

**Cosmological horizon modes and linear response in de Sitter spacetime**

Paul R. Anderson\*

*Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27109, USA  
and Departamento de Física Teórica and IFIC, Universidad de Valencia-CSIC, C. Dr. Moliner 50, Burjassot-46100, Valencia, Spain*

Carmen Molina-París†

*Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, United Kingdom*

Emil Mottola‡

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

(Received 6 July 2009; published 6 October 2009)

Linearized fluctuations of quantized matter fields and the spacetime geometry around de Sitter space are considered in the case that the matter fields are conformally invariant. Taking the unperturbed state of the matter to be the de Sitter invariant Bunch-Davies state, the linear variation of the stress tensor about its self-consistent mean value serves as a source for fluctuations in the geometry through the semiclassical Einstein equations. This linear response framework is used to investigate both the importance of quantum backreaction and the validity of the semiclassical approximation in cosmology. The full variation of the stress tensor  $\delta\langle T^a_b \rangle$  contains two kinds of terms: (1) those that depend explicitly upon the linearized metric variation  $\delta g_{cd}$  through the  $\langle [T^a_b, T^{cd}] \rangle$  causal response function; and (2) state dependent variations, independent of  $\delta g_{cd}$ . For perturbations of the first kind, the criterion for the validity of the semiclassical approximation in de Sitter space is satisfied for fluctuations on all scales well below the Planck scale. The perturbations of the second kind contain additional massless scalar degrees of freedom associated with changes of state of the fields on the cosmological horizon scale. These scalar degrees of freedom arise necessarily from the local auxiliary field form of the effective action associated with the trace anomaly, are potentially large on the horizon scale, and therefore can lead to substantial nonlinear quantum backreaction effects in cosmology.

DOI: [10.1103/PhysRevD.80.084005](https://doi.org/10.1103/PhysRevD.80.084005)

PACS numbers: 04.62.+v, 04.60.-m, 95.36.+x, 98.80.Qc

**I. INTRODUCTION**

Although matter is undeniably quantum, in general relativity its gravitational effects are treated classically. In order to include the quantum effects of matter, a natural next step from the purely classical theory is a semiclassical treatment, in which the classical matter stress tensor  $T^a_b$  in Einstein's equations is replaced by its expectation or average value  $\langle T^a_b \rangle$ , but the spacetime geometry is still treated classically. This is clearly an approximation to a more complete treatment, whose validity depends upon a number of assumptions, chief among them that the quantum fluctuations of the matter energy-momentum-stress tensor  $T^a_b$  about its average value are “small,” at least on macroscopic distance scales.

This heuristic condition can be given a precise meaning in the semiclassical or mean field limit of an effective field theory approach to gravity. In the limit,  $\hbar \rightarrow 0$  but the number of quantum fields  $N \rightarrow \infty$  with  $\hbar N$  fixed, it becomes permissible to replace the conserved  $T^a_b$  of a quantum theory by its expectation value  $\langle T^a_b \rangle$ , as the source term for the semiclassical Einstein equations,

$$R^a_b - \frac{R}{2} \delta^a_b + \Lambda \delta^a_b = 8\pi G_N \langle T^a_b \rangle_R. \quad (1.1)$$

The corrections to the renormalized expectation value  $\langle T^a_b \rangle_R$  are suppressed by  $1/N$ , and negligible in the large  $N$  limit [1]. In this way and in this limit the treatment of the metric geometry of spacetime as classical, sharply peaked about its mean value, can be obtained from an underlying quantum theory of matter.

Replacing the classical source of Einstein's equations by a quantum expectation value has a number of important implications. Since  $T^a_b(x)$  is an operator that is at least bilinear in local quantum fields at the same spacetime point, any direct evaluation of its expectation value is divergent. The expectation value  $\langle T^a_b \rangle$  must be renormalized, i.e. its divergent parts identified and removed by means of suitable local counterterms in the effective action of low energy gravity.

It is essential that the renormalization procedure maintain the covariant conservation of  $\langle T^a_b \rangle_R$ , for it is to act as a source of the gravitational field through the semiclassical Einstein Eqs. (1.1); otherwise, (1.1) would be inconsistent. Covariant methods for identifying and removing the ultraviolet divergences  $\langle T^a_b \rangle$  by point splitting, heat kernel expansions, or dimensional regularization, which satisfy the requirements of general covariance, have been devel-

\*anderson@wfu.edu

†carmen@maths.leeds.ac.uk

‡emil@lanl.gov

oped, and lead to consistent well-defined finite results for  $\langle T^a_b \rangle_R$  [2,3].

Equations (1.1) also provide consistency conditions for the finite terms in the renormalized expectation value  $\langle T^a_b \rangle_R$ . To evaluate this quantity in the usual Minkowski vacuum state of flat spacetime, one is *required* to define the renormalized expectation value of  $T^a_b$  to vanish identically, when  $\Lambda = 0$ , or (1.1) will not be satisfied for  $R^a_b = 0$ . Equivalently, by transposing the cosmological term  $\Lambda$  to the right side of (1.1), the existence of a flat spacetime solution to the semiclassical Einstein equations requires that the difference,  $\langle T^a_b \rangle_R - \frac{\Lambda}{8\pi G_N} \delta^a_b$  vanish for  $R^a_b = 0$ , so that any cosmological term is exactly canceled by a finite renormalization of  $\langle T^a_b \rangle_R$  in flat space. Thus, neither the vacuum expectation value  $\langle T^a_b \rangle_R$  nor the cosmological term  $\Lambda \delta^a_b$  alone have direct physical significance, but only the combination,  $\langle T^a_b \rangle_R - \frac{\Lambda}{8\pi G_N} \delta^a_b$ , which acts as the source of curvature.

The condition that flat spacetime be a solution to the semiclassical Eqs. (1.1) with all matter in its Lorentz invariant Minkowski vacuum state is a renormalization condition on the low energy effective theory of gravity. This condition serves to define the effective theory by fixing the subtractions of infinite and finite vacuum terms, and makes the semiclassical framework a predictive one in spaces other than flat space and/or in states other than the vacuum state. At present this renormalization condition cannot be justified by any fundamental theory, but solely by appeal to the experimental fact that the very high frequency fluctuations responsible for the divergences in  $\langle T^a_b \rangle$  do not produce any appreciable mean curvature of space. The semiclassical theory described by (1.1) must be regarded then as a low energy effective theory [4], on a par with pion effective theories of nuclear forces or the Fermi theory of the weak interactions, whose UV divergences are subsumed into effective low energy parameters, and whose predictions are reliable only at sufficiently low energies or long wavelengths, pending a more complete, but unspecified theory at high energies and short wavelengths.

Classically, with  $\Lambda > 0$  and in the absence of any other sources, there is a maximally symmetric nonflat solution to Einstein's equations, namely, de Sitter spacetime with  $R^a_b = \Lambda \delta^a_b$ . In the classical theory  $\Lambda$  is a fixed constant with absolutely no dynamics, and de Sitter space is a stable solution to Einstein's equations with a positive cosmological constant. On the other hand, in the semiclassical framework of (1.1), as just observed, the cosmological constant term cannot be divorced from the definition of  $\langle T^a_b \rangle_R$  itself, as  $\Lambda$  can be reabsorbed into a finite redefinition of  $\langle T^a_b \rangle_R$ . Since  $\langle T^a_b \rangle_R$  depends upon the state in which it is evaluated, the energy of the vacuum is no longer a fixed nondynamical constant of the Lagrangian, as in the classical theory, but instead depends upon the infrared choice of "vacuum" state in curved space, in which the expectation value  $\langle T^a_b \rangle_R$  is defined. Which infrared state and which

spacetime are preferred then becomes a dynamical question, which can be investigated by probing the response of the system to perturbations of the state of the fields and the spacetime geometry together on length scales much greater than the Planck scale. In the semiclassical theory (1.1) the stability of de Sitter space to these infrared perturbations depends upon the behavior of the quantum fluctuations of the matter sources on cosmological length scales of the order of  $H^{-1} = \sqrt{3/\Lambda}$ , and the dynamical question can be studied in a systematic way with the same effective action that gives rise to (1.1) by the method of linear response.

The linear response of a system to a small perturbation is familiar from other branches of physics, and has been formulated for semiclassical gravity in Refs. [5,6]. The fluctuations of the matter fields lead to fluctuations in the expectation value  $\langle T^a_b \rangle_R$ , which is probed by considering the response of  $\langle T^a_b \rangle_R$  to a small perturbation of the geometry  $g_{ab} \rightarrow g_{ab} + \delta g_{ab}$ . The quantum action principle tells us that this variation is given formally by

$$\begin{aligned} \delta \langle T^a_b(x) \rangle &= \frac{1}{2} \int d^4 x' \sqrt{-g(x')} \Pi^a_{b\ cd(\text{ret})}(x, x') \delta g_{cd}(x') \\ &= -\frac{i\hbar}{2} \int d^4 x' \sqrt{-g(x')} \theta(t, t') \\ &\quad \times \langle in|[T^a_b(x), T^{cd}(x')]|in \rangle \delta g_{cd}(x'), \end{aligned} \quad (1.2)$$

up to contact terms proportional to delta functions and derivatives thereof with support at  $x = x'$  [5,6]. The local contact terms are determined by the renormalization of the retarded polarization  $\Pi^a_{b\ cd(\text{ret})}(x, x')$  of the matter field(s), evaluated in the same state as the self-consistent background solution of (1.1).

The linear variation (1.2) is essentially equivalent to first order perturbation theory in the metric perturbation  $\delta g_{ab} \equiv h_{ab}$ , and assumes that the entire change in the quantum state and the expectation value,  $\langle T^a_b \rangle_R$  comes about directly from the change in the geometry. These variations, linear in the metric variation  $h_{ab}$ , we term variations of the first kind. This first class of linear perturbations are the ones upon which attention has been focused in the earlier literature on this subject [7–12]. However, it is also possible to consider variations of the state that lead to variations in  $\langle T^a_b \rangle_R$ , which are *not* directly driven by local variations in the metric. For example, one might consider variations of the temperature of a system coupled to a heat reservoir, or other variations of the boundary conditions imposed on the system. These variations may depend upon dynamical degrees of freedom contributing to the stress tensor other than the local metric geometry, and hence are not proportional to  $\delta g_{cd}$ . The corresponding state dependent variations in  $\langle T^a_b \rangle_R$  which do not depend upon the local  $\delta g_{cd}$ , we term variations of the second kind. We shall see that the trace anomaly parameterizes certain specific state dependent variations of the stress tensor, which are of

this second kind, and that they can have significant effects in de Sitter space.

The dynamical linear response equation which probes the self-consistent solution of (1.1) to small fluctuations of both kinds is the linear integro-differential equation for the self-consistent metric perturbation  $h_{ab}$ ,

$$\delta \left\{ R^a_b - \frac{R}{2} \delta^a_b + \Lambda \delta^a_b \right\} = 8\pi G_N \delta \langle T^a_b \rangle_R, \quad (1.3)$$

obtained when the local linearized variation of the left side of (1.1) is set equal to the properly renormalized right side. Once the ultraviolet divergences of this linear response equation are isolated and removed, (1.3) can be used to study the infrared fluctuations of matter and geometry of de Sitter space in a self-consistent and reliable way. This is our main purpose in this paper.

The most direct way of deriving (1.3) and determining the correct renormalization counterterms necessary to properly define the formal expressions for  $\langle T^a_b \rangle$  and  $\delta \langle T^a_b \rangle$  is to integrate out the matter fields, obtaining their one-loop effective action in a general gravitational background. In the large  $N$  limit, the contributions of higher order gravitational metric fluctuations themselves to  $\delta \langle T^a_b \rangle$  are suppressed by  $1/N$  and may be neglected at leading order. In this limit the metric may be treated as a (semi-)classical quantity. In particular, there are no stochastic or “noise” terms in the linear response equations, (1.3) at leading order in the  $1/N$  expansion. These can arise only at higher orders and are outside the domain of the semiclassical theory.

Renormalizing the one-loop matter effective action and taking its first variation yields the semiclassical Einstein Eqs. (1.1). In addition to fixing the background geometry, the solution of (1.1) involves also fixing the quantum state (or density matrix)  $|in\rangle\langle in|$  of the fields, with respect to which all expectation values are to be computed. Then a second variation of the effective action, or first variation of (1.1) is performed around this background metric to yield the linear response Eq. (1.3) for  $h_{ab}$ . The kernel of the polarization integral in (1.2) is computed using the background geometry and the Green’s functions for the quantum field in the specific quantum state (density matrix) fixed in the previous step. The form of the counterterms used to define the renormalized expectation value  $\langle T^a_b \rangle_R$  are the same used to renormalize its linear variation  $\delta \langle T^a_b \rangle_R$  [6].

In [6] we proposed a criterion for the validity of the semiclassical Einstein Eqs. (1.1), namely, that the solutions of the linearized Eqs. (1.3) should lead to linearized gauge invariant amplitudes that remain bounded for all times. The criterion is satisfied in flat spacetime. The present work extends the analysis and tests the criterion in de Sitter spacetime for conformal matter, where the variation of the stress tensor in (1.3) also can be evaluated in closed form.

The interest in applying this linear response method to de Sitter spacetime arises from several different considerations. First, de Sitter spacetime serves as a fundamental testing ground for semiclassical effects of quantum vacuum fluctuations in a curved space with a cosmological constant, in which analytic results can be obtained. Second, current cosmological models of structure formation in the Universe assume that this structure grew by gravitational instability from small quantum inhomogeneities in an otherwise featureless, primordial de Sitter-like inflationary phase. Finally, observations of Type Ia supernovae suggest that the Universe may be in a de Sitter-like accelerating phase in the present epoch, with some 70–75% of the energy density of the Universe today being of a hitherto undetected dark variety, with negative pressure,  $p \approx -\rho$  [13]. Since this is exactly the equation of state of a cosmological term or that of the quantum “vacuum,” the study of quantum fluctuations and their stability in de Sitter spacetime is of possible direct relevance to the cosmology of the early Universe, structure formation, and the cosmic acceleration of the present Universe.

In this paper we study the solutions of this dynamical linear response Eq. (1.3) under small fluctuations in de Sitter spacetime for the case of conformal matter fields, i.e. those that are classically conformally invariant, up to the quantum conformal or trace anomaly. The study of metric perturbations in cosmology due to quantum conformal matter in cosmology was initiated in Ref. [7], whose authors made use of an in-out effective action formalism, rather than the retarded boundary conditions appropriate for causal linear response. The in-in formalism for the effective action in cosmology was developed in Ref. [11], and extended and elaborated in Ref. [12]. The general form of the perturbed stress tensor in an arbitrary Friedmann-Robertson-Walker (FRW) cosmological spacetime obtained in [12], agrees with the earlier results of Refs. [8–10]. We study the solution of the resulting causal linear response Eqs. (1.3) with this general stress tensor source for inhomogeneous and anisotropic perturbations away from de Sitter space in Sec. VII A. Previous discussions of linearized perturbations of de Sitter space have focused instead on spatially homogeneous, isotropic perturbations, earlier in the functional Schrödinger initial state formalism [14], and more recently for a minimally coupled scalar field in the in-in effective action formalism [15].

For the more general case of spatially inhomogeneous and anisotropic perturbations, we find the analogous short distance (i.e. Planck scale) quantum fluctuations expected from the corresponding analysis in flat spacetime, which should be discarded as outside the range of validity of the semiclassical theory. Going further, when general inhomogeneous and anisotropic state dependent variations of the stress tensor are considered (i.e. variations of the second kind), we find additional modes arising from coherent macroscopic changes of state on every scale, characterized

by a spatial wave vector  $\vec{k}$ , including the scale of the de Sitter horizon itself. These cosmological horizon modes are associated with the effective action of the trace anomaly, not variations of the local metric geometry, and as such are different than anything in the purely classical theory, or encountered in previous studies of restricted classes of perturbations. The existence of this new set of scalar cosmological horizon modes arising from the quantum trace anomaly is a principal result of this paper.

To keep this result in perspective, and when considering quantum effects in cosmology generally, it is important to recognize that the intrinsically quantum phenomenon of phase coherence and consequently nonclassical effects may be present at any scale. Indeed macroscopic quantum states are encountered at low enough temperatures in virtually all branches of physics on a very wide variety of scales. In addition to low temperature and condensed matter laboratory systems, the chiral symmetry breaking expectation value of quark bilinears  $\langle \bar{q}q \rangle$  in QCD and the Higgs field  $\langle \Phi \rangle$  in electroweak theory are examples of macroscopic vacuum condensates, which extend over arbitrarily large distances. It is these condensates that provide the connection between a more microscopic approach to both the strong and electroweak interactions and the low energy effective theories that preceded them. Such coherence effects due to the quantum wavelike properties of matter, and its propensity to form phase correlated states over macroscopic distance scales cannot be treated in a purely classical description of gravity. However, the semiclassical treatment of the effective stress-energy tensor source  $\langle T^a_b \rangle_R$  for Einstein's equations already allows for such effects, in a mean field description [16].

A significant part of the effective action for conformally invariant fields is determined by the conformal or trace anomaly. Although intrinsically nonlocal in terms of the spacetime geometry, the effective action determined by the anomaly can be cast into a covariant local form by the introduction of one or more scalar auxiliary degree(s) of freedom [16–21]. Since there are two distinct cocycles in the nontrivial cohomology of the Weyl group in four dimensions [22], the most general representation of the anomaly action is in terms of two auxiliary scalar degrees of freedom, each satisfying fourth order linear differential equations of motion (4.8). The resulting stress tensor in terms of these scalar auxiliary fields has been used to study the nonlocal macroscopic quantum coherence effects contained in the semiclassical effective theory (1.1) [16,23]. The scalar auxiliary fields are related to the freedom to change the macroscopic quantum vacuum state, and supply additional infrared relevant terms in the low energy effective theory of gravity [22]. In perturbations about de Sitter spacetime we shall show they give rise to new cosmological horizon modes, not contained in classical perturbation theory, that can lead to large backreaction effects at the very largest Hubble scale of a de Sitter universe. Thus,

scalar degrees of freedom in cosmology arise naturally from the effective action of the conformal anomaly of massless quantum fields in the standard model, without the *ad hoc* introduction of an inflaton field.

The paper is organized as follows: In order to fix ideas and notations we review in the next section the linear response fluctuation analysis in flat spacetime with a vanishing cosmological constant. In Sec. III, we consider linear response for conformal matter fields in conformally flat spacetimes, where the stress tensor variation of the first kind (1.2) can be expressed in closed form. We show that for long wavelength solutions of the equations within the range of validity of the semiclassical theory, it is sufficient to study only the time-time component of the linear response equations, in a general FRW spacetime. In Sec. IV, we consider the stress tensor and linear response equations that follow from the trace anomaly part of the effective action only, in the auxiliary field description, comparing the result to the previous general approach in conformally flat spacetimes. In Sec. V, we derive the linear response equations in both approaches for fluctuations in de Sitter spacetime, and show how the auxiliary field degrees of freedom in the anomaly action describe cosmological horizon scale modes. In Sec. VI, we discuss coordinate transformations and recast our linear response equations in a completely gauge invariant form, showing, in particular, that the new cosmological horizon modes are gauge invariant. We give the quadratic gauge invariant action corresponding to the linear response equations in these variables. In Sec. VII, we give a qualitative and quantitative discussion of the solutions of the linear response equations in de Sitter space, showing that the auxiliary field anomaly action captures all the infrared quantum effects correctly. In Sec. VIII, the cosmological horizon modes, their interpretation as fluctuations in the Hawking-de Sitter temperature, and their possible implications for the stability of de Sitter space are discussed. Section IX contains a full summary and discussion of our results. There are two appendices, the first containing an in-depth discussion of the properties of the nonlocal kernel  $K$ , which arises in the linear response equations in conformally flat spaces, and in the second, some of the details of the second variation of the anomaly action. Except where indicated the units used are  $\hbar = c = 1$  and the metric and curvature conventions are the same as those of Misner, Thorne, and Wheeler [24].

## II. LINEAR RESPONSE IN FLAT SPACETIME

In this section the derivation and solutions to the linear response equations in a flat space background are reviewed. In flat spacetime Lorentz invariance permits the full decomposition of perturbations into scalar and tensor components, with respect to the background Minkowski four-metric  $\eta_{ab} = \text{diag}(-+++)$ . In Ref. [6] explicit results for these components were obtained for a quantum

scalar field with general mass  $m$  and curvature coupling  $\xi$ . Making use of those results as a specific example, we generalize the discussion here to any set of underlying quantum fields. We emphasize that the four dimensional tensor decomposition in Minkowski spacetime is distinct from the decomposition of the perturbations into scalar, vector, and tensor components with respect to the spatial three-metric  $g_{ij}$  appropriate for general FRW spacetimes, and employed in the succeeding sections.

The starting point of a linear response analysis of quantum fluctuations is a self-consistent solution to the semiclassical Einstein equations (1.1). In flat spacetime this requires that  $\langle T^a_b \rangle_R = 0$  in the usual Minkowski vacuum state with  $\Lambda = 0$ . Since it is a dimension-four operator, the renormalization of  $\langle T^a_b \rangle$  requires fourth order counterterms in the effective action. These gives rise to additional geometric terms in Eqs. (1.1), fourth order in derivatives, which are usually displayed on the left side of the semiclassical Eqs. (1.1). We shall instead adopt the convention of including such geometric terms in the definition of the renormalized expectation value,  $\langle T^a_b \rangle_R$  itself: cf. Eqs. (2.1) and (2.2) below. Viewing the semiclassical theory as an effective field theory, these local higher derivative terms multiplied by finite coefficients are in any case negligibly small at macroscopic length scales much greater than the extreme microscopic Planck scale,  $L_{\text{Pl}} = \sqrt{\hbar G_N/c^3} \simeq 1.616 \times 10^{-33}$  cm, or at energy scales much less than the Planck energy  $M_{\text{Pl}}c^2 = \sqrt{\hbar c^5/G_N} \simeq 1.223 \times 10^{19}$  GeV.

In flat spacetime, the nonlocal integro-differential Eq. (1.2) in coordinate space is easily handled by Fourier transforming to momentum space. With the self-consistent solution of (1.1) for  $\Lambda = 0$  just the usual Minkowski vacuum state, the linear response equations around flat spacetime take the form [6]

$$\begin{aligned} \delta G_{ab} &= 8\pi G_N \delta \langle T_{ab} \rangle_R \\ &= 8\pi G_N \left\{ \left( \alpha_R + \frac{F^{(T)}}{2} \right) \delta^{(C)} H_{ab} \right. \\ &\quad \left. + \left( \beta_R + \frac{F^{(S)}}{12} \right) \delta^{(1)} H_{ab} \right\}. \end{aligned} \quad (2.1)$$

Here, the two local, conserved geometric tensors,  ${}^{(C)}H_{ab}$  and  ${}^{(1)}H_{ab}$ , also denoted as  $-A_{ab}$  and  $-B_{ab}$ , respectively, are

$$\begin{aligned} {}^{(C)}H_{ab} &\equiv \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{ab}} \int d^4x \sqrt{-g} C_{abcd} C^{abcd} \\ &= 4\nabla_c \nabla_d C_{(a}{}^c{}_{b)}{}^d + 2C_{a}{}^c{}_{b}{}^d R_{cd} = -A_{ab}, \end{aligned} \quad (2.2a)$$

$$\begin{aligned} {}^{(1)}H_{ab} &\equiv \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{ab}} \int d^4x \sqrt{-g} R^2 \\ &= 2g_{ab} \square R - 2\nabla_a \nabla_b R + 2RR_{ab} - \frac{g_{ab}}{2} R^2 \\ &= -B_{ab}. \end{aligned} \quad (2.2b)$$

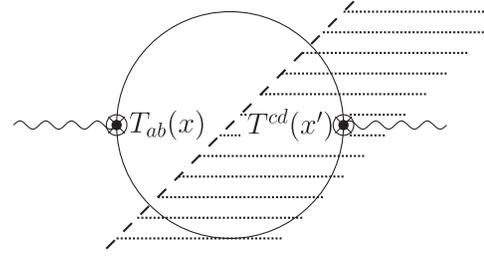


FIG. 1. The one-loop vacuum polarization  $\Pi_{ab}{}^{cd}(x, x')$ . Its imaginary part, obtained by cutting the internal lines as illustrated and placing them on shell gives the spectral functions  $\rho^{(T,S)}$ .

These tensors, which are derived from the two dimension-four local invariants,  $C_{abcd}C^{abcd}$  and  $R^2$ , respectively, are required for renormalization of the expectation value  $\langle T_{ab} \rangle$ . The parameters  $\alpha_R$ , and  $\beta_R$  are the finite, renormalized coefficients of these local terms. The causal response functions,

$$F^{(T,S)}(K^2) \equiv K^2 \int_0^\infty \frac{ds}{s^3} \frac{\rho^{(T,S)}(s)}{s + K^2 - i\epsilon \text{sgn} \omega}, \quad (2.3)$$

in (2.1) have been renormalized by subtracting the dispersion integral three times at  $s = 0$ .<sup>1</sup> They are the nonlocal part of  $\delta \langle T_{ab} \rangle_R$  arising from the matter fluctuations. The renormalization effected by subtracting the Taylor series expansion in  $K^2$  of the unrenormalized response function up to order  $(K^2)^2$  fixes the value of the cosmological constant term, Newtonian constant (order  $K^2$ ) and fourth order terms  $A_{ab}$  and  $B_{ab}$  at order  $(K^2)^2$  to be their prescribed renormalized values at  $K^2 = 0$ . The renormalization prescription and local counterterms of the retarded polarization tensor  $\Pi_{ab}{}^{cd(\text{ret})}(x, x')$  at  $x = x'$  are defined in momentum space by this procedure at  $K^2 = 0$ . The modification of the renormalization procedure necessary for massless quantum fields is discussed below.

The superscripts  $T$  and  $S$  in (2.3) denote quantities associated with tensor (spin-2) and scalar (spin-0) perturbations, which are defined with respect to the four dimensional background Minkowski metric by Eqs. (2.14) below. As discussed in Ref. [6] these causal response functions are completely determined by their imaginary parts in terms of the spectral functions  $\rho^{(T,S)}$  of the matter fields. The spectral functions  $\rho^{(T,S)}$  are finite and may be computed directly from the one-loop cut diagram of Fig. 1 in flat spacetime.

It is essential that the linear response Eq. (2.1) is causal, i.e. that the stress tensor response at  $(t, \vec{x})$  is sensitive to variations of the geometry only within the past light cone

<sup>1</sup>In this section only we employ the notation  $K^a$  for a four-vector in flat spacetime, and  $K^2 \equiv \eta_{ab} K^a K^b = -\omega^2 + k^2$  to distinguish it from  $k \equiv |\vec{k}|$ , the magnitude of its three spatial components used throughout the paper.

of  $(t, \vec{x})$ . The  $-i\epsilon\text{sgn}\omega$  prescription (where  $K^2 = -\omega^2 + k^2$ ,  $k \equiv |\vec{k}|$  and  $\omega \equiv K^0$ ) enforces retarded boundary in-in conditions, appropriate for this causal linear response.

Note also that both terms on the right side of (2.1) are geometrical, involving the linearized metric fluctuation  $h_{ab}$  itself, i.e. they are homogeneous variations of the first kind only. It is clear that nongeometrical terms could always be added to  $\delta\langle T^a_b \rangle_R$ , by allowing states other than the Minkowski vacuum to contribute to the variation, even in the absence of any metric variation. These *state dependent* contributions will be considered in the next section.

The expressions (2.1), (2.2), and (2.3) are finite and general for the linear response of a quantum field theory in the usual Minkowski vacuum perturbed around flat spacetime. To linear order all indices are raised and lowered with the flat space metric  $\eta_{ab}$ . The dependence on the particular quantum matter theory enters only through the spectral functions  $\rho^{(T,S)}$ , which must be positive definite on general grounds of unitarity. Explicit calculations of the spectral functions and causal response functions of a free scalar field theory with mass  $m$  and curvature coupling  $\xi$  are given in Ref. [6]. The results are

$$\rho^{(S)}(s) = \frac{\theta(s - 4m^2)}{24\pi^2} \sqrt{1 - \frac{4m^2}{s}} \left[ m^2 + \frac{(1 - 6\xi)s}{2} \right]^2, \quad (2.4a)$$

$$\rho^{(T)}(s) = \frac{\theta(s - 4m^2)}{60\pi^2} \sqrt{1 - \frac{4m^2}{s}} \left( \frac{s}{4} - m^2 \right)^2. \quad (2.4b)$$

For conformal scalars,  $m = 0$  and  $\xi = \frac{1}{6}$ , the scalar spectral function  $\rho^{(S)}$  vanishes for all  $s > 0$ . However, the expression  $F^{(S)}$  in (2.3) is finite in the limit  $m \rightarrow 0$  limit. The two response functions in this limit are

$$F^{(S)} = \frac{1}{1440\pi^2} = \frac{4b}{3}, \quad (2.5a)$$

$$F^{(T)} \rightarrow \frac{1}{960\pi^2} \ln\left(\frac{K^2 - i\epsilon\text{sgn}\omega}{m^2}\right) = 2b \ln\left(\frac{K^2 - i\epsilon\text{sgn}\omega}{m^2}\right), \quad (2.5b)$$

where  $b = \frac{1}{(4\pi)^2} \frac{1}{120}$  is the coefficient of the  $C_{abcd}C^{abcd}$  term of the trace anomaly for the conformal scalar field in an arbitrary curved background. The logarithmic divergence in (2.5b) in the limit of zero mass is the result of the renormalization of  $\alpha_R$  at  $K^2 = 0$ . If the renormalization is performed instead at an arbitrary but nonzero mass scale  $\mu^2$ , and we define

$$\alpha_R(\mu^2) \equiv \alpha_R + b \ln\left(\frac{\mu^2}{m^2}\right), \quad (2.6)$$

then the combinations appearing in the linear response Eq. (2.1) are

$$\alpha_R + \frac{F^{(T)}}{2} = \alpha_R(\mu^2) + b \ln\left(\frac{K^2 - i\epsilon\text{sgn}\omega}{\mu^2}\right), \quad (2.7a)$$

$$\beta_R + \frac{F^{(S)}}{12} = \beta_R + \frac{b}{9}. \quad (2.7b)$$

These are finite in the conformal limit  $m \rightarrow 0$ , provided  $\alpha_R(\mu^2)$  remains finite in that limit. Note also that since

$$\mu^2 \frac{d}{d\mu^2} \left[ \alpha_R(\mu^2) + b \ln\left(\frac{K^2 - i\epsilon\text{sgn}\omega}{\mu^2}\right) \right] = 0, \quad (2.8)$$

the physical linear response does not depend on the arbitrary renormalization scale  $\mu^2$ , i.e. this combination is renormalization group invariant. For the scalar response function  $F^{(S)}$  of a conformal matter field, there is no logarithmic term and no  $\mu^2$  dependence at one-loop order, and the last term in (2.7b) is simply an additional finite renormalization of the coefficient  $\beta_R$  of the  $R^2$  term in the effective action.

Equations (2.5) and (2.7) are valid in fact for any conformal matter field, provided one uses the appropriate anomaly  $b$  coefficient for that field. This may be seen from the general form of the trace anomaly for conformal matter fields in an arbitrary curved spacetime, viz.,

$$\begin{aligned} \langle T^a_a \rangle_R &= bF + b' \left( E - \frac{2}{3} \square R \right) + b'' \square R \\ &= b \left( F + \frac{2}{3} \square R \right) + b'E + \left( b'' - \frac{2b + 2b'}{3} \right) \square R, \end{aligned} \quad (2.9)$$

where

$$E \equiv {}^*R_{abcd} {}^*R^{abcd} = R_{abcd}R^{abcd} - 4R_{ab}R^{ab} + R^2, \quad (2.10a)$$

and

$$F \equiv C_{abcd}C^{abcd} = R_{abcd}R^{abcd} - 2R_{ab}R^{ab} + \frac{R^2}{3}, \quad (2.10b)$$

with  $R_{abcd}$  the Riemann curvature tensor,  ${}^*R_{abcd} = \frac{1}{2} \varepsilon_{abef} R^{ef}_{cd}$  its dual, and  $C_{abcd}$  the Weyl conformal tensor. The coefficients  $b$ ,  $b'$ , and  $b''$  are computed at one-loop order, proportional to  $\hbar$ , and determined by the number of massless conformal fields and their spin via

$$b = \frac{\hbar}{120(4\pi)^2} (N_S + 6N_F + 12N_V), \quad (2.11a)$$

$$b' = -\frac{\hbar}{360(4\pi)^2} (N_S + 11N_F + 62N_V), \quad (2.11b)$$

with  $N_S$  the number of spin 0 fields,  $N_F$  the number of spin  $\frac{1}{2}$  Dirac fields, and  $N_V$  the number of spin 1 fields [3]. The  $\hbar$  in (2.11) emphasizes that these coefficients appear at one-loop order, and that the anomaly terms are present in the leading order semiclassical limit,  $\hbar \rightarrow 0$ ,  $N \rightarrow \infty$ , with  $\hbar N$  finite (where  $N$  is any of  $N_S$ ,  $N_F$ ,  $N_V$ ). With this understood, we shall henceforth set  $\hbar = 1$ .

The  $b$  coefficient is also the one that determines the logarithmic scale dependence of the  $C_{abcd}C^{abcd}$  term in the effective action. The coefficient  $b''$  is scheme dependent, corresponding as it does to the trace of the tensor  ${}^{(1)}H_{ab} = -B_{ab}$ , which is derived from a local action, whose coefficient can be shifted by a finite shift of  $\beta_R$ . For classically conformally invariant theories there is no logarithmic scale dependence in the coefficient of the  $R^2$  term in the effective action at one-loop order and one finds  $b'' = 2(b + b')/3$  in covariant regularization schemes, such as dimensional regularization, so that the last term in (2.9) vanishes [25].

From these considerations we deduce that the renormalized stress tensor variation in (2.1) in Fourier space for conformal matter in flat spacetime may be written in the form

$$\delta\langle T_{ab} \rangle_{R,\text{conf}} = -\left[ \alpha_R(\mu^2) + b \ln\left(\frac{K^2 - i\epsilon \text{sgn}\omega}{\mu^2}\right) \right] \delta A_{ab} - \left( \beta_R + \frac{b}{9} \right) \delta B_{ab}, \quad (2.12)$$

which is obtained by substituting Eq. (2.7) into Eq. (2.1). Since both the invariants  $E$  and  $F$  are quadratic in the curvature tensor, their linear variations away from flat spacetime vanish, and only the  $\frac{2}{3}b\Box\delta R$  term survives in the variation of the trace,

$$\delta\langle T^a{}_a \rangle_{R,\text{conf}} = \left( 6\beta_R + \frac{2b}{3} \right) \Box(\delta R). \quad (2.13)$$

All of the dependence of the linear response on the conformal matter theory in flat spacetime is contained therefore in the single parameter  $b$ , determined by the trace anomaly (2.9) and given by (2.11a) for free conformal matter fields of any spin. The  $b'$  anomaly coefficient of the Euler-Gauss-Bonnet density  $E$  does not enter the linear variation of the stress tensor around flat spacetime.

The tensor and scalar components in (2.1) or (2.12) are conveniently separated with the aid of the set of orthogonal projectors

$$P_{ab}^{(T)cd} = \frac{1}{2}(\theta_a^c \theta_b^d + \theta_a^d \theta_b^c) - \frac{1}{3}\theta_{ab}\theta^{cd}, \quad (2.14a)$$

$$P_{ab}^{(V)cd} = \frac{1}{2}(\delta_a^c \delta_b^d + \delta_a^d \delta_b^c - \theta_a^c \theta_b^d - \theta_a^d \theta_b^c), \quad (2.14b)$$

$$P_{ab}^{(S)cd} = \frac{1}{3}\theta_{ab}\theta^{cd}, \quad (2.14c)$$

$$\theta_{ab} \equiv \eta_{ab} - \partial_a \frac{1}{\Box} \partial_b \rightarrow \eta_{ab} - \frac{K_a K_b}{K^2} \quad (2.14d)$$

in momentum space, so that the linearized tensor and scalar metric perturbations around flat spacetime are

$$h_{ab}^{(T,S)} \equiv P_{ab}^{(T,S)cd} h_{cd}, \quad (2.15)$$

respectively. Both perturbations are transverse in the four dimensional sense,

$$\partial^b h_{ab}^{(T,S)} = 0 = K^b h_{ab}^{(T,S)}, \quad (2.16)$$

while the tensor perturbation is also tracefree in the four dimensional sense,

$$\eta^{ab} h_{ab}^{(T)} = 0. \quad (2.17)$$

The linear variations around flat spacetime of the three local geometric tensors appearing in (2.1) may be expressed then as

$$\begin{aligned} \delta G_{ab} &= \delta G_{ab}^{(T)} + \delta G_{ab}^{(S)} \\ &= -\frac{1}{2}\Box h_{ab}^{(T)} + \Box h_{ab}^{(S)} \rightarrow \frac{K^2}{2} h_{ab}^{(T)} - K^2 h_{ab}^{(S)}, \end{aligned} \quad (2.18a)$$

$$\delta A_{ab} = \Box^2 h_{ab}^{(T)} \rightarrow (K^2)^2 h_{ab}^{(T)}, \quad (2.18b)$$

$$\delta B_{ab} = 6\Box^2 h_{ab}^{(S)} \rightarrow 6(K^2)^2 h_{ab}^{(S)}, \quad (2.18c)$$

in terms of this decomposition into tensor and scalar metric components in momentum space. The remaining projector onto linearized vector perturbations (2.14b) is not needed because  $h_{ab}^{(V)} = P_{ab}^{(V)cd} h_{cd}$  is gauge dependent and nontransverse, and hence cannot appear in the linear variations of the conserved tensors in (2.1).

Making use of the projection operators (2.14), we find that the linear response equation around flat spacetime (2.1) separates finally into two independent, orthogonal components, viz.,

$$\begin{aligned} \delta G_{ab}^{(T)} &= \frac{K^2}{2} h_{ab}^{(T)} \\ &= -8\pi G_N \left\{ \alpha_R(\mu^2) + b \ln\left(\frac{K^2 - i\epsilon \text{sgn}\omega}{\mu^2}\right) \right\} (K^2)^2 h_{ab}^{(T)}, \end{aligned} \quad (2.19a)$$

$$\begin{aligned} \delta G_{ab}^{(S)} &= -\frac{1}{3}\theta_{ab}\delta R = -K^2 h_{ab}^{(S)} \\ &= -8\pi G_N \left( 6\beta_R + \frac{2b}{3} \right) (K^2)^2 h_{ab}^{(S)}. \end{aligned} \quad (2.19b)$$

The solutions of the linear response equations in flat spacetime were discussed in Ref. [6] for scalar fields with arbitrary mass and curvature coupling  $\xi$ . In Fourier space it is straightforward to show that there are no low energy solutions to (2.19) other than the usual propagating transverse, traceless, spin-2 gravitational wave modes at  $K^2 = 0$  of the pure Einstein theory with no sources. Since these transverse degrees of freedom are massless, there are only two helicity states that propagate, the other three spin-2 states and the one apparent scalar mode in (2.19b) being eliminated by the four diffeomorphism constraints coming from the  $ti$  and  $tt$  components of the equations in a canonical analysis splitting spacetime into space + time.

Additional nontrivial solutions to both the semiclassical tensor and scalar equations appear at  $K^2 \sim M_{\text{Pl}}^2$  (for  $\alpha_R$  and  $\beta_R$  of order unity). However, these solutions are outside the range of applicability of the semiclassical effective theory, which requires  $K^2 \ll M_{\text{Pl}}^2$ , and hence should be discarded. Physically, this must be the case since there is no indepen-

dent length scale against which the wavelength of the metric perturbations of a conformally invariant quantum field can be compared in flat spacetime other than the Planck length. The situation will be quite different for cosmological spacetimes where the horizon scale plays an important role. From (2.13) or (2.19b) it is clear that the Planck scale solutions in the scalar or trace sector involve  $\delta R \neq 0$ . These Planck scale solutions will be present in the linear response equations around arbitrary spacetimes as well. Hence, we shall find it convenient to set  $\delta R = 0$ , in order to exclude these Planck scale solutions in the effective field theory description, and focus on the remaining solutions (if any) that are determined by the low energy curvature or horizon scale in cosmology, which is assumed many orders of magnitude greater than the Planck scale.

### III. LINEAR RESPONSE OF CONFORMAL MATTER IN COSMOLOGY

In this section some general properties of conformally invariant fields in conformally flat spacetimes are briefly reviewed along with some previous computations of the perturbed stress-energy tensor for these fields in these spacetimes. Then the specific form of the linear response equations in a general spatially flat FRW cosmological spacetime is derived and the infrared physical content of the equations is shown to reside purely in the  $tt$  component.

A conformally flat spacetime is one for which the metric is

$$g_{ab}(x) = \Omega^2(x)\eta_{ab}, \quad (3.1)$$

with  $\eta_{ab}$  the Minkowski metric. The function  $\Omega(x)$  is called the conformal factor. Conformally invariant fields are ones whose classical action is invariant under conformal transformations when the field is multiplied by some power of the conformal factor. If the spacetime under consideration is conformally flat and the matter fields are conformally invariant, then the state of the fields that is most often considered is called the conformal vacuum [3]. It is obtained by mapping the Minkowski vacuum and all of its Green's functions in flat spacetime to the conformally flat spacetime using the same transformation of the fields that keeps the action invariant under conformal transformations. In this special state the expectation value of the stress tensor of the conformally invariant fields is completely determined by its trace anomaly (2.9) and can be written in terms of local curvature invariants and their derivatives in the geometry specified by (3.1) in the form [3,26],

$$\langle T_{ab} \rangle_{R,\text{conf}} = -2b^{(3)}H_{ab} - \frac{b}{9}B_{ab}, \quad (3.2)$$

where

$$\begin{aligned} {}^{(3)}H_{ab} \equiv & -R_a{}^c R_{cb} + \frac{2}{3}RR_{ab} + \frac{1}{2}R_{cd}R^{cd}g_{ab} - \frac{1}{4}R^2g_{ab} \\ & - 2C_{abcd}R^{cd}, \end{aligned} \quad (3.3)$$

and  $B_{ab}$  is defined by (2.2b). Of course, just as there are many states in flat space that are of interest besides the Minkowski vacuum, the conformal state is simply one choice of allowed state among many, even if we restrict ourselves to the subclass of states, which are spatially homogeneous and isotropic, consistent with the symmetries of (3.4). Applying the term, ‘‘vacuum’’ to the conformal state is also somewhat misleading, since  $\langle T_{ab} \rangle_{R,\text{conf}} \neq 0$ .

Some time ago several authors considered the linear variation of the expectation value (3.2) for conformal matter fields by conformally transforming the linear variation (2.12) in flat space [9,10]. It was assumed in that work that the fields are in the conformal vacuum state and the variation considered was of the first kind, namely, it was assumed that the quantum state of the fields follows the metric variation in a prescribed way, with no other state dependent variations considered. The results of Refs. [9,10] are used in this section to write the specific form of the linear response equations in a spatially flat FRW spacetime, with the addition of the state dependent terms, which are omitted in [9,10], but certainly allowed on general grounds [3].

The line elements for spatially flat FRW spacetimes can be written in the alternate forms

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2 = \Omega^2(\eta)(-d\eta^2 + d\vec{x}^2), \quad (3.4)$$

where  $\eta$  is the conformal time and  $t$  is the comoving time. These cosmological spacetimes are conformally flat with the conformal factor  $\Omega(\eta) = a(t)$ , the FRW scale factor.

In the flat space version of the linear response equations it was useful to perform a full Fourier transform. For a FRW spacetime it is useful to Fourier transform only in space, but not in time. The inverse Fourier transform can be used to undo the time part of the logarithmic kernel in (2.12) with the result,

$$\begin{aligned} K(\eta - \eta'; k; \mu) = & \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(\eta - \eta')} \\ & \times \ln \left[ \frac{-\omega^2 + k^2 - i\epsilon \text{sgn}\omega}{\mu^2} \right]. \end{aligned} \quad (3.5)$$

This is a distribution, which may be defined by its action upon integration against a suitably smooth, restricted class of test functions  $f(\eta')$  for which  $\int d\eta' K(\eta - \eta')f(\eta')$  exists. Because of the  $-i\epsilon \text{sgn}\omega$  prescription,  $K$  has support only for times  $\eta'$  earlier than the time  $\eta$ , consistent with causal linear response. At  $\eta = \eta'$  the distribution (3.5) is singular and requires regularization. The properties of this distribution and two possible regularizations are studied in detail in Appendix A.

In terms of this distribution the linearized perturbation of  $\langle T_{ab} \rangle_R$  for a conformal field in the conformal vacuum around a conformally flat spacetime with metric (3.4) is [9,10]

$$\begin{aligned} \delta \langle T_{ab}(\eta; \vec{k}) \rangle_R &= -\frac{b}{\Omega^2} \int_{\eta_0}^{\eta} d\eta' K(\eta - \eta'; k; \mu^2) \delta \bar{A}_{ab}(\eta'; \vec{k}) \\ &+ L_{ab} + \frac{1}{\Omega^2} \delta \langle \bar{T}_{ab}(\eta; \vec{k}) \rangle_R. \end{aligned} \quad (3.6)$$

Here,  $L_{ab}$  is the local contribution,

$$\begin{aligned} L_{ab} &= -\alpha_R(\mu^2) \delta A_{ab} - \left( \beta_R + \frac{b}{9} \right) \delta B_{ab} - 2b' \delta^{(3)} H_{ab} \\ &- \frac{8b}{\Omega^2} \partial_c \partial_d [(\ln \Omega) \delta \bar{C}_{a \ b}^{\ c \ d}], \end{aligned} \quad (3.7)$$

and the final term in (3.6) is the variation of the stress tensor arising from the variation of the state away from the conformal ‘‘vacuum,’’ which was omitted in [9,10], but which is present in the more general context of state dependent variations [3]. Except for the lower limit of the integral in (3.6) introduced to allow for the possibility of starting the linear response system at some finite initial conformal time  $\eta_0$ , and the addition of the variation of the final state dependent term in (3.6), these are the results for the variation of the stress tensor of conformal matter around a conformally flat space as given by the authors of Refs. [9,10].

In (3.6) both  $\delta \bar{A}_{ab}$  and  $\delta \bar{C}_{a \ b}^{\ c \ d}$  are evaluated on the linearized perturbation

$$\bar{h}_{ab} \equiv \Omega^{-2} h_{ab} \quad (3.8)$$

from flat spacetime, with  $h_{ab}$  the full linearized metric perturbation about the conformally flat spacetime  $g_{ab} = \Omega^2 \eta_{ab}$  of (3.4). The linearized variation around flat space of  ${}^{(C)}H_{ab} = -A_{ab}$  is given in momentum space by Eqs. (2.14), (2.15), (2.16), (2.17), and (2.18) of the previous section, while  $\delta \bar{C}_{a \ b}^{\ c \ d}$  is given by the completely traceless part of

$$\delta \bar{R}_{a \ b}^{\ c \ d} = \frac{1}{2} (\partial_b \partial^c \bar{h}_a^{\ d} + \partial_a \partial^d \bar{h}_b^{\ c} - \partial_a \partial_b \bar{h}^{cd} - \partial^c \partial^d \bar{h}_{ab}) \quad (3.9)$$

in the usual way, all derivatives and indices referring to flat Minkowski coordinates.

As in the corresponding expression for flat spacetime (2.12), the  $\alpha_R(\mu^2) \delta A_{ab}$  term in (3.7) comes from varying the local  $C_{abcd} C^{abcd}$  Weyl invariant term in the effective action, while the  $\beta_R \delta B_{ab}$  term is a finite renormalization of the local  $R^2$  term in the effective action, which we have chosen to include in the variation of  $\langle T_{ab} \rangle_R$ . The  $b \delta B_{ab}$  and  $b' \delta^{(3)} H_{ab}$  terms are the variations of the background geometric terms in (3.2), while the  $\partial_c \partial_d [(\ln \Omega) \delta \bar{C}_{a \ b}^{\ c \ d}]$  term has been obtained by the authors of Refs. [9,10] in different ways, and is necessary for consistency. This term will be shown in the next section to be a necessary consequence of

the covariant effective action functional determined by the trace anomaly (2.9).

Since  $\delta A_{ab}$ ,  $\delta \bar{A}_{ab}$  and the last term in (3.7) involving the Weyl tensor variation are traceless, the trace of the linear response equations for conformal matter perturbed away from a FRW background for which  $C_{abcd} = 0$  simplifies considerably. In fact, for variations of the first kind, the trace is determined completely by the trace anomaly (2.9), which is given in terms of local curvature terms and is independent of the quantum state of the field. Hence, the trace of the variation is given by the local geometric equation

$$\begin{aligned} -\delta R &= 8\pi G_N \delta \langle T^a_{\ a} \rangle_R \\ &= 16\pi G_N \left\{ -2b' \left( R^a_{\ c} \delta R^c_{\ a} - \frac{1}{3} R \delta R \right) \right. \\ &\quad \left. + \left( 3\beta_R + \frac{b}{3} \right) \delta(\square R) \right\}. \end{aligned} \quad (3.10)$$

By assumption, in the semiclassical approximation we require the Ricci curvature and its variations to be much smaller than the Planck scale. Thus all the terms on the right side of (3.10) are small compared to the left side. Hence, we may restrict ourselves to only those solutions of the full linear response equations for which

$$\delta R = 0. \quad (3.11)$$

The only solutions eliminated by this condition are Planck scale solutions, such as the ones we have already encountered in flat spacetime, with  $k^2 \sim M_{\text{Pl}}^2$ , which lie outside the range of validity of the semiclassical approximation under consideration. Henceforth we will restrict ourselves therefore to the tracefree or nonconformal solutions of the linear response Eqs. (1.3) around FRW spacetimes. Earlier authors have studied the conformal variations of the metric and stress tensor for conformally invariant fields in de Sitter space in more detail and found no interesting gauge invariant modes other than those expected on the Planck scale [14]. The condition (3.11) eliminates these trace or conformal Planck scale solutions from the start, and simplifies the analysis of the remaining tracefree, non-conformal perturbations. This condition also eliminates the class of quantum inflationary solutions considered in Refs. [27,28].

With  $\delta R = 0$ , we are restricted to the sector of tracefree metric perturbations (in the four dimensional sense). Further, in expanding about a FRW background one can decompose the metric perturbations into scalar, vector and tensor perturbations with respect to the conformally flat three dimensional metric  $g_{ij} = a^2 \eta_{ij}$ , and focus on the scalar sector (in the three dimensional sense). The vector and tensor perturbations for conformal matter around a conformally flat spacetime are not expected to show any nontrivial degrees of freedom in linear response other than those in the standard classical theory.

The metric perturbations that are scalar with respect to the background three-metric can be parameterized in terms of four functions,  $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{E})$ , of the form [29,30]

$$h_{tt} = -2\mathcal{A}, \quad (3.12a)$$

$$h_{tj} = a\partial_j\mathcal{B} \rightarrow iak_j\mathcal{B}, \quad (3.12b)$$

$$h_{ij} = 2a^2\left[\eta_{ij}\mathcal{C} + \left(\frac{\eta_{ij}}{3}k^2 - k_ik_j\right)\mathcal{E}\right] \quad (3.12c)$$

in momentum space. As is well known, and elaborated further in Sec. VI, only two linear combinations of these four functions are gauge invariant. Since we have also required (3.11), this means that there remains only one dynamical gauge invariant metric function to be determined by linear response in this scalar sector. The information about this remaining metric degree of freedom is contained completely in the  $tt$  component of the full linear response equations.

To show this, let the background semiclassical Einstein Eqs. (1.1) be written in the form

$$\mathcal{T}^a{}_b \equiv \langle T^a{}_b \rangle_R - \frac{1}{8\pi G_N} \left( R^a{}_b - \frac{R}{2} \delta^a{}_b + \Lambda \delta^a{}_b \right) = 0. \quad (3.13)$$

Their covariant conservation

$$\nabla_a \mathcal{T}^a{}_b = \partial_a \mathcal{T}^a{}_b + \Gamma^a{}_{ac} \mathcal{T}^c{}_b - \Gamma^c{}_{ab} \mathcal{T}^a{}_c = 0, \quad (3.14)$$

and first variation imply

$$\partial_t(\delta\mathcal{T}^t{}_t) + \partial_i(\delta\mathcal{T}^i{}_t) + \Gamma^j{}_{jt}(\delta\mathcal{T}^t{}_t) - \Gamma^i{}_{jt}(\delta\mathcal{T}^j{}_i) = 0, \quad (3.15a)$$

$$\partial_t(\delta\mathcal{T}^t{}_i) + \partial_j(\delta\mathcal{T}^j{}_i) + \Gamma^j{}_{jt}(\delta\mathcal{T}^t{}_i) - \Gamma^j{}_{it}(\delta\mathcal{T}^t{}_j) - \Gamma^t{}_{ij}(\delta\mathcal{T}^j{}_i) = 0 \quad (3.15b)$$

for the time and space components, respectively. Here, the background Eqs. (3.13) have been used along with the spatial homogeneity and isotropy of the background. This leaves the Christoffel symbols shown in (3.15) as the only remaining nonzero ones. In coordinates (3.4) the nonvanishing Christoffel symbols are

$$\Gamma^i{}_{jt} = \Gamma^i{}_{tj} = \frac{\dot{a}}{a} \delta^i{}_j, \quad (3.16a)$$

$$\Gamma^t{}_{ij} = \frac{\dot{a}}{a} \eta_{ij}, \quad (3.16b)$$

where the overdot denotes differentiation with respect to the comoving proper time variable  $t$ . Note that because of the self-consistent background Eqs. (3.13), one may equally well consider the variations,  $\delta\mathcal{T}^a{}_b$ ,  $\delta\mathcal{T}_{ab}$  or  $\delta\mathcal{T}^{ab}$ , freely raising and lowering indices with the background metric.

Using relations (3.16) in (3.15a), and imposing the time-time component of the linear response equations,

$$\delta\mathcal{T}^t{}_t = 0 \quad (3.17)$$

and the total trace of the linear response equations,

$$\delta\mathcal{T}^a{}_a = \delta\mathcal{T}^t{}_t + \delta\mathcal{T}^i{}_i = 0, \quad (3.18)$$

equivalent to (3.10), gives

$$\partial_i(\delta\mathcal{T}^i{}_t) = 0. \quad (3.19)$$

For scalar perturbations  $\delta\mathcal{T}^i{}_t = g^{ij}\partial_j\mathcal{P}$  for some scalar function  $\mathcal{P}$ , there being no way for a spatial vector to appear in the scalar sector other than by differentiation with respect to  $x^i$ . Then (3.19) implies that  $\vec{\nabla}^2\mathcal{P} = 0$  or  $\mathcal{P} = 0$  up to a possible constant. Since a constant in  $\mathcal{P}$  is eliminated in  $\delta\mathcal{T}^i{}_t = g^{ij}\partial_j\mathcal{P}$ , (3.19) implies

$$\delta\mathcal{T}^i{}_t = 0, \quad (3.20)$$

or in other words, the time-space components of the linear response equations around a general FRW space are automatically satisfied for scalar metric perturbations of the form (3.12), provided the self-consistent background Eq. (3.13), the time-time component (3.17), and the trace Eq. (3.18), and are all satisfied.

To complete the proof one then notes from (3.17) and (3.18), that the spatial trace  $\delta\mathcal{T}^i{}_i$  vanishes, leaving only a possible traceless part, in  $\delta\mathcal{T}^i{}_j$ , which for scalar perturbations can be expressed in the form,

$$\delta\mathcal{T}^i{}_j = \left( \frac{1}{3} \delta^i{}_j - g^{il} \partial_l \frac{1}{\vec{\nabla}^2} \partial_j \right) \mathcal{Q} \quad (3.21)$$

for some scalar function  $\mathcal{Q}$ . Substituting this form into (3.15b), and using (3.17) and (3.20) gives  $\partial_i\mathcal{Q} = 0$ , which by an argument identical to that for the time-space component implies

$$\delta\mathcal{T}^i{}_j = 0. \quad (3.22)$$

Thus, (3.20) and (3.22) show that the sufficient conditions for *all* of the semiclassical linear response equations to be satisfied around a general FRW background for scalar perturbations of the metric (3.12) are

- (i) the background is self-consistent, so that Eq. (3.13) is satisfied;
- (ii) the time-time component (3.17) is imposed;
- (iii) the trace Eq. (3.18) is imposed.

Since the trace equation can be imposed by (3.10), for non-Planck scale fluctuations, the  $tt$  component contains the only remaining gauge invariant dynamical information about scalar perturbations parameterized by (3.12) around a self-consistent solution of the semiclassical Einstein Eqs. (3.13). Being able to focus solely on the time-time component will simplify our analysis considerably.

Since the variation of  $\langle T_{ab} \rangle$  of conformal matter in a conformally flat spacetime is determined by the form of the trace anomaly (2.9), another route to the linear response equations in this case is afforded by the effective action and stress tensor corresponding to the anomaly, which has been

derived in several previous works. In the next section this alternative derivation from the effective action and stress tensor of the anomaly is given and compared with the above derivation based on Refs. [9,10].

#### IV. ANOMALY ACTION AND STRESS TENSOR

For massless matter or radiation fields, there is no intrinsic length scale associated with their quantum fluctuations, and the conformal variation of the one-loop effective action in a general curved background is determined completely by the trace anomaly. Although the anomaly determines uniquely only the part of the effective action that responds to conformal variations, and leaves the remaining conformally invariant part undetermined, it has been argued elsewhere that the anomaly action contains the only effective field theory corrections to Einstein's theory, which can be relevant in the infrared, i.e. at macroscopic

distance scales much greater than the Planck scale [22,31]. The availability of the alternative formulation of linear response for conformal matter around conformally flat backgrounds reviewed in the previous section allows this argument to be tested in FRW spacetimes by comparing directly the energy-momentum tensor and linear response equations derived from the anomaly action with the previous exact formulation.

The general form of the effective action that records the effects of the trace anomaly is nonlocal in the metric. However, in several earlier works a local form was obtained by the introduction of two scalar auxiliary fields,  $\varphi$  and  $\psi$  [16,23]. This auxiliary field effective action is of the form

$$S_{\text{anom}} = b' S_{\text{anom}}^{(E)}[g; \varphi] + b S_{\text{anom}}^{(F)}[g; \varphi, \psi], \quad (4.1)$$

with

$$S_{\text{anom}}^{(E)}[g; \varphi] \equiv \frac{1}{2} \int d^4x \sqrt{-g} \left\{ -(\square\varphi)^2 + 2 \left( R^{ab} - \frac{R}{3} g^{ab} \right) (\nabla_a \varphi) (\nabla_b \varphi) + \left( E - \frac{2}{3} \square R \right) \varphi \right\}, \quad (4.2a)$$

$$S_{\text{anom}}^{(F)}[g; \varphi, \psi] \equiv \int d^4x \sqrt{-g} \left\{ -(\square\varphi)(\square\psi) + 2 \left( R^{ab} - \frac{R}{3} g^{ab} \right) (\nabla_a \varphi) (\nabla_b \psi) + \frac{1}{2} F \varphi + \frac{1}{2} \left( E - \frac{2}{3} \square R \right) \psi \right\}. \quad (4.2b)$$

By varying this coordinate invariant effective action functional with respect to the spacetime metric, we obtain the stress-energy tensor

$$T_{ab}^{\text{anom}} = b' E_{ab} + b F_{ab}, \quad (4.3)$$

where the two separately conserved tensors are given by Eqs. (3.41) and (3.42) of [16], namely,

$$\begin{aligned} E_{ab} = & -2(\nabla_{(a}\varphi)(\nabla_{b)}\square\varphi) + 2\nabla^c[(\nabla_c\varphi)(\nabla_a\nabla_b\varphi)] - \frac{2}{3}\nabla_a\nabla_b(\nabla\varphi)^2 + \frac{2}{3}R_{ab}(\nabla\varphi)^2 - 4R^c{}_{(a}(\nabla_{b)}\varphi)(\nabla_c\varphi) + \frac{2}{3}R(\nabla_a\varphi) \\ & \times (\nabla_b\varphi) + \frac{1}{6}g_{ab}\{-3(\square\varphi)^2 + \square(\nabla\varphi)^2\} + 2(3R^{cd} - Rg^{cd})(\nabla_c\varphi)(\nabla_d\varphi) - \frac{2}{3}\nabla_a\nabla_b\square\varphi - 4C_a{}^c{}_b{}^d\nabla_c\nabla_d\varphi \\ & - 4R^c{}_{(a}\nabla_{b)}\nabla_c\varphi + \frac{8}{3}R_{ab}\square\varphi + \frac{4}{3}R\nabla_a\nabla_b\varphi - \frac{2}{3}(\nabla_{(a}R)\nabla_{b)}\varphi + \frac{1}{3}g_{ab}\{2\square^2\varphi + 6R^{cd}\nabla_c\nabla_d\varphi - 4R\square\varphi + (\nabla^cR)\nabla_c\varphi\}, \end{aligned} \quad (4.4)$$

and

$$\begin{aligned} F_{ab} = & -2(\nabla_{(a}\varphi)(\nabla_{b)}\square\psi) - 2(\nabla_{(a}\psi)(\nabla_{b)}\square\varphi) + 2\nabla^c[(\nabla_c\varphi)(\nabla_a\nabla_b\psi) + (\nabla_c\psi)(\nabla_a\nabla_b\varphi)] - \frac{4}{3}\nabla_a\nabla_b[(\nabla_c\varphi)(\nabla^c\psi)] \\ & + \frac{4}{3}R_{ab}(\nabla_c\varphi)(\nabla^c\psi) - 4R^c{}_{(a}[(\nabla_{b)}\varphi)(\nabla_c\psi) + (\nabla_b)\psi)(\nabla_c\varphi)] + \frac{4}{3}R(\nabla_{(a}\varphi)(\nabla_{b)}\psi) \\ & + \frac{1}{3}g_{ab}\{-3(\square\varphi)(\square\psi) + \square[(\nabla_c\varphi)(\nabla^c\psi)] + 2(3R^{cd} - Rg^{cd})(\nabla_c\varphi)(\nabla_d\psi)\} - 4\nabla_c\nabla_d(C_{(a}{}^c{}_b{}^d\varphi) - 2C_a{}^c{}_b{}^dR_{cd}\varphi \\ & - \frac{2}{3}\nabla_a\nabla_b\square\psi - 4C_a{}^c{}_b{}^d\nabla_c\nabla_d\psi - 4R^c{}_{(a}\nabla_{b)}\nabla_c\psi + \frac{8}{3}R_{ab}\square\psi + \frac{4}{3}R\nabla_a\nabla_b\psi - \frac{2}{3}(\nabla_{(a}R)\nabla_{b)}\psi \\ & + \frac{1}{3}g_{ab}\{2\square^2\psi + 6R^{cd}\nabla_c\nabla_d\psi - 4R\square\psi + (\nabla^cR)(\nabla_c\psi)\}, \end{aligned} \quad (4.5)$$

where  $(\nabla\varphi)^2 = (\nabla_a\varphi)(\nabla^a\varphi)$ . By varying (4.1) with respect to the local auxiliary fields one obtains their equations of motion,

$$\Delta_4\varphi = \frac{E}{2} - \frac{\square R}{3}, \quad (4.6a)$$

$$\Delta_4\psi = \frac{F}{2} = \frac{1}{2}C_{abcd}C^{abcd}, \quad (4.6b)$$

which are linear in  $\varphi$  and  $\psi$ . Here,

$$\Delta_4 \equiv \square^2 + 2R^{ab}\nabla_a\nabla_b - \frac{2}{3}R\square + \frac{1}{3}(\nabla^aR)\nabla_a. \quad (4.7)$$

Since the terms quadratic in the auxiliary fields in (4.2) are conformally invariant, they lead to only traceless terms in the corresponding stress tensors (4.4) and (4.5). Therefore, the terms in the trace are linear in  $\varphi$  and  $\psi$ , and in fact are proportional to (4.6), yielding the local geometrical traces,

$$E^a{}_a = 2\Delta_4\varphi = E - \frac{2}{3}\square R, \quad (4.8a)$$

$$F^a{}_a = 2\Delta_4\psi = F = C_{abcd}C^{abcd}, \quad (4.8b)$$

Thus, the *classical* trace of the stress tensor following from the local effective action (4.2) reproduces the quantum trace anomaly (2.9).

Moreover, it was shown in Refs. [21,22,32] by considering its conformal variation under  $g_{ab} \rightarrow \bar{g}_{ab} = e^{-2\sigma} g_{ab}$  that

$$\begin{aligned} T_{ab}^{\text{WZ}}[\bar{g}; \sigma] &\equiv -\frac{2}{\sqrt{-\bar{g}}} \frac{\delta}{\delta g^{ab}} \Gamma^{\text{WZ}}[g; \sigma] \\ &= 2^{(3)}\bar{H}_{ab} + \frac{1}{9}{}^{(1)}\bar{H}_{ab}, \end{aligned} \quad (4.9)$$

where  $T_{ab}^{\text{WZ}}[g; \sigma]$  is the stress tensor of the Wess-Zumino effective action,

$$\begin{aligned} \Gamma^{\text{WZ}}[g; \sigma] &= \int d^4x \sqrt{-g} \left[ 2\sigma \Delta_4 \sigma + \left( E - \frac{2}{3} \square R \right) \sigma \right] \\ &= -S_{\text{anom}}^{(E)}[g; \varphi = -2\sigma] \end{aligned} \quad (4.10)$$

obtained by varying the metric  $g_{ab}$ , keeping  $\sigma$  fixed. Considering Eqs. (4.1), (4.2), (4.3), and (4.4), dropping the overbar on (4.9) and using the notation of (2.2) one finds that

$$E_{ab}|_{\varphi=2\ln\Omega} = -T_{ab}^{\text{WZ}} \left[ g; \sigma = -\frac{\varphi}{2} \right] = -2^{(3)}H_{ab} + \frac{1}{9}B_{ab}, \quad (4.11)$$

in the full metric  $g_{ab}$ , provided

$$g_{ab} = e^\varphi \eta_{ab} = \Omega^2 \eta_{ab} \quad (4.12)$$

is conformally flat, with the auxiliary field  $\varphi = 2 \ln \Omega$ . It is easily checked from the conformal variation,

$$E - \frac{2}{3} \square R = 4 \Delta_4 \ln \Omega = 2 \Delta_4 \varphi \quad (4.13)$$

in the full metric, that  $\varphi = 2 \ln \Omega$  satisfies (4.8a). These relations may be proven directly as well, by the use of formulae for the conformal variations of the Ricci tensor, such as

$$R_{ab} = -\nabla_a \nabla_b \varphi - \frac{1}{2} (\nabla_a \varphi) (\nabla_b \varphi) + \frac{g_{ab}}{2} [-\square \varphi + (\nabla \varphi)^2], \quad (4.14a)$$

$$R = -3 \square \varphi + \frac{3}{2} (\nabla \varphi)^2 \quad (4.14b)$$

in the full metric  $g_{ab} = e^\varphi \eta_{ab}$ .

Thus, provided that the auxiliary field is fixed in terms of the geometric conformal factor via  $\varphi = 2 \ln \Omega$ , the tensor  $E_{ab}$  reduces to the purely local combination of geometric tensors given by (4.11). Taking the second auxiliary field  $\psi = 0$ , and the Weyl tensor  $C_a{}^c{}_b{}^d = 0$ , gives

$$T_{ab}^{\text{anom}} = b' E_{ab} = -2b'^{(3)}H_{ab} + \frac{b'}{9}B_{ab} \quad (4.15)$$

for the full contribution of the anomaly effective action (4.1) to the expectation value of the energy-momentum tensor of conformal matter in an arbitrary conformally

flat spacetime. We note that for a conformally flat spacetime of the form (3.4) and  $\varphi(t) = 2 \ln a(t) + c_0$ , the tensor  $E^a{}_b$  derived from the auxiliary field effective action gives rise to precisely the geometric tensor  ${}^{(3)}H^a{}_b$  discussed in Ref. [22], which is not derivable from a local geometric action and for that reason was called ‘‘accidentally conserved’’ in [3]. The actual origin of this tensor is the form of the anomaly action (4.2).

Equation (4.15) differs from the known exact result (3.2) only by the addition of a  $b'' \square R$  term to the anomaly and corresponding local  $R^2$  term to the effective action with  $b'' = 2(b + b')/3$ , the same choice of  $b''$  expected from the flat space result: cf. Eqs. (2.9), (2.10), (2.11), (2.12), and (2.13). Since such local terms with arbitrary coefficients should be added to the effective action in any case, we conclude that the action (4.1) plus these additional local terms gives the correct general form for the stress tensor of conformal matter in its conformal vacuum state in an arbitrary conformally flat spacetime.

To study linear response next we wish to vary the anomaly stress tensor (4.3) away from the conformally flat background. Since the relations (4.11) and (4.15) hold for arbitrary conformal transformations, it follows that they hold also for linearized conformal transformations about a conformally flat FRW spacetime. Hence, it is clear that in the variation of (4.15) we must expect to obtain local geometrical terms, which are just the variations of the local geometrical terms in (4.15). By adding the local  $F = C_{abcd} C^{abcd}$  and  $R^2$  terms to the effective action, it is clear that one should expect the variation of the local geometrical tensors  $A_{ab}$  and  $B_{ab}$ , which arise from them. Thus, the origin of all the local terms in (3.6) and (3.7) is clear, except perhaps for the last term in (3.7), whose origin we now discuss.

Inspection of the  $\varphi$  dependent terms in (4.5) shows that

$$\begin{aligned} \delta F_{ab}|_{\varphi=2\ln\Omega, \psi=0} &= -8 \nabla_c \nabla_d [(\ln \Omega) \delta C_{(a}{}^c{}_{b)}{}^d] \\ &\quad - 4(\ln \Omega) \delta C_{(a}{}^c{}_{b)}{}^d R_{cd} \end{aligned} \quad (4.16)$$

around any conformally flat spacetime with  $C_a{}^c{}_b{}^d = 0$ . Since the Lagrangian density  $\sqrt{-g} F \varphi$  from which this term is derived is conformally invariant under  $g_{ab} = \Omega^2 \eta_{ab} \rightarrow \eta_{ab}$ ,  $\delta C_{(a}{}^c{}_{b)}{}^d$  and the covariant derivatives in (4.16) may be replaced by their values,  $\delta \bar{C}_{(a}{}^c{}_{b)}{}^d$  and  $\partial_c \partial_d$  on the perturbation (3.12) in flat spacetime by a simple overall scaling with  $\Omega^{-2}$ . That is, one may check explicitly that in general

$$\begin{aligned} &2 \nabla_c \nabla_d [\varphi C_{(a}{}^c{}_{b)}{}^d] + \varphi C_{(a}{}^c{}_{b)}{}^d R_{cd} \\ &= \frac{1}{\Omega^2} \{ 2 \bar{\nabla}_c \bar{\nabla}_d [\varphi \bar{C}_{(a}{}^c{}_{b)}{}^d] + \varphi \bar{C}_{(a}{}^c{}_{b)}{}^d \bar{R}_{cd} \} \end{aligned} \quad (4.17)$$

for any two metrics conformally related by  $g_{ab} = \Omega^2 \bar{g}_{ab}$ . Varying this relation keeping  $\varphi$  and  $\Omega$  fixed, with  $\delta g_{ab} = \Omega^2 \delta \bar{g}_{ab}$ , and taking  $\bar{g}_{ab} = \eta_{ab}$  to be flat gives upon mak-

ing use of (4.16),

$$\delta F_{ab}|_{\varphi=2\ln\Omega, \psi=0} = -\frac{8}{\Omega^2} \partial_c \partial_d [(\ln\Omega) \bar{C}_{(a}{}^c{}_{b)}{}^d]. \quad (4.18)$$

Hence the last local term in (3.7) of the variation of the stress tensor expectation value of conformal matter in a conformally flat spacetime may be regarded as a direct result of the anomaly effective action (4.1) and (4.2) in the auxiliary field formalism. Since the  $A_{ab}$  tensor is just proportional to (4.17) with  $\varphi$  replaced by a constant, we also deduce that

$$A^a{}_b = \frac{1}{\Omega^4} \bar{A}^a{}_b, \quad (4.19)$$

upon raising one index, for the two metrics related by (4.12).

Thus, the anomaly action and stress tensor reproduces all the local terms denoted by  $L_{ab}$  in (3.7) in the exact result (3.6). However, it does not reproduce the nonlocal term in the variation (3.6), involving the logarithmic distribution  $K$ . This term comes from the part of the one-loop effective action, which is invariant under conformal transformations  $g_{ab} \rightarrow \Omega^2 g_{ab}$ , and does not contribute to the anomaly or the anomaly action. It may be argued from the addition of the local  $\alpha_R(\mu^2) C_{abcd} C^{abcd}$  term, renormalized at an arbitrary mass scale  $\mu$  that a nonlocal logarithmic term involving  $K$  with the correct coefficient must be present in the effective action in order for the scale  $\mu^2$  to drop out of the low energy dynamics according to (2.8). Whether the nonlocal term is essential or not in perturbations around de Sitter space will be studied in detail in Sec. VII.

Through the anomaly effective action and its auxiliary field description, it is possible to express the variation of the stress tensor in a completely local form, and to gain the advantage of examining a wider class of state variations parameterized by the auxiliary fields as well. By specializing to de Sitter space in the next section it is shown that the independent variation of the metric and auxiliary fields of the anomaly stress tensor (4.3) exhibits specific additional state dependent horizon scale degrees of freedom, which are not present in the classical Einstein theory.

## V. LINEAR RESPONSE OF CONFORMAL MATTER IN DE SITTER SPACETIME

In this section we specialize the discussion to de Sitter spacetime, and derive the linear response equations in the anomaly action auxiliary field formulation, comparing it in detail to the general expression for the linearized stress tensor variation in (3.6).

In FRW coordinates with flat spatial sections (3.4) the scale factor of de Sitter spacetime is

$$a(t) = e^{Ht} = -\frac{1}{H\eta} = \Omega(\eta), \quad \eta \in (-\infty, 0). \quad (5.1)$$

Although the coordinates (3.4) cover only half of the fully

analytically extended de Sitter spacetime, they exhibit the spatial homogeneity and isotropy and standard FRW form explicitly, and are convenient for admitting the standard Fourier mode decomposition in the flat spatial coordinate  $\vec{x}$ . This allows the linear response equations to be expressed as ordinary differential equations in the time variable  $t$  (or  $\eta$ ), for each spatial Fourier mode separately.

In the maximally symmetric de Sitter spacetime,

$$R^a{}_c{}^b{}_d = H^2(g^{ab}g_{cd} - \delta^a{}_d\delta^b{}_c), \quad (5.2a)$$

$$R^a{}_b = 3H^2\delta^a{}_b, \quad H = \frac{\dot{a}}{a}, \quad (5.2b)$$

$$R = 12H^2, \quad \nabla_a R = 0 = \square R, \quad (5.2c)$$

$$E_{ab} = -2^{(3)}H_{ab} = 6H^4 g_{ab}, \quad E = 24H^4, \quad (5.2d)$$

$$C^a{}_c{}^b{}_d = 0, \quad A_{ab} = B_{ab} = 0, \quad F = 0. \quad (5.2e)$$

Because of the  $O(4, 1)$  maximal symmetry, with 10 Killing generators, the conformal vacuum of a conformal field in de Sitter spacetime also is maximally symmetric. The linear response approach requires a self-consistent solution of Eqs. (1.1), around which we perturb the metric and stress tensor together. Hence, the choice of this maximally symmetric state, known as the Bunch-Davies (BD) state is the natural one about which to consider perturbations. From the point of view of a free falling observer this conformal vacuum is in fact a thermally populated state with the temperature  $T_H = H/2\pi$  [3]. In this state  $O(4, 1)$  de Sitter symmetry implies that  $\langle T^a{}_b \rangle$  is also proportional to  $\delta^a{}_b$ . The self-consistent value of the scalar curvature  $R$  including the quantum contribution from  $\langle T^a{}_b \rangle$  is given by the trace equation

$$-R + 4\Lambda = 8\pi G_N b' E = \frac{4\pi G_N b'}{3} R^2, \quad \text{or} \quad (5.3a)$$

$$H^2 = \frac{\Lambda}{3} \left[ 1 - \frac{16\pi G_N b' \Lambda}{3} + \dots \right] \quad (5.3b)$$

in an expansion around the classical de Sitter solution with  $8\pi G_N \Lambda |b'| \ll 1$  [33]. Thus, in this limit the stress tensor source of the semiclassical Einstein's Eqs. (1.1) in the BD state is a small finite correction to the classical cosmological term in the self-consistent de Sitter solution. The fractional size of this correction is  $\hbar G_N \Lambda / c^3 \ll 1$ , if the de Sitter horizon scale  $H^{-1}$  is much greater than the Planck scale  $L_{\text{Pl}} \equiv \sqrt{\frac{\hbar G}{c^3}} = 1.616 \times 10^{-33}$  cm.

Given the smallness of the quantum correction in a self-consistent solution of (1.1), at first sight it would appear unlikely for the quantum fields to have any significant effects at the macroscopic scales of cosmology. However, these effects will become apparent only when both the geometry and quantum state of the matter are not fixed but allowed to vary dynamically. Since a quantum state specified over a macroscopic region of space is highly nonlocal, variations of the state are independent of the magnitude of local curvature variations. Although there

is a very small coupling  $L_{\text{Pl}}^2 H^2 \ll 1$  controlling the loop expansion around the classical de Sitter background, the expansion in this coupling may be nonuniform or even infrared divergent due to state dependent variations on the horizon scale  $H^{-1}$ .

In the auxiliary field effective action approach, information about the quantum state of the underlying conformal matter fields is contained in the particular solution of the auxiliary field Eqs. (4.6). In de Sitter spacetime the conformal differential operator  $\Delta_4$  of (4.7) factorizes,

$$\Delta_4|_{dS} = -\square(-\square + 2H^2), \quad (5.4)$$

and the equation satisfied by  $\varphi$  is inhomogeneous, whereas the equation for  $\psi$  is homogeneous. Thus, it is consistent to set  $\psi = 0$  in order to obtain a de Sitter invariant state. Taking  $\psi \neq 0$  will not lead to a de Sitter invariant stress tensor. In any spatially homogeneous and isotropic state  $\varphi$  can be a function only of time. With  $\varphi = \bar{\varphi}(t)$  one easily finds that the general homogeneous solution of either of Eqs. (4.6) is the linear combination,  $c_0 + c_1 a^{-1} + c_2 a^{-2} + c_3 a^{-3}$ . All of these behaviors except the first lead to de Sitter noninvariant stress tensors, with the constant  $c_0$  dropping out of the stress tensor entirely. Hence, to obtain a de Sitter invariant state all the coefficients of these homogeneous solutions of Eqs. (4.6) must be set to zero. The remaining inhomogeneous solution to (4.6a) is easily found to be  $2Ht$ . If the solutions

$$\bar{\varphi} = 2 \ln \Omega = 2Ht, \quad (5.5a)$$

$$\bar{\psi} = 0 \quad (5.5b)$$

are substituted into the stress tensors (4.4) and (4.5) the result is

$$T^a_b[\bar{\varphi}, \bar{\psi} = 0]_{dS} = b' E^a_b[g^{dS}; \bar{\varphi}] = 6b'H^4 \delta^a_b, \quad (5.6)$$

which is exactly the value of the renormalized stress tensor of a conformally invariant field of any spin in the BD state, found either by direct computation or the methods of [26]. Thus, the solutions (5.5) correspond to the Bunch-Davies state.

With the self-consistent BD de Sitter solution (3.4), (5.1), (5.3), (5.5), and (5.6) one may consider the linear response variation

$$\begin{aligned} \delta R^a_b - \frac{\delta R}{2} \delta^a_b &= 8\pi G_N \delta\{(T^a_b)^{\text{loc}} + (T^a_b)^{\text{anom}}\} \\ &= 8\pi G_N \{-\alpha_R \delta A^a_b - \beta_R \delta B^a_b \\ &\quad + b' \delta E^a_b + b \delta F^a_b\}, \end{aligned} \quad (5.7)$$

where all terms to linear order in  $\delta g_{ab} = h_{ab}$  and in the variations of the auxiliary fields

$$\delta \varphi \equiv \varphi - \bar{\varphi} \equiv \phi, \quad (5.8a)$$

$$\delta \psi \equiv \psi - \bar{\psi} = \psi \quad (5.8b)$$

are to be retained. All indices are raised and lowered by the

background de Sitter metric  $g_{ab}$  at linear order in the perturbations.

For the first local tensor  $A^a_b$  it is simplest to vary the alternative form

$$\begin{aligned} A^a_b &= -4C^a_{\text{cbd}} R^{cd} - 2\square R^a_b + \frac{2}{3}\nabla^a \nabla_b R + \frac{1}{3}\delta^a_b \square R \\ &\quad + 4R^a_c R^c_b - \frac{4}{3}R^a_b R - \delta^a_b R^c_d R^d_c + \frac{1}{3}\delta^a_b R^2, \end{aligned} \quad (5.9)$$

which follows from the Bianchi identities. The variations are simplified by using the symmetry properties of de Sitter space catalogued in (5.2), and by imposing the  $\delta R = 0$  condition discussed in the previous section. It is easily demonstrated that this implies that  $\delta(\nabla^a \nabla_b R) = 0$  and  $\delta(\square R) = 0$ , as well, while  $\delta(\square R^a_b) = \square \delta R^a_b$ . Hence,

$$\delta A^a_b|_{dS, \delta R=0} = 2\left(-\square + \frac{R}{3}\right)\delta R^a_b. \quad (5.10)$$

For the second local tensor  $B^a_b$  one finds

$$\delta B^a_b|_{dS, \delta R=0} = -2R \delta R^a_b. \quad (5.11)$$

The variation of the stress tensors obtained from the anomaly action,  $E^a_b$  and  $F^a_b$  depend on both the variations of the metric and the variations of the auxiliary fields. It is shown in Appendix B that

$$\delta F^t_t|_{dS, \delta R=0} = 2Ht \delta A^t_t - \frac{2}{3} \frac{\vec{\nabla}^2}{a^2} \left[ \partial_t^2 + H \partial_t - \frac{\vec{\nabla}^2}{a^2} \right] \psi. \quad (5.12)$$

From (4.15) one expects the variation of  $E^a_b$  to contain the local geometric term

$$\begin{aligned} -2\delta^{(3)}H^a_b + \frac{1}{9}\delta B^a_b &= -\frac{R}{3}\delta R^a_b - \frac{2R}{9}\delta R^a_b \\ &= -\frac{20H^2}{3}\delta R^a_b. \end{aligned} \quad (5.13)$$

This has been verified by direct calculation using the algebraic manipulation programs MATHEMATICA and MATHTECTOR. The variation of the auxiliary field equations, (4.6) gives

$$\begin{aligned} \delta(\square^2 \varphi) - \frac{R}{6}\delta(\square \varphi) &= \left(\square - \frac{R}{6}\right)\delta(\square \varphi) \\ &= -2(\nabla_a \nabla^b \bar{\varphi})\delta R^a_b, \end{aligned} \quad (5.14a)$$

$$\delta(\square^2 \psi) - \frac{R}{6}\delta(\square \psi) = \left(\square - \frac{R}{6}\right)\square \psi = 0. \quad (5.14b)$$

Both this variation and that of the additional terms in  $\delta E^t_t$  dependent upon  $\delta \varphi = \phi$  can be simplified somewhat by a convenient choice of gauge. Consider

$$\begin{aligned} \delta(\square \varphi) &= \delta(g^{ab} \nabla_a \nabla_b \varphi) \\ &= -h^{ab} \nabla_a \nabla_b \bar{\varphi} - g^{ab} (\delta \Gamma^c_{ab}) \nabla_c \bar{\varphi} + \square \phi, \end{aligned} \quad (5.15)$$

where

$$\delta\Gamma^c_{ab} = \frac{1}{2}(-\nabla^c h_{ab} + \nabla_a h^c_b + \nabla_b h^c_a), \quad (5.16a)$$

$$g^{ab}\delta\Gamma^c_{ab} = \nabla_a h^{ac} - \frac{1}{2}\nabla^c h \quad (5.16b)$$

is the variation of the Christoffel symbol and its trace (with  $h \equiv g^{ab}h_{ab}$ ). The fact that  $\bar{\varphi}$  is a linear function of  $t$  allows the variation to be taken inside one wave operator, i.e.  $\delta(\square^2\varphi) = \square\delta(\square\varphi)$  as in (5.14a), but from (5.15),  $\delta(\square\varphi) \neq \square\phi$  in general. However, the additional terms in (5.15) may be set to zero by making the *gauge choice*,

$$h^{ab}\nabla_a\nabla_b\bar{\varphi} = 0, \quad (5.17a)$$

$$g^{bc}\delta\Gamma^a_{bc}\nabla_a\bar{\varphi} = (\nabla_b h^{ab} - \frac{1}{2}\nabla^a h)\nabla_a\bar{\varphi} = 0. \quad (5.17b)$$

Because of the simple form of  $\bar{\varphi}$  from (5.5), these two gauge conditions are also equivalent to

$$h^{ij}g_{ij} = h^a_a + h_{tt} = 0, \quad (5.18a)$$

$$\nabla^i h_{ti} = g^{ij}\nabla_i h_{tj} = \frac{1}{2}\partial_t h_{tt}, \quad (5.18b)$$

or in terms of the decomposition (3.12),

$$-\vec{\nabla}^2\mathcal{B} = a(\dot{\mathcal{A}} + 6H\mathcal{A}), \quad (5.19a)$$

$$\mathcal{C} = 0. \quad (5.19b)$$

With this choice of gauge the variation in (5.15),  $\delta(\square\varphi) = \square\phi$ , and the auxiliary field Eqs. (5.14) can be written in Fourier space in the simple form

$$\left(\frac{d^2}{dt^2} + 5H\frac{d}{dt} + 6H^2 + k^2e^{-2Ht}\right)v = 0, \quad (5.20a)$$

$$\left(\frac{d^2}{dt^2} + 5H\frac{d}{dt} + 6H^2 + k^2e^{-2Ht}\right)(w - 2h_{tt}) = 0, \quad (5.20b)$$

with

$$H^2v \equiv \left(\frac{d^2}{dt^2} + H\frac{d}{dt} + k^2e^{-2Ht}\right)\psi = \frac{1}{a^2}\left(\frac{d^2}{d\eta^2} + k^2\right)\psi. \quad (5.21a)$$

$$H^2w \equiv \left(\frac{d^2}{dt^2} + H\frac{d}{dt} + k^2e^{-2Ht}\right)\phi = \frac{1}{a^2}\left(\frac{d^2}{d\eta^2} + k^2\right)\phi. \quad (5.21b)$$

In terms of these quantities, with the condition  $\delta R = 0$  and the gauge choice (5.17), the full variation of  $E^t_t$  is

$$\begin{aligned} \delta E^t_t &= -\frac{20H^2}{3}\delta R^t_t - \frac{2}{3}\frac{\vec{\nabla}^2}{a^2} \\ &\quad \times \left[ \left( \partial_t^2 + H\partial_t - \frac{\vec{\nabla}^2}{a^2} \right) \phi - 2H^2 h_{tt} \right] \\ &= -\frac{20H^2}{3}\delta R^t_t - \frac{2H^2}{3a^2}\vec{\nabla}^2(w - 2h_{tt}), \end{aligned} \quad (5.22)$$

while that for  $F^t_t$  in (5.12) becomes

$$\delta F^t_t|_{dS, \delta R=0} = 4Ht(-\square + 4H^2)\delta R^t_t - \frac{2H^2}{3a^2}\vec{\nabla}^2v. \quad (5.23)$$

The gauge choice (5.17) is useful for simplifying the scalar auxiliary field contributions to (5.22) and (5.23), while the local geometric terms involving  $\delta R^t_t$  in both expressions are independent of this gauge choice. In the next section it is shown that in fact the quantities  $w - 2h_{tt}$  and  $v$  are gauge invariant.

All the variations in (5.10), (5.11), (5.22), and (5.23) involve the local geometric variation  $\delta R^t_t$ . Thus, the linear response equation in de Sitter space derived from the auxiliary field effective action takes the form

$$\begin{aligned} \delta R^t_t &= -\frac{\bar{\alpha}_R}{H^2}(-\square + 4H^2)\delta R^t_t + (\bar{\beta}_R + 5\varepsilon')\delta R^t_t \\ &\quad + \frac{\varepsilon t}{H}(-\square + 4H^2)\delta R^t_t + \frac{\varepsilon'}{2a^2}\vec{\nabla}^2(w - 2h_{tt}) \\ &\quad - \frac{\varepsilon}{6a^2}\vec{\nabla}^2v, \end{aligned} \quad (5.24)$$

where the coupling constants have been written in terms of the dimensionless small numbers

$$\bar{\alpha}_R \equiv 16\pi G_N H^2 \alpha_R, \quad (5.25a)$$

$$\bar{\beta}_R \equiv 192\pi G_N H^2 \beta_R, \quad (5.25b)$$

$$\varepsilon \equiv 32\pi G_N H^2 b, \quad (5.25c)$$

$$\varepsilon' \equiv -\frac{32\pi}{3}G_N H^2 b'. \quad (5.25d)$$

Defining the dimensionless variable  $q$  by

$$q \equiv -\frac{2a^2}{H^2}\delta G^t_t = -\frac{2a^2}{H^2}\left(\delta R^t_t - \frac{1}{2}\delta R\right), \quad (5.26)$$

it is straightforward to show that under the condition  $\delta R = 0$  and with the gauge choice (5.17),

$$\begin{aligned} (-\square + 4H^2)\delta R^t_t &= -\left[ \partial_t^2 + 7H\partial_t + 12H^2 - \frac{\vec{\nabla}^2}{a^2} \right] \frac{qH^2}{2a^2} \\ &= -\frac{H^2}{2a^2}\left[ \partial_t^2 + 3H\partial_t + 2H^2 - \frac{\vec{\nabla}^2}{a^2} \right] q. \end{aligned} \quad (5.27)$$

Thus, multiplying (5.24) through by  $-2a^2/H^2$  and Fourier transforming the spatial variables gives

$$\begin{aligned} (1 - \bar{\beta}_R - 5\varepsilon')q &= \varepsilon H t \mathcal{D}q - \bar{\alpha}_R \mathcal{D}q + \varepsilon' \frac{k^2}{H^2}(w - 2h_{tt}) \\ &\quad - \frac{\varepsilon}{3} \frac{k^2}{H^2} v, \end{aligned} \quad (5.28)$$

with

$$\mathcal{D}q \equiv \frac{1}{H^2} \left( \frac{d^2}{dt^2} + 3H \frac{d}{dt} + 2H^2 + k^2 e^{-2Ht} \right) q. \quad (5.29)$$

Equations (5.20) and (5.28), with (5.21), (5.25), and (5.26) are the linear response equations for de Sitter space derived from the anomaly action and stress tensor. Since by assumption  $\bar{\beta}_R$  and  $\varepsilon'$  are very small quantities they may be neglected with respect to unity on the left side of (5.28). Note that in the strict classical limit, where all the small quantities in (5.25) first order in  $\hbar$  are set to zero, the classical constraint equation  $q = 0$  or  $\delta R^t_t = 0$  is recovered from (5.28). This is consistent with the absence of any dynamics in the pure Einstein theory in the sector of scalar perturbations without matter sources.

This form can be compared now to those that are obtained by the method of Refs. [9,10] reviewed in the previous sections. From (1.3), (3.6), and (3.7), upon raising one index with the inverse metric  $g^{ab} = \Omega^{-2} \bar{g}^{ab}$ , one finds

$$\begin{aligned} \left( 1 - \bar{\beta}_R - \frac{5\varepsilon}{3} \right) q &= -\frac{2\Omega^2}{H^2} \delta \langle T^t_t \rangle_R \\ &= \varepsilon H t \mathcal{D}q - \bar{\alpha}_R \mathcal{D}q - \frac{\varepsilon}{2\Omega^2} \int_{\eta_0}^{\eta} d\eta' K(\eta - \eta'; k; \mu) [\Omega^2 \mathcal{D}q]_{\eta'} - \frac{2}{\Omega^2} \delta \langle \bar{T}^t_t \rangle_R. \end{aligned} \quad (5.32)$$

Comparing to (5.28), it is apparent that the linear response Eq. (5.32) based on the methods of [9,10] differs from that based on the anomaly action in four respects:

- (i) The coefficient of  $\delta B^t_t$  in (5.30) and (5.31),  $-\beta_R - \frac{b}{9}$  is replaced by  $-\beta_R + \frac{b'}{9}$  in (5.7) and (5.28). This is an inessential difference, amounting to a different coefficient of the finite local  $R^2$  term in the action, already discussed, and which, as can be seen from (5.25), is any case negligible.
- (ii) The linear response form (5.30) or (5.32) contains the general state dependent term,  $\delta \langle \bar{T}^a_b \rangle_R$ , whereas (5.28) contains the specific  $w - 2h_{tt}$  and  $v$  auxiliary field terms.
- (iii) The auxiliary field action gives rise to additional equations of motion (5.20) for the state dependent terms, absent in the approach of [9,10].
- (iv) Linear response based on the anomaly action, (5.28) lacks the term with the nonlocal kernel  $K$  in (5.30) and (5.32).

In [9,10], the variation of the quantum state of the field was (implicitly) constrained by the metric variation, and not allowed to vary independently. This is equivalent to setting  $v = w - 2h_{tt} = 0$  in the anomaly auxiliary field approach. Since these fields obey the homogeneous Eqs. (5.20), the particular solution in which they both vanish is allowed. However, the local auxiliary field approach allows for a wider class of variations of the state than that of [9,10]. That the  $w - 2h_{tt}$  and  $v$  auxiliary field terms in (5.7) are the result of allowing the state to vary over and above the local metric variation follows from the

$$\begin{aligned} &\delta \left\{ R^t_t - \frac{R}{2} \delta^t_t + \Lambda \delta^t_t \right\} \\ &= 8\pi G_N \left\{ -\frac{b}{\Omega^4} \int_{-\eta_0}^{\eta} d\eta' K(\eta - \eta'; k; \mu^2) \delta \bar{A}^t_t(\eta'; \vec{k}) \right. \\ &\quad \left. + L^t_t + \frac{1}{\Omega^4} \delta \langle \bar{T}^t_t \rangle_R \right\}, \end{aligned} \quad (5.30)$$

with

$$\begin{aligned} L^t_t &= -\alpha_R \delta A^t_t - \left( \beta_R + \frac{b}{9} \right) \delta B^t_t - 2b' \delta^{(3)} H^t_t \\ &\quad - \frac{8b}{\Omega^4} \partial_i \partial_j [(\ln \Omega) \delta C^{ij}]. \end{aligned} \quad (5.31)$$

Upon making use of (4.18), (5.10), (5.27), (5.26), (5.29), (5.25), and (4.19), Eq. (5.30) becomes

fact that the freedom to vary the auxiliary fields amounts to freedom to vary the boundary conditions on the Green's function of  $\Delta_4$  entering the nonlocal form of the trace anomaly effective action [16]. From (5.24) one sees that these terms scale with the conformal factor exactly as expected for the state dependent variation  $\delta \langle \bar{T}^t_t \rangle_R$  in the general form (5.32). In (5.7) these terms take a specific form related to the variation of the scalar auxiliary fields, which obey their own independent equations of motion (5.20b) and (5.20a). The important feature of the auxiliary field formulation of the anomaly effective action (4.1) and (4.2) is that these additional state dependent terms are expressed in terms of local field variations that have their own independent equations of motion (5.14). Although such state dependent variations are certainly allowed on general grounds, the specific form of these variations are determined by the local auxiliary field form of the effective action.

## VI. GAUGE INVARIANT VARIABLES AND ACTION

As already remarked in Sec. III the scalar sector perturbations involve the four functions  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ ,  $\mathcal{E}$ , only two linear combinations of which are gauge invariant. In this section gauge transformations are discussed and a set of gauge invariant variables are constructed. The linear response equations around de Sitter space are written in terms of these gauge invariant variables, and the gauge invariant quadratic action functional corresponding to these equations is given.

The linearized coordinate (gauge) transformation of the metric perturbation is

$$h_{ab} \rightarrow h_{ab} + \nabla_a \xi_b + \nabla_b \xi_a. \quad (6.1)$$

Under this gauge transformation the scalar auxiliary fields transform as

$$\varphi \rightarrow \varphi + \xi^a \nabla_a \varphi, \quad (6.2a)$$

$$\psi \rightarrow \psi + \xi^a \nabla_a \psi. \quad (6.2b)$$

With the definitions

$$\xi^t = aT, \quad (6.3a)$$

$$\xi^i = \eta^{ij} \partial_j L \quad (6.3b)$$

appropriate for scalar perturbations, it is found that the linearized metric functions in the decomposition (3.12) transform as

$$\mathcal{A} \rightarrow \mathcal{A} + \dot{a}T + a\dot{T}, \quad (6.4a)$$

$$\mathcal{B} \rightarrow \mathcal{B} + a\dot{L} - T, \quad (6.4b)$$

$$\mathcal{C} \rightarrow \mathcal{C} + \frac{1}{3} \bar{\nabla}^2 L + \dot{a}T, \quad (6.4c)$$

$$\mathcal{E} \rightarrow \mathcal{E} + L. \quad (6.4d)$$

These are the linearized coordinate transformations possible in the scalar sector [29,30]. By considering the first order variations of Eqs. (6.2) and making use of the notation introduced in Eq. (5.8), one obtains the transformation for both auxiliary field perturbations:

$$\phi \rightarrow \phi + 2\dot{a}T, \quad (6.5a)$$

$$\psi \rightarrow \psi. \quad (6.5b)$$

Thus  $\psi$  is already gauge invariant, as is the quantity  $v$  defined by (5.21a), while  $\phi$  transforms due to the nonzero  $\bar{\varphi} = 2Ht$  in the BD state. It is easily checked from the transformation of the  $\phi$  field (6.5a) that

$$\Phi \equiv \phi + 2\dot{a}\mathcal{B} - 2a\dot{a}\dot{\mathcal{E}} \quad (6.6)$$

is gauge invariant. This is similar to the gauge invariant variable that can be constructed from the scalar field in scalar inflaton models of slow roll inflation [34]. Finally, the metric variables

$$Y_{\mathcal{A}} \equiv \mathcal{A} + \partial_t(a\mathcal{B}) - \partial_t(a^2 \partial_t \mathcal{E}), \quad (6.7a)$$

$$Y_{\mathcal{C}} \equiv \mathcal{C} - \frac{\bar{\nabla}^2 \mathcal{E}}{3} + \dot{a}\mathcal{B} - a\dot{a}\dot{\mathcal{E}} \quad (6.7b)$$

are invariant under the linearized gauge transformations (6.4). These are the gauge invariant Bardeen-Stewart potentials denoted by  $\Phi_{\mathcal{A}}$  and  $\Phi_{\mathcal{C}}$  in Refs. [29,30]

The two quantities  $\delta R$  and  $q$  encountered in the linear response analysis can be written in terms of the metric gauge invariant variables  $Y_{\mathcal{A}}$  and  $Y_{\mathcal{C}}$ . Indeed, using the variation for  $\delta R$  [35] one finds

$$\begin{aligned} \delta R &= -\square h + \nabla_a \nabla_b h^{ab} - R^{ab} h_{ab} \\ &= 6(\ddot{Y}_{\mathcal{C}} - H\dot{Y}_{\mathcal{A}}) + 24H(\dot{Y}_{\mathcal{C}} - HY_{\mathcal{A}}) \\ &\quad - \frac{2}{a^2} \bar{\nabla}^2 Y_{\mathcal{A}} - \frac{4}{a^2} \bar{\nabla}^2 Y_{\mathcal{C}}. \end{aligned} \quad (6.8)$$

Hence condition (3.11),  $\delta R = 0$ , is gauge invariant, and provides one constraint between the two gauge invariant potentials  $Y_{\mathcal{A}}$  and  $Y_{\mathcal{C}}$ . Next one can verify that  $q$ , defined in Eq. (5.26), can be written in the form,

$$q \equiv -\frac{2a^2}{H^2} \delta G^t_t = 12a^2 \left( \frac{1}{H} \dot{Y}_{\mathcal{C}} - Y_{\mathcal{A}} \right) - \frac{4}{H^2} \bar{\nabla}^2 Y_{\mathcal{C}}, \quad (6.9)$$

which is also gauge invariant.

In the linear response equations the quantity  $w - 2h_{tt}$  is also encountered in the gauge defined by the conditions (5.19). The gauge invariant quantity, which reduces to  $w - 2h_{tt}$  in this gauge, is

$$\begin{aligned} H^2 u &= \left[ \frac{d^2}{dt^2} + H \frac{d}{dt} + \frac{k^2}{a^2} \right] \Phi + 6H \frac{dY_{\mathcal{C}}}{dt} - 2H \frac{dY_{\mathcal{A}}}{dt} \\ &\quad - 8H^2 Y_{\mathcal{A}}. \end{aligned} \quad (6.10)$$

This can be seen by evaluating  $Y_{\mathcal{A}}$ ,  $Y_{\mathcal{C}}$ , and  $\Phi$  in the gauge (5.19) and using the definitions (6.6) and (6.7).

Then the linear response equations around de Sitter space can be written in terms of the gauge invariant variables  $u$ ,  $v$ ,  $\delta R$ , and  $q$ . With the condition  $\delta R = 0$  imposed, the equations for  $u$  and  $v$  are

$$\left( \frac{d^2}{dt^2} + 5H \frac{d}{dt} + 6H^2 + \frac{k^2}{a^2} \right) u = 0, \quad (6.11a)$$

$$\left( \frac{d^2}{dt^2} + 5H \frac{d}{dt} + 6H^2 + \frac{k^2}{a^2} \right) v = 0, \quad (6.11b)$$

while that for  $q$  is

$$\begin{aligned} q &= -\frac{2\Omega^2}{H^2} \delta \langle T^t_t \rangle_R \\ &= \varepsilon H t \mathcal{D} q - \bar{\alpha}_R \mathcal{D} q - \frac{\varepsilon}{2\Omega^2} \int_{\eta_0}^{\eta} d\eta' K(\eta - \eta'; k; \mu) [\Omega^2 \mathcal{D} q]_{\eta'} + \varepsilon' \frac{k^2}{H^2} u - \frac{\varepsilon}{3} \frac{k^2}{H^2} v. \end{aligned} \quad (6.12)$$

Here, the differential operator  $\mathcal{D}$  is defined by (5.29). In (6.12) the  $\bar{\beta}_R$ ,  $\varepsilon$ , and  $\varepsilon'$  terms on the left side of (5.28) or (5.32) which differ in the two linear response approaches but which in any case are small compared with unity, have been dropped. The nonlocal term from the analysis of [9,10] has been retained and the specific state dependent terms from the

auxiliary field anomaly action have been written in terms of the gauge invariant variables  $u$  and  $v$ . In arriving at (6.12) we have reconciled the exact but formal calculations of [9,10] with those following from the anomaly effective action by adding the nonlocal term from the former to the specific state dependent variations determined by the auxiliary fields of the latter approach. By the arguments of Sec. III, all the gauge invariant information about scalar perturbations around the self-consistent de Sitter invariant BD state and all components of the curvature tensor variations can be obtained from the solution(s) of (6.11) and (6.12). This is the final gauge invariant form of the linear response equations for the only remaining metric degree of freedom in the tracefree but spatial scalar sector of perturbations of de Sitter spacetime whose solutions we will analyze in the next section.

If the classical Einstein-Hilbert action is added to the effective action of the anomaly, Eqs. (4.1) and (4.2), and both are expanded to quadratic order about the self-consistent BD de Sitter solution given by (5.3), (5.5), and (5.6), the resulting quadratic action can be expressed in terms of the gauge invariant variables  $u$ ,  $v$ ,  $Y_A$  and  $Y_C$  in the relatively simple form,

$$S^{(2)} = \left(1 - 5\varepsilon'\right)S_G + b' \int d^3\vec{x} dt a^3 \left\{ -\frac{H^4 u^2}{2} + \frac{H^2 u \delta R}{3} \right\} + b \int d^3\vec{x} dt a^3 \left\{ -H^4 uv + \frac{H^2 v \delta R}{3} + \frac{4 \ln a}{3a^4} \left[ \vec{\nabla}^2 (Y_A - Y_C) \right]^2 \right\} \quad (6.13)$$

where

$$S_G = \frac{1}{8\pi G_N} \int d^3\vec{x} dt a^3 \left[ -3 \left( \frac{\partial Y_C}{\partial t} \right)^2 + 6H Y_A \frac{\partial Y_C}{\partial t} + \frac{2}{a^2} (\vec{\nabla} Y_A) \cdot (\vec{\nabla} Y_C) + \frac{(\vec{\nabla} Y_C)^2}{a^2} - 3H^2 Y_A^2 \right] \quad (6.14)$$

is the Einstein-Hilbert part of the action, and  $\delta R$  is given by (6.8). Varying (6.13) with respect to  $\Phi$ ,  $\psi$ ,  $Y_A$  and  $Y_C$ , and setting  $\delta R = 0$  yields the gauge invariant linear response Eqs. (6.11) and (6.12), without the non-local contribution involving  $K$ , and without the  $\bar{\alpha}_R$  and  $\bar{\beta}_R$  terms, which would require adding to (6.13) the contributions of the purely local  $C^2$  and  $R^2$  actions also expanded to quartic order in the perturbations about de Sitter space.

## VII. SOLUTIONS TO THE LINEAR RESPONSE EQUATIONS

### A. Solutions of the first kind: $u = v = 0$

Since  $u$  and  $v$  satisfy the same homogeneous equation, cf. (6.11), we first consider the case  $u = v = 0$ . Equation (6.12) is then homogeneous in  $q$  and includes only those variations of  $\langle T^a_b \rangle_R$ , which are driven by the metric fluctuations, i.e. variations of the first kind. The homoge-

neous equation possesses the trivial solution  $q = 0$ . This is the unique classical solution in pure classical gravity, for when  $\hbar = 0$ , all effects of the stress tensor  $\langle T^a_b \rangle_R$ , and its quantum fluctuations are set to zero on the right side of (6.12), and there is no dynamics at all in the scalar sector. The classical equation  $q = 0$  is an equation of constraint.

Next it is straightforward to find nontrivial solutions of (6.12) if  $u = v = 0$ , and the nonlocal term involving  $K$  is neglected. One may then check the consistency of this local approximation by inserting the previous solution into the integral and evaluating the nonlocal term. Since at late times the  $\varepsilon H t \mathcal{D}q$  term dominates over the  $\bar{\alpha}_R \mathcal{D}q$  term, which in any case just shifts the origin of the time variable  $t$ , the  $\bar{\alpha}_R \mathcal{D}q$  term can be neglected as well without loss of generality.

If the nonlocal term in Eq. (6.12) is ignored and  $u = v = 0 = \bar{\alpha}_R = 0$  then the linear response equation reduces to

$$q = \varepsilon H t \mathcal{D}q. \quad (7.1)$$

The Fourier components with  $k \neq 0$  are redshifted away exponentially rapidly by the de Sitter expansion,  $k^2 e^{-2Ht} \rightarrow 0$ . Hence, after a few expansion times this term may be neglected, effectively setting  $k = 0$ . Then,

$$\frac{d^2 q}{dt^2} + 3H \frac{dq}{dt} + 2H^2 q - \frac{qH}{\varepsilon t} \simeq 0. \quad (7.2)$$

This equation describes an exponential growth on the time scale  $H/\sqrt{\varepsilon} \sim M_{\text{Pl}}$ . Substituting the exponential form

$$q(t) = \exp\left(\int^{Ht} W(\tau) d\tau\right), \quad (7.3)$$

one finds

$$\frac{1}{H} \frac{dW}{dt} + W^2 + 3W + 2 - \frac{1}{\varepsilon H t} = 0. \quad (7.4)$$

Ignoring the  $dW/dt$  term, the zeroth order WKB solution is

$$W \simeq -\frac{3}{2} + \frac{1}{2} \sqrt{1 + \frac{4}{\varepsilon H t}}. \quad (7.5)$$

Starting from an initial time  $t_0$  such that  $Ht_0$  is of order unity, the WKB approximation (7.5) is valid up until the turning point where  $W$  vanishes. This turning point is at  $Ht_{\text{max}} = 1/2\varepsilon$ , and this is where  $q$  achieves the maximum value

$$q_{\text{max}} = q(t_{\text{max}}) \simeq \exp\left(\frac{\ln 2}{\varepsilon} - \frac{2Ht_0}{\sqrt{\varepsilon}}\right), \quad \varepsilon \ll 1. \quad (7.6)$$

Here,  $t_0$  is the initial time, which is assumed to be much less than  $(H\varepsilon)^{-1}$ . Thus, if the nonlocal term is ignored, then the perturbation becomes exponentially large in  $\varepsilon^{-1} \sim M_{\text{Pl}}^2/H^2 \gg 1$ . The additional factor of the time  $t$

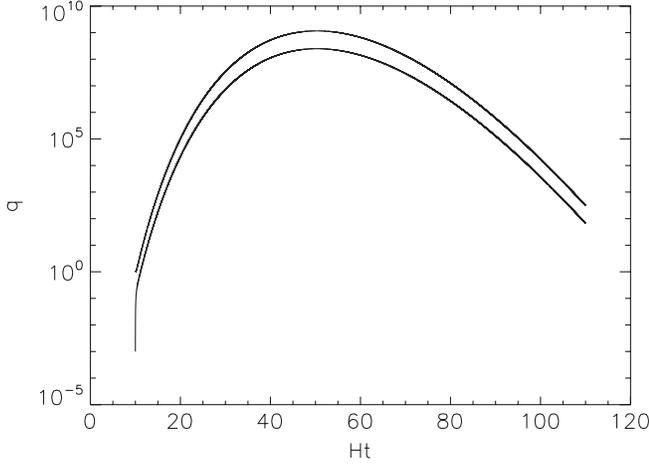


FIG. 2. The plot shows solutions to the linear response Eq. (6.12) in the approximation (7.1) when the nonlocal term is neglected,  $u = v = 0$ ,  $\varepsilon = 0.01$ , and  $k = 100$ . The initial time is  $Ht = 10$ . The solid curve is the solution, which begins with  $q = 1$ ,  $\dot{q} = 0$ , and the dashed curve is the solution, which begins with  $q = 0$ ,  $\dot{q} = 1$ . The solutions grow exponentially according to (7.3) and (7.5) up until  $Ht_{\max} \approx \frac{1}{2\varepsilon} = 50$ .

reduces this growth at later times, and after a time of order  $\varepsilon^{-1}$  eventually turns off the exponential growth.

Equation (7.1) was also solved numerically for  $\varepsilon = 0.01$ , and the results are shown in Fig. 2. Note that the solutions grow rapidly for a long time before reaching a maximum and finally decreasing again. The exponential growth is well described by the analytic WKB approximation of (7.3) and (7.5).

To test whether this general behavior survives in the full linear response equation including the nonlocal term, one can first solve (7.1) and then substitute the solution into the nonlocal term to see if its effects are significant or not. The nonlocal term in Eq. (6.12) can be written in the form

$$I = -\frac{1}{2\Omega^2} \int_{\eta_0}^{\eta} d\eta' K(\eta - \eta'; k; \mu) f(\eta'), \quad (7.7)$$

with  $K$  given by (3.5) and

$$f(\eta) \equiv \varepsilon \Omega^2(\eta) \mathcal{D}q(\eta). \quad (7.8)$$

The behavior of the kernel  $K$  is discussed in Appendix A. For an initial value formulation of linear response, the lower limit of the time integral in (7.7) should be taken as an arbitrary but finite  $\eta_0$ . As discussed in Appendix A, a proper regularization of the nonlocal kernel  $K$ , such as by means of a Pauli-Villars regulator mass, produces transient terms that fall off at late times, and can be neglected. Then Eq. (A13) can be used with the lower limit of the integral replaced by  $\eta_0 \ll \eta$  to obtain the approximate form

$$I \simeq \frac{1}{\Omega^2(\eta)} \int_{\eta_0}^{\eta} d\eta' \text{ci}[k(\eta - \eta')] \frac{df(\eta')}{d\eta'} + \ln\left(\frac{\mu}{k}\right) \frac{f(\eta)}{\Omega^2(\eta)}. \quad (7.9)$$

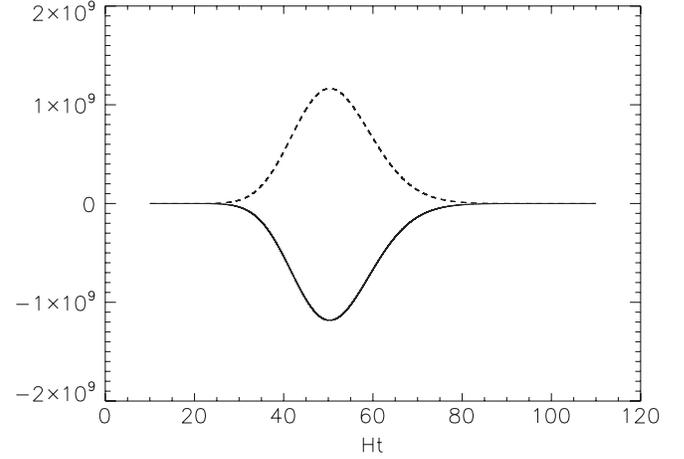


FIG. 3. Shown are the nonlocal and local terms for the case  $k = 100$ . The dashed line corresponds to the local term  $\varepsilon Ht \mathcal{D}q = q$ . The initial conditions for the solution shown for  $q$  are specified at  $Ht = 10$ .

For the solution of (7.1),

$$f(\eta) = \frac{\Omega^2(\eta)q(\eta)}{\ln(\Omega(\eta))} \simeq \frac{-1}{H^2 \eta^2 \ln(-H\eta)} \exp\left(\int^{-\ln(-H\eta)} W(\tau) d\tau\right). \quad (7.10)$$

In Figs. 3 and 4 the nonlocal and local terms on the right-hand side of the linear response equations are shown for the case in which  $q$  is the solution to the local linear response equations displayed in Fig. 2. It is clear from the figures that for the solutions shown the contribution of the local and nonlocal terms on the right-hand side of (6.12) are not only comparable but tend to partially cancel. Hence, we conclude that it is *not* correct to ignore the nonlocal term

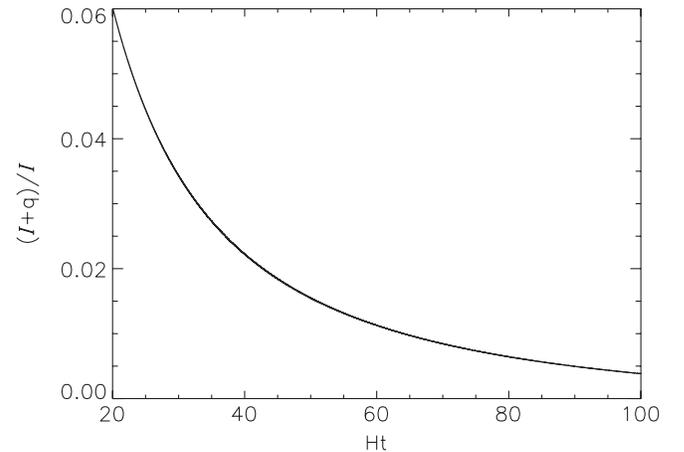


FIG. 4. Shown is the approximate cancellation, which occurs when the local and nonlocal terms,  $\varepsilon Ht \mathcal{D}q = q$  and  $I$ , respectively, are added together for the case  $k = 100$ .

relative to the local terms in (7.1) for these homogeneous solutions for which  $u = v = 0$ .

It is possible to understand the behavior of the nonlocal term analytically for the case that  $f(\eta)$  is a rapidly increasing function, as it is for (7.10). In this case the largest contribution comes from the region near the endpoint of the integral (7.7) as  $\eta' \rightarrow \eta$ . Consider a time  $\eta_1 < \eta$  close to the upper limit of the nonlocal integral, so that the corresponding interval  $\Delta\eta = \eta - \eta_1$  satisfies

$$k\Delta\eta \ll 1. \quad (7.11)$$

If this condition is satisfied then from (A11),  $\text{ci}[k(\eta - \eta')] \approx C + \ln[k(\eta - \eta')]$  and (7.9) becomes

$$\begin{aligned} I &\approx I_1 \\ &= \frac{1}{\Omega^2(\eta)} \left\{ \int_{\eta-\Delta\eta}^{\eta} d\eta' \ln(k(\eta - \eta')) \frac{df(\eta')}{d\eta'} \right. \\ &\quad \left. + \left[ Cf(\eta) - Cf(\eta - \Delta\eta) + f(\eta) \ln\left(\frac{\mu}{k}\right) \right] \right\} \\ &\approx \frac{1}{\Omega^2(\eta)} \left\{ [f(\eta) - f(\eta - \Delta\eta)] [\ln(k\Delta\eta) + C] \right. \\ &\quad \left. + f(\eta) \ln\left(\frac{\mu}{k}\right) \right\} \\ &\approx \left\{ \frac{1}{\Omega^2(\eta)} \Delta\eta \frac{df(\eta)}{d\eta} [\ln(k\Delta\eta) + C] + f(\eta) \ln\left(\frac{\mu}{k}\right) \right\}. \end{aligned} \quad (7.12)$$

Now from (7.3), (7.5), and (7.10), during the period of exponential growth,

$$\frac{df}{d\eta} \approx f \frac{W}{|\eta|} \approx \frac{f}{|\eta| \sqrt{\varepsilon} |\ln(-H\eta)|}. \quad (7.13)$$

Hence, if

$$\Delta\eta \approx |\eta| \sqrt{\varepsilon} |\ln(-H\eta)|, \quad (7.14)$$

for which (7.11) is well satisfied, with  $\varepsilon \ll 1$ , then

$$\begin{aligned} I &\approx I_1 \approx \frac{f(\eta)}{\Omega^2(\eta)} \ln(e^C \mu \Delta\eta) \approx \varepsilon \mathcal{D}q \ln\left(\frac{\mu}{H} e^{C-Ht} \sqrt{\varepsilon Ht}\right) \\ &= -\varepsilon(Ht) \mathcal{D}q + \varepsilon \mathcal{D}q \ln\left(\frac{\mu}{H} e^{C} \sqrt{\varepsilon Ht}\right). \end{aligned} \quad (7.15)$$

The first term of (7.15) exactly cancels the local term,  $\varepsilon(Ht) \mathcal{D}q$  on the right-hand side of the linear response Eq. (6.12), thus giving an analytic approximation, which reproduces the cancellation found numerically, and exhibited in Fig. 4. The second term in (7.15) gives an estimate for the remainder. Examination of the steps leading to (7.15) shows that this approximation is valid as long as  $\mathcal{D}q$  is a rapidly increasing function. In this case, the nonlocal integral (7.7) receives its dominant contribution from the very short time ultraviolet region (7.14) close to its upper endpoint.

From both this analytic estimate and the numerical studies we conclude that for the homogeneous solutions of the linear response Eq. (6.12), the nonlocal term cannot be ignored. Thus, the exponential growth found in (7.3), (7.4), (7.5), and (7.6) by neglecting the nonlocal integral in (6.12) is not reliable. Actually this could have been anticipated by returning to the origin of these terms in Sec. III. The nonlocal term, and the local  $\bar{\alpha}_R \mathcal{D}q$  and  $Ht \mathcal{D}q$  terms, all of which involve  $\mathcal{D}q$  have the same origin in the logarithmic distribution in Fourier space. Equation (2.8) shows that the  $\mathcal{D}q$  dependent terms should be considered together, since only the sum is independent of the arbitrary renormalization scale  $\mu$ . The  $Ht \mathcal{D}q = \ln \Omega \mathcal{D}q$  term arises from the fact that the logarithmic distribution is defined in flat conformal coordinates, in terms of the conformal frequency and momentum  $K^a = (\omega, \vec{k})$ , whereas the frequency and momentum relative to the physical line element (3.4) is  $K^a/\Omega$  (for slowly varying  $\Omega$ ). Thus, one should expect the combination

$$\begin{aligned} &\ln \left[ \frac{-\omega^2 + k^2 - i\varepsilon \text{sgn} \omega}{\mu^2} \right] - 2 \ln \Omega \\ &= \ln \left[ \frac{-\omega^2 + k^2 - i\varepsilon \text{sgn} \omega}{\mu^2 \Omega^2} \right] \end{aligned} \quad (7.16)$$

always to appear together. Heuristically, the local  $\ln \Omega \mathcal{D}q$  term is just the term needed to insert  $\Omega$  in the denominator of the logarithm in (7.16), and convert the conformal frequency and wave number to their values in the physical metric. Since it is known from the linear response equations in flat space that the only nontrivial modes occur at the Planck scale, these same Planck scale solutions should persist in the de Sitter background when all three of the terms involving  $\mathcal{D}q$  in (6.12) are considered together and solved self-consistently. For these homogeneous solutions of (6.12) for which  $u = v = 0$  the nonlocal integral  $I$  cannot be neglected. This is the significance of the integral being dominated by its extreme short distance logarithm near its upper limit. To find these modes accurately one should solve the full nonlocal Eq. (6.12). However, these ultraviolet Planck scale solutions should be excluded from consideration in de Sitter space in any case for the same reason as in flat space, because they lie outside of the range of validity of the low energy semiclassical description of gravity. Therefore, we do not pursue the solutions of the homogeneous Eq. (6.12) further, and turn instead to the inhomogeneous solutions of the second kind.

## B. Solutions of the second kind: $u \neq 0$ or $v \neq 0$

Next, consider the case in which the inhomogeneous gauge invariant state dependent terms  $u$  or  $v$  are different from zero. The general solution of (6.11) for either  $u$  or  $v$  is easily found. In either case the solutions are

$$\begin{aligned} u_{\pm} &= v_{\pm} = \frac{1}{H^2 a^2} \exp\left(\pm \frac{ik}{Ha}\right) e^{i\vec{k}\cdot\vec{x}} \\ &= \eta^2 \exp(\mp ik\eta + i\vec{k}\cdot\vec{x}). \end{aligned} \quad (7.17)$$

Note that this solution and the Eq. (6.11) it satisfies involve the cosmological horizon scale  $H$ , but not the Planck scale. Thus, we term these new solutions *cosmological horizon modes*. Taking e.g.  $u = 0$ , neglecting all the  $\mathcal{D}q$  terms in (6.12), we have

$$q \simeq \frac{\varepsilon}{2} \vec{k}^2 H^2 \eta^2 e^{\mp ik\eta + i\vec{k}\cdot\vec{x}}. \quad (7.18)$$

This corresponds to a linearized stress tensor perturbation of

$$\delta\langle T^{tt} \rangle_R = \frac{H^2 q}{16\pi G_N a^2} = b \frac{k^2 H^2}{a^4} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}}. \quad (7.19)$$

The solutions with  $u$  of the form (7.17) give linearized stress tensor perturbations similar to (7.19) with  $b$  replaced by  $b'$ .

We will now show that it is legitimate to neglect the terms involving  $\mathcal{D}q$  in (6.12) for these solutions. Note that

$$\mathcal{D}q = \pm \frac{2ik}{Ha} q = \pm 2i \frac{k_{\text{phys}}}{H} q, \quad (7.20)$$

where  $k_{\text{phys}} = k/a = H|k\eta|$  is the redshifting physical momentum of the mode. Recalling the definitions (5.25) we have

$$|\bar{\alpha}_R \mathcal{D}q| \sim |\alpha_R G_N H k_{\text{phys}}| |q| \ll |q|, \quad (7.21)$$

provided

$$G_N H^2 \ll 1, \quad \text{or} \quad H \ll M_{\text{Pl}}, \quad (7.22a)$$

$$G_N k_{\text{phys}}^2 \ll 1, \quad \text{or} \quad k_{\text{phys}} \ll M_{\text{Pl}}, \quad (7.22b)$$

and  $|\alpha_R|$  is of order unity. Thus, from (7.21) the  $\bar{\alpha}_R \mathcal{D}q$  term in (6.12) can be neglected.

For the nonlocal term note that (7.20) falls off with time, which has the consequence that the function (7.8) appearing in the integral also falls with time. In this case the nonlocal integral  $I$  in Eq. (7.7) receives contributions over the entire cosmological expansion time scale  $H^{-1}$ , and remains bounded. This is in contrast to the nonlocal term evaluated on the steeply rising function considered in the homogeneous case when  $u = v = 0$ , where (7.7) was dominated by the short time behavior (7.14) near its upper endpoint. For the solution in (7.18),  $I$  may even be calculated analytically using the Pauli-Villars form of the kernel in Eq. (A21), for  $M$  large and setting  $M = \mu$  in the terms that remain after this limit is taken. In this way one finds that all the terms in the nonlocal integral are small and may be neglected, provided the conditions (7.22) are satisfied,

except for one possibly large logarithmic term which is of order

$$b G_N H \frac{k}{a} \ln\left(\frac{\mu^2}{Hk}\right) q. \quad (7.23)$$

Because of (7.16) with  $\Omega = a$ , this combines with the remaining local term involving  $\mathcal{D}q$ , i.e.

$$\varepsilon H t \mathcal{D}q \sim b G_N H \frac{k}{a} \ln(a) q \quad (7.24)$$

to give

$$b G_N H k_{\text{phys}} \ln\left(\frac{\mu^2}{H k_{\text{phys}}}\right) q, \quad (7.25)$$

which is also of negligible magnitude with respect to  $|q|$  provided  $k_{\text{phys}} < \mu \lesssim M_{\text{Pl}}$ ,  $b$  is of order unity, and conditions (7.22) are satisfied. Since these are necessary conditions for the applicability of the semiclassical effective theory (1.1) in the first place, we conclude that all the  $\mathcal{D}q$  terms in the linear response Eq. (6.12) may be neglected compared to (7.18), and (7.18) is a nonvanishing solution to the *full* linear response equations, to a very high degree of accuracy if the conditions (7.22) are satisfied.

The other components of the stress tensor perturbation can be found for this solution from (7.19) by a general tensor decomposition for scalar perturbations analogous to Eqs. (3.12) for the metric. That is, the general perturbation of the stress tensor  $\delta\langle T^{ab} \rangle_R$  in the scalar sector can be expressed in terms of  $\delta\langle T^{tt} \rangle_R$  plus three additional functions. These three functions are determined in terms of  $\delta\langle T^{tt} \rangle_R$  by the conditions of covariant conservation

$$\nabla_b \delta\langle T^{ab} \rangle_R = 0 \quad (7.26)$$

for  $a = t$  and  $a = i$  (two conditions), plus the tracefree condition

$$\delta\langle T^a_a \rangle_R = 0, \quad (7.27)$$

imposed as a result of the  $\delta R = 0$  condition. A straightforward calculation using these conditions and the Christoffel coefficients (3.16) gives then

$$\delta\langle T^{tt} \rangle_R = b H^2 \frac{k^2}{a^4} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}} = -\frac{b H^2}{a^4} \vec{\nabla}_{\vec{x}}^2 e^{\mp ik\eta + i\vec{k}\cdot\vec{x}}, \quad (7.28a)$$

$$\delta\langle T^{ti} \rangle_R = \pm b H^2 \frac{k^i k}{a^5} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}} = \frac{b H^2}{a^4} \frac{\partial^2}{\partial x^i \partial t} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}}, \quad (7.28b)$$

$$\delta\langle T^{ij} \rangle_R = b H^2 \frac{k^i k^j}{a^6} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}} = -\frac{b H^2}{a^6} \frac{\partial^2}{\partial x^i \partial x^j} e^{\mp ik\eta + i\vec{k}\cdot\vec{x}} \quad (7.28c)$$

for the other components of the stress tensor variation for these modes in the flat FRW coordinates of de Sitter space.

Thus, the auxiliary fields of the anomaly action yield the nontrivial gauge invariant solutions (7.28) for the stress tensor and corresponding linearized Ricci tensor perturbations  $\delta R^{ab}$ . Being solutions of (6.11), which itself is independent of the Planck scale, these solutions vary instead on arbitrary scales determined by the wave vector  $\vec{k}$ , and are therefore genuine low energy modes of the semiclassical effective theory. The Newtonian gravitational constant  $G_N$  and the Planck scale enter Eq. (6.12) only through the small coupling parameters  $\varepsilon$  and  $\varepsilon'$  between the auxiliary fields and the metric perturbation  $q$ . Thus, in the limit of either flat space, or arbitrarily weak coupling  $G_N H^2 \rightarrow 0$  these modes decouple from the metric perturbations at linear order.

The result (7.28) would be obtained if the anomaly action (4.1) *alone* is used to generate the linear response Eqs. (5.28), and the nonlocal term is neglected completely. This demonstrates the relevance of the anomaly action for describing physical infrared fluctuations in the effective semiclassical theory of gravity, on macroscopic or cosmological scales unrelated to the Planck scale.

### VIII. COSMOLOGICAL HORIZON MODES

The modes found in Sec. VII B when the gauge invariant variables  $u$  and/or  $v$  are nonzero can vary on any scale rather than the Planck scale, which characterizes modes of the first kind found in Sec. VII A when  $u = v = 0$ . The second set of modes arise naturally from the scalar auxiliary fields, which render the nonlocal anomaly action local, and are therefore implied by the form of the trace anomaly at one-loop order. This is a nontrivial result since these modes appear in the *tracefree* sector of the semiclassical Einstein equations, with  $\delta R = 0$ , and hence cannot be deduced from the form of the trace anomaly itself, but only with the help of the covariant action functional and the additional scalar degrees of freedom, which the anomaly implies. These additional modes, which couple to the scalar sector of metric perturbations in a gauge invariant way, are due to a quantum effect because the auxiliary scalar fields from which they arise are part of the one-loop effective action for conformally invariant quantum fields, rather than a classical action for an inflaton field usually considered in inflationary models [34].

In Sec. V, it was shown that the background solutions  $\bar{\varphi} = 2 \ln \Omega = 2Ht$  and  $\bar{\psi} = 0$  result in the stress-energy tensor for conformally invariant fields in the Bunch-Davies state. The additional modes arising from perturbations of the auxiliary fields  $\varphi$  and  $\psi$  from their background values correspond to changes of state for the underlying conformal quantum fields from their de Sitter invariant BD state. Some physical intuition about these modes and the changes of state of the underlying quantum fields they correspond to may be gleaned from the form of their stress tensor in

(7.28). If one averages this form over the spatial direction of  $\vec{k}$ , a spatially homogeneous, isotropic stress tensor is obtained with pressure  $p = \rho/3$ . Thus, in FRW coordinates this averaging describes incoherent or mixed state thermal perturbations of the stress tensor, which are just those of massless radiation, which redshift with  $a^{-4}$ .

A second interpretation of this second set of modes emerges if we consider *coherent* linear superpositions of different  $\vec{k}$  solutions in different global coordinate systems. In Ref. [16] a class of solutions to the auxiliary field Eqs. (4.6) were found for de Sitter space in *static* rather than cosmological coordinates. In those background field calculations the spacetime geometry was held fixed, and there is no restriction on the state other than that required for renormalization of the stress tensor for the quantum fields. In the linear response equations the solutions to the perturbed Eqs. (6.11) for the auxiliary fields become source terms for the metric fluctuation  $q$  in Eq. (6.12) so their backreaction effects on the spacetime geometry are included to linear order in fluctuations about the de Sitter background geometry. However, in either case the Eqs. (4.6) for the gauge invariant perturbations of the auxiliary fields,  $u$  and  $v$ , being linear, are the *same* equations whether or not the background is varied. Thus, they correspond to the same changes of state as in the background field calculations of [16], and can be compared directly with the fixed de Sitter background solutions found for the auxiliary fields in the static coordinates in [16].

To make this comparison we first need to consider the relationship between the flat FRW and static coordinate systems in de Sitter space. The de Sitter line element in its static form

$$ds^2 = -(1 - H^2 r^2) d\hat{t}^2 + \frac{dr^2}{1 - H^2 r^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (8.1)$$

is identical to (3.4) with the de Sitter scale factor (5.1) if the static time  $\hat{t}$  and radius vector  $\vec{r}$  are related to the flat FRW coordinates  $(t, \vec{x})$  of (3.4) by

$$r = |\vec{x}| e^{Ht} \equiv \rho e^{Ht}, \quad (8.2a)$$

$$\hat{t} = t - \frac{1}{2H} \ln(1 - H^2 \rho^2 e^{2Ht}). \quad (8.2b)$$

The inverse transformations are

$$\rho \equiv |\vec{x}| = \frac{r e^{-Ht}}{\sqrt{1 - H^2 r^2}}, \quad (8.3a)$$

$$t = \hat{t} + \frac{1}{2H} \ln(1 - H^2 r^2). \quad (8.3b)$$

The Jacobian matrix of this  $2 \times 2$  transformation is

$$\begin{pmatrix} \frac{\partial \hat{t}}{\partial t} & \frac{\partial \hat{t}}{\partial \rho} \\ \frac{\partial r}{\partial t} & \frac{\partial r}{\partial \rho} \end{pmatrix} = \begin{pmatrix} 1 & \frac{Hr^2}{\rho(1-H^2r^2)} \\ Hr & \frac{r}{\rho} \end{pmatrix}. \quad (8.4)$$

Using these relations, one may express the action of the differential operators in Eqs. (6.11) in terms of the static coordinates (8.1) instead. We consider the case that the scalar auxiliary fields are functions of  $r$  only and focus attention on  $v$ . A short calculation using (8.4) shows that

$$-\frac{\tilde{\nabla}_{\hat{x}}^2}{a^2} v(r) = -\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dv}{dr} \right), \quad (8.5)$$

while

$$\begin{aligned} H^2 v(r) &= \left( \frac{\partial^2}{\partial \hat{t}^2} + H \frac{\partial}{\partial \hat{t}} - \frac{\tilde{\nabla}_{\hat{x}}^2}{a^2} \right) \psi(r) \\ &= -(1 - H^2 r^2) \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\psi(r)}{dr} \right), \end{aligned} \quad (8.6)$$

operating on functions  $v = v(r)$ , which are functions only of  $r$  and not  $\hat{t}$ .

In [16] the static solutions to

$$\Delta_4 \psi(r) = 0 \quad (8.7)$$

are given. The four solutions are

$$\psi = \frac{1}{Hr} \ln\left(\frac{1 - Hr}{1 + Hr}\right), \quad \ln\left(\frac{1 - Hr}{1 + Hr}\right), \quad 1, \quad \frac{1}{r}. \quad (8.8)$$

The last of these is singular at the origin and so was not considered in [16]. In any case, this solution and the third constant solution to (8.7) give vanishing contribution to  $v$  in (8.6), while from (8.6) the first and second solutions give for  $v$ ,

$$\frac{4}{1 - H^2 r^2}, \quad \frac{4}{Hr} \frac{1 + H^2 r^2}{1 - H^2 r^2}, \quad (8.9)$$

respectively. The second gives a singular contribution to  $v$  and the stress tensor at  $r = 0$ , and we consider only the first solution, taking

$$v = \frac{4}{1 - H^2 r^2}. \quad (8.10)$$

Then, from (5.23), (5.24), (8.5), and (8.10), we obtain

$$\delta\langle T^u \rangle_R = \frac{2bH^2}{3} \frac{\tilde{\nabla}_{\hat{x}}^2}{a^2} v = b \left[ \frac{4H^4}{(1 - H^2 r^2)^2} + \frac{16H^6 r^2}{3(1 - H^2 r^2)^3} \right], \quad (8.11a)$$

$$\begin{aligned} \delta R^t_t &= 8\pi G_N \delta\langle T^t_t \rangle_R \\ &= -\varepsilon H^2 \left[ \frac{1}{(1 - H^2 r^2)^2} + \frac{4H^2 r^2}{3(1 - H^2 r^2)^3} \right]. \end{aligned} \quad (8.11b)$$

The solutions for  $u$  are exactly analogous. This shows that a linear superposition of solutions of the linear response equations of the second kind in static coordinates can lead

to gauge invariant perturbations, which *diverge* on the de Sitter horizon.

To see what (8.11) corresponds to in the static  $(\hat{t}, r)$  coordinates, we use the form of the other components in (7.28) in FRW coordinates

$$\delta\langle T^{ii} \rangle_R = \frac{bH^2}{a^2} \left( \frac{\partial}{\partial \hat{t}} + 2H \right) \frac{\partial v}{\partial x^i}, \quad (8.12a)$$

$$\delta\langle T^{ij} \rangle_R = -\frac{bH^2}{a^2} \frac{\partial^2 v}{\partial x^i \partial x^j}, \quad (8.12b)$$

and the transformation relation for tensors

$$T^{\hat{t}\hat{t}} = \left( \frac{\partial \hat{t}}{\partial t} \right)^2 T^{tt} + 2 \left( \frac{\partial \hat{t}}{\partial t} \right) \left( \frac{\partial \hat{t}}{\partial x^i} \right) T^{ti} + \left( \frac{\partial \hat{t}}{\partial x^i} \right) \left( \frac{\partial \hat{t}}{\partial x^j} \right) T^{ij}, \quad (8.13)$$

with (8.4) and (8.5), to obtain

$$\begin{aligned} \delta\langle T^{\hat{t}\hat{t}} \rangle_R &= -(1 - H^2 r^2) \delta\langle T^{\hat{t}\hat{t}} \rangle_R \\ &= bH^4 \left[ \frac{1}{r} \frac{d}{dr} \left( r \frac{dv}{dr} \right) \right] \\ &\quad + \frac{bH^4}{(1 - H^2 r^2)} \left[ \frac{1}{r} \frac{dv}{dr} - 5H^2 r \frac{dv}{dr} \right] \\ &= -\frac{16bH^4}{(1 - H^2 r^2)^2}. \end{aligned} \quad (8.14)$$

Here, use has been made of the following identities:

$$\frac{\partial \hat{t}}{\partial x^i} = \frac{\partial \hat{t}}{\partial \rho} \frac{\partial \rho}{\partial x^i} = \frac{\partial \hat{t}}{\partial \rho} \frac{x_i}{\rho}, \quad (8.15a)$$

$$\frac{\partial}{\partial \hat{t}} = \frac{\partial r}{\partial \hat{t}} \frac{\partial}{\partial r} + \frac{\partial \hat{t}}{\partial \hat{t}} \frac{\partial}{\partial \hat{t}} = Hr \frac{d}{dr}, \quad (8.15b)$$

$$x_i \partial^i = x^i \partial_i = \rho \frac{\partial}{\partial \rho} = \rho \left( \frac{\partial r}{\partial \rho} \frac{\partial}{\partial r} + \frac{\partial \hat{t}}{\partial \rho} \frac{\partial}{\partial \hat{t}} \right) = r \frac{d}{dr}, \quad (8.15c)$$

$$x_i x_j \partial^i \partial^j = (x^i \partial_i)(x^j \partial_j) - x^i \partial_i = r \frac{d}{dr} \left( r \frac{d}{dr} \right) - r \frac{d}{dr}, \quad (8.15d)$$

valid when operating on functions of  $r$  only. Likewise, we find

$$\delta\langle T^r_r \rangle_R = \delta\langle T^\theta_\theta \rangle_R = \delta\langle T^\phi_\phi \rangle_R = \frac{16bH^4}{3(1 - H^2 r^2)^2}, \quad (8.16)$$

corresponding to a perfect fluid with  $p = \rho/3$ , in *static* coordinates. The contributions proportional to  $b'$  instead of  $b$  from the  $u$  solutions are also of the same form.

The form of the stress tensor (8.14) and (8.16) is the form of a finite temperature fluctuation away from the Hawking-de Sitter temperature  $T_H = H/2\pi$  of the Bunch-Davies state in static coordinates [36]. Since the equation the solutions (8.8) satisfy is the same as (4.6), it follows that there exist linear combinations of the solutions (7.17)

found in Sec. VII B, which give (8.10) and the diverging behavior of the linearized stress tensor on the horizon, corresponding to this global fluctuation in temperature over the volume enclosed by the de Sitter horizon. Note that in static coordinates the stress tensor  $p = \rho/3$  does not require averaging over directions of  $\vec{k}$ , but a particular *coherent* linear superposition over modes (7.17) with different  $\vec{k}$  in order to obtain a particular isotropic but spatially inhomogeneous solution of (5.20a), which selects a preferred origin and corresponding horizon in static coordinates (8.1). The fluctuations in Hawking-de Sitter temperature thus preserve an  $O(3)$  subgroup of the de Sitter isometry group  $O(4, 1)$ .

To follow the diverging behavior (8.14) and (8.16) all the way to the horizon one would clearly require a linear combination of the solutions (7.17) with large Fourier components. However, once  $8\pi G_N$  times the perturbed stress tensor in (8.14) becomes of the same order as the classical background Ricci tensor  $H^2$ , the linear theory breaks down and nonlinear backreaction effects must be taken into account. This occurs at  $r = H^{-1} - \Delta r$  near the horizon, where

$$\Delta r \sim L_{\text{Pl}}, \quad (8.17)$$

or because of the line element (8.1) at the proper distance from the horizon of

$$\ell \sim \sqrt{\frac{L_{\text{Pl}}}{H}} \gg L_{\text{Pl}}. \quad (8.18)$$

Thus, in the background field calculation, one finds that the stress energy is relatively small well inside the horizon. This correspond to a maximum  $k_{\text{phys}} \sim 1/\ell \ll M_{\text{Pl}}$ , where the semiclassical description may still be trusted. At the distance (8.18) from the horizon, the state dependent contribution to the stress-energy tensor becomes comparable to the classical de Sitter background curvature, the linear approximation breaks down, and nonlinear backreaction effects may be expected. This is the same estimate that was obtained in Ref. [37], where a matching of de Sitter space within the horizon onto an exterior Schwarzschild spacetime was discussed.

## IX. SUMMARY AND CONCLUSIONS

In this paper the specific form of the linear response equations in a de Sitter space background for semiclassical gravity coupled to conformally invariant matter fields has been derived with the principal result being Eqs. (6.11) and (6.12). In terms of a decomposition of the perturbations into scalar, vector, and tensor parts with regard to spatial hypersurfaces [29,30], the analysis has been restricted to the scalar sector in (3.12), and it has been shown explicitly that the resulting linear response equations can be written in terms of gauge invariant variables through (6.7), (6.8), (6.9), and (6.10). The quadratic action in terms of these

gauge invariant variables corresponding to the linear response equations is given by Eqs. (6.13) and (6.14). A general condition  $\delta R = 0$  has been utilized, which eliminates conformal or trace solutions that vary on the Planck scale and are thus outside the range of validity of the semiclassical approximation. This condition is approximate in generic spacetimes but is an exact restriction on solutions in the maximally symmetric de Sitter spacetime. Remaining under consideration are only those solutions of the linear response Eqs. (1.3) around de Sitter space, which both have zero four dimensional trace, but which are scalars with respect to the FRW spatial sections. Owing to the freedom to make linearized coordinate transformations of the background de Sitter space, this leads to one and only one remaining gauge invariant scalar metric degree of freedom, whose dynamics is fully described by the  $tt$  component of the perturbed semiclassical Einstein equations in FRW coordinates (3.4), namely, (6.12).

The importance of the trace anomaly and the auxiliary fields used to render it local have been emphasized in our analysis. Among other things, the linear response analysis presented here provides an important test for the auxiliary field action. If the auxiliary field form of the anomaly effective action is used, then the linear response equations consist entirely of local linear partial differential equations with state dependent perturbations resulting from variations in the auxiliary fields. In the linear response equations derived from the exact effective action there are minor differences associated with the values of renormalization parameters, along with two more significant differences. One is that in the anomaly action approach the unspecified state dependent terms allowed in the general approach take quite specific forms in terms of the scalar auxiliary fields, which possess their own dynamical equations of motion. The second significant difference is that the anomaly action takes no account of the conformally invariant term associated with the nonlocal term with the kernel  $K$ .

The solutions to the linear response equations have been investigated in Sec. VII. Our main results can be divided in two parts, depending upon whether the perturbations of the stress tensor are driven by changes in the metric (solutions of the first kind), or additional state dependent perturbations such as those generated by the auxiliary fields of the anomaly action (solutions of the second kind) are considered. For perturbations of the first kind, the local linear response equation has been solved by first ignoring the nonlocal term involving  $K$  and then evaluating the nonlocal term using the solutions so obtained. It was found that the nonlocal term cannot be neglected for this first class of solutions to the linear response equations, and this is closely related to the fact that these modes are fundamentally short distance Planck scale modes, analogous to those already found in flat spacetime, which in any case lie outside the range of validity of the semiclassical effective theory.

The only remaining solution to the exact linear response equations for perturbations of the first kind is the constrained trivial solution  $q = 0$ . Since only this trivial solution to the linear response equation survives when the Planck scale solutions are eliminated, it means that our criterion for the validity of the semiclassical approximation in de Sitter space is satisfied [6]. As originally formulated this validity criterion applies only to gauge invariant fluctuations of the first kind, which depend upon the two-point retarded correlation function for the stress tensor,  $\langle [T^a_b, T^{cd}] \rangle$  and thus probe the response of the geometry to the quantum fluctuations of the stress-energy tensor about its mean value.

For state dependent perturbations of the second kind the situation is much more interesting. A new class of infrared modes has been found in the tracefree but scalar sector, given by Eqs. (7.17), (7.18), (7.19), and (7.28), which are associated with the scalar auxiliary fields of the anomaly effective action. Since these modes vary not on the Planck scale but on arbitrarily large scales, which can be comparable to the cosmological horizon scale in de Sitter space, we have termed them *cosmological horizon modes*. We emphasize that the existence of these scalar modes is a necessary consequence of the one-loop effective action and trace anomaly of quantum conformal matter or radiation. Thus, the conformal anomaly provides scalar degrees of freedom in cosmology without the *ad hoc* introduction of an inflaton.

These scalar cosmological horizon modes satisfy the full linear response equations to leading order in  $G_N H^2$ . In particular, the nonlocal term is always negligible for these modes, provided the semiclassical evolution is begun with modes whose physical wavelength is much larger than the Planck length  $L_{\text{Pl}}$ . Thus, the simple analytic solutions to the equations resulting from the anomaly action (7.17) give excellent approximations to the exact solutions of the linear response equations around de Sitter space to first order in  $H^2 L_{\text{Pl}}^2 \ll 1$ .

Although it is not the full quantum effective action obtained by integrating out the conformal matter or radiation fields exactly, the difference is a nonlocal term that is important for the Planck scale solutions of the first kind only. Ignoring this nonlocal term, the auxiliary field action gives the correct linear response equations valid on low energy scales far removed from the UV Planck scale, i.e. in the infrared limit where the semiclassical theory can be trusted, and for the second class of state dependent solutions. This is consistent with the general arguments that the anomaly action should capture the main infrared or macroscopic quantum effects discussed in earlier work [16,22]. The fact that there is agreement with the local part of the perturbed stress-energy tensor and the infrared relevant terms in a context where the results could be compared to exact calculations from entirely different methods provides a good indication of the usefulness of the anomaly

action as a tool for investigating infrared quantum effects due to conformally invariant fields in other contexts where exact results may not be readily available.

In the FRW coordinates the cosmological horizon modes for fixed comoving wave number  $k$  redshift away at late times like  $a^{-4}$ . The linear homogeneous equation these modes obey can also be considered in the static coordinates (8.1) of de Sitter spacetime. By recognizing the connection between these state dependent solutions to the linear response equations and the solutions to the same auxiliary field equations found from the background field analysis in de Sitter space in static coordinates in Ref. [16], we have found that the cosmological horizon modes can describe state dependent changes on the horizon scale with a large or even divergent stress tensor at  $r = H^{-1}$  (with respect to an arbitrarily chosen origin in de Sitter space). This may be understood as a thermal fluctuation of temperature of the underlying conformal quantum field away from its value in the Bunch-Davies state of  $T_H = H/2\pi$ . If the temperature differs from  $T_H$  by even a small amount, the corresponding stress tensor diverges on the horizon [36].

A diverging stress tensor on the cosmological horizon signals the breakdown of the linear approximation and the necessity to consider nonlinear backreaction effects in the vicinity of the horizon. Thus, our analysis of the cosmological horizon modes and associated stress tensor perturbations in the static coordinates suggests these infrared modes allow the temperature of the causal region interior to the de Sitter horizon to fluctuate away from the BD value, leading to large non-FRW fluctuations of the geometry in the vicinity of the cosmological horizon.

## ACKNOWLEDGMENTS

We, and especially E.M., are very grateful to many enlightening discussions with A. Roura during the course of this work. We would also like to thank B.-L. Hu for helpful conversations, and A. Starobinsky for reading a preliminary draft of the manuscript and giving us his comments. P.R.A. would like to thank J. Hartle for a helpful conversation and access to some of his unpublished calculations. P.R.A. acknowledges financial support from the Spanish Ministerio de Educación y Ciencia and from the National Science Foundation under Grant Nos. PHY-0556292 and PHY-0856050. Numerical computations were performed on the Wake Forest University DEAC Cluster with support from an IBM SUR grant and the Wake Forest University IS Department. Computational results were supported by an IBM SUR grant.

## APPENDIX A: THE LOGARITHMIC DISTRIBUTION IN FLAT SPACE

There are several ways to work with the distribution  $K$  defined by (3.5), which lead to similar results. The distribution has the same form in a FRW background in terms of

the conformal time  $\eta$  as it does in Minkowski spacetime in terms of the usual time coordinate  $t$ . In terms of the latter the distribution is

$$K(t - t'; k; \mu) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \times \ln \left[ \frac{-\omega^2 + k^2 - i\epsilon \operatorname{sgn} \omega}{\mu^2} \right]. \quad (\text{A1})$$

As it stands this distribution is well defined for  $t \neq t'$ , but is undefined for  $t = t'$ . Two different approaches may be considered. In one approach, which owes its origins to Hadamard [38,39], an appropriate strictly local distribution is added to (A1) to extend the definition to functions with nonvanishing support at  $t = t'$ . We shall also describe a second approach, more akin to the Pauli-Villars regulator method, which modifies the distribution (A1) at large frequencies and allows for a smooth limit at  $t = t'$ .

To begin, a mass  $m$  can be introduced into the argument of the logarithm in (A1) by replacing  $k^2$  by  $\omega_k^2 = k^2 + m^2$ . Then differentiating with respect to  $m^2$  gives

$$\begin{aligned} & \frac{\partial}{\partial m^2} K(t - t'; \omega_k; \mu) \\ &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \frac{1}{-\omega^2 + k^2 + m^2 - i\epsilon \operatorname{sgn} \omega} \\ &= \frac{\sin[\omega_k(t - t')]}{\omega_k} \theta(t - t'). \end{aligned} \quad (\text{A2})$$

This may be recognized as the spatial Fourier transform of the usual retarded Green's function for a scalar field with mass  $m$ . For  $t' > t$  the retarded Green's function vanishes and therefore is explicitly causal. For  $t = t'$  the distribution is also defined and vanishes at that point. Note also that after differentiation with respect to  $m^2$ , there is no longer any  $\mu$  dependence in (A2). The reason for this is that the  $\mu$  dependence of (A1) enters only through the purely local contribution

$$\begin{aligned} \mu \frac{d}{d\mu} K(t - t'; \omega_k; \mu) &= - \int_{-\infty}^{\infty} \frac{d\omega}{\pi} e^{-i\omega(t-t')} \\ &= -2\delta(t - t'), \end{aligned} \quad (\text{A3})$$

which is independent of  $k$  or  $m$ .

If (A2) is integrated with respect to  $dm^2 = d\omega_k^2 = 2\omega_k d\omega_k$  at fixed  $k$ , then one obtains

$$K(t - t'; \omega_k; \mu) = K_{\text{bare}}(t - t'; \omega_k) + K_{\text{local}}(t - t'; \mu), \quad (\text{A4})$$

where

$$K_{\text{bare}}(t - t'; \omega_k) = - \frac{2}{t - t'} \cos[\omega_k(t - t')] \theta(t - t'). \quad (\text{A5})$$

This determines the causal  $K$  distribution for all  $t' < t$ , up to a local contribution  $K_{\text{local}}(t - t'; \mu)$ , which is indepen-

dent of  $\omega_k$  and which satisfies (A3). In the form (A4) one can set  $m = 0$  and  $\omega_k = k$  to recover the distribution (3.5).

The need for the local contribution  $K_{\text{local}}$  in (A4) appears when  $K$  is integrated against smooth test functions  $f(t')$ , which are nonvanishing at  $t' = t$ , for in that case,

$$\begin{aligned} & \int_{-\infty}^{t-\xi} dt' K_{\text{bare}}(t - t'; k) f(t') \\ &= -2 \int_{-\infty}^{t-\xi} dt' \frac{\cos[k(t - t')]}{t - t'} f(t') \end{aligned} \quad (\text{A6})$$

diverges logarithmically at its upper endpoint as  $\xi \rightarrow 0$ . This logarithmic divergence can be removed if the  $\mu$  dependent local distribution satisfying (A3) is taken to be

$$K_{\text{local}}(t - t'; \mu) = -2[\ln(\mu\xi) + C]\delta(t - t') \quad (\text{A7})$$

and the limit  $\xi \rightarrow 0$  is taken after summing  $K_{\text{bare}} + K_{\text{local}}$ . Here,  $C = 0.577215\dots$  is Euler's constant, which is a finite additional term added for reason of normalization that will become clear shortly. Then one can define the action of the distribution  $K$  by the one parameter sequence

$$\begin{aligned} K(t - t'; k; \mu) \sim & -2 \left\{ [\ln(\mu\xi) + C] \delta(t - t') \right. \\ & \left. + \frac{1}{t - t'} \cos[k(t - t')] \theta(t - t' - \xi) \right\}_{\xi \rightarrow 0}, \end{aligned} \quad (\text{A8})$$

as  $\xi \rightarrow 0$ . The meaning of the symbol  $\sim$  in (A8) is that acting upon an arbitrary smooth function  $f(t')$ ,

$$\begin{aligned} & \int_{-\infty}^t dt' K(t - t'; k; \mu) f(t') \\ & \equiv -2 \lim_{\xi \rightarrow 0} \left\{ [\ln(\mu\xi) + C] f(t) \right. \\ & \left. + \int_{-\infty}^{t-\xi} \frac{dt'}{t - t'} \cos[k(t - t')] f(t') \right\}, \end{aligned} \quad (\text{A9})$$

that is, the limit  $\xi \rightarrow 0$  is to be taken after the integral over  $t'$  of a member of a class of suitably smooth, test functions  $f(t')$  has been computed, provided the integral in (A9) converges.

We note that the definition (A9) is not unique since an arbitrary local distribution  $K_{\text{local}}(t - t'; \mu)$  with the correct dimensions and the property (A3) could have been added to  $K_{\text{bare}}$ . The definition (A9) is in fact an adaptation of Hadamard's *partie finie* definition of distributions of this kind [38]. Since

$$\begin{aligned} & \frac{d}{d\xi} \int_{-\infty}^{t-\xi} \frac{dt'}{t - t'} \cos[k(t - t')] f(t') \\ &= - \frac{\cos(k\xi)}{\xi} f(t - \xi) = - \frac{f(t)}{\xi} + f'(t) + \mathcal{O}(\xi), \end{aligned} \quad (\text{A10})$$

for test functions which are continuous and differentiable at  $t' = t$ , it is clear that the integral's logarithmic depen-

dence on  $\xi$  is just cancelled by the  $\ln(\mu\xi)$  term, and the limit  $\xi \rightarrow 0$  of the sum in (A9) is finite.

This finite limit in (A9) can be demonstrated directly by making use of the cosine integral function,

$$\text{ci}(z) = \int_{-\infty}^{-z} dx \frac{\cos x}{x} = C + \ln z + \sum_{n=1}^{\infty} (-)^n \frac{z^{2n}}{2n \cdot (2n)!},$$

$$z > 0, \quad (\text{A11})$$

to integrate (A9) by parts. Since

$$\frac{d}{dt'} \text{ci}[\omega_k(t-t')] = -\frac{\cos[\omega_k(t-t')]}{t-t'}, \quad (\text{A12})$$

we have

$$\begin{aligned} & \int_{-\infty}^t dt' K(t-t'; k; \mu) f(t') \\ &= -2 \lim_{\xi \rightarrow 0} \left\{ [\ln(\mu\xi) + C] f(t) - \text{ci}(k\xi) f(t-\xi) \right. \\ & \quad \left. + \int_{-\infty}^{t-\xi} dt' \text{ci}[k(t-t')] \frac{df}{dt'} \right\} \\ &= -2 \int_{-\infty}^t dt' \text{ci}[k(t-t')] \frac{df}{dt'} + 2f(t) \ln\left(\frac{k}{\mu}\right), \end{aligned} \quad (\text{A13})$$

provided that  $f(t')$  is differentiable, and

$$\lim_{t' \rightarrow -\infty} \text{ci}[k(t-t')] f(t') = 0.$$

The definition (A9) of the finite part of the logarithmically divergent distribution is closely related to dimensional regularization and the subtraction of simple poles at  $n=4$ , corresponding to logarithmic divergences of Feynman integrals. As in dimensional regularization and subtraction, the introduction of some large but otherwise arbitrary mass scale  $\mu$  is required for dimensional reasons. The remaining dependence on the arbitrary renormalization scale  $\mu$  in the full linear response equations is removed by (2.8), which amounts to the freedom to shift the  $\mu$  dependence into the unknown coefficient of the local Weyl squared term in the full effective action and the coefficient of the local tensor  $A_{ab}$  in the semiclassical linear response equations.

It is instructive to apply (A9) or (A13) to a simple example of a suitable smooth test function, namely,

$$f(t) = e^{\gamma t}, \quad \gamma > 0, \quad (\text{A14})$$

for which the integral (A9) converges, and may be computed analytically. Indeed,

$$\begin{aligned} & -2 \int_{-\infty}^{t-\xi} \frac{dt'}{t-t'} \cos[\omega_k(t-t')] e^{\gamma t'} \\ &= e^{\gamma t} \text{Ei}[-(\gamma + i\omega_k)\xi] + e^{\gamma t} \text{Ei}[-(\gamma - i\omega_k)\xi], \end{aligned} \quad (\text{A15})$$

where the exponential integral function for complex arguments,  $\text{Ei}(z)$  is defined by the analytic continuation of

$$\text{Ei}(z) = \int_{-\infty}^z dx \frac{e^x}{x} = C + \ln(-z) + \sum_{n=1}^{\infty} \frac{z^n}{n \cdot n!}, \quad (\text{A16})$$

for negative real  $z$ . Using this in (A15) and (A9) for  $\xi \rightarrow 0$  gives

$$\int_{-\infty}^t dt' K(t-t'; k; \mu) e^{\gamma t'} = e^{\gamma t} \ln\left(\frac{\gamma^2 + k^2}{\mu^2}\right), \quad (\text{A17})$$

in which the dependence on Euler's constant has cancelled. The argument of the logarithm is exactly  $K^2/\mu^2$  for the function (A14), in which the frequency  $\omega^2$  is replaced by  $-\gamma^2$ . This is the reason for the introduction of Euler's constant  $C$  into the definition (A8). In terms of a Fourier transform in time, this just amounts to the requirement that the  $\mu$  introduced in (A7) be identical to the  $\mu$  in the Fourier transform (3.5). The fact that (A17) is strictly proportional to  $f(t)$  is a result of the simple Fourier transform of (A14), and will not hold in the case of general  $f(t)$ .

Note that if  $\gamma > 0$ , (A17) possesses a smooth limit as  $k \rightarrow 0$ , and indeed the finiteness of this limit can be demonstrated from (A13) for a class of test functions, satisfying

$$\lim_{t' \rightarrow -\infty} \ln[\mu(t-t')] f(t') = 0, \quad (\text{A18})$$

which excludes  $f(t) = \text{const}$ , i.e.  $\gamma = 0$ , in which case (A17) diverges logarithmically for  $k = 0$ . By repeating the steps leading from (A9) to (A13) one obtains for  $k = 0$ ,

$$\begin{aligned} & \int_{-\infty}^t dt' K(t-t'; k=0; \mu) f(t') \\ &= -2 \int_{-\infty}^t dt' \ln[\mu(t-t')] \frac{df}{dt'} - 2Cf(t) \\ &= -2 \int_{-\infty}^t dt' \ln[\mu e^C(t-t')] \frac{df}{dt'}, \end{aligned} \quad (\text{A19})$$

for differentiable functions satisfying (A18).

In this treatment of the logarithmic distribution  $K$  it has been assumed that the lower limit of the integration over  $t'$  has been extended to  $-\infty$ . However, for the application to linear response with a well-defined initial perturbation at a given finite time  $t_0$ , this definition requires modification. The need for a modification can be seen from the integration by parts needed to pass from (A9) to (A13). If the lower limit of the integral is replaced by  $t_0$ , there will be an additional contribution proportional to  $\text{ci}[k(t-t_0)]f(t_0)$ . From the definition (A11) this vanishes as  $t_0 \rightarrow -\infty$  provided that  $f(t_0)/t_0$  vanishes in this limit. However, with  $t_0$  fixed and finite, this surface contribution diverges logarithmically as  $t \rightarrow t_0$ .

This problem with finite  $t_0$  and indeed the previous divergence as  $t' \rightarrow t$  are encountered because the nonlocal kernel has been treated in isolation from the other components of the linear response equation, which must be solved self-consistently. In the self-consistent solution the Fourier transform has zero support for  $|\omega| \rightarrow \infty$ , and the full linear response equation contains no divergences for finite initial

conditions. Since the very large  $|\omega|$  components will drop out in any case, one may consider the associated distribution

$$K_{\text{PV}}(t - t'; \omega_k; M) \equiv \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \times \ln \left[ \frac{-\omega^2 + k^2 + m^2 - i\epsilon \text{sgn}\omega}{-\omega^2 + k^2 + M^2 - i\epsilon \text{sgn}\omega} \right], \quad (\text{A20})$$

in which the contributions at large  $|\omega|$  are subtracted against an equivalent distribution for a field with large Pauli-Villars mass  $M$ . Comparison with (A8) shows that when  $M \gg \omega_k$  and  $K$  is integrated against smooth test functions  $f(t')$  with bounded Fourier components possessing vanishing support for  $\omega \gg M$ , the result should become indistinguishable from that of the original distribution in (2.19a), with  $M$  playing the role of  $\mu$  in (A8).

The advantage of the explicitly subtracted Pauli-Villars definition (A20) is that its ultra high frequency components are suppressed, with (A20) vanishing for  $|\omega| \gg M$ . This leads to the integral being well defined for all  $t - t'$  with no purely local  $\delta(t - t')$  contribution of the form (A3). Indeed the integration of (A2) between  $m$  and  $M$  now gives

$$\begin{aligned} K_{\text{PV}} \circ f &\equiv \int_{t_0}^t dt' K_{\text{PV}}(t - t'; k; M) f(t') \\ &= -2 \int_{t_0}^t \frac{dt'}{t - t'} \{ \cos[k(t - t')] - \cos[\Omega_k(t - t')] \} f(t') \\ &= -2 \int_{-\infty}^t \frac{dt'}{t - t'} \{ \cos[k(t - t')] - \cos[\Omega_k(t - t')] \} f(t') + 2 \int_{-\infty}^{t_0} \frac{dt'}{t - t'} \{ \cos[k(t - t')] - \cos[\Omega_k(t - t')] \} f(t'), \end{aligned} \quad (\text{A22})$$

which is completely finite at both its upper and lower limits for suitably smooth test functions  $f(t)$ . In the last form one may recognize the  $t_0$  dependence as giving rise to a temporal transient that falls off for  $k(t - t_0) \gg 1$ .

In order to compare the Pauli-Villars definition of the integral kernel (A21) with the finite part prescription (A9), (A21) may be applied to the same test function (A14). Taking first the infinite range ( $t_0 \rightarrow -\infty$ ), one finds

$$\int_{-\infty}^t dt' K_{\text{PV}}(t - t'; k; M) e^{\gamma t'} = e^{\gamma t} \ln \left( \frac{\gamma^2 + k^2}{\gamma^2 + k^2 + M^2} \right). \quad (\text{A23})$$

If  $M^2 \gg \gamma^2 + k^2$  this result can be expanded as

$$\begin{aligned} K_{\text{PV}}(t - t'; \omega_k; M) \\ = -\frac{2}{t - t'} \{ \cos[\omega_k(t - t')] - \cos[\Omega_k(t - t')] \} \theta(t - t'), \end{aligned} \quad (\text{A21})$$

with  $\Omega_k \equiv \sqrt{k^2 + M^2}$ , instead of (A4). Unlike (A4) this function is strictly nonzero only for  $t' < t$ , vanishing identically at  $t' = t$ . It is clear that if  $M$  is large, the second cosine in (A21) oscillates very rapidly and can give a significant contribution only in the short time interval  $t - t' \lesssim 1/M$ . Thus, the Pauli-Villars regulator plays a significant role only in this very narrow interval of time where  $t' \rightarrow t$ , making  $K_{\text{PV}}$  essentially equivalent to  $K_{\text{bare}}$  outside the region  $t - t' \lesssim 1/M$ . Although there is no strictly local  $\delta(t - t')$  contribution in the Pauli-Villars method,  $K_{\text{PV}}$  has a *quasilocal* contribution in the interval  $t - t' \lesssim 1/M$ . In this narrow interval the role of the second contribution in (A21) is crucial since it removes the logarithmic divergence of  $K_{\text{bare}}$ , and the difference,  $K_{\text{PV}}$  vanishes identically at  $t = t'$ . Hence, the integral over test functions  $f(t')$  with nonvanishing support at  $t' = t$  becomes well defined.

For  $K_{\text{PV}}$  there is no problem with an initial value formulation, and we may define the action of the retarded Pauli-Villars kernel on test functions in the finite interval  $[t_0, t]$  by

$$\begin{aligned} \int_{-\infty}^t dt' K_{\text{PV}}(t - t'; k; M) e^{\gamma t'} \\ = e^{\gamma t} \left\{ \ln \left( \frac{\gamma^2 + k^2}{M^2} \right) - \left( \frac{\gamma^2 + k^2}{M^2} \right) + \frac{1}{2} \left( \frac{\gamma^2 + k^2}{M^2} \right)^2 + \dots \right\}. \end{aligned} \quad (\text{A24})$$

Thus, in the limit of large  $M$  the Pauli-Villars regularization of the nonlocal term becomes equivalent to the definition (A9), with the identification  $M \rightarrow \mu$ , at least as far as the logarithmic dependence upon  $M$  is concerned. This shows that the arbitrary mass scale  $\mu$  of (A9) is equivalent to an ultraviolet cutoff scale  $M$ , and the additional terms of the expansion in (A24) may be recognized in position space as an expansion of higher derivative local terms  $(-\square/M^2)^n$  in ascending inverse powers of the UV cutoff  $M^2$ , typical of a low energy expansion of an effective field theory. These finite terms may be neglected for sufficiently

large  $M^2$ , and are absent entirely in the Hadamard finite part regularization (A9).

For finite  $t_0$  the integral formula

$$\begin{aligned} & 2 \int_{-\infty}^{t_0} \frac{dt'}{t-t'} \cos[\omega_k(t-t')] e^{\gamma t'} \\ &= -e^{\gamma t} \{ \text{Ei}[-(\gamma + i\omega_k)(t-t_0)] \\ & \quad + \text{Ei}[-(\gamma - i\omega_k)(t-t_0)] \} \end{aligned} \quad (\text{A25})$$

can be used in (A22) to obtain

$$\begin{aligned} & \int_{t_0}^t dt' K_{\text{PV}}(t-t'; k; M) e^{\gamma t'} \\ &= e^{\gamma t} \left\{ \ln\left(\frac{\gamma^2 + k^2}{M^2}\right) - \text{Ei}[-(\gamma + ik)(t-t_0)] \right. \\ & \quad \left. - \text{Ei}[-(\gamma - ik)(t-t_0)] + 2\text{ci}[M(t-t_0)] \right\} + \dots, \end{aligned} \quad (\text{A26})$$

where all terms that vanish for  $M \gg (\gamma, k)$  have been dropped. This expression is completely finite as  $t \rightarrow t_0$ , and in fact vanishes in that limit. Hence, there are no unphysical divergences at the arbitrary initial time  $t_0$ . This is due to the fact that unlike (A1), which requires a careful limiting procedure (A9) to remove its short time divergence as  $t' \rightarrow t$ , the very high frequency response of (A20) is suppressed at all times, cf. (A23) for  $k \rightarrow \infty$ . Moreover, since  $\text{Ei}(z) \rightarrow e^z/z$  for large  $|z|$ , the last three initial state dependent terms in the bracket of (A26) become

$$\begin{aligned} & -\frac{2e^{-\gamma(t-t_0)}}{t-t_0} \frac{[\gamma \cos(k(t-t_0)) - k \sin(k(t-t_0))]}{\gamma^2 + k^2} \\ & + 2 \frac{\sin(M(t-t_0))}{M(t-t_0)}. \end{aligned} \quad (\text{A27})$$

The first term decays exponentially for  $t-t_0 \gg \gamma^{-1}$ , while the latter falls only as a power times a rapidly oscillating function. Thus, at late times

$$\int_{t_0}^t dt' K_{\text{PV}}(t-t'; k; M) e^{\gamma t'} \rightarrow e^{\gamma t} \ln\left(\frac{\gamma^2 + k^2}{M^2}\right), \quad (\text{A28})$$

for  $M \gg \gamma, k$ ,

which agrees with (A17). Even for constant functions with  $\gamma = 0$ , the initial state dependent transient terms in (A27) fall off linearly with large  $t-t_0$ , and (A28) approaches the Hadamard form (A17) at late times, albeit more slowly. This example shows that up to terms that vanish for large  $M$ , at late times or when operating on continuous and differentiable functions  $f(t)$  with bounded first derivative at all times, the Pauli-Villars regularization of the logarithmic kernel becomes identical to the Hadamard *partie finie* definition, with  $M \leftrightarrow \mu$ .

In a different prescription, one could also consider the Hadamard definition of the logarithmic kernel for finite  $t_0$  by replacing the lower time limit by  $t_0$ , *after* the integration by parts in (A13), i.e.

$$\begin{aligned} & \int_{t_0}^t dt' K(t-t'; k; \mu) f(t') \\ &= -2 \int_{t_0}^t dt' \text{ci}[k(t-t')] \frac{df}{dt'} + 2f(t) \ln\left(\frac{k}{\mu}\right). \end{aligned} \quad (\text{A29})$$

Computing this for the same test function (A14) gives

$$\begin{aligned} & \int_{t_0}^t dt' K(t-t'; k; \mu) e^{\gamma t'} \\ &= e^{\gamma t} \left\{ \ln\left(\frac{\gamma^2 + k^2}{\mu^2}\right) - \text{Ei}[-(\gamma + ik)(t-t_0)] \right. \\ & \quad \left. - \text{Ei}[-(\gamma - ik)(t-t_0)] + 2e^{-\gamma(t-t_0)} \text{ci}[k(t-t_0)] \right\}. \end{aligned} \quad (\text{A30})$$

This is similar to the corresponding Pauli-Villars form (A26), with  $\mu$  replacing  $M$  and no  $1/M$  terms, and is also finite as  $t \rightarrow t_0$ . It differs from (A26) only in the last transient term, which falls exponentially for  $\gamma > 0$  in (A30) but is given by the last term of (A27) in the Pauli-Villars method.

Unlike (A17) for  $k = 0$  and  $\gamma = 0$ , the result (A26) remains finite, becoming

$$\begin{aligned} & \int_{t_0}^t dt' K_{\text{PV}}(t-t'; k=0; M) \\ &= -2 \ln[M(t-t_0)] + 2\text{ci}[M(t-t_0)] - 2C. \end{aligned} \quad (\text{A31})$$

For times  $t-t_0 \gg 1/M$ , the logarithm dominates and (A31) grows without bound. Hence, the logarithmic divergence of (A9) or (A31) for  $k = \gamma = 0$  at all times is replaced by a logarithmic growth in time for the Pauli-Villars definition of the distribution (A20).

To see how the logarithmic kernel behaves in position space, one can return to the unregulated form of the distribution (A5), which is valid either if  $t' < t$ , or if  $K$  is integrated against test functions, which vanish at  $t = t'$ . Its form in position space is found to be

$$K_{\text{bare}}(t, \vec{x}) = \int \frac{d^3 \vec{k}}{(2\pi)^3} e^{i\vec{k} \cdot \vec{x}} \left\{ -\frac{2}{t} \theta(t) \cos(kt) \right\}. \quad (\text{A32})$$

The integral over  $\vec{k}$  can be performed by going to polar coordinates, and integrating over the angles, with the result

$$\begin{aligned}
K_{\text{bare}}(t, \vec{x}) &= -\frac{2\theta(t)}{(2\pi)^2 t} \int_0^\infty k^2 dk \left( \frac{e^{ikr} - e^{-ikr}}{ikr} \right) \cos(kt) \\
&= \frac{1}{4\pi^2} \frac{\theta(t)}{tr} \frac{\partial}{\partial r} \int_0^\infty dk [e^{ik(t+r)} + e^{-ik(t+r)} \\
&\quad + e^{ik(t-r)} + e^{-ik(t-r)}] \\
&= \frac{1}{2\pi} \frac{\theta(t)}{tr} \frac{\partial}{\partial r} [\delta(t+r) + \delta(t-r)], \\
t > 0, \quad r > 0. \tag{A33}
\end{aligned}$$

Because of  $\theta(t)$ , the argument of the first delta function cannot vanish, except possibly at  $t = 0$  and  $r = 0$ , where the bare distribution  $K_{\text{bare}}$  is undefined. Hence, the first delta function in (A33) can be dropped if one integrates against test functions, which vanish at  $t = t'$  and  $\vec{x} = \vec{x}'$ .

Reinstating  $t - t'$  and  $\vec{x} - \vec{x}'$ , the effect of the distribution on a test function,

$$K_{\text{bare}} \circ f \equiv \int dt' \int d^3 \vec{x}' K_{\text{bare}}(t - t', \vec{x} - \vec{x}') f(t', \vec{x}') \tag{A34}$$

can be computed by first going to polar coordinates centered at  $\vec{x}$ ,

$$\vec{x}' = \vec{x} + \hat{n}r, \tag{A35}$$

for some unit vector  $\hat{n}$ , and then introducing the standard retarded and advanced time coordinates with respect to  $(t, \vec{x})$ ,

$$u = t' - t - r, \tag{A36a}$$

$$v = t' - t + r. \tag{A36b}$$

In these coordinates

$$\begin{aligned}
K_{\text{bare}} \circ f &= \int_{-\infty}^\infty dt' \int_0^\infty r^2 dr \int \frac{d\Omega_{\hat{n}}}{2\pi} \frac{\theta(t-t')}{(t-t')r} \frac{\partial}{\partial r} \\
&\quad \times \delta(t' - t + r) f \\
&= \int \frac{d\Omega_{\hat{n}}}{4\pi} \int_{-\infty}^{-\epsilon} du \int_{-\infty}^\infty dv \left[ \frac{u-v}{u+v} \right] \frac{d\delta(v)}{dv} f \\
&= - \int \frac{d\Omega_{\hat{n}}}{4\pi} \int_{-\infty}^{-\epsilon} du \frac{\partial}{\partial v} \left[ \frac{u-v}{u+v} f \right]_{v=0} \\
&= - \int \frac{d\Omega_{\hat{n}}}{4\pi} \int_{-\infty}^{-\epsilon} du \left[ -\frac{2f}{u} + \frac{\partial f}{\partial v} \right]_{v=0}. \tag{A37}
\end{aligned}$$

The integral over  $\int_{-\infty}^{-\epsilon} \frac{du}{u} f$  diverges logarithmically as  $\epsilon \rightarrow 0^+$ . Thus, it may be handled by the same *partie finie* prescription as in (A9). In order to determine the local contribution at  $t' = t$ ,  $\vec{x}' = \vec{x}$ , a similar method can be used as that leading to (A19) for fixed  $k = 0$ . As in (A18) note that one may integrate (A37) by parts and drop the lower limit surface term for functions  $f$  for which

$$\lim_{u \rightarrow -\infty} \{\ln(-\mu u) f(u, v = 0)\} = 0, \tag{A38}$$

which excludes spacetime constant functions. Then we may take the limit  $\epsilon \rightarrow 0^+$  and obtain the finite result [40],

$$K \circ f = - \int \frac{d\Omega_{\hat{n}}}{2\pi} \int_{-\infty}^0 du \left[ \ln(-u/\lambda) \frac{\partial f}{\partial u} + \frac{1}{2} \frac{\partial f}{\partial v} \right]_{v=0}, \tag{A39}$$

for some constant  $\lambda$ . Taking into account a relative normalization factor of  $-1/2\pi$ , (A39) agrees with Eq. (20) of Ref. [8], where the notation  $H_\lambda$  was used for the distribution  $K$ . See also Ref. [41].

In (A39),  $f$  is to be viewed as a function of  $u$ ,  $v$  and  $\hat{n}$  according to

$$f(t', \vec{x}') = f\left(t + \frac{u+v}{2}, \vec{x} + \hat{n} \frac{v-u}{2}\right). \tag{A40}$$

If one takes the previous example of

$$f(t', \vec{x}') = e^{\gamma t'} = e^{\gamma t + \gamma(u+v)/2}, \tag{A41}$$

and substitutes this into (A39), using  $\int_0^\infty dx e^{-x} \ln x = -C$ , one finds  $2e^{\gamma t} \ln(\gamma/\mu)$ , in agreement with (A17) for  $k = 0$ , provided

$$\frac{1}{\lambda} = \frac{\mu}{2} e^{C-(1/2)}. \tag{A42}$$

Hence,

$$\begin{aligned}
K \circ f &\equiv \int dt' \int d^3 \vec{x}' K(t - t'; \vec{x} - \vec{x}'; \mu) f(t', \vec{x}') \\
&= - \int \frac{d\Omega_{\hat{n}}}{4\pi} \int_{-\infty}^0 du \left[ \ln\left(\frac{\mu^2 e^{2C} u^2}{4e}\right) \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v} \right]_{v=0}, \tag{A43}
\end{aligned}$$

for differentiable functions satisfying (A38).

Unlike (A9), the definition (A39) or (A43) is Lorentz invariant. In the Pauli-Villars regulated distribution (A20), which is also Lorentz invariant, the time limits are arbitrary and the lower limit can be taken to be finite. Hence condition (A38) may be relaxed. In the definition (A43) there is no *a priori* justification for replacing the lower limit with a finite time boundary. Hence, although equivalent when evaluated on functions possessing vanishing support as  $|\omega| \rightarrow \infty$ , which is the physically relevant case, the Pauli-Villars subtracted definition (A20) is both well defined over a wider class of test functions, and better suited to a finite initial value formulation of linear response.

## APPENDIX B: THE VARIATIONS IN DE SITTER SPACE

In a conformally flat spacetime with  $C^a_{bcd} = 0$ , and in the conformal vacuum state for which the background auxiliary fields have the values  $\bar{\psi} = 0$ ,  $\bar{\varphi} = 2 \ln \Omega$ , the variation of the  $F^a_b$  tensor (4.5) is given by

$$\begin{aligned}
\delta F^a_b = & -2(\nabla^a \bar{\varphi})(\nabla_b) \square \psi - 2(\nabla^a \psi)(\nabla_b) \square \bar{\varphi} + 2\nabla^c [(\nabla_c \psi)(\nabla^a \nabla_b \bar{\varphi}) + (\nabla_c \bar{\varphi})(\nabla^a \nabla_b \psi)] - \frac{4}{3} \nabla^a \nabla_b [(\nabla^c \bar{\varphi})(\nabla_c \psi)] \\
& + \frac{4}{3} R^a_b (\nabla_c \bar{\varphi})(\nabla^c \psi) - 4R^{ca} [(\nabla_b) \bar{\varphi})(\nabla_c \psi) + (\nabla_b) \psi)(\nabla_c \bar{\varphi})] + \frac{4}{3} R(\nabla_a) \bar{\varphi})(\nabla_b) \psi) - \delta^a_b (\square \bar{\varphi})(\square \psi) \\
& + \frac{1}{3} \delta^a_b \square [(\nabla^c \bar{\varphi})(\nabla_c \psi)] + 2\delta^a_b \left( R^{cd} - \frac{R}{3} g^{cd} \right) (\nabla_c \bar{\varphi})(\nabla_d \psi) - 4\nabla_c \nabla_d (\delta C^{(ac}{}_b{}^d \bar{\varphi}) - 2R_{cd} \bar{\varphi} \delta C^a{}_b{}^c{}_d \\
& - \frac{2}{3} \nabla_a \nabla_b \square \psi - 4R^{ca} (\nabla_b) \nabla_c \psi) + \frac{8}{3} R^a_b \square \psi + \frac{4}{3} R \nabla_a \nabla_b \psi - \frac{2}{3} (\nabla^a R)(\nabla_b) \psi) + \frac{2}{3} \delta^a_b \square^2 \psi + 2\delta^a_b R^{cd} \nabla_c \nabla_d \psi \\
& - \frac{4}{3} \delta^a_b R \square \psi + \frac{1}{3} \delta^a_b (\nabla^c R)(\nabla_c \psi). \tag{B1}
\end{aligned}$$

Specializing to de Sitter spacetime and using the form of  $\bar{\varphi} = 2Ht$  for the BD state in de Sitter space, one finds

$$\square \bar{\varphi} = -6H^2 = -\frac{R}{2}, \tag{B2a}$$

$$(\nabla \bar{\varphi})^2 = g^{ab} (\nabla_a \bar{\varphi})(\nabla_b \bar{\varphi}) = -4H^2 = -\frac{R}{3}. \tag{B2b}$$

Using (B2), the equation of motion (5.14b) for  $\psi$  and the condition (3.11)  $\delta R = 0$  one finds that (B1) becomes

$$\begin{aligned}
\delta F^a_b = & -2(\nabla^a \bar{\varphi})(\nabla_b) \square \psi + 2\nabla^c [(\nabla_c \psi)(\nabla^a \nabla_b \bar{\varphi}) + (\nabla_c \bar{\varphi})(\nabla^a \nabla_b \psi)] - \frac{4}{3} \nabla^a \nabla_b [(\nabla^c \bar{\varphi})(\nabla_c \psi)] \\
& + \frac{1}{3} \delta^a_b \square [(\nabla^c \bar{\varphi})(\nabla_c \psi)] + \frac{R}{6} \delta^a_b (\nabla^c \bar{\varphi})(\nabla_c \psi) - \frac{2R}{3} (\nabla^a \bar{\varphi})(\nabla_b) \psi) + \frac{4R}{9} \delta^a_b \square \psi - 4\nabla_c \nabla_d (\delta C^{(ac}{}_b{}^d \bar{\varphi}) \\
& - \frac{2}{3} \nabla^a \nabla_b \square \psi + \frac{R}{3} \nabla^a \nabla_b \psi. \tag{B3}
\end{aligned}$$

Evaluating the  $a = b = t$  component of this tensor in the flat FRW coordinates (3.4) of de Sitter space, using (2.2a) and (5.10) gives

$$\delta F^t_t = 2Ht \delta A^t_t - \frac{2}{3} \frac{\tilde{\nabla}^2}{a^2} \left[ \partial_t^2 + H \partial_t - \frac{\tilde{\nabla}^2}{a^2} \right] \psi = 4Ht (-\square + 4H^2) \delta R^t_t - \frac{2}{3} \frac{H^2}{a^2} \tilde{\nabla}^2 v, \tag{B4}$$

which is (5.23) of the text.

The variation of the  $E^a_b$  tensor is more involved. Omitting the detailed steps we find in the gauge (5.17),

$$\delta E^t_t = -\frac{20H^2}{3} \delta R^t_t - \frac{2}{3} \frac{\tilde{\nabla}^2}{a^2} \left[ \left( \partial_t^2 + H \partial_t - \frac{\tilde{\nabla}^2}{a^2} \right) \phi - 2H^2 h_{tt} \right] = -\frac{20H^2}{3} \delta R^t_t - \frac{2H^2}{3a^2} \tilde{\nabla}^2 (w - 2h_{tt}). \tag{B5}$$

Here,  $w$  and  $v$  are defined in Eq. (5.21). The algebraic manipulation programs MATHEMATICA and MATHTENSOR were used to help obtain this result.

- 
- |   |  |
|---|--|
| [1] E. Tomboulis, Phys. Lett. <b>70B</b> , 361 (1977).  | [9] G. T. Horowitz and R. M. Wald, Phys. Rev. D <b>21</b> , 1462 (1980); <b>25</b> , 3408 (1982).  |
| [2] B. S. DeWitt, Phys. Rep. <b>19</b> , 295 (1975).  | [10] A. A. Starobinsky, Piz. Eksp. Teor. Fiz. <b>34</b> , 460 (1981) [JETP Lett. <b>34</b> , 438 (1981)].  |
| [3] N. D. Birrell and P. C. W. Davies, <i>Quantum Fields in Curved Space</i> (Cambridge University Press, Cambridge, UK, 1982), and references therein. Sec. 6.3 is particularly relevant to this work. | [11] E. Calzetta and B. L. Hu, Phys. Rev. D <b>35</b> , 495 (1987); R. D. Jordan, Phys. Rev. D <b>36</b> , 3604 (1987).  |
| [4] J. F. Donoghue, Phys. Rev. D <b>50</b> , 3874 (1994).   | [12] A. Campos and E. Verdaguer, Phys. Rev. D <b>49</b> , 1861 (1994); <b>53</b> , 1927 (1996).  |
| [5] E. Mottola, Phys. Rev. D <b>33</b> , 2136 (1986).   | [13] A. G. Riess <i>et al.</i> , Astron. J. <b>116</b> , 1009 (1998); Astrophys. J. <b>607</b> , 665 (2004); S. Perlmutter <i>et al.</i> , Astrophys. J. <b>517</b> , 565 (1999); J. L. Tonry <i>et al.</i> , Astrophys. J. <b>594</b> , 1 (2003). |
| [6] P. R. Anderson, C. Molina-Paris, and E. Mottola, Phys. Rev. D <b>67</b> , 024026 (2003).  | [14] J. A. Isaacson and B. Rogers, Nucl. Phys. <b>B364</b> , 381 (1991); B. Rogers and J. A. Isaacson, Nucl. Phys. <b>B368</b> ,   |
| [7] J. B. Hartle and B. L. Hu, Phys. Rev. D <b>21</b> , 2756 (1980); J. B. Hartle, Phys. Rev. D <b>22</b> , 2091 (1980).  |  |
| [8] G. T. Horowitz, Phys. Rev. D <b>21</b> , 1445 (1980).   |  |

- 415 (1992); see also C. Busch, arXiv:0803.3204.
- [15] G. Pérez-Nadal, A. Roura, and E. Verdaguier, Phys. Rev. D **77**, 124033 (2008); Classical Quantum Gravity **25**, 154013 (2008).
- [16] E. Mottola and R. Vaulin, Phys. Rev. D **74**, 064004 (2006).
- [17] R. J. Riegert, Phys. Lett. B **134**, 56 (1984).
- [18] E. S. Fradkin and A. A. Tseytlin, Phys. Lett. B **134**, 187 (1984).
- [19] I. L. Shapiro and A. G. Jacksenaev, Phys. Lett. B **324**, 286 (1994).
- [20] R. Balbinot, A. Fabbri, and I. L. Shapiro, Phys. Rev. Lett. **83**, 1494 (1999); Nucl. Phys. **B559**, 301 (1999).
- [21] For a recent review see I. Antoniadis, P. O. Mazur, and E. Mottola, New J. Phys. **9**, 11 (2007).
- [22] P. O. Mazur and E. Mottola, Phys. Rev. D **64**, 104022 (2001).
- [23] P. R. Anderson, E. Mottola, and R. Vaulin, Phys. Rev. D **76**, 124028 (2007).
- [24] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
- [25] M. J. Duff, Nucl. Phys. **B125**, 334 (1977).
- [26] L. S. Brown and J. P. Cassidy, Phys. Rev. D **16**, 1712 (1977).
- [27] A. A. Starobinsky, Phys. Lett. B **91**, 99 (1980).
- [28] T. Azuma and S. Wada, Prog. Theor. Phys. **75**, 845 (1986).
- [29] J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).
- [30] J. M. Stewart, Classical Quantum Gravity **7**, 1169 (1990).
- [31] I. Antoniadis and E. Mottola, Phys. Rev. D **45**, 2013 (1992).
- [32] I. Antoniadis, P. O. Mazur, and E. Mottola, Phys. Rev. D **55**, 4756 (1997); **55**, 4770 (1997).
- [33] S. Wada and T. Azuma, Phys. Lett. **132B**, 313 (1983).
- [34] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, Phys. Rep. **215**, 203 (1992).
- [35] N. H. Barth and S. M. Christensen, Phys. Rev. D **28**, 1876 (1983).
- [36] P. R. Anderson, W. A. Hiscock, and D. A. Samuel, Phys. Rev. D **51**, 4337 (1995).
- [37] P. O. Mazur and E. Mottola, arXiv:gr-qc/0109035; Proc. Natl. Acad. Sci. U.S.A. **101**, 9545 (2004).
- [38] J. Hadamard, *Le Problème de Cauchy et les Équations aux Dérivées Partielles Linéaires Hyperboliques* (Hermann et Cie, Paris, 1932).
- [39] L. Schwartz, *Théorie des Distributions*, Actualités Scientifiques et Industrielles: 1122 (Hermann, Paris, 1966).
- [40] P.-D. Methée, Commentarii Mathematici Helvetici **28**, 225 (1954).
- [41] R. D. Jordan, Phys. Rev. D **36**, 3593 (1987).