The Schwinger effect in inflationary cosmology

A project report
submitted in partial fulfillment for the award of the degree of
Bachelor of Technology
in
Engineering Physics
by

Amit Vikram Anand

under the guidance of
Dr. L. Sriramkumar

Department of Physics
Indian Institute of Technology Madras
Chennai 600036, India

June 2018

CERTIFICATE

This is to certify that the project titled **The Schwinger effect in inflationary cosmology** is a bona fide record of work done by **Amit Vikram Anand** towards the partial fulfillment of the requirements of the Bachelor of Technology degree in Engineering Physics at the Indian Institute of Technology, Madras, Chennai 600036, India.

(L. Sriramkumar, Project supervisor)

ACKNOWLEDGEMENTS

I thank **Dr. L. Sriramkumar** for guiding me through an exciting study of quantum field theory in curved spacetimes, the subject at the foundation of this project. I am grateful for his discussions on this subject as well as for his general advice. I have also greatly benefited from many useful discussions with professors and peers, and thank them all for the same. I would additionally like to acknowledge the courses offered by various departments, and the general environment at IIT Madras, which have made for a very interesting undergraduate experience.

ABSTRACT

We study the quantum theory of a complex scalar field in a background electromagnetic field in curved spacetime. We are particularly concerned with the creation of particles from the vacuum in such backgrounds (i.e. the Schwinger effect), and will use canonical quantization and Bogolubov transformations to describe this phenomenon. A discussion of some general features of particle production is followed by examples in both flat and curved spacetimes. In particular, we study the constant and Sauter pulsed electric field in Minkowski spacetime, and a constant magnitude electric field in de Sitter and power law expanding universes. We finally discuss an application of the curved spacetime Schwinger effect to early universe cosmology, in particular, a model of inflationary magnetogenesis.

Contents

Chapter 1

Introduction

1.1 An overview

'The Schwinger effect' (named after Julian Schwinger for his work on this subject, Ref. [1]) is a term used to describe various phenomena that involve the production of charged particle-antiparticle pairs from the vacuum by strong electric fields. Very closely related to this is the production of such particle-antiparticle pairs in nontrivial spacetimes, such as in expanding universes (the first detailed study being attributed[1] to Leonard Parker, e.g. Ref. [3]) and in the well known phenomenon of Hawking radiation from black holes (Ref. [4]). These phenomena are currently best described in the framework of quantum field theory.

Our interest in this report will be at the interface of these two classes of phenomena, i.e. particle production by electromagnetic fields in curved spacetime. One can think of various astrophysical and cosmological applications involving strong electric fields and gravitational fields, but our discussion will primarily lead up to applications to electromagnetic fields at cosmological scales in an expanding universe. In particular, in Chap. 7, we will discuss how the Schwinger effect plays a role in models of electromagnetic field generation by inflation in the early universe. While there are a number of ways of describing these phenomena using the methods of quantum field theory, we will use the method of canonical quantization with Bogolubov transformations (see e.g. Refs. [5, 6]).

The broad structure of this report is as follows: we review the classical theory of scalar fields in (reasonably general) electromagnetic and gravitational backgrounds (Chap. 2), and then quantize it using the method of canonical quantization (Chap. 3). This will be followed

[1] e.g. in Ref. [2].

by a general discussion of particle production in these backgrounds (Chap. 4). We will then consider examples in flat spacetime (Chap. 5) and curved spacetime (Chap. 6), always allowing for non-vanishing electric fields. This leads on to a brief discussion on the application of the curved spacetime Schwinger effect to inflationary magnetogenesis (Chap. 7). We conclude with a summary of our discussion (Chap. 8).

1.2 Notation and conventions

We will use the following notation for some standard sets of numbers: $\mathbb{N}$ for the set of positive integers/natural numbers, $\mathbb{N}_0$ for the set of nonnegative integers, $\mathbb{R}$ for the set of real numbers and $\mathbb{C}$ for the set of complex numbers. $\mathrm{L}^2(\mathcal{M})$ will denote the usual Hilbert space of square integrable functions on the domain $\mathcal{M}$. We will also have occasion to use a number of special functions as solutions of various second order linear differential equations, and our notation, definitions and nomenclature for these functions is intended to be consistent with Ref. [7]. We are typically interested in a continuum of values for the parameters occurring in these functions, and will ignore most 'special cases' for these functions corresponding to a discrete subset of these values where one may expect a limiting form of results to hold.

In this report, we are interested exclusively in a pseudo-Riemannian $(1+3)$-D spacetime (i.e. with 1 time dimension and 3 spatial dimensions), with x^μ (or sometimes x) describing points in this spacetime. We will work with the proper-time metric $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ with signature $(+,-,-,-)$, where $g_{\mu\nu}$ denotes the components of the metric tensor. We use $\bar{\eta}_{\mu\nu}$ to denote a diagonal matrix with diagonal entries $(+1,-1,-1,-1)$, which equals the components of the Minkowski metric in the standard Minkowski coordinates. As far as indices for components of various vectors are concerned, we use lowercase Greek letters (μ,ν...) for spacetime indices, uppercase Latin letters (I,J...) for internal space indices for scalar fields (in the relevant sections), and lowercase Latin letters (i,j...) for most other indices. Elements of $\mathbb{R}^3$ (representing spatial vectors) are occasionally denoted in bold font e.g. $\mathbf{x}$ or $\mathbf{k}$, rather than in terms of their components, where convenient.

We will work in 'natural' units[2] $\hbar = c = 1$. Our convention for the charge q of a particle is such that the electron, for example, would have a positive q, and the (classical, non-relativistic) force on a charged particle in an electric field $\mathbf{E}$ is given by $\mathbf{F} = -q\mathbf{E}$.

[2] G is irrelevant as we will not discuss the dynamics of the gravitational field.

Chapter 2

Charged scalar fields in curved spacetimes

We begin by studying the quantization of a scalar field theory in a globally hyperbolic[1] curved spacetime, interacting with both the gravitational field and a gauge field (the latter two will be termed as 'background fields'). For generality, we will first consider a scalar field with a general gauge symmetry and then immediately specialize to the complex scalar field with $U(1)$ gauge symmetry, which will occupy our interests thereafter.

We will neglect the dynamics of the gravitational and gauge fields, so that our system is essentially (mathematically equivalent to) that of a free scalar field evolving in a fixed background of these non-dynamical fields. This corresponds to the approximation that the background fields are in the 'classical' regime (perhaps due to interaction with another system outside our present considerations), where any backreaction on the background from the scalar field is negligible.

After discussing a convenient mathematical tool called the Klein-Gordon inner product, we will make use of it to express our scalar field theory in a form that readily lends itself to quantization.

2.1 The action for the scalar field

We will consider a scalar field $\phi^I(x)$ with an N-dimensional (real Euclidean) internal space, where $I \in \{1, ..., N\}$ denotes the internal space index. We assume an internal gauge symme-

[1] This is a technical requirement for having well defined initial value problems that we will not pay too much attention to. See e.g. Ref. [2].

try with generators $(T_k)^I{}_J$ (see e.g. Ref. [8]), and demand that any local gauge transformation,

$$\phi^I(x) \to \phi'^I(x) = \left(e^{iq\sum_k f_k(x)T_k}\right)^I{}_J \phi^J(x), \tag{2.1}$$

does not change observable physics. Here, q is the 'charge' of the scalar field corresponding to these generators, which is of relevance as it is possible to have different fields with different charges but a shared (part of an) internal space with the same gauge transformations. As the $f_k(x)$ are in general position dependent, we must express the dynamics of the field in terms of the gauge covariant derivative D_μ, which for an internal space vector field is given by

$$(\mathrm{D}_\mu\phi)^I(x) = \nabla_\mu\phi^I(x) - iq\sum_k A^k_\mu(x)(T_k)^I{}_J\phi^J(x), \tag{2.2}$$

where ∇_μ is the spacetime covariant derivative (the components of which reduce to coordinate derivative ∂_μ when directly acting on a scalar field). The gauge fields transform under gauge transformations as follows:

$$A^k_\mu(x) \to A'^k_\mu(x) = A^k_\mu(x) + \nabla_\mu f_k(x), \tag{2.3}$$

which ensures that the gauge covariant derivative also transforms like the scalar field, as an internal space vector under gauge transformations

$$(\mathrm{D}_\mu\phi)^I(x) \to (\mathrm{D}_\mu\phi)'^I(x) = \left(e^{iq\sum_k f_k(x)T_k}\right)^I{}_J (\mathrm{D}_\mu\phi)^J(x). \tag{2.4}$$

From the assumption of local invariance, the action for the scalar field (more precisely, its functional derivatives with respect to the field) must also be invariant under these transformations. We write it in terms of a Lagrangian (density) $\mathcal{L}$,

$$S[\phi^I; A^k_\mu; g_{\mu\nu}] = \int\limits_{\mathcal{R}} \mathrm{d}^4x\,\sqrt{-g}\,\mathcal{L}(\phi^I, \mathrm{D}_\mu\phi^I; A^k_\mu; g_{\mu\nu}), \tag{2.5}$$

where $g_{\mu\nu}$ is the metric tensor, and $\mathcal{R}$ is the spacetime region of interest. We note that $\mathcal{L}$ has a dependence on the gauge field and the metric tensor (and potentially their derivatives, which we will not consider - i.e. we are interested in the minimally coupled case). When considering the dynamics of the background fields, we will need to include appropriate terms quadratic in their derivatives via the gauge field strength and spacetime curvature tensors, but in the 'classical background' approximation we may ignore these terms.

Before we write an explicit form for $\mathcal{L}$, we consider two types of field observables of interest: the conserved currents corresponding to global gauge invariance and the stress energy tensor. To obtain the former, we consider an infinitesimal gauge transformation,

$$\phi'^I(x) = \phi^I(x) - iq \sum_k \varepsilon f_k(x) (T_k)^I{}_J \phi^J(x) \tag{2.6}$$

(to $O(\varepsilon)$), which being a symmetry of the action, gives the following conservation law from Noether's theorem:

$$\nabla_\mu \left(-iq \sum_k f_k (T_k)^I{}_J \phi^J \frac{\partial \mathcal{L}}{\partial (\mathrm{D}_\mu \phi)^I} \right) = 0. \tag{2.7}$$

Specifying $f_k = \delta_{ka}$ (a specific global gauge transformation) for each a gives the conservation law $\nabla_\mu j_a^\mu = 0$, where the conserved currents are

$$j_k^\mu = -iq \frac{\partial \mathcal{L}}{\partial (\mathrm{D}_\mu \phi)^I} (T_k)^I{}_J \phi^J. \tag{2.8}$$

We note that these quantities are locally gauge invariant.

The stress energy tensor (which couples to gravity, as opposed to the canonical stress energy tensor) is given by

$$\begin{aligned} T_{\mu\nu}(x) &= \frac{2}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}(x)} \\ &= \frac{2}{\sqrt{-g}} \frac{\partial \sqrt{-g}\mathcal{L}}{\partial g^{\mu\nu}} \\ &= 2 \frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} - g_{\mu\nu}\mathcal{L}, \end{aligned} \tag{2.9}$$

where after the first line, the right hand side is implicitly assumed to be evaluated at x.

Now, we return to the form of the Lagrangian. We will choose the Klein-Gordon Lagrangian, which is a quadratic form in the scalar field (including derivative operators), and therefore results in a linear theory:

$$\mathcal{L} = \frac{1}{2} g^{\mu\nu} \delta_{IJ} (\mathrm{D}_\mu \phi)^I (\mathrm{D}_\nu \phi)^J - \frac{1}{2} m^2 \delta_{IJ} \phi^I \phi^J. \tag{2.10}$$

Here, δ_{IJ} is the Euclidean metric of the internal space. m is the mass of the field. We could more generally add a gauge invariant nonlinear potential energy term $V(\delta_{IJ}\phi^I\phi^J)$, but will avoid doing so at present. The Euler-Lagrange equations, which are the classical equations of motion of the field, correspond to the Klein-Gordon equation

$$(\mathrm{D}_\mu \mathrm{D}^\mu \phi)^I + m^2 \phi^I = 0 \tag{2.11}$$

(to interpret the double derivative, we note that $(\mathrm{D}_\mu\phi)^I$ is a vector both in spacetime and the internal space, as opposed to ϕ which is a spacetime scalar and an internal space vector, and the second spacetime covariant derivative must act accordingly). As noted above, this equation is linear in ϕ^I, which makes quantization relatively straightforward.

The currents and stress energy tensor for this Lagrangian are given by

$$j^\mu_k = -\frac{1}{2} i q g^{\mu\nu} \delta_{IL} (T_k)^L{}_J (\mathrm{D}_\nu\phi)^I \phi^J \tag{2.12}$$

and

$$T_{\mu\nu} = \delta_{IJ}(\mathrm{D}_\mu\phi)^I(\mathrm{D}_\nu\phi)^J - \frac{1}{2} g_{\mu\nu}\left(g^{\alpha\beta}\delta_{IJ}(\mathrm{D}_\alpha\phi)^I(\mathrm{D}_\beta\phi)^J - m^2\delta_{IJ}\phi^I\phi^J\right) \tag{2.13}$$

respectively.

2.2 The Hamiltonian framework

To apply the well-known procedure for canonical quantization, we must first express the scalar field theory in the Hamiltonian framework (see e.g. Ref. [9] for an analogous treatment in flat spacetime). To do this, we introduce a continuous family of non-intersecting spacelike hypersurfaces[2] $\Sigma(t)$ in spacetime, $t \in \mathbb{R}$ being a timelike parameter. We also introduce 'spatial' coordinates $r = (r^i) = (r^1, r^2, r^3) \in \mathbb{R}^3$ bijectively on each $\Sigma(t)$, such that the spatial coordinates are also continuous in t i.e.

$$\lim_{t\to t'} x^\mu(t, r_0) - x^\mu(t', r_0) = 0, \tag{2.14}$$

for all $r_0 \in \mathbb{R}^3$ and $t, t' \in \mathbb{R}$, where $x^\mu(t, r)$ denotes the point mapped to r on $\Sigma(t)$. Thus, (t, r) is a coordinate system in spacetime, in which we will denote the components of the metric tensor as $\tilde{g}_{\mu\nu}$. We will continue to use $g_{\mu\nu}$ for the metric tensor components in an arbitrary system.

For future use, we note that each surface $\Sigma(t)$ has a timelike unit normal vector field, $n^\mu(x), \forall x \in \Sigma(t)$, satisfying $g_{\mu\nu}n^\mu n^\nu = 1$; by extension, there is a unique timelike unit normal vector at each point in spacetime corresponding to this choice of spacelike surfaces.

It is convenient to define a vector field t^μ representing an instantaneous 'flow' along t at constant r:

$$t^\mu(t, r) = \lim_{\delta t\to 0} \frac{x^\mu(t + \delta t, r) - x^\mu(t, r)}{\delta t}. \tag{2.15}$$

[2]See e.g. Ref. [10] for a discussion of hypersurfaces

For a suitable set of surfaces $\Sigma(t)$ (with a suitable (local) scaling of the parameter t), it is possible to have $t^\mu(t,r) = n^\mu(x^\rho(t,r))\forall(t,r)$ i.e. the timelike flow vectors are the normals to these surfaces. The most important simplification resulting from this choice is that $\sqrt{-\tilde{g}}$ is now the determinant of the induced metric on $\Sigma(t)$, as the metric in the (t,r) coordinate system now has components $\tilde{g}_{tt} = 1$ and $\tilde{g}_{tr} = 0$. The volume element on $\Sigma(t)$ is therefore $\mathrm{d}^3r\sqrt{-\tilde{g}}$. We will often find it convenient to choose such a set of surfaces.

The field theory can now be interpreted as the classical dynamics of generalized coordinates $\left[q^I(r)\right](t) = \phi^I(x^\mu(t,r))$ as a function of t. This requires choosing a region $\mathcal{R} = \bigcup_{t\in[t_2,t_1]}\Sigma(t)$, between 'times' t_2 and t_1, so that the values of the field at the boundary $\partial\mathcal{R}$ may be directly translated to that of the $q^I(r)$ at t_2 and t_1. Defining general velocities the conventional way,

$$[q^I_t(r)](t) = \frac{\mathrm{d}q^I(r)}{\mathrm{d}t}(t) = t^\mu\partial_\mu\phi^I(t,r). \tag{2.16}$$

This is given in terms of the coordinate/spacetime derivative. But we can instead define a generalized velocity in terms of the gauge covariant derivative,

$$[\mathrm{D}_t q^I(r)](t) = t^\mu(\mathrm{D}_\mu\phi)^I(t,r). \tag{2.17}$$

Using this velocity does not alter the result of the variational principle. With this interpretation, we rewrite the action as

$$S[q^I(r); A^k_\mu; g_{\mu\nu}] = \int_{t_2}^{t_1} \mathrm{d}t\; L\left(q^I(r), \mathrm{D}_t q^I(r); A^k_\mu; g_{\mu\nu}\right). \tag{2.18}$$

To express L in terms of $\mathcal{L}$, we may switch to the coordinate system (t,r). Then, $\int \mathrm{d}^4x\sqrt{-g} = \int \mathrm{d}t\mathrm{d}^3r\sqrt{-\tilde{g}}$, and $\mathcal{L}$ is coordinate independent, leading to

$$L = \int \mathrm{d}^3r\, \sqrt{-\tilde{g}}\, \mathcal{L}. \tag{2.19}$$

It follows that the momenta corresponding to the $q^I(r)$ are given by (using functional derivatives as r is a continuous parameter, with d^3r as the integration measure)

$$p_I(r) = \frac{\delta L}{\delta \mathrm{D}_t q^I(r)} = \sqrt{-\tilde{g}}\frac{t_\mu}{g_{\alpha\beta}t^\alpha t^\beta}\frac{\partial\mathcal{L}}{\partial(\mathrm{D}_\mu\phi)^I}. \tag{2.20}$$

Using this, the Hamiltonian can be defined by the standard Legendre transform,

$$
\begin{aligned}
H(q^I(r), p_I(r); A^k_\mu; g_{\mu\nu}) &= \left(\int \mathrm{d}^3 r \;\; p_I(r)\mathrm{D}_t q^I(r)\right) - L(q_I(r), \mathrm{D}_t q^I(r); A^k_\mu; g_{\mu\nu}) \\
&= \int \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \, \left(\frac{p_I(r)}{\sqrt{-\tilde{g}}}\mathrm{D}_t q^I(r) - \mathcal{L}\right) \\
&= \int \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \frac{t_\mu t^\nu}{g_{\alpha\beta}t^\alpha t^\beta} \left(\frac{\partial \mathcal{L}}{\partial (\mathrm{D}_\mu \phi)^I}(\mathrm{D}_\nu \phi)^I - \delta^\mu_\nu \mathcal{L}\right) \\
&= \int \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \frac{t^\mu t^\nu}{g_{\alpha\beta}t^\alpha t^\beta} \\
&\left(\frac{1}{2}\frac{\partial \mathcal{L}}{\partial (\mathrm{D}^\mu \phi)^I}(\mathrm{D}_\nu \phi)^I + \frac{1}{2}\frac{\partial \mathcal{L}}{\partial (\mathrm{D}^\nu \phi)^I}(\mathrm{D}_\mu \phi)^I - \delta^\mu_\nu \mathcal{L}\right) \\
&= \int \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \frac{t^\mu t^\nu}{g_{\alpha\beta}t^\alpha t^\beta} T_{\mu\nu},
\end{aligned} \tag{2.21}
$$

which shows that that this Hamiltonian is related to the (symmetric) stress energy tensor of the field.

Now, consider two dynamical variables

$$A(q^I(r), p_I(r)) = \int \mathrm{d}^3 r \; \mathcal{A}(q^I(r), p_I(r); r), \tag{2.22}$$

$$B(q^I(r), p_I(r)) = \int \mathrm{d}^3 r \; \mathcal{B}(q^I(r), p_I(r); r), \tag{2.23}$$

which are functionals of the generalized coordinates and momenta with their own density functions $\mathcal{A}$ and $\mathcal{B}$. We define their Poisson bracket as the dynamical variable given by

$$[A, B]_{\mathrm{PB}} = \int \mathrm{d}^3 r \; \sum_I \left(\frac{\delta A}{\delta q^I(r)}\frac{\delta B}{\delta p_I(r)} - \frac{\delta A}{\delta p_I(r)}\frac{\delta B}{\delta q^I(r)}\right). \tag{2.24}$$

The Poisson brackets happen to be independent of any particular choice of generalized coordinates and momenta. Considered as an operation, it satisfies a number of useful properties, including antisymmetry and multilinearity in its arguments, distributivity over multiplication and the Jacobi identity (see, for instance, Ref. [11]).

The Poisson brackets between generalized coordinates and momenta are of special importance in both the study of canonical transformations and, more importantly for our purposes, canonical quantization. They are given by

$$[q^I(r), q^J(r')]_{\mathrm{PB}} = 0, \tag{2.25}$$

$$[p_I(r), p_J(r')]_{\mathrm{PB}} = 0, \tag{2.26}$$

$$[q^I(r), p_J(r')]_{\mathrm{PB}} = \delta^I_J \delta(r - r'). \tag{2.27}$$

We also note that the law of motion in the Hamiltonian picture can be expressed as a rather simple equation:

$$t^\mu \mathrm{D}_\mu A = [A, H]_{\mathrm{PB}}. \tag{2.28}$$

We will return to these Poisson brackets when quantizing the complex scalar field, in Sec. 3.1.

2.3 The complex scalar field and $U(1)$ gauge invariance

The $U(1)$ complex scalar field is of particular interest as it is the simplest example of a charged field and is also straightforward to quantize using the procedure of canonical quantization.

The internal space of the simplest $U(1)$ invariant scalar field has 2 real dimensions i.e. $N = 2$, and the field has two real components, namely ϕ^1 and ϕ^2. The $U(1)$ group has a single generator, $T^I{}_J$ (and therefore a single gauge field $A_\mu(x)$). As a general element of the group is given by $U(\theta) = e^{i\theta T}$, the requirement of unitarity on $U(\theta)$ translates to the requirement that T be Hermitian. Additionally, as the transformations preserve $(\phi^1)^2 + (\phi^2)^2$, we require that $T^I{}_J$ be an antisymmetric matrix. The only 2×2 antisymmetric Hermitian matrix (up to scaling) is the y-Pauli matrix σ_y. Thus, we choose

$$T = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{2.29}$$

Now, we define two new complex fields,

$$\phi = \phi^1 + i\phi^2, \tag{2.30}$$

$$\phi^* = \phi^1 - i\phi^2. \tag{2.31}$$

We have $T\phi = \phi$, $T\phi^* = -\phi^*$ which means that $T = r_z$ in the (ϕ, ϕ^*) representation i.e. T is diagonal. Consequentially, the covariant derivative is also diagonal in this representation:

$$\mathrm{D}_\mu \phi(x) = \nabla_\mu \phi(x) - iqA_\mu(x)\phi(x), \tag{2.32}$$

$$\mathrm{D}_\mu \phi^*(x) = \nabla_\mu \phi^*(x) + iqA_\mu(x)\phi^*(x). \tag{2.33}$$

It is then easily seen that the Klein-Gordon equation, which involves the covariant derivative as its only nontrivial internal space operator, is also diagonalized:

$$\left((\nabla_\mu - iqA_\mu)(\nabla^\mu - iqA^\mu) + m^2\right)\phi = 0, \tag{2.34}$$

$$\left((\nabla_\mu + iqA_\mu)(\nabla^\mu + iqA^\mu) + m^2\right)\phi^* = 0. \tag{2.35}$$

Going over into the complex scalar field description is therefore essentially a trick to diagonalize the Klein-Gordon equation for a 2-component $U(1)$-gauge invariant scalar field. On the other hand, the internal space metric is now necessarily off-diagonal in the (ϕ, ϕ^*) representation,

$$[\delta_{IJ}]_{(\phi,\phi^*)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.36}$$

i.e. $\delta_{IJ} \to (\sigma_x)_{IJ}$. This also means that finding the dual of an internal space vector amounts to taking the complex conjugate in this representation.

We can now express physical quantities of interest in terms of the complex scalar field. The Lagrangian density becomes:

$$\begin{aligned}\mathcal{L} &= g^{\mu\nu}\left((\nabla_\mu + iqA_\mu)\phi^*\right)\left((\nabla_\nu - iqA_\nu)\phi\right) - m^2\phi^*\phi \\ &= g^{\mu\nu}(\mathrm{D}_\mu\phi)^*(\mathrm{D}_\nu\phi) - m^2\phi^*\phi.\end{aligned} \tag{2.37}$$

The conserved current due to global $U(1)$ gauge invariance is

$$j^\mu = -iq(\phi(\mathrm{D}^\mu\phi)^* - \phi^*(\mathrm{D}^\mu\phi)), \tag{2.38}$$

and the stress energy tensor is

$$T_{\mu\nu} = (\mathrm{D}_\mu\phi)^*\mathrm{D}_\nu\phi + \mathrm{D}_\mu\phi(\mathrm{D}_\nu\phi)^* - g_{\mu\nu}\mathcal{L}. \tag{2.39}$$

The covariant derivative can be separated into a component along the local normal vector $n^\mu(x)$, and the component orthogonal to it, respectively given by

$$\mathrm{D}_t = n^\mu \mathrm{D}_\mu, \tag{2.40}$$

$$\mathrm{S}_\mu = \mathrm{D}_\mu - g_{\mu\nu}n^\nu n^\rho \mathrm{D}_\rho, \tag{2.41}$$

the latter of which satisfies $n^\mu \mathrm{S}_\mu = 0$ at all points. This decomposition relies only on the existence of the vector field $n^\mu(x)$, without requiring that this field actually corresponds to space-like surfaces. The covariant derivative itself can be expressed in terms of these operators as follows:

$$\mathrm{D}_\mu = g_{\mu\nu}n^\nu \mathrm{D}_t + \mathrm{S}_\mu. \tag{2.42}$$

We may now simplify the Hamiltonian Eq. (2.21) at a time t (i.e. on the surface $\Sigma(t)$) using D_t and S_μ, expressing it as

$$H(t) = \int\limits_{\Sigma(t)} \mathrm{d}^3r\,\sqrt{-\tilde{g}}\,\left((\mathrm{D}_t\phi)^*\mathrm{D}_t\phi - (\mathrm{S}^\mu\phi)^*\mathrm{S}_\mu\phi + m^2\phi^*\phi\right). \tag{2.43}$$

Integrating the second term by parts, assuming that the field vanishes (or an equivalent boundary condition) at spatial infinity, we get the more useful form

$$H(t) = \int\limits_{\Sigma(t)} \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \, \left((\mathrm{D}_t \phi)^* \mathrm{D}_t \phi + \phi^* \left(\mathrm{S}^\mu \mathrm{S}_\mu + m^2 \right) \phi \right). \tag{2.44}$$

2.3.1 An aside: Identifying electric and magnetic field backgrounds

Now, we will just clarify some terminology pertaining to the $U(1)$ gauge field, which will come in frequent use later when we discuss the behaviour of the scalar field in different backgrounds. In particular, we will discuss what we mean by an 'electric' or 'magnetic' field in curved spacetimes. The terminology of course has its origins in the historical development of the classical theory of electromagnetism in flat space and time, but it is not as natural in a curved spacetime. Here, we will discuss one way of consistently using this terminology.

The most natural gauge invariant entity one can construct from the $U(1)$ gauge field is the electromagnetic field strength tensor $F_{\mu\nu}$ (which is antisymmetric in its indices). One way of defining it is based on the action of the commutator of gauge covariant derivatives on the scalar field,

$$[\mathrm{D}_\mu, \mathrm{D}_\nu]\phi = -iqF_{\mu\nu}\phi. \tag{2.45}$$

From this we can obtain the explicit expression

$$F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu. \tag{2.46}$$

To identify the electric and magnetic field parts, we will first consider the nomenclature used in Minkowski spacetime. With coordinates (t, x, y, z), these are given by

$$E_i = F_{ti}, \; B_i = \epsilon_{ijk} F_{jk}. \tag{2.47}$$

In analogy with these, we can define electric and magnetic fields in curved spacetimes if there is a family of timelike vectors such as n^μ available (though they may be arbitrarily chosen), as follows (see e.g. Ref. [12]):

$$E_\mu = F_{\mu\nu} n^\nu, \; B_\mu = \frac{1}{2} \epsilon_{\mu\nu\alpha\beta} n^\nu F^{\alpha\beta}, \tag{2.48}$$

where $\epsilon_{\mu\nu\alpha\beta}$ is totally antisymmetric with $\epsilon_{0123} = \sqrt{-g}$. We clearly have $n^\mu E_\mu = 0$ and $n^\mu B_\mu = 0$, making the electric and magnetic fields so defined vectors in the tangent spaces (see e.g. Ref. [10]) of the $\Sigma(t)$.

2.4 The Klein-Gordon inner product

It is worthwhile to briefly study what is known as the Klein-Gordon inner product and its properties as an indefinite inner product, independent of its connection to the dynamics of the scalar field. In this section, we will use $\mathcal{M}$ to denote spacetime. We will also implicitly treat continuous parameters as a limiting case of discrete parameters (for example, we will argue about a Hilbert space as if it were an n-dimensional vector space, with $n \to \infty$, without explicitly stating so). We begin with the Hilbert space of complex functions on spacetime, $\mathrm{L}^2(\mathcal{M})$ (which have an implicit interpretation as being in the space of the gauge field ϕ), with the standard inner product

$$\langle f_1, f_2 \rangle = \int \mathrm{d}^4x \sqrt{-g(x)}\; f_2^*(x) f_1(x). \tag{2.49}$$

Consider a spacelike surface Σ like the ones used in the Hamiltonian formulation, with coordinates r. We define a Hermitian operator $\eta_{\mathrm{KG}}^{\Sigma}$ (via its matrix elements) as a function of Σ, with $n^\mu(r)$ being the unit future directed normal vector field to Σ, and $h^{\Sigma}_{r_i,r_j}$ the (r-coordinate system) components of the induced metric on Σ:

$$\begin{aligned}\eta_{\mathrm{KG}}^{\Sigma}(x_2, x_1) = i \int_{\Sigma} \mathrm{d}^3r \sqrt{|h^{\Sigma}|}\; n^\mu \Big[&\delta(x_2 - x(t,r)) \mathrm{D}_\mu|_{x(t,r)} \delta(x(t,r) - x_1) \\ &- \left(\mathrm{D}_\mu^*|_{x_1} \delta(x_2 - x(t,r))\right) \delta(x(t,r) - x_1) \Big].\end{aligned} \tag{2.50}$$

Here, the δ functions are with respect to the spacetime volume measure $\mathrm{d}^4x\sqrt{-g}$. To explain the conjugation on the covariant derivatives, we note that x_1 must be interpreted as a conjugate index (in the space of ϕ^*) and x_2 as an index in the space of ϕ, so that η_{KG} maps the space of ϕ to itself.

The action of the above operator may be used to define a new kind of product of functions (essentially the corresponding matrix element), which we will call the KG product for now:

$$\langle f_1, f_2\rangle_{\mathrm{KG}}^{\Sigma} = \langle \eta_{\mathrm{KG}}^{\Sigma} f_1, f_2\rangle \tag{2.51}$$
$$= -i \int\limits_{\Sigma} \mathrm{d}^3 r\ \sqrt{|h^{\Sigma}|} n^{\mu} \left(f_1 (\mathrm{D}_{\mu} f_2)^* - f_2^* (\mathrm{D}_{\mu} f_1)\right). \tag{2.52}$$

The operator $\eta_{\mathrm{KG}}^{\Sigma}$ is indefinite (as in neither positive nor negative (semi-)definite) as $\langle f, f\rangle_{\mathrm{KG}}$ of functions f ranges across all of $\mathbb{R}$, including negative values. In other words, the eigenvalues of $\eta_{\mathrm{KG}}^{\Sigma}$ are not restricted to positive/nonnegative real numbers or negative/nonpositive real numbers alone.

We would like to interpret $\langle f_1, f_2\rangle_{\mathrm{KG}}^{\Sigma}$ as an (indefinite) inner product, for later convenience. This will allow us to use much of the machinery associated with inner products, particularly orthonormal bases, to simplify our study of the dynamics and canonical quantization of the scalar field.

A map $\langle \cdot, \cdot\rangle_{\mathrm{I}} : \mathrm{V} \times \mathrm{V} \to \mathbb{C}$ on a vector space V defined over the complex numbers $\mathbb{C}$ is called an indefinite inner product if it satisfies the following defining properties, $\forall\, f_a, f_b, f_c \in \mathrm{V},\ \alpha, \beta \in \mathbb{C}$ (see, for instance, Ref. [13])

- Conjugate symmetry[3]:

$$\langle f_a, f_b\rangle_{\mathrm{I}} = \langle f_b, f_a\rangle_{\mathrm{I}}^*. \tag{2.53}$$

- Linearity:

$$\langle \alpha f_a + \beta f_b, f_c\rangle_{\mathrm{I}} = \alpha\langle f_a, f_c\rangle_{\mathrm{I}} + \beta\langle f_b, f_c\rangle_{\mathrm{I}}. \tag{2.54}$$

- Nondegeneracy:

$$(\langle f_a, f\rangle_{\mathrm{I}} = 0\ \forall\ f \in \mathrm{V}) \implies f_a = 0. \tag{2.55}$$

It is trivial to show that the map given by $\langle f_1, f_2\rangle_{\mathrm{KG}}^{\Sigma}$ satisfies the conditions of antisymmetry and linearity. However, we may show that it is not nondegenerate - a function f_a such that it is not identically zero in $\mathcal{M}$ but satisfies $f_a(\Sigma) = 0$, $n^{\mu}\mathrm{D}_{\mu} f_a(\Sigma) = 0$ must also satisfy $\langle f_a, f\rangle_{\mathrm{KG}}^{\Sigma} = 0\ \forall\ f \in \mathrm{L}^2(\mathcal{M})$. Therefore, $\langle f_1, f_2\rangle_{\mathrm{KG}}^{\Sigma}$ is not an indefinite inner product on $\mathrm{L}^2(\mathcal{M})$. However, we will find it useful to consider instead a subspace $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M}) \subset \mathrm{L}^2(\mathcal{M})$ consisting only of those functions that are solutions to a second order linear (partial) differential

[3]This is referred to as 'antisymmetry' in Ref. [13].

equation (which we will just call the PDE) in $\mathcal{M}$. These functions are completely specified by boundary values for the function and its normal derivative on a Cauchy surface such as Σ i.e. by fixing $f(x \in \Sigma)$ and $n^\mu \mathrm{D}_\mu f(x \in \Sigma)$.

We may represent a function $f \in \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$ by two functions $u_f, v_f : \Sigma \to \mathbb{C}$ which are uniquely related to f via the boundary conditions

$$f(x \in \Sigma) = u_f(x), \tag{2.56}$$

$$n^\mu \mathrm{D}_\mu f(x \in \Sigma) = v_f(x). \tag{2.57}$$

We will use the notation $f \leftrightarrow (u_f, v_f)$ for such a representation. Evidently, from the linearity of the boundary conditions, the null (additive identity) element of $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$ is given by $0 \leftrightarrow (0,0)$ (which need not in the most general case necessarily be the same as $0 \in \mathrm{L}^2(\mathcal{M})$, which is the function that is identically zero everywhere. For example, consider the $(0,0)$ function in $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$ for the gauge potential $A_\mu = \alpha\delta^0_\mu$ - this implies a non-vanishing time derivative ∇_t of $(0,0)$ at Σ, which could be a non-vanishing solution of the unspecified PDE).

Now, we redefine the product $\langle f_1, f_2\rangle^\Sigma_{\mathrm{KG}}$ to apply only within the subspace of solutions to the differential equation i.e. $\langle\cdot,\cdot\rangle^\Sigma_{\mathrm{KG}} : \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M}) \times \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M}) \to \mathbb{C}$ by

$$\langle f_1, f_2\rangle^\Sigma_{\mathrm{KG}} = \langle \eta^\Sigma_{\mathrm{KG}} f_1, f_2\rangle \; \forall f_1, f_2 \in \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M}). \tag{2.58}$$

The formal expression Eq. (2.52) itself does not change, just the domain of validity. In the (u_f, v_f) representation, we have

$$\langle f_1, f_2\rangle^\Sigma_{\mathrm{KG}} = -i \int\limits_\Sigma \mathrm{d}^3 r\, \sqrt{|h^\Sigma|}\left(u_{f_1} v^*_{f_2} - u^*_{f_2} v_{f_1}\right). \tag{2.59}$$

For this product, antisymmetry and linearity continue to hold, as before. But this restricted product also satisfies non-degeneracy. We show this explicitly by first assuming $f_1 \leftrightarrow (u_{f_1}, v_{f_1})$ to be given, and choosing $f_2 \leftrightarrow (u_{f_2}, v_{f_2}) = \tilde{f}_1 \leftrightarrow (iv_{f_1}, -iu_{f_1})$ i.e. $u_{f_2} = iv_{f_1}$, $v_{f_2} = -iu_{f_1}$). This may always be done because u,v may be any of the most general (sufficiently well-behaved) complex valued functions on Σ. In that case,

$$\langle f_1, \tilde{f}_1\rangle^\Sigma_{\mathrm{KG}} = \int\limits_\Sigma \mathrm{d}^3 r\, \sqrt{|h^\Sigma|}\left(|u_{f_1}|^2 + |v_{f_1}|^2\right). \tag{2.60}$$

This is always positive and therefore nonzero unless $f_1 = 0 \in \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$. It is also trivial to show that $\langle 0, f\rangle^\Sigma_{\mathrm{KG}} = 0 \;\forall\; f \in \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$. Thus, we have succeeded in constructing a function which has a non-vanishing KG product with any given function other than 0 in the space of solutions of the PDE, and 0 has a vanishing KG product with all functions in this space, implying that the KG product is non-degenerate in $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$.

The KG-product is therefore an indefinite inner product in $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$, which we will call the Klein-Gordon inner product (which is as of now still Σ-dependent). Therefore (from Proposition 2.2.2 and the subsequent discussion in Ref. [13]) we may construct a complete orthonormal basis $\{u(k;x),\ \forall\, k\}$ in $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$ in the sense of the Klein-Gordon inner product, such that

$$\langle u(k;x), u(k';x)\rangle^{\Sigma}_{\mathrm{KG}} = \pm\delta(k-k'), \tag{2.61}$$

where $\delta(k-k')$ is a Dirac delta function associated with a suitable integration measure $\widetilde{\mathrm{d}k}$, satisfying

$$\delta(k-k') = 0 \text{ for } k \neq k', \tag{2.62}$$

$$\int \widetilde{\mathrm{d}k}\ \delta(k-k') = 1. \tag{2.63}$$

We will call this the u-basis, for short. In such a basis, the 'matrix' elements of the Klein-Gordon indefinite inner product are either $+\delta(k-k)$ (or $+1$ in a discrete case) and $-\delta(k-k)$. We also note that in any L^2 space, any two elements f_1,f_1' are treated as equivalent if the standard norm of their difference vanishes, $\langle f_1 - f_1', f_1 - f_1'\rangle = 0$ (without requiring pointwise equality), and it is also in this sense that a basis is required to be complete - any element of such a space is equivalent to some linear combination of the basis elements.

Transforming to an orthogonal basis in the Klein-Gordon inner product sense involves diagonalizing η_{KG}. To normalize the basis vectors so that they are also orthonormal involves subsequently scaling each of them by a complex number $c(k)$ (say), a transformation that leaves the $\mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M}) \to \mathrm{L}^2_{\mathrm{PDE}}(\mathcal{M})$ linear operator $\eta^{\Sigma}_{\mathrm{KG}}$ invariant (due to an operator A transforming as SAS^{-1} under a vector transformation S) but scales the 'matrix elements' of the standard Euclidean 'metric tensor' (which gives the inner product) by $|c(k)|^{-2}$ so that the inner product itself is preserved. The 'matrix elements' of the Klein-Gordon inner product, which is given by a contraction of the standard metric tensor with $\eta^{\Sigma}_{\mathrm{KG}}$ (see Eq. (2.51)), are therefore scaled by the positive number $|c(k)|^{-2}$. Thus, this shows that the $+\delta(k-k)$ entries along the diagonal of the matrix representation of the Klein-Gordon inner product correspond to positive eigenvalues of $\eta^{\Sigma}_{\mathrm{KG}}$, and the $-\delta(k-k)$ entries to negative eigenvalues. We also note that if 0 were an eigenvalue, then non-degeneracy would not be satisfied.

It is convenient to define two complementary subsets $K^+(u)$ and $K^-(u)$ of the space of values of k corresponding respectively to the positive and negative norm elements in the u-basis as follows:

$$K^{\pm}(u) = \{k : \langle u(k;x), u(k;x)\rangle_{\mathrm{KG}}^{\Sigma} = \pm\delta(k-k)\}. \tag{2.64}$$

Now, we consider the subspace $\mathrm{L}^2_{K^+(u)}(\mathcal{M}) \subset \mathrm{L}^2_{\mathrm{PDF}}(\mathcal{M})$ spanned by the positive norm basis vectors $\{u(k;x) : k \in K^+(u)\}$ and $\mathrm{L}^2_{K^-(u)}(\mathcal{M}) \subset \mathrm{L}^2_{\mathrm{PDF}}(\mathcal{M})$ spanned by the negative norm vectors $\{u(k;x) : k \in K^-(u)\}$ (not to be confused with exhaustive sets of positive norm and negative norm vectors respectively, which do not form vector spaces by themselves - these subspaces are analogous to lower dimensional surfaces in the vector space). Because the positive norm basis modes have non-vanishing norms themselves and any nonzero linear combination of them must have a non-vanishing inner product with at least one of them, there is no element of $\mathrm{L}^2_{K^+(u)}(\mathcal{M})$ other than 0 which has a vanishing inner product with all modes in the subspace, making it non-degenerate. A similar reasoning holds for the subspace spanned by the negative norm modes. The Klein-Gordon inner product is therefore also an inner product in these subspaces.

This brings us to a useful result (which is a special case of Proposition 2.2.3 in Ref. [13]). As each of the vectors in $\mathrm{L}^2_{K^+(u)}(\mathcal{M})$ corresponds to a positive eigenvalue of $\eta^{\Sigma}_{\mathrm{KG}}$,

$$\int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\ \langle u(k;x), u(k;x)\rangle_{\mathrm{KG}}^{\Sigma} = n_+\left(\eta^{\Sigma}_{\mathrm{KG}}\right), \tag{2.65}$$

where the quantity on the right hand side is the number of positive eigenvalues of $\eta^{\Sigma}_{\mathrm{KG}}$ (we will treat it as a well-defined quantity, following our implicit discrete interpretation, though it is infinite in the limit). Similarly, we have

$$\int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\ \langle u(k;x), u(k;x)\rangle_{\mathrm{KG}}^{\Sigma} = n_-\left(\eta^{\Sigma}_{\mathrm{KG}}\right), \tag{2.66}$$

with $n_-\left(\eta^{\Sigma}_{\mathrm{KG}}\right)$ being the number of negative eigenvalues.

As the eigenvalues of a Hermitian operator are an invariant set (independent of the choice of basis vectors), the quantities in Eqs. (2.65) and (2.66) are basis-independent. Then, for any two orthonormal bases, say a u-basis and a v-basis, we must have

$$\int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\ \langle u(k;x), u(k;x)\rangle^{\Sigma}_{\mathrm{KG}} = \int\limits_{K^+(v)} \widetilde{\mathrm{d}k}\ \langle v(k;x), v(k;x)\rangle^{\Sigma}_{\mathrm{KG}}, \tag{2.67}$$

$$\int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\ \langle u(k;x), u(k;x)\rangle^{\Sigma}_{\mathrm{KG}} = \int\limits_{K^-(v)} \widetilde{\mathrm{d}k}\ \langle v(k;x), v(k;x)\rangle^{\Sigma}_{\mathrm{KG}}. \tag{2.68}$$

This result will find application in Sec. 3.2, where it will be used to establish that particle production by a combined electromagnetic and gravitational field background must indeed create equal numbers of particles and antiparticles.

2.5 Solutions to the Klein-Gordon equation

Now, we will study in some more detail the equation of motion for the complex scalar field, the Klein-Gordon equation Eq. (2.34), reproduced here in more concise notation:

$$\left(\mathrm{D}_\mu \mathrm{D}^\mu + m^2\right)\phi = 0. \tag{2.69}$$

This equation, as commented earlier in the general case, is linear in ϕ - any linear combination of solutions is also a solution. This immediately suggests that its solutions may be considered elements of an infinite dimensional vector space (as with any linear differential equation). For this equation, the Klein-Gordon inner product is a particularly natural inner product to use, as we will now see.

2.5.1 Uniqueness of the Klein-Gordon inner product

Consider any two (complex-valued) solutions $f(x)$,$g(x)$ to Eq. (2.69). The Klein-Gordon inner product Eq. (2.52) for these two solutions, given a spacelike surface Σ, is

$$\langle f, g\rangle^{\Sigma}_{\mathrm{KG}} = -i \int\limits_{\Sigma} \mathrm{d}^3 r\ \sqrt{|h^\Sigma|} n^\mu \left(f(\mathrm{D}_\mu g)^* - g^*(\mathrm{D}_\mu f)\right). \tag{2.70}$$

Now, we proceed to show that the inner product is actually independent of the choice of Σ. This means that there is a unique Klein-Gordon inner product for solutions of the Klein-Gordon equation, and the question of the choice of Σ to be made, so that one may use the properties of the inner product (Sec. 2.4) to simplify the problem at hand, becomes irrelevant.

We first write the Klein-Gordon equation as satisfied by f and g, multiply the former by g^*, conjugate the latter and then multiply it by f and take the difference (with a factor of i thrown in) to get:

$$-i\left(f(\mathrm{D}_\mu \mathrm{D}^\mu g)^* - g^*(\mathrm{D}_\mu \mathrm{D}^\mu f)\right) = 0. \tag{2.71}$$

Using the product rule for gauge covariant derivatives (schematically $\mathrm{D}(fg) = (\mathrm{D}f)g + f(\mathrm{D}g)$, which can also be explicitly seen to hold by expanding it in terms of the gauge field), we get a divergence equation

$$\nabla_\mu\left(-i\left(f(\mathrm{D}_\mu g)^* - g^*(\mathrm{D}_\mu f)\right)\right) = 0. \tag{2.72}$$

As an aside, we note that for $g = f$, this reduces to the conservation of j^μ in Eq. (2.38) i.e. $\nabla_\mu j^\mu = 0$, which was initially derived from Noether's theorem.

Now, we consider a finite region $\mathcal{R}$. Integrating Eq. (2.72) over $\mathcal{R}$, using Gauss' theorem to reduce it to a boundary integral gives

$$-i\int\limits_{\partial\mathcal{R}} \mathrm{d}S_{\partial\mathcal{R}} n^\mu_{\partial\mathcal{R}}\left(f(\mathrm{D}_\mu g)^* - g^*(\mathrm{D}_\mu f)\right) = 0, \tag{2.73}$$

where $\mathrm{d}S_{\partial\mathcal{R}}$ represents the surface volume elements and $n^\mu_{\partial\mathcal{R}}$ the outward unit normals (i.e. normalized so that $|g_{\mu\nu}n^\mu n^\nu| = 1$, whether spacelike or timelike) to the boundary $\partial\mathcal{R}$ of $\mathcal{R}$. If $\partial\mathcal{R}$ were made up of $\Sigma(t_1)$ and $\Sigma(t_2)$, which may have no boundary themselves (e.g. a single period in problems with spatial periodicity) or may be taken together with a boundary/spatial infinity where all solutions and/or their normal derivatives vanish (so this corresponds to a range of boundary value problems), then we get (recalling that the n^μ are future directed, but the $n^\mu_{\partial\mathcal{R}}$ are directed 'outward' instead, and are past directed for $\Sigma(\min(t_1, t_2))$)

$$-i\int\limits_{\Sigma(t_1)} \mathrm{d}^3r\sqrt{|h^\Sigma|}\, n^\mu\left(f(\mathrm{D}_\mu g)^* - g^*(\mathrm{D}_\mu f)\right) = -i\int\limits_{\Sigma(t_2)} \mathrm{d}^3r\sqrt{|h^\Sigma|}\, n^\mu\left(f(\mathrm{D}_\mu g)^* - g^*(\mathrm{D}_\mu f)\right), \tag{2.74}$$

which establishes that the number $\langle f, g\rangle^\Sigma_{\mathrm{KG}}$ is independent of the choice of Σ, and we will henceforth treat it as a unique, invariant inner product of the two functions (and call it the Klein-Gordon inner product), and denote it as $\langle f, g\rangle_{\mathrm{KG}}$.

2.5.2 Orthonormal mode expansions

Now, we consider an orthonormal (in the Klein-Gordon sense) basis in the space of solutions of the Klein-Gordon equation, $\{u(k;x),\ \forall\, k\}$ (sometimes called modes, in analogy with waves), with $K^+(u)$ and $K^-(u)$ denoting the set of values of k for which the Klein-Gordon norm of the corresponding basis function $u(k;x)$ is positive and negative respectively. We will write the orthonormality relation as

$$\langle u(k;x), u(k';x)\rangle_{\text{KG}} = s_u(k)\delta(k-k'), \tag{2.75}$$

where

$$s_u(k) = \begin{cases} +1, & \text{for } k \in K^+(u) \\ -1, & \text{for } k \in K^-(u) \end{cases} . \tag{2.76}$$

Any solution configuration of the field ϕ can be expressed as a linear combination of the basis functions as follows:

$$\phi(x) = \int \widetilde{\mathrm{d}k}\ \varphi_u(k) u(k;x). \tag{2.77}$$

It is straightforward to invert this 'mode expansion' using the orthonormality relations, to express the coefficients $\varphi_u(k)$ in terms of the field,

$$\varphi_u(k) = s_u(k)\langle \phi(x), u(k;x)\rangle_{\text{KG}}. \tag{2.78}$$

Recall that we *a priori* require the field only to be an arbitrary configuration of complex numbers in spacetime, but the Klein-Gordon equation introduces a constraint that significantly restricts the actual number of allowed configurations. The coefficients $\varphi_u(k)$, due to the completeness (rather than 'overcompleteness') of the orthonormal basis in the space of solutions, form a 'minimal' representation of the field consisting of only the free parameters of its evolution. In this sense, specifying the $\varphi_u(k)$ is equivalent to specifying boundary conditions for a particular solution, such as specifying $\phi(x \in \Sigma)$ and $n^\mu \mathrm{D}_\mu \phi(x \in \Sigma)$ on a spacelike surface Σ.

It is convenient to split the mode expansion into a positive norm part and a negative norm part. To do this, we first define new coefficients $a_u(k)$, $b_u(k)$ via what amounts to a trivial re-labelling of the coefficients $\varphi_u(k)$,

$$\varphi_u(k) = \begin{cases} a_u(k), & \text{for } k \in K^+(u) \\ b_u^*(k), & \text{for } k \in K^-(u) \end{cases} . \tag{2.79}$$

The mode expansion of the field so split is then

$$\phi(x) = \int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\; a_u(k) u(k;x) + \int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\; b_u^*(k) u(k;x), \tag{2.80}$$

and for the complex conjugate of the field,

$$\phi^*(x) = \int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\; b_u(k) u^*(k;x) + \int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\; a_u^*(k) u^*(k;x). \tag{2.81}$$

This form is especially convenient for canonical quantization.

There is one more quantity of interest: the charge derived from the $U(1)$ Noether current j^μ (Eq. (2.38)) which is conserved on account of Noether's theorem,

$$Q = \int\limits_{\Sigma(t)} \mathrm{d}^3 r\, \sqrt{-\tilde{g}}\; n^\mu j_\mu(t,r) \tag{2.82}$$

$$= q\langle \phi(x), \phi(x) \rangle_{\mathrm{KG}}. \tag{2.83}$$

The second line shows an alternate way to interpret its being conserved: it is proportional to the Klein-Gordon inner product of the field with itself. Substituting the mode expansion Eq. (2.77) for ϕ, and using the orthonormality of the modes, we get a relatively simple expression,

$$\begin{aligned} Q &= q \int \widetilde{\mathrm{d}k}\; s_u(k) |\varphi_u(k)|^2 \\ &= q \left(\int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\; |a_u(k)|^2 - \int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\; |b_u(k)|^2 \right). \end{aligned} \tag{2.84}$$

This expression will prove useful after quantization.

In principle, Eq. (2.77) has already solved the problem of the (classical) dynamics of the scalar field - with the $\varphi(k)$ being constants of the motion. On the other hand, these constants play a key role in understanding the structure of the state space of the field in the quantum mechanical description, which we will turn to next.

Chapter 3

Scalar field quantization

We will have two primary points of discussion in this chapter. First, we will apply the standard method of canonical quantization (see e.g. Ref. [11]) and interpret the field in terms of eigenvalues of discrete observables, which will lead to picture of the field in terms of particles (as a mathematical notion, with no immediate physical interpretation) and corresponds to a Fock decomposition of the Hilbert space of states of the field. Subsequently, we will discuss transformations that relate two such decompositions of the state space, which will play a central role in our discussion of particle production.

3.1 Canonical quantization of the complex scalar field

We now consider the Hamiltonian formalism for the complex scalar field, choosing the surfaces $\Sigma(t)$ so that the timelike flow vectors are along the future-directed unit normals, $t^\mu(x) = n^\mu(x)$ (and $-\tilde{g} = |h^\Sigma|$). In that case, we may take the generalized coordinates to be $q^1(r) = \phi(r)$ and $q^2(r) = \phi^*(r)$ and the respective momenta $p_1(r) = \pi(r)$ and $p_2(r) = \pi^*(r)$ are given by

$$\pi(r) = \sqrt{-\tilde{g}}(\mathrm{D}_t\phi)^*(r); \; \pi^*(r) = \sqrt{-\tilde{g}}\mathrm{D}_t\phi(r). \tag{3.1}$$

Their independent Poisson brackets are given by (as the two degrees of freedom are complex conjugates of each other, the full set of Poisson brackets may be obtained by taking the complex conjugates of these)

$$[\phi(r), \phi(r')]_{\mathrm{PB}} = 0, \qquad [\phi(r), \phi^*(r')]_{\mathrm{PB}} = 0, \tag{3.2}$$

$$[\pi(r), \pi(r')]_{\mathrm{PB}} = 0, \qquad [\pi(r), \pi^*(r')]_{\mathrm{PB}} = 0, \tag{3.3}$$

$$[\phi(r), \pi(r')]_{\mathrm{PB}} = \delta(r - r'), \qquad [\phi(r), \pi^*(r')]_{\mathrm{PB}} = 0. \tag{3.4}$$

Following the standard methods of quantum mechanics, we promote dynamical variables such as $\phi(r)$, $\pi(r)$ to operators $\hat{\phi}(r)$, $\hat{\pi}(r)$ on a newly introduced linear vector space called the state space. The state of any system is described by a vector in this space, which determines the probability of obtaining specific values on measurement of various observables, and these values are further determined by the properties of the corresponding operator (e.g. Ref. [11]).

We will work in the Heisenberg picture, where it is the operators that evolve in time, with the state vectors remaining time-independent. The time evolution of any operator corresponding to a dynamical variable satisfies the Heisenberg equations of motion (with $\mathrm{D}_t = n^\mu \mathrm{D}_\mu$):

$$i\mathrm{D}_t \hat{A}(t_0) = [\hat{A}(t_0), \hat{H}(t_0)] \tag{3.5}$$

where $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is the commutator and $\hat{H}$ is the Hamiltonian operator. We note that the commutator is between operators at the same time t_0, as defined by our spacelike surfaces $\Sigma(t)$. The time-dependent field operators are now $\hat{\phi}(t, r) = \hat{\phi}(x)$, $\hat{\pi}(t, r) = \hat{\pi}(x)$.

The primary role of the Poisson brackets Eq. (3.2),(3.3),(3.4) in quantum mechanics is to *suggest* a form for the equal-time canonical commutation relations for the field observables, taking inspiration from the similarity between Eqs. (2.28) and (3.5) and requiring the quantized theory to give the right classical limit. This is given by the correspondence $[\hat{A}, \hat{B}] \leftrightarrow i[A, B]_{\mathrm{PB}}$, and we therefore postulate the canonical commutation relations (the independent ones; we must take the Hermitian conjugate of these relations to obtain a complete set)

$$[\hat{\phi}(t, r), \hat{\phi}(t, r')] = 0, \qquad [\hat{\phi}(t, r), \hat{\phi}^\dagger(t, r')] = 0, \tag{3.6}$$

$$[\hat{\pi}(t, r), \hat{\pi}(t, r')] = 0, \qquad [\hat{\pi}(t, r), \hat{\pi}^\dagger(t, r')] = 0, \tag{3.7}$$

$$[\hat{\phi}(t, r), \hat{\pi}(t, r')] = i\delta(r - r'), \qquad [\hat{\phi}(t, r), \hat{\pi}^\dagger(t, r')] = 0. \tag{3.8}$$

The Hamiltonian operator for the system is then (note that on account of Eqs. (3.6) and (3.7), the ordering of operators in each term is immaterial)

$$H(t) = \int\limits_{\Sigma(t)} \mathrm{d}^3r \sqrt{-\tilde{g}} \left(\frac{1}{(-\tilde{g})} \hat{\pi}^\dagger(t,r)\hat{\pi}(t,r) + \hat{\phi}^\dagger(t,r) \left(\mathrm{S}_\mu \mathrm{S}^\mu + m^2 \right) \hat{\phi}(t,r) \right), \tag{3.9}$$

and the Heisenberg equations of motion Eq. (3.5) for $\hat{A} = \hat{\phi}$ and $\hat{A} = \hat{\pi}^\dagger$ then respectively give

$$\mathrm{D}_t \hat{\phi} = \frac{1}{\sqrt{-\tilde{g}}} \hat{\pi}^\dagger, \tag{3.10}$$

$$\mathrm{D}_t \hat{\pi}^\dagger = \sqrt{-\tilde{g}} \left(\mathrm{S}_\mu \mathrm{S}^\mu + m^2 \right) \hat{\phi}. \tag{3.11}$$

It is straightforward to eliminate $\hat{\pi}^\dagger$ by substituting the first line in the second, leading to the operator version of the Klein-Gordon equation

$$\left(\mathrm{D}_\mu \mathrm{D}^\mu + m^2 \right) \hat{\phi}(x) = 0. \tag{3.12}$$

We already know that a mode expansion of the form Eq. (2.77) will completely solve this equation - however, the coefficients here must be operators. Thus, we just use Eq. (2.77) and promote the $\varphi_u(k)$ to operators $\hat{\varphi}_u(k)$ as well, so that we have

$$\hat{\phi}(x) = \int \widetilde{\mathrm{d}k}\ \hat{\varphi}_u(k) u(k;x). \tag{3.13}$$

It is important to note that the $\hat{\varphi}_u(k)$ are constants, rather than 'dynamical' variables, and do not obey the Heisenberg equation of motion Eq. (3.5). The operator version of Eq. (2.78) is

$$\hat{\varphi}_u(k) = s_u(k) \langle \hat{\phi}(x), u(k;x) \rangle_{\mathrm{KG}}. \tag{3.14}$$

We can now obtain commutators for the $\hat{\varphi}_u(k)$. Writing out the explicit expression for the Klein-Gordon inner products Eq. (2.52) in the Eq. (3.14), we may use Eq. (3.1) to convert the derivatives in the expression to momenta; using the appropriate relation from Eq. (3.6),(3.7),(3.8) and then the orthonormality of the mode functions gives the following (independent) commutators:

$$[\hat{\varphi}_u(k), \hat{\varphi}_u(k')] = 0, \tag{3.15}$$

$$[\hat{\varphi}_u(k), \hat{\varphi}_u^\dagger(k')] = s_u(k) \delta(k - k'). \tag{3.16}$$

In terms of the operators $\hat{a}_u(k)$ and $\hat{b}_u(k)$, defined via the operator version of Eq. (2.79), these commutators become

$$\hat{\varphi}_u(k) = \begin{cases} \hat{a}_u(k), \text{ for } k \in K^+(u) \\ \hat{b}_u^\dagger(k), \text{ for } k \in K^-(u) \end{cases}, \tag{3.17}$$

the non-vanishing independent commutators are:

$$[\hat{a}_u(k), \hat{a}_u^\dagger(k')] = \delta(k-k'), \text{ for } k, k' \in K^+(u), \tag{3.18}$$

$$[\hat{b}_u(k), \hat{b}_u^\dagger(k')] = \delta(k-k'), \text{ for } k, k' \in K^-(u). \tag{3.19}$$

Two observations are in order. Firstly, any two of the $\hat{a}$ or $\hat{b}$ operators associated with different values of k must commute; they may therefore be considered to act on independent subspaces of the state space. This allows us to write an arbitrary state vector $|\psi\rangle$ for the field as a tensor product of vectors with one in each such subspace,

$$|\psi\rangle = \bigotimes_k |\psi(k)\rangle_k, \tag{3.20}$$

where the subscript k denotes a vector in the subspace corresponding to k, and $\psi(k)$ labels the corresponding vector. Secondly, the operators (a or b) corresponding to a certain value of k satisfy the standard harmonic oscillator commutation relations, showing that each subspace k is spanned by eigenvectors of the operator $\hat{N}_u(k)$, defined as (for the full argument, see e.g. Refs. [8, 9])

$$\hat{N}_u(k) = \begin{cases} \hat{a}_u^\dagger(k)\hat{a}_u(k), \; k \in K^+(u) \\ \hat{b}_u^\dagger(k)\hat{b}_u(k), \; k \in K^-(u) \end{cases}. \tag{3.21}$$

The eigenvalue relation for these eigenvectors is given by

$$\hat{N}_u(k)|n;u\rangle_k = n\delta(k-k)|n;u\rangle_k, \; \forall\, n \in \mathbb{N}_0, \tag{3.22}$$

where $\mathbb{N}_0 = \{0, 1, 2, ...\}$. Thus, the k-th subspace is spanned by the orthonormal (due to being eigenstates of a Hermitian operator) basis $\{|n;u\rangle_k, \; \forall\, n \in \mathbb{N}_0\}$, often called a Fock basis; an arbitrary vector in this subspace can be written as

$$|\psi(k)\rangle_k = \sum_{n \in \mathbb{N}_0} c_n(k)|n;u\rangle_k, \tag{3.23}$$

for some constants $c_n(k)$ representing the components of the vector in the Fock basis corresponding to k.

The state $|\hat{N}_u(k); n; u\rangle_k$ is typically considered to represent n particles (if $k \in K^+(u)$) or antiparticles (if $k \in K^-(u)$) in the k-th mode, an interpretation that works well for a discrete set of modes. In fact, even for a continuous set of modes, we may always define an operator $\hat{n}_u(k)$ such that $\hat{N}_u(k) = \hat{n}_u(k)\delta(k-k)$ (see e.g. Refs. [8, 9]) whose eigenvalues are directly given by n. On the other hand, we will also define a 'total particle number' $\hat{\mathcal{N}}_u^+$ and 'total antiparticle number' $\hat{\mathcal{N}}_u^-$ by

$$\hat{\mathcal{N}}_u^\pm = \int\limits_{K^\pm(u)} \widetilde{\mathrm{d}k}\; \hat{N}_u(k). \tag{3.24}$$

To justify this, consider a Fock basis state $|\psi\rangle = \bigotimes_k |n(k); u\rangle$, where only a discrete set of subspaces $k \in K_D = \{k_1, k_2, ...\}$ have nonzero n, and the rest have $n(k) = 0$ i.e. $n(k)\delta(k-k) = \sum_i n(k_i)\delta(k-k_i)$. Then, we have

$$\hat{\mathcal{N}}_u^\pm|\psi\rangle = \left(\sum_{k_i \in K_D \cap K^\pm(u)} n(k_i) \right) |\psi\rangle, \tag{3.25}$$

i.e. the eigenvalue for each of $\hat{\mathcal{N}}^\pm$ is a finite number in $\mathbb{N}_0$, which lends itself better to an interpretation as a 'number of particles/antiparticles', but loses information about the specific mode the particle/antiparticle is associated with. The latter issue can be remedied by considering distribution functions $f(k) > 0$, and defining particle and antiparticle numbers weighted by this distribution:

$$\hat{\mathcal{N}}_u^\pm[f(k)] = \int\limits_{K^\pm(u)} \widetilde{\mathrm{d}k}\; f(k)\hat{N}_u(k). \tag{3.26}$$

For the aforementioned state $|\psi\rangle = \bigotimes_k |n(k); u\rangle$ with $n(k)\delta(k-k) = \sum_i n(k_i)\delta(k-k_i)$, we have for these operators

$$\hat{\mathcal{N}}_u^\pm[f(k)]\,|\psi\rangle = \left(\sum_{k_i \in K_D \cap K^\pm(u)} f(k_i)n(k_i) \right) |\psi\rangle. \tag{3.27}$$

There are of course the special cases

$$\hat{\mathcal{N}}_u^\pm = \hat{\mathcal{N}}_u^\pm[f(k) = 1], \tag{3.28}$$

$$\hat{N}_u(k') = \hat{\mathcal{N}}_u^\pm\left[f(k) = \delta(k-k'), k' \in K^\pm(u)\right]. \tag{3.29}$$

From this point of view, the $\hat{N}(k)$ are operators representing the 'density' of particles in k-space (with respect to the volume element $\widetilde{\mathrm{d}k}$), and the functions $f(k)$ represent our sensitivity to each k (chosen as suitable for the problem at hand).

The reason particle and antiparticle number (or equivalently, subspaces corresponding to positive norm and negative norm mode functions) are considered separately have to do with their contribution to the invariant Noether charge corresponding to $U(1)$ gauge transformations Eq. (2.84). The operator version of this charge has an ambiguity in its ordering of the $\hat{a}$ and $\hat{b}$ operators and their conjugates in each term. We choose what is called the normal ordering prescription (described in, for instance, Refs. [8, 9]), indicated by $: :_u$, where all the creation/conjugate operators $\hat{a}_u^\dagger(k)$,$\hat{b}_u^\dagger(k)$ will be to the left of the annihilation operators $\hat{a}_u(k)$, $\hat{b}_u(k)$. In that case, the charge defined in Eq. (2.84) reduces to the rather simple expression

$$: \hat{Q} :_u \; = q\hat{\mathcal{N}}_u^+ - q\hat{\mathcal{N}}_u^- . \tag{3.30}$$

Therefore, each particle contributes q and each antiparticle $-q$ to the total charge.

3.2 The non-uniqueness of the Fock space decomposition

We have seen that the state space of a complex scalar field can be separated into Fock spaces by an orthonormal mode expansion (we use the subscript $_u$ for operators and a label $|...;u\rangle$ for kets to indicate association with the mode functions $\{u(k;x)\forall k\}$),

$$\hat{\phi}(x) = \int\limits_{K^+(u)} \widetilde{\mathrm{d}k}\; \hat{a}_u(k)u(k;x) + \int\limits_{K^-(u)} \widetilde{\mathrm{d}k}\; \hat{b}_u^\dagger(k)u(k;x), \tag{3.31}$$

where each $\hat{a}_u(k)$ or $\hat{b}_u(k)$ is associated with a state subspace spanned by the eigenstates of $\hat{N}_u(k \in K^+(u)) = \hat{a}_u^\dagger(k)\hat{a}_u(k)$ and $\hat{N}_u(k \in K^-(u)) = \hat{b}_u^\dagger(k)\hat{b}_u(k)$ respectively, namely

$$|n;u\rangle_k \forall\, n \in \mathbb{N}_0 : \hat{N}_u(k)|n;u\rangle_k = n\delta(k-k)|n;u\rangle_k. \tag{3.32}$$

In a top-down interpretation, we may consider this a decomposition of the Hilbert space of states into these subspaces, referred to as a Fock space decomposition (or just a Fock space). The 'vacuum state' $|0;u\rangle$ for this particular choice of mode functions is

$$|0;u\rangle = \bigotimes_k |0;u\rangle_k, \tag{3.33}$$

which is uniquely identified by the following property:

$$\hat{a}_u(k)|0;u\rangle = 0 \;\forall\; k \in K^+(u),\; \hat{b}_u(k)|0;u\rangle = 0 \;\forall\; k \in K^-(u). \tag{3.34}$$

Clearly, a different choice of modes would have yielded a different 'Fock decomposition' of the Hilbert space. To study this further, we take another set of orthonormal mode functions $v(k;x)$, where the field expansion is instead

$$\hat{\phi}(x) = \int\limits_{K^+(v)} \widetilde{\mathrm{d}k}\; \hat{a}_v(k) v(k;x) + \int\limits_{K^-(v)} \widetilde{\mathrm{d}k}\; \hat{b}_v^\dagger(k) v(k;x). \tag{3.35}$$

This has its own Fock space decomposition of the Hilbert space with the vacuum state $|0;v\rangle$. The two sets of modes $u(k;x)$ and $v(k;x)$ can be related by an invertible linear transformation (called a Bogolubov transformation),

$$u(k;x) = \int\limits_{K(v)} \widetilde{\mathrm{d}k'}\; B_{uv}(k,k') v(k';x), \tag{3.36}$$

$$v(k;x) = \int\limits_{K(u)} \widetilde{\mathrm{d}k'}\; B_{vu}(k,k') u(k';x), \tag{3.37}$$

where the Bogolubov coefficients $B_{uv}(k,k')$ and $B_{vu}(k,k')$ are 'matrix inverses' of each other:

$$\int\limits_{K(v)} \widetilde{\mathrm{d}k'}\; B_{uv}(k,k') B_{vu}(k',k'') = \int\limits_{K(u)} \widetilde{\mathrm{d}k'}\; B_{vu}(k,k') B_{uv}(k',k'') = \delta(k-k''). \tag{3.38}$$

We can also find explicit expressions for these coefficients using the Klein-Gordon inner product:

$$B_{uv}(k,k') = s_v(k') \langle u(k;x), v(k';x)\rangle_{\mathrm{KG}}. \tag{3.39}$$

This leads to a useful relation with the inverse coefficients. As the Klein-Gordon inner product satisfies $\langle f,g\rangle_{\mathrm{KG}} = \langle g,f\rangle^*_{\mathrm{KG}}$, we have:

$$B_{uv}(k,k') = s_u(k) s_v(k') B^*_{vu}(k',k). \tag{3.40}$$

Bogolubov transformations must clearly preserve the orthonormality of modes (as the u and v modes both form complete orthonormal sets). Enforcing the orthonormality relations,

$$\langle u(k \in K^\pm(u);x), u(k';x)\rangle_{\mathrm{KG}} = \pm\delta(k-k'), \tag{3.41}$$

$$\langle v(k \in K^\pm(v);x), v(k';x)\rangle_{\mathrm{KG}} = \pm\delta(k-k'), \tag{3.42}$$

(or substituting equation Eq. (3.40) in Eq. (3.38)) gives the following constraint on the Bogolubov coefficients (and similarly with $(u \leftrightarrow v)$):

$$\int\limits_{K^+(v)} \widetilde{\mathrm{d}k'}\; B_{uv}(k,k')B^*_{uv}(k'',k') - \int\limits_{K^-(v)} \widetilde{\mathrm{d}k'}\; B_{uv}(k,k')B^*_{uv}(k'',k') = s_u(k)\delta(k-k''). \tag{3.43}$$

A corresponding relation between the operators can be obtained by equating the two expansions of $\hat{\phi}$ (equations Eqs. (3.31) and (3.35)),

$$\int\limits_{K^+(u)} \widetilde{\mathrm{d}k'}\; \hat{a}_u(k')B_{uv}(k',k) + \int\limits_{K^-(u)} \widetilde{\mathrm{d}k'}\; \hat{b}^\dagger_u(k')B_{uv}(k',k) = \begin{cases} \hat{a}_v(k) & , \; k \in K^+(v) \\ \hat{b}^\dagger_v(k) & , \; k \in K^-(v) \end{cases}. \tag{3.44}$$

Evidently,

$$|0;u\rangle = |0;v\rangle \Leftrightarrow B_{uv}(k,k') = 0 \;\forall\; (k,k') \in (K^+(u) \times K^-(v)) \cup (K^-(u) \times K^+(v)). \tag{3.45}$$

The two vacuum states are identical ($\hat{a}_u(k)|0;v\rangle = 0 \;\forall\; k \in K^+(u),\; \hat{b}_u(k)|0;v\rangle = 0 \;\forall\; k \in K^-(u)$ and $u \leftrightarrow v$) if and only if the elements of B_{uv} between $K^+(u)$ and $K^-(v)$ or between $K^-(u)$ and $K^+(v)$ are all zero i.e. the transformation does not mix positive and negative norm modes.

More specifically, we are often interested in the expectation value of the particle number density in the k-space of one set of modes in the vacuum state corresponding to another set of modes. If the former is in the v-basis and the latter in the u-basis, it is straightforward to show from Eq. (3.44) and the commutators Eq. (3.15),(3.16) that

$$\langle 0;u|\hat{N}_v(k \in K^\pm(v))|0;u\rangle = \int\limits_{K^\mp(u)} \widetilde{\mathrm{d}k'}\; |B_{vu}(k,k')|^2. \tag{3.46}$$

This is the primary feature of Bogolubov transformations that results in particle production by classical backgrounds, as we will see in Chap. 4.

Now, we will consider some features pertaining to particle numbers in Fock spaces related by Bogolubov transformations. First, we consider the particle and antiparticle numbers in an arbitrary expansion of $\hat{\phi}$ in terms of orthonormal modes $f(k;x)$:

$$\hat{\mathcal{N}}^+_f = \int\limits_{K^+(f)} \widetilde{\mathrm{d}k}\; a^\dagger_f(k)a_f(k), \tag{3.47}$$

$$\hat{\mathcal{N}}^-_f = \int\limits_{K^-(f)} \widetilde{\mathrm{d}k}\; b^\dagger_f(k)b_f(k). \tag{3.48}$$

For a vacuum state preserving Bogolubov transformation between modes u and v, we have, from using equations Eqs. (3.40) and (3.43),

$$\int_{K^+(u)} \widetilde{\mathrm{d}k}\; B_{uv}(k,k')B^*_{uv}(k,k'') = \delta(k'-k''). \tag{3.49}$$

Attempting to express $\hat{\mathcal{N}}_u^+$ in terms of the v operators then gives:

$$\begin{aligned}\hat{\mathcal{N}}_u^+ &= \int_{K^+(u)} \widetilde{\mathrm{d}k} \int_{K^+(v)} \widetilde{\mathrm{d}k'} \int_{K^+(v)} \widetilde{\mathrm{d}k''}\; B^*_{uv}(k,k')B_{uv}(k,k'')a_v^\dagger(k')a_v(k'') \\ &= \int_{K^+(v)} \widetilde{\mathrm{d}k}\; a_f^\dagger(v)a_f(v) \\ &= \hat{\mathcal{N}}_v^+. \end{aligned} \tag{3.50}$$

Similarly, we can show that $\hat{\mathcal{N}}_u^- = \hat{\mathcal{N}}_v^-$.

We see that if two Fock space decompositions have a common vacuum state, they must necessarily agree on the total number of particles as well as the total number of antiparticles in any state. There are no 'particle creation' phenomena between the two decompositions. An example of a Bogolubov transformation of this type is that between positive frequency plane wave modes and positive frequency spherical wave modes, which, as we have seen, must preserve the total particle and antiparticle content of any eigenstate; however, the distribution of particles in each 'mode' depends on the specific choice of modes.

A more general conclusion concerns the difference between the particle and antiparticle number

$$\Delta\hat{\mathcal{N}}_f = \hat{\mathcal{N}}_f^+ - \hat{\mathcal{N}}_f^-. \tag{3.51}$$

This can be related to an invariant observable, the Klein-Gordon inner product of the field with itself, $\langle\hat{\phi},\hat{\phi}\rangle_{\mathrm{KG}}$ (with the operators ordered so that $\hat{\phi}$ is to the left of $\hat{\phi}^\dagger$). Using the f-mode expansion,

$$\hat{\phi}(x) = \int_{K^+(f)} \widetilde{\mathrm{d}k}\; \hat{a}_f(k)f(k;x) + \int_{K^-(f)} \widetilde{\mathrm{d}k}\; \hat{b}_f^\dagger(k)f(k;x), \tag{3.52}$$

and the orthonormality relations

$$\langle f(k;x), f(k';x)\rangle_{\mathrm{KG}} = s_f(k)\delta(k-k'), \tag{3.53}$$

we find

$$\langle\hat{\phi}(x),\hat{\phi}(x)\rangle_{\mathrm{KG}} = \int\limits_{K^+(f)} \widetilde{\mathrm{d}k}\ \hat{a}_f\hat{a}_f^\dagger - \int\limits_{K^-(f)} \widetilde{\mathrm{d}k}\ \hat{b}_f^\dagger\hat{b}_f. \tag{3.54}$$

From the commutator $[a_f(k), a_f^\dagger(k')] = \delta(k-k') = \langle f(k;x), f(k';x)\rangle_{\mathrm{KG}}$, we get for $\Delta\hat{\mathcal{N}}_f$

$$\Delta\hat{\mathcal{N}}_f = \langle\hat{\phi}(x),\hat{\phi}(x)\rangle_{\mathrm{KG}} - \int\limits_{K^+(f)} \widetilde{\mathrm{d}k}\ \langle f(k;x), f(k;x)\rangle_{\mathrm{KG}}. \tag{3.55}$$

$\Delta\hat{\mathcal{N}}_f$ differs from the invariant inner product of the field with itself by a term equal to the integral of the norm of all positive norm modes in the chosen basis. From Eq. (2.67) in Sec. 2.4, this integral is invariant across all choices of an orthonormal set of basis mode functions.

Therefore, given two different orthonormal bases u and v related by a Bogolubov transformation, we have

$$\Delta\hat{\mathcal{N}}_u = \Delta\hat{\mathcal{N}}_v. \tag{3.56}$$

For this reason, we will denote this difference as simply $\Delta\hat{\mathcal{N}}$, without any reference to the choice of basis.

On a side note, this also leads to a constraint on the Bogolubov coefficients; using

$$\langle 0;u|\Delta\hat{\mathcal{N}}|0;u\rangle = 0, \tag{3.57}$$

with $\Delta\hat{\mathcal{N}}$ expressed as $\Delta\hat{\mathcal{N}}_v$, we find, using the Bogolubov transformation Eq. (3.44) to the u operators,

$$\int\limits_{K^-(u)} \widetilde{\mathrm{d}k'} \int\limits_{K^+(v)} \widetilde{\mathrm{d}k}\ |B_{uv}(k',k)|^2 = \int\limits_{K^+(u)} \widetilde{\mathrm{d}k'} \int\limits_{K^-(v)} \widetilde{\mathrm{d}k}\ |B_{uv}(k',k)|^2. \tag{3.58}$$

The orthonormality condition Eq. (3.43) also implies that

$$\int\limits_{K^+(u)} \widetilde{\mathrm{d}k'} \int\limits_{K^+(v)} \widetilde{\mathrm{d}k}\ |B_{uv}(k',k)|^2 = \int\limits_{K^-(u)} \widetilde{\mathrm{d}k'} \int\limits_{K^-(v)} \widetilde{\mathrm{d}k}\ |B_{uv}(k',k)|^2. \tag{3.59}$$

This means that any two arbitrary Fock space decompositions of the state space of a complex scalar field in a combined electromagnetic and gravitational background must therefore necessarily agree on the *difference* between the number of particles and antiparticles. Any state of the field as viewed in one Fock space decomposition where it has a precise number of particles and antiparticles may then at most differ from its interpretation in another decomposition by the presence of (superpositions of states with) additional particle-antiparticle pairs.

The role this plays in particle production is analogous to the conservation of charge in the dynamics of the field. Indeed, the normal-ordered charge operator for this theory is : $Q := q\Delta\hat{\mathcal{N}}$ (which differs from the field charge operator as such by q times the above integral over positive norm modes), and the above statement is equivalent to the statement that all Fock space decompositions 'see' the same charge (of course, with particles and antiparticles having charges $+q$ and $-q$ respectively) in any state.

Chapter 4

Particle production by classical backgrounds

In this chapter we will study particle production for the complex scalar field with $U(1)$ gauge symmetry, considering a general electromagnetic and gravitational background.

In the framework of Bogolubov transformations, discussed in Sec. 3.2, the description of particle production is as follows: Describing spacetime as a family of spacelike surfaces $\{\Sigma(t)\}$, we associate a particular Fock space decomposition (via a choice of some appropriate mode functions) with each t. The modes corresponding to different values of t are in general related by a Bogolubov transformation, and may have differing vacuum states. We typically assume the system begins in the vacuum state associated with early times, and evaluate the particle content corresponding to this state in the Fock decomposition at other times.

We will now review some special choices of mode functions that arguably make the necessary calculations easier, and subsequently discuss two different criteria for associating Fock spaces with each time t - instantaneous Hamiltonian diagonalization and the adiabatic vacuum criterion. We will also briefly describe how the results of these sections may be directly used to numerically evaluate the particle content in different modes.

4.1 Positive and negative frequency modes

4.1.1 The general positive and negative frequency modes

It is convenient to work with modes that behave, in some sense, analogous to the positive and negative frequency modes $e^{\mp i\omega t}$ in Minkowski spacetime. These modes respectively have positive and negative Klein-Gordon norm, which will be an important criterion for the discussion in this section. Here, we will define this class of modes, and show that they are orthonormal and complete in the Klein-Gordon sense.

We begin by considering $\mathrm{L}^2(\Sigma(t))$, the space of square integrable complex functions defined on a particular $\Sigma(t)$, with the usual inner product

$$\langle f(r), g(r)\rangle_{\Sigma(t)} = \int\limits_{\Sigma(t)} \mathrm{d}^3 r \, \sqrt{-\tilde{g}} \, f(r) g^*(r), \tag{4.1}$$

for $f, g \in \mathrm{L}^2(\Sigma(t))$, which is associated with a positive-definite norm for non-null functions. Let $\{\chi_t(\mathbf{k}; r), \ \forall \mathbf{k}\}$ be a complete set of orthonormal basis functions in $\mathrm{L}^2(\Sigma(t))$. The orthonormality condition reads

$$\langle \chi_t(\mathbf{k}; r), \chi_t(\mathbf{k}'; r)\rangle_{\Sigma(t)} = \delta(\mathbf{k} - \mathbf{k}'), \tag{4.2}$$

while completeness means that the projection operators on each of these basis functions sum to the identity operator

$$\int \mathrm{d}\mathbf{k} \, \chi_t(\mathbf{k}, r) \chi_t(\mathbf{k}, r') = \frac{\delta(r - r')}{\sqrt{-\tilde{g}(r)}}. \tag{4.3}$$

Here, $\mathrm{d}\mathbf{k}$ is a suitable integration measure in $\mathbf{k}$-space, with $\delta(\mathbf{k} - \mathbf{k}')$ being the corresponding Dirac-δ function, satisfying

$$\int \mathrm{d}\mathbf{k} \, h(\mathbf{k}) \delta(\mathbf{k} - \mathbf{k}') = h(\mathbf{k}'), \tag{4.4}$$

for any function $h(\mathbf{k})$.

Now, the mode functions for $\hat{\phi}$ satisfy the Klein-Gordon equation

$$\left(\mathrm{D}_\mu \mathrm{D}^\mu + m^2\right) u(k; x) = 0. \tag{4.5}$$

One of the ways of specifying a solution to such an equation is to specify the values of the mode function and its derivatives along the normal throughout a particular spacelike surface. In terms of the basis functions $\chi_t(\mathbf{k}, r)$ defined earlier, we now define the modes $u(k; x) = u_t^\pm(\mathbf{k}; x)$ via the initial conditions

$$\left.\begin{aligned} u_t^\pm(\mathbf{k}; x(t, r)) &= M^\pm(\mathbf{k}) \chi_t(\mathbf{k}; x), \\ \mathrm{D}_t u_t^\pm(\mathbf{k}; x(t, r)) &= -i\Omega_t^\pm(\mathbf{k}) M^\pm(\mathbf{k}) \chi_t(\mathbf{k}; x) \end{aligned}\right\} \forall r \in \Sigma(t), \tag{4.6}$$

where $\Omega_t^\pm(\mathbf{k})$ are arbitrary functions of $\mathbf{k}$, and loosely represent the 'frequencies' corresponding to these modes at time t. $M^\pm(\mathbf{k})$ is a normalization constant for each of these modes. We have replaced the generic single index k that we've used to index the modes with the pair $(s, \mathbf{k})$, where $s \in \{+, -\}$. We will show that the latter is a complete representation of the former i.e. that the $u_t^{\pm}(\mathbf{k}; x)$ generally form a complete set of basis modes for solutions to the Klein-Gordon equation.

To show that the modes are complete, we must establish that any solution of Eq. (4.5) must have a unique expression as a linear combination of these modes. We will see that this follows from the completeness of the $\chi_t(\mathbf{k}; x)$ in $\mathrm{L}^2(\Sigma(t))$. To begin with, consider a general solution $w(x)$, which is specified by the arbitrary initial conditions:

$$\left.\begin{aligned} w(x) &= w_1(x) \\ \mathrm{D}_t w(x) &= w_2(x) \end{aligned}\right\} \forall x \in \Sigma(t), \tag{4.7}$$

where $w_1, w_2 \in \mathrm{L}^2(\Sigma(t))$. We write $w(x)$ as a linear combination of the positive and negative frequency modes of $\Sigma(t)$:

$$w(x) = \int \mathrm{d}\mathbf{k} \, \left(c^+(\mathbf{k}) u_t^+(\mathbf{k}; x) + c^-(\mathbf{k}) u_t^-(\mathbf{k}; x)\right), \tag{4.8}$$

where $\mathrm{d}\mathbf{k}$ is a suitable integration measure.

The boundary conditions then reduce to relations between elements of $\mathrm{L}^2(\Sigma(t))$,

$$\left.\begin{aligned} w_1(x) &= \int \mathrm{d}\mathbf{k} \, \left(M^+(\mathbf{k}) c^+(\mathbf{k}) + M^-(\mathbf{k}) c^-(\mathbf{k})\right) \chi_t(\mathbf{k}; x), \\ w_2(x) &= -i \int \mathrm{d}\mathbf{k} \, \left(\Omega_t^+(\mathbf{k}) M^+(\mathbf{k}) c^+(\mathbf{k}) + \Omega_t^-(\mathbf{k}) M^-(\mathbf{k}) c^-(\mathbf{k})\right) \chi_t(\mathbf{k}; x) \end{aligned}\right\} \forall x \in \Sigma(t). \tag{4.9}$$

The completeness of the $\chi_t(\mathbf{k}; x)$ ensures that unique values are obtained for $(M^+(\mathbf{k}) c^+(\mathbf{k}) + M^-(\mathbf{k}) c^-(\mathbf{k}))$ and $\left(\Omega_t^+(\mathbf{k}) M^+(\mathbf{k}) c^+(\mathbf{k}) + \Omega_t^-(\mathbf{k}) M^-(\mathbf{k}) c^-(\mathbf{k})\right)$ i.e. $c^+(\mathbf{k})$ and $c^-(\mathbf{k})$ for all $\mathbf{k}$, as long as $\Omega_t^+(\mathbf{k}) \neq \Omega_t^-(\mathbf{k})$. Thus, the basis formed by $u_t^+(\mathbf{k}; x)$ and $u_t^-(\mathbf{k}; x)$ is complete in the space of solutions of Eq. (4.5) as required.

For the chosen modes to yield a Fock space, they must also be orthonormal in the Klein-Gordon sense. The orthogonality requirement is:

$$\langle u_t^{s_1}(\mathbf{k}; x), u_t^{s_2}(\mathbf{k}'; x) \rangle_{\mathrm{KG}} \propto \delta_{s_1, s_2} \delta(\mathbf{k} - \mathbf{k}'), \tag{4.10}$$

We have,

$$\langle u_t^{s_1}(\mathbf{k};x), u_t^{s_2}(\mathbf{k}';x)\rangle_{\mathrm{KG}} = -i\int_{\Sigma(t)} \mathrm{d}^3r\,\sqrt{-\tilde{g}}\,\left(u_t^{s_1}(\mathbf{k};x)\left(\mathrm{D}_t u_t^{s_2}(\mathbf{k}';x)\right)^* - u_t^{s_2*}(k';x)\mathrm{D}_t u_t^{s_1}(\mathbf{k};x)\right) \tag{4.11}$$

$$= \left[|M^{s_1}(\mathbf{k})|^2\Omega_t^{s_1}(\mathbf{k}) + |M^{s_2}(\mathbf{k})|^2(\Omega_t^{s_2}(\mathbf{k}))^*\right]\int_{\Sigma(t)} \mathrm{d}^3r\,\sqrt{-\tilde{g}}\,\chi_t^*(\mathbf{k}';x)\chi_t(\mathbf{k};x) \tag{4.12}$$

$$= \left[|M^{s_1}(\mathbf{k})|^2\Omega_t^{s_1}(\mathbf{k}) + |M^{s_2}(\mathbf{k})|^2(\Omega_t^{s_2}(\mathbf{k}))^*\right]\delta(\mathbf{k}-\mathbf{k}'). \tag{4.13}$$

To ensure orthonormality as in Eq. (4.10), we must satisfy the following condition

$$\Omega_t^+(\mathbf{k}) = -\left(\Omega_t^-(\mathbf{k})\right)^* = \Omega_t(\mathbf{k})\text{ (say)}. \tag{4.14}$$

This also means that $u_t^+(\mathbf{k};x)$ and $u_t^-(\mathbf{k};x)$ must have norms differing in sign. We choose the normalization so that the former has positive norm i.e.

$$\langle u_t^{s_1}(\mathbf{k};x), u_t^{s_2}(\mathbf{k}';x)\rangle_{\mathrm{KG}} = s_1\delta_{s_1,s_2}\delta(\mathbf{k}-\mathbf{k}'). \tag{4.15}$$

Imposing this normalization yields the additional conditions:

$$\mathrm{Re}\,\Omega_t(\mathbf{k}) > 0, \tag{4.16}$$

$$|M^+(\mathbf{k})|^2 = |M^-(\mathbf{k})|^2 = \frac{1}{2\,\mathrm{Re}\,\Omega_t(\mathbf{k})}. \tag{4.17}$$

Note that $\Omega_t(\mathbf{k})$ cannot be purely imaginary (which is contained in Eq. (4.16)) as that would correspond to modes of vanishing norm, preventing us from obtaining a Fock space structure with such modes.

The most general positive/negative frequency modes are then given by the boundary conditions

$$\left.\begin{aligned} u_t^{\pm}(\mathbf{k};x(t,r)) &= \frac{1}{\sqrt{2\,\mathrm{Re}\,\Omega_t(\mathbf{k})}}\chi_t(\mathbf{k};x) \\ \mathrm{D}_t u_t^{+}(\mathbf{k};x(t,r)) &= \frac{-i\Omega_t(\mathbf{k})}{\sqrt{2\,\mathrm{Re}\,\Omega_t(\mathbf{k})}}\chi_t(\mathbf{k};x) \\ \mathrm{D}_t u_t^{-}(\mathbf{k};x(t,r)) &= \frac{i\Omega_t^*(\mathbf{k})}{\sqrt{2\,\mathrm{Re}\,\Omega_t(\mathbf{k})}}\chi_t(\mathbf{k};x) \end{aligned}\right\}\ \forall r\in\Sigma(t). \tag{4.18}$$

Different choices of $\Omega_t(\mathbf{k})$ will yield different Fock spaces, even within this class of positive/negative frequency modes.

4.1.2 Eigenfunctions of the spatial Klein-Gordon operator

The eigenfunctions of the operator $(\mathrm{S}_\mu \mathrm{S}^\mu + m^2)$ (which we will call the 'spatial Klein-Gordon operator') occurring in the Hamiltonian Eq. (2.44) provide a convenient choice for the $\chi(k;r)$, which we will now discuss. As $(\mathrm{S}_\mu \mathrm{S}^\mu + m^2)$ is a local operator, its action on any element of $\mathrm{L}^2(\Sigma(t))$ is simply another element of $\mathrm{L}^2(\Sigma(t))$. We define the operator $\mathrm{K}_t : \mathrm{L}^2(\Sigma(t)) \to \mathrm{L}^2(\Sigma(t))$ so that

$$\mathrm{K}_t f = \left(\mathrm{S}_\mu \mathrm{S}^\mu + m^2\right) f, \ \forall f \in \mathrm{L}^2(\Sigma(t)). \tag{4.19}$$

We may now consider the eigenvalues $\lambda_t(\mathbf{k})$ and linearly independent eigenfunctions $\chi_t(\mathbf{k};r)$ of K_t ($\mathbf{k}$ is essentially an index to keep track of these eigenvalues and eigenfunctions), which satisfy the equation

$$\mathrm{K}_t \chi_t(\mathbf{k};r) = \lambda_t(\mathbf{k}) \chi_t(\mathbf{k};r), \ \forall r \in \Sigma(t). \tag{4.20}$$

The eigenvalues $\lambda_t(\mathbf{k})$ are real on account of K_t being a Hermitian operator. We will now show that the eigenvalues are also non-negative. To do this, multiply the above equation by $\chi_t^*(\mathbf{k};r)$ and integrate over $\Sigma(t)$, to get

$$\int\limits_{\Sigma(t)} \mathrm{d}^3 r \ \sqrt{-\tilde{g}} \ \chi_t^*(\mathbf{k};r) \mathrm{K}_t \chi_t(\mathbf{k};r) = \lambda_t(\mathbf{k}) \int\limits_{\Sigma(t)} \mathrm{d}^3 r \ \sqrt{-\tilde{g}} |\chi_t(\mathbf{k};r)|^2. \tag{4.21}$$

We express K_t in terms of the derivatives, and integrate by parts (assuming that the boundary term vanishes), and use the fact that the $\chi_t(\mathbf{k},x)$ are normalized to obtain

$$\lambda_t(\mathbf{k}) - m^2 = -\int\limits_{\Sigma(t)} \mathrm{d}^3 r \ \sqrt{-\tilde{g}} \ g^{\mu\nu}(t,r) \left(\mathrm{S}_\mu \chi_t(\mathbf{k};r)\right)^* \left(\mathrm{S}_\nu \chi_t(\mathbf{k};r)\right). \tag{4.22}$$

As the components $\mathrm{S}_\mu \chi_t(\mathbf{k};r)$ are orthogonal to the timelike n^μ, they are in a non-positive (we allow them to vanish) norm subspace of the local tangent space[1], and the $U(1)$ internal space norm (i.e. $\phi^*\phi$) is positive definite. Therefore, the right hand side is the negative of an integral over non-positive values, making it non-negative, which implies

$$\lambda_t(\mathbf{k}) - m^2 \geq 0. \tag{4.23}$$

[1] We note that the pseudo-Riemannian metrics we are working with correspond to indefinite inner products on tangent spaces.

This is a stronger condition than $\lambda_t(\mathbf{k}) \geq 0$ (we are assuming that m^2 is non-negative), so in any case, there exists a unique $\omega_t(\mathbf{k}) \in \mathbb{R}_{\geq 0}$ (i.e. non-negative) such that

$$\lambda_t(\mathbf{k}) = (\omega_t(\mathbf{k}))^2. \tag{4.24}$$

The eigenvalue equation Eq. (4.20) now reads:

$$\mathrm{K}_t \chi_t(\mathbf{k}; r) = \omega_t^2(\mathbf{k}) \chi_t(\mathbf{k}; r). \tag{4.25}$$

It will also be useful to express the Klein-Gordon equation in terms of the spatial Klein-Gordon operator. In general, we have

$$\left(\mathrm{D}_t^2 + (n_\mu \mathrm{D}_t n^\mu + \mathrm{S}_\mu n^\mu)\mathrm{D}_t + n^\mu \mathrm{D}_t \mathrm{S}_\mu + \mathrm{K}_t\right) \phi = 0. \tag{4.26}$$

This is a rather complicated equation. However, if the unit normals to the surfaces lie along geodesics (equivalently, we choose a single Cauchy surface and generate geodesics from its unit normal field, and the other surfaces are determined as those orthogonal to the resulting geodesics), then $\mathrm{D}_t n^\mu = 0$, and this reduces to the much simpler form

$$\left(\mathrm{D}_t^2 + (\mathrm{D}_\mu n^\mu)\mathrm{D}_t + \mathrm{K}_t\right) \phi = 0, \tag{4.27}$$

in which all spatial derivatives of ϕ are contained in K_t. We will see that choosing such surfaces is always possible in the examples we will be concerned with.

4.1.3 Spatial Klein-Gordon eigenfunctions and gauge invariance

The eigenfunctions of the spatial Klein-Gordon operator may be used for the boundary conditions Eq. (4.18), with the additional restriction that the frequency is a function of only the eigenvalues, $\Omega_t(\mathbf{k}) = \Omega(\omega_t(\mathbf{k}))$, so that they now read

$$u_t^{\pm}(\mathbf{k}; x(t, r)) = \frac{1}{\sqrt{2\,\mathrm{Re}\,\Omega(\omega_t(\mathbf{k}))}} \chi_t(\mathbf{k}; r), \tag{4.28}$$

$$\mathrm{D}_t u_t^{+}(\mathbf{k}; x(t, r)) = \frac{-i\Omega(\omega_t(\mathbf{k}))}{\sqrt{2\,\mathrm{Re}\,\Omega(\omega_t(\mathbf{k}))}} \chi_t(\mathbf{k}; r), \tag{4.29}$$

$$\mathrm{D}_t u_t^{-}(\mathbf{k}; x(t, r)) = \frac{i\Omega^*(\omega_t(\mathbf{k}))}{\sqrt{2\,\mathrm{Re}\,\Omega(\omega_t(\mathbf{k}))}} \chi_t(\mathbf{k}; r). \tag{4.30}$$

We will now consider the behaviour of such modes under local $U(1)$ gauge transformations,

$$\phi(x) \to e^{-i\Lambda(x)}\phi(x), \tag{4.31}$$

$$\mathrm{D}_\mu\phi(x) \to e^{-i\Lambda(x)}\mathrm{D}_\mu\phi(x), \tag{4.32}$$

for any $\phi(x)$ in the internal space, including the modes $u(k;x)$.

The way we have defined the particle modes above is gauge independent. This follows from the fact that the Klein-Gordon equation Eq. (4.5), the spatial eigenfunction equation Eq. (4.20), and the boundary conditions Eq. (4.6) are all gauge covariant (i.e. the equations have the same form in any gauge) as the eigenvalues $\omega_t^2(\mathbf{k})$ are gauge invariant.

In essence, under a gauge transformation by $e^{-i\Lambda(x)}$, the instantaneous positive and negative frequency modes transform only as vectors in the internal space

$$u_t^\pm(\mathbf{k};x) \to e^{-i\Lambda(x)}u_t^\pm(\mathbf{k};x). \tag{4.33}$$

The creation/annihilation (chosen as convenient) operators are given by:

$$\hat{a}_t(\mathbf{k}) = \langle\hat{\phi}(x), u_t^+(\mathbf{k};x)\rangle_{\mathrm{KG}}, \tag{4.34}$$

$$\hat{b}_t^\dagger(\mathbf{k}) = \langle\hat{\phi}(x), u_t^-(\mathbf{k};x)\rangle_{\mathrm{KG}}, \tag{4.35}$$

Therefore, under a gauge transformation, due to the gauge invariance of the Klein-Gordon inner product,

$$\hat{a}_t(\mathbf{k}) \to \langle\hat{\phi}(x)e^{-i\Lambda(x)}, u_t^+(\mathbf{k};x)e^{-i\Lambda(x)}\rangle_{\mathrm{KG}} = \langle\hat{\phi}(x), u_t^+(\mathbf{k};x)\rangle_{\mathrm{KG}} = \hat{a}_t(\mathbf{k}), \tag{4.36}$$

$$\hat{b}_t^\dagger(\mathbf{k}) \to \langle\hat{\phi}(x)e^{-i\Lambda(x)}, u_t^-(\mathbf{k};x)e^{-i\Lambda(x)}\rangle_{\mathrm{KG}} = \langle\hat{\phi}(x), u_t^-(\mathbf{k};x)\rangle_{\mathrm{KG}} = \hat{b}_t^\dagger(\mathbf{k}), \tag{4.37}$$

i.e. the operators $\hat{a}_t(\mathbf{k}), \hat{b}_t^\dagger(\mathbf{k})$ are gauge invariant, and the notion of instantaneous particles we are working with is therefore also gauge invariant (essentially because we forced the modes to transform appropriately under gauge transformations).

This fact will be relevant when considering specific examples for particle creation in an electromagnetic background, where as a consequence of this gauge invariance we may freely choose whatever gauge is convenient for the problem at hand.

4.1.4 Separable mode functions and the temporal gauge condition

In many physical problems of interest, it turns out (with a suitable choice of the $\Sigma(t)$ and the coordinates r) that $\chi_t(\mathbf{k},r) = \chi_{t'}(\mathbf{k},r)\ \forall\ t,t'$ (again, with a suitable choice of $\mathbf{k}$ over different times t). In that case, the mode functions can be separated into a purely time(t)-dependent factor and the time-independent spatial eigenfunctions:

$$u^{\pm}(\mathbf{k};t,r) = f^{\pm}(\mathbf{k};t)\chi(\mathbf{k};r). \tag{4.38}$$

The Klein-Gordon equation Eq. (4.27) then reduces to an ordinary differential equation for the time evolution of $f^{\pm}(\mathbf{k};t)$, of the form:

$$\left(\mathrm{D}_t^2 + (\mathrm{D}_\mu n^\mu)\mathrm{D}_t + \omega_t^2(\mathbf{k})\right) f^{\pm}(\mathbf{k};t) = 0, \tag{4.39}$$

where $(\mathrm{D}_\mu n^\mu)$ must now be a purely time dependent scalar, and D_t must also be purely time-dependent operator. Thus, this will work only if it is possible to choose some $\{\Sigma(t)\}$ in which the metric and gauge potential are homogeneous. We note that the eigenvalues $\omega_t(\mathbf{k})$ corresponding to the spatial eigenfunctions may still be time-dependent.

The initial conditions for positive and negative frequency modes Eq. (4.18) corresponding to the time t_0 are now

$$f_{t_0}^{\pm}(\mathbf{k};t_0) = \frac{1}{\sqrt{2\,\mathrm{Re}\,\Omega_{t_0}(\mathbf{k})}} \tag{4.40}$$

$$\mathrm{D}_t f_{t_0}^{+}(\mathbf{k};t_0) = \frac{-i\Omega_{t_0}(\mathbf{k})}{\sqrt{2\,\mathrm{Re}\,\Omega_{t_0}(\mathbf{k})}}, \tag{4.41}$$

$$\mathrm{D}_t f_{t_0}^{-}(\mathbf{k};t_0) = \frac{i\Omega_{t_0}^{*}(\mathbf{k})}{\sqrt{2\,\mathrm{Re}\,\Omega_{t_0}(\mathbf{k})}}, \tag{4.42}$$

with frequencies $\Omega_{t_0}(\mathbf{k})$. Irrespective of the particular initial conditions, it is straightforward to evaluate the Bogolubov coefficients for such separated modes. Between two times t_2 and t_1, we have

$$B_{t_2,t_1}((s_2,\mathbf{k}),(s_1,\mathbf{k}')) = s_1\langle f_{t_2}^{s_2}(\mathbf{k};t)\chi(\mathbf{k};r), f_{t_1}^{s_1}(\mathbf{k};t)\chi(\mathbf{k};r)\rangle_{\mathrm{KG}}, \tag{4.43}$$

$$= -is_1\left[f_{t_2}^{s_2}(\mathbf{k};t)(\mathrm{D}_t f_{t_1}^{s_1}(\mathbf{k};t))^* - (f_{t_1}^{s_1}(\mathbf{k};t))^*\mathrm{D}_t f_{t_2}^{s_2}(\mathbf{k},t)\right]\delta(\mathbf{k}-\mathbf{k}'). \tag{4.44}$$

Thus, separable modes corresponding to different spatial eigenfunctions do not 'mix' under Bogolubov transformations.

Yet another simplification may be achieved if we enforce (an adaptation of) the so-called temporal gauge condition (see e.g. Ref. [9]), $n^\mu A_\mu = 0$. In that case, D_t has purely real components, and the real and imaginary parts of each $f^{\pm}(\mathbf{k};t)$ do not mix in Eq. (4.39).

Consequentially, the functions $f^+_{t_0}(\mathbf{k},t)$ and $f^-_{t_0}(\mathbf{k},t)$ are complex conjugates of each other given the initial conditions Eq. (4.40) (which for $(+,\mathbf{k})$ and $(-,\mathbf{k})$ are also related by complex conjugation).

For each $\mathbf{k}$, this results in only two independent Bogolubov coefficients, namely

$$B_{t_2,t_1}((+,\mathbf{k}),(+,\mathbf{k}')) = B^*_{t_2,t_1}((-,\mathbf{k}),(-,\mathbf{k}')) = \alpha_{\mathbf{k}}(t_2,t_1)\,\delta(\mathbf{k}-\mathbf{k}'), \tag{4.45}$$

$$B_{t_2,t_1}((+,\mathbf{k}'),(-,\mathbf{k})) = B^*_{t_2,t_1}((-,\mathbf{k}),(+,\mathbf{k}')) = \beta_{\mathbf{k}}(t_2,t_1)\,\delta(\mathbf{k}-\mathbf{k}'), \tag{4.46}$$

where the α- and β- coefficients are given by

$$\alpha_{\mathbf{k}}(t_2,t_1) = -i\left[f^+_{t_2}(\mathbf{k};t)(\mathrm{D}_t f^+_{t_1}(\mathbf{k};t))^* - (f^+_{t_1}(\mathbf{k};t))^*\mathrm{D}_t f^+_{t_2}(\mathbf{k},t)\right], \tag{4.47}$$

$$\beta_{\mathbf{k}}(t_2,t_1) = i\left[f^+_{t_2}(\mathbf{k};t)(\mathrm{D}_t f^-_{t_1}(\mathbf{k};t))^* - (f^-_{t_1}(\mathbf{k};t))^*\mathrm{D}_t f^+_{t_2}(\mathbf{k},t)\right]. \tag{4.48}$$

This can of course be further simplified by replacing the complex conjugates with the conjugate functions. The condition Eq. (3.43) becomes

$$|\alpha_{\mathbf{k}}(t_2,t_1)|^2 - |\beta_{\mathbf{k}}(t_2,t_1)|^2 = 1. \tag{4.49}$$

The t_2 particle/antiparticle number ($\mathbf{k}$-density) in the t_1 vacuum is now

$$\langle 0;t_1|\hat{N}^{\pm}_{t_2}(\mathbf{k})|0;t_1\rangle = |\beta_{\mathbf{k}}(t_2,t_1)|^2\delta(\mathbf{k}-\mathbf{k}). \tag{4.50}$$

This will be the main quantity of interest in many of the examples to be considered. As the Bogolubov coefficients depend multilinearly on the mode functions (via the KG inner product), we may, if we wish, scale the mode functions so that they are no longer normalized (but for each Fock decomposition maintain their relative normalization), leading to scaled coefficients $\tilde{\alpha}_{\mathbf{k}}(t_2,t_1)$ and $\tilde{\beta}_{\mathbf{k}}(t_2,t_1)$ and compensate for this scaling in Eq. (4.50) by dividing by $|\tilde{\alpha}_{\mathbf{k}}(t_2,t_1)|^2 - |\tilde{\beta}_{\mathbf{k}}(t_2,t_1)|^2$ (which reduces to unity for properly normalized mode functions) i.e. by using

$$|\beta_{\mathbf{k}}(t_2,t_1)|^2 = \frac{|\tilde{\beta}_{\mathbf{k}}(t_2,t_1)|^2}{|\tilde{\alpha}_{\mathbf{k}}(t_2,t_1)|^2 - |\tilde{\beta}_{\mathbf{k}}(t_2,t_1)|^2}. \tag{4.51}$$

Alternatively, we can normalize the mode functions corresponding to t by scaling them so that $|\alpha_{\mathbf{k}}(t,t)| = 1$

There is also an intuitive picture that these assumptions allow us to achieve. The expectation value in Eq. (4.50) has a $\delta(\mathbf{k}-\mathbf{k})$ uniformly for all $\mathbf{k}$ rather than a delta function at a finite number of points, making not only the k-space particle/antiparticle number density infinite, but the overall particle content at t_2 as well. This infinite may be attributed to the infinite extent of the spatial slices $\Sigma(t)$. Intuitively, we would like a measure of particle number per unit spatial volume, so that the trivial infinity due to spatial extent can be ignored. As we will see below, there is a natural (though rather heuristic) way of doing this with separable modes.

If we evaluate the expectation value in the t_1-vacuum of the total particle/antiparticle number at t_2,

$$\begin{aligned} \langle 0; t_1 | \hat{\mathcal{N}}^{\pm}_{t_2} | 0; t_1 \rangle &= \int \mathrm{d}\mathbf{k}\ |\beta_{\mathbf{k}}(t_2, t_1)|^2 \\ &= \int\limits_{\Sigma(t_2)} \mathrm{d}^3 r\ \sqrt{-\tilde{g}} \int \mathrm{d}\mathbf{k}\ |\beta_{\mathbf{k}}(t_2, t_1)|^2 |\chi(\mathbf{k}; r)|^2, \end{aligned} \tag{4.52}$$

where in the second line, we have used the orthonormality of the $\chi(\mathbf{k}; r)$ in $\mathrm{L}^2(\Sigma(t_2))$. This suggests that we may associate a spatial density of particles or antiparticles with each point (t_2, r) on $\Sigma(t_2)$, given by

$$n^{\pm}(t_2, r) = \int \mathrm{d}\mathbf{k}\ |\beta_{\mathbf{k}}(t_2, t_1)|^2 |\chi(\mathbf{k}; r)|^2. \tag{4.53}$$

We may also say that $|\beta_{\mathbf{k}}(t_2, t_1)|^2 |\chi(\mathbf{k}; r)|^2$ is the particle/antiparticle density per unit volume of $\Sigma(t_2)$ and per unit volume of $\mathbf{k}$-space (also see e.g. Ref. [14]).

It is unclear if this interpretation can be given a general meaning by directly considering localized field operators, and seems to be of a similar nature to the Wigner function (see e.g. Refs. [15, 16] and the corresponding references therein) defined in phase space in standard quantum mechanics, rather than a quantity that emerges more 'naturally' from the theory. However, we will be content to make use of this interpretation where relevant - in particular, this is the interpretation we will have in mind when preferentially concerning ourselves with $|\beta_{\mathbf{k}}(t_2, t_1)|^2$ as opposed to the actual expectation value of the $\hat{N}^{\pm}(\mathbf{k})$ in later examples.

Finally, we remark that though the assumptions made to get to this point appear (and are) highly restrictive, many problems of interest are compatible with these assumptions and we will extensively use these results in the discussion of specific examples.

4.2 Choosing a vacuum

Now, we will attempt to assign a specific Fock basis to each of the surfaces $\Sigma(t)$ (introduced with the Hamiltonian framework), that we will sometimes call the instantaneous Fock basis. This would give us a definition of particles at each instant of time t, though by no means a unique or even natural one, as it is at the least dependent on the choice of surfaces $\Sigma(t)$, for which there are typically no natural candidates in the most general curved spacetimes. There are multiple ways to do this, and we will discuss two of the popular ones (to which some 'physical meaning' may be attached, to some extent).

4.2.1 Instantaneous Hamiltonian diagonalization

The first way of assigning an instantaneous Fock basis (and arguably one of the more obvious ones) is called the method of 'Instantaneous Hamiltonian Diagonalization' (discussed in e.g. Refs. [2, 5, 6]). While this may be done in a number of ways, we will proceed by constructing a set of mode functions associated with each t, and then show that the Hamiltonian (as obtained from the stress-energy tensor) is diagonal in the corresponding Fock space representation.

Now, we will define the positive/negative frequency modes for this method using the eigenvalues and eigenfunctions of K_t. As we will show, to diagonalize the Hamiltonian, we need to choose $\Omega_t(\mathbf{k}) = \omega_t(\mathbf{k})$, so that these modes are now defined by the boundary conditions

$$\left.\begin{aligned} u_t^{\pm}(\mathbf{k}; x(t,r)) &= \frac{1}{\sqrt{2\omega_t(\mathbf{k})}}\chi_t(\mathbf{k}; x) \\ \mathrm{D}_t u_t^{\pm}(\mathbf{k}; x(t,r)) &= \mp i\sqrt{\frac{\omega_t(\mathbf{k})}{2}}\chi_t(\mathbf{k}; x) \end{aligned}\right\} \forall r \in \Sigma(t). \tag{4.54}$$

The scalar field operator can be expanded in terms of these instantaneous positive and negative frequency modes on some $\Sigma(t)$,

$$\hat{\phi}(x) = \int \mathrm{d}\mathbf{k}\, \left(\hat{a}_t(\mathbf{k})u_t^{+}(\mathbf{k}; x) + \hat{b}_t^{\dagger}(\mathbf{k})u_t^{-}(\mathbf{k}; x)\right). \tag{4.55}$$

The (Hermitian-ized) stress-energy tensor is:

$$\hat{T}_{\mu\nu}(x) = \mathrm{D}_\mu\hat{\phi}^{\dagger}(x)\mathrm{D}_\nu\hat{\phi}(x) + \mathrm{D}_\nu\hat{\phi}^{\dagger}(x)\mathrm{D}_\mu\hat{\phi}(x) - g_{\mu\nu}\left(\mathrm{D}^{\rho}\hat{\phi}^{\dagger}(x)\mathrm{D}_\rho\hat{\phi}(x) - m^2\hat{\phi}^{\dagger}(x)\hat{\phi}(x)\right). \tag{4.56}$$

The Hamiltonian at t is therefore given by the operator version of Eq. (2.44), i.e.

$$\hat{H}(t) = \int\limits_{\Sigma(t)} \mathrm{d}^3 r\, \sqrt{-\tilde{g}}\, \left[\mathrm{D}_t\hat{\phi}^\dagger \mathrm{D}_t\hat{\phi} + \hat{\phi}^\dagger \left(\mathrm{S}^\rho \mathrm{S}_\rho + m^2\right)\hat{\phi}\right]. \tag{4.57}$$

In this, we may substitute the expansion Eq. (4.55) and use the boundary conditions for the modes, then use the eigenvalue equation Eq. (4.25) to get:

$$\begin{aligned}\hat{H}(t) = \int \mathrm{d}\mathbf{k} \int \mathrm{d}\mathbf{k}' \int\limits_{\Sigma(t)} \mathrm{d}^3 r\, \sqrt{-\tilde{g}}\, \chi_t^*(\mathbf{k};x)\chi_t(\mathbf{k}';x) \\ \Big[\left(\hat{a}_t^\dagger(\mathbf{k})\hat{a}_t(\mathbf{k}') + \hat{b}_t(\mathbf{k})\hat{b}_t^\dagger(\mathbf{k}')\right)\omega_t(\mathbf{k}')\left(\omega_t(\mathbf{k}) + \omega_t(\mathbf{k}')\right) \\ + \left(\hat{a}_t^\dagger(\mathbf{k})\hat{b}_t^\dagger(\mathbf{k}') + \hat{b}_t(\mathbf{k})\hat{a}_t(\mathbf{k}')\right)\omega_t(\mathbf{k}')\left(-\omega_t(\mathbf{k}) + \omega_t(\mathbf{k}')\right)\Big]\end{aligned} \tag{4.58}$$

The integral over $\Sigma(t)$ gives $\delta(\mathbf{k}-\mathbf{k}')/2\omega_t(\mathbf{k})$, and $\mathbf{k}'$ can then be integrated over to get the expression:

$$\hat{H}(t) = \int \mathrm{d}\mathbf{k}\, \left(\hat{a}_t^\dagger(\mathbf{k})\hat{a}_t(\mathbf{k}) + \hat{b}_t(\mathbf{k})\hat{b}_t^\dagger(\mathbf{k})\right)\omega_t(\mathbf{k}). \tag{4.59}$$

In terms of the instantaneous particle and antiparticle number operators $\hat{N}_t^+(\mathbf{k}) = \hat{a}_t^\dagger(\mathbf{k})\hat{a}_t(\mathbf{k})$ and $\hat{N}_t^-(\mathbf{k}) = \hat{b}_t^\dagger(\mathbf{k})\hat{b}_t(\mathbf{k})$,

$$\hat{H}(t) = \int \mathrm{d}\mathbf{k}\, \left(\hat{N}_t^+(\mathbf{k}) + \hat{N}_t^-(\mathbf{k}) + \delta(\mathbf{k}-\mathbf{k})\right)\omega_t(\mathbf{k}). \tag{4.60}$$

The Hamiltonian is therefore diagonal in the Fock space decomposition corresponding to $\hat{N}_t^+(\mathbf{k})$ as defined here; this expression is also reminiscent of the classical understanding of particles like photons as 'quanta of energy'.

4.2.2 The adiabatic vacuum for separable modes

While Hamiltonian diagonalization certainly appears reasonable, it is sometimes known to give an infinite density of particles even for slowly varying backgrounds, in some anisotropic metrics (see, for instance, the discussion in Ref. [2]). In such cases, it is argued that Hamiltonian diagonalization does not really furnish a useful definition of particle number, and one might try to define particles another way that gives finite answers that one can use to understand the system under consideration better.

The main candidate is called the adiabatic particle number, associated with an adiabatic vacuum (see e.g. Refs. [5, 6] and references therein). The starting point is to define the so-called adiabatically evolved modes, or (necessarily) approximate solutions to the Klein-Gordon equation assuming that the background changes slowly in time (as given by t). For separable modes Eq. (4.38) (with the spatial part being Klein-Gordon spatial eigenfunctions), the levels of approximation are usually treated systematically using the WKB method.

We will now quickly review this method (largely following the presentation in Ref. [17]). The time-dependent factor $f(t)$ in the mode functions satisfies Eq. (4.39)

$$\left(\mathrm{D}_t^2 + (\mathrm{D}_\mu n^\mu)\mathrm{D}_t + \omega_t^2(\mathbf{k})\right) f^\pm(\mathbf{k};t) = 0. \tag{4.61}$$

It is convenient to transform this equation to the following form[2] by writing $f^\pm(\mathbf{k};t) = p(t)\tilde{f}^\pm(\mathbf{k};t)$ for some suitable $p(t)$:

$$\left(\mathrm{d}_t^2 + \widetilde{\omega}_t^2(\mathbf{k})\right) \tilde{f}^\pm(\mathbf{k};t) = 0, \tag{4.62}$$

where $\mathrm{d}_t = \mathrm{d}/\mathrm{d}t$, so that the first derivative term disappears and the usual WKB method (as in Ref. [17]) for an oscillator with a time-dependent frequency is readily applied. It is useful to note that $\mathrm{D}_t = n^t \mathrm{d}_t$ in a temporal gauge, showing a relative scaling factor equal to the 'time' component n^t of the normal vector between these two forms. Additionally, as $p(t) = 1$ for Minkowski spacetime, and $(n^t)\widetilde{\omega}_t(\mathbf{k}) = \omega_t(\mathbf{k})$

Now, we are especially interested in the case where $\widetilde{\omega}_t(\mathbf{k})$ varies slowly as a function of the time t. For bookkeeping purposes, we replace the time derivatives d_t by $\varepsilon\mathrm{d}_t$ (and consequently time differentials $\mathrm{d}t$ by $\mathrm{d}t/\varepsilon$; equivalently, we scale $t \to t/\varepsilon$), where powers of $\varepsilon \ll 1$ will be used to keep track of how 'slowly' something is varying. Of course, after we have made all the arguments we need to using orders of ε, we will formally set $\varepsilon = 1$. The above differential equation then becomes

$$\left(\varepsilon^2\mathrm{d}_t^2 + \widetilde{\omega}_t^2(\mathbf{k})\right) \tilde{f}^\pm(\mathbf{k};t) = 0. \tag{4.63}$$

If we had $\mathrm{d}_t\widetilde{\omega}_t(\mathbf{k}) = 0$ i.e. $\widetilde{\omega}_t(\mathbf{k}) = \widetilde{\omega}_0(\mathbf{k})$, then the positive frequency solution is (up to a constant phase factor):

$$\tilde{f}_0^+(\mathbf{k};t) = \frac{1}{\sqrt{2\widetilde{\omega}_0(\mathbf{k})}} e^{-i\widetilde{\omega}_0(\mathbf{k})t/\varepsilon} = \frac{1}{\sqrt{2\widetilde{\omega}_0(\mathbf{k})}} e^{-\frac{i}{\varepsilon}\int^t \mathrm{d}t\ \widetilde{\omega}_0(\mathbf{k})}. \tag{4.64}$$

[2] We may always find a transformation that does this for any (well-behaved) second order linear differential equation. See e.g. Exercise 9.6.11 in Ref. [18].

where the lower limit of the integral is irrelevant and can be absorbed into the overall phase of this solution. Inspired by this form, one tries what is known as the WKB ansatz

$$\tilde{f}_n^+(\mathbf{k};t) = \frac{1}{\sqrt{2W_n(\mathbf{k};t)}} \exp\left(-\frac{i}{\varepsilon}\int_{-\infty}^{t} \mathrm{d}t'\, W_n(\mathbf{k};t')\right). \tag{4.65}$$

The $n \in \mathbb{N}_0$ is to keep track of the order of approximation, such that $W_n(\mathbf{k};t) = O(\varepsilon^n)$. Substituting this ansatz in the differential equation for the mode functions gives the following relation for $W_n(\mathbf{k};t)$:

$$W_n^2(\mathbf{k};t) + \varepsilon^2 \left(\frac{\mathrm{d}_t^2 W_n(\mathbf{k};t)}{2W_n(\mathbf{k};t)} - \frac{3}{4}\left(\frac{\mathrm{d}_t W_n(\mathbf{k};t)}{W_n(\mathbf{k};t)}\right)^2\right) = \widetilde{\omega}_t^2(\mathbf{k}). \tag{4.66}$$

Starting with $W_0(\mathbf{k};t) = \widetilde{\omega}_t(\mathbf{k})$, this can be used to recursively obtain $W_n(\mathbf{k};t)$ for $n \in \mathbb{N}$, and each recursion gives a term $O(\varepsilon^2)$ higher. Thus, we have $W_{2n+1}(\mathbf{k};t) = 0 \forall n \in \mathbb{N}_0$, and

$$W_{2n+2}(\mathbf{k};t) = \left.\sqrt{\widetilde{\omega}_t^2(\mathbf{k}) - \varepsilon^2 \left(\frac{\mathrm{d}_t^2 W_n(\mathbf{k};t)}{2W_n(\mathbf{k};t)} - \frac{3}{4}\left(\frac{\mathrm{d}_t W_n(\mathbf{k};t)}{W_n(\mathbf{k};t)}\right)^2\right)}\,\right|_{O(\varepsilon^{2n+2})}. \tag{4.67}$$

It is worth noting that the WKB approximation does not converge to actual solutions beyond a certain order (see e.g. Ref. [17]) and we must not exceed this order to obtain a reasonable description of the solutions.

The time derivative of $\tilde{f}_n^+(\mathbf{k},t)$ is of interest; we find that

$$\varepsilon \mathrm{d}_t \tilde{f}_n^+(\mathbf{k},t) = -i\left(W_n(\mathbf{k};t) - i\varepsilon\frac{\mathrm{d}_t W_n(\mathbf{k};t)}{2W_n(\mathbf{k};t)}\right)\tilde{f}_n^+(\mathbf{k},t). \tag{4.68}$$

Therefore, the normal time derivative of the corresponding original mode function $f^+(\mathbf{k};t) = p(t)\tilde{f}^+(\mathbf{k};t)$ is, using $\mathrm{D}_t = n^t \mathrm{d}_t$,

$$\varepsilon \mathrm{D}_t f_n^+(\mathbf{k},t) = -in^t\left(W_n(\mathbf{k};t) - i\varepsilon\frac{\mathrm{d}_t W_n(\mathbf{k};t)}{2W_n(\mathbf{k};t)} - i\varepsilon\frac{\mathrm{d}_t p(t)}{p(t)}\right) f_n^+(\mathbf{k},t). \tag{4.69}$$

We define the modes corresponding to the n-th order adiabatic vacuum at a time t by requiring them to have a frequency identical to the instantaneous frequency of the (at least n-th order) WKB approximation to $O(\varepsilon^n)$ (see e.g. Ref. [6]) i.e.

$$\Omega_t^{(n)}(\mathbf{k}) = n^t\left(W_n(\mathbf{k};t) - i\varepsilon\frac{\mathrm{d}_t W_n(\mathbf{k};t)}{2W_n(\mathbf{k};t)} - i\varepsilon\frac{\mathrm{d}_t p(t)}{p(t)}\right)_{O(\varepsilon^n)}. \tag{4.70}$$

We trivially have

$$\Omega_t^{(0)}(\mathbf{k}) = n^t \left(\widetilde{\omega}_t(\mathbf{k}) + i\varepsilon \frac{\mathrm{d}_t p(t).}{p(t)} \right) \tag{4.71}$$

We note that for large eigenvalues $\omega_t^2(\mathbf{k})$ of the mode of interest (compared to $\mathrm{D}_\mu n^\mu$), we may essentially neglect contributions to $\Omega_t^{(0)}(\mathbf{k})$ from $p(t)$ and its time derivatives and obtain

$$\Omega_t^{(0)}(\mathbf{k}) \approx \omega_t(\mathbf{k}), \tag{4.72}$$

which shows that the zeroth order adiabatic vacuum prescription virtually agrees with Hamiltonian diagonalization for these modes, and the difference is appreciable only for low-eigenvalue modes (which tend to have a low rate of spatial variation, which is detectable only at the largest scales). Typically, this would mean that the two prescriptions (at zeroth order) are indistinguishable at length scales or inverse mass scales[3] much smaller than the length scale of the curvature of spacetime. Of course, there is an exact equality in Minkowski spacetime as $p(t) = 1$.

It is also straightforward to write down $\Omega_t^{(1)}(\mathbf{k})$, whereas the higher order terms have increasingly complicated expressions in terms of $\widetilde{\omega}_t(\mathbf{k})$. The only contribution to this order is from the $O(\varepsilon^0)$ term in $W_n(\mathbf{k}; t)$, i.e. $W_0(\mathbf{k}; t) = \widetilde{\omega}_t(\mathbf{k})$, so we get

$$\Omega_t^{(1)}(\mathbf{k}) = n^t \left(\widetilde{\omega}_t(\mathbf{k}) - i\varepsilon \frac{\mathrm{d}_t \widetilde{\omega}_t(\mathbf{k})}{2\widetilde{\omega}_t(\mathbf{k})} - i\varepsilon \frac{\mathrm{d}_t p(t)}{p(t)} \right). \tag{4.73}$$

The adiabatic vacuum is especially important when considering expectation values of local field operators, where one subtracts the expectation value of the operator in the adiabatic vacuum for that time to regularize divergences in what is called the method of adiabatic regularization (as discussed in, for instance, Ref. [5]). A more comprehensive comparison of the Hamiltonian diagonalization and adiabatic prescriptions, which argues strongly in favour of the latter, is presented in Ref. [2].

4.3 Numerical evaluation of particle production

We will assume that the simplifying assumptions discussed in Sec. 4.1.4 hold, and will be concerned with the numerical evaluation of $|\beta_{\mathbf{k}}(t_2, t_1)|^2$. This is readily done by numerically solving the differential equation Eq. (4.39) with the appropriate initial conditions corresponding to the desired choice of vacuum states. We will consider both the Hamiltonian

[3] We recall that $\omega_t(\mathbf{k}) \geq m$, see Eq. (3.39).

diagonalizing and adiabatic vacuum prescriptions. The Bogolubov coefficients can then be obtained using Eqs. (4.47) and (4.48), normalizing the functions corresponding to time t by dividing them by $\sqrt{|\alpha_{\mathbf{k}}(t,t)|}$ (the phase being irrelevant).

The plots corresponding to various examples in subsequent chapters have been obtained by implementing the above outline in Mathematica[4].

[4]Software citation: Ref. [19].

Chapter 5

The Schwinger effect in flat spacetime

We will begin with the Schwinger effect in Minkowski spacetime (which exhibits translation invariance, in addition to rotational and Lorentz boost symmetries), with the standard Minkowski metric $\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}\mathbf{x}^2$ in 'inertial coordinates' $(t, \mathbf{x})$. The most straightforward problems involve spatially homogeneous electric fields (we are obviously choosing a particular frame of reference in which this holds). In particular, the case of the constant electric field and a Sauter ($E \sim \mathrm{sech}^2(at)$) pulsed field are known to have exact solutions, and we will discuss both of them. We will also briefly discuss the effect of a pure magnetic field background.

Throughout this chapter, the surfaces $\Sigma(t)$ will be chosen as surfaces of constant inertial time t (so the two ts virtually refer to the same thing here), with the coordinates r corresponding to the inertial spatial coordinates $\mathbf{x}$. The normal vectors then correspond to the direction of inertial time, $n^\mu = \delta^\mu_t$. Thus, the adiabatic vacuum prescription may be applied directly without transforming the equation of motion, as discussed in Sec. 4.2.2.

5.1 A constant electric field

The simplest (and classic) example of the Schwinger effect corresponds to a constant electric field background $\mathbf{E} = (E, 0, 0)$ (say, after a suitable global 3-rotation) with a vanishing magnetic field $\mathbf{B} = \mathbf{0}$ (see e.g. Refs. [20, 21]). We have specified the precise components of the electric field for definiteness - general results may be readily recovered from rotation invariance.

5.1.1 Exact solution in a temporal gauge

We may choose any among a family of potentials (related by gauge transformations) that lead to the above field strength, in attempting a concrete solution to the problem. But as we have seen in Sec. 4.1.4, imposing the temporal gauge condition $n^\mu A_\mu = 0$ may lead to a considerable simplification of the Bogolubov coefficients.

We then choose (the covariant components of) our potential to be, in the inertial coordinates, $A_\mu = (0, Et, 0, 0)$. The Klein-Gordon operator for a massive scalar field for this choice of gauge is then

$$\left(\mathrm{D}_\mu \mathrm{D}^\mu + m^2\right) = \left(\frac{\partial^2}{\partial t^2} - \left(\frac{\partial}{\partial x} - iqEt\right)^2 - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + m^2\right). \tag{5.1}$$

The eigenfunctions of the spatial Klein-Gordon operator,

$$\mathrm{K}_t = \left(-\left(\frac{\partial}{\partial x} - iqEt\right)^2 - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} + m^2\right), \tag{5.2}$$

turn out to be the standard (static) wave exponentials (normalized with integration measure d^3k),

$$\chi(\mathbf{k}; \mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \forall\, \mathbf{k} \in \mathbb{R}^3, \tag{5.3}$$

with time-dependent eigenvalues,

$$\omega_t^2(\mathbf{k}) = (k_x - qEt)^2 + k_y^2 + k_z^2 + m^2. \tag{5.4}$$

The positive and negative frequency modes for a time t_0 may therefore be written in a product form

$$u_{t_0}^{\pm}(\mathbf{k}; (t, \mathbf{x})) = f_{t_0}^{\pm}(\mathbf{k}; t)\chi(\mathbf{k}; \mathbf{x}). \tag{5.5}$$

The combination of the facts that the temporal gauge condition is satisfied and the positive/negative frequency mode functions are separable allow us to take advantage of the simplifications derived in Sec. 4.1.4.

To begin with, the functions $f_{t_0}^{\pm}(\mathbf{k}; t)$ are solutions (in place of the more generic $f(t)$) of the following differential equation derived from Eq. (4.39),

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} + \left[(k_x - qEt)^2 + k_y^2 + k_z^2 + m^2\right] f = 0. \tag{5.6}$$

Being a second order differential equation, this has two linearly independent solutions. It is conventional to define the two dimensionless quantities (as done in e.g. Ref. [20, 21]):

$$\tau = \sqrt{|qE|}\left(t - \frac{k_x}{qE}\right), \tag{5.7}$$

$$\lambda = \frac{k_y^2 + k_z^2 + m^2}{|qE|}, \tag{5.8}$$

which gives

$$\frac{\mathrm{d}^2 f}{\mathrm{d}\tau^2} + (\tau^2 + \lambda) f = 0. \tag{5.9}$$

Instead of considering just two independent solutions, it will be convenient to consider four different standard solutions, given by the Whittaker forms of the parabolic cylinder functions (see e.g. 9.255(2) of Ref. [7], and Ref. [20])

$$(D_p((1+i)\tau), D_{p^*}((1-i)\tau)); (D_p(-(1+i)\tau), D_{p^*}(-(1-i)\tau)), \tag{5.10}$$

where $p = -\frac{1+i\lambda}{2}$. The functions $D_p(z)$ are entire functions in both z and p, and this will be used implicitly in subsequent arguments (see e.g. 14.2 of Ref. [22]). We have written these solutions out in pairs - as τ is real valued, each pair of solutions is a pair of complex conjugates.

Of particular interest is the asymptotic forms of these functions for large $|\tau|$. In general, we have (see e.g. 9.246 (1) of Ref. [7])

$$D_p(z) \sim z^p e^{-\frac{z^2}{4}}, \text{ for } |z| \to \infty, |\arg z| < \frac{3\pi}{4}. \tag{5.11}$$

The restriction on the phase of z allows us to use this asymptotic form only for $\tau \to +\infty$ for $D_p((1+i)\tau)$, and for $\tau \to -\infty$ for $D_p(-(1+i)\tau)$.

$$D_p((1+i)\tau) \sim ((1+i)\tau)^{-\frac{1+i\lambda}{2}} e^{-\frac{i}{2}\tau^2}, \ \tau \to \infty, \tag{5.12}$$

$$D_p(-(1+i)\tau) \sim (-(1+i)\tau)^{-\frac{1+i\lambda}{2}} e^{-\frac{i}{2}\tau^2}, \ \tau \to -\infty. \tag{5.13}$$

It is also of interest to consider the asymptotic behaviour of the eigenvalues of the spatial Klein-Gordon operator,

$$\omega_t(\mathbf{k}) = \sqrt{|qE|}\sqrt{\tau^2 + \lambda} \to |\tau|, \text{ for } \tau \to \pm\infty. \tag{5.14}$$

This means that (since $\mathrm{d}\tau = \sqrt{qE}\mathrm{d}t$),

$$\frac{\mathrm{d}D_p((1+i)\tau)}{\mathrm{d}t} \sim -i\omega_t(\mathbf{k})D_p((1+i)\tau), \text{ for } \tau \to \infty, \tag{5.15}$$

$$\frac{\mathrm{d}D_p(-(1+i)\tau)}{\mathrm{d}t} \sim i\omega_t(\mathbf{k})D_p(-(1+i)\tau), \text{ for } \tau \to -\infty, \tag{5.16}$$

where we have retained only the leading order contribution in τ (which is due to the exponential phase factor $e^{-\frac{i}{2}\tau^2}$).

As $\frac{1}{\omega_t(\mathbf{k})}\frac{\mathrm{d}\omega_t(\mathbf{k})}{\mathrm{d}t} \ll \omega_t(\mathbf{k})$ for $|\tau| \to \infty$, we see that $D_p((1+i)\tau)$ is the appropriate positive frequency mode for $\tau \to \infty$ and $D_p(-(1+i)\tau)$ is the negative frequency mode for $\tau \to -\infty$ according to both the Hamiltonian diagonalization and adiabatic particle number criteria for choosing a vacuum. Correspondingly, their respective complex conjugates give appropriate modes in these limits with the opposite sign of frequency. For a particular $\mathbf{k}$, $\tau \to \infty$ also corresponds respectively to $t \to \infty$, and these modes correspond to suitable Fock spaces at asymptotically early and late times. We will soon also consider the situation at intermediate times.

The (asymptotic) Bogolubov coefficients $\alpha_{\mathbf{k}}(\infty, -\infty)$ and $\beta_{\mathbf{k}}(\infty, -\infty)$ may be read off (up to the same undetermined factor multiplying each of them) directly from the following relation (see e.g. 9.248(1) of Ref. [7])

$$D_p(z) = e^{i\pi p}D_p(-z) + \frac{\sqrt{2\pi}}{\Gamma(-p)}e^{i\pi(p+1)/2}D_{-p-1}(-iz). \tag{5.17}$$

Substituting $z = -(1+i)\tau$, noting that for our $-1-p = -(1-i\lambda)/2 = p^*$ and re-arranging the terms a bit, we get a relation expressing the late time positive frequency mode as a linear combination of the early time modes,

$$D_p((1+i)\tau) = -\frac{\sqrt{2\pi}}{\Gamma(-p)}e^{i\pi(1-p)/2}D_{p^*}(-(1-i)\tau) + e^{-i\pi p}D_p(-(1+i)\tau). \tag{5.18}$$

Thus,

$$\alpha_{\mathbf{k}}(\infty, -\infty) = -c\frac{\sqrt{2\pi}}{\Gamma(-p)}e^{i\pi(1-p)/2} = c\frac{\sqrt{2\pi}}{\Gamma\left(\frac{1+i\lambda}{2}\right)}e^{-\pi\lambda/4}, \tag{5.19}$$

$$\beta_{\mathbf{k}}(\infty, -\infty) = ce^{-i\pi p} = ice^{-\pi\lambda/2}, \tag{5.20}$$

where c is the undetermined constant referred to.

Using $|\Gamma\left(\frac{1}{2} + ix\right)|^2 = \frac{\pi}{\cosh \pi x}$ (see e.g. 5.4.3 in Ref. [22]),

$$|\alpha_{\mathbf{k}}(\infty, -\infty)|^2 = |c|^2\left(1 + e^{-\pi\lambda}\right) \text{ and } |\beta_{\mathbf{k}}(\infty, -\infty)|^2 = |c|^2e^{-\pi\lambda} \tag{5.21}$$

Clearly, the condition $|\alpha_{\mathbf{k}}(\infty,-\infty)|^2 - |\beta_{\mathbf{k}}(\infty,-\infty)|^2$ implies that $|c| = 1$. Each of the modes for a specific $\mathbf{k}$ (at most) differ from their correct normalization by a factor of the same magnitude, leading to correctly normalized Bogolubov coefficients. The particle/antiparticle density in $\mathbf{k}$ in the asymptotic future, for the vacuum state of the asymptotic past, is then

$$\begin{aligned}\langle 0,-\infty|\hat{N}^{\pm}_{+\infty}(\mathbf{k})|0,-\infty\rangle &= |\beta_{\mathbf{k}}(\infty,-\infty)|^2\delta(\mathbf{k}-\mathbf{k}) \\ &= e^{-\pi\lambda}\delta(\mathbf{k}-\mathbf{k}) \\ &= \exp\left(-\pi\frac{k_y^2+k_z^2+m^2}{|qE|}\right)\delta(\mathbf{k}-\mathbf{k}). \end{aligned} \tag{5.22}$$

5.1.2 An approximate picture of the pair creation process

We have obtained in Sec. 5.1.1 an (asymptotically) exact solution for the number of particles (in principle) generated from a vacuum at the infinite past till the infinite future by a constant electric field. However, the same solutions also provide some insight about the process at intermediate times, which we will now discuss.

In Minkowski spacetime, we can define a meaningful global 4-momentum for the field,

$$\hat{P}^{\mu}(t) = \int_t \mathrm{d}^3\mathbf{x}\,\hat{T}^{\mu t}. \tag{5.23}$$

The $\mu = t$ component is just the Hamiltonian, the expression for which depends on our criterion for defining the modes at each time. In any case, it turns out that the spatial components depend on the particle number (for plane wave modes) without explicitly depending on the choice of instantaneous vacua. Expanding the field in terms of the plane wave modes appropriate for each t, $u_t^{\pm}(\mathbf{k};(t,\mathbf{x})) = f_t^{\pm}(\mathbf{k};t)(2\pi)^{-3/2}e^{i\mathbf{k}\cdot\mathbf{x}}$, and normal-ordering with respect to the chosen vacuum at t, we get, for the spatial components

$$:P^i:_t(t) = \int \mathrm{d}\mathbf{k}\,(k^i - qA^i(t))\left(\hat{N}_t^{+}(\mathbf{k}) - \hat{N}_t^{-}(\mathbf{k})\right). \tag{5.24}$$

Our main takeaway from this expression is that each particle corresponding to a mode $\mathbf{k}$ contributes $p^i(\mathbf{k}) = k^i - qA^i$, and each antiparticle corresponding to $\mathbf{k}$ has the opposite momentum $-p^i(\mathbf{k})$.

Returning to the parabolic cylinder functions $D_p((1+i)\tau)$, $D_p(-(1+i)\tau)$ and their complex conjugates, we will consider the problem of treating the asymptotic forms Eqs. (5.12) and

(5.13) as approximations of the functions themselves. Let us just assume that we have a quantitative sense of what constitutes a good approximation, in which case we know that the asymptotic forms will be reasonable approximations in some interval $\tau \in \mathbb{R} - (\tau_\Delta^-, \tau_\Delta^+) : \tau_\Delta^- < 0 < \tau_\Delta^+)$ i.e. except in some neighbourhood $\Delta = (\tau_\Delta^-, \tau_\Delta^+)$ of $\tau = 0$ (the size of this neighbourhood depends on our tolerance for the approximation). In that case, for each $\mathbf{k}$, the functions $D_p((1+i)\tau), D_{p^*}((1-i)\tau)$ are (as good as) the appropriate positive and negative frequency modes for $\tau < \tau_\Delta^-$, and the functions $D_{p^*}(-(1-i)\tau), D_p(-(1+i)\tau)$ are (as good as) the appropriate positive and negative frequency modes for $\tau > \tau_\Delta^+$.

In what follows, it is useful to remember that τ depends linearly on the time t, and while we will discuss evolution in τ for mathematical convenience, it can be readily translated (as we soon will) to evolution in t.

For notational brevity, we use $n(\mathbf{k}) = |\beta_{\mathbf{k}}(+\infty, -\infty)|^2$, corresponding to $\mathbf{k}$ in the infinite future given a vacuum state in the infinite past. Due to the separability of the mode functions, there is no 'mixing' between instantaneous Fock basis elements in the state space for different $\mathbf{k}$s. The approximation discussed above then paints the following picture: the $\mathbf{k}$ subspace is in the instantaneous vacuum state until $\tau = \tau_\Delta^-$, pair creation occurs in the neighbourhood Δ of $\tau = 0$, and for $\tau > \tau_\Delta^+$ has evolved into a fixed superposition of the vacuum and other elements of the instantaneous Fock basis in the $\mathbf{k}$ subspace, with average particle and antiparticle number $n(\mathbf{k})$ - essentially a superposition of states with various numbers of created particle/antiparticle pairs (of course, in the Heisenberg picture, the state itself does not change with time (via τ), just its components in the instantaneous Fock basis).

Now, we may translate this picture to the time t rather than the dimensionless parameter τ. The main result of this translation is that particle creation in the $\mathbf{k}$ subspace occurs around $t = t_c(\mathbf{k}) = k_x/(qE)$, within an interval of time $|\Delta_t| = |qE|^{-\frac{1}{2}}|\Delta|$ ($|\Delta|$ could itself be a function of qE). If we are interested in large enough time scales compared to Δ_t, we may treat the particle creation as if it happens almost instantaneously around $t = t_c(\mathbf{k})$.

The components of the momentum of $\mathbf{k}$-particles at time t is given by

$$\begin{aligned} \mathbf{p}_+(\mathbf{k}) &= (k_x - qA_x(t), k_y, k_z) \\ &= (-qE(t - t_c(\mathbf{k})), k_y, k_z). \end{aligned} \tag{5.25}$$

This can be interpreted as these particles experiencing an acceleration of $-qE/m$, corresponding to the classical motion of a particle in a uniform electric field. The momentum of the antiparticle is of course oppositely directed. The components (k_x, k_y, k_z) of $\mathbf{k}$ therefore mean different things: while k_y and k_z do correspond to the momentum components of the particle in their respective directions, k_x measures the time at which particle/antiparticle pairs will be created for $\mathbf{k}$. As we can observe only gauge invariant quantities, what an observer might see *on average* is a uniform rate of generation of particle/antiparticle pairs at rest in the x-direction, which are subsequently accelerated according to the classical law of motion.

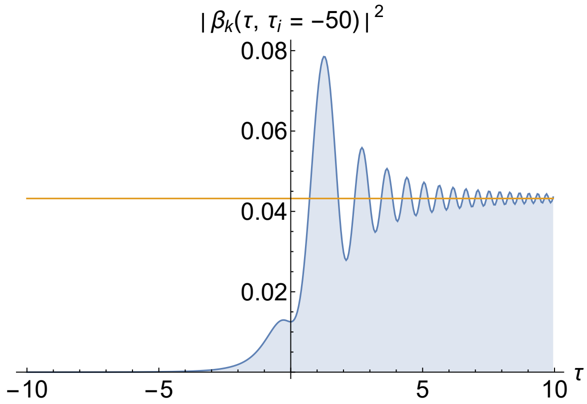
While this is a fairly intuitive picture, like most quantum mechanical effects it has its non-intuitive aspects. While the pairs are *typically* at rest in the x direction when they are created, they may have nonzero k_y,k_z. However, as $n(\mathbf{k}) = \exp\left(-\pi \frac{k_y^2+k_z^2+m^2}{|qE|}\right)$, nonzero 'transverse' momenta are suppressed, occurring in a Gaussian distribution centered at $k_y = k_z = 0$, with width $\sim \sqrt{|qE|}$. There is also already some particle creation (in the sense of a superposition of instantaneous vacuum and non-vacuum states) much before $t_c(\mathbf{k})$ and much after, as the asymptotic forms are only approximations after all - these particles, though detected only with low probability, have nonzero initial momenta along the x-direction as well.

5.1.3 Particle content at intermediate times

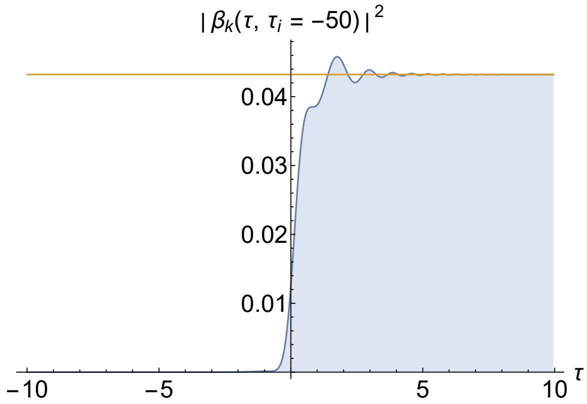
Now, we will discuss some plots obtained by numerically evaluating $|\beta_{\mathbf{k}}(t_2, t_1)|^2$ for intermediate times (i.e. within Δ), with vacua chosen from the zeroth order adiabatic/Hamiltonian diagonalizing vacuum (see also Ref. [15]) to the third order adiabatic vacuum, cf. Fig. 5.1.

We see that while there are oscillations in the average particle number density with the Hamiltonian diagonalizing Fock space, these oscillations are highly suppressed in the higher order adiabatic Fock spaces, almost fitting an intuitive picture one might have of the particle number monotonically rising to its asymptotic late time value.

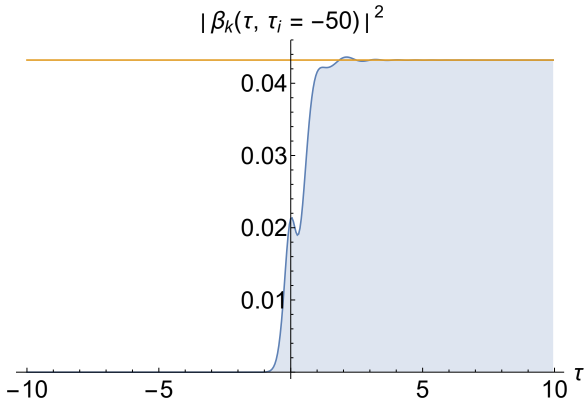
On the other hand, we see the growth of a peak near $\tau = 0$ with successive orders of the adiabatic definition. We may speculate that this is probably due to the appropriate adiabatic condition (i.e. slow variation of the frequency) not holding near $\tau = 0$ for some higher derivatives of $\omega_t(\mathbf{k})$, or the WKB approximation no longer converging in this region. As the variation is required to be 'slow' in comparison with the value of $\omega_t(\mathbf{k})$, we may compare the particle content with a mode that has a higher eigenvalue for the spatial Klein-Gordon operator. One such (rather qualitative) comparison, Fig. 5.2, shows that the peak is not as



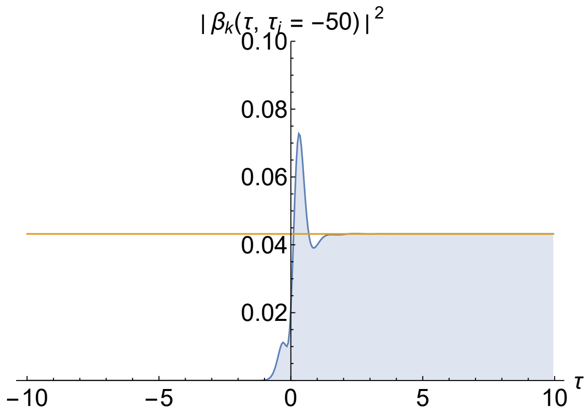
(a) Hamiltonian diagonalization



(b) Adiabatic, first order

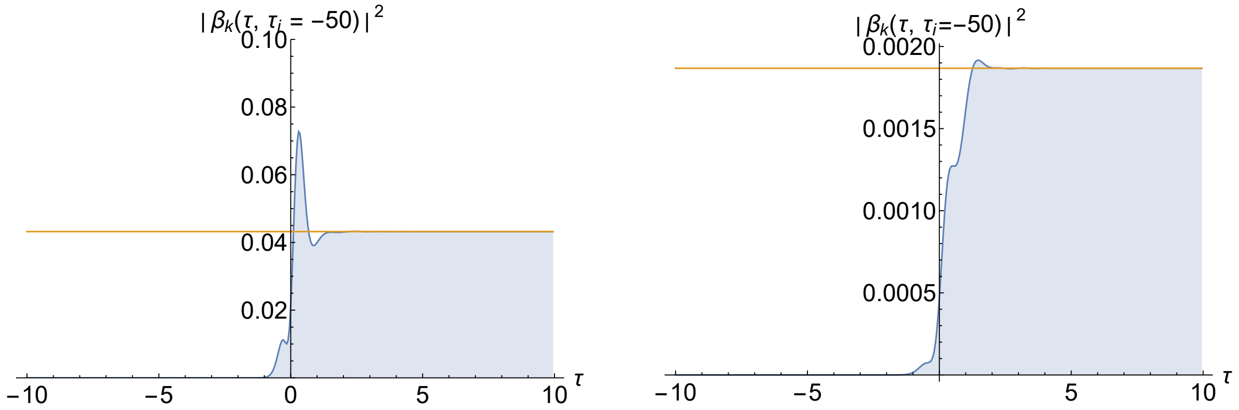


(c) Adiabatic, second order



(d) Adiabatic, third order

Figure 5.1: Particle content at intermediate times for modes with $k_x = 0$, $k_y^2 + k_z^2 + m^2 = 1$, and $qE = 1$ for the constant background electric field, all in arbitrary units. The filled curve (blue) is the curve of interest, while the horizontal line shows the asymptotic late time value of $|\beta_{\mathbf{k}}(t_2, t_1)|^2$ for the corresponding mode.



(a) $k_y^2 + k_z^2 + m^2 = 1$

(b) $k_y^2 + k_z^2 + m^2 = 2$

Figure 5.2: Comparison of particle content in intermediate times with third order adiabatic Fock spaces in two different modes, with $k_x = 0$ and $qE = 1$

prominent in the mode with a larger eigenvalue.

5.2 The Sauter pulsed electric field

5.2.1 Solutions and asymptotic particle content

A spatially homogeneous Sauter pulse (with a suitable choice of coordinates) is an electric field with a sech^2 time dependence i.e. of the form $\mathbf{E} = (E \operatorname{sech}^2(t/\tau), 0, 0)$, with τ here being a constant that represents a measure of the width of the pulse in time (see, for instance, Ref. [23] and the relevant references therein). We choose a potential that satisfies the temporal gauge condition: $A = (0, E\tau \tanh(t/\tau), 0, 0)$. Then, the eigenvalues of the spatial Klein-Gordon equation are $\omega_t^2(k) = \left[(k_x - qE\tau \tanh(t/\tau))^2 + k_y^2 + k_z^2 + m^2\right]$. We note that as $t \to \pm\infty$, these eigenvalues approach a constant value (following the behaviour of $\tanh(t/\tau)$) and both the Hamiltonian diagonalization and adiabatic vacuum prescriptions for choosing a Fock basis agree in these limits.

Once again, we may choose plane waves $\chi(\mathbf{k}; \mathbf{x}) = (2\pi)^{-\frac{3}{2}} e^{i\mathbf{k}\cdot\mathbf{x}}$, and the differential equation Eq. (4.39) resulting for the modes $f^{\pm}(\mathbf{k}; t)$ (written generically as f) is then

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} + \left[(k_x - qE\tau \tanh(t/\tau))^2 + k_y^2 + k_z^2 + m^2\right] f = 0. \tag{5.26}$$

It is convenient to expand this out and use the (hyperbolic) trigonometric relation $\operatorname{sech}^2 x =$

$1 + \tanh^2 x$, to get

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} + (k^2 + m^2 + (qE\tau)^2) f - \left(2k_x qE\tau \tanh(t/\tau) + q^2E^2\tau^2 \operatorname{sech}^2(t/\tau)\right) f = 0. \tag{5.27}$$

We note that, except for the particulars of the parameters in the equation, this has the same form as e.g. Equation 12.3.22 in Ref. [24]. We therefore try the same method of solution used in Ref. [24].

We introduce complex constants a and b which satisfy

$$a^2 + b^2 = -(k^2 + m^2 + q^2E^2\tau^2)\tau^2, \tag{5.28}$$

$$2ab = 2k_x qE\tau^3. \tag{5.29}$$

(as to why will become apparent soon). Adding and subtracting these two equations gives

$$\begin{aligned}(a \pm b)^2 &= -\tau^2 \left((k_x \mp qE\tau)^2 + k_y^2 + k_z^2 + m^2\right) \\ &= -\tau^2 \omega_{\pm\infty}^2(\mathbf{k}).\end{aligned} \tag{5.30}$$

We choose

$$a = \frac{1}{2} i\tau \left(\omega_{+\infty}(\mathbf{k}) + \omega_{-\infty}(\mathbf{k})\right), \tag{5.31}$$

$$b = \frac{1}{2} i\tau \left(\omega_{+\infty}(\mathbf{k}) - \omega_{-\infty}(\mathbf{k})\right). \tag{5.32}$$

Now, substituting $f(t) = e^{-at/\tau} \operatorname{sech}^b(t/\tau) \tilde{f}(t/\tau)$ gives, after some extensive simplification,

$$\tau^2 \frac{\mathrm{d}^2 \tilde{f}}{\mathrm{d}t^2} - 2\tau \left(a + b\tanh(t/\tau)\right) \frac{\mathrm{d}\tilde{f}}{\mathrm{d}t} - \left(b(b+1) + q^2E^2\tau^4\right) \operatorname{sech}^2(t/\tau) \tilde{f} = 0. \tag{5.33}$$

Further substituting $u = \frac{1}{2}\left(1 - \tanh(t/\tau)\right)$ gives a well known differential equation,

$$u(1-u)\frac{\mathrm{d}^2 \tilde{f}}{\mathrm{d}u^2} + \left(a + b + 1 - 2(b+1)u\right) \frac{\mathrm{d}\tilde{f}}{\mathrm{d}u} - \left(b(b+1) + q^2E^2\tau^4\right) \tilde{f} = 0. \tag{5.34}$$

This is the hypergeometric equation (see e.g. 9.151 and 9.153 (1) of Ref. [7]), with the standard solutions,

$${}_2F_1\left(\frac{1}{2} + b - \mu, \frac{1}{2} + b + \mu; 1 + a + b; u\right), \text{ and} \tag{5.35}$$

$$u^{-a-b}\, {}_2F_1\left(\frac{1}{2} - a - \mu, \frac{1}{2} - a + \mu; 1 - a - b; u\right), \tag{5.36}$$

where $\mu = \sqrt{\frac{1}{4} - q^2E^2\tau^2}$.

Now, we will study the asymptotic behaviour ($t \to \pm\infty$) of these functions, to obtain the particle content of the field at late times. To do this, we first rewrite $f(t)$ as

$$f(t) = e^{-a(t/\tau)-b\ln(\cosh(t/\tau))}\tilde{f}(t/\tau). \tag{5.37}$$

The two solutions above are appropriate for late times. To see this, we note that as $t \to \infty$, $u \to 0$. The Gauss hypergeometric function is given by a power series (see e.g. 9.100 of Ref. [7]) such that ${}_2F_1(a, b; c; z) = 1 + O(z)$. Also, in the $t \to \infty$ limit, we have $u \sim \frac{1}{4}\operatorname{sech}^2(t/\tau)$ (where we have used $\tanh(t/\tau) = +\sqrt{1 - \operatorname{sech}^2(t/\tau)}$ for $t > 0$), which we can rewrite as $u \sim e^{-2\ln(\cosh(t/\tau))+\text{const.}}$, where the const. is irrelevant as it is always multiplied in the above expressions by a pure imaginary number (b or $a + b$) and therefore only contributes to the overall phase. Thus, the leading behaviour of the two solutions for $\tilde{f}$ above are, respectively (up to phase factors)

$$ {}_2F_1\left(\frac{1}{2} + b - \mu, \frac{1}{2} + b + \mu; 1 + a + b; u\right) \sim 1, \tag{5.38}$$

$$u^{-a-b}\, {}_2F_1\left(\frac{1}{2} - a - \mu, \frac{1}{2} - a + \mu; 1 - a - b; u\right) \sim e^{2(a+b)\ln(\cosh(t/\tau))}. \tag{5.39}$$

Before substituting this into $f(t)$, we make a further simplification. As $t \to \infty$, $\cosh(t/\tau) \to \frac{1}{2}e^{t/\tau}$, and we therefore have $\ln\cosh(t/\tau) \to (t/\tau) + \text{const.}$, with the const. being once again irrelevant. We then have, corresponding to Eqs. (5.35) and (5.36) respectively (and omitting any phase factors),

$$f(t) \sim e^{-(a+b)(t/\tau)} = e^{-i\omega_{+\infty}(\mathbf{k})t}, \tag{5.40}$$

$$f(t) \sim e^{(a+b)(t/\tau)} = e^{i\omega_{+\infty}(\mathbf{k})t}. \tag{5.41}$$

Thus, we have identified the solutions (normalized by including a factor of $(2\omega_{+\infty}(\mathbf{k}))^{-\frac{1}{2}}$) with the correct positive frequency behaviour at late times, namely

$$\begin{aligned} f^{+}_{+\infty}(\mathbf{k}; t) =& \frac{e^{-a(t/\tau)-b\ln(\cosh(t/\tau))}}{\sqrt{2\omega_{+\infty}(\mathbf{k})}} \\ & {}_2F_1\left(\frac{1}{2} + b - \mu, \frac{1}{2} + b + \mu; 1 + a + b; \frac{1 - \tanh(t/\tau)}{2}\right), \end{aligned} \tag{5.42}$$

$$\begin{aligned} f^{-}_{+\infty}(\mathbf{k}; t) =& \frac{e^{-a(t/\tau)-b\ln(\cosh(t/\tau))}}{\sqrt{2\omega_{+\infty}(\mathbf{k})}} e^{(a+b)\ln(1-\tanh(t/\tau))} \\ & {}_2F_1\left(\frac{1}{2} - a - \mu, \frac{1}{2} - a + \mu; 1 - a - b; \frac{1 - \tanh(t/\tau)}{2}\right). \end{aligned} \tag{5.43}$$

There is another way of obtaining a hypergeometric equation from Eq. (5.33): we substitute $\bar{u} = 1 - u = \frac{1}{2}(1 + \tanh(t/\tau))$ instead. We obtain, in terms of $\bar{u}$,

$$\bar{u}(1-\bar{u})\frac{\mathrm{d}^2\tilde{f}}{\mathrm{d}\bar{u}^2} + (1 + b - a - 2(b+1))\frac{\mathrm{d}\tilde{f}}{\mathrm{d}\bar{u}} - \left(b(b+1) + q^2E^2\tau^4\right)\tilde{f} = 0. \tag{5.44}$$

The standard solutions (i.e. corresponding to 9.153(1) in Ref. [7]) to this equation are

$$_2F_1\left(\frac{1}{2} + b - \mu, \frac{1}{2} + b + \mu; 1 + b - a; \bar{u}\right), \text{ and} \tag{5.45}$$

$$\bar{u}^{a-b}\ {}_2F_1\left(\frac{1}{2} + a - \mu, \frac{1}{2} + a + \mu; 1 + a - b; \bar{u}\right). \tag{5.46}$$

An important thing to note here is that whereas we had $u \to 0$ for $t \to \infty$, now $\bar{u} \to 0$ for $t \to -\infty$. The limiting forms of these functions used earlier also apply here (independent of the parameters α, β, γ of the hypergeometric function ${}_2F_1(\alpha, \beta; \gamma; z)$), but for $\bar{u} \to 0$. In this $t \to -\infty$ limit, we also have $\bar{u} \sim \frac{1}{4}\operatorname{sech}^2(t/\tau)$ (this time, using $\tanh(t/\tau) = -\sqrt{1 - \operatorname{sech}^2(t/\tau)}$ for $t < 0$) or $\bar{u} \sim e^{2\ln\cosh(t/\tau) + \text{const.}}$, with the const. yet again contributing only to phase factors. Ignoring phase factors, the early time limiting forms are therefore

$$_2F_1\left(\frac{1}{2} + b - \mu, \frac{1}{2} + b + \mu; 1 + b - a; \bar{u}\right) \sim 1, \tag{5.47}$$

$$\bar{u}^{a-b}\ {}_2F_1\left(\frac{1}{2} + a - \mu, \frac{1}{2} + a + \mu; 1 + a - b; \bar{u}\right) \sim e^{2(a-b)\ln(\cosh(t/\tau))}. \tag{5.48}$$

For $t \to -\infty$, we have $\cosh(t/\tau) \to \frac{1}{2}e^{-t/\tau}$. Consequently the factor $e^{-a(t/\tau) - b\ln(\cosh(t/\tau))} \sim e^{-(a-b)(t/\tau)}$, and the early time behaviour of solutions corresponding to Eqs. (5.45) and (5.46) respectively are

$$f(t) \sim e^{-(a-b)t/\tau} = e^{-i\omega_{-\infty}(\mathbf{k})t}, \tag{5.49}$$

$$f(t) \sim e^{(a-b)t/\tau} = e^{i\omega_{-\infty}(\mathbf{k})t}. \tag{5.50}$$

The early time positive and negative frequency solutions (normalized by including a factor of $(2\omega_{-\infty}(\mathbf{k}))^{-\frac{1}{2}}$) are then:

$$\begin{aligned}
f^{+}_{-\infty}(\mathbf{k};t) =& \frac{e^{-a(t/\tau)-b\ln(\cosh(t/\tau))}}{\sqrt{2\omega_{-\infty}(\mathbf{k})}} \\
& {}_2F_1\left(\frac{1}{2}+b-\mu, \frac{1}{2}+b+\mu; 1+b-a; \frac{1+\tanh(t/\tau)}{2}\right), \qquad (5.51)\\
f^{-}_{-\infty}(\mathbf{k};t) =& \frac{e^{-a(t/\tau)-b\ln(\cosh(t/\tau))}}{\sqrt{2\omega_{-\infty}(\mathbf{k})}} e^{(a-b)\ln(1+\tanh(t/\tau))} \\
& {}_2F_1\left(\frac{1}{2}+a-\mu, \frac{1}{2}+a+\mu; 1+a-b; \frac{1+\tanh(t/\tau)}{2}\right). \qquad (5.52)
\end{aligned}$$

In finding the Bogolubov transformation relating the early time and late time modes, we may ignore the common factor of $e^{-a(t/\tau)-b\ln\cosh(t/\tau)}$; then the coefficients may be obtained using the following relation (from e.g. 9.131(2) of Ref. [7]):

$$\begin{aligned}
{}_2F_1(\alpha,\beta;\gamma;z) =& \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}\, {}_2F_1(\alpha,\beta;\alpha+\beta-\gamma+1;1-z) \\
& + (1-z)^{\gamma-\alpha-\beta}\frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)}\, {}_2F_1(\gamma-\alpha,\gamma-\beta;\gamma-\alpha-\beta+1;1-z).
\end{aligned} \qquad (5.53)$$

Making the substitutions $\alpha = \frac{1}{2}+b-\mu$, $\beta = \frac{1}{2}+b+\mu$, $\gamma = 1+a+b$ and $z = u$, (and noting that ${}_2F_1(\alpha,\beta;\gamma;z) = {}_2F_1(\beta,\alpha;\gamma;z)$ in general)

$$\begin{aligned}
& {}_2F_1\left(\frac{1}{2}+b-\mu, \frac{1}{2}+b+\mu; 1+a+b; u\right) = \\
& \frac{\Gamma(1+a+b)\Gamma(a-b)}{\Gamma\left(\frac{1}{2}+a+\mu\right)\Gamma\left(\frac{1}{2}+a-\mu\right)}\, {}_2F_1\left(\frac{1}{2}+b-\mu, \frac{1}{2}+b+\mu; 1+b-a; \bar{u}\right) \\
& + \frac{\Gamma(1+a+b)\Gamma(b-a)}{\Gamma\left(\frac{1}{2}+b-\mu\right)\Gamma\left(\frac{1}{2}+b+\mu\right)}\bar{u}^{a-b}\, {}_2F_1\left(\frac{1}{2}+a-\mu, \frac{1}{2}+a+\mu; 1+a-b; \bar{u}\right).
\end{aligned} \qquad (5.54)$$

Including the normalization factor, we obtain the Bogolubov coefficients between early and late times:

$$\alpha_{\mathbf{k}}(+\infty,-\infty) = \left(\frac{\omega_{-\infty}(\mathbf{k})}{\omega_{+\infty}(\mathbf{k})}\right)^{\frac{1}{2}} \frac{\Gamma(1+a+b)\Gamma(a-b)}{\Gamma\left(\frac{1}{2}+a+\mu\right)\Gamma\left(\frac{1}{2}+a-\mu\right)}, \qquad (5.55)$$

$$\beta_{\mathbf{k}}(+\infty,-\infty) = \left(\frac{\omega_{-\infty}(\mathbf{k})}{\omega_{+\infty}(\mathbf{k})}\right)^{\frac{1}{2}} \frac{\Gamma(1+a+b)\Gamma(b-a)}{\Gamma\left(\frac{1}{2}+b-\mu\right)\Gamma\left(\frac{1}{2}+b+\mu\right)}. \qquad (5.56)$$

We can now evaluate the late time particle/antiparticle number (**k**-density)

$$\langle 0,-\infty|\hat{N}^{\pm}_{+\infty}(\mathbf{k})|0,+\infty\rangle = |\beta_{\mathbf{k}}(+\infty,-\infty)|^2\delta(\mathbf{k}-\mathbf{k}), \qquad (5.57)$$

$$
\begin{aligned}
|\beta_{\mathbf{k}}(+\infty,-\infty)|^2 &= \frac{(\omega_{-\infty}(\mathbf{k})/\omega_{+\infty}(\mathbf{k}))\left|\Gamma\left(1+i\omega_{+\infty}(\mathbf{k})\tau\right)\Gamma\left(-i\omega_{-\infty}(\mathbf{k})\tau\right)\right|^2}{\left|\Gamma\left(\frac{1-2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\Gamma\left(\frac{1+2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\right|^2} \\
&= \frac{\tau^2\omega_{-\infty}(\mathbf{k})\omega_{+\infty}(\mathbf{k})\left|\Gamma\left(i\omega_{+\infty}(\mathbf{k})\tau\right)\Gamma\left(-i\omega_{-\infty}(\mathbf{k})\tau\right)\right|^2}{\left|\Gamma\left(\frac{1-2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\Gamma\left(\frac{1+2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\right|^2},
\end{aligned}
\tag{5.58}
$$

where we have used $\Gamma(1+x) = x\Gamma(x)$ in the last line. This can be further simplified using $|\Gamma(ix)|^2 = \pi(\mathrm{cosech}\,\pi x)/x$ (see e.g. 5.4.3 of Ref. [22]), giving

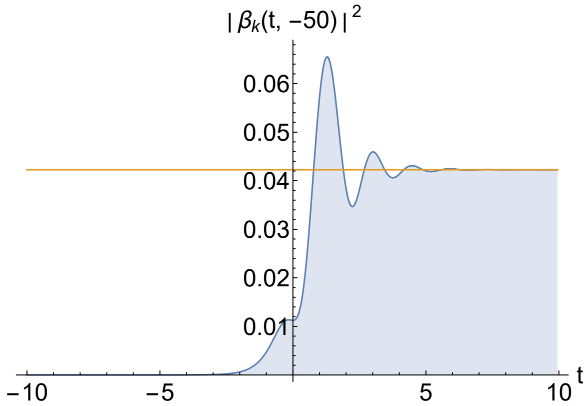
$$
|\beta_{\mathbf{k}}(+\infty,-\infty)|^2 = \frac{\tau^2\pi^2\,\mathrm{cosech}(\pi\tau\omega_{+\infty}(\mathbf{k}))\,\mathrm{cosech}(\pi\tau\omega_{-\infty}(\mathbf{k}))}{\left|\Gamma\left(\frac{1-2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\Gamma\left(\frac{1+2\mu+i\omega_{+\infty}(\mathbf{k})\tau-i\omega_{-\infty}(\mathbf{k})\tau}{2}\right)\right|^2}.
\tag{5.59}
$$

As $\mu = \sqrt{\frac{1}{4} - q^2E^2\tau^4}$, $\mu = i|\mu|$ for $|qE\tau^2| > 1/2$. If this condition is satisfied, we can simplify the above further (using $|\Gamma(\frac{1}{2}+ix)|^2 = \pi\,\mathrm{sech}(\pi x)$, see e.g. 5.4.4. of Ref. [22]) to get

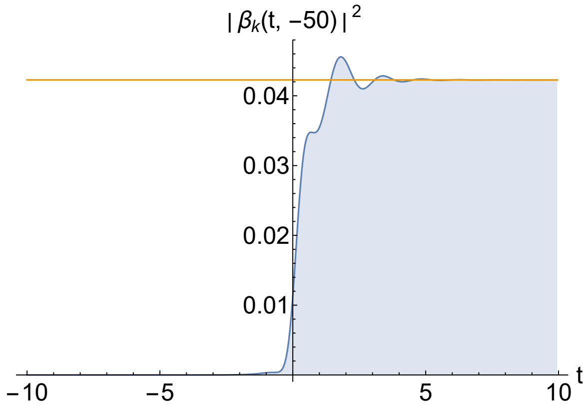
$$
|\beta_{\mathbf{k}}(+\infty,-\infty)|^2 = \frac{\tau^2\begin{pmatrix}\cosh\left(\pi\tau\omega_{+\infty}(\mathbf{k}) - \pi\tau\omega_{-\infty}(\mathbf{k}) + 2\pi|\mu|\right) \\ \times\cosh\left(\pi\tau\omega_{+\infty}(\mathbf{k}) - \pi\tau\omega_{-\infty}(\mathbf{k}) - 2\pi|\mu|\right)\end{pmatrix}}{\sinh(\pi\tau\omega_{+\infty}(\mathbf{k}))\sinh(\pi\tau\omega_{-\infty}(\mathbf{k}))}.
\tag{5.60}
$$

5.2.2 Particle content at intermediate times

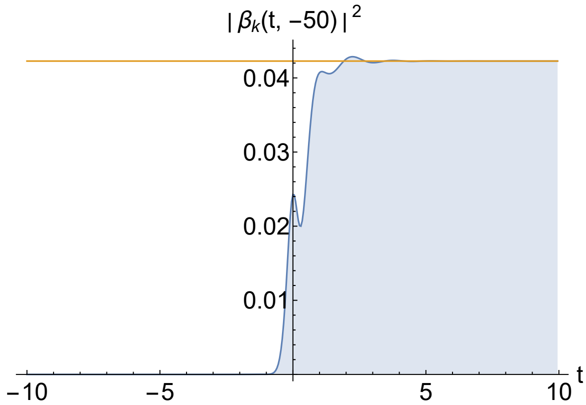
Now, we will consider plots pertaining to cases similar to the ones discussed in Sec. 5.1.3 - see Fig. 5.3. The oscillations in particle number density with time obviously last only for the rough 'duration' of the pulse, after which we have a (virtually) free flat spacetime background and do not expect any particle creation effects. Once again, these oscillations are suppressed if one chooses an adiabatic vacuum. We also see a peak near $t = 0$ (the center/point of maximum of the pulse) in higher order adiabatic Fock spaces, and Fig. 5.4 again shows that this effect is weaker for modes with larger spatial Klein-Gordon eigenvalues.



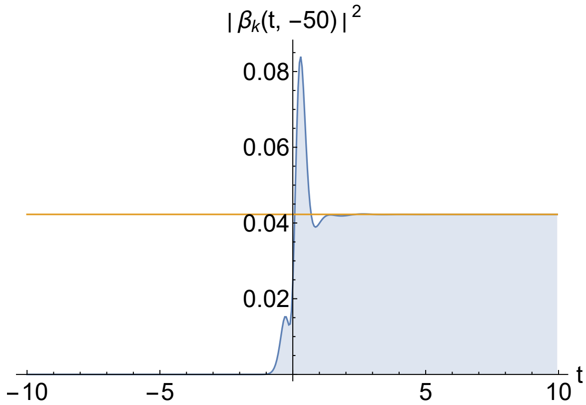
(a) Hamiltonian diagonalization



(b) Adiabatic, first order

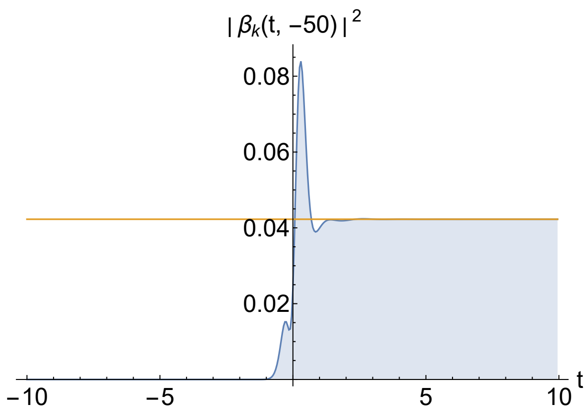


(c) Adiabatic, second order

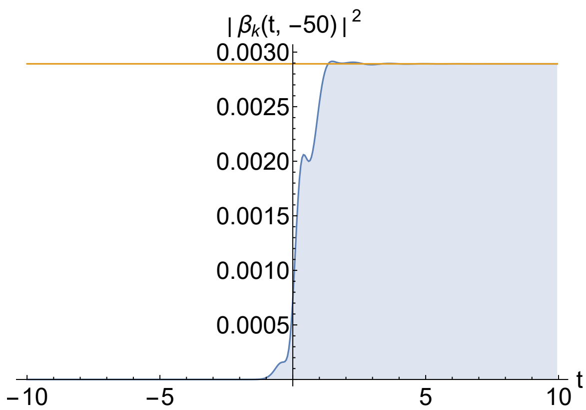


(d) Adiabatic, third order

Figure 5.3: Particle content at intermediate times for modes with $k_x = 0$, $k_y^2+k_z^2+m^2 = 1$, and $qE = 1$,$\tau = 2$ for the Sauter pulse, all in arbitrary units. The filled curve (blue) is the curve of interest, while the horizontal line shows the asymptotic late time value of $|\beta_{\mathbf{k}}(t_2, t_1)|^2$ for the corresponding mode.



(a) $k_y^2 + k_z^2 + m^2 = 1$



(b) $k_y^2 + k_z^2 + m^2 = 2$

Figure 5.4: Comparison of particle content in intermediate times with third order adiabatic Fock spaces in two different modes, with $k_x = 0$, $qE = 1$ and $\tau = 2$

Chapter 6

The Schwinger effect in FLRW spacetimes

The spatially flat FLRW metric, one of the simplest known exact solutions to the Einstein field equations (see e.g. Ref. [10]), is given by

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}\mathbf{x}^2, \tag{6.1}$$

where $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^3$, and $a(t)$ is called the scale factor. We will not concern ourself with the stress energy tensor that leads to this metric, apart from assuming that the effect of the scalar field of interest is negligible compared to the (for all practical purposes, classical rather than quantum) source stress-energy tensor of this metric, so that the background geometry remains fixed.

It will often be convenient to work with the conformal time coordinate η, defined up to an additive constant by

$$\mathrm{d}\eta = \int \frac{\mathrm{d}t}{a(t)}. \tag{6.2}$$

Expressing $\mathrm{d}s^2$ in terms of η rather than t, the components of the metric tensor are conformally related to the Minkowski metric in the $(\eta, \mathbf{x})$ coordinates:

$$\mathrm{d}s^2 = a^2(\eta)\left(\mathrm{d}\eta^2 - \mathrm{d}\mathbf{x}^2\right) = C(\eta)\left(\mathrm{d}\eta^2 - \mathrm{d}\mathbf{x}^2\right). \tag{6.3}$$

This is the form of the metric that we will use extensively in this chapter.

6.1 General considerations in FLRW spacetimes

The pure (conformal) time dependence of the conformal factor $C(\eta) = a^2(\eta)$ (i.e. the spatial homogeneity of the metric) suggests choosing the spacelike surfaces $\Sigma(t)$ to be surfaces of constant cosmic time t (or, equivalently, η) for convenience. Thus, we will work with surfaces $\Sigma(\eta) = \{(\eta, \mathbf{x}) : \forall\, \mathbf{x} \in \mathbb{R}^3\}$. The future-directed unit normals η^μ to these surfaces are in the η direction, i.e. $n^\mu \propto \delta^\mu_\eta$, and the normalization condition $g_{\mu\nu} n^\mu n^\nu = 1$ gives

$$n^\mu(\eta, \mathbf{x}) = \frac{1}{a(\eta)} \delta^\mu_\eta \tag{6.4}$$

We will be interested in a spatially homogeneous electric field, described by a potential in the temporal gauge, corresponding to

$$A_i(\eta, \mathbf{x}) = A_i(\eta),\ A_\eta(\eta, \mathbf{x}) = 0. \tag{6.5}$$

Recall that the field modes $u(k; x)$ satisfy the Klein-Gordon equation,

$$(\mathrm{D}_\mu \mathrm{D}^\mu + m^2) u(k; x) = 0. \tag{6.6}$$

For a spacetime (and internal space) vector V^μ,

$$\mathrm{D}_\mu V^\mu = \frac{1}{\sqrt{-g}} \left[(\partial_\mu - iqA_\mu) \sqrt{-g} V^\mu \right]. \tag{6.7}$$

whereas for a scalar field, the spacetime covariant derivative is equivalent to the coordinate derivative. Also, in (1+3)-D,

$$g = \det g_{\mu\nu} = C^4(\eta) \det \bar{\eta}_{\mu\nu} \implies \sqrt{-g} = C^2(\eta), \tag{6.8}$$

where $\bar{\eta}_{\mu\nu}$ represents the Minkowski metric components.

These facts can be used to rewrite the Klein-Gordon equation as follows:

$$\frac{1}{C^2(\eta)} \bar{\eta}^{\mu\nu} (\partial_\mu - iqA_\mu(\eta)) \left[C(\eta)(\partial_\nu - iqA_\nu(\eta)) \right] u(k; x) + m^2 u(k; x) = 0. \tag{6.9}$$

Substituting in the values of $\bar{\eta}_{\mu\nu}$

$$\left[\frac{1}{C(\eta)} \partial_\eta^2 + \frac{1}{C^2(\eta)} \frac{\mathrm{d}C(\eta)}{\mathrm{d}\eta} \partial_\eta - \sum_{j \neq \eta} \frac{1}{C(\eta)} (\partial_j - iqA_j(\eta))^2 + m^2 \right] u(k; x) = 0. \tag{6.10}$$

The spatial Klein-Gordon operator in this case is

$$\mathrm{K}_\eta = - \sum_{j \neq \eta} \frac{1}{C(\eta)} (\partial_j - iqA_j(\eta))^2 + m^2. \tag{6.11}$$

Its eigenfunctions are given by plane wave modes $\chi(\mathbf{k};\mathbf{x}) = (2\pi)^{-\frac{3}{2}} e^{i\mathbf{k}\cdot\mathbf{x}}$, with eigenvalues

$$\omega_\eta^2(\mathbf{k}) = \sum_{j\neq\eta} \frac{\left(k_j - qA_j(\eta)\right)^2}{a^2(\eta)} + m^2. \tag{6.12}$$

Eq. (6.10) is separable, and in a temporal gauge, in the sense of Sec. 4.1.4 and we can use the same methods to arrive at the particle number at various times. Choosing modes corresponding to eigenfunctions of K_η, we obtain Klein-Gordon corresponding to Eq. (4.39)

$$\left[\frac{1}{C(\eta)}\mathrm{d}_\eta^2 + \frac{1}{C^2(\eta)}\frac{\mathrm{d}C(\eta)}{\mathrm{d}\eta}\mathrm{d}_\eta + \omega_\eta^2(\mathbf{k})\right] f(\mathbf{k};\eta) = 0. \tag{6.13}$$

To transform this equation to the form in Eq. (4.62), for convenience in working with the adiabatic vacuum prescription, we write $\tilde{f}(\mathbf{k};\eta) = a(\eta) f(\mathbf{k};\eta)$, obtaining

$$\left[\mathrm{d}_\eta^2 + \left(a^2(\eta)\omega_\eta^2(\mathbf{k}) - \frac{1}{a(\eta)}\frac{\mathrm{d}^2 a(\eta)}{\mathrm{d}\eta^2}\right)\right] \tilde{f}(\mathbf{k};\eta) = 0. \tag{6.14}$$

Thus, we have

$$\widetilde{\omega}_\eta(\mathbf{k}) = \sqrt{a^2(\eta)\omega_\eta^2(\mathbf{k}) - \frac{1}{a(\eta)}\frac{\mathrm{d}^2 a(\eta)}{\mathrm{d}\eta^2}}. \tag{6.15}$$

For a significantly accelerated expansion, we see that it is possible for $\widetilde{\omega}_\eta(\mathbf{k})$ to be purely imaginary for some spatial eigenfunctions, making their norm zero; we will only use the adiabatic prescription for large enough $\omega_\eta^2(\mathbf{k})$ such that this situation is avoided (For a complete Fock space decomposition, we may always use Hamiltonian diagonalization (say) for the remaining eigenfunctions).

Now we turn to two examples that are relevant for inflation - the exactly solvable (asymptotically) de Sitter case and the case of a power law universe where exact analytical solutions are unknown, but we may still apply the methods of Sec. 4.1.4 numerically (see Sec. 4.3).

6.2 de Sitter spacetime

6.2.1 Traditional analysis in a limiting case

The metric for de Sitter spacetime is (for $(t,\mathbf{x}) = (t,x,y,z) \in \mathbb{R}^4$)

$$\mathrm{d}s^2 = \mathrm{d}t^2 - e^{2Ht}\mathrm{d}\mathbf{x}^2, \tag{6.16}$$

where $H > 0$ is the Hubble parameter. In terms of the conformal time $\eta = \int \frac{\mathrm{d}t}{e^{Ht}} = -\frac{1}{He^{Ht}}$, $\eta \in (-\infty, 0)$

$$\mathrm{d}s^2 = \frac{1}{H^2\eta^2}\left(\mathrm{d}\eta^2 - \mathrm{d}\mathbf{x}^2\right). \tag{6.17}$$

The conformal factor is then $C(\eta) = \frac{1}{H^2\eta^2}$. For this metric, H^{-1} determines the length scale of the curvature of spacetime[1]. We also now require (schematically) $a^2\omega^2 > H^2$ for the adiabatic positive frequency modes to make sense.

We choose the electromagnetic potential as[2]

$$A_\mu = -\frac{E}{H^2\eta}\delta^x_\mu, \tag{6.18}$$

so that the invariant quantity $F_{\mu\nu}F^{\mu\nu} = -2E^2$ corresponds to a constant electric field, in the x-direction. We note that this satisfies the temporal gauge condition, and we can proceed with the standard method of evaluating particle production discussed in Sec. 4.1.4 (an essentially similar calculation is carried out in, for instance, Refs. [12, 14, 25]).

The Klein-Gordon equation for the modes then becomes

$$\left[\partial_\eta^2 - \frac{2}{\eta}\partial_\eta - \sum_{a\in\{x,y,z\}}\left(\partial_a + \frac{iqE}{H^2\eta}\delta^x_a\right)^2 + \frac{m^2}{H^2\eta^2}\right]u(k;x) = 0. \tag{6.19}$$

As usual, we choose plane wave modes $u^\pm(\mathbf{k};x) = f^\pm(\mathbf{k};\eta)(2\pi)^{-\frac{3}{2}}e^{i\mathbf{k}\cdot\mathbf{x}}$, and replace $f^\pm(\mathbf{k};\eta)$ with the more generic $f(\eta)$ for notational convenience:

$$\left[\partial_\eta^2 - \frac{2}{\eta}\partial_\eta + \left(k^2 + \frac{2k_x}{\eta}\frac{qE}{H^2} + \frac{q^2E^2}{H^4\eta^2} + \frac{m^2}{H^2\eta^2}\right)\right]f = 0. \tag{6.20}$$

It is useful to define the dimensionless quantities

$$L = \frac{qE}{H^2}, M = \frac{m}{H}. \tag{6.21}$$

The eigenvalues of the spatial Klein-Gordon operator are then given by

$$\omega_\eta^2(\mathbf{k}) = H^2\eta^2k^2 + 2H^2\eta k_x L + H^2(L^2 + M^2). \tag{6.22}$$

[1] See also the discussion following Eq. (4.70).

[2] We remark that while this is a convenient choice for this problem, we must make a trivial gauge transformation by shifting the x component by $-E/H$ to avoid a divergent constant in the $H \to 0$ limit.

The asymptotic behaviour of $\omega_\eta(\mathbf{k})$ will be of relevance when considering particle production between the infinite past ($\eta \to -\infty$) and the infinite future ($\eta \to 0$), and is given by, in these limits, (obviously choosing the positive square root)

$$\omega_\eta(\mathbf{k}) \to -H\eta k, \qquad \eta \to -\infty, \tag{6.23}$$

$$\omega_\eta(\mathbf{k}) \to H(L^2+M^2)^{\frac{1}{2}} + 2H^2\eta k_x L, \quad \eta \to 0. \tag{6.24}$$

We have retained the $O(\eta)$ term in the late time limit as it is the leading contribution to the following quantity. As a measure of how fast these eigenvalues are changing in conformal time, we consider the dimensionless quantity $\omega_\eta^{-2}(\mathbf{k})(\mathrm{d}\omega_t(\mathbf{k})/\mathrm{d}\eta)$, and evaluate it in the above limits

$$\lim_{\eta\to-\infty} \omega_\eta^{-2}(\mathbf{k})\frac{\mathrm{d}\omega_t(\mathbf{k})}{\mathrm{d}\eta} = \lim_{\eta\to-\infty} -\frac{1}{Hk\eta^2} = 0, \tag{6.25}$$

$$\lim_{\eta\to 0} \omega_\eta^{-2}(\mathbf{k})\frac{\mathrm{d}\omega_t(\mathbf{k})}{\mathrm{d}\eta} = \frac{2k_x L}{L^2+M^2}. \tag{6.26}$$

While we expect particle creation to be negligible at asymptotically early times due to the slowly varying eigenvalue, it does not cease at late times at all, even for an individual mode (unlike the problems we have encountered so far), unless $L^2 + M^2 \to \infty$.

Substituting $f(\eta) = a^{-1}(\eta)\tilde{f}(\eta) = -H\eta\tilde{f}(\eta)$, Eq. (6.20) becomes

$$\frac{\mathrm{d}^2\tilde{f}}{\mathrm{d}\eta^2} + \left(k^2 + \frac{2k_x}{\eta}L + \frac{L^2+M^2-2}{\eta^2}\right)\tilde{f} = 0. \tag{6.27}$$

$\widetilde{\omega}_\eta(\mathbf{k})$ can be directly read off from this equation,

$$\widetilde{\omega}_\eta^2(\mathbf{k}) = k^2 + \frac{2k_x}{\eta}L + \frac{L^2+M^2-2}{\eta^2}. \tag{6.28}$$

With the further substitutions $z = 2ik\eta$, $\xi = -\frac{ik_x}{k}L$ and $\nu = +\sqrt{\frac{9}{4} - L^2 - M^2}$, this reduces to the defining equation for the Whittaker functions (see e.g. 9.220 in Ref. [7])

$$\frac{\mathrm{d}^2 w}{\mathrm{d}z^2} + \left(-\frac{1}{4} + \frac{\xi}{z} + \frac{\frac{1}{4}-\nu^2}{z^2}\right)w = 0. \tag{6.29}$$

It will be useful to note the following: z is purely imaginary with a negative imaginary part (since $\eta < 0$) and ξ is purely imaginary and proportional to qE. Further, we take ν to be a positive real number when $\frac{9}{4} - L^2 - M^2 > 0$ and a purely imaginary number with a positive real part i.e. $\nu = i|\nu|$ when $\frac{9}{4} - L^2 - M^2 < 0$.

We are interested in two pairs of solutions involving the Whittaker functions

$$(W_{\xi,\nu}(z), W_{-\xi,\nu}(-z))\,;(M_{\xi,\nu}(z), M_{\xi,-\nu}(z))\,. \tag{6.30}$$

These functions are analytic everywhere in the complex plane corresponding to z except for a branch cut along (conventionally) $|\arg z| = \pi$ and a singular point at $z = \infty$ (see e.g. 13.14 in Ref. [22], 9.220 in Ref. [7]). Analyticity in the neighbourhood of any point corresponding to finite, nonzero $\eta < 0$ (corresponding to $\arg z = -\frac{\pi}{2}$), which are the points of physical interest, readily follows from this. Also, for $z \neq 0$, $M_{\xi,\nu}(z)/\Gamma(2\nu+1)$ and $W_{\xi,\nu}(z)$ are entire in ξ and ν (see e.g. 13.14 of Ref. [22]; we don't really have to be concerned about the role of the Gamma function as its poles are in $\nu < 0$ (see e.g. 5.2 in Ref. [22]) which is not in the allowed range of values of ν). These facts will be used implicitly henceforth.

That the first pair $(W_{\xi,\nu}(z), W_{-\xi,\nu}(-z))$ consists of complex conjugates follows from z and ξ being purely imaginary, ν being either purely real or purely imaginary (depending on L and M), and the following relation (useful when ν is purely imaginary),

$$W_{\xi,\nu}(z) = W_{\xi,-\nu}(z). \tag{6.31}$$

The second pair are not related by complex conjugation, but we may make them complex conjugates of each other by multiplying one of them by an overall factor when ν is purely imaginary. To show this, we directly evaluate the complex conjugate of $M_{\xi,\nu}(z)$:

$$M^*_{\xi,\nu}(z) = M_{-\xi,\nu^*}(e^{i\pi}z), \tag{6.32}$$

where we have chosen the phase of $-z$ to avoid the branch cut, noting that $\operatorname{Im} z < 0$. Now, we use the relation (see e.g. 13.14.10 of Ref. [22])

$$M_{\xi,\nu}(e^{\pm i\pi}z) = \pm i e^{\pm i\pi\nu} M_{-\xi,\nu}(z), \tag{6.33}$$

from which it follows that, when ν is purely imaginary

$$M^*_{\xi,\nu}(z) = i e^{i\pi\nu} M_{\xi,-\nu}(z). \tag{6.34}$$

This condition on ν may seem restrictive, but as we will later see, the particle number is easy to obtain from these solutions only for $\nu \to i\infty$ (which corresponds to extremely strong electric fields or large masses, relative to a suitable power of the Hubble parameter), which is also the regime we will restrict ourselves to.

Similar to the Minkowski spacetime examples, we will look at the asymptotic behaviour of these standard solutions to identify the appropriate positive frequency mode functions.

One of the asymptotic forms of interest is

$$W_{\xi,\nu}(z) \sim e^{-\frac{z}{2}} z^{\xi} = e^{-\frac{z}{2}-\xi \ln(z)}, \text{ for } |z| \to \infty, |\arg z| < \pi. \tag{6.35}$$

As ξ is purely imaginary, the logarithmic term contributes only to the phase; we can neglect the slower growing logarithmic term (which leads to slower phase oscillations) in the exponential compared to the linear term (which has the dominant contribution to the phase) in the $\eta \to -\infty$ limit, leading to the simpler expression (in terms of η)

$$W_{\xi,\nu}(z) \sim e^{-ik\eta}. \tag{6.36}$$

Now, the mode corresponding to this function is $f_1(\eta) = (-H\eta)W_{\xi,\nu}(2ik\eta)$ and we see that $f_1 \to -H\eta e^{-ik\eta}$ as $\eta \to -\infty$. More importantly, we consider the normal derivative $\mathrm{D}_\eta f$ in this limit (for which the leading order contribution comes from the derivative of the exponential), which gives (using $n^\mu = -H\eta\delta^\mu_\eta$)

$$n^\mu \mathrm{D}_\mu f_1(-\infty) = ik\eta H f_1(-\infty) = -i\omega_{-\infty}(\mathbf{k}) f_1(-\infty). \tag{6.37}$$

Noting that the derivative $\mathrm{d}\omega_\eta(\mathbf{k})/\mathrm{d}\eta$ is negligible as $\eta \to -\infty$ compared to $\omega_{-\infty}(\mathbf{k})$, both Hamiltonian diagonalization and the adiabatic particle definition correspond to $\Omega_{-\infty}(\mathbf{k}) \approx \omega_{-\infty}(\mathbf{k})$. The early time positive frequency modes are then

$$f^+_{-\infty}(\mathbf{k};\eta) = c^+(\mathbf{k})(-H\eta)W_{\xi,\nu}(2ik\eta), \tag{6.38}$$

where $c^+(\mathbf{k})$ is a normalization constant. The corresponding negative frequency modes are given by the complex conjugates of the right hand side of this expression. It is worth noting that these modes also correspond to the Bunch-Davies vacuum in de Sitter spacetime (see e.g. Ref. [6]).

The other relevant asymptotic form is

$$M_{\xi,\nu}(z) \sim z^{\nu+\frac{1}{2}}, \text{ as } z \to 0. \tag{6.39}$$

At late times, $\eta \to 0$, in which limit

$$M_{\xi,\nu}(2ik\eta) \sim (2ik\eta)^{\nu+\frac{1}{2}} = (-2ik)^{\nu+\frac{1}{2}}(-\eta)^{\nu+\frac{1}{2}}. \tag{6.40}$$

Now, we consider the mode $f_2(\eta) = (-H\eta)M_{\xi,\nu}(\eta)$. Its normal derivative at late times is (using $\mathrm{d}z = 2ik\mathrm{d}\eta$),

$$n^\mu \mathrm{D}_\mu f_2(0) = -Hz\frac{\mathrm{d}f_2}{\mathrm{d}z}(0) = -H\left(\nu + \frac{1}{2}\right) f_2(0). \tag{6.41}$$

Comparing this with the eigenvalue at late times, $\omega_0(\mathbf{k}) = H\sqrt{L^2+M^2}$, we see that the f_2 modes do not have the right positive frequency behaviour at late times. However, if we consider a strong electric field or large mass such that $L^2 + M^2 \gg \frac{9}{4}$, then $\nu \approx i\sqrt{L^2+M^2}$, and in this limit alone, we have

$$n^\mu \mathrm{D}_\mu f_2(0) \approx -i\omega_0(\mathbf{k}) f_2(0). \tag{6.42}$$

For large electric fields/masses, the appropriate late time positive frequency modes are then given by (with $d^+(\mathbf{k})$ playing the role of a normalization constant)

$$f_0^+(\mathbf{k};\eta) = (-H\eta)M_{\xi,\nu}(\eta). \tag{6.43}$$

To obtain the Bogolubov transformation between the two sets of modes, we use the relation (see e.g. 9.233(1) in Ref. [7], which is valid for $\arg z = -\frac{\pi}{2}$ as required):

$$M_{\xi,\nu}(2ik\eta) = \frac{\Gamma(2\nu+1)e^{i\pi(\xi-\nu-\frac{1}{2})}}{\Gamma(\nu+\xi+\frac{1}{2})}W_{\xi,\nu}(2ik\eta) + \frac{\Gamma(2\nu+1)e^{i\pi\xi}}{\Gamma(\nu-\xi+\frac{1}{2})}W_{-\xi,\nu}(-2ik\eta). \tag{6.44}$$

This gives the non-normalized coefficients:

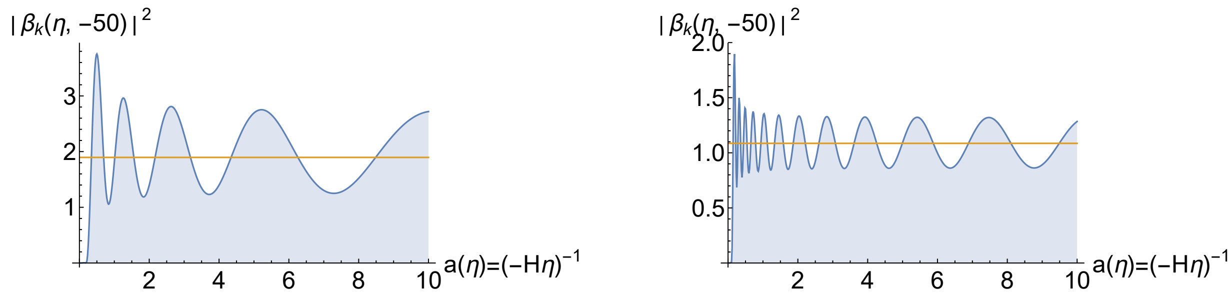
$$\tilde{\alpha}_{\mathbf{k}}(0,-\infty) = \frac{\Gamma(2\nu+1)}{\Gamma(\nu+\xi+\frac{1}{2})}e^{i\pi(\xi-\nu-\frac{1}{2})}, \tag{6.45}$$

$$\tilde{\beta}_{\mathbf{k}}(0,-\infty) = \frac{\Gamma(2\nu+1)}{\Gamma(\nu-\xi+\frac{1}{2})}e^{i\pi\xi}. \tag{6.46}$$

Using Eq. (4.51) we get the normalized β-coefficient:

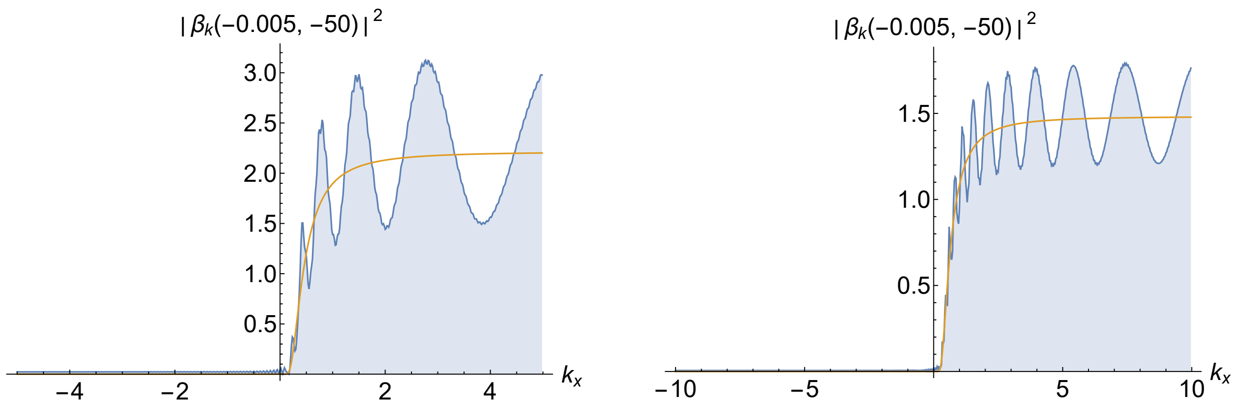
$$|\beta_{\mathbf{k}}(0,-\infty)|^2 = \frac{\cosh(\pi|\nu| + \pi\frac{k_x}{k}L)}{e^{2\pi|\nu|}\cosh(\pi|\nu| - \pi\frac{k_x}{k}L) - \cosh(\pi|\nu| + \pi\frac{k_x}{k}L)}. \tag{6.47}$$

It is straightforward to show that Eq. (6.47) reduces to the flat spacetime result Eq. (5.22) in the $H \to 0$ limit, after making a gauge transformation such that $A_\mu \to A_\mu + qE/H$ (which does not affect L) to remove a divergent constant, which must be accompanied by $k_x \to k_x + qE/H$.



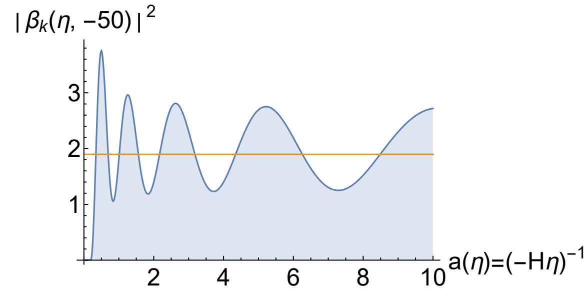
(a) L = 5 (b) L = 10

Figure 6.1: Particle content based on Hamiltonian diagonalization at intermediate times for modes with $k_x = 1$, $k_y^2 + k_z^2 = 0.01$ and $M = 1$ for a constant magnitude electric field in a de Sitter spacetime ($H = 1$), all in arbitrary units. The filled curve (blue) is the curve of interest, while the horizontal line shows the asymptotic late time value of $|\beta_{\mathbf{k}}(\eta_2, \eta_1)|^2$ for the corresponding mode.

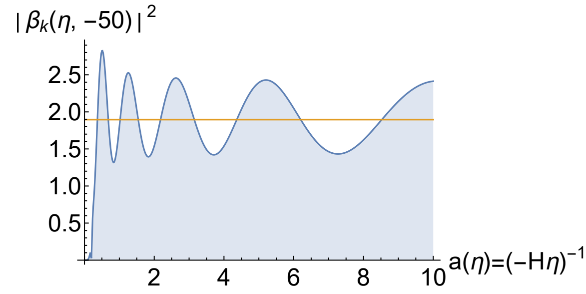


(a) L = 5 (b) L = 10

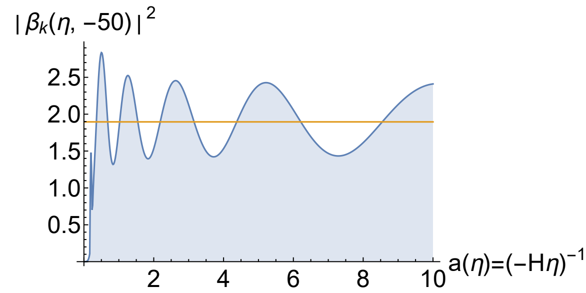
Figure 6.2: Particle content based on Hamiltonian diagonalization as a function of k_x for modes with $k_y^2 + k_z^2 = 0.01$ and $M = 1$ for a constant magnitude electric field in a de Sitter spacetime ($H = 1$), all in arbitrary units. The filled curve (blue) is the curve of interest, which oscillates around the analytical $L^2 + M^2 \gg (9/4)$ result



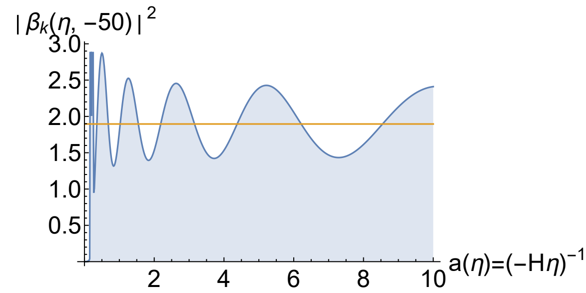
(a) Adiabatic, zeroth order



(b) Adiabatic, first order



(c) Adiabatic, second order



(d) Adiabatic, third order

Figure 6.3: Particle content with adiabatic Fock spaces at intermediate times for modes with $k_x = 1$, $k_y^2 + k_z^2 = 0.01$, $L = 5$ and $M = 1$ for a constant magnitude electric field in a de Sitter spacetime ($H = 1$), all in arbitrary units. The filled curve (blue) is the curve of interest, while the horizontal line shows the asymptotic late time value of $|\beta_{\mathbf{k}}(\eta_2, \eta_1)|^2$ for the corresponding mode.

6.2.2 Discussion on more general cases

We recall that the steps leading up to the above result, Eq. (6.47), are only (approximately) valid in the $L^2 + M^2 \gg \frac{9}{4}$ regime. We will now discuss the more general case, with the aid of numerical plots obtained as described in Sec. 4.3 (see also Ref. [16], where the general case is analyzed comprehensively in the Schrödinger picture).

First, we look at the particle content at intermediate times using the Hamiltonian diagonalization prescription, Fig. 6.1. We see that the particle content at late times indeed keeps oscillating rather than saturating to a constant value, as the eigenvalue $\omega_\eta^2(\mathbf{k})$ does not typically have a negligible rate of change. However, the magnitude of these oscillations relative to the (say) time-averaged particle content at late times is seen to decrease for a higher value of L; this is in line with our expectation that these oscillations become negligible as $L^2 + M^2 \to \infty$. Similar oscillations are observed in k_x as well, in Fig. 6.2, where we may readily attribute the asymmetry in the particle content to the directionality of the electric field.

The adiabatic Fock spaces Fig. 6.3 also see these oscillations at late times, with the only difference compared to Hamiltonian diagonalization in this limit being the size of about the first oscillation, and the emergence of a peak like in the Minkowski cases for higher adiabatic orders.

6.3 Power law spacetime

A general power law spacetime has a scale factor of the form $a(t) = \alpha t^p$. We restrict our interest to accelerating expansion (such as in inflation) for which $p > 1$. Expressing the metric in terms of the conformal time, we have

$$\mathrm{d}s^2 = \alpha^2 \left(\frac{-1}{\alpha(p-1)\eta} \right)^{\frac{2p}{p-1}} \left(\mathrm{d}\eta^2 - \mathrm{d}\mathbf{x}^2 \right). \tag{6.48}$$

Of particular interest is the scale factor in terms of conformal time,

$$a(\eta) = \alpha \left(\frac{-1}{\alpha(p-1)\eta} \right)^{\frac{p}{p-1}}. \tag{6.49}$$

For a constant electric field of magnitude E, we may choose the potential

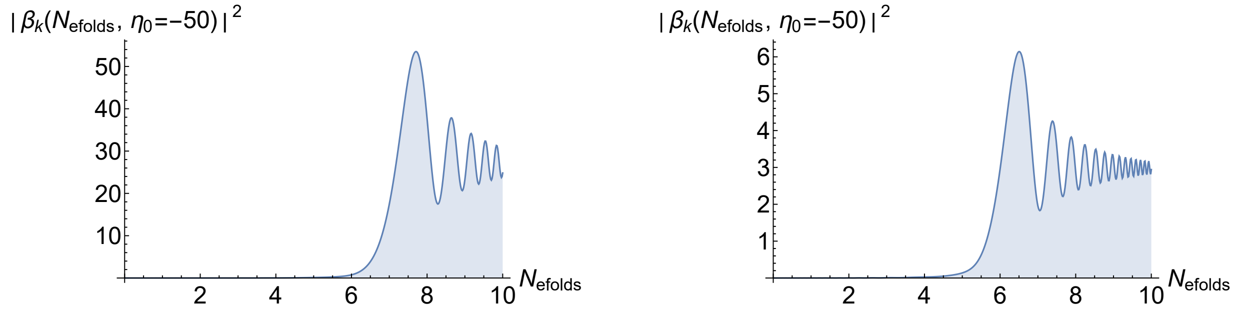
$$A_\mu(\eta) = \frac{E\alpha}{p+1} \left(\frac{-1}{\alpha(p-1)\eta} \right)^{\frac{p+1}{p-1}} \delta^x_\mu. \tag{6.50}$$

The spatial Klein-Gordon eigenvalues are then

$$\omega^2_\eta(\mathbf{k}) = \frac{k^2}{\alpha^2} \left(-\alpha(p-1)\eta\right)^{\frac{2p}{p-1}} - 2\left(\frac{p-1}{p+1}\right) qEk_x\eta + \frac{q^2E^2}{(p+1)^2} \left(-\alpha(p-1)\eta\right)^{-\frac{2}{p-1}} + m^2. \tag{6.51}$$

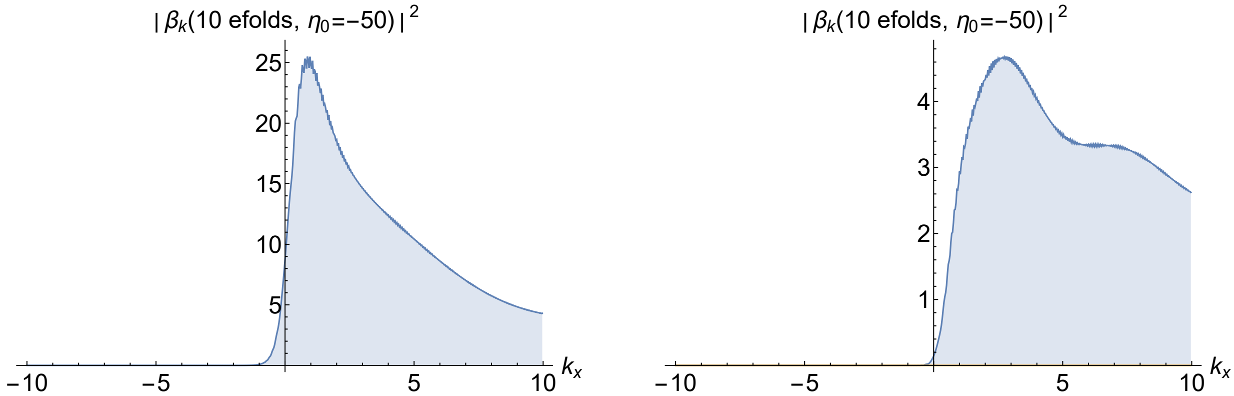
To connect with the conventions used in the theory of inflation (see e.g.Ref. [26]), we measure time evolution using the number of efolds, $N_{\text{efolds}}(\eta_2, \eta_1) = \ln(a(\eta_2)/a(\eta_1))$, which is a logarithmic measure of how much the scale factor has multiplied (more precisely, the number of times $a(\eta)$ has scaled up by e, the base of natural logarithms). If it is to account for the horizon problem, inflation is required to last for around 60 efolds.

It turns out that the solutions to the Klein-Gordon equation Eq. (6.10) with this scale factor and gauge potential have no known expression in terms of well known special functions (for a general p; evidently $p \to \infty$ with a suitable scaling of η corresponds to de Sitter spacetime, for which we do know the solutions, and there may be other special cases). We will



(a) $qE = 20$ (b) $qE = 50$

Figure 6.4: Particle content based on Hamiltonian diagonalization at intermediate times for modes with $k_x = 1$,$k_y^2 + k_z^2 = 1$ and $m = 1$ for a constant magnitude electric field in a power law spacetime ($\alpha = 1$, $p = 3$), all in arbitrary units.

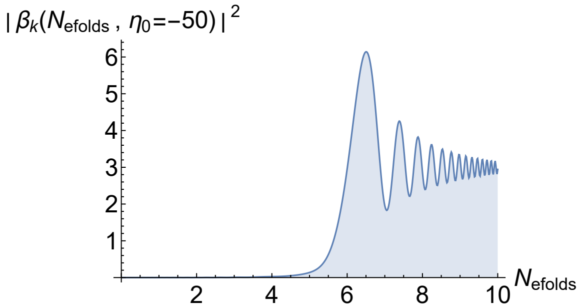


(a) $qE = 20$, $N_{\text{efolds}} = 10$ (b) $qE = 50$, $N_{\text{efolds}} = 10$

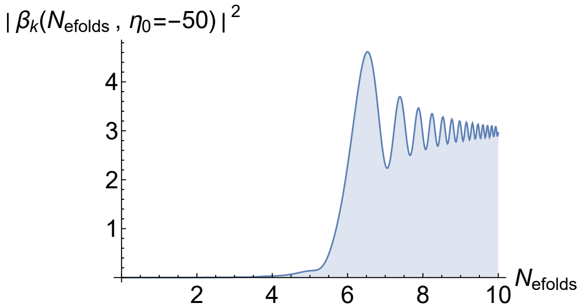
Figure 6.5: Particle content based on Hamiltonian diagonalization as a function of k_x for modes with $k_y^2 + k_z^2 = 1$ and $m = 1$, after 10 efolds for a constant magnitude electric field in a power law spacetime ($\alpha = 1$, $p = 3$), all in arbitrary units.

therefore have to work with numerical plots for specific cases, and a sampling of these is given in Figs.6.4, Fig. 6.5 and Fig. 6.6. For computational convenience, we restrict ourselves to 10 efolds and work with a low power law, $p = 3$.

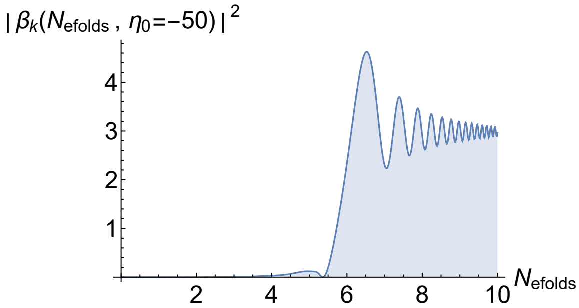
We note that the particle content at intermediate times (Fig. 6.4) as a function of efolds qualitatively resembles the flat spacetime plots Fig. 5.1a and Fig. 5.3a. Like in the de Sitter case (Fig. 6.3), the adiabatic Fock spaces (Fig. 6.6) do not seem to suppress the oscillations in late times relative to Hamiltonian diagonalization; however, the first oscillation (or so) is suppressed, and an additional peak emerges at higher adiabatic orders.



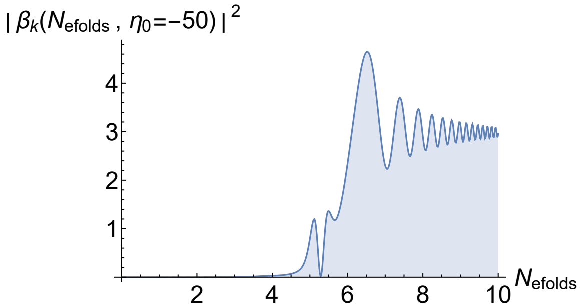
(a) Adiabatic, zeroth order



(b) Adiabatic, first order



(c) Adiabatic, second order



(d) Adiabatic, third order

Figure 6.6: Particle content with adiabatic Fock spaces at intermediate times for modes with $k_x = 1$, $k_y^2 + k_z^2 = 1$, for a constant magnitude electric field $qE = 50$ in a power law spacetime ($\alpha = 1$, $p = 3$), all in arbitrary units

Chapter 7

An application to cosmology: Inflationary magnetogenesis

In this chapter, we will briefly discuss the role of the Schwinger effect in a model of inflationary magnetogenesis (largely following the treatment in Ref. [12]), as an illustration of an application to cosmology. We will continue to treat the gauge field as a classical field.

7.1 The generation of electromagnetic fields during inflation

The Maxwell action for the electromagnetic field in curved spacetimes is given by

$$S_{\text{EM}} = \int \mathrm{d}^4x \, \sqrt{-g} \, \left(-\frac{1}{4} F_{\mu\nu}F^{\mu\nu}\right). \tag{7.1}$$

This action, being conformally invariant, does not result in the generation of electromagnetic fields in an expanding universe (see e.g. Ref. [6]). To achieve such a generation, we modify this action to

$$S_{\text{EM}} = \int \mathrm{d}^4x \, \sqrt{-g} \, \left(-\frac{f^2}{4} F_{\mu\nu}F^{\mu\nu}\right), \tag{7.2}$$

which is not conformally invariant. Dynamically, f^2 introduces a coupling of the $U(1)$ gauge field to the inflaton field (which is directly associated with the scale factor; see e.g. Ref. [26] for a discussion on the inflaton) - we will however assume that inflation as such is unaffected by the electromagnetic field, and therefore, it will be sufficient for our purposes to treat f as a function of the scale factor $f = f(a(\eta))$.

We will now be concerned with the equation of motion governing the field A_μ with this new action, which will play a central role when we discuss the role of the Schwinger effect.

The temporal gauge condition in FLRW spacetimes reads $A_\eta = 0$, which we will impose together with the Coulomb gauge condition $(\nabla_\mu - n_\mu n^\nu \nabla_\nu)A^\mu \propto \partial_i A^i = 0$. The equation of motion for the A_i is then

$$\frac{\mathrm{d}^2 A_i}{\mathrm{d}\eta^2} + \frac{2}{f}\frac{\mathrm{d}f}{\mathrm{d}\eta}\frac{\mathrm{d}A_i}{\mathrm{d}\eta} - a^2(\eta)\partial_k\partial^k A_i = 0. \tag{7.3}$$

We may restrict our attention to fields that vary slowly in space (the potential cannot be precisely homogeneous for a non-vanishing magnetic field), allowing us to drop the spatial derivative term:

$$\frac{\mathrm{d}^2 A_i}{\mathrm{d}\eta^2} = -\frac{2}{f}\frac{\mathrm{d}f}{\mathrm{d}\eta}\frac{\mathrm{d}A_i}{\mathrm{d}\eta}. \tag{7.4}$$

At this stage, we may conclude that for this to lead to growing fields, we require $\mathrm{d}f/\mathrm{d}\eta < 0$. The equation is readily integrated, giving the solution

$$A_i(\eta_2) - A_i(\eta_1) = c_i \int_{\eta_1}^{\eta_2} \mathrm{d}\eta\, f^{-2}(a(\eta)). \tag{7.5}$$

where c_i are constants of integration.

One common choice (see e.g. Refs. [12, 27]) is $f(a(\eta)) = a^s(\eta) \propto \eta^{-qs}$ where we have assumed a general power law $a(\eta) \propto \eta^{-q}$, with $q = 1$ corresponding to de Sitter spacetime. This leads to the potential evolving as

$$A_i(\eta) \propto \text{const.} + \eta^{-2qs+1}, \tag{7.6}$$

with

$$\frac{\mathrm{d}A_i}{\mathrm{d}\eta} = \frac{2qs-1}{-\eta}A_i. \tag{7.7}$$

This leads to fields whose magnitude grows as $\eta \to 0$ when $s > 1/(2q)$. In particular, the fields are given by

$$E_i = \frac{1}{a(\eta)}\frac{\mathrm{d}A_i}{\mathrm{d}\eta}, \tag{7.8}$$

$$B_i = \frac{1}{a(\eta)}(\epsilon_{ijl}\partial_j A_l), \tag{7.9}$$

where ϵ_{ijl} is the three-dimensional totally antisymmetric Levi-Civita symbol with $\epsilon_{123} = 1$. To find their magnitudes, we use the spatial metric, to get

$$E^2 = \frac{1}{a^4(\eta)}\left(\frac{\mathrm{d}A_i}{\mathrm{d}\eta}\right)^2, \tag{7.10}$$

$$B^2 = \frac{1}{a^4(\eta)}(\delta_{jm}\delta_{ln} - \delta_{jn}\delta_{lm})(\partial_j A_l)(\partial_m A_n). \tag{7.11}$$

Assuming that the spatial variation of the potential takes the form of a plane wave mode of wave vector k with the time evolution as given above, and enforcing the Coulomb gauge condition, we have (restricting ourselves to a de Sitter spacetime, $q = 1$ with Hubble parameter H for simplicity)

$$E^2 = H^4 \eta^2 (2s - 1) A^2, \tag{7.12}$$

$$B^2 = H^4 \eta^4 k^2 A^2. \tag{7.13}$$

In particular,

$$\frac{E^2}{B^2} = \frac{2s - 1}{k^2 \eta^2}. \tag{7.14}$$

7.2 The role of the Schwinger effect

In the presence of charged fields, we must consider the source term in the equation of motion for the electromagnetic potential. We will assume that this emerges as the classical limit of an expectation value relation from, say, the Heisenberg equations of motion for both fields, but with the limit being taken only for the electromagnetic field. In practice, this amounts to replacing the current j^μ in the classical Maxwell equations with the expectation value $\langle 0, -\infty | j^\mu | 0, -\infty \rangle$, assuming that the charged field is in the early time vacuum state (according to some appropriate Fock space prescription).

It is useful to write $\langle 0, -\infty | j_i | 0, -\infty \rangle = \sigma E_i$, where σ essentially plays the role of (typically nonlinear) conductivity, in the chosen vacuum state. The resulting equation of motion can be simplified as earlier (assuming slow spatial variation etc.) which now gives (Ref. [12])

$$\frac{\mathrm{d}^2 A_i}{\mathrm{d}\eta^2} + \left(\frac{2}{f} \frac{\mathrm{d}f}{\mathrm{d}\eta} + \frac{a(\eta)\sigma}{f^2} \right) \frac{\mathrm{d}A_i}{\mathrm{d}\eta} = 0. \tag{7.15}$$

Now, one makes the qualitative argument that any significant current from the scalar field will impede the growth of the electric and magnetic fields (as σ is positive). Thus, inflationary magnetogenesis proceeds to give strong electric and magnetic fields more or less only when this term is negligible, which translates into the condition:

$$\frac{a(\eta)\sigma}{f^2} \ll \left| \frac{2}{f} \frac{\mathrm{d}f}{\mathrm{d}\eta} \right|. \tag{7.16}$$

To determine the constraints that follow from this condition, we have to explicitly evaluate the expectation value of the current in the early time vacuum state. The calculation is rather involved, and results in a divergent expression that may be regularized using, say, adiabatic regularization (as done in Ref. [12] for de Sitter spacetime).

The result obtained in Ref. [12] is easier to state in the limiting case of a strong electric field ($L \gg 1$), for which

$$\sigma \approx \frac{q^3 E}{12\pi^3 H} e^{-\frac{\pi m^2}{|qE|}}. \tag{7.17}$$

Inverting this equation (using non-elementary functions) gives E as a function of σ, which allows us to derive from Eq. (7.16) a constraint on the magnitude of the electric field, that can be readily interpreted as a constraint on the magnetic field as well, using Eq. (7.14) at any given time.

There is also a related effect that comes into play post-inflation (see e.g. Ref. [28]), which we mention for the sake of completeness. After the inflationary era, we assume that f^2 approaches a constant, reducing to the conformally invariant Maxwell electromagnetic theory of everyday experience, and terminating the generation of the fields. The electric and magnetic fields that remain after inflationary magnetogenesis now come under the influence of charged particles created by processes including the Schwinger effect, which have a net conductivity σ associated with them. As the expansion of the universe is now slower, the σ term in Eq. (7.15) dominates, admitting a decaying solution (in time) and a time-independent solution. The former represents a decaying electric field, while the latter accounts for a left over magnetic field. This is one possible explanation for the observed magnetic fields in the universe at cosmological scales (see e.g. Refs. [12, 27, 28]).

Chapter 8

Summary

We will now summarize what has been discussed in this report.

We discussed the classical theory of scalar fields in a general gravitational and gauge field background (governed by a generalized Klein-Gordon equation), before specializing to a complex scalar field with a $U(1)$ gauge field (i.e. electromagnetic) background. We defined the Klein-Gordon inner product (as a function of a spacelike surface) and showed that it was an invariant inner product for solutions of the Klein-Gordon equation. This allowed us to define a unique natural notion of orthonormality in the space of solutions, enabling an expansion of an arbitrary solution as a linear combination of orthonormal basis solutions.

Such an expansion allows us to arrive at a Fock space decomposition of the state space of the field and identify two types of 'particle number' operators - particle numbers, and antiparticle numbers. However, the number of ways of doing so are as varied as the number of orthonormal bases one may choose in the space of solutions (i.e. infinitely many) necessitating some physically 'reasonable' criterion for selecting a Fock basis, though purely for convenience in choosing a point of view for describing the system. Every such decomposition identifies a corresponding vacuum state, which is an eigenstate of all the particle/antiparticle number operators in the basis with the eigenvalue 0. This state plays a useful role, as we saw that any two choices have differing total particle and antiparticle content if and only if they do not share the vacuum state. However, the difference between the particle number and the antiparticle number is a more universal notion, being related to the invariant Klein-Gordon inner product of the field with itself.

The criteria for selecting a Fock decomposition are typically based on the choice of a spacelike surface, with different spacelike surfaces yielding different vacuum states according to the chosen criterion. Given a family of spacelike surfaces (equivalently, a 'time coordinate') that we have chosen to study the dynamics of the field in, the particle/antiparticle content therefore varies with time, which is the 'phenomenon' of particle production. We discussed a class of modes, namely, the positive and negative frequency modes at a given spacelike surface, that are of particular convenience in describing particle production. We then described two commonly used criteria for choosing a vacuum at any time: instantaneous Hamiltonian diagonalization, and the adiabatic vacuum prescription. For separable modes, these criteria lend themselves readily to numerical calculation. This was the basic set-up we needed for studying various examples.

We began with examples in flat spacetime, firstly the classic case of the constant electric field. While our method relied on the asymptotic behaviour of mode functions at future and past infinity, we saw using some general arguments how the time-translation invariance of this problem is respected by the particle creation process. Numerical computations allowed us to determine the particle content (using both vacuum criteria discussed) at intermediate times, and we found that the adiabatic vacuum suppressed the magnitude of oscillations of the particle content in time, but at higher orders, led to the emergence of a new feature (an intermediate time peak in particle content), possibly caused by poor applicability of the WKB approximation in these regions to the corresponding order. Similar asymptotic calculations as well as numerical computations were done for the spatially homogeneous Sauter pulsed electric field.

After developing some familiarity with the flat spacetime Schwinger effect, we tackled the same for constant magnitude electric fields in spatially flat FLRW universes, with the de Sitter and power law universes with accelerating expansion being of interest. The de Sitter problem lends itself readily to the analytical asymptotic treatment applied to the flat spacetime examples, but only in a particular limit (nevertheless, we could obtain exact forms for the mode functions, and therefore exact solutions in principle), and we made use of numerical computation to reliably study the other cases. For the Schwinger effect in a power law spacetime, not even the exact solutions for mode functions are known, and we had to restrict ourselves completely to a numerical analysis of the problem.

Finally, we briefly reviewed an application of the curved spacetime Schwinger effect to inflationary cosmology - the phenomenon of inflationary magnetogenesis. We described how pair creation via the Schwinger effect can impede the process of magnetogenesis, and referred to computations of the limit this enforces on the magnitude of the fields. We concluded with a qualitative description of a shorting-out of the electric field after inflation, resulting in a leftover large-scale magnetic field in the universe.

Bibliography

[1] Julian Schwinger. "On gauge invariance and vacuum polarization." Physical Review 82, no. 5 (1951): 664.

[2] Stephen A. Fulling. *Aspects of quantum field theory in curved spacetime*. London Mathematical Society Student Texts. Cambridge University Press, 1989.

[3] Leonard Parker. "Particle creation in expanding universes." Physical Review Letters 21, no. 8 (1968): 562.

[4] Stephen W. Hawking. "Particle creation by black holes." Communications in mathematical physics 43, no. 3 (1975): 199-220.

[5] Nicholas D. Birrell, and Paul C. W. Davies. *Quantum fields in curved space*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1984.

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

[7] I. S. Gradshteyn, and I. M. Ryzhik. *Table of integrals, series, and products*. Seventh Edition. Academic Press, 2007.

[8] Lewis H. Ryder. *Quantum field theory*. Cambridge University Press, 1996.

[9] Walter Greiner, and Joachim Reinhardt. *Field quantization*. Springer Science & Business Media, 2013.

[10] Sean M. Carroll. *Spacetime and geometry. An introduction to general relativity*. Pearson, 2003.

[11] Paul Adrien Maurice Dirac. *The principles of quantum mechanics*. The International Series of Monographs on Physics. Oxford University Press, 1981.

[12] Takeshi Kobayashi, and Niayesh Afshordi. "Schwinger effect in 4D de Sitter space and constraints on magnetogenesis in the early universe." Journal of High Energy Physics 2014, no. 10 (2014): 166. arxiv:1408.4141 [hep-th].

[13] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Indefinite linear algebra and applications*. Springer Science & Business Media, 2006.

[14] Jaume Garriga. "Pair production by an electric field in (1+ 1)-dimensional de Sitter space." Physical Review D 49, no. 12 (1994): 6343.

[15] Gaurang Mahajan, and T. Padmanabhan. "Particle creation, classicality and related issues in quantum field theory: II. Examples from field theory." General Relativity and Gravitation 40, no. 4 (2008): 709-747. arxiv:0708.1237 [gr-qc].

[16] Ramkishor Sharma, and Suprit Singh. "Multifaceted Schwinger effect in de Sitter space." Physical Review D 96, no. 2 (2017): 025012. arxiv:1704.05076 [gr-qc].

[17] Sergei Winitzki. "Cosmological particle production and the precision of the WKB approximation." Physical Review D 72, no. 10 (2005): 104011. arXiv:0510001 [gr-qc].

[18] George B. Arfken, and Hans-Jurgen Weber. *Mathematical methods for physicists*. Sixth Edition. Academic Press, 2005.

[19] Wolfram Research, Inc., Mathematica, Version 11.0.1.0, Champaign, IL (2016).

[20] L. Sriramkumar. "Quantum Fields in Non-trivial Backgrounds." Ph. D thesis, IUCAA, 1997.

[21] T. Padmanabhan. "Quantum theory in external electromagnetic and gravitational fields: A comparison of some conceptual issues." Pramana 37, no. 3 (1991): 179-233.

[22] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.18 of 2018-03-27. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders, eds.

[23] Karthik Rajeev, Sumanta Chakraborty, and T. Padmanabhan. "A comment on generalized Schwinger effect." arXiv preprint arXiv:1712.06621 [gr-qc] (2017).

[24] Philip M. Morse, and Herman Feshbach. *Methods of theoretical physics, Vol. II*. International Student Edition. McGraw-Hill Book Company, 1953.

[25] Markus B. Fröb, Jaume Garriga, Sugumi Kanno, Misao Sasaki, Jiro Soda, Takahiro Tanaka, and Alexander Vilenkin. "Schwinger effect in de Sitter space." Journal of Cosmology and Astroparticle Physics 2014, no. 04 (2014): 009. arxiv:1401.4137 [hep-th].

[26] L. Sriramkumar. "An introduction to inflation and cosmological perturbation theory." Current science (2009): 868-886. arXiv:0904.4584 [astro-ph.CO].

[27] Ramkishor Sharma, Sandhya Jagannathan, T. R. Seshadri, and Kandaswamy Subramanian. "Challenges in inflationary magnetogenesis: Constraints from strong coupling, backreaction, and the Schwinger effect." Physical Review D 96, no. 8 (2017): 083511. arXiv:1708.08119 [astro-ph.CO].

[28] Jerome Martin, and Jun'ichi Yokoyama. "Generation of large scale magnetic fields in single-field inflation." Journal of Cosmology and Astroparticle Physics 2008, no. 01 (2008): 025. arXiv:0711.4307 [astro-ph].