PHYSICAL REVIEW D **89,** 104039 (2014)

Quantum vacuum instability of "eternal" de Sitter space

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 6 October 2013; published 23 May 2014)

The Euclidean or Bunch-Davies $O(4,1)$ invariant state for quantum fields in global de Sitter space is shown to be unstable to small perturbations, even for a massive free field with no self-interactions. There are perturbations of this state with energy density that is arbitrarily small at early times, is exponentially blueshifted in the contracting phase of "eternal" de Sitter space, and becomes large enough to disturb the classical geometry through the semi-classical Einstein equations at later times. In the closely analogous case of a constant, uniform electric field, a time symmetric state equivalent to the de Sitter invariant one is constructed, which is also not a stable vacuum state under perturbations. The role of a quantum anomaly in the growth of perturbations and symmetry breaking is emphasized in both cases. In de Sitter space, the same results are obtained either directly from the renormalized stress tensor of a massive scalar field, or for massless conformal fields of any spin, more directly from the effective action and stress tensor associated with the conformal trace anomaly. The anomaly stress tensor shows that states invariant under the $O(4)$ subgroup of the de Sitter group are also unstable to perturbations of lower spatial symmetry, implying that both the $O(4,1)$ isometry group and its $O(4)$ subgroup are broken by quantum fluctuations. Potential consequences of this result for cosmology and the problem of vacuum energy are discussed.

DOI: 10.1103/PhysRevD.89.104039 PACS numbers: 04.62.+v, 95.36.+x, 98.80.Jk, 98.80.Qc

I. VACUUM STATES IN DE SITTER SPACE

The existence of a ground state as the state of lowest energy is fundamental to all quantum mechanical systems. For quantum field theory (QFT) in flat Minkowski spacetime, the vacuum state is defined as the eigenstate of the Hamiltonian operator of the system with the lowest eigenvalue. The existence of a Hamiltonian generator of time translational symmetry, with a non-negative eigenvalue spectrum, bounded from below is crucial to the existence and determination of the vacuum ground state.

This definition of the vacuum in flat spacetime makes use of an essential property of the Poincaré group, namely that positive and negative (particle and antiparticle) halves of the Hamiltonian spectrum do not mix, remaining distinct under any of the continuous generators of the group. Hence the vacuum state in flat space QFT is the same for all inertial frames related to each other by translations, rotations and Lorentz boosts, and the vacuum enjoys complete invariance under Poincaré symmetry.

As is well known, none of these properties hold in a general curved spacetime, in time-dependent background fields, nor even in flat spacetime under general coordinate transformations which are not Poincaré symmetries. In these circumstances the definitions of "vacuum" and "particles" become much more subtle. Related to this, whereas the infinite zero point energy associated with the QFT vacuum may be disregarded as unobservable in flat space QFT, the energy of the quantum vacuum in curved spacetime cannot be neglected when the coupling to gravity is taken into account.

These issues come to the fore in the important special case of de Sitter space, the classical spacetime with a positive cosmological constant $\Lambda > 0$, which itself may be regarded as a vacuum energy density uniformly curving space. The geodesically complete full de Sitter manifold may be represented as a single sheeted hyperboloid of revolution embedded in five dimensional flat Minkowski spacetime, cf. Fig. 1 [1]. It has the isometry group $O(4,1)$ with 10 continuous symmetry generators, the same number as the Poincaré group of Minkowski space, and the maximal number possible for any solution of the vacuum Einstein field equations in 4 spacetime dimensions. This maximal symmetry is evident from the constant and uniform Riemann curvature tensor and Ricci scalar of de Sitter space, which are respectively,

$$R^{ab}{}_{cd} = H^2(\delta^a{}_c\delta^b{}_d - \delta^a{}_d\delta^b{}_c), \tag{1.1a}$$

$$R^a{}_b = 3H^2\delta^a{}_b = \Lambda\delta^a{}_b, \tag{1.1b}$$

$$R = 12H^2 = 4\Lambda. \tag{1.1c}$$

A natural attempt to generalize the QFT vacuum of flat space to de Sitter space makes use of this geometrical

*anderson@wfu.edu
†emil@lanl.gov

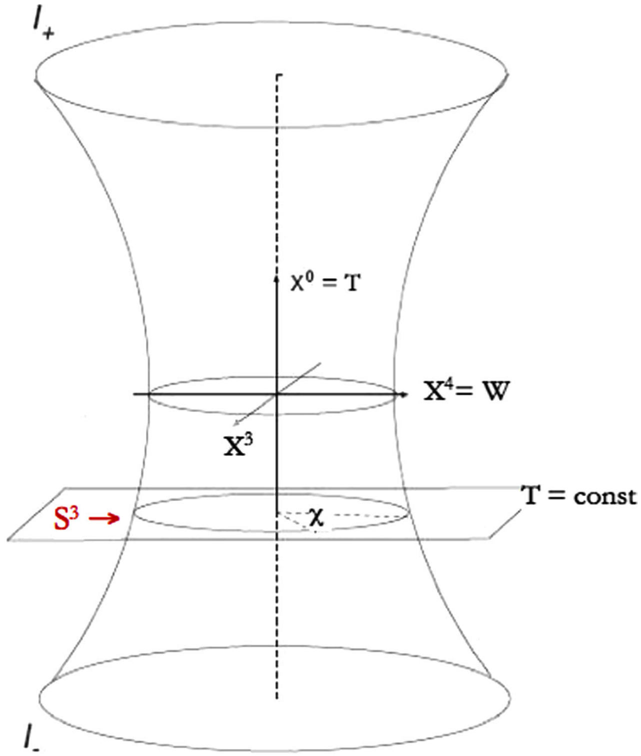

FIG. 1 (color online). The de Sitter manifold represented as a single sheeted hyperboloid of revolution about the $T = X^0$ axis, embedded in five dimensional flat spacetime (X^0, X^a), $a = 1, \ldots 4$, in which the X^1, X^2 coordinates are suppressed. The hypersurfaces at constant $T = X^0 = H^{-1}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

symmetry of de Sitter space to define the de Sitter invariant "vacuum" $|v\rangle$ as the state possessing the same maximal $O(4,1)$ symmetry in the Hilbert space of states. Introduced by Chernikov and Tagirov (CT) [2], this state is commonly known also as the Bunch-Davies (BD) state [3,4], or the Euclidean vacuum, because its Green's functions are those obtained by analytic continuation from the Euclidean $\mathbb{S}^4$, at least for massive fields where no obvious infrared issues arise [5].

It is important to recognize that unlike in flat space, the construction of the CTBD state is not based on diagonalization of any Hamiltonian nor any minimization of energy. In fact no suitable Hamiltonian operator with a spectrum bounded from below exists at all in de Sitter space, even for free QFT. In the globally complete coordinates of the de Sitter hyperboloid,

$$ds^2 = H^{-2}(-du^2 + \cosh^2 u \, d\Sigma^2), \tag{1.2}$$

with

$$d\Sigma^2 \equiv d\hat{N} \cdot d\hat{N} = d\chi^2 + \sin^2\chi d\hat{n} \cdot d\hat{n} = d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\phi^2), \tag{1.3}$$

the standard round metric on $\mathbb{S}^3$, the de Sitter metric is dependent on the time u. Thus translation in u is not a symmetry of de Sitter space and the generator of u time translations is not conserved. As a consequence, the vacuum defined by Hamiltonian diagonalization at one instant of u time will contain particles at any other u time. This is equally true in the flat spatial slicing of de Sitter space

$$ds^2 = -d\tau^2 + e^{2H\tau} d\vec{x} \cdot d\vec{x}, \tag{1.4}$$

used most frequently in cosmology, which is similarly dependent on the time τ.

The nonexistence of a conserved Hamiltonian generator bounded from below in de Sitter space is a consequence of the de Sitter symmetry group $O(4,1)$ itself. Unlike the Poincaré group, any de Sitter symmetry generator chosen for the role of the Hamiltonian has a spectrum of both positive and negative eigenvalues which are mixed by the action of other generators of the group [6]. One of the four noncompact Lorentz boost generators of the $O(4,1)$ symmetry group may be selected (arbitrarily) as the Hamiltonian of the system, generating time translations $t \to t + \Delta t$ in the static coordinates of de Sitter space, where the line element takes the form

$$ds^2 = -(1 - H^2 r^2) dt^2 + \frac{dr^2}{1 - H^2 r^2} + r^2(d\theta^2 + \sin^2\theta \, d\phi^2). \tag{1.5}$$

In these coordinates the geometry is independent of the time t. However the event horizon at $r = H^{-1}$ relative to the origin $r = 0$ is now manifest, and the static coordinates cover only one quarter of the full de Sitter manifold. The Killing symmetry $\partial/\partial t$ is not globally timelike, and changes its orientation from one quadrant to another, as may be seen from the Carter-Penrose conformal diagram of de Sitter space: Fig. 2. A direct consequence of this is that the corresponding Hamiltonian symmetry generator across any complete Cauchy surface is not positive definite, but rather unbounded from below, as Lorentz boosts are. Hence its eigenstates or expectation values cannot be used to select a global minimum energy vacuum state. The choice of $\partial/\partial t$ is also arbitrary and the separation into positive and negative energies with respect to $\partial/\partial t$ is noninvariant under de Sitter group transformations. The particle concept is likewise affected, as the CTBD de Sitter invariant vacuum state $|v\rangle$ is actually a state with a thermal distribution of particles with respect to the Killing Hamiltonian generator ∂_t of (1.5) with the Hawking de Sitter temperature [7],

$$T_H = \frac{\hbar H}{2\pi k_B}, \tag{1.6}$$

and in that sense is not a vacuum state at all.

The horizon and causal structure of de Sitter space raises the question of how a vacuum state can be prepared

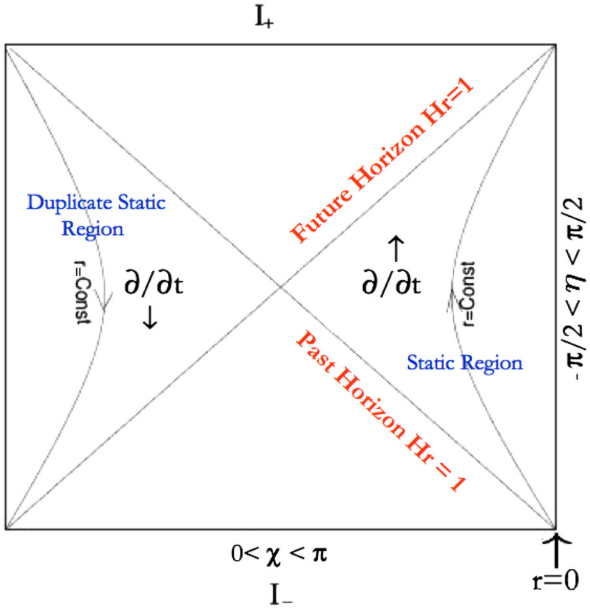

FIG. 2 (color online). The Carter-Penrose conformal diagram for de Sitter space, in which light rays emanating from any point are at 45°, and the angular coordinates θ, ϕ are suppressed. The quarter of the diagram labeled as the static region is covered by the static coordinates of (1.5). The orbits of the static time Killing field $\partial/\partial t$, $r = 0$ and curves of constant $r > 0$ are shown. In contrast, the surfaces of the constant u time coordinate in (1.2) are horizontal straight lines across the diagram with $\chi \in [0, \pi]$. The label on the right is the conformal time coordinate $\eta \equiv \sin^{-1}(\tanh u) \in (-\frac{\pi}{2}, \frac{\pi}{2})$. The coordinate u ranges from $-\infty$ to $+\infty$ at past (I_-) and future (I_+) infinity, respectively.

operationally, even in principle. Inspection of the conformal diagram in Fig. 2 shows that points on a Cauchy surface at $u_0 < 0$ with widely different $\hat{N}$ (for example at $\chi = 0$ and its antipodal point $\chi = \pi$) could never have been linked by any causal signal in the past. As the initial time u_0 is taken earlier and earlier, this causal disconnection affects more and more of the initial $u = u_0$ Cauchy surface. Since past infinity I_- is *spacelike*, as $u_0 \to -\infty$ in this limit no two different points on $\mathbb{S}^3$ could have been in any causal contact whatsoever. Thus any global initial data on $\mathbb{S}^3$, including that necessary to construct the CTBD state $|v\rangle$ cannot have been provided at any initial time $u_0 < 0$ by any causal process within de Sitter space itself. Instead initial data has simply to be posited over the full spacelike $\mathbb{S}^3$, at points outside of the causal horizon of any local agent who might have prepared it at early times. This is equivalent to the existence of a particle horizon [1], and is completely unlike that of flat Minkowski spacetime, where Cauchy data on a fixed time slice $t = \text{const}$ may be prepared in principle by transmitting signals causally from a single point early enough in the past. The existence of horizons and the absence of a global time coordinate connected with any symmetry reveal the essential difficulties with defining a global vacuum state for QFT in eternal de Sitter space.

They also imply that the mathematical requirement of global de Sitter invariance cannot be realized by any local physics within de Sitter space itself, requiring instead an acausal fine tuning of initial data at spacelike past infinity I_-, with a view to the entire future manifold, which is presumed to be known in advance in order to specify a globally $O(4,1)$ invariant state. Although maximal $O(4,1)$ symmetry may seem natural mathematically, or by analytic continuation from the Euclidean $\mathbb{S}^4$ where no causal relations apply, it is quite unnatural with respect to the physical principles of locality and causality in real time, as well as lacking any de Sitter invariant Hamiltonian minimization principle.

For these reasons it is important to study the sensitivity of physical quantities in de Sitter space to fluctuations and/or perturbations of states away from precisely the 'right' one for global $O(4,1)$ invariance, rather than simply assuming this symmetry. The calculation of the imaginary part of the effective action of a simple scalar QFT in de Sitter space due to particle creation [8,9] already shows that de Sitter space is unstable to spontaneous creation of particle pairs from the vacuum, just as is an 'eternal' uniform electric field $\mathbf{E} = E\hat{\mathbf{z}}$ permeating all of space [10]. This electric field analogy and the close relation between fluctuations and dissipation in any causal theory suggests that the 'shorting of the vacuum' should result in the classical energy of de Sitter space converting itself into standard matter and radiation, thus providing a route to a dynamical solution to the cosmological 'constant' problem [11,12].

As in electrodynamics, interactions in de Sitter space are certainly relevant to understanding of the detailed evolution and final state, particularly since spontaneous pair creation should be accompanied by induced emission processes which can create an avalanche of particles that will inevitably interact and thermalize, leading to the final dissipation of vacuum energy into matter and radiation. A fuller understanding of these nonequilibrium processes may well lead to a satisfactory resolution of the cosmological "constant" problem, and be relevant to observational cosmology through the residual dark energy in the present epoch [11,12]. However, this dynamics has not been fully solved in four dimensions even in flat space electrodynamics. Moreover a number of questions persist about QFT in de Sitter space, even in the noninteracting case, and these should be settled definitively first, because they depend through the energy-momentum-stress tensor T_{ab} only upon the universal coupling to the gravitational field, independently of any matter self-interactions.

In the preceding paper [9], we focused on the instability of global de Sitter space to particle creation, delineating in particular the close analogy to the Schwinger calculation of the decay rate per unit volume of a constant, uniform electric field permeating all of space [10]. In this paper we study the behavior of the renormalized energy-momentum tensor $\langle T_{ab}\rangle$ of QFT under perturbations of the $|v\rangle$ state to nearby states of lower symmetry, without regard to particle definitions. In the expanding part of de Sitter space $u > 0$ of (1.2) or in the Poincaré coordinates of flat spatial sections (1.4), it has been shown that in a fixed de Sitter background, the expectation value $\langle T_{ab} \rangle$ for a scalar field with effective mass $M^2 = m^2 + \xi R > 0$ approaches the $O(4,1)$ de Sitter invariant value at late times, for all *spatially homogeneous* UV allowed perturbations [13]. Calculations including the backreaction of the perturbations of the stress tensor on the de Sitter metric have also been done, with similar results [14,15]. Physically this result may seem intuitively obvious, since all deviations from the expectation value in the de Sitter invariant state $|v\rangle$ are redshifted in the de Sitter expansion and vanish in the $u \to \infty$ limit. Since global or 'eternal' de Sitter space is time reversal invariant, the attractor behavior in the expanding phase implies just the opposite behavior under time reversal in the contracting phase. That is, very small changes in the initial state in the very distant past $u_0 \to -\infty$ of eternal de Sitter space with initially very small $\langle T_{ab} \rangle$ must necessarily produce larger and larger effects in $\langle T_{ab} \rangle$ as the contraction proceeds towards $u = 0$. This is just the case where the aforementioned issues with causality at spacelike I_- arise, and this sensitivity to initial conditions at I_- is the source of the instability.

By studying the general behavior of the renormalized $\langle T_{ab} \rangle$ in states with lower symmetry, we show in this paper that the CTBD de Sitter invariant state $|v\rangle$ is unstable, in the sense that there is a large class of initial state perturbations which have exponentially small energy density in the infinite past $u_0 \to -\infty$ but which grow large enough through exponential blueshifting proportional to a^{-4}, where $a = H^{-1} \cosh u$ is the scale factor in (1.2), to exceed the classical background energy $\Lambda/8\pi G$ and hence significantly disturb the de Sitter geometry at $u = 0$. In fact, there are such states with $\langle T_{ab} \rangle$ larger than it for any fixed finite value at $u = 0$.

This extreme sensitivity to initial conditions as $u_0 \to -\infty$ implies that $O(4,1)$ de Sitter invariance is broken, and the spacetime will generally depart from de Sitter space when the backreaction of $\langle T_{ab} \rangle$ of any matter or radiation on the geometry is taken into account, through the semiclassical Einstein equations,

$$R^a{}_b - \frac{R}{2}\delta^a{}_b + \Lambda \delta^a{}_b = 8\pi G \langle T^a{}_b \rangle_R, \tag{1.7}$$

and quite apart from any matter self-interactions or higher loop effects. Although in a fixed de Sitter background the energy density of spatially homogeneous perturbations will begin to decrease again for $u > 0$, perturbations of the CTBD state and their backreaction through (1.7) will have already drastically altered the geometry in the contracting phase and broken the de Sitter symmetry by $u = 0$, rendering further evolution ignoring backreaction moot. This large backreaction of the energy-momentum tensor for perturbations of the CTBD state is independent of any definition of particles.

Although for definiteness we study this growth of $\langle T_{ab} \rangle$ explicitly in a scalar field theory, the result is clearly much more general. A very useful tool for characterizing the behavior of the stress tensor in any coordinates is the one-loop effective action of the trace anomaly and the stress tensor derived from it [16–18]. The nonlocal form of this effective action, cf. (5.1) already indicates infrared de Sitter breaking effects, and sensitivity to initial and/or boundary conditions for conformal QFTs of any spin. The corresponding stress tensor may be found in closed form in de Sitter space in any coordinates by solving a classical, linear Eq. (5.6) for a scalar condensate effective field, whose solutions necessarily break de Sitter invariance, and allow wide classes of initial state perturbations for fields of any spin to be surveyed at once. Because de Sitter space is conformally flat, this anomaly stress tensor is a complete description of the full QFT stress tensor for conformal fields linearized around the CTBD state $|v\rangle$ at all length scales much larger than the Planck length L_{Pl}, where semiclassical methods should apply [15].

The a^{-4} blueshifting of the energy density of even massive fields to eventually ultrarelativistic behavior shows that the conformal anomaly stress tensor is relevant for long time evolution even if the underlying QFT is not conformally invariant. When the scalar perturbations are spatially *in*homogeneous new effects may also become apparent. In Ref. [15] we studied spatially inhomgeneous scalar perturbations in linear response of conformal QFTs around de Sitter space and found a class of gauge invariant perturbations, which do *not* redshift away but instead give diverging energy-momentum components at $r = H^{-1}$ in static coordinates (1.5). These may be interpreted as fluctuations in the Hawking de Sitter temperature (1.6) at the de Sitter horizon with respect to some arbitrary but fixed choice of origin, and clearly respect only rotational $O(3)$ invariance around $r = 0$ and static time t translational invariance. This result suggests that fluctuations on the horizon scale H^{-1} may produce significant backreaction in de Sitter space, and that the $O(4,1)$ symmetry is unstable to such spatially inhomogeneous scalar fluctuations in the Hawking de Sitter temperature [19]. Tensor perturbations have been studied recently in [20].

Using the anomaly form of T_{ab} we shall show that there is even greater sensitivity to spatially inhomogeneous non-$O(4)$ invariant initial data in the distant past $u_0 \to -\infty$ of global de Sitter space, so that $O(4)$ invariance is broken as well as the full $O(4,1)$ de Sitter invariance. This strongly suggests that spatial inhomogeneities are more important in QFT in de Sitter space than previously suspected, supporting the results of [15]. Such spatially inhomogeneous perturbations clearly are relevant even in the expanding Poincaré patch (1.4). The interesting questions of the behavior of the stress tensor in states of lower symmetry, such as the $O(3)$ symmetry evident in static coordinates (1.5), and consequences for spatially inhomogeneous cosmologies will be taken up in future publications. An accompanying and closely related paper gives a fuller treatment of the instability of global de Sitter space to particle creation [9].

The paper is organized as follows. In the next section we construct the time symmetric invariant state analogous to the CTBD state in de Sitter space, in the case of a uniform, constant electric field background $\mathbf{E} = E\hat{\mathbf{z}}$, and show that it also is unstable to perturbations for which the mean current $\langle j_z \rangle$ grows with time. This growth of the current and breaking of background symmetries can be understood by consideration of a quantum anomaly, in this case the chiral anomaly of massless fields in two spacetime dimensions. The reader interested primarily in de Sitter space proper may skip this section upon first reading and proceed directly to Sec. III where we begin discussion of the CTBD state and general states of $O(4)$ symmetry in de Sitter space. In Sec. IV we construct the renormalized expectation value of the stress tensor of a massive scalar field with conformal coupling $\xi = \frac{1}{6}$ in general $O(4)$ invariant states in the global hyperboloid coordinates (1.2) of de Sitter space, and explicitly exhibit the class of states with large backreaction at $u = 0$. In Sec. V we consider the effective action and stress tensor associated with the trace anomaly of conformal fields in de Sitter space and show how the strong infrared effects, sensitivity to initial conditions, and breaking of de Sitter symmetry is inherent in the conformal anomaly for QFTs of any spin. In Sec. VI we extend the analysis of the anomaly stress tensor to states of lower than $O(4)$ symmetry, showing that these spatially inhomogeneous perturbations grow even more rapidly to larger values at $u = 0$ than $O(4)$ symmetric states. Section VII contains our conclusions and a discussion of their possible consequences for cosmology and the problem of cosmological vacuum energy.

II. CONSTANT UNIFORM ELECTRIC FIELD: INVARIANT STATE AND INSTABILITY

A. Time Symmetric Invariant State

The example of a charged quantum field in the background of a constant uniform electric field has many similarities with the de Sitter case. Although this problem has been considered by many authors [10,21–27], the existence of a time symmetric state analogous to the CTBD state in de Sitter space does not appear to have received previous attention, and is particularly relevant to our study of vacuum states in de Sitter space, so we consider this case first in some detail.

Analogous to choosing global time-dependent coordinates (1.2) in de Sitter space, one may choose the time-dependent gauge,

$$A_z = -Et, \qquad A_t = A_x = A_y = 0, \tag{2.1}$$

in which to describe a fixed constant and uniform electric field $\mathbf{E} = E\hat{\mathbf{z}}$ in the z direction. 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 $\mathbf{\Phi}$ is

$$[-(\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu) + m^2]\mathbf{\Phi} = 0 \tag{2.2}$$

in the classical electromagnetic potential (2.1).

The solutions of (2.2) may be decomposed into Fourier modes $\Phi \sim e^{i\mathbf{k}\cdot\mathbf{x}} f_{\mathbf{k}}(t)$ with

$$\left[\frac{d^2}{dt^2} + \omega_{\mathbf{k}}^2(t)\right] f_{\mathbf{k}}(t) = 0, \tag{2.3}$$

where the frequency function $\omega_{\mathbf{k}}(t)$ is defined by

$$\omega_{\mathbf{k}}(t) \equiv [(k_z + eEt)^2 + k_\perp^2 + m^2]^{\frac{1}{2}} = \sqrt{2eE}\sqrt{\frac{u^2}{4} + \lambda}. \tag{2.4}$$

We have defined the dimensionless variables,

$$u \equiv \sqrt{\frac{2}{eE}}(k_z + eEt), \qquad \lambda \equiv \frac{k_\perp^2 + m^2}{2eE} > 0, \tag{2.5}$$

and chosen the sign of eE to be positive without loss of generality. With $f_{\mathbf{k}}(t) \to f_\lambda(u)$, the dimensionless mode Eq. (2.3) becomes

$$\left[\frac{d^2}{du^2} + \frac{u^2}{4} + \lambda\right] f_\lambda(u) = 0, \tag{2.6}$$

the solutions of which may be expressed in terms of confluent hypergeometric functions ${}_1F_1(a;c;z)$ or parabolic cylinder functions D_ν [28].

Since (2.6) is real and symmetric under $u \to -u$, it is clear that its real solutions can be classified into those which are either even or odd under this discrete reflection symmetry. Let us define two fundamental real solutions of (2.6), $f_\lambda^{(i)}(u)$, $i = 0, 1$ by the conditions

$$[f_\lambda^{(i)}(u)]^* = f_\lambda^{(i)}(u), \qquad i = 0, 1 \qquad \text{(real)} \tag{2.7a}$$

$$f_\lambda^{(0)}(u) = f_\lambda^{(0)}(-u) \qquad \text{(even)} \tag{2.7b}$$

$$f_\lambda^{(1)}(u) = -f_\lambda^{(1)}(-u) \qquad \text{(odd)}, \tag{2.7c}$$

which are even or odd, respectively, and which satisfy the initial data,

$$f_\lambda^{(0)}(0) = 1, \qquad f_\lambda^{(0)\prime}(0) = 0, \tag{2.8a}$$

$$f_\lambda^{(1)}(0) = 0, \qquad f_\lambda^{(1)\prime}(0) = 1, \tag{2.8b}$$

at $u = 0$, where the primes denote differentiation with respect to u. These fundamental real solutions of (2.6) are most concisely expressed in terms of the confluent hypergeometric (Kummer) function,

$$\Phi(a,c;z) \equiv {}_1F_1(a;c;z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} \frac{z^n}{n!},$$
$$(a)_n \equiv \frac{\Gamma(a+n)}{\Gamma(a)}, \tag{2.9}$$

which has the integral representation [28]

$$\Phi(a,c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 dx\, e^{xz} x^{a-1} (1-x)^{c-a-1},$$
$$\mathrm{Re}\ c > \mathrm{Re}\ a > 0 \tag{2.10}$$

with the result that

$$f_\lambda^{(0)}(u) = e^{-\frac{iu^2}{4}} \Phi\left(\frac{1}{4} + \frac{i\lambda}{2}, \frac{1}{2}; \frac{iu^2}{2}\right)$$
$$= e^{\frac{iu^2}{4}} \Phi\left(\frac{1}{4} - \frac{i\lambda}{2}, \frac{1}{2}; -\frac{iu^2}{2}\right), \tag{2.11a}$$

$$f_\lambda^{(1)}(u) = u e^{-\frac{iu^2}{4}} \Phi\left(\frac{3}{4} + \frac{i\lambda}{2}, \frac{3}{2}; \frac{iu^2}{2}\right)$$
$$= u e^{\frac{iu^2}{4}} \Phi\left(\frac{3}{4} - \frac{i\lambda}{2}, \frac{3}{2}; -\frac{iu^2}{2}\right). \tag{2.11b}$$

These functions are clearly even and odd respectively, and are real by the Kummer transformation which yields the second forms in (2.11), and satisfy the initial data (2.8).

It is also possible to express these fundamental real solutions $f_\lambda^{(i)}$ as linear combinations of parabolic cylinder functions D_ν in the forms [28]

$$f_\lambda^{(0)}(u) = 2^{\frac{i\lambda}{2} - \frac{3}{4}} \frac{\Gamma\left(\frac{3}{4} + \frac{i\lambda}{2}\right)}{\sqrt{\pi}} \left[D_{-\frac{1}{2}-i\lambda}(e^{\frac{i\pi}{4}} u) + D_{-\frac{1}{2}-i\lambda}(-e^{\frac{i\pi}{4}} u)\right], \tag{2.12a}$$

$$f_\lambda^{(1)}(u) = 2^{\frac{i\lambda}{2} - \frac{5}{4}} e^{-\frac{i\pi}{4}} \frac{\Gamma\left(\frac{1}{4} + \frac{i\lambda}{2}\right)}{\sqrt{\pi}} \left[-D_{-\frac{1}{2}-i\lambda}\left(e^{\frac{i\pi}{4}} u\right) + D_{-\frac{1}{2}-i\lambda}\left(-e^{\frac{i\pi}{4}} u\right)\right], \tag{2.12b}$$

the representations of which are useful for identifying their relationship to the *in* and *out* positive frequency scattering solutions defined as $u \to \mp\infty$, respectively, in [9,21–23].

From the fundamental real solutions (2.8)–(2.11) one can construct the complex mode functions

$$\upsilon_\lambda(u) \equiv (8eE\lambda)^{-\frac{1}{4}} [f_\lambda^{(0)}(u) - i\lambda^{\frac{1}{2}} f_\lambda^{(1)}(u)]$$
$$= 2^{-\frac{1}{2}} (k_\perp^2 + m^2)^{-\frac{1}{4}} e^{-\frac{iu^2}{4}} \left[\Phi\left(\frac{1}{4} + \frac{i\lambda}{2}, \frac{1}{2}; \frac{iu^2}{2}\right) - i\lambda^{\frac{1}{2}} u \Phi\left(\frac{3}{4} + \frac{i\lambda}{2}, \frac{3}{2}; \frac{iu^2}{2}\right)\right], \tag{2.13}$$

which are normalized according to the Wronskian condition

$$i\left(\upsilon_\lambda^* \frac{d}{dt} \upsilon_\lambda - \upsilon_\lambda \frac{d}{dt} \upsilon_\lambda^*\right) = 1, \tag{2.14}$$

and which satisfy the time reversal conjugation property

$$\upsilon_\lambda^*(u) = \upsilon_\lambda(-u). \tag{2.15}$$

These υ_λ mode functions satisfy the initial data,

$$\upsilon_\lambda(0) = 2^{-\frac{1}{2}} (k_\perp^2 + m^2)^{-\frac{1}{4}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\bigg|_{u=0},$$
$$\frac{d\upsilon_\lambda}{dt}\bigg|_{u=0} = -\frac{i}{\sqrt{2}} (k_\perp^2 + m^2)^{\frac{1}{4}} = -i\omega_{\mathbf{k}} \upsilon_\lambda(u), \tag{2.16}$$

which coincides with the definition of the lowest order adiabatic frequency mode functions at the symmetric point $u = 0$. The solution of (2.6) satisfying conditions (2.14)–(2.16) is unique. Because of relations (2.12) and the simple asymptotic forms of the D_ν functions, the symmetric mode function υ_λ is a coherent superposition of positive and negative frequency (particle and antiparticle) solutions as $u \to \pm\infty$, just as the CTBD mode function is in de Sitter space [9].

The existence of such a time reversal invariant solution to (2.6) is related to the existence of a maximally symmetric state constructed along the lines of the maximally $O(4,1)$ invariant CTBD state in the de Sitter background. If the charged quantum field $\mathbf{\Phi}$ is expanded in terms of these symmetric basis functions in a finite volume V,

$$\mathbf{\Phi}(t, \mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} [a_{\mathbf{k}}^{\upsilon} \upsilon_\lambda(u) e^{i\mathbf{k}\cdot\mathbf{x}} + b_{\mathbf{k}}^{\upsilon\dagger} \upsilon_\lambda^*(u) e^{-i\mathbf{k}\cdot\mathbf{x}}], \tag{2.17}$$

with u and λ defined by (2.5), then a symmetric state $|\upsilon\rangle$ in a background constant uniform electric field may be defined by

$$a_{\mathbf{k}}^{\upsilon} |\upsilon\rangle = b_{\mathbf{k}}^{\upsilon} |\upsilon\rangle = 0. \tag{2.18}$$

The symmetry in this case is isomorphic to the full Poincaré symmetry group of zero electric field in flat Minkowski space. This is due to the remarkable fact that a canonical transformation exists that transforms the algebra of position and momentum operators, x^μ and $p_\nu = -i\partial/\partial x^\nu$, in a constant, uniform $\mathbf{E}$ field background to new position and momentum operators, X^μ and $P_\nu = -i\partial/\partial X^\nu$, such that the Klein-Gordon operator (2.2),

$$\begin{aligned} &-(\partial_\mu - ieA_\mu)(\partial^\mu - ieA^\mu) + m^2 \\ &\quad\to -p_t^2 + (p_z + eEt)^2 + p_x^2 + p_y^2 + m^2 \\ &\quad= -P_T^2 + P_Z^2 + P_X^2 + P_Y^2 + m^2 = 0, \end{aligned} \tag{2.19}$$

(with $P_X = p_x$, $P_Y = p_y$ and $P_Z = p_z$) becomes that of flat space with zero field [29]. The existence of this transformation and symmetry may be less surprising when it is recognized that there are two quantities,

$$P_T = (p_t^2 - 2eEt - e^2E^2t^2)^{\frac{1}{2}} = (p_z^2 + p_x^2 + p_y^2 + m^2)^{\frac{1}{2}} \tag{2.20a}$$

$$TP_Z + ZP_T = \frac{P_T}{eE}(eEz + p_t - P_T), \tag{2.20b}$$

that are conserved by virtue of the equation of motion (2.19), and (together with $P_Z = p_z$ which generates space Z translations) they generate T time translations and Lorentz boosts in the Z direction in the transformed (T, Z) coordinates. This dynamical maximal Poincaré symmetry in the constant, uniform $\mathbf{E}$ field is analogous to the maximal $O(4,1)$ point symmetry group of de Sitter space. In each case the existence of a maximally symmetric state $|v\rangle$ which enjoys the full symmetries of the background follows.

The expectation value of the electric current operator is given in the symmetric state $|v\rangle$ by

$$\langle v|j_z|v\rangle_R = 2e\int \frac{d^3\mathbf{k}}{(2\pi)^3}(k_z + eEt)\left[|v_\lambda(u)|^2 - \frac{1}{2\omega_\mathbf{k}(t)}\right] \tag{2.21}$$

where the second term is the lowest order adiabatic vacuum subtraction sufficient for the constant E field background [27]. Actually by changing integration variables from k_z to u and using the fact that both $|v_\lambda(u)|^2$ and $\omega_\mathbf{k}(t)$ are even functions of u, it is clear that both terms in the integrand of (2.21) are odd under $u \to -u$ and thus give vanishing contributions if integrated symmetrically in u. Hence as a consequence of time reversal invariance (or charge conjugation symmetry), the symmetric state $|v\rangle$ has exactly zero electric current expectation value

$$\langle v|j_z|v\rangle_R = \langle v|\mathbf{j}_\perp|v\rangle_R = 0 \tag{2.22}$$

at all times, by the symmetry of this state. Likewise the mean charge density $\langle v|\rho|v\rangle_R$ vanishes in this charge symmetric state. Thus the state $|v\rangle$ defined by (2.13)–(2.18) in a constant, uniform electric field background is an exact self-consistent solution of the semiclassical Maxwell equations,

$$\nabla\cdot\mathbf{E} = \langle v|\rho|v\rangle_R \tag{2.23a}$$

$$\nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t} = \langle v|\mathbf{j}|v\rangle_R \tag{2.23b}$$

with both sides vanishing identically. This is analogous to the maximally $O(4,1)$ symmetric and time reversal invariant CTBD state $|v\rangle$ which satisfies the semiclassical Einstein equations (1.7) in de Sitter space with a simple redefinition of Λ, since $\langle v|T^a_{\ b}|v\rangle_R = -\varepsilon_v\delta^a_{\ b}$, cf. Sec. IV.

B. Instability of the maximally symmetric state: Electric current

The existence of a 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 the imaginary part of the effective action and spontaneous decay rate of the electric field into particle/antiparticle pairs was first calculated by Schwinger [10]. By time reversal invariance the imaginary part of the effective action (which changes sign under time reversal) corresponding to the symmetric $|v\rangle$ state vanishes, in disagreement with Schwinger's result. As a precise coherent superposition of particle and antiparticle pairs for all modes, the time symmetric state defined by (2.13)–(2.18) is a very curious state indeed, corresponding to the rather unphysical boundary condition of each pair creation event being exactly balanced by its time reversed pair annihilation event, these pairs having been arranged with precisely the right phase relations to come from great distances at early times in order to effect just such a cancellation everywhere at all times. While mathematically allowed in a time reversal invariant background, it would be difficult to arrange such an artificial construction and fine tuning of initial and/or boundary conditions on the quantum state of the charged field with any macroscopic physical apparatus, and certainly it would not be produced with a more realistic adiabatic switching on and off of the electric field background in either finite time or over a finite region of space [26]. Nor does the state $|v\rangle$ minimize the Hamiltonian of the system which is time dependent in the gauge (2.1), or unbounded from below in the static gauge $A_0 = Ez$.

The above physical considerations and Schwinger's earlier result suggest that there should be an instability of the time symmetric state to nearby states in which the fine tuned cancellation between particle/antiparticle creation and annihilation events is slightly perturbed. In order to probe these nearby states we return to (2.3), and express its general solution in the form

$$f_\mathbf{k}(t) = A_\mathbf{k}v_\lambda(u) + B_\mathbf{k}v^*_\lambda(u), \tag{2.24}$$

with the (strictly time-independent) Bogoliubov coefficients required to obey

$$|A_\mathbf{k}|^2 - |B_\mathbf{k}|^2 = 1 \quad \text{for all } \mathbf{k}, \tag{2.25}$$

in order for the Wronskian condition

$$i\left(f^*_\mathbf{k}\frac{d}{dt}f_\mathbf{k} - f_\mathbf{k}\frac{d}{dt}f^*_\mathbf{k}\right) = 1 \tag{2.26}$$

to be satisfied. The Bogoliubov coefficients $(A_{\mathbf{k}}, B_{\mathbf{k}})$ may be regarded as specified by initial data $f_{\mathbf{k}}(t_0)$ and $\dot{f}_{\mathbf{k}}(t_0)$ at $t = t_0$ according to

$$A_{\mathbf{k}}(t_0) = i\left(v_\lambda^* \frac{d}{dt} f_{\mathbf{k}} - f_{\mathbf{k}} \frac{d}{dt} v_\lambda^*\right)\bigg|_{t=t_0} \tag{2.27a}$$

$$B_{\mathbf{k}}(t_0) = i\left(f_{\mathbf{k}} \frac{d}{dt} v_\lambda - v_\lambda \frac{d}{dt} f_{\mathbf{k}}\right)\bigg|_{t=t_0}. \tag{2.27b}$$

The quantized charged scalar field operator (2.17) may just as well be expressed in terms of these general mode functions (2.24) as

$$\mathbf{\Phi}(t,\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} [a_{\mathbf{k}}^f f_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} + b_{\mathbf{k}}^{f\dagger} f_{-\mathbf{k}}^*(t) e^{-i\mathbf{k}\cdot\mathbf{x}}], \tag{2.28}$$

where upon setting the Fourier components of (2.17) and (2.28) equal, the corresponding Fock space operators $a_{\mathbf{k}}^f, b_{\mathbf{k}}^{f\dagger}$ are related to the previous ones by

$$\begin{pmatrix} a_{\mathbf{k}}^v \\ b_{-\mathbf{k}}^{v\dagger} \end{pmatrix} = \begin{pmatrix} A_{\mathbf{k}} & B_{\mathbf{k}}^* \\ B_{\mathbf{k}} & A_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}}^f \\ b_{-\mathbf{k}}^{f\dagger} \end{pmatrix} \tag{2.29}$$

or its inverse

$$\begin{pmatrix} a_{\mathbf{k}}^f \\ b_{-\mathbf{k}}^{f\dagger} \end{pmatrix} = \begin{pmatrix} A_{\mathbf{k}}^* & -B_{\mathbf{k}}^* \\ -B_{\mathbf{k}} & A_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}}^v \\ b_{-\mathbf{k}}^{v\dagger} \end{pmatrix} \tag{2.30}$$

Hence if we define the state $|f\rangle$ by the condition

$$a_{\mathbf{k}}^f |f\rangle = b_{\mathbf{k}}^f |f\rangle = 0, \tag{2.31}$$

this state contains a nonzero expectation value

$$\langle f| a_{\mathbf{k}}^{v\dagger} a_{\mathbf{k}}^v |f\rangle = |B_{\mathbf{k}}|^2 = \langle f| b_{-\mathbf{k}}^{v\dagger} b_{-\mathbf{k}}^v |f\rangle \tag{2.32}$$

of v quanta. Conversely the $|v\rangle$ state contains a nonzero expectation value of f quanta. Since both the $|v\rangle$ and general $|f\rangle$ states are pure states, and each can be expressed as a coherent, squeezed state with respect to the other, it is best not to use the term particles for either of these expectation values, nor can one decide *a priori* which among them is the "correct" vacuum. This illustrates the fact that the questions of particle definition or which vacuum state to choose are not limited to de Sitter space or gravitational backgrounds only, but rather are characteristic of QFT in time-dependent and persistent classical background fields more generally.

The most general state which is both spatially homogeneous and charge symmetric is the mixed state with a density matrix $\rho_{f,N}$ and a finite expectation value of f quanta [27], which we denote by

$$\mathrm{Tr}(a_{\mathbf{k}}^{f\dagger} a_{\mathbf{k}}^f \rho_{f,N}) = N_{\mathbf{k}} = \mathrm{Tr}(b_{-\mathbf{k}}^{f\dagger} b_{-\mathbf{k}}^f \rho_{f,N}). \tag{2.33}$$

Computing the renormalized mean value of the electric current in this charge symmetric state, we find

$$\begin{aligned} &\mathrm{Tr}(j_z \rho_{f,N}) \\ &= 2e \int \frac{d^3\mathbf{k}}{(2\pi)^3} (k_z + eEt) \left[|f_{\mathbf{k}}(t)|^2 (1 + 2N_{\mathbf{k}}) - \frac{1}{2\omega_{\mathbf{k}}(t)} \right] \\ &= 4e \int \frac{d^3\mathbf{k}}{(2\pi)^3} (k_z + eEt) \Big\{ [N_{\mathbf{k}} + |B_{\mathbf{k}}|^2 (1 + 2N_{\mathbf{k}})] \\ &\quad \times |v_\lambda(u)|^2 + (1 + 2N_{\mathbf{k}}) \mathrm{Re}[A_{\mathbf{k}} B_{\mathbf{k}}^* v_\lambda^2(u)] \Big\}, \end{aligned} \tag{2.34}$$

where we have used (2.22) and (2.24)–(2.25) in arriving at the second expression. Charge asymmetric states or spatially inhomogeneous states with lower symmetry could be considered as well. In a general state with $B_{\mathbf{k}} \neq 0$ or $N_{\mathbf{k}} \neq 0$, the charge conjugation and time reversal symmetry of the background is broken and the current $\mathrm{Tr}(j_z \rho_{f,N}) \neq 0$. Because such states correspond to charged particle/antiparticle excitations that are rapidly accelerated to ultrarelativistic energies by the background electric field, they lead to persistent currents that do not decay and which destabilize the constant electric field background through the semiclassical Maxwell equation (2.23b).

To see this requires only a qualitative understanding of the integrand in (2.34). The three terms $u|v_\lambda|^2$, $u\,\mathrm{Re}\,(v_\lambda^2)$, and $u\,\mathrm{Im}\,(v_\lambda^2)$ appearing in the integrand of (2.34) are shown as functions of u for several values of λ in Figs. 3–5. In Fig. 3 the saturation of the function $u|v_\lambda|^2$ at large u is the result of acceleration of charged scalar particles to ultrarelativistic energies by the electric field, where they make a constant contribution to the current integrand. Hence if $N_{\mathbf{k}}$ and/or $|B_{\mathbf{k}}|^2$ in (2.34) is nonzero for any range of $\mathbf{k}$, such modes will make a contribution

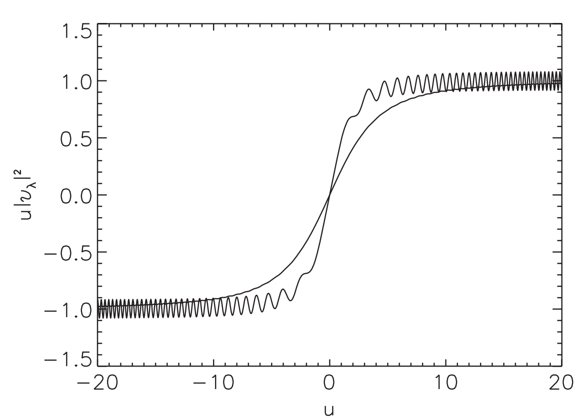

FIG. 3. The $u|v_\lambda|^2$ integrand in the current in (2.21) or (2.34) in units of $2eE$ as a function of u for two different values of λ, cf. (2.5) and (2.13). The curve with large oscillations is for $\lambda = 1$ while the oscillations are much smaller in the curve for $\lambda = 5$.

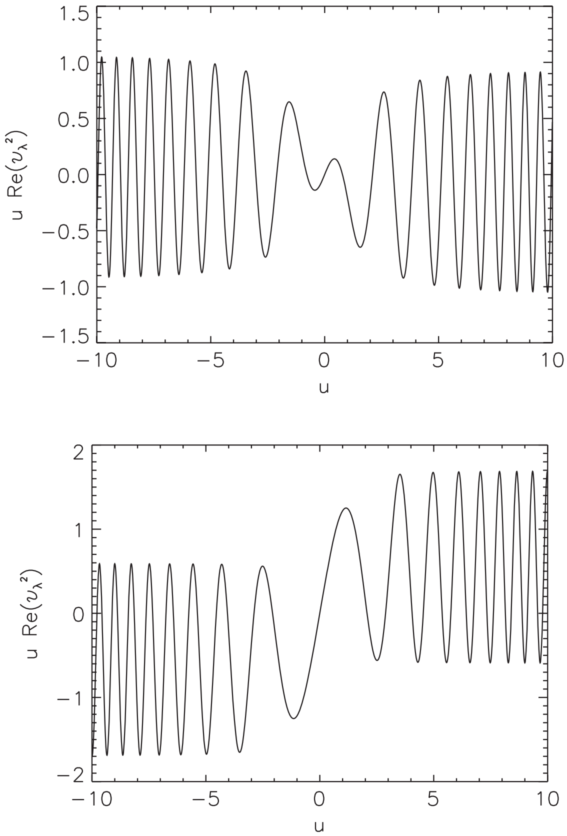

FIG. 4. The $u\,\mathrm{Re}(\upsilon_\lambda^2)$ integrand in the current in (2.34) in units of $2eE$ as a function of u for two different values of λ. The upper panel is for $\lambda = 1$ while the lower panel is for $\lambda = 0.1$ chosen to accentuate the asymmetry in $u \leftrightarrow -u$.

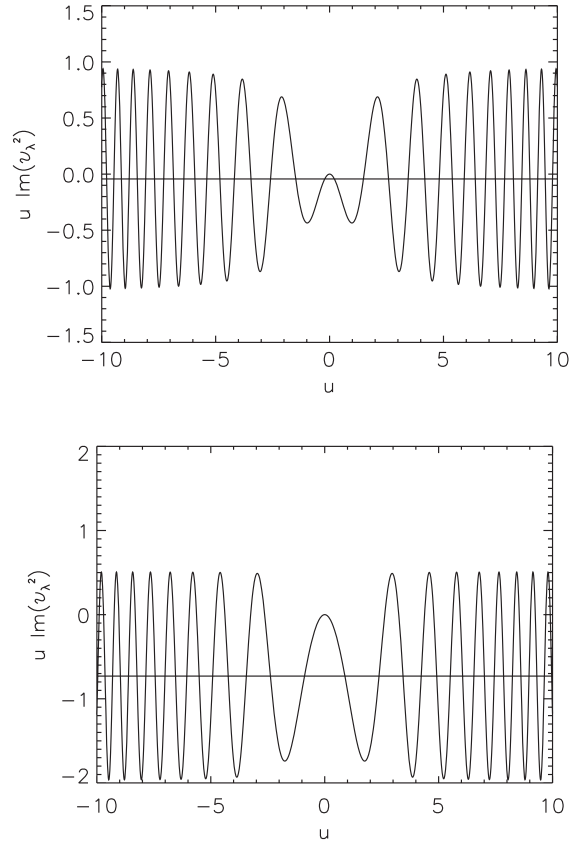

FIG. 5. The $u\,\mathrm{Im}(\upsilon_\lambda^2)$ integrand in the current in (2.34) in units of $2eE$ as a function of u for the same values of $\lambda = 1$ (upper) and $\lambda = 0.1$ (lower) as in Fig. 4. Shown also is the horizontal line at $-\exp(-\pi\lambda)$ around which the average amplitude of the oscillations are displaced.

to the current proportional to the occupied phase space volume $\int d^3\mathbf{k}$ which can therefore give an arbitrarily large $\langle j_z \rangle$ at late times, $u \gg 1$.

The oscillatory terms in the real and imaginary parts of $u\upsilon_\lambda^2$ are shown in Figs. 4–5. The envelope of the oscillations shows a saturation behavior at large $|u|$ similar to Fig. 3. For smaller λ the oscillations are significantly offset from the horizontal axis, by $\pm\exp(-\pi\lambda)$, showing that there will also be a net contribution to the current from modes with $A_\mathbf{k} B_\mathbf{k}^* \neq 0$. Hence these contributions to $\langle j_z \rangle$ can also become arbitrarily large if the range of $\mathbf{k}$ for which $A_\mathbf{k} B_\mathbf{k}^*$ is nonzero is large.

An interesting special case in which to evaluate (2.34) is the adiabatic vacuum state of initial data,

$$\begin{aligned} f_\mathbf{k}(t_0) &= \frac{1}{\sqrt{2\omega_\mathbf{k}(t_0)}}, \\ \dot{f}_\mathbf{k}(t_0) &= \left(-i\omega_\mathbf{k} - \frac{\dot{\omega}_\mathbf{k}}{2\omega_\mathbf{k}}\right) f_\mathbf{k}\Big|_{t=t_0}, \end{aligned} \tag{2.35}$$

and $N_\mathbf{k} = 0$. We denote this pure state which matches the lowest order adiabatic vacuum state at the particular time $t = t_0$ by $|t_0\rangle$. With these initial conditions it is shown in [9] that the Bogoliubov coefficients are given approximately by

$$A_\mathbf{k}(t_0) \simeq A_\lambda^{\mathrm{in}} \theta(-k_z - eEt_0) + A_\lambda^{\mathrm{out}} \theta(k_z + eEt_0) \tag{2.36a}$$

$$B_\mathbf{k}(t_0) \simeq B_\lambda^{\mathrm{in}} \theta(-k_z - eEt_0) + B_\lambda^{\mathrm{out}} \theta(k_z + eEt_0), \tag{2.36b}$$

where

$$A_\lambda^{\mathrm{in}} = A_\lambda^{\mathrm{out}*} = \sqrt{\frac{\pi}{2}} \left(\frac{2}{\lambda}\right)^{\frac{i\lambda}{2}} \left[\left(\frac{\lambda}{2}\right)^{\frac{1}{4}} \frac{1}{\Gamma(\frac{3}{4} - \frac{i\lambda}{2})} + \left(\frac{2}{\lambda}\right)^{\frac{1}{4}} \frac{e^{\frac{i\pi}{4}}}{\Gamma(\frac{1}{4} - \frac{i\lambda}{2})}\right] \times \exp\left(\frac{i\lambda}{2} - \frac{i\pi}{8} - \frac{\pi\lambda}{4}\right) \tag{2.37a}$$

$$B_\lambda^{\mathrm{in}} = B_\lambda^{\mathrm{out}*} = \sqrt{\frac{\pi}{2}} \left(\frac{2}{\lambda}\right)^{\frac{i\lambda}{2}} \left[\left(\frac{\lambda}{2}\right)^{\frac{1}{4}} \frac{1}{\Gamma(\frac{3}{4} - \frac{i\lambda}{2})} - \left(\frac{2}{\lambda}\right)^{\frac{1}{4}} \frac{e^{\frac{i\pi}{4}}}{\Gamma(\frac{1}{4} - \frac{i\lambda}{2})}\right] \times \exp\left(\frac{i\lambda}{2} - \frac{i\pi}{8} - \frac{\pi\lambda}{4}\right). \tag{2.37b}$$

In a better approximation, the step functions in (2.36) would be smooth functions which interpolate between the two limits, but this simple approximation is sufficient to illustrate the main features of the current expectation value which is its linear growth in time from the initial time t_0.

Substituting (2.36) into (2.34) with $N_{\mathbf{k}} = 0$, changing variables from $(k_z, k_\perp)$ to (u, λ), and making use of the fact that the functions $u|v_\lambda(u)|^2$ and $u\,\mathrm{Re}[v_\lambda^2(u)]$ are odd functions of u (cf. Figs. 3–4), while $u\,\mathrm{Im}[v_\lambda^2(u)]$ is even (cf. Fig. 5), we obtain

$$\begin{aligned}\langle t_0|j_z|t_0\rangle_R &\simeq \frac{4e}{(2\pi)^2}\int_0^\infty k_\perp dk_\perp\Big\{\int_{-\infty}^{-eEt_0} dk_z(k_z+eEt)[|B_\lambda^{\rm in}|^2|v_\lambda(u)|^2 + \mathrm{Re}[A_\lambda^{\rm in}B_\lambda^{\rm in*}v_\lambda^2(u)]]\\ &\quad + \int_{-eEt_0}^\infty dk_z(k_z+eEt)[|B_\lambda^{\rm out}|^2|v_\lambda(u)|^2 + \mathrm{Re}[A_\lambda^{\rm out}B_\lambda^{\rm out*}v_\lambda^2(u)]]\Big\}\\ &= -\frac{e^2E}{\pi^2}\int_{m^2/2eE}^\infty d\lambda\,\mathrm{Im}[A_\lambda^{\rm in}B_\lambda^{\rm in*}]\int_0^{\sqrt{2eE}(t-t_0)} du\,u\,\mathrm{Im}[v_\lambda^2(u)]\\ &= \frac{e^3E^2}{2\pi^2}\frac{1}{\sqrt{2eE}}\int_{m^2/2eE}^\infty d\lambda e^{-\pi\lambda}\int_0^{\sqrt{2eE}(t-t_0)} du[uf_\lambda^{(0)}(u)f_\lambda^{(1)}(u)], \end{aligned} \tag{2.38}$$

since from (2.37), or Eq. (4.20b) of Ref. [9] and (2.13),

$$|B_\lambda^{\rm in}|^2 = |B_\lambda^{\rm out}|^2 \tag{2.39a}$$

$$\mathrm{Re}[A_\lambda^{\rm in}B_\lambda^{\rm in*}] = \mathrm{Re}[A_\lambda^{\rm out}B_\lambda^{\rm out*}] \tag{2.39b}$$

$$\mathrm{Im}[A_\lambda^{\rm in}B_\lambda^{\rm in*}] = -\mathrm{Im}[A_\lambda^{\rm out}B_\lambda^{\rm out*}] = -\frac{1}{2}\mathrm{Im}B_\lambda^{\rm tot} = \frac{1}{2}e^{-\pi\lambda} \tag{2.39c}$$

$$\mathrm{Im}[v_\lambda^2(u)] = -\frac{1}{\sqrt{2eE}}f_\lambda^{(0)}(u)f_\lambda^{(1)}(u). \tag{2.39d}$$

Because of the offset from the u axis of $uf_\lambda^{(0)}(u)f_\lambda^{(1)}(u) = -u\mathrm{Im}[v_\lambda^2(u)]$ by $e^{-\pi\lambda}$, around which the oscillations average to zero (cf. Fig. 5), for large $t - t_0 \to \infty$ the u integral in (2.38) is

$$\int_0^{\sqrt{2eE}(t-t_0)} du[uf_\lambda^{(0)}(u)f_\lambda^{(1)}(u)] \to \int_0^{\sqrt{2eE}(t-t_0)} due^{-\pi\lambda} = \sqrt{2eE}(t-t_0)e^{-\pi\lambda} \tag{2.40}$$

and hence (2.38) gives for late times

$$\langle t_0|j_z|t_0\rangle_R \to \frac{e^3E^2}{4\pi^3}e^{-\pi m^2/eE}(t-t_0), \tag{2.41}$$

which is the same result as that of Eq. (5.22) in Ref. [9], which was obtained much more naturally in the adiabatic particle basis by consideration of particle creation events. That treatment makes it clear that the growth of the current is a cumulative effect of particle creation from the quantum vacuum which continues unabated as long as the constant electric field is maintained.

Thus there are states for which the current grows linearly with time related to the steady rate of particle creation in a constant electric field background. Moreover it is clear from the penultimate line of (2.38) that any perturbation of the symmetric $|v\rangle$ state with Bogoliubov coefficients $A_{\mathbf{k}}, B_{\mathbf{k}}$ of the form (2.36) obeying the conditions $A_\lambda^{\rm in} = A_\lambda^{\rm out*}$, $B_\lambda^{\rm in} = B_\lambda^{\rm out*}$ of (2.37), having constant but nonzero support for arbitrarily large and negative k_z will produce a cumulative effect on the current similar to (2.41), so that $\langle j_z\rangle$ continues to grow linearly with time for arbitrarily long times. This linear growth with time implies that however small the coupling e and the coefficient $\mathrm{Im}\,[A_\lambda^{\rm in}B_\lambda^{\rm in*}]$ (which can be enhanced by taking $N_{\mathbf{k}} > 0$), the current must eventually influence the background field through the semiclassical Maxwell equation (2.23b). Thus the symmetric $|v\rangle$ state in a fixed constant uniform electric field background is unstable to perturbations of the kind (2.36). If $B_{\mathbf{k}}$ has nonzero support up to some large but finite negative value $(k_z)_{\rm min} = -K_z$ the linear growth in (2.41) will be cut off at $t - t_0 = K_z/eE$ but still be large and produce a large backreaction through (2.23b).

If one goes beyond the simple mean field approximation considered here, it is also clear on physical grounds that the introduction of a single electrically charged particle into the $|v\rangle$ state will cause it to be accelerated by the electric field to arbitrarily large energies, which would allow it to emit photons and produce additional charged pairs resulting in an electromagnetic avalanche. Allowing these additional channels opened up by self-interactions makes the physical instability of the symmetric $|v\rangle$ state to small perturbations more obvious, although that instability already exists even without self-interactions, in the mean field approximation, as (2.41) and (2.23b) show.

This example of the quantum states in a constant, uniform external electric field shows quite clearly that the most symmetric state, with the full symmetry group of the background need not be the stable ground state of the system. In this case it is well known that the background is unstable to particle creation. In the accompanying paper [9] we have shown how the same conclusion follows in de

Sitter space, for essentially the same reasons. The treatment above shows that one need not be committed to any definition of particles to discover the instability of the electric field background by perturbations of the symmetric $|v\rangle$ state which have support at large canonical momentum $|k_z|$. For $k_z < 0$ this may correspond to small physical kinetic momentum $k_z + eEt_0$ at some early initial time t_0. The unlimited growth of the physical momentum $k_z + eEt$ with time for fixed canonical momentum k_z in terms of which the initial state is specified is the essential feature, and this feature is found in gravitational backgrounds such as de Sitter space as well.

C. Relation to quantum chiral anomaly in two dimensions

The linear secular growth of the current in a background constant electric field can also be understood through the Schwinger anomaly in $1+1$ dimensions [30]. For that comparison we drop the $d^2\mathbf{k}_\perp/(2\pi)^2$ integral in (2.34) to reduce to $1+1$ dimensions, and further set the mass $m=0$. We have then

$$\langle j_z\rangle_{2d} \to \frac{e^2 E}{\pi}(t-t_0) \tag{2.42}$$

at late times. Since scalars are essentially the same as fermions in $1+1$ dimensions one can use the bosonization results [31] for fermionic QED to express the current in the form

$$\langle j^\mu\rangle = \frac{e}{\sqrt{\pi}}\epsilon^{\mu\nu}\partial_\nu\chi, \tag{2.43}$$

where $\epsilon^{\mu\nu}$ is the antisymmetric symbol in two dimensions and χ is a pseudoscalar field whose derivative is the chiral current

$$\langle j^{\mu 5}\rangle = \frac{1}{\sqrt{\pi}}\partial^\mu\chi. \tag{2.44}$$

This current has the well-known chiral anomaly [32]

$$\partial_\mu\langle j^{\mu 5}\rangle_{2d} = \frac{1}{\sqrt{\pi}}\Box\chi = \frac{e}{2\pi}\epsilon^{\mu\nu}F_{\mu\nu} = \frac{eE}{\pi} \tag{2.45}$$

in a background electric field. The second-order Eq. (2.45) for χ with the anomaly source in a constant, uniform field has solutions independent of z of the form

$$\frac{1}{\sqrt{\pi}}\chi = \frac{eE}{2\pi}(t-t_0)^2. \tag{2.46}$$

Substituting this value of χ into the electric current (2.43) gives

$$\langle j_z\rangle_{2d} = \frac{e}{\sqrt{\pi}}\dot{\chi} = \frac{e^2 E}{\pi}(t-t_0), \tag{2.47}$$

which recovers (2.42). Thus the linear secular growth of the current with time in the massless limit is related to the two-dimensional chiral anomaly and the particular z independent solution (2.46) to the pseudoscalar field Eq. (2.45). This particular solution to (2.45) is associated with the spatially homogeneous initial state condition (2.35) and state specified by the mode functions (2.24) and (2.36).

It is interesting to note that although the anomaly Eq. (2.45) is Lorentz invariant—because it is an inhomogeneous equation—*none* of its solutions are Lorentz invariant. Thus the maximal Poincaré symmetry of the fixed electric field background is necessarily broken by the solutions to the anomaly Eq. (2.45), which leads to a spontaneous breaking of symmetry of the background, at least in the semiclassical approximation and neglecting backreaction [33]. This may be seen also from the effective action corresponding to the two-dimensional chiral anomaly [30,34],

$$\begin{aligned} S^{2D}_{\rm anom}[\chi] &= \frac{e^2}{8\pi}\int d^2x\int d^2x'[\epsilon^{\mu\nu}F_{\mu\nu}]_x\Box^{-1}(x,x')[\epsilon^{\alpha\beta}F_{\alpha\beta}]_{x'} \\ &= \frac{1}{2}\int d^2x\left[-\chi\Box\chi + \frac{e}{\sqrt{\pi}}\chi\epsilon^{\mu\nu}F_{\mu\nu}\right], \end{aligned} \tag{2.48}$$

where $\Box^{-1}(x,x')$ is the Green’s function inverse of the scalar wave operator $\Box$ in two dimensions. As is well known, the usual construction of the Feynman Green’s function for a massless scalar in two dimensions is infrared divergent due to the constant $k=0$ mode, and consequently no Lorentz invariant Feynman function exists in this case. Green’s functions $\Box^{-1}(x,x')$ obeying different boundary conditions exist, but these necessarily break some of the continuous or discrete symmetries of the background. Thus the form of the effective action of the two-dimensional chiral anomaly (2.48), together with the absence of a Lorentz invariant Feynman Green’s function $\Box^{-1}(x,x')$ due to infrared divergences is sufficient to conclude that the maximally symmetric state in a uniform constant background $\frac{1}{2}\epsilon^{\mu\nu}F_{\mu\nu} = E$ is sensitive to noninvariant initial and/or boundary conditions which break that maximal symmetry. The linear growth of the current found in (2.41) and reproduced by the solution (2.46) in (2.47) is symptomatic of that necessary breaking of the maximal symmetry of the classical background by the quantum chiral anomaly.

It is also interesting that this connection with the anomaly of massless fields in two dimensions survives in four dimensions and even if the field has a nonzero mass m, whose main effect is to suppress the coefficient of the linear growth by the Schwinger tunneling factor $\exp(-\pi m^2/eE)$. We shall see there is also an interesting connection to a quantum anomaly of massless fields in four dimensional de Sitter space, a local condensate bilinear of the underlying quantum field(s) analogous to χ, and simple arguments analogous to (2.43)–(2.47) which lead directly to the analogous conclusion of instability of the symmetric state and breaking of maximal de Sitter invariance in that case as well.

III. $O(4)$ INVARIANT STATES IN DE SITTER SPACE

Turning to our primary topic of de Sitter space, we develop the quantization and discussion of possible vacuum states in de Sitter space analogously to the electric field case of the previous section. For an uncharged scalar field Φ satisfying the free wave equation,

$$(-\Box + M^2)\Phi \equiv \left[-\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x^a}\left(\sqrt{-g}g^{ab}\frac{\partial}{\partial x^b}\right) + M^2\right]\Phi = 0, \tag{3.1}$$

in a gravitational background, with ξ the curvature coupling. The effective mass

$$M^2 \equiv m^2 + \xi R = m^2 + 12\xi H^2 \tag{3.2}$$

is a constant since the Ricci scalar $R = 12H^2$ is a constant in de Sitter spacetime. In the geodesically complete coordinates (1.2) the wave equation (3.1) may be separated into a complete basis of functions of cosmological time $y_k(u)$ times $Y_{klm_l}(\hat{N})$, the spherical harmonics on $\mathbb{S}^3$. A unit vector on $\mathbb{S}^3$ is denoted by $\hat{N}$ with coordinates

$$\hat{N}(\chi,\theta,\phi) = (\sin\chi\hat{\mathbf{n}},\cos\chi) = (\cos\chi,\sin\chi\cos\theta,\sin\chi\sin\theta\sin\phi,\sin\chi\sin\theta\cos\phi). \tag{3.3}$$

The $Y_{klm}(\hat{N})$ harmonics are eigenfunctions of the scalar Laplacian on the unit $\mathbb{S}^3$ satisfying

$$-\Delta_3 Y_{klm_l} = -\frac{1}{\sin^2\chi}\left[\frac{\partial}{\partial\chi}\sin^2\chi\frac{\partial}{\partial\chi} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]Y_{klm_l} = (k^2-1)Y_{klm_l}, \tag{3.4}$$

with the range of the integer $k = 1, 2, \ldots$ taken to be strictly positive, conforming to the notation of [13] and [35]. These $\mathbb{S}^3$ harmonics are given in terms of Gegenbauer functions $C^{l+1}_{k-l-1}(\cos\chi)$ and the familiar $\mathbb{S}^2$ spherical harmonics $Y_{lm_l}(\hat{\mathbf{n}})$ in the form [28]

$$Y_{klm_l}(\hat{N}) = 2^l l!\sqrt{\frac{2k(k-l-1)!}{\pi(k+l)!}}(\sin\chi)^l C^{l+1}_{k-l-1}(\cos\chi)Y_{lm_l}(\hat{\mathbf{n}}), \tag{3.5}$$

with $l = 0, 1, \ldots k-1$ and $m_l = -l, \ldots, l$, normalized so that

$$\int_{\mathbb{S}^3} d^3\Sigma Y^*_{k'l'm'_l}Y_{klm_l} = \int_0^\pi d\chi \sin^2\chi \int_0^\pi d\theta\sin\theta \times \int_0^{2\pi} d\phi Y^*_{k'l'm'_l}Y_{klm_l} = \delta_{k'k}\delta_{l'l}\delta_{m'_l m_l}. \tag{3.6}$$

Note also that $Y^*_{klm_l}(\hat{N}) = Y_{kl-m_l}(\hat{N})$.

The time-dependent functions $y_k(u)$ satisfy

$$\left[\frac{d^2}{du^2} + 3\tanh u\frac{d}{du} + (k^2-1)\mathrm{sech}^2 u + \left(\gamma^2 + \frac{9}{4}\right)\right]y_k = 0, \tag{3.7}$$

where the dimensionless parameter γ is defined by

$$\gamma \equiv \sqrt{\frac{M^2}{H^2} - \frac{9}{4}} \equiv i\nu. \tag{3.8}$$

In the massive case $M^2 > \frac{9}{4}H^2$ (the *principal series*), γ is real and positive. With the change of variables to $z = (1 - i\sinh u)/2$, the mode equation (3.7) can be recast in the form of the hypergeometric equation. The fundamental complex solution $y_k \to \upsilon_{k\gamma}(u)$ may be taken to be

$$\upsilon_{k\gamma}(u) \equiv Hc_{k\gamma}(\mathrm{sech}u)^{k+1}(1 - i\sinh u)^k \times F\left(\frac{1}{2} + i\gamma, \frac{1}{2} - i\gamma; k+1; \frac{1 - i\sinh u}{2}\right), \tag{3.9}$$

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

$$c_{k\gamma} \equiv \frac{1}{k!}\left[\frac{\Gamma(k+\frac{1}{2}+i\gamma)\Gamma(k+\frac{1}{2}-i\gamma)}{2}\right]^{\frac{1}{2}} \tag{3.10}$$

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

$$iHa^3(u)\left[\upsilon^*_{k\gamma}\frac{d}{du}\upsilon_{k\gamma} - \upsilon_{k\gamma}\frac{d}{du}\upsilon^*_{k\gamma}\right] = 1 \tag{3.11}$$

for all k, where $a(u) = H^{-1}\cosh u$ is the scale factor in coordinates (1.2). Note that under time reversal $u \to -u$, the mode function (3.9) goes to its complex conjugate,

$$\upsilon_{k\gamma}(-u) = \upsilon^*_{k\gamma}(u), \tag{3.12}$$

for all $M^2 > 0$.

If $0 < M^2 \le \frac{9}{4}H^2$, (3.8)–(3.12) continue to hold by analytic continuation to pure imaginary $\gamma \equiv i\nu$, with $\upsilon_{k\gamma} \to \upsilon_{k,\nu}$. The mode functions (3.9) reduce to elementary functions in the massless, conformally coupled case,

$$m=0,\qquad \xi=\frac{1}{6},\qquad \nu=\frac{1}{2}:$$
$$\upsilon_{k,\frac{1}{2}}=\frac{H}{\sqrt{2k}}\mathrm{sech}u(\mathrm{sech}u-i\tanh u)^k = \frac{H}{\sqrt{2k}}\cos\eta\; e^{-ik\eta}, \tag{3.13}$$

and in the massless, minimally coupled case,

$$m=0,\qquad \xi=0,\qquad \nu=\frac{3}{2}:$$
$$\upsilon_{k,\frac{3}{2}}=\frac{H(k\mathrm{sech}u+i\tanh u)}{\sqrt{2k(k^2-1)}}(\mathrm{sech}u-i\tanh u)^k = \frac{H(k\cos\eta+i\sin\eta)}{\sqrt{2k(k^2-1)}}e^{-ik\eta},\qquad k=2,3,\ldots, \tag{3.14}$$

where the conformal time variable η is given by, cf. Fig. 2,

$$\eta\equiv\sin^{-1}(\tanh u)\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right),\qquad \cos\eta=\mathrm{sech}u,\qquad \tan\eta=\sinh u. \tag{3.15}$$

The complex positive frequency modes $\upsilon_{k,\nu}$ of (3.9), (3.14) are undefined for the case $\nu=\frac{3}{2}$, $k=1$ since the solutions of (3.7) are nonoscillatory in this case, and must be treated separately [13,36]. This leads to the nonexistence of a de Sitter invariant vacuum state or Feynman Green's function $\square^{-1}(x,x')$ for a massless, minimally coupled scalar in de Sitter space [36,37], that is similar to that for a massless scalar in two dimensional flat space discussed in Sec. II C.

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

$$\mathbf{\Phi}(u,\hat{N})=\sum_{k=1}^{\infty}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}\{a^{\upsilon}_{klm_l}\upsilon_{k\gamma}(u)Y_{klm_l}(\hat{N}) + a^{\upsilon\dagger}_{klm_l}\upsilon^*_{k\gamma}(u)Y^*_{klm_l}(\hat{N})\} \tag{3.16}$$

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

$$[a^{\upsilon}_{klm_l},a^{\upsilon\dagger}_{k'l'm'_l}]=\delta_{kk'}\delta_{ll'}\delta_{m_lm'_l}. \tag{3.17}$$

With (3.6), (3.11), and (3.17) the canonical equal time field commutation relation,

$$[\mathbf{\Phi}(u,\hat{N}),\mathbf{\Pi}(u,\hat{N}')]=i\delta_\Sigma(\hat{N},\hat{N}'), \tag{3.18}$$

is satisfied, where $\mathbf{\Pi}=\sqrt{-g}\dot{\mathbf{\Phi}}=Ha^3\frac{\partial\mathbf{\Phi}}{\partial u}$ is the field momentum operator conjugate to $\mathbf{\Phi}$, the overdot denotes the time derivative $H\partial/\partial u$ 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 $|\upsilon\rangle$ [2,3,6] is defined by

$$a^{\upsilon}_{klm_l}|\upsilon\rangle=0\quad\forall\ k,l,m_l, \tag{3.19}$$

and is invariant under the full $O(4,1)$ isometry group of the complete de Sitter manifold, including under the discrete inversion symmetry of all coordinates in the embedding space, $X^A\to -X^A$ (cf. Fig. 1), or $(u,\hat{N})\to(-u,-\hat{N})$, which is not continuously connected to the identity. The Feynman Green's function in this maximally symmetric state is invariant under $O(4,1)$ and also coincides with that obtained by analytic continuation from the Euclidean $\mathbb{S}^4$ for $M^2>0$ with full $O(5)$ symmetry [5]. As in the electric field example of Sec. II the existence or construction of a maximally symmetric $O(4,1)$ invariant state does not imply that this state is a stable vacuum.

Alternative Fock representations in real u time are clearly possible. For example, since the general solution of (3.7) may be written as the linear combination

$$y_k(u)=A_k\upsilon_{k\gamma}(u)+B_k\upsilon^*_{k\gamma}(u), \tag{3.20}$$

and normalized by (3.11) in the same way by requiring

$$iHa^3(u)\left[y^*_k\frac{d}{du}y_k-y_k\frac{d}{du}y^*_k\right]=iH\left[f^*_k\frac{d}{du}f_k-f_k\frac{d}{du}f^*_k\right] = |A_k|^2-|B_k|^2=1, \tag{3.21}$$

the general functions $y_k\equiv a^{-\frac{3}{2}}f_k$ may just as well be chosen as a basis of quantization of the $\mathbf{\Phi}$ field by

$$\mathbf{\Phi}(u,\hat{N})=\sum_{k=1}^{\infty}\sum_{l=0}^{k-1}\sum_{m=-l}^{l}[a^{f}_{klm_l}y_k(u)Y_{klm_l}(\hat{N}) + a^{f\dagger}_{klm_l}y^*_k(u)Y^*_{klm_l}(\hat{N})]. \tag{3.22}$$

The Bogoliubov transformation between the corresponding Fock space operators is

$$\begin{pmatrix}a^{\upsilon}_{klm_l}\\ a^{\upsilon\dagger}_{kl-m_l}\end{pmatrix}=\begin{pmatrix}A_k & B^*_k\\ B_k & A^*_k\end{pmatrix}\begin{pmatrix}a^{f}_{klm_l}\\ a^{f\dagger}_{kl-m_l}\end{pmatrix}, \tag{3.23}$$

and its inverse is

$$\begin{pmatrix}a^{f}_{klm_l}\\ a^{f\dagger}_{kl-m_l}\end{pmatrix}=\begin{pmatrix}A^*_k & -B^*_k\\ -B_k & A_k\end{pmatrix}\begin{pmatrix}a^{\upsilon}_{klm_l}\\ a^{\upsilon\dagger}_{kl-m_l}\end{pmatrix}, \tag{3.24}$$

analogous to (2.24)–(2.30).

The mode function $f_k=a^{\frac{3}{2}}y_k$ also satisfies the equation of an harmonic oscillator,

$$\frac{d^2f_k}{du^2}+\left[\left(k^2-\frac{1}{4}\right)\mathrm{sech}^2u+\gamma^2\right]f_k=0, \tag{3.25}$$

analogous to (2.6), and (3.25), which is the starting point for an adiabatic or WKB analysis of particle creation in [9]. Here we note that because of (3.21) the commutation relations (3.17) are also satisfied by $a^f_{klm_l}, a^{f\dagger}_{klm_l}$, as is the canonical field commutation relation (3.18). Hence we may define a state $|f\rangle$ corresponding to the general solution (3.20) of (3.7) or (3.25) by

$$a^f_{klm_l}|f\rangle = 0 \quad \forall\ k, l, m_l, \tag{3.26}$$

for any set of complex coefficients $\{A_k, B_k\}$ satisfying (3.21). Since the solutions $y_k(u)Y_{klm_l}(\hat{N})$ at fixed k form an irreducible representation of the group $O(4)$ for any $\{A_k, B_k\}$, these states are invariant under $O(4)$ rotations of $\mathbb{S}^3$, but not the full $O(4,1)$ de Sitter group (unless $A_k = 1$, $B_k = 0$ for all k).

The $O(4)$ invariant states are associated with a preferred u time slicing which breaks the $O(4,1)$ symmetry. The Bogoliubov coefficients $\{A_k, B_k\}$ and hence the particular $|f\rangle$ state may be regarded as specified by initial data $y_k(u_0)$ and $\dot{y}_k(u_0)$ on the $u = u_0$ Cauchy surface according to

$$A_k(u_0) = iHa^3\left(v^*_{k\gamma}\frac{d}{du}y_k - y_k\frac{d}{du}v^*_{k\gamma}\right)\Big|_{u=u_0} \tag{3.27a}$$

$$B_k(u_0) = iHa^3\left(y_k\frac{d}{du}v_{k\gamma} - v_{k\gamma}\frac{d}{du}y_k\right)\Big|_{u=u_0}. \tag{3.27b}$$

States with lower symmetry than $O(4)$ may be obtained by considering Bogoliubov transformations more general than (3.24), mixing a^v and $a^{v\dagger}$ of different (klm_l). For example if the relation (3.24) is generalized to $a_{klm_l} = A^*_{kk'}a^v_{k'lm_l} - B^*_{kk'}a^{v\dagger}_{k'l-m_l}$ so that the Bogoliubov coefficients are (non-diagonal) matrices in k, k' (but still diagonal in l, m_l), the corresponding states (3.26) are $O(3)$ invariant only. These are appropriate for the static coordinates of de Sitter space. All states related to $|v\rangle$ by exact Bogoliubov transformations of this kind are pure states and related to each other by a unitary transformation [8], whether they involve different (klm_l) or not.

The expectation value of $a^{v\dagger}_{klm_l}a^v_{klm_l}$ is nonvanishing in the general $|f\rangle$ state defined by (3.20), (3.24) and (3.26),

$$\langle f|a^{v\dagger}_{klm_l}a^v_{klm_l}|f\rangle = |B_k|^2 \tag{3.28a}$$

$$\langle f|a^v_{klm_l}a^v_{kl-m_l}|f\rangle = A_kB^*_k = \langle f|a^{v\dagger}_{klm_l}a^{v\dagger}_{kl-m_l}|f\rangle^*. \tag{3.28b}$$

Hence the general $|f\rangle$ vacuum state apparently contains particles defined with respect to the de Sitter invariant state $|v\rangle$. However, the converse is also true as the expectation values,

$$\langle v|a^{f\dagger}_{klm_l}a^f_{klm_l}|v\rangle = |B_k|^2 \tag{3.29a}$$

$$\langle v|a^f_{klm_l}a^f_{kl-m_l}|v\rangle = -A^*_kB_k = \langle v|a^{f\dagger}_{klm_l}a^{f\dagger}_{kl-m_l}|v\rangle^*, \tag{3.29b}$$

are also nonzero, so that the de Sitter invariant vacuum may equally well be said to contain particles with respect to the general $O(4)$ basis states $|f\rangle$. Since each of the $|f\rangle$ states in either case is in fact a coherent, squeezed pure state with respect to the others, with all exact quantum phase correlations maintained, it is better not to attach the label of "particles" to either set of expectation values (3.28) or (3.29), or the term "vacuum" to any particular state in de Sitter space at this point. As in the electric field case, mixed states which are $O(4)$ invariant can be defined through a density matrix $\rho_{f,N}$ with [38]

$$\mathrm{Tr}(\rho_{f,N}a^{f\dagger}_{klm_l}a^f_{klm_l}) = N_k, \tag{3.30}$$

and $N_k = 0$ reducing to the pure $|f\rangle$ state defined in (3.26). A nonzero N_k above the arbitrary $O(4)$ invariant $|f\rangle$ vacuum is also best not identified with any physical particle number. In previous works and in a companion paper to this one [9,35], we give a definition of physical particle number in de Sitter space based on an adiabatic or slowly varying positive frequency basis [4].

IV. ENERGY-MOMENTUM TENSOR OF O(4) INVARIANT STATES

The behavior of perturbations of the CTBD $O(4,1)$ symmetric state in de Sitter space may be studied through the energy-momentum-stress tensor, and the potential backreaction effects on the background geometry through the semiclassical Einstein equations (1.7), analogous to perturbations of the symmetric $|v\rangle$ state and backreaction effects of the electric current through the semiclassical Maxwell equation (2.23b).

The conserved energy-momentum-stress tensor of the free scalar field is

$$T_{ab} = (\nabla_a\Phi)(\nabla_b\Phi) - \frac{g_{ab}}{2}\left[g^{cd}(\nabla_c\Phi)(\nabla_d\Phi) + m^2\Phi^2\right] + \xi\left[R_{ab} - \frac{g_{ab}}{2}R - \nabla_a\nabla_b + g_{ab}\Box\right]\Phi^2. \tag{4.1}$$

If the Heisenberg field operator in the general $O(4)$ basis (3.22) is substituted into this expression, and (3.30) is used, the expectation value of T_{ab} in the general $|f\rangle$ state may be expressed as a sum over modes. Since these states are spatially homogeneous and isotropic, and $O(4)$ invariant, we find that

$$\mathrm{Tr}(\rho_{f,N}T^u{}_u) = -\varepsilon_{f,N} \tag{4.2a}$$

$$\mathrm{Tr}(\rho_{f,N}T^i{}_j) = \delta^i{}_j p_{f,N} \qquad i, j = \chi, \theta, \phi, \tag{4.2b}$$

are the only nonvanishing components of the renormalized expectation value in coordinates (1.2). Since the renormalization counterterms are state independent, they may be subtracted from the mode sum for the de Sitter invariant state with $A_k = 1$, $B_k = 0$ once and for all. The renormalized expectation value $\langle v|T_{ab}|v\rangle_R$ has been computed in the CTBD state [3]. Because of its de Sitter invariance this expectation value satisfies (4.2) with $p_v = -\varepsilon_v$. Collecting then the remaining finite terms which differ from this when $A_k \neq 1$, $B_k \neq 0$ in the general $O(4)$ invariant mixed state, one obtains [38]

$$\varepsilon_{f,N} = \varepsilon_v + \frac{1}{2\pi^2}\sum_{k=1}^{\infty} k^2\{(1+2N_k)\mathrm{Re}[A_k B_k^* \varepsilon_k^A] + [N_k + |B_k|^2(1+2N_k)]\varepsilon_k^B\} \tag{4.3a}$$

$$p_{f,N} = p_v + \frac{1}{2\pi^2}\sum_{k=1}^{\infty} k^2\{(1+2N_k)\mathrm{Re}[A_k B_k^* p_k^A] + [N_k + |B_k|^2(1+2N_k)]p_k^B\}, \tag{4.3b}$$

where we have defined

$$\varepsilon_k^A \equiv \dot{v}_k^2 + 2hv_k\dot{v}_k + (\omega_k^2 + h^2)v_k^2 \equiv 3p_k^A + 2m^2 v_k^2, \tag{4.4a}$$

$$\varepsilon_k^B \equiv |\dot{v}_k|^2 + 2h\mathrm{Re}[v_k^*\dot{v}_k] + (\omega_k^2 + h^2)|v_k|^2 \equiv 3p_k^B + 2m^2|v_k|^2, \tag{4.4b}$$

$$\omega_k^2 \equiv \frac{k^2}{a^2} + m^2, \qquad h \equiv \frac{\dot{a}}{a} = H\tanh u. \tag{4.4c}$$

Here an overdot denotes Hd/du and $a = H^{-1}\cosh u$. We have suppressed the γ subscript and also set $\xi = \frac{1}{6}$ (but kept $m \neq 0$) in order to simplify the expressions. This is already sufficiently general for our purposes, as the general case $\xi \neq \frac{1}{6}$ adds no essentially new features. By using the mode equation (3.7) satisfied by v_k one may readily check that the renormalized stress tensor is covariantly conserved,

$$H\frac{d\varepsilon_{f,N}}{du} + 3h(\varepsilon_{f,N} + p_{f,N}) = \frac{H}{a^3}\frac{d}{du}(a^3\varepsilon_{f,N}) + 3hp_{f,N} = 0, \tag{4.5}$$

so that it is sufficient to focus attention on the energy density for the general $O(4)$ invariant state.

Since the renormalization subtractions have already been performed in defining the finite $p_v = -\varepsilon_v$ in the $O(4,1)$ invariant state $|v\rangle$, the additional state-dependent mode sums in (4.3) must not give rise to any new UV divergences. This implies that the Bogoliubov coefficients B_k and numbers N_k must satisfy

$$\lim_{k\to\infty}[k^4|B_k|] = \lim_{k\to\infty}[k^4|B_k|^2] = \lim_{k\to\infty}[k^4 N_k] = 0, \tag{4.6}$$

so that all of the sums over k for the remaining state-dependent terms in (4.3) converge. States $|f\rangle$ whose Bogoliubov coefficients satisfy (4.6) in addition to (3.21) are UV allowed or UV finite $O(4)$ invariant states [38]. Finiteness and conservation are clearly necessary conditions for the expectation value $\mathrm{Tr}\,(\rho_{f,N}T^a{}_b)$ to be used as a source for the semiclassical Einstein equations (1.7). These properties of $\langle T^a{}_b\rangle_R$ remain valid for all UV finite states, including those of lower symmetry, provided only that the Bogoliubov coefficients fall off rapidly enough at large k, as in (4.6). The UV finiteness conditions (4.6) are also both necessary and sufficient conditions for the two-point function to have Hadamard short distance behavior.

As in the current expectation value of Sec. II we seek a qualitative understanding of the terms contributing to the energy density in (4.3a) and (4.4). There are three kinds of terms for a given k in a general $O(4)$ invariant UV finite state, namely those multiplying the factors $|B_k|^2$, $\mathrm{Re}(A_k B_k^*)$, and $\mathrm{Im}(A_k B_k^*)$, respectively. These are plotted in Figs. 6–11. In Fig. 6 the three summands in (4.3a), namely $k^2\varepsilon_k^B/(2\pi^2)$, $k^2\mathrm{Re}\varepsilon_k^A/(2\pi^2)$ and $-k^2\mathrm{Im}\varepsilon_k^A/(2\pi^2)$ are shown in units of H^4 for the case $m = H$ and $k = 10$. The ε_k^A terms multiplying the complex $A_k B_k^*$ coefficient in (4.3a) are oscillatory, while the ε_k^B function multiplying the real coefficient $N_k + |B_k|^2(1+2N_k)$ is nonoscillatory. The main difference between the coefficients of the real and the imaginary parts of $A_k B_k^*$ is that the former is symmetric about $u = 0$ while the latter is antisymmetric. The plots also show that the maxima of the two oscillatory functions occur for $|u|$ of order one, while the maximum of the third, nonoscillatory function is at the symmetric point $u = 0$ and is much larger in magnitude. In all three cases the functions fall off for large values of the time $|u|$ where the scale factor $a(u)$ is large.

Since the field is massive one might expect that at large values of the scale factor the contributions to the energy density would scale like a^{-3}. To illustrate the power dependence on the scale factor we plot in Figs. 7 and 8 the coefficients of the real and imaginary parts of $(1 + 2N_k)A_k B_k^*$ multiplied by a^3 for $k = 1, 10, 100$ and $m = H$ and $m = 10H$, respectively. We observe that the oscillations have an envelope which does scale like a^{-3} for large $|u|$ and large $a(u)$. The envelope also scales like k^2 independently of m, so that if we were to sum modes up to a large but finite K, we would expect a K^3/a^3 behavior characteristic of a nonrelativistic gas. However the rapid oscillations, particularly for larger values of m and k, highlight the fact that these are highly coherent quantum states, and the energy density is not that of quasiclassical particles in any sense.

The a^{-3} behavior of the envelope of the oscillations also does not hold for small $|u|$. As shown in detail by a WKB analysis of the mode equation (3.7) in the accompanying paper [9], the mode functions and adiabatic vacuum state change character around the times $u = \mp u_{k\gamma}$, where

$$u_{k\gamma} = \ln\left[\frac{\sqrt{k^2 - \frac{1}{4}} + \sqrt{\gamma^2 + k^2 - \frac{1}{4}}}{\gamma}\right]. \tag{4.7}$$

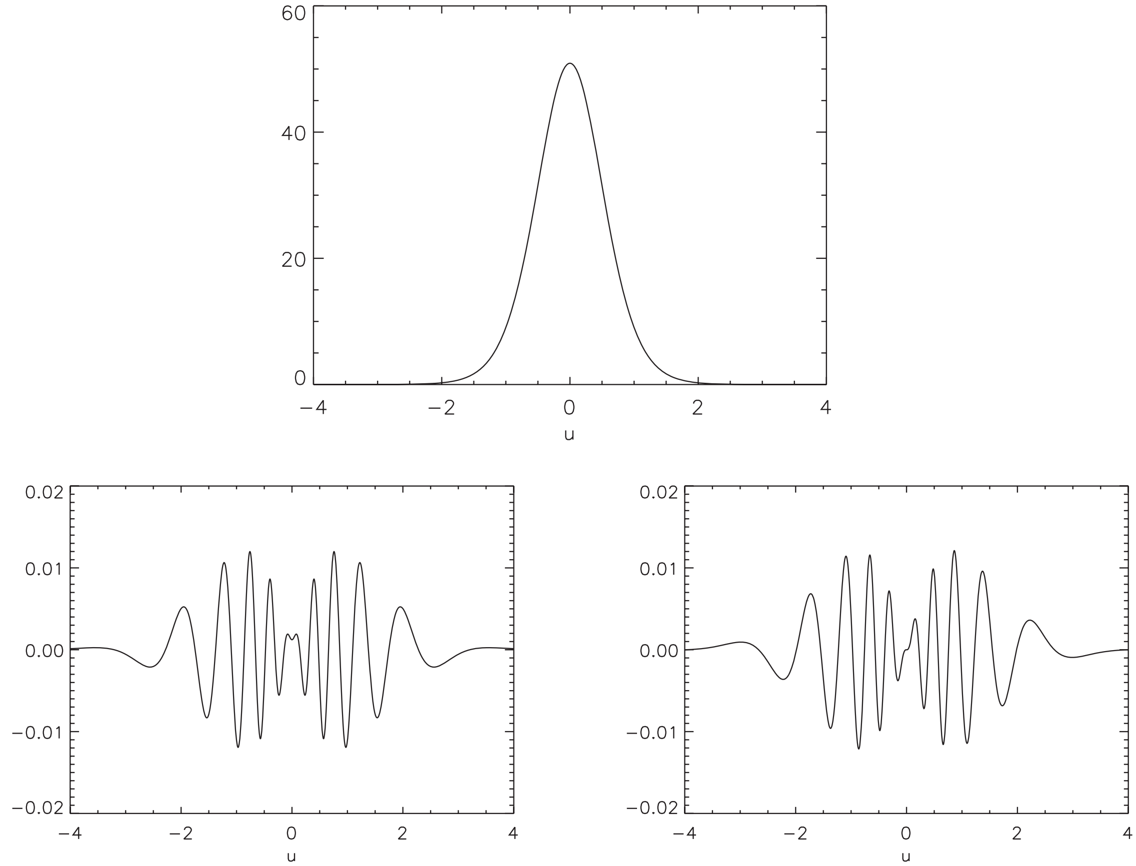

FIG. 6. The top panel shows the coefficient of the $|B_k|^2$ term in the energy density $k^2\varepsilon_k^B/(2\pi^2)$ of (4.3a), for $m = H$ and $k = 10$ in units of H^4. The bottom left panel shows the real part and the bottom right panel the imaginary part of the coefficient of the $A_k B_k^*$ term in the energy density, namely $k^2\mathrm{Re}\varepsilon_k^A/(2\pi^2)$ and $-k^2\mathrm{Im}\varepsilon_k^A/(2\pi^2)$, respectively, of (4.3a) and (4.4a) again for $m = H$ and $k = 10$, in the same units.

The modes are nonrelativistic for $|u| > u_{k\gamma}$, but relativistic for $|u| < u_{k\gamma}$. For a conformal massless field $m = 0$, with $v_k = v_{k,\frac{1}{2}}$ of (3.13), ε_k^A of (4.4a) vanishes identically. This accounts for the much smaller values of the energy densities in Figs. 7 and 8 in the central regions where $-u_{k\gamma} < u < +u_{k\gamma}$, where there is no simple behavior of the envelope of the quantum coherent oscillations. The maximum of the oscillatory terms occurs for all values of k and m investigated at $|u| \sim 1$ in the central region. This maximum saturates at a value of order one in H^4 units for large $k \gg 1$, as shown in Fig. 9.

The nonoscillatory $k^2\varepsilon_k^B/(2\pi^2)$ term in the energy density is shown in Figs. 10 and 11 for $m = H$ and $m = 10H$, respectively, for $k = 1$, 10, and 100. In the left panels the term is multiplied by $a^3(u)$ and in the right panels the term is multiplied by $a^4(u)$. It is clear that in all cases the contribution from this ε_k^B term is proportional to a^{-3} at large values of the scale factor and is proportional to a^{-4} near $u = 0$. In other words, it blueshifts in the contracting phase of de Sitter space (and redshifts in the expanding phase) as a nonrelativistic fluid for large $|u|$ but as a relativistic fluid for smaller $|u|$. At $|u| = u_{k\gamma}$ given by (4.7), the energy density transitions from nonrelativistic to relativistic behavior and for smaller $|u|$ the physical momentum k/a dominates the mass term in the mode equation (3.25). In the nonrelativistic region $u > |u_{k\gamma}|$ the k dependence is k^2. However the maximum of the $k^2\varepsilon_k^B$ term always occurs in the relativistic region $-u_{k\gamma} < u < +u_{k\gamma}$, where the k dependence is k^3, so that this maximum value grows unbounded for $k \gg 1$, in contrast to the oscillatory terms which are bounded for large k (Fig. 9).

For the strictly massless conformally invariant scalar field, there are no oscillatory ε_k^A terms since $\varepsilon_k^A = 0$ identically for $v_k = v_{k,\frac{1}{2}}$, and hence there are no terms linear in B_k in the energy density or pressure of a general $O(4)$ invariant UV allowed state. The only contributions come instead from the ε_k^B terms quadratic in the perturbation B_k from the de Sitter invariant CTBD state $|v\rangle$. Substituting $v_{k,\frac{1}{2}}$ into (4.4b) with $m = 0$ gives

$$\frac{k^2}{2\pi^2}\,\varepsilon_k^B|_{m=0} = \frac{k^3}{2\pi^2 a^4}, \tag{4.8}$$

showing that the relativistic behavior observed in Figs. 10–11 in the relativistic region $-u_{k\gamma} < u < u_{k\gamma}$ holds

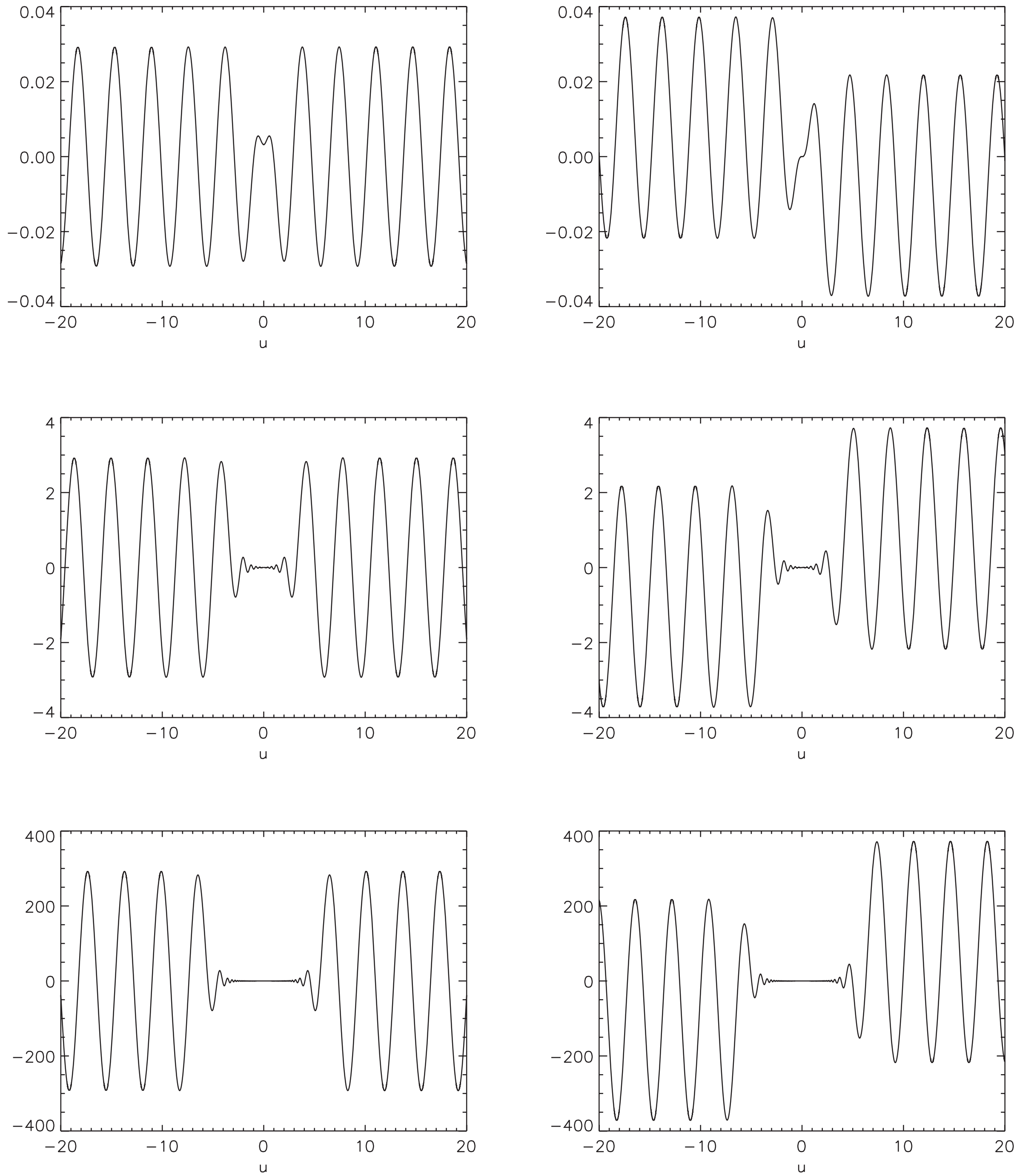

FIG. 7. The panels at left show, for $m = H$, the summand $k^2 \mathrm{Re}\varepsilon_k^A/(2\pi^2)$ in the energy density (4.3a) and (4.4b) in units of H^4 multiplied by a factor of a^3, with $a = H^{-1} \cosh u$ the scale factor. The panels at right show $-a^3 k^2 \mathrm{Im}\varepsilon_k^A/(2\pi^2)$. From top to bottom the values of k are $k = 1$, $k = 10$, and $k = 100$. The values of $u_{k\gamma}$ from (4.7) are 0.88, 3.14, and 5.44, respectively. The plots show that the envelope of the oscillations is proportional to a^{-3} for $|u| > u_{k\gamma}$, but that the behavior changes markedly for $|u| < u_{k\gamma}$.

for all u in the massless case. This result can also be obtained by conformally transforming from flat space to de Sitter space the exact stress tensor for a conformal field in a state other than the Minkowski vacuum [4]. Although this is exactly the behavior one would expect for a gas of relativistic particles, we emphasize that these are still coherent quantum excitations of the pure $|f\rangle$ vacuum state, in which the exact phase relations of (3.28)–(3.29) are maintained.

In all cases the perturbations from the CTBD state $|v\rangle$ fall off in the expanding half $u > 0$ of de Sitter space but grow in the contracting half $u < 0$. The maximum value at the symmetric point from (4.3), (4.4), and (4.8) is given by

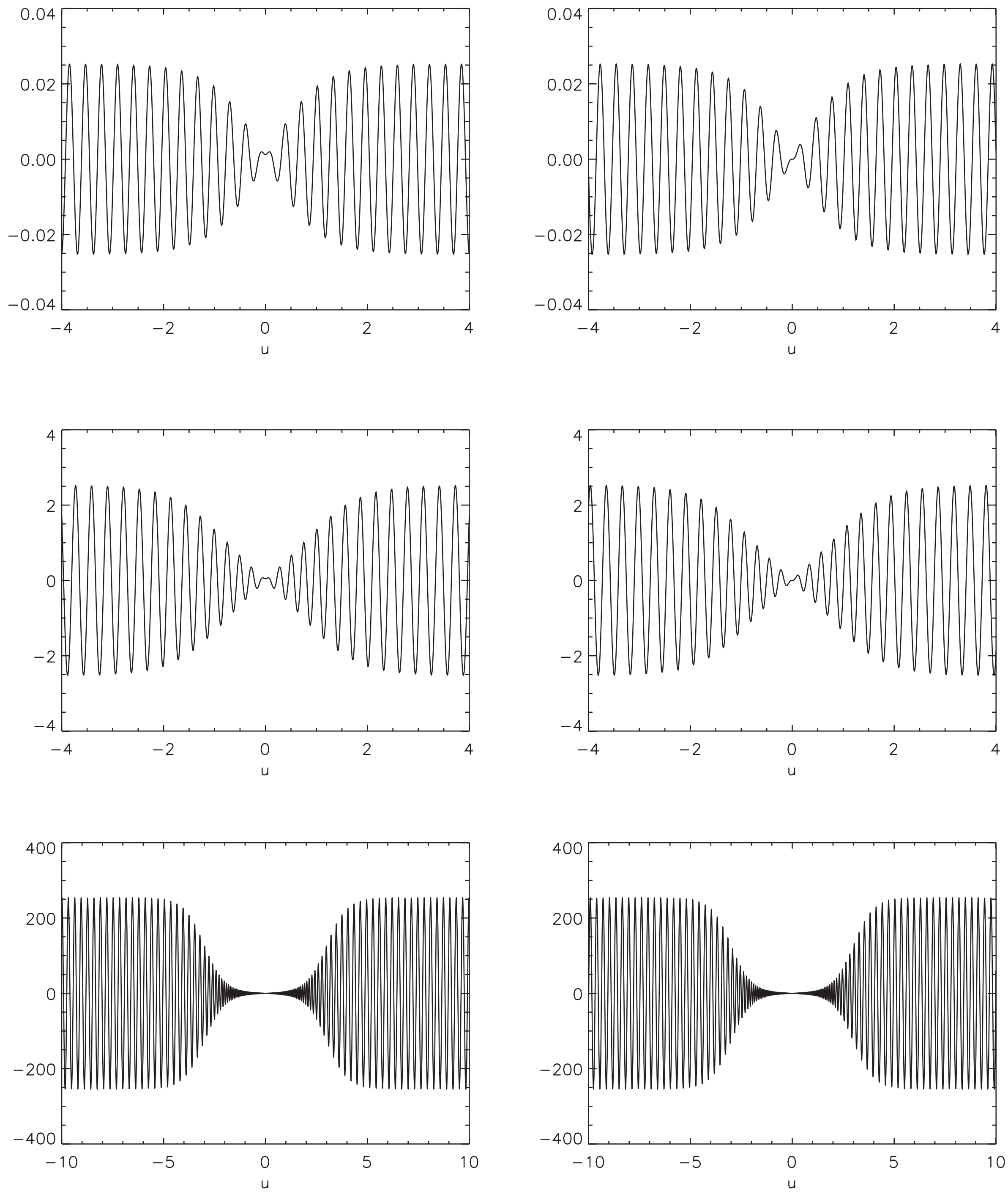

FIG. 8. The panels at left show, for $m = 10H$, the summand $k^2 \mathrm{Re}\ \varepsilon_k^A/(2\pi^2)$ in the energy density (4.3a) and (4.4a) in units of H^4 multiplied by a factor of a^3, with $a = H^{-1} \cosh u$ the scale factor. The panels at right show $-a^3 k^2 \mathrm{Im}\ \varepsilon_k^A/(2\pi^2)$. From top to bottom the values of k are $k = 1$, $k = 10$, and $k = 100$, with the corresponding values of $u_{k\gamma}$ from (4.7) 0.087, 0.88, and 3.00, respectively, where the behavior changes markedly. It is also observed that the envelopes for $|u| > u_{k\gamma}$ scale like k^2.

$$\varepsilon_{f,N\,\mathrm{max}} = \frac{1}{2\pi^2 a^4} \sum_{k=1}^{\infty} k^3 [N_k + |B_k|^2 (1 + 2N_k)] \simeq \frac{H^4}{8\pi^2} K^4 [N_K + |B_K|^2 (1 + 2N_K)], \tag{4.9}$$

where we have approximated the sum by an integral valid for large $k_{\mathrm{max}} = K$, the maximum value of k for which $|B_k|^2$ and/or N_k has support consistent with the UV finiteness conditions (4.6). This summarizes the results plotted in Figs. 6–11 as the estimate of the largest contribution to the

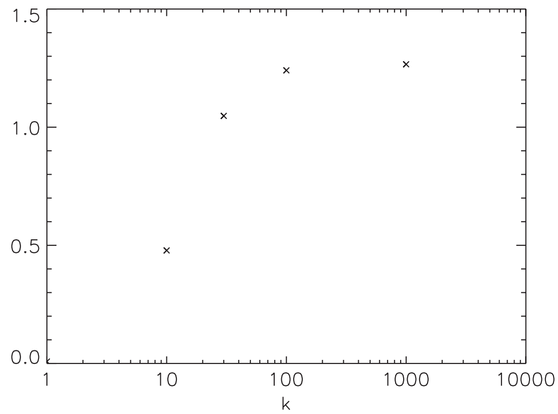

FIG. 9. The maxima of the oscillations of $k^2\mathrm{Re}\,\varepsilon_k^A/(2\pi^2)$ in units of H^4 are plotted for several values of k, for $m = 10H$. The saturation at large values of k is apparent.

energy density (4.3a). For this to produce a significant backreaction on the classical de Sitter geometry through the semiclassical Einstein equations, it is necessary for this to be larger than the background cosmological energy density, i.e.,

$$8\pi G\varepsilon_{f,N} \gtrsim \Lambda = 3H^2 \quad \text{or}$$
$$\frac{GH^2}{3\pi}[N_K + |B_K|^2(1+2N_K)]\left(\frac{K}{\cosh u}\right)^4 \gtrsim 1. \qquad (4.10)$$

Clearly no matter how small $GH^2 \ll 1$, or the state-dependent perturbation in square brackets is, as long as their product is nonzero there is always a large enough (but still finite K) for which the inequality (4.10) is satisfied at the maximum at $u = 0$. Since all finite k modes are redshifted in physical momentum as $k/a \to 0$ for $a(u) \to \infty$, perturbations satisfying (4.10) at $u = 0$ have vanishingly small energy densities at early times $u \to -\infty$. Hence for any finite $GH^2 > 0$ there is a large class of $O(4)$ invariant but de Sitter noninvariant states satisfying (4.10), which give rise to energy densities that are large enough to produce significant de Sitter noninvariant backreaction effects at the symmetric point $u = 0$, all of which have exponentially vanishing de Sitter noninvariant energy densities at times in the infinite past at I_- of ‘eternal’ de Sitter space. The physical momentum HK/a corresponding to the condition (4.10) for significant backreaction is of order $\sqrt{HM_{\rm Pl}} \ll M_{\rm Pl}$, far less than the Planck scale $M_{\rm Pl}$, so that the semiclassical approximation is still reliable.

To see how the general condition (4.10) for large backreaction and de Sitter instability is realized in a specific physical state, which is the one determined by adiabatically switching on of the background analogous to switching on of the electric field in the infinite past [39], one can choose Bogoliubov coefficients corresponding to the $O(4)$ invariant $|in\rangle$ state of [9] prepared at the initial time u_0. In the contracting phase of de Sitter space $u < 0$ this corresponds to choosing the mode functions according to the initial data at $u = u_0$,

$$A_k = A^{\rm in}_{k\gamma}\theta(-u_{k\gamma} - u_0) + \theta(u_0 + u_{k\gamma}) \qquad (4.11a)$$

$$B_k = B^{\rm in}_{k\gamma}\theta(-u_{k\gamma} - u_0), \qquad (4.11b)$$

with $u_{k\gamma}$ defined in (4.7). The step functions are again simple approximations to the actual smooth but rapid change at $-u_{k\gamma}$. Since $B_k = 0$ for $u_{k\gamma} > -u_0$ the mode sums in (4.3a) are cut off for $k > K_\gamma(u_0)$, where

$$K_\gamma(u_0) = \sqrt{\gamma^2\sinh^2 u_0 + \frac{1}{4}} \simeq \frac{\gamma}{2}e^{|u_0|} \qquad (4.12)$$

for $|u_0| \gg 1$. Also,

$$A^{\rm in}_{k\gamma} = \frac{i}{\sqrt{2\sinh(\pi\gamma)}}e^{-\frac{ik\pi}{2}}e^{\frac{\pi\gamma}{2}}, \qquad B^{\rm in}_{k\gamma} = \frac{1}{\sqrt{2\sinh(\pi\gamma)}}e^{\frac{ik\pi}{2}}e^{-\frac{\pi\gamma}{2}} \qquad (4.13)$$

so that

$$A^{\rm in}_{k\gamma}B^{\rm in*}_{k\gamma} = \frac{i(-)^k}{2\sinh(\pi\gamma)}\theta(-u_{k\gamma} - u_0) \qquad (4.14)$$

oscillates in k. Because the ε_k^A term is bounded in k for any state as shown in Fig. 11 and its coefficient (4.14) oscillates in k for this $|in\rangle$ state, its contributions tend to cancel when summed over k in (4.3a), and are negligible compared to the nonoscillatory ε_k^B term in the energy density for large $K_\gamma(u_0)$, hence large $|u_0|$. Retaining only the latter, we then have approximately

$$\begin{aligned}\varepsilon_{\rm in} &\simeq \varepsilon_v + \frac{1}{2\pi^2}\sum_{k=1}^{K_\gamma(u_0)} k^2|B^{\rm in}_{k\gamma}|^2\varepsilon_k^B \\ &\simeq \frac{1}{4\pi^2}\frac{e^{-\pi\gamma}}{\sinh(\pi\gamma)}\int_1^{K_\gamma(u_0)} dk\,\frac{k^3}{a^4(u)} \\ &\simeq \frac{H^4}{8\pi^2}\frac{1}{e^{2\pi\gamma}-1}\left(\frac{K_\gamma(u_0)}{\cosh u}\right)^4,\end{aligned} \qquad (4.15)$$

which agrees with Eq. (8.17) of [9] and (4.9) for the particular choice of $|B_K|^2$ from (4.13) and $N_k = 0$. This energy density, which has an arbitrarily small value for $u \to -\infty$, is blueshifted as a relativistic fluid and by $u = 0$ can grow large enough to be comparable to or even far exceed the background de Sitter energy density and hence significantly affect the background de Sitter geometry. It satisfies the inequality (4.10) for significant backreaction if

$$K_\gamma(u_0) \gtrsim \left[\frac{3\pi}{GH^2}(e^{2\pi\gamma} - 1)\right]^{\frac{1}{4}} \qquad (4.16)$$

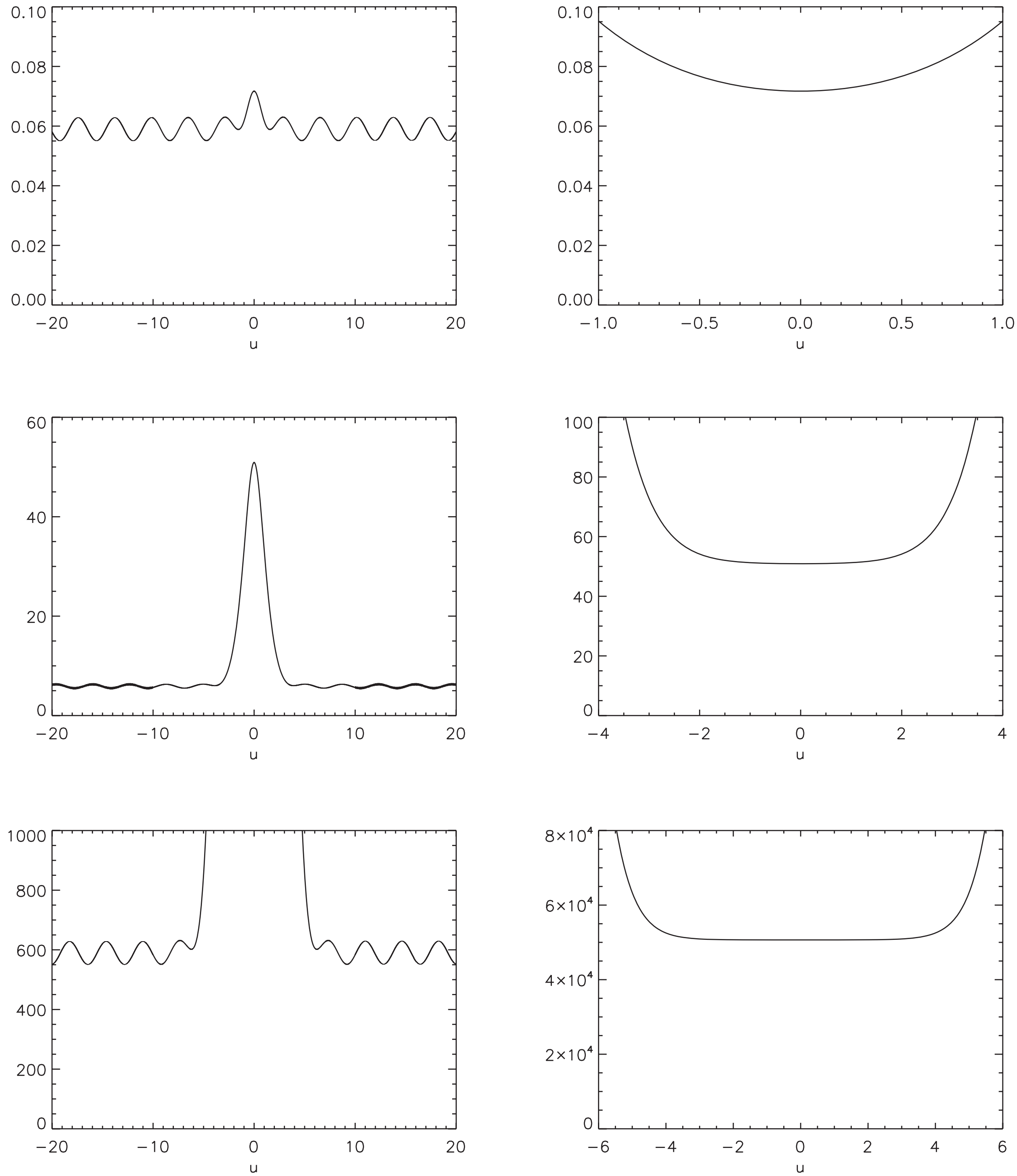

FIG. 10. The panels on the left show for $m = H$, the $k^2 \varepsilon_k^B/(2\pi^2)$ term in the energy density of (4.3a) and (4.4b) in units of H^4 multiplied by a factor of a^3. The panels on the right show this same term multiplied by a factor of a^4. From top to bottom the value of k which corresponds to each set of plots is $k = 1$, $k = 10$, and $k = 100$, with values of $u_{k\gamma}$ of 0.88, 3.14, and 5.44, respectively, where the behavior changes from nonrelativistic to relativistic.

or from (4.12),

$$|u_0| > \ln\left[\frac{2}{\gamma}\left(\frac{3\pi}{GH^2}\right)^{\frac{1}{4}}(e^{2\pi\gamma} - 1)^{\frac{1}{4}}\right], \tag{4.17}$$

which can always be satisfied for early enough u_0, and nonzero GH^2 and γ in eternal de Sitter space. Even for $GH^2 \simeq 10^{-122}$, $\gamma \simeq 1$, corresponding to the present value of the cosmological constant inferred from the SNIa data [40], this value of $|u_0|$ is quite moderate, of order 70.

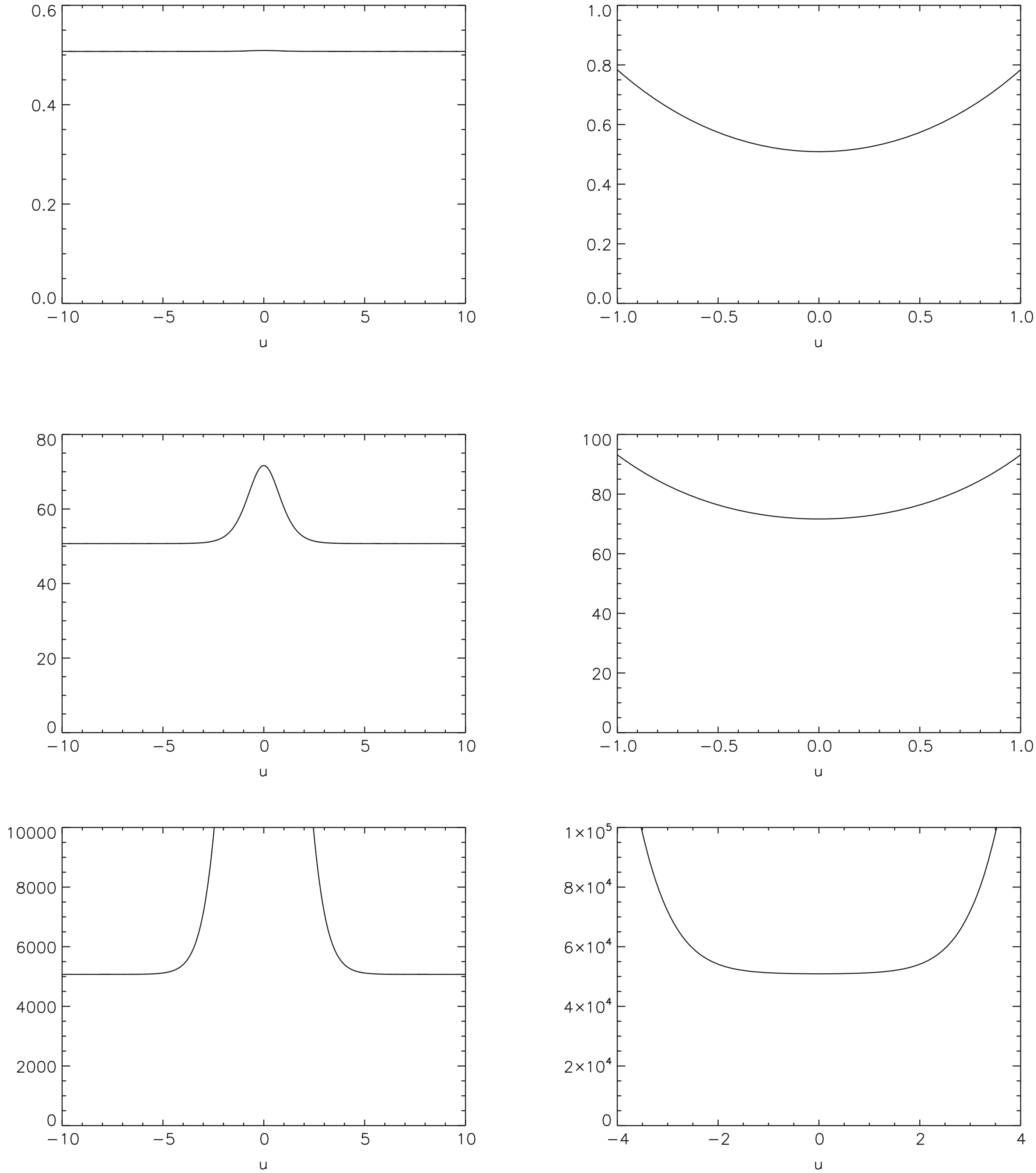

FIG. 11. The panels on the left show for $m = 10H$, the $k^2\varepsilon_k^B/(2\pi^2)$ term in the energy density of (4.3a) and (4.4b) in units of H^4 multiplied by a factor of a^3. The panels on the right show this same term multiplied by a factor of a^4. From top to bottom the value of k which corresponds to each set of plots is $k = 1$, $k = 10$, and $k = 100$. The values of $u_{k\gamma}$ where the behavior changes from nonrelativistic to relativistic are 0.087, 0.88, and 3.00, respectively, for this m.

V. CONFORMAL ANOMALY, STRESS TENSOR AND DE SITTER SYMMETRY BREAKING

As in the electric field example, the quantum vacuum instability in de Sitter space is illuminated by consideration of a quantum anomaly, in this case the conformal trace anomaly of the energy-momentum tensor [4,41]. The nonlocal covariant effective action that gives the conformal anomaly in four dimensions is [12,16–18,42]

$$S_{\rm anom}[g] = \frac{1}{2}\int d^4x\sqrt{-g}\int d^4x'\sqrt{-g'}\left(\frac{E}{2}-\frac{\Box R}{3}\right)_x \times \Delta_4^{-1}(x,x')\left[bF + b'\left(\frac{E}{2}-\frac{\Box R}{3}\right)\right]_{x'}. \qquad (5.1)$$

This nonlocal effective action (5.1) is the analog of (2.48) for the chiral anomaly in two dimensions and the $\int d^2x\sqrt{-g}\int d^2x'\sqrt{-g'}R_x\Box^{-1}(x,x')R'_x$ effective action for the conformal trace anomaly in two dimensions [43]. In four dimensions there are two invariants $E \equiv {}^*R_{abcd}{}^*R^{abcd} = R_{abcd}R^{abcd} - 4R_{ab}R^{ab} + R^2$ and $F \equiv C_{abcd}C^{abcd} = R_{abcd}R^{abcd} - 2R_{ab}R^{ab} + \frac{1}{3}R^2$ contributing to the nonlocal anomaly with corresponding dimensionless coefficients b and b' proportional to $\hbar$ in the notation of [41]. Being nonlocal in terms of the curvature invariants E and F, the one-loop effective action (5.1) contains information about nonlocal and global quantum effects, i.e. sensitivity to initial and/or boundary conditions, through the Green's function inverse $\Delta_4^{-1}(x,x')$ of the conformally covariant differential operator

$$\Delta_4 \equiv \Box^2 + 2R^{ab}\nabla_a\nabla_b - \frac{2}{3}R\Box + \frac{1}{3}(\nabla^a R)\nabla_a = \nabla_a\left(\nabla^a\nabla^b + 2R^{ab} - \frac{2}{3}Rg^{ab}\right)\nabla_b. \qquad (5.2)$$

To (5.1) it is possible to add any conformally invariant action (nonlocal or local) which does not affect the anomaly. However, only the conformal breaking (5.1) term in the effective action needs be retained in a low energy classification of operators in the effective action [16] and only this term can have relevant infrared effects.

Moreover, the effective action of the anomaly (5.1) is distinguished by being responsible for additional massless scalar degree(s) of freedom in low energy gravity, not present in the classical theory [44], as seen also in two dimensions by the shifting of the central charge from $N-26$ to $N-25$ [45]. In four dimensions this is made explicit by rewriting (5.1) in the local form

$$S_{\rm anom} = b'S^{(E)}_{\rm anom} + bS^{(F)}_{\rm anom}, \qquad (5.3)$$

by the introduction of at least one additional scalar field, where for example

$$S^{(E)}_{\rm anom}[g;\varphi] \equiv \frac{1}{2}\int d^4x\sqrt{-g}\left\{-(\Box\varphi)^2 + 2\left(R^{ab} - \frac{1}{3}Rg^{ab}\right) \times (\nabla_a\varphi)(\nabla_b\varphi) + \left(E - \frac{2}{3}\Box R\right)\varphi\right\} \qquad (5.4)$$

is the term related to E in terms of the additional scalar field φ. This scalar (analogous to χ of Sec. II C) is a new effective degree of freedom, not to be confused with the original scalar field Φ, which describes two-particle correlations or bilinears (relativistic Cooper pairs) of the underlying scalar, fermion or vector QFT [18,44]. QFTs of different spin may all be studied via the effective action (5.4), since the only dependence upon spin for free fields is through the trace anomaly coefficients b' and b, where

$$b' = -\frac{1}{360(4\pi)^2}(N_S + 11N_F + 62N_V) \qquad (5.5)$$

is the coefficient for the E term in the conformal anomaly for noninteracting scalar (S), fermion (F), or vector (V) fields, respectively. The b term in the anomaly proportional to F gives rise to an effective action $S^{(F)}_{\rm anom}$ similar to (5.4) but is less important in de Sitter space where $F = C_{abcd}C^{abcd} = 0$ [15–18].

Formally solving the Euler-Lagrange equation for φ that results from varying (5.4) in a general metric background requires inverting the differential operator Δ_4, i.e. finding its Green's function $\Delta_4^{-1}(x,x')$, and substituting the solution for φ into (5.4). This returns the b' term of the nonlocal form (5.1), up to surface terms. In de Sitter space we also have $E = 24H^4$, $\Box R = 0$, and the operator Δ_4 factorizes, so that the variation of (5.4) with respect to φ yields the linear equation of motion,

$$\Delta_4\varphi|_{dS} = -\Box(-\Box + 2H^2)\varphi = \frac{E}{2} - \frac{\Box R}{3} = 12H^4, \qquad (5.6)$$

with a constant source. Because of this constant source, analogous to (2.45), and the fact that the only invariant scalar in de Sitter space is a constant, it is clear that no de Sitter invariant constant solution to (5.6) for φ exists. In the local form of the effective action (5.4), the freedom to add homogeneous solutions to (5.6) is equivalent to that of specifying the particular Green's function inverse $\Delta_4^{-1}(x,x')$ dependent upon initial/boundary conditions in the nonlocal form (5.1). It is also clear from the factorized form (5.6) of Δ_4 in de Sitter space that its inverse,

$$\Delta_4^{-1}|_{dS} = \frac{1}{2H^2}[(-\Box)^{-1} - (-\Box + 2H^2)^{-1}], \qquad (5.7)$$

cannot be de Sitter invariant, since it is proportional to the difference of the inverses of a massless, minimally coupled ($\xi = 0$) scalar and a massless, conformally coupled ($\xi = \frac{1}{6}$) scalar, and no de Sitter invariant form of the former exists [36]. Thus the breaking of de Sitter invariance and infrared sensitivity to initial/boundary conditions is already apparent from either the nonlocal one-loop effective action (5.1) and nonexistence of a de Sitter invariant Feynman Green's function (5.7), or equivalently from the noninvariance of the solutions to (5.6) and hence those of the local effective action (5.4).

The form of the breaking of de Sitter invariance may be studied through the stress tensor corresponding to the local effective action (5.4), whose variation with respect to the metric gives the tensor

$$E_{ab}[\varphi]|_{dS} \equiv -\frac{2}{\sqrt{-g}}\frac{\delta S^{(E)}_{\rm anom}}{\delta g^{ab}}\bigg|_{dS} = -2(\nabla_{(a}\varphi)(\nabla_{b)}\Box\varphi) + 2(\nabla_c\varphi)(\nabla^c\nabla_a\nabla_b\varphi) + 2(\Box\varphi)(\nabla_a\nabla_b\varphi) - \frac{2}{3}\nabla_a\nabla_b[(\nabla\varphi)^2] - 4H^2(\nabla_a\varphi)(\nabla_b\varphi) + \frac{1}{2}g_{ab}\left[-(\Box\varphi)^2 + \frac{1}{3}\Box[(\nabla\varphi)^2] + 2H^2(\nabla\varphi)^2\right] - \frac{2}{3}\nabla_a\nabla_b\Box\varphi + 4H^2\nabla_a\nabla_b\varphi - \frac{2}{3}H^2 g_{ab}\Box\varphi + 8H^4 g_{ab}, \tag{5.8}$$

which is covariantly conserved by the use of (5.6). In (5.8) we have evaluated E_{ab} in de Sitter space and used the notation $(\nabla\varphi)^2 \equiv g^{ab}(\nabla_a\varphi)(\nabla_b\varphi)$. The stress tensor $T^{(E)}_{ab} = b' E_{ab}$ evaluated on solutions φ satisfying the classical linear equation (5.6) may be used to evaluate the renormalized expectation value $\langle T_{ab}\rangle_R$ of the underlying QFT. This is exact up to state-dependent (but curvature-independent) terms if the spacetime is conformally flat as is de Sitter space, and the QFT is classically conformally invariant [46]. We show below in particular that (5.8) reproduces the CTBD state value exactly for classically conformally invariant fields of any spin with an appropriate choice of φ.

Since the fourth order linear operator Δ_4 in (5.6) factorizes into two second order wave operators for a conformally coupled and minimally coupled massless scalar in de Sitter space, the general homogeneous solution of (5.6) in coordinates (1.2) is easily found in terms of $v_{k,\frac{1}{2}}Y_{klm_l}$ and $v_{k,\frac{3}{2}}Y_{klm_l}$ and their complex conjugates. Inspection of these solutions, (3.13)–(3.15) shows that the functions $v_{k,\frac{1}{2}}$ and $v_{k,\frac{3}{2}}$ may also be written as a linear combination of $\exp[-i(k\pm 1)\eta]$. The conformal covariance property of the operator Δ_4 is what makes this rearrangement of the solutions possible. In de Sitter spacetime (and in fact any conformally flat spacetime) there is a second factorization of Δ_4 into two second order operators, reflecting the fact that for a fixed k (5.6) may also be written [47]

$$\Delta_4|_{dS}\varphi_k(\eta)Y_{klm_l}(\hat{N}) = H^4\cos^4\eta\left[\frac{d^2}{d\eta^2} + (k-1)^2\right] \times \left[\frac{d^2}{d\eta^2} + (k+1)^2\right]\varphi_k(\eta)Y_{klm_l}(\hat{N}) \tag{5.9}$$

in conformal time η, where $H^4\cos^4\eta = a^{-4}$. Thus the homogeneous solutions of (5.6) are clearly linear combinations of $\exp[-i(k\pm 1)\eta]Y_{klm_l}(\hat{N})$ and their complex conjugates. To these one must add a particular solution of the inhomogeneous equation (5.6), which is easily found in coordinates (1.2) to be

$$\varphi_0 \equiv 2\ln(\cosh u). \tag{5.10}$$

This particular solution is $O(4)$ invariant but not $O(4,1)$ invariant. Other choices correspond to states of lower symmetry, but some choice must be made since the inhomogeneous term in (5.6) disallows the de Sitter invariant choice of constant φ. Then we may express the general solution of (5.6) in the form

$$\varphi = \varphi_1(u) + \frac{1}{2}\sum_{k=2}^{\infty}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}\left[\frac{a_{klm_l}}{\sqrt{2k(k-1)}}e^{-i(k-1)\eta}Y_{klm_l} + \frac{b_{klm_l}}{\sqrt{2k(k+1)}}e^{-i(k+1)\eta}Y_{klm_l} + \text{c.c.}\right], \tag{5.11}$$

where $\varphi_1(u)$ is the general solution of (5.6) for $k=1$, constant on $\mathbb{S}^3$, given by

$$\varphi_1(u) = \varphi_0(u) + c_0 + c_1\sin^{-1}(\tanh u) + c_2\text{sech}^2 u + c_3\tanh u\,\text{sech} u = 2\ln(\sec\eta) + c_0 + c_1\eta + c_2\cos^2\eta + c_3\sin\eta\cos\eta, \tag{5.12}$$

with the c_i arbitrary constants multiplying the 4 homogeneous solutions which are functions only of u or conformal time η defined in (3.15). The normalizations of the $k>1$ solutions in (5.11) are chosen to correspond to a previous canonical analysis on the conformally related Einstein static cylinder $R\otimes S^3$ where the $\varphi = 2\sigma$ field was quantized and the (a_{klm_l}, b_{klm_l}) obey canonical commutation relations (the $b_{klm_l}, b^\dagger_{klm_l}$ with negative metric) [47]. Here we treat all the expansion coefficients $(c_i, a_{klm_l}, b_{klm_l})$ of the general solution (5.11) to (5.6) for the effective action in de Sitter space as c numbers.

For $O(4)$ invariant states the stress tensor can only be a function of u. Because of the terms linear in φ in (5.8) this corresponds to choosing all the coefficients $a_{klm_l} = b_{klm_l} = 0$ in (5.11) for $k>1$. With $\varphi = \varphi_1(u)$ substituted into (5.8) we obtain the energy density

$$-E^u{}_u[\varphi_1(u)] = \dddot{\varphi}_1\dot{\varphi}_1 - \frac{1}{2}\ddot{\varphi}_1^2 + 2h\ddot{\varphi}_1\dot{\varphi}_1 + 3\left(2\dot{h} + \frac{3}{2}h^2 - H^2\right)\dot{\varphi}_1^2 - 2h\dddot{\varphi}_1 + 2(H^2 - 3h^2)\ddot{\varphi}_1 - 6h(\dot{h} + H^2)\dot{\varphi}_1, \tag{5.13}$$

where a dot denotes the derivative $H^{-1}d/du$. Substituting (5.12) into this expression gives

$$\varepsilon = -b' E^u{}_u[\varphi_1(u)] = -6b'H^4 + \frac{2b'}{a^4}(c_1^2 - c_2^2 - c_3^2 + 4). \tag{5.14}$$

The first term gives the constant value of the renormalized $\varepsilon_v = -6b'H^4$ for the de Sitter invariant state of a free conformal field of any spin, with the corresponding pressure $p_v = -\varepsilon_v = 6b'H^4$. The second a^{-4} term shows that exactly the term corresponding to the relativistic limit, obtained in Sec. IV from detailed analysis of the renormalized expectation value of the stress tensor of a quantum field in the general $O(4)$ invariant state, is reproduced by the anomaly stress tensor (5.8) with a classical effective field $\varphi = \varphi_1(u)$. The spatial components,

$$E_{ij}[\varphi_1(u)] = 6H^4 g_{ij} + \frac{2}{3a^4} g_{ij}(c_1^2 - c_2^2 - c_3^2 + 4), \quad (5.15)$$

and equation of state $p = \varepsilon/3$ for the second term are just that required by covariant conservation (4.5) for this general $O(4)$ invariant state.

In (5.14) the arbitrary coefficients c_i of the homogenous solution in (5.12) appear and may be related to the sum over the state-dependent coefficients N_k, B_k and K of (4.9). The de Sitter invariant expectation value for $\langle T_{ab}\rangle_R$ is recovered if

$$c_1^2 + 4 = c_2^2 + c_3^2 \qquad (\text{de Sitter invariant}\langle T_{ab}\rangle), \quad (5.16)$$

so that no relativistic radiation a^{-4} term is present. Any solution of (5.6) of the form (5.12) with the condition (5.16) on the coefficients c_i may be taken as corresponding to the CTBD state and gives a de Sitter invariant stress tensor with $\varepsilon_v = -p_v$. It is interesting to note in passing that for the particular values $c_1 = -2i$ and $c_0 = c_2 = c_3 = 0$, $\exp[\varphi_1(u)]$ is just the (complex) conformal transformation that maps flat space and its Minkowski vacuum to de Sitter space and the CTBD invariant state $|v\rangle$. However, if we restrict $\varphi_1(u)$ of (5.12) to be real, and invariant under time reversal $u \leftrightarrow -u$, corresponding to the discrete inversion symmetry of the CTBD state, then $c_1 = c_3 = 0$, and from (5.16) $c_2 = \pm 2$ so that

$$\begin{aligned}\bar{\varphi}(u) &= 2\ln(\cosh u) + c_0 \pm 2\mathrm{sech}^2 u \\ &= -2\ln(\cos\eta) + c_0 \pm 2\cos^2\eta \end{aligned} \quad (5.17)$$

is the background solution to (5.6) with $\varepsilon_v = -6b'H^4 = -p_v$ most closely corresponding to $|v\rangle$. Since the stress tensor (5.8) depends only upon derivatives of φ, the constant c_0 is irrelevant and may be set to zero, so that the choice of solution (5.17) is determined up to the sign of the last term.

This $\bar{\varphi}(u)$ in (5.17) is a kind of mean value condensate of the φ effective field in de Sitter. Although itself not de Sitter invariant, it gives a stress tensor corresponding to the de Sitter invariant CTBD state of the underlying QFT. It seems that one has to consider more complicated expectation values such as $\langle T_{ab}T_{cd}\rangle$ in order to see directly the de Sitter breaking effects of the inhomogeneous solution to (5.6). This is similar to the de Sitter invariant stress tensor $\langle T_{ab}\rangle$ obtained for a massless, minimally coupled field in de Sitter space despite the non–de Sitter invariant vacuum state [36].

A small variation of c_2 away from ± 2 produces a de Sitter noninvariant stress tensor of the form (5.14)–(5.15) which is infinitesimally small at asymptotic past infinity I_- because of its a^{-4} dependence upon the scale factor, but which grows to finite values at the symmetric time $u = 0$. The c_i satisfying (5.16) are clearly a subset of a wider class of a three parameter family corresponding to $O(4)$ invariant but non-$O(4,1)$ invariant states. In this parameterization the condition (4.10) that the perturbations of the CTBD state produce a large enough backreaction at $u = 0$ to affect the classical geometry is

$$16\pi GH^2 |b'(c_1^2 - c_2^2 - c_3^2 + 4)| \gtrsim 1. \quad (5.18)$$

Clearly there is a large class of such states all of which have exponentially vanishing de Sitter noninvariant energy densities at times $u \to -\infty$ in the infinite past. Since a perturbation of the CTBD state with infinitesimally small energy density at I_- with coefficients c_i satisfying (5.18) produces a large backreaction on the geometry at $u = 0$, we conclude that the de Sitter invariant $|v\rangle$ state is unstable to such state perturbations in the initial data of eternal de Sitter space.

Thus the anomaly effective action and stress tensor gives the same result of instability of the CTBD state to perturbations, obtained previously for massive scalar fields, without any need of renormalization subtractions or mode sums, although the connection to the large K cutoff in (4.9) or (4.10), or to particle creation in the $|in\rangle$ state of (4.15) or [9] is no longer transparent in (5.18). The anomaly derivation of the instability condition (5.18) emphasizes its generality, independent of the particular case of a noninteracting scalar field, so that (5.18) holds for fields of any spin simply by changing b' according to (5.5), or more generally for interacting QFTs as well with the appropriate b'. This result and the composite effective field φ is similar to the generality of the axial anomaly derivation of the linear growth of the current in a persistent electric field background in terms of the bosonized effective field χ in (2.47).

VI. STATES OF LOWER SYMMETRY: SPATIALLY INHOMOGENEOUS STRESS TENSOR

The expansion (5.11) of the anomaly scalar φ also enables a general study of states of lower than $O(4)$ symmetry simply by allowing any of the parameters a_{klm_l} or b_{klm_l} in the general solution (5.11) to be different from zero. Substituting that general solution for φ in (5.8) gives a T_{ab} which is a function of directions $\hat{N}$ on $\mathbb{S}^3$ as well as u. In order to study the effect of these $O(4)$ breaking terms, we linearize the anomaly stress tensor around the solution $\bar{\varphi}(u)$ of (5.17) with a de Sitter invariant stress tensor by

$$\varphi = \bar{\varphi}(u) + \phi(u, \hat{N}) \quad (6.1)$$

for ϕ a general solution of (5.6) with $\Delta_4\phi = 0$, expressed as the sum of modes (5.11). To first order in ϕ,

$$E_{ab}^{(1)} = -2(\nabla_{(a}\bar{\varphi})(\nabla_{b)}\Box\phi) - 2(\nabla_{(a}\Box\bar{\varphi})(\nabla_{b)}\phi) + 2(\nabla_c\bar{\varphi})(\nabla^c\nabla_a\nabla_b\phi) + 2(\nabla^c\nabla_a\nabla_b\bar{\varphi})(\nabla_c\phi)$$
$$+ 2(\Box\bar{\varphi})(\nabla_a\nabla_b\phi) + 2(\nabla_a\nabla_b\bar{\varphi})(\Box\phi) - \frac{4}{3}\nabla_a\nabla_b[g^{cd}(\nabla_c\bar{\varphi})(\nabla_d\phi)] - 8H^2(\nabla_{(a}\bar{\varphi})(\nabla_{b)}\phi)$$
$$+ g_{ab}\left\{-(\Box\bar{\varphi})(\Box\phi) + \frac{1}{3}\Box[g^{cd}(\nabla_c\bar{\varphi})(\nabla_d\phi)] + 2H^2g^{cd}(\nabla_c\bar{\varphi})(\nabla_d\phi)\right\}$$
$$-\frac{2}{3}\nabla_a\nabla_b\Box\phi + 4H^2\nabla_a\nabla_b\phi - \frac{2}{3}H^2g_{ab}\Box\phi, \tag{6.2}$$

which is both covariantly conserved and traceless. Using the identities

$$\nabla_\tau\nabla_\tau\varphi = \ddot{\varphi} \tag{6.3a}$$

$$\nabla_\tau\nabla_i\varphi = \nabla_i(\dot{\varphi} - h\varphi) \tag{6.3b}$$

$$\nabla_\tau\nabla_\tau\nabla_\tau\varphi = \dddot{\varphi} \tag{6.3c}$$

$$\nabla_\tau\nabla_\tau\nabla_i\varphi = \nabla_i(\ddot{\varphi} - 2h\dot{\varphi} - \dot{h}\varphi + h^2\varphi), \tag{6.3d}$$

which are valid for any scalar function φ in the de Sitter metric in coordinates (1.2), and the fact that $\bar{\varphi}(u)$ is a function only of $u = H\tau$, we obtain

$$\delta E_u^{(1)u} = 2\dot{\bar{\varphi}}(\Box\phi)^{\cdot} + 2(\Box\bar{\varphi})^{\cdot}\dot{\phi} + 2\dot{\bar{\varphi}}\dddot{\phi} + 2\dddot{\bar{\varphi}}\dot{\phi} - 2(\Box\bar{\varphi})\ddot{\phi}$$
$$- 2\ddot{\bar{\varphi}}(\Box\phi) - (\Box\bar{\varphi})(\Box\phi) - \frac{4}{3}(\dddot{\bar{\varphi}}\dot{\phi} + 2\ddot{\bar{\varphi}}\ddot{\phi} + \dot{\bar{\varphi}}\dddot{\phi})$$
$$+ 6H^2\dot{\bar{\varphi}}\dot{\phi} - \frac{1}{3}\Box(\dot{\bar{\varphi}}\dot{\phi}) + \frac{2}{3}(\Box\phi)^{\cdot\cdot} - 4H^2\ddot{\phi}$$
$$-\frac{2}{3}H^2\Box\phi \tag{6.4}$$

for the $a = b = u$ component of this linearized stress tensor. Since the off-diagonal metric components, g_{ui}, are zero, it is slightly easier to compute in this case

$$E_{ui}^{(1)} \equiv \nabla_i V^{(1)} = \partial_i V^{(1)}, \tag{6.5}$$

where

$$V^{(1)} = -\dot{\bar{\varphi}}(\Box\phi) - (\Box\bar{\varphi})^{\cdot}\phi - 2\dot{\bar{\varphi}}[\ddot{\phi} - 2h\dot{\phi} + (h^2 - \dot{h})\phi]$$
$$+ 2(\Box\bar{\varphi})(\dot{\phi} - h\phi) + \frac{4}{3}(\ddot{\bar{\varphi}}\dot{\phi} + \dot{\bar{\varphi}}\ddot{\phi} - h\dot{\bar{\varphi}}\dot{\phi})$$
$$- 4H^2\dot{\bar{\varphi}}\phi - \frac{2}{3}[(\Box\phi)^{\cdot} - h\Box\phi] + 4H^2(\dot{\phi} - h\phi). \tag{6.6}$$

The linearized energy density perturbation in a given (klm_l) mode can be obtained from this component by using the conservation equation $\nabla_a T^{(1)au} = 0$ with $T^{(1)ab} = b'E^{(1)ab}$, or

$$H\frac{\partial\varepsilon^{(1)}}{\partial u} + 3h(\epsilon^{(1)} + p^{(1)}) = b'\frac{\Delta_3}{a^2}V^{(1)} \tag{6.7}$$

together with the condition,

$$p^{(1)} = \frac{1}{3}\varepsilon^{(1)}, \tag{6.8}$$

following from the vanishing of the trace, with the result that

$$\varepsilon_{klm_l}^{(1)} = \frac{b'}{Ha^4}\int_{-\infty}^{u} du a^2\Delta_3 V_{klm_l}^{(1)}$$
$$= -b'\frac{(k^2-1)}{Ha^4}\int_{-\infty}^{u} du a^2 V_{klm_l}^{(1)}. \tag{6.9}$$

Substituting (5.11) and (5.17) into (6.6), we obtain in particular the contributions

$$-\frac{2}{3}\dot{\bar{\varphi}}\ddot{\phi} = \frac{2H^3}{3}k^2\mathrm{sech}^2u\tanh u(1\mp 2\mathrm{sech}^2u)$$
$$\times\left[\frac{a_{klm_l}}{\sqrt{2k}}e^{-i(k-1)\eta}Y_{klm_l} + \frac{b_{klm_l}}{\sqrt{2k}}e^{-i(k+1)\eta}Y_{klm_l} + \mathrm{c.c.}\right]$$
$$+\cdots \tag{6.10}$$

and

$$-\frac{2}{3}(\Box\phi)^{\cdot} = \frac{2H^3}{3}k^2\mathrm{sech}^2u\left[-i\frac{a_{klm_l}}{\sqrt{2k}}e^{-ik\eta}Y_{klm_l}\right.$$
$$\left. + i\frac{b_{klm_l}}{\sqrt{2k}}e^{-ik\eta}Y_{klm_l} + \mathrm{c.c.}\right] + \cdots, \tag{6.11}$$

both of which are shown to leading order in k. The ellipsis and all other terms in (6.6) are subleading in k for $k \gg 1$. Since these terms are linear in k for large k, and because of the additional factor of k^2 from (6.9), these leading terms in k in $\varepsilon_k^{(1)}$ are proportional to k^3. Next taking into account the time dependence, we observe that since $e^{\pm i\eta} = \mathrm{sech}u \pm i\tanh u \to \mp i$ as $u \to -\infty$, the leading sech^2u behavior of $V_{klm_l}^{(1)}$ cancels in the sum of (6.10) and (6.11), so that the integrand of (6.9) vanishes at its lower limit, making the integral convergent. The surviving subleading term then gives a contribution to (6.9) of

$$\begin{aligned}\varepsilon^{(1)}_{klm_l} &\simeq -\frac{b'k^3}{a^4}\frac{\sqrt{2}}{3}\int_{-\infty}^{u} du\,\mathrm{sech}u\,\tanh u[(a_{klm_l}+b_{klm_l})\\ &\quad\times e^{-ik\eta}Y_{klm_l}+\mathrm{c.c.}]\\ &\simeq \frac{b'\sqrt{2}}{3}H^4k^3\mathrm{sech}^5u[(a_{klm_l}+b_{klm_l})i^kY_{klm_l}+\mathrm{c.c.}]\end{aligned} \tag{6.12}$$

as $u\to-\infty$. The integral in (6.12) can be computed exactly for $u=0$ with the result

$$\varepsilon^{(1)}_{klm_l}|_{u=0}\to -b'H^4\frac{\sqrt{2}}{3}k^2[(a_{klm_l}+b_{klm_l})i^{k+1}Y_{klm_l}+\mathrm{c.c.}] \tag{6.13}$$

for $k\gg 1$. Since the contribution of this $O(4)$ breaking leading term in k to the total linearized energy density is

$$\varepsilon^{(1)}=\sum_{k=1}^{\infty}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}\varepsilon^{(1)}_{klm_l} \tag{6.14}$$

it falls off proportional to a^{-5} from (6.12) and hence a factor of a^{-1} faster than the $O(4)$ symmetric terms in (5.14) as $u\to-\infty$. From (6.13) at $u=0$ its maximum value grows with the maximum momentum K for which the coefficients a_{klm_l} or b_{klm_l} are nonzero, proportional to

$$\begin{aligned}\varepsilon^{(1)}|_{u=0} &\sim -b'H^4\sum_{k=1}^{K}\sum_{l=0}^{k-1}\sum_{m_l=-l}^{l}k^2\sim -b'H^4\int_{k=1}^{K}dk k^4\\ &\simeq -b'H^4\frac{K^5}{5}\end{aligned} \tag{6.15}$$

and hence can easily exceed (5.14) at the symmetric point $u=0$ if $K\gg 1$. The backreaction of these $O(4)$ breaking terms in the stress tensor becomes significant when

$$8\pi GH^2|b'|\frac{K^5}{5}|a_{Klm_l}|\sim 8\pi GH^2|b'|\frac{K^5}{5}|b_{Klm_l}|\gtrsim 1 \tag{6.16}$$

which is even easier to satisfy for a larger range of state coefficients. Thus these k-dependent $O(4)$ breaking terms begin smaller and, for large enough K, grow larger to dominate the a^{-4} de Sitter breaking $O(4)$ symmetric terms (5.14) in the stress tensor as a decreases from infinity at I_-.

We conclude that the general $O(4)$ invariant vacuum states $|f\rangle$ defined by (3.20)–(3.26) are dynamically unstable to producing large deviations in the stress tensor, even more so than the $O(4,1)$ invariant $|v\rangle$. Hence the $O(4)$ symmetry subgroup and spatial homogeneity is also spontaneously broken in eternal de Sitter space. This conclusion which follows from the stress tensor of the anomaly could also be obtained by calculating the expectation value $\langle T_{ab}\rangle_R$ for $O(4)$ noninvariant states in the underlying QFT in de Sitter space.

VII. CONCLUSIONS

The main conclusion of our analysis of possible states in both de Sitter space and in a constant, uniform electric field is that the most symmetric state in such persistent background fields is *not* the stable vacuum state. Unlike flat Minkowski space where the Poincaré invariant vacuum is determined by a physical minimization of the energy, no such conserved Hamiltonian bounded from below exists in either de Sitter space or in a constant, uniform electric field. Instead both of these systems are characterized by a mixing of particle and antiparticle modes with respect to any proposed Hamiltonian generator, and are therefore unstable to spontaneous particle pair creation from the vacuum [8–10]. In each case the persistent or eternal background classical field provides an inexhaustible supply of energy to create pairs at a finite rate and subsequently accelerate them to ultrarelativistic particle energies. In this situation one should expect the symmetric state to be unstable to perturbations and capable of generating large backreaction effects, even in a semiclassical mean field approximation in which particle self-interactions are neglected. The study of particle creation in real time and the resulting vacuum decay rate given in an accompanying publication [9] is perhaps the clearest path to the instability of eternal de Sitter space.

In this work we have provided two additional approaches to an analysis of the instability. These are both based not on any particular definition of particles but on the study of perturbations of the maximally symmetric $|v\rangle$ states and the conserved currents they produce. In the electric field background this state is constructed in Sec. II and is a self-consistent solution of the semiclassical Maxwell equations (2.23b) just as the $O(4,1)$ invariant CTBD state is a self-consistent solution of the semiclassical Einstein equations (1.7), with a shifted cosmological constant. In each case there are large classes of perturbations of the symmetric state that produce an electric current $\langle j_z\rangle$ or stress tensor $\langle T_{ab}\rangle$ that are initially zero or negligibly small, but which grow larger than any prescribed finite value. In each case this is due to the blueshifting of field modes to ultrarelativistic energies. In de Sitter space this occurs clearly in the contracting phase $u<0$, and requires that backreaction of the energy-momentum through the semiclassical Einstein equations (1.7) be taken into account. Hence the assumption of a fixed de Sitter background is violated, eternal de Sitter space is unstable to perturbations satisfying (4.10), which produce large backreaction effects, and the classical $O(4,1)$ symmetry is broken by quantum fluctuations.

The second approach to instability of the maximally symmetric state in both the electric field and de Sitter backgrounds is through the relation to a quantum anomaly. The chiral anomaly and bosonization method in the two-dimensional Schwinger model shows that the invariance of the electric field background is broken by quantum effects. In the approximation of a fixed background field the solutions of the anomaly equation (2.45) for the effective boson field which are spatially homogeneous predict the linear growth with time (2.47), found also by direct study of the perturbations of the symmetric state. Even at the level of the effective action (2.48), the appearance of the Green's function $\square^{-1}$ of the two-dimensional wave operator makes it clear that there will be infrared sensitivity to boundary and/or initial conditions associated with the anomaly.

In the gravitational case it is the conformal trace anomaly which produces long lived infrared effects sensitive to either boundary or initial conditions. It is important that the kinematics of the persistent de Sitter background will always produce ultrarelativistic energies for large enough k so that the stress tensor eventually behaves like that of a massless conformal field which is described by the stress tensor of the anomaly. From the nonlocal form (5.1) and the infrared properties of the conformal operator Δ_4 and its inverse, it is already clear without detailed calculation that de Sitter invariance is broken. As for the two-dimensional chiral anomaly one can introduce a composite effective bosonic field φ whose equation of motion (5.6) has a constant source and therefore possesses no de Sitter invariant solutions. Since the Eq. (5.6) is de Sitter invariant but none of its solutions are, the anomaly provides a mechanism for spontaneous breaking of de Sitter symmetry [48]. The behavior of the $O(4)$ symmetric solutions which break de Sitter invariance is easily found and the same conclusion of the instability of global de Sitter space to these state perturbations follows. The anomaly approach is quite general and shows that the same large backreaction effect and instability to initial state perturbations occurs in de Sitter space for fields of any spin.

In both cases it is essential that moderate or small physical momenta are blueshifted to very large physical momenta, arbitrarily large if backreaction is turned off and the background field persists indefinitely. This unboundedness and relation to anomalies is a direct consequence of the infinite reservoir of arbitrarily high momentum or short distance modes in any vacuum state of QFT with no UV cutoff. The physical momentum which first produces large backreaction effects is of order $\sqrt{HM_{\rm Pl}}$. There is thus an interesting interplay of UV and IR physics in these effects, as has been noticed by other authors [49,50].

The instability of the de Sitter invariant state to large backreaction effects shows that the $O(4,1)$ symmetry of global de Sitter space is broken by quantum perturbations of the state. Of course, one can still construct fully $O(4,1)$ invariant theories in eternal fixed de Sitter spacetime mathematically by continuation from the Euclidean $\mathbb{S}^4$, order by order in perturbation theory [51,52]. By this construction the very state perturbations responsible for the instability of de Sitter space in real time are disallowed by the Euclidean regularity conditions. If one requires these regularity conditions, either explicitly by analytic continuation from $\mathbb{S}^4$, or implicitly in an *in-in* formalism in the Poincaré patch [53], and fixes the geometry to be de Sitter exactly, also disallowing the possibility of any dynamical backreaction through the semiclassical Einstein equations, it is not surprising then to find no sign of the instability we have discussed in this paper, in which these very restrictive assumptions are relaxed. This shows that it is not matter self-interactions *per se* which are critical for the instability, but rather the initial or boundary conditions imposed on states and Green's functions. If Euclidean boundary conditions are imposed, matter interactions lead to no apparent instability [51,52]. Conversely, instability is seen once those restricted boundary/initial conditions are relaxed, even in free QFT, and the backreaction effects of the energy-momentum tensor on the background de Sitter geometry are considered.

In [51,52] it has been further argued that any correlation function of an interacting massive scalar field theory approaches the expected CTBD value at late times or large spacetime separation at any order of perturbation theory, for an appropriately dense set of states. The approach to $O(4,1)$ invariance at late times is similar to that found earlier in [13], where it was proven that for all fourth order UV finite (and therefore Hadamard) adiabatic initial states with a spatially homogeneous and isotropic stress-energy tensor, the renormalized stress-energy expectation value $\langle T^a{}_b\rangle$ for a free scalar field with $m^2 + \xi R > 0$ asymptotically approaches its CTBD value in the expanding Poincaré patch of de Sitter space (1.4). In this sense the CTBD state is a late time attractor for $\langle T^a{}_b\rangle$ in de Sitter space for a free, massive scalar QFT, and appears to be quite stable. The attractor or “cosmic no-hair” behavior at late times, i.e. asymptotic future infinity I_+ in Fig. 2, found in [13] or at large separations in [51,52] is clearly a result of the cosmological redshift, which applies both in the expanding Poincaré patch and the $u > 0$ half of the full manifold in coordinates (1.2).

In this paper we have focused on the extreme sensitivity to initial conditions and instability of the stress tensor to perturbations due to the converse blueshifting effect in the contracting half $u < 0$ of the full de Sitter manifold. Due to this initial state sensitivity the backreaction must be taken into account long before the expanding phase even begins. If one nevertheless simply begins the Universe's evolution with only an expanding section of de Sitter space, the question naturally arises as to whether and how the global de Sitter instability we have demonstrated in this and the preceding paper [9] are relevant to inflation, cosmological vacuum energy, or cosmology more generally.

In this connection we make the following observations:

(i) de Sitter space is a homogeneous space, 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 conditions on a definite time slicing are imposed, which break $O(4,1)$ invariance.

(ii) The present study shows that $O(4,1)$ invariance is necessarily broken by quantum fluctuations, the larger k values with larger spatial inhomogeneities producing the larger energy-momentum tensor deviations from the de Sitter invariant equation of state $p = -\rho$, so that the assumption of de Sitter invariance and relevance to cosmology of the $O(4,1)$ invariant CTBD state is open to question.

(iii) The anomaly stress tensor shows that there are also spatially inhomogeneous perturbations of the initial state that break $O(4)$ symmetry and that vanish more rapidly in the infinite past and blueshift to even larger values at later times than the $O(4)$ symmetric ones. This shows that there is no spatially homogeneous $O(4)$ invariant stable vacuum state in global de Sitter space either.

(iv) In the expanding Poincaré patch a small amplitude deviation of the state from the CTBD state in sufficiently high $|\mathbf{k}|$ modes also produces large deviations of the stress tensor (proportional to $a^{-4} = e^{-4H\tau}$) at sufficiently early times $\tau \to -\infty$, highlighting the potentially extreme sensitivity of inflation to its UV initial conditions.

(v) In static coordinates (1.5) which cover one quarter of de Sitter space, cf. Fig. 2, entirely contained within the expanding Poincaré patch, individual field modes are infinitely blueshifted in the vicinity of the horizon $r = H^{-1}$ relative to $r = 0$. Since the instability studied in this paper arises when slight perturbations become strongly blueshifted, a similar instability to spatially inhomogeneous perturbations on the horizon scale should be expected.

(vi) Fluctuations in the stress tensor $\langle T_{ab}(x)T_{cd}(x')\rangle$ are certainly spatially inhomogeneous for spacelike separations on the horizon scale H^{-1}.

(vii) In a previous study of linear response in de Sitter space [15], incorporating these fluctuations in the stress tensor away from its mean value, we have found correspondingly large stress tensor perturbations in the vicinity of the cosmological horizon in the static coordinates of de Sitter space, suggesting that both de Sitter invariance and spatial homogeneity are broken at the horizon scale H^{-1}.

(viii) These new scalar cosmological horizon modes associated with the anomaly effective field and stress tensor are capable of generating fluctuations on the horizon which describe the observed anisotropies in the CMB [54].

These considerations taken together lead to the conclusion that de Sitter space is neither eternal nor able to preserve its spatial homogeneity under quantum state perturbations. They suggest instead that spatially inhomogeneous models of cosmological dark energy on the Hubble horizon scale H^{-1}, possessing at most only rotational $O(3)$ symmetry, may be required for a stable quantum vacuum state, and for determining the magnitude of cosmological dark energy in the universe.

ACKNOWLEDGMENTS

P. R. A. would like to thank Dillon Sanders for help with the early stages of this project. 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.

[1] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, England, 1973).

[2] N. A. Chernikov and E. A. Tagirov, Ann. Poincaré Phys. Theor. A **9**, 109 (1968).

[3] T. S. Bunch and P. C. W. Davies, Proc. R. Soc. A **360**, 117 (1978).

[4] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982), and references therein.

[5] J. S. Dowker and R. Critchley, Phys. Rev. D **13**, 224 (1976).

[6] O. Nachtman, Commun. Math. Phys. **6**, 1 (1967).

[7] G. W. Gibbons and S. W. Hawking, Phys. Rev. D **15**, 2738 (1977); A. S. Lapedes, J. Math. Phys. (N.Y.) **19**, 2289 (1978).

[8] E. Mottola, Phys. Rev. D **31**, 754 (1985).

[9] P. R. Anderson and E. Mottola, arXiv:1310.0030.

[10] J. Schwinger, Phys. Rev. **82**, 664 (1951).

[11] E. Mottola, Phys. Rev. D **33**, 2136 (1986); *Physical Origins of Time Asymmetry*, edited by (Cambridge University Press, Cambridge, Engand, 1993), p. 504.

[12] See I. Antoniadis, P. O. Mazur and E. Mottola, New J. Phys. **9**, 11 (2007) for a recent review.

[13] P. R. Anderson, W. Eaker, S. Habib, C. Molina-París, and E. Mottola, Phys. Rev. D **62**, 124019 (2000).

[14] G. Perez-Nadal, A. Roura, and E. Verdaguer, Classical Quantum Gravity **25**, 154013 (2008); M. B. Fröb, D. B. Papadoupoulos, A. Roura, and E. Verdaguer, Phys. Rev. D **87**, 064019 (2013).

[15] P. R. Anderson, C. Molina-París, and E. Mottola, Phys. Rev. D **80**, 084005 (2009).
[16] P. O. Mazur and E. Mottola, Phys. Rev. D **64**, 104022 (2001).
[17] E. Mottola and R. Vaulin, Phys. Rev. D **74**, 064004 (2006).
[18] E. Mottola, Acta Phys. Pol. B **41**, 2031 (2010).
[19] E. Mottola, Phys. Rev. D **33**, 1616 (1986).
[20] J.-T. Hsiang, L. H. Ford, D.-S. Lee, H.-L. Yu, Phys. Rev. D **83**, 084027 (2011).
[21] N. B. Narozhnyi, Zh. Eksp. Teor. Fiz. **54**, 676 (1968) [Sov. Phys. JETP **27**, 360 (1968)].
[22] A. I. Nikishov, Zh. Eksp. Teor. Fiz. **57**, 1210 (1970) [Sov. Phys. JETP **30**, 660 (1970)].
[23] N. B. Narozhnyi and A. I. Nikishov, Yad. Fiz. **11**, 1072 (1970) [Sov. J. Nuc. Phys. **11**, 596 (1970)].
[24] E. S. Fradkin, D. M. Gitman, and Sh. M. Shvartsman, *Quantum Electrodynamics with Unstable Vacuum* (Springer-Verlag, Berlin, 1991).
[25] 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).
[26] S. O. Gavrilov and D. M. Gitman, Phys. Rev. D **53**, 7162 (1996).
[27] Y. Kluger, E. Mottola, and J. M. Eisenberg, Phys. Rev. D **58**, 125015 (1998).
[28] *Higher Transcendental Functions*, Vols. I and II, Bateman Manuscript Project, edited by A. Erdélyi (McGraw-Hill, New York, 1953).
[29] B. Beers and H. H. Nickle, Lett. Nuovo Cimento **4**, 320 (1972); H. H. Nickle and B. L. Beers, J. Phys. A: Gen. Phys. **5**, 1658 (1972).
[30] J. Schwinger, Phys. Rev. **125**, 397 (1962); **128**, 2425 (1962).
[31] S. Coleman, Phys. Rev. D **11**, 2088 (1975); J. Fröhlich and E. Seiler, Helv. Phys. Acta **49**, 889 (1976).
[32] K. Johnson, Phys. Lett. **5**, 253 (1963).
[33] Y.-Z. Chu and T. Vachaspati, Phys. Rev. D **81**, 085020 (2010); See also: J. Garriga, S. Kanno, M. Sasaki, J. Soda, and A. Vilenkin, J. Cosmol. Astropart. Phys. 12 (2012) 006.
[34] See e.g. R. Jackiw, in *Relativity, Groups and Topology*, Vol. II, edited by B. DeWitt and R. Stora (North-Holland, Amsterdam, 1983).
[35] S. Habib, C. Molina-París and E. Mottola, Phys. Rev. D **61**, 024010 (1999).
[36] B. Allen and A. Folacci, Phys. Rev. D **35**, 3771 (1987).
[37] L. H. Ford and L. Parker, Phys. Rev. D **16**, 245 (1977); B. Allen, Phys. Rev. D **32**, 3136 (1985).
[38] P. R. Anderson, C. Molina-París, and E. Mottola, Phys. Rev. D **72**, 043515 (2005).
[39] P. R. Anderson, E. Mottola, and D. H. Sanders (unpublished).
[40] 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).
[41] M. Duff, Nucl. Phys. **B125**, 334 (1977); **B234**, 269 (1984); Classical Quantum Gravity **11**, 1387 (1994).
[42] R. J. Riegert, Phys. Lett. B **134**, 56 (1984); E. S. Fradkin and A. A. Tseytlin, Phys. Lett. B **134**, 187 (1984).
[43] A. M. Polyakov, Phys. Lett. B **103**, 207 (1981).
[44] M. Giannotti and E. Mottola, Phys. Rev. D **79**, 045014 (2009).
[45] J. Polchinski, Nucl. Phys. **B324**, 123 (1989).
[46] L. S. Brown and J. P. Cassidy, Phys. Rev. D **16**, 1712 (1977).
[47] I. Antoniadis, P. O. Mazur, and E. Mottola, Phys. Rev. D **55**, 4756 (1997); **55**, 4770 (1997).
[48] P. O. Mazur and E. Mottola, Nucl. Phys. **B278**, 694 (1986).
[49] A. M. Polyakov, Nucl. Phys. **B797**, 199 (2008); **B834**, 316 (2010).
[50] D. Krotov and A. M. Polyakov, Nucl. Phys. **B849**, 410 (2011).
[51] D. Marolf and I. A. Morrison, Phys. Rev. D **82**, 105032 (2010); **84**044040 (2011); Gen. Relativ. Gravit. **43**, 3497 (2011).
[52] S. Hollands, Commun. Math. Phys. **319**, 1 (2013).
[53] A. Higuchi, D. Marolf, and I. A. Morrison, Phys. Rev. D **83**, 084029 (2011).
[54] I. Antoniadis, P. O. Mazur and E. Mottola, J. Cosmol. Astropart. Phys. 09 (2012) 024.