Vacuum catastrophe: An elementary exposition of the cosmological constant problem

Ronald J. Adler, Brendan Casey, and Ovid C. Jacob
Department of Physics, San Francisco State University, San Francisco, California 94132[a)]

(Received 7 July 1994; accepted 16 November 1994)

Quantum field theory predicts a very large energy density for the vacuum, and this density should have large gravitational effects. However these effects are not observed, and the discrepancy between theory and observation is an incredible 120 orders of magnitude. There is no generally accepted explanation for this discrepancy, although numerous papers have been written about it. As usually stated the problem requires a knowledge of quantum field theory and general relativity, topics not normally studied by undergraduates. We have tried to make the problem accessible to undergraduates by using only the simplest ideas of quantum theory, such as the uncertainty principle and the theory of the harmonic oscillator, and classical gravitational theory. We believe that such simplification is not only an amusing pedagogical exercise but clarifies how basic is the conflict between quantum theory and gravitational theory. We do not here discuss various proposed solutions to the problem, beyond the trivial and unsatisfactory one of assuming an ad hoc canceling term in the Hamiltonian or field equations. © *1995 American Association of Physics Teachers.*

I. INTRODUCTION

This work began when we were asked to give a talk to the physics club at San Francisco State University on the cosmological constant problem, a famous and spectacular problem that some of the students had heard comments about. Our first reaction was that it was not an appropriate topic for undergraduates, but after some consideration we did make an attempt. It was at least partially successful, and we present here an expanded version of the attempt. It is mainly addressed to teachers of undergraduate physics students, but we of course hope that it is directly accessible to the students themselves.

Quantum field theory is notorious for its divergences. The most fundamental one concerns the energy density of the vacuum. Theory clearly predicts that every mode or degree of freedom of the electromagnetic field has an associated ground state or zero point energy of $\hbar\omega/2$, where ω is the frequency or energy of the mode.[1,2] When summed over all modes up to some reasonable maximum this leads to an enormous energy density for the vacuum.[3] Such an energy density presents serious problems when gravity is considered, since its gravitational field has not been observed; indeed the nonobservation places an upper limit on the energy density of the vacuum, which is in extreme conflict with the theoretical estimate: about 120 orders of magnitude! One must conclude that there is a deep-seated inconsistency between the basic tenets of quantum field theory and gravity.

This problem is normally discussed from a fairly advanced level, that of canonical quantum field theory and the general relativity theory of gravity. Indeed the name usually used, the cosmological constant problem, derives from a term in the field equations of general relativity theory that acts like an energy density of the vacuum.[4] Our purpose is to present the problem without reference to the full formalism of quantum field theory or general relativity theory, but using instead elementary quantum mechanical ideas and classical gravitational theory. This is why we prefer to label the problem, provocatively, as the vacuum catastrophe, in analogy with the ultraviolet catastrophe associated with the blackbody radiation problem that was solved by Planck in 1900 to give rise to quantum theory.[5,6]

II. A WORD ON NATURAL UNITS

In virtually all work in particle physics natural units are used.[7,8] They relieve us from writing seemingly endless combinations of $\hbar$ and c in our equations, which are thereby made clearer to the eye, and presumably to the brain also.

Let us begin with c, which in mks units is 3.00×10^8 m/s. For doing particle physics we adopt a more reasonable unit of length, the Fermi, $1\ \text{f}=10^{-15}$ m. As a unit of time we temporarily adopt what we will call a "jiffy", $1\ \text{j}=(1/3.00)\times10^{-23}$ s, so that the velocity of light is 1 f/j. Indeed it is clear that, consistent with the basic ideas of relativity, there is really no fundamental difference between the Fermi unit of length and the jiffy unit of time, and we may in fact think of them as completely equivalent so that c is also dimensionless.

In the same manner consider Planck's constant $\hbar$, which is 6.58×10^{-16} eV s or 197 MeV j. As a convenient unit of energy we temporarily adopt a "blip," 1 b=197 MeV, so that the value of $\hbar$ is 1 bj. Then as above we can consider 1 blip to be equivalent to 1 inverse jiffy, which in turn is equivalent to 1 inverse Fermi, so that $\hbar$ is dimensionless.

An important fact should be stressed; it is not necessary to restore the factors of $\hbar$ and c in the answer after a calculation is completed. For example suppose we do a calculation using $c=1$ and obtain an answer of 2.06; then if the answer is supposed to be a length it is 2.06 f; if it is supposed to be a time it is $2.06\ \text{j}=0.69\times10^{-23}$ s.

Our unit system now has $\hbar=c=1$, both numerically and in terms of dimensions. The unit of distance is a Fermi, the unit of time is a Fermi, and the units of energy, mass, and frequency are all an inverse Fermi. Energy density has dimensions of energy/length^3 or units of inverse Fermi,^{4} which also is 197 MeV/f.^{3} Note that we need not refer to our fanciful time and energy units, the jiffy and blip, again.

Other similar natural unit systems are in use. One popular choice is to use 0.197 f as a distance unit, so that the energy unit is 1 GeV.

III. BASIC FACTS ABOUT THE HARMONIC OSCILLATOR IN QUANTUM MECHANICS

Most undergraduates have studied the quantum mechanical harmonic oscillator and are familiar with its energy levels.[9] We here review the spectrum with emphasis on the ground state; we stress that the nonzero ground state energy is a consequence of the uncertainty principle, and in fact may be estimated from it.

The energy (in terms of x and $\dot{x}$) or the Hamiltonian (in terms of x and p) for a one-dimensional harmonic oscillator is given by

$$H=\frac{m\dot{x}^2}{2}+\frac{m\omega^2x^2}{2}=\frac{p^2}{2m}+\frac{m\omega^2x^2}{2}. \tag{1}$$

Here ω is the classical frequency, related to the spring constant k by $\omega^2=k/m$. In an elementary course the problem is usually solved using the Schrödinger equation, or occasionally the creation and annihilation operator formalism, to give a spectrum

$$E=(n+1/2)\omega. \tag{2}$$

The nonzero ground state energy $E_g=\omega/2$ can be explained by saying that the particle in its parabolic potential well cannot be at the bottom of the well, $x=0$, and also have zero momentum, $p=0$, since this is not consistent with the uncertainty principle; thus it must have some positive ground state energy, that is a zero point energy.

A simple argument makes this explicit and semiquantitative. We suppose the particle is constrained away from being at rest at the bottom of the potential well at $x=0$ only by the uncertainty principle in the form $xp=1$ or $p=1/x$. Substituting this in the energy expression (1), with x, p, and H interpreted as c numbers, we have

$$H=\frac{1}{2mx^2}+\frac{m\omega^2x^2}{2}. \tag{3}$$

We choose x to minimize this expression and find

$$x_{\min}=1/\sqrt{m\omega}, \quad E_{\min}=\omega. \tag{4}$$

This argument works for diverse other potentials such as the Coulomb potential and the linear potential, and makes it clear that as a consequence of the uncertainty principle quantum mechanical systems in general do not have zero energy in the ground state.

For a harmonic oscillator with two degrees of freedom, for example a particle in a two-dimensional well or 2 particles in a one-dimensional well, there is a complexified form for the energy which is elegant and which will be useful later. If the coordinates are x and y we introduce the complex coordinate $z=x+iy$ and find that the energy may be expressed in terms of x and y, or in terms of z, as

$$H=\frac{m(\dot{x}^2+\dot{y}^2)}{2}+\frac{m\omega^2(x^2+y^2)}{2}=\frac{m\dot{z}\dot{z}^*}{2}+\frac{m\omega^2zz^*}{2},$$

$$=\frac{p_x^2+p_y^2}{2m}+\frac{m\omega^2(x^2+y^2)}{2}=\frac{2p_zp_{z*}}{m}+\frac{m\omega^2zz^*}{2}. \tag{5}$$

The spectrum of this is of course the sum of 2 one-dimensional oscillator spectra, and the ground state energy is twice the ground state energy for each degree of freedom, or $\omega=2(\omega/2)$.

IV. THE ENERGY DENSITY OF ELECTROMAGNETIC RADIATION

A key element in our discussion is to show an equivalence between electromagnetic radiation and the harmonic oscillator, so that we can infer the energy spectrum of the radiation field, in particular its ground state energy. First courses in electricity and magnetism give the energy density of the electromagnetic field,

$$\rho=(\mathbf{E}^2+\mathbf{B}^2)/2 \quad \text{(unrationalized units in which } c=1\text{)}, \tag{6}$$

and the **E** and **B** fields for a monochromatic plane wave,[10]

$$\mathbf{E}=\mathbf{E}_0\sin(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta), \quad \mathbf{B}=\mathbf{B}_0\sin(\mathbf{k}\cdot\mathbf{x}-\omega t+\delta). \tag{7}$$

These are clearly solutions of the wave equation. Here **k** is the wave vector, a vector in the direction of propagation with magnitude equal to the frequency ω, δ is an arbitrary phase, and the magnitudes of the **E** and **B** fields are equal. The orientations are shown in Figs. 1 for the 2 polarization states; in one the electric field is vertical, and in the other it is rotated through 90° to be horizontal.

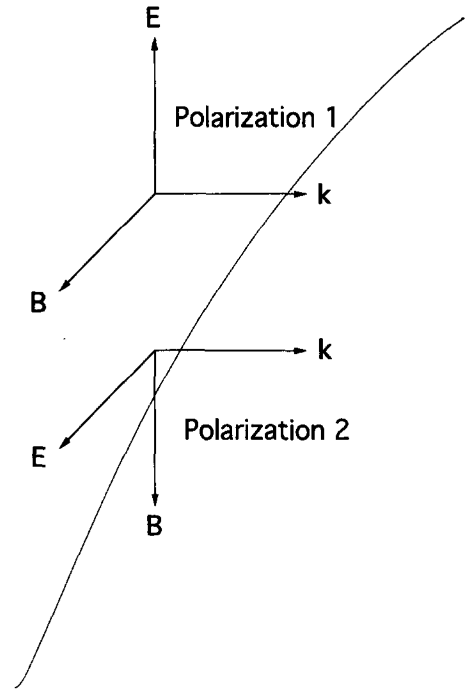

Fig. 1. Plane electromagnetic wave configurations.

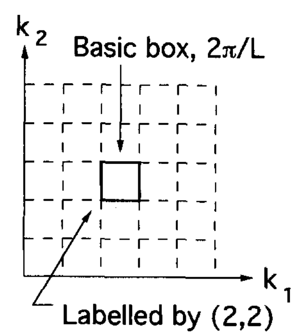

Fig. 2. Discretization of momentum space, illustrated in two dimensions, with basic box being $2\pi/L$ on a side.

To make an analogy between the radiation field and the harmonic oscillator we use the radiation gauge, in which the scalar potential is zero, and the **E** and **B** fields are given in terms of the vector potential **A** by

$$\mathbf{E} = -\dot{\mathbf{A}}, \quad \mathbf{B} = \nabla \mathbf{X} \mathbf{A}. \tag{8}$$

An elegant artifice is to introduce a complex **A** field and thus complex **E** and **B** fields. We write **A** as

$$\mathbf{A} = \mathbf{A}_0 e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}}, \tag{9}$$

so that the real and imaginary parts are solutions of the wave equation and give fields as in Eq. (7). It represents a *pair* of classical waves, and has a form analogous to the free particle de Broglie wave of elementary quantum mechanics: the energy or frequency is ω, and the momentum or wave vector is **k**. We also introduce a complexified form for the energy density (6) appropriate to such a complex field,

$$\begin{aligned} \rho &= (\mathbf{E}\cdot\mathbf{E}^* + \mathbf{B}\cdot\mathbf{B}^*)/2 \\ &= (\mathrm{Re}\ \mathbf{E}^2 + \mathrm{Re}\ \mathbf{B}^2)/2 + (\mathrm{Im}\ \mathbf{E}^2 + \mathrm{Im}\ \mathbf{B}^2)/2. \end{aligned} \tag{10}$$

This is the obvious generalization of the expression (6) and gives the sum of the energies of both the real and imaginary parts of the complex field. Using Eq. (9) for the radiation field and Eq. (10) for the energy density we may express the energy density in suggestive form as

$$\rho = (\dot{\mathbf{A}}\cdot\dot{\mathbf{A}}^* + \mathbf{k}^2\mathbf{A}\cdot\mathbf{A}^*)/2. \tag{11}$$

This expression for the energy density of two classical waves has the same form as the energy for the harmonic oscillator with 2 degrees of freedom, as given in terms of complex z in the second expression in Eq. (5). The correspondence is that m becomes 1, the complex particle displacement becomes the **A** field, and time remains time. This correspondence is the basis of the formalism of quantum field theory, in which the field at each point in space corresponds to a displacement, so that the field becomes a system with an infinite number of degrees of freedom. However without going further we see immediately that the spectrum for each mode **k** of the radiation field is that of a harmonic oscillator (2). The integer n is interpreted as the number of photons with wave vector **k** and frequency ω. In particular each mode will have a ground state or zero point energy of $\omega/2$ even when no photons are present.

In summary, we have cast the energy density of the electromagnetic field into the same form as that of a harmonic oscillator to show that for each possible momentum **k** and corresponding frequency ω the radiation field has a ground state energy which is not equal to zero but to $\omega/2$. This is true for each polarization state. The nonzero energy is traceable, as with the harmonic oscillator, to the uncertainty principle.

V. COUNTING MODES AND THE ENERGY DENSITY OF THE ELECTROMAGNETIC VACUUM

The wave vector or momentum **k** labels a mode of the radiation field. To get the energy of the radiation field we need to sum the energies of such modes. The problem of counting modes labeled by a continuous index is an old one and a number of elegant artifices have been developed. We will use the method of periodic boundary conditions.[11]

We suppose that the universe is a large cube of size L, and demand that a wave of the form $\exp(i\mathbf{k}\cdot\mathbf{x})$, such as the complex vector potential in Eq. (9), has the same value along the edges of the cube. Thus for the x-dependent factor the values at $x=0$ and $x=L$ must be the same, or

$$e^{ik_1L} = e^{ik_10} = 1, \quad k_1L = 2\pi n_1, \quad k_1 = (2\pi/L)n_1, \tag{12}$$

where n_1 is any integer, positive or negative. The wave vector component k_1 thereby has discrete values instead of continuous ones, but as L becomes large they become very dense. The x and y dependent factors are similar, so the wave is labeled by a set of three integers and the wave vector is given by

$$\mathbf{k} = (2\pi/L)(n_1, n_2, n_3). \tag{13}$$

We may think of momentum space as a cubic lattice with lattice points labeled by three integers, and the volume of each little cube is $(2\pi/L)^3$. (See Fig. 2.) A volume of momentum space V_k thus contains a number of modes

$$N = V_k/(2\pi/L)^3 = V_kL^3/(2\pi)^3. \tag{14}$$

The volume density of such modes is this number divided by the total volume of space L^3, which gives the well-known result,

$$n = V_k/(2\pi)^3. \tag{15}$$

We apply this result to calculate the total energy of the radiation field ground state. The energy of each mode is $\omega/2$, and only depends on the magnitude of **k**. The number of modes dn between ω and $\omega + d\omega$ is the volume of momentum space in a thin spherical shell, divided by $(2\pi)^3$, or

$$dn = 4\pi\omega^2 d\omega/(2\pi)^3. \tag{16}$$

Since the energy of each mode is $\omega/2$ the total energy is the integral of ωdn over all allowed modes, or[3]

$$\rho = \int_0^{\omega_{\max}} \frac{4\pi\omega^3 d\omega}{(2\pi)^3} = \frac{\omega_{\max}^4}{8\pi^2}. \tag{17}$$

There is a factor of 2 in this for the two polarization states of each mode. This is our main result, the energy density of the ground state of the electromagnetic radiation field, that is the electromagnetic vacuum.

Of course for $\omega_{\max} = \infty$ this result is formally infinite. However if we do not allow arbitrarily large frequencies of electromagnetic waves then we may specify a cutoff, and the energy density becomes finite. Photons of several hundred GeV are routinely produced in accelerator laboratories, so we may certainly suppose that $\omega_{\max}$ exceeds about 1 $\mathrm{TeV} = 10^3$ GeV. If we were to adopt this as a maximum then in natural units $\omega_{\max} = 5.08\times 10^3\ \mathrm{f}^{-1}$, and the vacuum energy density would be

$$\rho > 8.4\times 10^{12}\ \mathrm{f}^{-4} \quad (10^3\ \mathrm{GeV\ cutoff}). \tag{18}$$

Note that nuclear density is roughly 100 in these units, so this density is vastly greater.

The above is most certainly far too low a cutoff. For example primary cosmic ray photons are measured (indirectly) to have energies of up to about 10^{14} GeV. Many theorists speculate that a reasonable maximum energy for photons is the Planck energy of about 10^{19} GeV, the energy at which the gravitational interaction is roughly as strong as the electromagnetic and strong interactions.[12] With this taken as the maximum photon energy we obtain a vacuum energy density of

$$\rho \approx 8.4 \times 10^{76}\ \mathrm{f}^{-4} \quad (10^{19}\ \text{GeV cutoff, Planck energy}). \tag{19}$$

This is a truly large number. Either of the estimates (18) or (19) is in fact far too large to be consistent with observation as we will show in Sec. VIII; this constitutes the vacuum catastrophe.

VI. OTHER FIELDS

Our primary goal in this paper is to study the electromagnetic vacuum and its energy density. A full discussion of the other fields in nature is beyond our present scope and moreover does not change the nature of the problem. But we must briefly mention the other fields since they should contribute to the vacuum catastrophe. In the standard model of elementary particles and fields the strong force between quarks is produced by the gluon field, which is quite analogous to the electromagnetic field in that its quanta are massless and have two polarization states, but whereas there is only one photon there are eight types of gluons.[13] As such they should contribute to the vacuum energy density like the electromagnetic field, $\omega/2$ for each mode and polarization and type of gluon. This gives a density like Eq. (17) but larger by a factor of 8.

Similarly the weak force involves the vector bosons W^+, W^-, and Z^0, which are analogous to the photon but are quite massive, about 82 and 92 GeV, respectively, and have three polarization states instead of two. These should also contribute to the vacuum energy density, but due to their nonzero masses the result should be only approximately given by an equation like Eq. (17) if the cutoff energy is much larger than the masses. Of course since we are concerned only with rough magnitudes this is adequate, and we therefore expect a contribution like Eq. (17) but larger by a factor of 9/2. All of the fundamental force fields, the fundamental bosons, together give a result

$$\begin{aligned}\rho &= \frac{\omega_{\max}^4}{16\pi^2}(2+16+9)\\ &= \frac{27\omega_{\max}^4}{16\pi^2} \quad \text{(photons, gluons, } W\text{, and } Z \text{ bosons)},\end{aligned} \tag{20}$$

where the three terms in the parentheses are the contributions from the electromagnetic, gluon, and weak force fields. In light of the size of the result and the uncertainty in the cutoff energy this is not significantly different than the electromagnetic result (17).

An interesting effect occurs if one studies the fermion fields describing the fundamental quarks and leptons. These contribute a negative amount, $-\omega/2$, for each degree of freedom to the vacuum energy density and tend to cancel the effect of the bosons. Indeed if there were a pairing of fermions and bosons in nature, with each pair having the same mass, the contributions to the vacuum energy density would cancel and the vacuum energy density would vanish. Just such a pairing has been proposed and studied, and is called supersymmetry.[14] But alas, no supersymmetry partners of the known particles have been observed experimentally despite intense effort. This means that if supersymmetry is valid then it must be badly broken and, for example, the supersymmetric partner of the electron would have to be much more massive than the electron—thereby having escaped experimental detection to date. If supersymmetric partners were to exist with large masses of order M, then the vacuum energy would cancel for frequencies much greater than M and be finite. However it would not cancel for smaller frequencies and an expression like Eq. (17) would still result, with an effective cutoff of order M. The M might be speculated to be of order 1 TeV, giving an energy density comparable to that in Eq. (18). An outgrowth of supersymmetry is called superstring theory, in which small pieces of string are supposed to describe all the particles and forces of nature, including gravity. Here too the vacuum energy density is enormous and the catastrophe remains.[15]

Finally we might ask about the zero point energy of the gravitational field. There is at present no successful quantum theory of gravity, but we might expect on the basis of studies of weak gravitational waves in general relativity that the field would also have a ground state energy of $\omega/2$ for each mode and the two polarization states of the waves.[16] This would add yet another two to the parentheses in Eq. (20). We can say nothing about the behavior of the quantized gravitational field when the energy density becomes as large as we have been discussing, but we have no reason to expect that the vacuum energy density would be magically cancelled by some gravitational effect.

Let us emphasize that the remarkable property of the vacuum energy density is not that it is divergent, since the divergence can be understood on physical grounds. It is that it is so large even when a reasonable cutoff on frequency is used or when theories such as supersymmetry are invoked.

VII. OBSERVATIONAL VERIFICATION: THE CASIMIR EFFECT

Despite the conceptual oddity of the energy density of the vacuum there is ample experimental evidence for it. We here outline the derivation of the Casimir effect, in which a shift in the vacuum energy density produces a force on closely spaced conducting plates.[17,18]

On the interior surface of the plates the electric field must be zero, which provides a constraint on the modes of the field between the plates; a mode with a wavelength greater than $2d$ perpendicular to the plates cannot fit inside. Thus when computing the vacuum energy density in the region between the plates the sum over modes given by the integral in Eq. (17) must be modified. The result of the modified summation is[18]

$$\rho = \frac{\omega_{\max}^4}{8\pi^2} - \frac{\pi^2}{720 d^4}. \tag{21}$$

(The d^{-4} dependence is expected on dimensional grounds.) Thus the lack of low frequency modes between the plates lowers the vacuum energy density in the interior. The total energy difference is (for plates of area L^2):

$$U = -\frac{\pi^2 L^2}{720 d^3}. \tag{22}$$

This will give rise to a pressure on the plates which is easy to calculate. If the separation of the plates is increased by Δd, the work done by the pressure P is $P\Delta V = PL^2 \Delta d$. This changes the energy in the interior by $-dU$. We equate these to find the pressure,

$$P = -\frac{dU}{d(d)}\frac{1}{L^2} = \frac{-\pi^2}{240 d^4}. \tag{23}$$

This corresponds to an inward force, the famous Casimir force. One may view the force as arising due to radiation pressure of the vacuum; "zero point photons" of momentum $\omega/2$ exert a pressure, and since there are more of them outside the plates a net inward force results.

For a small but macroscopic system the Casimir force is quite small (about 0.013 dyn/cm^2 for a separation of 1 μm). It was first measured long ago and found to be consistent with the prediction, to about a factor of 2.[19] We thus conclude that the vacuum energy density is a real and measurable phenomenon. Of course the total vacuum energy density is not measured by such an experiment, but only the shift due to the presence of the conductor, which is independent of the upper cutoff $\omega_{\max}$.

The force discussed above is the simplest one of a host of similar effects due to shifts in the vacuum energy density, now collectively termed Casimir forces.[18] These include forces between dielectric plates, atoms and metal plates, atoms and dielectric plates, neutral atoms (the van der Waals force), etc. An abundance of experimental evidence supports the reality of these forces.[18] Even the Lamb shift, a small and precisely measured correction to the spectrum of hydrogen,[20] may be viewed as a consequence of the vacuum energy shift. Indeed the quantum vacuum appears to be quite lively and much more than empty space.[18,21]

VIII. OBSERVATIONAL CONFLICT: ABSENCE OF GRAVITATIONAL EFFECT

While the measurements of the Casimir effect give a verification of the small shift in the vacuum energy density, they do not verify the total size. This appears to require a gravitational probe, which is easily found. Indeed we will show that there is an upper limit on the density provided by solar system dynamics, which is amusing since it is so easily understood using elementary ideas. A much more stringent limit is provided by cosmology, as is to be expected.

Suppose that the universe were indeed filled with a uniform vacuum energy. The density would be homogeneous and isotropic, that is, it would appear the same everywhere and in all directions. With regard to the vacuum energy density all points in the universe would be equivalent, and any one point could serve as the center of a radially symmetric universe. (Precisely the same situation occurs in theoretical cosmology, where, on a very large scale, the universe is assumed to be homogeneous and isotropic, so any point can be considered the center.) Next consider the outer planets of the solar system. Their orbits have been studied for centuries and are well understood on the basis of Newtonian theory in which the sun produces the dominant force and the other planets produce small corrections. The solar force on a planet of mass m at distance r from the sun is

$$F_s = \frac{GM_0 m}{r^2}, \tag{24}$$

where G is the gravitational constant and M_0 is the mass of the sun. If there were a significant energy or mass density ρ in the vacuum in the solar system then there would be an additional force affecting planetary orbits. According to classical mechanics this would be due to the total mass $M(r)$ *inside* a sphere of radius r and would be given by

$$F_v = \frac{GM(r)m}{r^2} = \frac{4\pi\rho Gmr}{3}. \tag{25}$$

The uniform density *outside* the sphere would not produce a force on the planet. Since Newtonian theory provides a very good description of the planetary orbits we know that F_v must be very much less than F_s. This leads to

$$\rho \ll M_0 \Big/ \left(\frac{4\pi}{3}\right) r^3. \tag{26}$$

For a numerical estimate we substitute into this $M_0 = 1.99\times10^{30}$ kg and use the orbital radius of Pluto, $r = 5.9\times10^{12}$ m, to find

$$\rho \ll 2.3\times10^{-9}\ \text{kg/m}^3 = 6.6\times10^{-27}\ \text{f}^{-4}. \tag{27}$$

This limit is remarkable for its simplicity and lack of assumptions—and the disagreement of many orders of magnitude with the theoretical estimates (18) or (19).

General relativity theory provides a perfect analog of the above result.[22] The field equations including the cosmological term (see Sec. IX below) may be solved exactly for the spherically symmetric case to yield a solution which is a generalization of the well-known Schwarzschild solution; in the classical limit of weak fields it gives precisely the results (24) and (25) for the Newtonian and cosmological parts of the potential and force.

A much more stringent upper limit on the density can be obtained by considering larger scales in the universe. The observed average mass density of the universe is unfortunately poorly known. For matter in visible stars it is estimated to be about 10^{-28} kg/m^3.[23] However many astronomers believe that the total mass density may be much larger, due to so-called dark matter, which is not visible but can be inferred from the gravitational behavior of galaxies and clusters of galaxies. Such matter may have a density as much as 100 times the visible matter, or about 10^{-26} kg/m^3. If the energy or mass density of the vacuum were much larger than this then the behavior of the universe on the cosmological scale would be greatly affected. Since the present cosmological model seems to describe the universe rather well we can use this as a rough upper limit estimate for the vacuum energy density, or

$$\rho < 10^{-26}\ \text{kg/m}^3 = 2.9\times10^{-44}\ \text{f}^{-4}. \tag{28}$$

Comparing this upper limit with the theoretical estimate (19) based on the Planck cutoff energy we see that the two differ by 120 orders of magnitude. This is the vacuum catastrophe in its extreme form. It certainly represents the largest discrepancy between theory and experiment in all of science.

IX. AN AD HOC CANCELING TERM: THE COSMOLOGICAL CONSTANT

Our present purpose is to present an introduction and not a solution to the vacuum catastrophe. However we will briefly discuss an ad hoc canceling term, which is allowed by the mathematics. In most of classical physics and quantum mechanics the absolute value of the energy scale is not important, so one may add a constant term to potentials and the Hamiltonian without any effect on the physics. The critical exception is gravity, which according to general relativity is produced by *all* energy and momentum. Thus in textbooks on quantum field theory the zero point energy which occurs when fields are quantized is simply discarded, with the justification that one may cancel it by subtracting a constant (even an enormous or infinite constant) from the Hamiltonian with no physical effect—*if* one neglects gravity.[2]

From our discussion we see that the constant must be chosen to be enormous and must cancel the vacuum energy density to 120 digits, so it must be chosen very accurately! Both the ad hoc and unjustified use of such a constant and its incredible size and accuracy make this a manifestly poor solution, unacceptable in the view of most physicists.[3]

It is remarkable that general relativity also allows such a canceling constant term, which is called the cosmological constant. This was introduced by Einstein in 1917 into the original field equations of general relativity in order to allow a static cosmological solution to the equations, one in which the universe is static and does not expand.[12] After the observational discovery of the expansion of the universe it was abandoned by Einstein as unjustified and unacceptable. However it is allowed by the mathematics and indeed corresponds to a constant vacuum energy density with an arbitrary magnitude and sign. As such it may be chosen to cancel the vacuum energy density due to quantum effects. That is, it plays exactly the same role as an ad hoc constant term in the Hamiltonian of quantum field theory, and is no more justified.

In summary the use of an ad hoc constant term in the quantum field Hamiltonian, or equivalently a cosmological constant term in the field equations of general relativity, can cancel the vacuum energy density, but its use has no theoretical justification and many theorists consider that it really amounts to ignoring the problem.

X. CONCLUSION

Quantum field theory predicts without ambiguity that the vacuum has an energy density many orders of magnitude greater than nuclear density. Measurements of the Casimir force between conducting plates and related forces verify that the shift in this energy is real, but considerations of gravity in the solar system and in cosmology imply stringent upper limits on the magnitude, which are in extreme conflict with the theoretical estimate, by some hundred orders of magnitude! Unless one considers an ad hoc constant canceling term an adequate explanation then there appears to be a serious conflict between our concepts of the quantum vacuum and gravity; that is, there is a vacuum catastrophe.

A few brief comments on the ultraviolet catastrophe, the discovery of the quantum nature of light, and the first suggestion of a zero point energy, both made by Planck, are appropriate here; original references are given in the books by Milonni[5] and by Kuhn.[6] A number of people (notably Boltzmann, Wien, and Planck) had done work leading to a theoretical expression for the general form of the blackbody radiation spectrum from thermodynamic arguments and experimental data, some of it obtained in 1900. This led Planck to "give it a real physical meaning" in his theoretical derivation of 1900, which invoked the assumption that radiation is absorbed and emitted by a blackbody in quanta. This he termed a "desperate solution." (Some authors such as Kuhn[6] however disagree with this scenario and argue that Planck in 1900 really used the quanta as a formal device devoid of physical meaning.) In the same year the Rayleigh–Jeans law for the blackbody spectrum was obtained using purely classical radiation theory and the equipartition theorem, and was in disagreement with the data and in fact diverged for large frequencies. This situation was only later in 1911 called by Ehrenfest the ultraviolet catastrophe; thus the ultraviolet catastrophe was solved before it was given that name. The relevant fact is that the blackbody spectrum required going beyond the standard physics of the time to obtain a result free of divergence and in agreement with experiment. Planck was also responsible for the first suggestion of a zero point energy, long before the development of quantum mechanics. After further work to understand the rather ad hoc introduction of the quantum, he suggested in 1912 that when the temperature of a blackbody goes to zero the atomic oscillators still retain a zero point energy $\omega/2$, although he believed that it would have no experimental consequences. The first suggestion that the electromagnetic field itself contains zero point energy was made later, by Ernst in 1916.

It is no surprise that the cosmological constant problem, or vacuum catastrophe, has attracted a huge amount of attention from theorists since it may be that such a conflict could provide a clue to the unification of gravity and quantum theory. Many interesting and ingenious and desperate attempts at a solution have been proposed. None has been accepted as clearly correct. We refer the reader to the review article of Weinberg[3] for a discussion of such attempts before about 1988. References 24 and 25 contain elementary discussions and the technical Ref. 26 is an example of current work.

ACKNOWLEDGMENTS

We wish to thank L. Susskind of Stanford University who first mentioned this problem to us many years ago. L. Williams and the physics club of San Francisco State University made us aware that it might be understood by undergraduates. J. Greensite of San Francisco State University and M. Peskin of the Stanford Linear Accelerator Center were most helpful with their comments on the physics.

[a]The authors are also visitors at Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309.

[1]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965), pp. 74–77.

[2]C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980), pp. 127–133.

[3]S. Weinberg, "The Cosmological Constant Problem," Rev. Mod. Phys. **61**, (1), 1–82 (1989). This contains an extensive historical and technical discussion of work up to about 1988.

[4]R. J. Adler, M. Bazin, and M. M. Schiffer, *Introduction to General Relativity* (McGraw-Hill, New York, 1975), 2nd ed., pp. 344, 345.

[5]P. W. Milonni, *The Quantum Vacuum* (Academic Press, San Diego, CA, 1994), pp. 1–9.

[6]T. H. Kuhn, *Black-Body Theory and the Quantum Discontinuity 1894–1912* (Oxford University Press, New York, 1978), pp. 92–110.

[7]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964), p. 46.

[8]F. Halzen and A. Martin, *Quarks and Leptons* (Wiley, New York, 1984), pp. 12, 13.

[9]P. A. Goswami, *Quantum Mechanics* (Brown, Dubuque, IA, 1992) pp. 130–151.

[10]P. A. Tipler, *Physics*, 3rd ed. (Worth, New York, 1991), Vol. 2, pp. 947–957.

[11]D. Saxon, *Elementary Quantum Mechanics* (McGraw-Hill, New York, 1968), pp. 215, 216 and 267.

[12]R. J. Adler, "General Relativity," in *Encyclopedia of Physics*, edited by S. Parker (McGraw-Hill, New York, 1992), pp. 287–295.

[13]D. Griffiths, *Introduction To Elementary Particles* (Wiley, New York, 1987), pp. 279–301.

[14]M. Kaku, *Quantum Field Theory* (Oxford University Press, New York, 1993), pp. 663–699.

[15]M. A. Green, J. H. Schwarz, and E. Witten, *Superstring Theory* (Cambridge University Press, New York, 1987), Vol. 2, Chaps. 8 and 15.

[16]See Ref. 4, p. 317.

[17]See Ref. 2, p. 138.

[18]P. W. Milonni and M.-L. Shih, "Casimir Forces," Contemporary Physics **33**, (5), 313–322 (1992).

[19] M. J. Sparnaay, "Measurements of Attractive Forces Between Flat Plates," Physica **24**, 751–764 (1958).
[20] See Ref. 7, p. 52.
[21] T. H. Boyer, "The Classical Vacuum," Scientific American, August 1985, pp. 70–78. This contains a discussion of various properties of the vacuum which are due to quantum effects; one may therefore consider the title to be somewhat inappropriate.
[22] See Ref. 4, p. 396.
[23] See Ref. 4, p. 432.
[24] L. Abbott "The Mystery of the Cosmological Constant," Scientific American, May 1988, pp. 106–113.
[25] S. Weinberg, *Dreams of a Final Theory* (Pantheon, Random House, New York, 1992), pp. 225–229.
[26] M. D. Maia and G. S. Silva, "Geometrical Constraints on the Cosmological Constant," Fermilab Pub **94**, 002 A (1994).

The Cartesian rank two representation of spherical tensors with an application to Raman scattering

Eddy H. van Kleef
Dipartimento di Chimica, Universita' di Perugia, Via Elce di Sotto 8, 06100 Perugia, Italy

(Received 27 September 1993; accepted 27 January 1995)

The concept of Cartesian representations of lower rank spherical tensors is reviewed. To facilitate the introduction of the latter, basis vectors and tensors that transform according to the laws for spherical tensors are introduced, as well as a fast projection of a rank two Cartesian tensor onto its spherical components. From this the rotational invariants of the tensors are deduced. The method is applied to the problem of orientational averaging of the Raman polarizability tensor. General, complex, expressions for the depolarization ratio and the reversal ratio are derived. The entire treatment is carried out in a classical context.

I. INTRODUCTION

Tensor algebra offers mathematical tools well-known to the physicist. Whenever *directional* relationships need to be expressed, the standard formulations in terms of scalar equations no longer suffice, and vector equations appear, e.g.,

$$\mathbf{a}=\mathbf{f}(\mathbf{b}), \tag{1}$$

where a and b are vectors and $\mathbf{f}(\,)$ is a vector function. When the relationship between the two vectors is *linear*, it can be expressed by means of a Cartesian[1] tensor of rank two $\mathscr{A}$. Symbolically, it can then be written as

$$\mathbf{a}=\mathscr{A}\cdot\mathbf{b}.$$

Choosing a coordinate system (i.e., an appropriate basis), the tensor can be written as a matrix of rank two, and the vector relationship can be expressed as

$$a_i=\sum_j \mathscr{A}_{ij}b_j,$$

where i,j are indices running over the Cartesian axes: x,y,z and the individual $\mathscr{A}_{ij}$ are called the *elements* of the tensor. These equations can easily be extended to higher orders where, e.g., the dependence of a vector property on *several* other vectors needs to be expressed.[2,3] In this case Cartesian tensors and matrices of higher rank will appear.

Tensors in general are *defined* by the properties of their components under transformation, i.e., change of coordinate system. The Cartesian tensors introduced above transform according to certain well-known laws, which relate the change of the tensor to the change in the basis vectors. In the simplest case of a rank two Cartesian tensor (as above), this leads to the transformation law for vector operators:

$$\mathscr{A}'_{ij}=\sum_{k,l}\mathscr{B}^{-1}_{ik}\mathscr{A}_{kl}\mathscr{B}_{lj},$$

or, symbolically,

$$\mathscr{A}'=\mathscr{B}^{-1}\cdot\mathscr{A}\cdot\mathscr{B}, \tag{2}$$

where $\mathscr{A}'$ is the transformed tensor and $\mathscr{B}$ is the transformation matrix that expresses the new basis vectors in terms of the old ones.

Transformations of Cartesian tensors are readily carried out under stretching or reflection operations. Transformations involving rotations are, in general, much more cumbersome. This is because they introduce off-diagonal elements in the transformation matrices $\mathscr{B}$. This implies that, for the most general rotation operations, any element of the "new tensor" will depend upon all the elements of the "old tensor." The operation "mixes all the elements of the tensor."

Spherical tensors[4–10] deal with rotational transformations in a much more friendly way. Their elements are generally denoted as $T(k,q)$, where k is the *rank* and q is the *index* of the tensor. The index q runs from $-k$ to k, which means that a spherical tensor of rank two has five elements. Spherical tensors transform according to a law that, under rotations, mixes a minimum number of tensor elements:

$$T'(k,q)=\sum_{q'=-k}^{k}D^k_{q'q}(R)T(k,q'), \tag{3}$$

where the coefficients $D^k_{q'q}(R)$ are called the *rotation matrix elements*. These have been tabulated for low k.[10] The symbol R indicates the set of angles that characterize the rotation applied. These are often taken as the *Euler angles* ϕ,θ,χ.[4,5,7,10–12] The important point is that, for a given rank

