Elementary Introduction to Quantum Fields in Curved Spacetime

Lecture notes by Sergei Winitzki

Heidelberg, April 18-21, 2006

Contents

Preface

This course is a brief introduction to Quantum Field Theory in Curved Spacetime (QFTCS)—a beautiful and fascinating area of fundamental physics. The application of QFTCS is required in situations when both gravitation and quantum mechanics play a significant role, for instance, in early-universe cosmology and black hole physics. The goal of this course is to introduce some of the most accessible aspects of quantum theory in nontrivial backgrounds and to explain its most unexpected and spectacular manifestations—the Casimir effect (uncharged metal plates attract), the Unruh effect (an accelerated observer will detect particles in vacuum), and Hawking's theoretical discovery of black hole radiation (black holes are not completely black).

This short course was taught in the framework of *Heidelberger Graduiertenkurse* at the Heidelberg University (Germany) in the Spring of 2006. The audience included advanced undergraduates and beginning graduate students. Only a basic familiarity with quantum mechanics, electrodynamics, and general relativity is required. The emphasis is on concepts and intuitive explanations rather than on computational techniques. The relevant calculations are deliberately simplified as much as possible, while retaining all the relevant physics. Some remarks and derivations are typeset in smaller print and can be skipped at first reading.

These lecture notes are freely based on an early draft of the book [MW07] with some changes appropriate for the purposes of the Heidelberg course. The present text may be freely distributed according to the GNU Free Documentation License.[1]

Sergei Winitzki, April 2006

Suggested literature

The following more advanced books may be studied as a continuation of this introductory course:

[BD82] N. D. BIRRELL and P. C. W. DAVIES, *Quantum fields in curved space* (Cambridge University Press, 1982).

[F89] S. A. FULLING, *Aspects of quantum field theory in curved space-time* (Cambridge University Press, 1989).

[GMM94] A. A. GRIB, S. G. MAMAEV, and V. M. MOSTEPANENKO, *Vacuum quantum effects in strong fields* (Friedmann Laboratory Publishing, St. Petersburg, 1994).

The following book contains a significantly more detailed presentation of the material of this course, and much more:

[MW07] V. F. MUKHANOV and S. WINITZKI, *Quantum Effects in Gravity* (to be published by Cambridge University Press, 2007).[2]

1 Quantization of harmonic oscillator

This section serves as a very quick reminder of quantum mechanics of harmonic oscillators. It is assumed that the reader is already familiar with such notions as Schrödinger equation and Heisenberg picture.

1.1 Canonical quantization

A classical harmonic oscillator is described by a coordinate $q(t)$ satisfying

$$\ddot{q} + \omega^2 q = 0, \tag{1}$$

where ω is a real constant. The general solution of this equation can be written as

$$q(t) = ae^{i\omega t} + a^* e^{-i\omega t},$$

where a is a (complex-valued) constant. We may identify the "ground state" of the oscillator as the state without motion, i.e. $q(t) \equiv 0$. This is obviously the lowest-energy state of the oscillator.

The quantum theory of the oscillator is obtained by the standard procedure known as **canonical quantization**. Canonical quantization does not apply directly to an equation of motion. Rather, we first need to describe the system using the Hamiltonian formalism, which means that we must start with the Lagrangian action principle. The classical equation of motion (1) is reformulated as a condition to extremize the action,

$$\int L(q, \dot{q}) dt = \int \left[\frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2 \right] dt,$$

[1] See www.gnu.org/copyleft/fdl.html

[2] An early, incomplete draft is available at www.theorie.physik.uni-muenchen.de/~serge/T6/

where the function $L(q, \dot{q})$ is the **Lagrangian**. (For simplicity, we assumed a unit mass of the oscillator.) Then we define the **canonical momentum**

$$p \equiv \frac{\partial L(q, \dot{q})}{\partial \dot{q}} = \dot{q},$$

and perform a Legendre transformation to find the **Hamiltonian**

$$H(p, q) \equiv [p\dot{q} - L]_{\dot{q} \to p} = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2.$$

The Hamiltonian equations of motion are

$$\dot{q} = p, \quad \dot{p} = -\omega^2 q.$$

Finally, we replace the classical coordinate $q(t)$ and the momentum $p(t)$ by Hermitian operators $\hat{q}(t)$ and $\hat{p}(t)$ satisfying the same equations of motion,

$$\dot{\hat{q}} = \hat{p}, \quad \dot{\hat{p}} = -\omega^2 \hat{q},$$

and additionally postulate the Heisenberg commutation relation

$$[\hat{q}(t), \hat{p}(t)] = i\hbar. \tag{2}$$

The Hamiltonian $H(p, q)$ is also promoted to an operator,

$$\hat{H} \equiv H(\hat{p}, \hat{q}) = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2 \hat{q}^2.$$

All the quantum operators pertaining to the oscillator act in a certain vector space of **quantum states** or **wavefunctions**. (This space must be a Hilbert space; see Appendix A for details.) Vectors from this space are usually denoted using Dirac's "bra-ket" symbols: vectors are denoted by $|a\rangle$, $|b\rangle$, and the corresponding covectors by $\langle a|$, $\langle b|$, etc. Presently, we use the **Heisenberg picture**, in which the operators depend on time but the quantum states are time-independent. This picture is more convenient for developing quantum field theory than the **Schrödinger picture** where operators are time-independent but wavefunctions change with time. Therefore, we shall continue to treat the harmonic oscillator in the Heisenberg picture. We shall not need to use the coordinate or momentum representation of wavefunctions.

1.2 Creation and annihilation operators

The "classical ground state" $\hat{q}(t) \equiv 0$ is impossible in quantum theory because in that case the commutation relation (2) could not be satisfied by any $\hat{p}(t)$. Hence, a quantum oscillator cannot be completely at rest, and its lowest-energy state (called the **ground state** or the **vacuum state**) has a more complicated structure. The standard way of describing quantum oscillators is through the introduction of the creation and annihilation operators.

From now on, we use the units where $\hbar = 1$. The Heisenberg commutation relation becomes

$$[\hat{q}(t), \hat{p}(t)] = i. \tag{3}$$

We now define the **annihilation operator** $\hat{a}^-(t)$ and its Hermitian conjugate, **creation operator** $\hat{a}^+(t)$, by

$$\hat{a}^{\pm}(t) = \sqrt{\frac{\omega}{2}} \left[\hat{q}(t) \mp \frac{i}{\omega}\hat{p}(t) \right].$$

These operators are not Hermitian since $(\hat{a}^-)^\dagger = \hat{a}^+$. The equation of motion for the operator $\hat{a}^-(t)$ is straightforward to derive,

$$\frac{d}{dt}\hat{a}^-(t) = -i\omega\hat{a}^-(t). \tag{4}$$

(The Hermitian conjugate operator $\hat{a}^+(t)$ satisfies the complex conjugate equation.) The solution of Eq. (4) with the initial condition $\hat{a}^-(t)|_{t=0} = \hat{a}_0^-$ can be readily found,

$$\hat{a}^-(t) = \hat{a}_0^- e^{-i\omega t}. \tag{5}$$

It is helpful to introduce time-independent operators $\hat{a}_0^\pm \equiv \hat{a}^\pm$ and to write the time-dependent phase factor $e^{i\omega t}$ explicitly. For instance, we find that the canonical variables $\hat{p}(t)$, $\hat{q}(t)$ are related to $\hat{a}^\pm$ by

$$\hat{p}(t) = \sqrt{\omega}\frac{\hat{a}^- e^{-i\omega t} - \hat{a}^+ e^{i\omega t}}{i\sqrt{2}}, \quad \hat{q}(t) = \frac{\hat{a}^- e^{-i\omega t} + \hat{a}^+ e^{i\omega t}}{\sqrt{2\omega}}. \tag{6}$$

From now on, we shall only use the *time-independent* operators $\hat{a}^\pm$. Using Eqs. (3) and (6), it is easy to show that

$$[\hat{a}^-, \hat{a}^+] = 1.$$

Using the relations (6), the operator $\hat{H}$ can be expressed through the creation and annihilation operators $\hat{a}^\pm$ as

$$\hat{H} = \left(\hat{a}^+\hat{a}^- + \frac{1}{2}\right)\omega. \tag{7}$$

1.3 Particle number eigenstates

Quantum states of the oscillator are described by vectors in an appropriate (infinite-dimensional) Hilbert space. A complete basis in this space is made of vectors $|0\rangle$, $|1\rangle$, ..., which are called the occupation number states or particle number states. The construction of these states is well known, and we briefly review it here for completeness.

It is seen from Eq. (7) that the eigenvalues of $\hat{H}$ are bounded from below by $\frac{1}{2}\omega$. It is then *assumed* that the ground state $|0\rangle$ exists and is unique. Using this assumption and the commutation relations, one can show that the state $|0\rangle$ satisfies

$$\hat{a}^- |0\rangle = 0.$$

(This derivation is standard and we omit it here.) Then we have $\hat{H}|0\rangle = \frac{1}{2}\omega|0\rangle$, which means that the ground state $|0\rangle$ indeed has the lowest possible energy $\frac{1}{2}\omega$.

The **excited states** $|n\rangle$, where $n = 1, 2, ...$, are defined by

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^+)^n |0\rangle. \tag{8}$$

The factors $\sqrt{n!}$ are needed for normalization, namely $\langle m|n\rangle = \delta_{mn}$. It is easy to see that every state $|n\rangle$ is an eigenstate of the Hamiltonian,

$$\hat{H}|n\rangle = \left(n + \frac{1}{2}\right)\omega|n\rangle.$$

In other words, the energy of the oscillator is *quantized* (not continuous) and is measured in discrete "quanta" equal to ω. Therefore, we might interpret the state $|n\rangle$ as describing the presence of n "quanta" of energy or n "particles," each "particle" having the energy ω. (In normal units, the energy of each quantum is $\hbar\omega$, which is the famous Planck formula for the energy quantum.) The operator $\hat{N} \equiv \hat{a}^+\hat{a}^-$ is called the **particle number** operator. Since $\hat{N} |n\rangle = n |n\rangle$, the states $|n\rangle$ are also called **particle number** eigenstates. This terminology is motivated by the applications in quantum field theory (as we shall see below).

To get a feeling of what the ground state $|0\rangle$ looks like, one can compute the expectation values of the coordinate and the momentum in the state $|0\rangle$. For instance, using Eq. (6) we find

$$\langle 0| \hat{q}(t) |0\rangle = 0, \quad \langle 0| \hat{p}(t) |0\rangle = 0,$$
$$\langle 0| \hat{q}^2(t) |0\rangle = \frac{1}{2\omega}, \quad \langle 0| \hat{p}^2(t) |0\rangle = \frac{\omega}{2}.$$

It follows that the ground state $|0\rangle$ of the oscillator exhibits fluctuations of both the coordinate and the momentum around a zero mean value. The typical value of the fluctuation in the coordinate is $\delta q \sim (2\omega)^{-1/2}$.

2 Quantization of scalar field

The quantum theory of fields is built on two essential foundations: the classical theory of fields and the quantum mechanics of harmonic oscillators.

2.1 Classical field

A **classical field** is described by a function of spacetime, $\phi(\mathbf{x}, t)$, characterizing the local strength or intensity of the field. Here $\mathbf{x}$ is a three-dimensional coordinate in space and t is the time (in some reference frame). The function $\phi(\mathbf{x}, t)$ may have real values, complex values, or values in some finite-dimensional vector space. For example, the electromagnetic field is described by the 4-potential $A_\mu(\mathbf{x}, t)$, which is a function whose values are 4-vectors.

The simplest example of a field is a **real scalar** field $\phi(\mathbf{x}, t)$; its values are real numbers. A **free, massive** scalar field satisfies the Klein-Gordon equation[3]

$$\frac{\partial^2 \phi}{\partial t^2} - \sum_{j=1}^{3} \frac{\partial^2 \phi}{\partial x_j^2} + m^2\phi \equiv \ddot{\phi} - \Delta\phi + m^2\phi \equiv \partial_\mu \partial^\mu \phi + m^2\phi = 0. \tag{9}$$

The parameter m is the **mass** of the field. The solution $\phi(\mathbf{x}, t) \equiv 0$ is the classical vacuum state ("no field").

To simplify the equations of motion, it is convenient to use the spatial Fourier decomposition,

$$\phi(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \phi_{\mathbf{k}}(t), \tag{10}$$

where we integrate over all three-dimensional vectors $\mathbf{k}$. After the Fourier decomposition, the partial differential equation (9) is replaced by infinitely many ordinary differential equations, with one equation for each $\mathbf{k}$:

$$\ddot{\phi}_{\mathbf{k}} + \left(k^2 + m^2\right) \phi_{\mathbf{k}} = \ddot{\phi}_{\mathbf{k}} + \omega_k^2 \phi_{\mathbf{k}} = 0, \quad \omega_k \equiv \sqrt{|\mathbf{k}|^2 + m^2}.$$

In other words, each function $\phi_{\mathbf{k}}(t)$ satisfies the harmonic oscillator equation with the frequency ω_k. The complex-valued functions $\phi_{\mathbf{k}}(t)$ are called the **modes** of the

[3] To simplify the formulas, we shall (almost always) use the units in which $\hbar = c = 1$.

field ϕ (abbreviated from "Fourier modes"). Note that the modes $\phi_{\mathbf{k}}(t)$ of a real field $\phi(\mathbf{x},t)$ satisfy the relation $(\phi_{\mathbf{k}})^* = \phi_{-\mathbf{k}}$.

The equation of motion (9) can be found by extremizing the action

$$S[\phi] = \frac{1}{2}\int d^4x \left[\eta^{\mu\nu}\left(\partial_\mu\phi\right)\left(\partial_\nu\phi\right) - m^2\phi^2\right]$$
$$\equiv \frac{1}{2}\int d^3\mathbf{x}\, dt \left[\dot\phi^2 - \left(\nabla\phi\right)^2 - m^2\phi^2\right], \tag{11}$$

where $\eta^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$ is the Minkowski metric (in this chapter we consider only the flat spacetime) and the Greek indices label four-dimensional coordinates: $x^0 \equiv t$ and $(x^1,x^2,x^3) \equiv \mathbf{x}$. Using Eq. (10), one can also express the action (11) directly through the (complex-valued) modes $\phi_{\mathbf{k}}$,

$$S = \frac{1}{2}\int dt\, d^3\mathbf{k}\left[\dot\phi_{\mathbf{k}}\dot\phi^*_{\mathbf{k}} - \omega_k^2\phi_{\mathbf{k}}\phi^*_{\mathbf{k}}\right]. \tag{12}$$

2.2 Quantization of scalar field

The action (12) is analogous to that of a collection of infinitely many harmonic oscillators. Therefore, we may quantize each mode $\phi_{\mathbf{k}}(t)$ as a separate (complex-valued) harmonic oscillator.

Let us begin with the Hamiltonian description of the field $\phi(\mathbf{x},t)$. The action (11) must be thought of as an integral of the Lagrangian over *time* (but not over space), $S[\phi] = \int L[\phi]\, dt$, so the Lagrangian $L[\phi]$ is

$$L[\phi] = \int \mathcal{L} d^3\mathbf{x}; \quad \mathcal{L} \equiv \frac{1}{2}\eta^{\mu\nu}\left(\partial_\mu\phi\right)\left(\partial_\nu\phi\right) - \frac{1}{2}m^2\phi^2,$$

where $\mathcal{L}$ is the **Lagrangian density**. To define the canonical momenta and the Hamiltonian, one must use the Lagrangian $L[\phi]$ rather than the Lagrangian density $\mathcal{L}$. Hence, the momenta $\pi(\mathbf{x},t)$ are computed as the functional derivatives

$$\pi(\mathbf{x},t) \equiv \frac{\delta L[\phi]}{\delta\dot\phi(\mathbf{x},t)} = \dot\phi(\mathbf{x},t),$$

and then the classical Hamiltonian is

$$H = \int \pi(\mathbf{x},t)\,\dot\phi(\mathbf{x},t)\, d^3\mathbf{x} - L = \frac{1}{2}\int d^3\mathbf{x}\left[\pi^2 + (\nabla\phi)^2 + m^2\phi^2\right]. \tag{13}$$

To quantize the field, we introduce the operators $\hat\phi(\mathbf{x},t)$ and $\hat\pi(\mathbf{x},t)$ with the standard commutation relations

$$\left[\hat\phi(\mathbf{x},t), \hat\pi(\mathbf{y},t)\right] = i\delta(\mathbf{x}-\mathbf{y}); \quad \left[\hat\phi(\mathbf{x},t), \hat\phi(\mathbf{y},t)\right] = \left[\hat\pi(\mathbf{x},t), \hat\pi(\mathbf{y},t)\right] = 0. \tag{14}$$

The modes $\phi_{\mathbf{k}}(t)$ also become operators $\hat\phi_{\mathbf{k}}(t)$. The commutation relation for the modes can be derived from Eq. (14) by performing Fourier transforms in $\mathbf{x}$ and $\mathbf{y}$. After some algebra, we find

$$\left[\hat\phi_{\mathbf{k}}(t), \hat\pi_{\mathbf{k}'}(t)\right] = i\delta\left(\mathbf{k}+\mathbf{k}'\right). \tag{15}$$

Note the *plus* sign in $\delta(\mathbf{k}_1+\mathbf{k}_2)$: this is related to the fact that the variable which is conjugate to $\hat\phi_{\mathbf{k}}$ is not $\hat\pi_{\mathbf{k}}$ but $\hat\pi_{-\mathbf{k}} = \hat\pi^\dagger_{\mathbf{k}}$.

Remark: complex oscillators. The modes $\phi_{\mathbf{k}}(t)$ are complex variables; each $\phi_{\mathbf{k}}$ may be thought of as a pair of real-valued oscillators, $\phi_{\mathbf{k}} = \phi_{\mathbf{k}}^{(1)} + i\phi_{\mathbf{k}}^{(2)}$. Accordingly, the operators $\hat{\phi}_{\mathbf{k}}$ are *not* Hermitian and $(\hat{\phi}_{\mathbf{k}})^{\dagger} = \hat{\phi}_{-\mathbf{k}}$. In principle, one could rewrite the theory in terms of Hermitian variables such as $\phi_{\mathbf{k}}^{(1,2)}$ and $\pi_{\mathbf{k}}^{(1,2)}$ with standard commutation relations,

$$\left[\phi_{\mathbf{k}}^{(1)}, \pi_{\mathbf{k}'}^{(1)}\right] = i\delta(\mathbf{k} - \mathbf{k}'), \quad \left[\phi_{\mathbf{k}}^{(2)}, \pi_{\mathbf{k}'}^{(2)}\right] = i\delta(\mathbf{k} - \mathbf{k}')$$

but it is technically more convenient to keep the complex-valued modes $\phi_{\mathbf{k}}$. The non-standard form of the commutation relation (15) is a small price to pay.

For each mode $\phi_{\mathbf{k}}$, we proceed with the quantization as in Sec. 1.2. We first introduce the time-dependent creation and annihilation operators:

$$\hat{a}_{\mathbf{k}}^{-}(t) \equiv \sqrt{\frac{\omega_k}{2}}\left(\hat{\phi}_{\mathbf{k}} + \frac{i\hat{\pi}_{\mathbf{k}}}{\omega_k}\right); \quad \hat{a}_{\mathbf{k}}^{+}(t) \equiv \sqrt{\frac{\omega_k}{2}}\left(\hat{\phi}_{-\mathbf{k}} - \frac{i\hat{\pi}_{-\mathbf{k}}}{\omega_k}\right).$$

Note that $(\hat{a}_{\mathbf{k}}^{-})^{\dagger} = \hat{a}_{\mathbf{k}}^{+}$. The equations of motion for the operators $\hat{a}_{\mathbf{k}}^{\pm}(t)$,

$$\frac{d}{dt}\hat{a}_{\mathbf{k}}^{\pm}(t) = \pm i\omega_k \hat{a}_{\mathbf{k}}^{\pm}(t),$$

have the general solution $\hat{a}_{\mathbf{k}}^{\pm}(t) = {}^{(0)}\hat{a}_{\mathbf{k}}^{\pm} e^{\pm i\omega_k t}$, where the time-independent operators ${}^{(0)}\hat{a}_{\mathbf{k}}^{\pm}$ satisfy the relations (note the signs of $\mathbf{k}$ and $\mathbf{k}'$)

$$\left[\hat{a}_{\mathbf{k}}^{-},\, \hat{a}_{\mathbf{k}'}^{+}\right] = \delta\left(\mathbf{k} - \mathbf{k}'\right); \quad \left[\hat{a}_{\mathbf{k}}^{-},\, \hat{a}_{\mathbf{k}'}^{-}\right] = \left[\hat{a}_{\mathbf{k}}^{+},\, \hat{a}_{\mathbf{k}'}^{+}\right] = 0. \tag{16}$$

In Eq. (16) we omitted the superscript ${}^{(0)}$ for brevity; below we shall always use the *time-independent* creation and annihilation operators and denote them by $\hat{a}_{\mathbf{k}}^{\pm}$.

The Hilbert space of field states is built in the standard fashion. We postulate the existence of the **vacuum state** $|0\rangle$ such that $\hat{a}_{\mathbf{k}}^{-}|0\rangle = 0$ for all $\mathbf{k}$. The state with particle numbers n_s in each mode with momentum $\mathbf{k}_s$ (where $s = 1, 2, ...$ is an index that enumerates the excited modes) is defined by

$$|n_1, n_2, ...\rangle = \left[\prod_s \frac{\left(\hat{a}_{\mathbf{k}_s}^{+}\right)^{n_s}}{\sqrt{n_s!}}\right]|0\rangle. \tag{17}$$

We write $|0\rangle$ instead of $|0, 0, ...\rangle$ for brevity. The Hilbert space of quantum states is spanned by the vectors $|n_1, n_2, ...\rangle$ with all possible choices of the numbers n_s. This space is called the **Fock space**.

The quantum Hamiltonian of the free scalar field can be written as

$$\hat{H} = \frac{1}{2}\int d^3\mathbf{k}\left[\hat{\pi}_{\mathbf{k}}\hat{\pi}_{-\mathbf{k}} + \omega_k^2\hat{\phi}_{\mathbf{k}}\hat{\phi}_{-\mathbf{k}}\right],$$

and expressed through the creation and annihilation operators as

$$\hat{H} = \int d^3\mathbf{k}\frac{\omega_k}{2}\left[\hat{a}_{\mathbf{k}}^{-}\hat{a}_{\mathbf{k}}^{+} + \hat{a}_{\mathbf{k}}^{+}\hat{a}_{\mathbf{k}}^{-}\right] = \int d^3\mathbf{k}\frac{\omega_k}{2}\left[2\hat{a}_{\mathbf{k}}^{+}\hat{a}_{\mathbf{k}}^{-} + \delta^{(3)}(0)\right]. \tag{18}$$

Derivation of Eq. (18)

We use the relations

$$\hat{\phi}_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_k}}\left(\hat{a}_{\mathbf{k}}^{-}e^{-i\omega_k t} + \hat{a}_{-\mathbf{k}}^{+}e^{i\omega_k t}\right), \ \hat{\pi}_{\mathbf{k}} = i\sqrt{\frac{\omega_k}{2}}\left(\hat{a}_{-\mathbf{k}}^{+}e^{i\omega_k t} - \hat{a}_{\mathbf{k}}^{-}e^{-i\omega_k t}\right).$$

(Here $\hat{a}_{\mathbf{k}}^{\pm}$ are time-independent operators.) Then we find

$$\frac{1}{2}\left(\hat{\pi}_{\mathbf{k}}\hat{\pi}_{-\mathbf{k}} + \omega_k^2\hat{\phi}_{\mathbf{k}}\hat{\phi}_{-\mathbf{k}}\right) = \frac{\omega_k}{2}\left(\hat{a}_{\mathbf{k}}^{-}\hat{a}_{\mathbf{k}}^{+} + \hat{a}_{-\mathbf{k}}^{+}\hat{a}_{-\mathbf{k}}^{-}\right).$$

When we integrate over all $\mathbf{k}$, the terms with $-\mathbf{k}$ give the same result as the terms with $\mathbf{k}$. Therefore $\hat{H} = \frac{1}{2} \int d^3\mathbf{k} \left(\hat{a}^+_\mathbf{k} \hat{a}^-_\mathbf{k} + \hat{a}^-_\mathbf{k} \hat{a}^+_\mathbf{k} \right) \omega_k$.

Thus we have quantized the scalar field $\phi(\mathbf{x}, t)$ in the Heisenberg picture. Quantum observables such as $\hat{\phi}(\mathbf{x}, t)$ and $\hat{H}$ are now represented by linear operators in the Fock space, and the quantum states of the field ϕ are interpreted in terms of particles. Namely, the state vector (17) is interpreted as a state with n_s particles having momentum $\mathbf{k}_s$ (where $s = 1, 2, ...$). This particle interpretation is consistent with the relativistic expression for the energy of a particle, $E = \sqrt{p^2 + m^2}$, if we identify the 3-momentum $\mathbf{p}$ with the wavenumber $\mathbf{k}$ and the energy E with ω_k.

2.3 Mode expansions

We now give a brief introduction to mode expansions, which offer a shorter and computationally convenient way to quantize fields. A more detailed treatment is given in Sec 5.

The quantum mode $\hat{\phi}_\mathbf{k}(t)$ can be expressed through the creation and annihilation operators,

$$\hat{\phi}_\mathbf{k}(t) = \frac{1}{\sqrt{2\omega_k}} \left(\hat{a}^-_\mathbf{k} e^{-i\omega_k t} + \hat{a}^+_{-\mathbf{k}} e^{i\omega_k t} \right).$$

Substituting this into Eq. (10), we obtain the following expansion of the field operator $\hat{\phi}(\mathbf{x}, t)$,

$$\hat{\phi}(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_k}} \left[\hat{a}^-_\mathbf{k} e^{-i\omega_k t + i\mathbf{k}\cdot\mathbf{x}} + \hat{a}^+_{-\mathbf{k}} e^{i\omega_k t + i\mathbf{k}\cdot\mathbf{x}} \right],$$

which we then rewrite by changing $\mathbf{k} \to -\mathbf{k}$ in the second term to make the integrand manifestly Hermitian:

$$\hat{\phi}(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_k}} \left[\hat{a}^-_\mathbf{k} e^{-i\omega_k t + i\mathbf{k}\cdot\mathbf{x}} + \hat{a}^+_\mathbf{k} e^{i\omega_k t - i\mathbf{k}\cdot\mathbf{x}} \right]. \qquad (19)$$

This expression is called the **mode expansion** of the quantum field $\hat{\phi}$.

It is easy to see that the Klein-Gordon equation (9) is identically satisfied by the ansatz (19) with arbitrary time-independent operators $\hat{a}^\pm_\mathbf{k}$. In fact, Eq. (19) is a *general solution* of Eq. (9) with operator-valued "integration constants" $\hat{a}^\pm_\mathbf{k}$. On the other hand, one can verify that the commutation relations (15) between $\hat{\phi}_\mathbf{k}$ and $\hat{\pi}_\mathbf{k}$ are equivalent to Eq. (16). Therefore, we may view the mode expansion (19) as a convenient shortcut to quantizing the field $\phi(\mathbf{x}, t)$. One simply needs to postulate the commutation relations (16) and the mode expansion (19), and then the operators $\hat{\phi}_\mathbf{k}$ and $\hat{\pi}_\mathbf{k}$ do not need to be introduced explicitly. The Fock space of quantum states is constructed directly through the operators $\hat{a}^\pm_\mathbf{k}$ and interpreted as above.

3 Casimir effect

3.1 Zero-point energy

The **zero-point energy** is the energy of the vacuum state. We saw in Sec. 2.2 that a quantum field is equivalent to a collection of infinitely many harmonic oscillators. If the field ϕ is in the vacuum state, each oscillator $\phi_\mathbf{k}$ is in the ground state and has the energy $\frac{1}{2}\omega_k$. Hence, the total zero-point energy of the field is the sum of $\frac{1}{2}\omega_k$ over all wavenumbers $\mathbf{k}$. This sum may be approximated by an integral in the following way: If one quantizes the field in a box of large but finite volume V, one will obtain the result that the zero-point energy *density* is equal to

$$\frac{E_0}{V} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{2}\omega_k.$$

(A detailed computation can be found in the book [MW07].) Since ω_k grows with k, it is clear that the integral diverges. Taken at face value, this would indicate an infinite energy density of the vacuum state. Since any energy density leads to gravitational effects, the presence of a nonzero energy density in the vacuum state contradicts the experimental observation that empty space does not generate any gravitational force. The standard way to avoid this problem is to subtract this infinite quantity from the energy of the system ("renormalization" of zero-point energy). In other words, the ground state energy $\frac{1}{2}\omega_k$ is subtracted from the Hamiltonian of each oscillator $\phi_{\mathbf{k}}$. A justification for this subtraction is that the ground state energy cannot be extracted from an oscillator, and that only the change in the oscillator's energy can be observed.

3.2 Casimir effect

The Casimir effect is an experimentally verified prediction of quantum field theory. It is manifested by a force of attraction between two *uncharged* conducting plates in a vacuum. This force cannot be explained except by considering the zero-point energy of the quantized electromagnetic field. The presence of the conducting plates makes the electromagnetic field vanish on the surfaces of the plates. This boundary condition changes the structure of vacuum fluctuations of the field, which would normally be nonzero on the plates. This change causes a finite shift ΔE of the zero-point energy, compared with the zero-point energy in empty space without the plates. The energy shift $\Delta E = \Delta E(L)$ depends on the distance L between the plates, and it turns out that ΔE grows with L. As a result, it is energetically favorable for the plates to move closer together, which is manifested as the **Casimir force** of attraction between the plates,

$$F(L) = -\frac{d}{dL}\left[\Delta E(L)\right].$$

This theoretically predicted force has been confirmed by several experiments.[4]

3.3 Zero-point energy between plates

A realistic description of the Casimir effect involves quantization of the electromagnetic field in the presence of conductors having certain dielectric properties; thermal fluctuations must also be taken into account. We shall drastically simplify the calculations by considering a *massless scalar* field $\phi(t,x)$ in the flat 1+1-dimensional spacetime. To simulate the presence of the plates, we impose the following boundary conditions:

$$\phi(t,x)|_{x=0} = \phi(t,x)|_{x=L} = 0. \tag{20}$$

The equation of motion for the classical field is $\partial_t^2\phi - \partial_x^2\phi = 0$, and the general solution for the chosen boundary conditions is of the form

$$\phi(t,x) = \sum_{n=1}^{\infty} \left(A_n e^{-i\omega_n t} + B_n e^{i\omega_n t}\right) \sin \omega_n x, \quad \omega_n \equiv \frac{n\pi}{L}.$$

[4] For example, a recent measurement of the Casimir force to 1% precision is described in: U. Mohideen and A. Roy, Phys. Rev. Lett. **81** (1998), p. 4549.

We shall now use this solution as a motivation for finding the appropriate mode expansion for the field ϕ.

To quantize the field $\phi(t,x)$ in flat space, one would normally use the mode expansion

$$\hat{\phi}(t,x)=\int\frac{dk}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\omega_k}}\left[\hat{a}_k^- e^{-i\omega_k t+ikx}+\hat{a}_k^+ e^{i\omega_k t-ikx}\right].$$

However, in the present case the only allowed modes are those satisfying Eq. (20), so the above mode expansion cannot be used. To expand the field $\hat{\phi}(t,x)$ into the allowed modes, we use the orthogonal system of functions

$$g_n(x)=\sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}$$

which vanish at $x=0$ and $x=L$. These functions satisfy the orthogonality relation

$$\int_0^L g_m(x)g_n(x)dx=\delta_{mn}.$$

(The coefficient $\sqrt{2/L}$ is necessary for the correct normalization.) An arbitrary function $F(x)$ that vanishes at $x=0$ and $x=L$ can be expanded through the functions $g_n(x)$ as

$$F(x)=\sum_{n=1}^{\infty}F_n g_n(x),$$

where the coefficients F_n are found as

$$F_n=\int_0^L F(x)g_n(x)dx.$$

Performing this decomposition for the field operator $\hat{\phi}(t,x)$, we find

$$\hat{\phi}(t,x)=\sum_{n=1}^{\infty}\hat{\phi}_n(t)g_n(x),$$

where $\hat{\phi}_n(t)$ is the n-th mode of the field. The mode $\hat{\phi}_n(t)$ satisfies the oscillator equation

$$\ddot{\hat{\phi}}_n+\left(\frac{\pi n}{L}\right)^2\hat{\phi}_n\equiv\ddot{\hat{\phi}}_n+\omega_n^2\hat{\phi}_n=0, \tag{21}$$

whose general solution is

$$\hat{\phi}_n(t)=\hat{A}e^{-i\omega_n t}+\hat{B}e^{i\omega_n t},$$

where $\hat{A}$, $\hat{B}$ are operator-valued integration constants. After computing the correct normalization of these constants (see below), we obtain the the mode expansion for $\hat{\phi}$ as

$$\hat{\phi}(t,x)=\sqrt{\frac{2}{L}}\sum_{n=1}^{\infty}\frac{\sin\omega_n x}{\sqrt{2\omega_n}}\left[\hat{a}_n^- e^{-i\omega_n t}+\hat{a}_n^+ e^{i\omega_n t}\right]. \tag{22}$$

We need to compute the energy of the field only between the plates, $0<x<L$. After some calculations (see below), one can express the zero-point energy *per unit length* as

$$\varepsilon_0\equiv\frac{1}{L}\left\langle 0\right|\hat{H}\left|0\right\rangle=\frac{1}{2L}\sum_k\omega_k=\frac{\pi}{2L^2}\sum_{n=1}^{\infty}n. \tag{23}$$

Derivation of Eqs. (22) and (23)

We use the following elementary identities which hold for integer m, n:

$$\int_0^L dx \, \sin\frac{m\pi x}{L} \sin\frac{n\pi x}{L} = \int_0^L dx \, \cos\frac{m\pi x}{L} \cos\frac{n\pi x}{L} = \frac{L}{2}\delta_{mn}. \tag{24}$$

First, let us show that the normalization factor $\sqrt{2/L}$ in the mode expansion (22) yields the standard commutation relations $\left[\hat{a}_m^-, \hat{a}_n^+\right] = \delta_{mn}$. We integrate the mode expansion over x and use the identity (24) to get

$$\int_0^L dx \, \hat{\phi}(x,t) \sin\omega_n x = \frac{1}{2}\sqrt{\frac{L}{\omega_n}} \left[\hat{a}_n^- e^{-i\omega_n t} + \hat{a}_n^+ e^{i\omega_n t}\right].$$

Then we differentiate this with respect to t and obtain

$$\int_0^L dx' \, \hat{\pi}(y,t) \sin\omega_n x' = \frac{i}{2}\sqrt{L\omega_n} \left[-\hat{a}_n^- e^{-i\omega_n t} + \hat{a}_n^+ e^{i\omega_n t}\right].$$

Now we can evaluate the commutator

$$\left[\int_0^L dx \, \hat{\phi}(x,t) \sin\omega_n x, \int_0^L dy \, \frac{d}{dt}\hat{\phi}(x',t) \sin\omega_{n'} x'\right] = i\frac{L}{2}\left[\hat{a}_n^-, \hat{a}_{n'}^+\right]$$
$$= \int_0^L dx \int_0^L dx' \, \sin\frac{n\pi x}{L} \sin\frac{n'\pi x'}{L} i\delta(x - x') = i\frac{L}{2}\delta_{nn'}.$$

In the second line we used $\left[\hat{\phi}(x,t), \hat{\pi}(x',t)\right] = i\delta(x - x')$. Therefore the standard commutation relations hold for $\hat{a}_n^\pm$.

The Hamiltonian for the field (restricted to the region between the plates) is

$$\hat{H} = \frac{1}{2}\int_0^L dx \left[\left(\frac{\partial\hat{\phi}(x,t)}{\partial t}\right)^2 + \left(\frac{\partial\hat{\phi}(x,t)}{\partial x}\right)^2\right].$$

The expression $\langle 0|\, \hat{H}\, |0\rangle$ is evaluated using the mode expansion above and the relations

$$\langle 0|\, \hat{a}_m^- \hat{a}_n^+ \,|0\rangle = \delta_{mn}, \quad \langle 0|\, \hat{a}_m^+ \hat{a}_n^+ \,|0\rangle = \langle 0|\, \hat{a}_m^- \hat{a}_n^- \,|0\rangle = \langle 0|\, \hat{a}_m^+ \hat{a}_n^- \,|0\rangle = 0.$$

The first term in the Hamiltonian yields

$$\langle 0|\, \frac{1}{2}\int_0^L dx \left(\frac{\partial\hat{\phi}(x,t)}{\partial t}\right)^2 |0\rangle$$
$$= \langle 0|\, \frac{1}{2}\int_0^L dx \left[\sqrt{\frac{2}{L}}\sum_{n=1}^{\infty}\frac{\sin\omega_n x}{\sqrt{2\omega_n}} i\omega_n \left(-\hat{a}_n^- e^{-i\omega_n t} + \hat{a}_n^+ e^{i\omega_n t}\right)\right]^2 |0\rangle$$
$$= \frac{1}{L}\int_0^L dx \sum_{n=1}^{\infty}\frac{(\sin\omega_n x)^2}{2\omega_n}\omega_n^2 = \frac{1}{4}\sum_n \omega_n.$$

The second term gives the same result, and we find

$$\langle 0|\, \hat{H}\, |0\rangle = \frac{1}{2}\sum_{n=1}^{\infty}\omega_n.$$

Therefore, the energy density (the energy per unit length) is given by Eq. (23).

3.4 Regularization and renormalization

The zero-point energy density ε_0 is divergent. However, in the presence of the plates the energy density diverges in a different way than in free space because $\varepsilon_0 = \varepsilon_0(L)$ depends on the distance L between the plates. The zero-point energy density in free space can be thought of as the limit of $\varepsilon_0(L)$ at $L \to \infty$,

$$\varepsilon_0^{(\text{free})} = \lim_{L \to \infty} \varepsilon_0\left(L\right).$$

When the zero-point energy is renormalized in free space, the infinite contribution $\varepsilon_0^{(\text{free})}$ is subtracted. Thus we are motivated to subtract $\varepsilon_0^{(\text{free})}$ from the energy density $\varepsilon_0(L)$ and to expect to find a *finite* difference $\Delta\varepsilon$ between these formally infinite quantities,

$$\Delta\varepsilon\left(L\right) = \varepsilon_0\left(L\right) - \varepsilon_0^{(\text{free})} = \varepsilon_0\left(L\right) - \lim_{L \to \infty} \varepsilon_0\left(L\right). \tag{25}$$

In the remainder of the chapter we calculate this energy difference $\Delta\varepsilon(L)$.

Taken at face value, Eq. (25) is meaningless because the difference between two infinite quantities is undefined. The standard way to deduce reasonable answers from infinities is a regularization followed by a renormalization. A **regularization** means introducing an extra parameter into the theory to make the divergent quantity finite unless that parameter is set to (say) zero. Such regularization parameters or **cutoffs** can be chosen in many ways. After the regularization, one derives an asymptotic form of the divergent quantity at small values of the cutoff. This asymptotic may contain divergent powers and logarithms of the cutoff as well as finite terms. **Renormalization** means removing the divergent terms and leaving only the finite terms in the expression. (Of course, a suitable justification must be provided for subtracting the divergent terms.) After renormalization, the cutoff is set to zero and the remaining terms yield the final result. If the cutoff function is chosen incorrectly, the renormalization procedure will not succeed. It is usually possible to motivate the correct choice of the cutoff by physical considerations.

We shall now apply this procedure to Eq. (25). As a first step, a cutoff must be introduced into the divergent expression (23). One possibility is to replace ε_0 by the regularized quantity

$$\varepsilon_0\left(L;\alpha\right) = \frac{\pi}{2L^2}\sum_{n=1}^{\infty} n \exp\left[-\frac{n\alpha}{L}\right], \tag{26}$$

where α is the cutoff parameter. The regularized series converges for $\alpha > 0$, while the original divergent expression is recovered in the limit $\alpha \to 0$.

> **Remark: choosing the cutoff function.** We regularize the series by the factor $\exp(-n\alpha/L)$ and not by $\exp(-n\alpha)$ or $\exp(-nL\alpha)$. A motivation is that the physically significant quantity is $\omega_n = \pi n/L$, therefore the cutoff factor should be a function of ω_n. Also, renormalization will fail if the regularization is chosen incorrectly.

Now we need to evaluate the regularized quantity (26) and to analyze its asymptotic behavior at $\alpha \to 0$. A straightforward computation gives

$$\varepsilon_0\left(L;\alpha\right) = -\frac{\pi}{2L}\frac{\partial}{\partial\alpha}\sum_{n=1}^{\infty} \exp\left[-\frac{n\alpha}{L}\right] = \frac{\pi}{2L^2}\frac{\exp\left(-\frac{\alpha}{L}\right)}{\left[1-\exp\left(-\frac{\alpha}{L}\right)\right]^2}.$$

At $\alpha \to 0$ this expression can be expanded in a Laurent series,

$$\varepsilon_0\left(L;\alpha\right) = \frac{\pi}{8L^2}\frac{1}{\sinh^2\frac{\alpha}{2L}} = \frac{\pi}{2\alpha^2} - \frac{\pi}{24L^2} + O\left(\alpha^2\right). \tag{27}$$

The series (27) contains the singular term $\frac{\pi}{2}\alpha^{-2}$, a finite term, and further terms that vanish as $\alpha \to 0$. The crucial fact is that the singular term in Eq. (27) does not depend on L. (This would not have happened if we chose the cutoff e.g. as $e^{-n\alpha}$.) The limit $L \to \infty$ in Eq. (25) is taken *before* the limit $\alpha \to 0$, so the divergent term $\frac{\pi}{2}\alpha^{-2}$ cancels and the renormalized value of $\Delta\varepsilon$ is finite,

$$\Delta\varepsilon_{ren}(L) = \lim_{\alpha\to 0}\left[\varepsilon_0\left(L;\alpha\right) - \lim_{L\to\infty}\varepsilon_0\left(L;\alpha\right)\right] = -\frac{\pi}{24L^2}. \tag{28}$$

The formula (28) is the main result of this chapter; the zero-point energy density is nonzero in the presence of plates at $x = 0$ and $x = L$. The Casimir force between the plates is

$$F = -\frac{d}{dL}\Delta E = -\frac{d}{dL}\left(L\Delta\varepsilon_{ren}\right) = -\frac{\pi}{24L^2}.$$

Since the force is negative, the plates are pulled toward each other.

> **Remark: negative energy.** Note that the zero-point energy density (28) is *negative*. Quantum field theory generally admits quantum states with a negative expectation value of energy.

4 Oscillator with varying frequency

A gravitational background influences quantum fields in such a way that the frequencies ω_k of the modes become time-dependent, $\omega_k(t)$. We shall examine this situation in detail in chapter 5. For now, let us consider a single harmonic oscillator with a time-dependent frequency $\omega(t)$.

4.1 Quantization

In the classical theory, the coordinate $q(t)$ satisfies

$$\ddot{q} + \omega^2(t)q = 0. \tag{29}$$

An important example of a function $\omega(t)$ is shown Fig. 1. The frequency is approximately constant except for a finite time interval, for instance $\omega \equiv \omega_0$ for $t \le t_0$ and $\omega \equiv \omega_1$ for $t \ge t_1$. It is usually impossible to find an exact solution of Eq. (29) in such cases (of course, an approximate solution can be found numerically). However, the solutions in the regimes $t \le t_0$ and $t \ge t_1$ are easy to obtain:

$$q(t) = Ae^{i\omega_0 t} + Be^{-i\omega_0 t}, \quad t \le t_0;$$
$$q(t) = Ae^{i\omega_1 t} + Be^{-i\omega_1 t}, \quad t \ge t_1.$$

We shall be interested only in describing the behavior of the oscillator in these two regimes[5] which we call the "**in**" and "**out**" regimes.

The classical equation of motion (29) can be derived from the Lagrangian

$$L\left(t,q,\dot{q}\right) = \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega(t)^2q^2.$$

The corresponding canonical momentum is $p = \dot{q}$, and the Hamiltonian is

$$H(p,q) = \frac{p^2}{2} + \omega^2(t)\frac{q^2}{2}, \tag{30}$$

[5] In the physics literature, the word **regime** stands for "an interval of values for a variable." It should be clear from the context which interval for which variable is implied.

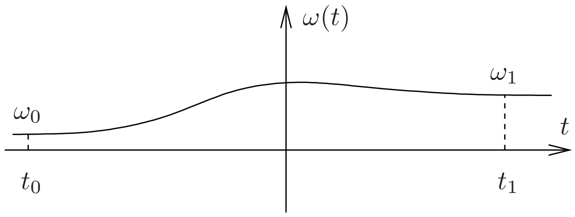

Figure 1: A frequency function $\omega(t)$ with "in" and "out" regimes (at $t \leq t_0$ and $t \geq t_1$).

which depends explicitly on the time t. Therefore, we do not expect that energy is conserved in this system. (There is an external agent that drives $\omega(t)$ and may exchange energy with the oscillator.)

A time-dependent oscillator can be quantized using the technique of creation and annihilation operators. By analogy with Eq. (6), we try the ansatz

$$\hat{q}(t) = \frac{1}{\sqrt{2}}\left(v(t)\hat{a}^+ + v^*(t)\hat{a}^-\right), \quad \hat{p}(t) = \frac{1}{\sqrt{2}}\left(\dot{v}(t)\hat{a}^+ + \dot{v}^*(t)\hat{a}^-\right), \tag{31}$$

where $v(t)$ is a complex-valued function that replaces $e^{i\omega t}$, while the operators $\hat{a}^{\pm}$ are time-independent. The present task is to choose the function $v(t)$ and the operators $\hat{a}^{\pm}$ in an appropriate way. We call $v(t)$ the **mode function** because we shall later apply the same decomposition to modes of a quantum field.

Since $\hat{q}(t)$ must be a solution of Eq. (29), we find that $v(t)$ must satisfy the same equation,

$$\ddot{v} + \omega^2(t)v = 0. \tag{32}$$

Furthermore, the canonical commutation relation $[\hat{q}(t), \hat{p}(t)] = i$ entails

$$\left[\hat{a}^-, \hat{a}^+\right] = \frac{2i}{\dot{v}v^* - \dot{v}^*v}.$$

Note that the expression

$$\dot{v}v^* - \dot{v}^*v \equiv W[v, v^*]$$

is the Wronskian of the solutions $v(t)$ and $v^*(t)$, and it is well known that $W =$ const. We may therefore normalize the mode function $v(t)$ such that

$$W[v, v^*] = \dot{v}v^* - \dot{v}^*v = 2i, \tag{33}$$

which will yield the standard commutation relations for $\hat{a}^{\pm}$,

$$\left[\hat{a}^-, \hat{a}^+\right] = 1.$$

We can then postulate the existence of the ground state $|0\rangle$ such that $\hat{a}^- |0\rangle = 0$. Excited states $|n\rangle$ ($n = 1, 2, ...$) are defined in the standard way by Eq. (8).

With the normalization (33), the creation and annihilation operators are expressed through the canonical variables as

$$\hat{a}^- \equiv \frac{\dot{v}(t)\hat{q}(t) - v(t)\hat{p}(t)}{i\sqrt{2}}, \quad \hat{a}^+ \equiv -\frac{\dot{v}^*(t)\hat{q}(t) - v^*(t)\hat{p}(t)}{i\sqrt{2}}. \tag{34}$$

(Note that the l.h.s. of Eq. (34) are time-independent because the corresponding r.h.s. are Wronskians.) In this way, a choice of the mode function $v(t)$ defines the operators $\hat{a}^{\pm}$ and the states $|0\rangle$, $|1\rangle$, ...

It is clear that different choices of $v(t)$ will in general define different operators $\hat{a}^{\pm}$ and different states $|0\rangle$, $|1\rangle$, ... It is not clear, *a priori*, which choice of $v(t)$ corresponds to the "correct" ground state of the oscillator. The choice of $v(t)$ will be studied in the next section.

Properties of mode functions

Here is a summary of some elementary properties of a time-dependent oscillator equation

$$\ddot{x} + \omega^2(t)x = 0. \tag{35}$$

This equation has a two-dimensional space of solutions. Any two linearly independent solutions $x_1(t)$ and $x_2(t)$ are a basis in that space. The expression

$$W[x_1, x_2] \equiv \dot{x}_1 x_2 - x_1 \dot{x}_2$$

is called the **Wronskian** of the two functions $x_1(t)$ and $x_2(t)$. It is easy to see that the Wronskian $W[x_1, x_2]$ is time-independent if $x_{1,2}(t)$ satisfy Eq. (35). Moreover, $W[x_1, x_2] \neq 0$ if and only if $x_1(t)$ and $x_2(t)$ are two linearly independent solutions.

If $\{x_1(t), x_2(t)\}$ is a basis of solutions, it is convenient to define the complex function $v(t) \equiv x_1(t) + ix_2(t)$. Then $v(t)$ and $v^*(t)$ are linearly independent and form a basis in the space of *complex* solutions of Eq. (35). It is easy to check that

$$\mathrm{Im}(\dot{v}v^*) = \frac{\dot{v}v^* - \dot{v}^*v}{2i} = \frac{1}{2i}W[v, v^*] = -W[x_1, x_2] \neq 0,$$

and thus the quantity $\mathrm{Im}(\dot{v}v^*)$ is a nonzero real constant. If $v(t)$ is multiplied by a constant, $v(t) \to \lambda v(t)$, the Wronskian $W[v, v^*]$ changes by the factor $|\lambda|^2$. Therefore we may normalize $v(t)$ to a prescribed value of $\mathrm{Im}(\dot{v}v^*)$ by choosing the constant λ, as long as v and v^* are linearly independent solutions so that $W[v, v^*] \neq 0$.

A complex solution $v(t)$ of Eq. (35) is an admissible mode function if $v(t)$ is normalized by the condition $\mathrm{Im}(\dot{v}v^*) = 1$. It follows that any solution $v(t)$ normalized by $\mathrm{Im}(\dot{v}v^*) = 1$ is necessarily complex-valued and such that $v(t)$ and $v^*(t)$ are a basis of linearly independent complex solutions of Eq. (35).

4.2 Choice of mode function

We have seen that different choices of the mode function $v(t)$ lead to different definitions of the operators $\hat{a}^{\pm}$ and thus to different "canditate ground states" $|0\rangle$. The true ground state of the oscillator is the lowest-energy state and not merely some state $|0\rangle$ satisfying $\hat{a}^- |0\rangle = 0$, where $\hat{a}^-$ is some arbitrary operator. Therefore we may try to choose $v(t)$ such that the mean energy $\langle 0| \hat{H} |0\rangle$ is minimized.

For any choice of the mode function $v(t)$, the Hamiltonian is expressed through the operators $\hat{a}^{\pm}$ as

$$\hat{H} = \frac{|\dot{v}|^2 + \omega^2 |v|^2}{4} \left(2\hat{a}^+\hat{a}^- + 1\right) + \frac{\dot{v}^2 + \omega^2 v^2}{4}\hat{a}^+\hat{a}^+ + \frac{\dot{v}^{*2} + \omega^2 v^{*2}}{4}\hat{a}^-\hat{a}^-. \tag{36}$$

Derivation of Eq. (36)

In the canonical variables, the Hamiltonian is

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2(t)\hat{q}^2.$$

Now we expand the operators $\hat{p}, \hat{q}$ through the mode functions using Eq. (31) and the commutation relation $\left[\hat{a}^+, \hat{a}^-\right] = 1$. For example, the term $\hat{p}^2$ gives

$$\begin{aligned}\hat{p}^2 &= \frac{1}{\sqrt{2}}\left(\dot{v}(t)\hat{a}^+ + \dot{v}^*(t)\hat{a}^-\right)\frac{1}{\sqrt{2}}\left(\dot{v}(t)\hat{a}^+ + \dot{v}^*(t)\hat{a}^-\right) \\ &= \frac{1}{2}\left(\dot{v}^2\hat{a}^+\hat{a}^+ + \dot{v}\dot{v}^*\left(2\hat{a}^+\hat{a}^- + 1\right) + \dot{v}^{*2}\hat{a}^-\hat{a}^-\right).\end{aligned}$$

The term $\hat{q}^2$ gives

$$\hat{q}^2 = \frac{1}{\sqrt{2}} \left(v(t)\hat{a}^+ + v^*(t)\hat{a}^- \right) \frac{1}{\sqrt{2}} \left(v(t)\hat{a}^+ + v^*(t)\hat{a}^- \right)$$
$$= \frac{1}{2} \left(v^2 \hat{a}^+ \hat{a}^+ + vv^* \left(2\hat{a}^+ \hat{a}^- + 1 \right) + v^{*2} \hat{a}^- \hat{a}^- \right).$$

After some straightforward algebra we obtain the required result.

It is easy to see from Eq. (36) that the mean energy at time t is given by

$$E(t) \equiv \langle 0 | \hat{H}(t) | 0 \rangle = \frac{|\dot{v}(t)|^2 + \omega^2(t) |v(t)|^2}{4}. \tag{37}$$

We would like to find the mode function $v(t)$ that minimizes the above quantity. Note that $E(t)$ is time-dependent, so we may first try to minimize $E(t_0)$ at a *fixed* time t_0.

The choice of the mode function $v(t)$ may be specified by a set of initial conditions at $t = t_0$,

$$v(t_0) = q, \quad \dot{v}(t_0) = p,$$

where the parameters p and q are complex numbers satisfying the normalization constraint which follows from Eq. (33),

$$q^* p - p^* q = 2i. \tag{38}$$

Now we need to find such p and q that minimize the expression $|p|^2 + \omega^2(t_0) |q|^2$. This is a straightforward exercise (see below) which yields, for $\omega(t_0) > 0$, the following result:

$$v(t_0) = \frac{1}{\sqrt{\omega(t_0)}}, \quad \dot{v}(t_0) = i\sqrt{\omega(t_0)} = i\omega(t_0) v(t_0). \tag{39}$$

If, on the other hand, $\omega^2(t_0) < 0$ (i.e. ω is imaginary), there is no minimum. For now, we shall assume that $\omega(t_0)$ is real. Then the mode function satisfying Eq. (39) will define the operators $\hat{a}^\pm$ and the state $|_{t_0} 0\rangle$ such that the instantaneous energy $E(t_0)$ has the lowest possible value $E_{\text{min}} = \frac{1}{2}\omega(t_0)$. The state $|_{t_0} 0\rangle$ is called the **instantaneous ground state** at time $t = t_0$.

Derivation of Eq. (39)

If some p and q minimize $|p|^2 + \omega^2 |q|^2$, then so do $e^{i\lambda} p$ and $e^{i\lambda} q$ for arbitrary real λ; this is the freedom of choosing the overall phase of the mode function. We may choose this phase to make q real and write $p = p_1 + ip_2$ with real $p_{1,2}$. Then Eq. (38) yields

$$q = \frac{2i}{p - p^*} = \frac{1}{p_2} \Rightarrow 4E(t_0) = p_1^2 + p_2^2 + \frac{\omega^2(t_0)}{p_2^2}. \tag{40}$$

If $\omega^2(t_0) > 0$, the function $E(p_1, p_2)$ has a minimum with respect to $p_{1,2}$ at $p_1 = 0$ and $p_2 = \sqrt{\omega(t_0)}$. Therefore the desired initial conditions for the mode function are given by Eq. (39).

On the other hand, if $\omega^2(t_0) < 0$ the function $E_{\mathbf{k}}$ in Eq. (40) has no minimum because the expression $p_2^2 + \omega^2(t_0) p_2^{-2}$ varies from $-\infty$ to $+\infty$. In that case the instantaneous lowest-energy ground state does not exist.

4.3 "In" and "out" states

Let us now consider the frequency function $\omega(t)$ shown in Fig. 1. It is easy to see that the lowest-energy state is given by the mode function $v_{\text{in}}(t) = e^{i\omega_0 t}$ in the "in" regime ($t \leq t_0$) and by $v_{\text{out}}(t) = e^{i\omega_1 t}$ in the "out" regime ($t \geq t_1$). However, note that $v_{\text{in}}(t) \neq e^{i\omega_0 t}$ for $t > t_0$; instead, $v_{\text{in}}(t)$ is a solution of Eq. (29) with the initial conditions (39) at $t = t_0$. Similarly, $v_{\text{out}}(t) \neq e^{i\omega_1 t}$ for $t < t_1$. While exact solutions for $v_{\text{in}}(t)$ and $v_{\text{out}}(t)$ are in general not available, we may still analyze the relationship between these solutions in the "in" and "out" regimes.

Since the solutions $e^{\pm i\omega_1 t}$ are a basis in the space of solutions of Eq. (35), we may write

$$v_{\text{in}}(t) = \alpha v_{\text{out}}(t) + \beta v^*_{\text{out}}(t), \tag{41}$$

where α and β are *time-independent* constants. The relationship (41) between the mode functions is an example of a **Bogolyubov transformation** (see Sec. 2.2). Using Eq. (33) for $v_{\text{in}}(t)$ and $v_{\text{out}}(t)$, it is straightforward to derive the property

$$|\alpha|^2 - |\beta|^2 = 1. \tag{42}$$

For a general $\omega(t)$, we will have $\beta \neq 0$ and hence there will be no single mode function $v(t)$ matching both $v_{\text{in}}(t)$ and $v_{\text{out}}(t)$.

Each choice of the mode function $v(t)$ defines the corresponding creation and annihilation operators $\hat{a}^{\pm}$. Let us denote by $\hat{a}^{\pm}_{\text{in}}$ the operators defined using the mode function $v_{\text{in}}(t)$ and $v_{\text{out}}(t)$, respectively. It follows from Eqs. (34) and (41) that

$$\hat{a}^{-}_{\text{in}} = \alpha \hat{a}^{-}_{\text{out}} - \beta \hat{a}^{+}_{\text{out}}.$$

The inverse relation is easily found using Eq. (42),

$$\hat{a}^{-}_{\text{out}} = \alpha^* \hat{a}^{-}_{\text{in}} + \beta \hat{a}^{+}_{\text{in}}. \tag{43}$$

Since generally $\beta \neq 0$, we cannot define a single set of operators $\hat{a}^{\pm}$ which will define the ground state $|0\rangle$ for all times.

Moreover, in the intermediate regime where $\omega(t)$ is not constant, an instantaneous ground state $|_t 0\rangle$ defined at time t will, in general, not be a ground state at the next moment, $t + \Delta t$. Therefore, such a state $|_t 0\rangle$ cannot be trusted as a physically motivated ground state. However, if we restrict our attention only to the "in" regime, the mode function $v_{\text{in}}(t)$ defines a perfectly sensible ground state $|0_{\text{in}}\rangle$ which remains the ground state for all $t \leq t_0$. Similarly, the mode function $v_{\text{out}}(t)$ defines the ground state $|0_{\text{out}}\rangle$.

Since we are using the Heisenberg picture, the quantum state $|\psi\rangle$ of the oscillator is time-independent. It is reasonable to plan the following experiment. We prepare the oscillator in its ground state $|\psi\rangle = |0_{\text{in}}\rangle$ at some early time $t < t_0$ within the "in" regime. Then we let the oscillator evolve until the time $t = t_1$ and compare its quantum state (which remains $|0_{\text{in}}\rangle$) with the true ground state, $|0_{\text{out}}\rangle$, at time $t > t_1$ within the "out" regime.

In the "out" regime, the state $|0_{\text{in}}\rangle$ is not the ground state any more, and thus it must be a superposition of the true ground state $|0_{\text{out}}\rangle$ and the excited states $|n_{\text{out}}\rangle$ defined using the "out" creation operator $\hat{a}^{+}_{\text{out}}$,

$$|n_{\text{out}}\rangle = \frac{1}{\sqrt{n!}} \left(\hat{a}^{+}_{\text{out}}\right)^n |0_{\text{out}}\rangle, \quad n = 0, 1, 2, \ldots$$

It can be easily verified that the vectors $|n_{\text{out}}\rangle$ are eigenstates of the Hamiltonian for $t \geq t_1$ (but not for $t < t_1$):

$$\hat{H}(t) |n_{\text{out}}\rangle = \omega_1 \left(n + \frac{1}{2}\right) |n_{\text{out}}\rangle, \quad t \geq t_1.$$

Similarly, the excited states $|n_{\text{in}}\rangle$ may be defined through the creation operator $\hat{a}^{+}_{\text{in}}$. The states $|n_{\text{in}}\rangle$ are interpreted as n-particle states of the oscillator for $t \leq t_0$, while for $t \geq t_1$ the n-particle states are $|n_{\text{out}}\rangle$.

Remark: interpretation of the "in" and "out" states. We are presently working in the Heisenberg picture where quantum states are time-independent and operators depend on time. One may prepare the oscillator in a state $|\psi\rangle$, and the state of the oscillator remains the same throughout all time t. However, the physical interpretation of this state changes with time because the state $|\psi\rangle$ is interpreted with help of the time-dependent operators $\hat{H}(t)$, $\hat{a}^-(t)$, etc. For instance, we found that at late times ($t \geq t_1$) the vector $|0_{\text{in}}\rangle$ is not the lowest-energy state any more. This happens because the energy of the system changes with time due to the external force that drives $\omega(t)$. Without this force, we would have $\hat{a}^-_{\text{in}} = \hat{a}^-_{\text{out}}$ and the state $|0_{\text{in}}\rangle$ would describe the physical vacuum at all times.

4.4 Relationship between "in" and "out" states

The states $|n_{\text{out}}\rangle$, where $n = 0, 1, 2, ...$, form a complete basis in the Hilbert space of the harmonic oscillator. However, the set of states $|n_{\text{in}}\rangle$ is another complete basis in the same space. Therefore the vector $|0_{\text{in}}\rangle$ must be expressible as a linear combination of the "out" states,

$$|0_{\text{in}}\rangle = \sum_{n=0}^{\infty} \Lambda_n |n_{\text{out}}\rangle , \tag{44}$$

where Λ_n are suitable coefficients. If the mode functions are related by a Bogolyubov transformation (41), one can show that these coefficients Λ_n are given by

$$\Lambda_{2n} = \left[1 - \left|\frac{\beta}{\alpha}\right|^2\right]^{1/4} \left(\frac{\beta}{\alpha}\right)^n \sqrt{\frac{(2n-1)!!}{(2n)!!}}, \quad \Lambda_{2n+1} = 0. \tag{45}$$

The relation (44) shows that the early-time ground state is a superposition of excited states at late times, having the probability $|\Lambda_n|^2$ for the occupation number n. We thus conclude that the presence of an external influence leads to excitations of the oscillator. (Later on, when we consider field theory, such excitations will be interpreted as particle production.) In the present case, the influence of external forces on the oscillator consists of the changing frequency $\omega(t)$, which is formally a parameter of the Lagrangian. For this reason, the excitations arising in a time-dependent oscillator are called **parametric excitations**.

Finally, let us compute the expected particle number in the "out" regime, assuming that the oscillator is in the state $|0_{\text{in}}\rangle$. The expectation value of the number operator $\hat{N}_{\text{out}} \equiv \hat{a}^+_{\text{out}}\hat{a}^-_{\text{out}}$ in the state $|0_{\text{in}}\rangle$ is easily found using Eq. (43):

$$\langle 0_{\text{in}}| \hat{a}^+_{\text{out}}\hat{a}^-_{\text{out}} |0_{\text{in}}\rangle = \langle 0_{\text{in}}| \left(\alpha\hat{a}^+_{\text{in}} + \beta^*\hat{a}^-_{\text{in}}\right) \left(\alpha^*\hat{a}^-_{\text{in}} + \beta\hat{a}^+_{\text{in}}\right) |0_{\text{in}}\rangle = |\beta|^2 .$$

Therefore, a nonzero coefficient β signifies the presence of particles in the "out" region.

Derivation of Eq. (45)

In order to find the coefficients Λ_n, we need to solve the equation

$$0 = \hat{a}^-_{\text{in}} |0_{\text{in}}\rangle = \left(\alpha\hat{a}^-_{\text{out}} - \beta\hat{a}^+_{\text{out}}\right) \sum_{n=0}^{\infty} \Lambda_n |n_{\text{out}}\rangle .$$

Using the known properties

$$\hat{a}^+ |n\rangle = \sqrt{n+1}\, |n+1\rangle , \quad \hat{a}^- |n\rangle = \sqrt{n}\, |n-1\rangle ,$$

we obtain the recurrence relation

$$\Lambda_{n+2} = \Lambda_n \frac{\beta}{\alpha}\sqrt{\frac{n+1}{n+2}}; \quad \Lambda_1 = 0.$$

Therefore, only even-numbered Λ_{2n} are nonzero and may be expressed through Λ_0 as follows,

$$\Lambda_{2n} = \Lambda_0 \left(\frac{\beta}{\alpha}\right)^n \sqrt{\frac{1 \cdot 3 \cdot ... \cdot (2n-1)}{2 \cdot 4 \cdot ... \cdot (2n)}} \equiv \Lambda_0 \left(\frac{\beta}{\alpha}\right)^n \sqrt{\frac{(2n-1)!!}{(2n)!!}}, \quad n \geq 1.$$

For convenience, one defines $(-1)!! = 1$, so the above expression remains valid also for $n = 0$.

The value of Λ_0 is determined from the normalization condition, $\langle 0_{\text{in}} | 0_{\text{in}} \rangle = 1$, which can be rewritten as

$$|\Lambda_0|^2 \sum_{n=0}^{\infty} \left| \frac{\beta}{\alpha} \right|^{2n} \frac{(2n-1)!!}{(2n)!!} = 1.$$

The infinite sum can be evaluated as follows. Let $f(z)$ be an auxiliary function defined by the series

$$f(z) \equiv \sum_{n=0}^{\infty} z^{2n} \frac{(2n-1)!!}{(2n)!!}.$$

At this point one can guess that this is a Taylor expansion of $f(z) = \left(1 - z^2\right)^{-1/2}$; then one obtains $\Lambda_0 = 1/\sqrt{f(z)}$ with $z \equiv |\beta/\alpha|$. If we would like to avoid guessing, we could manipulate the above series in order to derive a differential equation for $f(z)$:

$$\begin{aligned} f(z) &= 1 + \sum_{n=1}^{\infty} z^{2n} \frac{2n-1}{2n} \frac{(2n-3)!!}{(2n-2)!!} = 1 + \sum_{n=1}^{\infty} z^{2n} \frac{(2n-3)!!}{(2n-2)!!} - \sum_{n=1}^{\infty} \frac{z^{2n}}{2n} \frac{(2n-3)!!}{(2n-2)!!} \\ &= 1 + z^2 f(z) - \sum_{n=1}^{\infty} \frac{z^{2n}}{2n} \frac{(2n-3)!!}{(2n-2)!!}. \end{aligned}$$

Taking d/dz of both parts, we have

$$\frac{d}{dz} \left[1 + z^2 f(z) - f(z) \right] = \sum_{n=1}^{\infty} z^{2n-1} \frac{(2n-3)!!}{(2n-2)!!} = z f(z),$$

hence $f(z)$ satisfies

$$\frac{d}{dz} \left[(z^2 - 1) f(z) \right] = \left(z^2 - 1 \right) \frac{df}{dz} + 2z f(z) = z f(z); \quad f(0) = 1.$$

The solution is

$$f(z) = \frac{1}{\sqrt{1 - z^2}}.$$

Substituting $z \equiv |\beta/\alpha|$, we obtain the required result,

$$\Lambda_0 = \left[1 - \left| \frac{\beta}{\alpha} \right|^2 \right]^{1/4}.$$

4.5 Quantum-mechanical analogy

The time-dependent oscillator equation (35) is formally similar to the stationary Schrödinger equation for the wave function $\psi(x)$ of a quantum particle in a one-dimensional potential $V(x)$,

$$\frac{d^2 \psi}{dx^2} + (E - V(x)) \psi = 0.$$

The two equations are related by the replacements $t \to x$ and $\omega^2(t) \to E - V(x)$.

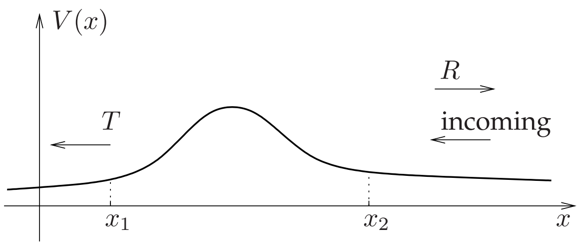

Figure 2: Quantum-mechanical analogy: motion in a potential $V(x)$.

To illustrate the analogy, let us consider the case when the potential $V(x)$ is almost constant for $x < x_1$ and for $x > x_2$ but varies in the intermediate region (see Fig. 2). An incident wave $\psi(x) = \exp(-ipx)$ comes from large positive x and is scattered off the potential. A reflected wave $\psi_R(x) = R\exp(ipx)$ is produced in the region $x > x_2$ and a transmitted wave $\psi_T(x) = T\exp(-ipx)$ in the region $x < x_1$. For most potentials, the reflection amplitude R is nonzero. The conservation of probability gives the constraint $|R|^2 + |T|^2 = 1$.

The wavefunction $\psi(x)$ behaves similarly to the mode function $v(t)$ in the case when $\omega(t)$ is approximately constant at $t \leq t_0$ and at $t \geq t_1$. If the wavefunction represents a pure incoming wave $x < x_1$, then at $x > x_2$ the function $\psi(x)$ will be a superposition of positive and negative exponents $\exp(\pm ikx)$. This is the phenomenon known as **over-barrier reflection**: there is a small probability that the particle is reflected by the potential, even though the energy is above the height of the barrier. The relation between R and T is similar to the normalization condition (42) for the Bogolyubov coefficients. The presence of the over-barrier reflection ($R \neq 0$) is analogous to the presence of particles in the "out" region ($\beta \neq 0$).

5 Scalar field in expanding universe

Let us now turn to the situation when quantum fields are influenced by strong gravitational fields. In this chapter, we use units where $c = G = 1$, where G is Newton's constant.

5.1 Curved spacetime

Einstein's theory of gravitation (**General Relativity**) is based on the notion of **curved spacetime**, i.e. a manifold with arbitrary coordinates $x \equiv \{x^\mu\}$ and a metric $g_{\mu\nu}(x)$ which replaces the flat Minkowski metric $\eta_{\mu\nu}$. The metric defines the **interval**

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu,$$

which describes physically measured lengths and times. According to the Einstein equation, the metric $g_{\mu\nu}(x)$ is determined by the distribution of matter in the entire universe.

Here are some basic examples of spacetimes. In the absence of matter, the metric is equal to the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}\,(1, -1, -1, -1)$ (in Cartesian coordinates). In the presence of a single black hole of mass M, the metric can be written in spherical coordinates as

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2 \left[d\theta^2 + \sin^2\theta d\phi^2\right].$$

Finally, a certain class of spatially homogeneous and isotropic distributions of matter in the universe yields a metric of the form

$$ds^2 = dt^2 - a^2(t)\left[dx^2 + dy^2 + dy^2\right], \tag{46}$$

where $a(t)$ is a certain function called the **scale factor**. (The interpretation is that $a(t)$ "scales" the flat metric $dx^2 + dy^2 + dz^2$ at different times.) Spacetimes with metrics of the form (46) are called Friedmann-Robertson-Walker (FRW) spacetimes with flat spatial sections (in short, **flat FRW** spacetimes). Note that it is only the *three-dimensional* spatial sections which are flat; the four-dimensional geometry of such spacetimes is usually curved. The class of flat FRW spacetimes is important in cosmology because its geometry agrees to a good precision with the present results of astrophysical measurements.

In this course, we shall not be concerned with the task of obtaining the metric. It will be assumed that a metric $g_{\mu\nu}(x)$ is already known in some coordinates $\{x^\mu\}$.

5.2 Scalar field in cosmological background

Presently, we shall study the behavior of a quantum field in a flat FRW spacetime with the metric (46). In Einstein's General Relativity, every kind of energy influences the geometry of spacetime. However, we shall treat the metric $g_{\mu\nu}(x)$ as fixed and disregard the influence of fields on the geometry.

A **minimally coupled**, free, real, massive scalar field $\phi(x)$ in a curved spacetime is described by the action

$$S = \int \sqrt{-g}d^4x \left[\frac{1}{2}g^{\alpha\beta}\left(\partial_\alpha\phi\right)\left(\partial_\beta\phi\right) - \frac{1}{2}m^2\phi^2\right]. \tag{47}$$

(Note the difference between Eq. (47) and Eq. (11): the Minkowski metric $\eta_{\mu\nu}$ is replaced by the curved metric $g_{\mu\nu}$, and the integration uses the covariant volume element $\sqrt{-g}d^4x$. This is the minimal change necessary to make the theory of the scalar field compatible with General Relativity.) The equation of motion for the field ϕ is derived straightforwardly as

$$g^{\mu\nu}\partial_\mu\partial_\nu\phi + \frac{1}{\sqrt{-g}}\left(\partial_\nu\phi\right)\partial_\mu\left(g^{\mu\nu}\sqrt{-g}\right) + m^2\phi = 0. \tag{48}$$

This equation can be rewritten more concisely using the covariant derivative corresponding to the metric $g_{\mu\nu}$,

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi + m^2\phi = 0,$$

which shows explicitly that this is a generalization of the Klein-Gordon equation to curved spacetime.

We cannot directly use the quantization technique developed for fields in the flat spacetime. First, let us carry out a few mathematical transformations to simplify the task.

The metric (46) for a flat FRW spacetime can be simplified if we replace the coordinate t by the **conformal time** η,

$$\eta(t) \equiv \int_{t_0}^{t}\frac{dt}{a(t)},$$

where t_0 is an arbitrary constant. The scale factor $a(t)$ must be expressed through the new variable η; let us denote that function again by $a(\eta)$. In the coordinates $(\mathbf{x}, \eta)$, the interval is

$$ds^2 = a^2(\eta)\left[d\eta^2 - d\mathbf{x}^2\right], \tag{49}$$

so the metric is **conformally flat** (equal to the flat metric multiplied by a factor): $g_{\mu\nu} = a^2\eta_{\mu\nu}$, $g^{\mu\nu} = a^{-2}\eta^{\mu\nu}$.

Further, it is convenient to introduce the auxiliary field $\chi \equiv a(\eta)\phi$. Then one can show that the action (47) can be rewritten in terms of the field χ as follows,

$$S = \frac{1}{2}\int d^3\mathbf{x}\, d\eta \left(\chi'^2 - (\nabla\chi)^2 - m_{\text{eff}}^2(\eta)\chi^2\right), \tag{50}$$

where the prime $'$ denotes $\partial/\partial\eta$, and m_{eff} is the time-dependent **effective mass**

$$m_{\text{eff}}^2(\eta) \equiv m^2a^2 - \frac{a''}{a}. \tag{51}$$

The action (50) is very similar to the action (11), except for the presence of the time-dependent mass.

Derivation of Eq. (50)

We start from Eq. (47). Using the metric (49), we have $\sqrt{-g} = a^4$ and $g^{\alpha\beta} = a^{-2}\eta^{\alpha\beta}$. Then

$$\sqrt{-g}\, m^2\phi^2 = m^2a^2\chi^2,$$
$$\sqrt{-g}\, g^{\alpha\beta}\phi_{,\alpha}\phi_{,\beta} = a^2\left(\phi'^2 - (\nabla\phi)^2\right).$$

Substituting $\phi = \chi/a$, we get

$$a^2\phi'^2 = \chi'^2 - 2\chi\chi'\frac{a'}{a} + \chi^2\left(\frac{a'}{a}\right)^2 = \chi'^2 + \chi^2\frac{a''}{a} - \left[\chi^2\frac{a'}{a}\right]'.$$

The total time derivative term can be omitted from the action, and we obtain the required expression.

Thus, the dynamics of a scalar field ϕ in a flat FRW spacetime is mathematically equivalent to the dynamics of the auxiliary field χ in the Minkowski spacetime. All the information about the influence of gravitation on the field ϕ is encapsulated in the time-dependent mass $m_{\text{eff}}(\eta)$ defined by Eq. (51). Note that the action (50) is explicitly time-dependent, so the energy of the field χ is generally not conserved. We shall see that in quantum theory this leads to the possibility of particle creation; the energy for new particles is supplied by the gravitational field.

5.3 Mode expansion

It follows from the action (50) that the equation of motion for $\chi(\mathbf{x}, \eta)$ is

$$\chi'' - \Delta\chi + \left(m^2a^2 - \frac{a''}{a}\right)\chi = 0. \tag{52}$$

Expanding the field χ in Fourier modes,

$$\chi(\mathbf{x}, \eta) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}}\chi_{\mathbf{k}}(\eta)e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{53}$$

we obtain from Eq. (52) the decoupled equations of motion for the modes $\chi_{\mathbf{k}}(\eta)$,

$$\chi_{\mathbf{k}}'' + \left[k^2 + m^2a^2(\eta) - \frac{a''}{a}\right]\chi_{\mathbf{k}} \equiv \chi_{\mathbf{k}}'' + \omega_k^2(\eta)\chi_{\mathbf{k}} = 0. \tag{54}$$

All the modes $\chi_{\mathbf{k}}(\eta)$ with equal $|\mathbf{k}| = k$ are complex solutions of the same equation (54). This equation describes a harmonic oscillator with a time-dependent frequency. Therefore, we may apply the techniques we developed in chapter 4.

We begin by choosing a mode function $v_k(\eta)$, which is a complex-valued solution of

$$v_k'' + \omega_k^2(\eta) v_k = 0, \quad \omega_k^2(\eta) \equiv k^2 + m_{\text{eff}}^2(\eta). \tag{55}$$

Then, the general solution $\chi_{\mathbf{k}}(\eta)$ is expressed as a linear combination of v_k and v_k^* as

$$\chi_{\mathbf{k}}(\eta) = \frac{1}{\sqrt{2}} \left[a_{\mathbf{k}}^- v_k^*(\eta) + a_{-\mathbf{k}}^+ v_k(\eta) \right], \tag{56}$$

where $a_{\mathbf{k}}^{\pm}$ are complex constants of integration that depend on the vector $\mathbf{k}$ (but not on η). The index $-\mathbf{k}$ in the second term of Eq. (56) and the factor $\frac{1}{\sqrt{2}}$ are chosen for later convenience.

Since χ is real, we have $\chi_{\mathbf{k}}^* = \chi_{-\mathbf{k}}$. It follows from Eq. (56) that $a_{\mathbf{k}}^+ = \left(a_{\mathbf{k}}^-\right)^*$. Combining Eqs. (53) and (56), we find

$$\begin{aligned} \chi(\mathbf{x}, \eta) &= \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2}} \left[a_{\mathbf{k}}^- v_k^*(\eta) + a_{-\mathbf{k}}^+ v_k(\eta) \right] e^{i\mathbf{k}\cdot\mathbf{x}} \\ &= \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2}} \left[a_{\mathbf{k}}^- v_k^*(\eta) e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^+ v_k(\eta) e^{-i\mathbf{k}\cdot\mathbf{x}} \right]. \end{aligned} \tag{57}$$

Note that the integration variable $\mathbf{k}$ was changed ($\mathbf{k} \to -\mathbf{k}$) in the second term of Eq. (57) to make the integrand a manifestly real expression. (This is done only for convenience.)

The relation (57) is called the **mode expansion** of the field $\chi(\mathbf{x}, \eta)$ w.r.t. the mode functions $v_k(\eta)$. At this point the choice of the mode functions is still arbitrary.

The coefficients $a_{\mathbf{k}}^{\pm}$ are easily expressed through $\chi_{\mathbf{k}}(\eta)$ and $v_k(\eta)$:

$$a_{\mathbf{k}}^- = \sqrt{2} \frac{v_k' \chi_{\mathbf{k}} - v_k \chi_{\mathbf{k}}'}{v_k' v_k^* - v_k v_k^{*\prime}} = \sqrt{2} \frac{W\left[v_k, \chi_{\mathbf{k}}\right]}{W\left[v_k, v_k^*\right]}; \quad a_{\mathbf{k}}^+ = \left(a_{\mathbf{k}}^-\right)^*. \tag{58}$$

Note that the numerators and denominators in Eq. (58) are time-independent since they are Wronskians of solutions of the same oscillator equation.

5.4 Quantization of scalar field

The field $\chi(x)$ can be quantized directly through the mode expansion (57), which can be used for quantum fields in the same way as for classical fields. The mode expansion for the field operator $\hat{\chi}$ is found by replacing the constants $a_{\mathbf{k}}^{\pm}$ in Eq. (57) by time-independent operators $\hat{a}_{\mathbf{k}}^{\pm}$:

$$\hat{\chi}(\mathbf{x}, \eta) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2}} \left(e^{i\mathbf{k}\cdot\mathbf{x}} v_k^*(\eta) \hat{a}_{\mathbf{k}}^- + e^{-i\mathbf{k}\cdot\mathbf{x}} v_k(\eta) \hat{a}_{\mathbf{k}}^+ \right), \tag{59}$$

where $v_k(\eta)$ are mode functions obeying Eq. (55). The operators $\hat{a}_{\mathbf{k}}^{\pm}$ satisfy the usual commutation relations for creation and annihilation operators,

$$\left[\hat{a}_{\mathbf{k}}^-, \hat{a}_{\mathbf{k}'}^+\right] = \delta(\mathbf{k} - \mathbf{k}'), \quad \left[\hat{a}_{\mathbf{k}}^-, \hat{a}_{\mathbf{k}'}^-\right] = \left[\hat{a}_{\mathbf{k}}^+, \hat{a}_{\mathbf{k}'}^+\right] = 0. \tag{60}$$

The commutation relations (60) are consistent with the canonical relations

$$[\chi(\mathbf{x}_1, \eta), \chi'(\mathbf{x}_2, \eta)] = i\delta(\mathbf{x}_1 - \mathbf{x}_2)$$

only if the mode functions $v_k(\eta)$ are normalized by the condition

$$\text{Im}\left(v_k' v_k^*\right) = \frac{v_k' v_k^* - v_k v_k'^*}{2i} \equiv \frac{W\left[v_k, v_k^*\right]}{2i} = 1. \tag{61}$$

Therefore, quantization of the field $\hat{\chi}$ can be accomplished by postulating the mode expansion (59), the commutation relations (60) and the normalization (61). The choice of the mode functions $v_k(\eta)$ will be made later on.

The mode expansion (59) can be visualized as the general solution of the field equation (52), where the operators $\hat{a}_{\mathbf{k}}^{\pm}$ are integration constants. The mode expansion can also be viewed as a *definition* of the operators $\hat{a}_{\mathbf{k}}^{\pm}$ through the field operator $\hat{\chi}(\mathbf{x}, \eta)$. Explicit formulae relating $\hat{a}_{\mathbf{k}}^{\pm}$ to $\hat{\chi}$ and $\hat{\pi} \equiv \hat{\chi}'$ are analogous to Eq. (58). Clearly, the definition of $\hat{a}_{\mathbf{k}}^{\pm}$ depends on the choice of the mode functions $v_k(\eta)$.

5.5 Vacuum state and particle states

Once the operators $\hat{a}_{\mathbf{k}}^{\pm}$ are determined, the vacuum state $|0\rangle$ is defined as the eigenstate of all annihilation operators $\hat{a}_{\mathbf{k}}^{-}$ with eigenvalue 0, i.e. $\hat{a}_{\mathbf{k}}^{-}|0\rangle = 0$ for all $\mathbf{k}$. An excited state $|m_{\mathbf{k}_1}, n_{\mathbf{k}_2}, ...\rangle$ with the occupation numbers $m, n, ...$ in the modes $\chi_{\mathbf{k}_1}$, $\chi_{\mathbf{k}_2}$, ..., is constructed by

$$|m_{\mathbf{k}_1}, n_{\mathbf{k}_2}, ...\rangle \equiv \frac{1}{\sqrt{m!n!...}} \left[\left(\hat{a}_{\mathbf{k}_1}^{+} \right)^m \left(\hat{a}_{\mathbf{k}_2}^{+} \right)^n ... \right] |0\rangle . \tag{62}$$

We write $|0\rangle$ instead of $|0_{\mathbf{k}_1}, 0_{\mathbf{k}_2}, ...\rangle$ for brevity. An arbitrary quantum state $|\psi\rangle$ is a linear combination of these states,

$$|\psi\rangle = \sum_{m,n,...} C_{mn...} |m_{\mathbf{k}_1}, n_{\mathbf{k}_2}, ...\rangle .$$

If the field is in the state $|\psi\rangle$, the probability for measuring the occupation number m in the mode $\chi_{\mathbf{k}_1}$, the number n in the mode $\chi_{\mathbf{k}_2}$, etc., is $|C_{mn...}|^2$.

Let us now comment on the role of the mode functions. Complex solutions $v_k(\eta)$ of a second-order differential equation (55) with one normalization condition (61) are parametrized by one complex parameter. Multiplying $v_k(\eta)$ by a constant phase $e^{i\alpha}$ introduces an extra phase $e^{\pm i\alpha}$ in the operators $\hat{a}_k^{\pm}$, which can be compensated by a constant phase factor $e^{i\alpha}$ in the state vectors $|0\rangle$ and $|m_{\mathbf{k}_1}, n_{\mathbf{k}_2}, ...\rangle$. There remains one real free parameter that distinguishes physically inequivalent mode functions. With each possible choice of the functions $v_k(\eta)$, the operators $\hat{a}_{\mathbf{k}}^{\pm}$ and consequently the vacuum state and particle states are different. As long as the mode functions satisfy Eqs. (55) and (61), the commutation relations (60) hold and thus the operators $\hat{a}_{\mathbf{k}}^{\pm}$ formally resemble the creation and annihilation operators for particle states. However, we do not yet know whether the operators $\hat{a}_{\mathbf{k}}^{\pm}$ obtained with some choice of $v_k(\eta)$ actually correspond to physical particles and whether the quantum state $|0\rangle$ describes the physical vacuum. The correct commutation relations alone do not guarantee the validity of the physical interpretation of the operators $\hat{a}_{\mathbf{k}}^{\pm}$ and of the state $|0\rangle$. For this interpretation to be valid, the mode functions must be *appropriately selected*; we postpone the consideration of this important issue until Sec. 5.8 below. For now, we shall formally study the consequences of choosing several sets of mode functions to quantize the field ϕ.

5.6 Bogolyubov transformations

Suppose two sets of isotropic mode functions $u_k(\eta)$ and $v_k(\eta)$ are chosen. Since u_k and u_k^* are a basis, the function v_k is a linear combination of u_k and u_k^*, e.g.

$$v_k^*(\eta) = \alpha_k u_k^*(\eta) + \beta_k u_k(\eta), \tag{63}$$

with η-independent complex coefficients α_k and β_k. If both sets $v_k(\eta)$ and $u_k(\eta)$ are normalized by Eq. (61), it follows that the coefficients α_k and β_k satisfy

$$|\alpha_k|^2 - |\beta_k|^2 = 1. \tag{64}$$

In particular, $|\alpha_k| \geq 1$.

Derivation of Eq. (64)

We suppress the index k for brevity. The normalization condition for $u(\eta)$ is

$$u^* u' - u u'^* = 2i.$$

Expressing u through v as given, we obtain

$$\left(|\alpha|^2 - |\beta|^2\right)\left(v^* v' - v v'^*\right) = 2i.$$

The formula (64) follows from the normalization of $v(\eta)$.

Using the mode functions $u_k(\eta)$ instead of $v_k(\eta)$, one obtains an alternative mode expansion which defines another set $\hat{b}^{\pm}_{\mathbf{k}}$ of creation and annihilation operators,

$$\hat{\chi}(\mathbf{x}, \eta) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2}} \left(e^{i\mathbf{k}\cdot\mathbf{x}} u^*_k(\eta) \hat{b}^-_{\mathbf{k}} + e^{-i\mathbf{k}\cdot\mathbf{x}} u_k(\eta) \hat{b}^+_{\mathbf{k}} \right). \tag{65}$$

The expansions (59) and (65) express the same field $\hat{\chi}(\mathbf{x}, \eta)$ through two different sets of functions, so the $\mathbf{k}$-th Fourier components of these expansions must agree,

$$e^{i\mathbf{k}\cdot\mathbf{x}} \left[u^*_k(\eta) \hat{b}^-_{\mathbf{k}} + u_k(\eta) \hat{b}^+_{-\mathbf{k}} \right] = e^{i\mathbf{k}\cdot\mathbf{x}} \left[v^*_k(\eta) \hat{a}^-_{\mathbf{k}} + v_k(\eta) \hat{a}^+_{-\mathbf{k}} \right].$$

A substitution of v_k through u_k using Eq. (63) gives the following relation between the operators $\hat{b}^{\pm}_{\mathbf{k}}$ and $\hat{a}^{\pm}_{\mathbf{k}}$:

$$\hat{b}^-_{\mathbf{k}} = \alpha_k \hat{a}^-_{\mathbf{k}} + \beta^*_k \hat{a}^+_{-\mathbf{k}}, \quad \hat{b}^+_{\mathbf{k}} = \alpha^*_k \hat{a}^+_{\mathbf{k}} + \beta_k \hat{a}^-_{-\mathbf{k}}. \tag{66}$$

The relation (66) and the complex coefficients α_k, β_k are called respectively the **Bogolyubov transformation** and the **Bogolyubov coefficients**.[6]

The two sets of annihilation operators $\hat{a}^-_{\mathbf{k}}$ and $\hat{b}^-_{\mathbf{k}}$ define the corresponding vacua $\left|_{(a)}0\right\rangle$ and $\left|_{(b)}0\right\rangle$, which we call the "$a$-vacuum" and the "$b$-vacuum." Two parallel sets of excited states are built from the two vacua using Eq. (62). We refer to these states as a-particle and b-particle states. So far the physical interpretation of the a- and b-particles remains unspecified. Later on, we shall apply this formalism to study specific physical effects and the interpretation of excited states corresponding to various mode functions will be explained. At this point, let us only remark that the b-vacuum is in general a superposition of a-states, similarly to what we found in Sec. 4.4.

5.7 Mean particle number

Let us calculate the mean number of b-particles of the mode $\chi_{\mathbf{k}}$ in the a-vacuum state. The expectation value of the b-particle number operator $\hat{N}^{(b)}_{\mathbf{k}} = \hat{b}^+_{\mathbf{k}} \hat{b}^-_{\mathbf{k}}$ in the state $\left|_{(a)}0\right\rangle$ is found using Eq. (66):

$$\begin{aligned}
\left\langle_{(a)}0\right| \hat{N}^{(b)} \left|_{(a)}0\right\rangle &= \left\langle_{(a)}0\right| \hat{b}^+_{\mathbf{k}} \hat{b}^-_{\mathbf{k}} \left|_{(a)}0\right\rangle \\
&= \left\langle_{(a)}0\right| \left(\alpha^*_k \hat{a}^+_{\mathbf{k}} + \beta_k \hat{a}^-_{-\mathbf{k}}\right) \left(\alpha_k \hat{a}^-_{\mathbf{k}} + \beta^*_k \hat{a}^+_{-\mathbf{k}}\right) \left|_{(a)}0\right\rangle \\
&= \left\langle_{(a)}0\right| \left(\beta_k \hat{a}^-_{-\mathbf{k}}\right) \left(\beta^*_k \hat{a}^+_{-\mathbf{k}}\right) \left|_{(a)}0\right\rangle = |\beta_k|^2 \, \delta^{(3)}(0).
\end{aligned} \tag{67}$$

The divergent factor $\delta^{(3)}(0)$ is a consequence of considering an infinite spatial volume. This divergent factor would be replaced by the box volume V if we quantized the field in a finite box. Therefore we can divide by this factor and obtain the mean *density* of b-particles in the mode $\chi_{\mathbf{k}}$,

$$n_k = |\beta_k|^2. \tag{68}$$

[6] The pronunciation is close to the American "bogo-lube-of" with the third syllable stressed.

The Bogolyubov coefficient β_k is dimensionless and the density n_k is the mean number of particles per spatial volume $d^3\mathbf{x}$ and per wave number $d^3\mathbf{k}$, so that $\int n_k d^3\mathbf{k}\, d^3\mathbf{x}$ is the (dimensionless) total mean number of b-particles in the a-vacuum state.

The combined mean density of particles in all modes is $\int d^3\mathbf{k}\, |\beta_k|^2$. Note that this integral might diverge, which would indicate that one cannot disregard the backreaction of the produced particles on other fields and on the metric.

5.8 Instantaneous lowest-energy vacuum

In the theory developed so far, the particle interpretation depends on the choice of the mode functions. For instance, the a-vacuum $\left|_{(a)}0\right\rangle$ defined above is a state without a-particles but with b-particle density n_k in each mode $\chi_{\mathbf{k}}$. A natural question to ask is whether the a-particles or the b-particles are the correct representation of the observable particles. The problem at hand is to determine the mode functions that describe the "actual" physical vacuum and particles.

Previously, we defined the vacuum state as the eigenstate with the lowest energy. However, in the present case the Hamiltonian explicitly depends on time and thus does not have time-independent eigenstates that could serve as vacuum states.

One possible prescription for the vacuum state is to select a particular moment of time, $\eta = \eta_0$, and to define the vacuum $|_{\eta_0}0\rangle$ as the lowest-energy eigenstate of the *instantaneous* Hamiltonian $\hat{H}(\eta_0)$. To obtain the mode functions that describe the vacuum $|_{\eta_0}0\rangle$, we first compute the expectation value $\left\langle_{(v)}0\right| \hat{H}(\eta_0) \left|_{(v)}0\right\rangle$ in the vacuum state $\left|_{(v)}0\right\rangle$ determined by arbitrarily chosen mode functions $v_k(\eta)$. Then we can minimize that expectation value with respect to all possible choices of $v_k(\eta)$. (A standard result in linear algebra is that the minimization of $\langle x| \hat{A} |x\rangle$ with respect to all normalized vectors $|x\rangle$ is equivalent to finding the eigenvector $|x\rangle$ of the operator $\hat{A}$ with the smallest eigenvalue.) This computation is analogous to that of Sec. 4.2, and the result is similar to Eq. (39): If $\omega_k^2(\eta_0) > 0$, the required initial conditions for the mode functions are

$$v_k(\eta_0) = \frac{1}{\sqrt{\omega_k(\eta_0)}}, \quad v_k'(\eta_0) = i\sqrt{\omega_k(\eta_0)} = i\omega_k v_k(\eta_0). \tag{69}$$

If $\omega_k^2(\eta_0) < 0$, the instantaneous lowest-energy vacuum state does not exist.

For a scalar field in the Minkowski spacetime, ω_k is time-independent and the prescription (69) yields the standard mode functions

$$v_k(\eta) = \frac{1}{\sqrt{\omega_k}} e^{i\omega_k \eta},$$

which remain the vacuum mode functions at all times. But this is not the case for a time-dependent gravitational background, because then $\omega_k(\eta) \neq \text{const}$ and the mode function selected by the initial conditions (69) imposed at a time η_0 will generally differ from the mode function selected at another time $\eta_1 \neq \eta_0$. In other words, the state $|_{\eta_0}0\rangle$ is not an energy eigenstate at time η_1. In fact, one can show that there are no states which remain instantaneous eigenstates of the Hamiltonian at all times.

5.9 Computation of Bogolyubov coefficients

Computations of Bogolyubov coefficients requires knowledge of solutions of Eq. (55), which is an equation of a harmonic oscillator with a time-dependent frequency, with specified initial conditions. Suppose that $v_k(\eta)$ and $u_k(\eta)$ are mode functions describing instantaneous lowest-energy states defined at times $\eta = \eta_0$ and $\eta = \eta_1$. To determine the Bogolyubov coefficients α_k and β_k connecting these mode functions, it is necessary to know the functions $v_k(\eta)$ and $u_k(\eta)$ and their derivatives at *only one* value of η, e.g. at $\eta = \eta_0$. From Eq. (63) and its derivative at $\eta = \eta_0$, we find

$$
\begin{aligned}
v_k^*(\eta_0) &= \alpha_k u_k^*(\eta_0) + \beta_k u_k(\eta_0), \\
v_k^{*\prime}(\eta_0) &= \alpha_k u_k^{*\prime}(\eta_0) + \beta_k u_k^{\prime}(\eta_0).
\end{aligned}
$$

This system of equations can be solved for α_k and β_k using Eq. (61):

$$
\alpha_k = \left.\frac{u_k' v_k^* - u_k v_k^{*\prime}}{2i}\right|_{\eta_0}, \quad \beta_k^* = \left.\frac{u_k' v_k - u_k v_k'}{2i}\right|_{\eta_0}. \tag{70}
$$

These relations hold at any time η_0 (note that the numerators are Wronskians and thus are time-independent). For instance, knowing only the asymptotics of $v_k(\eta)$ and $u_k(\eta)$ at $\eta \to -\infty$ would suffice to compute α_k and β_k.

A well-known method to obtain an approximate solution of equations of the type (55) is the **WKB approximation**, which gives the approximate solution satisfying the condition (69) at time $\eta = \eta_0$ as

$$
v_k(\eta) \approx \frac{1}{\sqrt{\omega_k(\eta)}} \exp\left[i \int_{\eta_0}^{\eta} \omega_k(\eta_1) d\eta_1\right]. \tag{71}
$$

However, it is straightforward to see that the approximation (71) satisfies the instantaneous minimum-energy condition at every other time $\eta \neq \eta_0$ as well. In other words, within the WKB approximation, $u_k(\eta) \approx v_k(\eta)$. Therefore, if we use the WKB approximation to compute the Bogolyubov coefficient between instantaneous vacuum states, we shall obtain the incorrect result $\beta_k = 0$. The WKB approximation is insufficiently precise to capture the difference between the instantaneous vacuum states defined at different times.

One can use the following method to obtain a better approximation to the mode function $v_k(\eta)$. Let us focus attention on one mode and drop the index k. Introduce a new variable $Z(\eta)$ instead of $v(\eta)$ as follows,

$$
v(\eta) = \frac{1}{\sqrt{\omega(\eta_0)}} \exp\left[i \int_{\eta_0}^{\eta} \omega(\eta_1) d\eta_1 + \int_{\eta_0}^{\eta} Z(\eta_1) d\eta_1\right]; \quad \frac{v'}{v} = i\omega(\eta) + Z(\eta).
$$

If $\omega(\eta)$ is a slow-changing function, then we expect that $v(\eta)$ is everywhere approximately equal to $\frac{1}{\sqrt{\omega}} \exp\left[i \int^{\eta} \omega(\eta) d\eta\right]$ and the function $Z(\eta)$ under the exponential is a small correction; in particular, $Z(\eta_0) = 0$ at the time η_0. It is straightforward to derive the equation for $Z(\eta)$ from Eq. (55),

$$
Z' + 2i\omega Z = -i\omega' - Z^2.
$$

This equation can be solved using perturbation theory by treating Z^2 as a small perturbation. To obtain the first approximation, we disregard Z^2 and straightforwardly solve the resulting linear equation, which yields

$$
Z_{(1)}(\eta) = -i \int_{\eta_0}^{\eta} d\eta_1 \omega'(\eta_1) \exp\left[-2i \int_{\eta_1}^{\eta} \omega(\eta_2) d\eta_2\right].
$$

Note that $Z_{(1)}(\eta)$ is an integral of a slow-changing function $\omega'(\eta)$ multiplied by a quickly oscillating function and is therefore small, $|Z| \ll \omega$. The first approximation $Z_{(1)}$ is sufficiently precise in most cases. The resulting *approximate* mode function is

$$v(\eta) \approx \frac{1}{\sqrt{\omega(\eta_0)}} \exp\left[i\int_{\eta_0}^{\eta} \omega(\eta_1) d\eta_1 + \int_{\eta_0}^{\eta} Z_{(1)}(\eta) d\eta\right].$$

The Bogolyubov coefficient β between instantaneous lowest-energy states defined at times η_0 and η_1 can be approximately computed using Eq. (70):

$$u(\eta_1) = \frac{1}{\sqrt{\omega(\eta_1)}}; \quad u'(\eta_1) = i\omega(\eta_1)u(\eta_1);$$

$$\beta^* = \frac{1}{2i\sqrt{\omega(\eta_1)}}\left[i\omega(\eta_1)v(\eta_1) - v'(\eta_1)\right] = v(\eta_1)\frac{\left(i\omega - \frac{v'}{v}\right)_{\eta_1}}{2i\sqrt{\omega(\eta_1)}} \approx -\frac{v(\eta_1)Z(\eta_1)}{2i\sqrt{\omega(\eta_1)}}.$$

Since $v(\eta_1)$ is of order $\omega^{-1/2}$ and $|Z| \ll \omega$, the number of particles is small: $|\beta|^2 \ll 1$.

6 Amplitude of quantum fluctuations

In the previous chapter the focus was on particle production. The main observable of interest was the average particle number $\langle \hat{N} \rangle$. Now we consider another important quantity—the amplitude of field fluctuations.

6.1 Fluctuations of averaged fields

The value of a field cannot be observed at a mathematical point in space. Realistic devices can only measure the value of the field averaged over some region of space. Spatial averages are also the relevant quantity in cosmology because structure formation in the universe is explained by fluctuations occurring over large regions, e.g. of galaxy size. Therefore, let us consider values of fields averaged over a spatial domain.

A convenient way to describe spatial averaging over arbitrary domains is by using window functions. A **window function** for scale L is any function $W(\mathbf{x})$ which is of order 1 for $|\mathbf{x}| \lesssim L$, rapidly decays for $|\mathbf{x}| \gg L$, and satisfies the normalization condition

$$\int W(\mathbf{x})\, d^3\mathbf{x} = 1. \tag{72}$$

A typical example of a window function is the **spherical Gaussian** window

$$W_L(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}L^3} \exp\left[-\frac{|\mathbf{x}|^2}{2L^2}\right],$$

which selects $|\mathbf{x}| \lesssim L$. This window can be used to describe measurements performed by a device that cannot resolve distances smaller than L.

We define the **averaged field** operator $\hat{\chi}_L(\eta)$ by integrating the product of $\hat{\chi}(\mathbf{x}, \eta)$ with a window function that selects the scale L,

$$\hat{\chi}_L(\eta) \equiv \int d^3\mathbf{x}\, \hat{\chi}(\mathbf{x}, \eta)\, W_L(\mathbf{x}),$$

where we used the Gaussian window (although the final result will not depend on this choice). The amplitude $\delta\chi_L(\eta)$ of fluctuations in $\hat{\chi}_L(\eta)$ in a quantum state $|\psi\rangle$ is found from

$$\delta\chi_L^2(\eta) \equiv \langle\psi| \left[\hat{\chi}_L(\eta)\right]^2 |\psi\rangle.$$

For simplicity, we consider the vacuum state $|\psi\rangle = |0\rangle$. Then the amplitude of vacuum fluctuations in $\hat{\chi}_L(\eta)$ can be computed as a function of L. We use the mode expansion (59) for the field operator $\hat{\chi}(\mathbf{x}, \eta)$, assuming that the mode functions $v_k(\eta)$ are given. After some straightforward algebra we find

$$\langle 0| \left[\int d^3\mathbf{x}\, W_L(\mathbf{x}) \hat{\chi}\left(\mathbf{x}, \eta\right) \right]^2 |0\rangle = \frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left|v_k\right|^2 e^{-k^2 L^2}.$$

Since the factor $e^{-k^2 L^2}$ is of order 1 for $|\mathbf{k}| \lesssim L^{-1}$ and almost zero for $|\mathbf{k}| \gtrsim L^{-1}$, we can estimate the above integral as follows,

$$\frac{1}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left|v_k\right|^2 e^{-k^2 L^2} \sim \int_0^{L^{-1}} k^2 \left|v_k\right|^2 dk \sim k^3 \left|v_k\right|^2 \Big|_{k=L^{-1}}.$$

Thus the amplitude of fluctuations $\delta\chi_L$ is (up to a factor of order 1)

$$\delta\chi_L^2 \sim k^3 \left|v_k\right|^2, \text{ where } k \sim L^{-1}. \tag{73}$$

The result (73) is (for *any* choice of the window function W_L) an order-of-magnitude estimate of the **amplitude of fluctuations on scale** L. This quantity, which we denote by $\delta\chi_L(\eta)$, is defined only up to a factor of order 1 and is a function of time η and of the scale L. Expressed through the wavenumber $k \equiv 2\pi L^{-1}$, the fluctuation amplitude is usually called the **spectrum of fluctuations**.

6.2 Fluctuations in Minkowski spacetime

Let us now compute the spectrum of fluctuations for a scalar field in the Minkowski space.

The vacuum mode functions are $v_k(\eta) = \omega_k^{-1/2} \exp\left(i\omega_k \eta\right)$, where $\omega_k = \sqrt{k^2 + m^2}$. Thus, the spectrum of fluctuations in vacuum is

$$\delta\chi_L(\eta) = k^{3/2} \left|v_k(\eta)\right| = \frac{k^{3/2}}{\left(k^2 + m^2\right)^{1/4}}. \tag{74}$$

This time-independent spectrum is sketched in Fig. 3. When measured with a high-resolution device (small L), the field shows large fluctuations. On the other hand, if the field is averaged over a large volume ($L \to \infty$), the amplitude of fluctuations tends to zero.

6.3 de Sitter spacetime

The **de Sitter spacetime** is used in cosmology to describe periods of accelerated expansion of the universe. The amplitude of fluctuations is an important quantity to compute in that context.

The geometry of the de Sitter spacetime may be specified by a flat FRW metric

$$ds^2 = dt^2 - a^2(t) d\mathbf{x}^2 \tag{75}$$

with the scale factor $a(t)$ defined by

$$a(t) = a_0 e^{Ht}. \tag{76}$$

The **Hubble parameter** $H = \dot{a}/a > 0$ is a fixed constant. For convenience, we redefine the origin of time t to set $a_0 = 1$, so that $a(t) = \exp(Ht)$. The exponentially growing scale factor describes an accelerating expansion (**inflation**) of the universe.

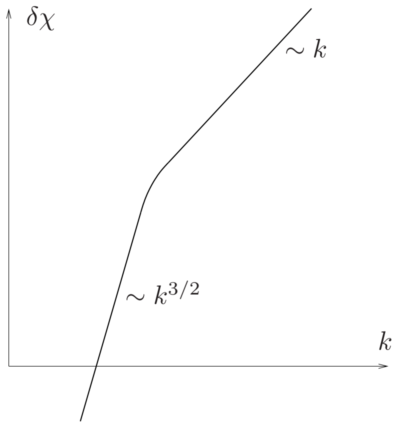

Figure 3: A sketch of the spectrum of fluctuations $\delta\chi_L$ in the Minkowski space; $L \equiv 2\pi k^{-1}$. (The logarithmic scaling is used for both axes.)

Horizons

An important feature of the de Sitter spacetime—the presence of horizons—is revealed by the following consideration of trajectories of lightrays. A null worldline $\mathbf{x}(t)$ satisfies $a^2(t)\dot{\mathbf{x}}^2(t) = 1$, which yields the solution

$$|\mathbf{x}(t)| = \frac{1}{H}\left(e^{-Ht_0} - e^{-Ht}\right)$$

for trajectories starting at the origin, $\mathbf{x}(t_0) = 0$. Therefore all lightrays emitted at the origin at $t = t_0$ asymptotically approach the sphere

$$|\mathbf{x}| = r_{\max}(t_0) \equiv H^{-1}\exp(-Ht_0) = (aH)^{-1}.$$

This sphere is the **horizon** for the observer at the origin; the spacetime expands too quickly for lightrays to reach any points beyond the horizon. Similarly, observers at the origin will never receive any lightrays emitted at $t = t_0$ at points $|\mathbf{x}| > r_{\max}$.

It is easy to verify that at any time t_0 the horizon is always at the same *proper* distance $a(t_0)r_{\max}(t_0) = H^{-1}$ from the observer. This distance is called the **horizon scale**.

6.4 Quantum fields in de Sitter spacetime

To describe a real scalar field $\phi(\mathbf{x}, t)$ in the de Sitter spacetime, we first transform the coordinate t to make the metric explicitly conformally flat:

$$ds^2 = dt^2 - a^2(t)d\mathbf{x}^2 = a^2(\eta)\left(d\eta^2 - d\mathbf{x}^2\right),$$

where the conformal time η and the scale factor $a(\eta)$ are

$$\eta = -\frac{1}{H}e^{-Ht}, \quad a(\eta) = -\frac{1}{H\eta}.$$

The conformal time η changes from $-\infty$ to 0 when the proper time t goes from $-\infty$ to $+\infty$. (Since the value of η is always negative, we shall sometimes have to write $|\eta|$ in the equations. However, it is essential that the variable η grows when t grows, so we cannot use $-\eta$ as the time variable. For convenience, we chose the origin of η so that the infinite future corresponds to $\eta = 0$.)

The field $\phi(\mathbf{x}, \eta)$ can now be quantized by the method of Sec. 2.2. We introduce the auxiliary field $\chi \equiv a\phi$ and use the mode expansion (59)-(55) with

$$\omega_k^2(\eta) = k^2 + m^2 a^2 - \frac{a''}{a} = k^2 + \left(\frac{m^2}{H^2} - 2\right)\frac{1}{\eta^2}. \tag{77}$$

From this expression it is clear that the effective frequency may become imaginary, i.e. $\omega_k^2(\eta) < 0$, if $m^2 < 2H^2$. In most cosmological scenarios where the early universe is approximated by a region of the de Sitter spacetime, the relevant value of H is much larger than the masses of elementary particles, i.e. $m \ll H$. Therefore, for simplicity we shall consider the massless field.

6.5 Bunch-Davies vacuum state

As we saw before, the vacuum state of a field is determined by the choice of mode functions. Let us now find the appropriate mode functions for a scalar field in de Sitter spacetime.

With the definition (77) of the effective frequency, where we set $m = 0$, Eq. (55) becomes

$$v_k'' + \left(k^2 - \frac{2}{\eta^2}\right) v_k = 0. \tag{78}$$

The general solution of Eq. (78) can be written as

$$v_k(\eta) = A_k \left(1 + \frac{i}{k\eta}\right) e^{ik\eta} + B_k \left(1 - \frac{i}{k\eta}\right) e^{-ik\eta}, \tag{79}$$

where A_k and B_k are constants. It is straightforward to check that the normalization of the mode function, $\mathrm{Im}\,(v_k^* v_k') = 1$, constrains the constants A_k and B_k by

$$|A_k|^2 - |B_k| = \frac{1}{k}.$$

The constants A_k, B_k determine the mode functions and must be chosen appropriately to obtain a physically motivated vacuum state. For fields in de Sitter spacetime, there is a preferred vacuum state, which is known as the **Bunch-Davies** (BD) **vacuum**. This state is defined essentially as the Minkowski vacuum in the early-time limit ($\eta \to -\infty$) of each mode.

Before introducing the BD vacuum, let us consider the prescription of the instantaneous vacuum defined at a time $\eta = \eta_0$. If we had $\omega_k^2(\eta_0) > 0$ for all k, this prescription would yield a well-defined vacuum state. However, there always exists a small enough k such that $k\,|\eta_0| \ll 1$ and thus $\omega_k^2(\eta_0) < 0$. We have seen that the energy in a mode $\chi_{\mathbf{k}}$ cannot be minimized when $\omega_k^2 < 0$. Therefore the instantaneous energy prescription cannot define a vacuum state of the entire quantum field (for all modes) but only for the modes $\chi_{\mathbf{k}}$ with $k\,|\eta_0| \gtrsim 1$. Note that these are the modes whose wavelength was shorter than the horizon length, $(aH)^{-1} = |\eta_0|$, at time η_0 (i.e. the **subhorizon** modes).

The motivation for introducing the BD vacuum state is the following. The effective frequency $\omega_k(\eta)$ becomes constant in the early-time limit $\eta \to -\infty$. Physically, this means that the influence of gravity on each mode $\chi_{\mathbf{k}}$ is negligible at sufficiently early (k-dependent) times. When gravity is negligible, there is a unique vacuum prescription—the Minkowski vacuum, which coincides with the instantaneous minimum-energy vacuum at all times. So it is natural to define the mode functions $v_k(\eta)$ by applying the Minkowski vacuum prescription in the limit $\eta \to -\infty$, separately for each mode $\chi_{\mathbf{k}}$. This prescription can be expressed by the asymptotic relations

$$v_k(\eta) \to \frac{1}{\sqrt{\omega_k}} e^{i\omega_k \eta}, \quad \frac{v_k'(\eta)}{v_k(\eta)} \to i\omega_k, \quad \text{as } \eta \to -\infty. \tag{80}$$

The vacuum state determined by the mode functions $v_k(\eta)$ satisfying Eq. (80) is called the **Bunch-Davies vacuum**. It follows that the mode functions of the BD vacuum are given by Eq. (79) with $A_k = 1/\sqrt{k}$ and $B_k = 0$, namely

$$v_k(\eta) = \frac{1}{\sqrt{k}} \left(1 + \frac{i}{k\eta}\right) e^{ik\eta}. \tag{81}$$

The Bunch-Davies vacuum prescription has important applications in cosmology where the de Sitter spacetime approximates the inflationary stage of the evolution of the universe. However, this approximation is valid only for a certain time interval, for instance $\eta_i < \eta < \eta_f$, while at earlier times, $\eta < \eta_i$, the spacetime is not de Sitter. Therefore the procedure of imposing the minimum-energy conditions at earlier times $\eta < \eta_i$ cannot be justified, and the BD vacuum state can be used only for modes $\chi_{\mathbf{k}}$ such that $k\,|\eta_i| \gg 1$. However, it is these modes that are important for cosmological predictions.

6.6 Spectrum of fluctuations in the BD vacuum

Let us now compute the fluctuation amplitude $\delta\phi_L(\eta)$ in the BD vacuum state and compare the result with the fluctuations in Minkowski spacetime.

According to the formula (73), the amplitude of fluctuations is determined by absolute values of the mode functions. Up to now we have been mostly working with the auxiliary field $\hat{\chi}(x) = a\hat{\phi}(x)$. The mode expansion for $\hat{\phi}(x)$ is simply $a^{-1}(\eta)$ times the mode expansion for $\hat{\chi}$. Therefore, the mode functions of the field $\hat{\phi}$ are $a^{-1}(\eta)v_k(\eta)$, where $v_k(\eta)$ are the mode functions of the field $\hat{\chi}$. Hence, the spectrum of fluctuations of $\hat{\phi}$ is

$$\delta\phi_L(\eta) = a^{-1}(\eta)k^{3/2}\,|v_k(\eta)| = H\sqrt{k^2\eta^2 + 1}. \tag{82}$$

This spectrum is to be compared with Eq. (74) with $m = 0$. The spectrum (82) for early times ($|\eta| \gg k$) is the same as in the Minkowski spacetime, while for late times or for superhorizon scales ($k \ll |\eta|^{-1}$) the spectrum (82) becomes almost independent of k (**scale-invariant**) and shows much larger fluctuations than the spectrum (74). The growth of fluctuations is due to the influence of gravity on the field $\hat{\phi}$.

The growth of quantum fluctuations is used in cosmology to explain the formation of large-scale structures (galaxies and clusters of galaxies) in the early universe. The theory of **cosmological inflation** assumes the existence of a de Sitter-like epoch in the history of the universe. During this epoch, vacuum fluctuations of the fields were significantly amplified. The resulting large quantum fluctuations acted as seeds for the inhomogeneities of energy density, which then grew by gravitational collapse and eventually caused the formation of galaxies. This theory is a practical application of quantum field theory in curved spacetime to astrophysics.

7 Unruh effect

The Unruh effect predicts that particles will be detected in a vacuum by an accelerated observer. In this chapter we consider the simplest case, in which the observer moves with constant acceleration through Minkowski spacetime and measures the number of particles in a massless scalar field. Even though the field is in the vacuum state, the observer finds a distribution of particles characteristic of a thermal bath of blackbody radiation.

7.1 Kinematics of uniformly accelerated motion

First we consider the trajectory of an object moving with constant acceleration in the Minkowski spacetime. A model of this situation is a spaceship with an infinite energy supply and a propulsion engine that exerts a constant force (but moves with the ship). The resulting motion of the spaceship is such that the acceleration of the ship in its own frame of reference (the **proper acceleration**) is constant. This is the natural definition of a uniformly accelerated motion in a relativistic theory. (An object cannot move with $d\mathbf{v}/dt = \text{const}$ for all time because its velocity is always smaller than the speed of light, $|\mathbf{v}| < 1$.)

We now introduce the reference frames that will play a major role in our considerations: the laboratory frame, the proper frame, and the comoving frame. The laboratory frame is the usual inertial reference frame with the coordinates (t, x, y, z). The proper frame is the accelerated system of reference that moves together with the observer; we shall also call it the accelerated frame. The comoving frame defined at a time t_0 is the *inertial* frame in which the accelerated observer is instantaneously at rest at $t = t_0$. (Thus the term **comoving frame** actually refers to a different frame for each t_0.)

By definition, the observer's proper acceleration at time $t = t_0$ is the 3-acceleration measured in the comoving frame at time t_0. We consider a uniformly accelerated observer whose proper acceleration is time-independent and equal to a given 3-vector $\mathbf{a}$. The trajectory of such an observer may be described by a worldline $x^\mu(\tau)$, where τ is the proper time measured by the observer. The proper time parametrization implies the condition

$$u^\mu u_\mu = 1, \quad u^\mu \equiv \frac{dx^\mu}{d\tau}. \tag{83}$$

It is a standard result that the 4-acceleration in the laboratory frame,

$$a^\mu \equiv \frac{du^\mu}{d\tau} = \frac{d^2 x^\mu}{d\tau^2},$$

is related to the three-dimensional proper acceleration $\mathbf{a}$ by

$$a^\mu a_\mu = -|\mathbf{a}|^2. \tag{84}$$

Derivation of Eq. (84). Let $u^\mu(\tau)$ be the observer's 4-velocity and let t_c be the time variable in the comoving frame defined at $\tau = \tau_0$; this is the time measured by an *inertial* observer moving with the constant velocity $u^\mu(\tau_0)$. We shall show that the 4-acceleration $a^\mu(\tau)$ in the comoving frame has components $\left(0, a^1, a^2, a^3\right)$, where a^i are the components of the acceleration 3-vector $\mathbf{a} \equiv d^2\mathbf{x}/dt_c^2$ measured in the comoving frame. It will then follow that Eq. (84) holds in the comoving frame, and hence it holds also in the laboratory frame since the Lorentz-invariant quantity $a^\mu a_\mu$ is the same in all frames.

Since the comoving frame moves with the velocity $u^\mu(\tau_0)$, the 4-vector $u^\mu(\tau_0)$ has the components $(1, 0, 0, 0)$ in that frame. The derivative of the identity $u^\mu(\tau)u_\mu(\tau) = 1$ with respect to τ yields $a^\mu(\tau)u_\mu(\tau) = 0$, therefore $a^0(\tau_0) = 0$ in the comoving frame. Since $dt_c = u^0(\tau)d\tau$ and $u^0(\tau_0) = 1$, we have

$$\frac{d^2 x^\mu}{dt_c^2} = \frac{1}{u^0}\frac{d}{d\tau}\left[\frac{1}{u^0}\frac{dx^\mu}{d\tau}\right] = \frac{d^2 x^\mu}{d\tau^2} + \frac{dx^\mu}{d\tau}\frac{d}{d\tau}\frac{1}{u^0}.$$

It remains to compute

$$\frac{d}{d\tau}\frac{1}{u^0(\tau_0)} = -\left[u^0(\tau_0)\right]^{-2}\left.\frac{du^0}{d\tau}\right|_{\tau=\tau_0} = -a^0(\tau_0) = 0,$$

and it follows that $d^2x^\mu/d\tau^2 = d^2x^\mu/dt_c^2 = (0, a^1, a^2, a^3)$ as required. (Self-test question: why is $a^\mu = du^\mu/d\tau \neq 0$ even though $u^\mu = (1,0,0,0)$ in the comoving frame?)

We now derive the trajectory $x^\mu(\tau)$ of the accelerated observer. Without loss of generality, we may assume that the acceleration is parallel to the x axis, $\mathbf{a} \equiv (a, 0, 0)$, where $a > 0$, and that the observer moves only in the x direction. Then the coordinates y and z of the observer remain constant and only the functions $x(\tau)$, $t(\tau)$ need to be computed. From Eqs. (83)-(84) it is straightforward to derive the general solution

$$x(\tau) = x_0 - \frac{1}{a} + \frac{1}{a}\cosh a\tau, \quad t(\tau) = t_0 + \frac{1}{a}\sinh a\tau. \tag{85}$$

This trajectory has zero velocity at $\tau = 0$ (which implies $x = x_0$, $t = t_0$).

Derivation of Eq. (85). Since $a^\mu = du^\mu/d\tau$ and $u^2 = u^3 = 0$, the components u^0, u^1 of the velocity satisfy

$$\left(\frac{du^0}{d\tau}\right)^2 - \left(\frac{du^1}{d\tau}\right)^2 = -a^2,$$
$$\left(u^0\right)^2 - \left(u^1\right)^2 = 1.$$

We may assume that $u_0 > 0$ (the time τ grows together with t) and that $du^1/d\tau > 0$, since the acceleration is in the positive x direction. Then

$$u^0 = \sqrt{1 + (u^1)^2}; \quad \frac{du^1}{d\tau} = a\sqrt{1 + (u^1)^2}.$$

The solution with the initial condition $u^1(0) = 0$ is

$$u^1(\tau) \equiv \frac{dx}{d\tau} = \sinh a\tau, \quad u^0(\tau) \equiv \frac{dt}{d\tau} = \cosh a\tau.$$

After an integration we obtain Eq. (85).

The trajectory (85) has a simpler form if we choose the initial conditions $x(0) = a^{-1}$ and $t(0) = 0$. Then the worldline is a branch of the hyperbola $x^2 - t^2 = a^{-2}$ (see Fig. 4). At large $|t|$ the worldline approaches the lightcone. The observer comes in from $x = +\infty$, decelerates and stops at $x = a^{-1}$, and then accelerates back towards infinity. In the comoving frame of the observer, this motion takes infinite proper time, from $\tau = -\infty$ to $\tau = +\infty$.

From now on, we drop the coordinates y and z and work in the 1+1-dimensional spacetime (t, x).

7.2 Coordinates in the proper frame

To describe quantum fields as seen by an accelerated observer, we need to use the **proper coordinates** (τ, ξ), where τ is the proper time and ξ is the distance measured by the observer. The proper coordinate system (τ, ξ) is related to the laboratory frame (t, x) by some transformation functions $\tau(t, x)$ and $\xi(t, x)$ which we shall now determine.

The observer's trajectory $t(\tau)$, $x(\tau)$ should correspond to the line $\xi = 0$ in the proper coordinates. Let the observer hold a rigid measuring stick of proper length ξ_0, so that the entire stick accelerates together with the observer. Then the stick is instantaneously at rest in the comoving frame and the far endpoint of the stick has the proper coordinates (τ, ξ_0) at time τ. We shall derive the relation between the coordinates (t, x) and (τ, ξ) by computing the laboratory coordinates (t, x) of the far end of the stick as functions of τ and ξ_0.

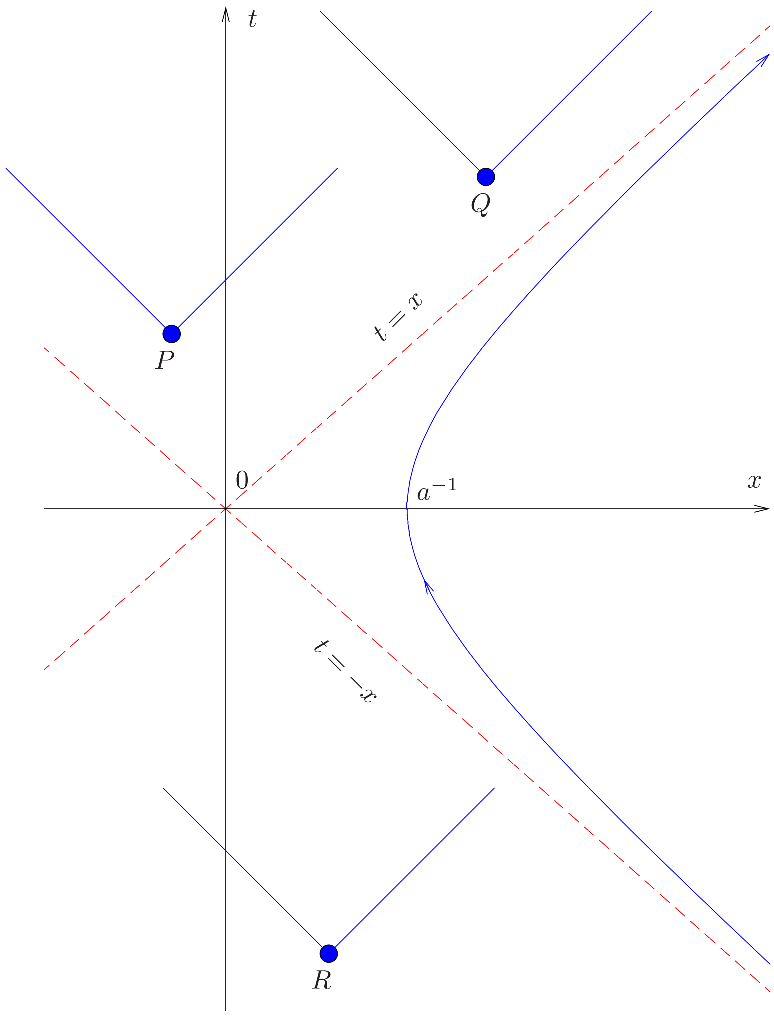

Figure 4: The worldline of a uniformly accelerated observer (proper acceleration $a \equiv |\mathbf{a}|$) in the Minkowski spacetime. The dashed lines show the lightcone. The observer cannot receive any signals from the events P, Q and cannot send signals to R.

In the comoving frame at time τ, the stick is represented by the 4-vector $s^{\mu}_{(\text{com})} \equiv (0, \xi_0)$ connecting the endpoints $(\tau, 0)$ and (τ, ξ_0). This comoving frame is an inertial system of reference moving with the 4-velocity $u^{\mu}(\tau) = dx^{\mu}/d\tau$. Therefore the coordinates $s^{\mu}_{(\text{lab})}$ of the stick in the laboratory frame can be found by applying the inverse Lorentz transformation to the coordinates $s^{\mu}_{(\text{com})}$:

$$\left[\begin{array}{c} s^0_{(\text{lab})} \\ s^1_{(\text{lab})} \end{array}\right] = \frac{1}{\sqrt{1-v^2}} \left(\begin{array}{cc} 1 & v \\ v & 1 \end{array}\right) \left[\begin{array}{c} s^0_{(\text{com})} \\ s^1_{(\text{com})} \end{array}\right] = \left(\begin{array}{cc} u^0 & u^1 \\ u^1 & u^0 \end{array}\right) \left[\begin{array}{c} s^0_{(\text{com})} \\ s^1_{(\text{com})} \end{array}\right] = \left[\begin{array}{c} u^1 \xi \\ u^0 \xi \end{array}\right],$$

where $v \equiv u^1/u^0$ is the velocity of the stick in the laboratory system. The stick is attached to the observer moving along $x^{\mu}(\tau)$, so the proper coordinates (τ, ξ) of the far end of the stick correspond to the laboratory coordinates

$$t(\tau, \xi) = x^0(\tau) + s^0_{(\text{lab})} = x^0(\tau) + \frac{dx^1(\tau)}{d\tau}\xi, \tag{86}$$

$$x(\tau, \xi) = x^1(\tau) + s^1_{(\text{lab})} = x^1(\tau) + \frac{dx^0(\tau)}{d\tau}\xi. \tag{87}$$

Note that the relations (86)-(87) specify the proper frame for *any* trajectory $x^{0,1}(\tau)$ in the 1+1-dimensional Minkowski spacetime.

Now we can substitute Eq. (85) into the above relations to compute the proper coordinates for a uniformly accelerated observer. We choose the initial conditions $x^0(0) = 0$, $x^1(0) = a^{-1}$ for the observer's trajectory and obtain

$$t(\tau, \xi) = \frac{1 + a\xi}{a} \sinh a\tau, \quad x(\tau, \xi) = \frac{1 + a\xi}{a} \cosh a\tau. \tag{88}$$

The converse relations are

$$\tau(t, x) = \frac{1}{2a} \ln \frac{x+t}{x-t}, \quad \xi(t, x) = -a^{-1} + \sqrt{x^2 - t^2}.$$

The horizon

It can be seen from Eq. (88) that the coordinates (τ, ξ) vary in the intervals $-\infty < \tau < +\infty$ and $-a^{-1} < \xi < +\infty$. In particular, for $\xi < -a^{-1}$ we would find $\partial t/\partial\tau < 0$, i.e. the direction of time t would be opposite to that of τ. One can verify that an accelerated observer cannot measure distances longer than a^{-1} in the direction opposite to the acceleration, for instance, the distances to the events P and Q in Fig. 4. A measurement of the distance to a point requires to place a clock at that point and to synchronize that clock with the observer's clock. However, the observer cannot synchronize clocks with the events P and Q because no signals can be ever received from these events. One says that the accelerated observer perceives a **horizon** at proper distance a^{-1}.

The existence of the horizon can be easily seen by the following qualitative considerations. The accelerated observer measures a constant gravitational field with acceleration a pointing in the negative x direction. Consider a photon of energy E propagating "upwards" in the gravitational field. After ascending to a height Δh, the photon loses energy, $\Delta E = -Ea\Delta h$. After ascending to the height $\Delta h = a^{-1}$, the photon would lose all its kinetic energy and will not be able to continue moving upwards. (This calculation is of course not rigorous but does give the correct answer.)

The coordinate system (88) is *incomplete* and covers only a "quarter" of the Minkowski spacetime, consisting of the subdomain $x > |t|$ (see Fig. 5). This is the subdomain of the Minkowski spacetime accessible to a uniformly accelerated observer. For instance, the events P, Q, R cannot be described by (real) values of τ and ξ. The past lightcone $x = -t$ corresponds to the proper coordinates $\tau = -\infty$ and $\xi = -a^{-1}$. The observer can see signals from the event R, however these signals appear to have originated not from R but from the horizon $\xi = -a^{-1}$ in the infinite past $\tau = -\infty$.

Another way to see that the line $\xi = -a^{-1}$ is a horizon is to consider a line of constant proper length $\xi = \xi_0 > -a^{-1}$. It follows from Eq. (88) that the line $\xi = \xi_0$ is a trajectory of the form $x^2 - t^2 = \text{const}$ with the proper acceleration

$$a_0 \equiv \frac{1}{\sqrt{x^2 - t^2}} = \left(\xi_0 + a^{-1}\right)^{-1}.$$

Therefore, the worldline $\xi = -a^{-1}$ would have to represent an infinite proper acceleration, which would require an infinitely large force and is thus impossible. It follows that an accelerated observer cannot hold a rigid measuring stick longer than a^{-1} in the direction opposite to acceleration. (A **rigid** stick is one that would keep its proper distance constant in the observer's reference frame.)

7.3 Rindler spacetime

It is straightforward to show that the Minkowski metric in the proper coordinates (τ, ξ) is

$$ds^2 = dt^2 - dx^2 = (1 + a\xi)^2 d\tau^2 - d\xi^2. \tag{89}$$

The spacetime with this metric is called the Rindler spacetime. The curvature of the Rindler spacetime is everywhere zero since it differs from the Minkowski spacetime merely by a change of coordinates.

To develop the quantum field theory in the Rindler spacetime, we first rewrite the metric (89) in a conformally flat form. This can be achieved by choosing the new spatial coordinate $\tilde{\xi}$ such that $d\xi = (1 + a\xi)d\tilde{\xi}$, because in that case both $d\tau^2$ and $d\tilde{\xi}^2$ will have a common factor $(1 + a\xi)^2$. (Note that the new coordinate $\tilde{\xi}$ is a "conformal distance" fully analogous to the "conformal time" variable η used in Sec. 5.4.) The necessary replacement is therefore

$$\tilde{\xi} \equiv \frac{1}{a} \ln(1 + a\xi).$$

Since the proper distance ξ is constrained by $\xi > -a^{-1}$, the conformal distance $\tilde{\xi}$ varies in the interval $-\infty < \tilde{\xi} < +\infty$. The metric becomes

$$ds^2 = e^{2a\tilde{\xi}}(d\tau^2 - d\tilde{\xi}^2). \tag{90}$$

The relation between the laboratory coordinates and the conformal coordinates is

$$t(\tau, \tilde{\xi}) = a^{-1} e^{a\tilde{\xi}} \sinh a\tau, \quad x(\tau, \tilde{\xi}) = a^{-1} e^{a\tilde{\xi}} \cosh a\tau. \tag{91}$$

7.4 Quantum field in Rindler spacetime

The goal of this section is to quantize a scalar field in the proper reference frame of a uniformly accelerated observer. To simplify the problem, we consider a massless scalar field in the 1+1-dimensional spacetime.

The action for a massless scalar field $\phi(t, x)$ is

$$S[\phi] = \frac{1}{2} \int g^{\alpha\beta} \phi_{,\alpha} \phi_{,\beta} \sqrt{-g} d^2x.$$

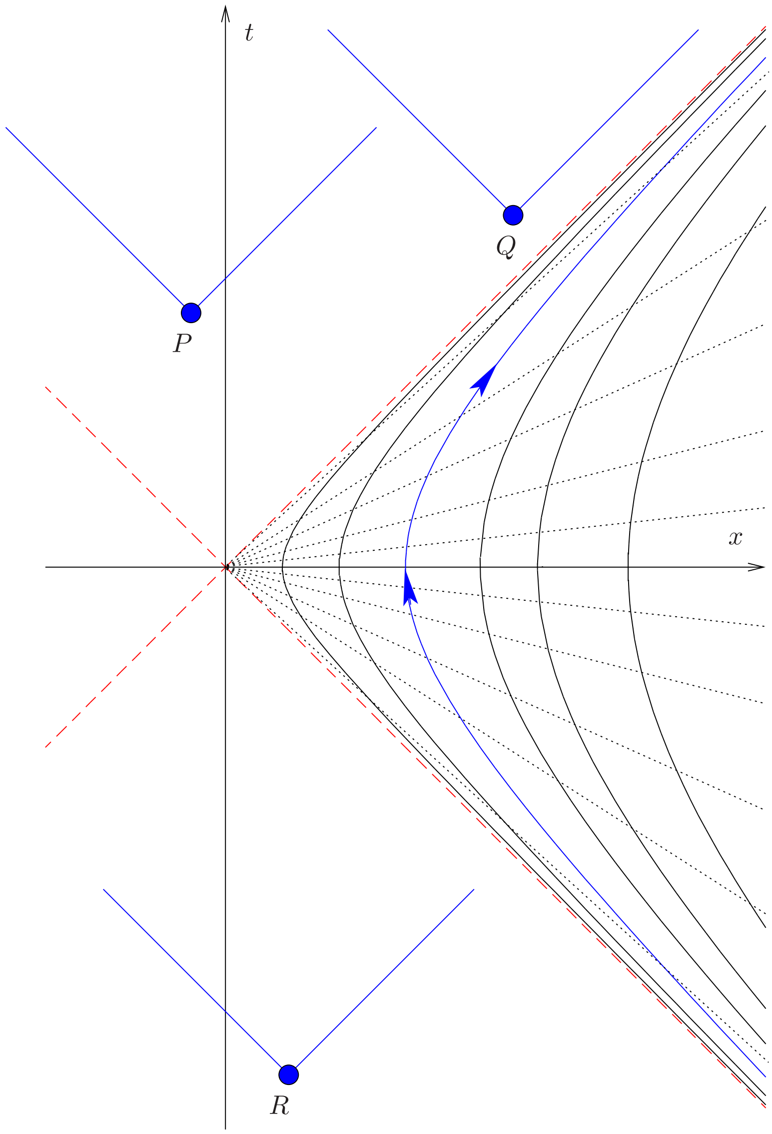

Figure 5: The proper coordinate system of a uniformly accelerated observer in the Minkowski spacetime. The solid hyperbolae are the lines of constant proper distance ξ; the hyperbola with arrows is the worldline of the observer, $\xi = 0$ or $x^2 - t^2 = a^{-2}$. The lines of constant τ are dotted. The dashed lines show the lightcone which corresponds to $\xi = -a^{-1}$. The events P, Q, R are not covered by the proper coordinate system.

Here $x^\mu \equiv (t, x)$ is the two-dimensional coordinate. It is easy to see that this action is conformally invariant: indeed, if we replace

$$g_{\alpha\beta} \to \tilde{g}_{\alpha\beta} = \Omega^2(t, x) g_{\alpha\beta},$$

then the determinant $\sqrt{-g}$ and the contravariant metric are replaced by

$$\sqrt{-g} \to \Omega^2 \sqrt{-g}, \quad g^{\alpha\beta} \to \Omega^{-2} g^{\alpha\beta}, \tag{92}$$

so the factors Ω^2 cancel in the action. The conformal invariance causes a significant simplification of the theory in 1+1 dimensions (a calculation for massive field in 3+1 dimensions would be more complicated).

In the laboratory coordinates (t, x), the action is

$$S[\phi] = \frac{1}{2} \int \left[(\partial_t \phi)^2 - (\partial_x \phi)^2 \right] dt\, dx.$$

In the conformal coordinates, the metric (90) is equal to the flat Minkowski metric multiplied by a conformal factor $\Omega^2(\tau, \tilde{\xi}) \equiv \exp(2a\tilde{\xi})$. Therefore, due to the conformal invariance, the action has the same form in the coordinates $(\tau, \tilde{\xi})$:

$$S[\phi] = \frac{1}{2} \int \left[(\partial_\tau \phi)^2 - (\partial_{\tilde{\xi}} \phi)^2 \right] d\tau\, d\tilde{\xi}.$$

The classical equations of motion in the laboratory frame and in the accelerated frame are

$$\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = 0; \quad \frac{\partial^2 \phi}{\partial \tau^2} - \frac{\partial^2 \phi}{\partial \tilde{\xi}^2} = 0,$$

with the general solutions

$$\phi(t, x) = A(t - x) + B(t + x), \quad \phi(\tau, \tilde{\xi}) = P(\tau - \tilde{\xi}) + Q(\tau + \tilde{\xi}).$$

Here A, B, P, and Q are arbitrary smooth functions. Note that a solution $\phi(t, x)$ representing a certain state of the field will be a very different function of τ and $\tilde{\xi}$.

We shall now quantize the field ϕ and compare the vacuum states in the laboratory frame and in the accelerated frame.

The procedure of quantization is formally the same in both coordinate systems (t, x) and $(\tau, \tilde{\xi})$. The mode expansion in the laboratory frame is found from Eq. (19) with the substitution $\omega_k = |k|$:

$$\hat{\phi}(t, x) = \int_{-\infty}^{+\infty} \frac{dk}{(2\pi)^{1/2}} \frac{1}{\sqrt{2\,|k|}} \left[e^{-i|k|t + ikx} \hat{a}_k^- + e^{i|k|t - ikx} \hat{a}_k^+ \right]. \tag{93}$$

The normalization factor $(2\pi)^{1/2}$ is used in 1+1 dimensions instead of the factor $(2\pi)^{3/2}$ used in 3+1 dimensions. The creation and annihilation operators $\hat{a}_k^\pm$ defined by Eq. (93) satisfy the usual commutation relations and describe particles moving with momentum k either in the positive x direction ($k > 0$) or in the negative x direction ($k < 0$).

The vacuum state in the laboratory frame (the **Minkowski vacuum**), denoted by $|0_M\rangle$, is the zero eigenvector of all the annihilation operators $\hat{a}_k^-$,

$$\hat{a}_k^- \,|0_M\rangle = 0 \ \text{ for all } k.$$

The mode expansion in the accelerated frame is quite similar to Eq. (93),

$$\hat{\phi}(\tau, \tilde{\xi}) = \int_{-\infty}^{+\infty} \frac{dk}{(2\pi)^{1/2}} \frac{1}{\sqrt{2\,|k|}} \left[e^{-i|k|\tau + ik\tilde{\xi}} \hat{b}_k^- + e^{i|k|\tau - ik\tilde{\xi}} \hat{b}_k^+ \right]. \tag{94}$$

Note that the mode expansions (93) and (94) are decompositions of the operator $\hat{\phi}(x,t)$ into linear combinations of two different sets of basis functions with operator-valued coefficients $\hat{a}_k^{\pm}$ and $\hat{b}_k^{\pm}$. So it is to be expected that the operators $\hat{a}_k^{\pm}$ and $\hat{b}_k^{\pm}$ are different, although they satisfy similar commutation relations.

The vacuum state in the accelerated frame $|0_R\rangle$ (the **Rindler vacuum**) is defined by

$$\hat{b}_k^{-} \, |0_R\rangle = 0 \text{ for all } k.$$

Since the operators $\hat{b}_k$ differ from $\hat{a}_k$, the Rindler vacuum $|0_R\rangle$ and the Minkowski vacuum $|0_M\rangle$ are two different quantum states of the field $\hat{\phi}$.

At this point, a natural question to ask is whether the state $|0_M\rangle$ or $|0_R\rangle$ is the "correct" vacuum. To answer this question, we need to consider the physical interpretation of the states $|0_M\rangle$ and $|0_R\rangle$ in a particular (perhaps imaginary) physical experiment. Let us imagine a hypothetical device for preparing the quantum field in the lowest-energy state; this device may work by pumping all energy, as much as possible, out of a certain volume. If mounted onto an accelerated spaceship, the device will prepare the field in the quantum state $|0_R\rangle$. Observers moving with the ship would agree that the field in the state $|0_R\rangle$ has the lowest possible energy and the Minkowski state $|0_M\rangle$ has a higher energy. Thus a particle detector at rest in the accelerated frame *will register particles* when the scalar field is in the state $|0_M\rangle$.

Neither of the two vacuum states is "more correct" if considered by itself, without regard for realistic physical conditions in the universe. Ultimately the choice of vacuum is determined by experiment: the correct vacuum state must be such that the theoretical predictions agree with the available experimental data. For instance, the spacetime near the Solar system is approximately flat (almost Minkowski), and we observe empty space that does not create any particles by itself. By virtue of this observation, we are justified to ascribe the vacuum state $|0_M\rangle$ to fields in the empty Minkowski spacetime. In particular, an accelerated observer moving through empty space will encounter fields in the state $|0_M\rangle$ and therefore will detect particles. This detection is a manifestation of the Unruh effect.

The rest of this chapter is devoted to a calculation relating the Minkowski frame operators $\hat{a}_k^{\pm}$ to the Rindler frame operators $\hat{b}_k^{\pm}$ through the appropriate Bogolyubov coefficients. This calculation will enable us to express the Minkowski vacuum as a superposition of excited states built on top of the Rindler vacuum and thus to compute the probability distribution for particle occupation numbers observed in the accelerated frame.

7.5 Lightcone mode expansions

It is convenient to introduce the lightcone coordinates[7]

$$\bar{u} \equiv t - x, \; \bar{v} \equiv t + x; \quad u \equiv \tau - \tilde{\xi}, \; v \equiv \tau + \tilde{\xi}.$$

The relation between the laboratory frame and the accelerated frame has a simpler form in lightcone coordinates: from Eq. (91) we find

$$\bar{u} = -a^{-1}e^{-au}, \quad \bar{v} = a^{-1}e^{av}, \tag{95}$$

so the metric is

$$ds^2 = d\bar{u}\, d\bar{v} = e^{a(v-u)} du\, dv.$$

[7] The chosen notation (u,v) for the lightcone coordinates in a uniformly accelerated frame and $(\bar{u},\bar{v})$ for the freely falling (unaccelerated) frame will be used in chapter 8 as well.

The field equations and their general solutions are also expressed more concisely in the lightcone coordinates:

$$\frac{\partial^2}{\partial\bar{u}\partial\bar{v}}\phi\left(\bar{u},\bar{v}\right)=0,\quad \phi\left(\bar{u},\bar{v}\right)=A\left(\bar{u}\right)+B\left(\bar{v}\right);$$
$$\frac{\partial^2}{\partial u\partial v}\phi(u,v)=0,\quad \phi(u,v)=P(u)+Q(v). \tag{96}$$

The mode expansion (93) can be rewritten in the coordinates $\bar{u},\bar{v}$ by first splitting the integration into the ranges of positive and negative k,

$$\hat{\phi}(t,x)=\int_{-\infty}^{0}\frac{dk}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\left|k\right|}}\left[e^{ikt+ikx}\hat{a}_k^-+e^{-ikt-ikx}\hat{a}_k^+\right]$$
$$+\int_{0}^{+\infty}\frac{dk}{(2\pi)^{1/2}}\frac{1}{\sqrt{2k}}\left[e^{-ikt+ikx}\hat{a}_k^-+e^{ikt-ikx}\hat{a}_k^+\right].$$

Then we introduce $\omega=|k|$ as the integration variable with the range $0<\omega<+\infty$ and obtain the **lightcone mode expansion**

$$\hat{\phi}\left(\bar{u},\bar{v}\right)=\int_{0}^{+\infty}\frac{d\omega}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\omega}}\left[e^{-i\omega\bar{u}}\hat{a}_\omega^-+e^{i\omega\bar{u}}\hat{a}_\omega^++e^{-i\omega\bar{v}}\hat{a}_{-\omega}^-+e^{i\omega\bar{v}}\hat{a}_{-\omega}^+\right]. \tag{97}$$

Lightcone mode expansions explicitly decompose the field $\hat{\phi}\left(\bar{u},\bar{v}\right)$ into a sum of functions of $\bar{u}$ and functions of $\bar{v}$. This agrees with Eq. (96) from which we find that $A(\bar{u})$ is a linear combination of the operators $\hat{a}_\omega^\pm$ with positive momenta ω, while $B(\bar{v})$ is a linear combination of $\hat{a}_{-\omega}^\pm$ with negative momenta $-\omega$:

$$\hat{\phi}\left(\bar{u},\bar{v}\right)=\hat{A}\left(\bar{u}\right)+\hat{B}\left(\bar{v}\right);$$
$$\hat{A}\left(\bar{u}\right)=\int_{0}^{+\infty}\frac{d\omega}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\omega}}\left[e^{-i\omega\bar{u}}\hat{a}_\omega^-+e^{i\omega\bar{u}}\hat{a}_\omega^+\right],$$
$$\hat{B}\left(\bar{v}\right)=\int_{0}^{+\infty}\frac{d\omega}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\omega}}\left[e^{-i\omega\bar{v}}\hat{a}_{-\omega}^-+e^{i\omega\bar{v}}\hat{a}_{-\omega}^+\right].$$

The lightcone mode expansion in the Rindler frame has exactly the same form except for involving the coordinates (u,v) instead of $(\bar{u},\bar{v})$. We use the integration variable Ω to distinguish the Rindler frame expansion from that of the Minkowski frame,

$$\hat{\phi}(u,v)=\hat{P}(u)+\hat{Q}(v)$$
$$=\int_{0}^{+\infty}\frac{d\Omega}{(2\pi)^{1/2}}\frac{1}{\sqrt{2\Omega}}\left[e^{-i\Omega u}\hat{b}_\Omega^-+e^{i\Omega u}\hat{b}_\Omega^++e^{-i\Omega v}\hat{b}_{-\Omega}^-+e^{i\Omega v}\hat{b}_{-\Omega}^+\right]. \tag{98}$$

As before, $\hat{P}(u)$ is expanded into operators $\hat{b}_\Omega^\pm$ with positive momenta Ω and $\hat{Q}(v)$ into the operators $\hat{b}_{-\Omega}^\pm$ with negative momenta $-\Omega$. (Note that the variables ω and Ω take only *positive* values. Also, the Rindler mode expansion is only valid within the domain $x>|t|$ covered by the Rindler frame; it is only within this domain that we can compare the two mode expansions.)

7.6 Bogolyubov transformations

The relation between the operators $\hat{a}_{\pm\omega}^\pm$ and $\hat{b}_{\pm\Omega}^\pm$, which we shall presently derive, is a Bogolyubov transformation of a more general form than that considered in Sec. 5.6.

Since the coordinate transformation (95) does not mix u and v, the identity

$$\hat{\phi}(u,v) = \hat{A}\left(\bar{u}(u)\right) + \hat{B}\left(\bar{v}(v)\right) = \hat{P}(u) + \hat{Q}(v)$$

entails two separate relations for u and for v,

$$\hat{A}\left(\bar{u}(u)\right) = \hat{P}(u), \quad \hat{B}\left(\bar{v}(v)\right) = \hat{Q}(v).$$

Comparing the expansions (97) and (98), we find that the operators $\hat{a}_{\omega}^{\pm}$ with positive momenta ω are expressed through $\hat{b}_{\Omega}^{\pm}$ with positive momenta Ω, while the operators $\hat{a}_{-\omega}^{\pm}$ are expressed through negative-momentum operators $\hat{b}_{-\Omega}^{\pm}$. In other words, there is no mixing between operators of positive and negative momentum. The relation $\hat{A}\left(\bar{u}\right) = \hat{P}(u)$ is then rewritten as

$$\begin{aligned}\hat{A}\left(\bar{u}\right) &= \int_0^{+\infty} \frac{d\omega}{(2\pi)^{1/2}} \frac{1}{\sqrt{2\omega}} \left[e^{-i\omega\bar{u}}\hat{a}_{\omega}^{-} + e^{i\omega\bar{u}}\hat{a}_{\omega}^{+}\right] \\ = \hat{P}(u) &= \int_0^{+\infty} \frac{d\Omega}{(2\pi)^{1/2}} \frac{1}{\sqrt{2\Omega}} \left[e^{-i\Omega u}\hat{b}_{\Omega}^{-} + e^{i\Omega u}\hat{b}_{\Omega}^{+}\right]. \end{aligned} \tag{99}$$

Here $\bar{u}$ is understood to be the function of u given by Eq. (95); both sides of Eq. (99) are equal as functions of u.

We can now express the positive-momentum operators $\hat{a}_{\omega}^{\pm}$ as explicit linear combinations of $\hat{b}_{\Omega}^{\pm}$. To this end, we perform the Fourier transform of both sides of Eq. (99) in u. The RHS yields

$$\int_{-\infty}^{+\infty} \frac{du}{\sqrt{2\pi}} e^{i\Omega u}\hat{P}(u) = \frac{1}{\sqrt{2\left|\Omega\right|}} \begin{cases} \hat{b}_{\Omega}^{-}, & \Omega > 0; \\ \hat{b}_{|\Omega|}^{+}, & \Omega < 0. \end{cases} \tag{100}$$

(The Fourier transform variable is denoted also by Ω for convenience.) The Fourier transform of the LHS of Eq. (99) yields an expression involving all $\hat{a}_{\omega}^{\pm}$,

$$\begin{aligned}\int_{-\infty}^{+\infty} \frac{du}{\sqrt{2\pi}} e^{i\Omega u}\hat{A}\left(\bar{u}\right) &= \int_0^{\infty} \frac{d\omega}{\sqrt{2\omega}} \int_{-\infty}^{+\infty} \frac{du}{2\pi} \left[e^{i\Omega u - i\omega\bar{u}}\hat{a}_{\omega}^{-} + e^{i\Omega u + i\omega\bar{u}}\hat{a}_{\omega}^{+}\right] \\ &\equiv \int_0^{\infty} \frac{d\omega}{\sqrt{2\omega}} \left[F(\omega,\Omega)\hat{a}_{\omega}^{-} + F(-\omega,\Omega)\hat{a}_{\omega}^{+}\right], \end{aligned} \tag{101}$$

where we introduced the auxiliary function[8]

$$F(\omega,\Omega) \equiv \int_{-\infty}^{+\infty} \frac{du}{2\pi} e^{i\Omega u - i\omega\bar{u}} = \int_{-\infty}^{+\infty} \frac{du}{2\pi} \exp\left[i\Omega u + i\frac{\omega}{a}e^{-au}\right]. \tag{102}$$

Comparing Eqs. (100) and (101) restricted to positive Ω, we find that the relation between $\hat{a}_{\omega}^{\pm}$ and $\hat{b}_{\Omega}^{-}$ is of the form

$$\hat{b}_{\Omega}^{-} = \int_0^{\infty} d\omega \left[\alpha_{\omega\Omega}\hat{a}_{\omega}^{-} + \beta_{\omega\Omega}\hat{a}_{\omega}^{+}\right], \tag{103}$$

where the coefficients $\alpha_{\omega\Omega}$ and $\beta_{\omega\Omega}$ are

$$\alpha_{\omega\Omega} = \sqrt{\frac{\Omega}{\omega}}F(\omega,\Omega), \quad \beta_{\omega\Omega} = \sqrt{\frac{\Omega}{\omega}}F(-\omega,\Omega); \quad \omega > 0, \Omega > 0. \tag{104}$$

[8] Because of the carelessly interchanged order of integration while deriving Eq. (101), the integral (102) diverges at $u \to +\infty$, so the definition of $F(\omega,\Omega)$ must be given more carefully if we desire full mathematical rigor.

The operators $\hat{b}^{+}_{\Omega}$ can be similarly expressed through $\hat{a}^{\pm}_{\omega}$ using the Hermitian conjugation of Eq. (103) and the identity

$$F^{*}(\omega,\Omega)=F(-\omega,-\Omega).$$

The relation (103) is a Bogolyubov transformation that mixes creation and annihilation operators with different momenta $\omega\neq\Omega$. In contrast, the Bogolyubov transformations considered in Sec. 5.6 are "diagonal," with $\alpha_{\omega\Omega}$ and $\beta_{\omega\Omega}$ proportional to $\delta(\omega-\Omega)$.

The relation between the operators $\hat{a}^{\pm}_{-\omega}$ and $\hat{b}^{\pm}_{-\Omega}$ is obtained from the equation $\hat{B}\left(\bar{v}\right)=\hat{Q}(v)$. We omit the corresponding straightforward calculations and concentrate on the modes with positive momentum; the results for negative momenta are completely analogous.

General Bogolyubov transformations

We now briefly consider the properties of a general Bogolyubov transformation,

$$\hat{b}^{-}_{\Omega}=\int_{-\infty}^{+\infty}d\omega\left[\alpha_{\omega\Omega}\hat{a}^{-}_{\omega}+\beta_{\omega\Omega}\hat{a}^{+}_{\omega}\right]. \tag{105}$$

The relation (103) is of this form except for the integration over $0<\omega<+\infty$ which is justified because the only nonzero Bogolyubov coefficients are those relating the momenta ω,Ω of equal sign, i.e. $\alpha_{-\omega,\Omega}=0$ and $\beta_{-\omega,\Omega}=0$. But for now we shall not limit ourselves to this case.

The relation for the operator $\hat{b}^{+}_{\Omega}$ is the Hermitian conjugate of Eq. (105).

> **Remark:** To avoid confusion in the notation, we always write the indices ω,Ω in the Bogolyubov coefficients in this order, i.e. $\alpha_{\omega\Omega}$, but never $\alpha_{\Omega\omega}$. In the calculations throughout this chapter, the integration is always over the first index ω corresponding to the momentum of a-particles.

Since the operators $\hat{a}^{\pm}_{\omega}$, $\hat{b}^{\pm}_{\Omega}$ satisfy the commutation relations

$$\left[\hat{a}^{-}_{\omega},\hat{a}^{+}_{\omega'}\right]=\delta(\omega-\omega'),\quad \left[\hat{b}^{-}_{\Omega},\hat{b}^{+}_{\Omega'}\right]=\delta(\Omega-\Omega'), \tag{106}$$

it is straightforward to check (by substituting Eq. (105) into the above relation for $\hat{b}^{\pm}_{\Omega}$) that the Bogolyubov coefficients are constrained by

$$\int_{-\infty}^{+\infty}d\omega\left(\alpha_{\omega\Omega}\alpha^{*}_{\omega\Omega'}-\beta_{\omega\Omega}\beta^{*}_{\omega\Omega'}\right)=\delta(\Omega-\Omega'). \tag{107}$$

This is analogous to the normalization condition $\left|\alpha_k\right|^2-\left|\beta_k\right|^2=1$ we had earlier.

Note that the origin of the δ function in Eq. (106) is the infinite volume of the entire space. If the field were quantized in a finite box of volume V, the momenta ω and Ω would be discrete and the δ function would be replaced by the ordinary Kronecker symbol times the volume factor, i.e. $\delta_{\Omega\Omega'}V$. The δ function in Eq. (107) has the same origin. Below we shall use Eq. (107) with $\Omega=\Omega'$ and the divergent factor $\delta(0)$ will be interpreted as the infinite spatial volume.

7.7 Density of particles

Since the vacua $\left|0_M\right\rangle$ and $\left|0_R\right\rangle$ corresponding to the operators $\hat{a}^{-}_{\omega}$ and $\hat{b}^{-}_{\Omega}$ are different, the a-vacuum is a state with b-particles and vice versa. We now compute the density of b-particles in the a-vacuum state.

The b-particle number operator is $\hat{N}_\Omega \equiv \hat{b}^+_\Omega \hat{b}^-_\Omega$, so the average b-particle number in the a-vacuum $|0_M\rangle$ is equal to the expectation value of $\hat{N}_\Omega$,

$$\begin{aligned}\langle \hat{N}_\Omega \rangle &\equiv \langle 0_M | \, \hat{b}^+_\Omega \hat{b}^-_\Omega \, |0_M\rangle \\ &= \langle 0_M | \int d\omega \left[\alpha^*_{\omega\Omega} \hat{a}^+_\omega + \beta^*_{\omega\Omega} \hat{a}^-_\omega \right] \int d\omega' \left[\alpha_{\omega'\Omega} \hat{a}^-_{\omega'} + \beta_{\omega'\Omega} \hat{a}^+_{\omega'} \right] |0_M\rangle \\ &= \int d\omega \, |\beta_{\omega\Omega}|^2 . \end{aligned} \tag{108}$$

This is the mean number of particles observed in the accelerated frame.

In principle one can explicitly compute the Bogolyubov coefficients $\beta_{\omega\Omega}$ defined by Eq. (104) in terms of the Γ function (see below). However, we only need to evaluate the RHS of Eq. (108) which involves an integral over ω, and we shall use a mathematical trick that allows us to compute just that integral and avoid other, more cumbersome calculations.

We first show that the function $F(\omega, \Omega)$ satisfies the identity

$$F(\omega, \Omega) = F(-\omega, \Omega) \exp\left(\frac{\pi\Omega}{a}\right), \quad \text{for } \omega > 0, \; a > 0. \tag{109}$$

Derivation of Eq. (109)

The function $F(\omega, \Omega)$ can be reduced to Euler's Γ function by changing the variable $u \to t$,

$$t \equiv -\frac{i\omega}{a} e^{-au}.$$

The result is

$$F(\omega, \Omega) = \frac{1}{2\pi a} \exp\left(i\frac{\Omega}{a} \ln\frac{\omega}{a} + \frac{\pi\Omega}{2a}\right) \Gamma\left(-\frac{i\Omega}{a}\right), \quad \omega > 0, \; a > 0.$$

However, it is not clear whether to take $\ln(-\omega) = \ln\omega + i\pi$ or some other phase instead of $i\pi$ in the above expression. To resolve this question, we need to analyze the required analytic continuation of the Γ function more carefully.

A direct approach (without using the Γ function) is to deform the contour of integration in Eq. (102). The contour can be shifted downwards by $-i\pi a^{-1}$ into the line $u = -i\pi a^{-1} + t$, where t is real, $-\infty < t < +\infty$ (see Fig. 6). Then $e^{-au} = -e^{-at}$ and we obtain

$$F(\omega, \Omega) = \int_{-\infty}^{+\infty} \frac{dt}{2\pi} \exp\left(i\Omega t + \frac{\pi\Omega}{a} - \frac{i\omega}{a} e^{-at}\right) = F(-\omega, \Omega) \exp\left(\frac{\pi\Omega}{a}\right).$$

It remains to justify the shift of the contour. The integrand has no singularities and, since the lateral lines have a limited length, it is enough to show that the integrand vanishes at $u \to \pm\infty - i\alpha$ for $0 < \alpha < \pi a^{-1}$. At $u = M - i\alpha$ and $M \to -\infty$ the integrand vanishes since

$$\lim_{u \to -\infty - i\alpha} Re\left(\frac{i\omega}{a} e^{-au}\right) = -\lim_{t \to -\infty} \frac{\omega}{a} e^{-at} \sin\alpha a = -\infty. \tag{110}$$

At $u \to +\infty - i\alpha$ the integral does not actually converge and must be regularized, e.g. by inserting a convergence factor $\exp\left(-bu^2\right)$ with $b > 0$:

$$F(\omega, \Omega) = \lim_{b \to +0} \int_{-\infty}^{+\infty} \frac{du}{2\pi} \exp\left(-bu^2 + i\Omega u + i\frac{\omega}{a} e^{-au}\right). \tag{111}$$

With this (or another) regularization, the integrand vanishes at $u \to +\infty - i\alpha$ as well. Therefore the contour may be shifted and our result is justified in the regularized sense.

Note that we cannot shift the contour to $u = -i(\pi + 2\pi n)a^{-1} + t$ with any $n \neq 0$ because Eq. (110) will not hold. Also, with $\omega < 0$ we will be unable to move the contour

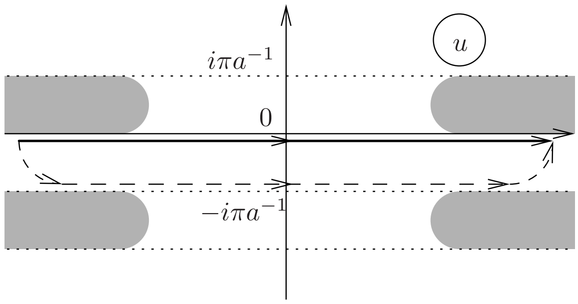

Figure 6: The original and the shifted contours of integration for Eq. (111) are shown by solid and dashed lines. The shaded regions cannot be crossed when deforming the contour at infinity.

in the negative imaginary direction. The shift of the contour we used is the only one possible.

We then substitute Eq. (104) into the normalization condition (107), use Eq. (109) and find

$$\delta(\Omega-\Omega') = \int_0^{+\infty} d\omega \frac{\sqrt{\Omega\Omega'}}{\omega}\left[F(\omega,\Omega)F^*(\omega,\Omega') - F(-\omega,\Omega)F^*(-\omega,\Omega')\right]$$
$$= \left[\exp\left(\frac{\pi\Omega+\pi\Omega'}{a}\right)-1\right]\int_0^{+\infty} d\omega \frac{\sqrt{\Omega\Omega'}}{\omega} F^*(-\omega,\Omega)F(-\omega,\Omega).$$

The last line above yields the relation

$$\int_0^{+\infty} d\omega \frac{\sqrt{\Omega\Omega'}}{\omega} F(-\omega,\Omega)F^*(-\omega,\Omega') = \left[\exp\left(\frac{2\pi\Omega}{a}\right)-1\right]^{-1}\delta(\Omega-\Omega'). \qquad (112)$$

Setting $\Omega' = \Omega$ in Eq. (112), we directly compute the integral in the RHS of Eq. (108),

$$\langle \hat{N}_\Omega\rangle = \int_0^{+\infty} d\omega \left|\beta_{\omega\Omega}\right|^2 = \int_0^{+\infty} d\omega \frac{\Omega}{\omega}\left|F(-\omega,\Omega)\right|^2 = \left[\exp\left(\frac{2\pi\Omega}{a}\right)-1\right]^{-1}\delta(0).$$

As usual, we expect $\langle \hat{N}_\Omega\rangle$ to be divergent since it is the total number of particles in the entire space. As discussed above, the divergent volume factor $\delta(0)$ represents the volume of space, and the remaining factor is the density n_Ω of b-particles with momentum Ω:

$$\int_0^{+\infty} d\omega \left|\beta_{\omega\Omega}\right|^2 \equiv n_\Omega \delta(0).$$

Therefore, the mean density of particles in the mode with momentum Ω is

$$n_\Omega = \left[\exp\left(\frac{2\pi\Omega}{a}\right)-1\right]^{-1}. \qquad (113)$$

This is the main result of this chapter.

So far we have computed n_Ω only for positive-momentum modes (with $\Omega > 0$). The result for negative-momentum modes is obtained by replacing Ω by $|\Omega|$ in Eq. (113).

7.8 Unruh temperature

A massless particle with momentum Ω has energy $E = |\Omega|$, so the formula (113) is equivalent to the Bose-Einstein distribution

$$n(E) = \left[\exp\left(\frac{E}{T}\right) - 1\right]^{-1}$$

where T is the **Unruh temperature**

$$T \equiv \frac{a}{2\pi}.$$

We found that an accelerated observer detects particles when the field $\hat{\phi}$ is in the Minkowski vacuum state $|0_M\rangle$. The detected particles may have any momentum Ω, although the probability for registering a high-energy particle is very small. The particle distribution (113) is characteristic of the thermal blackbody radiation with the temperature $T = a/2\pi$, where a is the magnitude of the proper acceleration (in Planck units). An accelerated detector behaves as though it were placed in a thermal bath with temperature T. This is the Unruh effect.

A physical interpretation of the Unruh effect as seen in the laboratory frame is the following. The accelerated detector is coupled to the quantum fields and perturbs their quantum state around its trajectory. This perturbation is very small but as a result the detector registers particles, although the fields were previously in the vacuum state. The detected particles are real and the energy for these particles comes from the agent that accelerates the detector.

Finally, we note that the Unruh effect is impossible to use in practice because the acceleration required to produce a measurable temperature is enormous. Here is an example calculation. Let us determine the acceleration corresponding to the Unruh temperature $T = 100°\mathrm{C}$; in that case, water will boil in an accelerated container due to the Unruh effect. We need to express all quantities in the SI units. The equation $T = a/(2\pi)$ becomes

$$kT = \frac{\hbar}{c}\frac{a}{2\pi}.$$

Here $k \approx 1.38 \cdot 10^{-23}\mathrm{J/K}$ is Boltzmann's constant. The boiling point of water is $T = 373\mathrm{K}$. The required acceleration is $a \sim 10^{22}\mathrm{m/s}^2$ which is clearly beyond any practical possibility. The Unruh effect is an extremely inefficient way to produce particles.

8 Hawking radiation

Classical general relativity describes black holes as massive objects with such a strong gravitational field that even light cannot escape their surface (the **black hole horizon**). However, quantum theory predicts that black holes emit particles moving away from the horizon. The particles are produced out of vacuum fluctuations of quantum fields present around the black hole. In effect, a black hole (BH) is not completely black but radiates a dim light as if it were an object with a low but nonzero temperature.

The main focus of this chapter is to compute the density of particles emitted by a static black hole, as registered by observers far away from the BH horizon.

8.1 Scalar field in Schwarzschild spacetime

We consider a scalar field in the presence of a single nonrotating black hole of mass M. The BH spacetime is described by the **Schwarzschild metric**,[9]

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2 \left(d\theta^2 + d\varphi^2 \sin^2\theta\right).$$

This metric is singular at $r = 2M$ which corresponds to the BH horizon, while for $r < 2M$ the coordinate t is spacelike and r is timelike. Therefore the coordinates (t, r) may be used with the normal interpretation of time and space only in the exterior region, $r > 2M$.

To simplify the calculations, we assume that the field ϕ is independent of the angular variables θ, φ and restrict our attention to a 1+1-dimensional section of the spacetime with the coordinates (t, r). The line element in 1+1 dimensions,

$$ds^2 = g_{ab} dx^a dx^b, \quad x^0 \equiv t,\ x^1 \equiv r,$$

involves the reduced metric

$$g_{ab} = \begin{bmatrix} 1 - \frac{2M}{r} & 0 \\ 0 & -\left(1 - \frac{2M}{r}\right)^{-1} \end{bmatrix}.$$

The theory we are developing is a toy model (i.e. a drastically simplified version) of the full 3+1-dimensional theory in the Schwarzschild spacetime. We expect that the main features of the full theory are preserved in the 1+1-dimensional model.

The action for a massless scalar field is

$$S[\phi] = \frac{1}{2} \int g^{ab} \phi_{,a} \phi_{,b} \sqrt{-g}\, d^2x.$$

As shown before, this action is in fact conformally invariant. Because of the conformal invariance, a significant simplification occurs if the metric is brought to a conformally flat form. This is achieved by a change of coordinates similar to that employed in previous section. Namely, we introduce a "conformal distance" r^* instead of r, where $r^* = f(r)$ and the function $f(r)$ is chosen such that

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 = A(r) \left[dt^2 - dr^{*2}\right],$$

where $A(r)$ is some function. It follows that we must choose

$$\frac{df(r)}{dr} = \left(1 - \frac{2M}{r}\right)^{-1}; \quad dr = \left(1 - \frac{2M}{r}\right) dr^*.$$

From this relation we find $r^*(r)$ up to an integration constant which we choose as $2M$ for convenience,

$$r^*(r) = r - 2M + 2M \ln\left(\frac{r}{2M} - 1\right). \tag{114}$$

The metric in the coordinates (t, r^*) is conformally flat,

$$ds^2 = \left(1 - \frac{2M}{r}\right) \left[dt^2 - dr^{*2}\right], \tag{115}$$

[9] In our notation here and below, the asimuthal angle is φ while the scalar field is ϕ.

where r must be expressed through r^* using Eq. (114). However, we shall not need an explicit formula for the function $r(r^*)$.

The coordinate $r^*(r)$ is defined only for $r > 2M$ and varies in the range $-\infty < r^* < +\infty$. It is called the "tortoise coordinate" because an object approaching the horizon $r = 2M$ needs to cross an infinite coordinate distance in r^*. From Eq. (115) it is clear that the tortoise coordinates (t, r^*) are asymptotically the same as the Minkowski coordinates (t, r) when $r \to +\infty$, i.e. in regions far from the black hole where the spacetime is almost flat.

The action for the scalar field in the tortoise coordinates is

$$S[\phi] = \frac{1}{2} \int \left[(\partial_t \phi)^2 - (\partial_{r^*} \phi)^2 \right] dt\, dr^*,$$

and the general solution of the equation of motion is of the form

$$\phi(t, r^*) = P(t - r^*) + Q(t + r^*),$$

where P and Q are arbitrary (but sufficiently smooth) functions.

In the lightcone coordinates (u, v) defined by

$$u \equiv t - r^*, \quad v \equiv t + r^*, \tag{116}$$

the metric is expressed as

$$ds^2 = \left(1 - \frac{2M}{r}\right) du\, dv. \tag{117}$$

Note that $r = 2M$ is a singularity where the metric becomes degenerate.

8.2 Kruskal coordinates

The coordinate system (t, r^*) has the advantage that for $r^* \to +\infty$ it asymptotically coincides with the Minkowski coordinate system (t, r) naturally defined far away from the BH horizon. However, the coordinates (t, r^*) do not cover the black hole interior, $r < 2M$. To describe the entire spacetime, we need another coordinate system.

It is a standard result that the singularity in the Schwarzschild metric (117) which occurs at $r = 2M$ is merely a **coordinate singularity** since a suitable change of coordinates yields a metric regular at the BH horizon. For instance, an observer freely falling into the black hole would see a normal, finitely curved space while crossing the horizon line $r = 2M$. Therefore one is motivated to consider a coordinate system $(\bar{t}, \bar{r})$ that describes the proper time $\bar{t}$ and the proper distance $\bar{r}$ measured by a freely falling observer. A suitable coordinate system is the **Kruskal frame**. We omit the construction of the Kruskal frame[10] and write only the final formulae. The Kruskal lightcone coordinates

$$\bar{u} \equiv \bar{t} - \bar{r}, \quad \bar{v} \equiv \bar{t} + \bar{r}$$

are related to the tortoise lightcone coordinates (116) by

$$\bar{u} = -4M \exp\left(-\frac{u}{4M}\right), \quad \bar{v} = 4M \exp\left(\frac{v}{4M}\right). \tag{118}$$

The parameters $\bar{u}$, $\bar{v}$ vary in the intervals

$$-\infty < \bar{u} < 0, \quad 0 < \bar{v} < +\infty. \tag{119}$$

[10]A detailed derivation can be found, for instance, in §31 of the book *Gravitation* by C.W. Misner, K. Thorne, and J. Wheeler (W. H. Freeman, San Francisco, 1973).

The inverse relation between $(\bar{u}, \bar{v})$ and the tortoise coordinates (t, r^*) is then found from Eqs. (114) and (118):

$$t = 2M \ln \left(-\frac{\bar{v}}{\bar{u}} \right),$$
$$\exp \left(-\frac{r^*}{2M} \right) = -\frac{\exp \left(1 - \frac{r}{2M} \right)}{1 - \frac{r}{2M}} = -\frac{16M^2}{\bar{u}\bar{v}}. \tag{120}$$

The BH horizon $r = 2M$ corresponds to the lines $\bar{u} = 0$ and $\bar{v} = 0$. To examine the spacetime near the horizon, we need to rewrite the metric in the Kruskal coordinates. With the substitution

$$u = -4M \ln \left(-\frac{\bar{u}}{4M} \right), \quad v = 4M \ln \frac{\bar{v}}{4M},$$

the metric (117) becomes

$$ds^2 = -\frac{16M^2}{\bar{u}\bar{v}} \left(1 - \frac{2M}{r} \right) d\bar{u}\, d\bar{v}.$$

Using Eqs. (114) and (120), after some algebra we obtain

$$ds^2 = \frac{2M}{r} \exp \left(1 - \frac{r}{2M} \right) d\bar{u}\, d\bar{v}, \tag{121}$$

where it is implied that the Schwarzschild coordinate r is expressed through $\bar{u}$ and $\bar{v}$ using the relation (120).

It follows from Eq. (121) that at $r = 2M$ the metric is $ds^2 = d\bar{u}\, d\bar{v}$, the same as in the Minkowski spacetime. Although the coordinates $\bar{u}, \bar{v}$ were originally defined in the ranges (119), there is no singularity at $\bar{u} = 0$ or at $\bar{v} = 0$, and therefore the coordinate system $(\bar{u}, \bar{v})$ may be extended to $\bar{u} > 0$ and $\bar{v} < 0$. Thus the Kruskal coordinates cover a larger patch of the spacetime than the tortoise coordinates (t, r^*). For instance, Eq. (120) relates r to $\bar{u}, \bar{v}$ also for $0 < r < 2M$, even though r^* is undefined for these r. Therefore, the Kruskal frame covers also the interior of the black hole.

Since the metric (121) is conformally flat, the action and the classical field equations for a conformally coupled field in the Kruskal frame have the same form as in the tortoise coordinates. For instance, the general solution for the field ϕ is $\phi(\bar{u}, \bar{v}) = A(\bar{u}) + B(\bar{v})$.

We note that Eq. (118) is similar to the definition (95) of the proper frame for a uniformly accelerated observer. The formal analogy is exact if we set $a \equiv (4M)^{-1}$. Note that a freely falling observer (with the worldline $\bar{r} =$ const) has zero proper acceleration. On the other hand, a spaceship remaining at a fixed position relative to the BH must keep its engine running at a constant thrust and thus has constant proper acceleration. To make the analogy with the Unruh effect more apparent, we chose the notation in which the coordinates $(\bar{u}, \bar{v})$ always refer to freely falling observers while the coordinates (u, v) describe accelerated frames.

8.3 Field quantization

In the previous section we introduced two coordinate systems corresponding to a locally inertial observer (the Kruskal frame) and a locally accelerated observer (the tortoise frame). Now we quantize the field $\phi(x)$ in these two frames and compare the respective vacuum states. The considerations are formally quite similar to those in Chapter 7.

To quantize the field $\phi(x)$, it is convenient to employ the lightcone mode expansions (defined in Sec. 7.5) in the coordinates (u, v) and $(\bar{u}, \bar{v})$. Because of the intentionally chosen notation, the relations (97) and (98) can be directly used to describe the quantized field $\hat{\phi}$ in the BH spacetime.

The lightcone mode expansion in the tortoise coordinates is

$$\hat{\phi}(u,v) = \int_0^{+\infty} \frac{d\Omega}{\sqrt{2\pi}} \frac{1}{\sqrt{2\Omega}} \left[e^{-i\Omega u} \hat{b}^-_\Omega + H.c. + e^{-i\Omega v} \hat{b}^-_{-\Omega} + H.c. \right],$$

where the "$H.c.$" denotes the Hermitian conjugate terms. The operators $\hat{b}^\pm_{\pm\Omega}$ correspond to particles detected by a stationary observer at a constant distance from the BH. The role of this observer is completely analogous to that of the uniformly accelerated observer considered in Sec. 7.1.

The lightcone mode expansion in the Kruskal coordinates is

$$\hat{\phi}\,(\bar{u},\bar{v}) = \int_0^{+\infty} \frac{d\omega}{\sqrt{2\pi}} \frac{1}{\sqrt{2\omega}} \left[e^{-i\omega\bar{u}} \hat{a}^-_\omega + H.c. + e^{-i\omega\bar{v}} \hat{a}^-_{-\omega} + H.c. \right].$$

The operators $\hat{a}^\pm_{\pm\omega}$ are related to particles registered by an observer freely falling into the black hole.

It is clear that the two sets of creation and annihilation operators $\hat{a}^\pm_{\pm\omega}$, $\hat{b}^\pm_{\pm\Omega}$ specify two different vacuum states, $|0_K\rangle$ ("Kruskal") and $|0_T\rangle$ ("tortoise"),

$$\hat{a}^-_{\pm\omega} \left|0_K\right\rangle = 0; \quad \hat{b}^-_{\pm\Omega} \left|0_T\right\rangle = 0.$$

Exactly as in the previous chapter, the operators $\hat{b}^\pm_{\pm\Omega}$ can be expressed through $\hat{a}^\pm_{\pm\omega}$ using the Bogolyubov transformation (103). The Bogolyubov coefficients are found from Eq. (104) if the acceleration a is replaced by $(4M)^{-1}$.

The correspondence between the Rindler and the Schwarzschild spacetimes is summarized in the following table. (We stress that this analogy is precise only for a conformally coupled field in 1+1 dimensions.)

Rindler	**Schwarzschild**
Inertial observers: vacuum $\|0_M\rangle$	Observers in free fall: vacuum $\|0_K\rangle$
Accelerated observers: $\|0_R\rangle$	Observers at $r=$ const: $\|0_T\rangle$
Proper acceleration a	Proper acceleration $(4M)^{-1}$
$\bar{u} = -a^{-1}\exp(-au)$	$\bar{u} = -4M\exp\left[-u/(4M)\right]$
$\bar{v} = a^{-1}\exp(av)$	$\bar{v} = 4M\exp\left[v/(4M)\right]$

8.4 Choice of vacuum

To find the expected number of particles measured by observers far outside of the black hole, we first need to make the correct choice of the quantum state of the field $\hat{\phi}$. In the present case, there are two candidate vacua, $|0_K\rangle$ and $|0_T\rangle$. We shall draw on the analogy with Sec. 7.4 to justify the choice of the Kruskal vacuum $|0_K\rangle$, which is the lowest-energy state for freely falling observers, as the quantum state of the field.

When considering a uniformly accelerated observer in the Minkowski spacetime, the correct choice of the vacuum state is $|0_M\rangle$, which is the lowest-energy state as measured by inertial observers. An accelerated observer registers this state as thermally excited. The other vacuum state, $|0_R\rangle$, can be physically realized by an accelerated vacuum preparation device occupying a very large volume of space. Consequently, the energy needed to prepare the field in the state $|0_R\rangle$ in the whole space is infinitely large. If one computes the mean energy density of the field $\hat{\phi}$ in the state $|0_R\rangle$, one finds (after subtracting the zero-point energy) that in the Minkowski frame the energy density diverges at the horizon.[11] On the other hand, the Minkowski vacuum state $|0_M\rangle$ has zero energy density everywhere.

It turns out that a similar situation occurs in the BH spacetime. At first it may appear that the field $\hat{\phi}$ should be in the state $|0_T\rangle$ which is the vacuum state selected by observers remaining at a constant distance from the black hole. However, the field $\hat{\phi}$ in the state $|0_T\rangle$ has an infinite energy density (after subtracting the zero-point energy) near the BH horizon.[12] Any energy density influences the metric via the Einstein equation. A divergent energy density indicates that the backreaction of the quantum fluctuations in the state $|0_T\rangle$ is so large near the BH horizon that the Schwarzschild metric is not a good approximation for the resulting spacetime. Thus the picture of a quantum field in the state $|0_T\rangle$ near an almost unperturbed black hole is inconsistent. On the other hand, the field $\hat{\phi}$ in the Kruskal state $|0_K\rangle$ has an everywhere finite and small energy density (when computed in the Schwarzschild frame after a subtraction of the zero-point energy). In this case, the backreaction of the quantum fluctuations on the metric is negligible. Therefore one has to employ the vacuum state $|0_K\rangle$ rather than the state $|0_T\rangle$ to describe quantum fields in the presence of a classical black hole.

Another argument for selecting the Kruskal vacuum $|0_K\rangle$ is the consideration of a star that turns into a black hole through the gravitational collapse. Before the collapse, the spacetime is almost flat and the initial state of quantum fields is the naturally defined Minkowski vacuum $|0_M\rangle$. It can be shown that the final quantum state of the field $\hat{\phi}$ after the collapse is the Kruskal vacuum.[13]

8.5 Hawking temperature

Observers remaining at $r = \text{const}$ far away from the black hole ($r \gg 2M$) are in an almost flat space where the natural vacuum state is the Minkowski one. The Minkowski vacuum at $r \gg 2M$ is approximately the same as the vacuum $|0_T\rangle$. Since the field $\hat{\phi}$ is in the Kruskal vacuum state $|0_K\rangle$, these observers would register the presence of particles.

The calculations of Sec. 7.7 show that the temperature measured by an accelerated observer is $T = a/(2\pi)$, and we have seen that the correspondence between the Rindler and the Schwarzschild cases requires to set $a = (4M)^{-1}$. It follows that observers at a fixed distance $r \gg 2M$ from the black hole detect a thermal spectrum of particles with the temperature

$$T_H = \frac{1}{8\pi M}. \tag{122}$$

This temperature is known as the **Hawking temperature**. (Observers staying closer to the BH will see a higher temperature due to the inverse gravitational redshift.)

Similarly, we find that the density of observed particles with energy $E = k$ is

$$n_E = \left[\exp\left(\frac{E}{T_H}\right) - 1\right]^{-1}.$$

[11] This result can be qualitatively understood if we recall that the Rindler coordinate $\tilde{\xi}$ covers an infinite range when approaching the horizon ($\tilde{\xi} \to -\infty$ as $\xi \to -a^{-1}$). The zero-point energy density in the state $|0_R\rangle$ is constant in the Rindler frame and thus appears as an infinite concentration of energy density near the horizon in the Minkowski frame; a subtraction of the zero-point energy does not cure this problem. We omit the detailed calculation, which requires a renormalization of the energy-momentum tensor of the quantum field.

[12] This is analogous to the divergent energy density near the horizon in the Rindler vacuum state. We again omit the required calculations.

[13] This was the pioneering calculation performed by S. W. Hawking.

This formula remains valid for massive particles with mass m and momentum k, after the natural replacement $E = \sqrt{m^2 + k^2}$. One can see that the particle production is significant only for particles with very small masses $m \lesssim T_H$.

The Hawking effect is in principle measurable, although the Hawking temperature for plausible astrophysical black holes is extremely small. For instance, a black hole of one solar mass $M = M_\odot = 2 \cdot 10^{30}$kg has the size of order 1km and the Hawking temperature $T_H \approx 6 \cdot 10^{-8}$K.

8.6 Black hole thermodynamics

In many situations, a static black hole of mass M behaves as a spherical body with radius $r = 2M$ and surface temperature T_H. According to the Stefan-Boltzmann law, a black body radiates the flux of energy

$$L = \gamma \sigma T_H^4 A,$$

where γ parametrizes the number of degrees of freedom available to the radiation, $\sigma = \pi^2/60$ is the Stefan-Boltzmann constant in Planck units, and

$$A = 4\pi R^2 = 16\pi M^2$$

is the surface area of the BH. The emitted flux determines the loss of energy due to radiation. The mass of the black hole decreases with time according to

$$\frac{dM}{dt} = -L = -\frac{\gamma}{15360\pi M^2}. \tag{123}$$

The solution with the initial condition $M|_{t=0} = M_0$ is

$$M(t) = M_0 \left(1 - \frac{t}{t_L}\right)^{1/3}, \quad t_L \equiv 5120\pi \frac{M_0^3}{\gamma}.$$

This calculation suggests that black holes are fundamentally unstable objects with the lifetime t_L during which the BH completely evaporates. A calculation shows that a black hole with a Solar mass $M = M_\odot = 2 \cdot 10^{30}$kg has a lifetime $t_L \sim 10^{74}$s, which is far greater than the age of the universe.

Laws of BH thermodynamics

Prior to the discovery of the BH radiation it was already known that black holes require a thermodynamical description involving a nonzero intrinsic entropy.

The entropy of a system is defined as the logarithm of the number of internal microstates of the system that are indistinguishable on the basis of macroscopically available information. Since the gravitational field of a static black hole is completely determined (both inside and outside of the horizon) by the mass M of the BH, one might expect that a black hole has only one microstate and therefore its entropy is zero. However, this conclusion is inconsistent with the second law of thermodynamics. A black hole absorbs all energy that falls onto it. If the black hole always had zero entropy, it could absorb some thermal energy and decrease the entropy of the world. This would violate the second law unless one assumes that the black hole has an intrinsic entropy that grows in the process of absorption.

Similar *gedanken* experiments involving classical general relativity and thermodynamics lead J. Bekenstein to conjecture in 1971 that a static black hole must have an intrinsic entropy S_{BH} proportional to the surface area $A = 16\pi M^2$. However, the coefficient of proportionality between S_{BH} and A could not be computed until the discovery of the Hawking radiation. The precise relation between the BH entropy and the horizon area follows from the first law of thermodynamics,

$$dE \equiv dM = T_H dS_{BH}, \tag{124}$$

where T_H is the Hawking temperature for a black hole of mass M. A simple calculation using Eq. (122) shows that

$$S_{BH} = 4\pi M^2 = \frac{1}{4}A. \tag{125}$$

A black hole of one solar mass has the entropy $S_\odot \sim 10^{76}$.

The thermodynamical law (124) suggests that in certain circumstances black holes behave as objects in thermal contact with their environment. This description applies to black holes surrounded by thermal radiation and to adiabatic processes of emission and absorption of heat.

For a complete thermodynamical description of black holes, one needs an equation of state. This is provided by the relation

$$E(T) = M = \frac{1}{8\pi T}.$$

It follows that the heat capacity of the BH is negative,

$$C_{BH} = \frac{\partial E}{\partial T} = -\frac{1}{8\pi T^2} < 0.$$

In other words, black holes become *colder* when they absorb heat. This unusual behavior leads to an instability of a BH surrounded by an *infinite* heat reservoir (i.e. a heat bath with a constant temperature). If the heat reservoir has a lower temperature $T < T_{BH}$, the BH would give heat to the reservoir and become even hotter. The process of evaporation will not be halted by the reservoir since its low temperature T remains constant. On the other hand, a black hole placed inside a reservoir with a higher temperature $T > T_{BH}$ will tend to absorb radiation from the reservoir and become colder. The process of absorption will continue indefinitely. In either case, no stable equilibrium is possible. A black hole can be stabilized with respect to absorption or emission of radiation only by a reservoir with a *finite* heat capacity.

The second law of thermodynamics now states that the combined entropy of all existing black holes and of all ordinary thermal matter never decreases,

$$\delta S_{total} = \delta S_{matter} + \sum_k \delta S_{BH}^{(k)} \geq 0.$$

Here $S_{BH}^{(k)}$ is the entropy (125) of the k-th black hole.

In classical general relativity it has been established that the combined area of all BH horizons cannot decrease in any interaction with classical matter (this is Hawking's "area theorem"). This statement applies not only to adiabatic processes but also to strongly out-of-equilibrium situations, such as a collision of black holes with the resulting merger. It is remarkable that this theorem, derived from a purely classical theory, assumes the form of the second law of thermodynamics when one considers quantum thermal effects of black holes. (The process of BH evaporation is not covered by the area theorem because it significantly involves quantum interactions.)

Moreover, there is a general connection between horizons and thermodynamics which has not yet been completely elucidated. The presence of a horizon in a spacetime means that a loss of information occurs, since one cannot observe events beyond the horizon. Intuitively, a loss of information entails a growth of entropy. It seems to be generally true in the theory of relativity that any event horizon behaves as a surface with a certain entropy and emits radiation with a certain temperature. For instance, the Unruh effect considered in Chapter 7 can be interpreted as a thermodynamical consequence of the presence of a horizon in the Rindler spacetime. A similar thermal effect (detection of particles in the Bunch-Davies vacuum state) is also present in de Sitter spacetime which also has a horizon.

These qualitative considerations conclude the present introductory course of quantum field theory in curved spacetime.

A Hilbert spaces and Dirac notation

Quantum operators such as $\hat{p}(t), \hat{q}(t)$ can be represented by linear transformations in suitable infinite-dimensional Hilbert spaces. In this section we summarize the properties of Hilbert spaces and also introduce the Dirac notation. We shall always consider complex vector spaces.

A.1 Infinite-dimensional vector spaces

A vector in a *finite*-dimensional space can be visualized as a collection of components, e.g. $\vec{a} \equiv (a_1, a_2, a_3, a_4)$, where each a_k is a (complex) number. To describe vectors in *infinite*-dimensional spaces, one must use infinitely many components. An important example of an infinite-dimensional complex vector space is the space L^2 of square-integrable functions, i.e. the set of all complex-valued functions $\psi(q)$ such that the integral

$$\int_{-\infty}^{+\infty} |\psi(q)|^2 \, dq$$

converges. One can check that a linear combination of two such functions, $\lambda_1\psi_1(q) + \lambda_2\psi_2(q)$, with constant coefficients $\lambda_{1,2}$, is again an element of the same vector space. A function $\psi \in L^2$ can be thought of as a set of infinitely many "components" $\psi_q \equiv \psi(q)$ with a continuous "index" q.

It turns out that the space of quantum states of a point mass in quantum mechanics is exactly the space L^2 of square-integrable functions $\psi(q)$, where q is the spatial coordinate of the particle. In that case the function $\psi(q)$ is called the **wavefunction**. Quantum states of a two-particle system belong to the space of functions $\psi(q_1, q_2)$, where $q_{1,2}$ are the coordinates of each particle. In quantum field theory, the "coordinates" are field configurations $\phi(x)$ and the wavefunction is a functional, $\psi\left[\phi(x)\right]$.

A.2 Dirac notation

Linear algebra is used in many areas of physics, and the Dirac notation is a convenient shorthand for calculations with vectors and linear operators. This notation is used for both finite- and infinite-dimensional vector spaces.

To denote a vector, Dirac proposed to write a symbol such as $|a\rangle$, $|x\rangle$, $|\lambda\rangle$, that is, a label inside the special brackets $|\rangle$. Linear combinations of vectors are written as $2\,|v\rangle - 3i\,|w\rangle$.

A linear operator $\hat{A} : V \to V$ acting in the space V transforms a vector $|v\rangle$ into the vector $\hat{A}\,|v\rangle$. (An operator $\hat{A}$ is **linear** if

$$\hat{A}\left(|v\rangle + \lambda\,|w\rangle\right) = \hat{A}\,|v\rangle + \lambda\hat{A}\,|w\rangle$$

for any $|v\rangle, |w\rangle \in V$ and $\lambda \in \mathbb{C}$.) For example, the identity operator $\hat{1}$ that does not change any vectors, $\hat{1}\,|v\rangle = |v\rangle$, is obviously a linear operator.

Linear forms acting on vectors, $f : V \to \mathbb{C}$, are **covectors** (vectors from the dual space) and are denoted by $\langle f|$. A linear form $\langle f|$ acts on a vector $|v\rangle$ and yields the number written as $\langle f|v\rangle$.

Usually a scalar product is defined in the space V. The scalar product of vectors $|v\rangle$ and $|w\rangle$ can be written as $(|v\rangle, |w\rangle)$ and is a complex number. The scalar product establishes a correspondence between vectors and covectors: each vector $|v\rangle$ defines a covector $\langle v|$ which is the linear map $|w\rangle \to (|v\rangle, |w\rangle)$. So the Dirac notation allows us to write scalar products somewhat more concisely as $(|v\rangle, |w\rangle) = \langle v|w\rangle$.

If $\hat{A}$ is a linear operator, the notation $\langle v|\,\hat{A}\,|w\rangle$ means the scalar product of the vectors $|v\rangle$ and $\hat{A}\,|w\rangle$. The quantity $\langle v|\,\hat{A}\,|w\rangle$ is also called the **matrix element** of the operator $\hat{A}$ with respect to the states $|v\rangle$ and $|w\rangle$.

The Dirac notation is convenient because the labels inside the brackets $|...\rangle$ are typographically separated from other symbols in a formula. So for instance one might denote specific vectors by $|0\rangle$, $|1\rangle$ (eigenvectors with integer eigenvalues), or by $|\psi\rangle$, $|a_i b_j\rangle$, or even by $\left|{}_{(\mathrm{out})}n_1, n_2, ...\right\rangle$, without risk of confusion. Note that the symbol $|0\rangle$ is the commonly used designation for the vacuum state, rather than the zero vector; the latter is denoted simply by 0.

If $|v\rangle$ is an eigenvector of an operator $\hat{A}$ with eigenvalue v, one writes

$$\hat{A}\,|v\rangle = v\,|v\rangle\,.$$

There is no confusion between the eigenvalue v (which is a number) and the vector $|v\rangle$ labeled by its eigenvalue.

A.3 Hermiticity

The scalar product in a complex vector space is **Hermitian** if $(\langle v|w\rangle)^* = \langle w|v\rangle$ for all vectors $|v\rangle$ and $|w\rangle$ (the asterisk * denotes the complex conjugation). In that case the **norm** $\langle v|v\rangle$ of a vector $|v\rangle$ is a real number.

A Hermitian scalar product allows one to define the **Hermitian conjugate** $\hat{A}^\dagger$ of an operator $\hat{A}$ via the identity

$$\langle v|\,\hat{A}^\dagger\,|w\rangle = \Big(\langle w|\,\hat{A}\,|v\rangle\Big)^*,$$

which should hold for all vectors $|v\rangle$ and $|w\rangle$. Note that an operator $\hat{A}^\dagger$ is uniquely specified if its matrix elements $\langle v|\,\hat{A}^\dagger\,|w\rangle$ with respect to all vectors $|v\rangle$, $|w\rangle$ are known. For example, it is easy to prove that $\hat{1}^\dagger = \hat{1}$.

The operation of Hermitian conjugation has the properties

$$(\hat{A}+\hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger; \quad (\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger; \quad (\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger.$$

An operator $\hat{A}$ is called **Hermitian** if $\hat{A}^\dagger = \hat{A}$, **anti-Hermitian** if $\hat{A}^\dagger = -\hat{A}$, and **unitary** if $\hat{A}^\dagger\hat{A} = \hat{A}\hat{A}^\dagger = \hat{1}$.

According to a postulate of quantum mechanics, the result of a measurement of some quantity is always an eigenvalue of the operator $\hat{A}$ corresponding to that quantity. Eigenvalues of a Hermitian operator are always real. This motivates an important assumption made in quantum mechanics: the operators corresponding to all observables are Hermitian.

> **Example:** The operators of position $\hat{q}$ and momentum $\hat{p}$ are Hermitian, $\hat{q}^\dagger = \hat{q}$ and $\hat{p}^\dagger = \hat{p}$. The commutator of two Hermitian operators $\hat{A}$, $\hat{B}$ is anti-Hermitian: $[\hat{A},\hat{B}]^\dagger = -[\hat{A},\hat{B}]$. Accordingly, the commutation relation for $\hat{q}$ and $\hat{p}$ contains the imaginary unit

i. The operator $\hat{p}\hat{q}$ is neither Hermitian nor anti-Hermitian: $(\hat{p}\hat{q})^\dagger = \hat{q}\hat{p} = \hat{p}\hat{q} + i\hbar\hat{1} \neq \pm\hat{p}\hat{q}$.

Eigenvectors of a Hermitian operator corresponding to different eigenvalues are always orthogonal. This is easy to prove: if $|v_1\rangle$ and $|v_2\rangle$ are eigenvectors of a Hermitian operator $\hat{A}$ with eigenvalues v_1 and v_2, then $v_{1,2}$ are real, so $\langle v_1|\hat{A} = v_1\langle v_1|$, and $\langle v_1|\hat{A}|v_2\rangle = v_2\langle v_1|v_2\rangle = v_1\langle v_1|v_2\rangle$. Therefore $\langle v_1|v_2\rangle = 0$ if $v_1 \neq v_2$.

A.4 Hilbert spaces

In an N-dimensional vector space one can find a finite set of basis vectors $|e_1\rangle$, ..., $|e_N\rangle$ such that any vector $|v\rangle$ is uniquely expressed as a linear combination

$$|v\rangle = \sum_{n=1}^{N} v_n |e_n\rangle .$$

The coefficients v_n are called **components** of the vector $|v\rangle$ in the basis $\{|e_n\rangle\}$. In an orthonormal basis satisfying $\langle e_m|e_n\rangle = \delta_{mn}$, the scalar product of two vectors $|v\rangle$, $|w\rangle$ is expressed through their components v_n, w_n as

$$\langle v|w\rangle = \sum_{n=1}^{N} v_n^* w_n .$$

By definition, a vector space is infinite-dimensional if no finite set of vectors can serve as a basis. In that case, one might expect to have an infinite basis $|e_1\rangle$, $|e_2\rangle$, ..., such that any vector $|v\rangle$ is uniquely expressible as an infinite linear combination

$$|v\rangle = \sum_{n=1}^{\infty} v_n |e_n\rangle . \qquad (126)$$

However, the convergence of this infinite series is a nontrivial issue. For instance, if the basis vectors $|e_n\rangle$ are orthonormal, then the norm of the vector $|v\rangle$ is

$$\langle v|v\rangle = \left(\sum_{m=1}^{\infty} v_m^* \langle e_n|\right)\left(\sum_{n=1}^{\infty} v_n |e_n\rangle\right) = \sum_{n=1}^{\infty} |v_n|^2 . \qquad (127)$$

This series must converge if the vector $|v\rangle$ has a finite norm, so the numbers v_n cannot be arbitrary. We cannot expect that e.g. the sum $\sum_{n=1}^{\infty} n^2 |e_n\rangle$ represents a well-defined vector. Now, if the coefficients v_n do fall off sufficiently rapidly so that the series (127) is finite, it may seem plausible that the infinite linear combination (126) converges and uniquely specifies the vector $|v\rangle$. However, this statement does not hold in all infinite-dimensional spaces. The required properties of the vector space are known in functional analysis as completeness and separability.[14]

A **Hilbert space** is a complete vector space with a Hermitian scalar product. When defining a quantum theory, one always chooses a separable Hilbert space as the space of quantum states. In that case, there exists a countable basis $\{|e_n\rangle\}$ and all vectors can be expanded as in Eq. (126). Once an orthonormal basis is chosen, all vectors $|v\rangle$ are unambiguously represented by collections $(v_1, v_2, ...)$ of their components. Therefore a separable Hilbert space can be visualized as the

[14] A normed vector space is **complete** if all norm-convergent (Cauchy) sequences in it converge to a limit; then all norm-convergent infinite sums always have an unique vector as their limit. The space is **separable** if there exists a countable set of vectors $\{|e_n\rangle\}$ which is everywhere dense in the space. Separability ensures that that all vectors can be approximated arbitrarily well by *finite* combinations of the basis vectors.

space of infinite rows of complex numbers, $|v\rangle \equiv (v_1, v_2, ...)$, such that the sum $\sum_{n=1}^{\infty} |v_n|^2$ converges. The convergence requirement guarantees that every scalar product $\langle v|w\rangle = \sum_{n=1}^{\infty} v_n^* w_n$ is finite.

> **Example:** The space $L^2[a,b]$ of square-integrable wave functions $\psi(q)$ defined on an interval $a < q < b$ is a separable Hilbert space, although it may appear to be "much larger" than the space of infinite rows of numbers. The scalar product of two wave functions $\psi_{1,2}(q)$ is defined by
>
> $$\langle\psi_1|\psi_2\rangle = \int_a^b \psi_1^*(q)\psi_2(q)dq.$$
>
> The canonical operators $\hat{p}$, $\hat{q}$ can be represented as linear operators in the space L^2 that act on functions $\psi(q)$ as
>
> $$\hat{p} : \psi(q) \to -i\hbar\frac{\partial\psi}{\partial q}, \quad \hat{q} : \psi(q) \to q\psi(q). \tag{128}$$
>
> It is straightforward to verify the commutation relation $[\hat{q}, \hat{p}] = i\hbar$.

B Mode expansions cheat sheet

We present a list of formulae relevant to mode expansions of free, real scalar fields. This should help resolve any confusion about the signs $\mathbf{k}$ and $-\mathbf{k}$ or similar technicalities.

All equations (except commutation relations) hold for operators as well as for classical quantities. The formulae for a field quantized in a box are obtained by replacing the factors $(2\pi)^3$ in the denominators with the volume V of the box. (Note that this replacement changes the physical dimension of the modes $\phi_{\mathbf{k}}$.)

$$\phi(\mathbf{x},t) = \int \frac{d^3\mathbf{k}\, e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}}\phi_{\mathbf{k}}(t); \quad \phi_{\mathbf{k}}(t) = \int \frac{d^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}}\phi(\mathbf{x},t)$$

$$a_{\mathbf{k}}^-(t) = \sqrt{\frac{\omega_k}{2}}[\phi_{\mathbf{k}} + \frac{i}{\omega_k}\pi_{\mathbf{k}}]; \quad a_{\mathbf{k}}^+(t) = \sqrt{\frac{\omega_k}{2}}[\phi_{-\mathbf{k}} - \frac{i}{\omega_k}\pi_{-\mathbf{k}}]$$

$$\phi_{\mathbf{k}}(t) = \frac{a_{\mathbf{k}}^-(t) + a_{-\mathbf{k}}^+(t)}{\sqrt{2\omega_k}}; \quad \pi_{\mathbf{k}}(t) = \sqrt{\frac{\omega_k}{2}}\frac{a_{\mathbf{k}}^-(t) - a_{-\mathbf{k}}^+(t)}{i}$$

Time-independent creation and annihilation operators $\hat{a}_{\mathbf{k}}^{\pm}$ are defined by

$$\hat{a}_{\mathbf{k}}^{\pm}(t) \equiv \hat{a}_{\mathbf{k}}^{\pm}\exp(\pm i\omega_k t)$$

Note that all $a_{\mathbf{k}}^{\pm}$ below are time-independent.

$$\phi^\dagger(x) = \phi(x); \quad (\phi_{\mathbf{k}})^\dagger = \phi_{-\mathbf{k}}; \quad \left(a_{\mathbf{k}}^-\right)^\dagger = a_{\mathbf{k}}^+$$

$$\pi(\mathbf{x},t) = \frac{d}{dt}\phi(\mathbf{x},t); \quad \pi_{\mathbf{k}}(t) = \frac{d}{dt}\phi_{\mathbf{k}}(t)$$

$$\left[\hat{\phi}(\mathbf{x},t),\, \hat{\pi}(\mathbf{x}',t)\right] = i\delta(\mathbf{x}-\mathbf{x}')$$
$$\left[\hat{\phi}_{\mathbf{k}}(t),\, \hat{\pi}_{\mathbf{k}'}(t)\right] = i\delta(\mathbf{k}+\mathbf{k}')$$
$$\left[\hat{a}_{\mathbf{k}}^-,\, \hat{a}_{\mathbf{k}'}^+\right] = \delta(\mathbf{k}-\mathbf{k}')$$

$$\hat{\phi}(\mathbf{x},t) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}}\frac{1}{\sqrt{2\omega_k}}\left[\hat{a}_{\mathbf{k}}^- e^{-i\omega_k t + i\mathbf{k}\cdot\mathbf{x}} + \hat{a}_{\mathbf{k}}^+ e^{i\omega_k t - i\mathbf{k}\cdot\mathbf{x}}\right]$$

Mode expansions may use anisotropic mode functions $v_{\mathbf{k}}(t)$. Isotropic mode expansions use scalar k instead of vector $\mathbf{k}$ because $v_{\mathbf{k}} \equiv v_k$ for all $|\mathbf{k}| = k$.

$$\hat{\phi}\left(\mathbf{x},t\right) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2}} \left[\hat{a}^-_{\mathbf{k}} v^*_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} + \hat{a}^+_{\mathbf{k}} v_{\mathbf{k}}(t) e^{-i\mathbf{k}\cdot\mathbf{x}}\right]$$

(Note: the factor $\sqrt{2}$ and the choice of $v^*_{\mathbf{k}}$ instead of $v_{\mathbf{k}}$ are for consistency with literature. This could have been chosen differently.)

$$v_{-\mathbf{k}} = v_{\mathbf{k}} \neq v^*_{\mathbf{k}}; \quad \ddot{v}_{\mathbf{k}} + \omega_k^2(t) v_{\mathbf{k}} = 0; \quad \dot{v}_{\mathbf{k}} v^*_{\mathbf{k}} - v_{\mathbf{k}} \dot{v}^*_{\mathbf{k}} = 2i$$

$$\phi_{\mathbf{k}}(t) = \frac{1}{\sqrt{2}} \left[a^-_{\mathbf{k}} v^*_{\mathbf{k}}(t) + a^+_{-\mathbf{k}} v_{\mathbf{k}}(t)\right]; \quad \pi_{\mathbf{k}}(t) = \frac{1}{\sqrt{2}} \left[a^-_{\mathbf{k}} \dot{v}^*_{\mathbf{k}}(t) + a^+_{-\mathbf{k}} \dot{v}_{\mathbf{k}}(t)\right]$$

Here the $a^\pm_{\mathbf{k}}$ are time-independent although $v_{\mathbf{k}}$ and $\phi_{\mathbf{k}}$, $\pi_{\mathbf{k}}$ depend on time:

$$a^-_{\mathbf{k}} = \frac{1}{i\sqrt{2}} \left[\dot{v}_{\mathbf{k}}(t)\phi_{\mathbf{k}}(t) - v_{\mathbf{k}}(t)\pi_{\mathbf{k}}(t)\right]; \quad a^+_{\mathbf{k}} = \frac{i}{\sqrt{2}} \left[\dot{v}^*_{\mathbf{k}}(t)\phi_{-\mathbf{k}}(t) - v^*_{\mathbf{k}}(t)\pi_{-\mathbf{k}}(t)\right]$$

Free scalar field mode functions in the flat space:

$$v_k(t) = \frac{1}{\sqrt{\omega_k}} e^{i\omega_k t}.$$

Bogolyubov transformations

Note: $\hat{a}^\pm_{\mathbf{k}}$ are defined by $v_{\mathbf{k}}(\eta)$ and $\hat{b}^\pm_{\mathbf{k}}$ are defined by $u_{\mathbf{k}}(\eta)$.

$$v^*_{\mathbf{k}}(\eta) = \alpha_{\mathbf{k}} u^*_{\mathbf{k}}(\eta) + \beta_{\mathbf{k}} u_{\mathbf{k}}(\eta); \quad \left|\alpha_{\mathbf{k}}\right|^2 - \left|\beta_{\mathbf{k}}\right|^2 = 1$$

$$\hat{b}^-_{\mathbf{k}} = \alpha_{\mathbf{k}} \hat{a}^-_{\mathbf{k}} + \beta^*_{\mathbf{k}} \hat{a}^+_{-\mathbf{k}}, \quad \hat{b}^+_{\mathbf{k}} = \alpha^*_{\mathbf{k}} \hat{a}^+_{\mathbf{k}} + \beta_{\mathbf{k}} \hat{a}^-_{-\mathbf{k}}$$
$$\alpha_{\mathbf{k}} = \alpha_{-\mathbf{k}}, \quad \beta_{\mathbf{k}} = \beta_{-\mathbf{k}}$$
$$\hat{a}^-_{\mathbf{k}} = \alpha^*_{\mathbf{k}} \hat{b}^-_{\mathbf{k}} - \beta^*_{\mathbf{k}} \hat{b}^+_{-\mathbf{k}}, \quad \hat{a}^+_{\mathbf{k}} = \alpha_{\mathbf{k}} \hat{b}^+_{\mathbf{k}} - \beta_{\mathbf{k}} \hat{b}^-_{-\mathbf{k}}$$