UMN-TH-908/90
DOE/ER/40423-19

A New Mechanism for Neutralizing the Cosmological Constant

M.J. Duncan[1]
Department of Physics, University of Minnesota, Minneapolis MN 55455
Lars Gerhard Jensen[2]
Department of Physics FM-15, University of Washington, Seattle WA 98195

Abstract

We propose a dynamical mechanism for neutralizing the cosmological constant in our universe. The analysis is wholly Lorentzian, unlike most other mechanisms which rely on the problematic Euclidean quantum gravity. Additionally, we do not have to assume any properties of a complete quantum gravity theory for our result. Quantum cosmology is an effective theory below the Planck scale, and it suffices for the purpose of our demonstration. The vacuum energy from particle physics is many orders of magnitude too great to explain the measured smallness of the cosmological constant. Any generic field that allows us to shift the vacuum energy has the possibility of neutralizing it. However, such a field has an infinity of possible vacuum states, and we have to be in exactly the right one to achieve the cancellation. We show that when we third quantize a cosmology containing such a field in its particle physics sector we do have the possibility of selecting one vacuum over another. (The field we use for the sake of demonstration is a four-form.) The essential result of our paper is that the dominant vacuum state for the field is indeed the one that exactly cancels the particle physics vacuum energy. This means that a large universe, such as ours, has a vanishing cosmological constant.

[1]E-mail: duncan@vx.acs.umn.edu
[2]E-mail: jensen@landau.phys.washington.edu

1 Introduction

A long-standing mystery in particle physics and gravitation is why Einstein's cosmological constant is so many orders of magnitude smaller than it ought to be [1]. The particle physics vacuum inevitably has some energy locked up in it. It is this residual energy density, measured at energy scales far below those of the underlying particle dynamics, that is the cosmological constant. To arrange for a cancellation by the necessary amount requires a fine tuning of the particle physics, and all its vacuum fluctuations, to an incredible degree of accuracy. Even if it is possible to succeed in this, all that has happened is that we have exchanged one problem for a host of others. In what follows we shall argue that there is a convenient mechanism, within the framework of third quantization, for neutralizing the vacuum energy. If we invoke this mechanism when investigating any particle theory we need not bother about the consequences for the cosmological constant.

Third quantization, as we shall use it, is a method for determining the state vector of the universe. We shall show that for a large universe, such as ours, the dominant state vector is one for which the lowest, stable, particle physics vacuum does not contribute a cosmological constant. The method itself is quite straightforward. Starting from the hamiltonian formulation of gravity we postulate canonical quantization rules for the gravitational degrees of freedom and their conjugate momenta. Our rationale for this is that whatever quantum gravity might be, it certainly ought to reproduce cosmology in some limit. Quantum cosmology is simply an effective theory for quantum gravity at energies below the Planck scale. The quantized degrees of freedom of the cosmology are the remnants of the degrees of freedom of the complete quantum theory of gravity that give rise to the large scale structure of the universe.

Gravity couples to matter fields in the usual way. The matter fields are already second quantized. The resulting differential equation for this system of first quantized gravity and second quantized matter is the Wheeler-DeWitt (WDW) equation [2] which can, in principle, be solved for the cosmological wave function. The WDW equation has similarities to the more familiar Klein-Gordon (KG) equation. Not surprisingly there is also a difficulty with normalization which is remedied by promoting the cosmological wave function to the status of a field operator [3, 4, 5]. The field operator satisfies the WDW equation. We have now second quantized the gravity. As the matter was previously second quantized this explains why the procedure has been dubbed 'third' quantization. We shall show that the quanta of the field operator correspond to different particle physics vacua.

The notion of using quantum cosmology to solve the cosmological constant problem goes back to an observation of Baum [6] and Hawking [7]. In Euclidean quantum gravity the path integral has a stationary value of $\exp(3\pi/G_N\Lambda)$. If the cosmological constant Λ can be made a dynamical variable then the path integral would be dominated by geometries for which $\Lambda \rightarrow 0_+$. Actually, the BH factor crops up quite often in quantum cosmology. The WDW equation typically has classically allowed and forbidden regions, where the solutions are oscillatory and non-oscillatory respectively. It is possible to calculate the tunneling probability through the classically forbidden region and, depending on the boundary conditions, it can be the BH factor [8] or its reciprocal [9]. Rubakov [3] has shown that the BH factor also appears in simple third quantized models, there interpreted as the probability of finding a large universe given a value for Λ. However, in the above cases Λ is a fixed parameter of the theory and therefore the BH mechanism is not directly applicable[1]. Our mechanism is the converse: given a large universe the most likely vacuum state is one with vanishing cosmological constant.

In passing we mention that wormholes [10], or topology changing configurations in Euclidean space, do give rise to a dynamical cosmological constant and could allow the use of the Baum-Hawking (BH) mechanism. It has been suggested that when both wormholes and universes other than ours are traced out of the theory that the cosmological constant in the resulting one-universe theory is driven to zero [11]. Unfortunately it has become clear that there are serious difficulties [12, 13] with the wormhole calculus. An alternative is to third quantize the WDW equation and use wormholes to vary Λ [5]. Unfortunately the wormholes have other dire consequences that ultimately invalidate the result. This is above and beyond the problems inherent in Euclidean space quantum gravity. In this paper we shall stay clear of wormholes and their pathology.

The analysis presented here is firmly rooted in Lorentzian spacetime. As we mentioned above, Euclidean quantum cosmology has its problems [14] and it is difficult to circumvent

[1]These models also suffer from being *too* simple. They have only one variable (the scale factor). This is akin to quantum mechanics with only a time coordinate and no space coordinates.

its pitfalls. To allow for the variation of the cosmological constant we shall introduce a four-form field which accomplishes the task admirably. It was realized some time ago that the presence of a four-form field gives rise to a variable cosmological constant [7, 15–18]. We are not claiming that there really is such a field in the universe. It is just that it is the simplest way to arrange for a dynamical cosmological constant, its physics is straightforward and, most important of all, we can actually calculate with it consistently. We shall find that we still recover the BH factor above, albeit by a different chain of arguments. To be precise, in what follows we shall third quantize a Friedmann-Robertson-Walker (FRW) universe containing a four-form and demonstrate that the cosmological constant is dynamically neutralized.

2 The Wheeler-DeWitt Equation

The total action for matter coupled to gravity in a spacetime with a Lorentzian signature is,

$$S_{\rm tot} = \frac{1}{16\pi G_N}\int_{\mathcal{M}} (R - 2\Lambda_0)\sqrt{-g}\, d^4x + S_{\partial\mathcal{M}} + S_{\rm matter} \tag{1}$$

This is the form of the effective action at energies much lower than the Planck scale $M_P = G_N^{-1/2}$ where quantum fluctuations in the spacetime are expected to invalidate (1). The 'bare' cosmological constant is denoted by Λ_0. What we mean by a bare cosmological constant will be discussed below. If the spacetime manifold has a boundary $\partial\mathcal{M}$ we are required to add the boundary correction [19],

$$S_{\partial\mathcal{M}} = \frac{1}{8\pi G_N}\int_{\partial\mathcal{M}} K\sqrt{h}\, d^3x \tag{2}$$

where h_{ij} is the induced three-metric on the boundary and K is the trace of the extrinsic curvature. The effect of this boundary term is to remove second time derivatives from the action upon integration by parts, and thus render the action a function of the gravitational degrees of freedom and their first time derivatives.

The choice of what matter to include depends on which model of particle dynamics we care to use below the Planck scale. We assume that we are working with a generic particle physics theory. Additionally we assume that there is a field present whose background configuration is undetermined by the theory itself, and which shifts the vacuum energy. The simplest example of such a field that we are aware of is a four-form field strength. Almost any other field with the aforementioned properties will suffice for our mechanism, but the advantage of the four-form for the sake of demonstration is that its physics is well-understood. Separate the matter action into this field and the rest of the particle physics, $S_{\text{matter}} = S_{\text{ff}} + S_0$.

A four-form field strength is the curl of a three-form potential, $\mathbf{F} = \mathbf{dA}$, and its action is similar to the Maxwell action for electromagnetism, but with a couple of extra indices,

$$S_{\text{ff}} = -\frac{1}{2}\int_{\mathcal{M}} \mathbf{F} \wedge \star\mathbf{F} = -\frac{1}{2}\frac{1}{4!}\int_{\mathcal{M}} F_{\mu\nu\lambda\rho}F^{\mu\nu\lambda\rho}\sqrt{-g}\, d^4x \tag{3}$$

Such a field arises naturally in theories with supergravity [15], although for our purposes it is simply an introduced field with no particular theory in mind. Its special property is that its vacuum configuration is undetermined by its dynamics [15, 16] and that this shifts the cosmological constant. The cosmological constant then defies its name by becoming a variable. The four-form has also been used in the framework of Euclidean quantum gravity as an example of the BH mechanism [7, 17, 18]. For a full discussion of this we refer the reader to ref. [18]. The mechanism we are proposing also takes advantage of its properties.

As for the rest of the matter, the generic particle physics model, it has a vacuum state, $|\Omega_0\rangle$ of lowest energy. All the particle quanta are excitations from this vacuum. The vacuum expectation value of the stress energy tensor (from S_0) can be written $\langle\Omega_0|T_{\mu\nu}|\Omega_0\rangle = -\rho_0 g_{\mu\nu}$, where ρ_0 is the particle physics vacuum energy density. We can absorb this into the bare cosmological constant $\Lambda_0 \to \Lambda_0 + 8\pi G_N \rho_0$ in (1). The matter action then becomes $S_0 \to S_0'$ whose stress energy tensor is purely due to the energy of the excitations *i.e.* we have moved the vacuum energy from the matter action into Λ_0, and $\langle\Omega_0|T'_{\mu\nu}|\Omega_0\rangle = 0$. The bare cosmological constant is *equivalent to the vacuum energy density arising from the particle physics.* (The nomenclature can sometimes lead to confusion, but we adopt it here as it is commonly used in the literature.) The matter excitations (particles) give the usual density and pressure terms in the Einstein equations, and if the four-form were not present the cosmological constant would be Λ_0. It is this latter quantity that we have to cancel, and so we may neglect the particle excitations in what follows.

In studying general relativistic cosmology it is convenient to express the metric in a form that distinguishes the true gravitational degrees of freedom from those that merely reflect the reparametrization of general coordinate invariance. The ADM formalism [20] is well-suited to this task. Moreover, assumptions about the symmetries of the spacetime greatly reduce the problem of solving the Einstein equations. Friedmann-Robertson-Walker cosmologies are based on the assumption of an homogeneous and isotropic universe, properties that our universe appears to possess on the large scale. It is for this reason that they are the most widely used of cosmological models. We shall assume we are dealing with a closed FRW cosmology. The spatial hypersurfaces are three-spheres. The FRW metric has the form,

$$ds^2 = \left(\frac{2G_N}{3\pi}\right)\left(-N^2(t)\,\mathbf{d}t\otimes\mathbf{d}t + a^2(t)\,\delta_{ij}\,\boldsymbol{\sigma}^i\otimes\boldsymbol{\sigma}^j\right) \tag{4}$$

The common prefactor is just a scaling for later convenience. We have chosen to write the line element of the unit three-sphere in terms of a homogeneous basis of one-forms $\boldsymbol{\sigma}^i$. They satisfy the algebra $\mathbf{d}\boldsymbol{\sigma}^i = -\epsilon_{ijk}\,\boldsymbol{\sigma}^j\wedge\boldsymbol{\sigma}^k$ and when integrated over the three-sphere give its volume $\int\boldsymbol{\sigma}^1\wedge\boldsymbol{\sigma}^2\wedge\boldsymbol{\sigma}^3 = 2\pi^2$. The only degree of freedom in the metric is the usual FRW scale parameter $a(t)$, a function only of time by virtue of the homogeneity of the three-space. The lapse function $N(t)$ simply represents the time reparametrization invariance.

If we expand the potential in the basis of this spacetime it must have the form,

$$\mathbf{A} = \chi(t)\,\boldsymbol{\sigma}^1\wedge\boldsymbol{\sigma}^2\wedge\boldsymbol{\sigma}^3 + \frac{1}{2}\zeta_{ij}(t)\,\mathbf{d}t\wedge\boldsymbol{\sigma}^i\wedge\boldsymbol{\sigma}^j \tag{5}$$

and we immediately see that $\mathbf{F}$ will be independent of ζ_{ij}. There is only one degree of freedom for a four-form in general, and in our particular case it is $\mathbf{F} = \dot{\chi}\,\mathbf{d}t\wedge\boldsymbol{\sigma}^1\wedge\boldsymbol{\sigma}^2\wedge\boldsymbol{\sigma}^3$. You can also see that the action (3) has only one degree of freedom in much the same way that you identify those of the Maxwell field in QED; essentially the ζ_{ij} may be gauged away.

If we now rescale the four-form and the bare cosmological constant according to,

$$\chi = (2\pi^2)^{-1/2}\left(\frac{2G_N}{3\pi}\right)\phi$$

$$\Lambda_0 = \left(\frac{9\pi}{2G_N}\right)\lambda_0$$

and evaluate the scalar curvature R arising from the metric (4) we find that the action (1) may be re-expressed as,

$$S = \frac{1}{2}\int dtN\left(-\frac{1}{N^2}\dot{a}^2a + a - \lambda_0a^3 + \frac{\dot{\phi}^2}{N^2a^3}\right) \tag{6}$$

where we have already performed the integral over the three-space. The lapse function does not have a kinetic term and therefore variation of the action with respect to it will yield a constraint. The scalings were chosen to make this action as simple as possible; everything is expressed in units of the Planck mass. We have reduced the problem to one of two time-dependent coordinates. The last term in the action is *not* equivalent to the kinetic term for a massless scalar particle. A scalar would have had the a^3 factor in the numerator rather than denominator. Rewriting the action in the form $S = \int (p_a \dot{a} + p_\phi \dot{\phi} - NH)dt$, after identifying the momenta, the Hamiltonian is found to be,

$$H = \frac{1}{2}\left(-\frac{p_a^2}{a} + a^3 p_\phi^2 - a + \lambda_0 a^3\right) \tag{7}$$

The classical equations of motion follow by varying (6) with respect to the coordinates N, a and ϕ. Variation with respect to N results in the Hamiltonian constraint, $H = 0$. Also from (6) it follows that ϕ is a cyclic coordinate, i.e. p_ϕ is conserved classically. This implies that (up to the scalings) $\mathbf{F} \sim p_\phi \boldsymbol{\eta}$ is the solution to the four-form equations of motion, where $\boldsymbol{\eta}$ is the invariant volume element of the spacetime (the canonical four-form). The one degree of freedom possessed by $\mathbf{F}$ is really an undetermined constant, different values of which correspond to different choices of vacuum. Solving the Hamiltonian constraint for $a(t)$ we find (upon setting $N = 1$),

$$a(t) = \frac{1}{\sqrt{\lambda_0 + p_\phi^2}} \cosh\left(\sqrt{\lambda_0 + p_\phi^2}\; t\right) \tag{8}$$

Note the similarity with the familiar solution of a closed FRW cosmology with a cosmological constant. The difference here is that the four-form has introduced an addition of p_ϕ^2 to the bare cosmological constant λ_0. As we shall see below, in the quantized version of the theory a similar phenomenon occurs.

The value of p_ϕ is arbitrary and characterizes the vacuum configuration $|p_\phi\rangle$ of the four-form. There is an infinity of possible configurations. The combined matter and four-form vacuum is $|\Omega_0; p_\phi\rangle$ which has an effective cosmological constant $\lambda = \lambda_0 + p_\phi^2$. As λ_0 is usually negative (particle physics vacua usually have a negative energy density from symmetry breaking) we observe that the states $|\Omega_0; \pm\sqrt{-\lambda_0}\rangle$ have vanishing effective cosmological constant. However, we have to be in exactly the right states for this to happen. We shall now third quantize the Hamiltonian (7) and show that we are indeed forced into some superposition of these special states.

The system described by (6) is quantized by imposing the usual canonical quantization relations, similar to the method used in ordinary quantum mechanics. We simply replace the canonical momenta with the corresponding differential operators, and construct the Hamiltonian operator which acts on a wave function $\Psi(a,\phi)$. We find that the quantum Hamiltonian is,

$$\mathcal{H} = \frac{1}{2a}\left[a^{-n}\frac{\partial}{\partial a}\left(a^{n}\frac{\partial}{\partial a}\right) - a^{4}\frac{\partial^2}{\partial\phi^2} + (\lambda_0 a^4 - a^2)\right] \tag{9}$$

alias the Wheeler-DeWitt operator. The parameter n reflects the fact that there is a factor ordering ambiguity in going from (7) to (9). We shall permit ourselves to choose it as we wish for the sake of calculational expediency. As we have seen above, in theories including gravity the total Hamiltonian is constrained, so we must also demand that the equation for the wave function $\Psi(a,\phi)$ must obey the Hamiltonian constraint, i.e. we must demand that,

$$\mathcal{H}\Psi = 0 \tag{10}$$

This is the Wheeler-DeWitt equation [2], and as the form of (9) attests it has some similarities to the KG equation. The scale parameter a and the four-form ϕ are analogous to time and a spatial coordinate in the KG equation respectively. The last term in (9) is analogous the the mass or potential of the KG equation. However, its dependence on the scale factor (the pseudo-time coordinate) means that it resembles more the KG equation on a curved background space [21].

It follows from (9) and (10) that there is a conserved current whose components are,

$$\begin{aligned} J^a &= i a^n \Psi^* \overleftrightarrow{\partial}_a \Psi \\ J^\phi &= -i a^{n+4} \Psi^* \overleftrightarrow{\partial}_\phi \Psi \end{aligned} \tag{11}$$

and which satisfy $\partial_a J^a + \partial_\phi J^\phi = 0$. We can therefore define an inner product of two wave functions according to,

$$(\Psi_1|\Psi_2) \equiv i a^n \int_{-\infty}^{\infty} \left(\Psi_1^* \overleftrightarrow{\partial}_a \Psi_2\right) d\phi \tag{12}$$

where both wave functions are evaluated at the same value of the variable a. The scale factor a in this case is naturally interpreted as the external order parameter of the wave function. In ordinary quantum mechanics the time coordinate is not an operator like $\hat{\mathbf{x}}$, but is rather the classical parameter that determines the evolution of the wave function from preparation to measurement. Of course time flows whether we like it or not and so the measurements there are described as cause and effect. In the case of the cosmological wave function there is no sense of direction to the scale factor and moreover it does not flow. Given the wave function (and its derivative) at one value of the scale we can evaluate it at any other using the WDW equation. However, as in the case of the KG equation, the inner product (12) need not be positive definite when $\Psi_1 = \Psi_2$. There is therefore a difficulty with a strict probabilistic interpretation of the WDW wave function, aside from the fact that the concept of probability in quantum cosmology is nebulous to begin with. This is remedied by third quantization.

Suppose we look for wave functions of the form $u_k(a,\phi) = e^{ik\phi} f_k(a)$. These are eigenstates of the four-form momentum operator. As we stated above, each eigenvalue k corresponds to a different vacuum state of the four-form, and therefore each of these eigenfunctions corresponds to a universe with a different vacuum. Setting the factor ordering parameter $n = -1$ for convenience the WDW equation reduces to one for $f_k(a)$,

$$\left[a\frac{d}{da}\left(\frac{1}{a}\frac{d}{da}\right) - a^2\left(1 - \lambda_k a^2\right)\right] f_k(a) = 0 \tag{13}$$

where $\lambda_k = \lambda_0 + k^2$ is the effective cosmological constant. Thus the presence of the four-form shifts the cosmological constant, in keeping with the classical theory [18]. In the quantum theory the shift depends on which eigenstate of p_ϕ we happen to be in. The solutions to (13) are Airy functions [22], solutions to the differential equation $w'' - zw = 0$, and are denoted $\mathrm{Ai}(z)$ and $\mathrm{Bi}(z)$. The general form of the $u_k(a,\phi)$ is therefore,

$$u_k(a,\phi) = \mathcal{N}_k\, e^{ik\phi}\left[\mathrm{Ai}(z) + \beta \mathrm{Bi}(z)\right] \tag{14}$$

where the argument is,

$$z = \frac{1 - \lambda_k a^2}{(4\lambda_k^2)^{1/3}} \tag{15}$$

The boundary conditions we impose in the next section will determine the parameter β.

It is then straightforward to construct a complete set of orthonormal solutions to the WDW equation, which under the inner product (12) satisfy,

$$\begin{aligned} (u_k|u_{k'}) &= \delta(k-k') \\ (u_k^*|u_{k'}) &= 0 \\ (u_k^*|u_{k'}^*) &= -\delta(k-k') \end{aligned} \tag{16}$$

The Wronskian of the Airy functions is $\mathrm{Ai}(z)\mathrm{Bi}'(z) - \mathrm{Ai}'(z)\mathrm{Bi}(z) = 1/\pi$ which fixes the normalization of the wave functions via the relation,

$$\frac{8|\mathcal{N}_k|^2}{(4\lambda_k^2)^{1/3}} [\lambda_k \Im\beta] = 1 \tag{17}$$

Immediately we see that $\Im\beta$, the imaginary part of β, must have the same sign as λ_k. As expected, our wave functions have to be complex otherwise they cannot be normalized.

3 Third Quantization

So far we have found a complete set of orthonormal solutions to the WDW equation. If we consider a general linear superposition of the states u_k and u_k^* this need not have a positive definite norm. To remedy this we are led to third quantization. We treat the wave function Ψ as an operator and expand it in terms of annihilation and creation operators for the modes u_k,

$$\hat{\Psi}(a,\phi) = \int_{-\infty}^{\infty} dk \left[\hat{c}(k)u_k(a,\phi) + \hat{c}^{\dagger}(k)u_k^*(a,\phi)\right] \tag{18}$$

and take these operators to satisfy the usual commutation relations,

$$\left[\hat{c}(k), \hat{c}^{\dagger}(k')\right] = \delta(k-k') \tag{19}$$

with all others vanishing. We would then find the equal scale commutation relations for the operator $\hat{\Psi}$ are,

$$\left[\frac{i}{a}\frac{\partial}{\partial a}\hat{\Psi}(a,\phi), \hat{\Psi}(a,\phi')\right] = \delta(\phi-\phi') \tag{20}$$

This is analogous to the quantization procedure for the KG scalar field. Actually, what we are doing is similar to quantizing a scalar on a *curved* background spacetime [21]. Instead of canonically quantizing from (20) we could also construct a path integral for the third quantized theory. The third quantized action is,

$$\mathcal{S}_3 = \frac{1}{2}\int a^n da d\phi \left[\left(\frac{\partial\Psi}{\partial a}\right)^2 - a^4\left(\frac{\partial\Psi}{\partial\phi}\right)^2 - \left(\lambda_0 a^4 - a^2\right)\Psi^2\right] \tag{21}$$

and we see that variation with respect to Ψ yields the WDW equation.

We fix the parameter β by demanding that the u_k correspond to positive frequency outgoing modes, frequency here referring to (negative) the rate of oscillation with respect to the scale. In classically allowed regions the positive frequency modes correspond to expanding universes [23]. As the potential in (9) is scale dependent so will the meaning of positive frequency. We therefore have many possible expansions (18) depending on at which scale we choose to fix the frequency. We shall consider the two choices of small scales ($a \gtrsim 1$) and large scales ($a \to \infty$), and impose the positive frequency conditions there.

The vacuum energy from the particle physics is λ_0 in Planck units and is therefore very small. Moreover, it is more than likely to be negative from the effects of symmetry breakings in these theories. For the electroweak symmetry breaking $\lambda_0 \sim -10^{-67}$, which is very small, although it is still many orders of magnitude greater than the observed upper bound on the cosmological constant 10^{-120} [1]. The wave number k is the momentum of the three-form in Planck units and is also very small. We can think of there being a cutoff k_c in (18) and also for integrals over k. Thus $|k| < k_c \ll 1$ and $|\lambda_0| \ll 1$. Hence $|\lambda_k| \ll 1$. In addition, $|\lambda_k a^2| \ll 1$ for a wide range of scales many times the Planck length. The effective cosmological constant can have either sign, depending on whether k^2 is greater than or less than $|\lambda_0|$.

For the moment let us consider $\lambda_k > 0$. In this case there are two different behaviours for $\lambda_k a^2 > 1$ and $\lambda_k a^2 < 1$ as there is a classical turning point for positive λ_k. In the limit $a \to \infty$ the argument of the Airy functions is $z \to -\infty$. In this limit the asymptotic behaviour of these functions is [22],

$$
\begin{aligned}
\mathrm{Ai}(-|z|) &\sim \pi^{-1/2}|z|^{-1/4}\sin(\zeta + \frac{\pi}{4}) \\
\mathrm{Bi}(-|z|) &\sim \pi^{-1/2}|z|^{-1/4}\cos(\zeta + \frac{\pi}{4})
\end{aligned}
\tag{22}
$$

where $\zeta = \frac{2}{3}|z|^{3/2}$. Thus the choice $\beta = i$ will ensure positive frequency outgoing modes for large scales. We shall call these modes which have positive frequency at large scales the 'out' modes. Setting $\beta = i$ in (17) to fix the normalization, the 'out' modes are then,

$$
u_k^{\mathrm{out}}(a,\phi) = \frac{e^{ik\phi}}{(128\lambda_k)^{1/6}}\left[\mathrm{Ai}(z) + i\mathrm{Bi}(z)\right]
\tag{23}
$$

when the wavenumber is such that $\lambda_k > 0$. Incidentally, the phase of this wave function in the large a limit is the same as if we had only used the semiclassical approximation, $\Psi \sim \exp(i\bar{S})$, where $\bar{S}$ is the action (6) evaluated on the classical path. This is just the positive frequency component of the Hartle-Hawking wave function [8] and the Vilenkin wave function [9].

Similarly we can impose the positive frequency conditions at small scales (we shall discuss the physical meaning of this in the next section). As $\lambda_k a^2 \to 0$, with positive λ_k, the argument of the Airy functions $z \to z_0 = (4\lambda_k^2)^{-1/3}$, which is very large from the discussion above. We can expand the wave function as a power series in the scale factor a. It is easily seen from (15) that both Airy functions have expansion,

$$w(z) = \sum_{n=0}^{\infty} \left(-\frac{a^2}{2}\right)^n \frac{1}{n!} \left[z_0^{-n/2} w^{(n)}(z_0)\right] \tag{24}$$

From the asymptotic form of the Airy functions for large positive z_0 [22] we find that their derivatives behave as,

$$\begin{aligned} z_0^{-n/2}\mathrm{Ai}^{(n)}(z_0) &\sim (-1)^n \frac{1}{2} \pi^{-1/2} z_0^{-1/4} e^{-\zeta_0} \left(1 + O(\zeta_0^{-1})\right) \\ z_0^{-n/2}\mathrm{Bi}^{(n)}(z_0) &\sim \pi^{-1/2} z_0^{-1/4} e^{\zeta_0} \left(1 + O(\zeta_0^{-1})\right) \end{aligned} \tag{25}$$

where again $\zeta_0 = \frac{2}{3} z_0^{3/2} = 1/3\lambda_k$. Thus the series (24) may be summed in both cases giving,

$$\begin{aligned} \mathrm{Ai}(z) &\sim \mathrm{Ai}(z_0) \exp\left(\frac{a^2}{2}\right) \\ \mathrm{Bi}(z) &\sim \mathrm{Bi}(z_0) \exp\left(-\frac{a^2}{2}\right) \end{aligned}$$

valid for $\lambda_k \ll 1$ with $\lambda_k a^2 < 1$, the latter inequality being satisfied for a wide range of small scale factors. Call the modes with positive frequency at small scales the 'in' modes. The asymptotic form of their wave function is then,

$$u_k^{\mathrm{in}}(a, \phi) \approx e^{ik\phi} \mathcal{N}_k \pi^{-1/2} z_0^{-1/4} \left(\frac{1}{2} e^{-\zeta_0} e^{a^2/2} + \beta e^{\zeta_0} e^{-a^2/2}\right) \tag{26}$$

for positive λ_k in the limit of small a. We can fix β by demanding that this have a positive frequency. However, there is another way we can determine β by demanding continuity of the 'in' modes when λ_k passes through zero for small wavevector.

As we have said, for $\lambda_k > 0$ there is a classical turning point and so the meaning of positive frequency is different at small and large scales. This is why we have two sets of modes above. However, for $\lambda_k \leq 0$ there is no turning point. Therefore the 'in' and 'out' modes are identical for $\lambda_k \leq 0$ as a positive frequency condition will hold over the complete range of a. It is only for $\lambda_k > 0$ that the 'in' and 'out' modes are distinct. The value of β can be determined by requiring continuity of the 'in' modes in the wavenumber k. The special case of $\lambda_k = 0$ is when the wavenumber $k = \pm k_0 = \pm|\lambda_0|^{1/2}$. The wave function for these two values may be obtained by solving (13) with $\lambda_k = 0$ and finding the positive frequency modes. The wave function at these critical wavenumbers is,

$$u^{\rm in}_{\pm k_0}(a,\phi) = u^{\rm out}_{\pm k_0}(a,\phi) = \frac{e^{\pm i k_0 \phi}}{\sqrt{8\pi}}\left(e^{a^2/2} + i e^{-a^2/2}\right) \tag{27}$$

The normalization is fixed by (16). With this choice of wave function the phase of (27) has a scale dependence of $\vartheta = -\arctan\left[\tanh\left(\frac{a^2}{2}\right)\right]$ leading to a positive frequency $\omega(a) = -\frac{1}{a}\frac{\partial\vartheta}{\partial a} = \mathrm{sech}(a^2)$.

We are now in a position to uniquely determine the value of β in (26), simply by demanding that (26) converges onto the states (27) in the limit of $\lambda_k \to 0$. The unique value is,

$$\beta = \frac{i}{2}e^{-2\zeta_0}$$

The important factor in this is the appearance of the exponential, necessary for correct normalization of the modes. The overall $e^{-\zeta_0}$ in (26) is canceled by a similar factor in $\mathcal{N}_k$. The numerical factor $i/2$ is uniquely determined by the factor of i in (27). This latter factor may be changed slightly, with a corresponding change in β, but must retain its positive imaginary part. All that happens, if we wish to do this, is that it introduces an overall numerical factor into our final results. The diligent reader can confirm that it does not alter our conclusions one bit. With this choice of β the wave function again has a positive frequency of $\omega(a) = \mathrm{sech}(a^2)$. With the normalization given by (17) the 'in' modes are explicitly,

$$u^{\rm in}_k(a,\phi) = \frac{e^{ik\phi}}{(16\lambda_k)^{1/6}}\left[e^{1/3\lambda_k}\mathrm{Ai}(z) + \frac{i}{2}e^{-1/3\lambda_k}\mathrm{Bi}(z)\right] \tag{28}$$

when the wavenumber is such that $\lambda_k > 0$.

The construction of the states above has been for $\lambda_k \geq 0$. We now return to the situation where the wavenumber k is small enough that $\lambda_k < 0$. In this case the argument z of (15) is always positive. With this sign of the effective cosmological constant the analysis is similar to that for the 'in' states above. The expansion of the Airy functions for all scales is as in (24), except for a positive sign to the $a^2/2$ terms inside the brackets. We go through the same procedure to find that the parameter $\beta = -\frac{i}{2}e^{-2\zeta_0}$ where now $\zeta_0 = 1/3|\lambda_k|$. There is only one common set of wave functions, valid for all scales,

$$u_k^{\rm in}(a,\phi) = u_k^{\rm out}(a,\phi) = \frac{e^{ik\phi}}{(16|\lambda_k|)^{1/6}}\left[ie^{1/3|\lambda_k|}{\rm Ai}(z) + \frac{1}{2}e^{-1/3|\lambda_k|}{\rm Bi}(z)\right] \tag{29}$$

when the wavevector is such that $\lambda_k < 0$. The frequency of these states is also $\omega(a) = \mathrm{sech}(a^2)$.

We have constructed two sets of complete, orthonormal modes, either of which may be used as a basis of states. The 'in' modes taken together are (29), (27) and (28) in order of increasing absolute wavenumber, whereas the 'out' modes are (29), (27) and (23). The classical cosmologies corresponding to the 'out' modes (23) in the limit of large scales are FRW with cosmological constant λ_k [9]. There is no direct interpretation of the modes (28), however, since $\lambda_k a^2 < 1$ is a classically forbidden region. Negative λ_k also does not correspond to a classical cosmology. All we have done is to construct positive frequency modes in regions of the (k, a) space. The wave operator can now be expanded in terms of either set of modes,

$$\begin{aligned}\hat{\Psi}(a,\phi) &= \int dk\left[\hat{c}_{\rm in}(k)u_k^{\rm in}(a,\phi) + \hat{c}_{\rm in}^\dagger(k)u_k^{\rm in*}(a,\phi)\right]\\ &= \int dk\left[\hat{c}_{\rm out}(k)u_k^{\rm out}(a,\phi) + \hat{c}_{\rm out}^\dagger(k)u_k^{\rm out*}(a,\phi)\right]\end{aligned} \tag{30}$$

each with their own set of annihilation and creation operators. Moreover, each set of modes has its own (third quantized) vacuum, which is annihilated by their respective operators,

$$\begin{aligned}\hat{c}_{\rm in}(k)\,|0,{\rm in}\rangle\!\rangle &= 0 \quad \forall k\\ \hat{c}_{\rm out}(k)\,|0,{\rm out}\rangle\!\rangle &= 0 \quad \forall k\end{aligned} \tag{31}$$

The double ket signifies that this is a third quantized Fock vacuum, not to be confused with the second quantized particle physics Fock vacuum $|\Omega_0\rangle$. The 'in' vacuum contains no positive frequency modes in the limit of small scales. The 'out' vacuum contains no positive frequency modes in the limit of large scales. We are now set to determine the state vector of the universe.

4 The Vacuum State of a Large Universe

It is worthwhile at this point to recap what we have done so far. We have considered a closed FRW cosmology with a matter sector containing a four-form. The matter is already described by a second quantized field theory. The vacuum energy from all fields, excluding the four-form, is λ_0 in Planck units and is very small and negative. The four-form vacuum state is undetermined: there is a continuum of possible background configurations for the four-form, each vacuum labeled by a real number k. The effective cosmological constant for the FRW universe is $\lambda_k = \lambda_0 + k^2$, a different one for each four-form vacuum. How do we then select a vacuum for the universe? To do this we third quantize the theory. We construct a complete set of solutions to the WDW equation, each mode corresponding to a different four-form vacuum state. There are two possible mode expansions, with positive frequency for small or large scales. It should be possible to calculate the density of k vacua for a certain third quantized state.

Given that the third quantized state of the theory is $|\Omega_3\rangle\rangle$ we can estimate the expectation value,

$$\mathcal{N}(k,k') = \langle\langle \Omega_3|\hat{c}^{\dagger}_{\text{out}}(k)\hat{c}_{\text{out}}(k')|\Omega_3\rangle\rangle \tag{32}$$

This is the distribution of k-vacua for a very large universe such as ours. A large universe has a scale factor $a \gg 1$ in Planck units, which is why we are interested in the number of 'out' modes. Our effective theory breaks down at scale factors of order the Planck length and so we should ensure that there is an absence of small universes in the third quantized state. For this reason we take $|\Omega_3\rangle\rangle = |0, \text{in}\rangle\rangle$. What we mean by a *small* universe is one having a scale factor $a \gtrsim 1$; for scale factors such as these, sufficiently far above the Planck length, the frequency $\omega(a) \ll 1$ and so the wave functions make sense as they do not have fluctuations of the order of the Planck length itself. We can now calculate the expectation (32).

For this we need the Bogolubov coefficients for the transformation between the two bases of states. The Bogolubov coefficients [24] are obtained by expanding the 'out' modes in terms of the complete set of 'in' modes,

$$u_k^{\text{out}}(a,\phi) = \int_{-\infty}^{\infty} dq \left[\alpha(k,q)u_q^{\text{in}}(a,\phi) + \beta(k,q)u_q^{\text{in}*}(a,\phi)\right] \tag{33}$$

By the orthonormality conditions (16) we identify the coefficients as the inner products,

$$\begin{aligned} \alpha(k,q) &= (u_k^{\text{out}}|u_q^{\text{in}}) \\ \beta(k,q) &= (u_k^{\text{out}}|u_q^{\text{in}*}) \end{aligned} \tag{34}$$

As we know the explicit form of the 'in' and 'out' states it is a simple job to evaluate these. The 'out' creation and destruction operators are similarly expressible in terms of the 'in' operators,

$$\hat{c}_{\text{out}}(k) = \int dq \left[\alpha(k,q)\hat{c}_{\text{in}}(q) + \beta(k,q)\hat{c}^{\dagger}_{\text{in}}(q)\right] \tag{35}$$

and likewise for the conjugate. Inserting this expansion into (32) with the 'in' vacuum we find that the number of states $\mathcal{N}(k,k')$ is,

$$\mathcal{N}(k,k') = \int dq \, \beta^*(k,q)\beta(k',q) \tag{36}$$

It should immediately be clear that this vanishes for any $|k|$ or $|k'|$ less than k_0. From (29) and (27) the 'in' and 'out' modes are identical when $\lambda_k \le 0$, and therefore $\beta(k,q)$ vanishes by the orthonormality relations (16). The only non-vanishing β coefficients are when $\lambda_k > 0$. From the forms of (28) and (23) we find that,

$$\alpha(k,q) = \delta(k-q)\frac{1}{\sqrt{2}}\left(e^{1/3\lambda_k} + \frac{1}{2}e^{-1/3\lambda_k}\right) \tag{37}$$

in this case. One of the relations amongst the Bogolubov coefficients is,

$$\int dq \left[\alpha^*(k,q)\alpha(k',q) - \beta^*(k,q)\beta(k',q)\right] = \delta(k-k') \tag{38}$$

from which we may directly obtain the number of states for positive effective cosmological constant. Altogether we find that,

$$\mathcal{N}(k,k') = \begin{cases} \frac{1}{2}\left(e^{1/3\lambda_k} - \frac{1}{2}e^{-1/3\lambda_k}\right)^2 \delta(k-k') & \text{if } \lambda_k > 0 \\ 0 & \text{if } \lambda_k \le 0 \end{cases} \tag{39}$$

There are two four-form vacua with wavenumber $\pm k$ for each effective cosmological constant λ_k. Removing the delta function and summing over both vacua we find that the density of vacuum states for a large universe is,

$$\rho(|k|) = \left(e^{1/3\lambda_k} - \frac{1}{2}e^{-1/3\lambda_k}\right)^2 \Theta(\lambda_k) \tag{40}$$

where Θ is the Heaviside theta function. This is strongly peaked around $\lambda_k \to 0_+$, where it has the form,

$$\rho(|k|) \sim e^{2/3\lambda_k} = \exp\left(\frac{3\pi}{G_N \Lambda(|k|)}\right) \tag{41}$$

precisely the Baum-Hawking factor. The dominant four-form vacua selected by the third quantization are the ones that cancel the vacuum energy of the rest of the particle physics, resulting in a vanishing cosmological constant for a large universe.

5 Discussion

This has been a long argument and it worthwhile recalling the salient points. If the cosmological constant were purely due to the vacuum energy of the universe that energy density would have to be on the order of $10^{-47}\mathrm{GeV}^4$, many orders of magnitude smaller than any value naturally derived from particle physics. We are in serious need of some mechanism for exactly canceling the vacuum energy of the universe. We have taken a rather pragmatic approach to the problem in this paper, by introducing the simplest method we know of for allowing the cosmological constant to vary and then showing that the value selected by third quantization is zero.

We have introduced an additional field whose special property is that its vacuum configuration, its background, is undetermined. This behaviour is unlike what normally occurs in a particle physics theory where the background is dictated by symmetries, and by being the configuration with the lowest vacuum energy. This would be the stable vacuum state of the theory, whose energy is expressible in terms of the parameters of the theory. The additional field, however, has a continuous infinity of possible background configurations. For the sake of demonstration we have taken this field to be a four-form, as it has the requisite properties. The advantage of using a four-form for demonstration purposes is that it has only gravitational couplings and has no effect on the rest of the particle physics. Moreover, we can actually perform sensible calculations with it.

Each vacuum, labeled by k in the preceding sections, really corresponds to a different theory (much like the θ vacuum of QCD) and there is no relaxation mechanism for the four-form to prefer one configuration over another. The background state can be denoted $|\Omega_0; k\rangle$. The $|\Omega_0\rangle$ is the stable vacuum state of the particle physics (*e.g.* the spontaneously broken vacuum of the electroweak theory) and has vacuum energy density ρ_0 say. In inflationary cosmology [25] it would be the true vacuum state. The $|k\rangle$ is whatever configuration the four-form happens to be in. The cosmological constant arising from the coupling of gravitation to the above particle theory plus four-form is $\Lambda(k) = 8\pi G_N \rho_0 + 9\pi k^2/2G_N$ for the choice of vacuum $|\Omega_0; k\rangle$. The question we have posed is the following: given a large universe, what is the most likely vacuum state and the most likely cosmological constant? To answer this we have quantized the cosmology. We do not have to rely on any deep understanding of what a quantum gravity theory might be. Below the Planck scale we may use quantum cosmology as an effective theory for the gravitational sector, and a generic particle physics model for the matter sector.

The field equation of quantum cosmology is the Wheeler-DeWitt equation. The solution to this equation is often referred to as the wave function of everything. In our approach it is interpreted as the wave function of nothing. The techniques of quantum cosmology allow us to determine the vacuum state of a large universe. The solutions to the WDW equation form a complete set of modes, each one corresponding to a different k-vacuum for the four-form. Positive frequency modes correspond to an expanding universe in classically allowed regions of the parameter space. Demanding that the modes be positive frequency at either small or large scales determines the form of the wave functions. When we third quantize the theory we find that there are two third quantized vacua, with an absence of positive frequency modes at small or large scales respectively.

In this picture the state vector in third quantization is that with no positive frequency modes at small scales. This is because we began with an effective description of gravity valid only on scales larger than the Planck length. We have required that the third quantized theory reflect this for consistency. It is then a straightforward task to evaluate the number of k-modes at large scales for this state vector. We have avoided the metaphysical interpretation of this as parthenogenesis or 'creation from nothing' and of the quanta as 'universes' being created or destroyed. Third quantization, as we have used it, is merely a method for determining the distribution of particle physics vacua when there is no other means for doing so. We have found that the dominant vacua for a large universe are the ones with vanishing cosmological constant.

In general the particle physics vacuum is a superposition of the k-vacua,

$$|\Omega\rangle = \int dk\, c(k)|\Omega_0; k\rangle$$

and the analysis of the third quantized theory has led us to the conclusion that, $|c(k)|^2 \sim \exp(3\pi/G_N\Lambda(k))$. Thus $|\Omega\rangle$ is saturated by the vacua possessing zero cosmological constant. If we are only interested in classical cosmology then we may regard this result as a superselection rule for the particle physics.

ACKNOWLEDGMENTS

We would like to thank Serge Rudaz for his comments on this work. MJD acknowledges support under grant DE-AC02-83ER40105, and LGJ under DE-AS06-88ER40423.

References

[1] S.Weinberg, Rev. Mod. Phys. **61** (1989) 1.

[2] B.S.DeWitt, Phys. Rev. **160** (1967) 1113;
J.A.Wheeler, in *Battelle Recontres*, eds. C.DeWitt and J.A.Wheeler (Benjamin, New York, 1968).

[3] V.A.Rubakov, Phys. Lett. **214B** (1988) 503.

[4] M.McGuigan, Phys. Rev. **D38** (1988) 3031;
A.Hosoya and M.Morikawa, Phys. Rev. **D39** (1989) 1123;
M.McGuigan, Phys. Rev. **D39** (1989) 2229.

[5] S.B.Giddings and A.Strominger, Nucl. Phys. **B321** (1989) 481;
W.Fischler, I.Klebanov, J.Polchinski and L.Susskind, Nucl. Phys. **B327** (1989) 157.

[6] E.Baum, Phys. Lett. **133B** (1983) 185.

[7] S.W.Hawking, Phys. Lett. **134B** (1984) 403.

[8] J.B.Hartle and S.W.Hawking, Phys. Rev. **D28** (1983) 2690;
S.W.Hawking, in *Relativity, Groups and Topology II*, eds. B.S.DeWitt and R.Stora (North-Holland, Amsterdam, 1984).

[9] A.Vilenkin, Phys. Rev. **D30** (1984) 509.

[10] S.W.Hawking, Phys. Lett. **195B** (1987) 337;
S.W.Hawking, Phys. Rev. **D37** (1988) 904;
S.B.Giddings and A.Strominger, Nucl. Phys. **B306** (1988) 890;
S.Coleman, Nucl. Phys. **B307** (1988) 867.

[11] S.Coleman, Nucl. Phys. **B310** (1988) 643;
I.Klebanov, T.Banks and L.Susskind, Nucl. Phys. **B317** (1989) 665.

[12] J.Polchinski, Phys. Lett. **219B** (1989) 251;
W.G.Unruh, Phys. Rev. **D40** (1989) 1053.

[13] W.Fischler and L.Susskind, Phys. Lett. **217B** (1989) 48;
J.Preskill, Nucl. Phys. **B323** (1989) 141;
S.Coleman and K.Lee, Phys. Lett. **221B** (1989) 242;
J.Polchinski, Nucl. Phys. **B325** (1989) 619.

[14] M.J.Duncan and L.G.Jensen, Nucl. Phys. **B328** (1989) 171.

[15] A.Aurelia, H.Nicolai and P.Townsend, Nucl. Phys. **B176** (1980) 509;
M.J.Duff and P.van Nieuwenhuizen, Phys. Lett. **94B** (1980) 179;
P.G.O.Freund and M.Rubin, Phys. Lett. **97B** (1980) 233.

[16] M.Henneaux and C.Teitelboim, Phys. Lett. **143B** (1984) 415;
E.Witten, in *Proceedings of the 1983 Shelter Island Conference on Quantum Field Theory and Fundamental Problems of Physics*, eds. R.Jackiw, N.Khuri, S.Weinberg and E.Witten (MIT Press, 1985);
J.D.Brown and C.Teitelboim, Phys. Lett. **195B** (1987) 177.

[17] M.J.Duff, Phys. Lett. **226B** (1989) 36.

[18] M.J.Duncan and L.G.Jensen, Nucl. Phys. **B336** (1990) 100.

[19] J.M.York, Phys. Rev. Lett. **28** (1972) 1082;
G.W.Gibbons and S.W.Hawking, Phys. Rev. **D15** (1977) 2752.

[20] R.Arnowitt, S.Deser and C.W.Misner, in *Gravitation, an Introduction to Current Research*, ed. L.Witten (Wiley, New York, 1962)

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

[22] M.Abramowitz and I.A.Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

[23] V.G.Lapchinsky and V.A.Rubakov, Acta. Phys. Polonica **B10** (1979) 1041;
T.Banks, W.Fischler and L.Susskind, Nucl. Phys. **B262** (1985) 159;
J.J.Halliwell and S.W.Hawking, Phys. Rev. **D31** (1985) 1777;
R.Brout and G.Venturi, Phys. Rev. **D39** (1989) 2436;
J.J.Halliwell, Santa Barbara preprint NSF-ITP-88-126 (1988).

[24] N.N.Bogolubov, Sov. Phys. JETP, **7** (1958) 51.

[25] A.H. Guth, Phys. Rev. **D23** (1981) 347;
A.D. Linde, Phys. Lett. **108B** (1982) 389;
A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. **48** (1982) 1220;
A.D. Linde, Phys. Lett. **129B** (1983) 177;
The following three references contain reviews on Inflation:
M.S. Turner, in *Proceedings of the Cargèse School on Fundamental Physics and Cosmology*, eds. J. Audouze and J. Tran Than Van (Editions Frontières, 1985);
R.H. Brandenberger, Rev. Mod. Phys. **57** (1985) 1;
K.A. Olive, Phys. Rep. **190** (1990) 307.