Cosmic Energy Density: Particles, Fields and the Vacuum

C. Armendariz-Picon[a] A. Diez-Tejedor[b]

[a]Physics Department
St. Lawrence University,
23 Romoda Dr., Canton, NY 13617, USA.

[b]Departamento de Física, División de Ciencias e Ingenierías
Campus León, Universidad de Guanajuato, León 37150, México

E-mail: carmendarizpicon@stlawu.edu, alberto.diez@fisica.ugto.mx

Abstract. We revisit the evolution of the vacuum energy density of a scalar field upon a transition between different cosmic epochs, and assess under what conditions the particle production formalism reproduces its correct value. Because the unrenormalized energy-momentum tensor diverges in the ultraviolet, it is necessary to frame our discussion within an appropriate regularization and renormalization scheme. Pauli-Villars avoids some of the drawbacks of adiabatic subtraction and dimensional regularization and is particularly convenient in this context. To illustrate our results we focus on transitions from inflation to radiation domination, discuss how the energy density depends on the details of the transition and emphasize those aspects that are model-independent. Along the way, we consider a transition that avoids some of the problems of the one commonly studied in the literature, and point out some instances in which the particle production formalism has led to the incorrect energy density.

Contents

1 Introduction

A dichotomy pervades many of the current and past analyses of the evolution of the universe and its content. Whereas we typically assume that the theories that describe nature are *field* theories, we often frame our analyses of cosmic phenomena in terms of *particles*. Yet cosmic expansion and structure formation are basically determined by the energy and pressure of matter, and it is hence irrelevant in this context whether matter consists of particles or not. In the end, what counts is the expectation of the energy-momentum tensor, and the notion of particle appears to be secondary if not superfluous. Say, all the existing predictions of the properties of the primordial perturbations, as well as many characterizations of cosmic structure, rely on *field* correlations. Certainly there are cases when a particle interpretation is convenient, such as when one sets up underground dark matter direct detection experiments, but these cases appear to be exceptions, rather than the rule. To some extent, this philosophy is also orthogonal to the one that prevails in particle physics, in which particles are the fundamental blocks of nature, and fields are simply a tool to construct an S-matrix with the appropriate properties [1].

In this manuscript we explore an aspect of cosmic evolution in which the particle concept, justifiably or not, also appears to play an important role: cosmic transitions. By "cosmic transitions" we mean the transition between different cosmic epochs, such as between inflation and radiation domination, or between radiation domination and matter domination. According to the conventional wisdom, departures from adiabaticity during such transitions cause the gravitational production of particles out of the vacuum, which leave an imprint on cosmic evolution through their energy density and pressure. The production is "gravitational" because the corresponding fields only interact gravitationally. These transitions are not only interesting from a theoretical perspective, but they may also have important phenomenological consequences. At this point, absent a direct detection of dark matter, an hypothesis that appears to be gaining ground is that dark matter only couples to gravitation [2]. If that were the case, there ought to be a mechanism to produce the density required by analyses of structure formation. Gravitational particle production at the end of inflation is one possibility [3–7]. Similarly, there appears to be a tension between the need to keep the couplings of the inflaton to matter small, in order to protect the flatness of its potential, and that of making them large enough in order for the inflaton to efficiently decay and reheat the universe [8]. If, instead, radiation could be produced gravitationally at the end of inflation [9], the inflaton would not need to couple to matter directly, and the problems related to the flatness of its potential might thus be ameliorated.

Our objective here is not only to bypass the particle description by directly computing the expectation of the energy density upon cosmic transitions like the above, but also to establish under which general conditions the particle interpretation yields a quantitatively correct estimate of the energy density, and under what circumstances the latter can be thought of as the contribution of a "classical" homogeneous scalar. The literature on gravitational particle production, particularly in the context of cosmology, is vast. Fortunately, a significant fraction of its foundations is reviewed in the now standard monographs [10–12]. Yet surprisingly, none of these references discusses the precise relation between the field and particle production formalisms, nor how renormalization enters the latter, and the resulting confusion has diffused into the literature on cosmic transitions and beyond. This is a gap that this manuscript also aims to close.

As far as particle production after inflation is concerned, the seminal article on the topic is that of Ford [9], in which particle production is used to estimate the energy density of a nearly conformally coupled massless field upon a cosmological transition from inflation to radiation domination. Ford's results were also critically revisited and extended by Yajnik [13], whose focus on the energy-momentum tensor resonates with the one of this manuscript, although with different conclusions. Particle production was also exploited in [3] to determine the energy density of the particles produced during inflation, and the initial particle production calculations were further generalized to other cosmological scenarios in [14]. As we shall see, our analysis uncovers serious shortcomings in some of the conventional implementations of particle production. At long wavelengths, previous approaches have missed or ignored important contributions to the energy-momentum tensor, whereas at short wavelengths ultraviolet divergent *renormalized* energy densities render some transitions unphysical.

Our analysis also overlaps with works that directly consider the renormalized energy density of a scalar after a transition from inflation, such as those in references [15–17]. On top of a different focus and scope, our treatment deviates form these in several important aspects. Whereas previous work appears to take highly idealized cosmic transitions rather literally, we emphasize here their model-independent features, and how these survive in more realistic cases and with an expanded parameter space. In addition, we rely on Pauli-Villars regularization [18], which offers notable advantages when compared with the commonly employed adiabatic scheme and dimensional regularization [19]. In particular, the adiabatic scheme does not render the role of the counterterms explicit and fails when applied to massless fields. Dimensional regularization is difficult to work with and it is not very useful when the mode functions are not analytically known, as happens in realistic cosmic transitions. These problems are avoided with Pauli-Villars, which also allows us to discuss the contribution to the energy density mode by mode, and to separate the infrared from the ultraviolet. It is worth stressing that, because the unrenormalized energy density diverges, it is crucial to properly regularize and renormalize the appropriate expectation value in order to obtain sensible results that we can connect with the particle production formalism.

Regarding the conventions that we follow in this paper, we use the mostly plus metric signature $(-,+,+,+)$ and work with natural units where $\hbar = c = 1$.

2 Formalism

Our main focus is the evolution of the energy density of a free, real, test scalar field ϕ coupled to gravity,

$$S_\phi = -\frac{1}{2} \int d^4x \sqrt{-g} \left(\partial_\mu \phi \partial^\mu \phi + m^2 \phi^2 + \xi R \phi^2 \right). \tag{2.1}$$

Gauge fields and massless fermions are conformally invariant, and do not react to changes in the expansion history. To avoid similar conclusions in the massless case $m = 0$, we assume that the scalar ϕ is minimally coupled to gravity ($\xi = 0$), as opposed to conformally coupled ($\xi = 1/6$). Even though we set $\xi = 0$, our results should remain qualitatively the same for nonzero ξ as long as it is not too close to the conformal value.

We also assume that the spacetime metric is that of a spatially flat homogeneous and isotropic universe,

$$ds^2 = a^2(\eta)\left(-d\eta^2 + d\vec{x}\cdot d\vec{x}\right), \tag{2.2}$$

where a is the scale factor and η labels conformal time. In these coordinates, the comoving Hubble parameter is $\mathcal{H} \equiv \dot{a}/a$, which is related to the "physical" Hubble constant H by $\mathcal{H} = aH$. Similarly, while m refers to the actual mass of the field, it will often be convenient to consider its comoving mass ma.

Our goal is to track changes in the scalar energy density as the universe transitions from an *in* region during which the universe inflates, to an *out* region during which the expansion is decelerating, say, during radiation domination. At this point *in* and *out* are to be regarded as labels for the two different epochs, and no connection with the particle production formalism is implied or required.

2.1 Mode Functions

In order to quantize the scalar field, we expand it into decoupled momentum modes,

$$\phi = \frac{1}{\sqrt{V}}\frac{1}{a}\sum_{\vec{k}}\left[a_{\vec{k}}\,\chi_k(\eta)e^{i\vec{k}\cdot\vec{x}} + a^\dagger_{\vec{k}}\,\chi^*_k(\eta)e^{-i\vec{k}\cdot\vec{x}}\right]. \tag{2.3}$$

In this expression V is the (monentarily finite) comoving volume of the universe. The creation operators $a^\dagger_{\vec{k}}$ can be interpreted as creating "particles" of comoving momentum $\vec{k}$, and constitute the only instance in which the particle notion shall sneak into our analysis. In fact, these states are actually eigenvectors of the momentum operator, and do not really represent localized particles.

The mode functions χ_k in the expansion (2.3) satisfy the mode equation

$$\ddot{\chi}_k + \hat{\omega}_k^2\,\chi_k = 0, \quad \text{where} \quad \hat{\omega}_k^2 \equiv \omega_k^2 - \frac{\ddot{a}}{a} \quad \text{and} \quad \omega_k^2 \equiv k^2 + m^2a^2. \tag{2.4}$$

The quantity $\hat{\omega}_k$ corresponds to the oscillation frequency of the mode functions χ_k, which is not the same as that of the actual modes of the scalar field, ω_k, since we have rescaled the latter by $1/a$. The canonical commutation relations imply the normalization condition

$$\chi_k\dot{\chi}^*_k - \dot{\chi}_k\chi^*_k = i, \tag{2.5}$$

where a dot denotes a derivative with respect to conformal time. In particular, because of the homogeneity and isotropy of the metric, we may assume that the mode functions only depend on $k \equiv |\vec{k}|$.

The zero mode $\vec{k} = 0$ needs special consideration, not just because symmetry allows it to have a nonzero expectation, but also because in general there is no preferred state for that mode. States typically considered in the literature are classical-like, with a nonzero expectation that evolves like a homogeneous classical field. Because the different modes of a free test field decouple, the zero mode can be treated separately.

In general, there are no exact analytical solutions to the mode equation (2.4), so we shall rely on approximate solutions instead. In the next two subsections we present high and low frequency expansions that will be useful in the remainder of this work.

2.1.1 High Frequencies

At high frequencies we can obtain approximate solutions of the mode equation (2.4) by adopting an "adiabatic" expansion in the number of time derivatives,

$$\boxed{\chi_k^{(n)}(\eta) = \frac{1}{\sqrt{2W_k^{(n)}(\eta)}} \exp\left(-i \int^{\eta} W_k^{(n)}(\tilde{\eta})\, d\tilde{\eta}\right).} \tag{2.6}$$

These adiabatic mode functions are n-derivative approximations to the actual solutions of the mode equation. In particular, the $W_k^{(n)}$ themselves are n-derivative approximate solutions of the differential equation

$$W_k^2 = \omega_k^2 - \frac{\ddot{a}}{a} - \frac{1}{2}\left(\frac{\ddot{W}_k}{W_k} - \frac{3}{2}\frac{\dot{W}_k^2}{W_k^2}\right), \tag{2.7}$$

which follows from the mode equation (2.4) upon substitution of the ansatz (2.6). The nature of the approximation thus guarantees that n is a positive even number. Up to four derivatives, the $W_k^{(n)}$ are

$$W_k^{(0)} = \omega_k, \quad W_k^{(2)} = W_k^{(0)} + {}^{(2)}W_k, \quad W_k^{(4)} = W_k^{(2)} + {}^{(4)}W_k, \tag{2.8a}$$

where

$${}^{(2)}W_k = \frac{3}{8}\frac{\dot{\omega}_k^2}{\omega_k^3} - \frac{\ddot{\omega}_k}{4\omega_k^2} - \frac{1}{2\omega_k}\frac{\ddot{a}}{a}, \tag{2.8b}$$

$$\begin{aligned} {}^{(4)}W_k = &-\frac{297}{128}\frac{\dot{\omega}_k^4}{\omega_k^7} + \frac{1}{4}\frac{\dot{a}^2\,\ddot{a}}{a^3\,\omega_k^3} + \frac{5}{8}\frac{\dot{a}\,\ddot{a}\,\dot{\omega}_k}{a^2\,\omega_k^4} + \frac{19}{16}\frac{\ddot{a}\,\dot{\omega}_k^2}{a\,\omega_k^5} - \frac{\ddot{a}^2}{4a^2\,\omega_k^3} + \frac{99}{32}\frac{\dot{\omega}_k^2\,\ddot{\omega}_k}{\omega_k^6} \\ &- \frac{3}{8}\frac{\ddot{a}\,\ddot{\omega}_k}{a\,\omega_k^4} - \frac{13}{32}\frac{\ddot{\omega}_k^2}{\omega_k^5} - \frac{\dot{a}\,\dddot{a}}{4a^2\,\omega_k^3} - \frac{5}{8}\frac{\dddot{a}\,\dot{\omega}_k}{a\,\omega_k^4} - \frac{5}{8}\frac{\dot{\omega}_k\,\dddot{\omega}_k}{\omega_k^5} + \frac{1}{8}\frac{a^{(4)}}{a\,\omega_k^3} + \frac{1}{16}\frac{\omega_k^{(4)}}{\omega_k^4}. \end{aligned} \tag{2.8c}$$

As we discuss further in Appendix A.1, it follows from this expansion that the adiabatic approximation generically (but not necessarily) applies when $\omega_k \gg \mathcal{H}$, that is, for sufficiently short wavelength modes or massive fields. But, strictly speaking, the "adiabatic regime" holds whenever equation (2.6) is a valid approximation to the solution of the mode equation. In subsection 3.3 we discuss an example in which modes are in the adiabatic regime even though their frequencies satisfy $\omega_k = k \ll \mathcal{H}$.

2.1.2 Low Frequencies

We shall rely on the solutions to the mode equation with $\vec{k}=0$ and $m=0$ as lowest order approximations when the mode frequencies ω_k are sufficiently small. For any scale factor $a(\eta)$, the mode function

$$\chi_0^{m=0}(\eta) = a\,(c_1 + c_2\, b)\,, \quad \text{where} \quad b \equiv \int^\eta \frac{d\tilde{\eta}}{a^2}, \tag{2.9}$$

solves equation (2.4) when $\omega_k = 0$, and also satisfies the normalization condition (2.5) when $c_1 c_2^* - c_1^* c_2 = i$. There are no natural conditions to otherwise fix the values of c_1 and c_2, which is related to the arbitrariness of the zero mode state that we previously mentioned. As we shall explicitly see below, this is why it is important that the *out* region be preceded by an inflationary *in* region. Note that we have not specified the lower limit of integration in equation (2.9), which simply shifts the value of c_1 by a constant. In integrals like that of equation (2.9) we shall omit the lower limit of integration when it is physically irrelevant.

If ω_k is nonzero, equation (2.9) is clearly not a solution of the mode equation. Instead, it can be regarded as the lowest order solution of the mode equation in the limit of small ω_k, with a correction $\Delta\chi_k \equiv \chi_k - \chi_0^{m=0}$ that is implicitly determined by

$$\Delta\chi_k = -\int^\eta d\tilde{\eta}\; G(\eta;\tilde{\eta})\,\omega_k^2(\tilde{\eta})\chi_k(\tilde{\eta}), \tag{2.10}$$

where $G(\eta;\tilde{\eta})$ is the Green's function of the zero-frequency equation, which can be readily constructed as a linear combination of the two solutions a and ab we just identified above. Equation (2.4) and its solution $\chi_0^{m=0} + \Delta\chi_k$ are analogous to those encountered in a scattering problem in nonrelativistic quantum mechanics, in which ω_k^2 plays the role of the potential, $\chi_0^{m=0}$ is the incoming wave function and $\Delta\chi_k$ the scattered one. Just as in the scattering problem, we can obtain an explicit solution of the mode equation by recursively expanding (2.10) in powers of ω_k^2. The n-th order in such an expansion thus contains n powers of ω_k and is related to the next one by

$$\chi_k^{(n+2)}(\eta) = \chi_k^{(0)}(\eta) - \int^\eta d\tilde{\eta}\;\left[a(\tilde{\eta})\,a(\eta)b(\eta) - a(\eta)\,a(\tilde{\eta})b(\tilde{\eta})\right]\omega_k^2(\tilde{\eta})\chi_k^{(n)}(\tilde{\eta}), \tag{2.11}$$

where $\chi_k^{(0)} \equiv \chi_0^{m=0}$, n is an even natural number, and we have used the explicit form of the retarded Green's function. In the Born approximation, one keeps just the leading order correction, $n=2$. As we discuss further in appendix A.2, it follows from this expansion that the low frequency approximation generically (but not necessarily) applies when $\omega_k \ll \mathcal{H}$, that is, for sufficiently long wavelength modes or light fields. An exception occurs close to de Sitter, where the expansion in powers of ω_k^2 may fail. At any rate, once the mode frequency becomes large, $\omega_k \gtrsim \mathcal{H}$, the mode functions χ_k start to oscillate and the zeroth order approximation (2.9) breaks down. This is precisely where the adiabatic approximation we discuss next begins to apply.

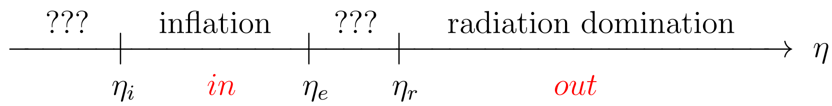

Figure 1. Different epochs of early cosmic evolution. What happened before inflation and between inflation and radiation domination is not fully understood. In a sharp transition $\eta_e = \eta_r$ and there is a discontinuity in the second derivative of the scale factor. In a smooth transition $\eta_e < \eta_r$ and the second derivative of the scale factor is continuous at $\eta = \eta_e$, though there is still a discontinuity in the higher order derivatives. In a differentiable transition $\eta_e < \eta_r$ and all the derivatives of the scale factor are continuous.

2.2 Cosmological Epochs

In order to single out the "initial" state of the quantum field, we shall consider a spacetime that proceeds from an initial inflationary *in* region of accelerated expansion to a subsequent decelerated *out* region where the universe expands as dominated by radiation, after going through an intermediate transition epoch during reheating. In this section we present the details of the two asymptotic regimes, as well as some models for the transition between them. Figure 1 sketches a timeline of our model universe.

2.2.1 *In* Region

The mode expansion of the field (2.3) allows us to introduce the notion of the number of quanta in each momentum mode. Yet the notion of vacuum, as well as the associated particle number, relies on the particular choice of mode functions, which are arbitrary up to the normalization condition (2.5). We are primarily interested here in transitions from an *in* region in which a unique notion of vacuum exists for all modes. This is the case if for any fixed value of k the mode equation admits solutions that in the asymptotic past match the zeroth order "positive frequency" adiabatic solutions we identified above,

$$\chi_k^{\text{in}} \to \frac{1}{\sqrt{2\omega_k}} \exp\left(-i\int^\eta \omega_k(\tilde{\eta})\, d\tilde{\eta}\right). \tag{2.12}$$

This condition is met in any universe that underwent an early period of inflation, not necessarily of the de Sitter type. Then $\ddot{a}/a$ tends to zero in the asymptotic past, and the mode equation admits solutions that approach (2.12) as $\eta \to -\infty$. Since as $k \to \infty$ the adiabatic approximation always remains valid, it also follows then that the mode functions approach (2.12) in the ultraviolet at any moment of cosmic history.

Strictly speaking, though, inflation cannot be past eternal [20]. When we set initial conditions for the field with equation (2.12) we act as if the *in* region extended all the way to the asymptotic past. But this is just a device to single out the quantum state of the field (the *in* mode functions), and it does not really presuppose that inflation extends indefinitely into the past. In practice, the finite duration of inflation does require the introduction of an infrared cutoff $\Lambda_{\text{IR}} \sim \mathcal{H}_i$, since there is no preferred quantum state for superhorizon modes at the beginning of inflation. In most cases, however, the total energy density of modes above $\Lambda_{\rm IR}$ is not sensitive to the actual value of the infrared cutoff.

When the mode functions in the mode expansion (2.3) satisfy the *in* initial conditions (2.12), the corresponding ladder operators $a^{\rm in}_{\vec{k}}$ can be identified with those of the *in* states. We define then the particle number operator associated with the mode $\vec{k}$ by $N^{\rm in}_{\vec{k}} \equiv a^{{\rm in}\dagger}_{\vec{k}} a^{\rm in}_{\vec{k}}$. The "*in* vacuum" is the state with no *in* quanta, $N^{\rm in}_{\vec{k}} |0_{\rm in}\rangle = 0$, while states with $N^{\rm in}_{\vec{k}} |\psi\rangle = n^{\rm in}_{\vec{k}} |\psi\rangle$ can be thought of as containing a definite number of quanta $n^{\rm in}_{\vec{k}}$. This is how inflation allows us to identify the quantum state of all modes with $k > \Lambda_{\rm IR}$, the *in* vacuum.

According to the inflationary paradigm, the early universe actually went through a stage of accelerated expansion during which the scale factor grew almost as in de Sitter spacetime. Although de Sitter is often a good approximation for inflation, its group of symmetries is larger than that of a generic (inflationary) Friedman-Robertson-Walker spacetime. We shall thus consider a generic power-law expansion, which allows us to work with a one parameter group of inflationary spacetimes that includes de Sitter,

$$a = a_0 \left(\frac{\eta}{\eta_0}\right)^p, \tag{2.13}$$

where $p \leq -1$ and $\eta_0 < 0$. The effective equation of state in such a universe is $w = (2-p)/3p$, which ranges from that of a cosmological constant at $p = -1$ to that of curvature at $p = -\infty$. Also note that there is a direct connection between p and the slow roll parameter $\epsilon_1 \equiv -\dot{H}/(aH^2) = (1+p)/p$, which is clearly constant during power-law inflation. The time η_0 in equation (2.13) is an arbitrary time during inflation that we shall choose by convenience, and may take different values in different contexts.

2.2.2 *Out* Region

The phenomenological success of the hot Big Bang model implies that inflation must have been followed by a radiation dominated epoch, albeit not necessarily right away. Hence, for phenomenological reasons we shall consider an *out* region that contains a radiation dominated universe starting at $\eta = \eta_r > \eta_e$. This choice is not only realistic, but also streamlines the analysis considerably, because in a radiation-dominated universe $\ddot{a}/a = 0$, which simplifies the mode equation (2.4) significantly.

In order to determine the mode functions in the *out* region, we need to find the solution of equation (2.4) that evolved through the transition from the one in the *in* region. For that purpose, it is often convenient to work with a set of solutions of (2.4) that do not necessarily satisfy the required initial conditions set by the preceding inflationary period. Let then $\chi^{\rm in}_k$ be the solution of (2.4) with the appropriate *in* boundary conditions, and let χ_k be an arbitrary solution of equation (2.4). We shall refer to the χ_k generically as the *out* mode functions, although they are not necessarily related to the *out* region. Assuming that χ_k and χ_k^* are linearly independent, we can express the solution $\chi^{\rm in}_k$ as a linear combination of the two,

$$\boxed{\chi^{\rm in}_k(\eta) = \alpha_k \chi_k(\eta) + \beta_k \chi^*_k(\eta).} \tag{2.14}$$

This equation has the form of a Bogolubov transformation. At this point the nature of the χ_k is irrelevant. We just demand that they solve (2.4) and that they satisfy the normalization condition (2.5). The Bogolubov coefficients α_k and β_k are then constant and, because of equation (2.5), they satisfy

$$|\alpha_k|^2 - |\beta_k|^2 = 1. \tag{2.15}$$

Combining equation (2.14) and its first derivative to solve for α_k and β_k, and making use of the normalization of the mode functions (2.5), we arrive at

$$i\alpha_k = \ \chi_k^{\text{in}}(\eta)\, \dot{\chi}_k^*(\eta) - \dot{\chi}_k^{\text{in}}(\eta)\, \chi_k^*(\eta), \tag{2.16a}$$

n_e 2.6

$$i\beta_k = -\chi_k^{\text{in}}(\eta)\, \dot{\chi}_k(\eta) + \dot{\chi}_k^{\text{in}}(\eta)\, \chi_k(\eta). \tag{2.16b}$$

The Bogolubov coefficients (2.16) are readily seen to satisfy equation (2.15). If, say, $\chi_k = \chi_k^{\text{in}}$, these equations imply that $\alpha_k = 1$ and $\beta_k = 0$, as expected. Note that, even though χ_k^{in} and χ_k are time-dependent, the Bogolubov coefficients are constant.

Substitution of equation (2.14) into (2.3) returns an analogous expansion, simply with all the mode functions and ladder operators replaced by their *out* counterparts, provided that we identify

$$\boxed{a_{\vec{k}} \equiv \alpha_k a_{\vec{k}}^{\text{in}} + \beta_k^* a_{-\vec{k}}^{\text{in}\dagger}.} \tag{2.17}$$

The two mode expansions allow us then to define the number of *in* and *out* particles in a given mode $\vec{k}$ as the expectation of the number operators $N_{\vec{k}}^{\text{in}} \equiv a^{\text{in}\dagger}_{\vec{k}} a_{\vec{k}}^{\text{in}}$ and $N_{\vec{k}} \equiv a_{\vec{k}}^\dagger a_{\vec{k}}$, respectively. By definition the *in* vacuum $|0_{\text{in}}\rangle$ contains no *in* particles, $a_{\vec{k}}^{\text{in}}|0_{\text{in}}\rangle = 0$, and the *out* vacuum $|0\rangle$ contains no *out* particles, $a_{\vec{k}}|0\rangle = 0$. But when β_k is nonzero, the *in* vacuum does contain *out* particles, $\langle 0_{\text{in}}|N_{\vec{k}}|0_{\text{in}}\rangle = |\beta_k|^2$, although it is not an eigenvector of the number operator, $\langle 0_{\text{in}}|a_{\vec{k}} a_{-\vec{k}}|0_{\text{in}}\rangle = \alpha_k \beta_k^*$. As a matter of fact, the *in* vacuum of a single mode $\vec{k}$ can be regarded as a two-mode squeezed state with an indefinite number of entangled *out* particles of momentum $\vec{k}$ and $-\vec{k}$ [21]. We shall extend the discussion of particle production to the scalar energy density in subsection 2.3.3.

In general, though, the Bogolubov coefficients do not have an independent meaning by themselves, since they are inherently linked to the arbitrary choice of mode functions χ_k. Only in the adiabatic regime, when the latter are chosen to approach the positive frequency approximations (2.6) (at a given order n), do they acquire a relatively context-independent significance. The *exact* solution of the mode equation that matches (2.6) and its first derivative at any chosen and fixed time η implicitly defines the n-th order adiabatic vacuum at that time. As stressed in reference [10], this defines a two-parameter family of vacua, characterized by the adiabatic order n and the time η.

2.2.3 Transitions

We have now established the nature of the *in* and *out* regions, but still need to specify how the transition between the two occurs. Because the transition itself clearly depends on the unknown details of reheating, we shall not make any specific assumptions about its properties. When specific examples are necessary, the following three transitions to radiation domination shall prove to be useful.

Sharp Transition. The simplest example consists in taking the limit in which the duration of the transition goes to zero, $\eta_r \to \eta_e$, by just matching the scale factor and its first derivative at the time of the transition, $\eta_e < 0$, as it is mostly done in the literature on the topic. If the energy-momentum tensor remains finite, the continuity of a and $\dot{a}$ is a consequence of the Darmois-Israel junction conditions [22]. By imposing the continuity of the scale factor and its derivative at time η_e we find

$$a = a_e + \dot{a}_e \left(\eta - \boxed{\eta_e}\right) \qquad (\eta > \eta_e), \tag{2.18}$$

where a_e and $\dot{a}_e$ respectively are the value of the scale factor and its first derivative at the end of inflation. This corresponds to a sharp change in the equation of state from $-1 \leq w \leq -1/3$ for $\eta < \eta_e$, to $w = 1/3$ for $\eta > \eta_e$, while the energy density remains continuous at $\eta = \eta_e$.

As we shall see, though, the main consequence of such an idealization is an unphysical behavior of the scalar energy density in the ultraviolet. The problem with equation (2.18) is that the transition is not sufficiently smooth: Since the second derivative of the scale factor is discontinuous the *renormalized* energy density diverges in the ultraviolet. Thus, strictly speaking, such a transition is not simply unrealistic; it is also unphysical.

Smooth Transition. We thus prefer to consider a scale factor with continuous second time derivatives,

$$a = c \left[K_0(e^{-r\eta}) + d\, I_0(e^{-r\eta})\right] \qquad (\eta > \eta_e), \tag{2.19}$$

where I_ν and K_ν respectively are the first and second kind modified Bessel functions of order ν, and r is a positive constant with dimensions of inverse time. The main advantage of equation (2.19) is that

$$\frac{\ddot{a}}{a} = r^2 e^{-2r\eta}, \tag{2.20}$$

which rapidly approaches zero, as in a radiation-dominated universe, and also simplifies the mode equation after the end of inflation. With this choice the energy density and the equation of state remain continuous at η_e. Imposing the continuity of $\ddot{a}/a$ at η_e fixes the value of the otherwise arbitrary r,

$$r\eta_e = -W_0\left(\sqrt{p(p-1)}\right). \tag{2.21}$$

Here, W_0 is the principal branch of the Lambert function. In the de Sitter limit, for instance, $r\eta_e = -W_0(\sqrt{2}) \approx -0.7$. Note in particular that r is of the order of the comoving Hubble constant at time η_e.

The constants c and d in equation (2.19) are determined by the continuity of the scale factor and its first derivative at the transition,

$$c = \frac{a_e}{r\eta_e}\left[pI_0 - \sqrt{p(p-1)}I_1\right], \quad d = -\frac{pK_0 + \sqrt{p(p-1)}K_1}{pI_0 - \sqrt{p(p-1)}I_1}, \tag{2.22}$$

where the Bessel functions are evaluated at $z = e^{-r\eta_e}$ and we have used their Wronskian to simplify c. It is easy to explicitly see that (2.19) approaches a radiation dominated universe at large η,

$$a \to c\,(r\,\eta + d + \log 2 - \gamma) \qquad (r\eta \to \infty), \tag{2.23}$$

where γ is the Euler-Mascheroni constant. By construction, both $\dot{a}(\eta)$ and $\ddot{a}(\eta)$ remain positive after the transition, so the scale factor and its derivative grow monotonically for any value of $p \leq -1$. An additional advantage of such a smooth transition is that the time between the end of inflation and the onset of radiation domination is finite, as expected in any realistic cosmology. In order to determine the time η_r after which it is safe to assume that the universe is radiation dominated, we shall impose $\sqrt{\ddot{a}/(a\mathcal{H}^2)} \leq 10^{-1}$. This condition for instance implies that the comoving Hubble scale at time η_r is $\mathcal{H}_r \approx 0.3\,\mathcal{H}_e$ after a transition from de Sitter.

As we shall see, even though the ultraviolet divergence that was present in the sharp transition is absent here, the jump in the third and higher derivatives of the scale factor will lead to substantial particle production in the ultraviolet. This, however, is also unrealistic, since we expect the scale factor and its derivatives to evolve smoothly across the transition, as we discuss next.

Differentiable Transition. Finally, to illustrate some of our results, we shall also consider a "realistic" example based on chaotic inflation with potential [23]

$$V(\varphi) = \frac{\lambda}{4}\varphi^4. \tag{2.24}$$

Although this inflationary potential was ruled out long ago by observations of the cosmic microwave background [24], it does provide a simple model to study a transition to a radiation-dominated universe in which all the derivatives of the scale factor remain continuous (hence the name "differentiable.") Indeed, after the end of inflation, around $H_e \approx 2\sqrt{\lambda}M_P$, the oscillations of the inflaton about the minimum of a quartic potential lead to an inflaton energy density that behaves on average like radiation [25]. At the time at which a radiation dominated universe becomes a good approximation for the scale factor evolution the Hubble parameter is $H_r \approx 0.1\sqrt{\lambda}M_P$ and the universe has expanded by a factor $a_r/a_e \approx 4$; hence, $\mathcal{H}_r \approx 0.2\mathcal{H}_e$.

As opposed to what we have assumed in the foregoing, in this model the equation of state during inflation does not remain constant. By matching the equation of state of the inflaton to the exponent in equation (2.13) we find that

$$p \equiv -1 + \Delta p, \quad \Delta p = -\left(\log\frac{a_e}{a} + \frac{1}{3}\right)^{-1}, \tag{2.25}$$

where we have used the slow-roll approximation and how ϕ depends on the scale factor [26]. Hence, the value of p deviates significantly from the de Sitter value $p = -1$ as the end of inflation nears. This is relevant because, as we shall see, the energy density of the scalar field depends on the value of p during inflation.

As we shall show below, the main difference with the two previous transitions is that in the differentiable case the production of particles is highly suppressed in the ultraviolet. This is what we expect to happen in a realistic transition.

Bogolubov Coefficients In order to determine the form of the *in* mode functions after a jump in the derivatives of the scale factor, we need to find the solution of equation (2.4) that matches the one in the *in* region at the transition time $\eta = \eta_e$. To do so, we shall use an arbitrary set of mode functions χ_k that solve (2.4) after the transition. Because the mode equation is second order, imposing continuity of the solution and its derivative at the future boundary of the *in* region we arrive at the same expressions for the Bogolubov coefficients as in equations (2.14) and (2.16), with the arbitrary η in (2.16) now replaced by the matching time η_e.

Provided that the mode function χ_k^{in} remains in the adiabatic regime throughout the *in* region, and under the assumption that the mode function χ_k is well approximated by the adiabatic expression (2.6), we can already estimate the Bogolubov coefficients after such a transition,

$$\alpha_k \approx 1 + i\frac{({}^{(2)}W_k^+/\omega_k)^\cdot - ({}^{(2)}W_k^-/\omega_k)^\cdot}{4\omega_k} + \frac{({}^{(2)}W_k^+ - {}^{(2)}W_k^-)^2}{8\omega_k^2} + \cdots , \tag{2.26a}$$

$$\beta_k \approx \frac{{}^{(2)}W_k^+ - {}^{(2)}W_k^-}{2\omega_k} - i\frac{({}^{(2)}W_k^+/\omega_k)^\cdot - ({}^{(2)}W_k^-/\omega_k)^\cdot}{4\omega_k} - \cdots , \tag{2.26b}$$

where we have used the continuity of the scale factor and its first derivative at η_e, and the plus and minus superscripts denote the limits in which η approaches η_e from above and below, respectively. Note that the different terms are organized by growing number of time derivatives. These formulas establish a link between the smoothness of the transition and the behavior of the Bogolubov coefficients in the adiabatic regime. If the transition remains differentiable (in the sense above), the coefficients β_k vanish in the adiabatic regime, as expected. As we have previously argued, the modulus square $|\beta_k|^2$ is the expected number of *out* particles in that mode, and particle production hence requires departures from adiabaticity.

In practice, any realistic transition is expected to be differentiable, so it cannot affect the ultraviolet, which is adiabatic at any time. On the other hand, on causality grounds one also expects that the long wavelength modes at the end of inflation are unaltered by the transition. Consequently, as far as the scalar energy density is concerned, the exact nature of the transition is relatively unimportant, as long as it is realistic. In subsequent sections we discuss this universality in more detail.

2.3 Energy Density

As we mentioned earlier, we are concerned here with the evolution of the energy density of the scalar ϕ as the universe transitions from the *in* to the *out* region. Just like the energy density of a classical field is determined by the time-time component of its energy-momentum tensor, in the semiclassical approximation the energy density of our scalar ϕ is related to the expectation of the same component, $\rho \equiv a^2 \langle T^{00} \rangle$. More rigorously, this identification follows, for instance, from the appropriate quantum-corrected equation of motion of the metric [8], where the expectation value of the energy-momentum tensor sources the semiclassical Einstein equations.

To further explore the energy density of the scalar, it is convenient to replace the mode sum in the expectation of the energy-momentum tensor by an integral, by taking the continuum limit $V \to \infty$. In that case, as in the analysis of Bose-Einstein condensation, the contribution of the zero mode ρ_0 needs to be treated separately,

$$\rho = \rho_0 + \int_0^\Lambda \frac{dk}{k} \frac{d\rho}{d\log k}, \tag{2.27a}$$

where

$$\rho_0 = \frac{1}{2a^2 V} \left\{ n_0 \left[\left| \left(\frac{\chi_0}{a} \right)^{\cdot} \right|^2 + m^2 |\chi_0|^2 \right] + m_0 \left[\left(\frac{\chi_0}{a} \right)^{\cdot 2} + m^2 \chi_0^2 \right] + \text{c.c.} \right\}, \tag{2.27b}$$

and

$$\frac{d\rho}{d\log k} \equiv \frac{k^3}{4\pi^2 a^4} \left\{ \left(n_k + \frac{1}{2} \right) \left[a^2 \left| \left(\frac{\chi_k}{a} \right)^{\cdot} \right|^2 + \omega_k^2 |\chi_k|^2 \right] + m_k \left[a^2 \left(\frac{\chi_k}{a} \right)^{\cdot 2} + \omega_k^2 \chi_k^2 \right] + \text{c.c.} \right\}. \tag{2.27c}$$

Note that in equation (2.27b) we have neglected the "quantum" $1/2$ that accompanies n_0, which disappears in the limit $V \to \infty$. Equation (2.27c), is a "spectral density" (per logarithmic interval), that can be interpreted as the energy density in a given mode $k \neq 0$. In these expressions the mode functions χ_k are arbitrary and "c.c." denotes complex conjugation of all the terms inside the curly brackets, including those that are manifestly real. We have also implicitly defined $\langle \hat{a}^\dagger_{\vec{k}} \hat{a}_{\vec{k}'} \rangle = n_k \, \delta(\vec{k} - \vec{k}')$ and $\langle \hat{a}_{\vec{k}} \hat{a}_{\vec{k}'} \rangle = m_k \, \delta(\vec{k} + \vec{k}')$, which are the only non-vanishing expectations allowed by homogeneity and isotropy, with n_k real and m_k complex.[1]

In practice, we will divide the integral in (2.27a) in two pieces

$$\rho_{<\Lambda_{\text{IR}}} \equiv \int_0^{\Lambda_{\text{IR}}} \frac{dk}{k} \frac{d\rho}{d\log k}, \quad \rho_{>\Lambda_{\text{IR}}} \equiv \int_{\Lambda_{\text{IR}}}^{\Lambda} \frac{dk}{k} \frac{d\rho}{d\log k}, \tag{2.28}$$

where $\Lambda_{\text{IR}} \sim \mathcal{H}_i$ is of the order the Hubble radius at the beginning of inflation and plays the role of an infrared cutoff in future analyses. As we discuss next, $\rho_{<\Lambda_{\text{IR}}}$ contains the contribution of those modes in the interval $0 < k < \Lambda_{\text{IR}}$ that were already outside the horizon at the beginning of inflation, and whose state hence remains undetermined. Modes in the interval $\Lambda_{\text{IR}} < k < \infty$, on the contrary, do have a preferred state, though we have limited their contribution up to those below an ultraviolet cutoff Λ that we have introduced for later convenience. Note that if inflation is responsible for the origin of cosmic structure within the horizon, then $\Lambda_{\text{IR}} < \mathcal{H}$ at any time $\eta > \eta_i$.

2.3.1 Modes $0 \leq k < \Lambda_{\text{IR}}$: The Classical Field

Zero Mode. As the volume of the universe approaches infinity the energy density of the zero mode (2.27b) goes to zero, unless the latter is populated by a macroscopic

[1] If the state is invariant under translations by $\vec{\Delta}$, and the latter are represented by the unitary operator U, then $\langle \hat{a}_{\vec{k}} \hat{a}_{\vec{k}'} \rangle = \langle U^\dagger \hat{a}_{\vec{k}} U U^\dagger \hat{a}_{\vec{k}'} U \rangle = e^{-i(\vec{k}+\vec{k}')\cdot\vec{\Delta}} \langle \hat{a}_{\vec{k}} \hat{a}_{\vec{k}'} \rangle$, which implies that the expectation is proportional to $\delta(\vec{k} + \vec{k}')$. Here, we have used the action of U on the annihilation operator, $U^\dagger a_{\vec{k}} U = e^{-i\vec{k}\cdot\vec{\Delta}} a_{\vec{k}}$. A similar argument with rotations indicates that the expectation can only depend on the magnitude of the wave vectors.

condensate with nonzero n_0/V and/or m_0/V. In that case its energy density (2.27b) matches that of a *complex* classical homogeneous scalar

$$\phi_{\rm cl} = \frac{1}{\sqrt{V}}\frac{1}{a}\left(\alpha_0\chi_0 + \beta_0\chi_0^*\right), \tag{2.29}$$

whose energy density is given by

$$\rho_{\rm cl} = \frac{1}{2a^2V}\left\{\frac{|\alpha_0|^2+|\beta_0|^2}{2}\left[\left|\left(\frac{\chi_0}{a}\right)^{\cdot}\right|^2 + m^2|\chi_0|^2\right] + \alpha_0\beta_0^*\left[\left(\frac{\chi_0}{a}\right)^{\cdot 2} + m^2\chi_0^2\right] + \text{c.c.}\right\}, \tag{2.30}$$

provided that we identify $2n_0 = |\alpha_0|^2 + |\beta_0|^2$ and $m_0 = \alpha_0\beta_0^*$. Only when $|m_0| = n_0$ it is possible to satisfy the previous relations by choosing $\beta_0 = \alpha_0^*$, which is the condition necessary for the classical field (2.29) to be real. This is the case, for example, if the zero mode finds itself in a coherent state $|\alpha_0\rangle$, with $a_0|\alpha_0\rangle = \alpha_0|\alpha_0\rangle$.

If the field is massless, the zero mode χ_0 is described by equation (2.9). In this case ρ_0 is determined by the kinetic energy density, which behaves like that of a stiff fluid with equation of state $w = 1$. In particular, the evolution of the energy density is determined by the mode proportional to ab, rather than the one that grows with the scale factor, which drops out of the time derivative of the field. If the mass is different from zero, on the other hand, there are two different regimes. As long as the field remains light, $m \ll H$, we can still approximate $\chi_0 \approx \chi_0^{m=0}$ in equation (2.30), but there is an additional nonvanishing contribution to the energy density stemming from the mass term. As a consequence, in addition to the one that mimics a stiff fluid, there also appears another "frozen field" contribution that behaves like a cosmological constant and dominates the energy density after an initial transient time (we have neglected the contribution of b to the potential energy, which decays in an expanding universe with $-1 \leq w < 1$.) After some time, when the field becomes heavy, $H \ll m$, the low frequency expansion (2.9) breaks down, and we need to make use of the adiabatic approximation, equation (2.6). If we introduce this expansion into equation (2.30), we obtain that at zeroth order in the adiabatic approximation the oscillating zero mode behaves like a pressureless fluid, whose overall density is proportional to the value of n_0. At any rate, since there is no natural way to specify the state of this mode, particularly when the scalar field is light, we shall not dwell on its contribution any further.

Continuum. A similar problem afflicts the energy density $\rho_{<\Lambda_{\rm IR}}$, which captures the contribution of modes whose state we also ignore.

In the massless case, we can set then $\chi_k \approx \chi_0^{m=0}$, as discussed in subsection 2.1.2. Substituting then the known form of $\chi_0^{m=0}$ into (2.28) we find that

$$\begin{aligned}\rho_{<\Lambda_{\rm IR}} \approx &\left[\frac{|c_2|^2}{2}\int_0^{\Lambda_{\rm IR}}\frac{dk}{k}\frac{k^3}{2\pi^2}\left(n_k+\frac{1}{2}\right) + \frac{c_2^2}{2}\int_0^{\Lambda_{\rm IR}}\frac{dk}{k}\frac{k^3}{2\pi^2}m_k + \text{c.c.}\right]\frac{1}{a^6}\\ &+\left[\frac{|c_1|^2}{2}\int_0^{\Lambda_{\rm IR}}\frac{dk}{k}\frac{k^5}{2\pi^2}\left(n_k+\frac{1}{2}\right) + \frac{c_1^2}{2}\int_0^{\Lambda_{\rm IR}}\frac{dk}{k}\frac{k^5}{2\pi^2}m_k + \text{c.c.}\right]\frac{1}{a^2},\end{aligned} \tag{2.31}$$

where in the second line we have approximated $c_1 + c_2 b$ by the constant c_1. Provided that we identify

$$\frac{n_0}{V} = \int_0^{\Lambda_{\rm IR}} \frac{dk}{k} \frac{k^3}{2\pi^2} \left(n_k + \frac{1}{2}\right), \quad \frac{m_0}{V} = \int_0^{\Lambda_{\rm IR}} \frac{dk}{k} \frac{k^3}{2\pi^2} m_k \tag{2.32}$$

the first line of equation (2.31), which contains the time derivatives of the field, can be cast as the energy density of the zero mode (2.27b), whose energy density behaves then like that of a fluid with equation of state $w = 1$. The second line contribution scales like spatial curvature and cannot be interpreted as that of a homogeneous scalar; it arises from the field gradients, which are absent if the scalar is homogeneous.

In the massive case we need to distinguish between two possible limits, When $\Lambda_{\rm IR} < ma$ all the relevant modes are nonrelativistic, so we can approximate the dispersion relation in equation (2.27c) by $\omega_k \approx ma$. This allows us to cast the energy density as that of the zero mode (2.27b), provided that we make the identifications in (2.32). In that case, the density evolution matches the well-known freezing vs. thawing dynamics of a scalar in an expanding universe: When $ma \ll \mathcal{H}$, the mode functions are well approximated by equation (2.9), so $\rho_{<\Lambda_{\rm IR}}$ effectively behaves like a cosmological constant, whereas when $\mathcal{H} \ll ma$, the mode functions are well approximated by the oscillatory (2.6), and the energy scales like that of nonrelativistic matter (see subsection 2.3.3.) Since ma grows monotonically, we expect $\Lambda_{\rm IR} < ma$, to hold at sufficiently late times, and this is the relevant limit then.

If the mass satisfies $ma \leq \Lambda_{\rm IR}$, modes in the interval $0 < k < ma$ are nonrelativistic, so their spectral density behaves like that of a massive case just discussed above. Similarly, modes in the interval $ma < k < \Lambda_{\rm IR}$ are relavistic, and their spectral density behave like that of a massless field. But since the boundary between the two regimes at $k = ma$ changes with time, we cannot in general make definite predictions about the time evolution of the integrated spectral densities. Hence, we shall not study this case explicitly here, though the methods we have discussed so far, along with those we present below, could be similarly deployed. It is nevertheless quite remarkable that, even though we ignore the state of modes with $k < \Lambda_{\rm IR}$, we can still make quite definite predictions about the behavior of their energy density.

2.3.2 Modes $\Lambda_{\rm IR} < k < \infty$: The Particle Interpretation

Modes in the range $\Lambda_{\rm IR} < k$ were inside the horizon at the beginning of inflation. These modes find themselves effectively in Minkowski space, where a preferred choice of state exists: the *in* vacuum. If, as opposed to a general state, the field is in the *in* vacuum, $\chi_k = \chi_k^{\rm in}$ and $n_k^{\rm in} = m_k^{\rm in} = 0$, the energy density $\rho_{>\Lambda_{\rm IR}}$ simplifies to

$$\rho^{\rm in}_{>\Lambda_{\rm IR}} = \int_{\Lambda_{\rm IR}}^{\Lambda} \frac{dk}{k} \frac{d\rho_{\rm in}}{d\log k}, \quad \text{where} \quad \frac{d\rho_{\rm in}}{d\log k} \equiv \frac{k^3}{4\pi^2 a^4}\left[a^2 \left|\left(\frac{\chi_k^{\rm in}}{a}\right)^{\cdot}\right|^2 + \omega_k^2 |\chi_k^{\rm in}|^2\right]. \tag{2.33a}$$

From now on, the energy density ρ and the spectral density $d\rho/d\log k$ will be those of modes above the infrared cutoff in the *in* vacuum, as in equation (2.33a), but for notational simplicity, we shall omit the labels "$> \Lambda_{\mathrm{IR}}$" and "*in*" from our expressions. If we replace the mode functions χ_k^{in} in equation (2.33a) by the generic ones χ_k, then the spectral density becomes that of what we define to be the "*out* vacuum." Clearly, at this point, the nature of this *out* state is arbitrary.

As it stands, the energy density (2.33a) diverges in the ultraviolet, when $\Lambda \to \infty$. This follows from equation (2.12), which implies that at large k the leading term in the spectral density is proportional to k^4. As described for instance in reference [27], it is possible to regularize this quantity while preserving diffeomorphism invariance by introducing a set of Pauli-Villars regulator fields. The contribution of these regulator fields and the counterterms lead to the renormalized energy density

$$\rho_{\mathrm{ren}} = \rho - \rho_{\mathrm{sub}}, \tag{2.33b}$$

where the subtraction terms are

$$\begin{aligned}\rho_{\mathrm{sub}} =& \frac{1}{2\pi^2}\left\{\frac{\Lambda^4}{8a^4} + \frac{m^2\Lambda^2}{8a^2} - \left[\delta\Lambda^f - \frac{m^4}{64}\left(1 - 2\log\frac{4\Lambda^2}{a^2\mu^2}\right)\right]\right.\\ &+ \frac{\Lambda^2\mathcal{H}^2}{8a^4} + \left[3(\delta M_P^2)^f - \frac{m^2}{6}\left(1 - \frac{3}{8}\log\frac{4\Lambda^2}{a^2\mu^2}\right)\right]\frac{\mathcal{H}^2}{a^2}\\ &\left. - \left[48\delta c^f - \frac{1}{8}\log\frac{4\Lambda^2}{a^2\mu^2}\right]\left(\frac{\mathcal{H}^2\ddot{a}}{a^5} + \frac{\ddot{a}^2}{4a^6} - \frac{\mathcal{H}\,\ddddot{a}}{2a^5}\right) - \frac{\mathcal{H}^4}{480a^4} - \frac{\mathcal{H}^2\ddot{a}}{30a^5} - \frac{19\ddot{a}^2}{480a^6} + \frac{19\mathcal{H}\,\ddddot{a}}{240a^5}\right\}.\end{aligned} \tag{2.33c}$$

Then, after the subtraction in equation (2.33b), as the cutoff Λ is sent to infinity the renormalized energy density remains finite and cutoff-independent by construction.

In expression (2.33c) μ is an arbitrary renormalization scale, and $\delta\Lambda^f$, $(\delta M_P^2)^f$ and δc^f are the finite pieces of the counterterms associated with a cosmological constant, the Einstein-Hilbert term and dimension four curvature invariants, respectively. Changes in the arbitrary renormalization scale μ effectively amount to changes in the finite values of the counterterms, which are determined by appropriate renormalization conditions. In that sense, observables that depend on the values of the counterterms are not predictions of the quantum theory.

The subtraction terms in (2.33c) arise from an adiabatic expansion of the vacuum energy of the regulators, and that of the original field in the ultraviolet. Note in particular that the first line contains no derivatives of the scale factor, the second line contains two, and the third line has four. On dimensional grounds, terms with a higher number of time derivatives vanish as the cutoff Λ is sent to infinity. Leaving the counterterms aside, it is in fact straight-forward (though somewhat tedious) to check that ρ_{sub} is just the integral of the adiabatic expansion to fourth order of the vacuum integrand in (2.33a),

$$
\begin{aligned}
\frac{d\rho}{d\log k} \approx \frac{1}{4\pi^2}\frac{k^3}{a^3}\Bigg\{ & \frac{\omega_k}{a} + \frac{\mathcal{H}^2}{2a^2}\left(\frac{a}{\omega_k}\right)^5\left[\frac{9m^4}{4} + \frac{3m^2k^2}{a^2} + \frac{k^4}{a^4}\right] \\
& + \frac{1}{2}\left(\frac{a}{\omega_k}\right)^{11}\Bigg[m^8\left(-\frac{189\mathcal{H}^4}{64a^4} + \frac{45\mathcal{H}^2\ddot{a}}{8a^5} + \frac{9\ddot{a}^2}{16a^6} - \frac{9\mathcal{H}\,\dddot{a}}{8a^5}\right) \\
& + \frac{m^6k^2}{a^2}\left(\frac{9\mathcal{H}^4}{8a^4} + \frac{111\mathcal{H}^2\ddot{a}}{8a^5} + \frac{15\ddot{a}^2}{8a^6} - \frac{15\mathcal{H}\,\dddot{a}}{4a^5}\right) \\
& + \frac{m^4k^4}{a^4}\left(\frac{43\mathcal{H}^4}{16a^4} + \frac{51\mathcal{H}^2\ddot{a}}{4a^5} + \frac{37\ddot{a}^2}{16a^6} - \frac{37\mathcal{H}\,\dddot{a}}{8a^5}\right) + \frac{m^2k^6}{a^6}\left(\frac{\mathcal{H}^4}{4a^4} + \frac{11\mathcal{H}^2\ddot{a}}{2a^5} + \frac{5\ddot{a}^2}{4a^6} - \frac{5\mathcal{H}\,\dddot{a}}{2a^5}\right) \\
& + \frac{k^8}{a^8}\left(\frac{\mathcal{H}^2\ddot{a}}{a^5} + \frac{\ddot{a}^2}{4a^6} - \frac{\mathcal{H}\,\dddot{a}}{2a^5}\right)\Bigg]\Bigg\}.
\end{aligned} \tag{2.34}
$$

Therefore, our regularization scheme reproduces and justifies the often employed adiabatic scheme [28], but it goes beyond it because it makes the role of the counterterms explicit and it also explains the origin of the subtraction terms.

Yet, from the perspective of Pauli-Villars regularization, the subtraction of adiabatic approximations to the spectral density is not fully justified [19]. In Pauli-Villars the masses of the regulators are assumed to be much larger than any accessible scale k, so their contribution to the *spectral* energy density at long distances is highly suppressed. For this reason, we shall not distinguish between the unrenormalized and renormalized spectral densities, as long as cosmological scales k are concerned. The regularization and renormalization afforded by the Pauli-Villars regulators is only of consequence in the ultraviolet, and only there does it play a role. Hence, we shall subtract equation (2.33c) from the energy density only when the mode integral includes the contributions of the ultraviolet.

Differences between Pauli-Villars and the adiabatic scheme renormalization are further underscored by an unphysical infrared singularity that appears in equation (2.34) in the massless limit. In that case, the adiabatic scheme leads to a renormalized energy density that diverges in the infrared, even when the unrenormalized spectral density and its renormalized Pauli-Villars counterpart are perfectly well-behaved there. In that limit, the adiabatic scheme fails.

Renormalizability places constraints on the physically allowed states of the field. Since the energy density is rendered finite by the fixed contributions of the subtraction terms, the occupation numbers of the high-momentum modes need to decay sufficiently fast. In particular, in order for the *renormalized* energy density to remain finite, n_k^{in} needs to decay faster than $1/k^4$. Thermal states with $n_k^{\text{in}} \propto \exp(-\omega_k/T)$ are thus physically allowed, while states with $n_k^{\text{in}} \propto 1/k^4$ are not. We are assuming that the occupation numbers only depend on the magnitude of $\vec{k}$ because of isotropy.

2.3.3 Particle Production

Equations (2.33) suffice to compute the energy density of the scalar ϕ in the *in* vacuum at any time in cosmic history. All one needs is an *in* region to single out the appropriate state of the field. This fixes the initial conditions for the mode functions χ_k^{in} in the asymptotic past, and equation (2.4) then determines its evolution all the way to the asymptotic future. However, the use of $\chi_k = \chi_k^{\text{in}}$ in equation (2.33a) is only a possible choice, and the same energy density can be expressed in any basis of mode functions.

Spectral Density. In order to obtain the spectral density of the *in* vacuum in terms of the arbitrary mode functions χ_k, it suffices to plug equation (2.14) into equation (2.33a). Clearly, by construction, the end result does not depend on the nature of the chosen mode functions χ_k, as long the state of the field remains unaltered. Carrying out the substitution, we thus find

$$\frac{d\rho}{d\log k} = \frac{k^3}{4\pi^2 a^4}\left\{\left(|\beta_k|^2 + \frac{1}{2}\right)\left[a^2\left|\left(\frac{\chi_k}{a}\right)^{\cdot}\right|^2 + \omega_k^2\,|\chi_k|^2\right] + \alpha_k\beta_k^*\left[a^2\left(\frac{\chi_k}{a}\right)^{\cdot 2} + \omega_k^2\,\chi_k^2\right] + \text{c.c.}\right\}, \tag{2.35}$$

where we have used that $|\alpha_k|^2 - |\beta_k|^2 = 1$. Comparing equations (2.27c) and (2.35) reveals that the *in* vacuum appears to effectively contain $n_k = |\beta_k|^2$ *out* particles that are not in an eigenvector of the *out* number operator, $m_k = \alpha_k\beta_k^*$. This formal similarity is behind what is referred to as particle production. In this approach, one would associate the spectral energy density

$$\frac{d\rho_{\text{p}}}{d\log k} \equiv \frac{k^3}{4\pi^2 a^4}\left\{|\beta_k|^2\left[a^2\left|\left(\frac{\chi_k}{a}\right)^{\cdot}\right|^2 + \omega_k^2\,|\chi_k|^2\right] + \alpha_k\beta_k^*\left[a^2\left(\frac{\chi_k}{a}\right)^{\cdot 2} + \omega_k^2\,\chi_k^2\right] + \text{c.c.}\right\} \tag{2.36}$$

to the "produced" particles. As it stands, equation (2.36) is just an approximation to (2.35) when the number of produced *out* particles in a given mode is large, $|\beta_k|^2 \gg 1/2$. Alternatively, one could regard equation (2.36) as the spectral density of the field (2.35) from which the spectral density of the *out* vacuum has been subtracted (recall that the latter is obtained by replacing $\chi_k^{\text{in}} \to \chi_k$ in equation (2.33a))

$$\frac{d\rho_{\text{out}}}{d\log k} \equiv \frac{k^3}{4\pi^2 a^4}\left[a^2\left|\left(\frac{\chi_k}{a}\right)^{\cdot}\right|^2 + \omega_k^2|\chi_k|^2\right]. \tag{2.37}$$

As a matter of fact, however, the *out* vacuum plays no role in our analysis, first because it depends on the choice of mode functions χ_k, and second because we assume that the field is in the *in* vacuum. In any case, since we are interested in the gravitational effects of the field, there is no physical basis for the removal of the vacuum energy density. In the adiabatic scheme, renormalization amounts to the subtraction of the adiabatic approximations to the spectral density in equation (2.34), but these subtractions correspond to the removal of the vacuum contribution only in the adiabatic regime, and only up to factors of sixth or higher adiabatic order.

Yet equation (2.36) is not what is usually associated with the particle production formalism. Rather, see for instance [13], the spectral density is often approximated by

$$\frac{d\rho_{\text{p}}}{d\log k} \approx \frac{k^3}{2\pi^2 a^4}|\beta_k|^2\left[a^2\left|\left(\frac{\chi_k}{a}\right)^{\cdot}\right|^2 + \omega_k^2\,|\chi_k|^2\right], \tag{2.38}$$

which, on top of neglecting the energy of the *out* vacuum, assumes that the *in* vacuum is an eigenvector of the *out* number operator, with eigenvalue $|\beta_k|^2$. Since in the limit of large $|\beta_k|$, the combination $|\alpha_k \beta_k^*|$ is of the same order as $|\beta_k|^2$, equation (2.38) does not immediately follow from (2.36). To explore the potential applicability of equation (2.38) it is useful to consider the spectral density when the corresponding modes are in the adiabatic regime. This does not generically hold, and applies, for instance, for massive fields at late times or sufficiently large wavenumbers. In that case, up to an arbitrary phase, the spectral density (2.36) reduces to

$$\frac{d\rho_{\rm p}}{d\log k} \approx \frac{k^3}{2\pi^2 a^4}\Bigg\{|\beta_k|^2\left[\frac{W_k}{2}+\frac{\omega_k^2+\mathcal{H}^2}{2W_k}+\frac{\mathcal{H}\dot{W}_k}{2W_k^2}+\frac{\dot{W}_k^2}{8W_k^3}\right] \\ +\frac{|\alpha_k\beta_k|}{2}\left[\left(\frac{\omega_k^2+\mathcal{H}^2}{W_k}-W_k+\frac{\mathcal{H}\dot{W}_k}{W_k^2}+\frac{\dot{W}_k^2}{4W_k^3}\right)\cos\left(2\int^\eta W_k\, d\tilde{\eta}\right)+\left(2\mathcal{H}+\frac{\dot{W}_k}{W_k}\right)\sin\left(2\int^\eta W_k\, d\tilde{\eta}\right)\right]\Bigg\}. \tag{2.39}$$

In this light, the approximation of the spectral density by (2.38) does receive some support when frequencies are large, $\omega_k \gg \mathcal{H}$. Then, the terms on the second line of equation (2.39) are doubly suppressed: First, because as opposed to those on the first line proportional to $W_k \approx \omega_k$, they are proportional to $\mathcal{H}$, since the terms of order ω_k cancel, and second, because they rapidly oscillate with time. In fact, on cosmological timescales of order $\mathcal{H}$, the evolution of the scale factor is only sensitive to the time average of the energy density, which is strongly suppressed when $W_k \gg \mathcal{H}$. The strength of the suppression depends on the particular details of the time average, and we shall simply assume for the time being that the average is such that the terms on the second line remain subdominant. In that case, the spectral density is well approximated by

$$\frac{d\rho_{\rm p}}{d\log k} \approx \frac{k^3}{2\pi^2 a^4}|\beta_k|^2\,\omega_k, \tag{2.40}$$

which possesses a clear interpretation in terms of particles, once we identify $|\beta_k|^2$ with the number of created *out* particles in the mode k. Since equations (2.38) and (2.40) are valid under the same conditions, but the latter is simpler and more intuitive, we are referring to (2.40) whenever we invoke the "particle production formalism." For massless particles the dispersion relation is $\omega_k = k$, and the spectral density (2.40) scales like radiation. For massive particles $\omega_k \approx ma$ at late times, and the spectral density would scale like dust. Although the meaning of the Bogolubov coefficients α_k and β_k is tied in general to the arbitrary choice of mode functions χ_k in equation (2.14), in order to arrive at (2.40) we have employed the adiabatic approximation (2.6). Hence, the β_k in equation (2.40) are uniquely determined by that choice of mode functions. Since (2.40) neglects terms with one derivative, it is inconsequential to calculate β_k beyond the zeroth order adiabatic approximation.

It is also worth pointing out that equation (2.40) fails at small frequencies even when the modes themselves are in the adiabatic regime, because to justify it we need to assume that $\mathcal{H} \ll \omega_k$ (see for instance the second term in the first line of equation

Eq.	$\lvert\beta_k\rvert \gg 1$	${}^{(n)}W_k \gg {}^{(n+2)}W_k$	$\omega_k \gg \mathcal{H}$
(2.36)	✓	✗	✗
(2.38)	✓	✓	✓
(2.39)	✓	✓	✗
(2.40)	✓	✓	✓

Table 1. Conditions under which the different particle production formulae are valid approximations to the actual spectral density of a scalar field in the *in* vacuum (2.35). A checkmark and a cross denote whether the corresponding condition is necessary or not, respectively. Note that $|\beta_k| \gg 1$ when particle production is effective, ${}^{(n)}W_k \gg {}^{(n+2)}W_k$ when modes are adiabatic, and $\omega_k \gg \mathcal{H}$ when frequencies are large. For minimally-coupled scalar fields, the condition $\omega_k \gg \mathcal{H}$ generically implies ${}^{(n)}W_k \gg {}^{(n+2)}W_k$.

(2.39).) Though this condition usually amounts to the validity of the adiabatic regime, there are cases in which modes are adiabatic even when their frequencies are small; see the appendix for details. Conversely, since the validity of the adiabatic approximation demands that ${}^{(n)}W_k \gg {}^{(n+2)}W_k$ for all n, it is conceivable for one of these conditions to be violated even when frequencies are large.

In conclusion, the particle production formula (2.40) is well-justified provided that

i) particle production is effective ($|\beta_k| \gg 1$),

ii) the relevant modes are in the adiabatic regime (${}^{(n)}W_k \gg {}^{(n+2)}W_k$),

iii) the mode frequencies are large ($\omega_k \gg \mathcal{H}$).

Even then one should recognize that the approximation (2.40) does not extend beyond the leading adiabatic order, since the terms that are neglected on the second line of equation (2.39) are of first order. Table 1 lists the conditions under which the various equations in this subsection are valid approximations to the scalar field energy density.

The adiabatic limit of the *out* mode functions in the ultraviolet also allows us to determine under what conditions the renormalized energy density after the transition remains finite. By construction, the terms in equation (2.35) that survive when β_k is set to zero give rise to a ultraviolet divergent integral that is regulated and renormalized by the subtraction terms in (2.33c). Therefore, the remaining terms must yield a finite contribution to the energy density. To estimate their behavior in the ultraviolet, we substitute the leading approximation $W_k \approx \omega_k \approx k$ into equation (2.39). At leading order we obtain

$$\frac{d\rho_{\mathrm{p}}^{\mathrm{UV}}}{d\log k} \approx \frac{1}{2\pi^2 a^4}\left(|\beta_k|^2 k^4 + \mathcal{H}|\alpha_k\beta_k| k^3 \sin(2k\eta+\varphi) + \cdots\right), \tag{2.41}$$

which implies that $|\beta_k|$ has to decay faster than $1/k^2$ in order for the integral to remain finite, since $\alpha_k \approx 1$ in the ultraviolet. This is the same as the restriction on the number of *in* particles n_k^{in} in the ultraviolet that we discussed previously. Looking back at equations (2.26) it means that the second derivative of the scale factor has to be continuous at the transition. Otherwise, it is not just that the energy density diverges; the structure of the divergencies is incompatible with our regularization scheme. With $|\beta_k| \sim 1/k^3$, the spectral density (2.40) is ultraviolet finite.

Total Energy Density. Even though the spectral density is particularly convenient to study the contributions of the different modes to the total energy density, it is not an actual observable. The gravitational equations are sourced by the total energy density, which is given by the integral of the spectral density. It follows from equations (2.33), (2.35), (2.36) and (2.37) that

$$\rho_{\text{ren}} = \rho_{\text{p}} + \rho_{\text{ren}}^{\text{out}}, \quad \text{where} \quad \rho_{\text{p}} \equiv \int_{\Lambda_{\text{IR}}}^{\infty} \frac{dk}{k} \frac{d\rho_{\text{p}}}{d\log k} \quad \text{and} \quad \rho_{\text{ren}}^{\text{out}} \equiv \rho_{\text{out}} - \rho_{\text{sub}}. \tag{2.42}$$

It is important to stress that the spectral density that enters the density of the produced particles ρ_{p} here is the one in (2.36), since only then is the equation $\rho_{\text{ren}} = \rho_{\text{p}} + \rho_{\text{ren}}^{\text{out}}$ exact. The energy density $\rho_{\text{ren}}^{\text{out}}$ can be interpreted as the renormalized energy density of the *out* vacuum, which again, remains arbitrary at this point. Since $|\beta_k|$ has to decay faster than $1/k^2$, the energy density of the produced particles ρ_p is ultraviolet finite, and we can directly set the cutoff Λ to infinity. On the other hand, both ρ_{out} and ρ_{sub} are ultraviolet divergent, and only their difference remains finite as $\Lambda \to \infty$.

When conditions *ii*) and *iii*) above are satisfied for those modes k that dominate the integral in (2.42), it is possible to interpret these equations under a new light, since adiabaticity singles out a preferred set of *out* mode functions. In that case the spectral density $d\rho_{\text{p}}/d\log k$ that enters (2.42) can be approximated by equation (2.39). Provided in addition that the corresponding β_k are large enough for the terms of higher adiabatic order in (2.39) to be negligible, we can finally approximate the particle spectral density of those modes just by (2.40). In some cases, the resulting energy density ρ_{p} may be negligible in comparison with the renormalized energy density of the *out* vacuum $\rho_{\text{ren}}^{\text{out}}$. If, on the other hand, $\rho_{\text{p}} \gg \rho_{\text{ren}}^{\text{out}}$, we can approximate $\rho_{\text{ren}} \approx \rho_{\text{p}}$. Although not necessarily so, this is expected to occur when the adiabatic approximation is violated for a sufficiently large period of time in the early universe, leading to large values of the coefficients β_k and hence large values of the integral ρ_{p}. Only under those circumstances it is then justified to write

$$\rho_{\text{ren}} \approx \int_{\Lambda_{\text{IR}}}^{\infty} \frac{dk}{k} \frac{k^3}{2\pi^2 a^4} |\beta_k|^2 \, \omega_k, \tag{2.43}$$

which can be interpreted as the integral over phase space of the distribution function $f(k) = |\beta_k|^2$ associated to an isotropic classical ensemble of particles of energy ω_k. Equation (2.43), usually with $\Lambda_{\text{IR}} = 0$, is hence the blanket "particle production" equation often used in the literature (see, for instance, [4, 42].) Alas, since one or several of the conditions stated above typically fails, (2.43) is often *not* a valid approximation, as we shall see below.

At this point it becomes clear that in most cases the particle production formalism is just an approximation at best. As far as the spectral density is concerned, equation (2.35) remains true no matter whether the notion of particle exists, and regardless of how the mode functions χ_k are chosen. Furthermore, if we knew the form of the *in* mode functions throughout cosmic history, $\chi_k^{\rm in}$, there would be no need to go though the process of introducing Bogolubov coefficients and evaluating (2.35) or its approximations, (2.36) to (2.40); it would just suffice to evaluate equations (2.33) at any desired time. The particle production formalism is useful at high frequencies, where the existence of an adiabatic regime enable us to single out an adiabatic vacuum, and the high frequencies and momenta allow us to interpret its excitations as actual particles. In that sense we can think of $\rho_{\rm p}$ as a renormalized energy density, from which the vacuum contribution has been subtracted. As opposed to ρ, which is divergent by itself, $\rho_{\rm p}$ hence remains finite in the ultraviolet.

Let us conclude by emphasizing that, although we have discussed the energy density of the scalar within the specific class of cosmic transitions discussed in subsection 2.2, many of results in subsection 2.3 are applicable to a much wider class of scenarios almost without modification. All that is essentially needed is for a subset of the scalar field modes to be in a preferred state $|0_{\rm in}\rangle$ such that $a_{\vec{k}}^{\rm in}|0_{\rm in}\rangle = 0$.

3 Massless Fields

We shall begin our illustration of the previous results with a massless field ϕ, which is easier to treat analytically. This is also relevant for the production of massless particles at the end of inflation and also facilitates our analysis of light fields later on. The vacuum energy is determined by the *in* mode functions $\chi_k^{\rm in}$ through equations (2.33). Qualitatively, the evolution of the modes throughout cosmic history is relatively simple when $m = 0$: In the asymptotic past the modes find themselves in the short wavelength regime, where they oscillate with positive frequency and constant known amplitude, as they would do in Minkowski space. Some of these modes are eventually pushed by inflation to superhorizon scales, where they stop oscillating and grow in proportion to the scale factor until a fraction reenters after the end of inflation. Once they do they oscillate again with an enhanced amplitude, but this time with positive and negative frequencies. These properties alone suffice to determine the shape of the spectral density, which we analyze below. In particular, the evolution of those modes that left during inflation and later reentered the horizon is behind the phenomenon of cosmological particle production. Superhorizon modes, on the other hand, do not admit a particle interpretation, as we clarify next. We show the behavior of the mode functions in the different regimes in figure 4. A summary of the results obtained in this section is presented in subsection 3.5.

3.1 Inflation

An important advantage of power-law inflation is that the mode functions that satisfy condition (2.12) are explicitly known in the massless case,

$$\chi_k^{\rm in}(\eta) = \frac{\sqrt{-\pi\eta}}{2} H_\nu^{(1)}(-k\eta), \quad \nu \equiv \frac{1-2p}{2}. \tag{3.1}$$

In de Sitter, these mode functions, which correspond to the Bunch-Davies vacuum [29], simplify considerably, so we consider this limit first. Substituting (3.1) into (2.33a) it is easy to see that when $p = -1$ the energy density is

$$\rho_{\rm dS} = \frac{1}{2\pi^2}\frac{1}{a^4}\left(\frac{\Lambda^4}{8} + \frac{\Lambda^2\mathcal{H}^2}{8}\right), \tag{3.2}$$

which precisely matches the cutoff-dependent subtraction terms in equation (2.33c) in the massless case, $m = 0$. This is of course no coincidence, as the subtraction terms are supposed to cancel the cutoff dependence by construction. Note, in addition, that the energy density is infrared finite. Subtracting equation (2.33c) from the unrenormalized energy density (3.2) returns the renormalized energy density

$$2\pi^2\rho^{\rm dS}_{\rm ren} \approx \delta\Lambda^f - 3H^2(\delta M_P^2)^f - \frac{119H^4}{480}. \tag{3.3}$$

Up to the counterterms, this is the result derived by Allen and Folacci [30]. Yet in de Sitter space the energy density is degenerate with the counterterms associated with the cosmological constant and the Planck mass, so it is not really an observable quantity.

At any rate, as far as inflation is concerned, de Sitter spacetime is not truly appropriate, particularly because in de Sitter there are no metric perturbations. Away from de Sitter there are no exact analytical expressions for the energy density. In order to estimate the latter, we split the integration range into the "infrared" ($k \leq \mathcal{H}$) and the "ultraviolet" ($k \geq \mathcal{H}$) and use small and large momentum expansions to determine the contribution of each domain to the integral.[2] In the ultraviolet, up to terms of fourth adiabatic order, the spectral density takes the form

$$\frac{d\rho_{\rm UV}}{d\log k} \approx \frac{k^4}{4\pi^2 a^4}\left[1 + \frac{1}{2}\left(\frac{aH}{k}\right)^2 + \frac{3}{8}\frac{p^2-1}{p^2}\left(\frac{aH}{k}\right)^4\right]. \tag{3.4}$$

As we have previously argued, equation (3.4) is only applicable up to wavenumbers k of the order of the regulators' masses, but those scales are not accessible to cosmological observations. At higher momenta the contribution of (3.4) to the energy density appears to diverge, but it is rendered finite by the subtraction terms (2.33c), which can be regarded as (minus) the energy density of the regulator fields. Integrating (3.4) over the ultraviolet modes and subtracting (2.33c) we then arrive at the renormalized energy density

$$\begin{aligned} 2\pi^2\rho^{\rm UV}_{\rm ren} \approx{}& \delta\Lambda^f - 3H^2(\delta M_P^2)^f + \frac{36(p^2-1)H^4}{p^2}\left[\delta c^f + \frac{1}{384}\log\frac{\mu^2}{4H^2}\right] \\ &-\frac{122p^2-60p+57}{480p^2}H^4. \end{aligned} \tag{3.5}$$

[2] When we refer to the ultraviolet we shall sometimes have the upper boundary at $k = \Lambda$ in mind, while others we may be simply referring to subhorizon modes.

We should point out that the error in our simple-minded approximations to the mode integrals is expected to be of order H^4, mostly from the lower boundary of the ultraviolet at $k \sim \mathcal{H}$, where the adiabatic approximation breaks down. As a consequence, equation (3.5) is just an order of magnitude estimate of the actual energy density, though it does imply that the renormalized coupling constant of the dimension four curvature invariants effectively runs logarithmically with time, as we could have guessed directly from equations (2.33c). A question we shall not address here is whether perturbation theory is still reliable when the logarithm is large, that is, when H is far from the renormalization scale μ.

In the infrared limit, the spectral density (2.33a) scales like

$$\frac{d\rho_{\text{IR}}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu} H^4 \left(\frac{k}{aH}\right)^{5-2\nu}, \tag{3.6}$$

since the leading terms in the infrared expansion cancel in the massless case (more on this below.) Therefore, the energy density is infrared divergent for $\nu \geq 5/2$ ($p \leq -2$) and infrared finite otherwise. Indeed, as we move away from $p = -1$ the long wavelength scale-invariant spectrum of field fluctuations becomes redder and redder, until an infrared divergence appears in the spectral density at equations of state $-2/3 < w < -1/3$ (i.e. $p < -2$). This range, however, is phenomenologically less relevant, because the spectral index of the primordial perturbations suggests that p is close to -1 (in the simplest inflationary models.) Infrared divergences are typical of massless theories, although we shall see that in some cases they persist when fields are sufficiently light, even close to de Sitter. In section 2.3 we pointed out that only modes with $k > \Lambda_{\text{IR}}$ can be assumed to be in the *in* vacuum. Although in some cases the infrared cutoff Λ_{IR} plays no role in the final expressions, in others it is the responsible of regulating the infrared divergencies.

Performing the integral of the spectral density (3.6) over the infrared, $\Lambda_{\text{IR}} \leq k \leq \mathcal{H}$, we obtain

$$\rho_{\text{IR}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{5-2\nu} H^4 \left[1 - \left(\frac{\Lambda_{\text{IR}}}{aH}\right)^{5-2\nu}\right]. \tag{3.7}$$

The cutoff has no practical impact on the energy density when $\nu > 5/2$ ($p > -2$), since in that case the energy density is dominated by the shorter modes (first term in the square brackets), but when $p < -2$ it is dominated by the infrared contribution (second term in square brackets.) In the de Sitter limit, the energy density (3.7) soon approaches the well-known result $\rho_{\text{IR}} = H^4/(16\pi^2)$.

The energy density (3.7) is positive. If $-2 < p < -1$ it grows from zero at the beginning of inflation and soon decays like H^4. On the other hand, if $p < -2$ it soon reaches a value of order H^4 shortly after the onset of inflation and subsequently decays in proportion to $1/a^2$. In either case the vacuum energy decays faster than the energy density of the background universe, so unless inflation begins at trans-Planckian energy densities it always remains subdominant.

Figure 2 shows the spectral density during inflation for a few values of p. The two distinct asymptotic regimes in the infrared and ultraviolet are clearly visible in the

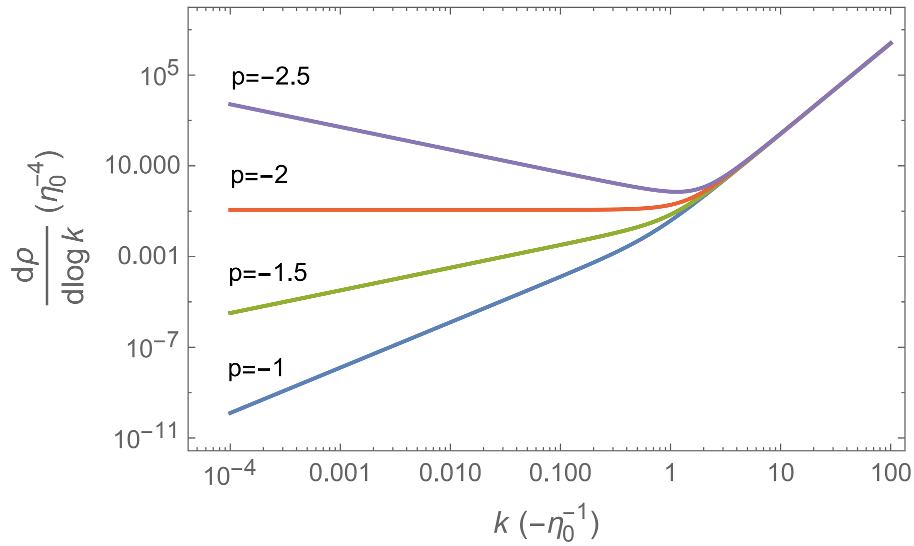

Figure 2. Spectral density during inflation for different values of the parameter p in equation (2.13). The time η_0 is arbitrary. The power-law behavior in the infrared is well approximated by equation (3.6), and that in the ultraviolet by equation (3.4). At $p \leq -2$ the energy density is infrared divergent.

spectral density, as well as the dependence of the infrared spectral index on the value of p. In figure 3 we also show the numerically determined renormalized energy density for various values of $p > -2$ and renormalization scale μ. The numerical results are in good agreement with our estimate of the total renormalized energy density, namely, the sum of ultraviolet and infrared contributions (3.5) and (3.7). In particular, the energy density remains of order H^4 modulo a running that originates from the decreasing ratio H/μ. As long as μ is not too far from H, the renormalized energy density near de Sitter inflation is close to the Allen-Folacci result (3.3), but when $-2 < p < -1$ we expect the energy density to eventually become positive, though this involves a large logarithmic correction.

The behavior of the energy density of a massless scalar field during inflation is summarized in table 2. Although we have not introduced any cosmic transition yet, notice that none of the results that we have obtained in this section can be recovered within the particle production formalism. In the ultraviolet modes are adiabatic, but the corresponding Bogolubov coefficients β_k are small, since these modes never leave the adiabatic regime. In the infrared the adiabatic mode functions are not approximate solutions, and the mode frequencies ω_k are in any case small. In either regime, at least one of the three conditions needed to justify (2.40) cannot be satisfied.

3.2 General Infrared Evolution After Inflation

Just as the adiabatic expansion allows us to analyze the ultraviolet regime on general grounds, we can describe the contribution from the infrared modes to the energy density after inflation in a model-independent way. The key here is that inflation directly

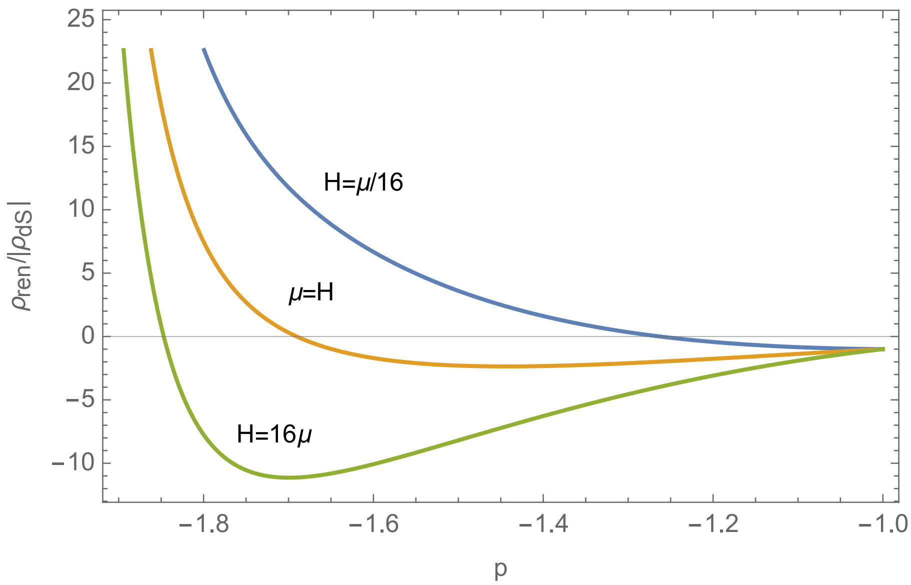

Figure 3. Renormalized energy density in the range $-2 < p \leq -1$ for different ratios of the renormalization scale μ to the Hubble constant H. The energy density is quoted in units of the magnitude of the Allen-Folacci value (3.3) at the corresponding time during inflation. We can think of this figure as one in which time has been fixed and μ is different in each curve, or one in which μ has been fixed and each curve corresponds to a different moment of time. In the latter case, time "proceeds" along lines of constant p upwards. Note that we have set the finite pieces of the counterterms to zero. The apparent infinite growth of the energy density as $p = -2$ is approached is controlled by the infrared cutoff $\Lambda_{\rm IR}$, that we have set to zero in this plot.

determines the properties of the superhorizon modes, whose evolution is insensitive to the specific details of the subsequent cosmic expansion and, in particular, to the transtion to radiation domination, in a sense that we make explicit below.

We shall study the infrared by inspecting the behavior of the mode functions around $\vec{k} = 0$, which, as we argued in subsection 2.1.2, is a good approximation as long as $\omega_k \ll \mathcal{H}$. Using the mode function (2.9) and its complex conjugate as basis, and the small argument expansion of (3.1), we can determine the Bogolubov coefficients in the limit where $-k\eta \ll 1$ by evaluating (2.16) at a time when both expressions are valid, e.g. at an arbitrary time η_0 during inflation when the mode is larger than the horizon,

$$\beta_k \approx \frac{\Gamma(\nu)}{\sqrt{2k\pi}} \left(-\frac{k\eta_0}{2}\right)^{1/2-\nu} \frac{1}{a_0} \left[c_2 - \frac{1-2p-2\nu}{2p} a_0^2 \mathcal{H}_0 \left(c_1 + c_2\, b_0\right)\right], \quad \alpha_k \approx -\beta_k^*. \tag{3.8}$$

As long as the scale factor and its first derivative are continuous at the transition, the evolution of the *in* mode functions on superhorizon scales does not depend on the nature of the transition. To see this, note that the mode function of the long wavelength modes during inflation (3.1) is proportional to the scale factor, and because its value

	power-law inflation		radiation domination	
$-2 < p \leq -1$	$\rho_{\rm ren} \propto H^4$	(IR&UV)	$\rho_{\rm ren} \propto a^{-2(3+p)}$	(IR&IM)
$p \leq -2$	$\rho_{\rm ren} \propto a^{-2}$	(IR)	$\rho_{\rm ren} \propto a^{-2(3+p)}$	(IM)

Table 2. Evolution of the vacuum energy of a massless scalar field during inflation and radiation domination. In parenthesis we indicate the modes that mostly contribute to the value of $\rho_{\rm ren}$. When $p \approx -1$ the ultraviolet modes can dominate the energy density if the transition-dependent factor $\dot{a}_r/\dot{a}_0$ is less than one, see equation (3.28). In some cases the density contains an additional logarithmic running that we skip for simplicity.

and first time derivative are continuous, after the transition it must be still proportional to a, which remains a particular solution of the mode equation when $\omega_k = 0$, no matter how the universe expands. This is exactly what follows from equations (3.8), which when combined with (2.14) and (2.9) yield a mode function that in the infrared equals

$$\chi_k^{\rm in} \approx -i\frac{\Gamma(\nu)}{\sqrt{2k\pi}}\left(-\frac{k\eta_0}{2}\right)^{1/2-\nu}\frac{a}{a_0}\left[1+\frac{1-2p-2\nu}{2p}a_0^2\mathcal{H}_0\left(b-b_0\right)\right]. \tag{3.9}$$

Since equation (2.9) remains a good approximation as long as $\omega_k = k \ll \mathcal{H}$, this expression is applicable for long wavelength modes at any time in cosmic history. For later purposes, we have assumed here that p and ν are unrelated, though if one imposes (3.1) the term proportional to b vanishes, as claimed. An alternative way to state the same is to say that $\chi_k^{\rm in}/a$ is conserved (i.e. the Fourier mode of the actual field ϕ is frozen) on superhorizon scales, no matter how the scale factor evolves. Observe that in order to arrive at (3.9) we only used equation (3.1) and $-k\eta_0 \ll 1$. In this context, then, η_0 may refer to any time during inflation at which the relevant mode is superhorizon-sized. In particular, when equation (2.13) is substituted into (3.9) the dependence on η_0 drops out. Note that in the approximation of equation (3.8) $|\alpha_k|^2 - |\beta_k|^2 = 0 \neq 1$, and we would need to include the subleading terms in equation (3.8) to obtain the correct difference. Since the Bogolubov coefficients (3.8) are large, the error in the difference is relatively small.

Since at lowest order in the small k expansion $\chi_k^{\rm in}/a$ is constant, its time derivative must be of order k^2, and its square proportional to k^4. Therefore, the leading contribution in the infrared stems from the non-derivative term in equation (2.33a),

$$\frac{d\rho_{\rm IR}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu}H_0^4\left(\frac{a_0}{a}\right)^2\left(\frac{k}{a_0H_0}\right)^{5-2\nu}, \tag{3.10}$$

which during inflation happens to be the same as equation (3.6). This agreement underscores the transition independence of the spectral density in this mode range and also illustrates that equation (3.10) holds for any form of the scale factor. Again, the conditions that we identified in section 2.3.3 are not satisfied here, so we cannot rely on the particle production formula (2.40). Even though the modes may be in the adiabatic regime (see the next section) and the numbers β_k are large, we are considering superhorizon modes, and the last of the three conditions in subsection 2.3.3 fails.

The time evolution of the infrared contribution to the energy density depends on the properties of inflation and, in particular, on the value of p. Carrying out the momentum integral of equation (3.10) over the infrared modes we arrive at

$$\rho_{\mathrm{IR}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{5-2\nu}H_e^4\left(\frac{a}{a_e}\right)^{3-2\nu}\left(\frac{H}{H_e}\right)^{5-2\nu}\left[1-\left(\frac{\Lambda_{\mathrm{IR}}}{aH}\right)^{5-2\nu}\right], \tag{3.11}$$

where, as in all the subsequent integrated spectral densities, we have chosen the arbitrary time η_0 to be the end of inflation, η_e, since this is when all the relevant modes are beyond the horizon. If $p \approx -1$, this energy density tracks the background evolution. Given that ρ_{IR} is of order H_e^4 at the end of inflation, the contribution of this component to the background density will remain negligible through cosmic history. On the other hand, if $p < -2$, the infrared component scales like curvature, as we pointed out earlier.

Just as a description of the scalar energy density in terms of particles is expected to work in the subhorizon limit, one may expect a description of the energy density in terms of a homogeneous scalar to be appropriate in the infrared. However, it is easy to see that this is not possible for a massless field. As we argued in subsection 2.3.1, the energy density of a massless homogeneous scalar field is proportional to a^{-6}. This is in direct contradiction with the actual behavior of the energy density in (3.11), which approximately decays like H^2 (when $p \approx -1$.) The origin of the disagreement is that the superhorizon modes find themselves in the mode proportional to a, as seen in equation (3.9), so the contribution of these modes to the energy density does not stem from the terms with time derivatives in equation (2.33a), which are the ones that would lead to the $1/a^6$ scaling, but from those that contain the gradient of the field, which vanish when the field is homogeneous. The energy density of the latter does not scale like curvature, as we found in equation (2.31), because the upper limit of integration at $k = \mathcal{H}$ implicit in equation (3.11) depends on time. Then, when $-2 < p < -1$, the integral is dominated by the upper boundary, which introduces an additional time dependence in the energy density. However, when $p < -2$ the infrared modes at $k \sim \Lambda_{\mathrm{IR}}$ dominate, and the integral displays the same curvature scaling as in section 2.3.1.

3.3 Radiation Domination

Modes that left the horizon during inflation eventually reenter after its end. Because the superhorizon modes are not sensitive to the details of the transition between inflation and radiation domination, we can also analyze the spectral density during radiation domination in a relatively model-independent way. To do so we postulate that there must be a time $\eta_r > \eta_e$ after which we can safely assume the universe to be radiation dominated, see figure 1. Then, our analysis applies in the range $k \ll \mathcal{H}_r$ (modes that reenter or remain outside the horizon during radiation domination) because the behavior of these modes during the transition is universal. At the same time, if inflation ends at η_e, our analysis also applies in the range $\mathcal{H}_e \ll k$ (modes that were inside the horizon at the end of inflation) because these remain in the adiabatic regime

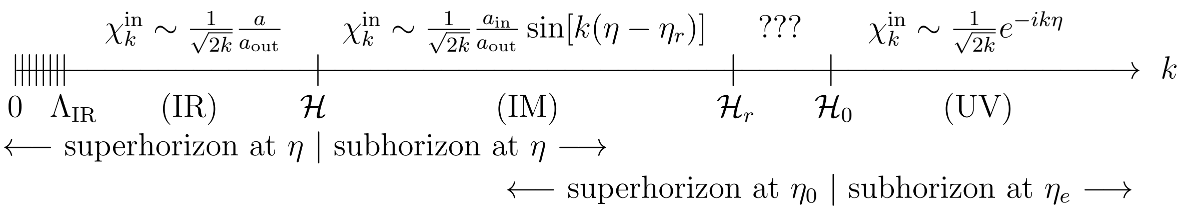

Figure 4. Modes as function of the scale at any time $\eta > \eta_r$ during radiation domination. UV (IR) modes are shorter (larger) than the horizon at the end of inflation η_e and also at η. On the contrary IM modes entered the horizon during radiation domination. The initial conditions for modes with $k < \Lambda_{\rm IR}$ are unknown, whereas those modes with $\mathcal{H}_r < k < \mathcal{H}_0$ depend on the details of reheating.

thoughout. What happens in the interval $\mathcal{H}_r \lesssim k \lesssim \mathcal{H}_e$ depends on the details of the transition, and is thus model dependent (note that $\mathcal{H}$ decreases after the end of inflation.) Nevertheless, this is expected to be a small window in comparison with the mode intervals for which we can make definite predictions.

In order to proceed, it shall prove to be convenient to divide the different modes into three distinct ranges. Following a similar convention as in subsection 3.1, we shall refer to the short wavelength regime $\mathcal{H}_e \ll k$ as the "ultraviolet." Modes that reenter before time η during radiation domination belong to the "intermediate" range, and those that remain outside the horizon by that time are part of the "infrared,"

$$\Lambda_{\rm IR} \leq k_{\rm IR} \ll \mathcal{H} \ll k_{\rm IM} \ll \mathcal{H}_r < \mathcal{H}_e \ll k_{\rm UV}. \tag{3.12}$$

These ranges, along with the behavior of the mode functions in each of them, are visually represented in figure 4. In the infrared, in particular, the spectal density is given by the "universal" equation (3.10).

Infrared and Intermediate Regimes. The prescribed evolution of the scale factor during radiation domination also fixes the mode functions of those modes that enter the horizon during that time. To see that, note that once the transition to radiation domination has been completed, the scale factor has to be of the form $a = a_r + \dot{a}_r(\eta - \eta_r)$, while equation (3.9) remains a solution of the mode equation at long wavelengths at the same time. The mode function $\chi_k^{\rm in}$ in that equation can be cast as the Bogolubov transformation in equation (2.14), with mode functions χ_k given by

$$\chi_k = \frac{e^{-ik(\eta-\eta_r)}}{\sqrt{2k}}, \tag{3.13}$$

which happen to solve the mode equation during radiation domination at any k. Also note that (3.13) is an adiabatic solution of the form (2.6), even though the field is massless and the frequency of the superhorizon modes is much smaller than the comoving Hubble constant, $\omega_k = k \ll \mathcal{H}$. This is a consequence of the conformal coupling of the scalar that appears in the massless case during radiation domination, when $R = 0$ (see appendix A.1.) Incidentally, an expansion of equation (3.13) to first order in $k\eta$ reveals that the mode functions are exactly of the form (2.9), since the latter is a linear combination of a constant and a term linear in η. This validates that (2.9) is indeed an approximate solution of the mode equation on superhorizon scales. Moreover, with $G_{\text{ret}} = \eta - \tilde{\eta}$ during the radiation era the correction (2.11) agrees with the next two terms in the series expansion of (3.13) at small $-k\eta$.

To express the *in* mode function (3.9) as a linear combination of (3.13) and its complex conjugate, we evaluate first the Bogolubov coefficients (2.16) at a time and mode range when both are valid solutions of the mode equation, say, at $\eta = \eta_r$ and $k \ll \mathcal{H}_r$,

$$\beta_k \approx -\frac{1}{2}\frac{\Gamma(\nu)}{\sqrt{\pi}}\left(-\frac{k\eta_0}{2}\right)^{1/2-\nu}\frac{a_r}{a_0}\left(\frac{\mathcal{H}_r}{k}+i\right), \quad \alpha_k \approx -\beta_k^*. \tag{3.14}$$

The phases of the Bogolubov coefficients obviously do not affect $|\alpha_k|^2$ and $|\beta_k|^2$ whatsoever, whereas common phases in α_k and β_k do not affect the terms proportional to $\alpha_k\beta_k^*$ in equation (2.35). Therefore, in order to simplify our formulas we shall often choose these common phases for convenience. Modes outside the horizon at η_r satisfy $\mathcal{H}_r/k \gg 1$, which fixes the dominant term inside the last parenthesis in equation (3.14), although it is necessary to keep the subominant one too to reproduce the correct spectral density in the infrared, equation (3.10).

Inserting the *out* mode functions (3.13) and the Bogolubov coefficients (3.14) into equation (2.14) we recover the *in* mode functions in the *out* region,

$$\chi_k^{\text{in}} \approx -i\frac{\Gamma(\nu)}{\sqrt{2k\pi}}\left(-\frac{k\eta_0}{2}\right)^{1/2-\nu}\frac{a_r}{a_0}\left[\frac{\mathcal{H}_r}{k}\sin\left[k(\eta-\eta_r)\right]+\cos\left[k(\eta-\eta_r)\right]\right], \tag{3.15}$$

which holds for modes in the infrared and intermediate regimes. This equation shows that superhorizon modes grow linearly with the scale factor, whereas subhorizon modes oscillate with positive and negative frequencies, with amplitudes determined by a_r and $\dot{a}_r$. Indeed, substitution of (3.15) into equation (2.33a) readily reproduces equation (3.10) in the infrared; then a_r drops out of the equation and a_0 can be taken to be the scale factor at an arbitrary time during inflation. In the intermediate regime, on the other hand, the spectral energy density becomes

$$\frac{d\rho_{\text{IM}}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu}\,H_0^2H^2\left(\frac{k}{a_0H_0}\right)^{3-2\nu}, \tag{3.16}$$

which is nearly scale-invariant if inflation is de Sitter-like. Yet this expression does not capture some of the subleading corrections that are relatively important for modes not too far inside the horizon. Going one order higher in the small wavelength expansion results in

$$\frac{d\Delta\rho_{\text{IM}}}{d\log k} \approx -\frac{d\rho_{\text{IM}}}{d\log k}\times\frac{\mathcal{H}}{k}\sin[2k(\eta-\eta_r)], \tag{3.17}$$

which we could have also recovered from the subleading term in equation (2.39) (up to the relative sign, due to the arbitrary phase in that equation.) Note that we have only used that the corresponding modes enter the horizon during radiation domination. The spectral density of modes that enter the horizon after the end of inflation but before the universe is fully radiation dominated ($\mathcal{H}_r \leq k \leq \mathcal{H}_0$) depends on the evolution during the transition. It is also worth stressing, as should be clear from the way in which they were derived, that η_0 and η_r in equations like (3.15) and (3.16) can be taken to be arbitrary times during inflation and radiation domination, respectively, as long as the mode is superhorizon-sized at both times.

We could have also used the particle production formalism to arrive at the spectral density (3.16), since in the intermediate regime the three conditions we quote in section 2.3.3 are satisfied. In particular, to arrive at (3.16) it suffices to substitute the Bogolubov coefficients (3.14) into the particle production formula (2.40). On the other hand, the subleading correction in the small wavelength expansion (3.17) goes beyond the particle production formalism.

Finally, in order to determine the total energy density of the intermediate modes, we simply need to integrate (3.16) over the corresponding range $\mathcal{H} \leq k \leq \mathcal{H}_r$. Ignoring the oscillatory corrections and specializing in the quoted limit we find

$$\rho_{\mathrm{IM}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu} H_e^4 \left(\frac{a}{a_e}\right)^{3-2\nu} \left(\frac{H}{H_e}\right)^{5-2\nu} \log\frac{a}{a_r}, \quad \text{when} \quad \log\frac{a}{a_r} \ll \frac{1}{2\nu-3}. \tag{3.18a}$$

Up to the logarithmic dependence, for a universe that underwent near de Sitter inflation, this scales like radiation and is of order H_e^4 when extrapolated to the end of inflation, as it also happens for the infrared modes. On the other hand, in the opposite limit the energy density approaches

$$\rho_{\mathrm{IM}} \approx -\frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{3-2\nu} H_e^4 \left(\frac{a}{a_e}\right)^{3-2\nu} \left(\frac{H}{H_e}\right)^{5-2\nu}, \quad \text{when} \quad \log\frac{a}{a_r} \gg \frac{1}{2\nu-3}. \tag{3.18b}$$

Remarkably, this contribution to the density decays slower than curvature when $p < -2$, and in fact effectively violates the weak energy condition when $p < -3$. It is well-known that quantum corrections can lead to violations of the standard energy conditions, see for instance [31] for an early reference.

Ultraviolet Regime. The ultraviolet regime encompasses subhorizon modes at the end of inflation. In order to estimate the spectral density in the ultraviolet we plug the Bogolubov transformation (2.14), with the *out* mode functions given by (3.13), into the general expression for the energy density in (2.33a) and obtain

$$\frac{d\rho_{\mathrm{UV}}}{d\log k} \approx \frac{k^4}{4\pi^2 a^4}\left[1 + \frac{1}{2}\left(\frac{aH}{k}\right)^2 + 2|\beta_k|^2 + 2|\alpha_k\beta_k|\left(\frac{aH}{k}\right)\sin\left[2k\eta + \varphi\right]\right]. \tag{3.19}$$

Because the mode functions (3.13) have the adiabatic form (2.6), these are just the dominant terms in equation (2.39), with $W_k = k$. The first two terms in (3.19) are characteristic of the *out* vacuum, whereas the details of the transition are encoded in the parameters α_k and β_k, the latter being highly suppressed at large k, as we anticipated in 2.2.3 and illustrate in more detail in the next subsection (these are the ones we considered in equation (2.41).) Thus, neglecting the contributions proportional to the Bogolubov coefficients, integrating over the ultraviolet modes and subtracting (2.33c) we arrive at

$$2\pi^2\rho_{\text{ren}}^{\text{UV}} \approx \delta\Lambda^f - 3H^2(\delta M_P^2)^f + \frac{H^4}{480} + \frac{H^2 H_e^2}{8}\left(\frac{a_e}{a}\right)^2 - \frac{H_e^4}{8}\left(\frac{a_e}{a}\right)^4, \tag{3.20}$$

which just corresponds to the renormalized energy density of the *out* vacuum in the corresponding mode range and, leaving the counterterms aside, quickly becomes negligible after the onset of radiation domination. Note that in a radiation-dominated universe, $p = 1$, the counterterm proportional to δc^f vanishes. Then, beyond the renormalization of the cosmological and Newton's constants, the term that dominates the energy density at late times and can be regarded as a prediction of the quantum theory is the last one. Unfortunately, though, as in the case of inflation, equation (3.20) should be regarded as a rough estimate of the energy density in the ultraviolet, with an error of order $H_e^4(a_e/a)^4$ that is comparable to our leading prediction.

3.4 Examples

As an illustration of our general analysis, we shall proceed to study the spectral density after the three kinds of transitions described in subsection 2.2.3. As we have previously argued, the energy density in the infrared and the intermediate regimes is not sensitive to the details of the transition, up to an overall transition-dependent factor that we identify below. Yet the details of the transition are particularly important in the ultraviolet, where the spectral density is sensitive to the jumps in the derivatives of the scale factor characteristic of the idealized models. For realistic transitions, however, this model dependence disappears, as we show explicitly below.

Sharp Transition. With the scale factor given by (2.18), the appropriately normalized positive-frequency solution of the mode equation (2.4) at $\eta \geq \eta_e$ is (3.13), where we choose $\eta_r = \eta_e$. The actual solution during radiation domination can be cast as a linear combination of the form (2.14), with the coefficients determined by (2.16). In the latter the *in* mode functions are those of equation (3.1) and the time η is $\eta = \eta_e$. We shall not write down the values of α_k and β_k explicitly, but merely study their asymptotic behavior.

The Bogolubov coefficients are dimensionless, so they depend on k and η_e only through the combination $k\eta_e$. Both in the infrared and intermediate regimes the wavenumbers obey $-k\eta_e \ll 1$. In such a limit, the Bogolubov coefficients can be readily seen to match the general result in equation (3.14), with $\eta_0 = \eta_r = \eta_e$, as pertains to a sharp transition. In the infrared, by combining equations (2.14), (3.13) and (3.14) we obtain precisely the mode function in equation (3.9), with the scale factor given by equation (2.18), as expected from our general analysis. Similarly, inserting the coefficients (3.14) and mode functions (3.13) into equation (2.35) we obtain the leading infrared contribution to the spectral density, which exactly matches that of equation (3.10) and thus validates our model-independent analysis. There is a similar agreement between our general analysis and the results of a sharp transition in the intermediate regime.

The main differences between a realistic transition and an idealized one, however, appear in the modes that are smaller than the horizon at the end of inflation. In the ultraviolet, up to a phase, we find that the Bogolubov coefficients approach

$$\beta_k \approx \frac{p(p-1)}{4(-k\eta_0)^2}, \quad \alpha_k \approx 1 + i\frac{p(p-1)}{2(-k\eta_0)} - \frac{p^2(p-1)^2}{8(-k\eta_0)^2}. \tag{3.21}$$

Observe that care must be exercised when comparing the Bogolubov coefficients defined using different sets of mode functions: The coefficients (3.21) do not agree exactly with the coefficients one gets from equations (2.26) because the corresponding mode functions χ_k do not match either. In any case, as noted earlier, the $1/k^2$ dependence of $|\beta_k|$ implies that the energy density is not renormalizable, in the sense that the renormalized energy density is itself logarithmically divergent. An analogous logarithmic divergence was noted by Ford for nearly conformally coupled massless scalars [9]. In the de Sitter limit, the coefficients in (3.21) are exact (the omitted terms vanish), resulting in $|\beta_k|^2 = (k\eta_0)^{-4}/4$, as found, for instance, in [13]. The non-renormalizability of the energy density implies that such a transition is unphysical: Presumably, the (infinite) backreaction due to the quantum field would prevent the sharp transition from actually happen. The expectation of the energy-momentum tensor after a sharp transition has been also discussed in references [15] and [16].

It is also instructive to determine the energy density within the particle production formalism. In this case we expect the total energy density to be the sum of the energy of each mode times the number density of particles in that mode, as in equation (2.40). However, in order to follow previous literature [13, 14], we shall rely instead on equation (2.38), which as we anticipated yields similar results. We shall simply ignore the ill-behaved ultraviolet and focus on the infrared and intermediate regimes. Substituting the Bogolubov coefficients (3.14) and the associated mode functions (3.13) into (2.38), for any value of p at late times, the dominant contribution to the energy density would presumably be

$$\frac{d\rho_\text{p}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu} H_0^4 \left(\frac{a_0}{a}\right)^4 \left[1 + \frac{\mathcal{H}^2}{2k^2}\right] \left(\frac{k}{a_0 H_0}\right)^{3-2\nu}. \tag{3.22}$$

This expression is correct in the intermediate regime, $\mathcal{H} \ll k \ll \mathcal{H}_r$, since it reduces to the spectral density (3.16) in that range. In fact, as we mentioned earlier, in this regime the particle production formalism applies, and we could have simply used the particle production formula (2.40) to find the correct answer. But the approximation utterly fails in the infrared, where the leading term in equation (3.22) drastically differs from equation (3.10). The reason behind the disagreement is that equation (3.22) misses the cross terms proportional to $\alpha_k \beta_k^*$ and their conjugates in equation (2.35), which are not subdominant in the infrared, because the condition $\omega_k \gg \mathcal{H}$ is not satisfied. In particular, in the limit $\omega_k = k \ll \mathcal{H}$, the latter cancel the leading terms proportional to $|\beta_k|^2$. For similar reasons, it is also quite clear at this point that, in general, equation (2.38) and the particle production formula (2.40) fail to estimate the contribution of the long wavelength modes of a massless field. Equation (3.22) is essentially the result derived in reference [13], and its surviving contribution in the limit $\eta \to \infty$ is the one derived in [14].

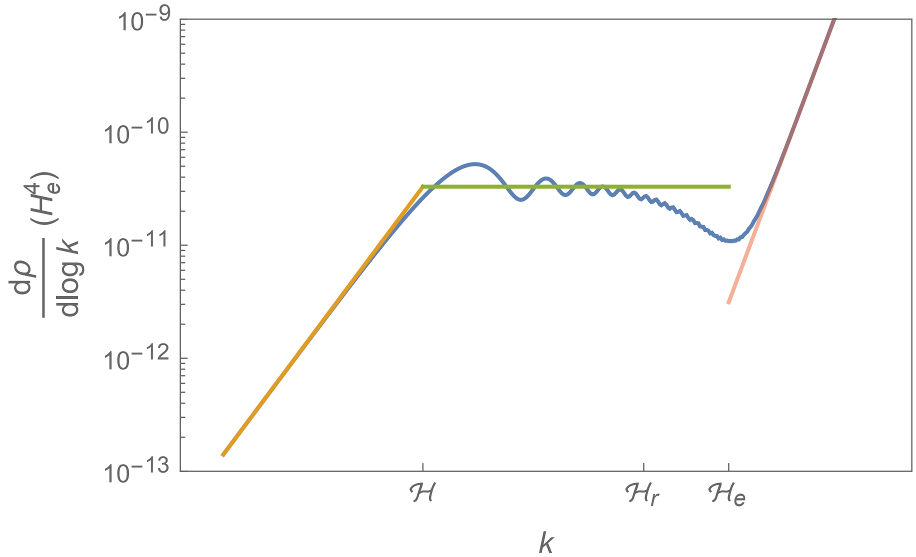

Figure 5. Spectral density of a massless scalar field after a smooth transition from de Sitter inflation to radiation domination (in blue). The superimposed power-law spectra (orange, green and red) are the corresponding analytical approximations in the infrared (3.10), intermediate (3.16) and ultraviolet regimes (3.19). Note that the oscillations in the intermediate regime arise from the corrections in equation (3.17), and decrease in amplitude as k increases, as expected. The oscillations in the far end of the ultraviolet that we expect from equation (3.27) are not perceptible in this figure. Note that the particle production formalism reproduces the correct behavior only in the intermediate range $\mathcal{H} < k < \mathcal{H}_r$. In this example $a/a_e = 3 \cdot 10^2$.

Smooth Transition. The ultraviolet divergence that we have encountered in the sharp transition prevents us from making meaningful statements about the ultimately observable renormalized energy density. Let us hence explore now the "smooth" transition to radiation domination described by (2.19). This transition is only smooth in the sense that $\ddot{a}$ remains continuous at $\eta = \eta_e$. The third and higher derivatives of a do still experience a jump, although this has no impact on the renormalizability of the energy density.

With the scale factor given by equation (2.19) the properly normalized solution of the mode equation (2.4) is

$$\chi_k = 2^{i\mu} e^{ik\eta_r} \frac{\Gamma(1+i\mu)}{\sqrt{2k}} I_{i\mu}\left(e^{-r\eta}\right), \quad \mu \equiv \frac{k}{r}, \tag{3.23}$$

where $I_{i\mu}$ is the modified Bessel function of imaginary order $i\mu$, and we have introduced a time η_r deep in the radiation era for later convenience. Using the power series of the $I_{i\mu}$ it is readily seen that in the limit of large η the mode function approaches the positive frequency solution (3.13).

In order to analyze the infrared and intermediate regimes we note that for modes that satisfy $k\eta_r \ll 1$, equation (2.21) implies that $\mu \ll 1$. Therefore, expanding the mode function (3.23) to first order in μ using the derivative of the Bessel function with respect to its order we arrive at the approximation

$$\chi_k \approx \frac{I_0(e^{-r\eta})}{\sqrt{2k}} + \frac{i}{\sqrt{2k}}\left[I_0(e^{-r\eta})\left(\log 2 + r\eta_r - \gamma\right) - K_0(e^{-r\eta})\right]\frac{k}{r}. \tag{3.24}$$

This is a linear combination of two linearly independent solutions of the mode equation at $k = 0$, just as in equation (2.9). In fact, with the scale factor given by (2.19), by changing integration variables and using the Wronskian of the Bessel functions it is seen that b in equation (2.9) is also a linear combination of $I_0(e^{-r\eta})$ and $K_0(e^{-r\eta})$. Just as in the case of the sharp transition, the two independent solutions of the mode equation at $k = 0$ are therefore contained in the expansion of the mode function to first order in k. Hence, by construction, to first order in k the mode function χ_k^{in} in equation (2.14) is proportional to the scale factor, as we established previously in equation (3.9).

Matching the small argument expansion of (3.1) to the approximate mode functions in equation (3.24) we arrive at the Bogolubov coefficients

$$\beta_k \approx -\frac{1}{2}\frac{\Gamma(\nu)}{\sqrt{\pi}}\left(-\frac{k\eta_0}{2}\right)^{1/2-\nu}\frac{1}{a_0}\left[\frac{cr}{k} + ic\left(r\eta_r + d + \log 2 - \gamma\right)\right], \quad \alpha_k \approx -\beta_k^*. \tag{3.25}$$

We obtained these expressions by evaluating (2.16) at $\eta = \eta_0$, and they are only valid in the superhorizon limit $k \ll \mathcal{H}_r$. Since the solutions (3.23) also approach equation (3.13) at late times, we can directly compare the coefficients (3.25) with those in (3.14). Inspection of the asymptotic behavior of the scale factor using equation (2.23) identifies the values of a_r and $\dot{a}_r$. Then, equation (3.25) reduces to (3.14) and we can reproduce the same results as in subsection 3.3. It follows, in particular, that the spectral density in the infrared limit is given by equation (3.10), whereas in the intermediate regime we recover equations (3.16) and (3.17).

In the ultraviolet regime, applying Hankel's expansion for large arguments to equation (3.1), and using the power series expansion for the modified Bessel function in equation (3.23), we find, to leading order in departures from the *out* vacuum ($\beta_k = 0$, $\alpha_k = 1$), that

$$\beta_k \approx i\,\frac{p(p-1)(1-r\eta_0)}{4(-k\eta_0)^3}, \quad \alpha_k \approx 1 + i\frac{p(p-1)}{2(-k\eta_0)}\left(1 - \frac{1}{2r\eta_0}\right). \tag{3.26}$$

The term in β_k proportional to $1/(k\eta_0)^2$ vanishes on account of equation (2.21), which follows from demanding that the second derivative of the scale factor be continuous at the transition, as we anticipated in equations (2.26). Then, with $|\beta_k|^2 \propto k^{-6}$, the energy density diverges in the ultraviolet just like for the vacuum state, and one can carry out its regularization and renormalization as described in subsection 3.3. In this particular case there is an additional finite ultraviolet contribution from the non-vanishing Bogolubov coefficients that could be attributed to particle production. In order to determine its magnitude, it just suffices to focus on the leading terms in

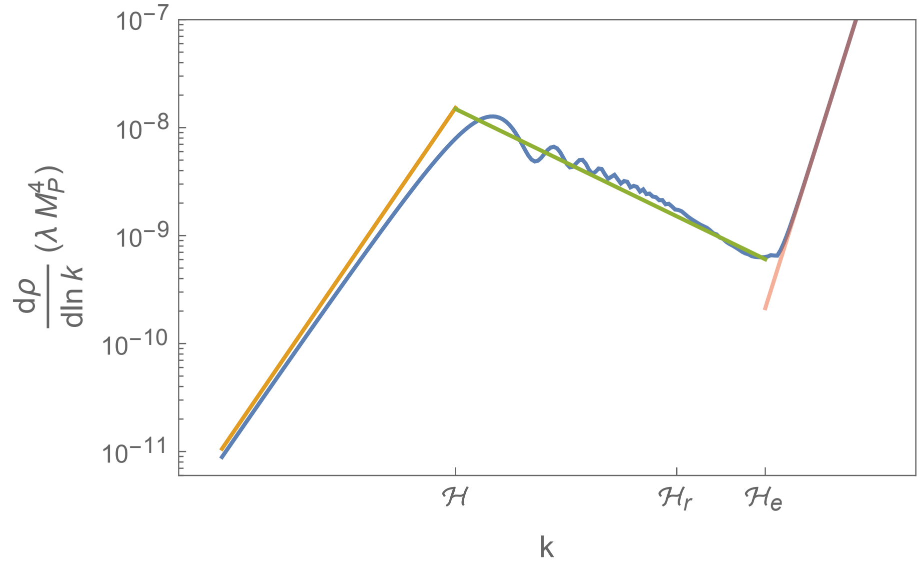

Figure 6. As in figure 5, but for the differentiable transition resulting from the inflationary potential (2.24). In this example $a/a_e = 10^2$.

equation (2.41). Bearing in mind that the Bogolubov coefficients are those in (3.26), we arrive at

$$\frac{d\rho_{\rm p}^{\rm UV}}{d\log k} \approx \frac{(p-1)^2(1-r\eta_0)^2}{32\pi^2 p^4} H_0^4 \left(\frac{a_0}{a}\right)^4 \left(\frac{a_0 H_0}{k}\right)^2 - \frac{(p-1)(1-r\eta_0)}{8\pi^2 p^2} H_0^3 H \left(\frac{a_0}{a}\right)^3 \sin[2k\eta+\varphi]. \tag{3.27}$$

The contribution from the first factor is what we would have obtained using the particle production approximation (2.40), but notice that, even though it oscillates, the second term has an amplitude that dominates throughout the ultraviolet. Therefore, on top of equation (3.20), there are two additional ultraviolet contributions to the energy density after the end of inflation, both of order H_e^4 around the end: One that scales like radiation, and another one that oscillates in time with frequency $\sim 2\mathcal{H}_e$ and a decaying amplitude proportional to a^{-6}.

A comparison of our general analytic predictions with the actual spectral density after a smooth transition is shown in figure 5. Strictly speaking equation (3.16) does not apply in the window $\mathcal{H}_r \lesssim k \lesssim \mathcal{H}_e$. But if we replace $\dot{a}_r$ by the value of $\dot{a}$ when a given mode enters the horizon, and note that $\dot{a}$ increases monotonically after the end of inflation we can readily explain the slight dip in the spectral density there.

The main problem of a smooth transition arguably is that in the ultraviolet it predicts a density of produce particles, of order H_e^4, that is of the same magnitude as the model-independent results we derived in the infrared and intermediate ranges. Therefore, if taken literally, the smooth transition would overestimate the scalar energy density after a realistic transition.

Differentiable Transition. The shape of the spectral density that we have described in our general analysis, and verified in the case of a smooth transition, survives in the realistic transition to radiation domination discussed in subsection 2.2.3. Figure 6 shows the numerically computed spectral density after the end of inflation. The spectral density in the infrared and intermediate regimes are well approximated by our analytical estimates, and the sharp increase in the ultraviolet is what we would expect from the leading k^4 behavior in the adiabatic approximation. In order to determine the value of p that appears in the analytical estimates, we have used equation (2.25) evaluated at a time a mode in the corresponding range left the horizon. Note that the main differences between the infrared and intermediate modes in figures 5 and 6 stem from inflation not being exactly de Sitter in our realization of a differentiable transition.

3.5 Overview

Overall we have found that the energy density of a massless scalar after a transition from inflation to radiation domination behaves as the superposition of three components, one for each mode range. When inflation is de Sitter-like ($p \approx -1$), all scale essentially like radiation,

$$\rho_{\rm IR} \approx \frac{H_e^4}{16\pi^2}\left(\frac{\dot a_r}{\dot a_e}\right)^2\left(\frac{a_e}{a}\right)^4, \quad \rho_{\rm IM} \approx \frac{H_e^4}{8\pi^2}\left(\frac{\dot a_r}{\dot a_e}\right)^2\left(\frac{a_e}{a}\right)^4 \log\frac{a}{a_r}, \quad \rho_{\rm UV} \approx -\frac{H_e^4}{16\pi^2}\left(\frac{a_e}{a}\right)^4, \tag{3.28}$$

where in the intermediate range we just quote the early time limit. These contributions are not only proportional to H_e^4, but they also crucially depend on the ratio $\dot a_r/\dot a_e$, which is the only factor that depends on the details of the transition and also determines the dominant component. Unfortunately, since the details of the transition from inflation to radiation domination remain unknown, the ratio $\dot a_e/\dot a_r$ is largely unconstrained. Note in particular that $\ddot a \propto \rho - 3p$, so whether $\dot a$ increases or decreases after the end of inflation depends on the equation of state of the universe after that time. We could hide this ratio in equations like (3.28) by invoking the identity

$$H = H_e\left(\frac{\dot a_r}{\dot a_e}\right)\left(\frac{a_e}{a}\right)^2, \tag{3.29}$$

which only applies during radiation domination, but this is only useful in equations (3.28) if one knows what $H(a)$ is, which is not typically the case. Let us emphasize again that among the three components in (3.28) only $\rho_{\rm IM}$ can be obtained within the particle production formalism, and that none of them can be cast as the energy density of a classical homogeneous scalar.

The total energy density of the scalar remains subdominant during radiation domination, unless the scale of inflation H_e was close to the Planck mass, which clashes with cosmic microwave background constraints on the amplitude of the primordial tensor modes. On the other hand, when inflation is far from de Sitter ($p < -2$), during radiation domination the infrared component scales like curvature, the intermediate one like dark energy (possibly violating the energy conditions), and the ultraviolet again like radiation H^4. Hence, we expect the intermediate modes to dominate in that case. The results of this section are further summarized in table 2.

It is also straightforward to extend our analysis to a transition to matter domination. In that case the spectral density in the infrared and ultraviolet is the same as during radiation domination. In the intermediate range, however, we have to distinguish between those modes that reenter the horizon during radiation domination, and those that reenter during matter domination. The former behave in exactly the same way as they do during radiation domination; the spectral index of the latter, however, changes from $3-2\nu$ to $1-2\nu$.

In addition to the previous contributions to the energy density we also have to consider those of the counterterms, which appear as a consequence of nonaccessible physics in the ultravioet. For a massless field they only renormalize the value of the cosmological constant, the Planck mass and the dimension four curvature scalars by constant values. During radiation domination only the first two survive.

4 Light Fields

Let us turn our attention to massive fields, in the limit in which their mass is much smaller than the Hubble constant during inflation, which we assume to be near de Sitter in this section. Although this case does share some of the properties of the massless limit, the presence of a mass term modifies the evolution of the *in* mode functions and introduces new features in the spectral density, both because of the appearance of the nonrelativistic regime and the breaking of a conformal symmetry that would be otherwise present during radiation domination. At the beginning of inflation the comoving mass of the field ma is lower than the infrared cutoff, and the field can be regarded in practice as massless, as the following analysis reveals. In this regime, the results of section 3 apply nearly without modification. As time goes by, however, the comoving mass grows, and it is possible for the field to become effectively massive, $ma > \Lambda_{\mathrm{IR}}$. The latter is the regime that we explore in this section, where we describe how the initial spectral density is increasingly distorted away from that of a massless scalar as the universe expands.

4.1 Inflation

When the scalar field ϕ is massive, simple analytical solutions of the mode equation (2.4) are known to us only during de Sitter inflation. Close to de Sitter, $p \approx -1$, we can substitute m^2a^2 in the dispersion relation (2.4) by the value it would have for $p=-1$, resulting in the following approximate solution for the *in* mode functions,

$$\chi_k^{\mathrm{in}}(\eta) \approx \frac{\sqrt{-\pi\eta}}{2} H_\nu^{(1)}(-k\eta), \quad \nu \equiv \sqrt{\frac{(1-2p)^2}{4} - p^2\frac{m^2}{H_0^2}}, \tag{4.1}$$

This expression differs from its massless counterpart (3.1) only in the value of the index ν, which agrees with that in (3.1) when $m=0$. By construction, the solution (4.1) is exact when $p=-1$, but is otherwise only valid for a limited period of time that in

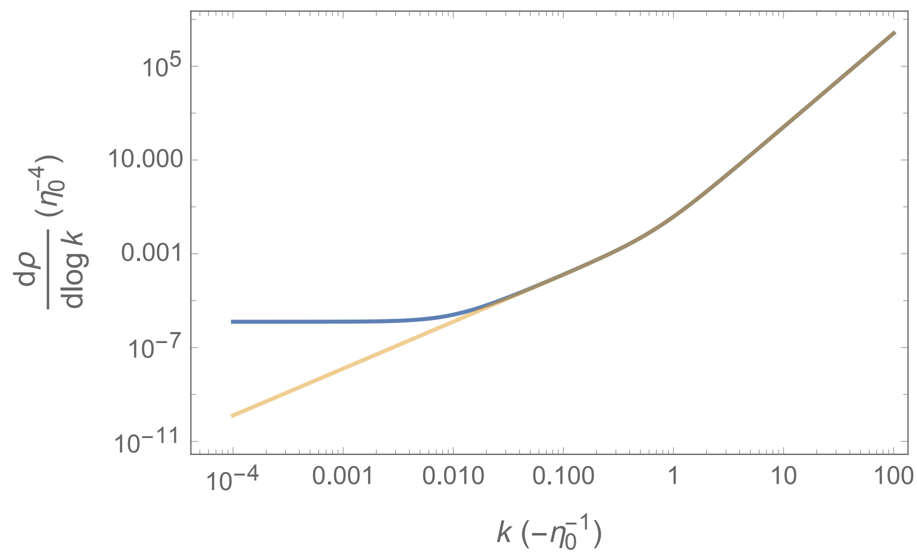

Figure 7. Spectral density during de Sitter inflation for a light field with $m = 10^{-2}H_0$ (blue). For comparison we also show the spectral density in the massless case (light orange.) The behavior in the infrared is well captured by equation (4.4). In particular, in the far infrared the spectral density is nearly flat, with a small blue tilt due to the mass term, and in the near infrared it matches that of a massless field. In the ultraviolet, the spectral density is seen to approach the leading term in equation (4.2), as expected. The energy density of the flat tail in the far infrared is negligible unless the number of e-folds of inflation is exponentially large.

general does not extend too far into the future or past. For that reason, the value of H_0 inside ν is that of the Hubble constant at the particular time η_0 around which we construct our approximation, which in principle can vary from mode to mode.

We analyze the validity of equation (4.1) in appendix B. For "ultra-light" fields, those whose mass satisfies $m^2 \ll |1+p|H^2$, equation (4.1) is a valid approximation at zeroth order in m/H_0, hence the choice of the expansion point η_0 remains arbitrary and cannot appear in our final expressions, where we can use $\nu \approx 3/2 + |1+p|$. On the other hand, for "merely-light" fields, those whose mass satisfies $|1+p|H^2 \ll m^2 \ll H^2$, the mode functions (4.1) remain valid at zeroth order in $1+p$. At this order the Hubble factor does not change in time, see e.g. equation (B.3), and we can write $\nu \approx 3/2 - m^2/(3H_0^2)$, with η_0 arbitrary within the regime of valifity of our approximation. **In phenomenologically realistic inflationary models, with $\epsilon_1 \sim |1+p| \sim 10^{-2}$ [35], the range of possible masses of a merely-light field is relatively restricted, whereas that of ultra-light field encompasses a much larger inteval in parameter space.**

In order to analyze the spectral density during inflation it is convenient to split the modes into the same ultraviolet and infrared ranges as in section 3.1. Using the adiabatic approximation, it takes some work then to check that in the ultraviolet the spectral density approaches

$$\frac{d\rho_{\rm UV}}{d\log k} \approx \frac{k^4}{4\pi^2 a^4}\left[1+\frac{1}{2}\left(\frac{aH}{k}\right)^2+\frac{1}{2}\left(\frac{am}{k}\right)^2 +\frac{3}{8}\frac{p^2-1}{p^2}\left(\frac{aH}{k}\right)^4+\frac{1}{4}\left(\frac{am}{k}\right)^2\left(\frac{aH}{k}\right)^2-\frac{1}{8}\left(\frac{am}{k}\right)^4\right]. \tag{4.2}$$

Because the mode function (4.1) is just an approximation, equation (4.2) does not quite agree with the (incorrect) form of the spectral density that we would find using (4.1). The latter yields equation (4.2) with the coefficient in front of the second-order term $a^4m^2H^2/k^4$ replaced by $-(2+p)/(4p)$. Both approaches then agree when (4.1) is exact, namely, when $p=-1$ or $m=0$. In fact, as we mention in appendix B, it is not difficult to see that the squared moduli of the mode functions (4.1) differ from ther adiabatic approximation counterparts (2.6) at second adiabatic order precisely when $p\neq -1$ or $m\neq 0$. Integrating the spectral density (4.2) over the subhorizon modes and subtracting equation (2.33c) we arrive at the renormalized energy density

$$2\pi^2\rho^{\rm UV}_{\rm ren} \approx \left[\delta\Lambda^f-\frac{m^4}{64}\left(1+2\log\frac{\mu^2}{4H^2}\right)\right]-3H^2\left[(\delta M_P^2)^f-\frac{m^2}{72}\left(1+\frac{3}{2}\log\frac{\mu^2}{4H^2}\right)\right] +\frac{36(p^2-1)H^4}{p^2}\left[\delta c^f+\frac{1}{384}\log\frac{\mu^2}{4H^2}\right]-\frac{122p^2-60p+57}{480p^2}H^4, \tag{4.3}$$

which is again finite and cutoff-independent. The reader can compare this expression with its counterpart in the massless case, equation (3.5). In the massive case, the renormalized Planck mass and cosmological constant run with time, since the presence of a physical quantity with the dimensions of mass allows these "constants" to be renormalized. Cosmologies with such running constants have been studied, for instance, in [33].

To find the spectral density in the infrared we expand the mode functions (4.1) to second order in $k\eta$ and substitute into equation (2.33a). At that order, and in the light field limit, we arrive at

$$\frac{d\rho_{\rm IR}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu}H^4\left[\frac{m^2a^2}{k^2}+1\right]\left(\frac{k}{aH}\right)^{5-2\nu}. \tag{4.4}$$

To further analyze the behavior of the spectral density (4.4), it shall prove convenient to distinguish the "far infrared," $k\ll ma$, where the mass term dominates the dispersion relation, from the "near infrared," $am\ll k\ll \mathcal{H}$, where the mass term is negligible and the mode is effectively relativistic. In the near infrared, equation (4.4) replicates equation (3.6), as expected. At this point it is important to hearken back to the results of appendix B, where we argue that it is advantageous to choose H_0 to be the Hubble constant at which the mode k crosses the horizon, H_k. Although the spectral index $5-2\nu$ has a slightly different (scale-dependent) value in the massive case then, since

$$\nu \approx \frac{3}{2}+|1+p|-\frac{1}{3}\frac{m^2}{H_k^2}, \tag{4.5}$$

the difference in the limit we consider here, $|1+p| \ll 1$ and $m^2/H_k^2 \ll 1$, is negligible. Therefore, in the near infrared, where the spectral density matches that of the massless case, the energy density is well approximated by (3.7), provided that we replace the second term in the square brackets there by the appropriate contribution of the lower boundary here, $\Lambda_{\rm IR} \to ma$. In any case, the lower boundary contribution soon becomes negligible after the onset of inflation. Hence, the near infrared energy density is approximately that of the de Sitter limit quoted after (3.7).

As shown in figure 8, the far infrared window $\Lambda_{\rm IR} < k < ma$ opens up when ma surpasses the cutoff $\Lambda_{\rm IR}$. Since in the limit of light fields $ma_i \ll \mathcal{H}_i \equiv \Lambda_{\rm IR}$, the far infrared window is always closed at the beginning of inflation, though it will eventually open, be it during inflation, or at any later, after a number of e-folds $N_{\rm fIR} = \log(H_i/m)$. In the far infrared the spectral density (4.4) is redder (closer to being singular) than its massless counterpart, which is perhaps unanticipated, because one would expect a mass term to regulate any existing infrared singularity in the massless theory. Nevertheless, the structure of (4.4) is what our approximations would suggest. Both in the infrared and the light field limits the spectral density is proportional to $(k^2+m^2a^2)|\chi_k^{\rm in}|^2$, because on superhorizon scales the derivative terms give subdominant contributions, since $\chi_k^{\rm in}/a$ is essentially frozen on those scales. In the massless case the leading infrared contribution is thus proportional to $k^2|\chi_k^{\rm in}|^2$, as in equation (3.10). But when $k^2 \ll m^2a^2$ the leading term is then proportional to $m^2a^2|\chi_k^{\rm in}|^2$, which explains the softer infrared behavior and the additional factors of m^2a^2 in equation (4.4).

By repeated integration by parts, the integral of the spectral density (4.4) over the far infrared modes can be expanded in powers of the small running of the spectral index, $d\log(3-2\nu)/d\log k \sim |1+p|$. At lowest order in the expansion we obtain

$$\rho_{\rm fIR} \approx \left[\frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{3-2\nu}m^2H^2\left(\frac{k}{aH}\right)^{3-2\nu}\right]_{\Lambda_{\rm IR}}^{ma}. \tag{4.6}$$

The behavior of this positive energy density is quite sensitive to the parameters of the model. In the merely-light limit, $|1+p|H^2 \ll m^2 \ll H^2$, after some work, the spectral density reduces to

$$\rho_{\rm fIR} \approx \frac{3H^4}{16\pi^2}\left[1-\left(\frac{a_i}{a}\right)^{\frac{2}{3}\frac{m^2}{H^2}}\right], \qquad \text{need to revisit this} \tag{4.7}$$

which remains finite as the infrared cutoff is removed, $a_i \to 0$. Initially, the energy density (4.7) grows like $\rho_{\rm fIR} \sim H^2m^2\log(a/a_i)$, until it eventually reaches $\rho_{\rm fIR} \sim H^4$. It is the presence of this additional contribution what causes the well-known discontinuity of the renormalized energy density in de Sitter spacetime as the massless limit is approached. As discussed in [34], the disagreement can be alternatively traced back to the absence of an appropriate de Sitter invariant *Fock* vacuum state for a massless scalar. We plot the spectral density during de Sitter inflation in figure 7.

In the opposite limit of ultra-light fields, in which $m^2 \ll |1+p|H^2$, the energy density is still given by (4.6), but the dependence of the spectral index on scale is

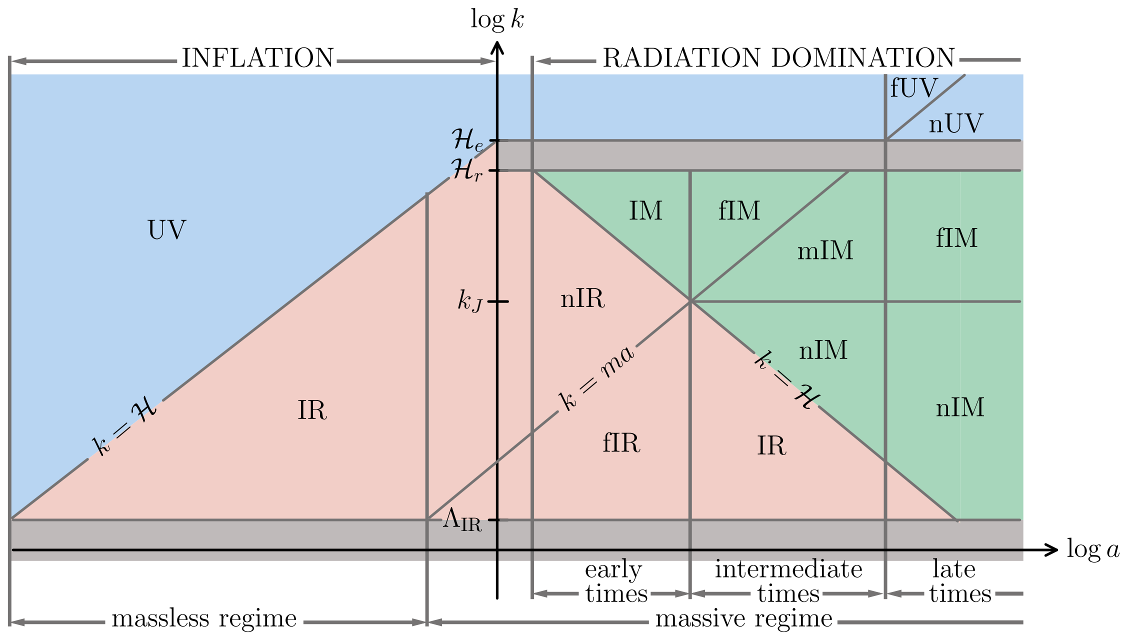

Figure 8. Different regimes of evolution of a massive scalar field in the light field limit, $ma_0 \ll \mathcal{H}_e$. The scales k_J and $\mathcal{H}_e$ remain constant in time during radiation domination, whereas ma grows linearly with time and $\mathcal{H}$ decreases as the inverse of time. The difference between the early, intermediate and late time regimes essentially relies on whether the scale ma that sets the onset of nonrelativistic effects lies on the infrared, intermediate and ultraviolet mode ranges, respectively. In this work we restrict our attention times such that $\mathcal{H} > \Lambda_{\rm IR}$, so all modes that eventually reenter the horizon left during inflation.

negligible, since $\nu \approx 3/2 + |1+p|$. Hence, in this case the energy density becomes

$$\rho_{\rm fIR} \approx \frac{m^2 H_i^2}{16\pi^2 |1+p|} \left[1 - \left(\frac{\Lambda_{\rm IR}}{ma} \right)^{2|1+p|} \right], \tag{4.8}$$

which grows like $m^2 H_i^2 \log(ma/\Lambda_{\rm IR})$ once the far infrared window opens. This growth is essentially the result of the random walk often attributed to a massless field during inflation, $\langle \phi^2 \rangle \sim H^2 N$ [43]. Alas, because typically $\rho_{\rm fIR} \ll \rho_{\rm nIR}$, the near infrared modes dominate the infrared energy density, and in practice, during inflation the infrared behavior is the same as in the massless case.

4.2 General Infrared Evolution After Inflation

Just as we did in the case of a massless field, it is possible to explore the behavior of the spectral density in the infrared in relatively full generality. Again, the key here is to utilize the approximate solution in equation (2.9), which generally remains valid at low frequencies, that is, when modes are long and the field is light, provided that the equation of state during the transition significantly deviates from $w = -1$, as argued in subsection A.2 of appendix A. In fact, during inflation, the last condition is not satisfied. Nevertheless, in appendix B we show that we can trust the approximation (2.9) when the field remains ultra-light, or if the additional condition (B.8) holds when the field is merely-light, as we shall implicitly assume for simplicity. Once inflation approaches its end, the value w significantly deviates from the one near de Sitter, and does not pose any obstacle to the applicability of the approximation (2.9).

We shall use then the solution (2.9) to join the known mode functions during inflation to the known solutions during radiation domination, even when the evolution of the scale factor during the intervening transition phase is unknown. This approximation only works when the low frequency approximation is applicable, while $ma \ll \mathcal{H}$, which we shall refer to as the limit of early times. Note that the latter extend past radiation domination if $m \gtrsim H_{\mathrm{eq}} \sim 10^{-28}\,\mathrm{eV}$, as happens, for instance, in late dark energy models with masses of order $m \sim 10^{-33}$ eV or lighter.

Because the mode functions during inflation (4.1) still has the same form as in the massless case, just with a different value of ν, and so does the mode function (2.9), we can readily borrow the result quoted in equation (3.9) as an approximate solution to the mode equation with the appropriate *in* boundary conditions. In this case the term proportional to $b - b_0$ does not vanish, but it is proportional to the small parameter $m^2/H_0^2 \ll 1$. It is also important to stress that the ν-dependent exponent in equation (3.9) already appears at zeroth order in the long wavelength expansion, so it is consistent to keep its mass dependence in subsequent approximations. Then, substituting equation (3.9) into the spectral density (2.33a) we find again that the time derivative terms are negligible. Therefore, in the infrared we arrive at

$$\frac{d\rho_{\mathrm{IR}}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu} H_0^4 \left(\frac{a_0}{a}\right)^2 \left[\frac{m^2 a^2}{k^2} + 1\right] \left(\frac{k}{a_0 H_0}\right)^{5-2\nu}. \tag{4.9}$$

This expression approximately agrees with the spectral density that we derived in equation (4.4) during inflation by different means. The origin of the difference, a factor of order $(a_0/a)^{2m^2/(3H_0)^2}$, stems from the somewhat different time evolution of the mode function (4.1), which we used to arrive at (4.4), and the mode function (3.9), which used to arrive at (4.9), after the mode has left the horizon. In fact, if we had kept the correction proportional to $b - b_0$ in (3.9) we would have found agreement to first order in m^2/H_0^2. In any case, as we discuss in appendix B, the difference between the two is small for ultra-light fields, and also remains small for merely-light fields when (B.8) is satisfied, which is the condition under which we can trust (B.8) to begin with. Otherwise equation (4.9) is universal, in the sense that it applies at early times, $ma \ll \mathcal{H}$, to all superhorizon modes, just as their massless counterpart (3.10) did. Recall, in particular, that the time η_0 is an arbitrary time during inflation, which actually drops out of equation (4.9) to first order in m^2/H_0^2 if the aforementioned correction is included. It ought to be clear from our derivation that at early times the particle production formalism does not apply in the infrared, where frequencies are low.

Just as we described in subsection 4.1, the precise contribution of the far infrared to the energy density delicately depends on the values of the different parameters of the model,

$$\rho_{\text{fIR}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{3-2\nu}m^2 H_e^2 \left[\left(\frac{m}{H_e}\frac{a}{a_e}\right)^{3-2\nu} - \left(\frac{\Lambda_{\text{IR}}}{a_e H_e}\right)^{3-2\nu}\right], \tag{4.10}$$

where we have chosen again $\eta_0 = \eta_e$. We may also return to the question whether this can be cast as the energy density of an appropriate homogeneous classical scalar. Indeed, because the energy density in equation (4.10) stems from the mass term in equation (2.33a), and in the far infrared we can approximate $\chi_k^{\text{in}} \approx \alpha_k \chi_0^{m=0} + \beta_k \chi_0^{m=0*}$, with α_k and β_k given by equation (3.8), $\chi_0^{m=0}$ by the low frequency expansion (2.9), and $\omega_k^2 \approx m^2 a^2$, we would obtain the same density if the classical scalar (2.29) had amplitudes

$$\beta_0 = -\alpha_0^* = \sqrt{\frac{V}{2\pi^2}\int_{\Lambda_{\text{IR}}}^{ma}\frac{dk}{k}k^3\,\beta_k^2}, \tag{4.11}$$

which are approximately constant because the integral in (4.11) is dominated by the lower boundary. Note that such a scalar would be purely imaginary, which agrees with the character of the *in* mode functions at long wavelengths (3.9). As we discussed in subsection 2.3.1, at early times such a scalar would remain "frozen," and its energy density would indeed behave like a dark energy component. The agreement, however, does not extend to the near infrared, because in that limit the gradient terms give the dominant contribution to the energy density, and, as in the massless case, those terms identically vanish when the field is homogeneous.

Equations (4.9) and (4.10) apply as long as the field remains light, $ma \ll \mathcal{H}$, no matter during which cosmological epoch. The approximately constant (and positive) ρ_{fIR} in this regime is what motivated reference [17] to identify the light scalar with a late time dark energy candidate. Although there is an additional nearly constant contribution to the vacuum energy stemming from the subtraction terms, the far infrared modes start to behave like nonrelativistic matter later on, so it is in principle possible to disentangle (4.10) from other contributions to an effective cosmological term. It is in fact conceivable, though beyond the scope of our present analysis, that an analogous form of dark energy, just with a heavier mass m, may resolve the much debated "Hubble tension" [35], as discussed in [36] (see [37], however.) To the extent that our quantized field shares some of the phenomenology of an oscillating classical scalar (see below for further details), this would agree with the analysis in [38]. A somewhat related explanation of the Hubble tension that relies on the field fluctuations of a light scalar during inflation has been recently proposed in [39].

In the near infrared interval $ma \ll k \ll \mathcal{H}$ the modes are relativistic and outside the horizon at the time of interest, and the spectral density (4.9) reduces to that of the massless case in the infrared, equation (3.10). Although the lower integration limit is different here, we can also borrow the expression for the energy density in the massless case (3.11) by replacing $\Lambda_{\text{IR}} \to ma$. In that case the lower boundary gives a negligible contribution, and the near infrared energy density tracks the background evolution, as it did when $m = 0$. This again underscores that when $ma < \Lambda_{\text{IR}}$ we can treat the field as effectively massless. In that limit, the far infrared window does not exist, and the contribution of the near infrared agrees with that of the massless case. As we noted in the context of inflation, this implies that for ultra-light fields the near modes dominate the energy density in the infrared.

Overall, then, we find that at early times the near infrared energy density in the modified equation (3.11) tracks that of the background and dominates over the nearly constant dark energy-like contribution of the far infrared in equation (4.10). But towards the end of early times, $ma = \mathcal{H}$, the near infrared range disappears, see figure 8, so after the end of inflation there must be a period during which the far infrared modes come to dominate the energy density. The far and near infrared densities are of the same order around $ma \sim \mathcal{H}$, so this period is rather short-lived. Once the universe expands past early times, the modes that used to be in the far infrared begin to redshift like nonrelativistic matter, as we shall see next.

4.3 Radiation Domination

As in the massless case, in order to compute the spectral density during radiation domination we shall simply assume that there must be a (smallest) time η_r after which we can safely assume the universe to be radiation dominated. Here, the evolution of the modes depends on the location of the comoving mass ma relative to the comoving horizon, so it shall prove convenient to split the radiation dominated era into early times, $m\,a \ll \mathcal{H}$, intermediate times, $\mathcal{H} \ll ma \ll \mathcal{H}_0$, and late times, $\mathcal{H}_0 \ll ma$. Similarly, during each epoch the location of a given mode within the different scales in the problem changes, so it is useful to split the different modes into distinct, non-overlapping, ranges. In the following we proceed to discuss the spectral density during these epochs at all mode regimes above $\Lambda_{\rm IR}$. We recommend that the reader consult figure 8 for a visual reference of the different mode intervals at the three different epochs.

Early Times ($m\,a \ll \mathcal{H}$). We argued in subsection 4.2 that, as long as the field remains light, the behavior of the spectral density in the infrared, $k \ll \mathcal{H}$, is universal. Hence, at early times during radiation domination the spectral density is given by equation (4.9), within the same infrared ranges identified in figure 8. In the intermediate interval $\mathcal{H} \ll k \ll \mathcal{H}_r$ the modes are relativistic at the end of inflation and remain that way at early times, so the spectral energy density behaves like that of a massless field, and thus agrees with the one quoted in equation (3.16). In the far and near infrared the energy density is given by equation (4.10) and the appropriately modified (3.11), respectively. The energy density in the intermediate regime is again (3.18).

Modes in the ultraviolet regime, $\mathcal{H}_e \ll k$, are inside the horizon also at the end of inflation, and thus remain adiabatic and relativistic throughout early times. In this range, given that the field mass is irrelevant, the same relation between the Bogolubov coefficients and the smoothness of the transition that we discussed in subsection 3.4 applies. If the transition is differentiable, then, we expect a sharp fall-off of the Bogolubov coefficients, and the spectral density approaches that in equation (4.2), provided we simply set $p = 1$, as appropriate during radiation domination. The contribution of the modes with $\mathcal{H}_e \ll k \sim \Lambda$ to the energy density is subtracted out in equation

(2.33b), so the renormalized energy density in this range stems from the *out* adiabatic vacuum,

$$2\pi^2\rho_{\rm ren}^{\rm UV} \approx \left[\delta\Lambda^f - \frac{m^4}{64}\left(1+2\log\frac{\mu^2}{4H_e^2}\frac{a^2}{a_e^2}\right)\right] - 3H^2\left[(\delta M_P^2)^f - \frac{m^2}{18}\left(1+\frac{1}{2}\log\frac{\mu^2}{4H_e^2}\frac{a^2}{a_e^2}\right)\right] + \frac{H^4}{480} + \frac{H^2H_e^2}{8}\left(\frac{a_e}{a}\right)^2 - \frac{H_e^4}{8}\left(\frac{a_e}{a}\right)^4. \tag{4.12}$$

Setting $m = 0$ in (4.12) we return to the energy density of a massless field in equation (3.20). The main difference between the two cases is that here the cosmological and Newton's "constants" run. Terms proportional to the curvature counterterm δc^f vanish during radiation domination, but the energy density nevertheless contains a contribution of fourth adiabatic order, proportional to H^4, that is of the same order as the one generally expected from the curvature invariants. Again, the adiabatic approximation breaks down around $k \sim \mathcal{H}_e$, so we expect an error in $\rho_{\rm ren}^{\rm UV}$ of same order as the leading prediction of the quantum theory, namely, the term proportional to H_e^4.

Figure 9 shows a plot of the spectral density at early times after a smooth transition from de Sitter inflation, which neatly displays the expected behavior in the four relevant mode ranges. As during inflation, see figure 7, at early times the spectral density differs from that of the massless case only in the far infrared, whose contribution to the energy density is negligible. In this limit we recover the same result as in the massless case; see subsection 3.5 for a discussion. Only as the universe approaches intermediate times do the far infrared modes begin to dominate, mimicking a dark energy component with equation of state $w \approx -1$. Like in the massless case, only intermediate modes admit an interpretation in terms of particles. In contrast, the far infrared admits an interpretation in terms of a classical field, in addition to that of the modes with $k < \Lambda_{\rm IR}$.

Intermediate Times ($\mathcal{H} \ll m\,a \ll \mathcal{H}_e$). When $\mathcal{H} \ll m\,a$, the low frequency approximation (2.9) we relied on at early times and long wavelengths breaks down. Still, the exact solutions of the mode equation during radiation domination are known to be the parabolic cylinder functions [10, 17],

$$\chi_k = \frac{1}{(2ma\mathcal{H})^{1/4}}\exp\left(\frac{\pi}{8}\frac{k^2}{ma\mathcal{H}}\right)D_\alpha\left[(1+i)\sqrt{\frac{m}{H}}\right], \quad \alpha \equiv -\frac{1}{2}\left(1+i\frac{k^2}{ma\mathcal{H}}\right), \tag{4.13}$$

which we have normalized by demanding that they approach the adiabatic limit (2.6) at "late times," when all modes are nonrelativistic. These solutions are the massive counterparts of equation (3.13), even though it is non-trivial to take the limit $m \to 0$ in the former. Inspection of the index α indicates the appearance of a new momentum scale in the problem, $k_J \equiv \sqrt{ma\mathcal{H}} = \sqrt{m\dot{a}_r} = const$. As we discuss in appendix A.1, this is related to the (broken) conformal invariance during radiation domination and the adiabaticity of the modes. Modes with $k \ll k_J$ are not adiabatic at early times, but modes with $k_J \ll k$ remain adiabatic throughout radiation domination, and it is then simpler to work in this range with the adiabatic approximation (2.6). Note

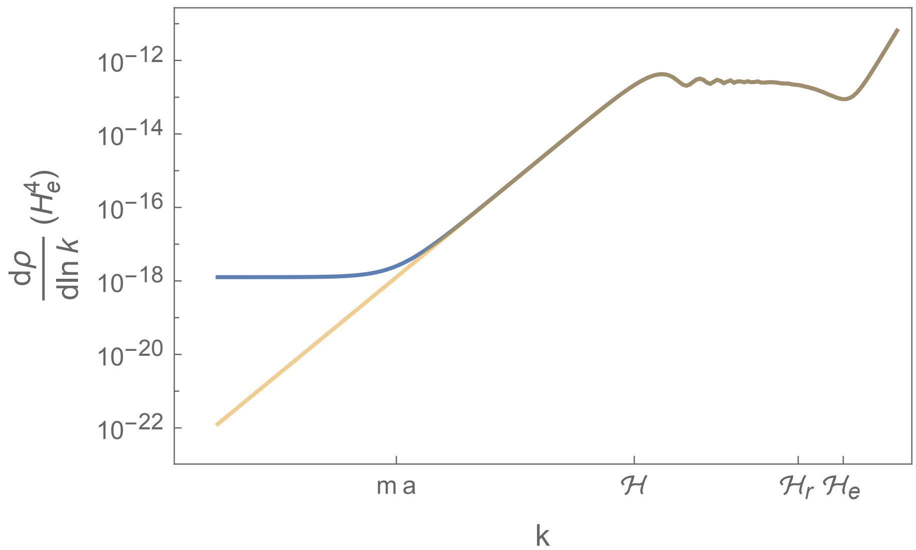

Figure 9. Spectral density at early times, $ma \ll \mathcal{H}$, after a smooth transition from de Sitter. In this example $m = 10^{-8}\, H_e$ and $a/a_e = 10^3$ (blue). For comparison we also plot the spectral density in the massless case (light orange). The spectral density is well approximated by (4.9) when $k \ll \mathcal{H}$, (3.16) when $\mathcal{H} \ll k \ll \mathcal{H}_r$ and (4.2) when $\mathcal{H}_e \ll k$. The only significant difference with the massless case is the presence of a flat tail in the far infrared, whose energy density behaves like a dark energy component. Its relevance depends on the parameters of the model. The particle production formalism only applies in the intermediate range $\mathcal{H} \ll k \ll \mathcal{H}_r$, though it does not capture the oscillations seen in the spectral density.

that this new scale is the inverse of the Jeans length of a self-gravitating scalar field identified in [40].

Since we already determined the universal solution of the mode equation that remains valid when both k and m are negligible, we can simply find the appropriate mode function χ_k^{in} throughout radiation domination by matching equation (3.9) to a linear combination of (4.13) at η_r. In the interval $k \ll k_J$ (which includes both the infrared and near intermediate modes, see figure 8) the Bogolubov coefficients are

$$\beta_k \approx -\frac{\Gamma(\nu)\Gamma(\frac{1}{4})}{4\pi}(-p)^{1/2}\left(-\frac{k\eta_0}{2}\right)^{-\nu}\left(\frac{\dot{a}_r}{\dot{a}_0}\right)^{3/4}\left(\frac{H_0}{m}\right)^{1/4}, \quad \alpha_k \approx -\beta_k^*, \tag{4.14}$$

where we have used the asymptotic form of the parabolic cylinder functions at early times, when $m/H \ll 1$. By the nature of the derivation, this analysis only applies if the field is still light at η_r, which is natural in this context, unless the mass of the field was close to the Hubble scale at the end of inflation. Plugging these Bogolubov coefficients into equation (2.14) and exploiting the form of the parabolic cylinder functions at late times (large arguments) we get

$$\chi_k^{\text{in}} \approx -i\frac{\Gamma(\nu)\Gamma(\frac{1}{4})}{2\pi}(-p)^{1/2}\left(-\frac{k\eta_0}{2}\right)^{-\nu}\left(\frac{\dot{a}_r}{\dot{a}_0}\right)^{3/4}\left(\frac{H_0}{m}\right)^{1/4}\frac{1}{\sqrt{2ma}}\sin\left(\frac{m}{2H}\right), \tag{4.15}$$

which happens to oscillate with time and decaying amplitude, as opposed to being proportional to the scale factor as during early times. Finally, substituting (4.15) into the spectral density (2.33a) we arrive at

$$\frac{d\rho_{\mathrm{IR+nIM}}}{d\log k} \approx \frac{\Gamma^2(\nu)\Gamma^2(\frac{1}{4})}{2^{5-2\nu}\pi^4}(-p)^{1-2\nu} m^2 H_0^2 \left(\frac{a_{\mathrm{osc}}}{a}\right)^3 \left(\frac{k}{a_0 H_0}\right)^{3-2\nu}. \tag{4.16}$$

Here, a_{osc} denotes the value of the scale factor at the threshold of intermediate times, when $\mathcal{H}_{\mathrm{osc}} = m a_{\mathrm{osc}}$, which is when superhorizon modes start to oscillate. In this case we could have also arrived at (4.16) using the particle production formalism. We argue in appendix A.1 that at intermediate times all modes are adiabatic, including those that are nonrelativistic. Therefore, since we are dealing with modes with frequencies $\omega_k \approx ma \gg \mathcal{H}$ and the coefficients β_k in equation (4.14) are large, the three conditions in subsection 2.3.3 are met. Indeed, if we substitute the Bogolubov coefficients (4.14) into the particle production formula (2.40) with $\omega_k \approx ma$ we precisely recover equation (4.16).

In order to obtain the energy density of the modes in the infrared and near intermediate ranges we need to integrate (4.16). Since we assume that inflation lasted long enough to predict the amplitude of all subhorizon modes, $\Lambda_{\mathrm{IR}} < \mathcal{H}$, the Jeans scale k_J is also above the infrared cutoff. In that case the mode interval is finite and the energy density equals

$$\rho_{\mathrm{IR+nIM}} \approx \frac{\Gamma^2(\nu)\Gamma^2(\frac{1}{4})}{2^{5-2\nu}\pi^4}\frac{(-p)^{1-2\nu}}{3-2\nu} m^2 H_e^2 \left(\frac{a_{\mathrm{osc}}}{a}\right)^3 \left[\left(\frac{m}{H_e}\frac{\dot{a}_r}{\dot{a}_e}\right)^{3/2-\nu} - \left(\frac{\Lambda_{\mathrm{IR}}}{a_e H_e}\right)^{3-2\nu}\right], \tag{4.17}$$

which scales like nonrelativistic matter, and whose magnitude again depends on the precise balance of two unrelated quantities. Note that the dark energy component we identified at early times morphs into dust at intermediate times, essentially because the constant spectral density in the early time infrared starts to decay like a presureless fluid at intermediate times. Indeed, at the intermediate threshold $a = a_{\mathrm{osc}}$ the far infrared energy density (4.10) is of the same order as the energy density in (4.17).

The origin of equation (4.17) is also readily explained in terms of the homogeneous scalar we introduced in equation (2.29). At early times, the scalar remains frozen, but once the universe enters intermediate times, it begins to oscillate.The energy density of such an oscillating field scales like dust, so at intemediate times it is of order $\rho = \rho_{\mathrm{osc}}(a_{\mathrm{osc}}/a)^3$, as we found in equation (4.17). Such an interpretation, however, would miss the contribution of the remaining mode ranges.

In the mid intermediate range $k_J \ll k \ll ma$ it is somewhat cumbersome to work with the parabolic cylinder functions in equation (4.13). As we discuss in appendix A.1, modes in this interval remain adiabatic throughout radiation domination. Therefore, it is simpler to rely on the adiabatic mode functions (2.6) here. The latter reduce to equation (3.13) at the time of the transition, so we can then lean here on the analysis that led to the Bogolubov coefficients (3.14) in the massless case. Yet at intermediate times the adiabatic mode functions do not reduce to (3.13), but to the nonrelativistic counrerparts $\chi_k \approx \exp(-im/2H)/\sqrt{2ma}$ instead. Therefore, by plugging (3.14) into (2.14) the *in* mode functions at intermediate times become

$$\chi_k^{\mathrm{in}} \approx -i\frac{\Gamma(\nu)}{\sqrt{\pi}}\left(-\frac{k\eta_0}{2}\right)^{1/2-\nu}\frac{a_r}{a_0}\frac{\mathcal{H}_r}{k}\frac{1}{\sqrt{2ma}}\sin\left(\frac{m}{2H}\right), \tag{4.18}$$

that substituted into (2.33a) finally leads to

$$\frac{d\rho_{\mathrm{mIM}}}{d\log k} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}(-p)^{1-2\nu}\left(\frac{\dot{a}_r}{\dot{a}_0}\right)^{1/2}\left(\frac{m}{H_0}\right)^{1/2} m^2 H_0^2\left(\frac{a_{\mathrm{osc}}}{a}\right)^3\left(\frac{k}{a_0 H_0}\right)^{2-2\nu}. \tag{4.19}$$

Integrating over the modes in the mid intermediate interval, then, we find that the energy density is dominated by the lower boundary of the integral, and, again, scales like nonrelativistic matter, as expected,

$$\rho_{\mathrm{mIM}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{2\nu-2}\left(\frac{\dot{a}_r}{\dot{a}_e}\right)^{3/2-\nu}\left(\frac{m}{H_e}\right)^{3/2-\nu} m^2 H_e^2\left(\frac{a_{\mathrm{osc}}}{a}\right)^3. \tag{4.20}$$

Even though we have derived the spectral densities in (4.16) and (4.19) by rather different means, note that they agree where they overlap, at $k \sim k_J$. As a consequence, we would expect both intermediate energy densities to be roughly of the same order, provided the infrared cutoff is not very small. Once again, we could have arrived at the same expression by plugging the Bogolubov coefficients (3.14) into the particle production formalism formula (2.40), with $\omega_k \approx m\,a$. The simplicity of this approach underscores the convenience of the particle production formalism in the cases in which it does apply. From this perspective, the origin of the transition-dependent powers of $\dot{a}_r/\dot{a}_e$ in equations (4.19) and (4.20) can be traced back to their appearance in the Bogolubov coefficients (3.14), which essentially set the oscillation amplitude of the intermediate modes as they enter the horizon.

Modes in the far intermediate range $m\,a \ll k \ll \mathcal{H}_r$ are adiabatic and relativistic both at the transition and intermediate times, so this time we can directly borrow the result quoted in equation (3.16). But since the interval boundaries in the far intermediate range are different from those in the massless case, the energy density is instead

$$\rho_{\mathrm{fIM}} \approx \frac{\Gamma^2(\nu)}{2^{4-2\nu}\pi^3}\frac{(-p)^{1-2\nu}}{3-2\nu}\left(\frac{m}{H_e}\right)^{3-2\nu} H_e^2 H^2\left(\frac{a_e}{a}\right)^{2\nu-1}\left[1-\left(\frac{H}{m}\right)^{3/2-\nu}\right]. \tag{4.21}$$

which approximately behaves like the energy density of a fluid with equation of state $w = 1$ and remains subdominant in comparison to (4.20). In the ultraviolet range $\mathcal{H}_e \ll k$ the spectral density has the same form as during early times, so the corresponding energy density is again given by equation (4.12).

Although the spectral density of a massless field in figure 5 looks very similar to that of a light field at early times, figure 9, as time goes by the presence of the growing mass term deforms and alters the spectral density of a light field. This is seen in figure 10, which shows a plot of the spectral density at intermediate times. Again, the distinct asymptotic behaviors in the four different mode ranges are clearly recognizable.

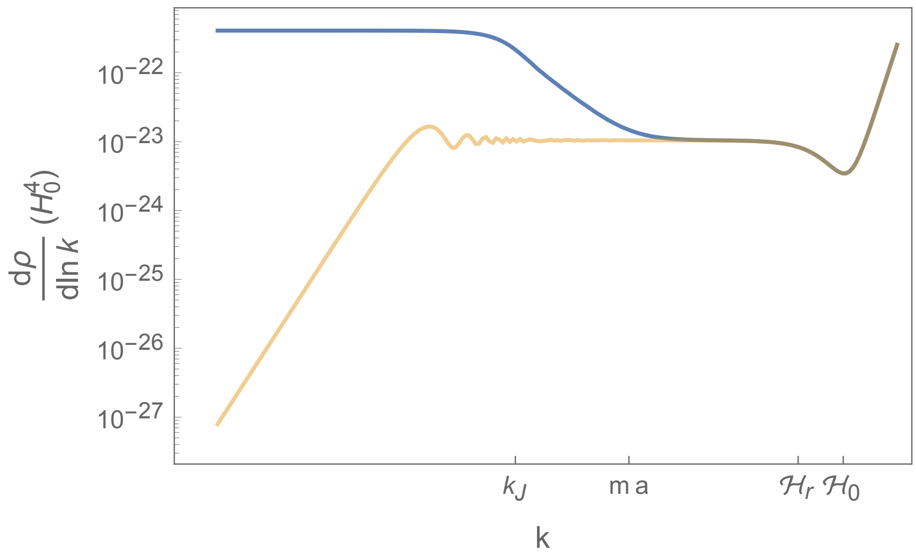

Figure 10. As in figure 9, but evaluated at an intermediate time $\mathcal{H} \ll ma \ll \mathcal{H}_e$, with $a/a_e = 4 \cdot 10^5$. The spectral density equals (4.16) when $k \ll k_J$, (4.19) when $k_J \ll k \ll ma$, (3.16) when $m\,a \ll k \ll \mathcal{H}_r$ and (4.2) when $\mathcal{H}_e \ll k$. At intermediate times, the particle production formalism reproduces the correct spectral density in the infrared and intermediate mode ranges, $k \ll \mathcal{H}_r$.

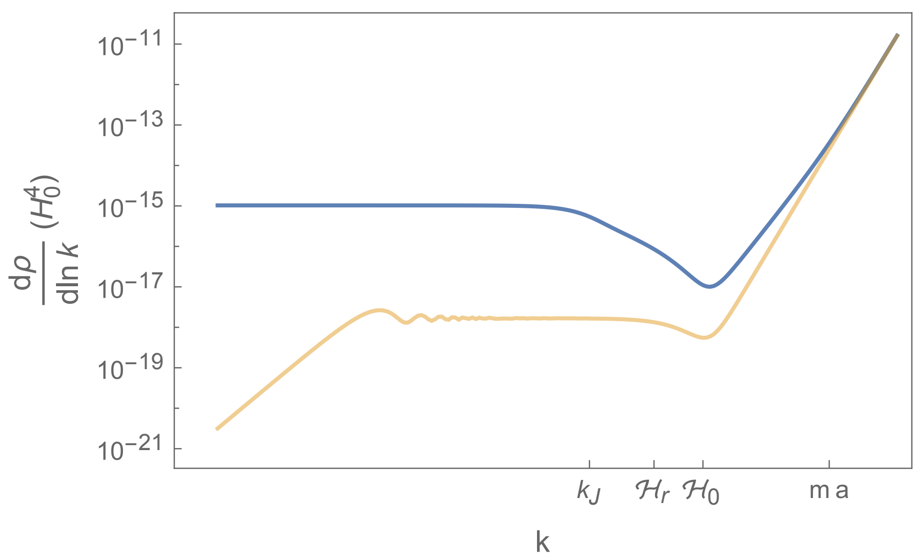

Figure 11. As in figures 9 and 10, but evaluated at a late time $\mathcal{H}_e \ll ma$. For numerical reasons in this example we have chosen $m = 10^{-3}H_e$ and $a/a_e = 2 \cdot 10^4$. The spectral density equals (4.16) when $k \ll k_J$, (4.19) when $k_J \ll k \ll \mathcal{H}_r$, (4.22) when $\mathcal{H}_e \ll k \ll m\,a$, and (4.2) when $m\,a \ll k$. As during intermediate times, the particle production formalism only applies when $k \ll \mathcal{H}_r$.

Late Times ($\mathcal{H}_e \ll ma$). The far intermediate range present at intermediate times disappears at late times, by definition, see figure 8. But, otherwise, all the expressions that we have derived for the the infrared and the two intermediate ranges at intermediate times apply here too. In the ultraviolet, $\mathcal{H}_e \ll k$, it is safe to assume that modes remain adiabatic, so the spectral density "reduces" to that of the *out* vacuum, which we quote in equation (2.34) at fourth order. In the near ultraviolet the mass terms dominate and the spectral density becomes

$$\frac{d\rho_{\mathrm{nUV}}}{d\log k} \approx \frac{1}{4\pi^2}\frac{k^3 m}{a^3}, \tag{4.22}$$

while in the far ultraviolet we recover again the relativistic result (4.2). Although the relevant modes are in the adiabatic regime, these expressions cannot be recovered from the particle production formalism because the coefficients β_k are small. The renormalized energy density in the ultraviolet in particular agrees with that of the *out* vacuum and therefore is

$$\begin{aligned} 2\pi^2 \rho^{\mathrm{UV}}_{\mathrm{ren}} \approx \left[\delta\Lambda^f + \frac{m^4}{32}\log\frac{m^2}{\mu^2}\right] - 3H^2\left[(\delta M_P^2)^f + \frac{m^2}{48}\log\frac{m^2}{\mu^2}\right] \\ -\frac{1}{6}\left(\frac{\dot{a}_r}{\dot{a}_e}\right)^{-3/2}\left(\frac{m}{H_e}\right)^{1/2} m^2 H_e^2 \left(\frac{a_{\mathrm{osc}}}{a}\right)^3. \end{aligned} \tag{4.23}$$

Leaving renormalization aside, at late times the leading contribution behaves like non-relativistic matter, with a density that is suppressed in comparison with those in equations (4.17) and (4.20).

Figure 11 shows the spectral density at late times after a smooth transition to radiation domination. As opposed to what happened at early times, the spectral density here is quite different from that of the massless case. As further illustration of our general analysis in a realistic case, we consider again the evolution of a scalar field in the quartic potential (2.24) that we analyzed in subsection 3.4. Figure 12 compares the numerically determined spectral density with our analytical estimates, and confirms again their validity.

4.4 Overview

During radiation domination, the behavior of a scalar field that was light during inflation displays a rich variety of behaviors that cannot be entirely captured by the particle production formalism or simply by a classical homogeneous scalar. At early times, when the scalar mass obeys $ma \ll \mathcal{H}$, its spectral density agrees with that of a massless scalar, up to the presence of a flat tail in the spectral density at $k \ll ma$, whose contribution behaves like a dark energy component (4.10) that is typically subdominant. Therefore, during this epoch the scalar essentially behaves like radiation with energy density $\rho \sim H_e(a_e/a)^4$. Once the field mass surpasses the Hubble constant, previously frozen modes begin to oscillate. The former dark energy morphs into a non-relavistic fluid that comes to dominate the energy density, on top of a subdominant contribution that scales like a fluid with equation of state $w = 1$. The magnitude of

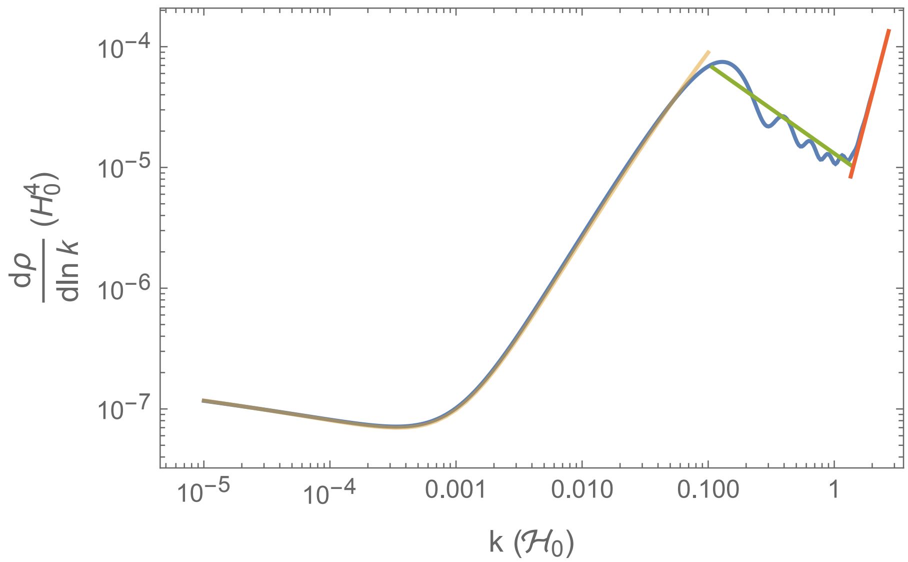

Figure 12. Spectral density in an inflationary model with chaotic potential (2.24) at early times. In this example $m = 10^{-4}H_e$ and $a/a_e = 10$. The superimposed curves follow equations (4.9), (3.16) and (4.2) in the corresponding mode ranges. With $ma = 10^{-3}\mathcal{H}_e$ and $\mathcal{H} \approx 7 \cdot 10^{-2}\mathcal{H}_e$ the transition between the different approximations happens where expected. In equation (4.9) we have taken η_0 to be the time at which the given mode crosses the horizon, with p determine by equation (2.25) at the same time. The approximation works so well that it overlaps with the actual spectral density.

the dominant contribution is quite sensitive to the particular parameters of the theory, though in the ultra-light limit it is expected to be of order $\rho \sim m^2 H_e^2 (a_{\rm osc}/a)^3$. The scalar field also renormalizes the cosmological constant and the Planck mass. As opposed to what happens in the massless case, here these "constants" experience a characteristic running with the logarithm of the scale factor that stops at late times, when ma surpasses $\mathcal{H}_e$. The corresponding change in these parameters as the mass crosses such threshold is a unique signature of this scenario that does not depend on the unknown values of the counterterms. We summarize the evolution of an ultra-light scalar in table 3.

5 Heavy Fields

We shall end our analysis with massive fields, in the limit in which the mass is much larger than the Hubble constant during inflation. In this limit, the adiabatic approximation is valid for all modes at all times, and we can reap the benefits of the general discussion in section 2. The downside of this regime is that it does not leave much room for "particle production" because the only dimensionless parameter that controls the latter is H/m, which is small by assumption.

By construction, the the renormalized *out* energy density is expected to be small, since it is of sixth adiabatic order. To estimate its magnitude, we expand the spectral

Time	Range I	Range II	Range III	UV
$m\,a \ll \mathcal{H}$	$k \ll m\,a$	$m\,a \ll k \ll \mathcal{H}$	$\mathcal{H} \ll k \ll \mathcal{H}_r$	$\mathcal{H}_e \ll k$
Early	$m^2 H_e^2 \log a$	$H_e^2 H^2$ (D)	-?-	H^4
$\mathcal{H} \ll m\,a \ll \mathcal{H}_r$	$k \ll k_J$	$k_J \ll k \ll m\,a$	$m\,a \ll k \ll \mathcal{H}_r$	$\mathcal{H}_e \ll k$
Intermediate	$m^2 H_e^2 \left(\frac{a_e}{a}\right)^3$ (D)	$m^2 H_e^2 \left(\frac{a_e}{a}\right)^3$ (D)	$H_e H^3$	H^4
$\mathcal{H}_e \ll m\,a$	$k \ll k_J$	$k_J \ll k \ll m\,a$	$\mathcal{H}_e \ll k \ll m\,a$	$m\,a \ll k$
Late	$m^2 H_e^2 \left(\frac{a_{\rm osc}}{a}\right)^3$	$m^2 H_e^2 \left(\frac{a_{\rm osc}}{a}\right)^3$	$H_e H^3$ (D)	$-\left(\frac{m}{H_e}\right)^{1/2}\left(\frac{a_{\rm osc}}{a}\right)^3$

Table 3. Energy density of a light scalar during radiation domination. Each entry shows the limits of each mode range and the energy density therein. For simplicity, we consider here the ultra-light field limit $m^2 \ll H_a^2 \ll |1+p|$, with $\Lambda_{\rm IR} \ll ma$. The "D" in parenthesis denotes the dominant component.

density of the *out* vacuum to sixth adiabatic order, and subtract the spectral density whose integral reproduces the subtraction terms, quoted in equation (2.34). Integrating over all modes below the cutoff and restoring the contribution of the counterterms we obtain

$$\begin{aligned}
2\pi^2 \rho_{\rm ren}^{\rm out} \approx {} & \left[\delta\Lambda^f + \frac{m^4}{32}\log\frac{m^2}{\mu^2}\right] - 3H^2\left[(\delta M_P^2)^f + \frac{m^2}{48}\log\frac{m^2}{\mu^2}\right] \\
& + 48\left[\delta c^f - \frac{1}{384}\log\frac{m^2}{\mu^2}\right]\left(\frac{\mathcal{H}^2 \ddot a}{a^5} + \frac{\ddot a^2}{4a^6} - \frac{\mathcal{H}\,\dddot a}{2a^5}\right) \\
& + \frac{1}{m^2}\left[\frac{211 H^6}{5040} - \frac{173}{280}\frac{H^4 \ddot a}{a^3} + \frac{1237}{2240}\frac{H^2 \ddot a^2}{a^6} + \frac{673}{10080}\frac{\ddot a^3}{a^9} + \frac{683}{1680}\frac{H^3\,\dddot a}{a^4}\right. \\
& \left. - \frac{673}{3360}\frac{H \ddot a\,\dddot a}{a^7} + \frac{17}{2240}\frac{\dddot a^2}{a^8} - \frac{17}{140}\frac{H^2 a^{(4)}}{a^5} - \frac{17}{1120}\frac{\ddot a\, a^{(4)}}{a^8} + \frac{17}{1120}\frac{H a^{(5)}}{a^6}\right].
\end{aligned} \tag{5.1}$$

The terms of zeroth, second and fourth adiabatic order in this equation depend on the contribution of the counterterms, so these are not predictions of the quantum theory. Those of fourth order renormalize the dimension four curvature invariants, which vanish in de Sitter and during radiation domination. Note that none of the coupling constants run with cosmic expansion. An expansion in powers of m of the standard renormalized energy density in de Sitter [41] reveals that the first non-vanishing term is indeed proportional to H^6/m^2, and exactly agrees with the predicted sixth-order term in (5.1).

Although $\rho_{\rm ren}^{\rm out}$ is small, when the field is heavy during inflation we expect it to provide the dominant contribution to the renormalized energy density. All we need to assume is that the modes remain adiabatic between inflation and radiation domination, and, in particular, that the transition is differentiable (we analyze the sharp and smooth transitions in the next subsection.) Then, the uninterrupted validity of the adiabatic approximation implies that the Bogolubov coefficients β_k vanish, so the associated energy density of the produced particles $\rho_{\rm p}$ vanishes too. Therefore, setting $\ddot a = 0$ in equation (5.1) yields the energy density during radiation domination,

$$2\pi^2 \rho_{\rm ren} \approx \left[\delta\Lambda^f + \frac{m^4}{32}\log\frac{m^2}{\mu^2}\right] - 3H^2\left[(\delta M_P^2)^f + \frac{m^2}{48}\log\frac{m^2}{\mu^2}\right] + \frac{211}{5040}\frac{H^6}{m^2}. \tag{5.2}$$

As we noted above, the contribution of order H^4 to the energy density is absent.

5.1 Examples

To arrive at equation (5.2) we have assumed that the scale factor and its time derivatives remain continuous during the transition. If the derivatives of the scale factor jump, the Bogolubov coefficients do not vanish, and the renormalized energy density receives additional contributions from the produced particles, $\rho_{\rm p}$, as on the examples we discuss next.

Sharp Transition. In a universe that expands like equation (2.18), the solutions of the mode equation (2.4) are still given by the parabolic cylinder functions. Yet these are somewhat unwieldy, particularly in the limit of large masses. Since we are interested in the limit of heavy fields anyway, it turns out to be simpler just to consider the adiabatic approximation, which remains valid for all modes throughout inflation and radiation domination. From equations (2.26), at two time derivatives the corresponding Bogolubov coefficients become

$$\beta_k \approx \frac{p(p-1)}{4}\left[\frac{\mathcal{H}_e}{p\omega_k(\eta_e)}\right]^2\left[1+\frac{m^2 a_e^2}{2\omega_k^2(\eta_e)}\right], \quad \alpha_k \approx 1. \tag{5.3}$$

The main difference with the massless case is that in the infrared limit, the mode functions and the Bogolubov coefficients remain finite. In this limit, the coefficient β_k is suppressed, as expected, by a power of $(H_e/m)^2$. The suppression extends to the ultraviolet, where β_k is of order $(\mathcal{H}_e/k)^2$. As we discussed earlier, this renders the energy density of the field non-renomalizable. In fact, it is readily seen that in the ultraviolet the value of β_k agrees with the one we found in the massless case, equation (3.21). Yet, again, since the energy density is not renormalizable, it is not strictly possible to obtain a sensible finite value that captures the ultraviolet contribution. We shall instead evaluate the renormalized energy density for a smooth transition.

Smooth Transition. With the scale factor given by (2.19), it is not possible to find analytical solutions of the mode equation (2.4). We shall hence focus again on the adiabatic approximation, which remains valid for all modes as long as the field is sufficiently heavy.

In order to find the Bogolubov coefficients in the adiabatic approximation, we shall use equations (2.26) again. Because the scale factor is continuous up to its second derivative, it suffices to focus on those terms that contain three time derivatives of the scale factor. Doing so we arrive at

$$\beta_k \approx i\frac{p(p-1)(1-r\eta_e)}{4}\left[\frac{\mathcal{H}_e}{p\omega_k(\eta_e)}\right]^3\left[1+\frac{m^2 a_e^2}{2\omega_k^2(\eta_e)}\right], \quad \alpha_k \approx 1+\beta_k^*. \tag{5.4}$$

In the infrared limit, the coefficient β_k is suppressed by $(H_e/m)^3$, one additional power than in the case of a sharp transition. This again illustrates that particle production is related to departures from adiabaticity: The smoother the transition, the higher the suppression. In the ultraviolet regime we find $\beta_k \sim (\mathcal{H}_e/k)^3$, which yields a renormalizable energy density and agrees with equation (3.26).

Because the Bogolubov coefficients β_k are mildly suppressed here, the renormalized energy density of the field may stray away from that of the field in the *out* vacuum in equation (5.2). In order to determine the contribution due to the nonzero β_k, we simply substitute equation (5.4) into equation (2.39). The integral proportional to $|\beta_k|^2$ can be evaluated exactly, but we only reproduce but we just reproduce here its limit when $a \gg a_e$,

$$\rho_{\mathrm{p}}^{(0)} \approx \frac{5p^2(p-1)^2(1-r\eta_0)^2}{2048\pi p^6} H_e^4 \frac{H_e^2}{m^2}\left(\frac{a_e}{a}\right)^3, \tag{5.5}$$

which scales like nonrelativistic matter, as expected, and dominates over the "vacuum" piece in equation (5.2). Note that equation (5.5) is what we would have obtained from the particle production formula (2.40). To estimate the contribution proportional to $|\alpha_k\beta_k|$, we restrict our attention to the nonrelativistic modes, which do not oscillate with k and ought to give the dominant contribution to the mode integral. In the same limit we find

$$\rho_{\mathrm{p}}^{(1)} \approx \frac{3(p^2-p)(1-r\eta_e)}{4\pi^2 p^3} H_e^4 \left(\frac{H_e}{m}\right)^{-1} \frac{H}{m} \log\left(\frac{a}{a_e}\right)\left(\frac{a_e}{a}\right)^3 \sin\left(m\int^{\eta} \tilde{a}\, d\tilde{\eta}\right). \tag{5.6}$$

Note that both energy densities $\rho_{\mathrm{p}}^{(0)}$ and $\rho_{\mathrm{p}}^{(1)}$ are essentially determined by two independent dimensionless expansion parameters, H_e/m and H/m. The former is the one that controls the magnitude of the Bogolubov coefficients, whereas the latter is the one that controls the accuracy of the derivative expansion. Although $\rho_{\mathrm{p}}^{(1)}$ indeed contains one more power of H/m than $\rho_{\mathrm{p}}^{(0)}$, the latter is suppressed by three additional powers of H_e/m. Hence, $\rho_{\mathrm{p}}^{(1)}$ may actually dominate the energy density of the "produced" particles if H_0/m is small enough. The time average of the energy density $\rho_{\mathrm{p}}^{(1)}$ over a Hubble time further suppresses $\rho_{\mathrm{p}}^{(1)}$ by an additional power of H/m, but it does not change that its relative enhancement by a factor of $(m/H_e)^3$. As we have repeatedly emphasized, these considerations illustrate that there are cases in which the particle production formula (2.40) is not the correct approximation to the energy density of the field.

6 Summary and Conclusions

In this article we have explored the conditions under which the energy density of a minimally coupled scalar can be computed using the particle production formalism, particularly after a transition between two cosmological epochs. In order for the particle production formula (2.40) to correctly approximate the energy density in a particular mode range, it is not only necessary for the corresponding modes to be in the adiabatic regime, but also that frequencies are large ($\omega_k \gg \mathcal{H}$) and particle production is effective ($|\beta_k| \gg 1$.) During radiation domination the first condition is satisfied, but the second fails on superhorizon scales. This is why work in the literature that estimated the energy density using the particle production formalism arrived at incorrect values in the infrared regime.

In cosmological spacetimes, Pauli-Villars regularization proves to be extremely useful in regulating and renormalizing the otherwise divergent vacuum energy density. In this approach, a set of massive regulator fields cancel the divergent energy density of the scalar in the ultraviolet and thus render it finite. This allows us to contentrate on the spectral density and to study the contribution of different mode regimes to the energy density separately. At scales lower than the regulator masses, the regulators can be largely ignored. Only at higher scales do the latter impact the total spectral density, by renormalizing the cosmological constant, the Planck mass and the coupling constant of an appropriate combination of curvature invariants. For massive fields, the running caused by this renormalization leads to changes in these "constants" that do not depend on the coutnerterms themselves; these are then testable predictions of the quantum theory.

We have also seen that the smoothness of the transition strongly influences the spectral density in the ultraviolet, where the field mass is negligible. If the first derivative of the scale factor changes discontinuously, particle production is so copious that the energy density diverges in the ultraviolet. Such a transition is therefore unphysical and should not be taken too literally. When only the second derivative of the scale factor changes discontinuously, the energy density of the produced particles does remain finite, but, still, such a discontinuous change is only to be regarded as a toy model.

We can make definite model-independent predictions because any transition from inflation is expected to be differentiable. In that case, the Bogolubov coefficients are highly suppressed in the ultraviolet, where the spectral density simply approaches that of the *out* vacuum This implies that the low-momentum modes typically provide the dominant contribution to the energy density. The spectral density of modes that remain frozen is essentially transition-independent, whereas that of the modes that oscillate, be it because they have entered the horizon or become effectively massive, only depends on the ratio $\dot{a}_\mathrm{r}/\dot{a}_e$.

The actual magnitude of the energy density, as well as the form of the spectral density, critically depends on the ratio of the mass of the scalar to the Hubble scale during inflation. We summarize the behavior of the spectral density of a massless scalar in subsection 3.5, and that of a massive scalar that is light during inflation in subsection 4.4. Although in some mode ranges it is possible to reproduce the corresponding energy densities by invoking the particle formalism or a classical homogeneous scalar, there is no case in which all of the spectral density can be captured by either. In particular, the spectral density typically displays a diversity of behaviors that cannot be simply captured by a single fluid.

Our results illustrate that care must be taken when using the particle production formalism, which we view as a simple shortcut to the spectral density when the appropriate conditions apply. Yet, in the end, other than convenience, we do not see a strong reason to introduce particles in a framework that ultimately deals with fields. In the field theories we use to describe the universe all that ought to be relevant is the expectation value of field operators like the energy density. Some of this work can be interpreted as the first step in the grand scheme of characterizing all of cosmology by appropriate expectation values.

A Validity of the Low and High Frequency Approximations

In this appendix we explore in more detail the regime of validity of the low and high frequency approximations of subsections 2.1.2 and 2.1.1 . Particular attention is paid to power-law expanding universes, equation (2.13).

A.1 High Frequency Approximation

As we mention in subsection 2.1.1 the *adiabatic expansion* is an expansion of the solutions of the mode equation in the number of time derivatives of the scale factor. The expansion is exact if the spacetime is non-expanding, but also in the presence of a conformal symmetry, since Friedmann-Robertson-Walker spacetimes are conformally related to Minkowski space. To the extent that we typically keep only a finite number of terms of the expansion, we shall refer to it as the adiabatic approximation.

The validity of the adiabatic approximation demands that the effective frequencies introduced in equation (2.8) satisfy $W_k^{(0)} \gg {}^{(2)}W_k \gg {}^{(4)}W_k \gg \dots$. The first inequality translates into

$$\frac{3}{8}\frac{\dot{\omega}_k^2}{\omega_k^4} - \frac{1}{4}\frac{\ddot{\omega}_k}{\omega_k^3} - \frac{1}{2}\frac{\ddot{a}}{\omega_k^2 a} \ll 1. \tag{A.1}$$

If $p \neq 1$ it is hence generally necessary to impose $\omega_k \gg \mathcal{H}$ in order for the last term on the left hand side of equation (A.1) to be smaller than one. This alone guarantees that the first two terms of this expression are also small, and in addition implies that the remaining inequalities above are generically satisfied too. At early times, $ma \ll \mathcal{H}$, the condition of adiabaticity imposes $k \gg \mathcal{H}$; only subhorizon modes are adiabatic. At late times, $ma \gg \mathcal{H}$, however, all modes are adiabatic.

Indeed, during radiation domination ($p = 1$) adiabaticity is less restrictive because of the conformal symmetry that appears when $m = 0$. In this case the last term on the left hand side of equation (A.1) vanishes and a necessary and sufficient condition for the validity of (A.1) is instead

$$\frac{m^2 a^2 \mathcal{H}^2}{\omega_k^4} \ll 1. \tag{A.2}$$

Higher orders in the adiabatic approximation lead to similar conclusions, since $\ddot{a} = 0$ and the scale factor always appears in conjunction with m. At early times, $ma \ll \mathcal{H}$, the inequality (A.2) demands $k \gg k_J$, where

$$k_J \equiv \sqrt{ma\mathcal{H}}. \tag{A.3}$$

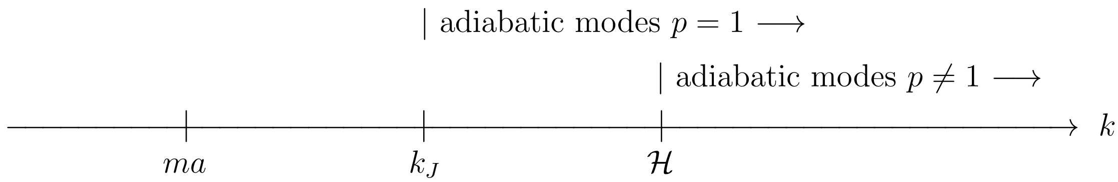

Figure 13. The adiabatic regime at early times, $ma \ll \mathcal{H}$, during radiation domination. Note that all adiabatic modes are relativistic, $k \gg ma$. If $p \neq 1$ only subhorizon modes $k \gg \mathcal{H}$ are adiabatic. If $p = 1$ this condition is relaxed, $k \gg k_J \equiv \sqrt{ma\mathcal{H}}$, and there are superhorizon modes which are also adiabatic. This is a consequence of the conformal symmetry that appears in a radiation dominated universe when $m = 0$ (where all modes are adiabatic). As time goes by ma ($\mathcal{H}$) moves to the right (left) in the figure until the transition between early and intermediate times at $ma \sim k_J \sim \mathcal{H}$, whereas k_J remains constant. At intermediate and late times, $ma \gg \mathcal{H}$, all modes are adiabatic.

Therefore, in addition to the subhorizon modes, there are superhorizon modes which are also adiabatic. Note that the scale k_J is constant during radiation domination, and it is superhorizon-sized at early times as indicated in figure 13. At late times, $ma \gg \mathcal{H}$, again, all modes are adiabatic. If we set $m = 0$, that is, if we restore the conformal symmetry, all modes satisfy (A.2) at any time and the zeroth order term in the adiabatic approximation is already exact; the solution of the mode equation is (3.13).

In a power-law expanding universe, some of the properties discussed above are manifest in the following adiabatic expansion of the mode functions to second order, which is valid in the limit of light fields, $ma \ll \mathcal{H}$, and short wavelengths, $\mathcal{H} \ll k$,

$$\begin{aligned}\chi_k^{(2)} \approx \frac{1}{\sqrt{2k}} e^{-ik\eta} \Bigg\{ 1 - i \left[\frac{p-1}{2} + \frac{p}{2(2p+1)} \left(\frac{m}{H}\right)^2 \right] \left(\frac{aH}{k}\right) \\ - \left[\frac{(1+p)(1-p)(2-p)}{8p} + \frac{p^2+p+1}{4(2p+1)} \left(\frac{m}{H}\right)^2 + \frac{p^2}{8(2p+1)^2} \left(\frac{m}{H}\right)^4 \right] \left(\frac{aH}{k}\right)^2 \Bigg\}.\end{aligned} \tag{A.4}$$

The corrections are suppressed by factors of $\mathcal{H}/k$, which are small inside the horizon. If the field is massless, these corrections disappear during radiation domination, when the adiabatic expansion is exact and reproduces (3.13). We shall use this expression in appendix B in order to determine the validity of the approximation (4.1) at short scales. The scale k_J defined in equation (A.3) arises from the expansion of $\chi_k^{(2)}$ to order $1/k^4$.

A.2 Low Frequency Approximation

The *low frequency expansion* in subsection 2.1.2 is an expansion of the solutions of the mode equation (2.4) in powers of ω_k^2. The expansion is exact for the zero mode of a massless field, no matter the form of the scale factor. To the extent that we typically keep only a finite number of terms of the expansion, we shall refer to it as the low frequency approximation.

The validity of the low-frequency approximation demands that the corrections to the mode functions decrease as the order of the series increases, $\chi_k^{(0)} \gg {}^{(2)}\chi_k \gg \dots$, where ${}^{(n+2)}\chi_k \equiv \chi_k^{(n+2)} - \chi_k^{(n)}$ denotes the correction of order ω_k^n in the expansion of the mode function χ_k, which can be obtained from equation (2.11) by replacing $\chi_k^{(n)}$ by ${}^{(n)}\chi_k$ and dropping the lowest order approximation $\chi_k^{(0)}$. Taking into account the definition of b we can roughly estimate the relative difference between two of the consecutive approximations generated by iteration of (2.11),

$$
{}^{(n+2)}\chi_k \sim \eta\, a^2 b\, \omega_k^2\, {}^{(n)}\chi_k \sim \eta^2 \omega_k^2\, {}^{(n)}\chi_k, \tag{A.5}
$$

implying that our approximate solution remains valid as long as $\omega_k \ll \mathcal{H} \sim 1/\eta$, that is, in the expected regime of long wavelengths and light fields. In particular, the condition for the validity of the approximate solution (2.9) is not $\omega_k^2 \ll \ddot{a}/a$, as one may have naively expected from (2.4).

It is in fact useful to evaluate the second order correction to $\chi_k^{(0)}$ during power-law expansion (2.13). For arbitrary fixed p, and setting $n = 0$, the integral in equation (2.11) can be explicitly evaluated,

$$
\begin{aligned}
\chi_k^{(2)}(\eta) = \bar{c}_1 a &\left[1 - \frac{p^2}{2(1+2p)}\left(\frac{k}{aH}\right)^2 + \frac{p^2}{2(1+4p)}\frac{m^2}{|1+p|H^2}\right] \\
&+ \frac{\bar{c}_2}{a\mathcal{H}}\left[1 - \frac{p^2}{2(3-2p)}\left(\frac{k}{aH}\right)^2 + \frac{p^2}{6}\frac{m^2}{|1+p|H^2}\right],
\end{aligned} \tag{A.6}
$$

where the constants $\bar{c}_1$ and $\bar{c}_2$ are rescaled versions of those in (2.9). The correction is generically small whenever $k \ll \mathcal{H}$ and $ma \ll \mathcal{H}$, which agrees with our previous condition $\omega_k \ll \mathcal{H}$. This hence supports the validity of (2.9) as a zeroth order approximation to the mode functions in the limit of light fields and superhorizon modes. Near de Sitter, however, we must distinguish between ultra-light fields, $m^2 \ll |1+p|H^2$, for which the series expansion works, and merely-light fields, $|1+p|H^2 \ll m^2 \ll H^2$, for which the expansion breaks down. We provide further details about this distinction in appendix B.

B Light Fields During Inflation

In the massive case, away from de Sitter, the mode equation (2.4) has no exact solution known to us. In order to construct an approximate solution, we note that the dispersion relation during power-law inflation reads

$$
\hat{\omega}_k^2 = k^2 + \left[p^2 \frac{m^2}{H_0^2}\left(\frac{\eta_0}{\eta}\right)^{-2(1+p)} - p(p-1)\right]\frac{1}{\eta^2}. \tag{B.1}
$$

Clearly, around the vicinity of the "expansion point" η_0 we can replace the power proportional to m^2 in the dispersion relation by the (constant) value it would have in de Sitter, which remains a valid approximation as long as $|(1+p)\log(\eta/\eta_0)| \ll 1$. Then, the solution of the mode equation is indeed equation (4.1).

To further assess the validity of this approximate solution we shall explore its short and long wavelength regimes. Let us compare first its short wavelength limit $-k\eta \gg 1$ with the adiabatic expansion (2.6), which we already know to be a valid approximation on subhorizon scales. For convenience, we shall focus directly on the spectral density. Let ρ^{in}_{UV} and $\rho^{(2)}_{UV}$ be the spectral densities computed using equation (4.1) expanded to second order in $(k\eta)^{-1}$ and the adiabatic approximation (A.4) respectively. Their relative difference is

$$\frac{\rho^{\text{in}}_{UV}-\rho^{(2)}_{UV}}{\rho^{(0)}_{UV}} = -\frac{1}{4p}\left[1+(2p+1)\frac{H^2}{H_0^2}\right]\left(\frac{ma}{k}\right)^2\left(\frac{aH}{k}\right)^2+\frac{1}{8}\left[1-\frac{H^4}{H_0^4}\right]\left(\frac{ma}{k}\right)^4, \quad \text{(B.2)}$$

which gives a measure of the error committed when using the approximate solution (4.1). Surely, equation (B.2) vanishes in de Sitter, $H = H_0$, and in the massless limit, $m = 0$, where the solution (4.1) is exact. Away from these two particular cases, in the neighborhood of the expansion point η_0 the error is proportional to $1+p$, which is small near de Sitter, since

$$\frac{H}{H_0} = \left(\frac{\eta}{\eta_0}\right)^{-(1+p)} \approx 1-(1+p)\log\frac{\eta}{\eta_0}. \quad \text{(B.3)}$$

As H significantly deviates from H_0 far into the past, the difference (B.2) approaches $\sim (m/H_0)^2(\mathcal{H}/k)^4$, which is heavily suppressed for light fields at short wavelengths. On the other hand, far into the future the difference tends to $\sim (m/H)^2(\mathcal{H}/k)^4$, which is also suppressed on the same grounds.

During inflation, however, the comoving Hubble factor increases in time, and some modes eventually leave the horizon. To determine to what extent we can still trust the approximation (4.1) in the long wavelength limit, and whether equation (2.9) indeed remains a valid approximation in that regime, it is convenient to consider the *exact* solution of the mode equation (2.4) at $k = 0$,

$$\chi_0(\eta) = \bar{c}_1\sqrt{-\eta}J_\gamma\left(\frac{-2p\gamma}{1-2p}\frac{m}{H}\right)+\bar{c}_2\sqrt{-\eta}J_{-\gamma}\left(\frac{-2p\gamma}{1-2p}\frac{m}{H}\right), \quad \gamma \equiv -\frac{1-2p}{2(1+p)}, \quad \text{(B.4)}$$

where J_γ is the Bessel function of the first kind. As the scale factor approaches that in de Sitter, $p \to -1$, the order and the argument of the Bessel functions diverges, and it is convenient to rely on an approximation that works regardless of the magnitude of m/H. Such an approximation is afforded by the uniform expansion 10.20 of [32]. Keeping only the leading terms in an expansion at small m/H and $1+p$ therein we find

$$\chi_0 \propto \frac{a}{a_0}\exp\left[\frac{p^2}{2(1-2p)}\left(\frac{m^2}{|1+p|H_0^2}-\frac{m^2}{|1+p|H^2}\right)\right], \quad \text{(B.5)}$$

where we have restricted our attention only to the dominant mode and which we have normalized to one at $\eta = \eta_0$. The behavior in equation (B.5) should be compared with that of our approximate solution (4.1) in the long wavelength limit,

$$\chi_k^{\rm in} \propto \left(\frac{a}{a_0}\right)^{\frac{1-2\nu}{2p}} \approx \frac{a}{a_0}\left(\frac{a}{a_0}\right)^{\frac{p}{1-2p}\frac{m^2}{H_0^2}}, \tag{B.6}$$

where we have only kept the leading term in the $(m/H_0)^2$ expansion in the exponent and which we have also normalized to one at $\eta = \eta_0$. Apart from that, the approximation (B.6) is exact in $1+p$.

At this point, it is convenient to distinguish between the ultra-light field limit, where $m^2 \ll |1+p|H^2$, and the merely-light field limit, where $|1+p|H^2 \ll m^2 \ll H^2$. If the field is ultra-light, we can expand the exponential of equation (B.5) in a power series, whereby we recover

$$\chi_0 \propto \frac{a}{a_0}\left[1 - \frac{p^2}{2(1-2p)}\left(\frac{m^2}{|1+p|H^2} - \frac{m^2}{|1+p|H_0^2}\right)\right]. \tag{B.7}$$

Incidentally, apart from the normalization factor, this equation agrees with (A.6) in the limit $k = 0$ and $p \to -1$, which further confirms the validity of the lowest order approximation (2.9). At this point, the advantage of (A.6) over (B.7) is that the former is exact in $1+p$, and for that reason we shall use this equation to compare with (B.6), which is also exact in $1+p$. Clearly, the time dependence of (A.6) agrees with that of (B.6) in the massless limit, that is, the approximation (4.1) is correct at zeroth order in $(m/H_0)^2$. Since equation (4.1) is exact in this case, this is what we expected. Yet away from the massless case, the evolution of (A.6) differs from that of (B.6) when a/a_0 is sufficiently large, although the $(m/H_0)^2$ correction in (B.6) remains negligible as long as

$$\left(\frac{a}{a_0}\right)^{\frac{p}{1-2p}\frac{m^2}{H_0^2}} \approx 1. \tag{B.8}$$

At the same time, the low frequency approximation (A.6) remains valid as long as

$$\frac{m^2}{|1+p|H^2} = \frac{m^2}{|1+p|H_0^2}\left(\frac{a}{a_0}\right)^{\frac{2(1+p)}{p}} \ll 1, \tag{B.9}$$

which guarantees that the field is ultra-light. Since in the ultra-light limit this last power of a is much bigger than that in equation (B.8), we can thus trust the approximate solution (4.1) at zeroth order in $(m/H_0)^2$, which is otherwise exact in $1+p$. This is equivalent to setting $m = 0$ in equation (4.1) from the beginning and working in terms of the massless solution (3.1), which, among other things, implies that any reference to the expansion point η_0 will disappear from our final expressions.

In the opposite limit, $|1+p|H^2 \ll m^2 \ll H^2$, our analysis in appendix A.2 already revealed that $\chi_0 \propto a$ is not a valid approximation, and the question is whether we can still trust (4.1) instead. In this case, the argument of the exponential in equation (B.5) is not small, and we cannot use of an expansion series of the form (B.7). Substitution of the expansion (B.3) into (B.5) yields instead

$$\chi_0 \propto \frac{a}{a_0}\exp\left[\frac{p}{1-2p}\frac{m^2}{H_0^2}\log\frac{a}{a_0}\left(1 + \frac{1+p}{p}\log\frac{a}{a_0}\right)\right], \tag{B.10}$$

whose time dependence agrees with that of (B.6) in the de Sitter limit, when $1+p=0$. Since equation (4.1) is also exact in this case, this is again what we expected. Yet away from de Sitter, the evolution of (B.10) differs from that of (B.6) when a/a_0 is sufficiently large, although the $1+p$ correction in (B.10) remains negligible as long as

$$\left(\frac{a}{a_0}\right)^{\frac{1+p}{p}} \approx 1. \tag{B.11}$$

This is the same condition needed to guarantee that the Hubble factor (B.3) does not evolve in time, as it happens in de Sitter, where there is no ambiguity in the value of H_0 that appears in equation (4.1). This corresponds to working at zeroth order in $1+p$ and leading nonvanishing order in $(m/H_0)^2$ (remember that the expansions (B.5) and (B.6) are only valid at this order), which is equivalent to setting $p=-1$ in equation (4.1) from the very beginning. Even if given the apparent limited range of validity of this approximation, in most cases of interest, such as in some dark matter and dark energy models, the mass of the field is much smaller than the Hubble factor during inflation, and condition (B.11) is satisfied during an extremely large number of e-folds, $\log(a/a_0) \gg 10^{56}(H_0/M_P)^2(m/\mathrm{eV})^{-2}$.

References

[1] S. Weinberg, "What is quantum field theory, and what did we think it is?," [arXiv:hep-th/9702027 [hep-th]].

[2] P. J. E. Peebles and A. Vilenkin, "Noninteracting dark matter," Phys. Rev. D **60**, 103506 (1999) doi:10.1103/PhysRevD.60.103506 [arXiv:astro-ph/9904396 [astro-ph]].

[3] M. S. Turner and L. M. Widrow, "Gravitational Production of Scalar Particles in Inflationary Universe Models," Phys. Rev. D **37**, 3428 (1988) doi:10.1103/PhysRevD.37.3428

[4] D. J. H. Chung, E. W. Kolb and A. Riotto, "Superheavy dark matter," Phys. Rev. D **59**, 023501 (1998) doi:10.1103/PhysRevD.59.023501 [arXiv:hep-ph/9802238 [hep-ph]].

[5] S. Hashiba and J. Yokoyama, "Gravitational particle creation for dark matter and reheating," Phys. Rev. D **99**, no.4, 043008 (2019) doi:10.1103/PhysRevD.99.043008 [arXiv:1812.10032 [hep-ph]].

[6] J. A. R. Cembranos, L. J. Garay and J. M. Sánchez Velázquez, "Gravitational production of scalar dark matter," JHEP **06**, 084 (2020) doi:10.1007/JHEP06(2020)084 [arXiv:1910.13937 [hep-ph]].

[7] M. A. G. Garcia, K. Kaneta, Y. Mambrini, K. A. Olive and S. Verner, "Freeze-in from preheating," JCAP **03**, no.03, 016 (2022) doi:10.1088/1475-7516/2022/03/016 [arXiv:2109.13280 [hep-ph]].

[8] C. Armendariz-Picon, "On the Expected Backreaction During Preheating," JCAP **05**, 035 (2020) doi:10.1088/1475-7516/2020/05/035 [arXiv:2003.01542 [gr-qc]].

[9] L. Ford, "Gravitational Particle Creation and Inflation," Phys. Rev. D **35**, 2955 (1987) doi:10.1103/PhysRevD.35.2955

[10] N. Birrell and P. Davies, “Quantum Fields in Curved Space,” doi:10.1017/CBO9780511622632

[11] L. E. Parker and D. Toms, “Quantum Field Theory in Curved Spacetime,” doi:10.1017/CBO9780511813924

[12] V. Mukhanov and S. Winitzki, “Introduction to quantum effects in gravity.”

[13] U. Yajnik, “Gravitational Particle Production in Inflation: A Fresh Look,” Phys. Lett. B **234**, 271-275 (1990) doi:10.1016/0370-2693(90)91926-3

[14] T. Damour and A. Vilenkin, “String theory and inflation,” Phys. Rev. D **53**, 2981-2989 (1996) doi:10.1103/PhysRevD.53.2981 [arXiv:hep-th/9503149 [hep-th]].

[15] D. Glavan, T. Prokopec and V. Prymidis, “Backreaction of a massless minimally coupled scalar field from inflationary quantum fluctuations,” Phys. Rev. D **89**, no.2, 024024 (2014) doi:10.1103/PhysRevD.89.024024 [arXiv:1308.5954 [gr-qc]].

[16] H. Aoki, S. Iso and Y. Sekino, “Evolution of vacuum fluctuations generated during and before inflation,” Phys. Rev. D **89**, no.10, 103536 (2014) doi:10.1103/PhysRevD.89.103536 [arXiv:1402.6900 [hep-th]].

[17] H. Aoki and S. Iso, “Evolution of Vacuum Fluctuations of an Ultra-Light Massive Scalar Field generated during and before Inflation,” PTEP **2015**, no.11, 113E02 (2015) doi:10.1093/ptep/ptv150 [arXiv:1411.5129 [gr-qc]].

[18] W. Pauli and F. Villars, “On the Invariant regularization in relativistic quantum theory,” Rev. Mod. Phys. **21**, 434-444 (1949) doi:10.1103/RevModPhys.21.434

[19] S. Weinberg, “Ultraviolet Divergences in Cosmological Correlations,” Phys. Rev. D **83**, 063508 (2011) doi:10.1103/PhysRevD.83.063508 [arXiv:1011.1630 [hep-th]].

[20] A. Borde, A. H. Guth and A. Vilenkin, “Inflationary space-times are incomplete in past directions,” Phys. Rev. Lett. **90**, 151301 (2003) doi:10.1103/PhysRevLett.90.151301 [arXiv:gr-qc/0110012 [gr-qc]].

[21] L. P. Grishchuk and Y. V. Sidorov, “Squeezed quantum states of relic gravitons and primordial density fluctuations,” Phys. Rev. D **42**, 3413-3421 (1990) doi:10.1103/PhysRevD.42.3413

[22] W. Israel, “Singular hypersurfaces and thin shells in general relativity,” Nuovo Cim. B **44S10**, 1 (1966) doi:10.1007/BF02710419

[23] A. D. Linde, “Chaotic Inflation,” Phys. Lett. B **129**, 177-181 (1983) doi:10.1016/0370-2693(83)90837-7

[24] Y. Akrami *et al.* [Planck], “Planck 2018 results. X. Constraints on inflation,” Astron. Astrophys. **641**, A10 (2020) doi:10.1051/0004-6361/201833887 [arXiv:1807.06211 [astro-ph.CO]].

[25] P. B. Greene, L. Kofman, A. D. Linde and A. A. Starobinsky, “Structure of resonance in preheating after inflation,” Phys. Rev. D **56**, 6175-6192 (1997) doi:10.1103/PhysRevD.56.6175 [arXiv:hep-ph/9705347 [hep-ph]].

[26] A. D. Linde, “Particle physics and inflationary cosmology,” Contemp. Concepts Phys. **5**, 1-362 (1990) [arXiv:hep-th/0503203 [hep-th]].

[27] C. Armendariz-Picon, "On the Expected Production of Gravitational Waves During Preheating," JCAP **08**, 012 (2019) doi:10.1088/1475-7516/2019/08/012 [arXiv:1905.05724 [astro-ph.CO]].

[28] L. Parker and S. A. Fulling, "Adiabatic regularization of the energy momentum tensor of a quantized field in homogeneous spaces," Phys. Rev. D **9**, 341-354 (1974) doi:10.1103/PhysRevD.9.341

[29] T. S. Bunch and P. C. W. Davies, "Quantum Field Theory in de Sitter Space: Renormalization by Point Splitting," Proc. Roy. Soc. Lond. A **360**, 117-134 (1978) doi:10.1098/rspa.1978.0060

[30] B. Allen and A. Folacci, "The Massless Minimally Coupled Scalar Field in De Sitter Space," Phys. Rev. D **35**, 3771 (1987) doi:10.1103/PhysRevD.35.3771

[31] H. Epstein, V. Glaser and A. Jaffe, "Nonpositivity of energy density in Quantized field theories," Nuovo Cim. **36**, 1016 (1965) doi:10.1007/BF02749799

[32] NIST Digital Library of Mathematical Functions. `http://dlmf.nist.gov/`, Release 1.1.5 of 2022-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[33] I. L. Shapiro, J. Sola and H. Stefancic, "Running G and Lambda at low energies from physics at M(X): Possible cosmological and astrophysical implications," JCAP **01**, 012 (2005) doi:10.1088/1475-7516/2005/01/012 [arXiv:hep-ph/0410095 [hep-ph]].

[34] K. Kirsten and J. Garriga, "Massless minimally coupled fields in de Sitter space: O(4) symmetric states versus de Sitter invariant vacuum," Phys. Rev. D **48**, 567-577 (1993) doi:10.1103/PhysRevD.48.567 [arXiv:gr-qc/9305013 [gr-qc]].

[35] N. Aghanim *et al.* [Planck], "Planck 2018 results. VI. Cosmological parameters," Astron. Astrophys. **641**, A6 (2020) [erratum: Astron. Astrophys. **652**, C4 (2021)] doi:10.1051/0004-6361/201833910 [arXiv:1807.06209 [astro-ph.CO]].

[36] V. Poulin, T. L. Smith, T. Karwal and M. Kamionkowski, "Early Dark Energy Can Resolve The Hubble Tension," Phys. Rev. Lett. **122**, no.22, 221301 (2019) doi:10.1103/PhysRevLett.122.221301 [arXiv:1811.04083 [astro-ph.CO]].

[37] K. Jedamzik, L. Pogosian and G. B. Zhao, "Why reducing the cosmic sound horizon alone can not fully resolve the Hubble tension," Commun. in Phys. **4**, 123 (2021) doi:10.1038/s42005-021-00628-x [arXiv:2010.04158 [astro-ph.CO]].

[38] T. L. Smith, V. Poulin and M. A. Amin, "Oscillating scalar fields and the Hubble tension: a resolution with novel signatures," Phys. Rev. D **101**, no.6, 063523 (2020) doi:10.1103/PhysRevD.101.063523 [arXiv:1908.06995 [astro-ph.CO]].

[39] E. Belgacem and T. Prokopec, "Quantum origin of dark energy and the Hubble tension," Phys. Lett. B **831**, 137174 (2022) doi:10.1016/j.physletb.2022.137174 [arXiv:2111.04803 [astro-ph.CO]].

[40] W. Hu, R. Barkana and A. Gruzinov, "Cold and fuzzy dark matter," Phys. Rev. Lett. **85**, 1158-1161 (2000) doi:10.1103/PhysRevLett.85.1158 [arXiv:astro-ph/0003365 [astro-ph]].

[41] J. S. Dowker and R. Critchley, "Effective Lagrangian and energy-momentum Tensor in de Sitter Space," Phys. Rev. D **13**, 3224 (1976) doi:10.1103/PhysRevD.13.3224

[42] L. Kofman, A. D. Linde and A. A. Starobinsky, "Towards the theory of reheating after inflation," Phys. Rev. D **56**, 3258-3295 (1997) doi:10.1103/PhysRevD.56.3258 [arXiv:hep-ph/9704452 [hep-ph]].

[43] A. Vilenkin and L. H. Ford, "Gravitational Effects upon Cosmological Phase Transitions," Phys. Rev. D **26**, 1231 (1982) doi:10.1103/PhysRevD.26.1231