PHYSICAL REVIEW D **89,** 104038 (2014)

# Instability of global de Sitter space to particle creation

Paul R. Anderson*

*Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27109, USA*

Emil Mottola†

*Theoretical Division, MS B285, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

(Received 30 September 2013; published 23 May 2014)

We show that global de Sitter space is unstable to particle creation, even for a massive free field theory with no self-interactions. The $O(4,1)$ de Sitter invariant state is a definite phase coherent superposition of particle and antiparticle solutions in both the asymptotic past and future and, therefore, is not a true vacuum state. In the closely related case of particle creation by a constant, uniform electric field, a time symmetric state analogous to the de Sitter invariant one is constructed, which is also not a stable vacuum state. We provide the general framework necessary to describe the particle creation process, the mean particle number, and dynamical quantities such as the energy-momentum tensor and current of the created particles in the de Sitter and electric field backgrounds respectively in real time, establishing the connection to kinetic theory. We compute the energy-momentum tensor for adiabatic vacuum states in de Sitter space initialized at early times in global $\mathbb{S}^3$ sections and show that particle creation in the contracting phase results in exponentially large energy densities at later times, necessitating an inclusion of their backreaction effects and leading to large deviation of the spacetime from global de Sitter space before the expanding phase can begin.

DOI: 10.1103/PhysRevD.89.104038 PACS numbers: 04.62.+v, 95.36.+x, 98.80.Jk, 98.80.Qc

## I. INTRODUCTION

The problem of vacuum zero-point energy and its effects on the curvature of space through Einstein's equations has been present in quantum theory since its inception, and was first recognized by Pauli [1]. Largely ignored and bypassed during the steady stream of successes of quantum mechanics and then quantum field theory (QFT) over a remarkable range of scales and conditions for five decades, the role of vacuum energy was raised to prominence by cosmological models of inflation. Inflation postulates a large vacuum energy density to drive exponential expansion of the universe, and invokes quantum fluctuations in the de Sitter epoch as the primordial seeds of density fluctuations that give rise both the observed cosmic microwave background anisotropies, and the formation of all observed structure in the universe [2]. The problem of quantum vacuum energy and the origin of structure are both strong motivations for the study of QFT in de Sitter space.

Further motivation comes from the discovery of dark energy in 1998 by measurements of the redshifts of distant type Ia supernovae [3]. This has led to the realization that cosmological vacuum energy may be some 70% of the energy density in the universe and be responsible for its accelerated Hubble expansion today. If correct, this implies that de Sitter space is actually a better approximation than flat Minkowski space to the geometry of the present observable universe. Accounting for the value of the apparent vacuum energy density today and elucidating its true nature and possible dynamics is widely viewed as one of the most important challenges at the intersection of quantum physics and gravitation theory, with direct relevance for observational cosmology.

Being a maximally symmetric solution of Einstein's equations with positive cosmological constant, which itself can be regarded as the energy of the vacuum, de Sitter space is the simplest setting for posing questions about the interplay of QFT, gravity, and cosmology. Progress toward a consistent theory of quantum vacuum energy and its gravitational effects, and the formation of structure in the universe almost certainly requires a thorough understanding of quantum processes in de Sitter space.

Although global de Sitter space is clearly an idealization, it is an important one amenable to exact analysis of quantum effects, which can serve as a basis for applications to cosmology. Current cosmological models of inflation and the late time expansion of the universe make use only of the expanding Poincaré patch of de Sitter space. However, de Sitter space is a homogeneous space, all points of which are on the same footing *a priori*. Hence one would expect that quantum processes taking place in global de Sitter space will have appropriate analogs in any sufficiently large coordinate patch of de Sitter space. One of the most basic of quantum processes that arise in curved spacetimes is the spontaneous creation of particles from the vacuum [4–8]. This process converts vacuum energy to ordinary matter and radiation, and therefore can

*anderson@wfu.edu
†emil@lanl.gov

 © 2014 American Physical Society

lead to the dynamical relaxation of vacuum energy with time [9,10]. An introduction to these quantum effects, summary of the earlier literature and the prospects for a dynamical theory of vacuum energy based on conversion of vacuum energy to particles may be found in [11]. Since that review, there has been further interesting work on various aspects of QFT in de Sitter space, particularly on quantum infrared and interaction effects [12,13].

Because of the mathematical appeal of maximal symmetry, much of both the earlier and more recent work assumes the stability of the de Sitter invariant state obtained by continuation from the Euclidean $\mathbb{S}^4$. This state of maximal $O(4,1)$ symmetry, known as the Chernikov-Tagirov or Bunch-Davies (CTBD) state [14–16], is also the one most often considered in inflationary models [2]. However the CTBD state raises a number of questions that seem still to require clarification. In flat Minkowski space the separation into positive and negative energies, hence particle and antiparticle solutions of any Lorentz invariant wave equation is itself Lorentz invariant. Minimizing the conserved Hamiltonian in flat space in any inertial frame produces a vacuum state that is invariant under the full Poincarè group. In de Sitter space there is no conserved Hamiltonian with a spectrum bounded from below available for this minimization [14]. Hence the dynamical stability of the maximally $O(4,1)$ symmetric CTBD state and the definition of the "vacuum" itself cannot be taken for granted. Since a freely falling detector in de Sitter space in the CTBD state will detect a nonzero, thermal distribution of particles at the Hawking–de Sitter temperature [17], the CTBD state is clearly not a vacuum state at all in the usual sense familiar from Minkowski spacetime.

Prior investigations of the stability of the CTBD state have addressed the late time behavior of the stress tensor or correlation functions. In [18] it was proven that for all fourth order adiabatic states that lead to a homogeneous and isotropic stress-energy tensor in the expanding spatially flat section of de Sitter space, the components of the renormalized stress-energy tensor expectation value $\langle T^a{}_b\rangle$ for a free scalar field asymptotically approach the CTBD value. In this sense the CTBD state is a late time attractor state for $\langle T^a{}_b\rangle$ in de Sitter space for a free, massive scalar QFT with $m^2 + \xi R > 0$. More recently in [19,20] it has been argued that any correlation function of an interacting massive scalar field theory also approaches the expected CTBD value at late times at any order of perturbation theory. This attractor behavior is clearly a result of the cosmological redshift of the de Sitter expansion.

Not addressed in any of these previous investigations are the effects of states other than the CTBD state on the early or intermediate time behavior of the stress-energy tensor or correlation functions in global de Sitter space, represented by the hyperboloid of Fig. 15 of the Appendix. It is only this geodesically complete full de Sitter hyperboloid that has the maximal symmetry group $O(4,1) = \mathbb{Z}_2 \otimes SO(4,1)$, including also the $\mathbb{Z}_2$ discrete reflection symmetry which maps all points on the hyperboloid to their antipodal points, cf. (A4). In spatially closed Robertson-Walker coordinates (A8)–(A9), the global de Sitter geometry begins at infinite size, contracts down to a minimum spatial radius, and then expands again time symmetrically to infinite size. Because the previously found attractor behavior is a consequence of the cosmological redshift as the universe expands, the same perturbations experience a cosmological blueshift in the contracting half of de Sitter space, and one should expect $\langle T^a{}_b\rangle$ or arbitrary correlation functions defined at some early initial time to deviate more and more from their CTBD values as the contraction proceeds. Thus a very small difference in the energy density from the CTBD state can become substantially magnified in the contracting phase of de Sitter space, and one should expect exactly the opposite of the attractor behavior found in the expanding phase. If enough magnification of these de Sitter breaking effects occurs and their backreaction effects are taken into account, then the universe may never reach the symmetric point, at which contraction ceases and expansion begins, and instead may evolve in a completely different way from de Sitter space for its entire future.

We treat the issue of perturbations of the CTBD state at early initial times in the contracting phase explicitly in an accompanying paper [21]. In this paper we show that the status of the vacuum in de Sitter space is very much analogous to that of a charged quantum field in the presence of a constant, uniform electric field $\mathbf{E} = E\hat{\mathbf{z}}$. Such an idealized static background field is completely invariant under time reversal and time translations. Yet in this case, as first shown by Schwinger, there can be little doubt that the vacuum is unstable to the spontaneous creation of charged particle/antiparticle pairs [22]. This spontaneous process breaks the time reversal and translational symmetry of the background, and leads to a positive imaginary part for the effective action of charged matter in the electric field background. Mathematically, this imaginary part is a consequence of the $m^2 \to m^2 - i0^+$ prescription in the Schwinger proper time treatment, or equivalently in the Feynman propagator, which distinguishes positive and negative frequency solutions as particles and antiparticles, respectively. It is this analytic continuation in mass (*not* global symmetries of the background or Euclidean continuation) that provides the physical definition of the vacuum and particle concepts for QFT in persistent background fields, including gravitational fields [23,24].

It is important to note that this prescription and Schwinger's calculation of the decay rate per unit volume of a constant, uniform electric field and hence its instability to particle creation already applies at the level of a noninteracting QFT. While interactions and the behavior of multiple point correlation functions are clearly important for the subsequent evolution of the created pairs, the instability of the background to creation of those particle/antiparticle pairs in the first place requires only the interaction of the quantum field with a classical background. In de Sitter space the absence of any minimum energy vacuum state, as well as the spontaneous particle creation rate calculated some time ago [9], analogous to that in a constant, uniform electric field, is evidence for the analogous instability of global de Sitter space to particle creation. We review that calculation in this paper and show that it can be fully justified by the proper definition of time dependent adiabatic Hadamard vacuum states. These adiabatic vacuum states provide the basis for describing particle excitations and their interactions in the transition to kinetic theory. The definition of these adiabatic vacua and particle production is restricted to massive fields, whereas it is apparent that the effects of light fields and gravitons are more subtle [25–27]. Settling the basic question of vacuum instability to massive particle creation of free QFT in de Sitter space is a necessary first step, and clearly relevant to the fundamental problem of cosmological vacuum energy, and its ultimate fate in a full quantum theory.

Although the problem of charged particle creation in an external electric field has a long history [22,28–36], it is quite illustrative, and the features directly relevant to de Sitter space are worth emphasizing. In particular, one can find a completely time symmetric state in a constant, uniform electric field background which has exactly zero decay rate by time reversal symmetry, and is thus the close analog of the maximally symmetric CTBD state in de Sitter space. The construction of such a time symmetric state, however appealing it may be mathematically, is an artificial coherent superposition of particle and antiparticle waves, which assures neither its nature as a true vacuum state, nor its stability. The spontaneous particle creation process in an electric field leads to an electric current that grows linearly with time and whose backreaction on the classical background electric field must eventually be taken into account [34,36].

Guided by the electric field analog, our main purpose in this paper is to present a detailed description of particle creation in de Sitter space in real time, extending and deepening previous analyses of its instability to particle creation [9], and computing the energy-momentum tensor of the created particles. Our study consists of two distinct but related parts. In the first part, Secs. II–III we compute the rate of the particle production in de Sitter space, and in Sec. IV review the analogous calculation for a constant uniform electric field. The standard Feynman-Schwinger prescription of particle excitations moving forward in time with negative energy modes interpreted as antiparticles propagating backwards in time provides the framework for defining $|in\rangle$ and $|out\rangle$ vacuum states in the infinite past $I_-$ and infinite future $I_+$ of de Sitter space. With this definition, it becomes evident that particles are created spontaneously, and the overlap $|\langle in|out\rangle|^2$ provides the vacuum decay probability, just as it does in the electric field analog. Both cases involve infinite time intervals in which the constant electric or gravitational field acts, with the result that $|\langle in|out\rangle|^2 = \exp(-\Gamma V_4) \to 0$, as the four-volume $V_4 \to \infty$. The finite decay rate $\Gamma$ for de Sitter space is obtained by a physical argument relating a cutoff in momentum mode sums to the cutoff in finite four-volume $V_4$ as both tend to infinity.

The infinite $V_4$ limit is somewhat subtle, and taken literally leads to non-Hadamard $|in\rangle$ and $|out\rangle$ states with zero overlap, corresponding to an infinite amount of particle creation, termed "pathological" in [8]. In fact, this should be expected for a persistent background producing particles at a finite *rate* per unit volume. To compute this finite rate of particle production rigorously, in the second part of the paper beginning in Sec. V, we define adiabatic Hadamard vacuum states and specify initial data for the mode functions on Cauchy hypersurfaces at a *finite* time. Then the time dependent Bogoliubov coefficients that describe the particle production mode by mode are computed and it is shown that in the infinite time, infinite $V_4$ limit they approach the time independent Bogoliubov coefficients connecting the $|in\rangle$ and $|out\rangle$ states. By thereby exposing the anatomy of particle production in real time, the infinite momentum and infinite time limits are shown not to commute, and the physical cutoff of the previous calculation is justified. This more careful treatment involving only UV finite adiabatic states removes a possible technical objection against the more heuristic approach of [9], reviewed in Sec. III. In addition, the second approach allows the finite but exponentially growing $\langle T^a{}_b\rangle$ in the contracting phase of de Sitter space to be computed, and its prospective large backreaction effects on the classical geometry to be estimated. This approach also makes possible a detailed investigation of the time dependence of the energy density, showing it to become quickly dominated by the particle production term in the contracting phase.

By investigating the particle creation process in real time and computing the corresponding energy momentum of the particles, the vacuum instability of global maximally extended de Sitter space to particle creation, the breaking of both time reversal and global de Sitter symmetries, and the necessity to include backreaction of the particles on the geometry become clear. The nature of the CTBD state as a particular coherent squeezed state combination of particle and antiparticle excitations is also clarified.

In the case of a uniform electric field, the Feynman-Schwinger prescription is known also to be equivalent to an adiabatic prescription of switching the electric field on and off again smoothly on a time scale $T$, evaluating the particles present in the final field free *out* region starting with the well-defined Minkowski vacuum in the initial field free *in* region [28,30,35]. We present evidence that for the analogous gentle enough switching on of the de Sitter phase from an initially static Einstein universe phase, for large values of $T$ and for modes with small enough momenta, the initial state produced for these modes is the asymptotic $|in\rangle$ state defined previously in eternal de Sitter space. Particle creation takes place for these modes after adiabatic switching on of de Sitter space and the particle spectrum produced is the same as in global de Sitter space.

The paper is organized as follows. In the next section we define the CTBD state for a non-self-interacting scalar quantum field theory in de Sitter space, fixing notation. In Sec. III we formulate the problem of particle creation in de Sitter space in terms of a one-dimensional time-independent scattering problem, review the construction of the $|in\rangle$ and $|out\rangle$ vacuum states, the nontrivial Bogoliubov transformation between them, and the de Sitter decay rate this implies. We show that the de Sitter invariant CTBD state is a definite phase coherent squeezed superposition of particle and antiparticle solutions in both the past and the future and therefore not a true vacuum state. In Sec. IV we digress to consider the case of particle creation in a constant, uniform electric field, showing the close analogy to the de Sitter case. In Sec. V we define the UV finite adiabatic states necessary to interpolate between the asymptotic $|in\rangle$ and $|out\rangle$ states and describe the particle creation process, the adiabatic particle number, and physical quantities such as the current and energy-momentum tensor of the created particles in real time. In Sec. VI we apply this adiabatic framework to global de Sitter space with closed spatial $\mathbb{S}^3$ sections, showing how it may be used to describe particle creation events in real time. In Sec. VII we show how these same methods may be applied equally well in the Poincaré coordinates of de Sitter space with flat spatial sections most often used in cosmology. In Sec. VIII we derive the form of the energy-momentum tensor for vacuum states set at finite initial times in the global $\mathbb{S}^3$ sections, and show for early initial times that particle creation in the contracting phase leads to exponentially large energy densities before the expanding phase even begins. In Sec. IX we present numerical results for the adiabatic turning on and off of de Sitter curvature on a time scale $T$ starting from a static space and back, showing that the $|in\rangle$ state is produced in this way and particle production proceeds just as in global de Sitter space. Section X contains a summary and discussion of our results. There is one Appendix which contains reference formulas for the de Sitter geometry and coordinates, included for completeness. The reader interested primarily in the results for de Sitter space may proceed from Sec. III directly to Secs. VI–VIII and the summary and discussion.

This paper may be read in conjunction with the closely related paper [21]. This first paper focuses almost exclusively on scalar particle creation and the resulting vacuum instability, while the second considers the instability to perturbations in a more general context, independent of specific fields or particle definitions, emphasizing instead the role of the effective action of the conformal anomaly and the behavior of the stress tensor derived from it. The more general analysis based on the anomaly makes it possible to draw more general conclusions about vacuum instability, the importance of inhomogeneous perturbations and sensitivity to initial conditions in de Sitter space and cosmology.

## II. WAVE EQUATION AND DE SITTER INVARIANT STATE

We consider in this paper a scalar field $\Phi$ with mass $m$ and conformal curvature coupling $\xi = \frac{1}{6}$ which in an arbitrary curved spacetime satisfies the free wave equation

$$(-\Box + M^2)\Phi \equiv \left[-\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^a}\left(\sqrt{-g}g^{ab}\frac{\partial}{\partial x^b}\right) + M^2\right]\Phi = 0 \tag{2.1}$$

with effective mass $M^2 \equiv m^2 + \xi R$. In cosmological spacetimes with $\mathbb{S}^3$ closed spatial sections the Robertson-Walker line element is

$$ds^2 = -d\tau^2 + a^2 d\Sigma^2. \tag{2.2}$$

Here $d\Sigma^2 = d\hat{N}\cdot d\hat{N}$ denotes the line element on $\mathbb{S}^3$ and $\hat{N}$, defined by Eq. (A10), is a unit vector on $\mathbb{S}^3$. Specializing to de Sitter spacetime and defining the dimensionless cosmological time $u \equiv H\tau$, the scale factor is $a(u) = H^{-1}\cosh u$, the Ricci scalar $R = 12H^2$ is a constant, and the effective mass is

$$M^2 = m^2 + 12\xi H^2 = m^2 + 2H^2 \tag{2.3}$$

for $\xi = \frac{1}{6}$.

The wave equation (2.1) is separable in coordinates (2.2) with solutions of the form $\Phi = y_k(u)Y_{klm_l}(\hat{N})$ and $Y_{klm_l}(\hat{N})$ a spherical harmonic on $\mathbb{S}^3$ given explicitly in [21]. The equation for $y_k(u)$ is

$$\left[\frac{d^2}{du^2} + 3\tanh u\frac{d}{du} + (k^2-1)\,\mathrm{sech}^2 u + \left(\frac{9}{4}+\gamma^2\right)\right]y_k = 0 \tag{2.4}$$

with the dimensionless parameter $\gamma$ defined by

$$\gamma \equiv \sqrt{\frac{M^2}{H^2} - \frac{9}{4}} = \sqrt{\frac{m^2}{H^2} - \frac{1}{4}} \tag{2.5}$$

with the latter expression valid for conformal coupling. The range of the integers $k = 1, 2, \ldots$ is taken to be strictly positive, so that the constant harmonic function on $\mathbb{S}^3$ corresponds to $k = 1$, conforming to the notation of [18] and [37]. In some works the sign under the square root in (2.5) is reversed and the quantity $\nu = i\gamma$ is defined, which is real for $M^2 \leq \frac{9}{4}H^2$ [8]. In this paper we shall be interested mainly in the massive case $M^2 > \frac{9}{4}H^2$ (the *principal series* representation of the de Sitter group) where $\gamma$ is real and positive, and for simplicity, in the case of conformal coupling $\xi = \frac{1}{6}$, so that $m^2 > \frac{1}{4}H^2$.

With the change of variables $z = (1 - i\sinh u)/2$, the mode equation (2.4) can be recast in the form of the hypergeometric equation. The fundamental complex solution $y_k = \upsilon_{k\gamma}(u)$ may be taken to be

$$\upsilon_{k\gamma}(u) \equiv Hc_{k\gamma}(\mathrm{sech}u)^{k+1}(1 - i\sinh u)^k \times F\left(\frac{1}{2} + i\gamma, \frac{1}{2} - i\gamma; k+1; \frac{1 - i\sinh u}{2}\right), \tag{2.6}$$

where $F \equiv {}_2F_1$ is the Gauss hypergeometric function and

$$c_{k\gamma} \equiv \frac{1}{k!}\left[\frac{\Gamma(k + \frac{1}{2} + i\gamma)\Gamma(k + \frac{1}{2} - i\gamma)}{2}\right]^{\frac{1}{2}} = \frac{|\Gamma(k + \frac{1}{2} + i\gamma)|}{\sqrt{2}\Gamma(k+1)} \tag{2.7}$$

is a real normalization constant, fixed so that $\upsilon_{k\gamma}$ satisfies the Wronskian condition

$$iHa^3(u)\left[\upsilon^*_{k\gamma}\frac{d}{du}\upsilon_{k\gamma} - \upsilon_{k\gamma}\frac{d}{du}\upsilon^*_{k\gamma}\right] = 1 \tag{2.8}$$

for all $k$. Note that under time reversal $u \to -u$ the mode function (2.6) goes to its complex conjugate

$$\upsilon_{k\gamma}(-u) = \upsilon^*_{k\gamma}(u) \tag{2.9}$$

for all $M^2 > 0$.

The scalar field operator $\Phi$ can be expressed as a sum over the fundamental solutions

$$\Phi(u, \hat{N}) = \sum_{k=1}^{\infty}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}\{a^{\upsilon}_{klm_l}\upsilon_{k\gamma}(u)Y_{klm_l}(\hat{N}) + a^{\upsilon\dagger}_{klm_l}\upsilon^*_{k\gamma}(u)Y^*_{klm_l}(\hat{N})\}, \tag{2.10}$$

with the Fock space operator coefficients $a^{\upsilon}_{klm_l}$ satisfying the commutation relations

$$[a^{\upsilon}_{klm_l}, a^{\upsilon\dagger}_{k'l'm'_l}] = \delta_{kk'}\delta_{ll'}\delta_{m_lm'_l}. \tag{2.11}$$

With (2.8), (2.11) and the standard unit normalization of harmonic functions on the unit sphere

$$\int_{\mathbb{S}^3} d^3\Sigma\, Y^*_{k'l'm'_l}Y_{klm_l} = \delta_{k'k}\delta_{l'l}\delta_{m'_lm_l} \tag{2.12}$$

the canonical equal time field commutation relation

$$[\Phi(u, \hat{N}), \Pi(u, \hat{N}')] = i\delta_\Sigma(\hat{N}, \hat{N}') \tag{2.13}$$

is satisfied, where $\Pi = \sqrt{-g}\dot{\Phi} = Ha^3\frac{\partial\Phi}{\partial u}$ is the field momentum operator conjugate to $\Phi$, the overdot denotes differentiation $H\partial/\partial u$ and $\delta_\Sigma(\hat{N}, \hat{N}')$ denotes the delta function on the unit $\mathbb{S}^3$ with respect to the canonical round metric $d\Sigma^2$.

The Chernikov-Tagirov or Bunch-Davies (CTBD) state $|\upsilon\rangle$ [14–16] defined by

$$a^{\upsilon}_{klm_l}|\upsilon\rangle = 0 \quad \forall\ k, l, m_l \tag{2.14}$$

is invariant under the full $O(4,1)$ isometry group of the complete de Sitter manifold (A1)–(A2), including under the $\mathbb{Z}_2$ discrete inversion symmetry of all coordinates in the embedding space, $X^A \to -X^A$, cf. Eq. (A4), which is not continuously connected to the identity. This de Sitter invariant state $|\upsilon\rangle$ is the one usually discussed in the literature, and the Green functions in this state are those obtained by analytic continuation to de Sitter spacetime from the Euclidean $\mathbb{S}^4$ with full $O(5)$ symmetry. The existence of an $O(4,1)$ invariant symmetric state does not imply that this state is a stable vacuum. In this and a closely related paper [21], we shall present the evidence that it is not.

## III. DE SITTER SCATTERING POTENTIAL, IN AND OUT STATES, AND DECAY RATE

The solutions (2.6) and de Sitter invariant state $|\upsilon\rangle$ are defined once and for all, globally in de Sitter space without any reference to a separation between positive and negative frequencies, which is axiomatic in flat spacetime to discriminate between particle and antiparticle states, and necessary to define a stable vacuum which is a minimum of a positive definite Hamiltonian. In flat spacetime such a separation into particle and antiparticle solutions of the wave equation is defined by positive and negative frequency solutions $e^{-i\omega_k t}$ and $e^{+i\omega_k t}$, which are analytic functions of $m^2$ in the lower and upper half complex $m^2$ plane, respectively, as $t \to +\infty$. For $t \to -\infty$ the analyticity in the two halves of the complex $m^2$ plane are reversed for the same positive and negative frequency modes. Clearly these are the same simple exponential functions for all times in flat Minkowski space. The fundamental CTBD solutions (2.6) do not have this property in de Sitter space. Correspondingly, there is no positive definite Hamiltonian operator to be minimized in global de Sitter space [14]. These important differences with flat space are responsible for the nontrivial features of quantum fields and the quantum vacuum in de Sitter space.

To appreciate the sharp distinction from flat space, it is useful to eliminate the factor of $a^3$ in the Wronskian condition (2.8) by defining the mode functions

$$f_k = a^{\frac{3}{2}}y_k \tag{3.1}$$

which satisfy the equation of a time dependent harmonic oscillator

$$\frac{d^2}{d\tau^2} f_k + \Omega_k^2 f_k = 0 \tag{3.2}$$

in each $k$ mode, with the time dependent frequency given by

$$\Omega_k^2 \equiv \omega_k^2 + \left(\xi - \frac{1}{6}\right) R - \frac{\dot{h}}{2} - \frac{h^2}{4}, \qquad \omega_k^2 \equiv \frac{k^2}{a^2} + m^2. \tag{3.3}$$

Here $h \equiv \dot{a}/a$ in a general RW spacetime with line element (2.2). Specializing to de Sitter space and again using $u = H\tau$, the time dependent harmonic oscillator frequency is

$$\Omega_k^2|_{\mathrm{dS}} = H^2 \left[\left(k^2 - \frac{1}{4}\right) \mathrm{sech}^2 u + \gamma^2\right], \tag{3.4}$$

so that we may rewrite (3.2) in dimensionless form as a stationary state scattering problem,

$$\left[-\frac{d^2}{du^2} + \mathcal{U}_k(u)\right] f_{k\gamma} = \gamma^2 f_{k\gamma}, \tag{3.5}$$

with the one-dimensional effective scattering potential,

$$\mathcal{U}_k(u) \equiv -\left(k^2 - \frac{1}{4}\right) \mathrm{sech}^2 u. \tag{3.6}$$

Here the "energy" $\gamma$ is defined by (2.5) and is both real and positive for the fields with $\xi = \frac{1}{6}$ and $m^2 > \frac{1}{4} H^2$ considered in this paper.

Since the scattering potential (3.6) is negative definite, and approaches zero exponentially as $|u| \to \infty$, the solutions of (3.5) for $\gamma^2 > 0$ describe over the barrier scattering and are everywhere oscillatory. The vanishing of the potential at large $|u|$ implies well-defined free asymptotic solutions as $u \to \mp\infty$, behaving like $e^{\pm i\gamma u}$. Because of the scattering by the potential, a positive frequency wave $e^{-i\gamma u}$ incident from the left (the past as $u \to -\infty$) will be partially transmitted to a positive frequency $e^{-i\gamma u}$ wave to the right (the future as $u \to +\infty$) and partially reflected to a negative frequency $e^{i\gamma u}$ wave to the left. Potential scattering of this kind and mixing of positive and negative energy solutions clearly does not occur in static spacetimes such as Minkowski spacetime.

Now the crucial point is that the asymptotic pure frequency scattering solutions behaving as $e^{-i\gamma u}$ have the required analyticity in $m^2$ to correspond exactly to the Feynman prescription of positive energy solutions as particles propagating forward in time, while $e^{+i\gamma u}$ are negative energy solutions corresponding to antiparticles propagating backward in time [24,38]. This leads to the covariant definition of the Feynman propagator as the boundary value of a function defined in the complex $m^2$ plane with the $m^2 - i0^+$ prescription specifying the limit in which the real axis is approached and pole contributions evaluated. This definition is easily generalized to nonvanishing background fields and curved spacetimes by the same generally covariant $m^2 - i0^+$ prescription, and is then completely equivalent to the Schwinger-DeWitt proper time method of defining the propagator and effective action functional in such situations [22,23]. This gives a rigorous definition of particles and antiparticles whenever the solutions of the time dependent mode equation (3.2) behave as pure oscillating exponential functions in the asymptotic past and the asymptotic future. This definition is physically based on the corresponding definitions in Minkowski space [24], free of any assumptions of analytic continuation from Euclidean time or $\mathbb{S}^4$, and generally quite different from that prescription. It is also the Feynman-Schwinger $m^2 - i0^+$ definition of the propagator, and *only* that definition that satisfies the composition rule for amplitudes defined by a single path integral [39]. This should be clear from the fact that only pure positive (or pure negative) frequency exponentials can satisfy the composition rule $e^{iS_{AC}} \sim \sum_B e^{iS_{AB}} e^{iS_{BC}}$ in the Feynman path integral, and that the composition rule will generally fail if superpositions of $e^{\pm iS}$ appear in the single particle proper time representation of the Feynman Green's function. Finally, in Sec. IX we provide evidence that this definition of asymptotic particle and antiparticle solutions to the wave equation is the also the unique one produced by adiabatically switching the background gravitational field on and off.

Let us therefore denote by $f_{k\gamma(+)}(u)$ the properly normalized positive frequency solution of (3.5) which behaves as $e^{-i\gamma u}$ as $u \to -\infty$ (the particle *in* solution), and by $f_{k\gamma}^{(+)}(u)$ the properly normalized positive frequency solution of (3.5) which behaves as $e^{-i\gamma u}$ as $u \to +\infty$ (the particle *out* solution). The corresponding negative frequency (or antiparticle) solutions $f_{k\gamma(-)}(u)$ and $f_{k\gamma}^{(-)}(u)$ which behave as $e^{i\gamma u}$ as $u \to \mp\infty$, respectively, are obtained from these by complex conjugation. Moreover since the potential $\mathcal{U}_k(u)$ is real and even under $u \to -u$, we have

$$f_{k\gamma(-)}(u) = [f_{k\gamma(+)}(u)]^* = f_{k\gamma}^{(+)}(-u) \tag{3.7a}$$

$$f_{k\gamma}^{(-)}(u) = [f_{k\gamma}^{(+)}(u)]^* = f_{k\gamma(+)}(-u), \tag{3.7b}$$

by which any one of the four solutions determines the other three. The proper normalization condition for each set of modes analogous to (2.8) is

$$iH\left(f_{k\gamma(+)}^* \frac{d}{du} f_{k\gamma(+)} - f_{k\gamma(+)} \frac{d}{du} f_{k\gamma(+)}^*\right) = 1$$
$$= iH\left(f_{k\gamma}^{(+)*} \frac{d}{du} f_{k\gamma}^{(+)} - f_{k\gamma}^{(+)} \frac{d}{du} f_{k\gamma}^{(+)*}\right). \tag{3.8}$$

The CTBD mode function (2.6), which we may write in terms of a Legendre function [40],

$$F_{k\gamma}(u) \equiv a^{\frac{3}{2}} v_{k\gamma}(u) = e^{-\frac{ik\pi}{2}\mathrm{sgn}(u)} \frac{|\Gamma(k+\frac{1}{2}+i\gamma)|}{\sqrt{2H}} \times (\cosh u)^{\frac{1}{2}} P^{-k}_{-\frac{1}{2}+i\gamma}(i\sinh u), \tag{3.9}$$

satisfies (3.5) and the normalization (3.8) by virtue of (2.7), (2.8) and (3.1). The dependence of the phase on the sign of $u$ enters to compensate for the discontinuity of the Legendre function $P^{-k}_{-\frac{1}{2}+i\gamma}(\zeta)$ as conventionally defined with a branch cut along the real axis from $\zeta = -\infty$ to $\zeta = +1$ [40], so that $F_{k\gamma}(u)$ is in fact continuous as $u$ crosses zero.

Since the *in*, the *out*, and the CTBD mode functions together with their complex conjugates are each a complete set of solutions to (3.5), which preserve the Wronskian relation (3.8), they are expressible in terms of each other by means of a Bogoliubov transformation. Specifically, the *in* mode functions are expressible in terms of the CTBD $F_{k\gamma}$ and $F^*_{k\gamma}$ by

$$\begin{pmatrix} f_{k\gamma(+)} \\ f_{k\gamma(-)} \end{pmatrix} = \begin{pmatrix} A^{in}_{k\gamma} & B^{in}_{k\gamma} \\ B^{in*}_{k\gamma} & A^{in*}_{k\gamma} \end{pmatrix} \begin{pmatrix} F_{k\gamma} \\ F^*_{k\gamma} \end{pmatrix} \tag{3.10}$$

and likewise for the *out* mode functions,

$$\begin{pmatrix} f^{(+)}_{k\gamma} \\ f^{(-)}_{k\gamma} \end{pmatrix} = \begin{pmatrix} A^{out}_{k\gamma} & B^{out}_{k\gamma} \\ B^{out*}_{k\gamma} & A^{out}_{k\gamma} \end{pmatrix} \begin{pmatrix} F_{k\gamma} \\ F^*_{k\gamma} \end{pmatrix}. \tag{3.11}$$

Each set of the (strictly time independent) $A_{k\gamma}$ and $B_{k\gamma}$ Bogoliubov coefficients satisfies the relation

$$|A_{k\gamma}|^2 - |B_{k\gamma}|^2 = 1. \tag{3.12}$$

By using (3.7) and $F_{k\gamma}(-u) = F^*_{k\gamma}(u)$ we immediately infer the relations

$$A^{out}_{k\gamma} = A^{in*}_{k\gamma} \quad \text{and} \quad B^{out}_{k\gamma} = B^{in*}_{k\gamma} \tag{3.13}$$

between the *in* and *out* Bogoliubov coefficients. Furthermore, by inverting (3.11) and substituting the result in (3.10), we obtain

$$\begin{pmatrix} f_{k\gamma(+)} \\ f_{k\gamma(-)} \end{pmatrix} = \begin{pmatrix} A^{in}_{k\gamma} & B^{in}_{k\gamma} \\ B^{in*}_{k\gamma} & A^{in*}_{k\gamma} \end{pmatrix} \begin{pmatrix} A^{out*}_{k\gamma} & -B^{out}_{k\gamma} \\ -B^{out*}_{k\gamma} & A^{out}_{k\gamma} \end{pmatrix} \begin{pmatrix} f^{(+)}_{k\gamma} \\ f^{(-)}_{k\gamma} \end{pmatrix} = \begin{pmatrix} A^{tot}_{k\gamma} & B^{tot}_{k\gamma} \\ B^{tot}_{k\gamma} & A^{tot}_{k\gamma} \end{pmatrix} \begin{pmatrix} f^{(+)}_{k\gamma} \\ f^{(-)}_{k\gamma} \end{pmatrix} \tag{3.14}$$

which with (3.13) gives

$$A^{\rm tot}_{k\gamma} = (A^{in}_{k\gamma})^2 - (B^{in}_{k\gamma})^2 \tag{3.15a}$$

$$B^{\rm tot}_{k\gamma} = A^{in*}_{k\gamma} B^{in}_{k\gamma} - A^{in}_{k\gamma} B^{in*}_{k\gamma} \tag{3.15b}$$

for the coefficients of the total Bogoliubov transformation relating the *in* and *out* bases.

To find these Bogoliubov coefficients explicitly we construct the de Sitter scattering solutions (3.7). From the asymptotic form of the Legendre functions for large arguments [40], the pure positive frequency solutions of (3.5) as $u \to \mp\infty$ are Legendre functions of the second kind, $Q^{-k}_{-\frac{1}{2}\pm i\gamma}$. Fixing the normalization by (3.8) these exact *in* and *out* solutions of (3.5) may be taken to be

$$f_{k\gamma(+)}|_{u<0} = \frac{e^{-\frac{\pi\gamma}{2}}}{|\Gamma(\frac{1}{2}-k+i\gamma)|} \left[\frac{\cosh u}{H\sinh(\pi\gamma)}\right]^{\frac{1}{2}} Q^{-k}_{-\frac{1}{2}-i\gamma}(i\sinh u) \tag{3.16a}$$

$$f^{(+)}_{k\gamma}|_{u>0} = \frac{e^{-\frac{\pi\gamma}{2}}}{|\Gamma(\frac{1}{2}-k+i\gamma)|} \left[\frac{\cosh u}{H\sinh(\pi\gamma)}\right]^{\frac{1}{2}} Q^{-k}_{-\frac{1}{2}+i\gamma}(i\sinh u) \tag{3.16b}$$

in the indicated regions of $u$, which have the required asymptotic behaviors [41]

$$f_{k\gamma(+)} \underset{u\to-\infty}{\longrightarrow} \frac{(-)^k}{\sqrt{2H\gamma}} e^{\frac{i\pi}{4}} e^{-i\eta_{k\gamma}} e^{-i\gamma u} \tag{3.17a}$$

$$f^{(+)}_{k\gamma} \underset{u\to\infty}{\longrightarrow} \frac{(-)^k}{\sqrt{2H\gamma}} e^{-\frac{i\pi}{4}} e^{i\eta_{k\gamma}} e^{-i\gamma u}, \tag{3.17b}$$

respectively, and where the phase $\eta_{k\gamma}$ here is defined by

$$\eta_{k\gamma} \equiv \arg\left\{\Gamma(1-i\gamma)\Gamma\left(k+\frac{1}{2}+i\gamma\right)\right\}. \tag{3.18}$$

Then by using Eq. (3.9) and Eq. 3.3.1 (11) of [40] relating the Legendre functions of the second kind to those of the first kind, we obtain

$$f_{k\gamma(+)} = \frac{1}{\sqrt{2\sinh(\pi\gamma)}} (ie^{-\frac{ik\pi}{2}} e^{\frac{\pi\gamma}{2}} F_{k\gamma} + e^{\frac{ik\pi}{2}} e^{-\frac{\pi\gamma}{2}} F^*_{k\gamma}) \tag{3.19a}$$

$$f^{(+)}_{k\gamma} = \frac{1}{\sqrt{2\sinh(\pi\gamma)}} (-ie^{\frac{ik\pi}{2}} e^{\frac{\pi\gamma}{2}} F_{k\gamma} + e^{-\frac{ik\pi}{2}} e^{-\frac{\pi\gamma}{2}} F^*_{k\gamma}), \tag{3.19b}$$

which are valid for all $u$. Making use of the definitions (3.10) and (3.11), we may read off the Bogoliubov coefficients,

$$A^{in}_{k\gamma} = \frac{i}{\sqrt{2\sinh(\pi\gamma)}} e^{-\frac{ik\pi}{2}} e^{\frac{\pi\gamma}{2}} = A^{out*}_{k\gamma} \tag{3.20a}$$

$$B^{in}_{k\gamma} = \frac{1}{\sqrt{2\sinh(\pi\gamma)}} e^{\frac{ik\pi}{2}} e^{-\frac{\pi\gamma}{2}} = B^{out*}_{k\gamma}, \tag{3.20b}$$

relating the *in* and *out* scattering solutions of (3.5) to the fundamental CTBD solution in de Sitter space.

Notice that if (3.10) or (3.11) with the Bogoliubov coefficients (3.20) are inverted, then it is clear that the CTBD solution (3.9) is a very particular phase coherent superposition of positive and negative frequency solutions at both $u \to \pm\infty$ [42]. Hence the $O(4,1)$ invariant state $|v\rangle$ they define through (2.14) contains particles in both the *in* and *out* bases and is not a particle vacuum state in either limit. A direct consequence of this is that the $O(4,1)$ invariant propagator function constructed from the CTBD modes and obtained also by analytic continuation from the Euclidean $\mathbb{S}^4$ manifold contains a superposition of phase increasing and decreasing exponentials $e^{\pm iS}$, and does not obey the composition rule of a Feynman propagator function [12].

Clearly the quantization of the scalar field $\Phi$ may be formally carried out in either the *in* or *out* bases and the corresponding Fock space operators introduced as in (2.10)–(2.11) for the CTBD basis. Since there is scattering in the de Sitter potential (3.5) and the *in* and *out* states are related by a nontrivial Bogoliubov transformation (3.15), which from (3.15) and (3.20) has coefficients

$$A^{\rm tot}_{k\gamma} = (-)^{k-1}\coth(\pi\gamma) \qquad (3.21a)$$

$$B^{\rm tot}_{k\gamma} = i(-)^{k-1}{\rm csch}(\pi\gamma), \qquad (3.21b)$$

the vacuum state $|in\rangle$ defined by absence of positive frequency particle excitations at early times is different from the corresponding vacuum state $|out\rangle$ defined by the absence of positive frequency particle excitations at late times. Equivalently the early time $|in\rangle$ state contains particle excitations relative to the late time $|out\rangle$ vacuum. The mean number density of particles of the *out* basis in the vacuum state defined by the *in* basis is

$$|B^{\rm tot}_{k\gamma}|^2 = {\rm csch}^2(\pi\gamma) \qquad (3.22)$$

in the mode labeled by $(klm_l)$. Also

$$w_\gamma \equiv \left|\frac{B^{\rm tot}_{k\gamma}}{A^{\rm tot}_{k\gamma}}\right|^2 = {\rm sech}^2(\pi\gamma) \qquad (3.23)$$

is the relative probability of creating a particle/antiparticle pair in this mode. Note that both (3.22) and (3.23) are independent of $(klm_l)$, depending only upon the mass of the field and its coupling to the scalar curvature. Equivalent results were found in earlier work [9] with a different choice of the arbitrary phases for the scattering solutions and Bogoliubov coefficients.

The overlap between the *in* and *out* bases yields the probability that no particles are created, or that the vacuum remains the vacuum, and is given by

$$|\langle out|in\rangle|^2 = \prod_{klm_l}(1-w_\gamma) = \exp\left\{\sum_{klm_l}\ln\left[\tanh^2(\pi\gamma)\right]\right\}. \qquad (3.24)$$

Because the summand in the last expression is independent of $(klm_l)$, the sum is quite divergent and the overlap between the $|in\rangle$ and $|out\rangle$ states strictly vanishes. This fact has led to questions about the physical meaning and appropriateness of these states [8,43]. Questions have also been raised by the closely related fact that the Wightman and Green's functions defined in the $|in\rangle$ and $|out\rangle$ states have non-Hadamard short distance behaviors, since they are in fact two members of the $\alpha$ vacuum family of states with particular nonzero values of the parameter $\alpha$ [9,44]. Although this entire family of states are formally de Sitter invariant under the $SO(4,1)$ subgroup of $O(4,1)$ continuously connected to the identity, they are not invariant under discrete $\mathbb{Z}_2$ inversion, and the two-point Wightman correlation function in all such states other than the CTBD $\alpha = 0$ state has short distance singularities as $x \to x'$ that differ from those in flat space. This would seem to imply a sensitivity of local short distance physics to global properties of the geometry, at odds with usual expectations of renormalization and effective field theory.

These difficulties are removed once one recognizes that the divergence of the sum in (3.24) and vanishing of the overlap $|\langle out|in\rangle|^2$ are due to the infinite four-volume $V_4$ of de Sitter space, and one should ask instead about the particle production probability per unit four-volume. As we shall see by detailed analysis of the particle creation process mode by mode in real time in Sec. VI, the unphysical non-Hadamard UV behavior of the $|in\rangle$ and $|out\rangle$ states is due to the noncommutivity of the infinite time $|u| \to \infty$ (and hence infinite $V_4$) and infinite momentum $k \to \infty$ limits. The short distance or UV properties of the state rely in momentum space on the vacuum matching the flat space or zero field vacuum to sufficiently high order at sufficiently high momentum or short distances, whereas these large $k$ short distance properties are lost if the infinite time limit is taken first. Thus both technical difficulties are eliminated when one considers first a finite time interval and relates the cutoff in $k$ in (3.24) properly to the finite four-volume $V_4$ over which the particle production takes place.

The finite particle production rate $\Gamma$ can be extracted from (3.24) by the following physical considerations, which we justify more rigorously in Sec. VI. First the sums in (3.24) are regulated by introducing a cutoff in the principal quantum number at $k_{\rm max} = K$, so that

$$\sum_{k=1}^{K}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l} 1 = \frac{K(K+1)(2K+1)}{6} \to \frac{K^3}{3} \qquad (3.25)$$

for $K \gg 1$. Then one recognizes that the cutoff in the mode sum corresponds to a time dependent cutoff in physical momenta at

$$P_K(u) = \frac{K}{a} = \frac{KH}{\cosh u}. \qquad (3.26)$$

Hence, for a *fixed* physical momentum cutoff $P_K$, an increase in time by $\Delta u$ results in an increase in $K$ such that

$$\frac{\Delta K}{K} = \frac{\Delta(\cosh u)}{\cosh u} \to \mathrm{sgn}(u)\Delta u = |\Delta u| \qquad (3.27)$$

as $K$ and $|u| \to \infty$. Thus the $K$ cutoff in the sum (3.25) may be traded for a cutoff in the time interval $u$ according to

$$\ln K \leftrightarrow |u| + \mathrm{const}, \qquad (3.28)$$

where the constant is dependent upon the finite fixed $P_K$ and is unimportant in the limit $K, |u| \to \infty$. Since the four-volume enclosed by the change of $u$ is

$$\Delta V_4 = \int d^4x\sqrt{-g}|_u^{u+\Delta u} = \frac{2\pi^2}{H^4}|\Delta u| \cosh^3 u \to \frac{\pi^2}{4H^4}e^{3|u|}\frac{\Delta K}{K} \to \frac{\pi^2}{4H^4}K^2\Delta K \qquad (3.29)$$

in this limit, the change in the sum in the exponent of (3.24) as the cutoff $K$ is changed,

$$\Delta\sum_{klm_l} \ln[\tanh^2(\pi\gamma)] = -2\ln[\coth(\pi\gamma)]K^2\Delta K \to -\frac{8H^2}{\pi^2}\ln[\coth(\pi\gamma)]\Delta V_4, \qquad (3.30)$$

may be regarded as giving rise to the finite decay rate per unit four-volume according to

$$|\langle out|in\rangle|^2_{V_4} = \exp(-\Gamma V_4) \qquad (3.31)$$

as $V_4 \to \infty$, with

$$\Gamma = \frac{8H^4}{\pi^2}\ln[\coth(\pi\gamma)] \qquad (3.32)$$

the decay rate of the vacuum $|in\rangle$ state due to particle creation in de Sitter space [9]. For $m \gg H$ the decay rate goes to zero exponentially

$$\Gamma \to \frac{16H^4}{\pi^2}e^{-2\pi m/H} \quad \text{for } m \gg H, \qquad (3.33)$$

while the divergence of (3.32) at $\gamma = 0$ indicates that the case of light masses must be treated differently.

The argument leading from (3.24) to (3.32) will be justified in Sec. VI by a more careful procedure based on an analysis of the real time particle creation process in de Sitter space. This requires evolving the system from a finite initial time to a finite final time and defining time dependent adiabatic vacuum states which interpolate smoothly between the $|in\rangle$ and $|out\rangle$ states, so that the infinite time infinite $V_4$ limit is taken only at the end. The analysis of particle creation in real time introduces the momentum dependence that is absent from the asymptotic Bogoliubov coefficients (3.21) at infinite times and which justifies the replacement (3.27). For finite elapsed $|u|$ and finite enclosed $V_4$ all states are Hadamard since their properties at $k \gg K$ are undisturbed from the UV finite adiabatic vacuum. This will also enable consideration of the finite renormalized stress tensor of the created particles and their backreaction on the classical geometry. Before embarking upon that more complete treatment of the particle creation process in de Sitter space, we review the analogous case of particle creation in a constant uniform electric field, which shares many of the same features, and for which the implication of an instability is clear.

## IV. IN/OUT STATES AND DECAY RATE OF A CONSTANT UNIFORM ELECTRIC FIELD

The case of a charged quantum field in the background of a constant uniform electric field has many similarities with the de Sitter case. Although this case has been considered by many authors [22,28–31,33–36], the aspects relevant to the de Sitter case are worth re-emphasizing, including the existence of a time symmetric state analogous to the CTBD state in de Sitter space, which apparently has not received previous attention.

Treating the electric field as a classical background field analogous to the classical gravitational field of de Sitter space, the wave equation of a non-self-interacting complex scalar field $\Phi$ is

$$[-(\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu) + m^2]\Phi = 0 \qquad (4.1)$$

in the background electromagnetic potential $A_\mu(x)$. Analogous to choosing global time dependent coordinates 2.2)) or (A8) in de Sitter space, one may choose the time dependent gauge,

$$A_z = -Et, \qquad A_t = A_x = A_y = 0, \qquad (4.2)$$

in which to describe a fixed constant and uniform electric field in the $\hat{\mathbf{z}}$ direction. Then the solutions of the field equation (4.1) may be separated in Fourier modes $\Phi \sim e^{i\mathbf{k}\cdot\mathbf{x}}f_{\mathbf{k}}(t)$ with

$$\left[\frac{d^2}{dt^2} + (k_z + eEt)^2 + k_\perp^2 + m^2\right]f_{\mathbf{k}}(t) = 0. \qquad (4.3)$$

This is again the form of a time-dependent harmonic oscillator analogous to (3.2), with the frequency function now given by

$$\omega_{\mathbf{k}}(t) \equiv [(k_z + eEt)^2 + k_\perp^2 + m^2]^{\frac{1}{2}} = \sqrt{2|eE|}\sqrt{\frac{u^2}{4} + \lambda} \tag{4.4}$$

instead of (3.3) of the de Sitter case. We have defined here the dimensionless variables

$$u \equiv \sqrt{\frac{2}{|eE|}}(k_z + eEt), \qquad \lambda \equiv \frac{k_\perp^2 + m^2}{2|eE|} > 0. \tag{4.5}$$

Without loss of generality we can take the sign of $eE$ to be positive. With $f_{\mathbf{k}}(t) \to f_\lambda(u)$, the wave equation (4.3) then becomes

$$\left[\frac{d^2}{du^2} + \frac{u^2}{4} + \lambda\right] f_\lambda(u) = 0, \tag{4.6}$$

whose solutions may be expressed in terms of confluent hypergeometric functions ${}_1F_1(a;c;z)$ or parabolic cylinder functions [40]

$$D_{-i\lambda-\frac{1}{2}}(e^{\frac{i\pi}{4}}u), \qquad D_{-i\lambda-\frac{1}{2}}(-e^{\frac{i\pi}{4}}u), \qquad D_{i\lambda-\frac{1}{2}}(e^{-\frac{i\pi}{4}}u), \qquad D_{i\lambda-\frac{1}{2}}(-e^{-\frac{i\pi}{4}}u). \tag{4.7}$$

Any two of the solutions (4.7) are linearly independent for general real $\lambda$.

It is important to recognize that questions relating to the definition of particles and the proper vacuum state arise in time-dependent background electromagnetic potentials such as (4.2), which are quite analogous to the same questions arising in gravitational backgrounds such as de Sitter space. As in the de Sitter case Eq. (4.6) may viewed as a one-dimensional stationary state scattering problem for the Schrödinger equation in the inverted harmonic oscillator potential $-u^2/4$, independent of $k$ in this case, with "energy" $\lambda$ (the analog of $\gamma^2$). We again have over the barrier scattering in a potential that is even under $u \to -u$, with no turning points on the real $u$ axis and the solutions (4.7) are everywhere oscillatory for positive $\lambda$. Although the potential $-u^2/4$ grows without bound as $|u| \to \infty$, pure positive frequency *in* and *out* particle modes can be defined by the requirement that they behave as $(2\omega_{\mathbf{k}})^{-\frac{1}{2}} e^{-i\Theta_\lambda(u)}$, where the adiabatic phase $\Theta_\lambda(u)$ is defined by

$$\begin{aligned}\Theta_\lambda(u) &\equiv \int_{t(u=0)}^{t(u)} dt\, \omega_{\mathbf{k}}(t) = \frac{1}{2}\int_0^u du\sqrt{u^2 + 4\lambda} = \frac{u}{4}\sqrt{u^2 + 4\lambda} + \lambda \ln\left(\frac{u + \sqrt{u^2 + 4\lambda}}{2\sqrt{\lambda}}\right)\\ &\to \mathrm{sgn}(u)\left\{\frac{u^2}{4} + \frac{\lambda}{2}\left[\ln\left(\frac{u^2}{\lambda}\right) + 1\right]\right\} + \mathcal{O}\left(\frac{\lambda^2}{u^2}\right)\end{aligned} \tag{4.8}$$

as $|u| \to \infty$. The fact that the phase (4.8) has a well-defined asymptotic form with small corrections means that well-defined positive and negative frequency mode functions exist in the limit of large $|u|$, although the potential (4.2) does not vanish (or even remain bounded) in this limit. Examining the asymptotic form of the various parabolic cylinder functions (4.7) one easily finds the exact solutions of (4.6) which behave as pure positive frequency adiabatic solutions of (4.3) or (4.6) [28–30], namely

$$f_{\lambda(+)}(u) = (2eE)^{-\frac{1}{4}} e^{-\frac{\pi\lambda}{4}} e^{i\eta_\lambda} D_{-\frac{1}{2}+i\lambda}(-e^{-\frac{i\pi}{4}}u) \tag{4.9a}$$

$$f_\lambda^{(+)}(u) = (2eE)^{-\frac{1}{4}} e^{-\frac{\pi\lambda}{4}} e^{-i\eta_\lambda} D_{-\frac{1}{2}-i\lambda}(e^{\frac{i\pi}{4}}u), \tag{4.9b}$$

and which satisfy the Wronskian normalization condition,

$$\begin{aligned} i\left(f_{\lambda(+)}^* \frac{d}{dt} f_{\lambda(+)} - f_{\lambda(+)} \frac{d}{dt} f_{\lambda(+)}^*\right) &= 1\\ &= i\left(f_\lambda^{(+)} \frac{d}{dt} f_\lambda^{(+)} - f_\lambda^{(+)} \frac{d}{dt} f_\lambda^{(+)*}\right),\end{aligned} \tag{4.10}$$

analogous to (3.8). These *in* and *out* scattering solutions are chosen to have the simple pure positive frequency asymptotic behaviors,

$$f_{\lambda(+)} \underset{u\to-\infty}{\longrightarrow} (2\omega_{\mathbf{k}})^{-\frac{1}{2}} e^{-i\Theta_\lambda(u)} \tag{4.11a}$$

$$f_\lambda^{(+)} \underset{u\to+\infty}{\longrightarrow} (2\omega_{\mathbf{k}})^{-\frac{1}{2}} e^{-i\Theta_\lambda(u)}, \tag{4.11b}$$

provided the arbitrary constant phase $\eta_\lambda$ in (4.9) is taken to be

$$\eta_\lambda \equiv \frac{\lambda}{2} - \frac{\lambda}{2}\ln\lambda - \frac{\pi}{8}. \tag{4.12}$$

The *in* and *out* particle mode solutions (4.9) and the corresponding complex conjugate antiparticle mode solutions are a set of four solutions of (4.6) which are related to each other by the precise analog of (3.7) in the de Sitter case. Here we have chosen to incorporate the phase $\eta_\lambda$ into the definition of the modes (4.9) rather than have it appear in the asymptotic forms (4.11), as the analogous phase $\eta_{k\gamma}$ does in (3.17) of the previous section.

Now an additional point of correspondence is the existence of a $u$-time symmetric solution to (4.6) analogous to the CTBD mode solution (2.6) or (3.9) in de Sitter space, and a corresponding maximally symmetric state of the charged quantum field in a constant, uniform electric field background. That such a mode solution to (4.6) obeying

$$\upsilon_\lambda(-u) = \upsilon_\lambda^*(u) \tag{4.13}$$

exists is clear from the $u \to -u$ symmetry of the real scattering potential $-u^2/4$. Since there is no expansion factor $a(u)$ in this case, this symmetric function is also the analog of $F_{k\gamma}(u)$ (3.9) in the de Sitter case. It is most conveniently expressed in terms of the confluent hypergeometric function defined by the confluent hypergeometric series,

$$\Phi(a,c;z) \equiv {}_1F_1(a;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n}\frac{z^n}{n!},$$
$$(a)_n \equiv \frac{\Gamma(a+n)}{\Gamma(a)}, \tag{4.14}$$

or the integral representation

$$\Phi(a,c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\int_0^1 dx e^{xz} x^{a-1}(1-x)^{c-a-1},$$
$$\mathrm{Re}\, c > \mathrm{Re}\, a > 0 \tag{4.15}$$

in the form

$$\upsilon_\lambda(u) = 2^{-\frac{1}{2}}(k_\perp^2 + m^2)^{-\frac{1}{4}} e^{-\frac{iu^2}{4}}\left[\Phi\left(\frac{1}{4}+\frac{i\lambda}{2},\frac{1}{2};\frac{iu^2}{2}\right) - iu\lambda^{\frac{1}{2}}\Phi\left(\frac{3}{4}+\frac{i\lambda}{2},\frac{3}{2};\frac{iu^2}{2}\right)\right], \tag{4.16}$$

which is correctly normalized by (4.10), and satisfies (4.13) by use of the Kummer transformation of the function $\Phi(a,c;z)$, cf. Eq. 6.3 (7) of [40]. By making use of the value $\Phi(a,c;0) = 1$ from (4.14) or (4.15), we find

$$\upsilon_\lambda(0) = 2^{-\frac{1}{2}}(k_\perp^2 + m^2)^{-\frac{1}{4}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\bigg|_{u=0} \tag{4.17a}$$

$$\frac{\partial \upsilon_\lambda}{\partial t}\bigg|_{u=0} = -i\sqrt{eE\lambda}(k_\perp^2 + m^2)^{-\frac{1}{4}} = \frac{-i\omega_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}}\bigg|_{u=0}, \tag{4.17b}$$

so that the symmetric solution $\upsilon_\lambda$ matches the adiabatic positive frequency form $(2\omega_{\mathbf{k}})^{-\frac{1}{2}}e^{-i\Theta_\lambda(u)}$ at the *symmetric* point of the potential $u=0$, halfway in between the asymptotic limits $u \to \pm\infty$. The solution of (4.6) with these properties is unique.

The existence of such a time reversal invariant solution to (4.6) implies the existence of a maximally symmetric state constructed along the lines of the maximally $O(4,1)$ invariant invariant state (2.14) in the de Sitter background. The existence of this state of maximal symmetry does not imply that it is the stable ground state of either the de Sitter or electric field backgrounds. In the electric field case this is well known and the decay rate of the electric field into particle/antiparticle pairs [22] is becoming close to being experimentally verified in the near future [45]. That result is easily recovered in the present formalism by calculations exactly parallel to those of the de Sitter case in the last section.

First the Bogoliubov transformation analogous to (3.10) relating the *in* state mode function to the symmetric one $\upsilon_\lambda(u)$ and its complex conjugate are determined from the relation between the parabolic cylinder function in $f_{\lambda(\pm)}(u)$ and the confluent hypergeometric functions, cf. Eqs. 6.9.2 (31) and 6.5 (7) of [40], which give

$$f_{\lambda(+)}(u) = A_\lambda^{in}\upsilon_\lambda(u) + B_\lambda^{in}\upsilon_\lambda^* \tag{4.18}$$

with

$$A_\lambda^{in} = \sqrt{\frac{\pi}{2}}\, 2^{\frac{i\lambda}{2}} e^{i\eta_\lambda} e^{-\frac{\pi\lambda}{4}}\left[\left(\frac{\lambda}{2}\right)^{\frac{1}{4}}\frac{1}{\Gamma(\frac{3}{4}-\frac{i\lambda}{2})} + \left(\frac{2}{\lambda}\right)^{\frac{1}{4}}\frac{e^{\frac{i\pi}{4}}}{\Gamma(\frac{1}{4}-\frac{i\lambda}{2})}\right] \tag{4.19a}$$

$$B_\lambda^{in} = \sqrt{\frac{\pi}{2}}\, 2^{\frac{i\lambda}{2}} e^{i\eta_\lambda} e^{-\frac{\pi\lambda}{4}}\left[\left(\frac{\lambda}{2}\right)^{\frac{1}{4}}\frac{1}{\Gamma(\frac{3}{4}-\frac{i\lambda}{2})} - \left(\frac{2}{\lambda}\right)^{\frac{1}{4}}\frac{e^{\frac{i\pi}{4}}}{\Gamma(\frac{1}{4}-\frac{i\lambda}{2})}\right]. \tag{4.19b}$$

Because the *in* and *out* mode functions satisfy the same relations as (3.7), and have the same relation to the symmetric mode function $\upsilon_\lambda(u)$ as the corresponding *in* and *out* mode functions have to the CTBD mode function (3.9) in the de Sitter case, the Bogoliubov coefficients defined by the analogs of (3.10)–(3.15), and the coefficients of the total Bogoliubov transformation from *in* to *out* states in the electric field case are given by the same relations as (3.15), namely

$$A_\lambda^{\rm tot} = (A_\lambda^{in})^2 - (B_\lambda^{in})^2 = \frac{\sqrt{2\pi}}{\Gamma(\frac{1}{2}-i\lambda)} e^{-\frac{\pi\lambda}{2}} e^{i\lambda - i\lambda\ln\lambda}, \tag{4.20a}$$

$$B_\lambda^{\rm tot} = A_\lambda^{in*}B_\lambda^{in} - A_\lambda^{in}B_\lambda^{in*} = -ie^{-\pi\lambda}. \tag{4.20b}$$

Thus the number density of *out* particles at late times in the mode labeled by $\mathbf{k}$ or $(k_z, k_\perp)$ if the system is prepared in the $|in\rangle$ vacuum is

$$|B_\lambda^{\rm tot}|^2 = e^{-2\pi\lambda} = \exp\left[-\frac{\pi(k_\perp^2+m^2)}{eE}\right], \tag{4.21}$$

and the relative probability of finding a particle/antiparticle charged pair in the mode characterized by $(k_z, \mathbf{k}_\perp)$ in the $|in\rangle$ vacuum is

$$w_\lambda \equiv \left|\frac{B_\lambda^{\rm tot}}{A_\lambda^{\rm tot}}\right|^2 = \frac{1}{e^{2\pi\lambda}+1}, \tag{4.22}$$

which is independent of $k_z$. The vacuum overlap or vacuum persistence probability is given then by the analog of (3.24),

$$|\langle out|in\rangle|^2 = \prod_{\mathbf{k}}(1-w_\lambda) = \exp\left\{-\sum_{\mathbf{k}} \ln(1+e^{-2\pi\lambda})\right\}. \tag{4.23}$$

Taking the infinite volume limit and converting the sum into an integral according to

$$\sum_{\mathbf{k}} \to \frac{V}{(2\pi)^3}\int dk_z \int d^2\mathbf{k}_\perp, \tag{4.24}$$

we see that the exponent in (4.23) both diverges in $V$ and diverges because the integrand is independent of $k_z$. Thus we encounter a divergence in this mode sum quite analogous to the de Sitter case (3.24). Again the reason for this is the infinite amount of particle production in an infinite four-volume and one should again define the decay rate by dividing the exponent in the vacuum persistence probability (4.23) by the four-volume $VT$, before taking the infinite time limit $T \to \infty$. In this case, we recognize that the physical (kinetic) longitudinal momentum of the particle in mode $k_z$ is $p = k_z + eEt$, so that for a fixed large $p = P$ cutoff we have

$$dk_z = -eEdt. \tag{4.25}$$

Thus the positive integral over $k_z$ in (4.24) may be replaced by $eET$, $T$ being the total elapsed time over which the electric field acts to create pairs. In this way we obtain from (4.23)–(4.25) the vacuum decay rate per unit three-volume $V$ per unit time $T$ to be

$$\begin{aligned}\Gamma &= \frac{eE}{(2\pi)^3}\int d^2\mathbf{k}_\perp \ln(1+e^{-2\pi\lambda})\\ &= \frac{eE}{(2\pi)^3}\int_0^\infty \pi dk_\perp^2 \sum_{n=1}^\infty \frac{(-)^{n+1}}{n}\exp\left[-\frac{\pi n(k_\perp^2+m^2)}{eE}\right]\\ &= \frac{(eE)^2}{(2\pi)^3}\sum_{n=1}^\infty \frac{(-)^{n+1}}{n^2}\exp\left(-\frac{\pi n m^2}{eE}\right),\end{aligned} \tag{4.26}$$

which is Schwinger's result for scalar QED. (Schwinger actually obtained his result for fermionic QED in which the alternating sign in the sum over $n$ is absent [22]).

Thus the definition of the $|in\rangle$ and $|out\rangle$ states which are purely positive frequency as $t \to \mp\infty$, respectively, according to (4.11) gives a nontrivial particle creation rate and imaginary part of the one-loop effective action which agrees with [22], notwithstanding the existence of a fully time symmetric state with mode functions (4.16). Clearly a nonzero imaginary part and decay rate breaks the time reversal symmetry of the background. Mathematically this is of course a result of initial boundary conditions on the vacuum, implemented in the present treatment by the definition of positive frequency solutions at early and late times, or in Schwinger's proper time original treatment by the $m^2 \to m^2 - i0^+$ prescription of avoiding a pole. As in the de Sitter case, the time symmetric modes (4.16) can be defined and have the maximal symmetry of the background $\mathbf{E}$ field. They do *not* describe a true vacuum state, but rather a specific coherent superposition of particles and antiparticles with respect to either the $|in\rangle$ or $|out\rangle$ vacuum states, "halfway between." The time symmetric state defined by the solution (4.16) is a very curious state indeed, corresponding to the rather unphysical boundary condition of each pair creation event (cf. Sec. V) being exactly balanced by its time reversed pair annihilation event, these pairs having been precisely arranged to come from great distances at early times in order to effect just such a cancellation.

We note also that taking the strict asymptotic states (4.9) in a constant uniform electric field leads to the same sort of divergence in the $k_z$ momentum integration we encountered in the $k$ sum in the de Sitter case, which can be handled by the replacement (4.25) based on similar considerations of a fixed physical momentum cutoff. The reason that the calculation leading to (4.23) together with a physical argument for the $k_z$ cutoff gives the identical answer to Schwinger's proper time method [22] is of course due to the fact that the definition of particles by the positive frequency solutions of the time dependent mode Eq. (4.6) is the same one selected by the covariant analyticity requirement of the $m^2 - i0^+$ prescription. For this correspondence to be unambiguous it is important that the adiabatic frequency function $\Theta_\lambda(u)$ in (4.8) have well-defined asymptotic behavior at large $|u| \gg \sqrt{\lambda}$, so that the *in* and *out* positive frequency mode functions may be identified by the asymptotic behaviors of the appropriate exact solutions of (4.6), even though the electric field does not vanish in these asymptotic regions at very early or very late times. Indeed exactly the same result (4.26) is obtained if the electric field is switched on and off smoothly [28,30,35] in a finite time $T$. Then the Bogoliubov coefficients have a nontrivial $k_z$ dependence and the integral over $k_z$ for finite $T$ is finite. Dividing by $T$ and taking the limit $T \to \infty$ one recovers exactly the decay rate (4.26) according to the replacement (4.25) above.

Presumably the $|in\rangle$ and $|out\rangle$ states in the constant, uniform electric field have Wightman functions with the same sort of non-Hadamard behavior as those in de Sitter space, and for the same reason, namely the noncommutivity of the infinite time $T \to \infty$ and infinite momentum $(k_z, k_\perp) \to \infty$ limits. In the electric field case the physical cutoff on $|k_z|$ is of order $eET$, so that the very high $|k_z|$ modes larger than this cutoff are undisturbed from the ordinary zero field vacuum and the unphysical ultraviolet behavior of matrix elements and Green's functions in the initial state is removed when $T$ is finite. The finite $T$ regulator eliminates all UV problems, and transfers the divergence instead to the question of the long time or infrared secular evolution of the system. Then time translational as well as time reversal symmetry is lost.

## V. ADIABATIC VACUUM STATES AND PARTICLE CREATION IN REAL TIME

All idealized calculations in background fields that persist for infinite times do not give much physical insight into the particle creation process itself in real time. In formulating a well-posed time dependent problem with UV finite initial data one needs to define states in which the momentum dependent particle creation process is started and can be followed at any finite time. This leads naturally to the introduction of adiabatic vacuum and particle states defined at arbitrary times, instead of just in the asymptotic past or future.

The *in* and *out* mode functions $f_{(+)}$ and $f^{(+)}$ are pure positive frequency particle modes in the asymptotic past and asymptotic future, respectively, while the time symmetric $v$ or $F$ is "halfway between" them and a positive frequency mode at $u = 0$. This suggests that it would be useful to introduce WKB mode functions,

$$\tilde{f}_{\mathbf{k}} = \frac{1}{\sqrt{2W_{\mathbf{k}}}} \exp\left(-i\int^t dt W_{\mathbf{k}}\right), \tag{5.1}$$

that are approximate adiabatic positive frequency modes at any intermediate time $t$, to interpolate between these limits. These approximate modes are related to any of the *exact* mode function solutions $f_{(+)}$ and $f^{(+)}$ or $v$ of the oscillator equation (3.2) or (4.3) in the de Sitter or electric field backgrounds (which we denote generically by $f_{\mathbf{k}}$) by a time-dependent Bogoliubov transformation

$$\begin{pmatrix} f_{\mathbf{k}} \\ f^*_{\mathbf{k}} \end{pmatrix} = \begin{pmatrix} \alpha_{\mathbf{k}}(t) & \beta_{\mathbf{k}}(t) \\ \beta^*_{\mathbf{k}}(t) & \alpha^*_{\mathbf{k}}(t) \end{pmatrix} \begin{pmatrix} \tilde{f}_{\mathbf{k}} \\ \tilde{f}^*_{\mathbf{k}} \end{pmatrix}, \tag{5.2}$$

where we require that

$$|\alpha_{\mathbf{k}}(t)|^2 - |\beta_{\mathbf{k}}(t)|^2 = 1 \tag{5.3}$$

be satisfied at all times. The time dependent real frequency function $W_{\mathbf{k}}$ in (5.1) is to be chosen to match the exact frequency function $\Omega_k$ or $\omega_{\mathbf{k}}$ of the time dependent harmonic oscillator equation (3.2) or (4.3), i.e. (3.3) or (4.4), to some order in the adiabatic expansion

$$W^2_{\mathbf{k}} \simeq \omega^2_{\mathbf{k}} - \frac{1}{2}\frac{\ddot{\omega}_{\mathbf{k}}}{\omega_{\mathbf{k}}} + \frac{3}{4}\frac{\dot{\omega}^2_{\mathbf{k}}}{\omega^2_{\mathbf{k}}} + \cdots, \tag{5.4}$$

obtained by substituting (5.1) into the oscillator equation and expanding in time derivatives of the frequency. The expansion (5.4) is adiabatic in the usual sense of slowly varying, in that it is clear that the approximate positive frequency mode (5.1) more and more accurately approaches an exact mode solution of the oscillator equation (3.2) or (4.3), as (3.3) or (4.4) becomes a more slowly varying function of time, which is controlled by the strength of the background gravitational or electric field.

An important property of the expansion (5.4) is that it is an asymptotic series (rather than a convergent series) which is nonuniform in time. The higher order terms fall off more and more rapidly at large $|\mathbf{k}|$, for any value of the background field $H$ or $E$, no matter how rapidly the background varies, and irrespective of its asymptotic behavior in time. In the literature the term adiabatic is most often used in this second sense of the large $|\mathbf{k}|$ behavior of vacuum modes for arbitrary (smooth) backgrounds, independently of whether or not they are slowly varying in time [8]. This guarantees that the adiabatic vacuum defined by (5.1) will match the usual Minkowski vacuum at sufficiently short distance scales as $|\mathbf{k}| \to \infty$, for any smoothly varying background field, and a sufficiently high order adiabatic vacuum leads to Green's functions with Hadamard behavior. This is essential to the renormalization program for currents and stress tensors, which is necessary to formulate the backreaction problem for time varying background fields [6–8]. However, as is generally the case with WKB methods and asymptotic series more generally, the adiabatic expansion misses exponentially small contributions in the vicinity of turning points where $\omega_{\mathbf{k}}$ vanishes. As a result, the adiabatic mode function (5.1) is not uniformly valid over all times, for any finite order truncation of the asymptotic series (5.4), and mixing with $\tilde{f}^*_{\mathbf{k}}$ generally occurs.

Because of the Wronskian normalization conditions (2.8) or (4.10), the coefficients of the time dependent Bogoliubov transformation (5.2) are completely defined only if the first time derivatives of the exact mode functions in terms of $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$ are also specified. The general form of $\dot{f}_{\mathbf{k}}$ in terms of the adiabatic modes $\tilde{f}_{\mathbf{k}}$ that preserves both the Wronskian condition (2.8) and (5.3) is [37]

$$\frac{d}{dt} f_{\mathbf{k}} = \left(-iW_{\mathbf{k}} + \frac{V_{\mathbf{k}}}{2}\right)\alpha_{\mathbf{k}}\tilde{f}_{\mathbf{k}} + \left(iW_{\mathbf{k}} + \frac{V_{\mathbf{k}}}{2}\right)\beta_{\mathbf{k}}\tilde{f}^*_{\mathbf{k}}, \tag{5.5}$$

where $V_{\mathbf{k}}$ is a second time dependent real function, with its own adiabatic expansion given by the time derivative of $W_{\mathbf{k}}$ from (5.4). For any real $(W_{\mathbf{k}}, V_{\mathbf{k}})$ the transformation of bases (5.2) may be viewed as a time dependent canonical transformation in the phase space of the coordinates $f_{\mathbf{k}}$ and their conjugate momenta $\dot{f}_{\mathbf{k}}$. The corresponding adiabatic particle and antiparticle creation and destruction operators may be defined by setting the Fourier components of the scalar field,

$$\begin{aligned}\varphi_{\mathbf{k}}(t) &\equiv a_{\mathbf{k}} f_{\mathbf{k}}(t) + b^{\dagger}_{\mathbf{k}} f^{*}_{-\mathbf{k}}(t)\\ &= \tilde{a}_{\mathbf{k}}(t)\tilde{f}_{\mathbf{k}}(t) + \tilde{b}^{\dagger}_{-\mathbf{k}}(t)\tilde{f}^{*}_{-\mathbf{k}}(t),\end{aligned} \tag{5.6}$$

equal so that the canonical transformation in the Fock space (of a charged scalar field) is

$$\begin{pmatrix}\tilde{a}_{\mathbf{k}}(t)\\ \tilde{b}^{\dagger}_{-\mathbf{k}}(t)\end{pmatrix} = \begin{pmatrix}\alpha_{\mathbf{k}}(t) & \beta^{*}_{\mathbf{k}}(t)\\ \beta_{\mathbf{k}}(t) & \alpha^{*}_{\mathbf{k}}(t)\end{pmatrix}\begin{pmatrix}a_{\mathbf{k}}\\ b^{\dagger}_{-\mathbf{k}}\end{pmatrix} \tag{5.7}$$

when referred to the time independent basis $(a_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}})$. For an uncharged Hermitian scalar field, $b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$ are replaced by $a_{\mathbf{k}}$ and $a^{\dagger}_{-\mathbf{k}}$, respectively. The time-dependent instantaneous mean adiabatic particle number in the mode $\mathbf{k}$ is defined in the $(\tilde{a}_{\mathbf{k}}, \tilde{b}^{\dagger}_{-\mathbf{k}})$ basis as

$$\begin{aligned}\mathcal{N}_{\mathbf{k}}(t) &\equiv \langle \tilde{a}^{\dagger}_{\mathbf{k}}(t)\tilde{a}_{\mathbf{k}}(t)\rangle\\ &= \langle \tilde{b}^{\dagger}_{-\mathbf{k}}(t)\tilde{b}_{-\mathbf{k}}(t)\rangle\\ &= |\alpha_{\mathbf{k}}(t)|^2\langle a^{\dagger}_{\mathbf{k}}a_{\mathbf{k}}\rangle + |\beta_{\mathbf{k}}(t)|^2\langle b_{-\mathbf{k}}b^{\dagger}_{-\mathbf{k}}\rangle\\ &= N_{\mathbf{k}} + (1+2N_{\mathbf{k}})|\beta_{\mathbf{k}}(t)|^2,\end{aligned} \tag{5.8}$$

where

$$N_{\mathbf{k}} \equiv \langle a^{\dagger}_{\mathbf{k}}a_{\mathbf{k}}\rangle = \langle b^{\dagger}_{-\mathbf{k}}b_{-\mathbf{k}}\rangle \tag{5.9}$$

is the number of particles (assumed equal to the number of antiparticles) referred to the time independent basis. This may be taken to be the particle number at the initial time $t = t_0$ provided that we initialize so that $|\beta_{\mathbf{k}}(t_0)|^2 = 0$.

With $V_{\mathbf{k}}$ defined in terms of $\dot{f}_{\mathbf{k}}$ by (5.5), the time dependent Bogoliubov coefficients may be found explicitly:

$$\alpha_{\mathbf{k}} = i\tilde{f}^{*}_{\mathbf{k}}\left[\dot{f}_{\mathbf{k}} - \left(iW_{\mathbf{k}} + \frac{V_{\mathbf{k}}}{2}\right)f_{\mathbf{k}}\right] \tag{5.10a}$$

$$\beta_{\mathbf{k}} = -i\tilde{f}_{\mathbf{k}}\left[\dot{f}_{\mathbf{k}} + \left(iW_{\mathbf{k}} - \frac{V_{\mathbf{k}}}{2}\right)f_{\mathbf{k}}\right]. \tag{5.10b}$$

and in particular,

$$|\beta_{\mathbf{k}}(t)|^2 = \frac{1}{2W_{\mathbf{k}}}\left|\dot{f}_{\mathbf{k}} + \left(iW_{\mathbf{k}} - \frac{V_{\mathbf{k}}}{2}\right)f_{\mathbf{k}}\right|^2 \tag{5.11}$$

is determined in terms of the adiabatic frequency functions $(W_{\mathbf{k}}, V_{\mathbf{k}})$ and the exact mode function solution $f_{\mathbf{k}}$ of the oscillator equation (3.2) or (4.3), which is specified by initial data $(f_{\mathbf{k}}, \dot{f}_{\mathbf{k}})$ at $t = t_0$. Although the choice of $(W_{\mathbf{k}}, V_{\mathbf{k}})$ is not unique, it is fairly tightly constrained by the requirements of matching the adiabatic behavior of the asymptotic expansion (5.4) to sufficiently high order, but not higher than is necessary to isolate the divergences of the current $\langle \mathbf{j}\rangle$ or stress tensor $\langle T^{a}{}_{b}\rangle$ operators in their vacuumlike contributions. We shall see that with these requirements, although the detailed time dependence of $\mathcal{N}_{\mathbf{k}}(t)$ depends on the precise choice of $(W_{\mathbf{k}}, V_{\mathbf{k}})$, the main features of the adiabatic particle number are largely independent of the specific choice of these functions.

Let us first apply this general adiabatic framework to the constant, uniform electric field example. Although it is sufficient to choose the lowest order adiabatic frequency functions,

$$W^{(0)}_{\mathbf{k}} = \omega_{\mathbf{k}} = \sqrt{\frac{eE}{2}}(u^2+4\lambda)^{\frac{1}{2}} \tag{5.12a}$$

$$V^{(1)}_{\mathbf{k}} = -\frac{\dot{\omega}_{\mathbf{k}}}{\omega_{\mathbf{k}}} = -\sqrt{2eE}\frac{u}{u^2+4\lambda}, \tag{5.12b}$$

in this case, we shall also study the second-order choice,

$$\begin{aligned}W^{(2)}_{\mathbf{k}} &= \omega_{\mathbf{k}} - \frac{1}{4}\frac{\ddot{\omega}_{\mathbf{k}}}{\omega^2_{\mathbf{k}}} + \frac{3}{8}\frac{\dot{\omega}^2_{\mathbf{k}}}{\omega^4_{\mathbf{k}}}\\ &= \sqrt{\frac{eE}{2}}(u^2+4\lambda)^{\frac{1}{2}}\left[1 - \frac{1}{(u^2+4\lambda)^2} + \frac{5}{2}\frac{u^2}{(u^2+4\lambda)^3}\right],\end{aligned} \tag{5.13}$$

for comparison purposes. The Bogoliubov coefficient $|\beta_{\mathbf{k}}|^2$ and the adiabatic mean particle number were studied in a constant electric field background with the choice $W^{(0)}_{\mathbf{k}}$ and $V_{\mathbf{k}} = 0$ in [36]. In Fig. 1 we plot $|\beta_{\mathbf{k}}|^2$ defined by (5.11) with $f_{\mathbf{k}}$ the *in* vacuum mode function $f_{\lambda(+)}(u)$ of (4.9) for both the lowest order and second order choices

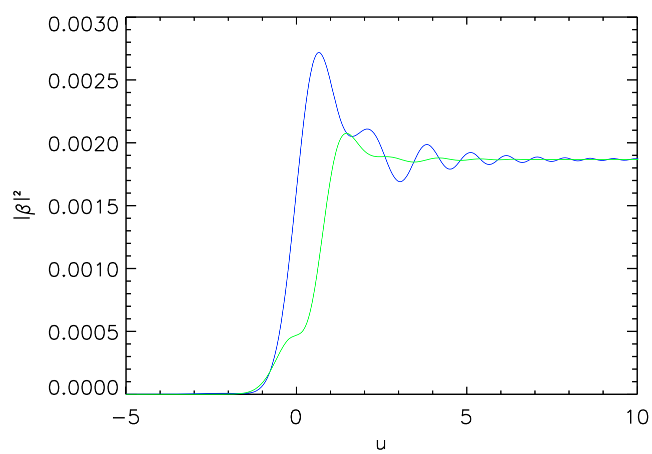


FIG. 1 (color online). The mean number of particles created from the vacuum $|in\rangle$ state, given by (5.11) with $f_{\mathbf{k}} = f_{\lambda(+)}(u)$ of (4.9), with $\lambda = 1$. The blue curve with larger oscillations is for the first order choice of $(W_{\mathbf{k}}, V_{\mathbf{k}})$ in (5.12), while the green curve is for the second order choice of $W_{\mathbf{k}}$ in (5.13) and the same $V_{\mathbf{k}}$. Both change rapidly around $u = 0$, and both tend to the same asymptotic value, $e^{-2\pi} = 0.001867$, of (5.15) as $u \to \infty$.

of adiabatic frequency $W_{\mathbf{k}}$, given by (5.12) and (5.13), respectively.

A continuous but sharp rise in $|\beta_{\mathbf{k}}|^2$ is observed in each $k_z$ mode around its "creation event," at $u = 0$, i.e. at the time when the kinetic momentum $p = k_z + eEt = 0$. Since the adiabatic mode functions are essentially WKB approximations to the time dependent harmonic oscillator equation (3.2) or (4.3), the particle creation process in real time and this rapid rise may be understood from a consideration of the WKB turning points in the complex $u$ plane [32,46,47]. These are defined by the values of $u$ where the frequency function $\omega_{\mathbf{k}}$ vanishes. Since the solutions are oscillatory on the real time axis, those turning points are located off the real line, and in the case of (4.3)–(4.4) the zeroes of the frequency are at

$$u = \pm u_\lambda \equiv \pm 2i\sqrt{\lambda} \tag{5.14}$$

as illustrated in Fig. 2.

Far from the turning points, for $|u| \gg 2\sqrt{\lambda}$, the exact mode functions are well approximated by the adiabatic WKB mode function (5.1) and hence $|\beta_{\mathbf{k}}(t)|^2$ defined by (5.11) will be approximately constant. For $u \ll -|u_\lambda| < 0$ the adiabatic vacuum is approximately the $|in\rangle$ vacuum discussed previously and $|\beta_{\mathbf{k}}(t)|^2$ is nearly zero if it is initialized so that $|\beta_{\mathbf{k}}(t_0)|^2 = 0$ for $(k_z + eEt_0) \ll -\sqrt{k_\perp^2 + m^2} < 0$. For $u \gg |u_\lambda| > 0$, i.e. for $(k_z + eEt) \gg \sqrt{k_\perp^2 + m^2} > 0$ the adiabatic vacuum is approximately the $|out\rangle$ vacuum. Again $|\beta_{\mathbf{k}}(t)|^2$ will be approximately constant in this region and given approximately by the total Bogoliubov coefficient $B_\lambda^{\rm tot}$ from *in* to *out*. In the region $u \in (-u_\lambda, u_\lambda)$, as $u$ crosses the Stokes' line of the function (4.8) along the imaginary axis joining the complex turning points (5.14), the exact mode function $f_{\mathbf{k}}(t)$ receives an increasing admixture of the negative frequency component, and $|\beta_{\mathbf{k}}|^2$ changes rapidly from its *in* to *out* value. This change in $|\beta_{\mathbf{k}}(t)|^2$ in the region $\Delta u \sim 4\sqrt{\lambda}$ or $\Delta t \sim 2\sqrt{k_\perp^2 + m^2}/eE$ around $u = 0$ (closest to the complex zeroes of $\omega_{\mathbf{k}}$) is given by (4.21) or

$$\Delta|\beta_{\mathbf{k}}|^2 = |B_\lambda^{\rm tot}|^2 = e^{-2\pi\lambda}. \tag{5.15}$$

This result can be derived also by analytic continuation of the adiabatic phase function of (4.8) in the complex $u$ plane, finding the lines of constant $\mathrm{Im}\,\Theta_\lambda(u)$ as they emanate from $u_\lambda$ and approach the real axis as $|u| \to \infty$ [46]. Since there is only one zero of $\omega_{\mathbf{k}}^2$ in the upper half complex plane and $\omega_{\mathbf{k}}^2$ vanishes linearly in $u$ as $u \to u_\lambda$, the linear turning point connection formulas for the WKB approximation extended to the complex plane apply, and one finds [32,46,47]

$$B_\lambda^{\rm tot} = -i\exp\{2i\Theta_\lambda(u_\lambda)\} = -ie^{-\pi\lambda} \tag{5.16}$$

from (4.8) together with (5.14), in agreement with (4.20b), and hence (5.15). The total Bogoliubov coefficient $B_\lambda^{\rm tot}$ is appropriate since the rise in $|\beta_{\mathbf{k}}|^2$ changes continuously in this region $u \in (-u_\lambda, u_\lambda)$ between the two complex zeroes (5.14) with no constant value halfway between. The numerical behavior of $|\beta_{\mathbf{k}}(u)|^2$ for various values of $\lambda$ is plotted in Fig. 3 showing the asymptotic value of the jump in particle number consistent with (5.15). Since this rise in $|\beta_{\mathbf{k}}|^2$ occurs around $u = 0$, the particle creation "event" occurs at a different time $t = -k_z/eE$ for modes with different values of $k_z$.

Consider now the adiabatic initial data at some finite time $t_0$,

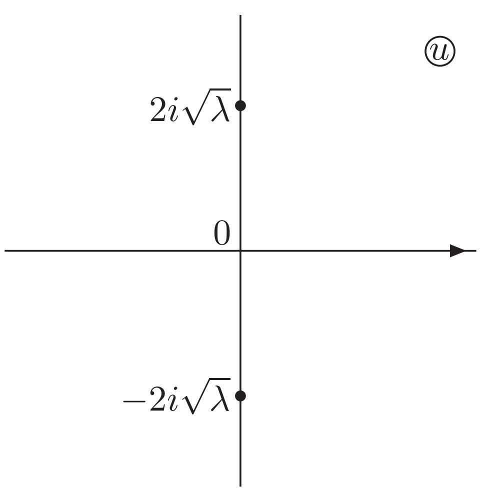


FIG. 2. Location of the zeroes of the frequency function $\omega_{\mathbf{k}}$ of (4.4) in the complex $u$ plane. Particle creation occurs as the real-time $u$ contour crosses $u = 0$, and the Stokes' line of the function (4.8) along the imaginary axis joining the pair of these zeroes (5.14).

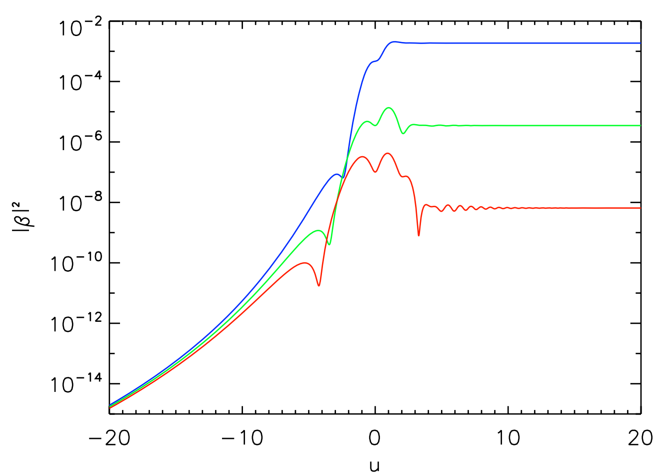


FIG. 3 (color online). The mean number of particles created from the vacuum $|in\rangle$ state, given by (5.11) with $f_{\mathbf{k}} = f_{\lambda(+)}(u)$ of (4.9), and the second order adiabatic frequency of (5.13) for $\lambda = 1, 2, 3$ (upper blue, middle green, lower red curves, respectively). Note the logarithmic scale. The asymptotic values for large $u$ are $1.87 \times 10^{-3}$, $3.49 \times 10^{-6}$, and $6.51 \times 10^{-9}$ for $\lambda = 1, 2, 3$, respectively, in agreement with (5.15).

$$f_{\mathbf{k}}(t_0) = \frac{1}{\sqrt{2\omega_{\mathbf{k}}(t_0)}} \tag{5.17a}$$

$$\left.\frac{df_{\mathbf{k}}}{dt}\right|_{t=t_0} = -\left(i\omega_{\mathbf{k}} + \frac{\dot{\omega}_{\mathbf{k}}}{2\omega_{\mathbf{k}}}\right) f_{\mathbf{k}}|_{t=t_0} \tag{5.17b}$$

with $\omega_{\mathbf{k}}(t)$ given by (4.4). This matches the adiabatic vacuum with (5.12) so that $\beta_{\mathbf{k}}(t_0) = 0$. Since the creation event occurs around $u = 0 = k_z + eEt$, with a finite starting time only those modes for which the initial kinetic momentum $p(t_0) = k_z + eEt_0 < 0$ can experience this creation event. They do so at the time when their kinetic momentum $p(t) = k_z + eEt = 0$, i.e. when the particle initially moving in the opposite direction to the electric field is brought to instantaneous rest $p(t) = 0$ by the constant positive acceleration of the field and begins to move in the direction of the electric field. On the other hand those modes for which $p(t_0) = k_z + eEt_0 > 0$ are already moving in the same direction as the electric field at the initial time and undergo no particle creation event at later times, being already approximately in the $|out\rangle$ vacuum state at the initial time $t_0$. Crudely approximating the creation event as a step function at $u = 0$ with step size (5.15), the number of particles in mode $\mathbf{k}$ at time $t > t_0$ may be estimated to be

$$\mathcal{N}_{\mathbf{k}}(t) = (1 + 2N_{\mathbf{k}})|\beta_{\mathbf{k}}(t)|^2 \approx \theta(p(t))\theta(-p(t_0))e^{-2\pi\lambda} \approx (1 + 2N_{\mathbf{k}})\theta(k_z + eEt)\theta(-k_z - eEt_0)e^{-2\pi\lambda}, \tag{5.18}$$

where the factor of $1 + 2N_{\mathbf{k}}$ accounts for the induced creation rate of particles if there are already particles $N_{\mathbf{k}} > 0$ in the initial state. From (5.18) there is a "window function" in $k_z$ for modes going through particle creation given by

$$-eEt < k_z < -eEt_0, \tag{5.19}$$

which grows linearly with elapsed time $t - t_0$. Modes with $k_z$ lying outside this range at any finite $t$ remain in the adiabatic vacuum. However as $t \to \infty$, modes with arbitrarily large $|k_z|$ experience a creation event. Hence it is clear that the large $t$ and large $|k_z|$ limits do not commute. This is a concrete expression of the nonuniformity in time of the single frequency adiabatic expansion (5.4). It is this noncommutivity of limits that leads to the non-Hadamard properties of the asymptotic $|in\rangle$ and $|out\rangle$ states, and the consequent vanishing of their overlap (3.31) in the infinite time, infinite volume limit.

The actual behavior of $|\beta_{\mathbf{k}}|^2$ is shown in detail in Figs. 1 and 3, and Figures 2-4 of Ref. [36], which rise smoothly on the time scale of $\Delta u \sim 4\sqrt{\lambda}$. This behavior can be accurately captured by the uniform asymptotic approximation of the parabolic cylinder functions even for moderately small $\lambda$ [36]. Replacing this smooth rise of the average particle number by a step function already gives a qualitatively correct picture of the semiclassical particle creation process mode by mode in real time, with the correct asymptotic density of particles. It is the window function (5.19) which justifies the replacement of the integral over $k_z$ in (4.24) by $eE$ times the total elapsed time $T = t - t_0$, which can then be divided out to obtain the decay rate (4.26). The window function (5.19) of the real time particle creation process also agrees with the analysis of adiabatically switching on and off of the background electric field, so that it acts only for a finite time [28,30,35,48]. It is this definition of particles created by the electric field in the adiabatic basis that forms the starting point in quantum theory for a kinetic description [36].

The adiabatic basis also furnishes a simple physically well-motivated method for defining renormalized expectation values of current and energy-momentum bilinears in the quantum field. In the approximation in which the electric field background is treated classically while the charged scalar matter field is quantized, the renormalized $j_z$ current expectation value is

$$\langle t_0|j_z(t)|t_0\rangle_R = 2e\int \frac{d^3\mathbf{k}}{(2\pi)^3}(k_z + eEt) \times \left[(1 + 2N_{\mathbf{k}})|f_{\mathbf{k}}(t)|^2 - \frac{1}{2\omega_{\mathbf{k}}(t)}\right], \tag{5.20}$$

where the leading divergence has been subtracted by the adiabatic vacuum term in which $|f_{\mathbf{k}}|^2$ has been replaced by $|\tilde{f}_{\mathbf{k}}|^2$ with (5.12) and $N_{\mathbf{k}}$ replaced by zero. It can be shown that this one subtraction removes all the UV divergences in the momentum integral for a constant $E$ field [34]. A logarithmic divergence proportional to $\ddot{E}$ can be removed by using the second adiabatic order approximation for $W_{\mathbf{k}}$ in the expansion (5.4). As this term can easily be reabsorbed into coupling renormalization in backreaction calculations and vanishes in any case for a constant $E$ field, the lowest order subtraction in (5.20) is sufficient for our present purposes.

Substituting (5.2) we obtain from (5.8) and (5.20)

$$\langle t_0|j_z(t)|t_0\rangle_R = 2e\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{(k_z + eEt)}{\omega_{\mathbf{k}}} \times [\mathcal{N}_{\mathbf{k}} + (1 + 2N_{\mathbf{k}})\mathrm{Re}(\alpha_{\mathbf{k}}\beta^*_{\mathbf{k}}e^{-2i\Theta_{\mathbf{k}}})], \tag{5.21}$$

where

$$\Theta_{\mathbf{k}} \equiv \int_{t_0}^{t} \omega_{\mathbf{k}} dt = \Theta_\lambda(u(t)) - \Theta_\lambda(u(t_0)) \tag{5.22}$$

is the adiabatic phase in (5.1), related to the function $\Theta_\lambda(u)$ defined in (4.8). Since $(k_z + eEt)/\omega_{\mathbf{k}} = p/\omega_{\mathbf{k}}$ is the $z$

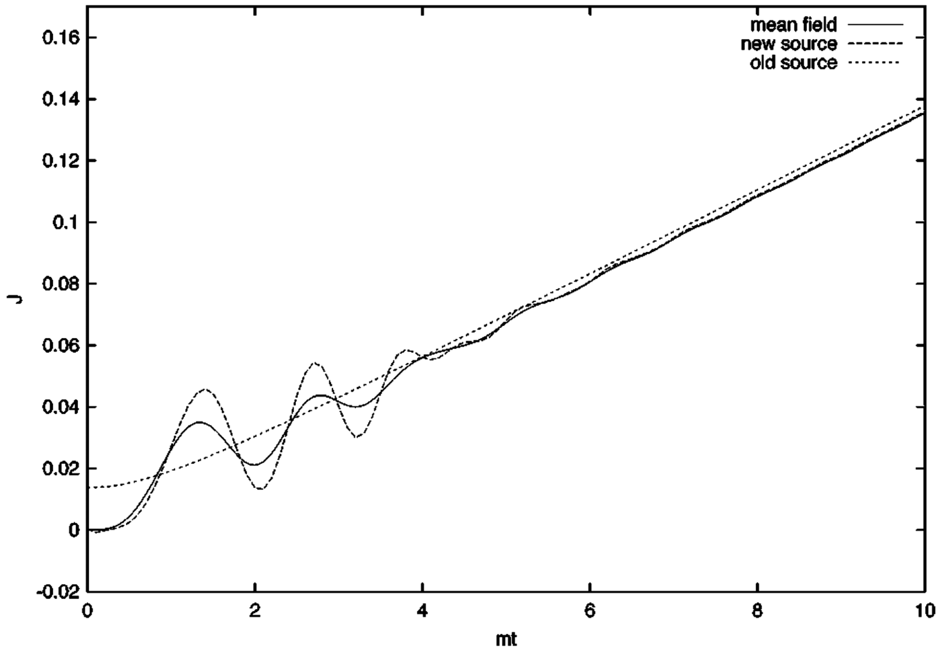


FIG. 4. The linear growth of the electric current $J = \langle j_z \rangle$ with time in the case of fixed constant background electric field in $1+1$ dimensions in units of $e^2 E$ for $eE/m^2 = 1$. The data and the plot were generated in Ref. [36]. The three curves shown are the current of the exact renormalized current expectation value (5.20), (solid, labelled *mean field*), a uniform approximation described in Ref. [36] (dashed, labelled *new source*), and that obtained from the simple window step function of particle creation in (5.18) and (5.23), (dotted, labelled *old source*).

component of the velocity of a classical particle in the electric field, the first term in the integral of (5.21) has a self-evident classical interpretation as the contribution to the electric current of the positive plus negatively charged particles with phase space number density $\mathcal{N}_{\mathbf{k}}$. The second $\alpha_{\mathbf{k}}\beta^*_{\mathbf{k}}$ term in (5.21) is a quantum interference term which has no classical analog. This term is both rapidly oscillating in time and rapidly oscillating in $|\mathbf{k}|$ for fixed time, so one would expect it to average out in the integral and give a relatively small contribution to the total current compared to the first term. For the semiclassical particle interpretation based on the adiabatic modes (5.1) to be most useful, this should be the case. If it is, one can also substitute the step approximation (5.18) for the particle density (assuming $N_{\mathbf{k}} = 0$, i.e. no particles in the initial state) and arrive at the simple result,

$$\begin{aligned}\langle t_0|j_z(t)|t_0\rangle_R &\approx 2e\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{(k_z+eEt)}{\omega_{\mathbf{k}}(t)}\theta(k_z+eEt)\\ &\quad\times\theta(-k_z-eEt_0)\exp\left[-\frac{\pi(k_\perp^2+m^2)}{eE}\right]\\ &=\frac{e}{\pi}\left[\sqrt{e^2E^2(t-t_0)^2+m^2}-m\right]\\ &\quad\times\int_0^\infty\frac{dk_\perp^2}{4\pi}\exp\left[-\frac{\pi(k_\perp^2+m^2)}{eE}\right]\\ &\to\frac{e^3E^2}{4\pi^3}(t-t_0)e^{-\frac{\pi m^2}{eE}},\end{aligned}\tag{5.23}$$

for the linear growth with time of the mean electric current of the created particles. This exhibits the secular effect coming from the window function (5.19) opening linearly with time so that more and more modes go through their particle creation event as time goes on, each becoming accelerated very rapidly to the speed of light, and making a constant contribution to the current.

One can also evaluate the exact expectation value (5.20) for a constant uniform electric field background starting with the initial adiabatic data (5.17) and compare it to the simple step function approximation (5.23). This comparison is shown in Fig. 4 [36]. The transient oscillations at early times are the effect of the second quantum interference term in (5.21), while the dominant *secular* effect of linear growth at late times is correctly captured by the simple approximation (5.23) based on the particle creation picture, labeled as *old source* in Fig. 4. The curve labeled *new source* is the uniform approximation of [36] that gives a slightly better approximation than the crude step function approximation of (5.18). Either gives correctly the coefficient of the linear secular growth with time, which implies that backreaction must eventually be taken into account, no matter how small $eE/m^2$ is, provided only that it is nonzero. This secular growth is a nonperturbative infrared *memory effect* in the sense of depending upon the time elapsed since the initial vacuum state was prepared at $t = t_0$. Note that this time dependence due to particle creation is a spontaneous breaking of the time translational and time reversal symmetry of the background constant **E** field [10,49]. The exponentially small tunneling factor associated with the spontaneous Schwinger particle creation rate from the vacuum shows that the effect is nonperturbative, but that however small, it can be overcome by a large initial state density of particles $N_{\mathbf{k}} \gg 1$ for which the induced particle creation and current is much larger. Even in the initial adiabatic vacuum case for $N_{\mathbf{k}} = 0$, particle creation eventually overcomes the small tunneling factor at late enough times.

## VI. ADIABATIC STATES AND INITIAL DATA IN DE SITTER SPACE

As in the electric field case, we introduce instantaneous adiabatic vacuum states in de Sitter space, defined by the adiabatic mode functions,

$$\tilde{f}_k = \frac{1}{\sqrt{2W_k}} \exp\left(-i\int^{\tau} d\tau W_k\right), \tag{6.1}$$

analogous to (5.1). Due to spatial homogeneity and isotropy in the cosmological case, these modes depend only upon the magnitude $k = |\mathbf{k}|$ which is the principal quantum number of the spherical harmonic on $\mathbb{S}^3$. The time dependent coefficients $\alpha_k(u)$ and $\beta_k(u)$ of the Bogoliubov transformation are defined by

$$f_k = \alpha_k \tilde{f}_k + \beta_k \tilde{f}_k^* \tag{6.2a}$$

$$H\frac{d}{du} f_k = \left(-iW_k + \frac{V_k}{2}\right)\alpha_k \tilde{f}_k + \left(iW_k + \frac{V_k}{2}\right)\beta_k \tilde{f}_k^*, \tag{6.2b}$$

where $f_k$ is an exact mode function solution of (3.2). They are given again by (5.10)

$$|\beta_{\mathbf{k}}(t)|^2 = \frac{1}{2W_{\mathbf{k}}}\left|\dot{f}_{\mathbf{k}} + \left(iW_{\mathbf{k}} - \frac{V_{\mathbf{k}}}{2}\right) f_{\mathbf{k}}\right|^2 \tag{6.3}$$

and (5.3) is satisfied, provided only that both $W_k$ and $V_k$ are arbitrary real functions of time. The analog of (5.7) is now

$$\begin{pmatrix} \tilde{a}_{klm_l}(u) \\ \tilde{a}^{\dagger}_{kl-m_l}(u) \end{pmatrix} = \begin{pmatrix} \alpha_k(u) & \beta_k^*(u) \\ \beta_k(u) & \alpha_k^*(u) \end{pmatrix} \begin{pmatrix} a_{klm_l} \\ a^{\dagger}_{kl-m_l} \end{pmatrix} \tag{6.4}$$

when referred to any time-independent basis $(a_{klm_l}, a^{\dagger}_{kl-m_l})$ for the Hermitian scalar field $\Phi$ (not necessarily the CTBD basis). The time dependent mean adiabatic particle number in the mode $(klm_l)$ is independent of $(lm_l)$ for $O(4)$ invariant adiabatic states and may be defined by the analog of (5.8) in de Sitter space to be

$$\mathcal{N}_k(u) = \langle \tilde{a}^{\dagger}_{klm_l}(u)\tilde{a}_{klm_l}(u)\rangle = N_k + (1 + 2N_k)|\beta_k(u)|^2, \tag{6.5}$$

where

$$N_k \equiv \langle a^{\dagger}_{klm_l} a_{klm_l}\rangle \tag{6.6}$$

is the number of particles at the initial time $u = u_0$, provided $|\beta_{\mathbf{k}}(u_0)|^2 = 0$ is initialized to zero at the initial time $u = u_0$. Note that with this initialization, the exact mode function solution of (3.2) satisfies the initial conditions,

$$f_{k\gamma}(u_0) = \frac{1}{\sqrt{2W_k}}\bigg|_{u=u_0},$$
$$\dot{f}_{k\gamma}(u_0) = \frac{1}{\sqrt{2W_k}}\left(-iW_k + \frac{V_k}{2}\right)\bigg|_{u=u_0}, \tag{6.7}$$

and hence is a certain linear combination (time independent Bogoliubov transformation) of the CTBD mode function $F_{k\gamma}$ and its complex conjugate $F^*_{k\gamma}$. Correspondingly, the time independent basis operators $a_{klm_l}, a^{\dagger}_{klm_l}$ in Fock space are certain linear combinations of the $a^{v}_{klm_l}, a^{v\dagger}_{klm_l}$ operators that define the de Sitter invariant state (2.14), which can be expressed in terms of each other by time independent Bogoliubov coefficients dependent upon the initial data (6.7).

As in the electric field example, the behavior of the solutions of the mode equation (3.2) is determined by the location of the zeroes of the frequency function, $\Omega_k = 0$ in (3.4) in the complex $u$ plane at $\cosh \bar{u} = \pm i\gamma^{-1}\sqrt{k^2 - 1/4}$, or

$$\bar{u} = u_R + iu_I = \pm u_{k\gamma} + i\pi\left(n + \frac{1}{2}\right) \quad \text{with} \quad n \in \mathbb{Z} \quad \text{and} \tag{6.8a}$$

$$\sinh u_{k\gamma} = \frac{\sqrt{k^2 - \frac{1}{4}}}{\gamma},$$
$$u_{k\gamma} = \ln\left[\frac{\sqrt{k^2 - \frac{1}{4}} + \sqrt{\gamma^2 + k^2 - \frac{1}{4}}}{\gamma}\right]. \tag{6.8b}$$

Thus there is an infinite line of zeroes of $\Omega_k$ in the complex $u$ plane along the two vertical axes at $u = \pm u_{k\gamma}$ for $\gamma^2 > 0$, cf. Fig. 5. The largest effect on the Bogoliubov coefficient $\beta_k(u)$ will occur when the real time contour passes closest to these lines of complex turning points at $u = \pm u_{k\gamma}$. Hence there are two "creation events" in global de Sitter space, one in the contracting and one in the expanding phase symmetric around $u = 0$. Because of the multiple zeroes the simple linear turning point formula which worked in the electric field case will not be exact in this case, so we rely on the

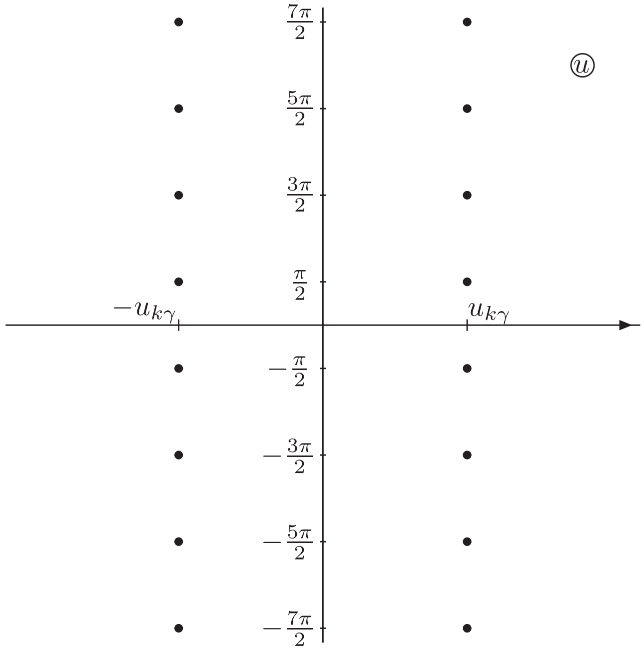


FIG. 5. Location of the zeroes (6.8) of the frequency function $\Omega_k$ (3.3) in the complex $u$ plane for $\gamma^2 > 0$. Particle creation occurs as the real time $u$ contour passes through the lines of these zeroes at $u = \mp u_{k\gamma}$.

Bogoliubov coefficients (3.20) and (3.21) computed from the exact *in* and *out* scattering solutions in Sec. III.

We may consider the two limits of (6.8b) for $u_{k\gamma}$:

$$u_{k\gamma} \to \frac{\sqrt{k^2 - \frac{1}{4}}}{\gamma} \to 0 \quad \text{for } \gamma \gg k \quad \text{or} \tag{6.9a}$$

$$u_{k\gamma} \to \ln\left(\frac{2\sqrt{k^2 - \frac{1}{4}}}{\gamma}\right) \to \ln\left(\frac{2k}{\gamma}\right) \to \infty \quad \text{for } k \gg \gamma. \tag{6.9b}$$

The first limit (6.9a) is the nonrelativistic limit of very heavy particles whose rest mass is much larger than their physical momentum $k/a$ at all times. These nonrelativistic particles are created nearly at rest close to the symmetric point $u = 0$ between the contracting and expanding de Sitter phases, so that the two events merge into one. The second limit (6.9b) is the relativistic limit of particles whose physical momentum is much larger than their rest mass for most of their history. These particles are created in two bursts, at $u = \mp u_{k\gamma}$, when their physical momentum $kH\text{sech}u_{k\gamma}$ is of the same order as their rest mass, so that they are moderately relativistic at creation. In the contracting phase of de Sitter space $u < 0$ these particles, created around $u = -u_{k\gamma}$, are blueshifted exponentially rapidly in $u$, and thus become ultrarelativistic. This contracting phase with the created particles becoming ultrarelativistic is therefore most analogous to the previous electric field example, and is the phase where we can expect the largest backreaction effects. Conversely, in the expanding phase, $u > 0$, the particles created around $u = +u_{k\gamma}$ will be subsequently exponentially redshifted in $u$, and therefore have a much smaller backreaction effect. We emphasize that the time $\pm u_{k\gamma}$ is of the order of the horizon crossing of the mode at $u \sim \pm \ln(2k)$ *only* for $\gamma \sim 1$. For large values of $\gamma$, when $\gamma \gg k$ the particle creation events occur when the wavelength of the mode is much smaller than the horizon, while for $\gamma \to 0$ the particle creation events occur when the wavelength of the mode is much greater than the horizon. Due to the different disposition of zeroes of the adiabatic frequency in the electric field and de Sitter cases, cf. Figs. 2 and 5, there is no analog of this second burst of particle creation in the electric field case.

For all values of $\gamma$, most of the $k$ modes fall into the second case (6.9b), and experience two well-separated creation events at large $u_{k\gamma} \gg 1$ in both the contracting and expanding phases of de Sitter space. In contrast to the electric field case considered previously we may therefore distinguish three distinct regions,

$$I\colon\ u < -u_{k\gamma}, \quad in \tag{6.10a}$$

$$II\colon\ -u_{k\gamma} < u < u_{k\gamma}, \quad CTBD \tag{6.10b}$$

$$III\colon\ u_{k\gamma} < u, \quad out, \tag{6.10c}$$

where we have indicated the character of the adiabatic vacuum in each region. If one takes the infinite time limits $u \to \mp\infty$ with $k$ and $\gamma$ and hence $u_{k\gamma}$ fixed, one is automatically in the first *in* region or the third *out* region, respectively. This corresponds to the *in/out* scattering problem considered in Sec. III. If on the other hand one takes the $k \to \infty$ limit for fixed $u, \gamma$ then Eq. (6.9b) shows that one is always in region II, where the CTBD state is the adiabatic vacuum. This shows explicitly the noncommutivity of the infinite $u$ and infinite $k$ limits, with the transition between the two limits occurring at $u = \pm u_{k\gamma}$.

Next we consider the mode function(s) and adiabatic vacuum state specified by the initial values (6.7) at an arbitrary finite time $u_0 < 0$. The modes for a given value of $k$ fall into two possible classes:

$$(i)\ -u_{k\gamma} < u_0 < 0 \tag{6.11a}$$

$$(ii)\ u_0 < -u_{k\gamma} < 0. \tag{6.11b}$$

For modes in the first class *(i)* the initial time $u_0$ is already later than the first creation event. For these modes in region II, the adiabatic initial condition is close to the CTBD state in the high $k$ limit, $\beta_k \approx 0$ and nothing further happens in the contracting phase, as they remain in region II for all $u_0 < u \leq 0$. In contrast, the $k$ modes in class *(ii)* are approximately in the $|in\rangle$ vacuum state initially. These modes have yet to go through their particle creation event which occurs at the later time $u = -u_{k\gamma} > u_0$ in the contracting phase. At that time, the adiabatic particle vacuum switches rapidly to approximately the CTBD state as $u$ increases past $-u_{k\gamma}$. Thus this mode sees its time dependent Bogoliubov coefficient change rapidly in a few expansion times ($\Delta u \sim 1$ since the imaginary part of the nearest complex zero of $\Omega_k$ is $\pi/2$ and independent of $k, \gamma$) from approximately zero to a nonzero plateau determined by the Bogoliubov coefficient (3.20b). Approximating the jump in particle number at these creation events by a step function as before, we have

$$\Delta|\beta_k|^2 = |B^{in}_{k\gamma}|^2 = \frac{e^{-\pi\gamma}}{2\sinh(\pi\gamma)} = \frac{1}{e^{2\pi\gamma} - 1} \tag{6.12a}$$

$$\mathcal{N}_k(u) \approx \theta(u + u_{k\gamma})\theta(-u_{k\gamma} - u_0)\Delta\mathcal{N}_{1,k\gamma} \quad \text{for } u < 0, \quad u_0 < 0 \tag{6.12b}$$

$$\Delta\mathcal{N}_{1,k\gamma} = (1 + 2N_k)\Delta|\beta_k|^2 = \frac{1 + 2N_k}{e^{2\pi\gamma} - 1} \tag{6.12c}$$

in the contracting phase. The first $\theta$ function in (6.12) specifies the time of the creation event when the step occurs, while the second $\theta$ function restricts the modes to class *(ii)* for which the step occurs at a later time $u = -u_{k\gamma} > u_0$ in the contracting phase. These two $\theta$ functions give the "window function" which is similar to that found in the electric field case (5.19), namely,

$$K_\gamma(u) \equiv \sqrt{\gamma^2\sinh^2 u + \frac{1}{4}} < k < \sqrt{\gamma^2\sinh^2 u_0 + \frac{1}{4}} = K_\gamma(u_0), \tag{6.13}$$

in the contracting phase of de Sitter space for which $u_0 < u \leq 0$. Like (5.19) this window function has an upper limit fixed by the initial time and a lower limit which decreases as time evolves (for $u < 0$).

If we continue the evolution past the symmetric point $u = 0$ into the expanding de Sitter phase, all of the modes of class *(ii)* have experienced the first particle creation event, and then begin (with the smallest value of $k$ first) to experience a second creation event at $u = +u_{k\gamma}$. Thus the modes of class *(ii)* which started in region I undergo two creation events with a total Bogoliubov transformation of (3.21), while the modes of class *(i)* which started in region II undergo only the second creation event in the expanding phase for which the single Bogoliubov transformation $B^{out}_{k\gamma}$ applies. Again approximating these creation events by step functions we obtain

$$\begin{aligned}\mathcal{N}_k(u) \approx [&\theta(u_{k\gamma} - u)\theta(-u_{k\gamma} - u_0) + \theta(u - u_{k\gamma})\theta(u_0 + u_{k\gamma})] \\ &\times \Delta\mathcal{N}_{1,k\gamma} + \theta(u - u_{k\gamma})\theta(-u_{k\gamma} - u_0)\Delta\mathcal{N}_{2,k\gamma} \\ &\text{for } u > 0, \quad u_0 < 0\end{aligned} \tag{6.14a}$$

$$\Delta\mathcal{N}_{2,k\gamma} = (1 + 2N_k)|B^{\text{tot}}_{k\gamma}|^2 = (1 + 2N_k)\text{csch}^2(\pi\gamma) \tag{6.14b}$$

in the expanding phase of de Sitter space. The window function for this second creation event in the expanding phase is now

$$k < \sqrt{\gamma^2\sinh^2 u + \frac{1}{4}} = K_\gamma(u) \tag{6.15}$$

for the modes undergoing the second creation event at $u = +u_{k\gamma}$. Those with $k < K_\gamma(u_0)$ undergo both the first and second creation events with $\Delta\mathcal{N} = \Delta\mathcal{N}_{2,k\gamma}$, while those with $k > K_\gamma(u_0)$ experience only the second creation event with $\Delta\mathcal{N} = \Delta\mathcal{N}_{1,k\gamma}$.

This analysis may be repeated if the initial time $u_0 > 0$ is in the expanding phase. In this case all modes initially in region II, with $u_0 < u_{k\gamma}$ undergo a single creation event at $u = +u_{k\gamma}$. Hence we have

$$\mathcal{N}_k(u) \approx \theta(u - u_{k\gamma})\theta(u_{k\gamma} - u_0)\Delta\mathcal{N}_{1,k\gamma} \quad \text{for } u, u_0 > 0, \tag{6.16}$$

replacing (6.14). The window function in $k$ is now the reverse of (6.13), namely,

$$K_\gamma(u_0) < k < K_\gamma(u), \tag{6.17}$$

which like (6.15) shows an upper limit that increases with time.

The various cases (6.12), (6.14) and (6.16) may be collected into one result,

$$\begin{aligned}\mathcal{N}_k(\tau) \approx [&\theta(u_{k\gamma} - |u|)\theta(-u_{k\gamma} - u_0) \\ &+ \theta(u - u_{k\gamma})\theta(u_{k\gamma} - |u_0|)]\left(\frac{1 + 2N_k}{e^{2\pi\gamma} - 1}\right) \\ &+ \theta(u - u_{k\gamma})\theta(-u_{k\gamma} - u_0)(1 + 2N_k)\text{csch}^2(\pi\gamma),\end{aligned} \tag{6.18}$$

valid for all values of $u$ and initial times $u_0$. From this or (6.14) it is clear that for fixed $k$, with $u_0 \to -\infty$, $u \to +\infty$, the mode experiences both particle creation events and we recover (3.22), while for a finite interval of time only those modes for which $u_0 < -u_{k\gamma}$ and $u_{k\gamma} < u$ experience both creation events. Thus taking the symmetric limit with $u = -u_0 > 0$, the values of $k$ satisfying both these conditions are cut off at the maximum value $K_\gamma(u)$, i.e.

$$k \le K_\gamma(u) \to \frac{\gamma}{2} e^{|u|} \quad \text{or} \quad \ln K_\gamma(u) \to |u| + \ln\left(\frac{\gamma}{2}\right) \tag{6.19}$$

as $|u| \to \infty$, which is exactly the cutoff (3.28) that we argued on physical grounds earlier in Sec. III (and in Ref. [9]) should be used in the $k$ sum of (3.24) to calculate the finite decay rate per unit volume (3.32) of de Sitter space to massive particle creation in the limit $V_4 \to \infty$. The constant in (3.28) has been determined to be $\ln(\gamma/2)$ by our detailed analysis of the particle creation process in real time. The non-Hadamard short distance behavior of the $|in\rangle$ and $|out\rangle$ states found in [9] has also been removed by regulating the large $k$ behavior with a finite initial and final time, since the modes for which $k > K_\gamma(u)$ remain in the CTBD state in region II for all $-|u_0| < u < |u_0|$ and the CTBD state is known to have the correct short distance behavior [16].

The actual smooth behaviors of $|\beta_k(u)|^2$ defined by (6.3) for various $k$ and $u_0 = -15$ and $u_0 = -5$ are shown in Figs. 6 and 7, respectively. The increases in $|\beta_k(u)|^2$ occur on a time scale $\Delta u \sim 1$ for all the modes. The values chosen for the adiabatic frequency functions $(W_k, V_k)$ are

$$\begin{aligned} W_k^{(2)} &= \Omega_k + \frac{3}{8}\frac{\dot\omega_k^2}{\omega_k^3} - \frac{1}{4}\frac{\ddot\omega_k}{\omega_k^2} \\ &= \Omega_k + \frac{h^2}{8\omega_k}\left(1 - \frac{6m^2}{\omega_k^2} + \frac{5m^4}{\omega_k^4}\right) \\ &\quad + \frac{\dot h}{4\omega_k}\left(1 - \frac{m^2}{\omega_k^2}\right), \end{aligned} \tag{6.20a}$$

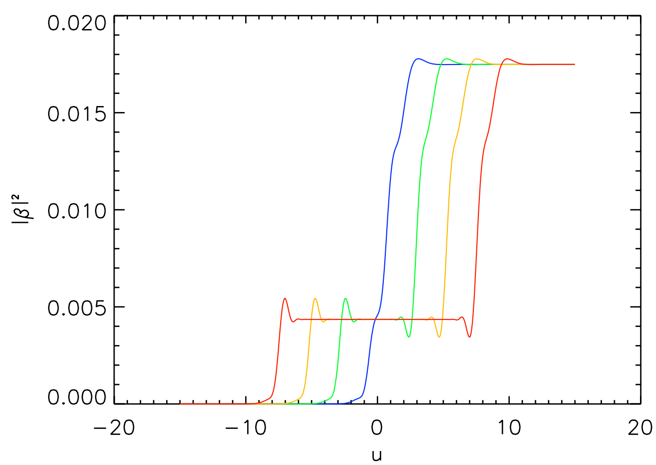


FIG. 6 (color online). Plotted is $|\beta_k(u)|^2$ for $m = H, \gamma = \sqrt{3}/2$ defined by (6.3) with the initial adiabatic matching time $u_0 = -15$ and the second order matching defined by (6.20). The innermost blue curve is for $k = 1$, the green for $k = 10$, the orange for $k = 100$ and the outermost red for $k = 1000$, the latter 3 values showing two clearly separated particle creation events. The values of $u_{k\gamma}$ given by (6.8) are 0.35, 2.31, 5.44, and 10.0, respectively, for these values of $k$ and $\gamma$. The asymptotic value of $|\beta_k|^2$ of all the curves for large $u$ is 0.01748 in agreement with (6.14b) for $N_k = 0$. The intermediate plateau is at 0.00435 in agreement with (6.12c).

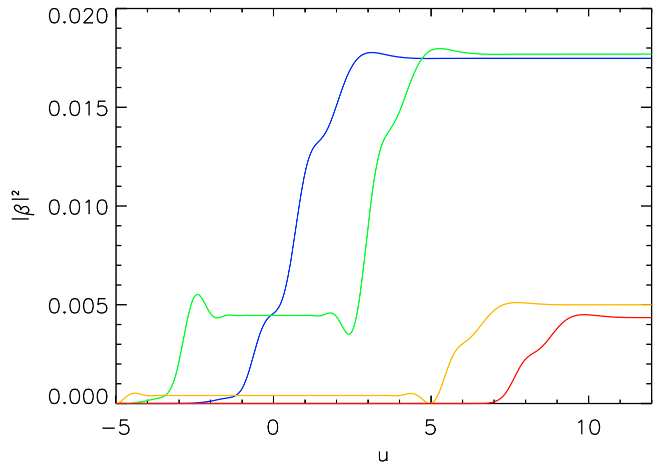


FIG. 7 (color online). Plotted is $|\beta_k(u)|^2$ for the same values of $\gamma$ and $k$ as in Fig. 6, but with the initial adiabatic matching time $u_0 = -5$. Note that for the two highest values of $k$ at 100 and 1000 (the lower orange and red curves), a marked first particle creation event does not occur since $u_0 > -u_{k\gamma}$ for these modes. The asymptotic value of the lower red curve at large $u$ is 0.00435 in agreement with the first term of (6.18). The yellow curve for $k = 100$ has a small contribution from the first creation event since $u_0$ and $-u_{k\gamma}$ are comparable.

$$V_k^{(1)} = -\frac{\dot\omega_k}{\omega_k} = h\left(1 - \frac{m^2}{\omega_k^2}\right) \tag{6.20b}$$

correct up to second order in the adiabatic expansion. A comparison of $|\beta_k(u)|^2$ for this choice and the simpler choice

$$W_k^{(0)} = \Omega_k = H\left[\left(k^2 - \frac{1}{4}\right)\text{sech}^2 u + \gamma^2\right]^{\frac{1}{2}} \tag{6.21a}$$

$$V_k^{(1)} = -\frac{\dot\omega_k}{\omega_k} = h\left(1 - \frac{m^2}{\omega_k^2}\right) \tag{6.21b}$$

for the $k = 10$ mode and $u_0 = -15$ is shown in Fig. 8.

As in the electric field case (cf. Fig. 1) the detailed time structure of the creation event is different with different choices of $(W_k, V_k)$, but the qualitative features and asymptotic values (and intermediate plateau value) are independent of the choice. The second order WKB choice (6.20) suppresses the oscillations observed with the choice (6.21) and comes closer to the approximate step function description. As predicted by the previous WKB analysis and (6.18), the modes with $k = 1$ and $k = 10$ in Fig. 7 go through both creation events with a rapid increase in $|\beta_k(u)|^2$ occurring for each at the appropriate value of $\mp u_{k\gamma}$. The modes with $k = 100$ and $k = 1000$ for which $u_{k\gamma} > |u_0|$ only go through a marked second creation event, although the orange curve for $k = 100$ has a small contribution from the first creation event, since $u_0 = -5$ and $-u_{k\gamma} = -5.44$ are comparable for this mode. The value of $|\beta_k(u)|^2$ after the first creation event is well approximated

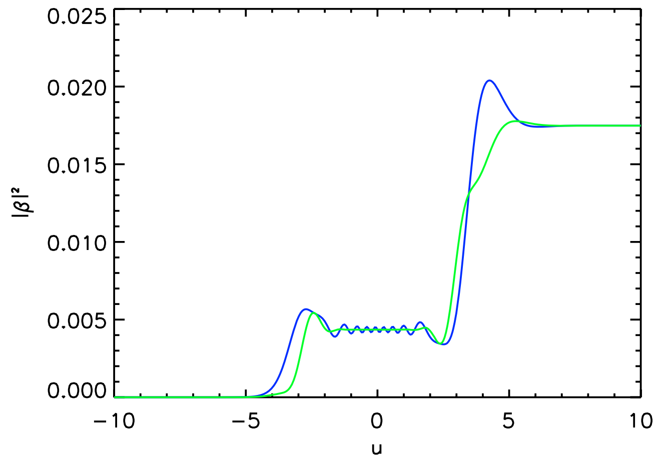


FIG. 8 (color online). Plotted is $|\beta_k(u)|^2$ for $k = 10$ and for the same value of $\gamma = \sqrt{3}/2$ as in Fig. 6, when the matching time is $u_0 = -15$. The blue curve with the larger oscillations corresponds to the zeroth order adiabatic vacuum state specified by $W_k = \Omega_k$ with $V_k$ given by (6.21). The green curve corresponds to the second order adiabatic vacuum state specified by (6.20).

by (6.12a) or (6.12c) with $N_k = 0$ for an initial vacuum, while for those modes undergoing two creation events the second plateau of $|\beta_k(u)|^2$ for $u > u_{k\gamma}$ is given by (6.14b). In Fig. 9 we also compare $|\beta_k(u)|^2$ for fixed $k = 1000$ and $u_0 = -15$, for three different values of the mass, showing the dependence of the time of the creation events on $\gamma$ given by (6.8b).

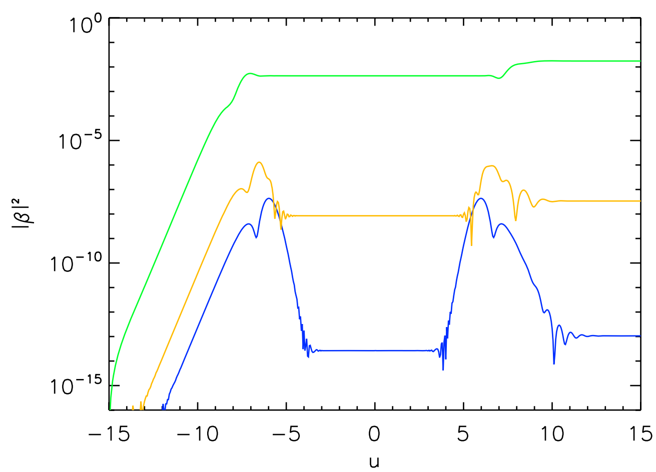


FIG. 9 (color online). Plotted is $|\beta_k(u)|^2$ for fixed $k = 1000$ and adiabatic matching time $u_0 = -15$, for three values of the mass: $m = H$ (upper, green), $m = 3H$ (middle, orange), $m = 5H$ (lower, blue) for $W_{\mathbf{k}}, V_{\mathbf{k}}$ specified by (6.20). Note the logarithmic scale. The asymptotic values of $|\beta_k|^2$ for large $u$ are $1.748 \times 10^{-2}$, $3.391 \times 10^{-8}$ and $1.062 \times 10^{-13}$, respectively, in agreement with (6.14b) for $N_k = 0$. The intermediate plateaux at $u = 0$ are at $4.35 \times 10^{-3}$, $8.48 \times 10^{-9}$, and $2.66 \times 10^{-14}$, respectively, in agreement with (6.12c). The particle creation events occur at $\mp u_{k\gamma}$ with $u_{k\gamma} = 7.74, 6.52$ and $6.00$, respectively, for the three values of $m$. The roughness of the blue and orange curves near the lower left corner is due to roundoff error.

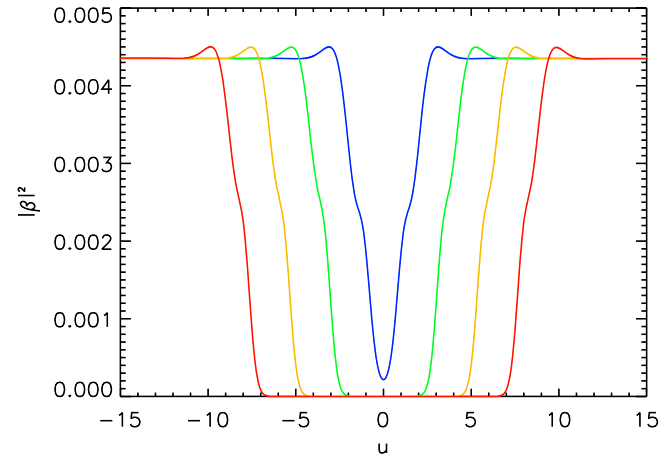


FIG. 10 (color online). Plotted is $|\beta_k(u)|^2$ for the CTBD state for various values of $k$: $k = 1$ (blue), $k = 10$ (green). $k = 100$ (orange), $k = 1000$ (red) for $W_{\mathbf{k}}, V_{\mathbf{k}}$ specified by (6.20). Note that the first event at $u = -u_{k\gamma}$ is a particle *annihilation* event in which the particle number decreases from 0.004352 to zero in each $k$ mode, rising again to the same value at $u = +u_{k\gamma}$ in a completely time symmetric manner.

For comparison we also plot the adiabatic particle number as a function of $u$ for the CTBD state in Fig. 10. This figure shows that the CTBD state contains particles in its initial condition and the first event at $u = -u_{k\gamma}$ is actually a particle *destruction* event. The phase coherent initial particles in modes with principal quantum number $k$ find each other and annihilate at the time $u = -u_{k\gamma}$, canceling each other precisely in region II. At the later time $u = u_{k\gamma}$, these particles are created again in a completely time symmetric manner. This is clearly a delicately balanced coherent process that is artificially arranged by initial conditions in the CTBD state. In an accompanying paper we show that a small perturbation of the CTBD state upsets this balance and leads again to instability [21].

## VII. PARTICLE CREATION IN SPATIALLY FLAT FLRW POINCARÉ COORDINATES

The analysis of particle creation in the spatially closed $\mathbb{S}^3$ coordinates of de Sitter space of the previous section can just as well be carried out in the spatially flat Poincaré coordinates of (A11), more commonly used in cosmology. The wave equation (2.1) again separates in the usual Fourier basis $\Phi \sim \phi_k(\tau) e^{i\mathbf{k}\cdot\mathbf{x}}$. Removing the scale factor by defining the mode function $f_k = a^{\frac{3}{2}}\phi_k$ as in (3.1) but with $a = \exp(H\tau)$ in this case gives the mode equation

$$\left[\frac{d^2}{d\tau^2} + k^2 e^{-2H\tau} + m^2 - \frac{H^2}{4}\right] f_k(\tau) = 0, \tag{7.1}$$

with $k \equiv |\mathbf{k}|$. This equation again has the form of an harmonic oscillator equation with a time dependent frequency which is given by

$$\omega_k^2(\tau) = k^2 e^{-2H\tau} + m^2 - \frac{H^2}{4}. \tag{7.2}$$

Thus with this change all of the methods employed in the spatially closed sections or the electric field background may be utilized again. In particular, for $\gamma^2 > 0$ we have over the barrier scattering in a nontrivial one-dimensional potential, and we should expect the stationary waves incident from the left as $\tau \to -\infty$ to be partially reflected and partially transmitted to the right as $\tau \to \infty$. This scattering will result again in a nontrivial Bogoliubov transformation between the positive frequency particle solutions at early times in the $|in\rangle$ vacuum and those at late times in the $|out\rangle$ vacuum, i.e. particle creation, just as in the electric field case.

By making the change of variables,

$$z \equiv \frac{k}{H} e^{-H\tau}, \tag{7.3}$$

Eq. (7.1) may be transformed into Bessel's equation with imaginary index $\nu = \pm i\gamma$. Thus the exact solutions are Bessel or Hankel functions with this index. The particular solution

$$\begin{aligned} \bar{v}_\gamma(z) &\equiv \frac{1}{2}\sqrt{\frac{\pi}{H}} e^{-\frac{\pi\gamma}{2}} e^{\frac{i\pi}{4}} H^{(1)}_{i\gamma}(z) \\ &= \sqrt{\frac{\pi}{H}} \frac{e^{\frac{\pi\gamma}{2}} e^{\frac{i\pi}{4}}}{e^{2\pi\gamma} - 1} [e^{\pi\gamma} J_{i\gamma}(z) - J_{-i\gamma}(z)] \end{aligned} \tag{7.4}$$

is the CTBD solution in flat coordinates, with the asymptotic behavior

$$\bar{v}_\gamma(z) \to \frac{1}{\sqrt{2Hz}} e^{iz} \quad \text{as} \quad z \to \infty. \tag{7.5}$$

Since $\omega_k = H\sqrt{z^2 + \gamma^2} \to Hz$ and

$$\begin{aligned} \Theta_\gamma(z) &\equiv \int^{\tau(z)} d\tau \omega_k(\tau) \\ &= -\int^z \frac{dz}{z}\sqrt{z^2+\gamma^2} \\ &= -\sqrt{z^2+\gamma^2} - \frac{\gamma}{2}\ln\left[\frac{\sqrt{z^2+\gamma^2}-\gamma}{\sqrt{z^2+\gamma^2}+\gamma}\right] \\ &\to -z \quad \text{as} \quad z \to \infty, \end{aligned} \tag{7.6}$$

the solution (7.4) is also the correctly normalized adiabatic *in* vacuum solution,

$$f_{k(+)}(\tau) = \bar{v}_\gamma(z) \to \frac{1}{\sqrt{2\omega_k}} e^{-i\Theta_\gamma} \quad \text{as} \quad \tau \to -\infty, \tag{7.7}$$

in the Poincaré coordinates.

This much is standard and may be found in standard references [8]. However, in the opposite limit of late times $\tau \to \infty$,

$$\omega_k \to H\gamma \quad \text{and} \quad \Theta_\gamma(z) \to -\gamma \ln z \quad \text{as} \quad z \to 0, \tag{7.8}$$

and therefore

$$\begin{aligned} f_k^{(+)}(\tau) &= \left(\frac{2H}{k}\right)^{i\gamma} \frac{\Gamma(1+i\gamma)}{\sqrt{2H\gamma}} J_{i\gamma}(z) \to \left(\frac{H}{k}\right)^{i\gamma} \frac{z^{i\gamma}}{\sqrt{2H\gamma}} \\ &= \frac{1}{\sqrt{2H\gamma}} e^{-i\gamma H\tau} \quad \text{as} \quad \tau \to \infty \end{aligned} \tag{7.9}$$

is the properly normalized positive frequency *out* solution, which agrees with the adiabatic form $e^{-i\Theta_\gamma}/\sqrt{2\omega_k}$ at late times. Comparison with the last form of (7.4) shows that indeed there is a nontrivial mixing of positive and negative frequencies at late times in the CTBD state. The Bogoliubov coefficients are

$$A_\gamma = \frac{\sqrt{2\pi\gamma} e^{\frac{i\pi}{4}}}{2^{i\gamma}\Gamma(1+i\gamma)} \frac{e^{\frac{3\pi\gamma}{2}}}{e^{2\pi\gamma} - 1} \tag{7.10a}$$

$$B_\gamma = -\frac{\sqrt{2\pi\gamma} e^{\frac{i\pi}{4}}}{2^{i\gamma}\Gamma(1+i\gamma)} \frac{e^{\frac{\pi\gamma}{2}}}{e^{2\pi\gamma} - 1}, \tag{7.10b}$$

with $|A_\gamma|^2 - |B_\gamma|^2 = 1$. Note that

$$|B_\gamma|^2 = \frac{1}{e^{2\pi\gamma} - 1} = |B^{out}_{k\gamma}|^2. \tag{7.11}$$

Thus $B_k$ has exactly the same magnitude as the corresponding Bogoliubov coefficient Eq. (3.20b) obtained previously in the closed $\mathbb{S}^3$ spatial sections. This equality is to be expected since in the asymptotic future the closed spatial sections have negligible spatial curvature and there is no local difference with the flat sections. This mode mixing and particle creation effect in the flat Poincaré coordinates, which follows from the second form of (7.4), seems to have been overlooked in [8], which states that "there is no particle production."

The particle creation process may be analyzed in real time in the flat Poincaré coordinates by using the methods of Sec. V. Indeed the zeroes of (7.2) in the complex $z$ plane occur at

$$z_\gamma = \pm i\gamma, \tag{7.12}$$

which represents an infinite series of zeroes at $H\tau = \ln(k/H\gamma) + i\pi(n + \frac{1}{2})$ similar to those in the complex $u$ plane in Eq. (6.8) and Fig. 5. Since the Poincaré coordinates cover only one half of the de Sitter manifold, where it is only expanding (or in the other half where it is only contracting), there is only one line of complex zeroes in Poincaré coordinates and only one creation event occurring at

$$\tau_{k\gamma} = \frac{1}{H}\ln\left(\frac{k}{H\gamma}\right) \tag{7.13}$$

for the mode with Fourier component $k = |\mathbf{k}|$: compare (6.9b). If one starts the evolution at some finite initial time $\tau_0$ then only those modes with $k$ in the range determined by

$$\tau_0 < \tau_{k\gamma} < \tau \tag{7.14}$$

will experience their creation event at a later time $\tau$. The number density of particles in modes with $k = |\mathbf{k}|$ at time $\tau$ is therefore

$$\mathcal{N}_k(\tau) \approx \theta(\tau_{k\gamma} - \tau_0)\theta(\tau - \tau_{k\gamma})(1 + 2N_k)|B_\gamma|^2 \tag{7.15}$$

in the approximation that the particles are created instantaneously when $\tau$ passes through $\tau_{k\gamma}$.

In the expanding phase of de Sitter space, whether described by closed $\mathbb{S}^3$ or flat $\mathbb{R}^3$ spatial sections, these particles will be redshifted in energy and make a decreasing contribution to the energy density and pressure at later times. In the next section we compute the energy density of the created particles which grow exponentially in the contracting part of de Sitter space due to their blueshifting toward the extreme ultrarelativistic limit. This does not occur in the Poincaré sections with only monotonic expansion, *for spatially homogeneous states*. In [18] we showed that the energy density and pressure relax to the values of the de Sitter invariant CTBD state for all such UV allowed spatially homogeneous states and for all $M^2 > 0$ in the expanding phase. Nevertheless because of the same kind of nontrivial Bogoliubov mixing between the $|in\rangle$ and $|out\rangle$ states in the Poincaré coordinates (7.10), a calculation of the decay rate analogous to that in Sec. III would show the same type of instability to particle creation as in the closed spatial $\mathbb{S}^3$ sections of the full hyperboloid.

## VIII. STRESS-ENERGY TENSOR OF CREATED PARTICLES

In this section we consider the stress-energy tensor of the created particles, and their ability to affect the background de Sitter spacetime by backreaction. The energy-momentum tensor of the scalar field with arbitrary curvature coupling $\xi$ is

$$T_{ab} = (\nabla_a\Phi)(\nabla_b\Phi) - \frac{g_{ab}}{2}(\nabla^c\Phi)(\nabla_c\Phi) - \frac{m^2}{2}g_{ab}\Phi^2 + \xi[g_{ab}\Box - \nabla_a\nabla_b + G_{ab}]\Phi^2, \tag{8.1}$$

where $G_{ab}$ is the Einstein tensor. Assuming a metric of the form (2.2) and spatial homogeneity and isotropy of the state on the $\mathbb{S}^3$ sections, the only nonvanishing components of the expectation value of $T_{ab}$ are the energy density $\varepsilon = \langle T_{\tau\tau}\rangle$ and the isotropic pressure $p = \frac{1}{3}\langle T^i{}_i\rangle$. The scalar field operator $\Phi$ can be expressed in terms of the exact mode function solutions of (3.2) such that

$$\Phi(u,\hat{N}) = a^{-\frac{3}{2}}\sum_{k=1}^{\infty}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}\{a^f_{klm_l}f_k(u)Y_{klm_l}(\hat{N}) + a^{f\dagger}_{klm_l}f^*_k(u)Y^*_{klm_l}(\hat{N})\}, \tag{8.2}$$

where the $a^f_{klm_l}$ now correspond to the state specified by the particular solution $f_k(u)$ of (3.5), fixed by specific initial data at a finite time, and not necessarily the *in*, *out* or CTBD states. Specializing to conformal coupling $\xi = \frac{1}{6}$, we find

$$\varepsilon|_{\xi=\frac{1}{6}} = \frac{1}{4\pi^2a^3}\sum_{k=1}^{\infty}(1+2N_k)\left[|\dot{f}_k|^2 - h\,\mathrm{Re}(f^*_k\dot{f}_k) + \left(\frac{k^2}{a^2} + m^2 + \frac{h^2}{4}\right)|f_k|^2\right] \tag{8.3a}$$

$$p|_{\xi=\frac{1}{6}} = \frac{1}{12\pi^2a^3}\sum_{k=1}^{\infty}(1+2N_k)\left[|\dot{f}_k|^2 - h\,\mathrm{Re}(f^*_k\dot{f}_k) + \left(\frac{k^2}{a^2} - m^2 + \frac{h^2}{4}\right)|f_k|^2\right]. \tag{8.3b}$$

The exact mode functions $f_k$ and their time derivatives can be expressed in terms of the adiabatic functions $\tilde{f}_k$ and the time dependent Bogoliubov coefficients $(\alpha_k, \beta_k)$ by (5.10) and (6.2). Thus (8.3) may be expressed in the general adiabatic basis as

$$\varepsilon|_{\xi=\frac{1}{6}} = \frac{1}{2\pi^2a^3}\sum_{k=1}^{\infty}k^2\left[\varepsilon^{\mathcal{N}}_k\left(\mathcal{N}_k + \frac{1}{2}\right) + \varepsilon^{\mathcal{R}}_k\mathcal{R}_k + \varepsilon^{\mathcal{I}}_k\mathcal{I}_k\right] \tag{8.4a}$$

$$p|_{\xi=\frac{1}{6}} = \frac{1}{2\pi^2a^3}\sum_{k=1}^{\infty}k^2\left[p^{\mathcal{N}}_k\left(\mathcal{N}_k + \frac{1}{2}\right) + p^{\mathcal{R}}_k\mathcal{R}_k + p^{\mathcal{I}}_k\mathcal{I}_k\right], \tag{8.4b}$$

where the three terms labeled by $\mathcal{N}$, $\mathcal{R}$, and $\mathcal{I}$ are

$$\varepsilon^{\mathcal{N}}_k|_{\xi=\frac{1}{6}} = \frac{1}{2W_k}\left[\omega_k^2 + W_k^2 + \frac{(V_k - h)^2}{4}\right] \tag{8.5a}$$

$$\varepsilon^{\mathcal{R}}_k|_{\xi=\frac{1}{6}} = \frac{1}{2W_k}\left[\omega_k^2 - W_k^2 + \frac{(V_k - h)^2}{4}\right] \tag{8.5b}$$

$$\varepsilon^{\mathcal{I}}_k|_{\xi=\frac{1}{6}} = \frac{V_k - h}{2} \tag{8.5c}$$

in the energy density, and

$$p_k^{\mathcal{N}}|_{\xi=\frac{1}{6}} = \frac{1}{6W_k}\left[\omega_k^2 - 2m^2 + W_k^2 + \frac{(V_k-h)^2}{4}\right] \tag{8.6a}$$

$$p_k^{\mathcal{R}}|_{\xi=\frac{1}{6}} = \frac{1}{6W_k}\left[\omega_k^2 - 2m^2 - W_k^2 + \frac{(V_k-h)^2}{4}\right] \tag{8.6b}$$

$$p_k^{\mathcal{I}}|_{\xi=\frac{1}{6}} = \frac{V_k-h}{6}, \tag{8.6c}$$

in the pressure, with $\mathcal{N}_k$ given by (6.5), and $\mathcal{R}_k, \mathcal{I}_k$ given by

$$\mathcal{R}_k = (1+2N_k)\,\mathrm{Re}(\alpha_k\beta_k^* e^{-2i\Theta_k}) \tag{8.7a}$$

$$\mathcal{I}_k = (1+2N_k)\,\mathrm{Im}(\alpha_k\beta_k^* e^{-2i\Theta_k}) \tag{8.7b}$$

in terms of the adiabatic phase

$$\Theta_k \equiv \int_{\tau_0}^{\tau} d\tau\, W_k. \tag{8.8}$$

The $\mathcal{N}_k$ terms have a quasi-classical interpretation as the energy density and pressure of particles with single particle energies $\varepsilon_k^{\mathcal{N}}$. The $\frac{1}{2}$ in these terms has the natural interpretation of the quantum zero point energy in the adiabatic vacuum specified by $(W_k, V_k)$. The $\mathcal{R}_k$ and $\mathcal{I}_k$ terms are oscillatory quantum interference terms that have no classical analog, analogous to the last term of (5.21).

The mode sums over $k$ in (8.4) are generally quartically divergent in four dimensions. It is in handling and removing these divergent contributions in the mode sums that the adiabatic method is most useful [4,6–8,37,50,51]. Although a fourth order adiabatic subtraction is needed in general, when $\xi = \frac{1}{6}$ it is sufficient to subtract only the second order adiabatic expressions

$$\varepsilon^{(2)} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2 \varepsilon_k^{(2)} \tag{8.9a}$$

$$p^{(2)} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2 p_k^{(2)} \tag{8.9b}$$

with

$$\varepsilon_k^{(2)}|_{\xi=\frac{1}{6}} = \omega_k + \frac{h^2 m^4}{8\omega_k^5} \tag{8.10a}$$

$$p_k^{(2)}|_{\xi=\frac{1}{6}} = \frac{1}{3}\left[\omega_k - \frac{m^2}{\omega_k} - \frac{m^4}{8\omega_k^5}(2\dot{h}+5h^2) + \frac{5m^6h^2}{8\omega_k^7}\right] \tag{8.10b}$$

to arrive at a finite, renormalized and conserved stress tensor. The reason for this is that it may be shown that the only possible remaining divergence is logarithmic and proportional to $(\xi - \frac{1}{6})^2$, and correspondingly there are no $\omega_k^{-3}$ terms in either of the expressions (8.10) for conformal coupling $\xi = \frac{1}{6}$. Moreover the logarithmic divergence is proportional to the tensor ${}^{(1)}H_{ab}$ [8,37] which vanishes in de Sitter space (similar to the vanishing of the counterterm proportionl to $\ddot{E}$ when $E$ is a constant).

The difference of the $\frac{1}{2}$ vacuum-like $\mathcal{N}$ terms in (8.4) and the subtraction terms are

$$\varepsilon_{vac} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2(\varepsilon_k^{\mathcal{N}} - \varepsilon_k^{(2)}) \tag{8.11a}$$

$$p_{vac} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2(p_k^{\mathcal{N}} - p_k^{(2)}) \tag{8.11b}$$

with the summands

$$\varepsilon_k^{\mathcal{N}} - \varepsilon_k^{(2)} = \frac{1}{2W_k}\left[(W_k-\omega_k)^2 + \frac{(V_k-h)^2}{4}\right] - \frac{h^2m^4}{8\omega_k^5} \tag{8.12a}$$

$$p_k^{\mathcal{N}} - p_k^{(2)} = \frac{1}{3}(\varepsilon_k^{\mathcal{N}} - \varepsilon_k^{(2)}) + \frac{m^2}{3W_k\omega_k}(W_k-\omega_k) + \frac{m^4}{12\omega_k^5}(\dot{h}+3h^2) - \frac{5m^6h^2}{24\omega_k^7}. \tag{8.12b}$$

In order for the sums in the renormalized energy-momentum tensor expectation value, subtracted as in (8.11) to converge, it is sufficient for the summands (8.12) to fall off as $k^{-5}$ or faster at large $k$. This is the important physical condition on the definition of the adiabatic mode functions $(W_k, V_k)$, which restricts the choice of the adiabatic vacuum state. The last term of (8.12a) and the last two terms of (8.12b) already satisfy this condition. Inspection of the other terms in (8.12) shows that in order to satisfy this condition it is sufficient for

$$W_k - \omega_k = \mathcal{O}(k^{-3}) \quad \text{and} \quad V_k - h = \mathcal{O}(k^{-2}) \quad \text{as} \quad k\to\infty, \tag{8.13}$$

and then the sums in (8.11) will converge quadratically. Either of the choices (6.20) or (6.21) of the last section satisfy this condition. Let us emphasize that the choice of $(W_k, V_k)$ as functions of time only affects how the individual $\mathcal{N}$, $\mathcal{R}$ and $\mathcal{I}$ terms contribute to the stress tensor expectation value in (8.4), while the sum of all the contributions and the subtraction terms (8.9)–(8.10) are independent of that choice, once the initial state is specified by $f_k(u_0)$ and $\dot{f}_k(u_0)$.

Thus with the vacuum contributions subtracted (8.4) becomes

$$\varepsilon_R = \frac{1}{2\pi^2 a^3}\sum_{k=1}^{\infty} k^2\left[\varepsilon_k^{\mathcal{N}}\mathcal{N}_k + \varepsilon_k^{\mathcal{R}}\mathcal{R}_k + \varepsilon_k^{\mathcal{I}}\mathcal{I}_k\right] + \varepsilon_{vac} \tag{8.14a}$$

$$p_R = \frac{1}{2\pi^2 a^3}\sum_{k=1}^{\infty} k^2[p_k^{\mathcal{N}}\mathcal{N}_k + p_k^{\mathcal{R}}\mathcal{R}_k + p_k^{\mathcal{I}}\mathcal{I}_k] + p_{vac}. \tag{8.14b}$$

It should make very little difference which of the choices for $(W_k, V_k)$ one uses to define the instantaneous adiabatic vacuum and time dependent Bogoliubov coefficients, since they all fall off at large $k$, and will give qualitatively the same behavior for the particle creation effects when passing through the lines of complex zeroes in Fig. 5. The change in the plateau values $\Delta\mathcal{N}_{1,k\gamma}$ or $\Delta\mathcal{N}_{2,k\gamma}$ obtained after one or two creation events are the same for all definitions and the only difference is in the detailed time dependence during the creation "event" itself, and only for the lower $k$ modes.

In our actual numerical calculations we have used the full fourth order adiabatic subtraction as described in Ref. [51] in which the renormalization counterterms are separated into terms which are divergent and terms which are finite. The latter can be integrated to form an analytic contribution to the stress-energy tensor that is separately conserved. The full renormalized stress-energy tensor is then given by (8.14) with

$$\varepsilon_{vac} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2\left(\varepsilon_k^{\mathcal{N}} - \frac{k}{a} - \frac{m^2 a}{2k} + \frac{m^4 a^3}{8k^3}\right) + \varepsilon_{\rm an} \tag{8.15a}$$

$$p_{vac} = \frac{1}{4\pi^2 a^3}\sum_{k=1}^{\infty} k^2\left(p_k^{\mathcal{N}} - \frac{ka}{3} + \frac{m^2 a^2 k}{6} - \frac{m^4 a^3}{8k}\right) + p_{\rm an}. \tag{8.15b}$$

The analytic terms are given by

$$\varepsilon_{\rm an} = \frac{1}{2880\pi^2}\left(\frac{6\dddot{a}\dot{a}}{a^2} + \frac{6\ddot{a}\dot{a}^2}{a^3} - \frac{3\ddot{a}^2}{a^2} - \frac{6\dot{a}^4}{a^4} + \frac{6}{a^4}\right) - \frac{m^2}{96\pi^2}\left(\frac{\dot{a}^2}{a^2} + \frac{1}{a^2}\right) - \frac{m^4}{64\pi^2}\left[\frac{1}{2} + \log\left(\frac{m^2 a^2}{4}\right) + 2C\right] \tag{8.16a}$$

$$p_{\rm an} = \frac{1}{2880\pi^2}\left(-\frac{2\ddddot{a}}{a} - \frac{4\dddot{a}\dot{a}}{a^2} + \frac{8\ddot{a}\dot{a}^2}{a^3} - \frac{3\ddot{a}^2}{a^2} - \frac{2\dot{a}^4}{a^4} + \frac{2}{a^4}\right) + \frac{m^2}{288\pi^2}\left(\frac{2\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{1}{a^2}\right) + \frac{m^4}{64\pi^2}\left[\frac{7}{6} + \log\left(\frac{m^2 a^2}{4}\right) + 2C\right], \tag{8.16b}$$

with $C$ Euler's constant. This differs from the vacuum subtraction in (8.12) by finite terms, which one can check remain small for all times in de Sitter space.

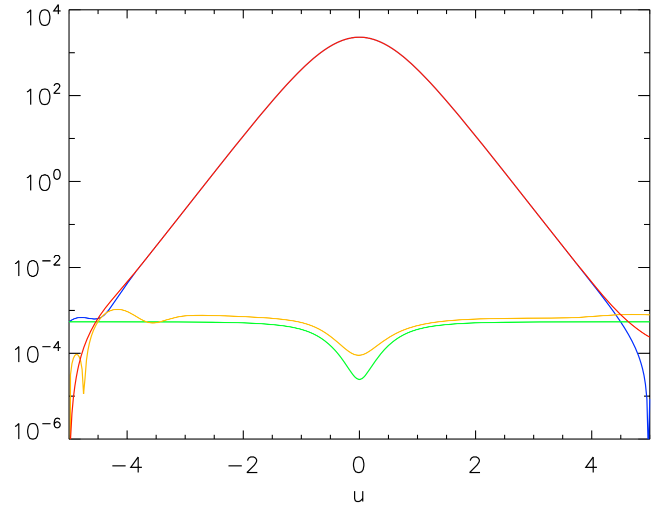


FIG. 11 (color online). The absolute values of various contributions in (8.14a) to the energy density for an adiabatic state when $m = H$ are shown in units of $H^4$ for a matching time of $u_0 = -5$ with $W_{\mathbf{k}}$, $V_{\mathbf{k}}$ specified by (6.20). The blue curve is the total energy density, while the red curve which is nearly coincident with it for most values of $u$, is the contribution from the $\mathcal{N}_k$ adiabatic particle term. Note the logarithmic scale and large growth of these at intermediate $u$. The lower two curves are the contributions to (8.14a) of the $\varepsilon_{vac}$ (in green), and the sum of the $\mathcal{R}_k$ and the $\mathcal{I}_k$ quantum interference terms (in orange), which remain small for all times shown.

One way to assess the usefulness of the particle description is to analyze its contribution to the stress-energy tensor. This is done in Fig. 11 where the full energy density and that due to the various terms in the energy density (8.14) are plotted. It is clear from the plot that near $u = 0$ the $\varepsilon_k^{\mathcal{N}}\mathcal{N}_k$ term provides by far the dominant contribution to the stress-energy tensor, whose total value depends only upon the initial data, while the interference $\mathcal{R}_k$ and $\mathcal{I}_k$ terms are very much smaller. At very early times and very late times this is not the case. At early times this is expected since the particle definition is designed to give a vacuum state at the matching time. At late times it is also expected since the energy density of the particles redshifts away.

From the window function (6.13) in the contracting phase of de Sitter space the energy density for the initial adiabatic vacuum matched by (6.7) at $u = u_0$ should behave like

$$\varepsilon_R \simeq \frac{1}{2\pi^2 a^3}\sum_{k=K_\gamma(u)}^{K_\gamma(u_0)} k^2 \varepsilon_k^{\mathcal{N}}\Delta\mathcal{N}_{1k\gamma} \simeq \frac{1}{8\pi^2}\frac{K_\gamma^4(u_0)}{a^4(u)}\frac{1}{e^{8\pi\gamma}-1} \simeq \frac{\gamma^4}{8\pi^2}\frac{1}{e^{2\pi\gamma}-1}\frac{\sinh^4 u_0}{a^4(u)} \tag{8.17}$$

and therefore grow exponentially with the fourth power of the scale factor as $a(u)$ decreases, consistent with our discussion above of the highest $k$ modes contributing in the window $K(u) < k < K(u_0)$ in the contracting phase, where their effects on the stress tensor are blueshifted, becoming highly relativistic. This is what accounts for the enormous growth of the particle contribution to the energy density $\varepsilon_k^{\mathcal{N}}$ in Fig. 11 as it rapidly dominates the quantum $\varepsilon_k^{\mathcal{R}}, \varepsilon_k^{\mathcal{I}}$ and vacuum terms in (8.14). This is completely analogous to the secular growth of the current in a constant uniform electric field starting from the adiabatic vacuum plotted in Fig. 4. To check the $a^{-4}$ relativistic dependence predicted by (8.17) we plot the energy density due to the particles multiplied by $a^4(u)$ in Fig. 12. As expected the resulting quantity is approximately constant for a large range of $u$, deviating from this behavior only for initial values of $u$ of order $u_0 < 0$ and again as the particles are redshifted away at $u$ of order $|u_0|$.

The estimate (8.17) also predicts that the maximum value of the energy density at the symmetric point $u = 0$ varies with the fourth power of $\sinh u_0$. In Fig. 13 the energy density is plotted for two different adiabatic matching times. It is clear from the plots that the maximum energy density at $u = 0$ is substantially larger for the earlier adiabatic matching time, consistent with (8.17). This is expected since an earlier matching time allows more modes to go through the first particle creation phase and increases the upper limit $K_\gamma(u_0)$ in (8.17).

These results show that the adiabatic particle definition is a very useful one, since its contribution to the energy density dominates when the particles become ultrarelativistic, that the energy density of the created particles grows exponentially in the contracting phase of de Sitter space, and most importantly that the maximum of the energy density also grows exponentially with the initial time as $u_0 \to -\infty$. This shows conclusively that global or "eternal" de Sitter space is unstable to particle creation, as the arbitrarily large energy densities of the particles will necessarily lead to a very large backreaction on the classical spacetime when used as a source for the semiclassical Einstein equations, especially as $u_0 \to -\infty$.

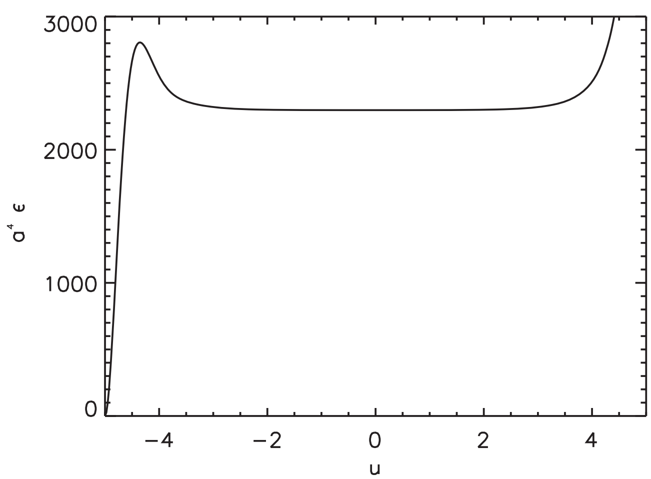


FIG. 12. The product of the fourth power of the scale factor and the energy density due to the particles for an adiabatic state when $m = H$ is shown. The matching time for the adiabatic state matching time of $u_0 = -5$ and $W_{\mathbf{k}}$ and $V_{\mathbf{k}}$ are specified by (6.20).

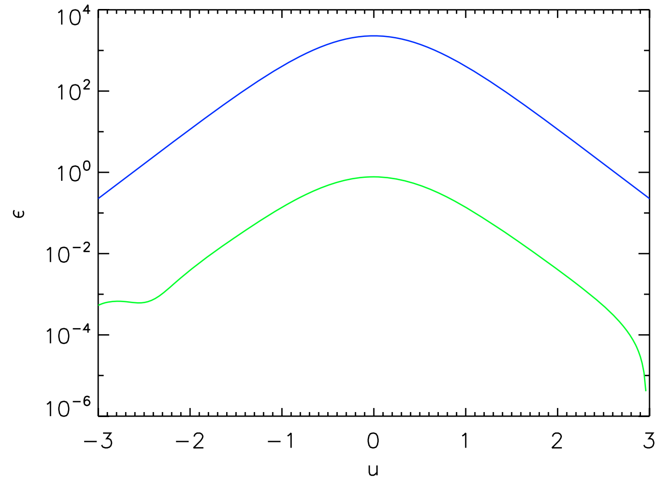


FIG. 13 (color online). The energy density for an adiabatic state when $m = H$ is shown in units of $H^4$ for an adiabatic matching time of $u_0 = -3$ (green lower curve) and a matching time of $u_0 = -5$ (blue upper curve). For both curves $W_{\mathbf{k}}$ and $V_{\mathbf{k}}$ are specified by (6.20). The exponential dependence on the time $u_0$, predicted by Eq. (8.17), is evident.

We note that this instability has a classical analog. The process of pair creation spontaneously from the vacuum is purely quantum in nature, but any spontaneous emission process is always accompanied by induced emission, which can be viewed in classical terms. Choosing the adiabatic vacuum, which is as empty as possible of excitations in the far past, shows that global de Sitter space is unstable to this spontaneous quantum process. The computation of the energy-momentum tensor in the adiabatic vacuum state, its growth like $a^{-4}$ in the contracting phase, and the dominance of the particle creation term $|\beta_k|^2$ in (6.5) and (8.4), cf. Fig. 11, show clearly the large de Sitter noninvariant energy density estimated in (8.17) that is generated by the spontaneous particle creation effect. The induced or classical effect is associated with the initial and constant $N_k$ term (6.6) in (6.5), quite apart from the spontaneous effect associated with $|\beta_k|^2$. Such an $N_k > 0$ term only makes the coefficient of the $a^{-4}$ growth and energy density at $u = 0$ larger. This would generate even larger deviations from de Sitter space if included in backreaction, before the expanding phase ever begins. However if $N_k = 0$, there is no classical perturbation, and the instability is purely quantum in nature.

That the effect grows in the contracting phase of de Sitter space, when modes are being blueshifted to the ultrarelativistic regime, and the large $k$ UV part of the mode sum dominates, as opposed to the expanding phase when the particles are redshifted and the low $k$ modes dominate is not surprising. To see this effect properly one needs to start with the adiabatic basis at a finite (early) initial time. We observe here one major difference between the electric field and de Sitter cases in the basic kinematics. The electric field is a vector field and uniformly accelerates all charged particles of a given charge in one direction. Particles (virtual or real) with initially negative kinetic momenta along the direction of the electric field are eventually brought momentarily to rest, and then turn around with continually increasing positive kinetic momentum ever after. It is these modes in the quantum theory that undergo particle creation at the turnaround time, and make the secular contribution to the current as their kinetic momentum and energy grow without bound and the corresponding particles approach the speed of light. Thus this late time contribution is clearly relativistic and UV dominated. There is only one creation event for each wave number mode.

On the other hand in the de Sitter case the physical or kinetic momentum is $p = k/a$ which is isotropic, depending only on the magnitude of $\mathbf{k}$ and not its direction in the spatially homogeneous states we are considering. There are two creation events for each $k$ mode, one in the contracting phase of de Sitter space, the second in the expanding phase. The first creation event is quite analogous to the electric field case in that once created the particles are blueshifted (exponentially in this case), rapidly becoming ultrarelativistic and making a contribution to the energy density and pressure that grows like $a^{-4}$, typical of ultrarelativistic particles in the contracting phase of de Sitter space. In the expanding phase of de Sitter space the situation is reversed, the created particles in each $k$ mode are redshifted, and make a decreasing contribution to the stress tensor, certainly for spatially homogeneous states, so that even a steady rate of particle creation cannot produce an effect in the stress tensor that is secularly growing in time. Instead the vacuum and other $\mathcal{R}$ and $\mathcal{I}$ interference terms in (8.14) remain of the same order as the particle creation term at late times and together their sum approaches the de Sitter invariant CTBD value [18]. There is no exact analog of this second behavior in the electric field case.

## IX. ADIABATIC SWITCHING ON OF DE SITTER SPACE AND THE IN VACUUM

In the case of the spatially uniform electric field there is an exactly soluble problem in which the electric field is adiabatically switched on and off according to the profile

$$A_z(t) = -ET\tanh(t/T), \qquad \mathbf{E}(t) = E\hat{\mathbf{z}}\,\mathrm{sech}^2(t/T). \quad (9.1)$$

By taking the limit $T \to \infty$ at the end of the calculation, the constant uniform field (4.2) is recovered [29,30,35]. On the other hand if $T$ is finite and $t \to \pm\infty$ the electric field goes to zero exponentially rapidly, and the particle and antiparticle solutions are the standard Minkowski ones (with $p = k_z \pm eET$), so there is no doubt that $|in\rangle$ and $|out\rangle$ states may be identified with the Minkowski vacuum and the excitations above that state correspond to the usual definition of particles.

No such analytically soluble model for de Sitter space is known. However there is no difficulty in studying the time profile of the scale factor,

$$a_T(u) = \frac{1}{H}\cosh\left[HT\tanh\left(\frac{u}{HT}\right)\right], \qquad (9.2)$$

in the line element (2.2) and the associated solutions of the scalar field mode equation (3.2) by numerical methods. With this profile at fixed $T$, in the infinite time limits $u \to \mp\infty$, $a(u)$ approaches the constant $H^{-1}\cosh(HT)$ so that the spacetime is static and the particle number is uniquely defined by the asymptotic constant positive and negative frequency functions. On the other hand for $HT \to \infty$ at fixed $u$, $a_T(u)$ approaches the de Sitter scale factor $H^{-1}\cosh u$. Thus (9.2) interpolates between static spacetimes in both the remote past and remote future with a symmetrical de Sitter contracting and expanding phase in between.

In Fig. 14 we show a comparison of the time dependent adiabatic particle number $|\beta_k(u)|^2$ for both the fixed de Sitter and the adiabatic switched metric of the form (9.2), for two representative values of $k$. The similarity of the curves for the two models with the same value of $k$ shows that essentially the *same* particle production process takes place in the fixed de Sitter space background or with the adiabatic switching on of the background from a static initial metric according to (9.2), at least for a large range of $k$ when $T$ is large enough. This supports the choice of the $|in\rangle$ state and positive frequency mode function (3.16a) in

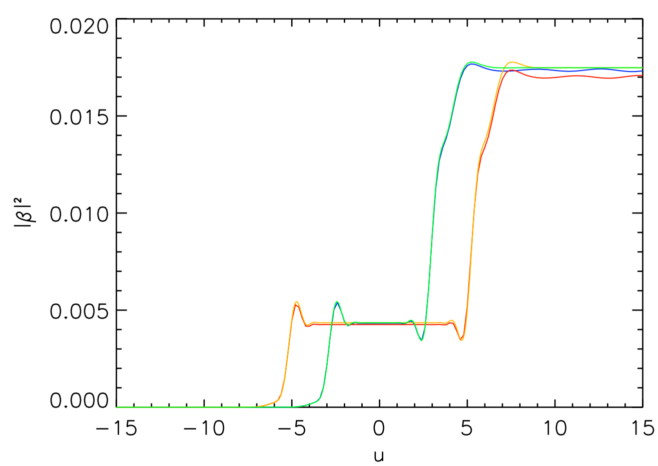


FIG. 14 (color online). Comparison of $|\beta_k(u)|^2$ for the $|in\rangle$ vacuum in exact de Sitter space for matching time $u_0 = -15$, and that for the smooth adiabatic switching on of de Sitter space according to the time profile (9.2) for $HT = 100$. Plotted is $|\beta_k(u)|^2$ for $k = 10$, $u_0 = -15$ (green, exact and blue, $HT = 100$), and $|\beta_k(u)|^2$ for $k = 100$, $u_0 = -15$ $HT = 100$ (orange, exact and red, $HT = 100$). If the curves for $k = 10$ and $k = 100$ were plotted for $HT = 1000$, they would be indistinguishable on the scale of the plot from the green and orange curves.

de Sitter space, as that one selected by turning on the de Sitter background adiabatically. Certainly the state produced in this way is very different from the maximally symmetric CTBD state defined by analytic continuation from Euclidean $\mathbb{S}^4$, in its low momentum modes, as expected by our analysis of adiabatic vacua and particle creation: compare Fig. 10. A detailed study of the state and subsequent evolution produced by (9.2) and other adiabatic profile functions will be presented in a subsequent publication [48].

## X. SUMMARY AND DISCUSSION

In this paper we have presented a detailed study of the spontaneous particle production of a massive free field theory in geodesically complete de Sitter space in real time. It is this spontaneous production of particles from the vacuum that is the basis for the instability of global de Sitter space. The formulation of particle creation as an harmonic oscillator with a time dependent frequency, or equivalently, as a one-dimensional stationary state scattering problem, determines the $|in\rangle$ and $|out\rangle$ positive energy particle states for massive fields in de Sitter space. This emphasizes the very close analogy with the spontaneous creation of charged particle/antiparticle pairs in a uniform, constant electric field. In each case the background gravitational or electric field configuration is symmetric under time reversal. Hence in each case it is possible to find a time reversal symmetric state in which no net particle creation occurs, and for which the imaginary part of the one-loop effective action and the decay rate vanish identically [43]. Such a maximally symmetric state allows a "self-consistent" solution to the semiclassical Maxwell or Einstein equations, with vanishing electric current or de Sitter invariant stress tensor. The artificiality of such a time symmetric state is apparent in the stationary scattering formulation, since it corresponds to choosing a very special coherent superposition of positive and negative frequency scattering solutions globally, which exactly cancels each spontaneous particle creation event by a time reversed particle annihilation event, cf. Fig. 10. This corresponds to adjusting the state of the quantum field to contain just as many pairs coming in from infinity and with precisely the right phase relations between them, so as to exactly cancel the electric currents or stress-energies of the pairs being spontaneously produced by the electric or de Sitter backgrounds. This is clearly not a true vacuum state in either case.

In situations such as these, the extension of the concepts of *particle* and *vacuum* from flat Minkowski space with no background fields must be reconsidered carefully. The essential generalization of the Feynman prescription of particles propagating forward in time and antiparticles propagating backward in time is to define $|in\rangle$ and $|out\rangle$ vacuum states corresponding to the choice of pure positive frequency modes (5.1) at intermediate times which are asymptotic to the exact particle *in* and *out* solutions (3.19) of the oscillator equation (3.2) in the remote past and remote future. Mathematically this is the condition that the positive frequency particle modes are analytic functions of $m^2$ in the lower or upper complex $m^2$ plane which are regular as $t \to \pm\infty$, respectively. This is the condition which also corresponds to the Schwinger-DeWitt method and choice of proper time contour. This should settle the question of whether the effective action in de Sitter space is real or imaginary due to particle creation effects [52].

We have provided evidence in Sec. IX that the $|in\rangle$ state is also the state obtained by turning the background fields on and off again according to a finite time $T$ parameter which may be taken to infinity at the end of all calculations. By any of these equivalent methods one obtains the standard Schwinger decay rate (4.26) for scalar charged particle creation in a constant, uniform electric field. By applying these same two methods to the background gravitational field of de Sitter space, one obtains the vacuum persistence amplitude and decay rate (3.32). Hence global de Sitter space is unstable to particle creation for the same reason as a constant, uniform electric field is in electrodynamics. This provides a mechanism for the relaxation of vacuum energy into matter or radiation and at least one possible route to the solution of the vacuum energy problem relying only upon known physics [10,11].

Although the definition of the adiabatic particle number (5.8) necessarily comes with some ambiguity in a time dependent background, and depends upon two frequency functions $(W_k, V_k)$ in (5.11) that are not unique, their choice is highly constrained by the requirements of rendering the vacuum zero-point contributions to physical currents, and the energy density and pressure (8.11) ultraviolet finite. The detailed time profile of the particle number depends upon the particular choice of adiabatic particle number through $(W_k, V_k)$, but the qualitative features and the asymptotic values in either the constant electric field or de Sitter cases do not depend upon this choice, cf. Figs. 1 and 8. The rapid change in the adiabatic particle number around "creation events" can be understood from the location of the zeroes of the adiabatic frequency function in the complex time plane, cf. (5.14) and Figs. 2–3 in the electric field case and (6.8) and Figs. 5–9 for de Sitter space. Some adiabatic particle number definition of this kind is certainly necessary to make the transition from QFT to kinetic theory, since the Boltzmann equation assumes that particle numbers and densities can be defined in the classical limit.

The usefulness of the particle concept is seen in the evaluation of expectation values of currents and stress tensors, and particularly in their secular terms, which are most important for backreaction. The "window function" (5.19) of pair creation in the electric field background accounts very well for the linear secular growth in time of the current of the produced pairs in Fig. 4 [36]. In the de Sitter case the corresponding window function (6.13) accounts very well for the exponential growth of the energy density and pressure of the created particles in the contracting phase, cf. Eq. (8.17) and Fig. 11. The linear growth in time in the first case and exponential $a^{-4}$ growth in time in the contracting de Sitter case, Fig. 12 are just what should be expected as the created particles are accelerated and rapidly become ultrarelativistic. Because of this secular growth backreaction effects on the background electric field must be taken into account through the semiclassical Maxwell equations. Likewise for the de Sitter geometry for early enough matching times $u_0$, cf. (8.17), backreaction effects similarly must be taken into account through the semiclassical Einstein equations.

This detailed study of particle creation removes a possible objection to the use of the strictly asymptotic $|in\rangle$ and $|out\rangle$ states in [9] to calculate the decay rate of de Sitter space, namely that these states are not Hadamard UV allowed states as defined e.g. in [53]. Instead they are members of the one parameter family of $\alpha$ vacua invariant under the $SO(4,1)$ subgroup of the de Sitter group continuously connected to the identity [9,44], although not the discrete $\mathbb{Z}_2$ inversion symmetry (A4). The difficulty is removed by the recognition that the large $|u|$ time and large $k$ limits do not commute. If one starts with UV allowed states such as the adiabatic initial state (6.7) at a finite initial time, and evolves forward for a finite time, only a finite number of modes experience particle creation events, according to the appropriate window function. The non-Hadamard $\alpha$ vacua are produced only in the improper limit of $|u| \to \infty$ in *eternal* de Sitter space, never in any finite time starting with UV finite initial data. Although the result for the decay rate (3.32) with an appropriate physical cutoff is the same, only a detailed analysis of the $k$ and time dependence allows a description free of any spurious UV problems and focuses attention on the resulting necessary breaking of de Sitter invariance instead [10,54].

In the electric field case it is generally accepted that particle creation will lead to eventual shorting of the electric field, although to date in four dimensions this process has only been studied in a large $N$ semiclassical approximation [34,55], which is not adequate to show the true long time behavior of the system even in QED. This depends upon self-interactions, and the long time behavior of correlation functions that are not accessible to the standard weak coupling approximations. Such processes involving multiple interactions in a medium, possibly very far from equilibrium, are generally described in many-body physics in the kinetic Boltzmann equation approximation, where all time reversal invariance properties of the underlying QFT are lost, and irreversible behavior is expected. What is perhaps less widely appreciated is that this breaking of time reversal invariance has its roots in the definition of the vacuum itself and the distinction between particles and antiparticles in QFT by the $m^2 - i0^+$ prescription for the Feynman propagator and Schwinger-DeWitt proper time method. When interactions are turned on, the bare mass becomes a dressed self-energy function $\Sigma - i\Gamma_p/2$ and the pole moves away from the real axis. The imaginary part is now finite and gives the quasiparticle lifetime in the medium. Causality fixes the sign of this imaginary part, and that same causal prescription is already present in the free propagator in the limit $\Gamma_p \to 0^+$ that the interactions are turned off. It is this causal boundary condition anticipating the inclusion of interactions, rather than the interactions themselves, which breaks time reversal symmetry.

Spontaneous particle creation vs the exact annihilation of particles in the CTBD state, cf. Fig. 10 raises another interesting point about the origins of time irreversibility, entropy and the second law in QFT. That particle creation is in some sense an irreversible process in which entropy increases [56] can be made precise by means of the quantum density matrix expressed in the adiabatic particle basis [57]. Since adiabatic particle number is by construction an adiabatic invariant of the evolution, the diagonal elements of the density matrix are slowly varying in this basis. In contrast, the off-diagonal elements are rapidly varying functions both of time and of momentum at a fixed time. Then it is reasonable to average over those rapidly varying phases and construct the *reduced* density matrix which shows general (though not strictly monotonic) increase in time as particles are created [57], much as Figs. 3, 6 and 7 for the particle number itself do. This is equivalent to the approximation of neglecting the oscillatory term(s) in the current (5.21) or stress-tensor (8.14) expectation values, which as we have seen is a very good approximation over long times when the secular effects of particle creation dominate. Clearly no such interpretation is possible for the time symmetric CTBD state in which phase correlations are exactly preserved, and particle annihilations represent a decrease in the effective entropy. One would expect such processes and such finely tuned states to be statistically disfavored.

The fact that particles can achieve arbitrarily high energies for persistent constant field backgrounds producing a secular effect in both the electric field and de Sitter cases underscores the interesting interplay of UV and IR aspects. Since anomalies perform exactly this function of connecting the UV to the IR, they can play an important role [58,59], a connection we explore in detail in [21]. In de Sitter space with $\mathbb{S}^3$ spatial sections the blueshifting toward ultrahigh energies is clear in the contracting phase. In the expanding phase of de Sitter space the created particles defined in the $\mathbb{S}^3$ sections are redshifted and do not produce any growing secular effect in spatially homogeneous states. In fact, one can prove that the energy density and pressure always tend to the de Sitter invariant Bunch-Davies value for fields with positive effective masses, $m^2 + \xi R > 0$, produced in any UV allowed $O(4)$ invariant state [18]. It is this redshifting of perturbations mode by mode due to the expansion in the flat FRW coordinates (A11) that leads to the impression that (one half of) de Sitter space is stable. As shown in Sec. VII, even in this case of continual expansion there is non-trivial mode mixing, particle creation and hence a nonvanishing vacuum decay rate.

This is clearly incorrect for global (“eternal”) de Sitter space which has both a contracting and an expanding phase. With the $\mathbb{S}^3$ time slicing chosen to cover all of de Sitter space, and for early matching times as discussed, the exponentially growing energy density and pressure of the created particles will necessarily produce an enormous backreaction on the geometry if taken into account even semiclassically, and even without self-interactions. Hence one may never arrive at the expanding or inflationary phase of de Sitter space at all. This emphasizes the importance of and potential sensitivity to initial conditions of inflation [60].

In this paper we have concentrated on massive particle creation in global or “eternal” de Sitter space. This idealized situation is amenable to an exact analysis, and is the necessary foundation to be established before full backreaction, interacting fields, or more subtle issues involving light fields or the gravitational field itself are tackled, upon which a full theory of dynamical vacuum energy undoubtedly depends. Our adiabatic method and formula for the decay rate (3.32) clearly break down when $\gamma^2 \leq 0$, as the mode Eq. (3.5) then exhibits turning points on the real $u$ axis, suggesting even more pronounced quantum effects for light fields and gravitons in de Sitter space, which we have not considered here. For massive fields among the many interesting open questions is what are the consequences of the instability of de Sitter space to spontaneous and induced particle creation processes for inflation, and cosmology more generally.

De Sitter space is widely believed to be relevant to cosmology as a good description of an epoch of inflation in the early universe and of our present (and future) era of accelerated expansion. Only expanding portions of de Sitter space are assumed to arise in these descriptions. Although our analysis has focused on global de Sitter space, and in particular, on the contracting phase where the blueshifting of created particles has the most obvious, significant effects, these effects are not limited to only that phase. Let us emphasize that de Sitter space is a homogeneous spacetime, all points of which are *a priori* equivalent. There is thus no invariant meaning to the contracting vs the expanding phase; these distinctions becoming meaningful only after initial and/or boundary conditions breaking $O(4,1)$ invariance are specified. Furthermore we have shown in Sec. VII that particle creation due to Bogoliubov mode mixing occurs also in the expanding Poincaré patch, leading to the same kind of decay rate and vacuum instability as found for global de Sitter space in Sec. III.

Here again the electric field example may be helpful. One can describe a constant, uniform electric field in either a time dependent or time independent, but spatially dependent gauge. Both are equally good for describing the idealized situation without boundaries in either space or time. However, what actually happens depends sensitively on boundary or initial conditions. Just as one can consider relaxing the constancy in time of the background to study the dependence upon vacuum initial conditions, adiabatically switching it on and then off in a finite time $T$, one could also consider the arguably more physical situation of relaxing strict spatial homogeneity, allowing the electric field to be established by some charge distribution at a large but finite distance away from the local region of interest. Sensitivity of the vacuum to initial conditions will likely then be accompanied by sensitivity to the spatial boundary conditions, and the final evolution may be quite different globally.

The distance scale over which the particle creation takes place, of order $2mc^2/eE$, is also the distance scale over which significant acceleration of particles to relativistic velocities takes place in an electric field background. Thus this is the scale at which one might expect spatial inhomogeneities to develop in a random particle creation process, spatial homogeneity becoming re-established only later after the particles interact. In de Sitter space the acceleration time scale is of order $H^{-1}$, which is independent of particle mass. Perturbations on the horizon scale are sensitive to the blueshifting kinematics which is critical for the instability discussed in this paper. Hence the natural scale for inhomogeneities to develop in de Sitter space is the horizon scale, which is also dictated by causality. We address spatially inhomogeneous perturbations of the CTBD state, irrespective of particle bases in an accompanying paper [21]. The behavior of the stress tensor due to these spatially inhomogeneous perturbations depending upon the direction of $\mathbf{k}$ as well as its magnitude again suggests that the assumption of spatial homogeneity on larger than horizon scales in de Sitter space may not hold. For these reasons, the implications of the instability we have discussed in global de Sitter space for the portions of de Sitter space generally assumed to be relevant to cosmology remain to be more fully explored.

## ACKNOWLEDGMENTS

P. R. A. would like to thank Dillon Sanders for help with the early stages of this project, and we also thank Ivan Argullo and Andrei Barvinski for careful reading and helpful comments on the manuscript. This work was supported in part by the National Science Foundation under Grants No. PHY-0856050 and No. PHY-1308325. Some of the plots were generated and some of the numerical computations were performed on the WFU DEAC cluster; we thank WFU’s Provost Office and Information Systems Department for their generous support.

## APPENDIX: GEOMETRY AND COORDINATES OF DE SITTER SPACE

The de Sitter manifold is most conveniently defined as the single sheeted hyperboloid

$$\eta_{AB}X^AX^B = -(X^0)^2 + \sum_{i=1}^{3} X^iX^i + (X^4)^2 = \frac{1}{H^2} \quad \text{(A1)}$$

embedded in five-dimensional flat Minkowski spacetime,

$$ds^2 = \eta_{AB}dX^AdX^B = -(dX^0)^2 + (dX^1)^2 + (dX^2)^2 + (dX^3)^2 + (dX^4)^2. \quad \text{(A2)}$$

This manifold has the isometry group $O(4,1)$ with the maximal number of continuous symmetry generators (10) for any solution of the vacuum Einstein field equations,

$$R^a{}_b - \frac{R}{2}\delta^a{}_b + \Lambda\delta^a{}_b = 0, \quad \text{(A3)}$$

in four dimensions. It also has the discrete inversion symmetry,

$$X^A \to -X^A, \quad \text{(A4)}$$

which is not continuously connected to the identity, making the isometry group of the full de Sitter manifold $O(4,1) = \mathbb{Z}_2 \otimes SO(4,1)$. The Riemann tensor, Ricci tensor, and scalar curvature are

$$R^{ab}{}_{cd} = H^2(\delta^a{}_c\delta^b{}_d - \delta^a{}_d\delta^b{}_c) \quad \text{(A5a)}$$

$$R^a{}_b = 3H^2\delta^a{}_b \quad \text{(A5b)}$$

$$R = 12H^2, \quad \text{(A5c)}$$

with the Hubble constant $H$ related to $\Lambda$ by

$$H = \sqrt{\frac{\Lambda}{3}}. \quad \text{(A6)}$$

The globally complete hyperbolic coordinates $(u,\chi,\theta,\phi)$ of de Sitter space are defined by

$$X^0 = \frac{1}{H}\sinh u, \quad \text{(A7a)}$$

$$X^i = \frac{1}{H}\cosh u\,\sin\chi\hat{n}^i, \qquad i = 1,2,3. \quad \text{(A7b)}$$

$$X^4 = \frac{1}{H}\cosh u\,\cos\chi \quad \text{(A7c)}$$

where $\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the unit vector on $\mathbb{S}^2$. In these coordinates the de Sitter line element takes the form

$$ds^2 = \frac{1}{H^2}(-du^2 + \cosh^2 u\, d\Sigma^2), \quad \text{(A8)}$$

where

$$d\Sigma^2 \equiv d\hat{N}\cdot d\hat{N} = d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\phi^2) \quad \text{(A9)}$$

is the standard round metric on $\mathbb{S}^3$. Hence in the coordinates of (A7) which cover the entire de Sitter manifold the de Sitter line element (A8) is a hyperboloid of revolution whose constant $u$ sections are three-spheres, cf. Fig. 15. In (A9) and the following we make use of the shorthand notation,

$$\hat{N} = (\sin\chi\hat{n}, \cos\chi), \quad \text{(A10)}$$

for the unit four-vector of $\mathbb{S}^3$ in the $(X^i, X^4)$ coordinates of the flat space embedding.

In cosmology it is more common to use instead the Friedmann-Lemaître-Robertson-Walker (FLRW) line element with flat $\mathbb{R}^3$ spatial sections,

$$ds^2 = -d\tau^2 + e^{2H\tau}d\mathbf{x}^2 = -d\tau^2 + e^{2H\tau}(d\varrho^2 + \varrho^2 d\Omega^2), \quad \text{(A11)}$$

with $\varrho \equiv |\mathbf{x}|$. De Sitter space (A1)–(A2) can be brought into the flat FLRW form (A11) by setting

$$X^0 = \frac{1}{H}\sinh(H\tau) + \frac{H\varrho^2}{2}e^{H\tau} \quad \text{(A12a)}$$

$$X^i = e^{H\tau}\varrho\hat{n}^i, \qquad i = 1,2,3 \quad \text{(A12b)}$$

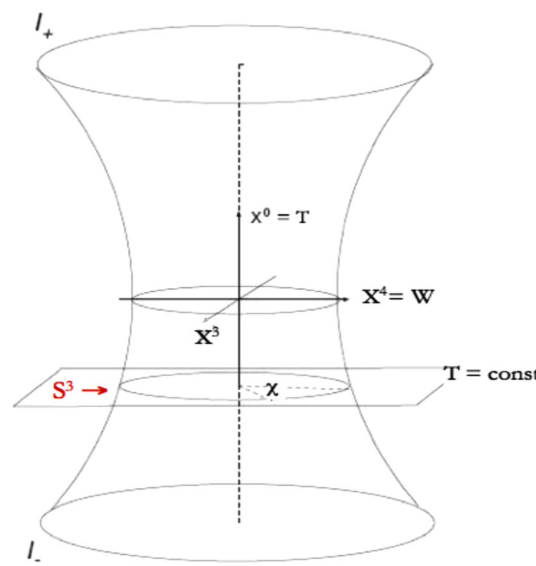


FIG. 15 (color online). The de Sitter manifold represented as a single sheeted hyperboloid of revolution about the $X^0$ axis, embedded in five-dimensional flat spacetime $(X^0, X^a)$, $a = 1, \ldots 4$, in which the $X^1$, $X^2$ coordinates are suppressed. The hypersurfaces at constant $X^0 = H^{-1}\sinh u$ are three-spheres, $\mathbb{S}^3$. The $\mathbb{S}^3$ at $X^0 = \pm\infty$ are denoted by $I_\pm$.

$$X^4 = \frac{1}{H}\cosh(H\tau) - \frac{H\varrho^2}{2}e^{H\tau}. \qquad \text{(A12c)}$$

From (A12) $T + W > 0$ in these coordinates for all $\tau \in (-\infty, \infty)$, with the hypersurfaces of constant FLRW time $\tau$ slicing the hyperboloid in Fig. 15 at a 45° angle. The null surface at $T + W = 0$ is approached in the limit $\tau \to -\infty$.

Hence the flat FLRW coordinates (A11) break the time inversion symmetry of global de Sitter space and cover only one half of the full de Sitter hyperboloid in which the spatial sections are always expanding as $\tau$ increases. The other half of the full de Sitter hyperboloid with $T + W < 0$ is obtained if $H\tau$ is replaced by $-H\tau' + i\pi$, so that $\exp(H\tau)$ is replaced by $-\exp(-H\tau')$, and the line element (A11) now takes the form

$$ds^2 = -d\tau'^2 + e^{-2H\tau'}d\mathbf{x}^2, \qquad \text{(A13)}$$

where the spatial sections are always contracting as $\tau'$ increases. The null surface $T + W = 0$ is approached in the limit $\tau' \to +\infty$.


[1] W. Pauli (unpublished); C. P. Enz and A. Thellung, Helv. Phys. Acta **33**, 839 (1960).
[2] See e.g. A. R. Liddle and D. H. Lyth, *Cosmological Inflation and Large-Scale Structure* (Cambridge University Press, Cambridge, England, 2000); or V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, Cambridge, England, 2005).
[3] A. G. Riess *et al.*, Astron. J. **116**, 1009 (1998); S. Perlmutter *et al.*, Astrophys. J. **517**, 565 (1999); J. L. Tonry *et al.*, Astrophys. J. **594**, 1 (2003).
[4] L. Parker, Phys. Rev. Lett.. **21**, 562 (1968); Phys. Rev. **183**, 1057 (1969); Phys. Rev. D **3**, 346 (1971); **3**2546(E) (1971); Phys. Rev. Lett. **28**, 705 (1972); **28**1497(E) (1972).
[5] Ya. B. Zel'dovich, Pis'ma Zh. Eksp. Teor. Fiz.. **12**, 443 (1970) [JETP Lett. **12**, 307 (1970)].
[6] L. Parker and S. A. Fulling, Phys. Rev. D **9**, 341 (1974); S. A. Fulling and L. Parker, Ann. Phys. (N.Y.) **87**, 176 (1974); S. A. Fulling, L. Parker, and B. L. Hu, Phys. Rev. D **10**, 3905 (1974).
[7] N. D. Birrell, Proc. R. Soc. B **361**, 513 (1978); T. S. Bunch, J. Phys. A **13**, 1297 (1980).
[8] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
[9] E. Mottola, Phys. Rev. D **31**, 754 (1985).
[10] E. Mottola, Phys. Rev. D **33**, 2136 (1986); *Physical Origins of Time Asymmetry*, edited by (Cambridge University Press, Cambridge, England, 1993), p. 504.
[11] I. Antoniadis, P. O. Mazur and E. Mottola, New J. Phys. **9**, 11 (2007).
[12] A. M. Polyakov, Nucl. Phys. **B797**, 199 (2008); **B834**316 (2010); arXiv:1209.4135; D. Krotov and A. M. Polyakov, Nucl. Phys. **B849**, 410 (2011).
[13] E. T. Akhmedov and P. V. Buividovich, Phys. Rev. D **78**, 104005 (2008).
[14] O. Nachtmann, Commun. Math. Phys. **6**, 1 (1967).
[15] N. A. Chernikov and E. A. Tagirov, Ann. Inst. H. Poincaré Phys. Theor. A **9**, 109 (1968); E. A. Tagirov, Ann. Phys. (N.Y.) **76**, 561 (1973).
[16] T. S. Bunch and P. C. W. Davies, Proc. R. Soc. A **360**, 117 (1978).
[17] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2738 (1977); A. S. Lapedes, J. Math. Phys. (N.Y.) **19**, 2289 (1978).
[18] P. R. Anderson, W. Eaker, S. Habib, C. Molina-París, and E. Mottola, Phys. Rev. D **62**, 124019 (2000).
[19] D. Marolf and I. A. Morrison, Phys. Rev. D **82**, 105032 (2010); **84**, 044040 (2011); Gen. Relativ. Gravit. **43**, 3497 (2011).
[20] S. Hollands, Commun. Math. Phys. **319**, 1 (2013).
[21] P. R. Anderson and E. Mottola, Phys. Rev. D **89**, 104039 (2014).
[22] J. Schwinger, Phys. Rev. **82**, 664 (1951).
[23] B. S. DeWitt *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965); Phys. Rep. **19**, 295 (1975).
[24] H. Rumpf, Phys. Lett. B **61**, 272 (1976); Nuovo Cimento Soc. Ital. Fis. **35**, 321 (1976); H. Rumpf and H. K. Urbantke, Ann. Phys. (N.Y.) **114**, 332 (1978).
[25] L. H. Ford, Phys. Rev. D **31**, 710 (1985).
[26] I. Antoniadis and E. Mottola, reportCERN-TH-4605/86, 1986; J. Math. Phys. (N.Y.) **32**, 1037 (1991).
[27] N. C. Tsamis and R. P. Woodard, Nucl. Phys. **B474**, 235 (1996); Ann. Phys. (N.Y.) **253**, 1 (1997); **267**145 (1998).
[28] N. B. Narozhnyi, Zh. Eksp. Teor. Fiz.. **54**, 676 (1968) [Sov. Phys. JETP **27**, 360 (1968)].
[29] A. I. Nikishov, Zh. Eksp. Teor. Fiz.. **57**, 1210 (1970) [Sov. Phys. JETP **30**, 660 (1970)].
[30] N. B. Narozhnyi and A. I. Nikishov, Yad. Fiz. **11**, 1072 (1970) [Sov. J. Nucl. Phys. **11**, 596 (1970)].
[31] E. Brezin and C. Itzykson, Phys. Rev. D **2**, 1191 (1970).
[32] M. S. Marinov and V. S. Popov, Fortschr. Phys. **25**, 373 (1977).
[33] E. S. Fradkin, D. M. Gitman, and Sh. M. Shvartsman, *Quantum Electrodynamics with Unstable Vacuum* (Springer-Verlag, Berlin, 1991).
[34] F. Cooper, E. Mottola, B. Rogers, and P. Anderson, in *Proceedings of the Santa Fe Workshop on Intermittency in High Energy Collisions*, edited by F. Cooper *et al.* (World Scientific, Singapore, 1991), p. 399; Y. Kluger, J. M. Eisenberg, B. Svetitsky, F. Cooper, and E. Mottola, Phys. Rev. Lett.. **67**, 2427 (1991); Phys. Rev. D **45**, 4659

(1992); F. Cooper, J. M. Eisenberg, Y. Kluger, E. Mottola, and B. Svetitsky, Phys. Rev. D **48**, 190 (1993).
[35] S. O. Gavrilov and D. M. Gitman, Phys. Rev. D **53**, 7162 (1996).
[36] Y. Kluger, E. Mottola, and J. M. Eisenberg, Phys. Rev. D **58**, 125015 (1998).
[37] S. Habib, C. Molina-París and E. Mottola, Phys. Rev. D **61**, 024010 (1999).
[38] R. P. Feynman, Phys. Rev., **76**, 749 (1949).
[39] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948); R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[40] *Higher Transcendental Functions*, Vols. I and II, Bateman Manuscript Project, edited by A. Erdélyi (McGraw-Hill, New York, 1953).
[41] M. Gutzwiller, Helv. Phys. Acta **29**, 213 (1956).
[42] H. Rumpf, Phys. Rev. D **24**, 275 (1981).
[43] J. S. Dowker and R. Crichley, Phys. Rev. D **13**, 224 (1976).
[44] B. Allen, Phys. Rev. D **32**, 3136 (1985).
[45] A. R. Bell and J. G. Kirk, Phys. Rev. Lett.. **101**, 200403 (2008); D. B. Blaschke, A. V. Prozorkevich, G. Ropke, C. D. Roberts, S. M. Schmidt, D. S. Shkirmanov, and S. A. Smolyansky, Eur. Phys. J. D **55**, 341 (2009); G. V. Dunne, H. Gies, and R. Schützhold, Phys. Rev. D **80**, 111301 (2009); G. V. Dunne, Int. J. Mod. Phys. A **25**, 2373 (2010).
[46] V. L. Pokrovskii and I. M. Khalatnikov, Zh. Eksp. Teor. Fiz. **40**, 1713 (1961) [Sov. Phys. JETP **13**, 1207 (1961)].
[47] J. Audretsch, J. Phys. A **12**, 1189 (1979).
[48] P. R. Anderson, E. Mottola, and D. H. Sanders (in preparation).
[49] J. Garriga, S. Kanno, M. Sasaki, J. Soda, and A. Vilenkin, J. Cosmol. Astropart. Phys. 12 (2012) 006.
[50] P. R. Anderson and L. Parker, Phys. Rev. D **36**, 2963 (1987).
[51] P. R. Anderson and W. Eaker, Phys. Rev. D **61**, 024003 (1999).
[52] E. T. Akhmedov, Mod. Phys. Lett. A **25**, 2815 (2010).
[53] P. R. Anderson, C. Molina-París, and E. Mottola, Phys. Rev. D **72**, 043515 (2005).
[54] P. O. Mazur and E. Mottola, Nucl. Phys. **B278**, 694 (1986).
[55] For a discussion of the two-dimensional case, see Y.-Z. Chu and T. Vachaspati, Phys. Rev. D **81**, 085020 (2010).
[56] B. L. Hu and D. Pavon, Phys. Lett. B **180**, 329 (1986); B. L. Hu and H. E. Kandrup, Phys. Rev. D **35**, 1776 (1987).
[57] S. Habib, Y. Kluger, E. Mottola, and J. P. Paz, Phys. Rev. Lett.. **76**, 4660 (1996); F. Cooper, S. Habib, Y. Kluger, and E. Mottola, Phys. Rev. D **55**, 6471 (1997).
[58] P. R. Anderson, C. Molina-París, and E. Mottola, Phys. Rev. D **80**, 084005 (2009).
[59] E. Mottola, Acta Phys. Pol. B **41**, 2031 (2010).
[60] A. Vilenkin, Phys. Rev. D **46**, 2355 (1992).