fact that coordinate covariant derivatives of scalar functions reduce to partial derivatives, and therefore they commute.

Null hypersurfaces have several interesting properties which are necessary to understand black hole event horizons. Given a null hypersurface Σ with orthogonal vector field $\vec{\zeta}$,

- $\vec{\zeta}$ is also tangent to Σ;

  Proof. *By definition of null hypersurface, $\vec{\zeta}$ is a null vector, and therefore it is orthogonal to itself having zero norm, and vectors that are orthogonal to $\vec{\zeta}$ necessarily are tangent to the hypersurface Σ.*

- all integral curves $x^\mu(\alpha)$ of $\vec{\zeta}$ defined by

  $$\frac{dx^\mu}{d\alpha} = \zeta^\mu \tag{1.3}$$

  stay inside Σ and satisfy the geodesic equation (α not necessarily being an affine parameter, and $\eta(\alpha)$ being a scalar function on the curve which vanishes for affine parameters)

  $$\zeta^\mu\nabla_\mu\zeta^\nu = \eta(\alpha)\zeta^\nu; \tag{1.4}$$

Proof. *If $f = const$ defines Σ we have, by* (1.1) *and by the fact that coordinate covariant derivatives of scalar functions commute,*

$$\zeta^\mu \nabla_\mu \zeta_\nu = \zeta^\mu \nabla_\mu \nabla_\nu f = \zeta^\mu \nabla_\nu \nabla_\mu f = \zeta^\mu \nabla_\nu \zeta_\mu = \frac{1}{2} \nabla_\nu (\zeta^\mu \zeta_\mu). \tag{1.5}$$

Being $\vec{\zeta}$ a vector field, the scalar quantity $\zeta^\mu \zeta_\mu$, which is zero on Σ being a null hypersurface, can be viewed as its defining function. Therefore the vector defined by $n^\mu = g^{\mu\nu} \nabla_\nu (\zeta^\sigma \zeta_\sigma)$ is orthogonal to Σ by (1.1)*. Now, because the vectors that are orthogonal to Σ must be proportional (they live in a one-dimensional space), we must have*

$$\nabla_\mu (\zeta^\nu \zeta_\nu) = C(\alpha) \nabla_\mu f = C(\alpha) \zeta_\mu,$$

where we used a proportionality constant $C(\alpha)$ which depends on the position on the curve. Therefore, by (1.5)*,*

$$\nabla_\mu (\zeta^\nu \zeta_\nu) = 2\zeta^\nu \nabla_\nu \zeta_\mu = C(\alpha) \zeta_\mu$$

and setting $\eta(\alpha) = \frac{1}{2} C(\alpha)$ we get the geodesic equation (1.4)*.*

- we can thus use vectors that are tangent to a geodesic on Σ as vectors orthogonal to Σ;
- since the vector field $\vec{\zeta}$ is defined everywhere on Σ, by the existence and unicity of the solution of the differential equation (1.3) given an initial point on Σ, the set of all null geodesics on Σ covers all the hypersurface, and the geodesics are called *generators* of the null hypersurface.

There is a particular class of vector fields called *hypersurface-orthogonal*, which are useful in general relativity:

Def. 1.2 (Hypersurface-orthogonality)**.** *A vector field $\vec{V}$ is said to be hypersurface orthogonal if there exists a foliation of the manifold into hypersurfaces whose normal vectors at each point are proportional to $\vec{V}$.*

In other words, if a vector field $\vec{V}$ is hypersurface-orthogonal, we can choose coordinates that cover the entire manifold and that define hypersurfaces by constraining one coordinate to be constant.

1.2 Killing vector fields and horizons

In spacetimes where Killing vector fields exist there are symmetries that can be exploited to obtain interesting results. Let us recall what a Killing vector field is.

Def. 1.3 (Killing vector field)**.** *A vector field $\vec{K}$ on a differentiable manifold is said to be a Killing vector field if the Lie derivative of the metric tensor along $\vec{K}$ vanishes:*

$$\pounds_{\vec{K}} g = 0. \tag{1.6}$$

In other words, the metric tensor stays constant along integral curves of $\vec{K}$. From this definition, we can obtain a useful identity that characterizes Killing vector fields, often referred to as *Killing equation*:

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = \nabla^\mu K^\nu + \nabla^\nu K^\mu = 0. \tag{1.7}$$

Proof. *For two arbitrary vector fields $\vec{A}$ and $\vec{B}$, since $g(\vec{A}, \vec{B})$ is a scalar, its Lie derivative coincides with its covariant derivative:*

$$\pounds_{\vec{V}} g(\vec{A}, \vec{B}) = \nabla_{\vec{V}} g(\vec{A}, \vec{B}).$$

Applying Leibniz rule on both sides and because $\nabla_{\vec{V}} g = 0$ because of metric compatibility we have

$$(\pounds_{\vec{V}} g)(\vec{A}, \vec{B}) + g(\pounds_{\vec{V}} \vec{A}, \vec{B}) + g(\vec{A}, \pounds_{\vec{V}} \vec{B}) = \cancel{(\nabla_{\vec{V}} g)(\vec{A}, \vec{B})} + g(\nabla_{\vec{V}} \vec{A}, \vec{B}) + g(\vec{A}, \nabla_{\vec{V}} \vec{B}).$$

Expliciting the components and isolating $\pounds_{\vec{V}} g$:

$$(\pounds_{\vec{V}} g)_{\mu\nu} = (\pounds_{\vec{V}} g)(\vec{e}_\mu, \vec{e}_\nu) = g(\nabla_{\vec{V}} \vec{e}_\mu, \vec{e}_\nu) + g(\vec{e}_\mu, \nabla_{\vec{V}} \vec{e}_\nu) - g(\pounds_{\vec{V}} \vec{e}_\mu, \vec{e}_\nu) - g(\vec{e}_\mu, \pounds_{\vec{V}} \vec{e}_\nu).$$

Expanding $\vec{V} = V^\alpha \vec{e}_\alpha$ we can write

$$\begin{aligned}
&\pounds_{\vec{V}} \vec{e}_\mu = [\vec{V}, \vec{e}_\mu] = -[\partial_\mu, V^\alpha \partial_\alpha] = -\left(\partial_\mu V^\alpha \partial_\alpha + \cancel{V^\alpha \partial_\mu \partial_\alpha} - \cancel{V^\alpha \partial_\alpha \partial_\mu}\right) = -\partial_\mu V^\alpha \vec{e}_\alpha; \\
&\nabla_{\vec{V}} \vec{e}_\mu = V^\alpha \nabla_\alpha \vec{e}_\mu = V^\alpha \Gamma^\sigma_{\mu\alpha} \vec{e}_\sigma.
\end{aligned}$$

If we substitute we get

$$\begin{aligned}
(\pounds_{\vec{V}} g)_{\mu\nu} &= g(V^\alpha \Gamma^\sigma_{\mu\alpha} \vec{e}_\sigma, \vec{e}_\nu) + g(\vec{e}_\mu, V^\alpha \Gamma^\sigma_{\nu\alpha} \vec{e}_\sigma) - g(-\partial_\mu V^\alpha \vec{e}_\alpha, \vec{e}_\nu) - g(\vec{e}_\mu, -\partial_\nu V^\alpha \vec{e}_\alpha) \\
&= (\partial_\mu V^\alpha g_{\alpha\nu} + V^\alpha \Gamma^\sigma_{\mu\alpha} g_{\sigma\nu}) + (\partial_\nu V^\alpha g_{\mu\alpha} + V^\alpha \Gamma^\sigma_{\nu\alpha} g_{\mu\sigma}) \\
&= (\partial_\mu V^\sigma + V^\alpha \Gamma^\sigma_{\mu\alpha}) g_{\sigma\nu} + (\partial_\nu V^\sigma + V^\alpha \Gamma^\sigma_{\nu\alpha}) g_{\mu\sigma} \\
&= g_{\nu\sigma} \nabla_\mu V^\sigma + g_{\mu\sigma} \nabla_\nu V^\sigma \\
&= \nabla_\mu V_\nu + \nabla_\nu V_\mu = 0,
\end{aligned}$$

where the last equality follows from (1.6)*, and because of metric compatibility we can also raise both indices together.*

Of course, a Killing vector field that encodes a symmetry of the spacetime (metric) is defined up to a proportionality factor, since the parameterization of the integral curves along which the metric stays constant does not matter physically.

If a spacetime admits a hypersurface-orthogonal timelike Killing vector field as in definition 1.2 it is said static, and the following theorem holds.

Theorem 1.1. *A spacetime admits a hypersurface-orthogonal timelike Killing vector field $\vec{K}$ if and only if $\vec{K}$ is defined everywhere and, choosing the time coordinate along $\vec{K}$ everywhere, the metric tensor is globally block-diagonal with time and space components that do not mix.*

Proof. *Suppose $\vec{K}$ is a hypersurface-orthogonal timelike Killing vector field and let us pick a coordinate system $x^\alpha = \{t, x^i\}$ where $\vec{e}_t = \vec{K} = \partial_t$. We can do this everywhere because, by definition 1.2, $\vec{K}$ is orthogonal to a spacelike foliation of the spacetime. Since the metric tensor is symmetric, $g_{\mu\nu} = g_{\nu\mu}$, and there are $n(n+1)/2$ independent components, so we can consider only the upper triangular part of the matrix $g_{\mu\nu}$ other than the diagonal elements. If we change coordinate system by flipping the sign of the time component $x^\alpha = \{t, x^i\} \leftrightarrow \{-t, x^i\} = y^\alpha$ we have a transformation matrix that looks like diag$(-1, 1, \dots, 1)$ and since the metric tensor is of type $(0, 2)$ its components transform like*

$$g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial y^\mu}\frac{\partial x^\beta}{\partial y^\nu} g_{\alpha\beta}.$$

Let us fix $\mu = 0$ and see what happens for $\nu = i = 1, \dots, n$:

$$g'_{0i} = \frac{\partial x^\alpha}{\partial y^0}\frac{\partial x^\beta}{\partial y^i} g_{\alpha\beta} = \frac{\partial x^0}{\partial y^0}\frac{\partial x^i}{\partial y^i} g_{0i} = (-1)(1)g_{0i} = -g_{0i},$$

where there is no sum over i. Now, since $\vec{K} = \partial_t$ is a Killing vector, the metric is invariant under the change of coordinates that we performed, because only the temporal coordinate changes. Therefore we must have

$$g_{0i} = g'_{0i} = -g_{0i} \quad \Rightarrow \quad g_{0i} = 0,$$

which proves that the off-diagonal temporal components of the metric tensor vanish globally.

Conversely, if $\vec{K}$ is defined on all the spacetime and the metric is globally block-diagonal with $\vec{K}$ defining the time coordinate, it means that each tangent space can be written as a direct sum of the one-dimensional subspace generated by $\vec{K}$ and a $(n-1)$-dimensional spacelike hyperplane with induced metric given by the spacelike block of *the metric. This in turn defines a spacelike foliation, because the aforementioned $(n-1)$-dimensional hyperplanes (one for each tangent space) can be viewed as tangent spaces of spacelike hypersurfaces that are everywhere orthogonal to the Killing vector field $\vec{K}$.*

Spacetimes can also have another interesting property, that is being *asymptotically flat*. In order to define this notion rigorously we would need to delve into conformal transformations, which exit the scope of this dissertation. We can however say that a spacetime is said to be asymptotically flat if there exists a coordinate system in which it is clear what spacial infinity means and the metric tensor is constant and equals the Minkowski one at spacial infinity. An example is the Schwarzschild spacetime, where we use spherical coordinates for the spacial position, and it is asymptotically flat because the metric tends to Minkowski for $r \to \infty$.

If a spacetime is static and asymptotically flat it is always possible to normalize the Killing vector field (since it is defined up to a constant multiplicative scalar factor):

$$K^\mu K_\mu(r \to \infty) = -1, \tag{1.8}$$

so that it represents the direction of the proper time of some observer at spacial infinity $\vec{K} = \partial_t$.

Let us now introduce Killing horizons.

Def. 1.4 (Killing horizon)**.** *Let Σ be a null hypersurface and $\vec{\chi}$ a Killing vector field defined at least on a neighbourhood of Σ. If $\vec{\chi}$ is null everywhere on Σ then Σ is said to be a Killing horizon.*

From this definition and the properties of null hypersurfaces discussed previously, we can immediately say that the constraint $\chi^\mu \chi_\mu = 0$ defines the hypersurface. In addition, the following theorem holds:

Theorem 1.2. *The Killing vector field $\vec{\chi}$ that defines a Killing horizon Σ is orthogonal to Σ (and also tangent).*

Proof. *By Σ being a null hypersurface, the normal vector $\zeta^\mu = g^{\mu\nu}\nabla_\nu(\chi^\sigma\chi_\sigma)$ given by* (1.1) *is also tangent to Σ. In addition, $\vec{\chi}$ and $\vec{\zeta}$ are orthogonal, in fact*

$$\chi^\mu \zeta_\mu = \chi^\mu \nabla_\mu(\chi^\nu \chi_\nu) = 2\chi^\mu \chi^\nu \nabla_\mu \chi_\nu = \chi^\mu \chi^\nu (\nabla_\mu \chi_\nu + \nabla_\nu \chi_\mu) = 0,$$

where we used the Killing equation (1.7) *in the last step. Therefore, $\vec{\chi}$ is tangent to Σ and by being null it is also orthogonal to itself, and hence to Σ too.*

1.3 Redshift factor

If we have a timelike Killing vector field $\vec{K}$ and a particle (even massless) with four-momentum $\vec{p}(\lambda)$ moving along a geodesic affinely parameterized by λ, the following quantity, called *Killing energy*, is conserved:

$$E_K = -p^\mu K_\mu. \tag{1.9}$$

Proof. *The (covariant) derivative of E_K along the direction given by the geodesic is*

$$\frac{D}{d\lambda} E_K = -\nabla_{\vec{p}}(K^\mu p_\mu) = -p^\nu \nabla_\nu (K^\mu p_\mu) = -K^\mu \underbrace{(p^\nu \nabla_\nu p_\mu)}_{=0} - \cancel{p^\nu p^\mu \nabla_\nu K_\mu} = 0,$$

where the first term is zero due to the geodesic equation, and the second vanishes because it is the product of a symmetric tensor and an antisymmetric tensor (due to the Killing equation (1.7)*).*

Consider a photon travelling along a geodesic with four-momentum $\vec{p}$ in a local inertial reference frame S. We know from special relativity that its energy measured locally by an observer O with four-velocity $\vec{U}$ is (we use natural units, so $\hbar = 1$ and ω is the frequency)

$$\omega = -p^\mu U_\mu. \tag{1.10}$$

Proof. *In the reference frame S_O where O is at rest, the first component of the photon's four-momentum is its energy in that frame, and taking the Minkowski scalar product with the four-velocity of O (which is* $(1,0,0,0)$ *since O is at rest) simply gives that energy with a minus sign, and because the result is Lorentz-invariant, this computation can be performed using four-vectors $\vec{p}$ and $\vec{U}$ in any frame of reference S, as the one we used for* (1.10)*.*

Now, let us consider an asymptotically flat static spacetime. We then have a normalized timelike Killing vector field $\vec{K} = \partial_t$ where t is the proper time of some stationary observer at spatial infinity. In general, *stationary observers* (also called *static*) are defined as travelling along orbits of $\vec{K}$ (not necessarily along geodesics), meaning that their four-velocity $\vec{U}$ is proportional to $\vec{K}$ with a proportionality factor $1/V(P)$ that depends on the position in spacetime:

$$K^\mu = V(P) U^\mu. \tag{1.11}$$

This yields a straightforward relation that can be used to compute V, which is also called the *redshift factor* for reasons we will see in a moment:

$$V = \sqrt{-K^\mu K_\mu}. \tag{1.12}$$

Proof. *Contracting both sides of* (1.11) *with* K_μ *we get*

$$K^\mu K_\mu = VU^\mu K_\mu = V^2 U^\mu U_\mu = -V^2,$$

where we used again (1.11) *with lowered index* μ *and the normalization of four-velocity* $U^\mu U_\mu = -1$*. We can then explicit* V *and obtain* (1.12).

Consider a photon travelling along a null geodesic affinely parameterized by λ with four-momentum $\vec{p}(\lambda)$ and two stationary observers O_1 and O_2, whose four-velocities are $\vec{U}_{(1)}$ and $\vec{U}_{(2)}$ with redshift factors $V_1(P_1)$ and $V_2(P_2)$ respectively. Their measurements ω_1 and ω_2 of the energy of the photon will then be related by the factors V_1 and V_2 as

$$\omega_1 V_1 = \omega_2 V_2. \tag{1.13}$$

Proof. *Using* (1.10), (1.11) *and* (1.9) *we can express the energy of the photon as measured by both observers in terms of the Killing energy and the redshift factor* $(i = 1, 2)$*:*

$$\omega_i = -p_\mu U^\mu_{(i)} = -\frac{p_\mu K^\mu}{V_i} = \frac{E_K}{V_i}.$$

Since the Killing energy E_K *is conserved, we can write*

$$E_K = \omega_1 V_1 = \omega_2 V_2.$$

In the case of Schwarzschild spacetime, whose metric gives the line element (in spherical spacial coordinates $\{t, r, \theta, \varphi\}$ and with M being the mass of the spherically symmetric gravitational source)

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2, \tag{1.14}$$

we can compute the redshift factor directly, knowing that the Killing vector field is, trivially, $\vec{K} = \partial_t$, yielding

$$V = \sqrt{1 - \frac{2GM}{r}}. \tag{1.15}$$

Proof. *Using* (1.12) *and expressing the Killing vector field using components* $\vec{K} = (1, 0, 0, 0)$ *in the coordinate basis* $\{\partial_t, \partial_r, \partial_\theta, \partial_\varphi\}$ *we have*

$$V = \sqrt{-K^\mu K_\mu} = \sqrt{-g_{tt}} = \sqrt{1 - \frac{2GM}{r}},$$

where we used the time diagonal component of the metric (1.14).

It is easy to see that in this case, V varies continuously between 0 (at $r = 2GM$) and 1 (for $r \to \infty$). If a photon travelling on an outgoing radial geodesic has energy ω_0 at a certain radial coordinate r_0 (with redshift factor V_0) according to a stationary observer standing there, using (1.12) we can compute its energy when it reaches a stationary observer at spacial infinity:

$$\omega_\infty = V_0 \omega_0 < \omega_0,$$

so we see that the energy decreases as the photon climbs up the gravitational field. This is an example of gravitational redshift.

Since stationary observers do not travel along geodesics in general, they have a four-acceleration defined by

$$a^\mu = U^\sigma \nabla_\sigma U^\mu, \tag{1.16}$$

where $\vec{U}$ is their four-velocity. The following relation holds:

$$a_\mu = \nabla_\mu \ln V = \frac{1}{V} \nabla_\mu V. \tag{1.17}$$

Proof. *By conservation of the modulus of four-velocity,*

$$\nabla_\nu (U^\mu U_\mu) = 0 \quad \Rightarrow \quad U_\mu \nabla_\nu U^\mu = U^\mu \nabla_\nu U_\mu = 0.$$

By using the above and the proportionality relation (1.11) *for stationary observers, we can manipulate the Killing equation* (1.7)*:*

$$\begin{aligned}
&\nabla_\mu (V U_\nu) + \nabla_\nu (V U_\mu) = 0 \\
&(\nabla_\mu V) U_\nu + V (\nabla_\mu U_\nu) + (\nabla_\nu V) U_\mu + V (\nabla_\nu U_\mu) = 0 \\
&(\nabla_\mu V) U^\nu U_\nu + V \underbrace{U^\nu (\nabla_\mu U_\nu)}_{=0} + (\nabla_\nu V) U^\nu U_\mu + V U^\nu (\nabla_\nu U_\mu) = 0 \\
&- \nabla_\mu V + (\nabla_\nu V) U^\nu U_\mu + V U^\nu (\nabla_\nu U_\mu) = 0 && (1.18) \\
&- U^\mu \nabla_\mu V + U^\nu U^\mu U_\mu (\nabla_\nu V) + V U^\nu \underbrace{U^\mu (\nabla_\nu U_\mu)}_{=0} = 0 \\
&- U^\mu \nabla_\mu V - U^\nu \nabla_\nu V = 0 \\
&U^\mu \nabla_\mu V = 0, && (1.19)
\end{aligned}$$

where we contracted with U^ν in the third row and with U^μ in the fifth. The second term of (1.18) *cancels because of* (1.19)*:*

$$-\nabla_\mu V + V U^\nu (\nabla_\nu U_\mu) = 0,$$

and therefore, by (1.16),

$$-\nabla_\mu V + V a_\mu = 0,$$

so finally, rearranging,

$$a_\mu = \frac{1}{V}\nabla_\mu V = \nabla_\mu \ln V.$$

Before moving to surface gravity, we derive a useful identity for Killing vector fields (the braces over the indices stand for anti-symmetrization):

$$3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]}) = \chi^\alpha\chi_\alpha(\nabla^\beta\chi^\gamma)(\nabla_\beta\chi_\gamma) - 2(\chi^\alpha\nabla^\beta\chi^\gamma)(\chi_\beta\nabla_\alpha\chi_\gamma). \tag{1.20}$$

Proof. *Let us compute the left-hand side directly, using also the Killing equation* (1.7)*:*

$$\begin{aligned}
&3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]})\\
&= \frac{3}{36}\begin{pmatrix}+\chi^\alpha\nabla^\beta\chi^\gamma - \chi^\alpha\nabla^\gamma\chi^\beta\\ +\chi^\beta\nabla^\gamma\chi^\alpha - \chi^\beta\nabla^\alpha\chi^\gamma\\ +\chi^\gamma\nabla^\alpha\chi^\beta - \chi^\gamma\nabla^\beta\chi^\alpha\end{pmatrix}\begin{pmatrix}+\chi_\alpha\nabla_\beta\chi_\gamma - \chi_\alpha\nabla_\gamma\chi_\beta\\ +\chi_\beta\nabla_\gamma\chi_\alpha - \chi_\beta\nabla_\alpha\chi_\gamma\\ +\chi_\gamma\nabla_\alpha\chi_\beta - \chi_\gamma\nabla_\beta\chi_\alpha\end{pmatrix}\\
&= \frac{1}{3}\big(\chi^\alpha\nabla^\beta\chi^\gamma + \chi^\beta\nabla^\gamma\chi^\alpha + \chi^\gamma\nabla^\alpha\chi^\beta\big)\big(\chi_\alpha\nabla_\beta\chi_\gamma + \chi_\beta\nabla_\gamma\chi_\alpha + \chi_\gamma\nabla_\alpha\chi_\beta\big)\\
&= \frac{1}{3}\big[3\chi^\alpha\chi_\alpha(\nabla^\beta\chi^\gamma)(\nabla_\beta\chi_\gamma) + 6\chi^\alpha\chi_\beta(\nabla^\beta\chi^\gamma)(\nabla_\gamma\chi_\alpha)\big]\\
&= \chi^\alpha\chi_\alpha(\nabla^\beta\chi^\gamma)(\nabla_\beta\chi_\gamma) - 2\chi^\alpha\chi_\beta(\nabla^\beta\chi^\gamma)(\nabla_\alpha\chi_\gamma),
\end{aligned}$$

where we have relabelled dummy indices multiple times in the multiplication in the second-last step.

1.4 Surface gravity

Each Killing horizon has an associated quantity called *surface gravity*, often indicated with κ. From what we have learned so far, a Killing horizon is a null hypersurface Σ which has a Killing vector field $\vec{\chi}$ as the orthogonal and tangent vector at each point. In addition, integral curves of $\vec{\chi}$ on Σ do not leave the hypersurface and are null geodesics, therefore they satisfy the geodesic equation

$$\chi^\mu\nabla_\mu\chi^\nu = -\kappa\chi^\nu, \tag{1.21}$$

where the right-hand side is non-zero because the geodesic might not be affinely parameterized due to the arbitrariness of the norm of $\vec{\chi}$, and we used $-\kappa$ as a scalar proportionality constant, which is by definition the opposite of surface gravity. The following relation for the computation of κ holds:

$$\kappa^2 = -\frac{1}{2}(\nabla_\mu \chi_\nu)(\nabla^\mu \chi^\nu). \tag{1.22}$$

Proof. *Since the Killing vector field $\vec{\chi}$ is orthogonal to the hypersurface Σ, we can use* (1.2)*, and by employing the Killing equation* (1.7)*,*

$$\begin{aligned} 6\chi_{[\alpha}\nabla_\beta\chi_{\gamma]} &= +\chi_\alpha\nabla_\beta\chi_\gamma - \chi_\gamma\nabla_\beta\chi_\alpha \\ &\quad + \chi_\beta\nabla_\gamma\chi_\alpha - \chi_\beta\nabla_\alpha\chi_\gamma \\ &\quad + \chi_\gamma\nabla_\alpha\chi_\beta - \chi_\alpha\nabla_\gamma\chi_\beta \\ &= +\chi_\alpha\nabla_\beta\chi_\gamma + \chi_\gamma\nabla_\alpha\chi_\beta \\ &\quad + \chi_\beta\nabla_\gamma\chi_\alpha + \chi_\beta\nabla_\gamma\chi_\alpha \\ &\quad + \chi_\gamma\nabla_\alpha\chi_\beta + \chi_\alpha\nabla_\beta\chi_\gamma \\ &= 2\big[\chi_\alpha\nabla_\beta\chi_\gamma + \chi_\beta\nabla_\gamma\chi_\alpha + \chi_\gamma\nabla_\alpha\chi_\beta\big] = 0, \end{aligned}$$

which becomes, again by using the Killing equation,

$$\chi_\gamma\nabla_\alpha\chi_\beta = -\chi_\alpha\nabla_\beta\chi_\gamma + \chi_\beta\nabla_\alpha\chi_\gamma.$$

If we contract both sides with $\nabla^\alpha\chi^\beta$ and reorder factors in each term we get

$$\chi_\gamma(\nabla^\alpha\chi^\beta)(\nabla_\alpha\chi_\beta) = -(\chi_\alpha\nabla^\alpha\chi^\beta)(\nabla_\beta\chi_\gamma) + \chi_\beta(\nabla^\alpha\chi^\beta)(\nabla_\alpha\chi_\gamma).$$

Now, if we use the geodesic equation (1.21) *and the Killing equation* (1.7) *repeatedly,*

$$\begin{aligned} \chi_\gamma(\nabla^\alpha\chi^\beta)(\nabla_\alpha\chi_\beta) &= \kappa(\chi^\beta\nabla_\beta\chi_\gamma) - (\chi_\beta\nabla^\beta\chi^\alpha)\nabla_\alpha\chi_\gamma \\ &= -\kappa^2\chi^\gamma + \kappa(\chi^\alpha\nabla_\alpha\chi^\gamma) \\ &= -\kappa^2\chi^\gamma - \kappa^2\chi^\gamma = -2\kappa^2\chi^\gamma, \end{aligned}$$

and, since the above holds for all χ^γ, (1.22) *follows.*

The physical interpretation of surface gravity is possible in an asymptotically flat static spacetime. In this case, κ represents the acceleration of a stationary observer O just above the event horizon, as measured by a stationary observer O' at spatial infinity. In other words it is the limit of the modulus a of proper acceleration of O multiplied by the redshift factor V (evaluated at the point where O is) as O approaches the Killing horizon Σ:

$$\kappa = \lim_{O\to\Sigma} Va. \tag{1.23}$$

Proof. *We need to compute the following derivative using* (1.2)*:*

$$\nabla_\mu\big[3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]})\big] = 6\,\underbrace{(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})}_{=0}\,\nabla_\mu(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]}) = 0.$$

In addition, we notice that, by using the Killing equation (1.7) *and the geodesic equation* (1.21)*, the following derivative is non-zero (unless* $\kappa = 0$*):*

$$\nabla_\mu(\chi^\alpha\chi_\alpha) = 2\chi^\alpha\nabla_\mu\chi_\alpha = -2\chi^\alpha\nabla_\alpha\chi_\mu = 2\kappa\chi_\mu \neq 0 \quad \Leftrightarrow \quad \kappa \neq 0.$$

Using these results we can compute the following limit by using L'Hôpital's rule (since it is an indeterminate form due to (1.2) *and the fact that* $\vec{\chi}$ *is null on the horizon):*

$$\lim_{O\to\Sigma}\frac{3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]})}{\chi^\alpha\chi_\alpha} = \lim_{O\to\Sigma}\frac{\nabla_\mu\big[3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]})\big]}{\nabla_\mu(\chi^\alpha\chi_\alpha)} = 0.$$

Therefore, by (1.20) *and* (1.22)*,*

$$\lim_{O\to\Sigma}\frac{3(\chi^{[\alpha}\nabla^\beta\chi^{\gamma]})(\chi_{[\alpha}\nabla_\beta\chi_{\gamma]})}{\chi^\alpha\chi_\alpha} = \lim_{O\to\Sigma}\left[\underbrace{(\nabla^\beta\chi^\gamma)(\nabla_\beta\chi_\gamma)}_{-2\kappa^2} - \frac{2(\chi^\alpha\nabla^\beta\chi^\gamma)(\chi_\beta\nabla_\alpha\chi_\gamma)}{\chi^\sigma\chi_\sigma}\right] = 0,$$

so we can write, by (1.16)*,* (1.11) *and* (1.12)*,*

$$\begin{aligned}\kappa^2 &= -\lim_{O\to\Sigma}\frac{(\chi^\beta\nabla_\beta\chi^\gamma)(\chi^\alpha\nabla_\alpha\chi_\gamma)}{\chi^\sigma\chi_\sigma}\\ &= \lim_{O\to\Sigma}\frac{V^4(U^\beta\nabla_\beta U^\gamma)(U^\alpha\nabla_\alpha U_\gamma)}{V^2}\\ &= \lim_{O\to\Sigma}V^2a^\gamma a_\gamma = \lim_{O\to\Sigma}V^2a^2.\end{aligned}$$

We can compute the surface gravity of a Schwarzschild black hole by first calculating the modulus of four-acceleration of a stationary observer ($R_H = 2GM$ is the Schwarzschild radius):

$$a = \frac{R_H}{2r^2\sqrt{1-\frac{R_H}{r}}}. \tag{1.24}$$

Proof. *Using* (1.17)*, the metric* (1.14) *and the Schwarzschild redshift factor* (1.15)*,*

$$a_\mu = \frac{1}{V}\nabla_\mu V = \left(\sqrt{1-\frac{R_H}{r}}\right)^{-1} \frac{1}{2\sqrt{1-\frac{R_H}{r}}} R_H r^{-2}\nabla_\mu r = \frac{R_H}{2r^2\left(1-\frac{R_H}{r}\right)}\delta^r_\mu .$$

Therefore, using the inverse metric to raise the index,

$$a = \sqrt{a_\mu a^\mu} = \sqrt{g^{rr}}a_r = \sqrt{1-\frac{R_H}{r}}\frac{R_H}{2r^2\left(1-\frac{R_H}{r}\right)} = \frac{R_H}{2r^2\sqrt{1-\frac{R_H}{r}}}.$$

Therefore the surface gravity is

$$\kappa = \frac{1}{2R_H} = \frac{1}{4GM}. \tag{1.25}$$

Proof. *Using* (1.23)*,* (1.15) *and* (1.24) *we get*

$$\kappa = \lim_{r\to R_H} Va = \lim_{r\to R_H}\sqrt{1-\frac{R_H}{r}}\frac{R_H}{2r^2\sqrt{1-\frac{R_H}{r}}} = \lim_{r\to R_H}\frac{R_H}{2r^2} = \frac{1}{2R_H}.$$

These are all the general relativity tools that are necessary to study the analogy of Hawking radiation with the Unruh effect for Schwarzschild black holes. What we need now is to better understand how particles are described mathematically as excitation of fields, and what the curved geometry of spacetime implies for the observation of particles. In the next chapter, we are going to take a brief tour of quantum field theory, and we will be able to answer these questions before applying them to the study of Hawking radiation.

Chapter 2

Quantum fields, particles and vacuum

Quantum field theory is a complex subject, whose aim is to study field quantities (scalar, spinor, and vector fields in particular are used in the standard model of particle physics) and quantize them with various approaches. One of them is postulating an expression for the lagrangian, promoting the field to a quantum operator, imposing canonical commutation relations between the field and its conjugate momentum, and studying the solutions of the equations of motion. From this procedure, we will see that there is a natural way to define the concepts of vacuum and particle, and this approach does indeed work since experiments that have been made in particle physics can be fully explained by this framework. For our purposes, we are going to focus on the case of a real scalar field. For a more detailed reference on quantum field theory see [6].

2.1 Klein-Gordon equation

Let us fix a coordinate system $\{t, \boldsymbol{x}\}$ in $\mathbb{R}^{1+n}$ (in our universe of course $n = 3$) and suppose we are given a real scalar field $\phi : \mathbb{R}^{1+n} \to \mathbb{R}$ subject to a harmonic potential

$$V(\phi) = \frac{1}{2} m^2 \phi^2. \tag{2.1}$$

Our task is now to derive the equations of motion for such a field. To do that we might want to rely on the stationary action principle, where the action S is defined as usual in terms of the integral over time of a lagrangian L yet to be defined:

$$S = \int_0^t L(\phi, \partial_\mu \phi, t') dt'. \tag{2.2}$$

In this case, it is useful to express the lagrangian L in terms of the lagrangian spacial density $\mathcal{L}$:

$$L = \int \mathcal{L}(\phi, \partial_\mu \phi, t') d^n x. \tag{2.3}$$

Then, by requiring that the action stays stationary $\delta S = 0$ under small variations of the field and its derivatives

$$\begin{aligned} &\phi \to \phi + \delta\phi \\ &\partial_\mu \phi \to \partial_\mu \phi + \delta(\partial_\mu \phi) \end{aligned}$$

we obtain the Euler-Lagrange equation for our field:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = 0. \tag{2.4}$$

Since the action is a scalar we need the lagrangian density to be Lorentz-invariant after being multiplied by $d^n x$ and dt, so it must be written in tensorial form using the field ϕ (which is already a scalar) and its first derivatives $\partial_\mu \phi$.

By analogy with the case of an elastic string in two dimensions extending on one spacial axis x and having each point constrained to move only along the y direction (fig. 2.1), we have three energy density contributions:

- **kinetic energy**, arising from the motion of a point through space (the y axis in the string analogy) over time:

$$\frac{1}{2}(\partial_t \phi)^2 \quad \longleftrightarrow \quad \frac{1}{2}(\partial_t y)^2;$$

- **gradient potential energy**, which comes from the elastic interaction of a point with its neighbors due to the elasticity of the string *itself*:

$$\frac{1}{2}(\boldsymbol{\nabla} \phi)^2 \quad \longleftrightarrow \quad \frac{1}{2}(\partial_x y)^2;$$

- **pure potential energy**, which describes the interaction of the field with the environment due to its setup and external conditions. In the analogy of the string, we can think that in addition to the elasticity of the string itself, there is a spring attached to each point of the string subjecting it to an elastic potential of the form (2.1) (as depicted in fig. 2.1):

$$V(\phi) = \frac{1}{2} m^2 \phi^2 \quad \longleftrightarrow \quad V(y) = \frac{1}{2} m^2 y^2.$$

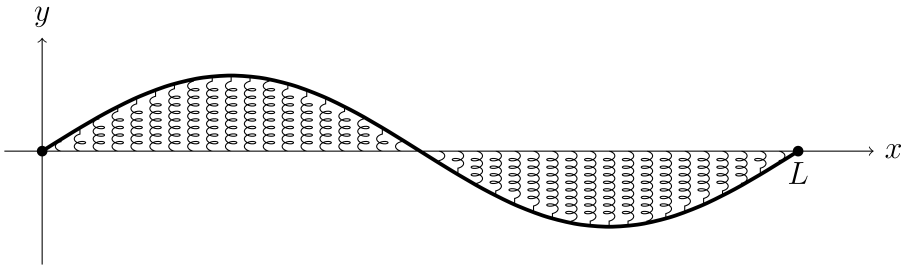


Fig. 2.1: Elastic string of rest length L. Each infinitesimal element of the string is constrained to move only along the y axis and feels a force that is proportional to the difference in height between itself and its neighbor elements. Furthermore, there is an additional elastic force represented by springs pulling each string element towards the equilibrium position $y = 0$.

Since the lagrangian can be written in this case as kinetic minus potential energy, we would need to be able to write, using tensor formalism, some expression that resembles the following:

$$\mathcal{L} = \frac{1}{2}(\partial_t \phi)^2 - \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 - \frac{1}{2}m^2\phi^2.$$

We observe that this is easily castable in tensorial form, yielding the Klein-Gordon lagrangian density in flat spacetime:

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}(\partial_\mu \phi)(\partial_\nu \phi) - \frac{1}{2}m^2\phi^2. \tag{2.5}$$

In natural units, we set $c = 1$, $\hbar = 1$ and $k_B = 1$, so dimensionally we have (by $E = mc^2$, $E = \hbar\omega$ and the equivalence of lengths and times due to the dimensionless velocity $c = 1$)

$$[energy] = [mass] = [length^{-1}] = [time^{-1}].$$

The action (which is $[energy \cdot time]$) becomes thus dimensionless. Since the lagrangian density is integrated in d^3x (in our universe) and then in dt (in natural units it is overall d^4x) to obtain the dimensionless action, and $[d^4x] = [length^4]$, it follows that the lagrangian density has the dimensions of $[mass^4]$. To meet this requirement, the field ϕ and the constant m both need to have the dimension of $[mass]$, and that is why m is called the mass constant of the field, which will become the mass of the particles upon quantization.

Now we are ready to use the Euler-Lagrange equation (2.4) to derive the equation of motion, which turns out to be the Klein-Gordon equation:

$$\Box\phi - m^2\phi = 0. \tag{2.6}$$

Proof. *We need to compute the two terms in* (2.4) *from* (2.5)*. The derivative of* $\mathcal{L}$ *with respect to* ϕ *is easy to compute:*

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2\phi,$$

whereas we treat the other term by steps:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_0\phi)} &= \partial_0\phi \quad \xrightarrow{\partial_0} \quad \partial_0^2\phi, \\ \frac{\partial \mathcal{L}}{\partial(\partial_i\phi)} &= -\partial_i\phi \quad \xrightarrow{\partial_i} \quad -\partial_i^2\phi. \end{aligned}$$

So the result is, by writing all the terms in the summation and using the signature $(-,+,+,+)$*,*

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)}\right) = \partial_0^2\phi - \sum_{i=1}^{n}\partial_i^2\phi = -\Box\phi.$$

By filling in the terms in the Euler-Lagrange equation we get

$$-\Box\phi + m^2\phi = 0,$$

which yields (2.6) *after a change of sign.*

2.1.1 Solutions

By its form (2.6), we immediately notice that the Klein-Gordon equation is linear in ϕ, hence we can reduce to find a basis of solutions and then write the general solution as a (possibly generalized, in case of an uncountable set of basic solutions) linear combination of elements of the basis with the boundary condition that the field vanishes at spacial infinity:

$$\phi(t, \boldsymbol{x}) \to 0 \qquad for\ |\boldsymbol{x}| \to \infty. \tag{2.7}$$

The basic solutions are given by plane waves:

$$\phi(t, \boldsymbol{x}) = \phi_0 e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{2.8}$$

where $\boldsymbol{k} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and the following dispersion relation holds:

$$\omega^2 = k^2 + m^2. \tag{2.9}$$

Proof. *Since there are no mixed derivatives, we can search for solutions in a factorized form*

$$\phi(t, \boldsymbol{x}) = \alpha(t)\beta(\boldsymbol{x}).$$

Substituting in (2.6) *we get*

$$-\alpha''\beta + \alpha\nabla^2\beta - m^2\alpha\beta = 0, \tag{2.10}$$

where α'' *denotes the second total derivative of* α*. By keeping* $\boldsymbol{x}$ *fixed and letting* t *vary we can multiply by* $-\beta^{-1}$ *(with the implicit assumption that* $\beta(\boldsymbol{x}) \neq 0$*, otherwise the solution for* ϕ *would trivially be* 0 *in that point):*

$$\alpha'' + \underbrace{\left(m^2 - \frac{1}{\beta}\nabla^2\beta\right)}_{\omega^2}\alpha = 0, \tag{2.11}$$

where we have called the quantity inside the parentheses ω^2 *since it is constant. The solutions of the above differential equation are given by linear combinations of positive and negative imaginary exponentials, and since* ω *has the freedom to be both positive and negative once its square is fixed we might as well choose one sign convention:*

$$\alpha(t) = e^{-i\omega t}. \tag{2.12}$$

Substituting $\alpha'' = -\omega^2\alpha$ *(given by* (2.11)*) into* (2.10)*, letting only* $\boldsymbol{x}$ *vary and* t *staying fixed, we get*

$$\omega^2\alpha\beta + \alpha\nabla^2\beta - m^2\alpha\beta = 0.$$

Dividing by α *(again assuming* $\alpha \neq 0$ *at that particular time, otherwise* ϕ *would trivially be* 0*) we obtain*

$$\nabla^2\beta + \underbrace{(\omega^2 - m^2)}_{k^2}\beta = 0,$$

whose solutions are linear combinations of plane waves with wave vector $\boldsymbol{k}$ *(provided* $\boldsymbol{k} \neq \boldsymbol{0}$*) such that* $k^2 = \omega^2 - m^2$ *(which is the dispersion relation* (2.9)*):*

$$\beta(\boldsymbol{x}) = e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{2.13}$$

If $\boldsymbol{k} = \boldsymbol{0}$*, the form of the solution would be*

$$\beta(\boldsymbol{x}) = \boldsymbol{a}\cdot\boldsymbol{x} + b,$$

but the only way to meet the boundary condition (2.7) *would be that both* $\boldsymbol{a}$ *and* b *equal* 0*, thus yielding* $\phi = 0$*, which we are not interested in because it would not be* *linearly independent from the other solutions in the basis. An analogous argument holds for $\omega = 0$ (only possible if $m = 0$). By combining* (2.12) *and* (2.13) *we get* (2.8) *with an arbitrary coefficient ϕ_0 allowed by linearity and to be determined by some normalization convention.*

We notice that if we require ω and $\boldsymbol{k}$ to be real and (restoring standard units) use $E = \hbar\omega$ and $\boldsymbol{p} = \hbar\boldsymbol{k}$, the dispersion relation (2.9) becomes the energy-momentum relation of special relativity, thus recovering a fundamental equation to describe the dynamics of our system (the choice of the lagrangian is cooked such that this equation is recovered):

$$E^2 = m^2c^4 + p^2c^2. \tag{2.14}$$

It is important to emphasize that the set of all the basic solutions of the Klein-Gordon equation is given by (2.8) for every possible value of $\boldsymbol{k} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$, and yet for each $\boldsymbol{k}$ there are two basic solutions distinguished by the sign of the temporal part in the exponential, where ω is bound to $\boldsymbol{k}$ by the dispersion relation (2.9) (where both ω and $\boldsymbol{k}$ appear squared, thus explaining the reason of the sign ambiguity). We can also write down the basic solutions (2.8) in an alternative way using tensor notation with $k^\mu = (\omega, \boldsymbol{k})$:

$$\phi(x^\mu) = \phi_0 e^{ik_\mu x^\mu}, \tag{2.15}$$

where ω is allowed to be both positive and negative once $\boldsymbol{k}$ is fixed (and so is ω^2). Furthermore, we shall stress that the basic solutions do not satisfy the boundary condition (2.7) although being bounded, and are complex-valued. This imposes constraints on the complex coefficients of linear combinations when expressing a general solution so that it is real and satisfies the boundary condition (2.7).

2.1.2 Positive and negative frequency modes

Let us define an indefinite inner product in the linear space of solutions of the Klein-Gordon equation. Given two solutions ϕ_1 and ϕ_2 and a hypersurface Σ_t with constant t in the currently used reference frame, we define their inner product as

$$\langle \phi_1, \phi_2 \rangle = -i \int_{\Sigma_t} (\phi_1 \partial_0 \phi_2^* - \phi_2^* \partial_0 \phi_1) d^n x, \tag{2.16}$$

and it does not depend on the value of t defining the hypersurface Σ_t.

Proof. *First, let us prove that the properties of indefinite inner products are satisfied. We need to show conjugate symmetry and linearity in the first argument. Indeed,* *linearity is proven by*

$$\begin{aligned}\langle a\phi_a + b\phi_b, \phi_2\rangle &= -i\int_{\Sigma_t}[a(\phi_a\partial_0\phi_2^* - \phi_2^*\partial_0\phi_a) + b(\phi_b\partial_0\phi_2^* - \phi_2^*\partial_0\phi_b)]d^3x \\ &= a\langle\phi_a, \phi_2\rangle + b\langle\phi_b, \phi_2\rangle,\end{aligned}$$

while conjugate symmetry follows from

$$\begin{aligned}\langle\phi_1, \phi_2\rangle^* &= i\int_{\Sigma_t}(\phi_1^*\partial_0\phi_2 - \phi_2\partial_0\phi_1^*)d^3x \\ &= -i\int_{\Sigma_t}(\phi_2\partial_0\phi_1^* - \phi_1^*\partial_0\phi_2)d^3x = \langle\phi_2, \phi_1\rangle.\end{aligned}$$

We now need to prove independence from t*. Let us define the quantity*

$$J_\mu(\phi_1, \phi_2) = -i(\phi_1\partial_\mu\phi_2^* - \phi_2^*\partial_\mu\phi_1). \tag{2.17}$$

By using the Leibniz rule, we can compute the four-divergence of J^μ *(changing the sign to the temporal component since we are raising the index):*

$$\begin{aligned}\partial_\mu J^\mu(\phi_1, \phi_2) = -i(&\cancel{-\partial_0\phi_1\partial_0\phi_2^*} - \phi_1\partial_0^2\phi_2^* \cancel{+\partial_0\phi_2^*\partial_0\phi_1} + \phi_2^*\partial_0^2\phi_1 \\ &\cancel{+\partial_1\phi_1\partial_1\phi_2^*} + \phi_1\partial_1^2\phi_2^* \cancel{-\partial_1\phi_2^*\partial_1\phi_1} - \phi_2^*\partial_1^2\phi_1 \\ &\cancel{+\partial_2\phi_1\partial_2\phi_2^*} + \phi_1\partial_2^2\phi_2^* \cancel{-\partial_2\phi_2^*\partial_2\phi_1} - \phi_2^*\partial_2^2\phi_1 \\ &\cancel{+\partial_3\phi_1\partial_3\phi_2^*} + \phi_1\partial_3^2\phi_2^* \cancel{-\partial_3\phi_2^*\partial_3\phi_1} - \phi_2^*\partial_3^2\phi_1),\end{aligned}$$

which turns out to be vanishing:

$$\partial_\mu J^\mu(\phi_1, \phi_2) = -i(\phi_1\Box\phi_2^* - \phi_2^*\Box\phi_1) = -i(\phi_1 m^2\phi_2^* - \phi_2^* m^2\phi_1) = 0,$$

since both ϕ_1 *and* ϕ_2^* *solve the Klein-Gordon equation* (2.6)*. This means by definition that* J^μ *is a conserved quantity. Now let us define the following hypersurfaces using spacial spherical coordinates and* $t_1 < t_2$*:*

$$\begin{aligned}\Sigma_{t_1,R} &= \{(t_1, r, \theta, \varphi)|0 \le r \le R, 0 \le \theta \le \pi, 0 \le \varphi \le 2\pi\}; \\ \Sigma_{t_2,R} &= \{(t_2, r, \theta, \varphi)|0 \le r \le R, 0 \le \theta \le \pi, 0 \le \varphi \le 2\pi\}; \\ V_R &= \{(t, R, \theta, \varphi)|t \in [t_1, t_2], 0 \le \theta \le \pi, 0 \le \varphi \le 2\pi\}; \\ \partial U_R &= \Sigma_{t_1,R} \cup \Sigma_{t_2,R} \cup V_R.\end{aligned}$$

We notice that ∂U is a closed hypersurface which is the border of a hypercylinder from t_1 to t_2. We can use Stokes' theorem which states that, being n^μ the unit vector normal to the hypersurface at each point,

$$\oint_{\partial U} J_\mu n^\mu d^3x = \int_U \partial_\mu J^\mu d^4x,$$

the right-hand side of which equals 0 because J^μ is a conserved quantity. The left-hand side can instead be written as (being $n^\mu = (1,0,0,0)$ and m^μ orthogonal to V_R)

$$\left(\int_{\Sigma_{t_1,R}} - \int_{\Sigma_{t_2,R}}\right) J_\mu n^\mu d^3x + \int_{V_R} J_\mu m^\mu d^3x = 0,$$

but in the limit $R \to \infty$ the integral over V_R vanishes because the fields ϕ_1 and ϕ_2^* vanish at spacial infinity due to the boundary condition (2.7). By the fact that $J_\mu n^\mu = J_0$ in the coordinates we have chosen, and since it is a scalar quantity, we have proven the following equality in the limit $R \to 0$:

$$\int_{\Sigma_{t_1,R}} J_0 d^3x = \int_{\Sigma_{t_2,R}} J_0 d^3x,$$

and this equivalence holds for any values of t_1 and t_2.

We can now compute the inner product of two arbitrary basic solutions (2.15) and obtain

$$\langle \phi_0 e^{ik_1^\mu x_\mu}, \phi_0 e^{ik_2^\mu x_\mu} \rangle = |\phi_0|^2(\omega_1 + \omega_2)(2\pi)^n \delta^n(\boldsymbol{k_1} - \boldsymbol{k_2}). \tag{2.18}$$

Proof. *If we choose $t = 0$ in the integral of (2.16),*

$$\begin{aligned}\langle e^{ik_1^\mu x_\mu}, e^{ik_2^\mu x_\mu} \rangle &= -i \int_{\Sigma_{t=0}} \left[e^{i\boldsymbol{k_1}\cdot\boldsymbol{x}}(i\omega_2)e^{-i\boldsymbol{k_2}\cdot\boldsymbol{x}} - e^{-i\boldsymbol{k_2}\cdot\boldsymbol{x}}(-i\omega_1)e^{i\boldsymbol{k_1}\cdot\boldsymbol{x}} \right] d^n x = \\ &= (\omega_1 + \omega_2) \int_{\Sigma_{t=0}} e^{i(\boldsymbol{k_1}-\boldsymbol{k_2})\cdot\boldsymbol{x}} d^n x,\end{aligned}$$

and the last integral is a Dirac delta multiplied by a factor of $(2\pi)^n$.

By the dispersion relation (2.9) we see that even if $\boldsymbol{k_1} = \boldsymbol{k_2} = \boldsymbol{k}$, we might have $\omega_1 = -\omega_2$, in which case the two elements of the basis are orthogonal. On the other hand if $\omega_1 = \omega_2 = \omega$ we simply get the square norm of the basic solution, which turns out to be $2\omega(2\pi)^n|\phi_0|^2$.

Now let us restrict only to the positive frequencies $\omega > 0$. Such basic solutions are parameterized by $\boldsymbol{k} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and have the following normalized form:

$$f_{\boldsymbol{k}}(x^\mu) = \frac{1}{\sqrt{2\omega(2\pi)^n}} e^{ik_\mu x^\mu}. \tag{2.19}$$

Therefore the following orthonormality relation holds (since $\omega > 0$):

$$\langle f_{\boldsymbol{k_1}}, f_{\boldsymbol{k_2}} \rangle = \delta^n(\boldsymbol{k_1} - \boldsymbol{k_2}). \tag{2.20}$$

The set of all $f_{\boldsymbol{k}}$ constitute all the possible positive frequency basic solutions, which are also called positive frequency modes. In order to recover the negative frequency ones, we would just need to switch the sign of the $i\omega t$ part from $-$ to $+$ in the exponential of (2.19) and keep $\omega > 0$. However, it is more convenient to take the complex conjugate so that we have

$$f^*_{\boldsymbol{k}}(x^\mu) = \frac{1}{\sqrt{2\omega(2\pi)^n}} e^{-ik_\mu x^\mu}. \tag{2.21}$$

In this way, both the $i\omega t$ and $i\boldsymbol{k} \cdot \boldsymbol{x}$ parts of the exponential change their sign, so that while the positive frequency mode has $(-\omega, +\boldsymbol{k})$ signs in the exponential, the negative frequency one has $(+\omega, -\boldsymbol{k})$, with the same ω and $\boldsymbol{k}$. Therefore the negative frequency equivalent of $f_{\boldsymbol{k}}$ having the same $\boldsymbol{k}$ is not $f^*_{\boldsymbol{k}}$ but rather $f^*_{-\boldsymbol{k}}$. This is an important fact to keep in mind in order to avoid confusion.

Another way to discern between positive and negative frequency modes is by taking the time derivative and checking whether it pulls out a negative imaginary factor or a positive one, respectively:

$$\begin{cases} \partial_t f_{\boldsymbol{k}} = i\omega f_{\boldsymbol{k}}, & \text{(2.22a)} \\ \partial_t f^*_{\boldsymbol{k}} = -i\omega f^*_{\boldsymbol{k}}. & \text{(2.22b)} \end{cases}$$

Since the frequency of $f^*_{\boldsymbol{k}}$ is $-\omega$, the orthonormality relation reads, in light of (2.18),

$$\langle f^*_{\boldsymbol{k_1}}, f^*_{\boldsymbol{k_2}} \rangle = -\delta^3(\boldsymbol{k_1} - \boldsymbol{k_2}). \tag{2.23}$$

The minus sign is somewhat expected because of the indefiniteness of the inner product (2.16). Lastly, since their frequencies are necessarily different, the inner product between a positive frequency mode and a negative frequency one is zero by (2.18):

$$\langle f_{\boldsymbol{k_1}}, f^*_{\boldsymbol{k_2}} \rangle = 0. \tag{2.24}$$

We have thus constructed a generalized orthonormal basis of solutions for the Klein-Gordon equation. Together, the set of all $f_{\boldsymbol{k}}$ and $f^*_{\boldsymbol{k}}$ with $\boldsymbol{k} \in \mathbb{R}^n \setminus \{0\}$ span all the linear space of solutions which is a generalized Hilbert space $\mathcal{H}$ representing the state space of the field ϕ.

One might worry about the inner product being non-positive-definite, but it turns out that most of the properties of Hilbert spaces still hold, except for considering the inner product between two negative frequency modes as redefined with a sign flip. For example, whenever we need to find a Fourier coefficient of a general solution ϕ corresponding to a negative frequency mode $f^*_{\boldsymbol{k}}$ we need to take their inner product as in (2.16) and flip the sign (see [7]):

$$\phi = \int \big(\langle \phi, f_{\boldsymbol{k}} \rangle f_{\boldsymbol{k}} - \langle \phi, f^*_{\boldsymbol{k}} \rangle f^*_{\boldsymbol{k}}\big) d^n k. \tag{2.25}$$

2.2 Quantum fields in flat spacetime

In the previous section, we have derived a generalized orthonormal set of modes for the Klein-Gordon equation. In order to quantize our real scalar field, we will impose the canonical commutation relations between the quantum operator-promoted versions of the field ϕ and its conjugate momentum π given by the derivative of the lagrangian density (2.5) with respect to $\partial_0 \phi$:

$$\pi(t, \boldsymbol{x}) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} = \partial_0 \phi. \tag{2.26}$$

Upon quantization, ϕ and π become respectively $\hat{\phi}$ and $\hat{\pi}$ which are operatorial fields that satisfy the canonical commutation relations (in natural units $\hbar = 1$)

$$[\hat{\phi}(t, \boldsymbol{x}), \hat{\phi}(t, \boldsymbol{x'})] = 0, \tag{2.27a}$$
$$[\hat{\pi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] = 0, \tag{2.27b}$$
$$[\hat{\phi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] = i\delta^n(\boldsymbol{x} - \boldsymbol{x'})\hat{1}. \tag{2.27c}$$

The Klein-Gordon equation becomes an operatorial equation:

$$\Box \hat{\phi} - m^2 \hat{\phi} = 0. \tag{2.28}$$

The general solution to this equation can be expanded as

$$\hat{\phi} = \int \left(\hat{a}_{\boldsymbol{k}} f_{\boldsymbol{k}} + \hat{a}^\dagger_{\boldsymbol{k}} f^*_{\boldsymbol{k}}\right) d^n k, \tag{2.29}$$

where $\hat{a}_{\boldsymbol{k}}$ and $\hat{a}^\dagger_{\boldsymbol{k}}$ constitute a pair of annihilation-creation operators satisfying the following commutation relations:

$$[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{k'}}] = 0, \tag{2.30a}$$
$$[\hat{a}^\dagger_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k'}}] = 0, \tag{2.30b}$$
$$[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k'}}] = \delta^n(\boldsymbol{k} - \boldsymbol{k'})\hat{1}. \tag{2.30c}$$

Proof. *By* (2.25)*, we can write a general solution* $\hat{\phi}$ *as expansion over the basic modes with operatorial coefficients* $\hat{a}_{\boldsymbol{k}}$ *and* $\hat{b}_{\boldsymbol{k}}$*:*

$$\hat{\phi} = \int \left(\hat{a}_{\boldsymbol{k}} f_{\boldsymbol{k}} + \hat{b}_{\boldsymbol{k}} f^*_{\boldsymbol{k}}\right) d^n k.$$

Since the scalar field ϕ *is real, its corresponding operator is hermitian* $\hat{\phi}^\dagger = \hat{\phi}$*, so*

$$\hat{\phi}^\dagger = \int \left(\hat{a}^\dagger_{\boldsymbol{k}} f^*_{\boldsymbol{k}} + \hat{b}^\dagger_{\boldsymbol{k}} f_{\boldsymbol{k}}\right) d^n k$$

gives $\hat{b}_{\boldsymbol{k}} = \hat{a}^\dagger_{\boldsymbol{k}}$*. Now we can express the Fourier coefficients* $\hat{a}_{\boldsymbol{k}}$ *and* $\hat{a}^\dagger_{\boldsymbol{k}}$ *using the inner product* (2.16) *as in* (2.25)*:*

$$\begin{aligned}
\hat{a}_{\boldsymbol{k}} &= \langle \hat{\phi}, f_{\boldsymbol{k}} \rangle = -i \int (\hat{\phi} \partial_0 f^*_{\boldsymbol{k}} - f^*_{\boldsymbol{k}} \hat{\pi}) d^n x, \\
\hat{a}^\dagger_{\boldsymbol{k}} &= -\langle \hat{\phi}, f^*_{\boldsymbol{k}} \rangle = i \int (\hat{\phi} \partial_0 f_{\boldsymbol{k}} - f_{\boldsymbol{k}} \hat{\pi}) d^n x,
\end{aligned}$$

where we used (2.26) *to cast* $\partial_0 \hat{\phi} = \hat{\pi}$*. Let us now compute the commutator:*

$$\begin{aligned}
&[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k'}}] = -i^2 \int d^n x \int d^n x' \big([\hat{\phi}(t, \boldsymbol{x}), \hat{\phi}(t, \boldsymbol{x'})] \partial_0 f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) \partial_0 f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) \\
&- [\hat{\phi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] \partial_0 f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) - [\hat{\pi}(t, \boldsymbol{x}), \hat{\phi}(t, \boldsymbol{x'})] f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) \partial_0 f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) \\
&+ [\hat{\pi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) \big).
\end{aligned}$$

Using the commutation relations (2.27) *we see that the first and last terms vanish, and the other two terms become* $i\delta^n(\boldsymbol{x} - \boldsymbol{x'})\hat{1}$ *with a plus sign and a minus sign respectively due to the antisymmetric property of the commutator. We thus get*

$$\begin{aligned}
[\hat{a}_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k'}}] &= i\hat{1} \int d^n x \int d^n x' \big(-\partial_0 f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) \\
&\qquad\qquad + f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) \partial_0 f_{\boldsymbol{k'}}(t, \boldsymbol{x'}) \big) \delta^n(\boldsymbol{x} - \boldsymbol{x'}) \\
&= -i\hat{1} \int d^n x \big(f_{\boldsymbol{k'}}(t, \boldsymbol{x}) \partial_0 f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) - f^*_{\boldsymbol{k}}(t, \boldsymbol{x}) \partial_0 f_{\boldsymbol{k'}}(t, \boldsymbol{x}) \big) \\
&= \langle f_{\boldsymbol{k'}}, f_{\boldsymbol{k}} \rangle \hat{1} = \delta^n(\boldsymbol{k} - \boldsymbol{k'}) \hat{1},
\end{aligned}$$

where we used (2.20). *Analogous steps using* (2.24) *lead us to*

$$[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{k'}}] = \langle f^*_{\boldsymbol{k'}}, f_{\boldsymbol{k}} \rangle = 0$$
$$[\hat{a}^\dagger_{\boldsymbol{k}}, \hat{a}^\dagger_{\boldsymbol{k'}}] = \langle f_{\boldsymbol{k'}}, f^*_{\boldsymbol{k}} \rangle = 0.$$

2.2.1 Vacuum and particles

The state labelled by $|0\rangle$ with the property of being annihilated by all annihilation operators is called the *vacuum state*:

$$\hat{a}_{\boldsymbol{k}} |0\rangle = 0 \quad \forall \boldsymbol{k} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}.$$

From basic quantum theory, we know that repeating $n_{\boldsymbol{k}}$ times the action of a creation operator corresponding to a given $\boldsymbol{k}$ on the vacuum state gives

$$\left(\hat{a}^\dagger_{\boldsymbol{k}}\right)^{n_{\boldsymbol{k}}} |0\rangle = |n_{\boldsymbol{k}}\rangle \sqrt{n_{\boldsymbol{k}}!},$$

where $|n_{\boldsymbol{k}}\rangle$ is the eigenstate of the number operator $\hat{n}_{\boldsymbol{k}} = \hat{a}^\dagger_{\boldsymbol{k}} \hat{a}_{\boldsymbol{k}}$ with eigenvalue $n_{\boldsymbol{k}} \in \mathbb{N}_0$:

$$\hat{n}_{\boldsymbol{k}} |n_{\boldsymbol{k}}\rangle = |n_{\boldsymbol{k}}\rangle \, n_{\boldsymbol{k}}.$$

We can obtain a state with different number eigenvalues for different values of $\boldsymbol{k}$. Let us illustrate this by choosing a set $\{\boldsymbol{k_1}, \boldsymbol{k_2}, \ldots, \boldsymbol{k_p}\}$ of p different values of the wave number $\boldsymbol{k}$. By applying n_i times the creation operator $\hat{a}_{\boldsymbol{k_i}}$ for $i \in \{1, \ldots, p\}$ on the vacuum state $|0\rangle$ we get

$$\left(\hat{a}^\dagger_{\boldsymbol{k_1}}\right)^{n_1} \left(\hat{a}^\dagger_{\boldsymbol{k_2}}\right)^{n_2} \cdots \left(\hat{a}^\dagger_{\boldsymbol{k_p}}\right)^{n_p} |0\rangle = |n_1, n_2, \ldots, n_p\rangle \sqrt{n_1! n_2! \cdots n_p!},$$

and the excited states behave as expected under creation and annihilation operators, as well as number operators:

$$\begin{aligned}
\hat{a}_{\boldsymbol{k_i}} |n_1, n_2, \ldots, n_i, \ldots, n_p\rangle &= |n_1, n_2, \ldots, n_i - 1, \ldots, n_p\rangle \sqrt{n_i}, \\
\hat{a}^\dagger_{\boldsymbol{k_i}} |n_1, n_2, \ldots, n_i, \ldots, n_p\rangle &= |n_1, n_2, \ldots, n_i + 1, \ldots, n_p\rangle \sqrt{n_i + 1}, \\
\hat{n}_{\boldsymbol{k_i}} |n_1, n_2, \ldots, n_i, \ldots, n_p\rangle &= |n_1, n_2, \ldots, n_i, \ldots, n_p\rangle \, n_i.
\end{aligned}$$

The set of all eigenstates of all possible number operators $\hat{n}_{\boldsymbol{k}} \; \forall \boldsymbol{k} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ constitutes a basis for the Hilbert space $\mathcal{H}$ of the states of our scalar field, which is called *Fock space*. We are now going to identify such eigenstates as states with a definite number of particles with various momenta given by the wave number $\boldsymbol{k}$. To do that, we need to obtain the hamiltonian operator and check that its eigenstates are indeed precisely the elements of the Fock basis.

The classical hamiltonian is given by performing a Legendre transformation on the lagrangian, which in the simple case we are considering just means taking kinetic plus potential energy. By recalling the three forms of energy involved in the development of our real scalar field classical theory, and by noticing that we used energy densities we want the hamiltonian to be the quantized version of

$$H = \int d^n x \left[\frac{1}{2}(\partial_0 \phi)^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2 \right], \tag{2.31}$$

which turns out to be

$$\hat{H} = \int d^n k \left[\hat{n}_{\boldsymbol{k}} + \frac{\hat{1}}{2}\delta^n(0) \right] \omega. \tag{2.32}$$

Proof. *To quantize* (2.31) *we employ the operatorial version of* ϕ*, and then we can express it in terms of the modes and creation and annihilation operators. Let us evaluate the hamiltonian term by term. The term with* ϕ^2*, by using* (2.25)*, becomes*

$$\begin{aligned}\int d^n x \frac{1}{2} m^2 \hat{\phi}^2 &= \frac{1}{2}m^2 \int d^n x \int d^n k \int d^n k' \left(\hat{a}_{\boldsymbol{k}} f_{\boldsymbol{k}} + \hat{a}^\dagger_{\boldsymbol{k}} f^*_{\boldsymbol{k}} \right)\left(\hat{a}_{\boldsymbol{k'}} f_{\boldsymbol{k'}} + \hat{a}^\dagger_{\boldsymbol{k'}} f^*_{\boldsymbol{k'}} \right) \\ &= \frac{1}{2}m^2 \int d^n x \int d^n k \int d^n k' \Big(\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k'}} f_{\boldsymbol{k}} f_{\boldsymbol{k'}} + \hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k'}} f_{\boldsymbol{k}} f^*_{\boldsymbol{k'}} + \\ &\qquad \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k'}} f^*_{\boldsymbol{k}} f_{\boldsymbol{k'}} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k'}} f^*_{\boldsymbol{k}} f^*_{\boldsymbol{k'}} \Big).\end{aligned}$$

Considering the first term and ignoring for a moment the integral over $\boldsymbol{k}$*, we can expand the expression of the modes* (2.19) *and get*

$$\begin{aligned}\int d^n x \int d^n k' \hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k'}} f_{\boldsymbol{k}} f_{\boldsymbol{k'}} &= \int d^n x \int d^n k' \hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k'}} \frac{e^{-i(\omega+\omega')t}}{2(2\pi)^n\sqrt{\omega\omega'}} e^{i(\boldsymbol{k}+\boldsymbol{k'})\cdot\boldsymbol{x}} \\ &= \int d^n k' \hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k'}} \frac{e^{-i(\omega+\omega')t}}{2\sqrt{\omega\omega'}} \delta^n(\boldsymbol{k}+\boldsymbol{k'}) = \frac{1}{2\omega}\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}} e^{-2i\omega t}.\end{aligned}$$

If we evaluate the other terms using (2.21) *also, we get that the last one has the same coefficient for the operators except for a sign flip in the exponential, while in the other two terms, the exponential becomes* 1 *because of the discordant signs. In addition, the subscripts of the operators are not discordant, so the potential energy is*

$$\int d^n x \frac{1}{2} m^2 \hat{\phi}^2 = \frac{1}{2}m^2 \int d^n k \frac{1}{2\omega} \left(\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}} e^{-2i\omega t} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k}} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}^\dagger_{-\boldsymbol{k}} e^{2i\omega t} \right).$$

The kinetic and gradient energy terms contain the derivatives of ϕ*, so, when we compute the kinetic energy term, a factor of* ω^2 *gets pulled down, while in the gradient* *energy, we get a k^2. In addition, the kinetic energy terms that contain the time exponential get a minus sign due to the double time derivation which pulls down an imaginary unit squared. We then get*

$$\int d^n x \frac{1}{2}(\partial_0\hat{\phi})^2 = \frac{1}{2}\int d^n k \frac{\omega}{2}\left(-\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}}e^{-2i\omega t} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k}} - \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}^\dagger_{-\boldsymbol{k}}e^{2i\omega t}\right),$$
$$\int d^n x \frac{1}{2}(\boldsymbol{\nabla}\hat{\phi})^2 = \frac{1}{2}\int d^n k \frac{k^2}{2\omega}\left(\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}}e^{-2i\omega t} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k}} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}^\dagger_{-\boldsymbol{k}}e^{2i\omega t}\right).$$

Putting all together and using the dispersion relation $k^2 = \omega^2 - m^2$, the potential energy simplifies with the $-m^2$ term of the gradient energy and the remaining terms with the exponentials simplify, so the result is

$$\hat{H} = \int d^n k \frac{\omega}{2}\left(\hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} + \hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k}}\right).$$

Now, since $\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{k}} = \delta^n(0)\hat{1} + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}}$ due to the creation-annihilation commutation relation and $\hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} = \hat{n}_{\boldsymbol{k}}$, we immediately obtain the expression (2.32) *of the hamiltonian.*

By looking at the expression of the hamiltonian (2.32), we notice that its eigenstates are the same as the number operators, and a state having a certain number of excitations for different values of $\boldsymbol{k}$ is interpreted as describing particles with definite momenta being present in the field in the form of plane waves (thus not localized). One might worry about the zero-point energy being a delta, but it can be shown that it is not a problem as long as we only care about the differences in the energy of different states, which is the case if we do not consider this energy acting as a gravitational source, that is an assumption we shall use henceforth. The technique of getting rid of infinities like this one is called *renormalization*, and it goes beyond the purpose of this discussion.

2.2.2 Lorentz invariance of the Fock basis

In our discussion, we always used the same coordinate system $\{t, \boldsymbol{x}\}$, but since we are dealing with relativity, we need to check how our system behaves under Lorentz transformations, in particular a Lorentz boost by velocity $\boldsymbol{v}$. The new coordinates are given by

$$t' = \gamma t - \gamma \boldsymbol{v}\cdot\boldsymbol{x}, \qquad \boldsymbol{x'} = \gamma\boldsymbol{x} - \gamma\boldsymbol{v}t,$$

and the inverse transformation is given by

$$t = \gamma t' + \gamma \boldsymbol{v}\cdot\boldsymbol{x'}, \qquad \boldsymbol{x} = \gamma\boldsymbol{x'} + \gamma\boldsymbol{v}t'.$$

To reveal the frequency of a mode in the new frame of reference, we can just take the time derivative with respect to the new frame and inspect the factor that is pulled down according to (2.22):

$$\partial_{t'} f_{\boldsymbol{k}} = \frac{\partial x^\mu}{\partial t'} \partial_\mu f_{\boldsymbol{k}} = \gamma(-i\omega) f_{\boldsymbol{k}} + \gamma \boldsymbol{v} \cdot (i\boldsymbol{k}) f_{\boldsymbol{k}}.$$

If we define $\omega' = \gamma\omega - \gamma \boldsymbol{v} \cdot \boldsymbol{k}$ we have

$$\partial_{t'} f_{\boldsymbol{k}} = -i\omega' f_{\boldsymbol{k}},$$

with $\omega > 0$ since we are considering positive-frequency modes $f_{\boldsymbol{k}}$. This means that a particle that had momentum $\boldsymbol{k}$ in the old reference frame will have a boosted energy ω' in the new one. By doing the same with the spacial derivative we can obtain the momentum in the new frame, which will be boosted to $\boldsymbol{k'}$. Therefore a Fock eigenstate $|n_1, n_2, \ldots, n_p; a\rangle$ describing n_1 particles with momentum $\boldsymbol{k_1}$, n_2 particles with momentum $\boldsymbol{k_2}$ and so on, will transform into another eigenstate $|n_1, n_2, \ldots, n_p; b\rangle$ describing n_1 particles with momentum $\boldsymbol{k'_1}$, n_2 particles with momentum $\boldsymbol{k'_2}$ and so on. This was to be expected because the four-momentum of a particle $p^\mu = (\omega, \boldsymbol{k})$ transforms in the same way as coordinates into $p^{\mu'} = (\omega', \boldsymbol{k'})$.

One key aspect is that the total number of particles, given by the eigenvalue of the total number operator

$$\hat{N} = \int \hat{n}_{\boldsymbol{k}} d^n k,$$

stays the same, since all the particles have just been boosted in energy and momentum, and no particles have been produced nor destroyed. Since the definition of the Fock space is built upon a specific coordinate system, the new reference frame will have a distinct Fock space $\mathcal{H}'$ and there will be a one-to-one correspondence between the states of $\mathcal{H}$ and $\mathcal{H}'$, which can thus be identified with one another. The only thing that changes in different inertial reference frames is the four-momentum of each particle, which will result in another Fock state with the same eigenvalue of the total number operator $\hat{N}$. We conclude that the Fock space is invariant under Lorentz transformations and, in flat spacetime, the concepts of vacuum and particles are *absolute* according to inertial observers. We are now going to discover that in curved spacetime this is no longer the case since inertial frames are no longer defined in general.

2.3 Quantum fields in curved spacetime

We will now retrace the steps of quantization in curved spacetime. First, we need the lagrangian density in curved spacetime:

$$\mathcal{L} = \sqrt{-g}\left(-\frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} m^2 \phi^2 - \frac{1}{2} \xi R \phi^2\right). \tag{2.33}$$

By comparing this expression with its flat-spacetime counterpart (2.5), we notice that there is a factor $\sqrt{-g}$, where g is the determinant of the matrix of components of the metric tensor because when we multiply $\mathcal{L}$ by $d^n x$ we need to obtain the volume element (that is invariant under changes of coordinates) so that the action keeps its scalar nature. Furthermore, we suppose that there is an additional form of elastic potential energy given by the curvature of spacetime. Since we need to obtain a scalar, the simplest quantity that can couple the field with the curvature is the Ricci curvature scalar R and the field ϕ appears at the second power because of the elasticity hypothesis. The factor ξ is an unspecified fixed constant representing the intensity of the coupling between the field and the curvature. The rest of the differences between flat and curved spacetime lagrangian density are due to the substitutions $\partial_\mu \leftrightarrow \nabla_\mu$ and $\eta^{\mu\nu} \leftrightarrow g^{\mu\nu}$, which are simply the application of the general covariance principle.

By the principle of least action, we have the Euler-Lagrange equation in curved spacetime:

$$\nabla_\mu \left(\frac{\partial \mathcal{L}}{\partial(\nabla_0 \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \tag{2.34}$$

which is the covariant version of (2.4). The conjugate momentum is

$$\pi = \frac{\partial \mathcal{L}}{\partial(\nabla_0 \phi)} = \sqrt{-g} \nabla_0 \phi, \tag{2.35}$$

where we used locally inertial coordinates ($g^{\mu\nu} = \eta^{\mu\nu}$ and $\Gamma^\sigma_{\mu\nu} = 0$) for the computation, and since the result is expressed in covariant form it holds for all coordinate systems. By a computation analogous to the flat-spacetime case, we see that the equation of motion becomes

$$\Box \phi - m^2 \phi - \xi R \phi = 0, \tag{2.36}$$

where the d'Alembertian is defined as

$$\Box = g^{\mu\nu} \nabla_\mu \nabla_\nu . \tag{2.37}$$

The curved-spacetime version of the Klein-Gordon equation (2.5) is indeed linear (as its flat-spacetime analog (2.6)), so we can introduce the generalized version of the inner product (2.16) in the space of solutions of this equation too. In fact, given a spacelike hypersurface Σ with induced metric $\gamma_{\mu\nu}$ and two solutions ϕ_1, ϕ_2 we define

$$\langle \phi_1, \phi_2 \rangle = -i \int_\Sigma \left(\phi_1 \nabla_\mu \phi_2^* - \phi_2^* \nabla_\mu \phi_1 \right) n^\mu \sqrt{\gamma} d^n x, \tag{2.38}$$

where n^μ is the normal vector to the hypersurface Σ at each point and the result is independent from the choice of the spacelike hypersurface Σ on which the integral is performed, as one can check just by retracing the same steps of the proof for the Minkowski case (2.16) using covariant derivatives, the generalized Stokes' theorem and two arbitrary hypersurfaces of a spacelike foliation of the spacetime manifold.

We can now proceed to quantize the fields ϕ and π, which become $\hat{\phi}$ and $\hat{\pi}$ respectively and satisfy the following commutation relations (in an arbitrary reference frame $\{t, \boldsymbol{x}\}$):

$$[\hat{\phi}(t, \boldsymbol{x}), \hat{\phi}(t, \boldsymbol{x'})] = 0, \tag{2.39a}$$

$$[\hat{\pi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] = 0, \tag{2.39b}$$

$$[\hat{\phi}(t, \boldsymbol{x}), \hat{\pi}(t, \boldsymbol{x'})] = \frac{i}{\sqrt{-g}} \delta^n(\boldsymbol{x} - \boldsymbol{x'})\hat{1}, \tag{2.39c}$$

where the factor $\sqrt{-g}$ appears because, in the integral defining the Dirac delta, it cancels with the measure, which is $\sqrt{-g}$ as well. Equation (2.36) then becomes an operatorial equation:

$$\Box\hat{\phi} - m^2\hat{\phi} - \xi R\hat{\phi} = 0. \tag{2.40}$$

2.3.1 Bogoljubov transformations

The solution of (2.36) cannot be expressed in terms of plane waves in general. However we can pick an orthonormal basis in the Hilbert space of solutions, say $\{f_\lambda, f^*_\lambda\}$, where λ is an arbitrary multi-index:

$$\langle f_\lambda, f_{\lambda'} \rangle = \delta(\lambda - \lambda'), \tag{2.41a}$$

$$\langle f^*_\lambda, f^*_{\lambda'} \rangle = -\delta(\lambda - \lambda'), \tag{2.41b}$$

$$\langle f_\lambda, f^*_{\lambda'} \rangle = 0. \tag{2.41c}$$

We can then expand an arbitrary solution $\hat{\phi}$ as

$$\hat{\phi} = \int d\mu(\lambda) \left(\hat{a}_\lambda f_\lambda + \hat{a}^\dagger_\lambda f^*_\lambda \right), \tag{2.42}$$

where $\int d\mu(\lambda)$ is a general expression indicating the linear combination for all possible values of λ, which reduces to a summation $\sum_\lambda$ in the discrete case. The reason why the coefficients are creation-annihilation operator pairs is the same as in the flat-spacetime case, where we used the quantization prescription and the inner product defined in the space of solutions, without ever expliciting the expression of the basis elements. Since the choice of the basis is arbitrary, we can choose another one $\{g_\lambda, g^*_\lambda\}$ such that

$$\hat{\phi} = \int d\mu(\lambda) \left(\hat{b}_\lambda g_\lambda + \hat{b}^\dagger_\lambda g^*_\lambda \right), \tag{2.43}$$

and the following orthonormality relations hold:

$$\langle g_\lambda, g_{\lambda'} \rangle = \delta(\lambda - \lambda'), \tag{2.44a}$$
$$\langle g^*_\lambda, g^*_{\lambda'} \rangle = -\delta(\lambda - \lambda'), \tag{2.44b}$$
$$\langle g_\lambda, g^*_{\lambda'} \rangle = 0. \tag{2.44c}$$

The two bases are related by a linear transformation, which is called *Bogoljubov transformation*:

$$g_\lambda = \int d\mu(\sigma) \left(\alpha_{\lambda\sigma} f_\sigma + \beta_{\lambda\sigma} f^*_\sigma \right), \tag{2.45}$$

whose inverse results to be

$$f_\lambda = \int d\mu(\sigma) \left(\alpha^*_{\sigma\lambda} g_\sigma - \beta_{\sigma\lambda} g^*_\sigma \right). \tag{2.46}$$

Proof. *As we said for* (2.25)*, we can view the coefficients of the expansion* (2.45) *as:*

$$\alpha_{\lambda\sigma} = \langle g_\lambda, f_\sigma \rangle,$$
$$\beta_{\lambda\sigma} = -\langle g_\lambda, f^*_\sigma \rangle.$$

By expliciting the coefficients in (2.46) *we get*

$$f_\lambda = \int d\mu(\sigma) \left(\langle f_\lambda, g_\sigma \rangle g_\sigma - \langle f_\lambda, g^*_\sigma \rangle g^*_\sigma \right).$$

Now we notice that

$$\langle f_\lambda, g_\sigma \rangle = \langle g_\sigma, f_\lambda \rangle^* = \alpha^*_{\sigma\lambda},$$

which is the first coefficient of the inverse transformation (2.46)*. Then we can explicit the second coefficient using* (2.38)*:*

$$\begin{aligned} -\langle f_\lambda, g^*_\sigma \rangle &= i \int_\Sigma \left(f_\lambda \nabla_\mu g_\sigma - g_\sigma \nabla_\mu f_\lambda \right) n^\mu \sqrt{\gamma} d^n x \\ &= -i \int_\Sigma \left(g_\sigma \nabla_\mu f_\lambda - f_\lambda \nabla_\mu g_\sigma \right) n^\mu \sqrt{\gamma} d^n x \\ &= \langle g_\sigma, f^*_\lambda \rangle = -\beta_{\sigma\lambda}. \end{aligned}$$

The Bogoljubov coefficients can be used to transform between creation-annihilation operators of the two different expansions (2.42) and (2.43):

$$\hat{a}_\lambda = \int d\mu(\sigma) \left(\alpha_{\sigma\lambda} \hat{b}_\sigma + \beta^*_{\sigma\lambda} \hat{b}^\dagger_\sigma \right), \tag{2.47a}$$
$$\hat{b}_\lambda = \int d\mu(\sigma) \left(\alpha^*_{\lambda\sigma} \hat{a}_\sigma - \beta^*_{\lambda\sigma} \hat{a}^\dagger_\sigma \right). \tag{2.47b}$$

Proof. *Since we are dealing with coefficients of the expansions* (2.42) *and* (2.43) *respectively, we can express* $\hat{a}_\lambda$ *as*

$$\begin{aligned}\hat{a}_\lambda &= \langle \hat{\phi}, f_\lambda \rangle \\ &= \int d\mu(\sigma) \left(\alpha_{\sigma\lambda} \langle \hat{\phi}, g_\sigma \rangle - \beta^*_{\sigma\lambda} \langle \hat{\phi}, g^*_\sigma \rangle \right) \\ &= \int d\mu(\sigma) \left(\alpha_{\sigma\lambda} \hat{b}_\sigma + \beta^*_{\sigma\lambda} \hat{b}^\dagger_\sigma \right) ,\end{aligned}$$

where we used (2.46), (2.43) *and conjugate-linearity in the second argument of* $\langle \cdot, \cdot \rangle$. *On the other hand,* $\hat{b}_\lambda$ *can be written as*

$$\begin{aligned}\hat{b}_\lambda &= \langle \hat{\phi}, g_\lambda \rangle \\ &= \int d\mu(\sigma) \left(\alpha^*_{\lambda\sigma} \langle \hat{\phi}, f_\sigma \rangle + \beta^*_{\lambda\sigma} \langle \hat{\phi}, f^*_\sigma \rangle \right) \\ &= \int d\mu(\sigma) \left(\alpha^*_{\lambda\sigma} \hat{a}_\sigma - \beta^*_{\lambda\sigma} \hat{a}^\dagger_\sigma \right) ,\end{aligned}$$

where we used (2.45), (2.42), *and again conjugate-linearity in the second argument of the inner product.*

The Bogoljubov transformation works also for transforming between different bases of modes in the flat spacetime case, and we will use this fact in the next chapter, where we will use two different sets of modes in Minkowski spacetime.

2.3.2 Relativity of vacuum and particles

Although we can always find a basis in the generalized Hilbert space the field states, there is no plane-wave basis, because the metric is in general not in diagonal form globally, so the d'Alembertian (2.37) contains mixed second-derivative terms and the factorization $\phi(t, \boldsymbol{x}) = \alpha(t)\beta(\boldsymbol{x})$ that we used to find the plane wave solutions in the flat-spacetime case cannot be performed. In flat spacetime, we expanded the field $\hat{\phi}$ as in (2.29) using the basis composed of positive (2.19) and negative (2.21) frequency modes. Then we showed that the quantization prescriptions implied that the coefficients of the expansion were creation-annihilation operator pairs, whose associated number operators appeared linearly combined in the hamiltonian (2.32). We then interpreted the eigenstates of the hamiltonian (shared with the number operators) as quantum states with a definite whole number of energy quanta ω, which we used as our notion of "particle". In curved spacetime, however, since in general it is not possible to factorize the field in time-dependent and space-dependent factors, there does not exist a global definition of "frequency" that appears to be the same in every point in spacetime, so there not exists a global notion of "particle" either. This is because we cannot reproduce the steps that we took to derive the hamiltonian (2.32) where we used the explicit form of the modes $f_{\boldsymbol{k}}$ and $f^*_{\boldsymbol{k}}$, which happened to be plane waves, that we do not have now. Although we could have chosen a different basis in the flat-spacetime case, the one consisting of plane waves is special in some sense, since it is the only one which leads to the hamiltonian being expressed in terms of number operators. This reflects the fact that in special relativity there exists a family of privileged reference frames, that are the inertial ones, while in general relativity this is no longer the case. In curved spacetime, the particular choice of basis that we make in the linear space of solutions of (2.36) does not matter and therefore the notions of "vacuum" and "particles" are not absolute. However, if a global hypersurface-orthogonal timelike Killing vector field exists, we will see that there is a way to define such concepts globally, exploiting the temporal symmetry and restoring the definiteness of frequencies in the modes we expand the field on.

We can now explore in more detail what we just said by using the two bases f_λ and g_λ which appeared in (2.42) and (2.43) respectively. Let us denote the vacuum states as $|0_f\rangle$ and $|0_g\rangle$ corresponding to the two different bases so that each vacuum state has eigenvalue 0 with respect to their corresponding number operators $\hat{n}_{f\lambda}$ and $\hat{n}_{g\lambda}$:

$$\begin{aligned} \hat{n}_{(f)\lambda} |0_f\rangle &= 0 \quad \forall\lambda, \\ \hat{n}_{(g)\lambda} |0_g\rangle &= 0 \quad \forall\lambda. \end{aligned}$$

If we calculate the expectation value of the number operator $\hat{n}_{g\lambda}$ for a fixed value of λ in the f-vacuum state $|0_f\rangle$, we get

$$\langle 0_f| \hat{n}_{(g)\lambda} |0_f\rangle = \int d\mu(\sigma) |\beta_{\lambda\sigma}|^2, \tag{2.48}$$

where $\beta_{\lambda\sigma}$ is the Bogoljubov coefficient.

Proof. *Since $\hat{n}_{(g)\lambda} = \hat{b}^\dagger_\lambda \hat{b}_\lambda$, by* (2.47b) *we have*

$$\begin{aligned} &\langle 0_f| \hat{n}_{(g)\lambda} |0_f\rangle = \langle 0_f| \hat{b}^\dagger_\lambda \hat{b}_\lambda |0_f\rangle \\ &= \langle 0_f| \left[\int d\mu(\sigma) \int d\mu(\rho) \left(\alpha_{\lambda\sigma} \hat{a}^\dagger_\sigma - \beta_{\lambda\sigma} \hat{a}_\sigma\right) \left(\alpha^*_{\lambda\rho} \hat{a}_\rho - \beta^*_{\lambda\rho} \hat{a}^\dagger_\rho\right) \right] |0_f\rangle \,. \end{aligned}$$

If we use the fact that the bra $\langle 0_f|$ gets annihilated by creation operators and the ket $|0_f\rangle$ gets annihilated by annihilation operators, and employ the commutation relation

(2.30c) *(where we substitute $\boldsymbol{k} \leftrightarrow \sigma$ and $\boldsymbol{k'} \leftrightarrow \rho$), we get*

$$\begin{aligned}
&\langle 0_f | \, \hat{n}_{g\lambda} \, |0_f\rangle = \\
&= \int d\mu(\sigma) \int d\mu(\rho) \Big(\cancel{\alpha_{\lambda\sigma}\alpha^*_{\lambda\rho} \langle 0_f | \, \hat{a}^\dagger_\sigma \hat{a}_\rho \, |0_f\rangle} - \cancel{\alpha_{\lambda\sigma}\beta^*_{\lambda\rho} \langle 0_f | \, \hat{a}^\dagger_\sigma \hat{a}^\dagger_\rho \, |0_f\rangle} \\
&\qquad\qquad - \cancel{\beta_{\lambda\sigma}\alpha^*_{\lambda\rho} \langle 0_f | \, \hat{a}_\sigma \hat{a}_\rho \, |0_f\rangle} + \beta_{\lambda\sigma}\beta^*_{\lambda\rho} \langle 0_f | \, \hat{a}_\sigma \hat{a}^\dagger_\rho \, |0_f\rangle \Big) \\
&= \int d\mu(\sigma) \int d\mu(\rho) \beta_{\lambda\sigma}\beta^*_{\lambda\rho} \langle 0_f | \left(\delta(\sigma,\rho)\hat{1} + \hat{a}^\dagger_\rho \hat{a}_\sigma \right) |0_f\rangle \\
&= \int d\mu(\sigma) \int d\mu(\rho) \beta_{\lambda\sigma}\beta^*_{\lambda\rho} \left(\delta(\sigma,\rho) \langle 0_f | 0_f\rangle + \cancel{\langle 0_f | \, \hat{a}^\dagger_\rho \hat{a}_\sigma \, |0_f\rangle} \right) \\
&= \int d\mu(\sigma) \beta_{\lambda\sigma}\beta^*_{\lambda\sigma} = \int d\mu(\sigma) |\beta_{\lambda\sigma}|^2,
\end{aligned}$$

where we also used the fact that the ket $|0_f\rangle$ is normalized.

In general, there is no guarantee that (2.48) vanishes, even though one expects such behaviour from the vacuum state. This shows that there is no unique way to identify "vacuum" and "particles" because they depend on which particular basis we choose for the expansion of the field.

2.3.3 Particle detection

At this point, one may ask how would a particle detector behave in general, since the presence or absence of particles seems to depend on the basis we choose for the expansion of the field. A particle detector, for our purposes, is just a localized apparatus that can detect the presence of a plane wave in the field and determine its frequency ω. Such an instrument, however, does not care about the modes we choose, because they are just a mathematical way to talk about the solutions of the equation of motion of the field, and do not have a physical significance in general.

Let us consider the flat spacetime case first, and suppose we have two particle detectors D and D', the former fixed at the origin of a reference frame $\{t, \boldsymbol{x}\}$ and the latter fixed at the origin of another reference frame $\{t', \boldsymbol{x'}\}$ moving with velocity $\boldsymbol{v}$ with respect to the former. As we discussed in section 2.2.2, changing inertial reference frame results in detecting a different frequency for the same particle. When we talk about frequency, we simply mean the ω factor that is pulled down by the corresponding definite-frequency mode when we take the time derivative with respect to the proper time of the detector, as in (2.22):

$$\partial_t f_{\boldsymbol{k},\omega}(t, \boldsymbol{x}) = -i\omega f_{\boldsymbol{k},\omega}(t, \boldsymbol{x}),$$

so in the reference frame of D', we have

$$\partial_{t'} f_{\boldsymbol{k'},\omega'}(t', \boldsymbol{x'}) = -i\omega' f_{\boldsymbol{k'},\omega'}(t', \boldsymbol{x'})$$

as we expected.

The behaviour of particle detectors in curved spacetime is non-trivial because one would need to express the set of definite-frequency modes corresponding to the local inertial reference frame of the detector (which are the Minkowski ones due to local flatness) in terms of the coordinate system used as the domain of the field, then expand the field using those modes, take the creation and annihilation operators that appear in the expansion, construct the respective number operators and apply them to the state of the field to obtain number eigenvalues representing the particles. However, if a timelike Killing vector field $\vec{K}$ exists, there is an easy way to do this as long as the four-velocity $\vec{U}$ of the detector is proportional to $\vec{K}$ (it may not be following a geodesic path).

In this case, by choosing the local inertial reference frame where the detector is at the origin, we can perform the factorization

$$f_\omega(t, \boldsymbol{x}) = e^{-i\omega t} \bar{f}_\omega(\boldsymbol{x}), \tag{2.49}$$

and the frequency is defined on the entire domain of the Killing vector field $\vec{K}$.

Proof. *By theorem 1.1 we know that if we choose the temporal coordinate along the timelike Killing vector field $\vec{K} = \partial_t$, the metric will be globally in block-diagonal form with no mixed time and space terms. Therefore, we can express the d'Alembertian* (2.37) *globally as*

$$\Box = g^{00}\nabla_0\nabla_0 + g^{ij}\nabla_i\nabla_j,$$

where we notice that the temporal and spatial derivatives do not mix, so we are now able to search for a basis of the solution space of (2.36) *in a globally factorized form. In fact, if we rewrite equation* (2.36)*, we get (knowing that, for a scalar, $\nabla_\mu\phi = \partial_\mu\phi$)*

$$g^{00}\partial_0^2\phi + (g^{ij}\partial_i\partial_j - m^2 - \xi R)\phi = 0.$$

If we factorize $\phi(t, \boldsymbol{x}) = \alpha(t)\beta(\boldsymbol{x})$ we obtain

$$\begin{gathered} g^{00}\alpha''\beta + \alpha(g^{ij}\partial_i\partial_j - m^2 - \xi R)\beta = 0 \\ \alpha'' + \underbrace{\frac{1}{g^{00}}\left(\frac{1}{\beta}g^{ij}\partial_i\partial_j\beta - m^2 - \xi R\right)}_{\omega^2}\alpha = 0, \end{gathered}$$

where the factor that multiplies α is constant in time for a fixed point in space and the double-primed superscript denotes the second total derivative. The basic solutions *are*

$$\alpha(t) = \alpha_0 e^{\pm i\omega t},$$

and therefore we can write $\phi(t, \boldsymbol{x})$ in factorized form:

$$\phi(t, \boldsymbol{x}) = \alpha_0 e^{\pm i\omega t} \beta_{\pm\omega}(\boldsymbol{x}).$$

As we did in flat spacetime, we can require ω to be real and turn it into a parameter for the basic solutions, so we write them in the general form

$$f_\omega(t, \boldsymbol{x}) = e^{-i\omega t} \bar{f}_\omega(\boldsymbol{x}),$$

with $\omega \in \mathbb{R} \setminus \{0\}$. The considerations about the case when $\omega = 0$ are the same as those we made when we found the solutions (2.8) *in the flat spacetime case.*

In this scenario, we can always find a basis with definite frequencies $\{f_\omega, f_\omega^*\}$. If we apply to them the covariant (directional) derivative with respect to $\vec{U}$, by doing calculations in the reference frame $\{t, \boldsymbol{x}\}$ that we have introduced, where the detector is locally inertial with four-velocity $\vec{U} = \partial_t$, we get

$$\nabla_{\vec{U}} f_\omega = U^\mu \nabla_\mu f_\omega = \nabla_t f_\omega = \partial_t f_\omega = -i\omega f_\omega,$$

where we used (2.49). We have thus defined the frequency for each mode in a coordinate-independent way since the left-hand side is a covariant derivative with respect to a vector of the spacetime manifold and the right-hand side is a scalar. This means that a detector whose four-velocity is proportional to the Killing vector field will measure a particle having energy ω if the scalar field has an excitation in the mode f_ω.

We have shown that, if the spacetime is static, there is a natural way to define "vacuum" and "particles" because there exists a natural operational procedure that tells an observer how to detect particles at each point in spacetime, that is preparing a detector in a way that it has four-velocity proportional to the Killing vector field (e.g. in Schwarzschild spacetime it means remaining at constant spacial coordinates, even if it means to overcome gravity). We are now ready to apply the notions we have learned so far to analyze the Unruh effect, in the next chapter.

Chapter 3

Unruh effect

To get rid of unnecessary complications, we are going to consider a massless real scalar quantum field in two-dimensional flat spacetime, so the theory of the previous chapter reduces to the $n = 1$ case with $\{t, x\}$ coordinates, and the Klein-Gordon equation becomes

$$\Box\phi = 0. \tag{3.1}$$

We are now going to study accelerated observers in Minkowski spacetime and see that they expand the field ϕ on a different basis of modes than inertial observers, and using Bogoljubov transformations (2.47) we will realize that the number of particles detected by the two observers is not the same.

3.1 Rindler observers

A Rindler observer is one whose worldline is described by

$$\begin{cases} t(\tau) = \dfrac{1}{\alpha}\sinh(\alpha\tau), & \text{(3.2a)} \\ x(\tau) = \dfrac{1}{\alpha}\cosh(\alpha\tau), & \text{(3.2b)} \end{cases}$$

where τ is the proper time and $\alpha \neq 0$. We can show that such an observer is accelerated, with $|\alpha|$ being the modulus of proper acceleration.

Proof. *By definition of four-acceleration, we need to compute the second total deriva-*

tives of (3.2)*:*

$$\frac{d^2t}{d\tau^2} = \alpha \sinh(\alpha\tau),$$
$$\frac{d^2x}{d\tau^2} = \alpha \cosh(\alpha\tau).$$

Therefore the modulus is

$$a = \sqrt{a^\mu a_\mu} = \sqrt{-\alpha^2 \sinh^2(\alpha\tau) + \alpha^2 \cosh^2(\alpha\tau)} = \sqrt{\alpha^2} = |\alpha|,$$

which is Lorentz-invariant, hence it corresponds to the modulus of proper acceleration.

The trajectory describes a hyperbola on the t-x Minkowski diagram as can be viewed in fig. 3.1, with equation

$$x^2 - t^2 = \frac{1}{\alpha^2}. \tag{3.3}$$

Proof. *Using* (3.2)*, it is a straightforward computation:*

$$x^2 - t^2 = \frac{1}{\alpha^2} \cosh^2(\alpha\tau) - \frac{1}{\alpha^2} \sinh^2(\alpha\tau) = \frac{1}{\alpha^2}.$$

We now introduce new coordinates $\{\eta, \xi\}$ with range $-\infty < \eta, \xi < +\infty$ (represented in fig 3.1) and such that they satisfy the following transformation relations (with $a > 0$):

$$\begin{cases} t = \dfrac{1}{a} e^{a\xi} \sinh(a\eta), & \text{(3.4a)} \\ x = \dfrac{1}{a} e^{a\xi} \cosh(a\eta). & \text{(3.4b)} \end{cases}$$

We shall notice that since we require $a > 0$, these new coordinates cover only the region $x > |t|$ where $\alpha > 0$, and we need to overload the definition of $\{\eta, \xi\}$ in order to make them work also in the region $x < -|t|$:

$$\begin{cases} t = -\dfrac{1}{a} e^{a\xi} \sinh(a\eta), & \text{(3.5a)} \\ x = -\dfrac{1}{a} e^{a\xi} \cosh(a\eta). & \text{(3.5b)} \end{cases}$$

In these coordinates, the trajectory of the Rindler observer becomes

$$\begin{cases} \eta(\tau) = \dfrac{\alpha}{a} \tau, & \text{(3.6a)} \\ \xi(\tau) = \dfrac{1}{a} \ln\left(\dfrac{a}{\alpha}\right). & \text{(3.6b)} \end{cases}$$

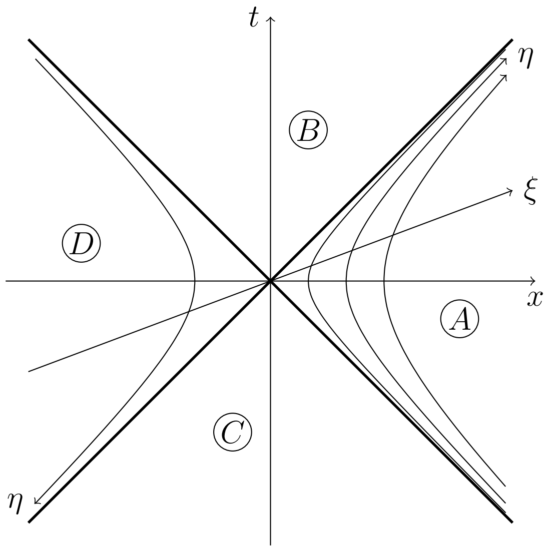


Fig. 3.1: Accelerated trajectories in Minkowski spacetime corresponding to fixed ξ and following η curves. The spacetime is divided by the light cone at the origin into four regions labelled A, B, C, D.

Proof. *Using* (3.3)*,*

$$x^2 - t^2 = \frac{1}{a^2} e^{2a\xi} \left(\cosh^2(a\eta) - \sinh^2(a\eta) \right) = \frac{1}{a^2} e^{2a\xi} = \frac{1}{\alpha^2},$$

which gives

$$e^{a\xi} = \frac{a}{\alpha} \qquad \Rightarrow \qquad \xi = \frac{1}{a} \ln \left(\frac{a}{\alpha} \right), \tag{3.7}$$

that is (3.6b)*. If we notice that* $\frac{1}{a} e^{a\xi} = \frac{1}{\alpha}$*, which immediately follows from the above* (3.7)*, and substitute in* (3.4)*, we get*

$$\begin{cases} t &= \frac{1}{\alpha} \sinh(a\eta) = \frac{1}{\alpha} \sinh(\alpha\tau), \\ x &= \frac{1}{\alpha} \cosh(a\eta) = \frac{1}{\alpha} \cosh(\alpha\tau). \end{cases}$$

where the second equalities come from (3.2) *since we are interested in the Rindler observer's trajectory. These give* (3.6a) *by comparison of the arguments of hyperbolic trigonometric functions,*

$$a\eta = \alpha\tau \qquad \Rightarrow \qquad \eta = \frac{\alpha}{a}\tau.$$

We notice that these coordinates simplify the description of the accelerated path, since η is proportional to the proper time and ξ is just constant. In particular, if $a = \alpha$

we immediately obtain

$$\begin{cases} \eta = \tau, & (3.8a) \\ \xi = 0. & (3.8b) \end{cases}$$

This also tells us that η is a timelike coordinate, while ξ is spacelike.

The metric in Rindler coordinates gives the following line element:

$$ds^2 = e^{2a\xi}(-d\eta^2 + d\xi^2). \tag{3.9}$$

Proof. *We can compute dt and dx using the chain rule with* (3.4) *and then substitute in $ds^2 = -dt^2 + dx^2$, which is the well-known line element for Minkowski metric in $\{t, x\}$ coordinates. We have*

$$\begin{aligned} dt &= \frac{\partial t}{\partial \eta} d\eta + \frac{\partial t}{\partial \xi} d\xi = e^{a\xi}\cosh(a\eta)d\eta + e^{a\xi}\sinh(a\eta)d\xi \\ dx &= \frac{\partial x}{\partial \eta} d\eta + \frac{\partial x}{\partial \xi} d\xi = e^{a\xi}\sinh(a\eta)d\eta + e^{a\xi}\cosh(a\eta)d\xi, \end{aligned}$$

and therefore

$$\begin{aligned} ds^2 = -dt^2 + dx^2 = e^{2a\xi}\big[&-\cosh^2(a\eta)d\eta^2 - \sinh^2(a\eta)d\xi^2 - \cancel{2\cosh(a\eta)\sinh(a\eta)d\eta d\xi} \\ &+ \sinh^2(a\eta)d\eta^2 + \cosh^2(a\eta)d\xi^2 + \cancel{2\sinh(a\eta)\cosh(a\eta)d\eta d\xi}\big], \end{aligned}$$

which gives (3.9) *by using the identity* $\cosh^2(a\eta) - \sinh^2(a\eta) = 1$ *twice.*

Since the metric does not depend on η and it is in diagonal form everywhere, as we can realize by inspecting (3.9) and invoking theorem 1.1, we can say that $\vec{K} = \partial_\eta$ is a hypersurface-orthogonal timelike Killing vector field, and we see that it relates to the coordinates $\{t, x\}$ as follows:

$$\vec{K} = \partial_\eta = a(x\partial_t + t\partial_x). \tag{3.10}$$

Proof. *Using the chain rule:*

$$\frac{\partial}{\partial \eta} = \frac{\partial x}{\partial \eta}\frac{\partial}{\partial x} + \frac{\partial t}{\partial \eta}\frac{\partial}{\partial t},$$

which leads to, using (3.4),

$$\partial_\eta = e^{a\xi}\sinh(a\eta)\partial_x + e^{a\xi}\cosh(a\eta)\partial_t,$$

and by employing again (3.4) *we can cast* $e^{a\xi}\sinh(a\eta) = at$ *and* $e^{a\xi}\cosh(a\eta) = ax$*, which gives* (3.10)*.*

The (hyper-) surfaces defined by $t = x$ and $t = -x$ (which are the asymptotes of Rindler trajectories, i.e. the light cone of the origin of the reference frame, as depicted in fig. 3.1) are Killing horizons for $\vec{K}$.

Proof. *By the definition 1.4 of Killing horizons, it is sufficient to show that* $\vec{K}$ *has zero norm on such surfaces. Let us use* $\{t, x\}$ *coordinates with the Minkowski metric* $diag(-1, 1)$ *to compute the square norm, and check that it is indeed zero:*

$$K^\mu K_\mu = -(K^0)^2 + (K^1)^2 = -a^2x^2 + a^2t^2 = 0$$

where we used (3.10) *to explicit the components of* $\vec{K}$ *and* $t = \pm x$ *in the last step to compute the result on the horizon.*

Since every Killing horizon has a surface gravity by definition, we can go on and compute it, obtaining

$$\kappa = a. \tag{3.11}$$

Proof. *We can use* (1.22) *evaluated at* $t = \pm x$*. Let us do the calculation explicitly, starting from* (3.10) *and lowering the index:*

$$K^\mu = (ax, at) \qquad\qquad K_\mu = (-ax, at).$$

Then we compute the derivatives:

$$\begin{array}{llll} \nabla_0 K_0 = 0 & \nabla_0 K_1 = a & \nabla_1 K_0 = -a & \nabla_1 K_1 = 0, \\ \nabla^0 K^0 = 0 & \nabla^0 K^1 = -a & \nabla^1 K^0 = a & \nabla^1 K^1 = 0, \end{array}$$

where we used the fact that $\nabla^0 = -\nabla_0$ *and* $\nabla^1 = \nabla_1$*. If we multiply same-index terms and sum we get* $-2a^2$ *which, multiplied by* $-1/2$ *and square rooted, gives exactly* a*.*

Despite the name, there is no actual gravity in this case, since the spacetime is flat. However, an accelerated observer feels like being in a gravitational field, as the equivalence principle states, and this is the reason why we will be able to apply the results obtained in this chapter to the case where an actual black hole with a physical Killing event horizon is present. Recall that the redshift factor is the modulus of the Killing vector field as in (1.12), which can easily be computed using $\vec{K} = \partial_\eta = 1\partial_\eta + 0\partial_\xi$ and the metric (3.9):

$$V = e^{a\xi}. \tag{3.12}$$

We can now proceed to analyze the solutions of equation (3.1).

3.2 Modes and frequencies for Rindler observers

Equation (3.1) in Rindler coordinates becomes

$$e^{-2a\xi}(-\partial_\eta^2 + \partial_\xi^2)\phi = 0. \tag{3.13}$$

Proof. *We already know* ∂_η *in terms of* ∂_t *and* ∂_x *through* (3.10)*. We need a similar relation for* ∂_ξ*. By using* (3.10) *and* (3.4) *we get*

$$\partial_\xi = \frac{\partial}{\partial\xi} = \frac{\partial t}{\partial\xi}\partial_t + \frac{\partial x}{\partial\xi}\partial_x = e^{a\xi}\sinh(a\eta)\partial_t + e^{a\xi}\cosh(a\eta)\partial_x = a(t\partial_t + x\partial_x).$$

Now we can substitute, and by employing (3.10) *we obtain*

$$\begin{aligned}
-\partial_\eta^2 + \partial_\xi^2 &= a^2\left[-(x\partial_t + t\partial_x)(x\partial_t + t\partial_x) + (t\partial_t + x\partial_x)(t\partial_t + x\partial_x)\right] \\
&= a^2\left[-x^2\partial_t^2 - \cancel{x\partial_x} - \cancel{xt\partial_t\partial_x} - \cancel{t\partial_t} - \cancel{tx\partial_x\partial_t} - t^2\partial_x^2\right. \\
&\qquad \left. + \cancel{t\partial_t} + t^2\partial_t^2 + \cancel{tx\partial_t\partial_x} + \cancel{xt\partial_x\partial_t} + \cancel{x\partial_x} + x^2\partial_x^2\right] \\
&= a^2\left[(t^2 - x^2)\partial_t^2 - (t^2 - x^2)\partial_x^2\right] \\
&= a^2(x^2 - t^2)(-\partial_t^2 + \partial_x^2) \\
&= e^{2a\xi}(-\partial_t^2 + \partial_x^2) \\
&= e^{2a\xi}\Box,
\end{aligned}$$

where we also used (3.4)*. By the chain of equalities,*

$$\Box = e^{-2a\xi}(-\partial_\eta^2 + \partial_\xi^2),$$

which proves (3.13) *from* (3.1)*.*

Since the exponential in (3.13) is always positive because $a > 0$, the equation is formally identical to $(-\partial_t^2 + \partial_x^2)\phi = 0$, which we solved in section 2.1.2, with $t \leftrightarrow \eta$, $x \leftrightarrow \xi$, $m = 0$ and $n = 1$, and therefore the basic positive frequency modes are given by (2.19):

$$g_k = \frac{1}{\sqrt{4\pi\omega}}e^{-i\omega\eta + ik\xi}, \tag{3.14}$$

with the trivial dispersion relation $\omega = |k|$. Of course, to get the negative frequency modes we just take the complex conjugate of positive frequency ones. We shall notice, however, that these modes are positive frequency only in the region $x > |t|$ (labelled by A in fig. 3.1). In fact, in this region, $\alpha > 0$, and by (3.6a), η has the same direction of the proper time τ. In the region defined by $x < -|t|$ (labeled D in fig. 3.1), we have $\alpha < 0$ and therefore η has a direction opposite to the proper time τ, so the modes (3.14) are negative-frequency. The right way to think about this is by invoking the definition (2.22) of positive and negative frequency modes, where we have to use the proper time τ of the Rindler observer as the quantity with respect to which partial derivatives are computed. By the chain rule and (3.6a), we have

$$\partial_\tau = \frac{\alpha}{a}\partial_\eta, \tag{3.15}$$

and applying it to a mode (3.14),

$$\partial_\tau g_k = -i\omega\frac{\alpha}{a}g_k.$$

We notice that other than ω, an additional factor of α/a appears, and its sign is determined by α because $a > 0$ by definition, so if $\alpha > 0$ the factor is negative and the mode is positive-frequency, while if $\alpha < 0$ the factor is positive and the mode is negative frequency. In order to organize our modes based on the positiveness or negativeness of frequency, we need to redefine the modes g_k in different ways in the two regions where α is positive or negative respectively:

$$g_k^{(1)} = \begin{cases} \frac{1}{\sqrt{4\pi\omega}}e^{-i\omega\eta+ik\xi} & in\ A \\ 0 & in\ D \end{cases}, \tag{3.16a}$$

$$g_k^{(2)} = \begin{cases} 0 & in\ A \\ \frac{1}{\sqrt{4\pi\omega}}e^{+i\omega\eta+ik\xi} & in\ D \end{cases}. \tag{3.16b}$$

In this way, we have made explicit the fact that $\{\eta, \xi\}$ are defined differently on the two distinct regions, by actually splitting the modes into two sets. Therefore the natural orthonormal basis that Rindler observers use for expanding ϕ is $g_k^{(1)}$, $g_k^{(2)}$, $g_k^{(1)*}$, $g_k^{(2)*}$, whose domain is the union of the regions A and D, corresponding to the constraint $|x| > |t|$. We are soon going to construct an analytic extension of these modes to all the Minkowski spacetime.

We thus have two sets of modes on which we can expand our field:

- **Minkowski modes** $\{f_k, f_k^*\}$, defined by (2.19), on which the field expands as in (2.29):

$$\hat{\phi} = \int_{-\infty}^{+\infty} dk\big(\hat{a}_k f_k + \hat{a}_k^\dagger f_k^*\big), \tag{3.17}$$

  and whose vacuum state is $|0_M\rangle$;

- **Rindler modes** $\{g_k^{(1)}, g_k^{(2)}, g_k^{(1)*}, g_k^{(2)*}\}$, on which the fields expands as

$$\hat{\phi} = \int_{-\infty}^{+\infty} dk\big(\hat{b}_k^{(1)}g_k^{(1)} + \hat{b}_k^{(1)\dagger}g_k^{(1)*} + \hat{b}_k^{(2)}g_k^{(2)} + \hat{b}_k^{(2)\dagger}g_k^{(2)*}\big), \tag{3.18}$$

  and whose vacuum state is $|0_R\rangle$, which is annihilated by all $\hat{b}_k^{(1)}$ and $\hat{b}_k^{(2)}$.

We could now employ Bogoljubov transformations (2.47) to compute the expectation value of Rindler number operators in the Minkowski vacuum. However, there is a trick due to Unruh that is more instructive, which consists in analytically extending the modes (3.16), defining an auxiliary set of modes that share the same vacuum state of Minkowski modes, and then employing the Bogoljubov transformation between Rindler modes and these new ones to compute the expectation value of number operators in the Minkowski vacuum.

3.3 Analytic extension of Rindler modes

We have seen that the domain of the modes $g_k^{(1)}$ and $g_k^{(2)}$ defined in (3.16) is not the whole Minkowski spacetime. To get their analytic extension we will make use of the following identities:

$$\textit{in region } A \quad \begin{cases} e^{-a(\eta-\xi)} = a(-t+x) & (3.19\text{a}) \\ e^{a(\eta+\xi)} = a(t+x) & (3.19\text{b}) \end{cases}$$

$$\textit{in region } D \quad \begin{cases} e^{-a(\eta-\xi)} = a(t-x) & (3.19\text{c}) \\ e^{a(\eta+\xi)} = a(-t-x). & (3.19\text{d}) \end{cases}$$

Proof. *In region A, by using* (3.4) *we get*

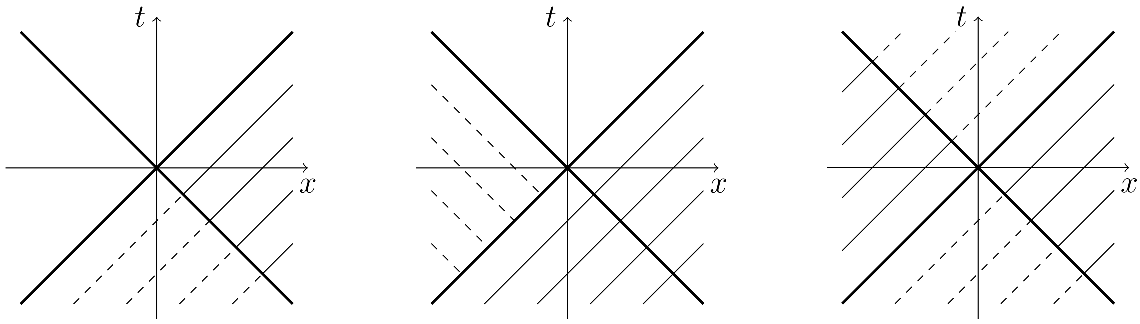
$$\begin{aligned} -t+x &= \frac{1}{a}e^{a\xi}\left(\frac{e^{a\eta}+e^{-a\eta}}{2} - \frac{e^{a\eta}-e^{-a\eta}}{2}\right) = \frac{1}{a}e^{a(\xi-\eta)} \;\Rightarrow\; e^{-a(\eta-\xi)} = a(-t+x), \\ t+x &= \frac{1}{a}e^{a\xi}\left(\frac{e^{a\eta}+e^{-a\eta}}{2} + \frac{e^{a\eta}-e^{-a\eta}}{2}\right) = \frac{1}{a}e^{a(\xi+\eta)} \;\Rightarrow\; e^{a(\eta-\xi)} = a(t+x), \end{aligned}$$

while in region D, by using (3.5) *we have*

$$\begin{aligned} t-x &= -\frac{1}{a}e^{a\xi}\left(\frac{e^{a\eta}-e^{-a\eta}}{2} - \frac{e^{a\eta}+e^{-a\eta}}{2}\right) = \frac{1}{a}e^{a(\xi-\eta)} \;\Rightarrow\; e^{-a(\eta-\xi)} = a(t-x), \\ -t-x &= -\frac{1}{a}e^{a\xi}\left(\frac{e^{-a\eta}-e^{a\eta}}{2} - \frac{e^{-a\eta}+e^{a\eta}}{2}\right) = \frac{1}{a}e^{a(\xi+\eta)} \;\Rightarrow\; e^{a(\eta+\xi)} = a(-t-x). \end{aligned}$$

We are now going to consider right-moving modes $k > 0$ and left-moving ones $k < 0$ separately, starting with $k > 0$. In this case, the dispersion relation becomes $\omega = k$, and the modes $g_k^{(1)}$ become, in region A,

$$g_k^{(1)} = \frac{1}{\sqrt{4\pi\omega}}[a(-t+x)]^{i\omega/a}, \tag{3.20}$$



(a) $-t+x = const$ contour plot of $g_k^{(1)}$ and analytical extension (dashed).

(b) Non-analytical union of $g_k^{(1)}$ (continuous line) and $g_k^{(2)}$ (dashed line) contour plots.

(c) Union of $g_k^{(1)}$ and $g_{-k}^{(2)*}$ contour plots with analytical extensions (dashed).

Fig. 3.2: Analytical extension process for the Rindler modes.

while the modes $g_k^{(2)}$, in region D become

$$g_k^{(2)} = \frac{1}{\sqrt{4\pi\omega}}[a(-t+x)]^{i\omega/a}. \tag{3.21}$$

Proof. *By substituting $k = \omega$ in* (3.16a) *and using* (3.19a) *we get, in region A,*

$$\sqrt{4\pi\omega}g_k^{(1)} = e^{-i\omega(\eta-\xi)} = [e^{-a(\eta-\xi)}]^{i\omega/a} = [a(-t+x)]^{i\omega/a},$$

whereas by (3.19d)*, in region D,*

$$\sqrt{4\pi\omega}g_k^{(2)} = e^{i\omega(\eta+\xi)} = [e^{a(\eta+\xi)}]^{i\omega/a} = [a(-t-x)]^{i\omega/a}.$$

We can make $g_k^{(1)}$ cover the C region just by extending the range of t and x allowing $x > t$ (regions A and C), and the resulting mode will be automatically analytical, as depicted in the contour plot $-t + x = const$ in fig. 3.2a. It is important to notice that it does not make sense to extend $g_k^{(1)}$ to region D because it is already defined there as identically vanishing. We could also extend to the region B, but we are going to cover this region using $g_k^{(2)}$, which is non-trivial only in the region D. However, if we inspect (3.21), we see that its expression is not compatible with (3.20) because there would be a discontinuity at the surface $t = x$ for $x < 0$, which can be thought as contour plots $-t + x = const$ and $-t - x = const$ forming cusps there (fig. 3.2b). To solve this problem, we are going to consider the complex conjugate of $g_k^{(2)}$ with a flipped sign in the subscript, which results in a compatible expression whose contour plot consists of $-t + x = const$ curves as desired (fig. 3.2c) and is easily extended on the B region by allowing $x < t$:

$$g_{-k}^{(2)*} = \frac{1}{\sqrt{4\pi\omega}}[e^{-i\pi}a(-t+x)]^{i\omega/a}. \tag{3.22}$$

Proof. *By complex conjugating* (3.21)*, using* $\omega = k > 0$ *and by* (3.19c)*, we have*

$$\sqrt{4\pi\omega}g_{-k}^{(2)*} = e^{-i\omega(\eta-\xi)} = [e^{-a(\eta-\xi)}]^{i\omega/a} = [a(t-x)]^{i\omega/a} = [e^{-i\pi}a(-t+x)]^{i\omega/a}.$$

The reason why we chose $-1 = e^{-i\pi}$ will be clear in a moment.

We therefore have a way to define a new mode that represents right-moving ($k > 0$) plane waves (as in fig. 3.2c) in all Minkowski spacetime:

$$h_k^{(1)} = N[a(-t+x)]^{i\omega/a}, \tag{3.23}$$

with N being a normalization factor. This yields

$$h_k^{(1)} = \frac{1}{\sqrt{2\sinh(\frac{\pi\omega}{a})}}\big(e^{\frac{\pi\omega}{2a}}g_k^{(1)} + e^{-\frac{\pi\omega}{2a}}g_{-k}^{(2)*}\big). \tag{3.24}$$

Proof. *From* (3.23) *and using* (3.20) *and* (3.22)*, we get*

$$h_k^{(1)} = N[a(-t+x)]^{i\omega/a} = N\big(g_k^{(1)} + e^{-\pi\omega/a}g_{-k}^{(2)*}\big),$$

where N *is a normalization factor that also absorbed* $\sqrt{4\pi\omega}$*. We can multiply by* $e^{\pi\omega/2a}$ *and absorb its inverse into the normalization factor, obtaining*

$$h_k^{(1)} = N\big(e^{\frac{\pi\omega}{2a}}g_k^{(1)} + e^{-\frac{\pi\omega}{2a}}g_{-k}^{(2)*}\big),$$

and then use the inner product (2.16) *to compute the normalization factor:*

$$\begin{aligned}
\langle h_{k_1}^{(1)}, h_{k_2}^{(2)}\rangle &= \big\langle N\big(e^{\frac{\pi\omega_1}{2a}}g_{k_1}^{(1)} + e^{-\frac{\pi\omega_1}{2a}}g_{-k_1}^{(2)*}\big), N\big(e^{\frac{\pi\omega_2}{2a}}g_{k_2}^{(1)} + e^{-\frac{\pi\omega_2}{2a}}g_{-k_2}^{(2)*}\big)\big\rangle \\
&= |N|^2\big(e^{\frac{\pi}{2a}(\omega_1+\omega_2)}\langle g_{k_1}^{(1)}, g_{k_2}^{(2)}\rangle + e^{-\frac{\pi}{2a}(\omega_1+\omega_2)}\langle g_{-k_1}^{(1)*}, g_{-k_2}^{(2)*}\rangle\big) \\
&= |N|^2\delta(k_1-k_2)(e^{\frac{\pi\omega}{a}} - e^{-\frac{\pi\omega}{a}}) \\
&= |N|^2\delta(k_1-k_2)2\sinh\left(\frac{\pi\omega}{a}\right),
\end{aligned}$$

where we set $\omega = \omega_1 = \omega_2$ *when the delta is present and we used orthonormality of the* $g_k^{(1)}$ *modes and their conjugate modes as in* (2.41)*. Since we require normalization to be* $\delta(k_1 - k_2)$*, we can choose* N *real such that*

$$N^2 2\sinh\left(\frac{\pi\omega}{a}\right) = 1 \quad \Rightarrow \quad N = \frac{1}{\sqrt{\sinh\left(\frac{\pi\omega}{a}\right)}}.$$

If we retrace the same steps for $k < 0$, we get the same result (3.24).

Proof. *For $k = -\omega < 0$, by* (3.16) *and* (3.19) *we have*

$$\sqrt{4\pi\omega}g_k^{(1)} = e^{-i\omega(\eta+\xi)} = [e^{a(\eta+\xi)}]^{-i\omega/a} = [a(t+x)]^{-i\omega/a},$$
$$\sqrt{4\pi\omega}g_{-k}^{(2)*} = e^{-i\omega(\eta+\xi)} = [e^{a(\eta+\xi)}]^{-i\omega/a} = [a(-t-x)]^{-i\omega/a} = [e^{-i\pi}a(t+x)]^{-i\omega/a},$$

where again we chose $-1 = e^{-i\pi}$ for a reason that we are going to clarify in a moment. Combining the two expressions, we get (because they vanish in each other's non-vanishing region)

$$[a(t+x)]^{-i\omega/a} = N\big(g_k^{(1)} + e^{-\frac{\pi\omega}{a}} g_{-k}^{(2)*}\big) = N\big(e^{\frac{\pi\omega}{2a}} g_k^{(1)} + e^{-\frac{\pi\omega}{2a}} g_{-k}^{(2)*}\big),$$

where in the last step we multiplied by $e^{\frac{\pi\omega}{2a}}$ and absorbed its inverse into the normalization factor N. The term in the parentheses turns out to be the same as the one in (3.24) *and therefore after normalization, the full expression is identical to it and analytical everywhere.*

To get a complete orthonormal basis for the Hilbert state space of the field, we also need to extend $g_k^{(2)}$ analytically, and if we retrace analogous steps as those we took for $g_k^{(1)}$ we get a new mode:

$$h_k^{(2)} = \frac{1}{\sqrt{2\sinh\left(\frac{\pi\omega}{a}\right)}}\big(e^{\frac{\pi\omega}{2a}} g_k^{(2)} + e^{-\frac{\pi\omega}{2a}} g_{-k}^{(1)*}\big). \tag{3.25}$$

Now we have that $h_k^{(1)}$ and $h_k^{(2)}$, along with their complex conjugates $h_k^{(1)*}$ and $h_k^{(2)*}$, form together an orthonormal basis of modes, called *Unruh modes.*

3.4 Particle number expectation value

The Unruh modes allow for an expansion of the field with coefficients given by creation-annihilation operators, again because of the commutation relations of the field and the orthonormality relations of the modes:

$$\hat{\phi} = \int dk\big(\hat{c}_k^{(1)}h_k^{(1)} + \hat{c}_k^{(1)\dagger}h_k^{(1)*} + \hat{c}_k^{(2)}h_k^{(2)} + \hat{c}_k^{(2)\dagger}h_k^{(2)*}\big). \tag{3.26}$$

We can now use the Bogoljubov transformation (2.47a) between Rindler and Unruh modes to relate their creation-annihilation operators, obtaining

$$\begin{cases} \hat{b}_k^{(1)} = \dfrac{1}{\sqrt{2\sinh(\frac{\pi\omega}{a})}}\big(e^{\frac{\pi\omega}{2a}}\hat{c}_k^{(1)} + e^{-\frac{\pi\omega}{2a}}\hat{c}_{-k}^{(2)\dagger}\big), & \text{(3.27a)} \\ \hat{b}_k^{(2)} = \dfrac{1}{\sqrt{2\sinh(\frac{\pi\omega}{a})}}\big(e^{\frac{\pi\omega}{2a}}\hat{c}_k^{(2)} + e^{-\frac{\pi\omega}{2a}}\hat{c}_{-k}^{(1)\dagger}\big), & \text{(3.27b)} \end{cases}$$

where we used the fact that in this case the Bogoljubov transformation should have four terms because the basis is $\{g_k^{(1)}, g_k^{(1)*}, g_k^{(2)}, g_k^{(2)*}\}$, but two of them are vanishing in each of the expressions (3.24) and (3.25), and the remaining coefficients are real.

We are now going to show that the Unruh vacuum state is the same as Minkowski (although excited states will not coincide in general):

$$|0_M\rangle = |0_U\rangle\,. \tag{3.28}$$

Proof. *Let us consider positive-frequency Minkowski modes* (2.19)

$$f_k = \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega t + ikx}.$$

If we restrict to right-moving modes, by $k = \omega > 0$ we get

$$\sqrt{4\pi\omega}\, f_k = e^{-i\omega(t-x)} = e^{-i\omega z},$$

where we have introduced the complex variable z. It is easy to realize that this expression is analytic everywhere and bounded in the region $\mathfrak{Im} z < 0$, while it is unbounded for $\mathfrak{Im} z > 0$. Now, consider that the Unruh mode $h_k^{(1)}$ with $k > 0$ is composed of Rindler modes $g_k^{(1)}$ and $g_{-k}^{(2)*}$. These latter are multivalued complex functions, in fact, if we use (3.16), (3.19) *and complexify $(t-x) \to z = \rho e^{i\theta}$ we get*

$$\begin{aligned}
\sqrt{4\pi\omega} g_k^{(1)} &= [a(-t+x)]^{\frac{i\omega}{a}} = a^{\frac{i\omega}{a}} (-z)^{\frac{i\omega}{a}} = (a\rho)^{\frac{i\omega}{a}} e^{\frac{\pi\omega}{a}} e^{-\frac{\theta\omega}{a}}, \\
\sqrt{4\pi\omega} g_{-k}^{(2)*} &= [e^{-i\pi} a(-t+x)]^{\frac{i\omega}{a}} = a^{\frac{i\omega}{a}} e^{\frac{\pi\omega}{a}} (-z)^{\frac{i\omega}{a}} = (a\rho)^{\frac{i\omega}{a}} e^{\frac{2\pi\omega}{a}} e^{-\frac{\theta\omega}{a}},
\end{aligned}$$

which are clearly not uniquely specified for $\theta \to \theta + 2m\pi$. As complex-valued functions, there needs to be a branch cut, and we choose it to be on the upper half plane $\mathfrak{Im} z > 0$. This justifies the choice $-1 = e^{-i\pi}$ we made earlier. If we set the range of θ, the two expressions above are bounded and since the branch cut is on the upper half plane they are analytical for $\mathfrak{Im} z < 0$.

Now, since the overall mode $h_k^{(1)}$ is therefore analytic and bounded in the lower half plane and not above, it can only be expressed through Minkowski modes that are analytic and bounded in the same region, which turn out to be f_k with $k > 0$. The same holds for $k = -\omega < 0$, in fact for the complexification $(t+x) \to z = \rho e^{i\theta}$ (and employing as usual (3.16) *and* (3.19)*):*

$$\begin{aligned}
\sqrt{4\pi\omega} f_k &= e^{-i\omega(t+x)} = e^{-i\omega z}, \\
\sqrt{4\pi\omega} g_k^{(1)} &= [a(t+x)]^{-\frac{i\omega}{a}} = a^{-\frac{i\omega}{a}} z^{-\frac{i\omega}{a}} = (a\rho)^{-\frac{i\omega}{a}} e^{\frac{\theta\omega}{a}}, \\
\sqrt{4\pi\omega} g_{-k}^{(2)*} &= [e^{-i\pi} a(t+x)]^{-\frac{i\omega}{a}} = a^{-\frac{i\omega}{a}} e^{-\frac{\pi\omega}{a}} z^{-\frac{i\omega}{a}} = (a\rho)^{-\frac{i\omega}{a}} e^{-\frac{\pi\omega}{a}} e^{\frac{\theta\omega}{a}},
\end{aligned}$$

and we can see that the Minkowski mode f_k is analytic and bounded for $\mathfrak{Im} z < 0$, and so we can choose the branch cut for $h_k^{(1)}$ in the upper half plane so that it can be expanded in terms of positive-frequency Minkowski modes only.

The same happens for $h_k^{(2)}$ modes. In fact, for the left-moving $h_k^{(2)}$ modes we have ($k = -\omega < 0$)

$$\begin{aligned}
\sqrt{4\pi\omega} f_{-k} &= e^{-i\omega(t-x)} = e^{-i\omega z}, \\
\sqrt{4\pi\omega} g_k^{(2)} &= [a(t-x)]^{-\frac{i\omega}{a}} = a^{-\frac{i\omega}{a}} z^{-\frac{i\omega}{a}} = (a\rho)^{-\frac{i\omega}{a}} e^{\frac{\theta\omega}{a}}, \\
\sqrt{4\pi\omega} g_{-k}^{(1)*} &= [e^{-i\pi} a(t-x)]^{-\frac{i\omega}{a}} = a^{-\frac{i\omega}{a}} e^{-\frac{\pi\omega}{a}} z^{-\frac{i\omega}{a}} = (a\rho)^{-\frac{i\omega}{a}} e^{-\frac{\pi\omega}{a}} e^{\frac{\theta\omega}{a}},
\end{aligned}$$

and for the right moving ones ($k = \omega > 0$)

$$\begin{aligned}
\sqrt{4\pi\omega} f_{-k} &= e^{-i\omega(t+x)} = e^{-i\omega z}, \\
\sqrt{4\pi\omega} g_k^{(2)} &= [e^{-i\pi} a(t+x)]^{\frac{i\omega}{a}} = a^{\frac{i\omega}{a}} e^{\frac{\pi\omega}{a}} z^{\frac{i\omega}{a}} = (a\rho)^{\frac{i\omega}{a}} e^{\frac{\pi\omega}{a}} e^{-\frac{\theta\omega}{a}}, \\
\sqrt{4\pi\omega} g_{-k}^{(1)*} &= [a(t-x)]^{\frac{i\omega}{a}} = a^{\frac{i\omega}{a}} z^{\frac{i\omega}{a}} = (a\rho)^{\frac{i\omega}{a}} e^{-\frac{\theta\omega}{a}}.
\end{aligned}$$

We have thus shown that the Unruh modes $\{h_k^{(1)}, h_k^{(1)*}, h_k^{(2)}, h_k^{(2)*}\}$ are fully expressable in terms of positive-frequency Minkowski modes f_k because of analytical considerations, so by the Bogoljubov transformation, it means that the Unruh annihilation operators $\hat{c}_k^{(1)}$ and $\hat{c}_k^{(2)}$ are expressed in terms of Minkowski annihilation operators $\hat{a}_k$ only. Therefore, the Minkowski vacuum $|0_M\rangle$ is annihilated by all Unruh annihilation operators, so it matches the definition of Unruh vacuum $|0_U\rangle$.

We are now able to compute the expectation value of the Rindler number operator $\hat{n}_R^{(1)}(k)$ corresponding to an accelerated observer in region A, which results

$$\langle 0_M | \hat{n}_R^{(1)}(k) | 0_M \rangle = \frac{1}{e^{\frac{2\pi\omega}{a}} - 1} \delta(0). \tag{3.29}$$

Proof. *By the definition of number operator and upon using the Bogoljubov transformation* (3.27)*, we have*

$$\langle 0_M | \hat{n}_R^{(1)}(k) | 0_M \rangle = \langle 0_M | \hat{b}_k^{(1)\dagger} \hat{b}_k^{(1)} | 0_M \rangle = \frac{e^{-\frac{\pi\omega}{a}}}{2\sinh\left(\frac{\pi\omega}{a}\right)} \left(\langle 0_M | \hat{c}_{-k}^{(2)} \right) \left(\hat{c}_{-k}^{(2)\dagger} | 0_M \rangle \right).$$

Since the state $\hat{c}_{-k}^{(2)\dagger} |0_M\rangle$ is normalized, we can write

$$\langle 0_M | \hat{n}_R^{(1)}(k) | 0_M \rangle = \frac{e^{-\frac{\pi\omega}{a}}}{2\sinh\left(\frac{\pi\omega}{a}\right)} \delta(0) = \frac{e^{-\frac{\pi\omega}{a}}}{e^{\frac{\pi\omega}{a}} - e^{-\frac{\pi\omega}{a}}} \delta(0),$$

which proves (3.29) *after multiplying both numerator and denominator by* $e^{\frac{\pi\omega}{a}}$.

The delta function in (3.29) is not something to worry about, because we have been using non-square-integrable plane waves as a generalized orthonormal basis for our Hilbert space, and that divergence is a consequence of this. Had we used square-integrable wave packets as basis modes, we would have obtained the same result.

We notice that the expectation value is a Bose-Einstein statistical distribution with temperature (we are using $k_B = 1$)

$$T = \frac{a}{2\pi}, \tag{3.30}$$

so there is a thermal spectrum of non-interacting particles detected by the Rindler accelerated observer: this is the Unruh effect. However, the frequency ω does not represent the energy that the Rindler observer measures using its proper time τ, but rather the frequency with respect to the coordinate η, so it has to be corrected by a factor of a/α according to (3.15), where α represents the modulus of proper acceleration:

$$\omega_\eta = \frac{a}{\alpha}\omega_\tau. \tag{3.31}$$

If we substitute in (3.29), we get a temperature of

$$T = \frac{\alpha}{2\pi}. \tag{3.32}$$

By (3.7), we have the relation between α and a:

$$\alpha = ae^{-a\xi}, \tag{3.33}$$

which is compatible with the redshift factor (3.12), meaning that a Rindler observer characterized by $\xi_1 = 0$ detects a temperature of $T_1 = a/2\pi$, while another Rindler observer with $\xi_2 = \xi$ detects a temperature of $T_2 = e^{-a\xi}a/2\pi = \alpha/2\pi$ because $T_1V_1 = T_2V_2$ (by (1.13) and the fact that temperature represents energy if $k_B = 1$) and $V_1 = 1$, $V_2 = e^{a\xi}$. We notice that if $\xi \to \infty$ the redshift factor goes to infinity, therefore the thermal radiation vanishes. This is to be expected, since $\xi \to \infty$ represents inertial observers (because the modulus of proper acceleration $\alpha = e^{-a\xi}a$ goes to zero), whose vacuum is exactly the Minkowski one.

We have thus proved that the observer with proper acceleration α in Minkowski vacuum detects thermal radiation with temperature $T = \alpha/2\pi$, which in standard units can be written as

$$T = \frac{\hbar\alpha}{2\pi k_B c}. \tag{3.34}$$

Before proceeding to derive the temperature of Hawking radiation, we shall stress the fact that our argument implicitly relied on the existence of the Killing horizon $t = x$, because that is where Rindler coordinates start not working anymore. This led us to look for the analytic extension of Rindler modes, combine them into Unruh modes, and obtain a vacuum state that is regular on the Killing horizon due to the analyticity of Unruh modes. Without a Killing horizon, our argument fails, so in some sense, we can think that it is the horizon itself that radiates. This is crucial for the understanding of Hawking radiation, which we are going to study in the last chapter.

Chapter 4

Hawking radiation

4.1 Temperature and entropy

Let us consider a Schwarzschild black hole with mass M in spherical spacial coordinates $\{t, r, \theta, \varphi\}$. Since the Schwarzschild metric is static and asymptotically flat, there exists a timelike surface-orthogonal Killing vector field that is proportional to the direction of proper time $\tau = t$ of stationary observers at infinity and can be normalized such that $\vec{K} = \partial_t$. For a stationary observer O_1 just outside the black hole, in the proximity of the event horizon $r_1 \simeq R_H = 2GM$, the proper acceleration has modulus $a_1 \gg R_H^{-1}$ and the local flatness requirement can be given a magnitude order, that is time and length scales (measured by proper time τ, which is the distance in spacetime) being much lower than the ones needed for the acceleration to become non-negligible, thus $\Delta\tau \ll R_H$.

Proof. *By manipulating the expression of the modulus of four-acceleration for a stationary observer* (1.24) *and using $r_1 - R_H \ll R_H$ (thus $r_1 \simeq R_H$),*

$$a = \frac{R_H}{2r^2\sqrt{1 - \frac{R_H}{r}}} = \frac{\sqrt{R_H}}{2r^{\frac{3}{2}}\sqrt{\frac{r-R_H}{R_H}}} \quad \Rightarrow \quad a_1 \gg \frac{1}{2R_H} \sim \frac{1}{R_H}.$$

To have local flatness, the variation of components of four-velocity must be negligible (i.e. the observer must be inertial), $\Delta u^\mu \ll |u| = 1$:

$$\Delta u^\mu \lesssim a\Delta\tau \ll 1,$$

which implies

$$\Delta\tau \ll \frac{1}{a} \sim R_H.$$

Since local flatness means that the spacetime can locally be approximated by Minkowski spacetime, it means that the vacuum state is the Minkowski one locally. However, since the observer has a proper acceleration with modulus a, the Unruh effect comes into play, and the observer detects thermal radiation with temperature $T = a_1/2\pi$ according to (3.30). In order to compute what temperature a distant stationary observer O_2 detects ($r_2 \gg R_H$ and the redshift factor is $V_2 = 1$ by plugging $r \to \infty$ in (1.15)) we cannot rely on any concrete concept of local flatness because we do not have a hypothesis like $r_1 - R_H \ll R_H$ which allowed us to assign an order of magnitude for that. We are able, however, to exploit the fact that stationary observers follow orbits of the Killing vector field $\vec{K} = \partial_t$, so the temperature detected by the observer O_2 at infinity is

$$T = \frac{\kappa}{2\pi}, \tag{4.1}$$

where κ is the surface gravity of the black hole.

Proof. *Using the redshift factor as in* (1.13) *with* $V_2 = 1$, *and the fact that temperatures are energies for* $k_B = 1$, *we have*

$$T_2 = \lim_{r_1 \to R_H^+} \frac{V_1}{V_2} T_1 = \lim_{r_1 \to R_H^+} \frac{V_1 a_1}{2\pi} = \frac{\kappa}{2\pi},$$

where we take the limit because at the event horizon $V_1 \to 0$ *and* $a_1 \to \infty$, *and their product is finite and defined as the surface gravity* (1.23).

If we substitute the surface gravity of a Schwarzschild black hole (1.25) and restore standard units, we get the famous formula for the temperature of a Schwarzschild black hole:

$$T = \frac{\hbar c^3}{8\pi k_B G M}. \tag{4.2}$$

Notice that in this case, the temperature does not vanish at spacial infinity as it did in the Unruh effect case. This is due to the particular form of the Schwarzschild metric and can be boiled down to the very existence of the black hole, which is not present in flat spacetime.

We should stress that our arguments rely on the existence of an event horizon, which allowed us to compute the order of magnitude of local flatness, which in turn was needed to make the analogy with the Unruh effect. Without the horizon, no such argument is possible since there is nothing that gives us a spacetime scale in which the Unruh effect can be applied. In some sense, the fictitious Killing horizon that we had in Minkowski spacetime for accelerated observers corresponds to the actual event horizon in Schwarzschild spacetime, and the vacuum state seen by free-falling observers near the horizon corresponds to the Minkowski vacuum seen by inertial observers in flat spacetime so that a stationary observer near the horizon corresponds to an accelerated observer in (locally) flat spacetime. This is nothing else than Einstein's equivalence principle at play. In addition, it is required that the field is regular on the event horizon, just as we did in the Unruh case when we extended the Rindler modes analytically over the Killing horizons. In this case, the field vacuum state can be shown to be regular on the Schwarzschild event horizon, because a free-falling observer can follow a geodesic path that crosses the horizon with continuity, and he detects local flatness everywhere, which is exactly what the field needs to be regular.

Another important consideration regards the validity of our approximation $r_1 - R_H \ll R_H$. This breaks down if we consider wavelengths comparable to R_H, since local flatness would not extend to such a scale and the Unruh formula for the temperature would not be valid anymore. However, the energy contributions of these wavelengths are negligible in the case of black holes characterized by a mass of the order of the sun's $M_\odot$.

Now, we can define the entropy of the black hole just by employing the thermodynamical definition $dS = \delta Q_{rev}/T$ and considering a quasi-static process of formation of a black hole from zero mass to mass M (it is just an approximation since we will see in a few moments that if the black hole is small it radiates away an enormous amount of energy in short periods of time, as can be viewed from the fact that the temperature is inversely proportional to the mass). The result is:

$$S = \frac{A}{4G}, \tag{4.3}$$

where $A = 4\pi R_H^2 = 16\pi G^2 M^2$ is the area of the event horizon.

Proof. *In natural units, $dQ = dM$, so we have $dM = TdS$. By using the formula for the temperature* (4.1) *with surface gravity $\kappa = 1/4GM$* (1.25)*, we have*

$$dS = \frac{dM}{T} = 8\pi GMdM.$$

If we integrate,

$$S = \int_0^S dS = \int_0^M 8\pi GMdM = 4\pi GM^2 = \frac{4\pi(2GM)^2}{4G} = \frac{A}{4G}.$$

In standard units, we have, by expliciting the area,

$$S = \frac{4\pi k_B GM^2}{\hbar c}. \tag{4.4}$$

Now that we have quantities of interest such as temperature and entropy, we can make some interesting considerations regarding black hole evaporation and information loss.

4.2 Black hole evaporation

We have already stated that the thermal radiation (3.29) in the Unruh effect, which is of the same type as the one in the Hawking effect, looks like a Bose-Einstein statistical distribution, and since the field we are considering is massless, it reduces to the particular case of Planck's law for spectral radiance of blackbody radiation, and therefore we can approximate the black hole as a blackbody and use the Stephan-Boltzmann law to compute the time needed for a black hole to completely disappear by radiating particles, according to a stationary observer at infinity.

Of course, this is just an approximation for the regime when the black hole has a mass comparable to that of the sun because we already noticed that our analogy with the Unruh effect breaks down when the wavelength is comparable to the Schwarzschild radius. In other words, if the black hole is big enough, there will be a first phase when our approximation holds, but as the black hole keeps radiating and shrinking, it will eventually reach a size for which the average wavelength of its radiation is comparable to its radius, and our approximation breaks down. However, in this second phase, the emission of energy is so fast that the time needed to radiate away the remaining mass is negligible in comparison to that of the first phase, so we still get a good estimate of the evaporation time.

Another consideration is that we are performing the calculation assuming that there is only one scalar quantum field that interacts with the curvature and radiates particles, but in the standard model there are multiple interacting fields, so to get a better estimate we would need to consider the contribution of all such fields. However, we shall expect that the order of magnitude that we will get is still indicative of the time scale of the black hole evaporation, because the energy radiated by each field can be assumed to be of the same order as the one we are considering, and in the standard model of particle physics there are only 17 fields.

Let us consider a Schwarzschild black hole with mass M and radius $R_H = 2GM$. Stefan-Boltzmann law states that the energy radiated per unit time per unit area by a blackbody is proportional to the fourth power of its temperature by

$$\frac{d^2E}{dtdA} = \frac{\pi^2}{60}T^4. \tag{4.5}$$

If we explicit the area of the event horizon $A = 4\pi R_H^2 = 16\pi G^2M^2/c^4$ (we are restoring standard units) and the temperature (4.2) we obtain, by a straightforward computation,

$$\frac{dE}{dt} = \frac{\hbar c^6}{15360\pi G^2M^2}. \tag{4.6}$$

Using this equation, we can compute the time Δt needed by the radiation to completely radiate away the black hole mass (in standard units):

$$\Delta t = \frac{5120\pi G^2M^3}{\hbar c^4}. \tag{4.7}$$

Proof. *Since* $dE = c^2 dM$*, by* (4.6) *and separating variables we get*

$$M^2 dM = \frac{\hbar c^4 dt}{15360\pi G^2},$$

which, after integration, gives

$$\frac{M^3}{3} = \frac{\hbar c^4 \Delta t}{15360\pi G^2},$$

where we can easily isolate Δt*.*

We see that the time needed for the evaporation increases as the third power of the mass of the black hole. We can plug some numbers into (4.7) to see what is the order of magnitude of such a time interval. The smallest black hole that has been detected by now is XTE J1650-500 in a binary system in the Milky Way, and its mass is $M \simeq 3.8 M_\odot$, where $M_\odot \simeq 2.0 \times 10^{30}\ kg$ is the mass of the sun. The computation of Δt for this black hole gives $\Delta t \simeq 3.7 \cdot 10^{76}\ s$, which, in years, is $\Delta t \simeq 1.2 \cdot 10^{69}\ y$. Since the estimated age of the universe is $\tau = 1.4 \cdot 10^{10}\ y$, what we got is almost 60 orders of magnitude greater than the age of the universe, which should give an idea of how long the process is. In addition, since the surroundings of this black hole contain a lot of matter, the time we obtained is an underestimation, because in our calculation we were assuming that spacetime was empty around the black hole, so its mass had no way to increase.

4.3 Information loss paradox

In statistical mechanics, we interpret the entropy as the logarithm of the number of microstates Γ that are accessible to our system for fixed macroscopic variables:

$$S = k_B \ln \Gamma \tag{4.8}$$

In the case of black holes, since we have an expression for the entropy (4.3), one could ask where the information about the microstate is stored. There is an important no-hair theorem that states that black holes in static asymptotically flat spacetimes like the ones we focused on are fully characterized by mass, spin, electric, and magnetic charge, and no other information is needed to describe them. The black hole (micro-) state therefore tells us no more than macroscopic variables do, so the entropy should be zero, in contrast with our result that tells us it is proportional to the area of the event horizon. Some interpretations of this contradiction are given by the *holographic principle*, which conjectures that no information is present inside the black hole, but rather it stays on the event horizon so that the no-hair theorem is not violated. If we cast the expression for the entropy (4.3) in Planck units we get $S = A/4$, and restoring standard units and using the Planck length $l_P = \sqrt{\frac{\hbar G}{c^3}}$ we get

$$S = k_B \frac{A}{4l_P^2}, \tag{4.9}$$

where we simply reintroduced the Boltzmann constant, which has the dimensions of entropy, and in the denominator, the Planck area appears, compatibly with Planck units. If we equate the entropy (4.9) with the information entropy

$$N = \frac{1}{\ln 2} \ln \Gamma, \tag{4.10}$$

which is the number of bits needed to store the information that is encoded in the system, we get

$$N = \frac{A}{4l_P^2 \ln 2}, \tag{4.11}$$

where we used (4.8) and (4.9) to say that $\ln \Gamma = A/4l_P^2$. This tells us that every region with area $4l_P^2 \ln 2$ (of the order of the Planck area) of the event horizon encodes one bit of information, and in the holographic interpretations this is thought as having some physical meaning which is hidden behind the Planck scale.

Out of curiosity, we could compute how many bits of information the black hole XTE J1650-500 mentioned in the previous section encodes. Using (4.11) with $A = 16\pi G^2 M^2/c^4$ and $l_P^2 = \hbar G/c^3$ we get

$$N = \frac{4\pi G M^2}{\hbar c \ln 2},$$

which gives $N \simeq 5.1 \cdot 10^{78}$ after plugging $M \simeq 3.8 M_\odot$. This means that XTE J1650-500 encodes around $6.4 \cdot 10^{56}$ zettabytes of data, which is more than 50 orders of magnitude greater than the estimated volume of data ever created, stored, copied, and consumed worldwide ever since the advent of digitalization according to [8], which in turn is no more than 10^3 zettabytes.

However, there is a problem with the holographic interpretation, that pops up with the disappearance of the event horizon once the black hole has completely evaporated through Hawking radiation. In fact, without an event horizon, there is no place to store information, and since the radiation is thermal, by definition it cannot encode much information, especially such an enormous amount like the one we have just computed. Therefore, information seems to be lost, in contrast with important theorems of conservation of information in quantum mechanics. If two completely different initial states were to form two distinct black holes, there would be no way to backtrack to the original states from the final states where both black holes have evaporated. This is the information loss paradox, which does not have a solidly accepted solution, for now.

Conclusion

We have concluded our derivation of Hawking radiation, effectively obtaining expressions for the temperature and entropy of a Schwarzschild black hole, which we have used to speculate about the evaporation time of black holes and the encoding of information on the event horizon, ultimately leading to the information loss paradox.

Retracing the steps we have taken to arrive at the results we aimed for, we first defined important quantities such as the redshift factor, the acceleration of stationary observers in asymptotically flat static spacetimes like the Schwarzschild one, Killing horizons as null hypersurfaces with orthogonal null Killing vector fields, and surface gravity as the limit of the product of redshift factor and stationary acceleration in the proximity of the event horizon.

After that, we took a detour into quantum field theory in flat and curved spacetime, highlighting analogies and differences. We learned that, in flat spacetime, particles are energy excitations of quantum fields, and the eigenstates of the hamiltonian of such fields are given by eigenstates of number operators defined by expanding the field on positive and negative frequency modes and interpreting the coefficients of the expansion as creation and annihilation operator pairs. We saw that the modes used as a basis for the solution space of the Klein-Gordon equation can be transformed into one another using Bogoljubov transformations, and each basis choice gives rise to different number operators. We realized that in curved spacetime it is not possible in general to define frequencies unless there exists a hypersurface-orthogonal timelike Killing vector field, which allows us to put the metric in block-diagonal form with space and time components that do not mix and to factorize the modes so that they have definite frequency. In this way we have a natural way to define vacuum and particles, that is in some sense shared by all observers through the Killing vector field.

After that, we moved to study accelerated observers in flat spacetime using Rindler coordinates, and we found a Killing horizon corresponding to the light cone of an inertial observer. Using this, we computed the surface gravity and the redshift factor, and then proceeded to explore the solutions of the massless two-dimensional Klein-Gordon equation expanded into Rindler modes instead of Minkowski ones. By analytically extending such modes exploiting the regularity of the field on the Killing horizon we obtained the Unruh modes, whose vacuum state is the same as Minkowski modes, and used this to compute the number expectation value of the Rindler accelerated observer in Minkowski vacuum. This turned out to be a Bose-Einstein statistical distribution, which allowed us to define the temperature of the radiated particles.

Finally, we realized that a stationary observer just above the event horizon of a Schwarzschild black hole can be viewed locally, for small enough space and time scales defined by the event horizon itself, as an accelerated observer in flat spacetime, thus allowing us to use the Unruh temperature for the resulting radiation. By redshifting this temperature towards spacial infinity, we obtained the temperature of the Hawking radiation as seen by a stationary observer at spacial infinity. From the temperature, we were able to define the entropy, which we also used for our last more speculative computations.

Hawking radiation is a fascinating phenomenon, and it gives us hope to perform, in the future, some experimental tests on black holes created in the laboratory, or could turn useful for the discovery of primordial black holes that are small enough to emit powerful radiation that is detectable on Earth. The path to a unified model of gravity and quantum mechanics, if exists, is still long, but every step we take could potentially be crucial, and Hawking radiation could be the key yet to be inserted into the lock that may be separating us from a more fundamental understanding of our universe.

Bibliography

[1] Sean Carrol. *Spacetime and geometry: An introduction to general relativity.* Addison Wesley, 2004.

[2] Robert Damian Adamson. "Black holes and thermodynamics". In: 2003. URL: https://api.semanticscholar.org/CorpusID:721740.

[3] Bernard Schutz. *A first course in general relativity.* Cambridge university press, 2009.

[4] Robert M. Wald. *General relativity.* Chicago: Univ. of Chicago Press, 2009.

[5] Roberto Casadio. *Elements of General Relativity.* Lecture notes for the course "Elementi di Teoria della Relatività Generale", Alma Mater Studiorum - Università di Bologna. 2023.

[6] S. Weinberg. "The Quantum Theory of Fields, Vol. 1: Foundations". In: (1995).

[7] Michael Dritschel and James Rovnyak. "Operators on indefinite inner product spaces". In: Dec. 1995, pp. 141–232. ISBN: 9780821804575. DOI: 10.1090/fim/003/03.

[8] John Rydning David Reinsel John Gantz. "The Digitization of the World from Edge to Core". In: 2018. URL: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf.