Advanced Science Letters
Vol. 2, 174–183, 2009

Dark Energy and Its Implications for Gravity

T. Padmanabhan

Inter University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India

The cosmological constant is the most economical candidate for dark energy. No other approach really alleviates the difficulties faced by the cosmological constant because, in all other attempts to model the dark energy, one still has to explain why the bulk cosmological constant (treated as a low-energy parameter in the action principle) is zero. I argue that the until the theory is made invariant under the shifting of the Lagrangian by a constant, one cannot obtain a satisfactory solution to the cosmological constant problem. This is impossible in any generally covariant theory with the conventional low-energy matter action, if the metric is varied in the action to obtain the field equations. I review an alternative perspective in which gravity arises as an emergent, long wavelength phenomenon, and can be described in terms of an effective theory using an action associated with null vectors in the spacetime. This action is explicitly invariant under the shift of the energy momentum tensor $T_{ab} \to T_{ab} + \Lambda g_{ab}$ and any bulk cosmological constant can be gauged away. Such an approach seems to be necessary for addressing the cosmological constant problem and can easily explain why its bulk value is zero. I describe some possibilities for obtaining the observed value from quantum gravitational fluctuations.

1. WHY DO WE BELIEVE IN DARK ENERGY?

The simplest possible universe one could imagine will contain just baryons and radiation. However, host of astronomical observations available since mid-70s indicated that the bulk of the matter in the universe is nonbaryonic and dark. Around the same time, the theoretical prejudice for $\Omega_{\rm tot} = 1$ gained momentum,[a] largely led by the inflationary paradigm.[1] During the eighties, this led many theoreticians to push (as usual wrongly!) for a model with $\Omega_{\rm tot} \approx \Omega_{\rm DM} \approx 1$ in spite of the fact that host of astronomical observations demanded that $\Omega_{\rm DM} \simeq 0.2 - 0.3$. The indications that the universe indeed has another component of energy density—so that $\Omega_{\rm tot} = 1$ can be reconciled with $\Omega_{\rm DM} \simeq 0.2 - 0.3$—started accumulating in the late eighties and early nineties. Early analysis of several observations[2] indicated that this component is unclustered and has negative pressure. This is confirmed dramatically by the supernova observations in the late nineties (see Ref. [3]; for a critical look at the data, see e.g., Ref. [4]). The observations suggest that the missing component contributes $\Omega_{\rm DE} \cong 0.60 - 0.75$ and has an *equation-of-state parameter* $w \equiv p/\rho \lesssim -0.78$. Treated as a fluid, this component has negative pressure (assuming positive energy density) and has been dubbed as *dark energy*.

The simplest choice for such dark energy with negative pressure is the cosmological constant which is a term that can be added to Einstein's equations. This term *acts like a fluid* with an equation of state $p_{\rm DE} = -\rho_{\rm DE}$. Combining this with all other observations,[5–8] we end up with a rather strange composition for the universe with $0.98 \lesssim \Omega_{\rm tot} \lesssim 1.08$ in which radiation (R), baryons (B), dark matter, made of weakly interacting massive particles (DM) and dark energy (DE) contributes $\Omega_R \simeq 5 \times 10^{-5}$, $\Omega_B \simeq 0.04$, $\Omega_{\rm DM} \simeq 0.26$, $\Omega_{\rm DE} \simeq 0.7$, respectively.

The key observational feature of dark energy—which dominates over everything else today in such a universe—is that it leads to an accelerated expansion of the universe. When treated as a fluid with a stress tensor $T^a_b = {\rm dia}\ (\rho, -p, -p, -p)$, it has an equation state $p = w\rho$ with $w \lesssim -0.8$ at the present epoch. In general relativity, the source of geodesic acceleration is $(\rho + 3p)$ and not ρ. As long as $(\rho + 3p) > 0$, gravity remains attractive while $(\rho + 3p) < 0$ can lead to 'repulsive' gravitational effects. In other words, dark energy with sufficiently negative pressure will accelerate the expansion of the universe, once it starts dominating over the normal matter. This is precisely what is established from the study of high redshift supernova, which can be used to determine the expansion rate of the universe in the past.[3,9]

While most physicists will look at such a weird composition for the universe with the suspicion it deserves, cosmologists have unhesitatingly accepted such a 'concordance model' over the last one decade. The reason is very simple. The concordance model is mandated by a host of observations and the cosmological paradigm based on such a composition is remarkably successful. This paradigm is a complex mix of several ingredients and works[10] broadly as follows:

(a) The basic idea is that if small fluctuations in the energy density existed in the early universe, then gravitational instability can amplify them leading to structures like galaxies etc. which exist today. The popular procedure for generating these fluctuations is

[a] It is convenient to measure the energy densities of the different species, which drive the expansion of the universe, in terms of this *critical density* using the dimensionless parameters $\Omega_i = \rho_i/\rho_c$ (with i denoting the different components like baryons, dark matter, radiation, etc.). The critical energy density is defined as $\rho_c = 3H_0^2/8\pi G$ where $H_0 = \dot{a}/a$ is the expansion rate of the universe.

1936-6612/2009/2/174/010 doi:10.1166/asl.2009.1024

based on the idea that if the very early universe went through an inflationary phase,[1] then the quantum fluctuations of the field driving the inflation can lead to energy density fluctuations.[11,12]
(b) While the inflationary models are far from unique and hence lacks predictive power, it is certainly possible to construct models of inflation such that these fluctuations are described by a Gaussian random field and are characterized by a power spectrum of the form $P(k) = Ak^n$ with $n \simeq 1$. The inflationary models cannot predict the value of the amplitude A in an unambiguous manner. But it can be determined from CMBR observations and the inflationary model parameters can be fine-tuned to reproduce the observed value. The CMBR observations are consistent with the inflationary model for the generation of perturbations and gives $A \simeq (28.3h^{-1}Mpc)^4$ and $n \lesssim 1$. (The first results were from COBE[13] and WMAP etc.[5] has re-confirmed them with far greater accuracy.)
(c) One can evolve the initial perturbations by a well understood linear perturbation theory when the perturbation is small. But when $\delta \approx (\delta\rho/\rho)$ is comparable to unity the perturbation theory breaks down and one has to resort to numerical simulations[14] or theoretical models based on approximate ansatz[15,16] to understand their evolution—especially the baryonic part, that leads to observed structures in the universe.

The resulting model, which is characterized essentially by seven numbers [$h \approx 0.7$ describing the current rate of expansion; $\Omega_{\rm DE} \simeq 0.7$, $\Omega_{\rm DM} \simeq 0.26$, $\Omega_B \simeq 0.04$, $\Omega_R \simeq 5 \times 10^{-5}$ giving the composition of the universe; the amplitude $A \simeq (28.3h^{-1}Mpc)^4$ and the index $n \simeq 1$ of the initial perturbations] seems to be quite successful in explaining the observations. This is a nontrivial measure of success since the paradigm could have been falsified by observations on several counts. For example, the values of $\Omega_B h^2$ could be constrained by CMBR observations as well as from the deuterium abundance from big bang nucleosynthesis. These two probe the universe at widely different epochs (a few hundred thousand years after big bang compared a few minutes) using very different techniques. The results, nevertheless lead to the same number. Such tests have led to faith in the concordance model including the existence of dark energy. So, even though we do not understand our universe, we have been quite successful in parametrising our ignorance in terms of well-chosen numbers.

2. WHAT IF DARK ENERGY IS JUST COSMOLOGICAL CONSTANT?

After such a brief overview, I will concentrate on the dark energy and the issues it rises (for a few of the recent reviews, see Ref. [17]). The concordance model, as defined above, uses the fact that the simplest model for a fluid with negative pressure is *not a fluid at all* but the cosmological constant with $w = -1, \rho = -p =$ constant. The cosmological constant introduces a fundamental length scale in the theory $L_\Lambda \equiv H_\Lambda^{-1}$, related to the constant dark energy density $\rho_{\rm DE}$ by $H_\Lambda^2 \equiv (8\pi G\rho_{\rm DE}/3)$. Though, in classical general relativity, based on G, c and L_Λ, it is not possible to construct any dimensionless combination from these constants, when one introduces the Planck constant, $\hbar$, it is possible to form the dimensionless combination $\lambda = H_\Lambda^2(G\hbar/c^3) \equiv (L_P^2/L_\Lambda^2)$. Observations then require $(L_P^2/L_\Lambda^2) \lesssim 10^{-123}$ requiring enormous fine tuning.[b]

In the earlier days, this was considered puzzling but most people believed that this number λ is actually zero. The cosmological constant problem in those days was to understand why it is strictly zero. Usually, the vanishing of a constant (which could have appeared in the low energy sector of the theory) indicates an underlying symmetry of the theory. For example, the vanishing of the mass of the photon is closely related to the gauge invariance of electromagnetism. No such symmetry principle is known to operate at low energies which made this problem very puzzling. There is a symmetry—called supersymmetry—which does ensure that $\lambda = 0$ but it is known that supersymmetry is broken at sufficiently high energies and hence cannot explain the observed value of λ.

Given the observational evidence for dark energy in the universe and the fact that the simplest candidate for dark energy, consistent with all observations today, is a cosmological constant with $\lambda \approx 10^{-123}$ the cosmological constant problem has got linked to the problem of dark energy in the universe. So, if we accept the simplest interpretation of the current observations, we need to explain why cosmological constant is non zero and has this small value. It should, however, be stressed that these two—the cosmological constant problem and the explaining of dark energy—are logically independent issues. *Even if all the observational evidence for dark energy goes away we still have a cosmological constant problem—viz., explaining why λ is zero.*

There is another, related, aspect to cosmological constant problem which need to be stressed. In conventional approach to gravity, one derives the equations of motion from a Lagrangian $\mathscr{L}_{\rm tot} = \mathscr{L}_{\rm grav}(g) + \mathscr{L}_{\rm matt}(g, \phi)$ where $\mathscr{L}_{\rm grav}$ is the gravitational Lagrangian dependent on the metric and its derivative and $\mathscr{L}_{\rm matt}$ is the matter Lagrangian which depends on both the metric and the matter fields, symbolically denoted as ϕ. In such an approach, the cosmological constant can be introduced via two different routes which are conceptually different but operationally the same. First, one may decide to take the gravitational Lagrangian to be $\mathscr{L}_{\rm grav} = (2\kappa)^{-1}(R - 2\Lambda_g)$ where Λ_g is a parameter in the (low energy effective) action just like the Newtonian gravitational constant κ. The second route is by shifting the matter Lagrangian by $\mathscr{L}_{\rm matt} \to \mathscr{L}_{\rm matt} - 2\lambda_m$. Such a shift is clearly equivalent to adding a cosmological constant $2\kappa\lambda_m$ to the $\mathscr{L}_{\rm grav}$. In general, what can be observed through gravitational interaction is the combination $\Lambda_{\rm tot} = \Lambda_g + 2\kappa\lambda_m$.

It is now clear that there are two distinct aspects to the cosmological constant problem. The first question is why $\Lambda_{\rm tot}$ is very small when expressed in natural units. Second, since $\Lambda_{\rm tot}$ could have had two separate contributions from the gravitational and matter sectors, why does the *sum* remain so fine tuned? This question is particularly relevant because it is believed that our universe went through several phase transitions in the course of its evolution, each of which shifts the energy momentum tensor of matter by $T^a_b \to T^a_b + L^{-4}\delta^a_b$ where L is the scale characterizing the transition. For example, the GUT and Weak Interaction scales are about $L_{\rm GUT} \approx 10^{-29}$ cm, $L_{\rm SW} \approx 10^{-16}$ cm respectively

[b]This is, of course, the party line. But it might help to get some perspective on how enormous, the 'enormous' really is. To begin with note that, the sensible particle physics convention considers ratios of length/energy scales and not their *squares*. This leads to $(L_P/L_\Lambda) \sim 10^{-61}$. In standard model of particle physics the ratio between Planck scale to neutrino mass scale is 10^{19} GeV/10^{-2} eV $\sim 10^{30}$ for which we have no theoretical explanation. So when we worry about the fine tuning of cosmological constant without expressing similar worries about standard model of particle physics, we are essentially assuming that 10^{30} is not a matter for concern but 10^{61} is. This subjective view is defensible but needs to be clearly understood.

which are tiny compared to L_Λ. Even if we take a more pragmatic approach, the observation of Casimir effect in the lab sets a bound that $L < \mathcal{O}(1)$ nanometer, leading to a ρ which is about 10^{12} times the observed value.[18]

Finally, I will comment on two other issues related to cosmological constant which appear frequently in the literature. The first one is what could be called the "why now" problem of the cosmological constant. How come the energy density contributed by the cosmological constant (treated as the dark energy) is comparable to the energy density of the rest of the matter at the *current epoch* of the universe? I do not believe this is an *independent* problem; if we have a viable theory predicting a particular numerical value for λ, then the energy density due to this cosmological constant will be comparable to the rest of the energy density at *some* epoch. So the real problem is in understanding the numerical value of λ; once that problem is solved the 'why now' issue will take care of itself. In fact, we do not have a viable theory to predict the current energy densities of *any* component which populates the universe, let alone the dark energy! For example, the energy density of radiation today is computed from its temperature which is an observed parameter—there is no theory which tells us that this temperature has to be 2.73 K when, say, galaxy formation has taken place for certain billion number of years. Neither do we have a theory which predicts the value of Ω_R/Ω_B at the present epoch. So the really important issue is to fix the numerical value of L_Λ in terms of Planck length.[c]

One also notices in the literature a discussion of the contribution of the zero point energies of the quantum fields to the cosmological constant which is often misleading, if not incorrect. What is usually done is to attribute a zero-point-energy $(1/2)\hbar\omega$ to each mode of the field and add up all these energies with an ultra violet cut-off. For an electromagnetic field, for example, this will lead to an integral proportional to

$$\rho_0 = \int_0^{k_{\max}} dk\ k^2 \hbar k \propto k_{\max}^4 \tag{1}$$

which will give $\rho_0 \propto L_P^{-4}$ if we invoke a Planck scale cut-off with $k_{\max} = L_P^{-1}$. It is then claimed that, this ρ_0 will contribute to the cosmological constant. There are several problems with such a naive analysis. First, the ρ_0 computed above can be easily eliminated by the normal ordering prescription in quantum field theory and what one really should compute is the *fluctuations* in the vacuum energy—not the vacuum energy itself. Second, even if we take the nonzero value of ρ_0 seriously, it is not clear this has anything to do with a cosmological constant. The energy momentum tensor due to the cosmological constant has a very specific form $T^a_b \propto \delta^a_b$ and its trace is nonzero. The electromagnetic field, for example, has a stress tensor with zero trace, $T^a_a = 0$; hence in the vacuum state the expectation value of the trace, $\langle \mathrm{vac}|T^a_a|\mathrm{vac}\rangle$, will vanish, showing that the equation of state of the bulk electromagnetic vacuum is still $\rho_0 = 3p_0$ which does not lead to a cosmological constant . (The trace anomaly will not work in the case of electromagnetic field.) So the naive calculation of vacuum energy density with a cutoff and the claim that it contributes to cosmological constant is not an accurate statement in many cases.

[c]For example, if some nonperturbative quantum gravity effect involving the exponential of a semiclassical action—or something similar—lead to a perfectly reasonable looking factor $L_\Lambda/L_P = \exp(\sqrt{2}\pi^4) \approx 10^{60}$, then all the issues are resolved.

3. WHAT IF DARK ENERGY IS NOT THE COSMOLOGICAL CONSTANT?

Based on some of these misgivings about the cosmological constant many people have tried to come up with alternative explanations for the dark energy. These attempts can be divided into two broad categories. The first set of ideas—which accounts for the bulk of the published research—assumes that the cosmological constant is zero for reasons unknown, and invokes some exotic physics (usually using scalar fields, higher dimensional models etc.). The second set also assumes that the cosmological constant is zero for reasons unknown and tries to explain the cosmological observations by some conventional, less esoteric physics.

The first set of approaches is conceptually no better compared to cosmological constant and it is very doubtful whether this—rather popular—approach, based on scalar fields, has helped us to understand the nature of the dark energy at any deeper level. These models, viewed objectively, suffer from several shortcomings: The most serious problem with them is that they have no predictive power. As can be explicitly demonstrated, virtually every form of expansion history $a(t)$ can be modeled[19,20] by a suitable "designer" scalar field potential $V(\phi)$. What is more, the scalar field potentials used in the literature have no natural field theoretical justification. All of them are non-normalizable in the conventional sense and have to be interpreted as a low energy effective potential in an ad hoc manner. Observationally, one key difference between cosmological constant and scalar field models is that the latter lead to a $(p/\rho) \equiv w(a)$ which varies with time. So they are worth considering if the observations have suggested a varying w, or if observations have ruled out $w = -1$ at the present epoch. However, all available observations are consistent with cosmological constant ($w = -1$) and—in fact—the possible variation of w is strongly constrained.[21] Further, it can be shown that even when $w(t)$ is determined by observations, it is not possible to proceed further and determine the nature of the scalar field Lagrangian. (See the first paper in Ref. [4] for an explicit example of such a construction.)[d]

Let us next consider the second set of approaches in which—again—we assume that the cosmological constant is zero because of some unknown reason but try to explain the observed acceleration of the universe in terms of reasonably conservative physics. One of the *least* esoteric ideas in this direction is that the cosmological constant term in the equations arises because we have not calculated the energy density driving the expansion of the universe correctly.

This idea arises as follows: The energy momentum tensor of the real universe, $T_{ab}(t, \mathbf{x})$ is inhomogeneous and anisotropic. If we could solve the exact Einstein's equations $G_{ab}[g] = \kappa T_{ab}$ with it as the source we will be led to a complicated metric g_{ab}. The metric describing the large scale structure of the universe should be obtained by averaging this exact solution over a large enough scale, leading to $\langle g_{ab}\rangle$. But since we cannot solve exact Einstein's equations, what we actually do is to average the stress tensor

[d]As an aside, let us note that in drawing conclusions from the observational data, one should be careful about the hidden assumptions in the statistical analysis. Claims regarding w depends crucially on the data sets used, priors which are assumed and possible parameterizations which are adopted. (For more details related to these issues, see the last reference in Ref. [21].) It is fair to say that all currently available data is consistent with $w = -1$. Further, there is some amount of tension between WMAP and SN-Gold data with the recent SNLS data[9] being more concordant with WMAP than the SN Gold data.

first to get $\langle T_{ab}\rangle$ and *then* solve Einstein's equations. But since $G_{ab}[g]$ is nonlinear function of the metric, $\langle G_{ab}[g]\rangle \neq G_{ab}[\langle g\rangle]$ and there is a discrepancy. This is most easily seen by writing

$$G_{ab}[\langle g\rangle] = \kappa[\langle T_{ab}\rangle + \kappa^{-1}(G_{ab}[\langle g\rangle] - \langle G_{ab}[g]\rangle)] \equiv \kappa[\langle T_{ab}\rangle + T_{ab}^{\text{corr}}] \quad (2)$$

If—based on observations—we take the $\langle g_{ab}\rangle$ to be the standard Friedman metric, this equation shows that it has, as its source, *two* terms: The first is the standard average stress tensor and the second is a purely geometrical correction term $T_{ab}^{\text{corr}} = \kappa^{-1}(G_{ab}[\langle g\rangle] - \langle G_{ab}[g]\rangle)$ which arises because of nonlinearities in the Einstein's theory that leads to $\langle G_{ab}[g]\rangle \neq G_{ab}[\langle g\rangle]$. If this term can mimic the cosmological constant at large scales there will be no need for dark energy and—as a bonus—one will solve the "why now" problem!

To make this idea concrete, we have to identify an effective expansion factor $a_{\text{eff}}(t)$ of an inhomogeneous universe (after suitable averaging), and determine the equation of motion satisfied by it. The hope is that it will be sourced by terms so as to have $\ddot{a}_{\text{eff}}(t) > 0$ while the standard matter (with $(\rho+3p) > 0$) leads to deceleration of standard expansion factor $a(t)$. Since any correct averaging of positive quantities in $(\rho+3p)$ will not lead to a negative quantity, the real hope is in defining $a_{\text{eff}}(t)$ and obtaining its dynamical equation such that $\ddot{a}_{\text{eff}}(t) > 0$. In spite of some recent attention this idea has received[22] it is doubtful whether it will lead to the correct result when implemented properly. The reasons for my skepticism are the following:

It is, of course, obvious that T_{ab}^{corr} is—mathematically speaking—non-zero (for an explicit computation, in a completely different context of electromagnetic plane wave, see Ref. [23]); the real question is how big is it compared to T_{ab}. When properly done, it seems unlikely that we will get a large effect for the simple reason that the amount of mass which is contained in the nonlinear regimes in the universe today is subdominant. Any calculation in linear theory or any calculation in which special symmetries are invoked will be inconclusive and untrustworthy in settling this issue. (Several papers on LTB models with mutually contradictory results can be cited as evidence for this!) There is also a serious issue of identifying a suitable analogue of expansion factor from an averaged geometry, which is nontrivial and it is not clear that the answer will be unique. To illustrate this point by an extreme example, suppose we decide to call $a(t)^n$ with, say $n > 2$ as the effective expansion factor i.e., $a_{\text{eff}}(t) = a(t)^n$; obviously $\ddot{a}_{\text{eff}}$ can be positive ('accelerating universe') even with $\ddot{a}$ being negative. So, unless one has a *unique* procedure to identify the expansion factor of the average universe, it is difficult to settle the issue. Finally this approach is strongly linked to explaining the acceleration as observed by SN. Even if we decide to completely ignore all SN data, we still have reasonable evidence for dark energy and it is not clear how this approach can tackle such evidence.

Another equally conservative explanation for the cosmic acceleration will be that we are located in a large underdense region in the universe; so that, locally, the underdensity acts like negative mass and produces a repulsive force. While there has been some discussion in the literature[24] as to whether observations indicate such a local 'Hubble bubble,' this does not seem to be a tenable explanation that one can take seriously at this stage. For one thing, it is not clear whether such a model—in which acceleration arises from a local underdensity—can be reconciled with baryonic acoustic oscillations[8] and the measured value of Hubble constant.[7] (But whether we are embedded in a local void or not is an important question we should find the answer to—purely observationally—irrespective of whether it explains *all* of the observed cosmic acceleration.)

Finally, note that, for any of these ideas to work (scalar field models or more conventional ones), we first need to find a mechanism which will make the cosmological constant vanish. All the scalar field potentials require fine tuning of the parameters in order to be viable. The same comment also applies to the more conventional approaches discussed above. Given this situation, it is certainly worthwhile to consider alternative paradigms in which one has a hope for explaining why cosmological constant is zero. Then, one can hope to get the small value of the cosmological constant from possibly quantum gravitational considerations. This is what we will discuss next.

4. THE COSMOLOGICAL CONSTANT PROBLEM DEMANDS AN ALTERNATIVE PERSPECTIVE ON GRAVITY

Even if all the evidence for dark energy disappears within a decade, *we still need to understand why cosmological constant is zero* and much of what I have to say in the sequel will remain relevant. I stress this because there is a recent tendency to forget the fact that the problem of the cosmological constant existed (and was recognized as a problem) long before the observational evidence for dark energy, accelerating universe etc. cropped up. In this sense, cosmological constant problem has an important theoretical dimension which is distinct from what has been introduced by the observational evidence for dark energy. So, it is worth examining this idea in detail and ask how its 'problems' can be tackled. I will now argue that the cosmological constant problem arises essentially because of our misunderstanding of the nature of gravity and that its solution *demands* an alternative perspective in which the metric tensor is not a dynamical variable and gravity is treated as an emergent phenomenon—like elasticity.[25] In the later sections I will explicitly describe such a model.

In the conventional approach to gravity, the gravitational field equations are obtained from a Lagrangian $\mathscr{L}_{\text{tot}} = \mathscr{L}_{\text{grav}}(g) + \mathscr{L}_{\text{matt}}(g, \phi)$ by varying the metric, where $\mathscr{L}_{\text{grav}}$ is the gravitational Lagrangian dependent on the metric and its derivative and $\mathscr{L}_{\text{matt}}$ is the matter Lagrangian which depends on both the metric and the matter fields, symbolically denoted as ϕ. This total Lagrangian is integrated over the spacetime volume with the covariant measure $\sqrt{-g}d^4x$ to obtain the action.

Suppose we now add a constant $(-2\lambda_m)$ to the matter Lagrangian thereby inducing the change $\mathscr{L}_{\text{matt}} \to \mathscr{L}_{\text{matt}} - 2\lambda_m$. The equations of motion for matter are invariant under such a transformation which implies that—in the absence of gravity—we cannot determine the value of λ_m. The transformation $\mathscr{L} \to \mathscr{L}_{\text{matt}} - 2\lambda_m$ is a symmetry of the matter sector (at least at scales below the scale of supersymmetry breaking; we shall ignore supersymmetry in what follows). But, in the conventional approach, gravity breaks this symmetry. *This is the root cause of the cosmological constant problem.* As long as gravitational field equations are of the form $E_{ab} = \kappa T_{ab}$ where E_{ab} is some geometrical quantity (which is G_{ab} in Einstein's theory) the theory cannot be invariant under the shifts of the form $T^a_b \to T^a_b + \rho\delta^a_b$. Since such shifts are allowed by the matter sector, it is very difficult to imagine a definitive solution to cosmological constant problem within the conventional approach to gravity.

More precisely, consider any model of gravity satisfying the following three conditions: (1) The metric is varied in the action to obtain the equations of motion. (2) We demand full general covariance of the equations of motion. (3) The equations of motion for matter sector is invariant under the addition of a constant to the matter Lagrangian. Then, we can prove 'no-go' theorem that the cosmological constant problem cannot be solved in such model.[26] The proof is elementary. Our demand (2) of general covariance requires the matter action to be an integral over $\mathscr{L}_{\text{matter}}\sqrt{-g}$. The demand (3) now allows us to add a constant Λ, say, to $\mathscr{L}_{\text{matter}}$ leading to a coupling $\Lambda\sqrt{-g}$ between Λ and the metric g_{ab}. By our demand (1), when we vary g_{ab} the theory will couple to Λ through a term proportional to Λg_{ab} thereby introducing an arbitrary cosmological constant into the theory.

The power of the above 'no-go theorem' lies in its simplicity! It clearly shows that we cannot solve cosmological constant problem unless we drop one of the three demands listed in the above paragraph. Of these, we do not want to sacrifice general covariance encoded in (2); neither do we have a handle on low energy matter Lagrangian so we cannot avoid (3). So the only hope we have is to introduce an approach in which gravitational field equations are obtained from varying some degrees of freedom other than g_{ab} in a maximization principle. When the new degrees of freedom are varied in the action, the field equations must remain invariant under the shift $\mathscr{L}_{\text{matt}} \to \mathscr{L}_{\text{matt}} + \lambda_m$ of the matter Lagrangian $\mathscr{L}_{\text{matt}}$ by a constant λ_m. This will give us some kind of 'gauge freedom' to absorb any λ_m while maintaining general covariance.

Once we obtain a theory in which gravitational action is invariant under the shift $T_{ab} \to T_{ab} + \Lambda g_{ab}$, we would have succeeded in making gravity decouple from the bulk vacuum energy. *This—by itself—is considerable progress*; for example, this will explain why bulk cosmological constant—treated as a parameter in the low energy Lagrangian—is irrelevant and can be taken to be zero.[e] Of course, given the observations, there still remains the second issue of explaining the observed value of the cosmological constant. Once the bulk value of the cosmological constant (or vacuum energy) decouples from gravity, *classical* gravity becomes immune to cosmological constant; that is, the bulk classical cosmological constant can be gauged away. Any observed value of the cosmological constant has to be necessarily a *quantum* phenomenon arising as a relic of microscopic spacetime fluctuations.[27] There must exist a deep principle in quantum gravity which leaves its non-perturbative trace even in the low energy limit that appears as the cosmological constant.

5. GRAVITY AS AN EMERGENT PHENOMENON

I will now provide an alternative perspective on gravity[28–30] which will lead to field equations which are invariant under the shift $T_{ab} \to T_{ab} + \Lambda g_{ab}$. I argue that we have misunderstood the true nature of gravity because of the way the ideas evolved historically and show that, when seen with the 'right side up,' the description of gravity becomes remarkably simple and explains features which we never thought needed explanation!

The historical development of some of the ideas we are interested in is indicated in Table I. Einstein started with the Principle of Equivalence and—with a few thought experiments—motivated why gravity should be described by a metric of spacetime. This approach gave the correct backdrop for the equality of inertial and gravitational masses and the *kinematics* of gravity. But then he needed to write down the field equations which govern the *dynamics* of g_{ab} and here is where the trouble started. There is no good guiding principle which Einstein could use that leads in a natural fashion to $G_{ab} = \kappa T_{ab}$ or to the corresponding action principle (explaining several false starts he had). Sure, one can obtain them from a series of postulates but they just do not have the same compelling force as, for example, the Principle of Equivalence.

Nevertheless, we all accepted general relativity, Einstein's equations and their solutions; after all, they agreed with observations so well! But—conceptually—strange things happen as soon as: (i) we let the metric to be dynamical and (ii) allow for arbitrary coordinate transformations or, equivalently, observers on any timelike curve examining physics. *Horizons are inevitable in such a theory and they are always observer dependent.* This arises as follows: (i) Principle of equivalence implies that trajectories of light will be affected by gravity. So in any theory which links gravity to spacetime dynamics, we can have nontrivial null surfaces which block information from certain class of observers. (ii) Similarly, one can construct timelike congruences (e.g., uniformly accelerated trajectories) such that all the curves in such a congruence have a horizon. You can't avoid horizons. What is more, the horizon is *always* an observer dependent concept, even when it can be given a purely geometrical definition. For example, the $r = 2M$ surface in Schwarzschild geometry acts operationally as a horizon *only* for the class of observers who choose to stay at $r > 2M$ and not for the observers falling into the black hole.

Once we have horizons—which are inevitable—we get into more trouble. It is an accepted dictum that *all observers have a right to describe physics using an effective theory*[31] *based only on the variables (s)he can access.* (This was, of course, the lesson from renormalization group theory. To describe physics at 10 GeV you shouldn't need to know what happens at 10^{14} GeV in "good" theories.) This raises the famous question first posed by Wheeler to Bekenstein: What happens if you mix cold and hot tea and pour it down a horizon, erasing all traces of "crime" in

Table I. Conventional perspective of gravity.

	Principle of equivalence (Einstein∼1908)
⇒	Gravity is described by the metric g_{ab} (Einstein∼1908)
?	Postulate Einstein's equations *without a real guiding principle! (Einstein∼1915)*
⇒	Black hole solutions with horizons allowing the entropy of hot tea to be hidden (Wheeler∼1971)
⇒	Entropy of black hole horizon (Bekenstein 1972)
⇒	Temperature of black hole horizon (Hawking 1975)
⇒	Temperature of the Rindler horizon (Davies, Unruh 1975–76)

[e] The issue, of course, is related to the fact non-gravitational physics does not care about the absolute zero of energy while gravity does. Curiously, we do not have theoretical formalism of even non-gravitational physics—say, in standard quantum mechanics— which is *manifestedly* invariant under shifting of the origin of energy and depends only on energy differences.

increasing the entropy of the world?[f] The answer to such thought experiments *demands* that horizons should have an entropy which should increase when energy flows across it.

With hindsight, this is obvious. The Schwarschild horizon—or for that matter any metric which behaves locally like Rindler metric—has a temperature which can be identified by the Euclidean continuation.[33] If energy flows across a hot horizon $dE/T = dS$ leads to the entropy of the horizon. Again, historically, nobody—including Wheeler and Bekenstein—looked at the periodicity in the Euclidean time (in Rindler or Schwarzschild metrics) *before* Hawking's result came! And the idea of Rindler temperature came *after* that of black hole temperature! So in summary, the history proceeded as indicated in Table I: This historical sequence raises a some serious issues for which there is no satisfactory answer in the conventional approach:

- *How can horizons have temperature without the spacetime having a microstructure?*
  They simply cannot. Recall that the thermodynamic description of matter at finite temperature provides a crucial window into the existence of the corpuscular substructure of solids. As Boltzmann taught us, heat is a form of motion and we will not have the thermodynamic layer of description if matter is a continuum all the way to the finest scale and atoms did not exist! *The mere existence of a thermodynamic layer in the description is proof enough that there are microscopic degrees of freedom.*—In a solid or in a spacetime. In the conventional approach, we are completely at a loss to understand why horizons are hot or what kind of 'motion' is this 'heat.'
  To tackle this issue, it is necessary to abandon the usual picture of treating the metric as the fundamental dynamical degrees of freedom of the theory and treat it as providing a coarse grained description of the spacetime at macroscopic scales, somewhat like the density of a solid—which has no meaning at atomic scales.[25, g]
- *Why is it that Einstein's equations reduces to a thermodynamic identity for virtual displacements of a horizon?*
  Here is the first algebraic mystery—which has no explanation in conventional approach—suggesting a deep connection between the dynamical equations governing the metric and the thermodynamics of horizons. The first example was provided in Ref. [36] in which it was shown that, in the case of spherically symmetric horizons, Einstein's equations can be interpreted as a thermodynamic relation $TdS = dE + PdV$ arising out of virtual radial displacements of the horizon. Further work showed that this result is valid in *all* the cases for which explicit computation can be carried out—as diverse as the Friedmann models as well as rotating and time dependent horizons in Einstein's theory.[37] Treating them as just some solutions to Einstein's field equations we cannot understand these results.
- *Why is Einstein-Hilbert action is holographic with a surface term that encodes same information as the bulk?*
  The Einstein-Hilbert Lagrangian has the structure $L_{\text{EH}} \propto R \sim (\partial g)^2 + \partial^2 g$. In the usual approach the surface term arising from $L_{\text{sur}} \propto \partial^2 g$ has to be ignored or canceled to get Einstein's equations from $L_{\text{bulk}} \propto (\partial g)^2$. But there is a peculiar (again unexplained) relationship between L_{bulk} and L_{sur}:

$$\sqrt{-g}L_{\text{sur}} = -\partial_a \left(g_{ij} \frac{\partial \sqrt{-g} L_{\text{bulk}}}{\partial(\partial_a g_{ij})} \right) \tag{3}$$

  This shows that the gravitational action is 'holographic'[38, 39] with the same information being coded in both the bulk and surface terms making either one of them to be sufficient. It is well known that varying g_{ab} in L_{bulk} leads to the standard field equations. More remarkable is the fact that one can also obtain Einstein's equations from an action principle which uses only the surface term and the virtual displacements of horizons[40] *without* treating the metric as a dynamical variable.
- *Why does the surface term in Einstein-Hilbert action give the horizon entropy?*
  Yet another algebraic result which defies physical understanding! You first throw away the surface term in the action, vary the rest to get the field equations, find a solution with a horizon, compute its entropy—only to discover that the surface term you threw away is intimately related to the entropy.
- *And, most importantly, why do all these results hold for a much wider class of theories than Einstein gravity, like Lanczos-Lovelock models?*
  There are more serious 'algebraic accidents' in store. Recent work has shown that *all the thermodynamic features described above extend far beyond Einstein's theory.* The connection between field equations and the thermodynamic relation $TdS = dE + PdV$ is not restricted to Einstein's theory (GR) alone, but is in fact true for the case of the generalized, higher derivative Lanczos-Lovelock gravitational theory in D dimensions as well.[41] The same is true for the holographic structure of the action functional:[42] the Lanczos-Lovelock action has the same structure and—again—the entropy of the horizons is related to the surface term of the action.

I believe these (and several related features) are not algebraic accidents but indicate that we have been looking at gravity the wrong way around. In the proper perspective, these features should emerge as naturally as the equivalence of inertial and gravitational masses emerges in the geometric description of the kinematics of gravity. *These results show that the thermodynamic description is far more general than just Einstein's theory* and occurs in a wide class of theories in which the metric determines the structure of the light cones and null surfaces exist blocking the information. So instead of the historical path, I will proceed as in Table II reversing most of the arrows:[30]

Let me elaborate. Take an event P and introduce a local inertial frame (LIF) around it with coordinates X^a. Go from the LIF to a local Rindler frame (LRF) coordinates x^a by accelerating along, say, X-axis with an acceleration κ. This LRF and its local horizon $\mathcal{H}(x = 0)$ will exist within a region of size $L \ll \mathcal{R}^{-1/2}$ as long as $\kappa^{-1} \ll \mathcal{R}^{-1/2}$ where $\mathcal{R}$ is a typical component of curvature tensor.

[f] This is based on what Wheeler told me in 1985, from *his* recollection of events; it is also mentioned in his book.[32] I have heard somewhat different versions from other sources.

[g] The unknown, microscopic degrees of freedom of spacetime (which should be analogous to the atoms in the case of solids), should normally play a role only when spacetime is probed at Planck scales (which would be analogous to the lattice spacing of a solid[34]). So we normally expect the microscopic structure of spacetime to manifest itself only at Planck scales or near singularities of the classical theory. However, in a manner which is not fully understood, the horizons—which block information from certain classes of observers—link[35] certain aspects of microscopic physics with the bulk dynamics, just as thermodynamics can provide a link between statistical mechanics and (zero temperature) dynamics of a solid. The reason is probably related to the fact that horizons lead to infinite redshift, which probes *virtual* high energy processes; it is, however, difficult to establish this claim in mathematical terms.

Table II. Alternative perspective of gravity.

	Principle of equivalence
$\Rightarrow$	Gravity is described by the metric g_{ab}
$\Rightarrow$	Existence of local Rindler frames (LRFs) with horizons $\mathcal{H}$ around any event
$\Rightarrow$	Temperature associated with $\mathcal{H}$ is obtainable from the Euclidean continuation
$\Rightarrow$	Virtual displacements of $\mathcal{H}$ allow for flow of energy across a hot horizon hiding an entropy $dS = dE/T$ as perceived by a given observer
$\Rightarrow$	The local horizon must have an entropy, $S_{\rm grav}$
$\Rightarrow$	The dynamics should arise from maximizing the total entropy of horizon ($S_{\rm grav}$) plus matter (S_m) for all LRF's *without* varying the metric
$\Rightarrow$	The field equations are those of Lanczos-Lovelock gravity with Einstein's gravity emerging as the lowest order term
$\Rightarrow$	The theory is invariant under the shift $T_{ab} \to T_{ab} + \Lambda g_{ab}$ allowing the bulk cosmological constantto be 'gauged away'

Now people can pour tea across $\mathcal{H}$ as suggested by Wheeler. Alternatively, one can consider a virtual displacement of the $\mathcal{H}$ normal to itself engulfing the tea. Either way, some entropy will be lost to the outside observers unless the horizon has an entropy. That is, displacing a piece of local Rindler horizon should cost some entropy $S_{\rm grav}$, say. It is then natural to demand that the dynamics should follow from the prescription $\delta[S_{\rm grav} + S_{\rm matt}] = 0$.

All we need is the expressions for $S_{\rm matt}$ and $S_{\rm grav}$. I will now write down the general expressions for both,[29,30] such that they have the correct interpretation in the LRF. Assuming a $D(\geq 4)$ dimensional spacetime for the later convenience, I take:

$$S_{\rm grav} = -4\int_{\mathcal{V}} d^D x\sqrt{-g}P_{ab}^{cd}\nabla_c n^a \nabla_d n^b \qquad (4)$$

$$S_{\rm matt} = \int_{\mathcal{V}} d^D x\sqrt{-g}T_{ab}n^a n^b$$

where n^a is vector field which will reduce to the null normal $[\partial_a(T-X)]$ on the horizon $\mathcal{H}$, T_{ab} is the matter energy momentum tensor and P_{ab}^{cd} is defined below. For mathematical convenience, it is better to treat n_a as a vector field with a fixed norm $n_a n^a \equiv \epsilon$ rather than as a strictly null vector. This allows us to obtain the horizon as a limiting case of a 'stretched horizon' which is a timelike surface. [This, as well as some other subtleties in the variational principle are described in Ref. [29].] The $S_{\rm matt}$ is easy to understand. In the LRF, (with $-g_{tt} = 2\kappa x = g^{xx}$, $\sqrt{-g} = 1$) an infinitesimal spacetime region will contribute $T_{ab}n^b n^b d^3x\ dt = \delta E\ dt$ which on integration over t in the range $(0, \beta)$ where $\beta^{-1} = T = (\kappa/2\pi)$ gives $\delta S_{\rm matter} = \beta\delta E = \beta T_{ab}n^a n^b d^3x$ when the energy flows across a surface with normal n^a. Integrating, we get $S_{\rm matt}$ to which the expression in Eq. (5) reduces to in LRF. (For example, if T_{ab} is due to an ideal fluid at rest in LIF, $T_{ab}n^a n^b$ will contribute $(\rho + P)$, which—by Gibbs-Duhem relation—is just Ts where s is the entropy density. Integrating over $\sqrt{-g}d^4x = dt d^3x$ with $0 < t < \beta$ gives $S_{\rm matt}$.)

The $S_{\rm grav}$, on the other hand, is a general quadratic functional of the derivatives of n_a which is the form of entropy of an elastic solid, say, if n^a is the displacement field. Here we interpret it as the entropy cost for virtual displacement of horizon. The crucial requirement is that, dynamics for the *background spacetime* should emerge when we set $(\delta S_{\rm tot}/\delta n_a) = 0$ for *all* null vectors n^a (rather than an equation for n^a). Incredibly enough, this can be achieved if (and only if) (i) the tensor P_{abcd} has the algebraic symmetries similar to the Riemann tensor R_{abcd} and (ii) we have $\nabla_a P^{abcd} = 0 = \nabla_a T^{ab}$. One can now show that[29] such a tensor can be constructed as a a series in the powers of the derivatives of the metric:

$$P^{abcd}(g_{ij}, R_{ijkl}) = c_1\, P_{(1)}^{abcd}(g_{ij}) + c_2\, P_{(2)}^{abcd}(g_{ij}, R_{ijkl}) + \cdots \qquad (5)$$

where $c_1, c_2, \cdots$ are coupling constants with the *unique* m-th order term being $P_{ab}^{cd} \propto \partial\mathcal{L}_m^{(D)}/\partial R_{cd}^{ab}$ where $\mathcal{L}_m^{(D)}$ is the m-th order Lanczos-Lovelock Lagrangian.[40,43] Then maximizing $(S_{\rm grav} + S_m)$ gives:[29]

$$16\pi\left[P_b^{ijk}R_{ijk}^a - \frac{1}{2}\delta_b^a\mathcal{L}_m^{(D)}\right] = 8\pi T_b^a + \Lambda\delta_b^a \qquad (6)$$

These are identical to the field equations for Lanczos-Lovelock gravity with a cosmological constant arising as an undetermined integration constant. The lowest order term $P_{cd}^{ab} = (1/32\pi)(\delta_c^a\delta_d^b - \delta_d^a\delta_c^b)$ leads to Einstein's theory while the first order term gives the Gauss-Bonnet correction. One can show, in the general case of Lanczos-Lovelock theory, Eq. (5) *does* give the correct gravitational entropy justifying our choice. Remarkably enough, we can derive not only Einstein's theory but even Lanczos-Lovelock theory from a dual description in terms on the normalized vectors in spacetime, *without varying g_{ab} in an action functional!*

The crucial feature of the coupling between matter and gravity through $T_{ab}n^a n^b$ is that, under the shift $T_{ab} \to T_{ab} + \rho_0 g_{ab}$, the ρ_0 term in the action in Eq. (5) decouples from n^a and becomes irrelevant:

$$\int_{\mathcal{V}} d^D x\sqrt{-g}T_{ab}n^a n^b \to \int_{\mathcal{V}} d^D x\sqrt{-g}T_{ab}n^a n^b + \int_{\mathcal{V}} d^D x\sqrt{-g}\epsilon\rho_0 \qquad (7)$$

Since $\epsilon = n_a n^a$ is not varied when n_a is varied there is no coupling between ρ_0 and the dynamical variables n_a and the theory is invariant under the shift $T_{ab} \to T_{ab} + \rho_0 g_{ab}$. Of course, when n_a is null—in the limit of stretched horizon becoming the horizon—the second term vanishes. We see that just the condition $n_a n^a =$ constant on the dynamical variables have led to a 'gauge freedom' which allows an arbitrary integration constant to appear in the theory which can absorb the bulk cosmological constant.

6. SURFACE DEGREES OF FREEDOM AND THE OBSERVED VALUE OF COSMOLOGICAL CONSTANT

The description of gravity given above provides a natural backdrop for gauging away the bulk value of the cosmological constant since it decouples from the dynamical degrees of freedom in the theory. Once the bulk term is eliminated, what is observable through gravitational effects, in the correct theory of quantum gravity, should be the *fluctuations* in the vacuum energy. These fluctuations will be non-zero if the universe has a DeSitter horizon which provides a confining volume. In this paradigm the vacuum structure can readjust to gauge away the bulk energy density $\rho_{\rm UV} \simeq L_P^{-4}$ while quantum *fluctuations* can generate the observed value $\rho_{\rm DE}$. This boils down to explaining the numerical value of $L_\Lambda/L_P (\simeq \exp(\sqrt{2}\pi^4)) \approx 10^{60}$. *I do not yet know how to do this* but will provide arguments as to why the approach with surface degrees of freedom might hold the key.

I shall argue that, the quantum fluctuations that lead to the observed value of cosmological constant should scale with the surface area rather than the bulk volume to give the correct answer. This, in turn, suggests that the relevant degrees of freedom (which replaces the metric) should be linked to surfaces in spacetime rather than bulk regions—something we have already seen in the emergent gravity approach. The observed cosmological constant should then be a relic of quantum gravitational physics and should arise from degrees of freedom which scale as the surface area.

Consider a 3-dimensional region of size L with a bounding area which scales as L^2. Let us assume that we associate with this region N microscopic cells of size L_P each having a Poissonian fluctuation in energy of amount $E_P \approx 1/L_P$. Then the mean square fluctuation of energy in this region will be $(\Delta E)^2 \approx N L_P^{-2}$ corresponding to the energy density $\rho = \Delta E/L^3 = \sqrt{N}/L_P L^3$. If we make the usual assumption that $N = N_{vol} \approx (L/L_P)^3$, this will give

$$\rho = \frac{\sqrt{N_{\rm vol}}}{L_P L^3} = \frac{1}{L_P^4}\left(\frac{L_P}{L}\right)^{3/2} \quad \text{(bulk fluctuations)} \qquad (8)$$

On the other hand, if we assume that the relevant degrees of freedom scale as the surface area of the region, then $N = N_{sur} \approx (L/L_P)^2$ and the relevant energy density is

$$\rho = \frac{\sqrt{N_{\rm sur}}}{L_P L^3} = \frac{1}{L_P^4}\left(\frac{L_P}{L}\right)^2 = \frac{1}{L_P^2 L^2} \quad \text{(surface fluctuations)} \quad (9)$$

If we take $L \approx L_\Lambda$, the surface fluctuations give precisely the geometric mean, of the two energy scales $\rho_{\rm UV} = 1/L_P^4$ and $\rho_{\rm IR} = 1/L_\Lambda^4$ in natural units ($c = \hbar = 1$), which is the observed value of the energy density contributed by the cosmological constant. On the other hand, the bulk *fluctuations* lead to an energy density which is larger by a factor $(L/L_P)^{1/2}$. Of course, if we do not take fluctuations in energy but coherently add them, we will get $N/L_P L^3$ which is $1/L_P^4$ for the bulk and $(1/L_P)^4 (L_P/L)$ for the surface. In summary, we have the following hierarchy:

$$\rho = \frac{1}{L_P^4} \times \left[1, \left(\frac{L_P}{L}\right), \left(\frac{L_P}{L}\right)^{3/2}, \left(\frac{L_P}{L}\right)^2, \left(\frac{L_P}{L}\right)^4 \cdots\cdots\right] \quad (10)$$

in which the first one arises by coherently adding energies $(1/L_P)$ per cell with $N_{vol} = (L/L_P)^3$ cells; the second arises from coherently adding energies $(1/L_P)$ per cell with $N_{\rm sur} = (L/L_P)^2$ cells; the third one is obtained by taking *fluctuations* in energy and using $N_{\rm vol}$ cells; the fourth from energy fluctuations with $N_{\rm sur}$ cells; and finally the last one is the thermal energy of the DeSitter space if we take $L \approx L_\Lambda$; clearly the further terms are irrelevant due to this vacuum noise. Of all these, the only viable possibility is what arises if we assume that: (a) The number of active degrees of freedom in a region of size L scales as $N_{\rm sur} = (L/L_P)^2$. (b) It is the *fluctuations* in the energy that contributes to the cosmological constant[27, 38] and the bulk energy does not gravitate.

The role of energy fluctuations contributing to gravity also arises, more formally, when we study the question of *detecting* the energy density using gravitational field as a probe. Recall that a detector with a linear coupling to the *field* ϕ actually responds to $\langle 0|\phi(x)\phi(y)|0\rangle$ rather than to the field itself.[44] Similarly, one can use the gravitational field as a natural "detector" of energy momentum tensor T_{ab} with the standard coupling $L = \kappa h_{ab} T^{ab}$. Such a model was analyzed in detail in Ref. [45] and it was shown that the gravitational field responds to the two point function $\langle 0|T_{ab}(x)T_{cd}(y)|0\rangle$. In fact, it is essentially this fluctuations in the energy density which is computed in the inflationary models[1] as the *source* for gravitational field, as stressed in Ref. [12]. All these suggest treating the energy fluctuations as the physical quantity "detected" by gravity, when one incorporates quantum effects.

Quantum theory, especially the paradigm of renormalization group has taught us that the concept of the vacuum state depends on the scale at which it is probed. The vacuum state which we use to study the lattice vibrations in a solid, say, is not the same as vacuum state of the QED and it is not appropriate to ask questions about the vacuum without specifying the scale. If the cosmological constant arises due to the fluctuations in the energy density of the vacuum, then one needs to understand the structure of the quantum gravitational vacuum at cosmological scales. If the spacetime has a cosmological horizon which blocks information, the natural scale is provided by the size of the horizon, L_Λ, and we should use observables defined within the accessible region. The operator $H(<L_\Lambda)$, corresponding to the total energy inside a region bounded by a cosmological horizon, will exhibit fluctuations ΔE since vacuum state is not an eigenstate of *this* operator. A rigorous calculation (see the first reference in Ref. [27]) shows that the fluctuations in the energy density of the vacuum in a sphere of radius L_Λ is given by

$$\Delta\rho_{\rm vac} = \frac{\Delta E}{L_\Lambda^3} \propto L_P^{-2} L_\Lambda^{-2} \qquad (11)$$

The numerical coefficient will depend on c_1 as well as the precise nature of infrared cutoff radius; but it is a fact of life that a fluctuation of magnitude $\Delta\rho_{\rm vac} \simeq H_\Lambda^2/G$ will exist in the energy density inside a sphere of radius H_Λ^{-1} if Planck length is the UV cut off. On the other hand, since observations suggest that there is a $\rho_{\rm vac}$ of similar magnitude in the universe it seems natural to identify the two. Our approach explains why there is a *surviving* cosmological constant which satisfies $\rho_{\rm DE} = \sqrt{\rho_{\rm IR}\rho_{\rm UV}}$.

It is, of course, possible to give all kinds of arguments (mostly based on some version of 'holography') to motivate an expression like Eq. (11). In most of these approaches, L_Λ will be identified with a *time dependent* length scale in the Friedmann universe—Hubble radius, past horizon, future horizon ... —and one will try to see whether the resulting model agrees with observations. There are two issues I want to briefly discuss in this context.

First, I am not sure this will lead to a satisfactory solution; it could very well be that some unknown quantum gravitational effect will actually give the ratio $L_\Lambda/L_P \approx 10^{61}$ just as we expect some theory to eventually tell us why $m_v/M_{Pl} \approx 10^{-30}$. (I have already indicated a non-perturbative numerology: $L_\Lambda/L_P = \exp(\sqrt{2}\pi^4) \approx 10^{60}$!) Cosmological constant is then interpreted as small because it is a nonperturbative quantum relic.

Second, and more important, I stress that invoking Eq. (11) with some cosmological length scale for L_Λ is completely meaningless in the models of gravity in which the metric couples to the bulk energy density. Until we have a paradigm in place which allows us to ignore the bulk cosmological constant—which most of the ad hoc, holographic dark energy type models do not have—one cannot invoke such a procedure. This is particularly true in any model which leads to Eq. (11) through any kind of quantum fluctuation. All such approaches will require a UV cutoff (at Planck scale) to give a finite answer; once it is imposed, one will always get a bulk contribution $\rho_{\rm UV} \approx L_P^{-4}$ with the usual problems. It is only because *we* have a way of decoupling the bulk term from contributing to the dynamical equations that, we have a right to look at the subdominant term $L_P^{-4}(L_P/L_\Lambda)^2$. Approaches in which the sub-dominant term is introduced by an ad hoc manner are conceptually flawed since the bulk term cannot be ignored in these usual approaches to gravity. Getting the correct value of the cosmological constant from the energy fluctuations is not as difficult as understanding why the bulk value (which is larger by 10^{120}!) can be ignored. Our approach provides a natural framework for ignoring the bulk term—and as a bonus—the *possibility* of obtaining the right value for the cosmological constant from the fluctuations.

7. CONCLUSIONS

The simplest choice for the negative pressure component in the universe is the cosmological constant; other models based on scalar fields (as well as those based on branes etc. which I have not discussed) do not alleviate the difficulties faced by cosmological constant and—in fact—makes them worse. I have shown that it is impossible to solve the cosmological constant problem unless the gravitational sector of the theory is invariant under the shift $T_{ab} \rightarrow T_{ab} + \lambda_m g_{ab}$. Any approach which does not address this issue cannot provide a comprehensive solution to the cosmological constant problem. But general covariance requires us to use the measure $\sqrt{-g}d^Dx$ in D-dimensions in the action which will couple the metric (through its determinant) to the matter sector. Hence, as long as we insist on metric as the fundamental variable that is varied in an action principle, one cannot address this issue. So we need to introduce some other degrees of freedom and an effective action which, however, is capable of constraining the background metric.

An action principle, based on the normalized vector fields in spacetime, satisfies all these criteria mentioned above. The new action does not couple to the bulk energy density and maintains invariance under the shift $T_{ab} \rightarrow T_{ab} + \lambda_m g_{ab}$. What is more, the on-shell value of the action is related to the entropy of horizons showing the relevant degrees of freedom scales as the area of the bounding surface.

Since our formalism ensures that the bulk energy density does not contribute to gravity—*and only because of that*—it makes sense to compute the next order correction due to fluctuations in the energy density. I have not been able to compute this rigorously with the machinery available, but a plausible case can be made as how this approach might lead to the correct, observed, value of the cosmological constant.

References and Notes

1. D. Kazanas, *Ap. J. Letts.* 241, 59 **(1980)**; A. A. Starobinsky, *JETP Lett.* 30, 682 **(1979)**; *Phys. Lett. B* 91, 99 **(1980)**; A. H. Guth, *Phys. Rev. D* 23, 347 **(1981)**; A. D. Linde, *Phys. Lett. B* 108, 389 **(1982)**; A. Albrecht and P. J. Steinhardt, *Phys. Rev. Lett.* 48, 1220 **(1982)**; For a review, see e.g., J. V. Narlikar and T. Padmanabhan, *Ann. Rev. Astron. Astrophys.* 29, 325 **(1991)**; L. Alabidi and D. H. Lyth.
2. G. Efstathiou, W. J. Sutherland, and S. Maddox, *Nature* 348, 705 **(1990)**; T. Padmanabhan and D. Narasimha, *MNRAS* 259, 41P **(1992)**; G. Efstathiou, J. R. Bond, and S. D. M. White, *MNRAS*, 258, 1 **(1992)**; J. P. Ostriker and P. J. Steinhardt, *Nature* 377, 600 **(1995)**; J. S. Bagla, T. Padmanabhan, and J. V. Narlikar, *Comments on Astrophysics* 18, 275 **(1996)**.
3. S. J. Perlmutter et al., *Astrophys. J.* 517, 565 **(1999)**; A. G. Reiss et al., *Astron. J.* 116, 1009 **(1998)**; J. L. Tonry et al., *Ap. J.* 594, 1 **(2003)**; B. J. Barris, *Astrophys. J.* 602, 571 **(2004)**; A. G. Reiss et al., *Astrophys. J.* 607, 665 **(2004)**.
4. T. Padmanabhan and T. Roy Choudhury, *MNRAS* 344, 823 **(2003)**; T. Roy Choudhury and T. Padmanabhan, *Astron. Astrophys.* 429, 807 **(2005)**; H. K. Jassal, J. S. Bagla, and T. Padmanabhan, *Phys. Rev. D* 72, 103503 **(2005)**; S. Nesseris and L. Perivolaropoulos, *JCAP* 0702, 025 **(2007)**; Y. Wang and P. Mukherjee, *Phys. Rev. D* 76, 103533 **(2007)**.
5. P. de Bernardis et al., *Nature* 404, 955 **(2000)**; A. Balbi et al., *Ap. J.* 545, L1 **(2000)**; S. Hanany et al., *Ap. J.* 545, L5 **(2000)**; T. J. Pearson et al., *Astrophys. J.* 591, 556 **(2003)**; C. L. Bennett et al, *Astrophys. J. Suppl.* 148, 1 **(2003)**; D. N. Spergel et al., *Ap.J.S.* 148, 175 **(2003)**; B. S. Mason et al., *Astrophys. J.* 591, 540 **(2003)**; E. Komatsu et al.
6. W. J. Percival et al., *Mon. Not. Roy. Astron. Soc.* 337, 1068 **(2002)**; *MNRAS* 327, 1297 **(2001)**; X. Wang et al., **(2002)**, *Phys. Rev. D* 65, 123001; T. Padmanabhan and Shiv Sethi, *Ap. J.* 555, 125 **(2001)**. For a review of BBN, see S. Sarkar, *Rept. Prog. Phys.* 59, 1493 **(1996)**.
7. W. Freedman et al., *Astrophys. J.* 553, 47 **(2001)**; J. R. Mould et al., *Astrophys. J.* 529, 786 **(2000)**.
8. D. J. Eisenstein et al., *Astrophys. J.* 633, 560 **(2005)**; R. Angulo et al., *Mon. Not. Roy. Astron. Soc.* 383, 755 **(2008)**.
9. P. Astier et al., *Astron. Astrophys.* 447, 31 **(2006)**.
10. For recent reviews of cosmological paradigm, see, e.g., T. Padmanabhan, *AIP Conference Proceedings* 843, 111 **(2006)**; T. Padmanabhan 100 Years of Relativity-Space-time Structure: Einstein and Beyond, A. Ashtekar (Ed.), World Scientific, Singapore **(2005)**, pp. 175–204.
11. S. W. Hawking, *Phys. Lett. B* 115, 295 **(1982)**; A. A. Starobinsky, *Phys. Lett. B* 117, 175 **(1982)**; A. H. Guth and S.-Y. Pi, *Phys. Rev. Lett.* 49, 1110 **(1982)**; J. M. Bardeen et al., *Phys. Rev. D* 28, 679 **(1983)**; L. F. Abbott and M. B. Wise, *Nucl. Phys. B* 244, 541 **(1984)**. For a recent discussion with detailed set of references, See L. Sriramkumar and T. Padmanabhan, *Phys. Rev. D* 71, 103512 **(2005)**.
12. T. Padmanabhan, *Phys. Rev. Letts.* 60, 2229 **(1988)**; T. Padmanabhan, T. R. Seshadri, and T. P. Singh, *Phys. Rev. D* 39, 2100 **(1989)**.
13. G. F. Smoot et al., *Ap. J.* 396, L1 **(1992)**;
14. For a pedagogical description, see J. S. Bagla; J. S. Bagla and T. Padmanabhan, *Pramana* 49, 161 **(1997)**; K. Dolag et al.
15. Ya. B. Zeldovich, *Astron. Astrophys.* 5, 84 **(1970)**; S. N. Gurbatov, A. I. Saichev, and S. F. Shandarin, *MNRAS* 236, 385 **(1989)**; T. G. Brainerd, R. J. Scherrer, and J. V. Villumsen, *Astrophys. J.* 418, 570 **(1993)**; J. S. Bagla and T. Padmanabhan, *MNRAS* 266, 227 **(1994)**; *MNRAS*, 286, 1023 **(1997)**; T. Padmanabhan and S. Engineer, *Ap. J.* 493, 509 **(1998)**; S. Engineer, N. Kanekar, and N. Padmanabhan, *MNRAS* 314, 279 **(2000)**; T. Tatekawa, *Recent Res. Devel. Astrophys.* 2, 1 **(2005)**.
16. A. J. S. Hamilton et al., *Ap. J.* 374, L1 **(1991)**; R. Nityananda and T. Padmanabhan, *MNRAS* 271, 976 **(1994)**; T. Padmanabhan, *MNRAS* 278, L29 **(1996)**; T. Padmanabhan et al., *Ap. J.* 466, 604 **(1996)**; D. Munshi and T. Padmanabhan, *MNRAS*, 290, 193 **(1997)**; J. S. Bagla, S. Engineer, and T. Padmanabhan, *Ap. J.* 495, 25 **(1998)**; N. Kanekar, and T. Padmanabhan, *MNRAS*, 324, 988 **(2001)**.
17. T. Padmanabhan, *Phys. Rept.* 380, 235 **(2003)**; *Current Science*, 88, 1057 **(2005)**; S. Nobbenhuis; J. S. Alcaniz; S. Hannestad, *Int. J. Mod. Phys.* A21, 1938 **(2006)**; L. Perivolaropoulos, *AIP Conf. Proc.* 848, 698 **(2006)**; E. J. Copeland, M. Sami, and S. Tsujikawa, *Int. J. Mod. Phys.* D15, 1753 **(2006)**; A. D. Dolgov, *Phys. Atom. Nucl.* 71, 651 **(2008)**.
18. G. Mahajan, S. Sarkar, and T. Padmanabhan, *Phys. Letts. B* 641, 6 **(2006)**.
19. T. Padmanabhan, *Phys. Rev. D* 66, 021301 **(2002)**; T. Padmanabhan and T. R. Choudhury, *Phys. Rev. D* 66, 081301 **(2002)**; J. S. Bagla, H. K. Jassal, and T. Padmanabhan, *Phys. Rev. D* 67, 063504 **(2003)**.
20. G. F. R. Ellis and M. S. Madsen, *Class. Quan. Grav.* 8, 667 **(1991)**; Also see F. E. Schunck and E. W. Mielke, *Phys. Rev. D* 50, 4794 **(1994)**.
21. H. K. Jassal, J. S. Bagla, and T. Padmanabhan, *MNRAS* 356, L11-L16 **(2005)**.
22. G. F. R. Ellis and T. Buchert, *Phys. Lett. A* 347, 38 **(2005)**; R. M. Zalaletdinov, *Gen. Rel. Grav.* 24, 1015 **(1992)**; A. Paranjape and T. P. Singh, *JCAP* 0803, 023 **(2008)**; A. Ishibashi and R. M. Wald, *Class. Quant. Grav.* 23, 235 **(2006)**; S. Rasanen, *JCAP* 0402, 003 **(2004)**; B. Losic and W. G. Unruh, *Phys. Rev. D* 72, 123510 **(2005)**; C. M. Hirata, U. Seljak, *Phys. Rev. D* 72, 083501 **(2005)**; E. W. Kolb, S. Matarrese and A. Riotto, *New J. Phys.* 8, 322 **(2006)**; A. Paranjape, *Phys. Rev. D* 78, 063522 **(2008)**; For a review, see T. Buchert, *Gen. Rel. Grav.* 40, 467 **(2008)**.
23. T. Padmanabhan, *Gen. Rel. Grav.* 19, 927 **(1987)**.
24. S. Jha, A. G. Riess, and R. P. Kirshner, *Astrophys. J.* 659, 122 **(2007)**; A. Conley, arXiv:0705.0367, unpublished.
25. Such an idea has a long history: A. D. Sakharov, *Sov. Phys. Dokl.* 12, 1040 **(1968)**; T. Jacobson, *Phys. Rev. Lett.* 75, 1260 **(1995)**; T. Padmanabhan, *Mod. Phys. Lett. A* 17, 1147 **(2002)**; *Int. Jour. Mod. Phys. D* 13, 2293 **(2004)**; 18, 2903 **(2003)**; *Class. Quan. Grav.* 21, 4485 **(2004)**; G. E. Volovik, *Phys. Rept.* 351, 195 **(2001)**; M. Visser, *Mod. Phys. Lett. A* 17, 977 **(2002)**; C. Barcelo et al., *Int. J. Mod. Phys. D* 10, 799 **(2001)**; G. E. Volovik; G. E. Volovik, The Universe in a Helium Droplet, Oxford University Press **(2003)**; Chao-Guang Huang and Jia-Rui Sun, *Commun. Theor. Phys.* 49, 928 **(2008)**; J. Makela, gr-qc/0701128, unpublished.

26. T. Padmanabhan, *Gravity's Immunity from Vacuum: The Holographic Structure of Semiclassical Action*, [Third prize essay; Gravity research Foundation Essay Contest, 2006] *Gen. Rel. Grav.* 38, 1547 **(2006)**; *Int. J. Mod. Phys. D* 15, 2029 **(2006)**.

27. T. Padmanabhan, *Class. Quan. Grav.* 22, L107-L110, **(2005)**; For earlier attempts in similar spirit, see T. Padmanabhan, *Class. Quan. Grav.* 19, L167 **(2002)**; D. Sorkin, *Int. J. Theor. Phys.* 36, 2759 **(1997)**; for related work, see G. E. Volovik, *Annalen Phys.* 14, 165 **(2005)**; J. V. Lindesay et al.; Y. S. Myung, *Phys. Lett. B* 610, 18 **(2005)**; J. D. Barrow, *Phys. Rev. D* 75, 067301 **(2007)**; E. Elizalde et al., *Phys. Rev. D* 71, 103504 **(2005)**.

28. T. Padmanabhan, *AIP Conference Proceedings* 989, 114 **(2007)**.

29. T. Padmanabhan, *Gen. Rel. Grav.* 40, 529 **(2008)**; T. Padmanabhan and A. Paranjape, *Phys. Rev. D* 75, 064004, **(2007)**.

30. T. Padmanabhan, *Gravity-the inside story*, [First Prize essay; Gravity Research Foundation Essay Contest, 2008] to appear in *Gen. Rel. Grav.*

31. T. Padmanabhan, *Phys. Rept.* 406, 49 **(2005)**; *Mod. Phys. Lett.* A17, 923 **(2002)**; T. Padmanabhan and A. D. Patel, gr-qc/0309053, unpublished.

32. J. A. Wheeler, *A Journey into Gravity and Spacetime*, Scientific American Library, NY **(1990)**, p. 221.

33. G. W. Gibbons and S. W. Hawking, *Phys. Rev. D* 15, 2738 **(1977)**; T. Padmanabhan, *Mod. Phys. Letts. A* 19, 2637 **(2004)**.

34. H. S. Snyder *Phys. Rev.* 71, 38 **(1947)**; B. S. DeWitt, *Phys. Rev. Lett.* 13, 114 **(1964)**; T. Yoneya *Prog. Theor. Phys.* 56, 1310 **(1976)**; T. Padmanabhan *Ann. Phys. NY* 165, 38 **(1985)**; *Class. Quantum Grav.* 4, L107 **(1987)**; A. Ashtekar, C. Rovelli, and L. Smolin, *Phys. Rev. Lett.* 69, 237 **(1992)**; T. Padmanabhan, *Phys. Rev. Lett.* 78, 1854 **(1997)**; *Phys. Rev. D* 57, 6206 **(1998)**; K. Srinivasan, L. Sriramkumar, and T. Padmanabhan, *Phys. Rev. D* 58, 044009 **(1998)**; X. Calmet, M. Graesser, and S. D. H. Hsu, *Phys. Rev. Lett.* 93, 211101 **(2004)**; M. Fontanini, E. Spallucci, and T. Padmanabhan, *Phys. Lett. B* 633, 627 **(2006)**; For a review, see L. J. Garay, *Int. J. Mod. Phys. A* 10, 145 **(1995)**.

35. See e.g., T. Padmanabhan, *Phys. Rev. Letts.* 81, 4297 **(1998)**; *Phys. Rev. D.* 59, 124012 **(1999)**, and references therein.

36. T. Padmanabhan, *Class. Quan. Grav.* 19, 5387 **(2002)**.

37. J. Zhou et al., *Phys. Lett. B* 652, 86 **(2007)**.; R.-G. Cai and L.-M. Cao, *Phys. Rev. D* 75, 064008 **(2007)**; M. Akbar, R.-G. Cai, *Phys. Rev. D* 75, 084003 **(2007)**; M. Akbar, *Chin. Phys. Lett.* 24, 1158 **(2007)**; D. Kothawala, S. Sarkar, and T. Padmanabhan, *Phys. Letts. B* 652, 338 **(2007)**; M. Akbar and R.-G. Cai, *Phys. Lett. B* 648, 243 **(2007)**; X.-H. Ge, *Phys. Lett. B* 651, 49 **(2007)**; A. Sheykhi, B. Wang, and Rong-Gen Cai, *Nucl. Phys. B* 779, 1 **(2007)**; G. Allemandi, M. Francaviglia, and M. Raiteri, *Class. Quant. Grav.* 20, 5103 **(2003)**.

38. T. Padmanabhan, *Int. J. Mod. Phys. D* 15, 1659 **(2006)**.

39. T. Padmanabhan, *Gen. Rel. Grav.* 34, 2029 **(2002)**; *Brazilian J. Phys. (Special Issue)* 35, 362 **(2005)**; *Class. Quan. Grav.* 21, L1 **(2004)**; *Gen. Rel. Grav.* 35, 2097 **(2003)**;

40. T. Padmanabhan, *AIP Conference Proceedings* 861, 858, **(2006)**; *Int. J. Mod. Phys. D* 14, 2263 **(2005)**.

41. A. Paranjape, S. Sarkar, and T. Padmanabhan, *Phys. Rev. D* 74, 104015 **(2006)**.

42. A. Mukhopadhyay and T. Padmanabhan, *Phys. Rev. D* 74, 124023 **(2006)**.

43. C. Lanczos, *Z. Phys.* 73, 147 **(1932)**; *Annals Math.* 39, 842 **(1938)**; D. Lovelock, *J. Math. Phys.* 12, 498 **(1971)**.

44. S. A. Fulling, *Phys. Rev. D*7, 2850 **(1973)**; W. G. Unruh, *Phys. Rev. D* 14, 870 **(1976)**; B. S. DeWitt, "General Relativity: An Einstein Centenary Survey" (S. W. Hawking and W. Israel, Ed.), Cambridge University Press **(1979)**, pp. 680–745; T. Padmanabhan, *Class. Quan. Grav.* 2, 117 **(1985)**; K. Srinivasan, and T. Padmanabhan, *Phys. Rev. D* 60, 24007 **(1999)**; L. Sriramkumar and T. Padmanabhan, *Int. J. Mod. Phys. D* 11, 1 **(2002)**.

45. T. Padmanabhan and T. P. Singh, *Class. Quan. Grav.* 4, 1397 **(1987)**.

Received: 15 July 2008. Accepted: 15 September 2008.

AMERICAN
SCIENTIFIC
PUBLISHERS