PHYSICAL REVIEW D **87,** 063010 (2013)

Scrutinizing the cosmological constant problem and a possible resolution

Denis Bernard[1] and André LeClair[1,2]

[1]*Laboratoire de physique théorique, Ecole Normale Supérieure and CNRS, Paris 75005, France*
[2]*Physics Department, Cornell University, Ithaca, New York 14850, USA*
(Received 14 December 2012; published 26 March 2013)

We suggest a new perspective on the cosmological constant problem by scrutinizing its standard formulation. In classical and quantum mechanics without gravity, there is no definition of the zero point of energy. Furthermore, the Casimir effect only measures how the vacuum energy *changes* as one varies a geometric modulus. This leads us to propose that the physical vacuum energy in a Friedmann-Lemaître-Robertson-Walker expanding universe only depends on the time variation of the scale factor $a(t)$. Equivalently, requiring that empty Minkowski space is gravitationally stable is a principle that fixes the ambiguity in the zero-point energy. On the other hand, if there is a meaningful bare cosmological constant, this prescription should be viewed as a fine-tuning. We describe two different choices of vacuum, one of which is consistent with the current universe consisting only of matter and vacuum energy. The resulting vacuum energy density ρ_{vac} is constant in time and approximately $k_c^2 H_0^2$, where k_c is a momentum cutoff and H_0 is the current Hubble constant; for a cutoff close to the Planck scale, values of ρ_{vac} in agreement with astrophysical measurements are obtained. Another choice of vacuum is more relevant to the early universe consisting of only radiation and vacuum energy, and we suggest it as a possible model of inflation.

DOI: 10.1103/PhysRevD.87.063010 PACS numbers: 95.36.+x, 04.62.+v, 98.80.Qc

I. INTRODUCTION

The cosmological constant problem (CCP) is now regarded as a major crisis of modern theoretical physics. For some reviews of the "old" CCP, see Refs. [1–4]. The problem is that simple estimates of the zero-point energy, or vacuum energy, of a single bosonic quantum field yield a huge value (the standard calculation is reviewed below). In the past, this led many theorists to suspect that it was zero, perhaps due to a principle such as supersymmetry. The modern version of the crisis is that astrophysical measurements reveal a very small positive value [5,6]:

$$\rho_\Lambda = 0.7 \times 10^{-29}\,\text{g}\,\text{cm}^{-3} = 2.8 \times 10^{-47}\ \text{Gev}^4/\hbar^3 c^5. \quad (1)$$

This value is smaller than the naive expectation by a factor of 10^{120}. This embarrassing discrepancy suggests a conceptual rather than computational error. The main point of this paper is to question whether the CCP as it is currently stated is actually properly formulated. As we will see, our line of reasoning leads to an estimate of the cosmological constant which is much more reasonable, and of the correct order of magnitude.

Let us begin by ignoring gravity and considering only quantum mechanics in Minkowski space. Wheeler and Feynman once estimated that there is enough zero-point energy in a teacup to boil all the Earth's oceans. This has led to the fantasy of tapping this energy for useful purposes; however, most physicists do not take such proposals very seriously, and in light of the purported seriousness of the CCP, one should wonder why. In fact, there is no principle in quantum mechanics that allows a proper definition of the zero of energy: as in classical mechanics, one can only measure changes in energy; i.e., all energies can be shifted by a constant with no measurable consequences. Similarly, the rules of statistical mechanics tell us that probabilities of configurations are ratios of (conditioned) partition functions, and these are invariant if the partition functions are multiplied by a common factor as induced by a global shift of the energies. Based on his understanding of quantum electrodynamics and his own treatment of the Casimir effect, Schwinger once said [7], "the vacuum is not only the state of minimum energy, it is the state of *zero* energy, zero momentum, zero angular momentum, zero charge, zero whatever." One should not confuse zero-point energy with "vacuum fluctuations" which refer to loop corrections to physical processes: photons do not scatter off the vacuum energy; otherwise they would be unable to traverse the Universe. All of this strongly suggests that it is impossible to harness vacuum energy in order to do work, which in turn calls into question whether it could be a source of gravitation.

The Casimir effect is often correctly cited as proof of the reality of vacuum energy. However it needs to be emphasized that what is actually measured is the *change* of the vacuum energy as one varies a geometric modulus, i.e., how it depends on this modulus, and this is unaffected by an arbitrary shift of the zero of energy. The classic experiment is to measure the force between two plates as one changes their separation; the modulus in question here is the distance ℓ between the plates and the force depends on how the vacuum energy varies with this separation. The Casimir force $F(\ell)$ is minus the derivative of the electrodynamic vacuum energy $E_{\text{vac}}(\ell)$ between the two plates, $F(\ell) = -dE_{\text{vac}}(\ell)/d\ell$. An arbitrary shift of the vacuum energy by a constant that is independent of ℓ does not affect the measurement. For the electromagnetic field, with two polarizations, the well-known result is that the energy density between the plates is $\rho_{\rm vac}^{\rm cas} = -\pi^2/720\ell^4$. Note that this is an attractive force; as we will see, in the cosmic context a repulsive force requires an overabundance of fermions. It is also clear that the Casimir effect is an infrared phenomenon that has nothing to do with Planck scale physics. Our cosmological proposal will actually involve a mixing of the infrared (IR) and ultraviolet (UV).

For reasons that will be clear, let us illustrate the above remark on the Casimir effect with another version of it: the vacuum energy in the finite size geometry of a higher dimensional cylinder. Namely, consider a massless quantum bosonic field on a Euclidean space-time geometry of $S^1 \otimes R^3$ where the circumference of the circle S^1 is β. Viewing the compact direction as spatial, the momenta in that direction are quantized and the vacuum energy density is

$$\rho_{\rm vac}^{\rm cyl} = \frac{1}{2\beta}\sum_{n\in\mathbb{Z}}\int\frac{d^2\mathbf{k}}{(2\pi)^2}\sqrt{\mathbf{k}^2 + (2\pi n/\beta)^2} = -\beta^{-4}\pi^{3/2}\Gamma(-3/2)\zeta(-3) + \text{const.} \quad (2)$$

Due to the different boundary conditions in the periodic versus finite size directions, $\rho_{\rm vac}^{\rm cas}(\ell) = 2\rho_{\rm vac}^{\rm cyl}(\beta = 2\ell)$, where the overall factor of 2 is because of the two photon polarizations. The above integral is divergent; however, if one is only interested in its β dependence, it can be regularized using the Riemann zeta function giving the above expression. Note that the (infinite) constant that has been discarded in the regularization is actually at the origin of the CCP. What is measurable is the β dependence. One way to convince oneself that this regularization is meaningful is to view the compactified direction as Euclidean time, where now $\beta = 1/T$ is an inverse temperature. The quantity $\rho_{\rm vac}^{\rm cyl}$ is now the free energy density of a single scalar field, and standard quantum statistical mechanics gives the convergent expression which is just the standard black-body formula:

$$\rho_{\rm vac}^{\rm cyl} = \frac{1}{\beta}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\log(1 - e^{-\beta k}) = -\beta^{-4}\frac{\zeta(4)}{2\pi^{3/2}\Gamma(3/2)} = -\frac{\pi^2}{90}T^4. \quad (3)$$

The two above expressions, (2) and (3), are equal due to a nontrivial functional identity satisfied by the ζ function: $\xi(\nu) = \xi(1-\nu)$ where $\xi(\nu) = \pi^{-\nu/2}\Gamma(\nu/2)\zeta(\nu)$. (See for instance the appendix in Ref. [8] in this context.) The comparison of Eqs. (2) and (3) strongly manifests the arbitrariness of the zero-point energy: whereas there is a divergent constant in (2), from the point of view of quantum statistical mechanics, the expression (3) is actually convergent. Either way of viewing the problem allows a shift of $\rho_{\rm vac}^{\rm cyl}$ by an arbitrary constant with no measurable consequences. For instance, such a shift would not affect thermodynamic quantities like the entropy or density since they are derivatives of the free energy; the only thing that is measurable is the β dependence.

We now include gravity in the above discussion. Before stating the basic hypotheses of our study, we begin with general motivating remarks. All forms of energy should be considered as possible sources of gravitation, including the vacuum energy. However, if one accepts the above arguments that the zero of energy is not absolutely definable in quantum mechanics, and that only the dependence of the vacuum energy on geometric moduli including the space-time metric is physically measurable, it then remains unspecified how to incorporate vacuum energy as a source of gravity. One needs an additional principle to fix the ambiguity.

The above observations on the Casimir energy were instrumental toward formulating such a principle, as we now describe. The cosmological Friedmann-Lemaître-Robertson-Walker (FLRW) metric has no modulus corresponding to a finite size analogous to β; however, it does have a time dependent scale factor $a(t)$:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2 d\mathbf{x}\cdot d\mathbf{x}. \quad (4)$$

(We assume the spatial curvature $k = 0$, as shown by recent astrophysical measurements.) When $a(t)$ is constant in time, the FLRW metric is just the Minkowski spacetime metric. This leads us to propose that the dependence of the vacuum energy on the time variation of $a(t)$ is all that is physically meaningful, in analogy with the β dependence of $\rho_{\rm vac}^{\rm cyl}$. This idea is stated as a principle below, in terms of the stability of empty Minkowski space, and is at the foundation of our conclusions.

Let us quickly review the standard cosmology. The Einstein equations are

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = 8\pi G T_{\mu\nu}, \quad (5)$$

where G is Newton's constant. The stress energy tensor $T_{\mu\nu} = \text{diag}(\rho, p, p, p)$ where ρ is the energy density and p the pressure. The nonzero elements of the Ricci tensor are $R_{00} = -3\ddot{a}/a$, $R_{ij} = (2\dot{a}^2 + a\ddot{a})\delta_{ij}$, and the Ricci scalar is $\mathcal{R} = g^{\mu\nu}R_{\mu\nu} = 6((\dot{a}/a)^2 + \ddot{a}/a)$, where overdots refer to time derivatives. The temporal and spatial Einstein equations (5) for the FLRW metric are then the Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho, \quad (6)$$

$$\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{\ddot{a}}{a} = -8\pi G p. \quad (7)$$

Taking a time derivative of the first equation and using the second, one obtains

$$\dot{\rho} = -3\left(\frac{\dot{a}}{a}\right)(\rho + p), \tag{8}$$

which expresses the usual energy conservation. The above three equations are thus not functionally independent, the reason being that Bianchi identities relate the two Friedmann equations to the energy conservation equation (8). The total energy density is usually assumed to consist of a mixture of three noninteracting fluids, radiation, matter, and dark energy, $\rho = \rho_{\rm rad} + \rho_m + \rho_\Lambda$, each of which satisfies Eq. (8) separately, with $p = w\rho$ for $w = 1/3$, 0, and -1 respectively. Then, Eq. (8) consistently implies $\dot{\rho}_\Lambda = 0$. The energy density is related to the classical cosmological constant as $\Lambda = 8\pi G \rho_\Lambda$.

In this paper we will assume that dark energy comes entirely from vacuum energy, $\rho_\Lambda = \rho_{\rm vac}$. The vacuum energy $\rho_{\rm vac}$ is a quantum expectation value,

$$\rho_{\rm vac} = \langle \mathcal{H} \rangle = \langle {\rm vac}|\mathcal{H}|{\rm vac}\rangle, \tag{9}$$

where $\mathcal{H}$ a quantum operator corresponding to the energy density, which is usually associated with T_{00}.

Apart from the ambiguity of the zero-point energy, several other points should be emphasized. We will be studying the semiclassical Einstein equations, where on the right-hand side we include the contribution of vacuum energy $\langle T_{\mu\nu}\rangle = \langle {\rm vac}|T_{\mu\nu}|{\rm vac}\rangle$ for some choice of vacuum state $|{\rm vac}\rangle$. Given the very low energy scale of expansion in the current universe, and the weakness of cosmological gravitational fields, it is very reasonable to assume that there is no need to quantize the gravitational field itself in the present epoch. One hypothesis of the standard formulation of the CCP is that the vacuum stress tensor is proportional to the metric [1]. In an expanding universe, the Hamiltonian is effectively time dependent, and there is not necessarily a unique choice of $|{\rm vac}\rangle$, and in contrast to flat Minkowski space, no Lorentz symmetry argument [9] enforces that $\langle T_{\mu\nu}\rangle \propto g_{\mu\nu}$. One needs extra information that characterizes $|{\rm vac}\rangle$. This implies that $\langle T_{\mu\nu}\rangle$ is not universal since it depends on $|{\rm vac}\rangle$, and thus, for example, cannot always be expressed in terms of purely geometric properties with no reference to the data of $|{\rm vac}\rangle$. One mathematical consistency condition is $D^\mu\langle T_{\mu\nu}\rangle = 0$ if the various components of the total energy are separately conserved, where D^μ is the covariant derivative, which is the statement of energy conservation. However this may not follow from $D^\mu T_{\mu\nu} = 0$ since $|{\rm vac}\rangle$ may be time dependent. Also, $\langle T_{\mu\nu}\rangle$ is not necessarily expressed in terms of manifestly covariantly conserved tensors such as $G_{\mu\nu}$; again because it depends on $|{\rm vac}\rangle$. In fact, the only convariantly conserved geometric tensor that is second order in time derivatives is $G_{\mu\nu}$, and if $\langle T_{\mu\nu}\rangle \propto G_{\mu\nu}$, this would just amount to a renormalization of Newton's constant G.

The second point is that if one includes $\langle T_{\mu\nu}\rangle$ as a source in Einstein's equations, then since it depends on $a(t)$ and its time derivatives, doing so can be thought of as studying the backreaction of this vacuum energy on the geometry. The resulting equations must be solved self-consistently and there is no guarantee that there is a solution consistent with energy conservation.

Having made these preliminary observations, let us state all of the hypotheses that this work is based upon, which specify either the vacuum states or the nature of their stress tensor. They are the following:

(i) As a criterion to identify possible vacuum states $|{\rm vac}\rangle$, we look for preferred quantization schemes such that $|{\rm vac}\rangle$ is an eigenstate of the Hamiltonian at all times, which implies there is no particle production.

(ii) We calculate a bare $\rho_{\rm vac,0}$ from the Hamiltonian, i.e., $\rho_{\rm vac,0} = \langle {\rm vac}|\mathcal{H}|{\rm vac}\rangle$ where $\mathcal{H}$ is the quantum Hamiltonian energy density operator. The calculation is regularized with a sharp cutoff k_c in momentum space in order to make contact with the usual statement of the CCP.

(iii) We propose the following principle which prescribes how to define a physical $\rho_{\rm vac}$ from $\rho_{\rm vac,0}$: Minkowski space that is empty of matter and radiation should be stable, that is, static. This requires that the physical $\rho_{\rm vac}$ equal zero when $a(t)$ is constant in time. This leads to a $\rho_{\rm vac}$ that depends on $a(t)$ and its derivatives, and also the cutoff.

(iv) Given this $\rho_{\rm vac}$, we assume the components of the vacuum stress energy tensor have the form of a cosmological constant:

$$\langle T_{\mu\nu}\rangle = -\rho_{\rm vac} g_{\mu\nu}. \tag{10}$$

We provide some support for this hypothesis in Sec. III, where we compare our calculation with manifestly covariant calculations performed in the past [10]. We are going to check the consistency of this assumption in the next point (v).

(v) We include $\langle T_{\mu\nu}\rangle$ in Einstein's equations and solve them self-consistently, assuming that vacuum energy and other forms of energy are separately conserved. In other words we study the consistency of the backreaction of the vacuum energy on the geometry. The consistency condition is $\dot{\rho}_{\rm vac} = 0$, which is equivalent to $D^\mu\langle T_{\mu\nu}\rangle = 0$. There is no guarantee there is such a solution since $\rho_{\rm vac}$ depends on $a(t)$ and its derivatives.

Certainly one may question the validity of these assumptions. However in our opinion, they are rather conservative in that they do not invoke symmetries, particles, or other, perhaps higher dimensional structures, which are not yet known to exist. The purpose of this paper is to work out the logical consequences of these modest hypotheses. Our main findings are the following:

(i) If there is a cutoff in momentum space k_c, then by dimensional analysis the vacuum energy density has symbolically the "adiabatic" expansion (up to constants):

$$\rho_{\rm vac,0} = k_c^4 + k_c^2 \widehat{R} + \widehat{R}^2 + \cdots, \quad (11)$$

where $\widehat{R}$ is related to the curvature and is a linear combination of $(\dot{a}/a)^2$ and $\ddot{a}/a$, depending on the choice of $|{\rm vac}\rangle$. The principle of the stability of empty Minkowski space (iii) leads us to discard the k_c^4 term, but not the other terms since they depend on time derivatives of $a(t)$. The vacuum energy is now viewed as a *low energy* phenomenon, like the Casimir effect. Other regularization schemes, based for example on point splitting [10], insist on a finite $\rho_{\rm vac}$ and thus discard the first two terms. According to our principles, the second term must be kept since it depends dynamically on the geometry. In the current universe $\widehat{R}$ it is approximately on the order of H_0^2, where H_0 is the Hubble constant, and if the cutoff k_c is on the order of the Planck energy, then the resulting value of $\rho_{\rm vac}$ is the right order of magnitude in comparison with the measured value (1), namely $\rho_{\rm vac} \approx (k_p H_0)^2 = 3.2 \times 10^{-46}\ {\rm Gev}^4$, using for H_0 the present value of the Hubble constant. The $\widehat{R}^2$ is much too small to explain the measured value. We emphasize that our $\rho_{\rm vac}$ is not simply proportional to H^2 [see Eqs. (24) and (35) below] and is in fact constant in time for the self-consistent solutions that we find. There is nothing special about H_0 here, since $\rho_{\rm vac}$ is constant in time; we are simply evaluating it at the present time which involves H_0. A practical point of view is that astrophysical observations are telling us that the k_c^4 term should be shifted away. More importantly, it remains to determine whether the term that we do keep, $k_c^2\widehat{R}$, has physical consequences in agreement with observations, which is the main purpose of our study. Is shifting away the k_c^4 term a fine-tuning? Let us address this in the context of the Aronowitt-Deser-Misner and Abbott-Deser framework [11,12]. In the latter work it was shown that in classical general relativity, once the value of the cosmological constant Λ is fixed, there is a unique choice of energy that is conserved; i.e., there is no more freedom to shift it. For asymptotically flat spacetimes, this energy was proven to be positive and only zero for Minkowski space [13,14]. Its "main importance is that it is related to the stability of Minkowski space as the ground state of general relativity," to quote Ref. [14]. For de Sitter space, there are similar statements, though with some restrictions [12]. Let us suppose that for some as yet unknown physical reason, perhaps due to quantum gravity, there is a meaningful bare cosmological constant Λ_0. Then the $\rho_{\rm vac,0}$ that we calculate leads to an effective cosmological constant $\Lambda_{\rm eff} = \Lambda_0 + 8\pi G \rho_{\rm vac,0}$ in semiclassical gravity, and this is what is actually measured. Our prescription (iii) amounts to setting $\Lambda_{\rm eff}$ to zero in Minkowski space where $\dot{a}$ vanishes, by adjusting Λ_0 or the manifestly constant part of $\rho_{\rm vac,0}$. This is a fine-tuning if Λ_0 is unambiguously defined. Our work has nothing more to say about this issue; rather, the main point of this work is to study the effects and consistency of what remains, i.e., the $\Lambda_{\rm eff} \sim k_c^2 \widehat{R}$ term which does not vanish in an expanding geometry. One may argue that, as in any field theory, the parameters in the effective Lagrangian must ultimately be chosen to match experiments. As we will see, under certain conditions, this term can mimic a cosmological constant in a way that is consistent with the current era of cosmology.

The analogy with the Casimir effect is clear both mathematically [compare Eqs. (2) and (21)] and physically. In the Casimir effect, as one pulls apart the plates in a controlled manner in an experiment, this induces a measurable force. In cosmology the analog of the growing separation of the plates is the expansion itself, which induces an acceleration; the complication is that the effect of this backreaction must be solved self-consistently, as we will do. Note that whereas the Casimir force is attractive, to describe the positive accelerated expansion of the Universe, one needs a positive $\rho_{\rm vac}$, which as we will explain, requires an overabundance of fermions.

(ii) For a universe consisting of only matter and vacuum energy, such as the present universe, there is a choice of $|{\rm vac}\rangle$ with the above $\rho_{\rm vac}$ that leads to a consistent solution if a specific relation between k_c and the Newton constant G is satisfied. By "consistent," we mean $\dot{\rho}_{\rm vac} = 0$. Our solution for $a(t)$ is consistent with present day astrophysical observations if one ignores the very small radiation component. In fact, as we will show, our solution $a(t)$, Eq. (30) below, once one matches the integration constants to their present measured values, is identical to the standard ΛCDM model of the present universe, i.e., a universe consisting only of a cosmological constant plus cold dark matter, and is thus not ruled out by observations up to fairly large redshift $z < 1000$; it certainly agrees with supernova observations at low z. To our knowledge this choice for $|{\rm vac}\rangle$ has not been considered before. Below, we also remark on the cosmic coincidence problem in light of our result. We speculate that the relation between G and k_c suggests that gravity itself arises from quantum fluctuations, and we provide an argument that "derives" gravity from quantum mechanics.

(iii) For a universe consisting of only radiation and vacuum energy, there is another *different* choice of vacuum, $|\widehat{\text{vac}}\rangle$, that also has a consistent solution, again only for a certain relation between k_c and G. This vacuum has been studied before and is referred to as the conformal vacuum in the literature. We suggest that this solution possibly describes inflation, without invoking an inflaton field, and speculate on a scenario to resolve the "graceful exit" problem. We also argue that when $H = \dot{a}/a$ is large, the first Friedmann equation sets the scale $H \sim k_c$, which is the right order of magnitude if k_c is the Planck scale.

It is worthwhile comparing our model with other, similar proposals. Based on "wave-function of the universe" arguments [15], or simply dimensional analysis [16], it was proposed that $\rho_{\text{vac}} \sim (k_p/d_H(t))^2$, where d_H is a dimensionful scale factor related to the cosmological horizon; roughly $d_H \sim a(t)t$. In the present universe $d_H(t_0) \approx t_0 \approx 1/H_0$, so this ρ_{vac} is also of the right magnitude. The problem with it is that it is time dependent, and ruled out by observations. Different arguments based on unimodular gravity [17,18] also led to the proposal that $\Lambda \sim 1/d_H^2$. The work that is closest to ours is by Maggiore and collaborators [19]. Our approach differs from all the above in that our ρ_{vac} is constant in time, in agreement with observations.

The next two sections simply describe these two choices of vacua and analyze the self-consistency of the backreaction. Our analysis is done using an adiabatic expansion. In the conclusion, we further discuss our results.

II. VACUUM ENERGY PLUS MATTER

A. Choice of vacuum and its energy density

We first review the standard version of the cosmological constant problem. Since a free quantum field is an infinite collection of harmonic oscillators for each wave vector $\mathbf{k}$, we first review simple quantum mechanical versions in order to point out the difference between bosons and fermions. Canonical quantization of a bosonic mode [24] of frequency ω yields to a pair of creation and annihilation operators, a, $a^\dagger$, with $[a, a^\dagger] = 1$, and a Hamiltonian $H = \frac{\omega}{2}(aa^\dagger + a^\dagger a) = \omega(a^\dagger a + \frac{1}{2})$. The boson zero-point energy is thus identified as $\omega/2$. For fermions, the zero-point energy has the opposite sign. Fermionic canonical quantization [25] yields to Grassmanian operators b, $b^\dagger$, with $\{b, b^\dagger\} = 1$, $b^2 = b^{\dagger 2} = 0$, and a Hamiltonian $H = \frac{\omega}{2}(b^\dagger b - bb^\dagger) = \omega(b^\dagger b - \frac{1}{2})$. The fermion zero-point energy is $-\omega/2$.

In a free relativistic quantum field theory with particles of mass m in three spatial dimensions, the above applies with $\omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m^2}$, where $\mathbf{k}$ is a three-dimensional wave vector. Thus the zero-point vacuum energy density is

$$\rho_{\text{vac}} = \frac{N_b - N_f}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \sqrt{\mathbf{k}^2 + m^2}, \tag{12}$$

where $N_{b,f}$ is the number of bosonic, fermionic particle species. Regularizing the integral with an ultraviolet cutoff k_c much larger than m, one finds $\rho_{\text{vac}} \approx (N_b - N_f)k_c^4/16\pi^2$. If k_c is taken to be the Planck energy k_p, then $k_c^4/16\pi^2 = 10^{75}$ Gev^4. The modern version of the cosmological constant problem is the fact that this is too large by a factor of 10^{122} in comparison with the measured value. One should also note that in the above calculation a positive value for ρ_{vac} requires more bosons than fermions, contrary to the currently known particle content of the Standard Model.

As explained in the introduction, we are interested in the vacuum energy of a free quantum field in the nonstatic FLRW background spacetime geometry. For simplicity we consider a single scalar field, with action [26]

$$S = \int dt d^3\mathbf{x} \sqrt{|g|} \left(-\frac{1}{2} g^{\mu\nu} \partial_\mu \Phi \partial_\nu \Phi - \frac{m^2}{2} \Phi^2 \right). \tag{13}$$

In order to simplify the explicit time dependence of the action, and thereby simplify the quantization procedure, we define a new field χ as $\Phi = \chi/a^{3/2}$. Then the action (13), after an integration by parts, becomes

$$S = \int dt d^3\mathbf{x} \frac{1}{2} (\partial_t \chi \partial_t \chi - \frac{1}{a^2} \vec{\nabla}\chi \cdot \vec{\nabla}\chi - m^2\chi^2 + \mathcal{A}(t)\chi^2), \tag{14}$$

where

$$\mathcal{A} \equiv \frac{3}{4}\left(\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{\ddot{a}}{a} \right). \tag{15}$$

The advantage of quantizing χ rather than Φ is that most of the time dependence is now in $\mathcal{A}$, so that there is no spurious time dependence in the canonical momenta, etc. The field can be expanded in modes:

$$\chi = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} (a_{\mathbf{k}} u_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^\dagger u_{\mathbf{k}}^* e^{-i\mathbf{k}\cdot\mathbf{x}}), \tag{16}$$

where the $a_{\mathbf{k}}$'s satisfy canonical commutation relations $[a_{\mathbf{k}}, a_{\mathbf{k}'}^\dagger] = \delta(\mathbf{k} - \mathbf{k}')$. The function $u_{\mathbf{k}}$ is time dependent and required to satisfy

$$(\partial_t^2 + \omega_{\mathbf{k}}^2) u_{\mathbf{k}} = 0, \qquad \omega_{\mathbf{k}}^2 \equiv (\mathbf{k}/a)^2 + m^2 - \mathcal{A}. \tag{17}$$

The solution is the formal expression

$$u_{\mathbf{k}} = \frac{1}{\sqrt{2W}} \exp\left(i \int^t W(s) ds \right), \tag{18}$$

where W satisfies the differential equation:

$$W^2 = \omega_{\mathbf{k}}^2 + \frac{3}{4}(\dot{W}/W)^2 - \frac{1}{2}\ddot{W}/W. \tag{19}$$

Let us assume that the time dependence is slowly varying, i.e., we make an adiabatic expansion. The above equation can be solved iteratively, where, to lowest approximation, W is the above expression with W replaced by $\omega_{\mathbf{k}}$ on the right-hand side of the differential equation. In other words, the "adiabatic condition" is $\dot{\omega}_{\mathbf{k}}/\omega_{\mathbf{k}} \ll \omega_{\mathbf{k}}$.

As we now explain, it appears one has to distinguish between massive versus massless particles. Consider for instance the term proportional to $(\dot{\omega}_{\mathbf{k}}/\omega_{\mathbf{k}})^2 = (\dot{a}/a)^2 \times (\mathbf{k}^2/(\mathbf{k}^2 + m^2 a^2))^2$. When $m = 0$ this gives a term which modifies $\mathcal{A}$, as does the $\ddot{\omega}/\omega$ term. The adiabatic condition is simply $\dot{a}/a \ll k$. The result is that the additional two terms on the right-hand side of Eq. (19) (with $W = \omega_{\mathbf{k}}$) give $W^2 = (\mathbf{k}/a)^2 - \mathcal{R}/6$; i.e., $\mathcal{A}$ is converted to $\mathcal{R}/6$. This dependence on the Ricci scalar $\mathcal{R}$ can be derived more directly using conformal time, as in the next section. In this section we will only be considering cosmological matter plus vacuum energy. When $m \neq 0$, the additional terms do not simply convert $\mathcal{A}$ to $\mathcal{R}/6$. In order to implement an adiabatic expansion in this case, we consider the opposite limit of m large. One way to perhaps justify this is as follows. We will ultimately be interested in this vacuum energy in the presence of a nonzero density of real matter. In cosmology, "matter" refers to nonrelativistic particles, and formally, the nonrelativistic limit corresponds to $m \to \infty$, e.g., $\sqrt{\mathbf{k}^2 + m^2} \approx m^2 + \mathbf{k}^2/2m$. More importantly, matter is defined as having zero pressure. For a relativistic fluid, the contribution of each mode $\mathbf{k}$ to the pressure is $p = n_{\mathbf{k}} \mathbf{k}^2/3\omega_{\mathbf{k}}$ where $n_{\mathbf{k}}$ is the density. One then sees that zero pressure corresponds to $m \to \infty$. Here the adiabatic condition is $\dot{a}/a \ll (\mathbf{k}^2 + m^2)^{3/2}/\mathbf{k}^2$ which is automatically satisfied in this limit. In the limit $m \to \infty$, the additional terms on the right-hand side of Eq. (19) actually vanish. Thus, to lowest order we simply take $W = \omega_{\mathbf{k}}$, and to this order $\dot{u}_{\mathbf{k}} = i\omega_{\mathbf{k}} u_{\mathbf{k}}$. As we will show in the next subsection, for a pressureless fluid this has a self-consistent backreaction.

With this choice of $u_{\mathbf{k}}$, and to lowest order in the adiabatic expansion, the Hamiltonian takes the standard form:

$$H = \frac{1}{2}\int d^3\mathbf{x}\left(\dot{\chi}^2 + \frac{1}{a^2}(\vec{\nabla}\chi)^2 + (m^2 - \mathcal{A})\chi^2\right) = \frac{1}{2}\int d^3\mathbf{k}\,\omega_{\mathbf{k}}(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a_{\mathbf{k}} a^\dagger_{\mathbf{k}}). \tag{20}$$

Importantly, there are no $a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}}$ terms, which implies the vacuum $|\text{vac}\rangle$ defined by $a_{\mathbf{k}}|\text{vac}\rangle = 0$ is an eigenstate of H for all times, i.e., there is no particle production, again to lowest order in the adiabatic expansion. By the translational invariance of the vacuum, for the bare vacuum energy we finally have

$$\rho_{\text{vac},0} = \frac{1}{V}\langle H\rangle = \frac{1}{2}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\sqrt{\mathbf{k}^2 + m^2 - \mathcal{A}}, \tag{21}$$

where V is the volume and we have used $\delta_{\mathbf{k}}(0) = V/(2\pi)^3$. In obtaining the above expression we have properly scaled by redshift factors: $V \to a^3 V$, the cutoff was scaled to k_c/a, and we made the change of variables $\mathbf{k} \to a\mathbf{k}$. Comparing the above equation with Eq. (2), the analogy with the Casimir effect is clear.

Introducing an ultraviolet cutoff k_c as before, one finds

$$\rho_{\text{vac},0} \approx \frac{k_c^4}{16\pi^2}\left[1 + \alpha + \frac{\alpha^2}{8}(1 + 2\log(\alpha/4))\right], \tag{22}$$

where $\alpha \equiv (m^2 - \mathcal{A})/k_c^2$ is assumed to be small and positive. Assuming that masses m are all much smaller than the cutoff, we approximate the above expression as

$$\rho_{\text{vac},0} \approx \frac{k_c^4}{16\pi^2}\left[1 + \frac{m^2}{k_c^2} - \frac{\mathcal{A}}{k_c^2} + \frac{\mathcal{A}^2}{8k_c^4}\right], \tag{23}$$

where we have neglected the logarithmic contribution. It should also be noticed that the $\mathcal{A}^2$ term is beyond the lowest order in the adiabatic expansion.

Now we apply the principle (iii) of the introduction. In empty Minkowski space, by definition $\dot{a} = \ddot{a} = 0$ and ρ_{vac} must be zero; otherwise empty Minkowski spacetime would not be static due to gravity. Thus, $\rho_{\text{vac},0}$ must be regularized to a physical ρ_{vac} by subtracting the first two constant terms in brackets:

$$\rho_{\text{vac}} \approx \Delta N\left[\frac{k_c^2}{16\pi^2}\mathcal{A} - \frac{1}{128\pi^2}\mathcal{A}^2\right], \tag{24}$$

such that $\rho_{\text{vac}} = 0$ when $\dot{a} = \ddot{a} = 0$. Above, we have included multiple species $\Delta N = N_f - N_b$ where $N_{f,b}$ are the numbers of species of fermions and bosons. It is important to observe that in the cylindrical version of the Casimir effect, Eq. (2), the analog of the first term above is proportional to $\zeta(-2)k_c/\beta^3 = 0$, so that there is no analog of it in the Casimir effect.

Before proceeding, let us first check that the above expression gives reasonable values. In the present universe, $\dot{a}/a = H_0 = 1.5 \times 10^{-42}$ Gev is the Hubble constant, and $(\dot{a}/a)^2 \approx \ddot{a}/a$. If k_c is taken to be the Planck energy k_p, then $\rho_{\text{vac}} \sim (k_p H_0)^2 = 3.2 \times 10^{-46}$ Gev^4, which at least is in the ballpark. Fortunately there are more fermions than bosons in the Standard Model of particle physics so that the above expression is positive. Each quark/antiquark has two spin states, and comes in three chromodynamic colors. The electron/positron has two spin states, whereas a neutrino has one. For three flavor generations, this gives $N_f = 90$. Each massless gauge boson has two polarizations, eight for QCD, and four for the electroweak theory, which leads to $N_f - N_b = 60$ including the four Higgs fields before spontaneous electroweak symmetry breaking and the two graviton polarizations. Incidentally, for one generation $N_f = N_b = 30$, so in order for the cosmological constant to be positive, one needs at least two generations. The measured value of the vacuum energy can be accounted for with a cutoff about an order of magnitude below the Planck energy [28], $k_c \approx 3 \times 10^{18}$ Gev. We have ignored interactions which modify the value of ρ_{vac}; however, we expect that they do not drastically change our results. One should also bear in mind that the sharp cutoff k_c is meant to represent a crossover from the effective theory valid at energy scale well below k_c to that (including gravity) valid above k_c.

B. Consistent backreaction

Let us suppose that the only form of vacuum energy is $\rho_{\rm vac}$ of the last section, Eq. (24), and that $a(t)$ is varying slowly enough in time that the $\mathcal{A}^2$ term can be neglected. Define the dimensionless constant

$$g = \frac{3\Delta N}{8\pi} G k_c^2 \tag{25}$$

such that

$$\rho_{\rm vac} = \frac{g}{6\pi G}\mathcal{A}. \tag{26}$$

Including $\rho_{\rm vac}$ in the total ρ, the first Friedmann equation can be written as

$$\left(1-\frac{g}{3}\right)\left(\frac{\dot a}{a}\right)^2 - \frac{2g}{3}\frac{\ddot a}{a} = \frac{8\pi G}{3}(\rho_m + \rho_{\rm rad}). \tag{27}$$

We emphasize that we have not modified the Friedmann equation; the extra terms on the left-hand side come from $\rho_{\rm vac}$ which were originally on the right-hand side of the first Friedmann equation.

As we now argue, there is only a consistent solution when $g = 1$. First consider the case where there is no radiation or matter. Then Eq. (27) implies $\ddot a/a = (3-g)(\dot a/a)^2/2g$. First, note that this implies a constant expansion, i.e., de Sitter space, only if $g = 1$. Second, the pressure can then be found from Eq. (7):

$$p_{\rm vac} = -\frac{1}{g}\rho_{\rm vac}. \tag{28}$$

Thus, the equation of state parameter $w = -1/g$ when there is only $\rho_{\rm vac}$. However energy conservation requires $\dot\rho_{\rm vac} = 0$, which requires $p_{\rm vac} = -\rho_{\rm vac}$, i.e., $g = 1$. The solution is $a(t) \propto e^{Ht}$ for an arbitrary constant H, and $\rho_{\rm vac}$ is independent of time, as a cosmological constant must be.

What is not immediately obvious is that a consistent solution can also be found when matter is included, again when $g = 1$. At the current time t_0, as usual, define the critical density $\rho_c = 3H_0^2/8\pi G$ where H_0 is the Hubble constant. The matter and radiation densities scale as $\rho_m/\rho_c = \Omega_m/a^3$ and $\rho_{\rm rad}/\rho_c = \Omega_{\rm rad}/a^4$, where Ω_m, $\Omega_{\rm rad}$ are the current fractions of the critical density at time $t = t_0$ where $a(t_0) = 1$. The first Friedmann equation becomes, when $g = 1$,

$$\frac{2}{3H_0^2}\left[\left(\frac{\dot a}{a}\right)^2 - \frac{\ddot a}{a}\right] = \frac{\Omega_m}{a^3} + \frac{\Omega_{\rm rad}}{a^4}. \tag{29}$$

When $\Omega_{\rm rad} = 0$, the general solution, up to a time translation, is

$$a(t) = \left(\frac{\Omega_m}{\mu}\right)^{1/3}[\sinh(3\sqrt{\mu}H_0 t/2)]^{2/3}. \tag{30}$$

The constant μ is fixed by $a(t_0) = 1$. One can check that $\rho_{\rm vac}$ is indeed constant in time:

$$\frac{\rho_{\rm vac}}{\rho_c} = \mu, \tag{31}$$

which implies that $\mu + \Omega_m = 1$, i.e., μ is just $\Omega_{\rm vac}$ However, when $\Omega_{\rm rad} \neq 0$, $\rho_{\rm vac}$ is no longer constant in time. This can be proven directly from the Friedmann equations or, if one wishes, numerically.

Thus, there is a choice of vacuum with a backreaction that is entirely consistent with the current era; namely our solution to $a(t)$ is identical to the ΛCDM model. At early times, $a(t) \propto t^{2/3}$, i.e., matter dominated, and at later times grows exponentially, $a(t) \propto \exp(\sqrt{\mu}H_0 t)$, i.e., is dominated by vacuum energy. Given Ω_m, then the equation $a(t_0)=1$ determines $t_H \equiv H_0 t_0$ and thus the age of the Universe. Observations indicate $\Omega_m = 0.266$, and Eq. (30) gives $t_H = 0.997$. The reason this is so close to the measured value of $t_H = 0.996$ is that radiation is nearly negligible.

It is interesting to compare our model with the Standard Model of cosmology when one includes radiation, since, as explained above, $\rho_{\rm vac}$ no longer behaves like a cosmological constant. In Fig. 1 we compare the expansion rate $H = \dot a/a$ as a function of redshift z. One sees that for redshifts $z = 1 - 1/a(t)$ up to at least 1000, there is only a small discrepancy between our model and the Standard Model of cosmology. Interestingly, our vacuum energy ceases to behave as a cosmological constant roughly around the time the cosmic microwave background was formed.

The condition $g = 1$ relates Newton's constant G to the cutoff k_c. There are a number of possible interpretations of this curious result. Recall the Planck scale k_p is simply the

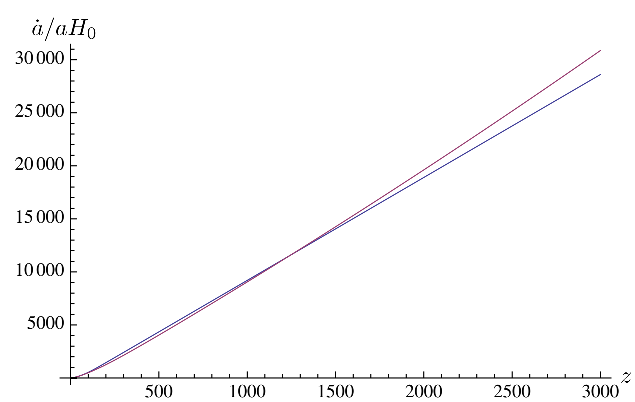

FIG. 1 (color online). The Hubble constant as a function of redshift z for the solution to our model Eq. (29) including radiation, versus the Friedmann equation Eq. (6) with radiation, matter, and a standard cosmological constant with $\Omega_m = 0.266$, $\Omega_{\rm rad} = 8.24 \times 10^{-5}$, $\Omega_\Lambda = 1 - \Omega_m - \Omega_{\rm rad}$.

scale one can define from G, but it is not *a priori* a physically meaningful energy scale; rather it is just the scale that one naively *expects* some form of quantization of the gravitational field to become important. Here, the relation $g = 1$ is a specific relation between the cutoff, Newton's constant G, and the number of particle species, and is unrelated to the quantization of gravity itself. One interpretation is simply that the cutoff k_c is the fundamental scale and that G is not fundamental, but rather is fixed by the cutoff from $g = 1$.

Allow us to speculate further: the relation $g = 1$ suggests that gravity itself originates from quantum vacuum fluctuations. Let us now argue how gravity can be heuristically "derived" from quantum mechanics. Processes in a closed universe are adiabatic, $dQ = 0$, and the first law of thermodynamics is $dE = -pdV$, where p is the total pressure due to all constituents regardless of their nature. Let us identify the internal energy with vacuum energy, $dE = \rho_{\rm vac} dV$. Recalling that $\rho_{\rm vac}$ is proportional to $\mathcal{A}$, Eq. (24) to lowest order, the first law is then nothing other than the second Friedmann equation (7), with Newton's constant identified as $G = \frac{8\pi}{3\Delta N k_c^2}$, which is the same as $g = 1$. If one assumes energy conservation $D^\mu T_{\mu\nu} = 0$, then this implies Eq. (8), from which, together with Eq. (7), one can derive the first Friedmann equation (6). It is well-known that the first Friedmann equation can be derived from nonrelativistic Newtonian gravity, so the above argument indirectly implies Newton's law of gravitation. From this point of view, the fundamental constants are $\hbar$, c, and k_c, and Newton's constant G is emergent. Curiously, in a universe with more bosons than fermions, gravity would actually be repulsive. The Planck scale has lost any real physical meaning here, and gravity is very weak simply because the cutoff k_c is large. If there is any truth to this idea, it renders the goal of quantizing gravity obsolete since it is already a quantum effect. Gravity would then be the ultimate macroscopic quantum mechanical phenomenon.

Finally we wish to make some observations on the so-called cosmic coincidence problem. Simply stated, the problem is that at the present time t_0 the densities of matter and vacuum energy are comparable, and since they evolved at different rates, their ratio would apparently have had to be fine-tuned to differ by many orders of magnitude in the very far past. From the point of view of the second order differential Eq. (29), $\mu = \Omega_{\rm vac}$ is just one of its arbitrary integration constants and we cannot predict it. However our construction has a bit more to it, based on the detailed Eq. (24). First of all, in our approach to the problem, $\rho_{\rm vac}$ is determined by current era physics, and is guaranteed to be of order 1 if k_c is close to the Planck scale; this is how our proposal involves UV/IR mixing. Using $\mathcal{A} \approx 9H_0^2/4$, which is an observational input, one finds $\rho_{\rm vac}/\rho_c \approx \frac{3\Delta N k_c^2}{8\pi k_p^2}$. Note also that if one rules out "phantom energy" with $w < -1$ based either on its strange cosmological properties [29] or general thermodynamic arguments [30], then this implies $g < 1$ which gives $\rho_{\rm vac}/\rho_c < 1$.

One can argue further. Let t_m be the time beyond which radiation can be neglected. In the solution Eq. (30), t should be replaced by $t - t_m$. To a good approximation, $t_0 - t_m \approx t_0$. Imposing then $a(t_0) = 1$ in Eq. (30) with $\mu + \Omega_m = 1$ determines $H_0 t_0$ as a function of Ω_m. One can show that $2/3 < H_0 t_0 < \infty$. As stated above, with the present data, $\Omega_m = 0.226$, this gives $H_0 t_0 = 0.997$. Thus the "coincidence" that $\rho_m/\rho_{\rm vac} \approx 0.37$ is now linked to the fact that current measurements give $H_0 t_0$ very close to 1. How would these numbers change if they had been measured in the past, say, at time $t' = t_0/2$, nearly 7 billion years ago when the Universe was half as old? Using Eq. (30) with $\Omega_m = 0.266$, one finds that at the time $t_0/2$ the Hubble constant was $H' = 1.52 H_0$ and $H't' = 0.76$. As we did, an observer at that time interprets the solution with $a(t') = 1$, and from Eq. (30) we can infer the value of Ω'_m; one finds $\Omega'_m = 0.68$ and $\rho'_m/\rho'_{\rm vac} = 2.13$. Thus for the entire duration of the second half of the Universe's history, the product $H't'$ only varied by a factor 3/4 and the ratio $\rho'_m/\rho'_{\rm vac}$ by less than a factor of 6. Actually the evolution of this ratio is very slow. For instance, if one would have measured it at time $t_N = t_0/N$, one would have obtained $(\rho_m/\rho_{\rm vac})(t_N) = 0.61 N^2$ for N large. Thus one has to go very deep into the past to have a huge difference between ρ_m and $\rho_{\rm vac}$. Furthermore, at these very early times the radiation plays a role and our model breaks down, perhaps with $|{\rm vac}\rangle$ being replaced by $|\widehat{\rm vac}\rangle$ of the next section.

An alternative way of summarizing what we have added to the discussion of the cosmic coincidence problem is the following: if one takes the point of view that the scale of vacuum energy $\rho_{\rm vac}$ is not determined by Planck scale physics, but rather by current day physics, as in our model, then there is much less of a need to explain any fine-tuning in the very far past. All one needs is a high energy cutoff, which is within the framework of low energy quantum field theory as we currently understand it. In our model, vacuum energy is a *low energy, IR* phenomenon, like the Casimir effect, but is also influenced by UV physics, via the cutoff.

III. VACUUM ENERGY PLUS RADIATION

A. Vacuum energy in the conformal time vacuum

In this section, we show that another choice of vacuum $|\widehat{\rm vac}\rangle$ is consistent with a universe consisting only of vacuum energy and radiation, i.e., massless particles. The quantization scheme is based on conformal time τ, defined as $dt = a d\tau$. Like the choice in the last section, this also simplifies the time dependence of the action (13), and is common in the literature. (See for instance Refs. [31,32] and references therein.) Rescaling the field $\Phi = \phi/a$, and integrating by parts, the action becomes

$$S = \int d\tau d^3x\left(\frac{1}{2}\partial_\tau\phi\partial_\tau\phi - \frac{1}{2}\vec{\nabla}\phi\cdot\vec{\nabla}\phi + \frac{\mathcal{R}a^2}{12}\phi^2\right), \quad (32)$$

with the Ricci scalar $\mathcal{R} = 6a''/a^3$, where primes indicate derivatives with respect to conformal time τ.

The field can be expanded in modes

$$\phi = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}}(a_\mathbf{k} v_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{x}} + a_\mathbf{k}^\dagger v_\mathbf{k}^* e^{-i\mathbf{k}\cdot\mathbf{x}}), \quad (33)$$

where the $a_\mathbf{k}$'s satisfy canonical commutation relations as before. The function $v_\mathbf{k}$ is now required to satisfy

$$(\partial_\tau^2 + \hat{\omega}_\mathbf{k}^2)v_\mathbf{k} = 0, \qquad \hat{\omega}_\mathbf{k}^2 \equiv \mathbf{k}^2 - \mathcal{R}a^2/6. \quad (34)$$

The analysis of the last section applies with $\mathcal{A}$ replaced by $\mathcal{R}/6$, which leads to

$$\rho_{\widehat{\rm vac}} \approx \Delta N\left[\frac{k_c^2}{96\pi^2}\mathcal{R} - \frac{1}{4608\pi^2}\mathcal{R}^2\right]. \quad (35)$$

It is clear that $|{\rm vac}\rangle \neq |\widehat{\rm vac}\rangle$ since the $v_\mathbf{k} \neq u_\mathbf{k}$.

It is instructive to compare the above result with the detailed point-splitting calculation performed in Ref. [10] for de Sitter space. The regularization utilized there insists on a finite answer and thus discards the k_c^4 and k_c^2 term:

$$\langle T_{\mu\nu}\rangle_{\rm ren} = -g_{\mu\nu}\left(\frac{1}{128\pi^2}(\xi - 1/6)^2\mathcal{R}^2 - \frac{1}{138240\pi^2}\mathcal{R}^2\right), \quad (36)$$

where ξ is an additional coupling to $\mathcal{R}\Phi$ in the original action. In our calculation $\xi = 0$, and one sees that our simple calculation reproduces the first $\mathcal{R}^2$ term. When $\xi = 1/6$ the theory is conformally invariant and the additional term is the conformal anomaly [33], which our simple calculation has missed. This is not surprising, since the anomaly depends on the spin of the field, and not simply of opposite sign for bosons versus fermions. In any case, in our approximation we are dropping the $\mathcal{R}^2$ terms. What this indicates is that the assumption (iv) in the introduction is essentially correct if one carefully constructs the full stress tensor in a covariant manner, such as by point splitting.

B. Consistent backreaction

In this case define the dimensionless constant

$$\hat{g} = \frac{\Delta N}{3\pi}Gk_c^2 \quad (37)$$

such that

$$\rho_{\widehat{\rm vac}} = \frac{\hat{g}}{32\pi G}\mathcal{R}. \quad (38)$$

Including $\rho_{\widehat{\rm vac}}$ in ρ, the first Friedmann equation then becomes

$$\left(1 - \frac{\hat{g}}{2}\right)\left(\frac{\dot{a}}{a}\right)^2 - \frac{\hat{g}}{2}\frac{\ddot{a}}{a} = \frac{8\pi G}{3}(\rho_m + \rho_{\rm rad}). \quad (39)$$

Similarly to what was found in the last section, a consistent solution only exists when $\hat{g} = 1$, but this time with no matter, $\rho_m = 0$. First consider the case where there is no radiation or matter. Then Eq. (39) implies $\ddot{a}/a = (2 - \hat{g})(\dot{a}/a)^2/\hat{g}$. Using this, the pressure can again be found from Eq. (7):

$$p_{\widehat{\rm vac}} = -\frac{(4 - \hat{g})}{3\hat{g}}\rho_{\widehat{\rm vac}}. \quad (40)$$

Consistency requires the equation of state parameter $w = -1$, i.e., $\hat{g} = 1$. The solution is $a(t) \propto e^{Ht}$ for some constant H, and $\rho_{\widehat{\rm vac}}$ is independent of time.

Now, let us include radiation. At a fixed time t_i define $\rho_i = 3H^2/8\pi G$ where H is a constant equal to $\dot{a}/a$ at the time t_i. Now we have to solve (when $\hat{g} = 1$)

$$\frac{1}{2H^2}\left[\left(\frac{\dot{a}}{a}\right)^2 - \frac{\ddot{a}}{a}\right] = \frac{\Omega_{\rm rad}}{a^4}, \quad (41)$$

where $\Omega_{\rm rad} = \rho_{\rm rad}/\rho_i$ at the time t_i where $a(t_i) = 1$. The general solution, up to a time shift, is

$$a(t) = \left(\frac{\Omega_{\rm rad}}{\nu}\right)^{1/4}\sqrt{\sinh(2H\sqrt{\nu}t)}, \quad (42)$$

where ν is a free parameter. Surprisingly, again $\rho_{\widehat{\rm vac}}$ is still a constant,

$$\frac{\rho_{\widehat{\rm vac}}}{\rho_i} = \nu. \quad (43)$$

However this is spoiled if there is matter present (see below). At early times, radiation dominates, $a(t) \propto t^{1/2}$, and at later times vacuum energy dominates, $a(t) \propto \exp(\sqrt{\nu}Ht)$.

This choice of vacuum and self-consistent backreaction is perhaps relevant to the very early universe which consists primarily of radiation and no matter. In fact, at the very earliest times, H is presumably set by the Planck time $t_p = 1/E_p$, which is a much larger scale than H_0 by many orders of magnitude. In fact, since the only scale in $\rho_{\rm vac}$ is k_c, we expect that higher orders in the adiabatic expansion give H/k_c of order 1 [34]. For k_c near the Planck scale, then H is roughly of the right scale for inflation. When vacuum energy dominates, $a(t)$ then grows exponentially on a time scale consistent with the inflationary scenario [35–37]. Here this is accomplished without invoking an inflaton field. Many models of inflation typically suffer from the "graceful exit problem," i.e., inflation must come to an end in a relatively short period of time. Based on our work, we suggest the following scenario. Initially there is only radiation and vacuum energy, which consistently leads to inflation. However as matter is produced, perhaps by particle creation from the vacuum energy, the

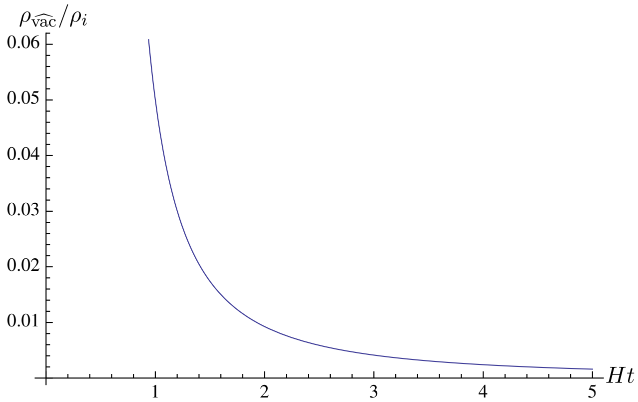

FIG. 2 (color online). The vacuum energy $\rho_{\widehat{\rm vac}}/\rho_i$ as a function of Ht for $\Omega_{\rm rad} = 0.8$ and $\Omega_m = 0.15$.

above solution is no longer consistent. Thus, $|\widehat{\rm vac}\rangle$ is no longer a consistent vacuum, which suggests that $\rho_{\widehat{\rm vac}}$ should somehow relax to zero. In support of this idea, we numerically solved Eq. (41) with an additional matter contribution on the right-hand side equal to Ω_m/a^3. As expected $\rho_{\widehat{\rm vac}}$ is no longer constant, but decreases in time as shown in Fig. 2. Of course, we are already aware that Eq. (41) with an additional $\Omega_m \neq 0$ is not consistent with Eq. (10) since for a time varying $\rho_{\widehat{\rm vac}}$ the pressure $p_{\widehat{\rm vac}} \neq -\rho_{\widehat{\rm vac}}$; however this plot does indeed show that $\rho_{\widehat{\rm vac}}$ decreases.

IV. DISCUSSION AND CONCLUSIONS

In this work we have proposed a different point of view on the cosmological constant problem. In analogy with the Casimir effect, we proposed the principle that empty Minkowski space should be gravitationally stable in order to fix the zero-point energy which is otherwise arbitrary. In a FLRW cosmological geometry, this leads to a prescription for defining a physical vacuum energy $\rho_{\rm vac}$ which depends on $\dot{a}$ and $\ddot{a}$. In the current era, this leads to a $\rho_{\rm vac}$ that is constant in time with $\rho_{\rm vac} \approx k_c^2 H_0^2$, which is the correct order of magnitude in comparison to the measured value if the cutoff k_c is on the order of the Planck scale.

We described two different choices of vacua, and studied the self-consistent backreaction of this vacuum energy on the geometry. One choice of vacuum is consistent with the current matter and dark energy dominated era. Another choice of vacuum is consistent with the early universe consisting of only radiation and vacuum energy, and we suggested that this perhaps describes inflation, and also a resolution to the graceful exit problem. Although our proposals could certainly be further improved, their consequences have at least survived a few checks. The role of higher orders of the adiabatic expansion on the backreaction should be better deciphered.

Both these consistent solutions require a relation between the cutoff k_c and Newton's constant G, and we speculated above on possible interpretations of this relation. It remains unclear how to apply the ideas of this work to the time period intermediate between inflation and the current era, where in our scenario, $|\widehat{\rm vac}\rangle$ would somehow evolve to $|{\rm vac}\rangle$, and this is clearly beyond the scope of this paper.

ACKNOWLEDGMENTS

We wish to thank the IIP in Natal, Brazil, for their hospitality during which time this work was initiated. This work is supported by the National Science Foundation under Grant No. NSF-PHY-0757868 and by the "Agence Nationale de la Recherche" Contract No. ANR-2010-BLANC-0414.

[1] S. Weinberg, Rev. Mod. Phys. **61**, 1 (1989).
[2] S. Nobbenhuis, Found. Phys. **36**, 613 (2006).
[3] J. Martin, C.R. Physique **13**, 566 (2012).
[4] S. E. Rugh and H. Zinkernagel, Studies in History and Philosophy of Science Part B **33**, 663 (2002).
[5] Astronomical data are taken from the most recent WMAP results. See e.g., http://pdg.lbl.gov/2012/reviews/rpp2012-rev-astrophysical-constants.pdf.
[6] We will mainly work in energy units. We remind the reader that an inverse length $1/L$ can be converted to an energy $1/L \to \hbar c/L$, and similarly for an inverse time $1/t \to \hbar/t$. The Planck energy E_p and momentum k_p are $E_p = \hbar c k_p = \sqrt{\hbar c^5/G} = 1.2 \times 10^{19}$ Gev. The current value of the Hubble constant is $H_0 = 2.3 \times 10^{-18}\ {\rm s}^{-1} = 1.5 \times 10^{-42}$ Gev$/\hbar$. Unless otherwise indicated we set $\hbar = c = 1$.
[7] J. Schwinger, in *The Physicist's Conception of Nature*, edited by J. Mehra (Reidel, Dordrecht, 1973), p. 413.
[8] A. LeClair, Int. J. Mod. Phys. A **23**, 1371 (2008).
[9] Of course, we can adopt a hydrodynamic point of view so that the cosmological dark energy fluid is locally characterized by the flat space vacuum energy. But then this vacuum energy is a constant and moduli independent, and thus ill-defined unless we invoke some principle fixing the zero-point energy in flat space.
[10] T. S. Bunch and P. C. W. Davies, Proc. R. Soc. A **360**, 117 (1978).
[11] R. L. Arnowitt, S. Deser, and C. W. Misner, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962).
[12] L. F. Abbott and S. Deser, Nucl. Phys. **B195**, 76 (1982).
[13] R. Schoen and S.-T. Yau, Commun. Math. Phys. **65**, 45 (1979).
[14] E. Witten, Commun. Math. Phys. **80**, 381 (1981).
[15] A. Strominger, Nucl. Phys. **B319**, 722 (1989).

[16] W. Chen and Y.-S. Wu, Phys. Rev. D **41**, 695 (1990).

[17] Y. J. Ng and H. van Dam, Int. J. Mod. Phys. Conf. Ser. **10**, 49 (2001).

[18] M. Ahmed, S. Dodelson, P. B. Green, and R. Sorkin, Phys. Rev. D **69**, 103523 (2004).

[19] After completion of this article we were made aware of some recent works where similar but not identical ideas were pursued [20–23]. The first works [20–22] considered $\rho_{\rm vac} \propto H^2$, which is time dependent and ruled out by observations. The later work [22] briefly considered the conformal time vacuum $|\widehat{\rm vac}\rangle$, described here in Sec. III; based on our understanding this vacuum is also inconsistent with the present universe, again because it is time dependent, in contrast to our vacuum $|{\rm vac}\rangle$ described in Sec. II.

[20] M. Maggiore, Phys. Rev. D **83**, 063514 (2011).

[21] M. Maggiore, L. Hollenstein, M. Jaccard, and E. Mitsou, Phys. Lett. B **704**, 102 (2011).

[22] L. Hollenstein, M. Jaccard, M. Maggiore, and E. Mitsou, Phys. Rev. D **85**, 124031 (2012).

[23] B. M. Deiss, arXiv:1209.5386.

[24] Consider the action for a single bosonic field ϕ in zero spatial dimensions: $S = \int dt(\frac{1}{2}\partial_t\phi\partial_t\phi - \frac{\omega^2}{2}\phi^2)$. The equation of motion is $\partial_t^2\phi = -\omega^2\phi$, and thus ϕ can be expanded as follows: $\phi = (ae^{-i\omega t} + a^\dagger e^{i\omega t})/\sqrt{2\omega}$. Canonical quantization gives $[a, a^\dagger] = 1$, and the Hamiltonian is $H = \frac{1}{2}(\partial_t\phi\partial_t\phi + \omega^2\phi^2) = \frac{\omega}{2}(aa^\dagger + a^\dagger a) = \omega(a^\dagger a + \frac{1}{2})$.

[25] Consider a zero-dimensional version of a Majorana fermion with action $S = \int dt(i\bar{\psi}\partial_t\bar{\psi} - i\psi\partial_t\psi - \omega\bar{\psi}\psi)$. The equation of motion is $\partial_t\psi = -i\omega\bar{\psi}$, and $\partial_t\bar{\psi} = -i\omega\psi$, and thus ψ, $\bar{\psi}$ have the following expansions: $\psi = \frac{1}{\sqrt{2}}(be^{i\omega t} + b^\dagger e^{-i\omega t})$, and $\bar{\psi} = \frac{1}{\sqrt{2}}(-be^{i\omega t} + b^\dagger e^{-i\omega t})$. Canonical quantization leads to $\{b, b^\dagger\} = 1$, $b^2 = b^{\dagger 2} = 0$, and the Hamiltonian is $H = \omega\bar{\psi}\psi = \frac{\omega}{2}(b^\dagger b - bb^\dagger) = \omega(b^\dagger b - \frac{1}{2})$.

[26] Shifting the potential $V(\Phi)$ of a bosonic field does not modify its equation of motion but shifts its stress tensor by a term proportional to $g_{\mu\nu}$. In a hydrodynamic description (see e.g., Ref. [27]), this corresponds to a shift of the vacuum energy density ρ by a constant but leaving $(\rho + p)$ unchanged.

[27] V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, Cambridge, England, 2005).

[28] If such a scenario has some truth to it, it suggests that not very much "new physics" occurs below the Planck scale.

[29] R. R. Caldwell, M. Kamionskowski, and N. N. Weinberg, Phys. Rev. Lett. **91**, 071301 (2003).

[30] A. LeClair, J. High Energy Phys. 05 (2005) 068.

[31] N. D. Birrell and P. C. W. Davies, *Quantum Field Theory in Curved Space* (Cambridge University Press, Cambridge, England, 1984).

[32] T. Jacobson, arXiv:gr-qc/0308048.

[33] J. S. Dowker and R. Critchley, Phys. Rev. D **13**, 3224 (1976).

[34] Above, as in the last section, we have assumed that $\mathcal{R}$ was small enough that we could drop the $\mathcal{R}^2$ term in $\rho_{\rm vac}$. However in the very early universe this is not necessarily the case. Although it is unclear whether it is justified to include the higher order term in the adiabatic expansion, if one does so, then the scale of H is set by the cutoff, as we now explain. Keeping the additional terms in Eq. (22), including the log, assuming only vacuum energy and that this leads to de Sitter space, one has $(\dot{a}/a)^2 = \ddot{a}/a = H^2 = $ const. The first Friedmann equation then becomes an algebraic equation for H when $\hat{g} = 1$, with the solution $H/k_c = \sqrt{2}e^{-1/4}$.

[35] A. Guth, Phys. Rev. D **23**, 347 (1981).

[36] A. Linde, Phys. Lett. **108B**, 389 (1982).

[37] A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).