ELSEVIER
SCIENCE DIRECT
PHYSICS LETTERS B
Physics Letters B 594 (2004) 8–12
www.elsevier.com/locate/physletb

Dark energy and the nature of the graviton

A. Zee

Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA

Received 23 April 2004; accepted 30 April 2004

Available online 9 June 2004

Editor: H. Georgi

Abstract

Does the existence of dark energy suggest that there is more to the graviton than we think we know?

The word paradox has been emasculated by indiscriminate usage in the physics literature. A real paradox should involve a major and clear-cut discrepancy between theoretical expectation and experimental measurement. The ultraviolet catastrophe, for example, is a paradox, the resolution of which around the dawn of the 20th century ushered in quantum physics. Surely, the most egregious paradox of physics around the dawn of the 21st century is the cosmological constant paradox [1].

The root of the paradox lies in a fundamental clash between Einstein's view and the particle theorist's view of gravity. To particle physicists, the graviton is just another particle, or a particular mode of the vibrating string. Indeed, given that a massless spin-2 particle couples to the stress-energy tensor, one can reconstruct Einstein's theory. According to Einstein, however, gravity has to do with the curvature of spacetime, the arena in which all fields and particles live in. The graviton is not just another particle.

The graviton is not just another particle—it knows too much. The electromagnetic force knows about the particles carrying charge, and the strong force knows about the particles carrying color. But the gravitational force knows about anything carrying energy and momentum, including an apparently innocuous constant shift in the Lagrangian density $\mathcal{L} \to \mathcal{L} - \Lambda$.

As is well known, the paradox can be easily described. The natural value of Λ in particle physics is expected by dimensional analysis to be $\mu^4 = \mu/(\mu^{-1})^3$ for some relevant mass scale μ where the second form of writing μ^4 reminds us that Λ is a mass or energy density. Whether one associates μ with grand unification, electroweak symmetry breaking, or the quark confinement transition and consequently has a value of order 10^{19} GeV, 10^2 GeV, and 1 GeV, respectively, is immaterial. The natural value $\Lambda \sim \mu^4 = \mu/(\mu^{-1})^3$ is outrageous even if we take the smallest value for μ. We do not even have to put in actual numbers to see that there is a humongous discrepancy between theoretical expectation and observational reality. We know the universe is not permeated with a mass density of order of 1 GeV on every cube of size 1 $(\text{GeV})^{-1}$.

E-mail address: zee@kitp.ucsb.edu (A. Zee).

0370-2693 © 2004 Published by Elsevier B.V. Open access under CC BY license.
doi:10.1016/j.physletb.2004.04.087

The cosmological constant paradox is basically an enormous mismatch between the units natural to particle physics and natural to cosmology. Measured in units of GeV^4 the cosmological constant is so incredibly tiny that particle physicists have traditionally assumed that it must be mathematically zero, and have looked in vain for a plausible mechanism to drive it to zero. One of the disappointments of string theory is its inability to resolve the cosmological constant paradox.

But Nature has a big surprise for us. While theorists racked their brains trying to come up with a convincing argument that $\Lambda = 0$, observational cosmologists [2,3] steadily refined their measurements and changed their upper bound to an approximate equality[1]

$$\Lambda \sim \left(10^{-3}\ \mathrm{eV}\right)^4\ (!!!). \tag{1}$$

The cosmological constant paradox deepens. Theoretically, it is easier to explain why some quantity is mathematically 0 than why it happens to be $\sim 10^{-124}$ in the units natural (if indeed they are) to the problem.

As is also well known, strictly speaking, we should refer to the observation of the cosmological constant as the observation of a hitherto unknown dark energy since we do not know the equation of state associated with the observed energy density $(10^{-3}\ \mathrm{eV})^4$.

To make things worse, the energy density $(10^{-3}\ \mathrm{eV})^4$ happens to be the same order of magnitude as the present matter density of the universe ρ_M. This is sometimes refer as the cosmic coincidence problem. The cosmological constant Λ is, within our traditional understanding, a parameter in the Lagrangian. On the other hand, since most of the mass density of the universe resides in the rest mass of baryons, as the universe expands $\rho_M(t)$ decreases like $(1/R(t))^3$, where $R(t)$ denotes the scale size of the universe. In the far past, ρ_M was much larger than Λ, and in the far future, it will be much smaller. It just so happens that in this particular epoch of the universe, when we are around, that $\rho_M \sim \Lambda$. Or to be less anthropocentric, the epoch when $\rho_M \sim \Lambda$ happens to be when galaxy formation has been largely completed. In their desperation, some theorists have even been driven to invoke anthropic selection [5–7].

My impression is that theoretical physicists outside the high energy community are not completely aware of how desperate the situation is, but people are grasping at straws and a number of outlandish suggestions have been aired. In this spirit I would like to offer a thought I have entertained for some time but did not "dare" to publish. When I wrote my recent field theory textbook I sketched [8] what I had in mind in passing. It may be worthwhile to elaborate on what I wrote there and to bring it to the attention of a broader audience.

In the development of physics, there have been numerous instances of reasoning by historical precedent or analogy. For example, when confronted with data showing that the energy of the electron in an atom is quantized physicists recalled that the vibrational frequency of a violin string is also quantized. As we all know, this turns out to be an apt analogy as both the energy of the electron and the vibrational frequency of a string are given by the eigenvalues of linear partial differential equations. I suggest that perhaps similarly we can ask if historically there have been cases of a physical quantity initially thought to be 0 but then turned out to be extremely small but not precisely 0. I suspect that the proton decay rate may be an apt example and that it may shed some light on the cosmological constant paradox.

Let us go through the story of proton decay. To make my point I will take some liberty with history. Suppose that in 1953 some theorists were to calculate the rate Γ for protons to decay in the natural mode $p \to e^+ + \pi^0$. The interaction of the pion with the proton and the neutron was known to be described by a term like $g\pi\bar{n}p$ in the Lagrangian with g a dimensionless coupling of order 1. These theorists would naturally construct a Lagrangian out of the available fields, namely the proton field p, the electron field e, and the pion field π, and thus write down something like $f\pi\bar{e}p$ with some constant f. Note that $\pi\bar{e}p$ has mass dimension 4 and hence f is dimensionless just like g. Since $\pi\bar{e}p$ violates isospin invariance, the theorists would expect f to be suppressed relative to g by some measure of isospin breaking, say the fine structure constant α. The natural

[1] The value 10^{-3} eV does not correspond to any known mass scale in particle physics. The differences in mass squared of the 3 neutrinos have been measured experimentally to be $|m_3^2 - m_2^2| \sim 3 \times 10^{-3}\ \mathrm{eV}^2$ and $|m_2^2 - m_1^2| \sim 7 \times 10^{-5}\ \mathrm{eV}^2$ and so it is conceivable that the lightest neutrino has mass $\sim 10^{-3}$ eV and that this may have something to do with the cosmological constant. In a recent paper, using a particular ansatz [4] determined neutrino masses to lie in the range 10^{-2} to 10^{-4} eV.

value for Γ would then come out to be many orders of magnitude larger than the experimental upper bound on Γ. The theorists would then set $\Gamma = 0$ and cast about for an explanation. After an enormous struggle, the theorists were unable to come up with a compelling explanation and this failure became known as the proton decay rate paradox.

Eventually, someone with great authority and prestige in the community, namely Wigner, decreed the law of baryon number conservation. Surely, even in the unthinkably primitive days of 1953 this would have been recognized as a pronouncement and not as an explanation. (The pronouncement could be dressed up formally by imposing a $U(1)$ transformation under which $p \to e^{i\theta} p$ while e and π do not change and requiring that the Lagrangian remains invariant.) But there would have been no deep understanding of this astonishing discrepancy between theoretical expectation and experimental upper bound.

Indeed, imagine an alternative history in which, while other important particle physics experiments were being performed in 1957, some intrepid experimentalist, ignoring conventional theoretical wisdom, actually went out and measured the proton decay rate to be some tiny but non-zero value. The proton decay rate paradox would have deepened, much as how the cosmological constant paradox deepened with the discovery of a tiny Λ.

Let us now review how the proton decay rate paradox was resolved historically. The first remark is that the eventual explanation did not emerge within the orthodox theory fashionable in 1957, nor did it come from an understanding of some kind of mechanism causing protons to decay, but rather it came totally from "left field", from a study of baryon spectroscopy, which led to the notion of quarks. The correct degrees of freedom are not given by the proton and pion fields p and π, but by the quark fields q. The effective Lagrangian $\mathcal{L}$ is to be constructed out of quark q and lepton l fields and must satisfy the symmetries that we know. Three quarks disappear, so we write down schematically qqq, but three spinors do not a Lorentz scalar make. We have to include a lepton field and write $qqql$. Since four fermion fields are involved, these terms have mass dimension 6 and so in $\mathcal{L}$ they have to appear as $(1/M^2)qqql$ with some mass M, corresponding to the mass scale of the physics responsible for proton decay. Thus, the probability of proton decay is proportional to $(1/M^2)^2 = 1/M^4$. By dimensional reasoning, we obtain the proton decay rate $\Gamma \sim (m_p/M)^4 m_p$. The absurdly small value of Γ is then naturally explained by the fourth power of the small number m_p/M for M big enough. No mystery left!

Note that in principle all of this could be done as soon as Gell-Mann introduced the notion of quarks in 1964, long before anybody even dreamed of a grand unified theory with proton decay.

As long we are discussing revisionist, but possible, history, we can imagine some brilliant theorist in another civilization far away puzzling over the proton decay rate paradox eventually realizing that the key to explaining an absurdly small number is to promote the dimension of the effective Lagrangian merely from 4 to 6. In hindsight, we can say that the extremely long lifetime of the proton could have pointed to the existence of quarks.

I would like to raise the question whether the cosmological constant paradox might not be solved in the same way. Perhaps the gravitational field $g_{\mu\nu}$ is the analog of the proton and pion field p and π. The high energy and more fundamental degrees of freedom in the gravitational field may not be the metric $g_{\mu\nu}$, but some mysterious analog of the quark field q. This may emerge as a construct in string or M-theory, or it could be something else completely. In the history of the proton decay paradox as recounted by me, there is an additional twist, namely that the degree of freedom q is confined and not physical. Before the advent of quantum chromodynamics, theorists could only write $p \sim qqq$, without any clear idea about what the symbol $\sim$ might mean. We are in a similar position here: the metric $g_{\mu\nu}$ might be a composite object, but I certainly do not know what it is a composite of, and the objects of which $g_{\mu\nu}$ is a composite may also be as observable or as unobservable as the quarks.

The cosmological term $\Lambda\sqrt{g}$ in the Lagrangian has mass dimension 0 and we somehow have to promote 0 to a higher number. One difficulty with this view is of course how we could possibly promote the dimension of the cosmological term without at the same time changing the mass dimension of the Einstein–Hilbert term $\frac{1}{G}\sqrt{g}R$. Our historical analogy may again be helpful: the 1953 view that the pion nucleon coupling term has dimension 4 turns out to be correct. While the dimension 4 term $\pi\bar{e}p$ was replaced by the dimension 6 term $qqql$ the dimension 4 term $\pi\bar{n}p$ was replaced by dimension 4 terms of the form $\bar{q}Aq$ with A a gluon potential. The dimension of one of the terms gets promoted while the dimension of the other term remains the same. So it is entirely conceivable to me that the cosmological constant term could end up with a higher dimension while the Einstein–Hilbert term either remains dimension 4 or is replaced by dimension 4 terms. Thus, suppose the cosmological constant term actually has dimension $p > 0$ so that it is given in the Lagrangian by a term of the form $\frac{1}{M^{p-4}}\mathcal{O}$ with M some mass scale characteristic of the deeper structure of the graviton, perhaps the same as the Planck mass, perhaps not. The observed cosmological constant would then be given by

$$\Lambda \sim \frac{1}{M^{p-4}}\langle\mathcal{O}\rangle = \left(\frac{m}{M}\right)^p M^4,$$

where the expectation value of the operator $\mathcal{O}$ in the physical universe $\langle\mathcal{O}\rangle = m^p$ is set by physics at some low energy scale m. With m small enough, and or p big enough, we could easily get the suppression factor we want.

As hinted above, I even suspect that the Lagrangian formalism, being a mathematical realization of the quasi-theological (at least historically) variational principle, may well be wrong. The cause of all our trouble is that the flat space Lagrangian $\mathcal{L}$ could always be shifted by a constant $\mathcal{L} \to \mathcal{L} - \Lambda$ without changing its variation. But this also points to the mystery of quantum mechanics[2] because without quantum mechanics we could have lived happily with equations of motion without ever bothering with the Lagrangian.

Another possible way to nullify the physical consequence of the shift $\mathcal{L} \to \mathcal{L} - \Lambda$ is to postulate that $g = \det g_{\mu\nu}$ is not a dynamical variable. This was proposed [9] 20 years ago as an explanation of why Λ is mathematically zero. But with Λ now known to be tiny but non-zero this avenue seems to me less promising.

I also could not resist mentioning another wild speculation [10]. Many years ago, inspired by the almost exact correspondence between Einstein's post-Newtonian equations of gravity and Maxwell's equations of motion I proposed the gravitipole in analogy with Dirac's magnetic monopole. After Dirac there was considerable debate on how a field theory of magnetic monopoles may be formulated. Eventually, 't Hooft and Polyakov showed that the magnetic monopole exists as an extended solution in certain non-abelian gauge theories. Most theorists now believe that electromagnetism is merely a piece of a grand unified theory and that magnetic monopoles exist. Might it not turn out that Einstein's theory is but a piece of a bigger theory and that gravitipoles exist? In grand unified theory the electromagnetic field is a component of a multiplet. Could it be that the gravitational field also somehow carries an internal index and that the field we observe is just a component of a multiplet? Throwing caution to the wind, I also asked in [10] if the gravitipole and the graviton might not form a representation under some dual group just as the magnetic monopole and the photon form a triplet under the dual group of Montonen and Olive [11].

Perhaps we do not know as much about the graviton as we think we do.

Acknowledgement

This work was supported in part by the National Science Foundation under grant number PHY 99-07949.

References

[1] S. Weinberg, Rev. Mod. Phys. 61 (1989) 1.
[2] A.G. Riess, et al., Astron. J. 116 (1998) 1009.
[3] S. Perlmutter, et al., Astrophys. J. 517 (1999) 565.
[4] X.G. He, A. Zee, hep-ph/0302201.
[5] T. Banks, Nucl. Phys. B 249 (1985) 332.
[6] S. Weinberg, Phys. Rev. Lett. 59 (1987) 2607;
S. Weinberg, Phys. Rev. D 61 (2000) 103505.
[7] A. Vilenkin, Phys. Rev. Lett. 74 (1995) 846.
[8] See p. 440 in: A. Zee, Quantum Field Theory in a Nutshell, Princeton Univ. Press, Princeton, NJ, 2003.
[9] F. Wilczek, A. Zee, unpublished;
A. Zee, Remarks on the Cosmological Constant Paradox, in: S.L. Mintz, A. Perlmutter (Eds.), High Energy Physics in

[2] When I discussed the speculative idea proposed in [9] with R. Feynman, he asked me if I knew of a formalism in quantum mechanics in which one could calculate the difference between two energy levels in say an atom directly (which is after all what experimentalists measure) without having to calculate the two energy levels separately and then subtract one from the other. Without gravity, only differences in energy matter.

Honor of P.A.M. Dirac in his Eightieth Year, Proceedings of the 20th Orbis Scientiae, Plenum, New York, 1983, p. 211. This possibility was analyzed further in Section VI of [1].

[10] A. Zee, Phys. Rev. Lett. 55 (1985) 2379.

[11] C. Montonen, D. Olive, Phys. Lett. B 72 (1977) 117.